LINEARE ALGEBRA

DIETER HOFFMANN

Für Louise, Luca, Nicolas, Gabriel, Étienne und Léon

Bei diesem Text handelt es sich um Notizen zu meiner einsemestrigen Vorlesung *Lineare Algebra* für Studierende der *Mathematischen Finanzökonomie* (MFÖ), der *Physik mit „großer Mathematik“* und der *Informatik* (Drittsemester) an der Universität Konstanz im Wintersemester 2012/13.

Inhaltsverzeichnis

Dieser Begleittext zur Vorlesung beruht in einigen Teilen auf dem Skript, das Herr Kollege Oliver Schnürer dankeswerterweise für seine Vorlesung im vorletzten Wintersemester auf seiner homepage — inklusive Quelle — zur Verfügung gestellt hat. Von meinem Skript des Wintersemesters 2011/12 weicht dieser Text nur geringfügig ab.

0. Vorbemerkungen

0.1. **Lineare Algebra in der Mathematik.**

Lineare Algebra ist eine der beiden Grundvorlesungen, deren sämtliche Inhalte später in allen mathematischen Gebieten und Anwendungen ständig auftauchen werden. Sie bietet daher einen guten Einstieg in die Mathematik.

Die lineare Algebra beschäftigt sich mit Objekten, die *linear,* d. h. irgendwie ‚gerade', aussehen, also wie eine Gerade oder eine Ebene. Die Realität ist aber in der Regel *nicht* linear: *It's a Nonlinear World*

`http://www.springer.com/mathematics/applications/book/978-0-387-75338-6`

Jedoch sehen Dinge aus der Nähe — also lokal — häufig linear aus, z. B. ein kleines Stück der Oberfläche des Bodensees.

0.2. **Was kann man mit Mathematik anfangen?**

Nach `http://www.maths.monash.edu.au/` eine ganz kleine Auswahl:

- Understanding our world (Bewegung der Sterne, schwarze Löcher, Wasserwellen, Windhose, Buschbrände)
- Modelling and improving systems (Verkehrsleitsysteme, Logistik für Containerschiffe, Börse, Produktion, *Medizin*, ...)
- Studying the beauty of perfection (Seifenblasen, Symmetrien in Sonnenblumen oder geometrischen Mustern, Fraktale, Wassertropfen)

0.3. **Literatur.**

Die meisten Bücher über lineare Algebra sind — ergänzend zu diesem Skript — geeignet, zum Beispiel die Bücher von U. Stammbach [13], Gerd Fischer [2], Jürgen Hausen [7], Max Koecher [10] und Serge Lang [11], alle etwa mit dem Titel „Lineare Algebra" und — mit Vorsicht! — auch Wikipedia (`http://www.wikipedia.org/`) für Definitionen. Für Studierende der Physik sind auch die entsprechenden Kapitel aus Helmut Fischer und Helmut Kaul [3] empfehlenswert.

1. Lineare Gleichungssysteme

1.1. **Lineare Gleichungssysteme, Einstieg.**

Beispiel 1.1.1. Wir wollen das folgende Gleichungssystem lösen:

$$\begin{array}{rcrcrcr} 2x & + & 6y & - & 4z & = & 10 \\ -x & + & 5y & + & 2z & = & 11 \\ 3x & + & 7y & + & z & = & -3 \end{array}$$

Dazu dividieren wir die erste Zeile durch 2 und erhalten:

$$\begin{array}{rcrcrcr} x & + & 3y & - & 2z & = & 5 \\ -x & + & 5y & + & 2z & = & 11 \\ 3x & + & 7y & + & z & = & -3 \end{array}$$

Wir addieren die nun erste Zeile zur zweiten Zeile und subtrahieren sie, mit 3 multipliziert, von der dritten Zeile:

$$\begin{array}{rcrcrcr} x & + & 3y & - & 2z & = & 5 \\ & & 8y & & & = & 16 \\ & - & 2y & + & 7z & = & -18 \end{array}$$

Dividiere die zweite Zeile durch 4 und addiere das Resultat zur dritten Zeile:

$$\begin{array}{rcrcrcr} x & + & 3y & - & 2z & = & 5 \\ & & 2y & & & = & 4 \\ & & & & 7z & = & -14 \end{array}$$

Hieraus ergibt sich von unten nach oben:

$$\begin{array}{rcr} z & = & -2 \\ y & = & 2 \\ x & = & -5 \end{array}$$

Wir haben also die Lösung $(x, y, z) = (-5, 2, -2)$. D. h. für $x = -5$, $y = 2$ und $z = -2$ sind sämtliche Gleichungen des obigen Gleichungssystems erfüllt. ◁

Wir wollen jedoch nicht in erster Linie spezielle Probleme lösen, sondern Aussagen über *allgemeine Probleme* machen.

Definition 1.1.2 (Lineares Gleichungssystem)**.**
Ein *reelles lineares Gleichungssystem* in n Unbekannten $x^1, \dots, x^n$ und m Gleichungen ist ein System von Gleichungen der Form

$$\begin{array}{ccccccccc} a_1^1 x^1 & + & a_2^1 x^2 & + & \cdots & + & a_n^1 x^n & = & b^1 \\ a_1^2 x^1 & + & a_2^2 x^2 & + & \cdots & + & a_n^2 x^n & = & b^2 \\ \vdots & & & & & & \vdots & & \vdots \\ a_1^m x^1 & + & a_2^m x^2 & + & \cdots & + & a_n^m x^n & = & b^m \end{array} \tag{1.1}$$

Dabei sind a_ν^μ und b^μ, $1 \le \mu \le m$, $1 \le \nu \le n$, vorgegebene reelle Zahlen, also $a_\nu^\mu, b^\mu \in \mathbb{R}$.
(Anders als in vielen Büchern über lineare Algebra schreiben wir den Zeilenindex der *Koeffizienten* a_ν^μ nach oben. (a_ν^μ hier entspricht $a_{\mu\nu}$ in solchen Büchern.) Dies bedeutet keine Potenzierung und ist eine besonders in der Differentialgeometrie und in der Physik (EINSTEINsche Summenkonvention) übliche Schreibweise.)
Das System heißt *homogen*, falls $b^1 = \cdots = b^m = 0$ gilt und sonst *inhomogen*.
Ein n-Tupel $(\xi^1, \dots, \xi^n)$ reeller Zahlen, also $(\xi^1, \dots, \xi^n) \in \mathbb{R}^n$, heißt genau dann *Lösung* des linearen Gleichungssystems, wenn sämtliche Gleichungen des Gleichungssystems erfüllt sind, wenn wir die Zahlen ξ^ν statt x^ν für $1 \le \nu \le n$ einsetzen.

Satz 1.1.3.
Es sei (H) ein homogenes *lineares Gleichungssystem wie in (1.1). Sind $(\xi^1,\dots,\xi^n)$ und $(\eta^1,\dots,\eta^n)$ Lösungen von (H) und $\lambda\in\mathbb{R}$, dann sind $(\xi^1+\eta^1,\dots,\xi^n+\eta^n)$ und $(\lambda\xi^1,\dots,\lambda\xi^n)$ ebenfalls Lösungen von (H).*

Beweis: Nach Voraussetzung gilt für alle $1\le\mu\le m$

$$a_1^\mu\xi^1+a_2^\mu\xi^2+\cdots+a_n^\mu\xi^n=0, \tag{1.2}$$

$$a_1^\mu\eta^1+a_2^\mu\eta^2+\cdots+a_n^\mu\eta^n=0. \tag{1.3}$$

Wir addieren (1.2) und (1.3) und erhalten

$$a_1^\mu(\xi^1+\eta^1)+a_2^\mu(\xi^2+\eta^2)+\cdots+a_n^\mu(\xi^n+\eta^n)=0.$$

Somit ist $(\xi^1+\eta^1,\cdots,\xi^n+\eta^n)$ ebenfalls eine Lösung von (H).
Multipliziere (1.2) mit λ und erhalte

$$a_1^\mu(\lambda\xi^1)+a_2^\mu(\lambda\xi^2)+\cdots+a_n^\mu(\lambda\xi^n)=0.$$

Daher ist $(\lambda\xi^1,\dots,\lambda\xi^n)$, wie behauptet, auch eine Lösung von (H). □

Bemerkung 1.1.4.
Es sei (H) ein *homogenes* Gleichungssystem. Dann besitzt (H) stets die *‚triviale' Lösung* $(0,\dots,0)$.

Es sei (L) ein lineares Gleichungssystem wie in (1.1). Dann bezeichnen wir das homogene lineare Gleichungssystem, das aus (1.1) entsteht, wenn wir die Koeffizienten b^μ alle durch 0 ersetzen, als das *zugehörige homogene System* $(L)_H$.

Satz 1.1.5.
Es seien (L) ein lineares Gleichungssystem wie in (1.1) und $(L)_H$ das zugehörige homogene lineare Gleichungssystem.

(i) Ist $(\zeta^1,\dots,\zeta^n)$ eine Lösung von (L) und $(\xi^1,\dots,\xi^n)$ eine Lösung von $(L)_H$, so ist $(\zeta^1+\xi^1,\dots,\zeta^n+\xi^n)$ eine Lösung von (L).

(ii) Sind $(\zeta^1,\dots,\zeta^n)$ und $(\eta^1,\dots,\eta^n)$ Lösungen von (L), so ist die Differenz $(\zeta^1-\eta^1,\dots,\zeta^n-\eta^n)$ eine Lösung von $(L)_H$.

Beweis:

(i) Addiere

$$a_1^\mu\zeta^1+a_2^\mu\zeta^2+\cdots+a_n^\mu\zeta^n=b^\mu$$

und

$$a_1^\mu\xi^1+a_2^\mu\xi^2+\cdots+a_n^\mu\xi^n=0$$

und erhalte

$$a_1^\mu(\zeta^1+\xi^1)+a_2^\mu(\zeta^2+\xi^2)+\cdots+a_n^\mu(\zeta^n+\xi^n)=b^\mu.$$

Das zeigt die erste Behauptung.

(ii) Subtrahiere

$$a_1^\mu \eta^1 + a_2^\mu \eta^2 + \cdots + a_n^\mu \eta^n = b^\mu$$

von

$$a_1^\mu \zeta^1 + a_2^\mu \zeta^2 + \cdots + a_n^\mu \zeta^n = b^\mu$$

und erhalte

$$a_1^\mu(\zeta^1 - \eta^1) + a_2^\mu(\zeta^2 - \eta^2) + \cdots + a_n^\mu(\zeta^n - \eta^n) = 0\,,$$

die zweite Behauptung. □

Das zeigt:

Folgerung 1.1.6.
Hat man eine spezielle (‚partikuläre') Lösung (des inhomogenen Systems), dann erhält man alle *Lösungen, indem man zu dieser alle Lösungen des zugehörigen homogenen Systems hinzuaddiert.*

Beweis: ∘ ∘ ∘ □[1]

Folgerung 1.1.7.
Das Gleichungssystem (L) sei lösbar. Dann ist es genau dann eindeutig *lösbar, wenn $(L)_H$ nur die triviale Lösung besitzt. Besitzt $(L)_H$ eine nicht-triviale Lösung, so besitzt (L)* unendlich viele *Lösungen.*

Beweis: ∘ ∘ ∘ □

Für n-Tupel verwenden wir die Abkürzungen

$$\eta := (\eta^1, \eta^2, \ldots, \eta^n),$$
$$\zeta := (\zeta^1, \zeta^2, \ldots, \zeta^n), \ldots,$$

und definieren

$$\eta + \zeta := (\eta^1 + \zeta^1, \eta^2 + \zeta^2, \ldots, \eta^n + \zeta^n)$$

sowie für $\lambda \in \mathbb{R}$

$$\lambda\eta := (\lambda\eta^1, \lambda\eta^2, \ldots, \lambda\eta^n).$$

Lemma 1.1.8.
Es sei (L) ein lineares Gleichungssystem. Dann ist die Lösungsmenge unter den folgenden Operationen invariant:

(i) Vertauschen von Gleichungen
(ii) Multiplikation einer Gleichung mit $0 \neq \lambda \in \mathbb{R}$
(iii) Addition des λ-fachen der k-ten Gleichung zur ℓ-ten Gleichung (für $\lambda \in \mathbb{R}$ und $k, \ell \in \{1, \ldots, m\}$ mit $k \neq \ell$)

[1] So notiere ich gelegentlich Stellen, an denen ich einen Routine-Beweis nur in der Vorlesung ausführe.

Beweis: Es seien L_1 die Lösungsmenge von (L) und L_2 die Lösungsmenge des Systems, das wir erhalten, wenn wir die jeweilige Operation anwenden.

(i) Klar
(ii) Für jede Zahl $\lambda \in \mathbb{R}$ gilt $L_1 \subset L_2$. Ist $\lambda \neq 0$, so erhalten wir durch Multiplikation derselben Gleichung mit $\frac{1}{\lambda}$ gerade wieder (L). Somit gilt auch $L_2 \subset L_1$, insgesamt also $L_1 = L_2$.
(iii) Auch hier ist $L_1 \subset L_2$ wieder klar. Addition des $(-\lambda)$-fachen der k-ten Gleichung zur ℓ-ten Gleichung liefert wieder (L), also gilt auch $L_2 \subset L_1$, und daher ist auch hier wieder $L_1 = L_2$. □

1.2. **Gaußsches Eliminationsverfahren.**

Das Gaußsche Eliminationsverfahren erlaubt es, allgemeine lineare Gleichungssysteme zu lösen oder auch nachzuweisen, daß sie keine Lösung besitzen.

Definition 1.2.1 (Zeilenstufenform)**.**
Es sei (L) ein lineares Gleichungssystem in der Form (1.1). Dann ist (L) genau dann in *Zeilenstufenform*, falls zunächst ‚in jeder Zeile links mehr Koeffizienten a^μ_ν Null sind als in der Vorgängerzeile' und dann höchstens noch Zeilen, in denen alle Koeffizienten Null sind, folgen. Genauer:
Es gibt eine Zahl $\mu_0 \in \{0, \dots, m\}$ derart, daß in den Zeilen mit Index 1 bis μ_0 jeweils nicht alle Koeffizienten Null sind und in den folgenden Zeilen alle Koeffizienten Null sind.
Für jedes $\mu \in \{1, \dots, \mu_0\}$ betrachten wir den niedrigsten Spaltenindex $n(\mu)$ dessen zugehöriger Koeffizient in der μ-ten Zeile ungleich Null ist, also

$$n(\mu) := \min\{j \in \mathbb{N} \colon a^\mu_j \neq 0\}\,.$$

Es gilt $1 \leq n(\mu) \leq n$, und die *Stufenbedingung* lautet:

$$n(1) < \cdots < n(\mu_0)$$

Zu beachten ist, daß der Fall $\mu_0 = 0$ zugelassen ist. In diesem Fall sind alle Koeffizienten gleich Null.
Die besonders ausgezeichneten — von Null verschiedenen — Koeffizienten $a^1_{n(1)}$ bis $a^{\mu_0}_{n(\mu_0)}$ heißen *Pivotelemente* oder kurz *Pivots* (deutsch Angel- oder Drehpunkte).

Besonders übersichtlich ist für viele Überlegungen der *Spezialfall*

$$n(1) = 1, \dots n(\mu_0) = \mu_0\,.$$

Durch ein Umnumerierung der Unbekannten des Gleichungssystems (Buchführung!) kann dies stets erreicht werden. *Für die folgenden Überlegungen habe nun das Gleichungssystems diese einfache Form:*

$$
\begin{array}{lllllllllllll}
a_1^1\,x^1 & + & a_2^1\,x^2 & + & \cdots & & & & + & a_n^1\,x^n & = & b^1 \\
0\,x^1 & + & a_2^2\,x^2 & + & \cdots & & & & + & a_n^2\,x^n & = & b^2 \\
\vdots & & & & & & & & & \vdots & & \vdots \\
0\,x^1 & + & 0\,x^2 & + & \cdots & + & a_{\mu_0}^{\mu_0}\,x^{\mu_0} + \cdots & & + & a_n^{\mu_0}\,x^n & = & b^{\mu_0} \\
0\,x^1 & + & 0\,x^2 & + & \cdots & & & & + & 0\,x^n & = & b^{\mu_0+1} \\
\vdots & & & & & & & & & \vdots & & \vdots \\
0\,x^1 & + & 0\,x^2 & + & \cdots & & & & + & 0\,x^n & = & b^m
\end{array}
$$

Bemerkung 1.2.2.
Gibt es ein $\mu \in \{\mu_0+1, \ldots, m\}$ *mit* $b^\mu \neq 0$*, dann ist das Gleichungssystem nicht lösbar.*

Beweis: ∘ ∘ ∘ □

Gilt hingegen $b^\mu = 0$ für $\mu \in \{\mu_0+1, \ldots, m\}$, dann lassen sich Lösungen einfach bestimmen:

x^{μ_0+1} bis x^n sind *freie Variablen*, sie können beliebige Werte annehmen. x^1 bis x^{μ_0} sind *gebundene Variablen*. Sie sind jeweils nach Festgelegung der Werte für die freie Variablen eindeutig bestimmt:
Mit $k := n-\mu_0$ seien zu k beliebigen ‚Parametern' λ^κ $(\kappa = 1, \ldots, k)$ die freien Variablen

$$x^{\mu_0+1} := \lambda^1, \ldots, x^n := \lambda^k$$

gewählt. Dann sind die gebundenen Variablen — für eine Lösung des Gleichungssystems — dadurch eindeutig festgelegt (von unten nach oben die Gleichungen auflösen ...).

Der wichtige *Spezialfall* $m = n$, bei dem man ebensoviele Gleichungen wie Unbekannte hat (quadratisches Koeffizientenschema), verdient besondere Beachtung: Hat man in diesem Fall eine Zeilenstufenform mit $\mu_0 = n$, so gibt es — wegen $k = n - \mu_0 = 0$ keine freien Variablen, also stets genau eine Lösung.

Satz 1.2.3 (Gaußsches Eliminationsverfahren)**.**
Es sei (L) ein lineares Gleichungssystem wie in (1.1)*. Dann lässt sich (L) mit endlich vielen Operationen aus Lemma 1.1.8 in ein lineares Gleichungssystem (G) der Form* (1.1) *so umformen, daß (G) in Zeilenstufenform ist. (L) und (G) haben dieselbe Lösungsmenge.*

Beweis: Die Gleichheit der Lösungsmengen folgt direkt nach Lemma 1.1.8. Sind alle Koeffizienten a_ν^μ gleich Null, so sind wir fertig; denn wir haben schon eine Zeilenstufenform mit $\mu_0 = 0$.
Im anderen Fall können wir davon ausgehen — nach eventueller Vertauschung von Zeilen, daß in der *ersten* Zeile ein Koeffizient a_ν^1 von

Null verschieden ist. Nach eventueller Umnumerierung der Unbekannten (Buchführung!) sei dies a_1^1. Wir addieren nun für $\mu = 2, \ldots, m$ das $-\frac{a_1^\mu}{a_1^1}$-fache der ersten Zeile zur μ-ten und haben dann das obige System auf eine Form mit $a_1^\mu = 0$ für $2 \leq \mu \leq m$ gebracht:

$$\begin{array}{ccccccccc} a_1^1x^1 & + & a_2^1x^2 & + & \cdots & + & a_n^1x^n & = & b^1 \\ & & \widetilde{a}_2^2x^2 & + & \cdots & + & \widetilde{a}_n^2x^n & = & \widetilde{b}^2 \\ & & \vdots & & & & \vdots & & \vdots \\ & & \widetilde{a}_2^mx^2 & + & \cdots & + & \widetilde{a}_n^mx^n & = & \widetilde{b}^m \end{array}$$

Wir behalten die ursprünglichen Bezeichnungen bei, schreiben also wieder a_ν^μ statt $\widetilde{a}_\nu^\mu$ und b^μ statt $\widetilde{b}^\mu$.

Jetzt verfährt man mit dem verbleibenden System (ab der zweiten Zeile und der zweiten Spalte) — mutatis mutandis — entsprechend, usw. □

Ein ‚richtiger' Mathematiker liest natürlich so ein „usw" mit Grauen und erwartet an dieser Stelle einen formal sauberen Induktionsbeweis. Aber in *dieser* Vorlesung dürfen wir uns gelegentlich solche mehr beschreibenden Beweise erlauben.

Nach Multiplikation der μ-ten Zeile mit $\frac{1}{a_{n(\mu)}^\mu}$ für $\mu \in \{1, \ldots, \mu_0\}$ erhalten wir $a_{n(\mu)}^\mu = 1$. Wir können also ohne Einschränkung[2] annehmen, daß das Gleichungssystem in Zeilenstufenform $a_{n(\mu)}^\mu = 1$ erfüllt, daß also der erste von Null verschiedene Koeffizient gleich 1 ist, falls die Gleichung nicht trivial ist. Wir sprechen in diesem Fall von *spezieller Zeilenstufenform.*

Folgerung 1.2.4.
Es sei (H) ein homogenes *lineares Gleichungssystem der Form* (1.1) *mit $m < n$, also mit weniger Zeilen als Variablen. Dann besitzt (H) eine nicht-triviale Lösung, d. h. eine Lösung $(\xi^1, \ldots, \xi^n) \neq (0, \ldots, 0)$.*

Das ‚sieht' man — fast — ohne *Beweis*; ‚etwas' genauer geht das wie folgt: Wir bringen das Gleichungssystem in Zeilenstufenform (wieder mit $n(\mu) = \mu$ für $\mu \in \{1, \ldots, \mu_0\}$). Wegen $\mu_0 \leq m < n$ ist $k := n - \mu_0 > 0$. So ist zumindest ξ^n frei wählbar. □

Folgerung 1.2.5.
Es sei (L) ein beliebiges lineares Gleichungssystem wie in (1.1) *mit $m = n$, also mit genauso vielen Gleichungen wie Variablen.*
Besitzt dann $(L)_H$ nur die triviale Lösung, so ist (L) eindeutig lösbar.

Beweis: Da $(L)_H$ nur die triviale Lösung besitzt, hat (L) — auf Zeilenstufenform gebracht — die Gestalt

$$\begin{array}{ccccccccc} a_1^1x^1 & + & a_2^1x^2 & + & \ldots & + & a_n^1x^n & = & b^1 \\ & & a_2^2x^2 & + & \ldots & + & a_n^2x^n & = & b^2 \\ & & & & \ddots & & \vdots & & \vdots \\ & & & & & & a_n^nx^n & = & b^n \end{array}$$

[2] Dafür schreiben wir im Folgenden kurz Œ.

mit $a^\nu_\nu \neq 0$ für $1 \leq \nu \leq n$, also „obere Dreiecksgestalt“. Daher besitzt (L) eine eindeutige Lösung. □

Wir verwenden oft die Bezeichnungen

$$\mathbb{N} := \{1, 2, 3, \ldots\},\ \mathbb{N}_0 := \{0, 1, 2, 3, \ldots\} \text{ und } \mathbb{N}_k := \{k, k+1, k+2, \ldots\}$$

für $k \in \mathbb{N}$.

1.3. **Matrizen.**

Definition 1.3.1. Eine $(m \times n)$-Matrix A ist ein Element in $\mathbb{R}^{m \cdot n}$. Wir schreiben $A \in \mathbb{R}^{m \times n}$. A hat $m \cdot n$ reelle Komponenten, die wir mit a^μ_ν, $1 \leq \mu \leq m$, $1 \leq \nu \leq n$, bezeichnen: $A = (a^\mu_\nu)$ oder genauer $A = (a^\mu_\nu)_{\substack{1 \leq \mu \leq m \\ 1 \leq \nu \leq n}}$. Wir stellen Matrizen wie folgt dar:

$$A = \begin{pmatrix} a^1_1 & a^1_2 & \ldots & a^1_n \\ a^2_1 & a^2_2 & \ldots & a^2_n \\ \vdots & \vdots & & \vdots \\ a^m_1 & a^m_2 & \ldots & a^m_n \end{pmatrix}.$$

Der obere Index bezieht sich auf die *Zeilen,* der untere auf die *Spalten* der Matrix. Das Element a^μ_ν steht daher in Zeile μ und Spalte ν.

Es seien A, B zwei $(m \times n)$-Matrizen mit Komponenten a^μ_ν bzw. b^μ_ν. Wir definieren die *Gleichheit* und *Summe* durch

$$A = B \text{ genau dann, wenn } a^\mu_\nu = b^\mu_\nu \text{ für alle } \mu = 1, \ldots, m \text{ und } \nu = 1, \ldots, n\,,$$

$$A + B := (a^\mu_\nu + b^\mu_\nu)\,,$$

also jeweils *komponentenweise.* Es sei $\lambda \in \mathbb{R}$. Wir definieren das Produkt einer Matrix mit einer reellen Zahl komponentenweise durch

$$\lambda A := (\lambda a^\mu_\nu)\,.$$

Wir können n-Tupel $\xi \in \mathbb{R}^n$ als $(1 \times n)$-Matrizen auffassen. Dann stimmen die Operationen Summieren und Produktbildung mit einer reellen Zahl mit den für n-Tupel bereits definierten Operationen überein.

Definition 1.3.2. Eine $(n \times n)$-Matrix heißt *quadratische* Matrix. Die speziell $(n \times n)$-Matrix

$$I = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix}$$

heißt *Einheitsmatrix* der Ordnung n. Man schreibt auch $I = \mathbb{1}$, genauer $\mathbb{1}_n$. Wir definieren das KRONECKERsymbol durch

$$\delta^\mu_\nu := \begin{cases} 1, & \nu = \mu, \\ 0, & \text{sonst.} \end{cases}$$

Dann gilt $I = (\delta^\mu_\nu)$.

Eine quadratische Matrix heißt *Diagonalmatrix,* falls $a^\mu_\nu = 0$ für $\mu \neq \nu$ gilt. Die Elemente a^ν_ν heißen *Diagonalelemente*.

Definition 1.3.3 (Matrixmultiplikation)**.**
Es seien — mit natürlichen Zahlen m, n und k — A eine $(m \times n)$-Matrix und B eine $(n \times k)$-Matrix, mit Einträgen (a^μ_ν) und (b^ν_κ). Wenn — wie angegeben — die Spaltenzahl von A mit der Zeilenzahl von B übereinstimmt, definieren wir das Produkt $C = AB$, eine $(m \times k)$-Matrix (c^μ_κ), für $1 \leq \mu \leq m$ und $1 \leq \kappa \leq k$ durch:

$$c^\mu_\kappa := \sum_{\nu=1}^{n} a^\mu_\nu b^\nu_\kappa = a^\mu_1 b^1_\kappa + a^\mu_2 b^2_\kappa + \cdots + a^\mu_n b^n_\kappa$$

(Physiker benutzen die EINSTEINsche Summenkonvention und schreiben kurz $c^\mu_\kappa = a^\mu_\nu b^\nu_\kappa$.)

Beispiele 1.3.4. Es gilt

$$\begin{pmatrix} 1 & 2 \\ -1 & 0 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ -1 & 0 \\ 3 & 1 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ -1 & 0 \\ 2 & 1 \end{pmatrix} \quad \text{ist } \textit{nicht} \text{ definiert},$$

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 1 & 0 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix}.$$

Wie das letzte Beispiel zeigt, gilt im allgemeinen nicht $AB = BA$, selbst, wenn beide Verknüpfungen definiert sind. Gilt $AB = BA$, so sagen wir, daß die beiden Matrizen *vertauschen* oder *kommutieren.*

Lemma 1.3.5.
A sei eine $(m \times n)$-Matrix, B, C seien $(n \times k)$-Matrizen, D eine $(k \times \ell)$-Matrix und $\varrho \in \mathbb{R}$. Dann gelten die Rechenregeln

$$\begin{aligned} A(B + C) &= AB + AC, \\ (\varrho A)B &= \varrho(AB) = A(\varrho B), \\ A(BD) &= (AB)D =: ABD. \end{aligned}$$

Beweis:

$$\begin{aligned}
A(B+C) &= \left(\sum_{\nu=1}^{n} a_\nu^\mu (b_\kappa^\nu + c_\kappa^\nu)\right)_{\substack{1\le\mu\le m\\ 1\le\kappa\le k}} = \left(\sum_{\nu=1}^{n} a_\nu^\mu b_\kappa^\nu\right)_{\substack{1\le\mu\le m\\ 1\le\kappa\le k}} + \left(\sum_{\nu=1}^{n} a_\nu^\mu c_\kappa^\nu\right)_{\substack{1\le\mu\le m\\ 1\le\kappa\le k}} \\
&= AB + AC, \\
(\varrho A)B &= \left(\sum_{\nu=1}^{n} (\varrho a_\nu^\mu) b_\kappa^\nu\right)_{\substack{1\le\mu\le m\\ 1\le\kappa\le k}} = \underbrace{\left(\sum_{\nu=1}^{n} a_\nu^\mu (\varrho b_\kappa^\nu)\right)_{\substack{1\le\mu\le m\\ 1\le\kappa\le k}}}_{=A(\varrho B)} = \underbrace{\varrho \left(\sum_{\nu=1}^{n} a_\nu^\mu b_\kappa^\nu\right)_{\substack{1\le\mu\le m\\ 1\le\kappa\le k}}}_{=\varrho(AB)}, \\
A(BD) &= \left(\sum_{\nu=1}^{n} a_\nu^\mu \left(\sum_{\kappa=1}^{k} b_\kappa^\nu d_\lambda^\kappa\right)\right)_{\substack{1\le\mu\le m\\ 1\le\lambda\le \ell}} = \left(\sum_{\kappa=1}^{k} \left(\sum_{\nu=1}^{n} a_\nu^\mu b_\kappa^\nu\right) d_\lambda^\kappa\right)_{\substack{1\le\mu\le m\\ 1\le\lambda\le \ell}} \\
&= (AB)D. \qquad \square
\end{aligned}$$

Bemerkung 1.3.6. Es sei

$$\sum_{\nu=1}^{n} a_\nu^\mu x^\nu = b^\mu, \quad 1 \le \mu \le m,$$

ein lineares Gleichungssystem. Wir definieren die Koeffizientenmatrix A und die *erweiterte Koeffizientenmatrix* $(A \,|\, b)$ zu diesem linearen Gleichungssystem durch

$$A := \begin{pmatrix} a_1^1 & a_2^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 \\ \vdots & \vdots & & \vdots \\ a_1^m & a_2^m & \dots & a_n^m \end{pmatrix} \quad \text{und} \quad (A \,|\, b) := \begin{pmatrix} a_1^1 & a_2^1 & \dots & a_n^1 & b^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 & b^2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_1^m & a_2^m & \dots & a_n^m & b^m \end{pmatrix}.$$

(In $(A \,|\, b)$ verwendet man gelegentlich auch einen senkrechten Strich vor der letzten Spalte.) Definiere weiterhin

$$b = \begin{pmatrix} b^1 \\ b^2 \\ \vdots \\ b^m \end{pmatrix} \quad \text{und} \quad x = \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix}.$$

Dann können wir das Gleichungssystem in der konzisen und dadurch übersichtlichen Form

$$Ax = b$$

schreiben.

Alle Aussagen dieses Kapitels lassen sich nicht nur für lineare Gleichungssysteme und Matrizen im Reellen zeigen, sondern auch für beliebige Körper (die wie im nächsten Kapitel definieren werden). Lediglich die Aussage über unendlich viele Lösungen ist abzuändern.

2. Körper und Vektorräume

2.1. **Körper.**

Dieser Abschnitt orientiert sich in der Darstellung an Dieter Hoffmann [8].

Bei den folgenden Axiomen kommen u. a. Begriffe wie *„Kommutativität“* und *„Assoziativität“* für Addition und Multiplikation vor. Man ist leicht geneigt, diese wichtigen Eigenschaften als banal anzusehen. Deshalb weise ich vorweg darauf hin, daß es ein besonderer Glücksfall ist, wenn Operationen kommutativ und assoziativ sind, weil dann der Umgang mit ihnen besonders einfach wird. Die *Kommutativität* bedeutet, daß es bei der Verknüpfung von je zwei Elementen nicht auf die Reihenfolge ankommt, zum Beispiel im Folgenden $a + b = b + a$ oder $a \cdot b = b \cdot a$ z. B. für reelle Zahlen a und b. Schon die Subtraktion liefert ein Beipiel für eine Operation, die nicht kommutativ ist; denn es ist $3 - 5 \neq 5 - 3$. Auch bei der Zusammensetzung von Wörtern hat man keine Kommutativität, es kommt auf die Reihenfolge an: Ein Hausmädchen ist etwas anderes als ein Mädchenhaus, ein Hosenlatz etwas anderes als Latzhosen.
Natürlich ist auch bei fast allen Vorgängen im täglichen Leben beim zeitlichen Ablauf die Reihenfolge wesentlich: Beim Autofahren (mit einem Auto mit Schaltgetriebe) tritt man zuerst die Kupplung und legt erst dann den Gang ein. (Der Versuch der umgekehrten Reihenfolge führt zu einem anderen Ergebnis.) Diejenigen, die sich vor Jahren mit dem Rubik-Cube („Zauberwürfel“) beschäftigt haben, wissen, daß die Schwierigkeit in der hochgradigen Nicht-Kommutativität der einzelnen Operationen liegt; es kommt entscheidend auf die Reihenfolge an.
Die *Assoziativität* bedeutet, daß bei der Verknüpfung von drei Elementen die Art der Klammersetzung keine Rolle spielt: Für reelle Zahlen etwa hat man zum Beispiel: $(a + b) + c = a + (b + c)$ (Assoziativität der Addition) und entsprechend $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ (Assoziativität der Multiplikation). Auch hier liefert wieder die Subtraktion ein einfaches und wichtiges Beispiel einer Operation für Zahlen, die *nicht* assoziativ ist: $(9 - 7) - 2 \neq 9 - (7 - 2)$. Beispiele aus dem sprachlichen Bereich, die deutlich machen, daß die Art der Klammerung ganz wesentlich ist, sind etwa: ‚Mädchenhandelsschule‘, ‚Kindergartenfest‘, ‚Urinsekten‘ und ‚Urinstinkt‘.

Definition 2.1.1 (Körper).

Ein Körper ist ein Tripel[a] $(\mathbb{K}, +, \cdot)$, bestehend aus einer nicht-leeren Menge $\mathbb{K}$ und zwei Abbildungen, *„Addition“* und *„Multiplikation“*,

$$\begin{array}{ccccc} + : & \mathbb{K}\times\mathbb{K} & \longrightarrow & \mathbb{K} & \\ & \cup & & \cup & \text{und} \\ & (a,b) & \longmapsto & a+b & \end{array} \qquad \begin{array}{ccccc} \cdot : & \mathbb{K}\times\mathbb{K} & \longrightarrow & \mathbb{K} & \\ & \cup & & \cup & \\ & (a,b) & \longmapsto & a\cdot b =: ab & \end{array}$$

derart, daß die folgenden Gesetze gelten:

(A1) $\forall\, a,b,c \in \mathbb{K} \quad (a+b)+c \;=\; a+(b+c) =: a+b+c$
„Assoziativität der Addition“

(A2) $\forall\, a,b \in \mathbb{K} \quad a+b \;=\; b+a$ *„Kommutativität der Addition“*

(A3) $\exists\, 0 \in \mathbb{K} \quad \forall a \in \mathbb{K} \quad a+0 = a$ *„Null“* oder *„Nullelement“*

(A4) $\forall\, a \in \mathbb{K} \quad \exists\, -a \in \mathbb{K} \quad a+(-a) \;=\; 0$
„Inverses Element zu a bezüglich $+$“, meist gelesen als *„minus a“*

Für die Multiplikation gelten ganz analog — bis auf die Sonderstellung der Null — :

(M1) $\forall\, a,b,c \in \mathbb{K} \quad (a\cdot b)\cdot c \;=\; a\cdot(b\cdot c) =: a\cdot b\cdot c =: abc$

(M2) $\forall\, a,b \in \mathbb{K} \quad a\cdot b \;=\; b\cdot a$ *„Kommutativität der Multiplikation“*

(M3) $\exists\, 1 \in \mathbb{K}\setminus\{0\} \quad \forall a \in \mathbb{K} \quad a\cdot 1 = a$ *„Eins“* oder *„Einselement“*

(M4) $\forall\, a \in \mathbb{K}\setminus\{0\} \quad \exists\, a^{-1} \in \mathbb{K} \quad a\cdot a^{-1} \;=\; 1$
„Inverses Element zu a bezüglich $\cdot$“, gelesen als *„a hoch minus 1“*

Für die Verbindung von Addition und Multiplikation gelte:

(D) $\forall a,b,c \in \mathbb{K} \quad a\cdot(b+c) \;=\; (a\cdot b) \;+\; (a\cdot c)$ *(„Distributivität“)*

[a] Oft notieren wir lax auch nur $\mathbb{K}\ldots$.

Zum Beispiel bei dem Distributivgesetz ist man von der Schule her gewohnt, die Klammern rechts wegzulassen, also kürzer $a\cdot b + a\cdot c$ oder nur $ab+ac$ statt $(a\cdot b)+(a\cdot c)$ zu notieren, weil man vereinbart: *Punktrechnung geht vor Strichrechnung.* Hier soll also die Multiplikation stärker ‚binden‘ als die Addition. Man erspart sich viele lästige Klammern, wenn man vereinbart, was zuerst ausgerechnet wird, falls keine Klammern gesetzt sind.

Wir beschränken uns zunächst einmal auf $(\mathbb{K}, +)$ mit den Axiomen **(A1) bis (A4)** und ziehen allein aus diesen vier Gesetzen Folgerungen. Mathematiker sagen dafür: *„$(\mathbb{K}, +)$ ist eine abelsche (*oder *kommutative) Gruppe“*.

Wir benutzen natürlich die vertrauten Sprechweisen: Für $a+b$ heißen a und b *Summanden* und $a+b$ *Summe*.

Bemerkung 2.1.2.
Das Nullelement ist — durch (A3) — eindeutig bestimmt.

Es gibt also nur *eine* Null, wir können daher von *dem* Nullelement sprechen.
Der *Beweis* ist ganz einfach: Ist auch $0' \in \mathbb{K}$ ein Nullelement, so gilt also insbesondere $0 + 0' = 0\,(*)$ und so $0 \underset{(*)}{=} 0 + 0' \underset{(A2)}{=} 0' + 0 \underset{(A3)}{=} 0'$. □

Bemerkung 2.1.3.
Zu gegebenem $a \in \mathbb{K}$ ist das Inverse bezüglich $+$ — durch (A4) — eindeutig bestimmt.

Wir können daher von *dem* Inversen (bezüglich $+$) zu a sprechen.

Auch hierzu ist der *Beweis* nicht schwierig: Ist auch a' ein Inverses bezüglich $+$ zu a, so gilt also $a + a' = 0\,(\diamondsuit)$. So hat man $-a \underset{(A3)}{=} (-a) + 0 \underset{(\diamondsuit)}{=} (-a) + (a + a') \underset{(A1)}{=} ((-a) + a) + a' \underset{(A2)}{=} (a + (-a)) + a' \underset{(A4)}{=} 0 + a' \underset{(A2)}{=} a' + 0 \underset{(A3)}{=} a'$. □

Bemerkung 2.1.4.
Zu gegebenen $a, b \in \mathbb{K}$ existiert eindeutig ein $x \in \mathbb{K}$, das die Gleichung $a + x = b$ erfüllt.

Dieses x notieren wir als $b - a$ und lesen *„b minus a“* .

Die Abbildung, die jedem Paar $(a, b) \in \mathbb{K} \times \mathbb{K}$ den Wert $b - a$ zuordnet, bezeichnen wir als *Subtraktion*. Den einzelnen Wert $b - a$ spricht man auch als *Differenz* an.

Beweis: Man rechnet sofort nach, daß $x := (-a) + b$ *eine* Lösung der angegebenen Gleichung ist, und hat so die *Existenz:* $a + x = a + \big((-a) + b\big) \underset{(A1)}{=} \big(a + (-a)\big) + b \underset{(A4)}{=} 0 + b \underset{(A2)}{=} b + 0 \underset{(A3)}{=} b$. Für die *Eindeutigkeit* geht man von einer Lösung x, also $a + x = b$, aus und rechnet: $x \underset{(A3)}{=} x + 0 \underset{(A2)}{=} 0 + x \underset{(A2),(A4)}{=} \big((-a) + a\big) + x \underset{(A1)}{=} (-a) + (a + x) = (-a) + b$ □

Wir haben gleichzeitig mitbewiesen:

Bemerkung 2.1.5.
$\forall\, a, b \in \mathbb{K} \quad b - a \;=\; (-a) + b \;=: -a + b \;=\; b + (-a) \quad (\diamondsuit)$,
speziell $\quad 0 - a = -a + 0 = -a$

Die Kommutativität und Assoziativität für die Addition vermerken wir bei den folgenden Beweisen nicht mehr gesondert, da der Gebrauch inzwischen vertraut sein dürfte.

Bemerkung 2.1.6.
$\forall\, a \in \mathbb{K} \quad -(-a) = a$
$\forall\, a, b \in \mathbb{K} \quad -(a + b) = -b + (-a) \underset{(\diamondsuit)}{=} -b - a$

Beweis: Einerseits hat man $(-a)+a = a+(-a) \underset{(A4)}{=} 0$, andererseits $(-a)+\big(-(-a)\big) \underset{(A4)}{=} 0$, also nach der vorangehenden Bemerkung die erste Behauptung.
$-(a+b)$ ist die — nach obigem eindeutig bestimmte — Lösung der Gleichung $a+b+x=0$. So genügt zu zeigen: $a+b+\big(-b+(-a)\big)=0$:
$\ell.S. = a+\big[b+\big(-b+(-a)\big)\big] = a+\big[\big(b+(-b)\big)+(-a)\big]$
$\underset{(A4)}{=} a+\big[0+(-a)\big] \underset{(A3)}{=} a+(-a) \underset{(A4)}{=} 0$ □

Wir betrachten jetzt $(\mathbb{K},+,\cdot)$ mit den Gesetzen **(A1) bis (A4), (M1) bis (M4) und (D)** und ziehen daraus weitere Folgerungen.
Wir benutzen auch hier die vertrauten Sprechweisen: Für $a\cdot b$ heißen a und b *Faktoren* und $a\cdot b$ *Produkt*.

Bemerkung 2.1.7.
Das Einselement ist — durch (M3) — eindeutig bestimmt.

Es gibt also nur *eine* Eins, wir können daher von *dem* Einselement sprechen.

Bemerkung 2.1.8.
Zu gegebenem $a\in\mathbb{K}\setminus\{0\}$ ist das Inverse bezüglich $\cdot$ — durch (M4) — eindeutig bestimmt.

Wir können daher von *dem* Inversen (bezüglich $\cdot$) zu a sprechen.

Die Beweise dieser beiden Bemerkungen entsprechen völlig denen zu 2.1.2 und 2.1.3; ich führe sie daher nicht noch einmal gesondert aus.

Bemerkung 2.1.9.
Zu gegebenen $a\in\mathbb{K}\setminus\{0\}$ und $b\in\mathbb{K}$ existiert eindeutig ein $x\in\mathbb{K}$, das die Gleichung $a\cdot x=b$ erfüllt.

Dieses x bezeichnen wir als $\frac{b}{a}$ oder $b:a$ und lesen *„b dividiert durch a“.*

Auch hier enspricht der Beweis dem zu 2.1.4 Gezeigten und liefert zusätzlich:

Bemerkung 2.1.10.
Für alle $a\in\mathbb{K}\setminus\{0\}$ und $b\in\mathbb{K}$ gilt $\frac{b}{a} = a^{-1}\cdot b = b\cdot a^{-1}$,
speziell $\frac{1}{a} = a^{-1}$.

Wir haben — wie allgemein üblich — keine Klammer um a^{-1} gesetzt.
Die Abbildung, die jedem Paar $(a,b)\in\mathbb{K}\times\mathbb{K}\setminus\{0\}$ den Wert $ab^{-1}=\frac{a}{b}$ zuordnet, bezeichnen wir als *Division*. Den einzelnen Wert $\frac{a}{b}$ spricht man als *Quotient* oder *Bruch* an, a als *Zähler* und b als *Nenner*.

Bemerkung 2.1.11. $\forall a\in\mathbb{K}\qquad a\cdot 0 = 0\cdot a = 0$

Beweis: Es genügt — wegen (M2) — die Gleichung $a \cdot 0 = 0$ zu zeigen: Man hat $a \cdot 0 \underset{(A3)}{=} a \cdot (0+0) \underset{(D)}{=} a \cdot 0 + a \cdot 0$, also mit (A3) und Bemerkung (2.1.4) die Behauptung. □

Bemerkung 2.1.12.
Für $a \in \mathbb{K} \setminus \{0\}$ *gilt* $a^{-1} \neq 0$ *und* $(a^{-1})^{-1} = a$.

Beweis: Nach (M4) und (M3) ist $aa^{-1} = 1 \neq 0$, also nach (2.1.11) $a^{-1} \neq 0$. $a^{-1}a \underset{(M2)}{=} aa^{-1} \underset{(M4)}{=} 1$ und $a^{-1}(a^{-1})^{-1} \underset{(M4)}{=} 1$ zeigen mit der Eindeutigkeit gemäß (2.1.9) $(a^{-1})^{-1} = a$. □

Bemerkung 2.1.13. *Für* $a, b \in \mathbb{K}$ *gilt:* $ab = 0 \iff a = 0 \vee b = 0$.

Das Gleiche noch einmal in Worten: *Ein Produkt von Elementen aus* $\mathbb{K}$ *ist genau dann Null, wenn mindestens ein Faktor Null ist.*

Beweis: Die Richtung von rechts nach links ist durch (2.1.11) gegeben. Für die andere Implikation ist man im Fall $a = 0$ fertig. Für $a \neq 0$ erhält man: $0 \underset{(2.1.11)}{=} a^{-1} \cdot 0 = a^{-1}(ab) \underset{(M1)}{=} (a^{-1}a)b \underset{(M2),(M4)}{=} 1b, \underset{(M2),(M3)}{=} b$ □

Bemerkung 2.1.14.
Für $a, b \in \mathbb{K} \setminus \{0\}$: $(ab)^{-1} = b^{-1}a^{-1} = a^{-1}b^{-1}$

Auch hier kann der *Beweis* wieder weggelassen werden, weil er genau nach dem Muster des Beweises zu (2.1.6) verläuft. Dabei ist nur zu berücksichtigen, daß nach (2.1.13) hier das Produkt ab auch von 0 verschieden ist.

Bemerkung 2.1.15.
Für $a, b \in \mathbb{K}$: $(-a)b = -(ab) = a(-b)$, $(-1)b = -b$, $(-a)(-b) = ab$

Beweis: $-(ab)$ ist die eindeutige Lösung der Gleichung $ab + x = 0$. Also muß für die erste Gleichung nur $ab + (-a)b = 0$ nachgewiesen werden: $ab + (-a)b \underset{(M2)}{=} ba + b(-a) \underset{(D)}{=} b\big(a + (-a)\big) \underset{(A4)}{=} b0 \underset{(2.1.11)}{=} 0.$
Der Beweis der zweiten Gleichung folgt daraus durch Vertauschung von a und b unter Berücksichtigung von (M2). Die dritte Gleichung ergibt sich nun mit $a := 1$.
Nach dem gerade Bewiesenen ist $(-a)(-b) = -\big(-(ab)\big) \underset{(2.1.6)}{=} ab$. □

Bemerkung 2.1.16. *Für* $a, b, c \in \mathbb{K}$: $c\,(b-a) = c\,b - c\,a$

Wir erinnern noch einmal an die Verabredung: *‚Punktrechnung‘ (Multiplikation und Division) geht vor ‚Strichrechnung‘ (Addition und Subtraktion).* Die rechte Seite ist also als $(cb) - (ca)$ zu verstehen.

Beweis: $\ell.S. \underset{(2.1.5)}{=} c\big(b+(-a)\big) \underset{(D)}{=} cb+c(-a) \underset{(2.1.15)}{=} cb+\big(-(ca)\big) \underset{(2.1.5)}{=} cb - ca$ □

Ausdrücklich betonen möchte ich, daß das Inverse Element zu 0 bezüglich $\cdot$, dessen Existenz oben gerade *nicht* gefordert wurde, gar nicht existieren kann; denn sonst hätte man $0 \cdot 0^{-1} = 1 \underset{(M3)}{\neq} 0$ im Widerspruch zu (2.1.11). In der Schule lernt man das meist in der Form *„Durch* 0 *darf man nicht dividieren!“* und hat dann Schwierigkeiten, dieses ‚Verbot‘ einzusehen, weil es nicht begründet wird.

2.2. „Bruchrechnen“.

Allein aus der Definition des Quotienten $\frac{a}{b}$ für $(a,b) \in \mathbb{K} \times \mathbb{K} \setminus \{0\}$ lassen sich einfach alle Regeln über das *„Bruchrechnen“* herleiten. Wir notieren die wichtigsten:

Für $a, e \in \mathbb{K}$ *und* $b, c, d \in \mathbb{K} \setminus \{0\}$ *gelten:*

$(\boldsymbol{\alpha})$ $\frac{1}{-b} = -\frac{1}{b}$ $(\boldsymbol{\beta})$ $\frac{b}{b} = 1$ $(\boldsymbol{\gamma})$ $\frac{a}{1} = a$

$(\boldsymbol{\delta})$ $\frac{a}{b} = \frac{e}{d} \iff ad = be$ $(\boldsymbol{\varepsilon})$ $\frac{a \cdot c}{b \cdot c} = \frac{a}{b}$ $(\boldsymbol{\zeta})$ $\frac{a}{b} \cdot \frac{e}{d} = \frac{a \cdot e}{b \cdot d}$

$(\boldsymbol{\eta})$ $\frac{a}{b} \pm \frac{e}{d} = \frac{ad \pm be}{bd}$ $(\boldsymbol{\vartheta})$ $\frac{\frac{a}{b}}{\frac{c}{d}} = \frac{ad}{bc}$

(δ) beschreibt die *Gleichheit von Brüchen: Brüche sind genau dann gleich, wenn die ‚Überkreuzprodukte‘ gleich sind.*

(ε) beschreibt von links nach rechts das *Kürzen*, von rechts nach links das *Erweitern* von Brüchen.

(ζ) und (η) zeigen, wie Brüche multipliziert, addiert und subtrahiert werden. Dabei sind Formeln mit $\pm$ natürlich immer so zu lesen, daß auf beiden Seiten gleichzeitig $+$ oder gleichzeitig $-$ zu nehmen ist. (ϑ) schließlich belegt — in Verbindung mit (ζ) — : *Ein Bruch (hier* $\frac{a}{b}$*) wird durch einen Bruch (hier* $\frac{c}{d}$*) dividiert, indem man mit dem ‚Kehrwert‘ (hier also* $\frac{d}{c}$*) multipliziert.*

Wir beweisen exemplarisch und ausführlich (δ), (ζ) und (η) und überlassen die Beweise der restlichen Aussagen, die alle nach dem gleichen Muster verlaufen, als *Übungsaufgabe.*

Beweis: $\underline{(\delta):}$ Nach (2.1.9) gilt $\frac{a}{b} = \frac{e}{d}$ genau dann, wenn $b \cdot \frac{e}{d} = a \quad \odot$. Hat man dies, so liefert die Multiplikation dieser Gleichung mit d — unter Berücksichtigung von (M1) und (2.1.9) — $be = ad$. Ausgehend von $be = ad$ ergibt sich — unter Berücksichtigung von (M1) und (2.1.9) — $\odot$ nach Multiplikation mit d^{-1}. $\underline{(\zeta):}$ $\frac{ae}{bd}$ ist die eindeutig bestimmte Lösung der Gleichung $(bd) \cdot x = ae$. (Nach (2.1.13) ist dabei $bd \neq 0$.) Andererseits gilt $bd \cdot \left(\frac{a}{b} \cdot \frac{e}{d}\right) \underset{(M1),(M2)}{=}$

$\left(b\cdot\frac{a}{b}\right)\left(d\cdot\frac{e}{d}\right) = ae$. (η): Die $r.S.$ ist die eindeutige Lösung der Gleichung $(bd)\,x = ad \pm be$. Die $\ell.S.$ erfüllt auch diese Gleichung; denn $(bd)\left[\frac{a}{b} \pm \frac{e}{d}\right] \underset{(M2),(D),(2.1.16)}{=} (db)\frac{a}{b} \pm (bd)\frac{e}{d} \underset{(M1),(M2)}{=} ad \pm be$ □

Wir haben in diesem und dem vorangehenden Teilabschnitt allein ausgehend von den Axiomen (A1) bis (A4), (M1) bis (M4) und (D) viele — von der Schule her — bereits vertraute Gesetze hergeleitet. Der entscheidende Vorteil dieses Vorgehens ist, daß all diese Folgerungen immer schon dann gelten, wenn nur die oben aufgeführten Axiome erfüllt sind, also insbesondere bei den rationalen Zahlen $\mathbb{Q}$, den reellen Zahlen $\mathbb{R}$ und den komplexen Zahlen $\mathbb{C}$. Deshalb kann man nun u. a. in all diesen drei Bereichen — $\mathbb{R}$, $\mathbb{Q}$ und besonders auch $\mathbb{C}$ — wie ‚gewohnt' rechnen.

Beispiele 2.2.1.

a) Die Mengen $\mathbb{R}, \mathbb{Q}, \mathbb{C}$ mit der üblichen Addition und Multiplikation bilden jeweils einen Körper.

b) $\mathbb{K} := \left\{a + b\sqrt{17} \colon a, b \in \mathbb{Q}\right\} =: \mathbb{Q}\left[\sqrt{17}\right] \subset \mathbb{R}$ mit der eingeschränkten Addition und Multiplikation von reellen Zahlen ist ein Körper.

c) Wir benötigen die Menge der *ganzen Zahlen*

$$\mathbb{Z} := \mathbb{N}_0 \cup \{-m : m \in \mathbb{N}\}.$$

Es sei p eine natürliche Zahl. Für $a \in \mathbb{Z}$ betrachten wir

$$\bar{a} := \{a + pm \colon m \in \mathbb{Z}\} =: a + p\mathbb{Z},$$

also $\bar{a} = \{\ldots, a-3p, a-2p, a-p, a, a+p, a+2p, a+3p, \ldots\}$. Wir setzen

$$\mathbb{Z}_p := \left\{\bar{0}, \bar{1}, \bar{2}, \ldots, \overline{p-1}\right\}$$

und definieren Addition und Multiplikation durch

$$\bar{a} + \bar{b} := \overline{a+b} \quad \text{und} \quad \bar{a}\cdot\bar{b} := \overline{a\cdot b}$$

für $a, b \in \mathbb{Z}$). (Hier ist noch die *Wohldefiniertheit* nachzuweisen, d. h. wir müssen zeigen, daß $\overline{a+kp} + \overline{b+\ell p} = \bar{a} + \bar{b}$ und $\overline{a+kp}\cdot\overline{b+\ell p} = \bar{a}\cdot\bar{b}$ für alle $k, \ell \in \mathbb{Z}$ gelten.) Wir addieren und multiplizieren also jeweils *vertreterweise.* $\bar{a}$ ist die Menge aller Zahlen, die bei Division durch p denselben Rest lassen wie a. Deshalb nennt man diese Mengen meist *Restklassen* (modulo p). Ein Element einer Restklasse bezeichnet man auch als *Repräsentant* der Restklasse. Häufig verwendet man die *Standardrepräsentanten* $0, \ldots, p-1$. Ist p eine *Primzahl,* dann ist $\mathbb{Z}_p$ ein *Körper.*

2.3. **Vektorräume.**

Die Lösungsmenge eines homogenen linearen Gleichungssystems ist ein erstes Beispiel für einen Vektorraum.

Definition 2.3.1 (Vektorraum)**.**

Es sei $\mathbb{K}$ ein Körper. Ein Vektorraum über $\mathbb{K}$ oder ein $\mathbb{K}$-Vektorraum ist ein Tripel $(\mathfrak{V}, +, .)$, bestehend aus einer nicht-leeren Menge $\mathfrak{V}$ und zwei Abbildungen, *„Addition“* und *„Multiplikation mit Skalaren“*,

$$\begin{array}{ccccc} +: & \mathfrak{V}\times\mathfrak{V} & \longrightarrow & \mathfrak{V} & \\ & \cup & & \cup & \text{und} \\ & (a,b) & \longmapsto & a+b & \end{array} \qquad \begin{array}{ccccc} .: & \mathbb{K}\times\mathfrak{V} & \longrightarrow & \mathfrak{V} & \\ & \cup & & \cup & \\ & (\alpha,a) & \longmapsto & \alpha.a =: \alpha a & \end{array}$$

derart, daß die folgenden Gesetze gelten:

(A1) $\forall\, a,b,c \in \mathfrak{V} \quad (a+b)+c \;=\; a+(b+c) =: a+b+c$

„Assoziativität der Addition“

(A2) $\forall\, a,b \in \mathfrak{V} \quad a+b \;=\; b+a$ *„Kommutativität der Addition“*

(A3) $\exists\, 0 \in \mathfrak{V} \quad \forall a \in \mathfrak{V} \quad a+0=a$ *„Null“* oder *„Nullelement“*

(A4) $\forall\, a \in \mathfrak{V} \quad \exists\, -a \in \mathfrak{V} \quad a+(-a) \;=\; 0$

„Inverses Element zu a bezüglich $+$“, wieder gelesen als *„minus a“*

Die Multiplikation mit Skalaren erfüllt:

(SM1) $\lambda.(a+b) = (\lambda.a)+(\lambda.b)$ für alle $\lambda \in \mathbb{K}$, $a,b \in \mathfrak{V}$

(SM2) $(\lambda+\mu).a = (\lambda.a)+(\mu.a)$ für alle $\lambda,\mu \in \mathbb{K}$, $a \in \mathfrak{V}$

(Distributivgesetze)

(SM3) $(\lambda\cdot\mu).a = \lambda.(\mu.a)$ für alle $\lambda,\mu \in \mathbb{K}$, $a \in \mathfrak{V}$

(SM4) $1.a = a$ für alle $a \in \mathfrak{V}$.

Beispiele 2.3.2.

a) $\mathbb{R}^n$ mit der üblichen Addition und Multiplikation mit Skalaren, d. h.

$$\begin{aligned}(\xi^1,\dots,\xi^n)+(\eta^1,\dots,\eta^n) &:= (\xi^1+\eta^1,\dots\xi^n+\eta^n),\\ \lambda.(\xi^1,\dots,\xi^n) &:= (\lambda\xi^1,\dots,\lambda\xi^n),\end{aligned}$$

ist ein $\mathbb{R}$-Vektorraum.

b) $\mathbb{C}^n$, der Raum der komplexen n-Tupel mit entsprechenden Verknüpfungen ist ein $\mathbb{C}$-Vektorraum.

c) $\mathbb{Q}^n$ ist ein $\mathbb{Q}$-Vektorraum.

d) Ist $\mathbb{K}$ ein Körper, Dann ist $\mathbb{K}^n$ ein $\mathbb{K}$-Vektorraum.

e) Die Menge der Folgen reeller Zahlen

$$\mathbb{R}^{\mathbb{N}_0} = \left\{\left(\xi^0,\xi^1,\xi^2,\dots\right) : \xi^i \in \mathbb{R} \text{ für alle } i \in \mathbb{N}_0\right\}$$

mit den Verknüpfungen

$$\begin{aligned}\left(\xi^0,\xi^1,\xi^2,\ldots\right)+\left(\eta^0,\eta^1,\eta^2,\ldots\right) &:= \left(\xi^0+\eta^0,\xi^1+\eta^1,\xi^2+\eta^2,\ldots\right),\\ \lambda\cdot\left(\xi^0,\xi^1,\xi^2,\ldots\right) &:= \left(\lambda\xi^0,\lambda\xi^1,\lambda\xi^2,\ldots\right)\end{aligned}$$

bildet einen $\mathbb{R}$-Vektorraum. Die Teilmengen der konvergenten Folgen oder der Nullfolgen bilden ebenfalls einen Vektorraum.

f) Die Menge der Lösungen eines linearen *homogenen* Gleichungssystems wie in (1.1) bildet einen Vektorraum.

g) Es seien A eine nicht-leere Menge, $\mathbb{K}$ ein Körper und $\mathfrak{V}$ eine $\mathbb{K}$-Vektorraum. Die Menge der Funktionen $f\colon A\to\mathfrak{V}$ bildet einen $\mathbb{K}$-Vektorraum, wenn man die Verknüpfungen *elementweise* erklärt:

$$\begin{aligned}(f+g)(x) &:= f(x)+g(x) \quad \text{für } x\in A,\\ (\lambda.f)(x) &:= \lambda.(f(x)) \quad \text{für } x\in A.\end{aligned}$$

h) Die Menge der stetigen Funktionen $f\colon[0,1]\to\mathbb{R}$ bildet einen $\mathbb{R}$-Vektorraum $C^0([0,1]) = C([0,1])$ mit wie für beliebige Funktionen definierter Addition und Multiplikation mit Skalaren.

Wir hatten allein aus den vier Gesetzen **(A1) bis (A4)** Folgerungen gezogen. Diese gelten also auch — mit den dortigen Notierungsweisen und Vereinbarungen — in jedem Vektorraum:

Bemerkung 2.3.3.

a) Das Nullelement ist — durch (A3) — eindeutig bestimmt.

b) Zu gegebenem $a\in\mathfrak{V}$ ist das Inverse bezüglich $+$ — durch (A4) — eindeutig bestimmt.

c) Zu gegebenen $a,b\in\mathfrak{V}$ existiert eindeutig ein $x\in\mathfrak{V}$, das die Gleichung $a+x=b$ erfüllt.

d) $\forall\, a,b\in\mathfrak{V}\quad b-a = (-a)+b =: -a+b = b+(-a)$,
speziell $0-a=-a+0=-a$

e) $\forall\, a\in\mathfrak{V}\quad -(-a)=a$
$\forall\, a,b\in\mathfrak{V}\quad -(a+b)=-b+(-a) = -b-a$

Bemerkung 2.3.4.

Es seien $\mathbb{K}$ ein Körper und $\mathfrak{V}$ ein $\mathbb{K}$-Vektorraum. Dann gelten für alle $a,b\in\mathfrak{V}$ und $\lambda,\mu\in\mathbb{K}$:

a) $0.a=0$

b) $\lambda.0=0$

c) Aus $\lambda.a=0$ folgt $a=0$ oder $\lambda=0$.

d) $(-\lambda).a=\lambda.(-a)=-(\lambda.a)$

e) $\lambda.(a-b) = \lambda.a-\lambda.b$ *und* $(\lambda-\mu).a = \lambda.a-\mu.a$

Beweis:

a) Nach (SM2) gilt $0.a = (0+0).a = 0.a + 0.a$. Andererseits gilt $0.a = 0 + 0.a$. Nach 2.3.3.c folgt die Behauptung.

b) $\lambda.0 = \lambda.(0+0) \underset{(SM1)}{=} \lambda.0 + \lambda.0$. Wieder liefert 2.3.3.c (mit (A3)) die Behauptung.

c) Es sei $\lambda.a = 0$ mit $\lambda \neq 0$. Zu zeigen ist also, daß $a = 0$ gilt: $0 \underset{b)}{=} \lambda^{-1}.0 = \lambda^{-1}.(\lambda.a) \underset{(SM3)}{=} (\lambda^{-1} \cdot \lambda).a = 1.a \underset{(SM4)}{=} a$

d) ∘ ∘ ∘

e) ∘ ∘ ∘

□

In einer abelschen Gruppe $(\mathfrak{G}, +)$ gilt das Assoziativgesetz (A1) entsprechend auch für mehr als drei Summanden, ebenso gilt das Kommutativgesetz (A2) für mehr als zwei Summanden. Der Beweis ist nicht kompliziert, aber durchaus umständlich aufzuschreiben. Man benutzt Induktion nach der Anzahl der Summanden.
Die *rekursive Definition*, auch *Definition* oder *Konstruktion durch vollständige Induktion* genannt, ist heute den meisten schon durch Programmier-Erfahrung bestens vertraut: Man legt fest, wie gestartet wird *(Anfangswert)* und zusätzlich, wie es weitergehen soll, wenn man schon bis zu einer bestimmten Stelle gelangt ist

Zum Beispiel ist der Ausdruck $1+2+\cdots+n$, besonders die Pünktchen darin, mathematisch keineswegs exakt, und vielleicht ist der eine oder andere Leser schon darüber gestolpert — denn, wie ist das beispielsweise für $n = 1$ zu lesen?
Dies läßt sich durch *„rekursive Definition“* präzisieren. Wir wollen dieses Definitionsprinzip aber nicht besonders begründen, sondern ‚naiv‘ rangehen, da es unmittelbar einsichtig zu sein scheint. Wer es an dieser Stelle doch genauer wissen will, kann zum Beispiel in Martin Barner und Friedrich Flohr [1] nachsehen.
Es seien $\mathbb{K}$ ein Körper, $\mathfrak{V}$ ein $\mathbb{K}$-Vektorraum und

$$\begin{array}{rccc} x\colon & \mathbb{N}_0 & \longrightarrow & \mathfrak{V} \\ & \cup & & \cup \\ & j & \longmapsto & x_j\,. \end{array}$$

Wir nennen allgemein solche auf $\mathbb{N}_0$ definierten Abbildungen *„Folgen“*, x_j das j-te *„Glied“* mit *„Index“* j (für $j \in \mathbb{N}_0$) und notieren sie oft auch in der Form $(x_0, x_1, x_2, \ldots)$.

Für ein (festes) $k \in \mathbb{N}_0$ wollen wir den Ausdruck $x_k + \cdots + x_n$, also die Summe der Folgenglieder mit Indizes k bis n, rekursiv definieren

und benutzen dafür das neue Zeichen $\sum_{\nu=k}^{n} x_\nu$, definiert durch:

$$\sum_{\nu=k}^{k} x_\nu := x_k\,, \quad \sum_{\nu=k}^{n+1} x_\nu := \left(\sum_{\nu=k}^{n} x_\nu\right) + x_{n+1} \qquad \big(k \le n \in \mathbb{N}_0\big)$$

Wir lesen dies als *„Summe der x_ν für $\nu = k$ bis n“* oder ähnlich. Für manche Zwecke ist noch nützlich,

$$\sum_{\nu=k}^{n} x_\nu := 0 \quad \text{für } \mathbb{N} \ni n < k \qquad \textit{(„leere Summe“)}^3$$

zu vereinbaren.

Der *„Summationsindex“* ν in $\sum_{\nu=k}^{n} x_\nu$ hat keine besondere Bedeutung. Er dient als Platzhalter und kann insbesondere durch irgendein anderes Zeichen (nur nicht gerade k und n) ersetzt werden, zum Beispiel:

$$\sum_{\nu=k}^{n} x_\nu = \sum_{j=k}^{n} x_j = \sum_{p=k}^{n} x_p = \sum_{\square=k}^{n} x_\square = \sum_{\heartsuit=k}^{n} x_\heartsuit = \sum_{Z=k}^{n} x_Z$$

Ich selbst habe die Angewohnheit, jeweils den ‚passenden‘ griechischen Buchstaben zu wählen, also hier ν zu n, an anderen Stellen beispielsweise κ zu k, λ zu ℓ oder μ zu m, aber das ist nicht mehr als eine persönliche Vorliebe und Systematik.
Bei solchen Summen kann man — aufgrund des Assoziativ- und Kommutativ-Gesetzes — *beliebig vertauschen und Klammern setzen*. Den Beweis dieser Aussage, der keineswegs trivial ist, lassen wir wieder weg. Dafür haben wir unsere Mathematiker!
In einem Vektorraum gelten dann die üblichen Rechenregeln:

Bemerkung 2.3.5.

a) $\lambda \cdot \sum_{\nu=1}^{n} a_\nu = \sum_{\nu=1}^{n} \lambda \cdot a_\nu$ *für alle* $\lambda \in \mathbb{K}$ *und* $a_1, \ldots, a_n \in \mathfrak{V}$,

b) $\left(\sum_{\nu=1}^{n} \lambda^\nu\right) \cdot a = \sum_{\nu=1}^{n} \lambda^\nu \cdot a$ *für alle* $\lambda^1, \ldots, \lambda^n \in \mathbb{K}$ *und* $a \in \mathfrak{V}$,

c) $\sum_{\nu=1}^{n} \lambda^\nu a_\nu + \sum_{\nu=1}^{n} \mu^\nu a_\nu = \sum_{\nu=1}^{n} (\lambda^\nu + \mu^\nu)\, a_\nu$ *für alle* $\lambda^1, \ldots, \lambda^n, \mu^1, \ldots, \mu^n$ *aus* $\mathbb{K}$ *und* $a_1, \ldots, a_n \in \mathfrak{V}$.

Beweis: ∘ ∘ ∘ □

2.4. **Linearkombinationen.**

Es sei wieder $\mathbb{K}$ ein Körper.

[3] Oft notieren wir auch $\sum_{\emptyset} x_\nu$ und ähnlich.

Definition 2.4.1. Es seien $\mathfrak{V}$ ein $\mathbb{K}$-Vektorraum und S eine nicht-leere Teilmenge von $\mathfrak{V}$. Gilt mit einem $n \in \mathbb{N}_0$

$$a = \sum_{\nu=1}^{n} \lambda^\nu a_\nu \quad \text{für geeignete } \lambda^1, \dots, \lambda^n \in \mathbb{K} \quad \text{und } a_1, \dots, a_n \in S,$$

so sagen wir, daß a als *Linearkombination* von Vektoren aus S dargestellt ist. (Beachte, daß die Summen in solchen Linearkombinationen stets *endlich* sind.)
Die Menge aller Vektoren, die sich als Linearkombination von Vektoren aus S darstellen lässt, heißt *lineare Hülle* von S; wir schreiben $\langle S \rangle$. Ist S eine endliche Menge, so schreiben wir auch $\langle a_1, \dots, a_m \rangle$ statt $\langle \{a_1, \dots, a_m\} \rangle$. Wir sagen dann auch, daß $\langle a_1, \dots, a_m \rangle$ aus den Linearkombinationen von $a_1, \dots, a_m$ besteht.

Beispiele 2.4.2.

(*i*) Es seien $\mathfrak{V}$ ein $\mathbb{K}$-Vektorraum und $S = \{a_1, \dots, a_p\} \subset \mathfrak{V}$. Dann ist

$$\langle S \rangle = \left\{ \lambda^1 a_1 + \dots + \lambda^p a_p \colon \lambda^1, \dots, \lambda^p \in \mathbb{K} \right\}.$$

(*ii*) Es seien $\mathfrak{V} = \mathbb{R}^3$ und $S = \{(1,1,0), (1,0,1)\}$. Es folgt

$$\begin{aligned} \langle S \rangle &= \{\alpha(1,1,0) + \beta(1,0,1) \colon \alpha, \beta \in \mathbb{R}\} \\ &= \{(\alpha+\beta, \alpha, \beta) \colon \alpha, \beta \in \mathbb{R}\} \\ &= \{(x,y,z) \in \mathbb{R}^3 \colon x = y + z\}. \end{aligned}$$

Ist $T = \{(1,1,0), (2,1,1)\}$, so gilt $\langle S \rangle = \langle T \rangle$.

2.5. Unterräume.

Es sei wieder $\mathbb{K}$ ein Körper.

Definition 2.5.1.
Es sei $\mathfrak{V}$ ein Vektorraum über $\mathbb{K}$. Eine nicht-leere Teilmenge $\mathfrak{U}$ von $\mathfrak{V}$ heißt genau dann *Unterraum* (von $\mathfrak{V}$), wenn $\mathfrak{U}$ mit der induzierten (eingeschränkten) Addition und Multiplikation mit Skalaren ein $\mathbb{K}$-Vektorraum ist.

Bemerkung 2.5.2.
Eine Teilmenge $\mathfrak{U}$ eines $\mathbb{K}$-Vektorraums $\mathfrak{V}$ ist genau dann ein Unterraum, *wenn $0 \in \mathfrak{U}$ und für alle $a, b \in \mathfrak{U}$ und $\lambda \in \mathbb{K}$ auch $\lambda.a + b \in \mathfrak{U}$ gilt.*

Beweis: ∘ ∘ ∘ □

Beispiele 2.5.3.

a) In jedem Vektorraum $\mathfrak{V}$ sind $\mathfrak{V}$ selbst und $\{0\}$ ‚triviale' Unterräume von $\mathfrak{V}$.
b) Es sei $a \in [0,1]$. Dann ist $\{f \in C([0,1]) \colon f(a) = 0\}$ ein Unterraum von $C([0,1])$.

c) Die Menge der Lösungen eines linearen *homogenen* Gleichungssystems in n Variablen ist ein Unterraum von $\mathbb{K}^n$.

Hier und im Folgenden betrachten wir Vektorräume über einem beliebigen aber festen Körper $\mathbb{K}$, falls dies nicht ausdrücklich anders angegeben ist. Dabei sind sicher für viele Dinge die Körper $\mathbb{R}$ und $\mathbb{C}$ die wichtigsten!

Satz 2.5.4.
Es seien $\mathfrak{V}$ ein Vektorraum und $S \subset \mathfrak{V}$. Dann ist $\langle S\rangle$ der kleinste Unterraum *von $\mathfrak{V}$, der S enthält.*

Beweis: Wegen $\sum_{\emptyset} = 0$ ist $\langle S\rangle \neq \emptyset$. $\langle S\rangle$ ist ein Unterraum, denn für $a, b \in \langle S\rangle$ mit Œ (‚notfalls' mit Nullen auffüllen...) $a = \sum_{i\in I} a^i s_i$ und $b = \sum_{i\in I} b^i s_i$ für eine endliche Menge I, ... und $\lambda \in \mathbb{K}$ gilt auch

$$\lambda a + b \;=\; \sum_{i\in I} \left(\lambda a^i + b^i\right) s_i \;\in\; \langle S\rangle.$$

Somit ist $\langle S\rangle$ ein Unterraum von $\mathfrak{V}$, der insbesondere die Vektoren von S enthält.
Ist $\mathfrak{W}$ ein beliebiger Unterraum von $\mathfrak{V}$ mit $S \subset \mathfrak{W}$. Dann enthält $\mathfrak{W}$ auch alle Linearkombinationen von Vektoren aus S, ist also eine Obermenge von $\langle S\rangle$. □

Satz 2.5.5.
Es seien $\mathfrak{V}$ ein Vektorraum, $S \subset \mathfrak{V}$. Für jedes $b \in \langle S\rangle$ gilt

$$\langle S \cup \{b\}\rangle = \langle S\rangle\,.$$

Beweis: Es gilt $S\cup\{b\} \subset \langle S\rangle$. $\langle S\rangle$ ist ein Vektorraum. Somit folgt nach Satz 2.5.4 $\langle S \cup \{b\}\rangle \subset \langle S\rangle$. $\langle S\rangle \subset \langle S \cup \{b\}\rangle$ ist klar. □

Definition 2.5.6.
Wir sagen, daß $\langle S\rangle$ von S erzeugt ist. Im Fall $\langle S\rangle = \mathfrak{V}$ heißt S ein *Erzeugendensystem* von $\mathfrak{V}$. Besitzt ein Vektorraum $\mathfrak{V}$ ein *endliches* Erzeugendensystem, so heißt $\mathfrak{V}$ *endlich erzeugt* bzw. *endlich erzeugbar*.

Beispiele 2.5.7.

a) Es sei $A \in \mathbb{K}^{m\times n}$. Fassen wir die ($n$) Spalten von A jeweils als Vektoren des $\mathbb{K}^m$ auf, so nennen wir den von ihnen erzeugten Unterraum den *Spaltenraum* von A. Ebenso ist der *Zeilenraum* von A der von den (m) Zeilen von A in $\mathbb{K}^n$ erzeugte Unterraum.

b) Es seien $\mathfrak{V} := \mathbb{R}^n$ und $S := \{e_1, \ldots, e_n\}$ mit

$$\begin{aligned} e_1 &:= (1, 0, 0, \ldots, 0),\\ e_2 &:= (0, 1, 0, \ldots, 0),\\ &\vdots\\ e_n &:= (0, 0, 0, \ldots, 1). \end{aligned}$$

(Für später wäre es eigentlich besser, die Vektoren aus dem $\mathbb{R}^n$ schon hier als Spaltenvektoren zu schreiben...) Dann ist $\langle S\rangle = \mathfrak{V}$, also S ein Erzeugendensystem von $\mathfrak{V}$. Somit ist $\mathbb{R}^n$ endlich erzeugbar. Für $\xi = (\xi^1, \dots, \xi^n)$ gilt dabei insbesondere:

$$\xi = \sum_{\nu=1}^{n} \xi^\nu e_\nu$$

Satz 2.5.8.
Es sei $(\mathfrak{W}_i)_{i\in I}$ eine beliebige ‚Familie' von Unterräumen von $\mathfrak{V}$. Dann ist auch $\mathfrak{W} := \bigcap_{i\in I} \mathfrak{W}_i$ ein Unterraum von $\mathfrak{V}$.

Beweis: Der Durchschnitt ist nicht leer, da jeder Unterraum die Null enthält. Für $a, b \in \mathfrak{W}$ gilt $a, b \in \mathfrak{W}_i$ für alle $i \in I$. Für $\lambda \in \mathbb{K}$ gilt dann $\lambda a + b \in \mathfrak{W}_i$ für alle $i \in I$ und somit $\lambda a + b \in \mathfrak{W}$. □

Im Allgemeinen ist die Vereinigung von Unterräumen eines Vektorraumes kein Unterraum mehr. Die folgende Konstruktion macht aus endlich vielen Unterräumen einen Unterraum, der jeden dieser Unterräume enthält. (Dies funktioniert auch für beliebig viele Unterräume; dann erklärt man die Summe als die Menge beliebiger endlicher Summen von Vektoren aus den Unterräumen).

Definition 2.5.9.
Für ein $k \in \mathbb{N}$ seien $\mathfrak{W}_\kappa$ Unterräume eines $\mathbb{K}$-Vektorraumes $\mathfrak{V}$ (für $\kappa = 1, \dots, k$). Definiere deren *Summe* durch

$$\mathfrak{W}_1 + \cdots + \mathfrak{W}_k := \{a_1 + \cdots + a_k \colon a_\kappa \in \mathfrak{W}_\kappa,\ \kappa = 1, \dots k\}.$$

Satz 2.5.10.
Es seien $\mathfrak{W}_1, \dots, \mathfrak{W}_k$ Unterräume eines $\mathbb{K}$-Vektorraums $\mathfrak{V}$. Dann ist $\mathfrak{W} := \mathfrak{W}_1 + \cdots + \mathfrak{W}_k$ der kleinste Unterraum von $\mathfrak{V}$, der jedes $\mathfrak{W}_\kappa$ enthält.

Beweis: Wir zeigen zunächst, daß $\mathfrak{W}$ ein Unterraum von $\mathfrak{V}$ ist:
$0 \in \mathfrak{W}$: ✓ Es seien $a = a_1 + \cdots + a_k$ und $b = b_1 + \cdots + b_k \in \mathfrak{W}$ mit $a_\kappa, b_\kappa \in \mathfrak{W}_\kappa$. Mit $\lambda \in \mathbb{K}$ ist dann

$$\lambda a + b = (\lambda a_1 + b_1) + \cdots + (\lambda a_k + b_k) \in \mathfrak{W}\,.$$

Somit ist $\mathfrak{W}$ ein Unterraum.
Offenbar gilt $(\circ\circ\circ)$ $\mathfrak{W}_\kappa \subset \mathfrak{W}$. Es seien $\mathfrak{U}$ ein weiterer Unterraum von $\mathfrak{V}$ mit $\mathfrak{W}_\kappa \subset \mathfrak{U}$ für alle κ und $s := a_1 + \cdots + a_k \in \mathfrak{W}$ (mit ...) beliebig. Da $a_\kappa \in \mathfrak{W}_\kappa \subset \mathfrak{U}$ ist, folgt auch $s \in \mathfrak{U}$, weil $\mathfrak{U}$ ein Unterraum ist. Somit haben wir: $\mathfrak{W} \subset \mathfrak{U}$. □

3. Struktur von Vektorräumen

3.1. **Lineare Unabhängigkeit.**

Es sei wieder — wie üblich — $\mathbb{K}$ ein Körper.

Definition 3.1.1.
Vektoren $a_1, \ldots, a_n$ — oder das n-Tupel[4] $(a_1, \ldots, a_n)$ oder auch (bei verschiedenen Vektoren) die Menge $\{a_1, \ldots, a_n\}$ — in einem $\mathbb{K}$-Vektorraum heißen genau dann *linear unabhängig*, wenn aus

$$\sum_{\nu=1}^{n} \lambda^\nu a_\nu = 0$$

für $\lambda^\nu \in \mathbb{K}$ stets $\lambda^1 = \cdots = \lambda^n = 0$ folgt (d. h. der Nullvektor erlaubt nur die ‚triviale' Darstellung.) Andernfalls heißt sie *linear abhängig*.
Eine *beliebige* Menge heißt genau dann *linear unabhängig*, wenn *je* endlich viele Vektoren aus ihr linear unabhängig sind. Andernfalls heißt sie *linear abhängig*.

Es ist für manche Dinge vorteilhaft, auch die *leere Menge*, die den Nullraum erzeugt, *linear unabhängig* zu nennen.

Beispiele 3.1.2.

a) Ein einzelner Vektor $\{a\}$ ist genau dann linear unabhängig, wenn $a \neq 0$ gilt.

b) Ist eine Menge linear abhängig, dann ist auch jede Obermenge von ihr linear abhängig.

c) Das Tripel $((2,-1,1),(1,0,0),(0,2,1)) =: (a_1, a_2, a_3)$ ist linear unabhängig; denn für eine Linearkombination der Null, also $0 = \lambda^1 a_1 + \lambda^2 a_2 + \lambda^3 a_3$ mit … , gilt

$$\begin{array}{rcrcrcl} 2\lambda^1 & + & \lambda^2 & & & = & 0 \\ -\lambda^1 & & & + & 2\lambda^3 & = & 0 \\ \lambda^1 & & & + & \lambda^3 & = & 0\,, \end{array}$$

und diese lineare Gleichungssystem besitzt offenbar nur die Null als Lösung.

d) Es sei wieder $e_\nu := (0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{K}^n$ mit Eintrag 1 an der Stelle ν. Dann sind die Vektoren $e_1, \ldots, e_n$ in $\mathbb{K}^n$ linear unabhängig.

e) Es seien für $x \in \mathbb{R}$ und $j \in \mathbb{N}$

$$p_0(x) := 1 \quad \text{und} \quad p_j(x) := x^j\,.$$

Dann ist die Menge der Funktionen $\{p_0, p_1, \ldots\}$ linear unabhängig.

Die Untersuchung auf lineare Abhängigkeit von n Vektoren des $\mathbb{K}^m$ führt auf ein homogenes Gleichungssystem von m Gleichungen in n Unbekannten.

[4] Man sagt in diesem Zusammenhang oft auch: (endliche) *Familie*

Bemerkung 3.1.3.
Eine Familie von Vektoren $(a_1, \ldots, a_n)$ ist genau dann linear abhängig, wenn sich (mindestens) einer der Vektoren als Linearkombination der anderen schreiben lässt.

Beweis:

a) Es sei zunächst Œ $a_1 = \sum_{\nu=2}^{n} \lambda^\nu a_\nu$ für geeignete $\lambda^\nu \in \mathbb{K}$, $\nu = 2, \ldots, n$. Mit $\lambda^1 := -1$ gilt dann $\sum_{\nu=1}^{n} \lambda^\nu a_\nu = 0$ mit $\lambda^1 = -1 \neq 0$. Somit ist die Familie $(a_1, \ldots, a_n)$ linear abhängig.

b) Es gelte $\sum_{\nu=1}^{n} \lambda^\nu a_\nu = 0$ und es sei Œ Einschränkung $\lambda^1 \neq 0$. Dann folgt $a_1 = -\frac{1}{\lambda^1} \sum_{\nu=2}^{n} \lambda^\nu a_\nu$. Somit haben wir a_1 als Linearkombination der übrigen Vektoren dargestellt. □

Bemerkung 3.1.4.
Eine Familie S von Vektoren ist genau dann linear unabhängig, wenn jedes $a \in \langle S \rangle$ nur auf genau eine Art und Weise linear aus den Vektoren aus S kombiniert werden kann.

Beweis:

a) Insbesondere ist $0 \in \langle S \rangle$. Der Nullvektor lässt sich als Linearkombination (mit Vektoren aus S) mit Koeffizienten 0 schreiben. Da dies nach Voraussetzung die einzige Möglichkeit ist, sind die Vektoren in S linear unabhängig.

b) Es sei ein Vektor auf zwei verschiedene Arten dargestellt, gelte also Œ

$$\sum_{\nu=1}^{n} \lambda^\nu a_\nu = \sum_{\nu=1}^{n} \mu^\nu a_\nu$$

mit $(\lambda^1, \ldots, \lambda^n) \neq (\mu^1, \ldots, \mu^n)$ und $a_\nu \in S$. Dann ist

$$\sum_{\nu=1}^{n} (\lambda^\nu - \mu^\nu)\, a_\nu \;=\; 0$$

eine nicht-triviale Linearkombination der Null. Die Vektoren sind also linear abhängig. □

3.2. **Basen.**

Es seien wieder $\mathbb{K}$ ein Körper und $\mathfrak{V}$ ein $\mathbb{K}$-Vektorraum.

Definition 3.2.1.
Eine Teilmenge S eines $\mathbb{K}$-Vektorraums $\mathfrak{V}$ heißt genau dann *Basis* von $\mathfrak{V}$, wenn S linear unabhängig ist und $\langle S \rangle = \mathfrak{V}$ gilt.

Beispiel 3.2.2.
$(e_1, \ldots, e_n)$ ist eine Basis des $\mathbb{K}^n$, die *Standardbasis* von $\mathbb{K}^n$.

Als Folgerung zu Bemerkung 3.1.4 erhalten wir

Folgerung 3.2.3.
$S \subset \mathfrak{V}$ ist genau dann eine Basis von $\mathfrak{V}$, wenn sich jeder Vektor aus $\mathfrak{V}$ in eindeutiger Weise als Linearkombination von Vektoren aus S schreiben lässt.

Bemerkung 3.2.4.
Es seien S eine linear unabhängige Teilmenge von $\mathfrak{V}$ und $b \in \mathfrak{V} \setminus \langle S \rangle$. Dann ist auch $S \cup \{b\}$ linear unabhängig.

Beweis: Für Vektoren $a_1, \ldots, a_n \in S$ ist zu zeigen, daß $(a_1, \ldots, a_n, b)$ linear unabhängig ist: Gelte

$$\mu b + \sum_{\nu=1}^{n} \lambda^\nu a_\nu = 0.$$

Dann folgt zunächst $\mu = 0$, denn sonst wäre $b \in \langle S \rangle$. Aufgrund der linearen Unabhängigkeit von S folgt dann aber auch $\lambda^\nu = 0$ für alle ν. Somit ist $S \cup \{b\}$ linear unabhängig. □

Mit Hilfe dieses Satzes kann man linear unabhängige Teilmengen sukzessive vergrößern, wenn sie nicht schon ‚maximal' sind.

Definition 3.2.5.
Es sei S eine linear unabhängige Teilmenge von $\mathfrak{V}$. Dann heißt S genau dann *maximal* oder genauer *maximal linear unabhängig*, wenn $S \cup \{b\}$ für jedes $b \in \mathfrak{V} \setminus S$ linear abhängig ist.

Satz 3.2.6.
Es sei $S \subset \mathfrak{V}$. Dann ist S genau dann eine Basis von $\mathfrak{V}$, wenn S linear unabhängig und maximal ist.

Beweis:

a) Es sei S linear unabhängig und maximal. Gäbe es $b \in \mathfrak{V} \setminus \langle S \rangle$, so wäre $S \cup \{b\}$ nach Bemerkung 3.2.4 auch linear unabhängig. Dies widerspricht aber der Maximalität. Somit erhalten wir $\langle S \rangle = \mathfrak{V}$.

b) Ist S eine Basis, so ist S — nach Definition — linear unabhängig. Da jedes $b \in \mathfrak{V}$ eine Linearkombination von Vektoren aus S ist, ist $S \cup \{b\}$ für beliebiges $b \in \mathfrak{V}$ nach Bemerkung 3.1.3 linear abhängig. Somit ist S bereits maximal. □

Das Lemma von Zorn, auf das wir in dieser Vorlesung *nicht* eingehen, erlaubt zu zeigen, daß jeder Vektorraum eine Basis besitzt. Wir lassen diesen — nicht konstruktiven — Beweis weg; denn das können wir getrost unseren Mathematikern überlassen! Einen Beweis findet man z. B. in W. H. Greub [5].

Satz 3.2.7.

(a) $\mathfrak{V}$ besitzt eine Basis.

(b) Je zwei Basen von $\mathfrak{V}$ können bijektiv aufeinander abgebildet werden.

(c) Ist A eine linear unabhängige Teilmenge von $\mathfrak{V}$, dann gibt es eine Teilmenge R von $\mathfrak{V}$ derart, daß $A \cup R$ eine Basis von $\mathfrak{V}$ ist.

Nach Konvention ist $\emptyset$ die Basis von $\{0\}$, da $\sum_{\emptyset} = 0$ gilt.

Definition 3.2.8.
Es sei S ein Erzeugendensystem von $\mathfrak{V}$. Dann heißt S genau dann *minimal,* falls es keine echte Teilmenge von S gibt, die schon $\mathfrak{V}$ erzeugt.

Satz 3.2.9.
Eine Familie S von Vektoren ist genau dann eine Basis von $\mathfrak{V}$, wenn S ein minimales Erzeugendensystem von $\mathfrak{V}$ ist.

Beweis:

a) Ist S *nicht* minimal (als Erzeugendensystem), so gibt es ein $b \in S$ derart, daß $S \setminus \{b\}$ schon ein Erzeugendensystem von $\mathfrak{V}$ ist. Insbesondere gilt dann aber auch

$$b = \sum_{\nu=1}^{n} \lambda^\nu a_\nu$$

für geeignete $\lambda^1, \ldots, \lambda^n \in \mathbb{K}$, $a_1, \ldots, a_n \in S \setminus \{b\}$, $n \in \mathbb{N}$. Nach Bemerkung 3.1.3 ist dann S linear abhängig, also keine Basis.

b) Es sei S ein minimales Erzeugendensystem. Ist S *linear abhängig,* so gibt es $a_1, \ldots, a_n \in S$ und $\lambda^1, \ldots, \lambda^n \in \mathbb{K}$, $n \in \mathbb{N}$, mit

$$\sum_{\nu=1}^{n} \lambda^\nu a_\nu = 0$$

und $(\lambda^1, \ldots, \lambda^n) \neq (0, \ldots, 0)$. Œ sei $\lambda^1 \neq 0$. Dann lässt sich a_1 als Linearkombination von $a_2, \ldots, a_n$ schreiben. Nach Satz 2.5.5 gilt $\mathfrak{V} = \langle S \rangle = \langle S \setminus \{a_1\} \rangle$. Somit ist S nicht minimal. *Widerspruch!* □

Für eine Teilmenge S des Vektorraums $\mathfrak{V}$ gilt nach den Sätzen 3.2.6 und 3.2.9:

Folgerung 3.2.10.
Die folgenden Aussagen sind äquivalent:

(1) S ist eine Basis von $\mathfrak{V}$.
(2) S ist eine maximale linear unabhängige Teilmenge von $\mathfrak{V}$.
(3) S ist ein minimales Erzeugendensystem von $\mathfrak{V}$.

Im endlich erzeugten Fall können wir auch absteigende Folgen von Erzeugendensystemen betrachten, um eine Basis zu bekommen.

Bemerkung 3.2.11.
Es sei S ein endliches *Erzeugendensystem von $\mathfrak{V}$. Dann gibt es eine Basis $B \subset S$ von $\mathfrak{V}$.*

Beweis: Ist S keine Basis, so ist S nach Folgerung 3.2.10 nicht minimal. Somit gibt es ein $S_1 \subsetneq S$, das ebenfalls ein Erzeugendensystem ist. Ist

S_1 keine Basis, so wiederholen wir diesen Schritt und erhalten $S_2 \subsetneq S_1$, ... Da S endlich ist, bricht die strikt absteigende Folge

$$S \supsetneq S_1 \supsetneq S_2 \supsetneq \cdots \supsetneq S_k$$

bei einem $k \in \mathbb{N}$ ab. $B := S_k$ ist minimal und somit nach Folgerung 3.2.10 eine Basis. □

3.3. **Dimension.**

Es seien wieder $\mathbb{K}$ ein Körper und $\mathfrak{V}$ ein $\mathbb{K}$-Vektorraum.

In diesem Abschnitt soll jedem Vektorraum die *Dimension* als eine charakteristische Größe zugeordnet werden. Dies wollen wir über die ‚Länge' einer Basis definieren. Zunächst einmal ist aber *nicht* klar, ob die Dimension so wohldefiniert ist, da es verschiedene Basen mit unterschiedlich vielen Elementen geben könnte. (Den von uns nicht bewiesenen Satz 3.2.7 wollen wir *nicht* heranziehen.) Wir werden mit Satz 3.3.1 zeigen, daß dies *nicht* der Fall ist, also alle Basen gleich lang sind.

Satz 3.3.1 (Austauschsatz von Steinitz)**.**
In $\mathfrak{V}$ seien $(v_1, \ldots, v_r)$ eine Basis und eine linear unabhängige Familie $(w_1, \ldots, w_n)$ gegeben. Dann gilt $n \leq r$, und — nach einer eventuellen Umnumerierung — liefert auch $B := (w_1, \ldots, w_n, v_{n+1}, \ldots, v_r)$ eine Basis von $\mathfrak{V}$.

Zum *Beweis* zeigen wir zunächst einen ersten wichtigen Schritt:

Lemma 3.3.2 (Austauschlemma)**.**
Es seien $(v_1, \ldots, v_r)$ eine Basis von $\mathfrak{V}$ und

$$w = \sum_{\varrho=1}^{r} \lambda^\varrho \, v_\varrho$$

mit $(\lambda^1, \ldots, \lambda^r) \in \mathbb{K}^r$ und $\lambda^k \neq 0$ für ein $k \in \{1, \ldots, r\}$. Dann ist

$$(v_1, \ldots, v_{k-1}, w, v_{k+1}, \ldots, v_r)$$

wieder eine Basis von $\mathfrak{V}$. Der Vektor v_k kann also in der Basis gegen w ausgetauscht werden.

Beweis: ℭ $k = 1$ (sonst Umnumerierung). Zu zeigen ist dann also: $B := (w, v_2, \ldots, v_r)$ ist eine Basis:
$\lambda^1 \neq 0$ liefert $v_1 \in \langle B \rangle$ und damit $\mathfrak{V} = \langle v_1, \ldots, v_r \rangle \subset \langle B \rangle$. B ist also ein Erzeugendensystem. Aus

$$0 = \mu w + \sum_{\varrho=2}^{r} \mu^\varrho \, v_\varrho = \mu\lambda^1 v_1 + \sum_{\varrho=2}^{r} (\mu\lambda^\varrho + \mu^\varrho) \, v_\varrho \quad (\text{für } \mu, \mu^\varrho \in \mathbb{K})$$

folgt zunächst $\mu = 0$, da $(v_1, \ldots, v_r)$ linear unabhängig ist mit $\lambda^1 \neq 0$, dann $\sum_{\varrho=2}^{r} \mu^\varrho \, v_\varrho = 0$, also $\mu^2 = \cdots = \mu^r = 0$. □

Beweis des Satzes (durch Induktion über n)*:* Für $n = 0$ ist nichts zu zeigen. Für ein $n \in \mathbb{N}$ sei der Satz schon für $n-1$ bewiesen. Da auch $(w_1, \ldots, w_{n-1})$ linear unabhängig ist, folgt also, daß $n-1 \leq r$ gilt und — nach einer eventuellen Umnumerierung —

$$(w_1, \ldots, w_{n-1}, v_n, \ldots, v_r)$$

eine Basis von $\mathfrak{V}$ ist. Zum Nachweis von $n \leq r$ ist also nur noch zu zeigen, daß $n-1 = r$ nicht gelten kann: Dann wäre ja schon $(w_1, \ldots, w_{n-1})$ eine Basis von $\mathfrak{V}$, was 3.2.6 widerspricht. Wir notieren $w_n = \lambda^1 w_1 + \cdots + \lambda^{n-1} w_{n-1} + \lambda^n v_n + \cdots + \lambda^r v_r$ mit geeigneten $\lambda^\varrho \in \mathbb{K}$. Wäre $\lambda^n = \cdots = \lambda^r = 0$, so hätte man einen Widerspruch zur vorausgesetzten linearen Unabhängigkeit von $w_1, \ldots, w_n$. Œ kann daher wieder $\lambda^n \neq 0$ angenommen werden. Dann kann nach dem Austauschlemma v_n gegen w_n so ausgetauscht werden, daß B eine Basis von $\mathfrak{V}$ ist. □

Nun können wir ernten:

Folgerung 3.3.3.

Hat $\mathfrak{V}$ eine endliche Basis, so ist jede *Basis von $\mathfrak{V}$ endlich.*

Beweis: ∘ ∘ ∘ □

Folgerung 3.3.4.

Je zwei endliche Basen von $\mathfrak{V}$ haben die gleiche Länge.

Beweis: ∘ ∘ ∘ □

Damit können wir nun definieren:

Definition 3.3.5.

$$\dim_{\mathbb{K}} \mathfrak{V} := \begin{cases} \infty, & \text{falls } \mathfrak{V} \text{ keine endliche Basis besitzt,} \\ r, & \text{falls } \mathfrak{V} \text{ eine Basis der Länge } r \in \mathbb{N}_0 \text{ besitzt.} \end{cases}$$

$\dim_{\mathbb{K}} \mathfrak{V}$ heißt die *Dimension* von $\mathfrak{V}$. Falls klar ist, welcher Körper gemeint ist, schreibt man auch nur $\dim \mathfrak{V}$.

Folgerung 3.3.6.

Es sei $\mathfrak{V}$ ein n-dimensionaler Vektorraum (für ein $n \in \mathbb{N}_0$). Dann gelten:

a) Jedes Erzeugendensystem von $\mathfrak{V}$ besitzt wenigstens n Vektoren.
b) Jede Familie von mehr als n Vektoren ist linear abhängig.
c) Jedes Erzeugendensystem aus n Vektoren ist eine Basis von $\mathfrak{V}$.
d) Je n linear unabhängige Vektoren bilden eine Basis von $\mathfrak{V}$.

Beweis.

a) Nach Bemerkung 3.2.11 gäbe es zu einem Erzeugendensystem mit weniger als n Elementen auch eine Basis mit weniger als n Elementen im Widerspruch zu Folgerung 3.3.4.

b) Nach Satz 3.3.1 könnte man sonst diese Menge zu einer Basis mit mehr als n Elementen ergänzen. Dies widerspricht wiederum Folgerung 3.3.4
c) Ein minimales Erzeugendensystem ist eine Basis.
d) Eine maximale linear unabhängige Familie ist eine Basis. □

Folgerung 3.3.7.
Es seien $\mathfrak{W}$ ein Unterraum von $\mathfrak{V}$ und $\dim \mathfrak{V} = n$. Dann gilt $\dim \mathfrak{W} \leq \dim \mathfrak{V}$, und Gleichheit gilt genau dann, wenn $\mathfrak{V} = \mathfrak{W}$ ist.

Beweis: Es seien S eine Basis von $\mathfrak{W}$ und T mit $S \subset T$ eine Basis von $\mathfrak{V}$. Dann gilt für die Anzahlen $|S| \leq |T|$. Bei Gleichheit gilt $S = T$, und die Erzeugnisse stimmen überein. Die Rückrichtung ist trivial. □

Beispiele 3.3.8.

a) Es sei wieder $e_\nu := (0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{K}^n$ mit Eintrag 1 an der Stelle ν. Dann bilden diese Vektoren $e_1, \ldots, e_n$ in $\mathbb{K}^n$ eine Basis (Standardbasis). Somit gilt: $\dim_{\mathbb{K}} \mathbb{K}^n = n$
b) Die Menge $C^0[0,1]$ der stetigen (reellwertigen) Funktionen bildet einen $\mathbb{R}$-Vektorraum mit $\dim_{\mathbb{R}} C^0[0,1] = \infty$. Denn nach Beispiel 3.1.2, *e*) wissen wir: Es seien für $x \in \mathbb{R}$ und $j \in \mathbb{N}$

$$p_0(x) := 1 \quad \text{und} \quad p_j(x) := x^j .$$

Dann ist die Menge der Funktionen $\{p_0, p_1, \ldots\}$ linear unabhängig.
c) $\dim_{\mathbb{R}} \mathbb{C} = 2$ (Denn $(1, i)$ ist eine Basis.)

Satz 3.3.9 (Dimensionsformel)**.**

Es seien $\mathfrak{U}, \mathfrak{V} \subset \mathfrak{W}$ endlich-dimensionale Unterräume eines Vektorraumes $\mathfrak{W}$. Dann gilt

$$\dim(\mathfrak{U} + \mathfrak{V}) = \dim \mathfrak{U} + \dim \mathfrak{V} - \dim(\mathfrak{U} \cap \mathfrak{V}).$$

Beweis: Mit $k := \dim(\mathfrak{U} \cap \mathfrak{V})$ sei $(w_1, \ldots, w_k)$ eine Basis von $\mathfrak{U} \cap \mathfrak{V}$. Mit $n := \dim \mathfrak{U} - k$ können dann $u_1, \ldots, u_n$ so gewählt, daß sie zusammen mit den w_κ's eine Basis von $\mathfrak{U}$ bilden. Mit $m := \dim \mathfrak{V} - k$ können entsprechend $v_1, \ldots, v_m$ so gewählt werden, daß sie zusammen mit den w_κ's eine Basis von $\mathfrak{V}$ bilden. Wir behaupten nun, daß alle diese Vektoren zusammen eine Basis von $\mathfrak{U} + \mathfrak{V}$ bilden. Es ist klar, daß sie $\mathfrak{U} + \mathfrak{V}$ erzeugen. Zum Nachweis der linearen Unabhängigkeit sei

$$\underbrace{\sum_\kappa \gamma^\kappa w_\kappa}_{=:w \in \mathfrak{U}\cap\mathfrak{V}} + \underbrace{\sum_\nu \alpha^\nu u_\nu}_{=:u \in \mathfrak{U}} + \underbrace{\sum_\mu \beta^\mu v_\mu}_{=:v \in \mathfrak{V}} = 0 \quad (\text{mit} \ldots).$$

Dann ist $v \in \mathfrak{V}$ und $v = -w - u \in \mathfrak{U}$, also $v \in \mathfrak{U} \cap \mathfrak{V}$. Die w_κ's bilden zusammen mit den v_μ's eine Basis von $\mathfrak{V}$, Vektoren in $\mathfrak{U} \cap \mathfrak{V}$ werden aber bereits eindeutig als Linearkombination von w_κ's dargestellt. Die

obige Darstellung von v enhält daher keine v_μ's, und somit ist $v = 0$. Ebenso ist $u = 0$. Das liefert die lineare Unabhängigkeit. □

Koordinaten

Definition 3.3.10 (Koordinaten)**.**
Es sei $(a_1, \ldots, a_n)$ eine (geordnete) Basis von $\mathfrak{V}$. Dann läßt sich nach Folgerung 3.2.3 jedes $a \in \mathfrak{V}$ in eindeutiger Weise aus den a_ν's linear kombinieren

$$a = \sum_{\nu=1}^{n} \lambda^\nu a_\nu$$

mit $\lambda^\nu \in \mathbb{K}$. Die Koeffizienten $\lambda^1, \ldots, \lambda^n$ heißen *Koordinaten* von a *bezüglich der Basis* $(a_1, \ldots, a_n)$.

3.4. Lineare Abbildungen, Teil I.

Es seien $\mathfrak{U}, \mathfrak{V}$ und $\mathfrak{W}$ drei Vektorräume über einem Körper $\mathbb{K}$.

Definition 3.4.1 (Lineare Abbildung)**.**

Eine Abbildung $f\colon \mathfrak{U} \longrightarrow \mathfrak{V}$ heißt genau dann *linear* — genauer auch $\mathbb{K}$*-linear* — oder *Homomorphismus* wenn
$f(a+b) = f(a) + f(b)$ für alle $a, b \in \mathfrak{U}$ (*additiv*) und
$f(\lambda a) = \lambda f(a)$ für alle $a \in \mathfrak{U}, \lambda \in \mathbb{K}$ (*homogen*) gelten.

Lineare Abbildungen sind also mit den Vektorraum-Operationen ‚verträglich' , sie respektieren die Vektorraumstruktur.

Lemma 3.4.2.
$f\colon \mathfrak{U} \longrightarrow \mathfrak{V}$ sei linear. Dann gelten die folgenden Aussagen:
a) $f(0) = 0$, $f(-a) = -f(a)$, $f(a-b) = f(a) - f(b)$ $(\ldots)$
b) $f\left(\sum_{\nu=1}^{n} \lambda^\nu a_\nu\right) = \sum_{\nu=1}^{n} \lambda^\nu f(a_\nu)$ *für alle* $\lambda^\nu \in \mathbb{K}$, $a_\nu \in \mathfrak{U}$.
Insbesondere ist

$$f(\lambda a + b) = \lambda f(a) + f(b)$$

für alle $\lambda \in \mathbb{K}$ *und* $a, b \in \mathfrak{U}$.
Diese letzte Beziehung ist äquivalent zur Linearität.

Beweis: ∘ ∘ ∘ □

Es ist im Folgenden zweckmäßig, *Vektoren* im $\mathbb{K}^n$ nicht — wie bisher — als Zeilenvektoren, sondern *als Spaltenvektoren* zu schreiben. D. h. , wir fassen den $\mathbb{K}^n$ als $\mathbb{K}^{n\times 1}$ auf.

Beispiele

(B0) Die *Nullabbildung* $0\colon \mathfrak{U} \longrightarrow \mathfrak{V}$, definiert durch

$$0(u) := 0$$

für alle $u \in \mathfrak{U}$, ist linear.

(B1) Die *Identität* $\mathrm{id} = \mathrm{id}_{\mathfrak{U}} \colon \mathfrak{U} \longrightarrow \mathfrak{U}$, definiert durch

$$\mathrm{id}(u) := u$$

für alle $u \in \mathfrak{U}$, ist linear.

(B2) Es sei $A \in \mathbb{K}^{m\times n}$. Dann ist $f_A \colon \mathbb{K}^n \longrightarrow \mathbb{K}^m$, definiert durch

$$f(x) := f_A(x) := Ax$$

für $x \in \mathbb{K}^n$, — nach Lemma 1.3.5 — linear.

Ist zudem noch $B \in \mathbb{K}^{n\times k}$, also $f_B \colon \mathbb{K}^k \longrightarrow \mathbb{K}^n$, dann gilt

$$f_{AB} = f_A \circ f_B \,.$$

Beweis: ∘ ∘ ∘ □

Dies zeigt, daß ‚wir' das Matrizenprodukt ‚vernünftig' definiert haben.

(B3) Es sei wieder $(a_1, \ldots, a_n)$ eine (geordnete) Basis von $\mathfrak{U}$. Dann lässt sich jedes $a \in \mathfrak{U}$ in eindeutiger Weise aus den a_ν's linear kombinieren

$$a = \sum_{\nu=1}^{n} \lambda^\nu a_\nu$$

mit $\lambda^\nu \in \mathbb{K}$. Die Abbildung $f \colon \mathfrak{U} \longrightarrow \mathbb{K}^n$, die jedem $a \in \mathfrak{U}$ den Vektor $(\lambda^1, \ldots, \lambda^n)$ seiner Koordinaten (bezüglich der Basis $(a_1, \ldots, a_n)$) zuordnet, ist linear.

(B4) Es sei $\mathfrak{W}$ ein weiterer $\mathbb{K}$-Vektorraum. Sind die beiden Abbildungen $f \colon \mathfrak{U} \longrightarrow \mathfrak{V}$ und $g \colon \mathfrak{V} \longrightarrow \mathfrak{W}$ linear. Dann ist auch

$$g \circ f \colon \mathfrak{U} \longrightarrow \mathfrak{W}$$

linear. *Beweis:* ∘ ∘ ∘ □

(B5) Mit einer nicht-leeren Menge M betrachten wir den $\mathbb{K}$-Vektorraum — vgl. g) der Beispiele 2.3.2 — der Abbildungen von M in $\mathfrak{U}$:

$$\mathfrak{F} := \mathfrak{F}(M, \mathfrak{U}) := \{f \mid f \colon M \longrightarrow \mathfrak{U}\}$$

Für ein $t \in M$ ist dann die *Auswertungsabbildung*

$$\begin{array}{rccc} A\colon & \mathfrak{F} & \longrightarrow & \mathfrak{U} \\ & \cup & & \cup \\ & f & \longmapsto & f(t) \end{array}$$

linear.

(B6) In der Analysis lernt man: *Grenzwertbildung, Differentiation* und *Integration* (und vieles mehr) *sind linear.*

◁

Bemerkung 3.4.3.
Es seien $f\colon \mathfrak{U} \longrightarrow \mathfrak{V}$ *eine lineare Abbildung,* $\mathfrak{U}_1$ *ein Unterraum von* $\mathfrak{U}$ *und* $\mathfrak{V}_1$ *ein Unterraum von* $\mathfrak{V}$*. Dann gelten:*

a) $f(\mathfrak{U}_1)$ *ist ein Unterraum von* $\mathfrak{V}$*.*

b) Speziell ist also das Bild von f

$$\operatorname{im} f := \{f(u) \mid u \in \mathfrak{U}\} = f(\mathfrak{U})$$

ein Unterraum von $\mathfrak{V}$*.*

c) $f^{-1}(\mathfrak{V}_1)$ *ist ein Unterraum von* $\mathfrak{U}$*.*

d) Speziell ist also der Kern von f

$$\ker f := \{u \in \mathfrak{U} \mid f(u) = 0\}$$

ein Unterraum von $\mathfrak{U}$*.*

Beweis: ∘ ∘ ∘ □

Wir bezeichnen

$\mathcal{L}_{\mathbb{K}}(\mathfrak{U}, \mathfrak{V}) := \mathcal{L}(\mathfrak{U}, \mathfrak{V}) := \{f \mid f\colon \mathfrak{U} \longrightarrow \mathfrak{V} \text{ linear}\}$
Manche Autoren notieren dafür auch $\mathrm{HOM}_{\mathbb{K}}(\mathfrak{U}, \mathfrak{V})$.

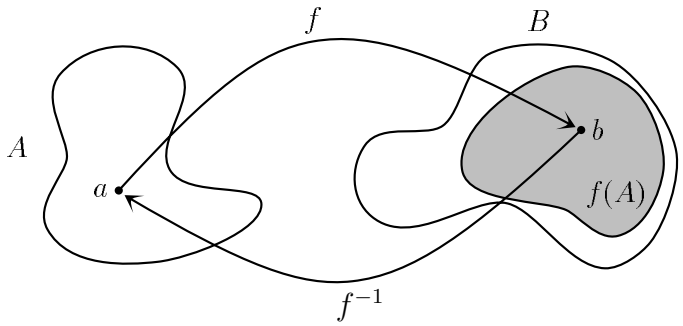
Nachdem wir uns bei dem Beweis dieser Bemerkung ausgeruht (und vielleicht gelangweilt) haben, können wir uns ein paar vornehme Vokabeln merken:

Definition 3.4.4.
Eine *lineare* Abbildung $f\colon \mathfrak{U} \longrightarrow \mathfrak{V}$ heißt genau dann:

- *Monomorphismus*, wenn f injektiv ist,
- *Epimorphismus*, wenn f surjektiv ist,
- *Isomorphismus*, wenn f bijektiv ist,
- *Endomorphismus*, wenn $\mathfrak{U} = \mathfrak{V}$ gilt,
- *Automorphismus*, wenn f bijektiv ist und $\mathfrak{U} = \mathfrak{V}$ gilt.

Umkehrabbildung

A, B seien Mengen und $f\colon A \longrightarrow B$ eine *injektive* Abbildung.
Für alle $b \in f(A)$ existiert dann *eindeutig* ein $a \in A$ mit $f(a) = b$. Die Abbildung, die jedem $b \in f(A)$ gerade dieses $a \in A$ zuordnet, bezeichnen wir mit f^{-1} und sprechen von der *Umkehrabbildung zu* f. Es gilt also insbesondere $D_{f^{-1}} = f(A)$.



Nach Definition von f^{-1} gilt offenbar:

(0) $f(a) = b \iff f^{-1}(b) = a \quad$ für $(a, b) \in A \times f(A)$.

Weiter erhält man:

Bemerkung 3.4.5.

(1) $\forall\ a \in A \quad f^{-1}\big(f(a)\big) = a$

(2) $\forall\ b \in f(A) \quad f\left(f^{-1}(b)\right) = b$

(3) $f^{-1}\colon\ f(A) \longrightarrow A$ ist bijektiv.

(4) $D_{(f^{-1})^{-1}} = D_f$ und für alle $x \in D_f \quad \left(f^{-1}\right)^{-1}(x) = f(x)$

Der *Beweis* von (1) und (2) ist unmittelbar durch (0) gegeben. (3): Nach (1) ist f^{-1} surjektiv, nach (2) ist f^{-1} injektiv; denn (1) zeigt, daß zu beliebigem $a \in A$ gerade $f(a)$ ein Urbild unter f^{-1} ist, und (2) liefert, daß für $b_1, b_2 \in f(A)$ aus $f^{-1}(b_1) = f^{-1}(b_2)$ (durch Anwendung von f) $b_1 = b_2$ folgt. (4): Nach Definition der Umkehrabbildung ist der Definitionsbereich von $\left(f^{-1}\right)^{-1}$ das Bild von $D_{f^{-1}} = f(A)$ unter f^{-1}, also nach (1) gerade $A = D_f$. Für $a \in A = D_f$ und $b := f(a)$ zeigt (0) zunächst $f^{-1}(b) = a$ und weiter (jetzt angewendet auf f^{-1} statt f) $b = \left(f^{-1}\right)^{-1}(a)$. □

Im Zusammenhang mit der Umkehrabbildung tritt eine — allgemein übliche — Doppelbezeichnung auf, die ich kurz erläutern möchte:

Einerseits bezeichnet $f^{-1}(B')$ für $B' \subset B$ das Urbild von B' unter f, andererseits — falls f injektiv ist und $B' \subset f(A)$ gilt — das Bild von B' unter der Umkehrabbildung f^{-1}. Zunächst sind das natürlich zwei verschiedene Dinge. Wir notieren jedoch hierzu die

Bemerkung 3.4.6.

Ist f injektiv *und* $B' \subset f(A)$*, dann stimmen das Bild von* B' *unter* f^{-1} *und das Urbild von* B' *bezüglich* f *überein.*

Die Bezeichnungen sind also, wenn beide Bildungen sinnvoll sind, konsistent; die Doppelbezeichnung ist daher nicht störend.
Für diejenigen Leser, die es ganz genau wissen wollen, notiere ich einen *Beweis:* Ein Element $a \in A$ liegt im Bild von B' unter f^{-1} genau dann, wenn es ein $b \in B'$ mit $f^{-1}(b) = a$ gibt. Nach (0) ist dies

gleichbedeutend dazu, daß es ein $b \in B'$ gibt mit $f(a) = b$, und das bedeutet gerade, daß a im Urbild von B' bezüglich f liegt. □

Bemerkung 3.4.7.

Es seien $f\colon \mathfrak{U} \longrightarrow \mathfrak{V}$ *linear und* $(a_1, \ldots, a_n) \in \mathfrak{U}^n$ *linear abhängig. Dann ist auch* $(f(a_1), \ldots, f(a_n))$ *linear abhängig.*

Beweis: Aus $\sum_{\nu=1}^n \lambda^\nu a_\nu = 0$ folgt $\sum_{\nu=1}^n \lambda^\nu f(a_\nu) = 0$ für .... □

Bemerkung 3.4.8.

Es seien $f\colon \mathfrak{U} \longrightarrow \mathfrak{V}$ *ein* Monomorphismus *und* $(a_1, \ldots, a_n) \in \mathfrak{U}^n$ *linear unabhängig. Dann ist auch* $(f(a_1), \ldots, f(a_n))$ *linear unabhängig.*

Beweis: Aus $\sum_{\nu=1}^n \lambda^\nu f(a_\nu) = 0$ folgt hier $\sum_{\nu=1}^n \lambda^\nu a_\nu = 0$ für .... □

Satz 3.4.9.

$\mathcal{L}_{\mathbb{K}}(\mathfrak{U}, \mathfrak{V})$ *ist ein Unterraum des* $\mathbb{K}$*-Vektorraums* $\mathfrak{F}(\mathfrak{U}, \mathfrak{V})$*.*

Beweis: $0 \in \mathcal{L}(\mathfrak{U}, \mathfrak{V})$: ✓ Es seien $f, g \in \mathcal{L}(\mathfrak{U}, \mathfrak{V})$ und $\lambda \in \mathbb{K}$: Für $x, y \in \mathfrak{U}$ und $\alpha \in \mathbb{K}$ gelten dann:
$(f+g)(\alpha x + y) = f(\alpha x + y) + g(\alpha x + y) = (\alpha f(x) + f(y)) + (\alpha g(x) + g(y)) = \alpha(f(x) + g(x)) + (f(y) + g(y)) = \alpha(f+g)(x) + (f+g)(y)$.
Das zeigt: $f + g \in \mathcal{L}(\mathfrak{U}, \mathfrak{V})$.
$(\lambda f)(\alpha x + y) = \lambda f(\alpha x + y) = \lambda(\alpha f(x) + f(y)) = \alpha\lambda f(x) + \lambda f(y) = \alpha(\lambda f)(x) + (\lambda f)(y)$, also $\lambda f \in \mathcal{L}(\mathfrak{U}, \mathfrak{V})$.
Bei der vorletzten Gleichung wird erstmals entscheidend benutzt, daß die Multiplikation in $\mathbb{K}$ kommutativ ist! □

Bemerkung 3.4.10.

Eine lineare Abbildung ist genau dann injektiv, *wenn ihr Kern nur aus der Null besteht.*

Beweis: ∘ ∘ ∘ □

Lemma 3.4.11 (Koordinatenabbildung)**.**

Es sei $\mathfrak{V}$ *ein* $\mathbb{K}$*-Vektorraum mit geordneter Basis* $B := (a_1, \ldots, a_n)$*. Dann ist die Abbildung* $f\colon \mathbb{K}^n \longrightarrow \mathfrak{V}$*, definiert durch*

$$\mathbb{K}^n \ni (\lambda^1, \ldots, \lambda^n) \longmapsto \sum_{\nu=1}^n \lambda^\nu a_\nu \in \mathfrak{V}$$

ein Isomorphismus. Wir bezeichnen ihn mit Φ_B.

Die Umkehrabbildung zu f hatten wir schon in (B3) betrachtet.

Beweis: Die *Linearität* von f ist klar. Die *Injektivität* ist durch die lineare Unabhängigkeit von B gegeben. f ist *surjektiv,* da B ein Erzeugendensystem ist. □

Es sei A eine $(m \times n)$-Matrix mit Elementen aus $\mathbb{K}$, also $A \in \mathbb{K}^{m\times n}$. Ich erinnere noch einmal daran, daß m die Anzahl der Zeilen und n die Anzahl der Spalten ist.

$$A = (a^\mu_\nu)_{\substack{1\le\mu\le m\\ 1\le\nu\le n}} = \begin{pmatrix} a^1_1 & a^1_2 & \dots & a^1_n \\ a^2_1 & a^2_2 & \dots & a^2_n \\ \vdots & \vdots & \ddots & \vdots \\ a^m_1 & a^m_2 & \cdots & a^m_n \end{pmatrix} .$$

Definition 3.4.12.
Für $\mu = 1, \dots, m$ bezeichne A^μ die μ-te Zeile von A, also

$$A^\mu = (a^\mu_1, a^\mu_2, \dots, a^\mu_n) .$$

Entsprechend bezeichne A_ν für $\nu = 1, \dots, n$ die ν-te Spalte von A:

$$A_\nu = \begin{pmatrix} a^1_\nu \\ a^2_\nu \\ \vdots \\ a^m_\nu \end{pmatrix} .$$

Für $B \in \mathbb{K}^{n\times k}$ kann damit die $(m \times k)$-Matrix AB in der Form

$$AB = (A^\mu B_\kappa)_{\substack{\mu=1,\dots,m\\ \kappa=1,\dots,k}}$$

geschrieben werden.
Für $x \in \mathbb{K}^n = \mathbb{K}^{n\times 1}$ speziell gilt also

$$Ax = (A^\mu x)_{\mu=1,\dots,m} = \begin{pmatrix} A^1 x \\ A^2 x \\ \vdots \\ A^m x \end{pmatrix} .$$

Mit den Einheitsvektoren $e_\nu \in \mathbb{K}^n$ gilt offenbar:

$$f_A(e_\nu) = Ae_\nu = (A^\mu e_\nu)_{\mu=1,\dots,m} = (a^\mu_\nu)_{\mu=1,\dots,m} = A_\nu$$

Diese Beziehung

$$\boxed{Ae_\nu = A_\nu}$$

ist an vielen Stellen hilfreich. Der ν-te Einheitsvektor ‚schneidet' aus A gerade die ν-te Spalte heraus.

Bemerkung 3.4.13.
Eine lineare Abbildung $f\colon \mathfrak{U} \longrightarrow \mathfrak{V}$ ist genau dann ein Isomorphismus, *wenn es eine Abbildung $g\colon \mathfrak{V} \longrightarrow \mathfrak{U}$ mit $g \circ f = \mathrm{id}_\mathfrak{U}$ und $f \circ g = \mathrm{id}_\mathfrak{V}$ gibt. Diese Abbildung g ist dann gleich f^{-1} und ebenfalls ein Isomorphismus.*

Beweis: Ist f ein Isomorphismus, dann erfüllt $g := f^{-1}$ nach Bemerkung 3.4.5 die beiden Beziehungen $g \circ f = \mathrm{id}_\mathfrak{U}$ und $f \circ g = \mathrm{id}_\mathfrak{V}$. Umgekehrt folgt aus $g \circ f = \mathrm{id}_\mathfrak{U}$ die Injektivität von f; denn für $x_1, x_2 \in \mathfrak{U}$

mit $f(x_1) = f(x_2)$ folgt: $x_1 = g(f(x_1)) = g(f(x_2)) = x_2$. Entsprechend folgt aus $f \circ g = \mathrm{id}_{\mathfrak{V}}$ die Surjektivität von f; denn zu $y \in \mathfrak{V}$ ist $x := g(y)$ ein Urbild von y unter f. Linearität von g: Zu $v, w \in \mathfrak{V}$ und $\lambda \in \mathbb{K}$ existieren $x, y \in \mathfrak{U}$ mit $f(x) = v, f(y) = w$. $f(\lambda x + y) = \lambda f(x) + f(y) = \lambda v + w$, also $\lambda f^{-1}(v) + f^{-1}(w) = \lambda x + y = f^{-1}(\lambda v + w)$. $g = f^{-1}$: ✓ □

Bemerkung 3.4.14.

Sind $f\colon \mathfrak{U} \longrightarrow \mathfrak{V}$ und $g\colon \mathfrak{V} \longrightarrow \mathfrak{W}$ Isomorphismen, dann ist auch $g \circ f\colon \mathfrak{U} \longrightarrow \mathfrak{W}$ ein Isomorphismus.

Beweis:
Nach (B4) ist $g \circ f$ linear. $g \circ f$ surjektiv: ✓ $g \circ f$ injektiv: ✓ □

Definition 3.4.15.

Zwei Vektorräume $\mathfrak{U}$ und $\mathfrak{V}$ heißen genau dann *isomorph,* wenn es einen Isomorphismus $f\colon \mathfrak{U} \longrightarrow \mathfrak{V}$ gibt. Wir schreiben dafür auch $\simeq$.

Folgerung 3.4.16.

Ist $\mathfrak{V}$ ein n-dimensionaler $\mathbb{K}$-Vektorraum, dann ist $\mathfrak{V}$ isomorph zu $\mathbb{K}^n$.

Beweis: Lemma 3.4.11 □

> Mäxchen SCHLAUMEIER meint dazu: Alle endlich-dimensionalen $\mathbb{K}$-Vektorräume sind im Wesentlichen $\mathbb{K}^n$ (für ein geeignetes $n \in \mathbb{N}_0$). Im $\mathbb{K}^n$ geht fast alles wie im $\mathbb{R}^n$. Den $\mathbb{R}^n$ kennen wir aber schon ausreichend. Die ‚Lineare Algebra' kann daher bald abgehakt werden!
> Für manche Gesichtspunkte hat er damit durchaus recht. *Aber:* Die ‚Identifizierung' bezieht sich auf eine spezielle Basis. Wichtige Eigenschaften der Lineare Algebra sind jedoch basisinvariant. Viele der noch anstehenden Probleme bestehen gerade darin, eine dem speziellen Problem angepaßte ‚schöne' Basis zu finden. . . .

Satz 3.4.17.

Es seien $\mathfrak{U}$ und $\mathfrak{V}$ endlich erzeugbar. Dann sind $\mathfrak{U}$ und $\mathfrak{V}$ genau dann isomorph, wenn $\dim \mathfrak{U} = \dim \mathfrak{V}$ gilt.

Beweis:
Hat man $n := \dim \mathfrak{U} = \dim \mathfrak{V}$, dann gibt es nach Lemma 3.4.11 Isomorphismen $\varphi\colon \mathfrak{U} \longrightarrow \mathbb{K}^n$ und $\psi\colon \mathfrak{V} \longrightarrow \mathbb{K}^n$. Nach den Bemerkungen 3.4.13 und 3.4.14 ist dann $\psi^{-1} \circ \varphi\colon \mathfrak{U} \longrightarrow \mathfrak{V}$ ein Isomorphismus.

Es seien $(a_1, \ldots, a_n)$ eine Basis von $\mathfrak{U}$ und $f\colon \mathfrak{U} \longrightarrow \mathfrak{V}$ ein Isomorphismus. Dann ist $(f(a_1), \ldots, f(a_n))$ linear unabhängig (nach Bemerkung 3.4.8). Es ist auch ein Erzeugendensystem von $\mathfrak{V}$: Für $b \in \mathfrak{V}$ gilt

$$\mathfrak{U} \ni f^{-1}(b) = \sum_{\nu=1}^{n} \lambda^\nu a_\nu$$

mit geeigneten $\lambda_\nu \in \mathbb{K}$. Somit ist $b = f\left(f^{-1}(b)\right) = \sum_{\nu=1}^{n} \lambda^\nu f(a_\nu)$. Insgesamt ist also $(f(a_1), \ldots, f(a_n))$ eine Basis von $\mathfrak{V}$. Da die Anzahl

von Basiselementen folglich in beiden Vektorräumen übereinstimmt, gilt $\dim \mathfrak{U} = \dim \mathfrak{V}$. □

Bemerkung 3.4.18.

a) $\mathfrak{U} \simeq \mathfrak{U}$
b) $\mathfrak{U} \simeq \mathfrak{V} \Longrightarrow \mathfrak{V} \simeq \mathfrak{U}$
c) $\mathfrak{U} \simeq \mathfrak{V} \wedge \mathfrak{V} \simeq \mathfrak{W} \Longrightarrow \mathfrak{U} \simeq \mathfrak{W}$

Beweis: a): $\mathrm{id}_{\mathfrak{U}}\colon \mathfrak{U} \longrightarrow \mathfrak{U}$ ist ein Isomorphismus.
b): nach 3.4.13 . *c)*: nach 3.4.14 . □

Die Isomorphie von Vektorräumen ist also eine *Äquivalenzrelation.*

Definition 3.4.19.

$$\mathrm{GL}(\mathfrak{V}) := \{f \mid f\colon \mathfrak{V} \longrightarrow \mathfrak{V} \text{ Isomorphismus}\}$$

heißt *„general linear group“*, *„Allgemeine lineare Gruppe“* oder *„Automorphismengruppe“* .

Bemerkung 3.4.20.

$(\mathrm{GL}(\mathfrak{V}), \circ)$ *ist eine Gruppe.*

Der *Beweis* ist ganz leicht. Nur haben wir dummerweise[5] bisher nur abelsche Gruppen definiert und betrachtet. Und diese Gruppe ist *nicht* kommutativ! (Dies sieht man etwa über die beiden Matrizen aus dem letzten Beispiel von Beispiel 1.3.4.) Den allgemeinen Gruppenbegriff sehen wir uns im übernächsten Kapitel an. Dementsprechend stellen wir den Beweis von Bemerkung 3.4.20 zurück.

3.5. Direkte Summen.

Es seien $\mathfrak{S}$, $\mathfrak{T}$ und $\mathfrak{V}$ drei Vektorräume über einem Körper $\mathbb{K}$.

Wir haben schon Summen $\mathfrak{S} + \mathfrak{T}$ von Vektorräumen betrachtet (siehe Seite 25). Den wichtigen Spezialfall, daß $\mathfrak{S} \cap \mathfrak{T} = \{0\}$ gilt, wollen wir uns noch etwas genauer ansehen:

Definition 3.5.1.

Sind $\mathfrak{S}$ und $\mathfrak{T}$ Unterräume von $\mathfrak{V}$ mit $\mathfrak{S} + \mathfrak{T} = \mathfrak{V}$ und $\mathfrak{S} \cap \mathfrak{T} = \{0\}$, dann heißt $\mathfrak{S} + \mathfrak{T}$ *direkte Summe* (von $\mathfrak{S}$ und $\mathfrak{T}$) und wird als $\mathfrak{S} \oplus \mathfrak{T}$ notiert. Zu vorgegebenem $\mathfrak{S}$ heißt jedes solche $\mathfrak{T}$ ein *Komplement* von $\mathfrak{S}$ in $\mathfrak{V}$.

Bemerkung 3.5.2.

Es seien $\mathfrak{S}$ und $\mathfrak{T}$ Unterräume von $\mathfrak{V}$. Dann gilt $\mathfrak{V} = \mathfrak{S} \oplus \mathfrak{T}$ genau dann, wenn jeder Vektor $v \in \mathfrak{V}$ sich eindeutig *in der Form $v = s + t$ mit $s \in \mathfrak{S}$ und $t \in \mathfrak{T}$ schreiben läßt.*

[5] Natürlich war das nicht dumm! Das haben wir aus didaktischen Gründen gemacht.

Beweis: Ausgehend von $\mathfrak{V} = \mathfrak{S} \oplus \mathfrak{T}$ ist nur die Eindeutigkeit zu zeigen: Hat man $s_1 + t_1 = s_2 + t_2$ für $s_\nu \in \mathfrak{S}$ und $t_\nu \in \mathfrak{T}$, dann gilt $s_1 - s_2 = t_2 - t_1 \in \mathfrak{S} \cap \mathfrak{T} = \{0\}$, also $s_1 = s_2$ und $t_1 = t_2$.
In der anderen Richtung ist nur $\mathfrak{S} \cap \mathfrak{T} = \{0\}$ zu zeigen: Für $v \in \mathfrak{S} \cap \mathfrak{T}$ gelten

$$\mathfrak{V} \ni 0 = \underbrace{0}_{\in \mathfrak{S}} + \underbrace{0}_{\in \mathfrak{T}} \quad \text{und} \quad 0 = \underbrace{v}_{\in \mathfrak{S}} + \underbrace{(-v)}_{\in \mathfrak{T}} ,$$

also $v = 0$. □

Bemerkung 3.5.3.

Es seien $\mathfrak{S}$ und $\mathfrak{T}$ Unterräume des endlich-dimensionalen *Vektorraums $\mathfrak{V}$. Dann sind die folgenden Aussagen äquivalent:*

a) $\mathfrak{V} = \mathfrak{S} \oplus \mathfrak{T}$

b) $\mathfrak{V} = \mathfrak{S} + \mathfrak{T}$ *und* $\dim \mathfrak{V} = \dim \mathfrak{S} + \dim \mathfrak{T}$

c) $\mathfrak{S} \cap \mathfrak{T} = \{0\}$ *und* $\dim \mathfrak{V} = \dim \mathfrak{S} + \dim \mathfrak{T}$

Beweis:
a) $\Rightarrow$ *b)*: Nach Dimensionsformel (3.3.9)
b) $\Rightarrow$ *c)*: Die Dimensionsformel liefert $\dim(\mathfrak{S} \cap \mathfrak{T}) = 0$, also $\mathfrak{S} \cap \mathfrak{T} = \{0\}$.
c) $\Rightarrow$ *a)*: $\dim(\mathfrak{S} + \mathfrak{T}) \underset{(3.3.9)}{=} \dim \mathfrak{S} + \dim \mathfrak{T} \underset{\text{Vor.}}{=} \dim \mathfrak{V}$, also $\mathfrak{V} = \mathfrak{S} + \mathfrak{T}$ (nach Folgerung 3.3.7) . □

Später betrachten wir auch direkte Summen aus mehr als zwei Unterräumen, was aber keine zusätzliche Schwierigkeit nach sich zieht.

Einschub: Die natürlichen Zahlen $\mathbb{N}$

Der Mathematiker L. Kronecker pflegte zu sagen: *„Die natürlichen Zahlen hat der liebe Gott geschaffen, alles andere ist Menschenwerk.“*

Kennt man die Menge $\mathbb{N}$ der natürlichen Zahlen schon, so wird man als eine wichtige und kennzeichnende Eigenschaft ansehen, daß 1 eine natürliche Zahl ist und innerhalb dieser Menge immer um 1 weitergezählt werden kann, mithin

$$(*) \qquad 1 \in \mathbb{N} \quad \text{und} \quad n \in \mathbb{N} \Longrightarrow n + 1 \in \mathbb{N}$$

gilt, und die Menge in diesem Sinne minimal ist.
Zunächst haben dann aber ein irgendwie schon gegebenes $\mathbb{N}$ und die (in der Analysis) axiomatisch eingeführte Menge der reellen Zahlen $\mathbb{R}$ wenig miteinander zu tun.
Falls man die natürlichen Zahlen *nicht* schon als gegeben ansehen will, so läßt sie sich innerhalb $\mathbb{R}$ wie folgt präzise definieren:
Teilmengen von $\mathbb{R}$, die die obige Eigenschaft $(*)$ haben, nennt man *‚induktiv‘* und führt dann $\mathbb{N}$ als ‚kleinste‘ induktive Menge ein.

Vielen wird dieses Vorgehen vermutlich äußerst umständlich erscheinen — und sie haben damit ja auch irgendwie recht! Doch haben Sie etwas Nachsicht mit den ach so ‚pingeligen' Mathematikern, die die Grundlagen gerne ‚gesichert' haben. Wenn nicht, dann überschlagen Sie die nächsten Zeilen bis zum Prinzip der vollständigen Induktion!

Definition

Eine Teilmenge M von $\mathbb{R}$ heiße genau dann *„induktiv"*, wenn gilt:
$1 \in M \quad \text{und} \quad a \in M \implies a+1 \in M$

Offenbar gibt es induktive Mengen, zum Beispiel $\mathbb{R}$ selbst und die Menge der positiven reellen Zahlen.

Definition

$\mathbb{N} := \{x \in \mathbb{R} : x \text{ gehört zu jeder induktiven Teilmenge von } \mathbb{R}\}$
Die Elemente aus $\mathbb{N}$ heißen *„natürliche Zahlen"*.

$\mathbb{N}$ ist dann offenbar die *kleinste induktive Teilmenge von* $\mathbb{R}$.

Einige einfache Eigenschaften von $\mathbb{N}$ sind scheinbar evident, erfordern aber bei diesem Zugang eigentlich einen Beweis. Dies führen wir jedoch nicht mehr aus und freuen uns, daß wir solche lästigen Arbeiten den Mathematikern überlassen dürfen.
Mit den *Bezeichnungen*

$2 := 1+1, \quad 3 := 2+1, \quad 4 := 3+1, \quad \ldots$

gilt dann $\mathbb{N} = \{1, 2, 3, 4, \ldots\}$

Die Tatsache, daß $\mathbb{N}$ die *kleinste* induktive Teilmenge von $\mathbb{R}$ ist, ergibt sofort die

Folgerung $T \subset \mathbb{N} \quad \wedge \quad T \quad \textit{induktiv} \quad \implies \quad T = \mathbb{N}.$

Vollständige Induktion.

Es sei nun $A(n)$ eine Aussageform $(n \in \mathbb{N})$.
Mit $T := \{n \in \mathbb{N} : A(n)\}$ liefert dann die Folgerung das wichtige

Prinzip der vollständigen Induktion:

$$\Big[A(1) \quad \wedge \quad \forall k \in \mathbb{N} \quad \big(A(k) \implies A(k+1) \big) \Big] \implies \forall n \in \mathbb{N}\ A(n)$$

In Worten: *Gilt* $A(1)$ *und folgt aus der Gültigkeit* $A(k)$ *für eine natürliche Zahl* k *stets die Gültigkeit* $A(k+1)$ *für die nachfolgende Zahl* $k+1$*, so gilt die Aussage* $A(n)$ *für alle natürlichen Zahlen* n*.*

Üblicherweise bezeichnet man hierbei $A(1)$ als *„Induktionsanfang"* oder auch *„Induktionsverankerung"*, den Schluß $\big(A(k) \implies A(k+1)\big)$ als *„Induktionsschritt"* und dabei $A(k)$ als *„Induktionsannahme"*.

Die vollständige Induktion ist nur eine Beweismethode, sie zeigt in der Regel *nicht*, wie man die entsprechende Behauptung findet! Ein schönes Standard-Beispiel, durch das die der vollständigen Induktion zugrundeliegende Idee besonders klar wird, ist das folgende: Stellt man **Dominosteine** nebeneinander hochkant, parallel und mit den Breitseiten zueinander gewandt auf, so kann man *alle* dadurch zum Umfallen bringen, daß man — in einer vorher festgelegten Richtung — den ersten Stein umwirft und bei der Aufstellung beachtet, daß der Abstand von einem Stein zum nächsten etwa immer kleiner als die halbe Länge eines Steines ist. Das Umwerfen des ersten Steines *(„Induktionsanfang“)* bewirkt, daß der Prozeß überhaupt in Gang kommt. Die Bedingung über den Abstand sichert, daß jeder fallende Stein seinen ‚Nachfolger‘ mit umwirft *(„Induktionsschritt“)*.

Wir sehen uns zwei erste *Beispiele* zur vollständige Induktion an:

Beispiele

(B1) $\forall\, n \in \mathbb{N} \qquad 1 + 2 + \cdots + n \;=\; \frac{1}{2}\, n\,(n+1)$

Zum *Beweis* bezeichnen wir für $n \in \mathbb{N}$ die entsprechende Aussage mit $A(n)$. $A(1):\quad \ell.S. = 1, \quad r.S. = \frac{1}{2}\cdot 1 \cdot 2 = 1$

$A(k) \Longrightarrow A(k{+}1):\quad 1{+}2{+}\cdots{+}k{+}(k{+}1) \underset{A(k)}{=} \frac{1}{2}k\,(k{+}1) + (k{+}1)$

$= (\frac{1}{2}k \,+\, 1)(k+1) \;=\; \frac{1}{2}(k+2)(k+1) \;=\; \frac{1}{2}(k+1)\big((k+1)+1\big)\,\square$

Natürlich ist auch mir bekannt, daß man diese Aussage auf andere Weise — wie schon der zehnjährige Gauss seinem Lehrer vormachte — einfacher gewinnen kann; doch hier wollen wir ja stattdessen gerade das Prinzip der vollständigen Induktion einüben.

(B2) $\forall\, n \in \mathbb{N} \qquad 1 + 3 + 5 + \cdots + (2n-1) \;=\; n \cdot n$

Man kann die Aussage von (B2) recht einfach aus (B1) herleiten; doch auch hier soll der Beweis ja stattdessen ein eigenständiges Beispiel zur vollständigen Induktion sein:

Wir bezeichnen wieder zum *Beweis* für $n \in \mathbb{N}$ die entsprechende Aussage mit $A(n)$. $A(1):\quad \ell.S. \,=\, 1 \,=\, r.S.$

$A(k) \Longrightarrow A(k+1):\; 1 + 3 + 5 + \cdots + (2k-1) + \big(2(k+1)-1\big)$

$\underset{A(k)}{=} k \cdot k \,+\, 2k \,+\, 1 \;=\; (k+1)\cdot(k+1)$ $\square$

$\triangleleft$

Bei den Dominosteinen ist unmittelbar klar, daß entsprechend nur alle Steine ab dem 7. umfallen, wenn man statt des ersten den 7. umstößt.

So kann auch allgemein an die Stelle von 1 eine beliebige natürliche Zahl oder auch 0 als Startelement ℓ für den Induktionsanfang treten, wobei dann natürlich die Behauptung entsprechend nur für alle $n \geq \ell$ gilt. Man erhält das

Modifiziertes Prinzip der vollständigen Induktion

$$\boxed{\Big[A(\ell) \quad \wedge \quad \forall k \in \mathbb{N}_\ell \quad \big(A(k) \implies A(k+1)\big)\Big] \implies \forall n \in \mathbb{N}_\ell \;\; A(n)}$$

dadurch, daß man es mit $B(n) := A(n+\ell-1) \quad (n \in \mathbb{N})$ unmittelbar auf die spezielle Version zurückführt.
Statt der ausführlichen Formulierung *„Beweis durch vollständige Induktion“* sagen wir — auch im modifizierten Fall — oft kürzer *„induktiver Beweis“* und ähnlich.

Gelegentlich höre ich von Studenten, die die ersten Beispiele für Beweise durch vollständige Induktion mit ungläubigem Staunen gesehen haben, daß man *damit* ja wohl *alles* beweisen könne. Um die Verwirrung voll zu machen, bringe ich das folgende Beispiel einer offensichtlich falschen Aussage, die ich induktiv ‚beweise‘:

Alle reellen Zahlen sind gleich. Nanu! Nanu?

‚Beweis‘: Es genügt offenbar, die folgende reduzierte Aussage zu zeigen: Für jedes $n \in \mathbb{N}$ gilt: *Je n reelle Zahlen sind gleich.*
Dies ‚beweisen‘ wir durch vollständige Induktion: Für $n = 1$ ist die Aussage sicher richtig. Der Schluß von k auf $k+1$ ergibt sich für beliebiges $k \in \mathbb{N}$ wie folgt: Hat man $k+1$ reelle Zahlen $r_1, r_2, \ldots, r_k, r_{k+1}$, so sind die ersten k dieser Zahlen $r_1, r_2, \ldots, r_k$ nach Induktionsvoraussetzung gleich und ebenso die k letzten $r_2, \ldots, r_k, r_{k+1}$. Dann sind aber auch alle $k+1$ Zahlen gleich:

$$\overbrace{r_1, \underbrace{r_2, \ldots, r_k}, r_{k+1}}^{\text{alle gleich}} \qquad \text{alle gleich} \qquad \square$$

Wo liegt der Fehler???

Sie haben natürlich sofort erkannt — auch ohne hier weitergelesen zu haben, daß der *Fehler im Schluß von 1 auf 2* liegt. Denn das oben so suggestiv aussehende Klammern klappt dann offensichtlich nicht, weil in diesem Fall kein Element mehr zu *beiden* Teilmengen gehört! Es ist natürlich andererseits auch klar, daß man damit alles bewiesen hätte: Wären je zwei Zahlen schon gleich, dann müßten auch alle Zahlen untereinander gleich sein.
Ich habe dieses Beispiel auch hier aufgeführt, um deutlich zu machen, daß es ganz wichtig ist, darauf zu achten, daß der Induktionsschritt von k auf $k+1$ wirklich für *jedes* k funktioniert! Auch das sieht man schon sehr schön an den Dominosteinen: Wenn nur an einer einzigen Stelle der Abstand zwischen zwei benachbarten Steinen zu groß ist, wird der gesamte Prozeß dort unterbrochen und die nachfolgenden Steine fallen nicht mehr um (wenn die Unterlage nicht gerade zu wackelig ist und ...).

3.6. **Gruppen.**

Der Begriff der „Gruppe“ ist von fundamentaler Bedeutung in der gesamten Mathematik — und nicht nur dort. Wir sind in Bemerkung 3.4.20 auf eine erste wichtige Gruppe gestoßen, die *nicht* kommutativ ist. Deshalb müssen wir uns mit *allgemeinen* Gruppen kurz beschäftigen.

Definition 3.6.1 (Gruppe)**.**

Eine Gruppe ist ein Paar[a] $(\mathfrak{G}, v)$, bestehend aus einer nicht-leeren Menge $\mathfrak{G}$ und einer Abbildung (Verknüpfung)

$$\begin{array}{rccc} v\colon & \mathfrak{G}\times\mathfrak{G} & \longrightarrow & \mathfrak{G} \\ & \cup & & \cup \\ & (a,b) & \longmapsto & a\cdot b =: ab \end{array}$$

derart, daß die folgenden Gesetze gelten:

(G1) $\forall\, a,b,c \in \mathfrak{G} \quad (a\cdot b)\cdot c = a\cdot(b\cdot c) =: a\cdot b\cdot c =: abc$ *„Assoziativität“*

(G2) $\exists\, e \in \mathfrak{G}\ \forall a \in \mathfrak{G}\ e\cdot a = a$ *„neutrales Element“*

(G3) $\forall\, a \in \mathfrak{G}\ \exists\, x \in \mathfrak{G}\ x\cdot a = e$ *„inverses Element zu a“*

[a] Ist aus dem Zusammenhang heraus klar, welche Verknüpfung v gemeint ist, dann sprechen wir auch oft (unpräzise) von der Gruppe $\mathfrak{G}$.

(G4) Gilt *zusätzlich* $a\cdot b = b\cdot a$ für alle $a, b \in G$, so heißt die Gruppe *kommutativ* oder *abelsch.*

Die Eigenschaft (G2) bedeutet: Es exisitiert (mindestens) ein *„Linkseinselement“* e. Die Eigenschaft (G3) besagt: Zu e existiert für jedes $a \in \mathfrak{G}$ (mindestens) ein *„Linksinverses“*.

Bemerkung 3.6.2.

Es sei $(\mathfrak{G}, \cdot)$ eine Gruppe mit einem Linkseinselement e. Dann gelten:

(1) $ca = cb \Longrightarrow a = b$ *(Linke Kürzungsregel)*

(2) $ac = bc \Longrightarrow a = b$ *(Rechte Kürzungsregel)*

(3) Für alle $a \in \mathfrak{G}$ ist $ae = a$.

(4) Es gibt in $\mathfrak{G}$ genau ein (Links-) Einselement.

(5) Für $a \in \mathfrak{G}$ und $x \in \mathfrak{G}$ mit $xa = e$ gilt auch $ax = e$.

(6) Für alle $a \in \mathfrak{G}$ ist das (Links-) Inverse eindeutig bestimmt. Wir notieren es mit a^{-1}.

(7) $\forall a, b \in \mathfrak{G}\ \dot\exists\, y \in \mathfrak{G}\ ay = b \wedge \dot\exists\, z \in \mathfrak{G}\ za = b$

(8) $(a^{-1})^{-1} = a, \quad (ab)^{-1} = b^{-1}a^{-1}$ $(\dots)$

Mit dem Zeichen $\dot\exists$ notieren wir die eindeutige Existenz.
a^{-1} lesen wir wieder oft als *„a hoch minus 1“*.

(3) besagt: Ein Linkseinselement ist auch Rechtseinselement.

(5) besagt: Ein Linksinverses zu a ist auch Rechtsinverses.

Nach (3) und (4) können wir von *dem Einselement* (oder neutralem Element) sprechen. Nach (5) und (6) können wir entsprechend von *dem Inversen* (zu einem $a \in \mathfrak{G}$) sprechen.

Die Beweise sind — bei fehlender Kommutativität — meist aufwendiger als im kommutativen Fall. Sie erfordern jedenfalls deutlich mehr Sorgfalt:

Beweise:

(1): Es existiert ein $y \in \mathfrak{G}$ mit $yc = e$:
$a = ea = (yc)a = y(ca) = y(cb) = (yc)b = eb = b$

(3): Zu $a \in \mathfrak{G}$ existiert ein $x \in \mathfrak{G}$ mit $xa = e$:
$x(ae) = (xa)e = ee = e = xa$; nach (1) also $ae = a$.

(4): *Falls* $f \in \mathfrak{G}$ mit $fa = a$ für alle $a \in \mathfrak{G}$: $f \underset{(3)}{=} fe = e$.

(5): $x(ax) = (xa)x = ex = x \underset{(3)}{=} xe$, nach (1) also $ax = e$

(2): Es existiert ein $y \in \mathfrak{G}$ mit $cy = e$ ((5) beachten):
$a \underset{(3)}{=} ae = a(cy) = (ac)y = (bc)y = b(cy) = be \underset{(3)}{=} b$

(6): Falls $a \in \mathfrak{G}$ und $x_1, x_2 \in \mathfrak{G}$ mit $x_1 a = e = x_2 a$:
Nach (2): $x_1 = x_2$

(7): Eindeutigkeit: nach (1) beziehungsweise (2)
Existenz: $y := a^{-1}b, z := ba^{-1}$

(8): $a^{-1}a = e = a^{-1}\left(a^{-1}\right)^{-1}$, nach (1) also: $a = \left(a^{-1}\right)^{-1}$
$\left(b^{-1}a^{-1}\right)(ab) = b^{-1}\left(a^{-1}a\right)b = b^{-1}eb = b^{-1}b = e$
Nach (6) folgt: $b^{-1}a^{-1} = (ab)^{-1}$ □

Definition 3.6.3.

Ist $(\mathfrak{G}, v)$ eine Gruppe mit Einselement e, dann bezeichnen wir — bei multiplikativer Schreibweise der Verknüpfung — :
$a^0 := e, a^1 := a$, $a^{n+1} := a^n a$ und $a^{-n} := \left(\left(a^{-1}\right)^n \underset{(8)}{=} \left(a^n\right)^{-1}\right) \quad (n \in \mathbb{N})$

Definition 3.6.4 (Untergruppe)**.**

Es sei $\mathfrak{G}$ eine Gruppe. Eine nicht-leere Teilmenge $\mathfrak{U}$ von $\mathfrak{G}$ heißt genau dann *Untergruppe* von $\mathfrak{G}$, wenn $\mathfrak{U}$ mit der eingeschränkten Verknüpfung eine Gruppe ist.

Bemerkung 3.6.5.

Eine nicht-leere Teilmenge $\mathfrak{U}$ *einer Gruppe* $\mathfrak{G}$ *ist genau dann* Untergruppe*, wenn gilt:*

$$\forall a, b \in \mathfrak{U} \quad ab^{-1} \in \mathfrak{U}$$

Beweis: Ist $\mathfrak{U}$ Untergruppe, dann gilt notwendigerweise die untere Aussage. Ausgehend von der unteren Beziehung hat man mit einem $a \in \mathfrak{U}$: $e = aa^{-1} \in \mathfrak{U}$, dann für ein $b \in \mathfrak{U}$ auch $b^{-1} = eb^{-1} \in \mathfrak{U}$ und somit $ab = a\,(b^{-1})^{-1} \in \mathfrak{U}$. Die Nicht-Existenz-Eigenschaft (G1) gilt trivialerweise in $\mathfrak{U}$, da sie ja in ganz $\mathfrak{G}$ gilt. □

Beispiele

(B1) $\{1, i, -1, -i\}$ bilden eine Untergruppe der multiplikativen Gruppe $\mathbb{C} \setminus \{0\}$ der komplexen Zahlen.

(B2) Die geraden Zahlen bilden eine additive Untergruppe der ganzen Zahlen $\mathbb{Z}$. ◁

Wir tragen den noch ausstehenden (einfachen) Beweis der Bemerkung 3.4.20 nach: *$(\mathrm{GL}(\mathfrak{V}), \circ)$ ist eine Gruppe.*

Beweis: Daß die Hintereinanderausführung $\circ$ von Abbildungen assoziativ ist, wissen wir ganz allgemein. Daß $f \circ g$ für $f, g \in \mathrm{GL}(\mathfrak{V})$ wieder zu $\mathrm{GL}(\mathfrak{V})$ gehört, liefert Bemerkung 3.4.14. Das Einselement ist hier natürlich $\mathrm{id}_{\mathfrak{V}}$. Nach Bemerkung 3.4.13 gilt $f^{-1} \in \mathrm{GL}(\mathfrak{V})$ für jedes $f \in \mathrm{GL}(\mathfrak{V})$. □

In diesem Abschnitt könnte noch sehr viel über Gruppen ergänzt werden. Ich stelle einige Überlegungen zurück, bis wir sie dann später benötigen.

3.7. Die allgemeine lineare Gruppe.

Es seien $n \in \mathbb{N}$ und $\mathbb{K}$ ein Körper.

In (B2) aus Abschnitt 3.4 hatten wir u. a. gesehen, daß jedes $A \in \mathbb{K}^{m\times n}$ eine Abbildung $f_A \in \mathcal{L}(\mathbb{K}^n, \mathbb{K}^m)$ induziert. Wir zeigen die ‚Umkehrung' dazu:

Bemerkung 3.7.1.
Zu jedem $f \in \mathcal{L}(\mathbb{K}^n, \mathbb{K}^m)$ existiert eindeutig ein $A \in \mathbb{K}^{m\times n}$ mit

$$f = f_A\,.$$

Beweis: Ist $f = f_A$, so wissen wir schon, daß $f(e_1), \ldots, f(e_n)$ gerade die Spalten von A sind. Es existiert also *höchstens* ein solches A. Bildet man andererseits eine Matrix mit den Spalten $f(e_1), \ldots, f(e_n)$ (in dieser Reihenfolge!), dann hat man offenbar ein solches A; denn mit $A := (A_1, \ldots, A_n) := \big(f(e_1), \ldots, f(e_n)\big)$ gilt

$$f(x) = f\left(\sum_{\nu=1}^{n} x^\nu e_\nu\right) = \sum_{\nu=1}^{n} x^\nu f(e_\nu) = \sum_{\nu=1}^{n} x^\nu A_\nu = Ax = f_A(x)$$

für $x = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} \in \mathbb{K}^n$. □

Bemerkung 3.7.2.

Für ein $A \in \mathbb{K}^{n\times n}$ *sind* äquivalent:

a) $\exists B \in \mathbb{K}^{n\times n} \quad BA = \mathbb{1}$

b) f_A *ist surjektiv.*

c) f_A *ist injektiv.*

d) f_A *ist bijektiv.*

e) $\dot{\exists}\ B \in \mathbb{K}^{n\times n} \quad BA = \mathbb{1} = AB$

Beweis:

a) $\Longrightarrow$ *c)*: $\mathrm{id}_{\mathbb{K}^n} = f_{\mathbb{1}} = f_{BA} = f_B \circ f_A$: Somit ist f_A injektiv.

Die Äquivalenz von *b)*, *c)* und *d)* wird in Übungsaufgabe 5.1.a gezeigt.

Das hätte ich doch fast — ganz aus Versehen — hier bewiesen. ☺

d) $\Longrightarrow$ *e)*: Zu der Umkehrabbildung $(f_A)^{-1}$ existiert ein $B \in \mathbb{K}^{n\times n}$ mit $(f_A)^{-1} = f_B$. $f_{AB} = f_A \circ f_B = \mathrm{id}_{\mathbb{K}^n}$ liefert $AB = \mathbb{1}$. Aus $f_{BA} = f_B \circ f_A = \mathrm{id}_{\mathbb{K}^n}$ folgt $BA = \mathbb{1}$. *Eindeutigkeit:* Aus $BA = \mathbb{1} = AC$ folgt $B = C$; denn $B = B\mathbb{1} = B(AC) = (BA)C = \mathbb{1}C = C$.

e) $\Longrightarrow$ *a)*: trivial □

Definition 3.7.3.

Erfüllt eine Matrix $A \in \mathbb{K}^{n\times n}$ eine — und damit alle — Bedingungen aus Bemerkung 3.7.2, dann nennen wir A *invertierbar*, *umkehrbar* oder *regulär*. Das dazu eindeutig existierende $B \in \mathbb{K}^{n\times n}$ mit

$$AB = \mathbb{1} = BA$$

zu A heißt *inverse Matrix* (zu A). Wir notieren diese inverse Matrix als A^{-1}.

Definition 3.7.4.

$$\mathrm{GL}(n,\mathbb{K}) := \{A \in \mathbb{K}^{n\times n} \mid A \text{ invertierbar}\}$$

Bemerkung 3.7.5.

Mit der Matrixmultiplikation $\cdot$ gilt:

$\big(\mathrm{GL}(n,\mathbb{K}), \cdot\big)$ *ist eine Gruppe.*

Sie heißt wieder *„general linear group“* beziehungsweise *„Allgemeine lineare Gruppe“*.

Beweis: Die Zurückführung auf Bemerkung 3.4.20 via

$$\mathrm{GL}(n,\mathbb{K}) \ni A \longmapsto f_A \in \mathrm{GL}(\mathbb{K}^n)$$

ist leicht möglich. Doch die Dinge sieht man hier auch ‚direkt‘ ganz einfach: Die *Assoziativität* der Matrixmultiplikation haben wir schon vor Urzeiten (vgl. Lemma 1.3.5) gezeigt. Die Matrix $\mathbb{1}$ gehört zu $\mathrm{GL}(n,\mathbb{K})$ und ist Einselement. Zu $A \in \mathrm{GL}(n,\mathbb{K})$ existiert — nach Definition von

$\mathrm{GL}(n,\mathbb{K})$ — A^{-1} und gehört selbst zu $\mathrm{GL}(n,\mathbb{K})$. Zu $A, B \in \mathrm{GL}(n,\mathbb{K})$ gehört auch AB zu $\mathrm{GL}(n,\mathbb{K})$; denn $B^{-1}A^{-1}$ ist offenbar die inverse Matrix zu AB. □

Transponierte Matrix

Für manche Überlegungen ist noch die transponierte Matrix A^T zu einer gegebenen Matrix A von Interesse: Es seien wieder $\mathbb{K}$ ein Körper und — mit $n, m \in \mathbb{N}$ — A eine $(m \times n)$-Matrix mit Elementen aus $\mathbb{K}$, also $A \in \mathbb{K}^{m\times n}$.

Die für $A = (a^\mu_\nu)_{\substack{1\le\mu\le m\\ 1\le\nu\le n}}$ durch $A^T := (b^\nu_\mu)_{\substack{1\le\nu\le n\\ 1\le\mu\le m}}$ mit $b^\nu_\mu := a^\mu_\nu$ definierte Matrix heißt *transponierte Matrix* (zu A). ‚Ausführlicher' :

$$A^T = \underbrace{\begin{pmatrix} a^1_1 & a^2_1 & \dots & a^m_1 \\ a^1_2 & a^2_2 & \dots & a^m_2 \\ \vdots & \vdots & & \vdots \\ a^1_n & a^2_n & \dots & a^m_n \end{pmatrix}}_{\in\mathbb{K}^{n\times m}} \quad \text{aus} \quad A = \underbrace{\begin{pmatrix} a^1_1 & a^1_2 & \dots & a^1_n \\ a^2_1 & a^2_2 & \dots & a^2_n \\ \vdots & \vdots & & \vdots \\ a^m_1 & a^m_2 & \dots & a^m_n \end{pmatrix}}_{\in\mathbb{K}^{m\times n}} .$$

Man findet auch die Notierungsweisen A^t und auch ${}^tA$.
Aus den Spalten von A werden — bei gleicher Reihenfolge — die Zeilen von A^T, und aus den Zeilen von A werden die Spalten von A^T. Insbesondere wird aus einem Zeilenvektor ein Spaltenvektor und umgekehrt.

Beispielchen

(B1) $A = \begin{pmatrix} 2 & 3 & 4 \\ 7 & 0 & 1 \end{pmatrix}$ liefert $A^T = \begin{pmatrix} 2 & 7 \\ 3 & 0 \\ 4 & 1 \end{pmatrix}$.

(B2) Zu $x = \begin{pmatrix} 1 & 9 & 4 & 3 \end{pmatrix}$ gehört der Spaltenvektor $x^T = \begin{pmatrix} 1 \\ 9 \\ 4 \\ 3 \end{pmatrix}$. ◁

Bemerkung 3.7.6. *Für $k \in \mathbb{N}, A, B \in \mathbb{K}^{m\times n}, C \in \mathbb{K}^{n\times k}$ und $\alpha \in \mathbb{K}$:*

a) $\mathbb{1}_n^T = \mathbb{1}_n$
b) $(\alpha\, A)^T = \alpha\, A^T$
c) $(A+B)^T = A^T + B^T$
d) $(AC)^T = C^T A^T$
e) $\left(A^T\right)^T = A$
f) $A \in \mathrm{GL}(n,\mathbb{K}) \implies A^T \in \mathrm{GL}(n,\mathbb{K}) \wedge \left(A^T\right)^{-1} = (A^{-1})^T$

Beweis:
a): $\delta^\mu_\nu = \delta^\nu_\mu$ $(\dots)$
b), c), e): ✓

d): Für $\kappa = 1, \ldots, k$ und $\mu = 1, \ldots, m$ ist $\sum_{\nu=1}^{n} a_\nu^\mu \, c_\kappa^\nu$ das Element in der μ-ten Zeile und κ-ten Spalte von AC gerade das Element in der κ-ten Zeile und μ-ten Spalte von $(AC)^T$. Dies ist offenbar gerade das entsprechende Element von $C^T A^T$.

f): $A^{-1} A = \mathbb{1}_n \implies \mathbb{1}_n = \mathbb{1}_n^T = (A^{-1} A)^T = A^T (A^{-1})^T$ □

4. Lineare Abbildungen und Matrizen

Beide Begriffe kamen schon vielfach und ausgiebig vor. In diesem Abschnitt sollen nun noch wesentliche Ergänzungen zu den bisherigen Überlegungen betrachtet werden.

Es seien $\mathfrak{U}, \mathfrak{V}$ und $\mathfrak{W}$ drei Vektorräume über einem Körper $\mathbb{K}$.

Wir erinnern nur noch einmal an die zentrale **Definition** einer linearen Abbildung:

Eine Abbildung $f\colon \mathfrak{U} \longrightarrow \mathfrak{V}$ heißt genau dann *linear* — genauer auch *$\mathbb{K}$-linear* — oder *Homomorphismus* wenn
$f(a+b) = f(a) + f(b)$ für alle $a, b \in \mathfrak{U}$ (*additiv*) und
$f(\lambda a) = \lambda f(a)$ für alle $a \in \mathfrak{U}, \lambda \in \mathbb{K}$ (*homogen*) gelten.

Uns sind zwar schon sehr viele Beispiele vertraut. Doch können zwei weitere sicher nicht schaden:

Beispiele

(B1) Es sei $\mathfrak{W} = \mathfrak{U} \oplus \mathfrak{V}$. Definiere $p_{\mathfrak{U}}\colon \mathfrak{W} \longrightarrow \mathfrak{W}$, die *Projektion* auf $\mathfrak{U}$, wie folgt: Zu $w \in \mathfrak{W}$ gibt es genau ein $u \in \mathfrak{U}$ und genau ein $v \in \mathfrak{V}$ mit $w = u + v$. Setze $p_{\mathfrak{U}}(w) := u$. $p_{\mathfrak{U}}$ ist eine lineare Abbildung.

(B2) $C^0([0,1]) \ni f \longmapsto \int_0^1 f(x)\,dx \in \mathbb{R}$ ist linear. ◁

4.1. Erzeugung linearer Abbildungen.

Eine lineare Abbildung ist durch ihre Werte auf einer Basis schon festgelegt:

Satz 4.1.1.

Es seien I eine beliebige Menge (‚Indexmenge'), $(u_i)_{i\in I}$ eine Basis von $\mathfrak{U}$ und $v_i \in \mathfrak{V}$ für $i \in I$ beliebig vorgegeben. Dann gibt es genau eine lineare Abbildung

$$f\colon \mathfrak{U} \longrightarrow \mathfrak{V}$$

mit $f(u_i) = v_i$ für alle $i \in I$. Für dieses f gilt:

a) $f(\mathfrak{U}) = \big\langle \{v_i \mid i \in I\} \big\rangle$

b) f ist genau dann injektiv, *wenn $\big(v_i \mid i \in I\big)$ linear unabhängig ist.*

c) f ist genau dann ein Isomorphismus, *wenn $\big(v_i \mid i \in I\big)$ eine Basis ist.*

Beweis:
Ein beliebiger Vektor $u \in \mathfrak{U}$ läßt sich mit einer endlichen Teilmenge J von I und geeigneten $\lambda^i \in \mathbb{K}$ ‚eindeutig' darstellen in der Form

$$u \;=\; \sum_{i\in J} \lambda^i u_i \,.$$

Falls f linear ist mit $f(u_i) = v_i$, muß

$$f(u) \;=\; \sum_{i\in J} \lambda^i v_i$$

gelten. Es kann also *höchstens ein solches* f geben. Zum Nachweis der *Existenz* ist somit aber auch das Vorgehen klar. Die erhaltene Beziehung ziehen wir zur Definition heran:

$$f(u) \;:=\; \sum_{i\in J} \lambda^i v_i$$

Dadurch ist f wohldefiniert, da $(u_i)_{i\in I}$ eine Basis ist. Die so definierte Abbildung f ist *linear:* Œ können wir für den Nachweis der Linearität von $a = \sum_{i\in J} \alpha^i u_i$ und $b = \sum_{i\in J} \beta^i u_i$ (beide Vektoren mit denselben Basisvektoren darstellen) und $\lambda \in \mathbb{K}$ ausgehen:

$$\begin{aligned} f(\lambda a + b) = f\left(\lambda \sum_{i\in J} \alpha^i u_i + \sum_{i\in J} \beta^i u_i\right) &= f\left(\sum_{i\in J} \left(\lambda\alpha^i + \beta^i\right) u_i\right) \\ = \sum_{i\in J} \left(\lambda\alpha^i + \beta^i\right) v_i = \lambda \sum_{i\in J}^{n} \alpha^i v_i + \sum_{i\in J} \beta^i v_i \;&=\; \lambda f(a) + f(b)\,. \end{aligned}$$

a): ✓
b): Die lineare Unabhängigkeit von $(\, v_i \mid i \in I \,)$, wenn f injektiv ist, liefert die Bemerkung 3.4.8. In der anderen Richtung müssen wir nach Bemerkung 3.4.10 zeigen, daß der Kern von f nur aus der Null besteht: Ist $f(u) = 0$ für ein $u \in \mathfrak{U}$, dann gilt mit einer Darstellung für u wie oben

$$0 \;=\; f(u) \;=\; \sum_{i\in J} \lambda^i v_i \,.$$

Damit gilt $\lambda^i = 0$ für $i \in J$, also $u = 0$.
c): folgt unmittelbar aus *a*) und *b*). □

Folgerung 4.1.2.
Es seien $(u_i)_{i\in I}$ *eine linear unabhängige Familie in* $\mathfrak{U}$ *und* $v_i \in \mathfrak{V}$ *für* $i \in I$ *beliebig. Dann gibt es eine lineare Abbildung* $f\colon \mathfrak{U} \longrightarrow \mathfrak{V}$ *mit* $f(u_i) = v_i$ *für alle* $i \in I$.

Beweis:
Ergänze $(u_i)_{i\in I}$ zu einer Basis von $\mathfrak{U}$, wähle die ‚restlichen' v_i's beliebig und wende den Satz 4.1.1 an. □

4.2. **Dimensionsformel für Homomorphismen.**

Satz 4.2.1 (Dimensionsformel für Homomorphismen)**.**
Es seien $f\colon \mathfrak{U} \longrightarrow \mathfrak{V}$ *eine lineare Abbildung und* $\dim \mathfrak{U} = n \in \mathbb{N}_0$. *Dann gilt:*

$$\boxed{\dim(\ker f) + \dim(\operatorname{im} f) \;=\; \dim \mathfrak{U}}$$

Wir bezeichnen die Dimension des Kernes von f auch als *1. Defekt*, also

$$\operatorname{def}_1 f \;:=\; \dim(\ker f)\,.$$

Dieser Defekt mißt, wieviel an Dimension unter der Abbildung f ‚verlorengeht'.

Beweis:
Œ $\mathfrak{U} \neq \{0\}$, also $n \in \mathbb{N}$. Es seien $r := \dim(\ker f)$ und $B = (u_1, \ldots, u_r)$, falls $r \in \mathbb{N}$, beziehungsweise $B = \emptyset$, falls $r = 0$, eine Basis von $\ker f$. B kann nach dem Basisergänzungssatz (vgl. (c) aus Satz 3.2.7) durch $R = (u_{r+1}, \ldots, u_n)$ — beziehungsweise $R = \emptyset$, falls $r = n$— zu einer Basis von $\mathfrak{U}$ ergänzt werden. Nach $a)$ aus Satz 4.1.1 wird das Bild von f von

$$f(u_1) = 0, \ldots, f(u_r) = 0, f(u_{r+1}), \ldots, f(u_n)$$

erzeugt. Wir zeigen, daß die Vektoren $f(u_{r+1}), \ldots, f(u_n)$ linear unabhängig sind, also eine Basis von $\operatorname{im} f$ bilden: Aus $\sum_{\nu=r+1}^{n} \lambda^\nu f(u_\nu) = 0$ für $\lambda^\nu \in \mathbb{K}$ folgt $f\left(\sum_{\nu=r+1}^{n} \lambda^\nu u_\nu\right) = 0$, also $\sum_{\nu=r+1}^{n} \lambda^\nu u_\nu \in \ker f$, folglich $\lambda^{r+1} = \cdots = \lambda^n = 0$. Somit gilt

$$\dim(\ker f) + \dim(\operatorname{im} f) = r + (n-r) \;=\; n \;=\; \dim \mathfrak{U}. \qquad \square$$

Wir bezeichnen noch die Dimension des Bildraumes einer linearen Abbildung $f\colon \mathfrak{U} \longrightarrow \mathfrak{V}$ als *Rang (von f)* und notieren

$$\operatorname{rang} f \;:=\; \dim(\operatorname{im} f)\,.$$

Trivialerweise gilt $\operatorname{rang} f \leq \dim \mathfrak{U}$; denn nach 3.4.7 gibt es in $f(\mathfrak{U}) = \operatorname{im} f$ höchstens so viele linear unabhängige Vektoren wie in $\mathfrak{U}$. Hier — und an vielen anderen Stellen — ist natürlich $\infty \leq \infty$ und $x \leq \infty$ für alle $x \in \mathbb{R}$ zu lesen.

Für ein $A \in \mathbb{K}^{m\times n}$ definieren wir damit den *Rang* über

$$\operatorname{rang} A \;:=\; \operatorname{rang} f_A\,.$$

Das ist gerade die Dimension des Spaltenraums von A; denn

$$\operatorname{im} f_A \;=\; f_A(\mathfrak{U}) \underset{(4.1.1.a)}{=} \langle f_A(e_1), \ldots, f_A(e_n)\rangle \;=\; \langle A_1, \ldots, A_n\rangle$$

Wir werden noch zeigen (siehe Satz 5.2.9), daß $\operatorname{rang} f$ auch die Dimension des Zeilenraums ist.

Hier notieren wir noch ein einfaches *Kriterium für die Lösbarkeit eines linearen Gleichungssystems*, das wir wieder kurz als Matrixgleichung schreiben:

Bemerkung 4.2.2.

Zu $b \in \mathbb{K}^m$ *betrachten wir die erweiterte Matrix*

$$(A, b) := (A_1, \ldots, A_n, b)\,.$$

Es existiert genau dann ein $x \in \mathbb{K}^n$ *mit* $Ax = b$*, wenn gilt:*

$$\operatorname{rang} A = \operatorname{rang}(A, b)$$

Die Matrix (A, b) entsteht also aus A, indem die rechte Seite b als $(n+1)$-te Spalte an A noch angehängt wird. Das Gleichungssystems $Ax = b$ ist also genau dann lösbar, wenn A und die erweiterte Matrix (A, b) den gleichen Rang haben.

Beweis: $b \in \operatorname{im} f_A = \langle A_1, \ldots, A_n\rangle \iff \langle A_1, \ldots, A_n, b\rangle = \langle A_1, \ldots, A_n\rangle$
$\iff \operatorname{rang}(A, b) = \operatorname{rang} A$ □

Aus der obigen Dimensionsformel liest man nun leicht ab, was wir zum Teil schon nach Übungsaufgabe 5.1.a wissen:

Bemerkung 4.2.3.

Es seien $\mathfrak{U}$ *ein* endlich-dimensionaler $\mathbb{K}$*-Vektorraum und* $f\colon \mathfrak{U} \longrightarrow \mathfrak{V}$ *eine lineare Abbildung. Dann gelten:*

a) f *ist genau dann* injektiv, *wenn* $\operatorname{rang} f = \dim \mathfrak{U}$ *gilt.*

b) f *ist genau dann* surjektiv, *wenn* $\operatorname{rang} f = \dim \mathfrak{V}$ *gilt.*

c) Falls $\dim \mathfrak{U} = \dim \mathfrak{V}$ $(\in \mathbb{N}_0)$ *gilt, hat man:*
f *Monomorphismus* $\iff$ f *Epimorphismus* $\iff$ f *Isomorphismus*

Beweis:

a): f injektiv $\underset{(3.4.10)}{\iff}$ $\ker f = \{0\}$ $\iff$ $\operatorname{def}_1 = 0 \underset{(4.2.1)}{\iff} \operatorname{rang} f = \dim \mathfrak{U}$

b): Ist f surjektiv, dann gilt $\operatorname{rang} f = \dim(\operatorname{im} f) = \dim \mathfrak{V}$.
Hat man $\mathbb{N}_0 \ni \operatorname{rang} f = \dim(\operatorname{im} f) = \dim \mathfrak{V}$, so gilt nach Bemerkung 3.3.7 $\operatorname{im} f = \mathfrak{V}$, also f surjektiv.

c): nach *a)* und *b)* gilt *in diesem Fall:* f injektiv $\iff$ f surjektiv □

Die Dimensionsformel für Homomorphismen bleibt auch für *beliebige* Vektorräume $\mathfrak{U}$ richtig, wie man durch leichte Modifikation des gegebenen Beweises sieht. Hingegen ist in der Bemerkung 4.2.3 die Voraussetzung, daß $\mathfrak{U}$ *endlich-dimensional* ist, *wesentlich:*

Beispiele

(B1) $\mathfrak{U} := \mathfrak{V} := \mathbb{R}^{\mathbb{N}}$; $f\colon \mathfrak{U} \ni (a_1, a_2, a_3, \ldots) \longmapsto (0, a_1, a_2, a_3, \ldots) \in \mathfrak{V}$
f ist *injektiv,* aber *nicht* surjektiv.
$\operatorname{rang} f = \infty = \dim \mathfrak{U}$

(B2) $\mathfrak{U} := \mathfrak{V} := \mathbb{R}^{\mathbb{N}}$; $f\colon \mathfrak{U} \ni (a_1, a_2, a_3, \ldots) \longmapsto (a_2, a_3, a_4, \ldots) \in \mathfrak{V}$
f ist *surjektiv,* aber *nicht* injektiv.
$\operatorname{rang} f = \infty = \dim \mathfrak{V}$, $\operatorname{def}_1 f = 1$ ◁

4.3. **Darstellende Matrix einer linearen Abbildung.**

Mit $n, m, r \in \mathbb{N}$ seien nun $\mathcal{A} = (u_1, \ldots, u_n)$ eine Basis von $\mathfrak{U}$, $\mathcal{B} = (v_1, \ldots, v_m)$ eine Basis von $\mathfrak{V}$ und $\mathcal{C} = (w_1, \ldots, w_r)$ eine Basis von $\mathfrak{W}$, also alle drei Räume insbesondere endlich-dimensional.

Wir benötigen im Folgenden oft die Koordinatenabbildungen gemäß Lemma 3.4.11, z. B. :

$$\Phi_{\mathcal{A}}\colon \mathbb{K}^n \ni \begin{pmatrix} \lambda^1 \\ \vdots \\ \lambda^n \end{pmatrix} \longmapsto \sum_{\nu=1}^{n} \lambda^\nu u_\nu \in \mathfrak{U}$$

Hier haben wir, was wir ja künftig machen wollten, die Elemente von $\mathbb{K}^n$ als Spaltenvektoren geschrieben.

Wir bezeichnen noch für ein $u \in \mathfrak{U}$ den Koordinatenvektor von u bezüglich $\mathcal{A}$ mit ${}_{\mathcal{A}}[u]$, also

$${}_{\mathcal{A}}[u] := (\Phi_{\mathcal{A}})^{-1}(u)\,.$$

Satz 4.3.1.

Zu jeder linearen Abbildung $f\colon \mathfrak{U} \longrightarrow \mathfrak{V}$ existiert genau eine $(m \times n)$-Matrix

$$A =: \mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(f) = (a_\nu^\mu)_{\substack{1\le\mu\le m \\ 1\le\nu\le n}}$$

mit

$$f(u_\nu) = \sum_{\mu=1}^{m} a_\nu^\mu v_\mu \quad \text{für alle } \nu = 1, \ldots, n\,.$$

Dabei gilt:

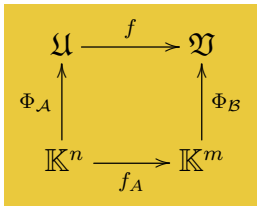
$$\Phi_{\mathcal{B}} \circ f_A = f \circ \Phi_{\mathcal{A}}$$

Der ν-te Spaltenvektor A_ν der darstellenden Matrix $A = \mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(f)$ ist gerade der Koordinatenvektor von $f(u_\nu)$ bezüglich $\mathcal{B}$ $(\nu = 1, \ldots, n)$, also

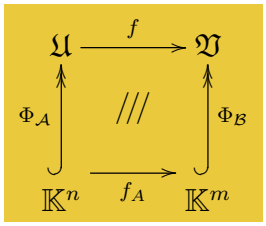
$$\mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(f)\, e_\nu = {}_{\mathcal{B}}[f(u_\nu)]\,.$$

In der Literatur ist die Bezeichnungsweise, welche Basis unten oder oben notiert wird, uneinheitlich ...

Die Beziehung $\Phi_{\mathcal{B}} \circ f_A = f \circ \Phi_{\mathcal{A}}$ kann man in einem *Diagramm* verdeutlichen. Solche Diagramme können gelegentlich helfen, komplizierte Sachverhalte recht übersichtlich darzustellen:



Wir schreiben auch



um die *Kommutativität* anzudeuten. Ein Diagramm aus ‚Räumen' und ‚Pfeilen', die Abbildungen zwischen diesen Räumen entsprechen, heißt *kommutativ,* wenn folgendes gilt: Starte in einem Raum und folge den Pfeilen (‚Wege' mit gleichem Start und Ziel). Dann ist das Ergebnis in einem anderen Raum unabhängig von der speziellen Wahl der Pfeile, hier also $\Phi_{\mathcal{B}} \circ f_A = f \circ \Phi_{\mathcal{A}}$.

Dabei verwenden wir gelegentlich die Pfeile $\hookrightarrow$ für eine *injektive*, $\twoheadrightarrow$ für eine *surjektive* Abbildung und $\xhookrightarrow{}$ für eine *bijektive* Abbildung. $\simeq$ an einem Pfeil soll andeuten, daß es sich um einen *Isomorphismus* handelt.

Beweis:
Die eindeutige Existenz der Matrix $\mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(f)$ ist klar.
Für den Nachweis der letzten Beziehung genügt es, die Übereinstimmung der beiden Abbildungen auf den Basisvektoren $e_1, \ldots, e_n \in \mathbb{K}^n$ zu zeigen: Für $\nu = 1, \ldots,$ hat man:

$$(\Phi_{\mathcal{B}} \circ f_A)(e_\nu) = \Phi_{\mathcal{B}}(f_A(e_\nu)) = \Phi_{\mathcal{B}}\, A_\nu = \Phi_{\mathcal{B}} \begin{pmatrix} a_\nu^1 \\ \vdots \\ a_\nu^m \end{pmatrix} = \sum_{\mu=1}^{m} a_\nu^\mu v_\mu$$

$$(f \circ \Phi_{\mathcal{A}})(e_\nu) = f\,(\Phi_{\mathcal{A}}(e_\nu)) \overset{\checkmark}{=} f(u_\nu) = \sum_{\mu=1}^{m} a_\nu^\mu v_\mu$$

□

Aus der Beziehung $\Phi_{\mathcal{B}} \circ f_A = f \circ \Phi_{\mathcal{A}}$ aus Satz 4.3.1 (mit $A = \mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(f)$) liest man noch ab:

Folgerung 4.3.2.
Für $u \in \mathfrak{U}$ und $f \in \mathcal{L}(\mathfrak{U}, \mathfrak{V})$ gilt:

$$_{\mathcal{B}}[f(u)] = \mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(f)_{\mathcal{A}}[u]$$

Beweis: Mit $\lambda := {}_{\mathcal{A}}[u] = (\Phi_{\mathcal{A}})^{-1}(u)$ liefert

$$\Phi_{\mathcal{B}}(A\,\lambda) = (\Phi_{\mathcal{B}} \circ f_A)(\lambda) = (f \circ \Phi_{\mathcal{A}})(\lambda) = f(u)$$

die Beziehung

$$\mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(f)\,{}_{\mathcal{A}}[u] = A\,\lambda = {}_{\mathcal{B}}[f(u)]\,. \qquad \square$$

Die Koordinaten von $f(u)$ — für eine lineare Abbildung $f \in \mathcal{L}(\mathfrak{U},\mathfrak{V})$ und $u \in \mathfrak{U}$ — bezüglich der Basis $\mathcal{B}$ erhält man also durch Multiplikation des Koordinatenvektors von u (bezüglich der Basis $\mathcal{A}$) von links mit der zu f gehörenden Matrix $\mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(f)$.

Ist speziell $\mathfrak{U} = \mathfrak{V}$ und $\mathcal{A} = \mathcal{B}$, dann notieren wir auch

$$\mathcal{M}_{\mathcal{A}}(f)$$

statt $\mathcal{M}_{\mathcal{A}}^{\mathcal{A}}(f)$.

Bemerkung 4.3.3.

Zu jedem $M \in \mathbb{K}^{m\times n}$ existiert genau ein $f \in \mathcal{L}(\mathfrak{U},\mathfrak{V})$ mit

$$M = \mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(f)$$

Die Abbildung

$$\mathcal{L}(\mathfrak{U},\mathfrak{V}) \ni f \longmapsto \mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(f) \in \mathbb{K}^{m\times n}$$

ist also bijektiv.

Beweis: Zu $M = (\gamma_\nu^\mu)_{\substack{1\le\mu\le m\\ 1\le\nu\le n}}$ betrachten wir $w_\nu := \sum_{\mu=1}^{m} \gamma_\nu^\mu\, v_\mu$ für $\nu = 1,\dots,n$. Nach Satz 4.1.1 existiert genau ein $f \in \mathcal{L}(\mathfrak{U},\mathfrak{V})$ mit $f(u_\nu) = w_\nu = \sum_{\mu=1}^{m} \gamma_\nu^\mu\, v_\mu$. Das aber bedeutet gerade $\mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(f) = M$. $\square$

Bemerkung 4.3.4.

Für $f,g \in \mathcal{L}(\mathfrak{U},\mathfrak{V})$ und $\lambda \in \mathbb{K}$ gilt:

$$\mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(\lambda f + g) = \lambda\,\mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(f) + \mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(g)$$

Die Abbildung

$$\mathcal{L}(\mathfrak{U},\mathfrak{V}) \ni f \longmapsto \mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(f) \in \mathbb{K}^{m\times n}$$

ist also ein Isomorphismus.

Beweis: Mit $\mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(f) = (a_\nu^\mu)$ und $\mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(g) = (b_\nu^\mu)$ hat man

$$f(u_\nu) = \sum_{\mu=1}^{m} a_\nu^\mu v_\mu \quad\text{und}\quad g(u_\nu) = \sum_{\mu=1}^{m} b_\nu^\mu v_\mu \quad \text{für alle } \nu = 1,\dots,n\,,$$

also

$$(\lambda f+g)(u_\nu) = \lambda f(u_\nu)+g(u_\nu) = \lambda \sum_{\mu=1}^{m} a_\nu^\mu v_\mu + \sum_{\mu=1}^{m} b_\nu^\mu v_\mu = \sum_{\mu=1}^{m} (\lambda a_\nu^\mu + b_\nu^\mu)\, v_\mu\,.$$

Somit gilt

$$\mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(\lambda f + g) \;=\; (\lambda a_\nu^\mu + b_\nu^\mu) \;=\; \lambda\,(a_\nu^\mu) + (b_\nu^\mu) \;=\; \lambda\mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(f) + \mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(g)$$

□

Wir wissen somit insbesondere, daß die Abbildung

$$\mathcal{M}_{\mathcal{A}}^{\mathcal{B}}\colon \mathcal{L}(\mathfrak{U},\mathfrak{V}) \ni f \longmapsto \mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(f) \in \mathbb{K}^{m\times n}$$

ein *Isomorphismus* ist. Daraus liest man unmittelbar die folgende Aussage ab, die wir aber doch noch einmal ‚direkt' beweisen.

Bemerkung 4.3.5.

$$\dim \mathcal{L}(\mathfrak{U},\mathfrak{V}) \;=\; \dim\mathfrak{U}\cdot\dim\mathfrak{V}$$

Beweis: Es existieren (nach Satz 4.1.1) eindeutig $f_\nu^\mu \in \mathcal{L}(\mathfrak{U},\mathfrak{V})$ mit

$$f_\nu^\mu(u_i) \;=\; \begin{cases} v_\mu, & i=\nu \\ 0, & \text{sonst} \end{cases}$$

für $i,\nu = 1,\dots,n$ und $\mu = 1,\dots,m$. Dann ist

$$(f_\nu^\mu \mid \nu = 1,\dots,n; \mu = 1,\dots,m)$$

eine Basis von $\mathcal{L}(\mathfrak{U},\mathfrak{V})$ (und damit ist die Behauptung gegeben):
Erzeugendensystem: Für $h \in \mathcal{L}(\mathfrak{U},\mathfrak{V})$ und $i = 1,\dots,n$ existiert eindeutig $\alpha_i^\mu \in \mathbb{K}$ mit $h(u_i) = \sum_{\mu=1}^m \alpha_i^\mu v_\mu$. Die Abbildung

$$g \;:=\; \sum_{\nu=1}^{n}\sum_{\mu=1}^{m} \alpha_\nu^\mu\, f_\nu^\mu \in \mathcal{L}(\mathfrak{U},\mathfrak{V})$$

erfüllt

$$g(u_i) \;=\; \sum_{\nu=1}^{n}\sum_{\mu=1}^{m} \alpha_\nu^\mu\, f_\nu^\mu(u_i) \;=\; \sum_{\mu=1}^{m} \alpha_i^\mu v_\mu \;=\; h(u_i)\,;$$

also $g = h$ nach Satz 4.1.1.
Lineare Unabhängigkeit: Für $\beta_\nu^\mu \in \mathbb{K}$ mit $\sum_{\nu=1}^n\sum_{\mu=1}^m \beta_\nu^\mu f_\nu^\mu = 0$ hat man speziell:

$$0 \;=\; \left(\sum_{\nu=1}^{n}\sum_{\mu=1}^{m} \beta_\nu^\mu\, f_\nu^\mu\right)(u_i) \;=\; \sum_{\nu=1}^{n}\sum_{\mu=1}^{m} \beta_\nu^\mu\, f_\nu^\mu(u_i) \;=\; \sum_{\mu=1}^{m} \beta_i^\mu\, v_\mu$$

folglich: $\beta_i^\mu = 0$ $(\dots)$. □

Bemerkung 4.3.6.
Für $f \in \mathcal{L}(\mathfrak{U},\mathfrak{V})$ und $h \in \mathcal{L}(\mathfrak{V},\mathfrak{W})$ gilt:

$$\mathcal{M}_{\mathcal{A}}^{\mathcal{C}}(h\circ f) \;=\; \mathcal{M}_{\mathcal{B}}^{\mathcal{C}}(h)\,\mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(f)$$

Die Hintereinanderausführung linearer Abbildung entspricht also gerade der Multiplikation der zugehörigen Matrizen.
Es paßt also wieder einmal alles zusammen. ☺

Beweis:

Die Basis $\mathcal{C}$ hat die Länge r.
Mit $A = \mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(f) = (a_\nu^\mu)_{\substack{1\le\mu\le m\\ 1\le\nu\le n}}$ und $B = \mathcal{M}_{\mathcal{B}}^{\mathcal{C}}(h) = \left(b_\mu^\varrho\right)_{\substack{1\le\varrho\le r\\ 1\le\mu\le m}}$ gilt

$$f(u_\nu) \;=\; \sum_{\mu=1}^{m} a_\nu^\mu v_\mu \quad \text{und} \quad h(v_\mu) \;=\; \sum_{\varrho=1}^{r} b_\mu^\varrho w_\varrho$$

und folglich

$$\begin{aligned}(h\circ f)(u_\nu) \;=\; h\left(f(u_\nu)\right) \;&=\; \sum_{\mu=1}^{m} a_\nu^\mu h(v_\mu)\\ &=\; \sum_{\mu=1}^{m} a_\nu^\mu \sum_{\varrho=1}^{r} b_\mu^\varrho w_\varrho \;=\; \sum_{\varrho=1}^{r}\left(\sum_{\mu=1}^{m} b_\mu^\varrho\, a_\nu^\mu\right) w_\varrho\,.\end{aligned}$$

In der letzten Klammer stehen aber gerade die Elemente von BA. □

Folgerung 4.3.7.
Ist $f\colon \mathfrak{U} \longrightarrow \mathfrak{V}$ ein Isomorphismus, *dann ist $\mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(f)$ regulär mit:*

$$\mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(f^{-1}) \;=\; \left(\mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(f)\right)^{-1}$$

Beweis: Mit $h := f^{-1} \in \mathcal{L}(\mathfrak{V},\mathfrak{U})$ liefert 4.3.6:

$$\mathbb{1}_n \overset{\checkmark}{=} \mathcal{M}_{\mathcal{A}}^{\mathcal{A}}(\mathrm{id}_{\mathfrak{U}}) \;=\; \mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(f^{-1})\,\mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(f) \qquad \square$$

4.4. **Dualraum.**

Wir sehen uns — aus Zeitgründen nur ganz kurz — den Begriff des *Dualraums* an, wir machen dabei insbesondere keine Dualitätstheorie.

Es seien $\mathbb{K}$ ein Körper, $\mathfrak{V}$ ein $\mathbb{K}$-Vektorraum und $n \in \mathbb{N}$.

Definition 4.4.1.
Abbildungen $f \in \mathcal{L}(\mathfrak{V},\mathbb{K})$ heißen (lineare) *Funktionale* oder *Linearformen* auf $\mathfrak{V}$. Wir schreiben

$$\mathfrak{V}^* \;:=\; \mathcal{L}(\mathfrak{V},\mathbb{K})\,.$$

$\mathfrak{V}^*$ heißt der zu $\mathfrak{V}$ *duale Raum* oder *Dualraum* von $\mathfrak{V}$.

Hierbei wird natürlich $\mathbb{K}$ als (ein-dimensionaler) Vektorraum über sich selbst betrachtet.

Bemerkung 4.4.2.
Ist $\mathfrak{V}$ endlich-dimensional, dann gilt — nach Bemerkung 4.3.5 —

$$\dim\mathfrak{V} \;=\; \dim\mathfrak{V}^*\,.$$

Die beiden Räume sind somit isomorph.

Ist $(v_1, \dots, v_n)$ *eine Basis von* $\mathfrak{V}$*, dann ist — nach dem Beweis der Bemerkung 4.3.5 — durch*

$$v_\nu^*(v_\mu) := \delta_\nu^\mu \quad (\nu, \mu = 1, \dots, n)$$

eine Basis $(v_1^*, \dots, v_n^*)$ *von* $\mathfrak{V}^*$*, die* duale Basis *zu* $(v_1, \dots, v_n)$*, gegeben.*

Beispiele

(B1) Für $\nu \in \{1, \dots, n\}$ die Projektion auf die ν-te Komponente

$$\begin{array}{rccc} \mathrm{pr}_\nu\colon & \mathbb{K}^n & \longrightarrow & \mathbb{K} \\ & \cup & & \cup \\ & (x^1, \dots, x^n)^T & \longmapsto & x^\nu \end{array}$$

(B2) Grenzwertbildungen

(B3) Auswertungsabbildungen, z. B. $\delta\colon C^0[0,1] \longrightarrow \mathbb{R}$ durch $\delta(f) := f(0)$, das Dirac-*Funktional*

(B4) $C^0[0,1] \ni f \longmapsto \int\limits_0^1 f(x)\,dx$ ◁

Da man jeden Vektor $v \in \mathfrak{V} \setminus \{0\}$ zu einer Basis ergänzen kann, folgt:

Folgerung 4.4.3.
Zu jedem Vektor $v \in \mathfrak{V} \setminus \{0\}$ *existiert ein* $f \in \mathfrak{V}^*$ *mit* $f(v) = 1$.

4.5. Basiswechsel.

Zu einer linearen Abbildung $f\colon \mathfrak{U} \longrightarrow \mathfrak{V}$ und (endlichen geordneten) Basen $\mathcal{A}$ von $\mathfrak{U}$ und $\mathcal{B}$ von $\mathfrak{V}$ hatten wir in Satz 4.3.1 die darstellende Matrix

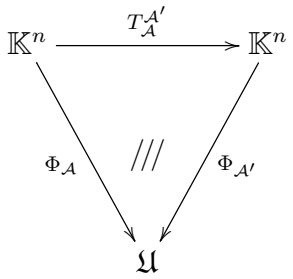
$$\mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(f)$$

erklärt. Hat man weitere (endliche geordnete) Basen $\mathcal{A}' = (u_1', \dots, u_n')$ von $\mathfrak{U}$ und $\mathcal{B}' = (v_1', \dots, v_m')$ von $\mathfrak{V}$, dann stellt sich naturgemäß die Frage nach dem Zusammenhang zwischen $\mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(f)$ und $\mathcal{M}_{\mathcal{A}'}^{\mathcal{B}'}(f)$.

Hierzu betrachten wir zunächst die *Transformationsabbildung*

$$T_{\mathcal{A}}^{\mathcal{A}'} := (\Phi_{\mathcal{A}'})^{-1} \circ \Phi_{\mathcal{A}}\colon \mathbb{K}^n \longrightarrow \mathbb{K}^n. \tag{4.1}$$

Einem Vektor $\lambda \in \mathbb{K}^n$ wird durch $T_{\mathcal{A}}^{\mathcal{A}'}$ also ${}_{\mathcal{A}'}[\Phi_{\mathcal{A}}(\lambda)]$ zugeordnet.



Man hat offenbar:

Bemerkung 4.5.1.

$$\left(T_{\mathcal{A}}^{\mathcal{A}'}\right)^{-1} = T_{\mathcal{A}'}^{\mathcal{A}}$$

Speziell für $\mathfrak{U} = \mathfrak{V}$ (also $n = m$) und $f = \mathrm{id}_{\mathfrak{U}}$ liefert die Matrix

$$A = \mathcal{M}_{\mathcal{A}}^{\mathcal{A}'}(f)$$

die Beschreibung der Basisvektoren u_ν durch die Basis $\mathcal{A}' = (u'_1, \ldots, u'_n)$:

$$u_\nu = \sum_{\mu=1}^{n} a_\nu^\mu \, u'_\mu \quad (\nu = 1, \ldots, n)$$

$$\Phi_{\mathcal{A}'} \circ f_A = \mathrm{id}_{\mathfrak{U}} \circ \Phi_{\mathcal{A}} = \Phi_{\mathcal{A}}$$

zeigt:

$$f_A = (\Phi_{\mathcal{A}'})^{-1} \circ \Phi_{\mathcal{A}} = T_{\mathcal{A}}^{\mathcal{A}'} \quad \text{für} \quad A = \mathcal{M}_{\mathcal{A}}^{\mathcal{A}'}(\mathrm{id}_{\mathfrak{U}}) \tag{4.2}$$

Einem $\lambda = (\lambda^1, \ldots, \lambda^n) \in \mathbb{K}^n$ wird also zunächst $\sum_{\nu=1}^n \lambda^\nu \, u_\nu$ und dann $(\beta^1, \ldots, \beta^n) \in \mathbb{K}^n$ mit $\sum_{\nu=1}^n \lambda^\nu \, u_\nu = \sum_{\nu=1}^n \beta^\nu \, u'_\nu$ zugeordnet.

Man sollte sich dabei merken:

Die Koordinaten transformieren sich gerade invers *zu den Basen!*

Damit wird die Eingangsfrage beantwortet durch:

Satz 4.5.2.
Für $f \in \mathcal{L}(\mathfrak{U}, \mathfrak{V})$ und endliche geordnete Basen $\mathcal{A}, \mathcal{A}'$ von $\mathfrak{U}$ und $\mathcal{B}, \mathcal{B}'$ von $\mathfrak{V}$ gilt:

$$\mathcal{M}_{\mathcal{A}'}^{\mathcal{B}'}(f) = T_{\mathcal{B}}^{\mathcal{B}'} \, \mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(f) \, \left(T_{\mathcal{A}}^{\mathcal{A}'}\right)^{-1}$$

Hier haben wir die Matrix A und dadurch gegebene lineare Abbildung f_A in der Notierung nicht mehr unterschieden

Mit $A := \mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(f), A' := \mathcal{M}_{\mathcal{A}'}^{\mathcal{B}'}(f), S := T_{\mathcal{B}}^{\mathcal{B}'}, T := T_{\mathcal{A}}^{\mathcal{A}'}$ lautet dieses Beziehung:

$$A' = S \, A \, T^{-1}$$

Beweis: Über

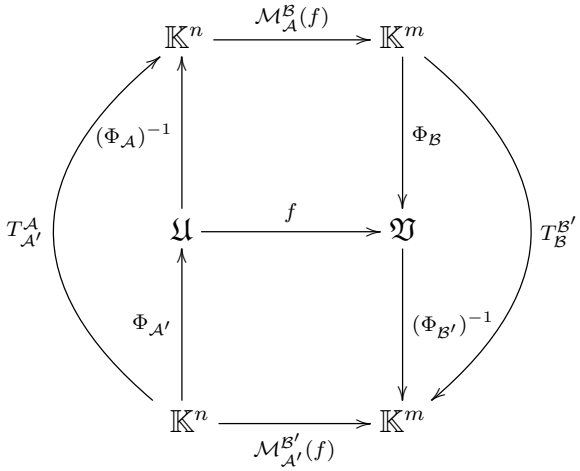
$$f = \mathrm{id}_{\mathfrak{V}} \circ f \circ \mathrm{id}_{\mathfrak{U}}$$

folgt mit Bemerkung 4.3.6 und der Beziehung (4.2):

$$\begin{aligned} \mathcal{M}_{\mathcal{A}'}^{\mathcal{B}'}(f) &= \mathcal{M}_{\mathcal{B}}^{\mathcal{B}'}(\mathrm{id}_{\mathfrak{V}}) \, \mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(f) \, \mathcal{M}_{\mathcal{A}'}^{\mathcal{A}}(\mathrm{id}_{\mathfrak{U}}) \\ &\underset{(4.2)}{=} T_{\mathcal{B}}^{\mathcal{B}'} \, \mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(f) \, T_{\mathcal{A}'}^{\mathcal{A}} \\ &\underset{(4.5.1)}{=} T_{\mathcal{B}}^{\mathcal{B}'} \, \mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(f) \, \left(T_{\mathcal{A}}^{\mathcal{A}'}\right)^{-1} \end{aligned}$$

□

Mit dem Diagramm von Seite 55 erhalten wir hier:



Auch hieraus liest man den Beweis unmittelbar ab.

Wir sehen uns ein kleines, ganz einfaches und dadurch gut überschaubares **Beispiel** an:

Mit $a_1' := \begin{pmatrix}3\\1\end{pmatrix}$ und $a_2' := \begin{pmatrix}1\\2\end{pmatrix}$ seien

$$\mathcal{A} := (e_1, e_2) \quad \text{und} \quad \mathcal{A}' := (a_1', a_2')\,.$$

Für $\begin{pmatrix}x\\y\end{pmatrix} \in \mathbb{R}^2$ gilt offenbar ${}_{\mathcal{A}}\left[\begin{pmatrix}x\\y\end{pmatrix}\right] = \begin{pmatrix}x\\y\end{pmatrix}$. Koordinaten bezüglich $\mathcal{A}'$ sind zum Beispiel:

$$ {}_{\mathcal{A}'}\left[\begin{pmatrix}3\\1\end{pmatrix}\right] = \begin{pmatrix}1\\0\end{pmatrix}, \quad {}_{\mathcal{A}'}\left[\begin{pmatrix}1\\2\end{pmatrix}\right] = \begin{pmatrix}0\\1\end{pmatrix}, \quad {}_{\mathcal{A}'}\left[\begin{pmatrix}4\\3\end{pmatrix}\right] = \begin{pmatrix}1\\1\end{pmatrix},$$

$$ {}_{\mathcal{A}'}\left[\begin{pmatrix}3\\0\end{pmatrix}\right] = \begin{pmatrix}6/5\\-3/5\end{pmatrix}, \quad {}_{\mathcal{A}'}\left[\begin{pmatrix}1\\0\end{pmatrix}\right] = \begin{pmatrix}2/5\\-1/5\end{pmatrix}, \quad {}_{\mathcal{A}'}\left[\begin{pmatrix}0\\2\end{pmatrix}\right] = \begin{pmatrix}-2/5\\6/5\end{pmatrix}$$

Die ersten drei Gleichungen sind ohne Rechnung zu sehen. Die vierte Gleichung folgt aus der fünften. Die fünfte und sechste Gleichung berechnen wir wie folgt:

$a_1' = 3\,e_1 + 1\,e_2$
$a_2' = 1\,e_1 + 2\,e_2$ ergibt $T_{\mathcal{A}'}^{\mathcal{A}} = \begin{pmatrix}3 & 1\\1 & 2\end{pmatrix}$ und folglich

$$T_{\mathcal{A}}^{\mathcal{A}'} \underset{(4.5.1)}{=} \left(T_{\mathcal{A}'}^{\mathcal{A}}\right)^{-1} \overset{\checkmark}{=} \frac{1}{5}\begin{pmatrix}2 & -1\\-1 & 3\end{pmatrix}.$$

Damit erhält man:

$$ {}_{\mathcal{A}'}\left[\begin{pmatrix}1\\0\end{pmatrix}\right] = \frac{1}{5}\begin{pmatrix}2 & -1\\-1 & 3\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}2/5\\-1/5\end{pmatrix},$$

$$ {}_{\mathcal{A}'}\left[\begin{pmatrix}0\\2\end{pmatrix}\right] = \frac{2}{5}\begin{pmatrix}2 & -1\\-1 & 3\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}-2/5\\6/5\end{pmatrix}$$

Auch eine beispielhafte *Probe* kann dem Verständnis nicht schaden:

$$\Phi_{\mathcal{A}'}\begin{pmatrix}-2/5\\6/5\end{pmatrix} = -2/5\,a'_1 + 6/5\,a'_2 = \begin{pmatrix}0\\2\end{pmatrix} \qquad \triangleleft$$

Im Spezialfall eines *Endomorphismus* ergibt sich:

Bemerkung 4.5.3.
Für $f \in \mathcal{L}(\mathfrak{U},\mathfrak{U})$ *und Basen* $\mathcal{A}$, $\mathcal{A}'$ *von* $\mathfrak{U}$ *gilt:*

$$\mathcal{M}_{\mathcal{A}'}(f) = T_{\mathcal{A}}^{\mathcal{A}'}\,\mathcal{M}_{\mathcal{A}}(f)\,\left(T_{\mathcal{A}}^{\mathcal{A}'}\right)^{-1}$$

Mit $A := \mathcal{M}_{\mathcal{A}}(f)$, $A' := \mathcal{M}_{\mathcal{A}'}(f)$ und $T := T_{\mathcal{A}}^{\mathcal{A}'}$ lautet dieses Beziehung:

$$A' = T\,A\,T^{-1}$$

Die beiden Transformationsweisen werden uns — bei Klassen von Matrizen — zu den Begriffen *äquivalent* und *ähnlich* und damit zu *Normalformenproblemen* führen.

4.6. Quotientenraum.

Neben den bisher schon behandelten Konstruktionsprinzipien von Vektorräumen — Unterräume, direkte Summen, Vektorräume von Abbildungen — ist die Bildung von Quotientenräumen durch Bildung von Äquivalenzklassen eine weitere wichtige Konstruktionsmethode. Hierbei wird alles — im Sinne des eingenommenen Standpunktes — Unwesentliche an den untersuchten Objekten abgestreift. Ein erstes Beispiel dazu lernt man schon in der Schule bei der Bruchrechnung kennen. In Beispiel 2.2.1.c) hatten wir mit der Einführung von $\mathbb{Z}_p$ diese Dinge gestreift. Auch in der Analysis-Vorlesung lernt man oft — bei der Einführung der reellen Zahlen über die Vervollständigung der rationalen Zahlen mittels Äquivalenzklassen von CAUCHY-Folgen rationaler Zahlen — ein sehr wichtiges Beispiel kennen.

In diesem Abschnitt seien $\mathbb{K}$ ein Körper, $\mathfrak{V},\mathfrak{W}$ $\mathbb{K}$-Vektorräume und $\mathfrak{U}$ ein Unterraum von $\mathfrak{V}$.

Definition 4.6.1.
Für $x,y \in \mathfrak{V}$ $\quad x \sim y :\Longleftrightarrow x \sim_{\mathfrak{U}} y :\Longleftrightarrow x - y \in \mathfrak{U}$
Wir lesen dies als: x ist *„äquivalent (bezüglich $\mathfrak{U}$)“* zu y.

Bemerkung 4.6.2.
$\sim$ *ist eine Äquivalenzrelation auf* $\mathfrak{V}$ *, d. h. für* $x,y,z \in \mathfrak{V}$ *gelten:*

$x \sim x$ *(Reflexivität)*

$x \sim y \implies y \sim x$ *(Symmetrie)*

$x \sim y \wedge y \sim z \implies x \sim z$ *(Transitivität)*

Beweis:
a): $x - x = 0 \in \mathfrak{U}$ zeigt $x \sim x$.
b): $x \sim y \Longrightarrow x - y \in \mathfrak{U} \Longrightarrow y - x = -(x-y) \in \mathfrak{U}$: $y \sim x$

c): $x \sim y \wedge y \sim z \Longrightarrow x-y, y-z \in \mathfrak{U} \Longrightarrow x-z \in \mathfrak{U} : x \sim z$ □

Bemerkung 4.6.3.

Für $a \in \mathfrak{V}$ *gilt:*

$$a+\mathfrak{U} := \{a+u \mid u \in \mathfrak{U}\} = \{x \in \mathfrak{V} \mid x \sim a\} =: \omega(a)$$

Wir bezeichnen $\omega(a)$ als *Äquivalenzklasse* oder *Nebenklasse* zu a und jedes $x \in \omega(a)$ als *Repräsentanten* oder *Vertreter* der Äquivalenzklasse $\omega(a)$.

Beweis: Für $x \in \mathfrak{V}$: $x \sim a \Longleftrightarrow x-a \in \mathfrak{U} \Longleftrightarrow x \in a+\mathfrak{U}$ □

Bemerkung 4.6.4.

Für $a, a', b, b' \in \mathfrak{V}$ *und* $\alpha \in \mathbb{K}$ *gelten:*

1. $\omega(a) = \omega(a') \Longleftrightarrow a \sim a'$

2. $a \sim a' \wedge b \sim b' \Longrightarrow \alpha a + b \sim \alpha a' + b'$

3. $\omega(a) = \omega(a') \wedge \omega(b) = \omega(b') \Longrightarrow \omega(\alpha a + b) = \omega(\alpha a' + b')$

Beweis:

1. „$\Longrightarrow$“: $a = a + 0 \in a + \mathfrak{U} = \omega(a) = \omega(a')$, also $a \sim a'$

Aus Symmetriegründen genügt für die andere Richtung, die Inklusion $\omega(a) \subset \omega(a')$ zu zeigen: $x \in \omega(a)$ bedeutet $x \sim a$. Das liefert mit $a \sim a'$ dann $x \sim a'$ und so $x \in \omega(a')$.

2. $a-a', b-b' \in \mathfrak{U} \Longrightarrow \alpha a + b - (\alpha a' + b') = \alpha(a-a') + (b-b') \in \mathfrak{U}$

3. ist — nach *1.* — die gleiche Aussage wie *2.* □

Die spezielle Äquivalenzrelation $\sim$ ist also verträglich mit den Vektorraumoperationen. Wir sprechen in einem solchen Fall von einer *„Kongruenzrelation“*.

Nach *3.* ist die folgende Festsetzung wohldefiniert:

Definition 4.6.5. Für $a, b \in \mathfrak{V}$ und $\alpha \in \mathbb{K}$:

$A(\omega(a), \omega(b)) := \omega(a) + \omega(b) := \omega(a+b)$

$\sigma(\alpha, \omega(a)) := \alpha\omega(a) := \omega(\alpha a)$

Wir bezeichnen noch

$$\mathfrak{V}/\mathfrak{U} := \{\omega(a) \mid a \in \mathfrak{V}\}$$

als *„Quotientenraum“* (manche sagen auch *„Faktorraum“*).

Satz 4.6.6.

a) $(\mathfrak{V}/\mathfrak{U}, A, \sigma)$ *ist ein* $\mathbb{K}$*-Vektorraum.*

b) $\omega : \mathfrak{V} \ni a \longmapsto \omega(a) \in \mathfrak{V}/\mathfrak{U}$ *ist ein Epimorphismus.*

c) $\dim \mathfrak{U} + \dim \mathfrak{V}/\mathfrak{U} = \dim \mathfrak{V}$, *falls* $\mathfrak{V}$ *endlich-dimensional ist.*

d) Zu $f \in \mathcal{L}(\mathfrak{V},\mathfrak{W})$ *existiert eindeutig eine lineare Abbildung*

$$\widetilde{f}\colon \mathfrak{V}/\ker f \longrightarrow \mathfrak{W}$$

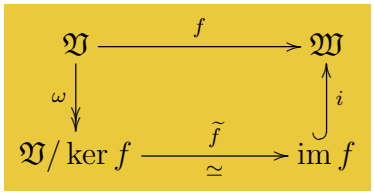
mit $f = \widetilde{f}\circ\omega\,.$

Dieses $\widetilde{f}$ *ist injektiv.*

e) Für $f \in \mathcal{L}(\mathfrak{V},\mathfrak{W})$ *ist* $\boxed{\operatorname{im} f \simeq \mathfrak{V}/\ker f}$. *(Isomorphiesatz)*

Zusatz: Teil *c*) gilt auch ohne die Voraussetzung, daß $\mathfrak{V}$ endlich-dimensional ist.

Ein *Diagramm* kann wieder den Kern der Aussagen *d*) und *e*) verdeutlichen:



Hierbei bezeichnen $\omega\colon \mathfrak{V} \longrightarrow \mathfrak{V}/\ker f$, $\mathfrak{V} \ni v \longmapsto v+\ker f$, die *kanonische Abbildung*, $i\colon \operatorname{im} f \longrightarrow \mathfrak{W}$ die Inklusionsabbildung $(z \longmapsto z)$ und $\widetilde{f}\colon \mathfrak{V}/\ker f \longrightarrow \mathfrak{W}$ die *durch* f *induzierte Abbildung*, definiert durch $\widetilde{f}(a+\ker f) := f(a)$. Dieses Diagramm ist kommutativ.

Beweis:

a): Die Vektorraum-Eigenschaften übertragen sich unmittelbar aus den entsprechenden Eigenschaften in $\mathfrak{V}$, z. B. :

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

Das Nullelement in $(\mathfrak{V}/\mathfrak{U}, A)$ ist $\omega(0) = \mathfrak{U}$.

b): ω ist *linear* nach Definition von A und σ. Die Surjektivität ist trivial.

c): Nach Satz 4.2.1 gilt $\dim(\ker\omega) + \dim(\operatorname{im}\omega) = \dim\mathfrak{V}$. Hier hat man $\operatorname{im}\omega \underset{b)}{=} \mathfrak{V}/\mathfrak{U}$ und $\ker\omega \overset{\checkmark}{=} \mathfrak{U}$.

d): *Existenz:* Wir setzen für $a \in \mathfrak{V}$ (notwendigerweise)

$$\widetilde{f}(\omega(a)) := f(a)\,.$$

Dann haben wir: 1. $\widetilde{f}$ ist eine Abbildung:
$\omega(a) = \omega(b) \implies a \sim b \implies a-b \in \ker f \implies f(a) = f(b)$.
2. $\widetilde{f}$ ist linear:
$\widetilde{f}(\alpha\omega(a)+\omega(b)) = \widetilde{f}(\omega(\alpha a+b) = f(\alpha a+b) = \alpha f(a)+f(b) = \alpha\widetilde{f}(\omega(a))+\widetilde{f}(\omega(b))$.
3. $\widetilde{f}\circ\omega = f$: Nach Definition von $\widetilde{f}$
4. $\widetilde{f}$ ist injektiv: $0 = \widetilde{f}(\omega(a)) = f(a) \implies a \in \ker f \implies a \sim 0$, also $\omega(a) = \omega(0)$.

Eindeutigkeit: Hat man $h\colon \mathfrak{V}/\ker f \longrightarrow \mathfrak{W}$ mit $h \circ \omega = f$, so gilt für alle $a \in \mathfrak{V}$: $h(\omega(a)) = f(a)$ also $h = \widetilde{f}$.

e): $\widetilde{f}\,(\mathfrak{V}/\ker f) \;=\; f(\mathfrak{V}) \;=\; \operatorname{im} f$. Nach *d*) folgt somit die Behauptung. □

Für $f \in \mathcal{L}(\mathfrak{V},\mathfrak{W})$ bezeichnen wir noch die Dimension von $\mathfrak{W}/\operatorname{im} f$ als *2. Defekt*, also

$$\operatorname{def}_2 f \;:=\; \dim\,(\mathfrak{W}/\operatorname{im} f)\;.$$

Dieser Defekt mißt, wieviel an Dimension in $\mathfrak{W}$ von f nicht ‚ausgefüllt' wird.

Folgerung 4.6.7.

Falls $\mathfrak{W}$ endlich-dimensional ist, gilt: $\operatorname{rang} f + \operatorname{def}_2 f \;=\; \dim \mathfrak{W}$

Beweis: Satz 4.6.6 *c*) □

Zusatz: Auch diese Folgerung gilt für *beliebige* $\mathfrak{W}$.

Nach Folgerung 4.6.7 und Satz 4.2.1 hat man unter der *Voraussetzung* $\dim \mathfrak{V} = \dim \mathfrak{W} =: n \in \mathbb{N}$ für $f \in \mathcal{L}(\mathfrak{V},\mathfrak{W})$: $\operatorname{def}_1 f \;=\; \operatorname{def}_2 f$

Für $\dim \mathfrak{V} = \dim \mathfrak{W} = \infty$ ist diese Defektaussage *nicht* richtig: Wir wählen für die folgenden Beispiele jeweils $\mathfrak{V} := \mathfrak{W} := \mathbb{R}^{\mathbb{N}}$:

Beispiele

(B1) $f\colon \mathfrak{V} \ni (a_1, a_2, a_3, \ldots) \longmapsto (a_2, a_3, \ldots) \in \mathfrak{W}$:
$\operatorname{def}_1 f = 1,\ \operatorname{def}_2 f = 0$

(B2) $f\colon \mathfrak{V} \ni (a_1, a_2, a_3, \ldots) \longmapsto (a_1, a_3, a_5, \ldots) \in \mathfrak{W}$:
$\operatorname{def}_1 f = \infty,\ \operatorname{def}_2 f = 0$

(B3) $f\colon \mathfrak{V} \ni (a_1, a_2, a_3, \ldots) \longmapsto (0, a_1, a_2, \ldots) \in \mathfrak{W}$:
$\operatorname{def}_1 f = 0,\ \operatorname{def}_2 f = 1$

(B4) $f\colon \mathfrak{V} \ni (a_1, a_2, a_3, \ldots) \longmapsto (0, a_1, 0, a_2, \ldots) \in \mathfrak{W}$:
$\operatorname{def}_1 f = 0,\ \operatorname{def}_2 f = \infty$ ◁

Aus rein mathematischen Gründen würde man jetzt ein Kapitel über Strukturtheorie linearer Abbildungen anschließen. Da dies jedoch oft als ‚etwas schwieriger' angesehen wird, wollen wir uns zunächst etwas ‚erholen' und erst etwas später — wenn wir etwas reifer sind — dies angreifen.

5. Vektorräume mit Skalarprodukt

Ist auf einem Vektorraum ein *Skalarprodukt* gegeben, so besitzt dieser weitaus mehr Eigenschaften als allgemeine Vektorräume. Es lassen sich *Winkel,* speziell *Orthogonalität,* und eine passende *Norm* definieren. Hat man bezüglich der zugeordneten Norm *Vollständigkeit,* spricht man von einem Hilbertraum. Hilberträume spielen eine zentrale Rolle in der mathematischen Beschreibung der Quantenmechanik. Typische Beispiele sind der Hilbert*sche Folgenraum* und viele *Funktionenräume.* Wenn wir Skalarprodukte verwenden, gehen wir stets davon aus, daß wir $\mathbb{R}$- oder $\mathbb{C}$-Vektorräume betrachten.

5.1. Definition und Grundeigenschaften.

Es seien also $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $\mathfrak{H}$ ein $\mathbb{K}$-Vektorraum

Definition 5.1.1.

$\langle\ ,\ \rangle\colon \underset{\cup}{\mathfrak{H}} \times \underset{\cup}{\mathfrak{H}} \longrightarrow \underset{\cup}{\mathbb{K}}$, $(x,y) \longmapsto \langle x,y\rangle$ heißt genau dann *„Semiskalarprodukt“* auf $\mathfrak{H}$, wenn (für alle $x,y,z \in \mathfrak{H}$ und $\alpha \in \mathbb{K}$)

$\langle \cdot\, , z\rangle$ linear (also $\langle \alpha x + y, z\rangle = \alpha\langle x,z\rangle + \langle y,z\rangle$)

$\langle x,y\rangle = \overline{\langle y,x\rangle}$ und

$\langle x,x\rangle \in [0,\infty)$ *(positiv semidefinit)* gelten. Ein Semiskalarprodukt $\langle\ ,\ \rangle$ heißt genau dann *„Skalarprodukt“*, wenn statt der letzten Eigenschaft sogar $\langle x,x\rangle \in (0,\infty)$ für $x \neq 0$ gilt. *(positive Definitheit)*

Ein Paar $(\mathfrak{H}, \langle\ ,\ \rangle)$, bestehend aus einem $\mathbb{K}$-Vektorraum $\mathfrak{H}$ und einem Skalarprodukt $\langle\ ,\ \rangle$, sprechen wir auch als *Prä-*Hilbert*-Raum* an.

Im Folgenden sei $\langle\ ,\ \rangle$ ein *Semiskalarprodukt* auf $\mathfrak{H}$.

Trivialität 5.1.2. $\langle z, \alpha x + y\rangle = \overline{\alpha}\langle z,x\rangle + \langle z,y\rangle$ (...)

Aus der Linearität bezüglich der ersten Komponente folgt die *konjugierte Linearität* bezüglich der zweiten.

Euklidische Vektorräume

Wir sehen uns kurz — aus didaktischen Gründen – zunächst den Fall $\mathbb{K} = \mathbb{R}$ separat an:

Ein *euklidischer Vektorraum* ist ein Vektorraum mit einem *reellen* Skalarprodukt, wir sprechen dann gelegentlich auch von einem *euklidischen Skalarprodukt*. Es sei nun $(\mathfrak{H}, \langle\ ,\ \rangle)$ ein euklidischer Vektorraum:
Hier hat man $\langle a,b\rangle = \langle b,a\rangle$ *(Symmetrie)* und damit Linearität bezüglich beider Argumente: *Bilinearität*

Bemerkung 5.1.3.

a) Per Induktion ergibt sich für beliebige Linearkombinationen

$$\left\langle \sum_{\mu=1}^{m} \lambda^\mu a_\mu, \sum_{\nu=1}^{n} \eta^\nu b_\nu \right\rangle = \sum_{\mu=1}^{m} \sum_{\nu=1}^{n} \lambda^\mu \eta^\nu \langle a_\mu, b_\nu \rangle \,.$$

Somit ist ein Skalarprodukt schon durch seine Werte auf einer Basis eindeutig bestimmt.

b) Es sei speziell $\mathfrak{H} = \mathbb{R}^n$. Für Vektoren $\xi = (\xi^1, \ldots, \xi^n)^T$ und $\eta = (\eta^1, \ldots, \eta^n)^T$ aus $\mathbb{R}^n$ definieren wir

$$\langle \xi, \eta \rangle := \sum_{\nu=1}^{n} \xi^\nu \eta^\nu \,.$$

Dies liefert ein Skalarprodukt auf dem $\mathbb{R}^n$, das *Standard-Skalarprodukt.*

c) Es sei $\mathfrak{H} = \mathbb{R}^2$. Definiere für $\xi = (\xi^1, \xi^2)^T$ und $\eta = (\eta^1, \eta^2)^T$

$$\langle \xi, \eta \rangle := \xi^1\eta^1 + 5\xi^1\eta^2 + 5\xi^2\eta^1 + 26\xi^2\eta^2.$$

Dies ist ein weiteres Skalarprodukt auf $\mathbb{R}^2$. Es stimmt *nicht* mit dem Standard-Skalarprodukt auf $\mathbb{R}^2$ überein.

d) Es sei $\mathfrak{H}$ der Vektorraum der auf $[a, b] \subset \mathbb{R}$ stetigen reellwertigen Funktionen, also $\mathfrak{H} = C^0([a, b])$. (Bei solchen Beispielen gelte stets $a, b \in \mathbb{R}$ mit $a < b$). Definiere für $f, g \in \mathfrak{H}$

$$\langle f, g \rangle := \int_a^b f(x)g(x)\,dx.$$

Dies ist ein Skalarprodukt auf $\mathfrak{H}$. (Das erledigt Kollege Freistühler für uns.)

Würde man im letzten Beispiel statt $C^0([a, b])$ etwa die Menge der Riemann-integrierbaren Funktionen nehmen, so wäre die resultierende Abbildung nur positiv semidefinit, nicht positiv definit. Da es viele solche — und durchaus wichtige — Beispiele gibt, ist es sinnvoll, soweit wie möglich mit Semiskalarprodukten zu arbeiten.

Unitäre Vektorräume

Ein Vektorraum über $\mathbb{C}$ mit einem Skalarprodukt heißt *unitärer Vektorraum.* Wir sprechen in diesem Fall gelegentlich auch von einem *unitären Skalarprodukt*. Wir wiederholen noch einmal die definierenden Eigenschaften:

Definition 5.1.4.

$\langle\ ,\ \rangle\colon \underset{\cup}{\mathfrak{H}} \times \underset{\cup}{\mathfrak{H}} \longrightarrow \underset{\cup}{\mathbb{C}}$, $(x,y) \longmapsto \langle x,y\rangle$ heißt genau dann *„Semiskalarprodukt“* auf $\mathfrak{H}$, wenn (für alle $x,y,z \in \mathfrak{H}$ und $\alpha \in \mathbb{C}$)

$\langle\,\cdot\,,z\rangle$ linear (also $\langle \alpha x + y, z\rangle = \alpha\langle x,z\rangle + \langle y,z\rangle$)

$\langle x,y\rangle = \overline{\langle y,x\rangle}$ *(hermitesch)* und

$\langle x,x\rangle \in [0,\infty)$ gelten. Ein Semiskalarprodukt $\langle\ ,\ \rangle$ heißt genau dann *„Skalarprodukt“*, wenn statt der letzten Eigenschaft sogar $\langle x,x\rangle \in (0,\infty)$ für $x \neq 0$ gilt. *(positive Definitheit)*

Linearität im ersten Argument und konjugierte Linearität im zweiten Argument (siehe 5.1.2) bezeichnet man auch als *Sesquilinearität.*

Es gibt auch die umgekehrte Konvention, d. h. man definiert ein unitäres Skalarprodukt so, daß es im zweiten Argument (statt im ersten Argument) linear ist und daß die übrigen Eigenschaften unverändert gelten. Im ersten Argument werden dann Skalare konjugiert nach außen gezogen.

Folgende Eigenschaften und Beispiele sind analog zum reellen Fall:

a) Per Induktion folgt wieder für beliebige Linearkombinationen

$$\left\langle \sum_{\mu=1}^{m} \lambda^\mu a_\mu, \sum_{\nu=1}^{n} \eta^\nu b_\nu \right\rangle = \sum_{\mu=1}^{m}\sum_{\nu=1}^{n} \lambda^\mu \overline{\eta^\nu} \langle a_\mu, b_\nu\rangle.$$

Daher ist auch ein unitäres Skalarprodukt durch seine Werte auf einer Basis bereits eindeutig bestimmt.

b) Es seien $x = (x^1,\dots,x^n)^T$ und $y = (y^1,\dots,y^n)^T$ Vektoren in $\mathbb{C}^n$, so ist durch

$$\langle x,y\rangle := \sum_{\nu=1}^{n} x^\nu\, \overline{y^\nu}$$

ein unitäres Skalarprodukt auf $\mathbb{C}^n$ definiert.

c) Es sei $\mathfrak{H}$ der komplexe Vektorraum der auf $[a,b]$ komplexwertigen stetigen Funktionen einer reellen Variablen. Dann definiert

$$\langle f,g\rangle := \int_a^b f(t)\,\overline{g(t)}\,dt$$

ein unitäres Skalarprodukt auf $\mathfrak{H}$.

(Auch das erledigt wieder Kollege Freistühler für uns.)

Definition 5.1.5.

$\mathfrak{H} \ni x,y$ *„orthogonal“* $:\Longleftrightarrow x \perp y :\Longleftrightarrow \langle x,y\rangle = 0$,

$\|z\| := \langle z,z\rangle^{1/2}$ $(z \in \mathfrak{H})$

Für eine nichtleere Menge I und $x_i \in \mathfrak{H}$ für $i \in I$:

$(x_i)_I$ *„orthonormal“ („Orthonormalsystem“, „ONS“)*

$:\Longleftrightarrow \langle x_i, x_j\rangle = \delta_i^j$ $(i,j \in I)$

Den Begriff *„Orthonormalbasis“* verwenden wir bewußt *nicht,* da er in algebraisch orientierten Lehrbüchern (z. B. KOWALSKY, HOLMANN, FISCHER) meist anders benutzt wird als in analytisch orientierten Lehrbüchern (z. B. FLORET, RUDIN, SCHÄFER)!

Bemerkung 5.1.6.

Es seien I eine nicht-leere Menge, $(x_i)_I$ ein ONS; $y \in \mathfrak{H}$; $I \supset J$ endlich und $\alpha_i \in \mathbb{K}$ $(i \in J)$. Dann gelten:

a) $\left\| y - \sum_J \alpha_i x_i \right\|^2 = \|y\|^2 - \sum_J |\langle y, x_i\rangle|^2 + \sum_J \left|\alpha_i - \langle y, x_i\rangle\right|^2$

b) $\ell.\,S.$ *(strikt) minimal* $\iff$ $\alpha_i = \langle y, x_i\rangle$ $(i \in J)$

c) $\left\| y - \sum_J \langle y, x_i\rangle x_i \right\|^2 = \|y\|^2 - \sum_J \left|\langle y, x_i\rangle\right|^2$

d) $\left\| \sum_J \alpha_i x_i \right\|^2 = \sum_J |\alpha_i|^2$

Beweis:

a) $\overset{\checkmark}{\Longrightarrow}$ *b), c), d)*

a): $\ell.\,S. = \|y\|^2 - \sum_J \alpha_i \langle x_i, y\rangle - \sum_J \overline{\alpha_i}\langle y, x_i\rangle + \sum_J |\alpha_i|^2 \overset{\checkmark}{=} r.\,S.$ □

Nach b) liefert

$$v := \sum_J \langle y, x_i\rangle x_i$$

die *beste Approximation* in $\langle x_i \mid i \in J\rangle$ zu einem gegebenen y.
Die Faktoren $\langle y, x_i\rangle$ werden oft als FOURIER-*Koeffizienten* bezeichnet.

Folgerung 5.1.7.

Für eine eine nicht-leere Menge I, ein ONS $(x_i)_I$ und $y \in \mathfrak{H}$ hat man:

a) $\sum_I |\langle y, x_i\rangle|^2 \le \|y\|^2$ (BESSEL-Ungleichung)

b) $\{i \in I : \langle y, x_i\rangle \neq 0\}$ *ist (höchstens) abzählbar.*

Dabei ist $\sum_I \gamma_i := \sup\left\{ \sum_J \gamma_i : I \supset J \text{ endlich} \right\}$ für $\gamma_i \in [0, \infty[$ gesetzt.

Beweis:

a): Nach Teil *c)* der Bemerkung 5.1.6

b): Nach *a)* ist für jedes $n \in \mathbb{N}$ ist die Menge $\left\{i \in I : |\langle y, x_i\rangle| \ge \frac{1}{n}\right\}$ endlich. Das impliziert offenbar die Behauptung.

□

Bemerkung 5.1.8.

Mit $x, y \in \mathfrak{H}$ gelten:

a) $\big|\langle x, y\rangle\big| \leq \|x\|\,\|y\|$ (SCHWARZ-**Ungleichung**)

b) $\|\ \|\colon \mathfrak{H} \longrightarrow [0, \infty[$ *ist eine Halbnorm.*

c) $\|\ \|$ *ist genau dann eine Norm, wenn* $\langle\ ,\ \rangle$ *ein Skalarprodukt ist.*

d) $\|x+y\|^2 + \|x-y\|^2 = 2\big(\|x\|^2 + \|y\|^2\big)$

(**„Parallelogrammgleichung“**)

In einem Raum mit Semiskalarprodukt betrachten wir — wenn nichts anderes gesagt — ‚immer‘ diese, die *zugeordnete,* Halbnorm.

‚Sicherheitshalber‘ notieren wir noch die Definition der Begriffe *Halbnorm* und *Norm:*

Definition 5.1.9.

$\mathfrak{V} := (\mathfrak{V}, a, s, \|\ \|)$ *„Normierter Vektorraum“* über $\mathbb{K} :\Longleftrightarrow$
$(\mathfrak{V}, a, s)$ Vektorraum über $\mathbb{K}$ und $\|\ \|$ *„Norm“* auf X, d. h.:

(N0) $\|\ \|\colon X \ni x \longmapsto \|x\| \in [0, \infty)$ mit

(N1) $\|x\| = 0 \implies x = 0$

(N2) $\|\alpha x\| = |\alpha|\,\|x\|$

(N3) $\|x+y\| \leq \|x\| + \|y\|$ (*für* $x, y \in \mathfrak{V}$ *und* $\alpha \in \mathbb{K}$)

Ohne Forderung *(N1)*:

„Halbnorm“, *„Halbnormierter Vektorraum“*

Beweis (der Bemerkung 5.1.8):

a): Falls $\|x\| = \|y\| = 0$: Mit $\alpha := \langle x, y\rangle$:
$0 \leq \|x - \alpha y\|^2 = \|x\|^2 - \overline{\alpha}\langle x, y\rangle - \alpha\langle y, x\rangle + |\alpha|^2\|y\|^2 = -2|\langle x, y\rangle|^2$
Sonst: Œ $\|x\| > 0$; in Teil a) der Bemerkung 5.1.6 betrachten wir $I := \{1\}$ und $x_1 := \frac{1}{\|x\|}x$. (Dieser Beweis ist eine Idee raffinierter als der übliche, da wir nur ein *Semi*skalarprodukt voraussetzen.)

b): $\|x\| \geq 0$ und $\|\alpha x\| = |\alpha|\,\|x\|$ für $\alpha \in \mathbb{K}$: ✓

$$\|x+y\|^2 = \langle x+y, x+y\rangle \overset{\checkmark}{\leq} \|x\|^2 + 2|\langle x, y\rangle| + \|y\|^2 \overset{a)}{\leq} \big(\|x\| + \|y\|\big)^2$$

c): ✓

Bildchen! *d):* $\ell.S$ ausrechnen ∘ ∘ ∘ □

Definition 5.1.10.

Es sei $\langle\ ,\ \rangle\colon \mathfrak{H} \times \mathfrak{H} \longrightarrow \mathbb{C}$ ein Skalarprodukt und $(a_1, \ldots, a_n)$ eine Basis von $\mathfrak{H}$. Dann heißt die Matrix $C = (\langle a_\nu, a_\mu\rangle)_{1 \leq \nu, \mu \leq n}$ *Matrix des Skalarproduktes* $\langle\ ,\ \rangle$ bezüglich der gegebenen Basis.

Satz 5.1.11.

Die Matrix C eines unitären *Skalarproduktes ist* hermitesch, *d. h. es gilt*

$$c^\mu_\nu \;=\; \overline{c^\nu_\mu} \quad (\dots)$$

Die eines euklidischen *Skalarproduktes ist* symmetrisch.

Beweis: ∘ ∘ ∘ □

Beispiele

(B1) Bezüglich der Standardbasis ist die Matrix des Standardskalarproduktes auf $\mathbb{R}^n$ gleich $\mathbb{1}_n \;=\; (\delta^\mu_\nu)_{1\le\nu,\mu\le n}$.

(B2) Das durch

$$\langle \xi, \eta\rangle \;:=\; \xi^1\eta^1 + 5\xi^1\eta^2 + 5\xi^2\eta^1 + 26\xi^2\eta^2$$

gegebene Skalarprodukt ist bezüglich der Standardbasis des $\mathbb{R}^2$ durch die Matrix

$$\begin{pmatrix} 1 & 5 \\ 5 & 26 \end{pmatrix}$$

und bezüglich der Basis aus den Vektoren $(1,0)^T$ und $(-5,1)^T$ durch die Matrix $\mathbb{1}_2$ dargestellt.

(B3) Es sei $\mathfrak{H}$ der Vektorraum der Polynome vom Grad ≤ 3 mit Basis $(1, x, x^2, x^3)$. Dann ist das durch

$$\langle p, q\rangle = \int_0^1 p(x)q(x)\,dx$$

gegebene Skalarprodukt durch die Matrix

$$\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} \end{pmatrix}$$

dargestellt.

◁

Bemerkung 5.1.12.

Jedes ONS $(x_i)_I$ ist linear unabhängig.

Beweis: Es seien $I \supset J$ endlich und $\alpha_j \in \mathbb{K}$ für $j \in J$ mit

$$\sum_{j\in J} \alpha_j\, x_j \;=\; 0\,.$$

Für $i \in J$ hat man dann

$$0 = \left\langle \sum_{j\in J} \alpha_j\, x_j, x_i \right\rangle = \sum_{j\in J} \alpha_j \,\langle x_j, x_i\rangle = \alpha_i\,.$$

□

Bemerkung 5.1.13.

Ist $(x_i)_J$ ein endliches ONS, dann gilt für $y \in \mathfrak{H}$ und

$$v := \sum_{j\in J} \langle y, x_j\rangle\, x_j$$

$y - v \perp x_i \quad (i \in J)$

Beweis:
$\langle y - v, x_i\rangle = \langle y, x_i\rangle - \sum_{j\in J}\langle y, x_j\rangle\,\langle x_j, x_i\rangle = \langle y, x_i\rangle - \langle y, x_i\rangle = 0$ □

5.2. Orthonormalisierung und Orthogonalraum.

Es seien $(\mathfrak{H}, \langle\ ,\ \rangle)$ ein Prä-HILBERT-Raum, $n \in \mathbb{N}$ und $(a_1, \ldots, a_n) \in \mathfrak{H}^n$ linear unabhängig.

Satz 5.2.1. *(GRAM-SCHMIDT-Verfahren)*

Es existieren $e_1, \ldots, e_n \in \mathfrak{H}$ mit $\langle e_i, e_j\rangle = \delta_i^j \quad (\ldots)$ und

$$\langle a_1, \ldots, a_\nu\rangle = \langle e_1, \ldots, e_\nu\rangle \quad \textit{für alle } \nu = 1, \ldots, n\,.$$

Beweis: n=1: Da — nach Voraussetzung — (a_1) linear unabhängig ist, gilt $a_1 \neq 0$ und damit $\|a_1\| > 0$. $e_1 := \frac{1}{\|a_1\|} a_1$ leistet das Gewünschte.

$n \rightsquigarrow n+1$: Sind $e_1, \ldots, e_n$ schon mit … bestimmt, dann betrachten wir

$$e'_{n+1} = a_{n+1} - \sum_{\nu=1}^{n} \langle a_{n+1}, e_\nu\rangle\, e_\nu\,.$$

Nach der vorangehenden Bemerkung gilt

$$e'_{n+1} \perp e_\nu \quad \text{für alle } \nu = 1, \ldots, n\,.$$

$a_{n+1} \notin \langle a_1, \ldots, a_n\rangle \underset{(n)}{=} \langle e_1, \ldots, e_n\rangle$ zeigt $e'_{n+1} \neq 0$.

$$e_{n+1} := \frac{1}{\|e'_{n+1}\|} e'_{n+1} \in \langle e_1, \ldots, e_n, a_{n+1}\rangle \underset{(n)}{=} \langle a_1, \ldots, a_n, a_{n+1}\rangle$$

Damit folgt $\langle e_1, \ldots, e_n, e_{n+1}\rangle \subset \langle a_1, \ldots, a_n, a_{n+1}\rangle$.
$a_{n+1} \in \langle e_1, \ldots, e_n, e'_{n+1}\rangle = \langle e_1, \ldots, e_n, e_{n+1}\rangle$ zeigt mit der Aussage für n:
$\langle a_1, \ldots, a_n, a_{n+1}\rangle \subset \langle e_1, \ldots, e_n, e_{n+1}\rangle$. □

Es seien M und N Teilmengen von $\mathfrak{H}$.

Definition 5.2.2.

M und N heißen genau dann *„orthogonal“*, notiert als $M \perp N$, wenn $x \perp y$ für alle $x \in M$ und $y \in N$ gilt. Wir schreiben auch $z \perp N$ statt $\{z\} \perp N$ für ein $z \in \mathfrak{H}$.

Bemerkung 5.2.3.

$$M \perp N \iff N \perp M \iff \langle M \rangle \perp N \iff \langle M \rangle \perp \langle N \rangle$$

Beweis: ○ ○ ○ □

Definition 5.2.4.

$$M^\perp := \{x \in \mathfrak{H} \mid x \perp M\}$$

heißt *„orthogonales Komplement (von M)“*

Bemerkung 5.2.5.

$M^\perp$ ist ein Unterraum *von $\mathfrak{H}$ mit $M^\perp = \langle M \rangle^\perp$.*

Beweis: Liest man unmittelbar aus Bemerkung 5.2.3 ab. □

Somit können wir uns auf orthogonale Komplemente von Unterräumen beschränken.

Definition 5.2.6.

Ist $\mathfrak{U}$ Unterraum von $\mathfrak{H}$ und $x \in \mathfrak{H}$, dann heißt ein Vektor $u \in \mathfrak{U}$ genau dann *„orthogonale Projektion“* von x in $\mathfrak{U}$, wenn $x - u \in \mathfrak{U}^\perp$ gilt.

Bildchen!

Bemerkung 5.2.7.

Es gibt höchstens eine orthogonale Projektion von x in $\mathfrak{U}$.

Beweis: ○ ○ ○ □

Satz 5.2.8.

Für einen endlich-dimensionalen *Unterraum $\mathfrak{U}$ von $\mathfrak{H}$ und $x \in \mathfrak{H}$ gelten:*

a) Es existiert eindeutig die orthogonale Projektion von x in $\mathfrak{U}$.

b) $\mathfrak{H} = \mathfrak{U} \oplus \mathfrak{U}^\perp$

c) $\left(\mathfrak{U}^\perp\right)^\perp = \mathfrak{U}$

Man findet für diese besondere direkte Summe auch das Zeichen $\mathfrak{U} \textcircled{\perp} \mathfrak{U}^\perp$.

Beweis: Œ $\mathfrak{U} \neq \{0\}$

a): Zu $n := \dim \mathfrak{U} (\in \mathbb{N})$ existiert — nach Satz 5.2.1 und Bemerkung 5.1.12 — eine Basis $(e_1, \ldots, e_n)$, die orthonormal ist. Dann ist

$$u := \sum_{\nu=1}^{n} \langle x, e_\nu \rangle \, e_\nu$$

nach (5.1.13), (5.2.3) und (5.2.7) *die* Projektion von x in $\mathfrak{U}$.

b): $\mathfrak{U} \cap \mathfrak{U}^\perp = \{0\}$: ✓

Zu $x \in \mathfrak{H}$ sei u die orthogonale Projektion von x in $\mathfrak{U}$. Dann ist $x = \underbrace{u}_{\in \mathfrak{U}} + \underbrace{(x-u)}_{\in \mathfrak{U}^\perp}$

c): $\mathfrak{U} \perp \mathfrak{U}^\perp$ zeigt: $\mathfrak{U} \subset \left(\mathfrak{U}^\perp\right)^\perp$. Zu $x \in \left(\mathfrak{U}^\perp\right)^\perp$ existieren nach *b*) ein $u \in \mathfrak{U}$ und ein $v \in \mathfrak{U}^\perp$ mit $x = u + v$. Dann hat man $x \perp v$ und so $v = x - u \perp v$, also $v = 0$ und somit $x = u \in \mathfrak{U}$. □

Wir beschließen diesen Abschnitt mit einem *einfachen* und durchsichtigen Beweis der zurückgestellten Aussage, daß die Dimension des Spaltenraums einer Matrix gleich der Dimension des Zeilenraums ist. Wir beschränken uns dabei hier auf den Fall $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, notieren die Dinge jedoch der besseren Übersichtlichkeit halber nur für $\mathbb{K} = \mathbb{R}$. Die erforderlichen kleinen Änderungen für $\mathbb{K} = \mathbb{C}$ wird der Leser (hoffentlich!) sofort übersehen.

Zur Erinnerung: Zu $n, m \in \mathbb{N}$ und $A = (a^\mu_\nu) \in \mathbb{K}^{m\times n}$ hatten wir

$$\operatorname{rang} A := \operatorname{rang} f_A = \dim\langle A_1, \ldots, A_n\rangle$$

mit den *Spalten* $A_1, \ldots, A_n$ definiert. Wir betrachten das homogene Gleichungssystem

$$(*) \qquad \begin{array}{ccccccc} a^1_1 x^1 & + & \cdots & + & a^1_n x^n & = & 0 \\ \vdots & & & & \vdots & & \vdots \\ a^m_1 x^1 & + & \cdots & + & a^m_n x^n & = & 0 \end{array}$$

Mit den *Zeilen* $A^1, \ldots, A^m$ sind für $x = (x^1, \ldots, x^n)^T \in \mathbb{K}^n$ offenbar *äquivalent:*

- x ist Lösung von $(*)$
- $\sum_{\nu=1}^{n} x^\nu A_\nu = 0$
- $x \perp \langle A^1, \ldots, A^m\rangle$

Für die lineare Abbildung $L\colon \begin{array}{ccc} \mathbb{K}^n & \longrightarrow & \mathbb{K}^m \\ \cup & & \cup \\ x & \longmapsto & \sum_{\nu=1}^{n} x^\nu A_\nu \end{array}$ hat man also

$\operatorname{im} L = \langle A_1, \ldots, A_n\rangle$ und $\ker L = \{x \in \mathbb{K}^n \mid x \text{ ist Lösung von } (*)\}$.

Nach Satz 4.2.1 gilt

$$\dim \ker L + \dim \operatorname{im} L = n\,.$$

Nach Teil *b*) aus Satz 5.2.8 weiß man:

$$\dim\langle A^1, \ldots, A^m\rangle + \dim\langle A^1, \ldots, A^m\rangle^\perp = n$$

Damit haben wir erhalten:

Zeilenrang $:= \dim\langle A^1, \ldots, A^m\rangle = \dim\langle A_1, \ldots, A_n\rangle =:$ *Spaltenrang*

Satz 5.2.9.

Für jede Matrix ist der Zeilenrang gleich dem Spaltenrang.

Winkel in euklidischen Räumen

Es sei nun $(\mathfrak{H}, \langle\ ,\ \rangle)$ ein euklidischer Vektorraum. Für $v, w \in \mathfrak{H} \setminus \{0\}$ gilt nach der SCHWARZschen Ungleichung $\big|\langle v, w\rangle\big| \leq \|v\|\,\|w\|$, also

$$-1 \leq \frac{\langle v, w\rangle}{\|v\|\,\|w\|} \leq 1\,.$$

Daher existiert genau ein $\varphi \in [0, \pi]$ mit

$$\cos\varphi = \frac{\langle v, w\rangle}{\|v\|\,\|w\|}\,.$$

Diesen Winkel bezeichnen wir mit

$$\measuredangle(v, w) \qquad (\text{„Winkel zwischen } v \text{ und } w\text{“})\,.$$

Hierbei ist die Funktion cos z. B. über ihre Potenzreihe definiert.

Über eine Determinantenfunktion kann auf dem $\mathbb{R}^2$ eine Orientierung festgelegt werden. Damit können dann auch *orientierte Winkel* definiert werden. Wir gehen darauf aber im Rahmen dieser Vorlesung nicht ein.

5.3. Lineare Abbildungen auf Räumen mit Skalarprodukt.

Es seien $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $(\mathfrak{H}, \langle\ ,\ \rangle), (\mathfrak{K}, \langle\ ,\ \rangle)$ Prä-HILBERT-Räume[6] über $\mathbb{K}$.

Aus Zeitgründen und der Übersichtlichkeit halber setzen wir in diesem Abschnitt durchweg voraus, daß beide Räume $\mathfrak{H}$ und $\mathfrak{K}$ *endlich-dimensional* sind.

Definition 5.3.1.

Zu $y \in \mathfrak{H}$ ist die Abbildung y^*: $\begin{array}{ccc} \mathfrak{H} & \longrightarrow & \mathbb{K} \\ \cup & & \cup \\ x & \longmapsto & \langle x, y\rangle \end{array}$ linear, also ein Element des Dualraums $\mathfrak{H}^*$. Wir notieren diese Abbildung auch suggestiv als $\langle\ , y\rangle$. Wir betrachten damit die Abbildung:

$$\Phi\colon \begin{array}{ccc} \mathfrak{H} & \longrightarrow & \mathfrak{H}^* \\ \cup & & \cup \\ y & \longmapsto & \langle\ , y\rangle = y^* \end{array}$$

Bemerkung 5.3.2.

a) Φ ist konjugiert linear, d. h. für alle $\alpha, \beta \in \mathbb{K}$ und $y, z \in \mathfrak{H}$ gilt:

$$\Phi(\alpha y + \beta z) = \overline{\alpha}\Phi(y) + \overline{\beta}\Phi(z)$$

b) Φ ist bijektiv.

Damit sind $\mathfrak{H}$ und $\mathfrak{H}^*$ im wesentlichen gleich.

[6] Wir unterscheiden also die beiden Skalarprodukte nicht in der Notierung.

Beweis:

a) $\ell.\mathrm{S.}(x) = \langle x, \alpha y + \beta z\rangle = \overline{\alpha}\langle x, y\rangle + \overline{\beta}\langle x, z\rangle = \mathrm{r.S.}(x)$

b) *Injektivität:* $\Phi(y) = \Phi(z)$ zeigt nach *a)* $\Phi(y - z) = 0$, insbesondere $0 = \Phi(y - z)(y - z) = \langle y - z, y - z\rangle$, damit $y - z = 0$, d. h. $y = z$.
Surjektivität: Es sei $f \in \mathfrak{H}^*$ und Œ $f \neq 0$. Dann ist $\dim \operatorname{im} f = 1$, also nach Satz 4.2.1 $\dim(\ker f) = \dim \mathfrak{H} - 1$ und so $\dim((\ker f)^\perp) = 1$ nach Satz 5.2.8. Für ein festes $v \in (\ker f)^\perp \setminus \{0\}$ und $\alpha \in \mathbb{K}$ gilt $\langle v, \alpha v\rangle = \overline{\alpha}\|v\|^2$, also speziell mit $\alpha := \overline{f(v)}/\|v\|^2$ und $y := \alpha v$: $(\Phi(y))(v) = y^*(v) = \langle v, y\rangle = f(v)$. Dies gilt, da $(\ker f)^\perp$ eindimensional ist, auch für alle $v \in (\ker f)^\perp$ und — trivialerweise — für $v \in \ker f$. Somit $\Phi(y) = f$. □

Betrachtet man für $n \in \mathbb{N}$ im $\mathbb{K}^n$ das ‚übliche' Skalarprodukt, dann existiert nach (5.3.2) zu jedem $f \in (\mathbb{K}^n)^*$ eindeutig ein $y \in \mathbb{K}^n$ mit $f = \Phi(y) = \langle\ , y\rangle$, also für alle $x \in \mathbb{K}^n$

$$f(x) = \langle x, y\rangle = \overline{y}^T x\,.$$

Schreibt man — wie vereinbart — die Vektoren des $\mathbb{K}^n$ als Spaltenvektoren, dann *entsprechen also die Elemente aus* $(\mathbb{K}^n)^*$ *genau den Zeilenvektoren (mit n Komponenten aus* $\mathbb{K}$*).*

Satz 5.3.3.

Zu einer linearen Abbildung $\varphi\colon \mathfrak{H} \longrightarrow \mathfrak{K}$ *existiert eindeutig* $\varphi^*\colon \mathfrak{K} \longrightarrow \mathfrak{H}$ *linear mit*

$$\langle \varphi v, w\rangle = \langle v, \varphi^* w\rangle$$

für alle $v \in \mathfrak{H}$ *und* $w \in \mathfrak{K}$*. Mit einer Basis* $(e_1, \ldots, e_n)$ *von* $\mathfrak{H}$*, die orthonormal ist, gilt für alle* $w \in \mathfrak{K}$*:*

$$\varphi^* w = \sum_{\nu=1}^{n} \langle w, \varphi e_\nu\rangle\, e_\nu \qquad (*)$$

Wir bezeichnen φ^* als *Adjungierte* zu φ.

Beweis: Für $v \in \mathfrak{H}$ hat man die Darstellung

$$v = \sum_{\nu=1}^{n} \langle v, e_\nu\rangle\, e_\nu\,.$$

Ausgehend von $\langle \varphi v, w\rangle = \langle v, \varphi^* w\rangle$ für $v \in \mathfrak{H}$ und $w \in \mathfrak{K}$, folgt

$$\langle v, \varphi^* w\rangle = \sum_{\nu=1}^{n} \langle v, e_\nu\rangle\langle e_\nu, \varphi^* w\rangle = \sum_{\nu=1}^{n} \langle v, e_\nu\rangle\langle \varphi e_\nu, w\rangle = \left\langle v, \sum_{\nu=1}^{n} \langle w, \varphi e_\nu\rangle e_\nu \right\rangle,$$

also die Beziehung $(*)$ und damit die *Eindeutigkeit.* Definiert man andererseits $\varphi^* w$ durch $(*)$, so ist φ^* *linear* mit

$$\langle \varphi v, w\rangle = \sum_{\nu=1}^{n} \langle v, e_\nu\rangle\langle \varphi e_\nu, w\rangle = \sum_{\nu=1}^{n} \overline{\langle w, \varphi e_\nu\rangle}\,\langle v, e_\nu\rangle = \langle v, \varphi^* w\rangle\,.$$

□

Wir stellen grundlegende Eigenschaften dieser neuen Begriffsbildung zusammen. Dazu seien $\varphi, \psi \in \mathcal{L}(\mathfrak{H}, \mathfrak{K}), \alpha \in \mathbb{K}$ und $\chi \in \mathcal{L}(\mathfrak{K}, \mathfrak{L})$ mit einem weiteren Prä-HILBERT-Raum $\mathfrak{L}$ gegeben:

Bemerkung 5.3.4.

a) $(\alpha\varphi + \psi)^* = \overline{\alpha}\varphi^* + \psi^*$

b) $(\chi\varphi)^* = \varphi^*\chi^*$

c) $(\mathrm{id}_{\mathfrak{H}})^* = \mathrm{id}_{\mathfrak{H}}$

d) Ist φ invertierbar, dann auch φ^* mit $(\varphi^{-1})^* = (\varphi^*)^{-1}$.

e) $\varphi^{**} := (\varphi^*)^* = \varphi$

f) $\ker \varphi^* = (\operatorname{im} \varphi)^\perp$

g) $\ker \varphi = (\operatorname{im} \varphi^*)^\perp$

h) φ ist genau dann surjektiv, wenn φ^* injektiv ist.

Beweis:

a), b), c), e): ∘ ∘ ∘

f): Für $w \in \mathfrak{K}$: $w \in \ker \varphi^* \iff \varphi^* w = 0 \iff \forall v \in \mathfrak{H}\ \langle v, \varphi^* w\rangle = 0$
$\iff \forall v \in \mathfrak{H}\ \langle \varphi v, w\rangle = 0 \iff w \in (\operatorname{im}\varphi)^\perp$

g): nach *e)* und *f)*

h): φ ist genau dann surjektiv, wenn $\operatorname{im}\varphi = \mathfrak{K}$ gilt. Das bedeutet aber gerade $\ker \varphi^* \underset{f)}{=} (\operatorname{im}\varphi)^\perp = \{0\}$ also die Injektivität von φ^*.

d): $\varphi^{-1}\varphi = \mathrm{id}_{\mathfrak{H}}$ liefert
$\mathrm{id}_{\mathfrak{H}} \underset{c)}{=} (\mathrm{id}_{\mathfrak{H}})^* = (\varphi^{-1}\varphi)^* \underset{b)}{=} \varphi^*(\varphi^{-1})^*$, somit: $(\varphi^{-1})^* = (\varphi^*)^{-1}$ □

Um einen *Bezug zu den darstellenden Matrizen* herzustellen, benötigen wir noch die *adjungierte Matrix*

$$A^* := \overline{A}^T$$

zu einer Matrix $A = (a_\nu^\mu)_{\substack{1\le\mu\le m\\ 1\le\nu\le n}} \in \mathbb{K}^{m\times n}$. Im Reellen gilt also $A^* = A^T$. Dabei bedeutet natürlich

$$\overline{A} := \left(\overline{a_\nu^\mu}\right)_{\substack{1\le\mu\le m\\ 1\le\nu\le n}}.$$

Die adjungierte Matrix entsteht also durch Vertauschen von Zeilen und Spalten und zusätzlicher Konjugation der Matrixelemente.

Bemerkung 5.3.5.

Es seien — mit $n, r \in \mathbb{N}$ — $\mathcal{E} := (e_1, \ldots, e_n)$ eine orthonormierte Basis von $\mathfrak{H}$, $\mathcal{F} := (f_1, \ldots, f_r)$ eine orthonormierte Basis von $\mathfrak{K}$ und $\varphi\colon \mathfrak{H} \longrightarrow \mathfrak{K}$ linear. Mit $A := \mathcal{M}_{\mathcal{E}}^{\mathcal{F}}(\varphi)$ gilt dann

$$\mathcal{M}_{\mathcal{F}}^{\mathcal{E}}(\varphi^*) = A^*.$$

Die darstellenden Matrizen zu den adjungierten Abbildungen sind also bezüglich eines beliebigen Paares von orthonormierten Basen ebenfalls zueinander adjungiert.

Es paßt also auch hier wieder alles zusammen. ☺

Beweis: Die Matrix $A = (a_\nu^\varrho)_{\substack{1\le\varrho\le r\\ 1\le\nu\le n}}$ ist bestimmt durch:

$$\varphi\, e_\nu \;=\; \sum_{\varrho=1}^{r} a_\nu^\varrho\, f_\varrho \qquad (\nu = 1,\ldots,n)$$

Entsprechend gilt für $B := \mathcal{M}_{\mathcal{F}}^{\mathcal{E}}(\varphi^*) = \left(b_\varrho^\nu\right)_{\substack{1\le\nu\le n\\ 1\le\varrho\le r}}$:

$$\varphi^*\, f_\varrho \;=\; \sum_{\nu=1}^{n} b_\varrho^\nu\, e_\nu \qquad (\varrho = 1,\ldots,r)$$

Damit hat man für $\nu = 1,\ldots,n$ und $\kappa = 1,\ldots,r$:

$$\langle \varphi e_\nu, f_\kappa\rangle \;=\; \left\langle \sum_{\varrho=1}^{r} a_\nu^\varrho f_\varrho, f_\kappa \right\rangle \;=\; \sum_{\varrho=1}^{r} a_\nu^\varrho\, \langle f_\varrho, f_\kappa\rangle \;=\; a_\nu^\kappa$$

und

$$\langle e_\nu, \varphi^* f_\kappa\rangle \;=\; \left\langle e_\nu, \sum_{\lambda=1}^{n} b_\kappa^\lambda e_\lambda \right\rangle \;=\; \sum_{\lambda=1}^{n} \overline{b_\kappa^\lambda}\, \langle e_\nu, e_\lambda\rangle \;=\; \overline{b_\kappa^\nu}$$

□

Es sei $\varphi \in \mathcal{L}(\mathfrak{H},\mathfrak{H})$.

Definition 5.3.6.

φ heißt genau dann *„normal“*, wenn $\varphi\varphi^* = \varphi^*\varphi$ gilt. φ heißt genau dann *„selbstadjungiert“* oder *„hermitesch“*, wenn $\varphi = \varphi^*$ ist.

φ heißt genau dann *„Isometrie“*, wenn für alle $v, w \in \mathfrak{H}$

$$\langle \varphi v, \varphi w\rangle \;=\; \langle v, w\rangle$$

gilt. Isometrien nennt man für $\mathbb{K} = \mathbb{R}$ auch *„orthogonal“*, im Falle $\mathbb{K} = \mathbb{C}$ auch *„unitär“*.

Natürlich kann man entsprechend Isometrien auch für lineare Abbildungen von $\mathfrak{H}$ nach $\mathfrak{K}$ definieren.

Trivialität 5.3.7.

Jeder selbstadjungierte Endomorphismus ist normal.

Bemerkung 5.3.8.

φ ist genau dann normal, wenn für alle $v, w \in \mathfrak{H}$ gilt

$$\langle \varphi v, \varphi w\rangle \;=\; \langle \varphi^* v, \varphi^* w\rangle\ .$$

Beweis: Ist φ normal, dann folgt:

$$\langle \varphi v, \varphi w\rangle \;=\; \langle v, \varphi^*\varphi w\rangle \;=\; \langle v, \varphi\,\varphi^* w\rangle \;=\; \langle \varphi^* v, \varphi^* w\rangle$$

Umgekehrt hat man

$$\langle \varphi^*\varphi v, w\rangle \;=\; \langle \varphi v, \varphi w\rangle \;=\; \langle \varphi^* v, \varphi^* w\rangle \;=\; \langle \varphi\varphi^* v, w\rangle\,,$$

folglich $\varphi^*\varphi v = \varphi\,\varphi^* v$, d. h. $\varphi^*\varphi = \varphi\,\varphi^*$. □

Bemerkung 5.3.9. *Die folgenden Aussagen sind* äquivalent:

a) φ *ist eine Isometrie.* *(Skalarprodukt erhaltend)*

b) $\|x\| = 1$ *impliziert* $\|\varphi x\| = 1$ *für* $x \in \mathfrak{H}$.

c) $\|\varphi x\| \;=\; \|x\|$ *für alle* $x \in \mathfrak{H}$. *(normerhaltend)*

d) Ist $(a_1, \ldots, a_k)$ *— für ein* $k \in \mathbb{N}$ *— ein ONS in* $\mathfrak{H}$*, dann ist auch* $(\varphi a_1, \ldots, \varphi a_k)$ *ein ONS.*

e) φ *ist ein Isomorphismus mit* $\varphi^{-1} = \varphi^*$.

Beweis: $a) \Longrightarrow b)$: ✓ $a) \Longrightarrow d)$: ✓

$b) \Longrightarrow c)$: Œ $x \neq 0$, dann x ‚normieren'

$c) \Longrightarrow a)$: vgl. Übung (9.4) (‚Polarisierungsformeln')

$d) \Longrightarrow b)$: $k = 1$ und $a_1 := x$ wählen

$e) \Longrightarrow a)$: Für $x, y \in \mathfrak{H}$: $\langle x, y\rangle = \langle x, \varphi^{-1}\varphi y\rangle = \langle x, \varphi^*\varphi y\rangle = \langle \varphi x, \varphi y\rangle$

$a) \Longrightarrow e)$: Für $x, y \in \mathfrak{H}$: $\langle x, y\rangle = \langle \varphi x, \varphi y\rangle = \langle x, \varphi^*\varphi y\rangle$. Da dies für alle $x \in \mathfrak{H}$ gilt, folgt $\varphi^*\varphi y = y$, also $\mathrm{id}_{\mathfrak{H}} = \varphi^*\varphi$.

□

Bemerkung 5.3.10.

Die Menge aller Isometrien von $\mathfrak{H}$ *bildet bezüglich der Komposition eine Gruppe.* Im Falle $\mathbb{K} = \mathbb{R}$ wird sie *„orthogonale Gruppe" von* $\mathfrak{H}$ genannt und mit $O(\mathfrak{H})$ notiert. Im Falle $\mathbb{K} = \mathbb{C}$ wird sie *„unitäre Gruppe" von* $\mathfrak{H}$ genannt und mit $U(\mathfrak{H})$ notiert.

Beweis: ∘ ∘ ∘ □

Definition 5.3.11. Es sei $n \in \mathbb{N}$.

Eine reelle $(n \times n)$-Matrix A heißt genau dann *„orthogonal"* , wenn

$$A^T A \;=\; A A^T \;=\; \mathbb{1}_n$$

gilt. Eine solche Matrix ist also invertierbar mit $A^{-1} = A^T$.

Entsprechend heißt eine Matrix $A \in \mathbb{C}^{n\times n}$ mit $A^* A = \mathbb{1}_n$ *„unitär"*.

$A \in \mathbb{K}^{n\times n}$ heißt genau dann *„hermitesch"* (vgl. dazu auch Satz 5.1.11), wenn $A = A^*$ gilt. Für $\mathbb{K} = \mathbb{R}$ bedeutet dies gerade $A = A^T$, man sagt dann auch: A ist *„symmetrisch"* .

Satz 5.3.12.

Es seien $\dim \mathfrak{H} =: n$, $(b_1, \ldots, b_n)$ *ein ONS in* $\mathfrak{H}$ — *nach Bemerkung 5.1.12 ist dann* $(b_1, \ldots, b_n)$ *eine Basis von* $\mathfrak{H}$ — *und*

$$d_\nu := \sum_{\mu=1}^{n} a_\nu^\mu \, b_\mu \qquad (\nu = 1, \ldots, n)$$

zu $A = (a_\nu^\mu)_{1 \le \nu, \mu \le n} \in \mathbb{K}^{n \times n}$ *gebildet. Dann ist* $(d_1, \ldots, d_n)$ *genau dann wieder ein ONS in* $\mathfrak{H}$, *wenn*

$$A^* A = \mathbb{1}_n$$

gilt.

Den Beweis für eine Richtung könnte man aus Bemerkung 5.3.9 ablesen. Es scheint mir hier jedoch angemessen und verständlicher zu sein, einen (einfachen) unabhängigen Beweis zu geben.

Beweis: Für $\nu, \varrho = 1, \ldots, n$ hat man:

$$\langle d_\nu, d_\varrho \rangle = \left\langle \sum_{\mu=1}^{n} a_\nu^\mu b_\mu, \sum_{\kappa=1}^{n} a_\varrho^\kappa b_\kappa \right\rangle = \sum_{\mu,\kappa=1}^{n} a_\nu^\mu \, \overline{a_\varrho^\kappa} \, \langle b_\mu, b_\kappa \rangle = \sum_{\mu=1}^{n} a_\nu^\mu \, \overline{a_\varrho^\mu} \, .$$

Das ist aber gerade das Element in der ϱ-ten Zeile und ν-ten Spalte von $A^* A$. Daraus liest man die Behauptung unmittelbar ab. □

Bemerkung 5.3.13.

Es seien wieder — mit $n \in \mathbb{N}$ — $\mathcal{E} := (e_1, \ldots, e_n)$ *eine orthonormierte Basis von* $\mathfrak{H}$ *und* $\varphi\colon \mathfrak{H} \longrightarrow \mathfrak{H}$ *linear. Mit* $A := \mathcal{M}_\mathcal{E}(\varphi)$ *gilt dann:*

a) Im Falle $\mathbb{K} = \mathbb{R}$: φ *ist genau dann orthogonal, wenn A orthogonal ist.*

b) Im Falle $\mathbb{K} = \mathbb{C}$: φ *ist genau dann unitär, wenn A unitär ist.*

c) φ *ist genau dann hermitesch, wenn A hermitesch ist.*

Beweis: a): φ orthogonal $\underset{(5.3.9)}{\iff}$ φ ist ein Isomorphismus mit $\varphi^{-1} = \varphi^*$ $\underset{(5.3.5)}{\iff}$ A ist invertierbar mit $A^{-1} = A^* = A^T \iff A$ ist orthogonal.

b): ‚ebenso'

c): Nach Bemerkung 5.3.5 gilt: $A = A^* \iff \varphi$ ist hermitesch □

Satz 5.3.14.

Eine reelle $(n \times n)$*-Matrix A ist genau dann* orthogonal, *wenn ihr Spalten ein ONS des* $\mathbb{R}^n$ *bilden, und genau dann, wenn ihr Zeilen ein ONS des* $\mathbb{R}^n$ *bilden (natürlich jeweils bezüglich des Standard-Skalarproduktes), und genau dann, wenn die zugehörige Abbildung* f_A *orthogonal ist.*

Beweis: Für $x, y \in \mathbb{R}^n$ gilt:

$$\langle f_A(x), f_A(y)\rangle \;=\; \langle Ax, Ay\rangle \;=\; \langle x, A^T A y\rangle$$

Die Abbildung f_A ist demgemäß genau dann orthogonal, wenn stets $\langle x, y\rangle = \langle x, A^T A y\rangle$ gilt. Die ist aber gleichbedeutend mit $\mathbb{1}_n = A^T A$, also der Orthogonalität von A.
Mit den Spalten $A_1, \ldots, A_n$ von A gelten

$$A = (A_1, \ldots, A_n) \quad \text{und} \quad A^T = \begin{pmatrix} A_1^T \\ \vdots \\ A_n^T \end{pmatrix}, \text{ also}$$

$$A^T A \;=\; \left(A_\mu^T A_\nu\right)_{\substack{1\le\mu\le n\\ 1\le\nu\le n}} \;=\; \left(\langle A_\mu, A_\nu\rangle\right)_{\substack{1\le\mu\le n\\ 1\le\nu\le n}} .$$

Daraus liest man ab: $A^T A = \mathbb{1}_n \iff \langle A_\mu, A_\nu\rangle = \delta_\nu^\mu$ für $1 \le \mu \le n$ und $1 \le \nu \le n$. Da A offenbar genau dann orthogonal ist, wenn A^T orthogonal ist, hat man auch die entsprechende Aussage für die Zeilen. □

In gleicher Weise erhält man:

Satz 5.3.15.

Eine komplexe $(n \times n)$-Matrix A ist genau dann unitär*, wenn ihr Spalten ein ONS des $\mathbb{C}^n$ bilden, und genau dann, wenn ihr Zeilen ein ONS des $\mathbb{C}^n$ bilden (natürlich jeweils bezüglich des Standard-Skalarproduktes), und genau dann, wenn die zugehörige Abbildung f_A unitär ist.*

Definition 5.3.16.

Mit A, B sind offenbar $A^T = A^{-1}$ und AB orthogonal. Die Menge $O(n)$ der orthogonalen $(n \times n)$-Matrizen bildet somit eine Gruppe, *orthogonale Gruppe (n-ten Grades)* genannt. Entsprechend bildet die Menge der unitären $(n \times n)$-Matrizen eine Gruppe $U(n)$, die *unitäre Gruppe (n-ten Grades)*.

Beispiel

(B1) Jede orthogonale (2×2)-Matrix ist von der Form

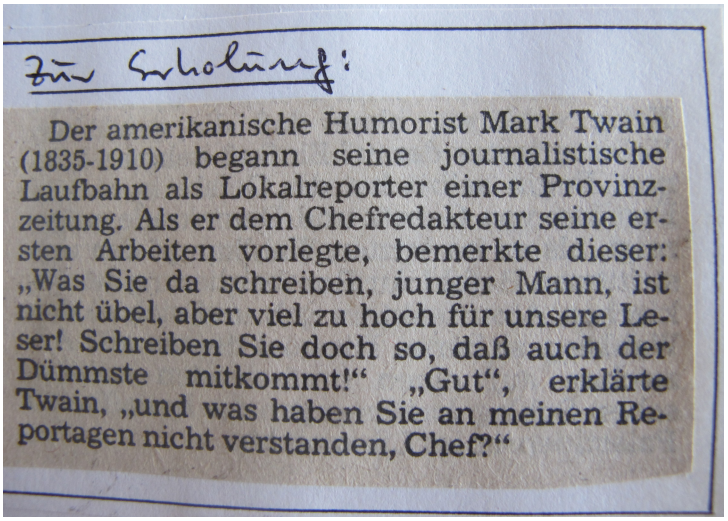
$$D(\vartheta) := \begin{pmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{pmatrix} \text{ oder } S(\vartheta) := \begin{pmatrix} \cos\vartheta & \sin\vartheta \\ \sin\vartheta & -\cos\vartheta \end{pmatrix}$$

für ein geeignetes reelles ϑ. Eine Matrix der linken Form wird als *Drehung* bezeichnet, da sie eine Rotation des $\mathbb{R}^2$ um den Ursprung um den Winkel ϑ beschreibt — vgl. hierzu Übung 10.4. Eine Matrix der rechten Form wird als *Spiegelung* bezeichnet. Sie bewirkt eine Spiegelung an der Geraden durch den Ursprung, die mit der x-Achse den Winkel $\vartheta/2$ einschließt.

Eine Matrix $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$ ist genau dann orthogonal, wenn:

$$a^2 + b^2 = 1 \qquad (1)$$
$$c^2 + d^2 = 1 \qquad (2)$$
$$ac + bd = 0 \qquad (3)$$

Nach (1) und (2) existieren eindeutig $\alpha, \beta \in [0, 2\pi)$ mit $a = \cos\alpha, b = \sin\alpha$ und $c = \sin\beta, d = \cos\beta$. Nach (3) ist $0 = \cos\alpha\sin\beta + \sin\alpha\cos\beta = \sin(\alpha+\beta)$: Damit gilt zunächst $\alpha+\beta = n\pi$ für ein $n \in \mathbb{N}_0$. Es verbleiben — wegen der Periode 2π der beiden Funktionen sin und cos — $\alpha+\beta = 0$ und $\alpha+\beta = \pi$. Das liefert im ersten Fall $A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$, im zweiten Fall $A = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}$ ◁


Zur Erholung:

Der amerikanische Humorist Mark Twain (1835-1910) begann seine journalistische Laufbahn als Lokalreporter einer Provinzzeitung. Als er dem Chefredakteur seine ersten Arbeiten vorlegte, bemerkte dieser: „Was Sie da schreiben, junger Mann, ist nicht übel, aber viel zu hoch für unsere Leser! Schreiben Sie doch so, daß auch der Dümmste mitkommt!“ „Gut“, erklärte Twain, „und was haben Sie an meinen Reportagen nicht verstanden, Chef?“

6. Determinanten

Die *Determinante* ist eine Abbildung (beziehungsweise der Wert dieser Abbildung), die — zu gegebenem $n \in \mathbb{N}$ — jeder $(n \times n)$-Matrix, also einer *quadratischen* Matrix, eine charakterisierende Zahl aus dem Grundkörper $\mathbb{K}$ zuordnet. Es gibt in der Literatur eine Vielzahl von Zugängen zum Determinantenbegriff. Wir sehen uns motivierend und begriffsklärend den einfachen Fall $n = 2$ an, der uns schon implizit in Übungsaufgabe (1.2) begegnet ist.

Grundfragen:
(1) Explizite Auflösungsformel für Gleichungssysteme mit invertierbarer Koeffizientenmatrix
(2) Berechnung der Inversen einer invertierbaren Matrix
(3) Orientierung im $\mathbb{R}^n$
(4) Volumenbestimmung von Parallelotopen (auch Parallelepiped genannt) $\rightsquigarrow$ Analysis: Integralrechnung für Funktionen mehrer Veränderlicher
(5) Eigenwerttheorie von Endomorphismen

Wir werden in dieser *ein*semestrigen Vorlesung jedoch nicht alle diese Themen behandeln.

6.1. **Vorüberlegungen.** Wir hatten in Übung (1.2) schon gesehen: Zu $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{K}^{2\times 2}$ und beliebiger rechter Seite $R = \begin{pmatrix} r \\ s \end{pmatrix} \in \mathbb{K}^2$ ist das Gleichungsssystem

$$Ax = R$$

genau dann eindeutig lösbar, wenn gilt:

$$\det A := \begin{vmatrix} a & b \\ c & d \end{vmatrix} := ad - bc \neq 0$$

Man kann die Determinante auffassen als Abbildungen auf den (2×2)-Matrizen oder auch als Abbildung auf den Spalten A_1, A_2, wir schreiben dann $\det A = \det(A_1, A_2)$. Im Fall $\det A \neq 0$ kann man die Lösung angeben durch:

$$x^1 := \frac{\det(R, A_2)}{\det A} \quad \text{und} \quad x^2 := \frac{\det(A_1, R)}{\det A}$$

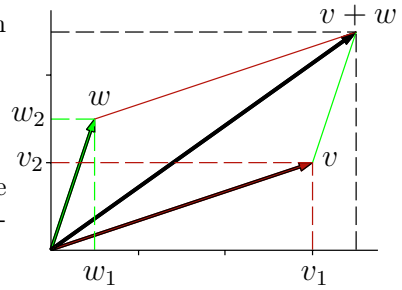
Das ist ein einfacher Spezialfall der Cramer*schen Regel*, die wir später allgemeiner kennenlernen werden.

Einfache Eigenschaften der Determinante lassen sich hier auch anschaulich *durch Flächeninhalt motivieren:*

Im $\mathbb{R}^2$ betrachten wir mit

$$P(v, w) := \{\lambda v + \mu w \mid 0 \leq \lambda, \mu \leq 1\}$$

das durch zwei gegebenen Vektoren v, w aufgespannte Parallelogramm. Das zugehörige ‚Volumen' (Flächeninhalt) läßt sich in ‚naiver Weise' wie folgt berechnen:



$$V(v,w) = (v_1+w_1)(v_2+w_2) - v_1v_2 - w_1w_2 - 2w_1v_2 = v_1w_2 - w_1v_2$$
$$= \det(v,w)$$

Es gelten:

(1) $V(\lambda v, w) = \lambda V(v,w) = V(v, \lambda w)$ $(\lambda \in \mathbb{R})$
$V(v+v', w) = V(v,w) + V(v',w)$ $(v' \in \mathbb{R}^2)$
$V(v, w+w') = V(v,w) + V(v,w')$ $(w' \in \mathbb{R}^2)$
(V ist linear in beiden Spalten.)

(2) $V(v,v) = 0$ (V ist alternierend.)

(3) $V(e_1, e_2) = 1$ (V ist normiert.)

Daß $V(v,w)$ auch negative Werte annehmen kann, sollte Sie nicht stören; das ist ja auch schon bei der Deutung des bestimmten Integrals als Flächeninhalt aufgetreten.

Die o. a. Eigenschaften (1), (2) und (3), die man in diesem einfachen Spezialfall auch aus der Definition leicht abliest, können als Grundlage zur Definition der Determinante genommen werden. Wir gehen also aus von

(1) $\det(\lambda v, w) = \lambda \det(v,w) = \det(v, \lambda w)$ $(\lambda \in \mathbb{K})$
$\det(v+v', w) = \det(v,w) + \det(v',w)$ $(v' \in \mathbb{K}^2)$
$\det(v, w+w') = \det(v,w) + \det(v,w')$ $(w' \in \mathbb{K}^2)$
(det ist linear in beiden Variablen.)

(2) $\det(v,v) = 0$ (det ist alternierend.)

(3) $\det(e_1, e_2) = 1$ (det ist normiert.)

und erhalten:

(4) $\det(\lambda A) = \lambda^2 \det A$ $(\lambda \in \mathbb{K})$

(5) *Addiert man ein Vielfaches einer Spalte zu einer anderen, dann ändert sich der Wert der Determinante nicht*, d. h. für $\lambda \in \mathbb{K}$ gilt

$$\det \begin{pmatrix} a+\lambda b & b \\ c+\lambda d & d \end{pmatrix} = \det \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

und entsprechend für die zweite Spalte.

(6) *Vertauscht man die beiden Spalten von A, dann ändert die Determinante ihr Vorzeichen:*

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = -\det \begin{pmatrix} b & a \\ d & c \end{pmatrix}$$

(7) $\det A = \det A^T$

(8) v, w linear abhängig $\iff \det(v,w) = 0$ $(v, w \in \mathbb{K}^2)$

(9) $\det A \neq 0 \iff A$ ist invertierbar

Ist $\det A \neq 0$, dann gilt: $A^{-1} = \dfrac{1}{\det A} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$

Beweis: (4) folgt unmittelbar aus (1)

(5) ist durch (1) und (2) gegeben:

$$\ell.S. = \det(A_1 + \lambda A_2, A_2) \underset{(1)}{=} \det(A_1, A_2) + \lambda \det(A_2, A_2) \underset{(2)}{=} \det(A_1, A_2)$$

(6) $0 \underset{(2)}{=} \det(A_1 + A_2, A_1 + A_2) \underset{(1),(2)}{=} \det(A_1, A_2) + \det(A_2, A_1)$

(7) $\ell.S. = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & c \\ b & d \end{vmatrix} = r.S.$

(8) $\Longrightarrow$: Man hat $\lambda v + \mu w = 0$ für $\lambda, \mu \in \mathbb{K}$ mit Œ$\lambda \neq 0$, so $v = -\lambda^{-1}\mu w$ und damit $\det(v, w) = 0$ nach (1) und (2).

$\Longleftarrow$: Es seien $0 = \det(v, w) = \begin{vmatrix} v_1 & w_1 \\ v_2 & w_2 \end{vmatrix} = v_1 w_2 - w_1 v_2$ und Œ $v \neq 0$, damit Œ $v_1 \neq 0$. Mit $\mu := -v_1$ und $\lambda := w_1$ folgt $\lambda v_1 + \mu w_1 = 0$ und $\lambda v_2 + \mu w_2 = 0$, also $\lambda v + \mu w = 0$.

(9) $\Longrightarrow$: $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} ad - bc & 0 \\ 0 & ad - bc \end{pmatrix} = \det A \cdot \mathbb{1}_2$

$\Longleftarrow$: Eind. Lösbarkeit der Gleichung $Ax = R$ durch $x = A^{-1}R$

□

Benötigt man — anders als wir! — zusätzlich *nur* noch *Determinaten für* (3×3)*-Matrizen*, so kann man diese zu

$$A := \begin{pmatrix} a_1^1 & a_2^1 & a_3^1 \\ a_1^2 & a_2^2 & a_3^2 \\ a_1^3 & a_2^3 & a_3^3 \end{pmatrix}$$

definieren durch *Entwicklung z. B. nach der ersten Zeile:*

$$\det A := a_1^1 \begin{vmatrix} a_2^2 & a_3^2 \\ a_2^3 & a_3^3 \end{vmatrix} - a_2^1 \begin{vmatrix} a_1^2 & a_3^2 \\ a_1^3 & a_3^3 \end{vmatrix} + a_3^1 \begin{vmatrix} a_1^2 & a_2^2 \\ a_1^3 & a_2^3 \end{vmatrix}$$

Diese Determinante läßt sich, wie man sofort nachrechnet, auch nach der folgenden *Regel von* SARRUS berechnen:
Man schreibt die erste und zweite Spalte noch einmal ‚hinter' die Matrix, erhält also

$$\begin{matrix} a_1^1 & a_2^1 & a_3^1 & a_1^1 & a_2^1 \\ a_1^2 & a_2^2 & a_3^2 & a_1^2 & a_2^2 \\ a_1^3 & a_2^3 & a_3^3 & a_1^3 & a_2^3 \end{matrix}$$

Nun bildet man die Summe der Produkte aller längs der *Hauptdiagonalen* (rot markiert) und ihrer Parallelen stehenden Elemente und subtrahiert die Summe der Produkte aller längs der ‚Nebendiagonalen' (grün markiert) und ihrer Parallelen stehenden Elemente:

$$\det A = a_1^1 a_2^2 a_3^3 + a_2^1 a_3^2 a_1^3 + a_3^1 a_1^2 a_2^3 - \left(a_3^1 a_2^2 a_1^3 + a_1^1 a_3^2 a_2^3 + a_2^1 a_1^2 a_3^3\right)$$

Den (einfachen) Nachweis, daß auch in diesem Fall die Eigenschaften (1), (2) und (3) — und dann auch (4) bis (9) — entsprechend gelten, verschieben wir auf den allgemeinen Fall.

6.2. **Permutationen.**

Definition 6.2.1. Für $n \in \mathbb{N}$ sei

$$\mathbb{N}_n := \{1, \ldots, n\}\,.$$

Eine *Permutation* der Zahlen $1, \ldots, n$ ist eine bijektive Abbildung

$$\sigma\colon\ \mathbb{N}_n \longrightarrow \mathbb{N}_n\,.$$

Schematisch können wir eine Permutation durch

$$\begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ \sigma(1) & \sigma(2) & \sigma(3) & \ldots & \sigma(n) \end{pmatrix}$$

oder kürzer durch

$$(\sigma(1)\ \sigma(2)\ \sigma(3)\ \ldots\ \sigma(n))$$

darstellen, im zweiten Fall auch durch Kommata getrennt.

Bemerkung 6.2.2.

a) Die Permutationen bilden mit der Hintereinanderausführung eine Gruppe, notiert als $\mathfrak{S}_n$ und als symmetrische Gruppe n-ten Grades angesprochen. Das neutrale Element ist $e := \mathrm{id}_{\mathbb{N}_n}$.

b) Es gibt genau $n! := 1 \cdot 2 \cdots n$ Permutationen von $1, \ldots, n$.

c) Genau für $n \geq 3$ ist $\mathfrak{S}_n$ nicht kommutativ.

Beweis: Die Schlußweise zu *a)* ist uns schon aus anderen ähnlichen Situationen vertraut. Wir führen sie nicht mehr aus.

b): Für das erste Bild hat man n Möglichkeiten, dann für das zweite noch $n-1$ usw. Insgesamt hat man so $n!$ Möglichkeiten. (Wenn man dies formal etwas sauberer haben will, kann man es natürlich durch Induktion beweisen.)

c): Übung (11.1)

□

Um nicht über den trivialen Fall $n = 1$ zu stolpern, setzen wir für den Rest des Abschnitts $n \geq 2$ voraus.

Definition 6.2.3.

Für $1 \leq r \leq n$ und paarweise verschiedene $i_1, \ldots, i_r \in \mathbb{N}_n$ bezeichnen wir die durch $f(i_\varrho) = i_{\varrho+1}$ $(\varrho = 1, \ldots, r-1)$, $f(i_r) = i_1$ und $f(j) = j$ für $j \in \mathbb{N}_n \setminus \{i_1, \ldots, i_r\}$ definierte Abbildung $f \in \mathfrak{S}_n$ kurz mit $[i_1, \ldots, i_r]$. Wir sprechen von einem *Zyklus,* genauer *r-Zyklus*.
Ein 2-Zyklus heißt *Transposition.* (Eine Transposition vertauscht also genau zwei Elemente miteinander.)

Für jede Transposition τ gilt offenbar: $\tau^2 = e$ und $\tau = \tau^{-1}$.

Definition 6.2.4.
Es sei σ eine Permutation der Zahlen $1, 2, \ldots, n$ und dazu $s = s(\sigma)$ die Anzahl der *„Inversionen“* oder *„Fehlstellungen“* , d. h. die Anzahl der Paare (i, j) mit $1 \le i < j \le n$ und $\sigma(i) > \sigma(j)$. Dann definieren wir das *Signum* der Permutation σ durch

$$\operatorname{sign}\sigma \;:=\; (-1)^{s(\sigma)}\,.$$

Die Permutation σ heißt *gerade,* wenn $\operatorname{sign}(\sigma) = 1$ gilt, sonst *ungerade*.

Bemerkung 6.2.5.
Für $\sigma \in \mathfrak{S}_n$ gilt:

$$\operatorname{sign}\sigma \;=\; \prod_{1\le i<j\le n} \frac{\sigma(j)-\sigma(i)}{j-i}$$

Beweis: Ist m die Anzahl der Inversionen von σ, dann gilt:

$$\begin{aligned}\prod_{1\le i<j\le n} (\sigma(j)-\sigma(i)) \;&=\; (-1)^m \prod_{(i,j)\ \text{Inversion}} |\sigma(j)-\sigma(i)| \cdot \prod_{\text{Rest}} (\sigma(j)-\sigma(i)) \\ &=\; (-1)^m \prod_{1\le i<j\le n} |\sigma(j)-\sigma(i)| \;=\; (-1)^m \prod_{1\le i<j\le n} (j-i)\end{aligned}$$

□

Bemerkung 6.2.6.
Für alle $\varrho, \sigma \in \mathfrak{S}_n$ gilt

$$\operatorname{sign}(\varrho\circ\sigma) \;=\; \operatorname{sign}\varrho \operatorname{sign}\sigma \text{ und so } \operatorname{sign}\left(\varrho^{-1}\right) \;=\; \operatorname{sign}\varrho\,.$$

Beweis:

$$\ell.S. \;= \prod_{i<j} \frac{\varrho(\sigma(j))-\varrho(\sigma(i))}{j-i} = \underbrace{\prod_{i<j} \frac{\varrho(\sigma(j))-\varrho(\sigma(i))}{\sigma(j)-\sigma(i)}}_{=:\,①} \cdot \underbrace{\prod_{i<j} \frac{\sigma(j)-\sigma(i)}{j-i}}_{=\operatorname{sign}\sigma}$$

$$\begin{aligned}① \;&=\; \prod_{\substack{i<j\\ \sigma(i)<\sigma(j)}} \frac{\varrho(\sigma(j))-\varrho(\sigma(i))}{\sigma(j)-\sigma(i)} \cdot \prod_{\substack{i<j\\ \sigma(i)>\sigma(j)}} \frac{\varrho(\sigma(j))-\varrho(\sigma(i))}{\sigma(j)-\sigma(i)} \\ &=\; \prod_{\substack{i<j\\ \sigma(i)<\sigma(j)}} \frac{\varrho(\sigma(j))-\varrho(\sigma(i))}{\sigma(j)-\sigma(i)} \cdot \prod_{\substack{i>j\\ \sigma(i)<\sigma(j)}} \frac{\varrho(\sigma(j))-\varrho(\sigma(i))}{\sigma(j)-\sigma(i)} \\ &=\; \prod_{\sigma(i)<\sigma(j)} \frac{\varrho(\sigma(j))-\varrho(\sigma(i))}{\sigma(j)-\sigma(i)} \overset{\checkmark}{=} \operatorname{sign}\varrho\end{aligned}$$

□

Folgerung 6.2.7.
Für jede Transposition τ gilt $\operatorname{sign}\tau = -1$.

Beweis: ∘ ∘ ∘ □

Satz 6.2.8.

Jedes $\sigma \in \mathfrak{S}_n$ läßt sich als endliches Produkt von Transpositionen *schreiben.*

Beweis. Ist $\sigma = e$, dann gilt $\sigma = \tau \circ \tau^{-1} = \tau \circ \tau$ mit irgendeiner Transposition τ. Also Œ $\sigma \neq e$: Es existiert

$$i_1 := \min\{j \in \mathbb{N}_n \mid \sigma(j) \neq j\}\,.$$

Dann ist $j_1 := \sigma(i_1) > i_1$. Mit $\tau_1 := [i_1, j_1]$ hat man dann für $\sigma_1 := \tau_1 \circ \sigma$, daß $\sigma_1(j) = j$ für $1 \leq j \leq i_1$ gilt. Ist $\sigma_1 = e$, so hat man $\sigma = \tau_1^{-1} = \tau_1$ und ist fertig. Sonst existiert

$$i_2 := \min\{j \in \mathbb{N}_n \mid \sigma_1(j) \neq j\} > i_1\,.$$

Man bildet dazu entsprechend τ_2 und σ_2 wie oben. Auf diese Weise erhält man nach $k \leq n$ Schritten: Es existieren Transpositionen $\tau_1, \ldots, \tau_k$ mit $\sigma_k = \tau_k \circ \cdots \circ \tau_1 \circ \sigma = e$, also $\sigma = \tau_1^{-1} \circ \cdots \circ \tau_k^{-1} = \tau_1 \circ \cdots \circ \tau_k$ □

Satz 6.2.9.

Es gibt genau $n!/2$ gerade und $n!/2$ ungerade Permutationen der Zahlen $1, 2, \ldots, n$.

Beweis. Es sei τ eine beliebige Transposition. Da die Permutationen eine Gruppe bilden, ist $\sigma \longmapsto \tau\sigma$ bijektiv. Die geraden Permutationen werden dabei auf die ungeraden und umgekehrt abgebildet. Folglich gibt es von jeder Sorte gleich viele. □

Bemerkung 6.2.10.

Eine Untergruppe der symmetrischen Gruppe $\mathfrak{S}_n$ ist die *alternierende Gruppe* $\mathfrak{A}_n$, bestehend aus den Elementen σ von $\mathfrak{S}_n$ mit $\operatorname{sign}\sigma = 1$.

6.3. Multilineare Abbildungen.

Es seien $\mathbb{K}$ ein Körper mit $1+1 \neq 0$ (Die ‚Charakteristik' von $\mathbb{K}$ ist verschieden von 2, kurz notiert: $\operatorname{char}_{\mathbb{K}} \neq 2$)[7], $n \in \mathbb{N}$ und $\mathfrak{V}_1, \ldots, \mathfrak{V}_n, \mathfrak{U}, \mathfrak{V}$ $\mathbb{K}$-Vektorräume.

[7] Dies vermeidet ein paar Feinsinnigkeiten, die wir getrost den Vollblutmathematikern überlassen können.

Definition 6.3.1.

Eine Abbildung $f\colon \mathfrak{V}_1 \times \cdots \times \mathfrak{V}_n \longrightarrow \mathfrak{V}$ heißt genau dann *multilinear*, genauer *n-linear*, wenn für alle $\nu \in \mathbb{N}_n$ und für alle $a_i \in \mathfrak{V}_i$ $(i \neq \nu)$ die Abbildung

$$\mathfrak{V}_\nu \ni x \longmapsto f(a_1, \ldots, a_{\nu-1}, x, a_{\nu+1}, \ldots, a_n) \in \mathfrak{V}\,^{8}$$

linear ist. Wir notieren diese Abbildung auch als

$$f(a_1, \ldots, a_{\nu-1}, \,\cdot\,, a_{\nu+1}, \ldots, a_n)\,.$$

$$\mathcal{L}(\mathfrak{V}_1, \ldots, \mathfrak{V}_n; \mathfrak{V}) := \{h \mid h\colon \mathfrak{V}_1 \times \cdots \times \mathfrak{V}_n \longrightarrow \mathfrak{V} \;\; n\text{-linear}\}$$

Ist $\mathfrak{V}_1 = \cdots = \mathfrak{V}_n = \mathfrak{U}$, dann schreiben wir

$$\mathcal{L}_n(\mathfrak{U}, \mathfrak{V}) := \mathcal{L}(\mathfrak{V}_1, \ldots, \mathfrak{V}_n; \mathfrak{V})\,.$$

Die Multilinearität für $f\colon \mathfrak{V}_1 \times \cdots \times \mathfrak{V}_n \longrightarrow \mathfrak{V}$ bedeutet also: Hält man $n-1$ Variable fest, dann erhält man eine lineare Abbildung bezüglich der verbleibenden Variablen.

Bemerkung 6.3.2.

$\mathcal{L}(\mathfrak{V}_1, \ldots, \mathfrak{V}_n; \mathfrak{V})$ ist ein Unterraum des $\mathbb{K}$-Vektorraums aller Abbildungen von $\mathfrak{V}_1 \times \cdots \times \mathfrak{V}_n$ in $\mathfrak{V}$.

Beweis: ✓ □

Beispiele

(B1) Euklidisches Skalarprodukt $(n = 2)$

(B2) Vektorprodukt im $\mathbb{R}^3$ $(n = 2)$

(B3) Spatprodukt $(n = 3)$ (kommt in Übung (11.2))

(B4) Für $p, q \in \mathbb{N}_0$ heißt eine $(p+q)$-lineare Abbildung

$$\begin{array}{ccccc} f\colon & \mathfrak{U}^{*p} \times \mathfrak{U}^q & \longrightarrow & \mathbb{K} \\ & \cup & & \cup \\ & (u^1, \ldots, u^p, x_1, \ldots, x_q) & \longmapsto & f(u^1, \ldots, u^p, x_1, \ldots, x_q) \end{array}$$

p-fach *kovarianter* und q-fach *kontravarianter Tensor* oder auch Tensor der *Stufe* $\binom{p}{q}$. $\left(\text{Wieder ‚richtig‘ zu lesen in den Fällen } p = 0 \text{ oder } q = 0. \text{ Tensor der } \textit{Stufe} \binom{0}{0} : \text{Skalare}\right)$

◁

Manche Physiker definierten zu meiner Studienzeit[9] — wenn überhaupt — Tensoren (für reelle Vektorräume niedriger Dimensionen) komplizierter. In der Physik

[8] Wieder ‚richtig‘ lesen in den Fällen $\nu = 1$ und $\nu = n$

[9] Heute ist das ja — zum Glück — alles anders … ☺

interessiert man sich meist für *Tensorfelder* (das sind auf Teilmengen von $\mathfrak{U}$ definierte tensorwertige Abbildungen); oft wird dabei nicht streng genug unterschieden zwischen Tensoren und Tensorfeldern, was nicht gerade zum Verständnis beiträgt.

Definition 6.3.3.
Eine Abbildung $f \in \mathcal{L}_n(\mathfrak{U}, \mathfrak{V})$ heißt genau dann *„alternierend“* , wenn $f(x_1, \ldots, x_n) = 0$ ist, falls irgendzwei der Vektoren $x_1, \ldots, x_n$ gleich sind. Für $\mathfrak{V} = \mathbb{K}$ heißt ein solches f *„alternierende (n)-Form“*.

$$\mathfrak{A}_n(\mathfrak{U}, \mathfrak{V}) := \{f \in \mathcal{L}_n(\mathfrak{U}, \mathfrak{V}) \mid f \text{ alternierend}\}, \quad \mathfrak{A}_n(\mathfrak{U}) := \mathfrak{A}_n(\mathfrak{U}, \mathbb{K})$$

Bemerkung 6.3.4.
$\mathfrak{A}_n(\mathfrak{U}, \mathfrak{V})$ *ist ein Unterraum von* $\mathcal{L}_n(\mathfrak{U}, \mathfrak{V})$.

Beweis: ✓ □

Definition 6.3.5.
Ein $f \in \mathcal{L}_n(\mathfrak{U}, \mathfrak{V})$ heißt genau dann *„antisymmetrisch“*, wenn für alle $\sigma \in \mathfrak{S}_n$ und $(x_1, \ldots, x_n) \in \mathfrak{U}^n$ gilt:

$$f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) = \operatorname{sign}\sigma \cdot f(x_1, \ldots, x_n)$$

Bemerkung 6.3.6.
Für $f \in \mathcal{L}_n(\mathfrak{U}, \mathfrak{V})$ *gilt:*

$$f \in \mathfrak{A}_n(\mathfrak{U}, \mathfrak{V}) \iff f \text{ ist antisymmetrisch}$$

Beweisskizze:

$\Longleftarrow$: ∘ ∘ ∘

$\Longrightarrow$: Dies zeigt man zuerst für den Spezialfall einer Transposition (man vergleiche dazu den Beweis von Eigenschaft (6) in Abschnitt 6.1), dann zieht man Satz 6.2.8 heran und schließt per Induktion. □

Determinantenfunktionen

Es seien wieder $\mathbb{K}$ ein Körper mit $1+1 \neq 0$, $\mathfrak{V}$ ein $\mathbb{K}$-Vektorraum mit $\dim \mathfrak{V} =: n \in \mathbb{N}$ und $(a_1, \ldots, a_n)$ eine (geordnete) Basis von $\mathfrak{V}$.

Definition 6.3.7.

Ein $\Delta \in \mathfrak{A}_n$, also eine *multilineare und alternierende Abbildung*

$$\Delta\colon \mathfrak{V}^n \longrightarrow \mathbb{K},$$

heißt *„Determinantenfunktion“*.

Es sei nun für den Rest dieses Abschnitts Δ eine — fest gewählte — Determinantenfunktion.

Satz 6.3.8. *(Eigenschaften von Determinantenfunktionen)*

(1) $\mathfrak{V} \ni v_1, \ldots, v_n$ *linear abhängig* $\Longrightarrow \Delta(v_1, \ldots, v_n) = 0$

(2) *Für* $\nu, \mu \in \mathbb{N}_n$ *mit* $\nu \neq \mu$, $\lambda \in \mathbb{K}$ *und* $v_1, \ldots, v_n \in \mathfrak{V}$ *gilt:*

$$\Delta(v_1, \ldots, v_{\nu-1}, v_\nu + \lambda v_\mu, v_{\nu+1}, \ldots, v_n) = \Delta(v_1, \ldots, v_n)$$

(3) *Mit* $(\alpha_\nu^\mu)_{1 \leq \nu, \mu \leq n} \in \mathbb{K}^{n \times n}$ *und* $v_\nu := \sum_{\mu=1}^{n} \alpha_\nu^\mu a_\mu$ *für* $\nu = 1, \ldots, n$:

$$\Delta(v_1, \ldots, v_n) = \Delta(a_1, \ldots, a_n) \cdot \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sign} \sigma \, \alpha_1^{\sigma(1)} \cdots \alpha_n^{\sigma(n)}$$

(4) *Ist* $0 \neq \Delta^*$ *eine weitere Determinantenfunktion, dann existiert eindeutig ein* $\varkappa \in \mathbb{K}$ *mit*

$$\Delta = \varkappa \cdot \Delta^* .$$

Beweis:

(1): Œ sei $v_1 = \sum_{\mu=2}^{n} \lambda_\mu v_\mu$ mit $\lambda_\mu \in \mathbb{K}$. Dann ist

$$\Delta(v_1, \ldots, v_n) = \sum_{\mu=2}^{n} \lambda_\mu \Delta(v_\mu, v_2, \ldots, v_n) = 0$$

(2): Δ ist linear in der ν-ten Variablen und alternierend.

(3): $$\Delta(v_1, \ldots, v_n) = \sum_{\mu_1=1}^{n} \cdots \sum_{\mu_n=1}^{n} \alpha_1^{\mu_1} \cdots \alpha_n^{\mu_n} \Delta(a_{\mu_1}, \ldots, a_{\mu_n})$$

Da Δ alternierend ist, genügt es, über alle *Permutationen* zu summieren. Das ergibt:

$$\sum_{\sigma \in \mathfrak{S}_n} \alpha_1^{\sigma(1)} \cdots \alpha_n^{\sigma(n)} \Delta(a_{\sigma(1)}, \ldots, a_{\sigma(n)}) ,$$

also mit Bemerkung 6.3.6 die Behauptung.

(4): Nach (3) ist $\Delta^*(a_1, \ldots, a_n) \neq 0$; $\varkappa := \dfrac{\Delta(a_1, \ldots, a_n)}{\Delta^*(a_1, \ldots, a_n)}$ leistet — wieder nach (3) — das Gewünschte.

□

Satz 6.3.9. *(Existenz einer normierten Determinantenfunktion)*

Es existiert eindeutig eine Determinantenfunktion Δ_0 *mit* $\Delta_0(a_1, \ldots, a_n) = 1$, *nämlich die durch*

(*) $$\Delta_0(v_1, \ldots, v_n) := \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sign} \sigma \, \alpha_1^{\sigma(1)} \cdots \alpha_n^{\sigma(n)}$$

für $v_\nu := \sum_{\mu=1}^{n} \alpha_\nu^\mu a_\mu$ $(\nu = 1, \ldots, n)$ *definierte.*

Beweis: Die *Eindeutigkeit* und $(*)$ folgen mit $\Delta_0(a_1,\dots,a_n)=1$ aus (3) von Satz 6.3.8. Zu zeigen bleibt, daß durch $(*)$ eine Determinantenfunktion definiert wird (, die dann offenbar $\Delta_0(a_1,\dots,a_n)=1$ erfüllt):

Δ_0 ist *multilinear:* Für $\nu\in\mathbb{N}_n$, $w_\nu := \sum\limits_{\mu=1}^{n}\beta_\nu^\mu a_\mu$ $(\dots)$ und $\lambda\in\mathbb{K}$:

$\Delta_0(v_1,\dots,v_{\nu-1},\lambda v_\nu+w_\nu,v_{\nu+1},\dots,v_n)$

$= \sum\limits_{\sigma\in\mathfrak{S}_n}\operatorname{sign}\sigma\,\alpha_1^{\sigma(1)}\cdots\alpha_{\nu-1}^{\sigma(\nu-1)}\left(\lambda\alpha_\nu^{\sigma(\nu)}+\beta_\nu^{\sigma(\nu)}\right)\alpha_{\nu+1}^{\sigma(\nu+1)}\cdots\alpha_n^{\sigma(n)}$

$= \cdots = \lambda\Delta_0(v_1,\dots,v_n)+\Delta_0(v_1,\dots,v_{\nu-1},w_\nu,v_{\nu+1},\dots,v_n)$

Δ_0 ist *alternierend:* Zu zeigen ist: Für $v_1,\dots,v_n\in\mathfrak{V}$ mit $v_k=v_\ell$ für ein Paar $(k,\ell)\in\mathbb{N}_n^2$ mit $k\neq\ell$ gilt $\Delta_0(v_1,\dots,v_n)=0$:

Aus $v_k=v_\ell$ folgt $\alpha_k^\mu=\alpha_\ell^\mu$ für $\mu=1,\dots,n$ $(\circledast)$. Mit der Transposition $\tau:=[k,\ell]$ gilt $\mathfrak{S}_n=\mathfrak{A}_n\uplus\mathfrak{A}_n\tau$, also

$$\Delta_0(v_1,\dots,v_n) = \sum_{\sigma\in\mathfrak{A}_n}\prod_{\nu=1}^{n}\alpha_\nu^{\sigma(\nu)} - \sum_{\sigma\in\mathfrak{A}_n}\prod_{\nu=1}^{n}\alpha_\nu^{\sigma(\tau(\nu))}\,.$$

Nun ist für jedes $\sigma\in\mathfrak{A}_n$

$$\prod_{\nu=1}^{n}\alpha_\nu^{\sigma(\tau(\nu))} = \left(\prod_{\substack{\nu=1\\ \nu\notin\{k,\ell\}}}^{n}\alpha_\nu^{\sigma(\nu)}\right)\left(\alpha_k^{\sigma(\tau(k))}\alpha_\ell^{\sigma(\tau(\ell))}\right) = \Big(\cdots\Big)\left(\alpha_k^{\sigma(\ell)}\alpha_\ell^{\sigma(k)}\right)$$

$$\underset{(\circledast)}{=} \Big(\cdots\Big)\left(\alpha_\ell^{\sigma(\ell)}\alpha_k^{\sigma(k)}\right) = \prod_{\nu=1}^{n}\alpha_\nu^{\sigma(\nu)}\,,\ \text{also}\ \Delta_0(v_1,\dots,v_n)=0\,. \qquad \square$$

Bemerkung 6.3.10.

Ist Δ eine nicht-triviale, d. h. $\Delta\neq 0$, Determinantenfunktion, dann gilt für $v_1,\dots,v_n\in\mathfrak{V}$:

$$\Delta(v_1,\dots,v_n)\neq 0 \iff v_1,\dots,v_n \text{ ist eine Basis}$$

Beweis:

$\Longrightarrow$: $v_1,\dots,v_n$ sind nach (1) von Satz 6.3.8 linear unabhängig, also eine Basis, da $\dim\mathfrak{V}=n$ ist.

$\Longleftarrow$: Es existieren $u_1,\dots,u_n\in\mathfrak{V}$ mit $\Delta(u_1,\dots,u_n)\neq 0$. Ist Δ_0 die — gemäß Satz 6.3.9 — zu $v_1,\dots,v_n$ gebildete Determinantenfunktion, dann gilt nach (4) aus Satz 6.3.8 $\Delta=\varkappa\Delta_0$ mit einem geeigneten $\varkappa\in\mathbb{K}$. $\Delta_0(v_1,\dots,v_n)=1$ zeigt $\Delta(v_1,\dots,v_n)=\varkappa$, also $0\neq\Delta(u_1,\dots,u_n)=\Delta(v_1,\dots,v_n)\,\Delta_0(u_1,\dots,u_n)$, somit $\Delta(v_1,\dots,v_n)\neq 0$.

$\square$

6.4. **Endomorphismen und quadratische Matrizen.**

Es seien $\mathbb{K}$ ein Körper mit $1+1 \neq 0$, $n \in \mathbb{N}$, $\mathfrak{V}$ ein $\mathbb{K}$-Vektorraum mit $\dim \mathfrak{V} = n$ und Δ eine *nicht-triviale Determinantenfunktion.*

Bemerkung 6.4.1.

Es sei $f \in \operatorname{End}(\mathfrak{V}) := \mathcal{L}(\mathfrak{V}, \mathfrak{V})$.

a)
$$\begin{array}{rccc} \Delta_f\colon & \mathfrak{V}^n & \longrightarrow & \mathbb{K} \\ & \cup & & \cup \\ & (v_1, \ldots, v_n) & \longmapsto & \Delta(f v_1, \ldots, f v_n) \end{array}$$
ist eine Determinantenfunktion. Nach Satz 6.3.8 existiert daher eindeutig $\varkappa(\Delta, f) = \varkappa \in \mathbb{K}$ *mit*
$$\Delta_f = \varkappa \cdot \Delta .$$

b) Ist δ *eine weitere nicht-triviale Determinantenfunktion, dann ist* $\varkappa(\delta, f) = \varkappa(\Delta, f) =: \det f$.

Beweis:

a): Δ_f ist *multilinear,* da Δ multilinear und f linear ist. Δ_f ist *alternierend,* da Δ alternierend ist. Da Δ nicht-trivial ist, ist der Faktor $\varkappa(\Delta, f)$ *eindeutig.*

b): Nach Satz 6.3.8 existiert ein $\alpha \in \mathbb{K}$ mit $\Delta = \alpha\delta$. Offenbar gilt dann auch $\Delta_f = \alpha\delta_f$. Damit hat man: $\Delta_f = \alpha\delta_f = \alpha\varkappa(\delta, f)\delta = \varkappa(\delta, f)\Delta$.

□

Definition 6.4.2. $\det f$ heißt *„Determinante von f“* .

Satz 6.4.3.

a) $f, g \in \operatorname{End}(\mathfrak{V}) \implies \det(f \circ g) = \det f \det g$

b) $\det(\mathrm{id}_{\mathfrak{V}}) = 1$

c) Für $f \in \operatorname{End}(\mathfrak{V})$: $f \in \mathrm{GL}(\mathfrak{V}) \iff \det f \neq 0$

d) Für $f \in \mathrm{GL}(\mathfrak{V})$: $\det(f^{-1}) = (\det f)^{-1}$

Beweis: Mit einer Basis $v_1, \ldots, v_n$ von $\mathfrak{V}$ gilt:

a): $\det(f \circ g) \underbrace{\Delta(v_1, \ldots, v_n)}_{\neq 0} = \Delta\big(f(g(v_1)), \ldots, f(g(v_n))\big)$

$= \det f \, \Delta\big(g(v_1), \ldots, g(v_n)\big) = \det f \det g \, \Delta(v_1, \ldots, v_n)$

b): ✓

d) und „$\Longrightarrow$“ von *c)*: $1 \underset{b)}{=} \det\left(f \circ f^{-1}\right) \underset{a)}{=} \det f \det(f^{-1})$

„$\Longleftarrow$“ von *c)*: $\Delta(f v_1, \ldots, f v_n) = \underbrace{\det f}_{\neq 0} \underbrace{\Delta(v_1, \ldots, v_n)}_{\neq 0}$; also ist

$(f v_1, \ldots, f v_n)$ eine Basis; daher ist f ein Isomorphismus. □

Definition 6.4.4.
Nach Satz 6.3.9 existiert eindeutig eine Determinantenfunktion

$$\Delta\colon (\mathbb{K}^n)^n \longrightarrow \mathbb{K}$$

mit $\Delta(e_1,\dots,e_n)=1$, nämlich die durch

$$\Delta(v_1,\dots,v_n) := \sum_{\sigma\in\mathfrak{S}_n} \operatorname{sign}\sigma\, \alpha_1^{\sigma(1)}\cdots\alpha_n^{\sigma(n)}$$

für $v_\nu := \begin{pmatrix}\alpha_\nu^1\\ \vdots\\ \alpha_\nu^n\end{pmatrix} \in \mathbb{K}^n \ (\nu=1,\dots,n)$ definierte.
Für eine Matrix $A=(\alpha_\nu^\mu)\in\mathbb{K}^{n\times n}$ mit den Spalten $A_1,\dots,A_n$, also $A=(A_1,\dots,A_n)$, sei die *„Determinante von A“* damit definiert durch:

$$\det A := \det(A_1,\dots,A_n) := \begin{vmatrix}\alpha_1^1 & \cdots & \alpha_n^1\\ \vdots & & \vdots\\ \alpha_1^n & \cdots & \alpha_n^n\end{vmatrix} := \Delta(A_1,\dots,A_n)$$

Man hat also:

$$\det A = \sum_{\sigma\in\mathfrak{S}_n} \operatorname{sign}\sigma\, \alpha_1^{\sigma(1)}\cdots\alpha_n^{\sigma(n)}$$

Bemerkung 6.4.5.
Für eine Matrix $A\in\mathbb{K}^{n\times n}$ ist $\det A = \det f_A$.

Beweis:
$\Delta(A_1,\dots,A_n) = \Delta(f_Ae_1,\dots,f_Ae_n) = \det(f_A)\Delta(e_1,\dots,e_n) = \det(f_A)$ □

Bemerkung 6.4.6.
Ist $\mathcal{B}:=(b_1,\dots,b_n)$ eine (geordnete) Basis von $\mathfrak{V}$ und $f\in\operatorname{End}(\mathfrak{V})$, dann gilt mit $A:=\mathcal{M}_\mathcal{B}(f)$:

$$\det(f) = \det(A)$$

Beweis: Ist Δ die zu $\mathcal{B}$ gemäß Satz 6.3.9 gebildete Determinantenfunktion, dann ist:

$$\det f = \det f\Delta(b_1,\dots,b_n) = \Delta(fb_1,\dots,fb_n)$$
$$\underset{(6.3.8)}{=} \sum_{\sigma\in\mathfrak{S}_n} \operatorname{sign}\sigma\, \alpha_1^{\sigma(1)}\cdots\alpha_n^{\sigma(n)} = \det(A)$$ □

Mit einer dieser beiden Bemerkungen können wir nun sofort die Eigenschaften von Determinanten von Endomorphismen ‚übersetzen‘ in entsprechende *Eigenschaften von Determinanten von Matrizen:*

Satz 6.4.7.

Für $A, B \in \mathbb{K}^{n\times n}$ *gelten:*

a) $\det(AB) = \det(A)\det(B)$

b) $\det(\mathbb{1}_n) = 1$

c) $A \in \mathrm{GL}(n,\mathbb{K}) \iff \det(A) \neq 0$

d) Für $A \in \mathrm{GL}(n,\mathbb{K})$: $\det(A^{-1}) = (\det(A))^{-1}$

Die Determinante von A ist also als Funktion der Spalten *n-linear* und *alternierend*. Wir stellen noch einmal die wichtigsten Eigenschaften zusammen:

(0) det *ist linear in jeder Spalte.* *(n-linear)*

(1) *Hat* A *zwei gleiche Spalten, so ist* $\det(A) = 0$. *(alternierend)*

(2) $\det(\lambda A) = \lambda^n \det(A) \quad (\lambda \in \mathbb{K})$

(3) *Entsteht eine Matrix* B *durch eine Spaltenvertauschung aus* A*, dann ist* $\det(B) = -\det(A)$.

(4) *Ist* $\lambda \in \mathbb{K}$ *und entsteht* B *aus* A *durch Addition des* λ*-fachen der* j*-ten Spalte zur* i*-ten Spalte für* $i \neq j$*, dann ist* $\det(B) = \det(A)$.

Die *Eigenschaften (0) bis (4) gelten entsprechend für die Zeilen*; denn man hat für $A \in \mathbb{K}^{n\times n}$:

Bemerkung 6.4.8.

$$\det(A) = \det(A^T)$$

Beweis: Wir bezeichnen die Elemente von A^T mit β^μ_ν, also mit den obigen Bezeichnungen für A: $\beta^\mu_\nu = \alpha^\nu_\mu$ $(\dots)$.

$$\det(A^T) = \sum_{\sigma\in\mathfrak{S}_n} \operatorname{sign}\sigma\, \beta_1^{\sigma(1)} \cdots \beta_n^{\sigma(n)} = \sum_{\sigma\in\mathfrak{S}_n} \operatorname{sign}\sigma\, \alpha^1_{\sigma(1)} \cdots \alpha^n_{\sigma(n)}$$

Mit σ ‚durchläuft' auch $\varrho := \sigma^{-1}$ die Gruppe $\mathfrak{S}_n$. Unter Beachtung von $\operatorname{sign}(\varrho^{-1}) = \operatorname{sign}\varrho$ ist die rechte Summe damit offenbar gleich $\sum_{\varrho\in\mathfrak{S}_n} \operatorname{sign}\varrho\, \alpha_1^{\varrho(1)} \cdots \alpha_n^{\varrho(n)}$, d. h. gleich $\det A$. □

6.5. Berechnungsverfahren für Determinanten von Matrizen.

Für $n = 1, 2, 3$ lassen sich Determinanten von Matrizen aus $\mathbb{K}^{n\times n}$ noch leicht über die definierende Gleichung berechnen. Für $A = (\alpha^\mu_\nu)$ hat man:

$n = 1$: $\det A = \alpha^1_1$

$n = 2$: $\det A = \alpha_1^1\alpha_2^2 - \alpha_2^1\alpha_1^2$; denn $\mathfrak{S}_2 = \{e, [1,2]\}$ mit $\operatorname{sign} e = 1$ und $\operatorname{sign}[1,2] = -1$.

$n = 3$: $\det A = \alpha_1^1\alpha_2^2\alpha_3^3 + \alpha_1^2\alpha_2^3\alpha_3^1 + \alpha_1^3\alpha_2^1\alpha_3^2 - \alpha_1^3\alpha_2^2\alpha_3^1 - \alpha_1^1\alpha_2^3\alpha_3^2 - \alpha_1^2\alpha_2^1\alpha_3^3$; denn hier hat man $\mathfrak{S}_3 = \{\underset{+}{e}, \underset{-}{[1,2]}, \underset{-}{[1,3]}, \underset{-}{[2,3]}, \underset{+}{[1,2,3]}, \underset{+}{[1,3,2]}\}$

Schon für $n = 4$ enthält die Summe 24 Summanden. Gesucht sind daher *bessere Hilfsmittel zur Berechnung ‚großer' Determinanten*: (vgl. auch die Anmerkung zur CRAMER-Regel ...)

Es sei im Folgenden $A = (\alpha_\nu^\mu) \in \mathbb{K}^{n\times n}$ mit einem $n \in \mathbb{N}_2$.

Satz 6.5.1.

Mit einem $n > p \in \mathbb{N}$ gelte $\alpha_\nu^\mu = 0$ für $\mu > p$ und $\nu \le p$. Mit $q := n - p$ hat A also die Blockgestalt

$$A = \begin{pmatrix} A_1^1 & A_2^1 \\ \mathbb{O} & A_2^2 \end{pmatrix}$$

mit $A_1^1 \in \mathbb{K}^{p\times p}$, $A_2^1 \in \mathbb{K}^{p\times q}$, $A_2^2 \in \mathbb{K}^{q\times q}$ und der Nullmatrix $\mathbb{O} \in \mathbb{K}^{q\times p}$. Dann gilt:

$$\det A = \det A_1^1 \cdot \det A_2^2$$

Beweis: Im Folgenden sei A_2^1 fest. Wir betrachten mit $X \in \mathbb{K}^{p\times p}$ und $Y \in \mathbb{K}^{q\times q}$ die durch

$$\delta(X;Y) := \begin{vmatrix} X & A_2^1 \\ \mathbb{O} & Y \end{vmatrix}$$

gegebene Abbildung. Bei festem Y ist δ p-linear und alternierend in den Spalten von X. Daher existiert ein $\mu = \mu(Y) \in \mathbb{K}$ mit $\delta(X;Y) = \mu(Y)\det(X)$. Für $X = \mathbb{1}_p$ also: $\delta(\mathbb{1}_p;Y) = \mu(Y)$. Dies zeigt, daß μ p-linear und alternierend in den Zeilen von Y ist. Somit existiert ein $\lambda \in \mathbb{K}$ mit $\mu(Y) = \lambda\det(Y)$. Speziell für $Y = \mathbb{1}_q$ ergibt sich: $\mu(\mathbb{1}_q) = \lambda$. Zusammen haben wir:

$$\delta(X;Y) = \lambda\det(X)\det(Y) = \delta(\mathbb{1}_p;\mathbb{1}_q)\det(X)\det(Y)$$

Zu zeigen bleibt daher $\delta(\mathbb{1}_p;\mathbb{1}_q) = 1$. Die Matrix $\begin{pmatrix} \mathbb{1}_p & A_2^1 \\ \mathbb{O} & \mathbb{1}_q \end{pmatrix}$ ergibt sich offenbar aus $\mathbb{1}_n$ durch Addition geeigneter Linearkombinationen der ersten p Spalten zu den letzten q Spalten. Nach (4) von Seite 95 ist somit:

$$\delta(\mathbb{1}_p;\mathbb{1}_q) = \begin{vmatrix} \mathbb{1}_p & A_2^1 \\ \mathbb{O} & \mathbb{1}_q \end{vmatrix} = \det(\mathbb{1}_n) = 1\,.$$

□

Induktiv erhält man dann das Resultat für Matrizen, die entsprechend aus endlich vielen Diagonalblöcken aufgebaut sind, speziell: *Für eine*

obere Dreiecksmatrix ist die Determinante gleich dem Produkt der Diagonalelemente. Nach Bemerkung 6.4.8 gilt dies dann entsprechend auch für *untere* Dreiecksmatrizen.

Wir können die Berechnung von Determinanten von $(n \times n)$-Matrizen auf die von $((n-1) \times (n-1))$-Matrizen zurückführen:

Satz 6.5.2 (Laplacescher Entwicklungssatz (Spezialfall))**.**

Zu $A \in \mathbb{K}^{n\times n}$ und $(\nu, \mu) \in \mathbb{N}_n^2$ betrachten wir die $((n-1) \times (n-1))$-Matrix A_ν^μ, die entsteht, wenn wir die μ-te Zeile und die ν-te Spalte streichen:

$$\left(\begin{array}{ccc|ccc} \alpha_1^1 & \dots & \alpha_{\nu-1}^1 & \alpha_{\nu+1}^1 & \dots & \alpha_n^1 \\ \vdots & & \vdots & \vdots & & \vdots \\ \alpha_1^{\mu-1} & \dots & \alpha_{\nu-1}^{\mu-1} & \alpha_{\nu+1}^{\mu-1} & \dots & \alpha_n^{\mu-1} \\ \hline \alpha_1^{\mu+1} & \dots & \alpha_{\nu-1}^{\mu+1} & \alpha_{\nu+1}^{\mu+1} & \dots & \alpha_n^{\mu+1} \\ \vdots & & \vdots & \vdots & & \vdots \\ \alpha_1^n & \dots & \alpha_{\nu-1}^n & \alpha_{\nu+1}^n & \dots & \alpha_n^n \end{array}\right)$$

Dann gilt für festes $\nu \in \mathbb{N}_n$

$$\det A = \sum_{\mu=1}^{n} (-1)^{\mu+\nu} \alpha_\nu^\mu \det A_\nu^\mu \qquad \textit{(Entwicklung nach ν-ter Spalte)}$$

und für festes $\mu \in \mathbb{N}_n$

$$\det A = \sum_{\nu=1}^{n} (-1)^{\mu+\nu} \alpha_\nu^\mu \det A_\nu^\mu \qquad \textit{(Entwicklung nach μ-ter Zeile)}.$$

Die ‚Vorzeichenverteilung' merkt man sich am einfachsten als Schachbrettmuster. Zum Entwickeln nimmt man natürlich zweckmäßigerweise eine Zeile oder Spalte, die möglichst viele Nullen enthält.

Beweis: Wegen Bemerkung 6.4.8 genügt es, die *erste Gleichung* zu beweisen: Hier hat man

$$\begin{aligned} \det A &= \Delta(A_1, \dots, A_n) = \Delta\Big(A_1, \dots, A_{\nu-1}, \sum_{\mu=1}^{n} \alpha_\nu^\mu e_\mu, A_{\nu+1}, \dots, A_n\Big) \\ &= \sum_{\mu=1}^{n} \alpha_\nu^\mu \Delta(A_1, \dots, A_{\nu-1}, e_\mu, A_{\nu+1}, \dots, A_n) \\ &= \sum_{\mu=1}^{n} \alpha_\nu^\mu (-1)^{\nu-1} \underbrace{\Delta(e_\mu, A_1, \dots, A_{\nu-1}, A_{\nu+1}, \dots, A_n)}_{=:\circledast} \end{aligned}$$

$$\circledast = \begin{vmatrix} 0 & \alpha_1^1 & \dots & \alpha_{\nu-1}^1 & \alpha_{\nu+1}^1 & \dots & \alpha_n^1 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & \alpha_1^{\mu-1} & \dots & \alpha_{\nu-1}^{\mu-1} & \alpha_{\nu+1}^{\mu-1} & \dots & \alpha_n^{\mu-1} \\ 1 & \alpha_1^{\mu} & \dots & \alpha_{\nu-1}^{\mu} & \alpha_{\nu+1}^{\mu} & \dots & \alpha_n^{\mu} \\ 0 & \alpha_1^{\mu+1} & \dots & \alpha_{\nu-1}^{\mu+1} & \alpha_{\nu+1}^{\mu+1} & \dots & \alpha_n^{\mu+1} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & \alpha_1^{n} & \dots & \alpha_{\nu-1}^{n} & \alpha_{\nu+1}^{n} & \dots & \alpha_n^{n} \end{vmatrix}$$

$$= (-1)^{\mu-1} \begin{vmatrix} 1 & \alpha_1^{\mu} & \dots & \alpha_{\nu-1}^{\mu} & \alpha_{\nu+1}^{\mu} & \dots & \alpha_n^{\mu} \\ 0 & & & & & & \\ \vdots & & & A_\nu^\mu & & & \\ 0 & & & & & & \end{vmatrix} \underset{(6.5.1)}{=} (-1)^{\mu-1} \det\left(A_\nu^\mu\right),$$

zusammen also die Behauptung. □

Satz 6.5.3.

Mit $B := (\beta_\nu^\mu) \in \mathbb{K}^{n\times n}$, *definiert durch*

$$\beta_\nu^\mu := (-1)^{\nu+\mu} \det\left(A_\mu^\nu\right) \qquad \text{für } \nu, \mu \in \mathbb{N}_n,$$

gelten:

a) $BA = \det A \mathbb{1}_n = AB$

b) Ist A *invertierbar, dann hat man also:* $A^{-1} = (\det A)^{-1} \cdot B$

Beweis:

a) : $AB =: (\gamma_\nu^\mu)$, also $\gamma_\nu^\mu = \sum_{\lambda=1}^{n} \alpha_\lambda^\mu \beta_\nu^\lambda = \sum_{\lambda=1}^{n} (-1)^{\nu+\lambda} \alpha_\lambda^\mu \det\left(A_\lambda^\nu\right)$. Nach der zweiten Gleichung von Satz 6.5.1 ist diese Summe für $\nu = \mu$ gleich $\det(A)$ und für $\nu \neq \mu$ gleich 0, weil sie gerade die Entwicklung nach der μ-ten Zeile der Determinante einer Matrix gibt, deren μ-te und ν-te Zeile übereinstimmen.

$BA = \det(A)\mathbb{1}_n$ erhält man ‚ebenso'.

b) : folgt unmittelbar aus *a)*. □

Es seien nun $A \in \mathrm{GL}(n, \mathbb{K})$ und $b = \begin{pmatrix} b^1 \\ \vdots \\ b^n \end{pmatrix} \in \mathbb{K}^n$. Die Gleichung $Ax = b$ (lineares Gleichungssystem) hat dann die eindeutige Lösung $x = A^{-1}b = (\det(A))^{-1} Bb$, also für $\nu \in \mathbb{N}_n$

$$x^\nu = (\det(A))^{-1} \sum_{\mu=1}^{n} \beta_\mu^\nu b^\mu = (\det(A))^{-1} \sum_{\mu=1}^{n} (-1)^{\nu+\mu} \det\left(A_\nu^\mu\right) b^\mu.$$

Die Matrix

$$A(\nu; b) := (A_1, \dots, A_{\nu-1}, b, A_{\nu+1}, \dots, A_n),$$

die also aus A entsteht, indem die ν-te Spalte durch die rechte Seite b ersetzt wird, hat als Determinante (Entwicklung nach ν-ter Spalte) gerade

$$\sum_{\mu=1}^{n}(-1)^{\nu+\mu}\, b^{\mu}\, \det\left(A^{\mu}_{\nu}\right) .$$

Zusammen also:

Satz 6.5.4. Cramer-Regel

Die eindeutige Lösung des linearen Gleichungssystems $Ax = b$ *ist gegeben durch* $x = (x^1, \ldots, x_n)^T$ *mit*

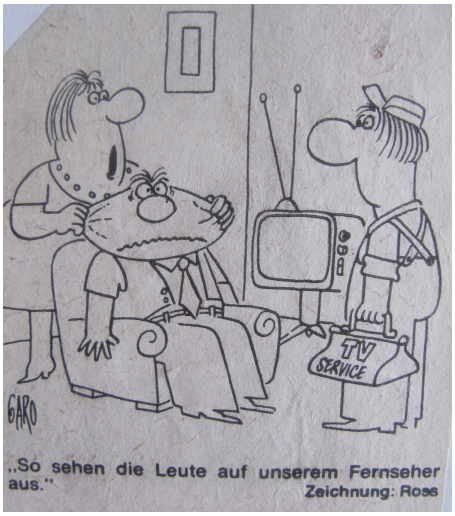
$$x^{\nu} \;=\; \frac{\det A(\nu; b)}{\det(A)}$$

für $\nu = 1, \ldots, n$.

Für große n *sind die beiden vorangehenden Sätze für die explizite (numerische) Berechnung der Inversen beziehungsweise der Lösung eines Gleichungssystems völlig unbrauchbar!* Neben rein numerischen Problemen (Rundungsfehler!) wird der *Aufwand exzessiv!* Aber man kann z. B. die stetige Abhängigkeit der Lösung von der vorgegebenen Matrix A und der rechten Seite b daraus ablesen. Also für theoretische Überlegungen nützlich ...
Für $n = 20$ wären nach der CRAMERschen Regel 21 Determinanten von (20×20)-Matrizen zu berechnen. Würde man noch die Dummheit begehen, diese nach der definierenden Formel zu berechnen, so wären allein $21 \cdot 20! \cdot 19 \approx 9.7 \cdot 10^{20}$ Multiplikationen durchzuführen. (Auch die Bestimmung von $\mathfrak{S}_{20}$ mit allen Vorzeichen ist nicht gerade eine Übungsaufgabe, die übermäßig Freude bereitet.) Ein Supercomputer, der 1 Milliarde Multiplikationen pro Sekunde (vgl. FLOPS: Floating Point Operations Per Second, Gleitkommaoperationen pro Sekunde) durchführt, würde also mindestens — pausenlos — $3 \cdot 10^4$ Jahre benötigen. Es gibt natürlich viele ‚vernünftigere' Methoden, Determinanten zu berechnen. Aber der Aufwand zur Lösung des Gleichungssystems $Ax = b$ ist bei einer *guten Methode* ungefähr so groß wie der zur Berechnung *einer* Determinante. Außerdem wird die Determinante gleich mitgeliefert ...

7. Eigenvektoren, Eigenräume und Normalformen

Eigenwerte und *Eigenvektoren* sind ein fundamentales Werkzeug zur Untersuchung von Endomorphismen und quadratischen Matrizen. *Sie kommen in unzähligen Anwendungen vor!* Beispielhaft seien nur genannt: *Systeme linearer Differentialgleichungen mit konstanten Koeffizienten, Schwingungsvorgänge, Hauptachsentransformation, Stabilitätsuntersuchungen.* Das Thema ist wichtig für viele Gebiete, wie etwa Physik, Elektrotechnik, Maschinenbau, Statik, Biologie, Informatik, Wirtschaftswissenschaften. Eigenwerte und Eigenvektoren beschreiben darin oft besondere Zustände von Systemen.
Aber selbst auf diesem Bildchen lassen sich Verbindungen zum Thema finden:



Beispielsweise werden quantenmechanische Zustände durch Wellenfunktionen, sogenannten *Eigenfunktionen,* beschrieben. Ihnen werden bestimmte quantisierte Werte an Energie, Impuls oder Drehimpuls zugeordnet. Man nennt sie die Eigenwerte der Zustände. Es wird jeder Observablen (d. h. jeder meßbaren Größe, wie z. B. Impuls oder Energie) ein hermitescher Operator zugeordnet. Die Eigenwerte ergeben die Meßwerte. Die wesentlichen Grundlagen für die mathematisch strenge Formulierung der Quantenmechanik wurden im Jahr 1932 durch John von Neumann formuliert. Demnach lässt sich ein physikalisches System allgemein durch drei wesentliche Bestandteile beschreiben: Seine Zustände, seine Observablen und seine Dynamik, d. h. durch seine zeitliche Entwicklung. Resultat der Messung einer physikalischen Größe ist einer der Eigenwerte der entsprechenden Observablen.

Die Bezeichnung „Eigen" ist auch im anglo-amerikanischen Bereich üblich, wo man beispielsweise von *„eigenvalue"* und *„eigenvector"* spricht.

7.1. **Definitionen und Grundlagen.**

Es seien $\mathbb{K}$ ein Körper, $n \in \mathbb{N}$, $\mathfrak{V}$ ein $\mathbb{K}$-Vektorraum mit $\dim \mathfrak{V} = n$, $T \in \mathrm{End}(\mathfrak{V})$ (also T ein Endomorphismus von $\mathfrak{V}$) und $A \in \mathbb{K}^{n\times n}$.

Wir erläutern vorbereitend eine wichtige **Anwendung**, *Systeme linearer Differentialgleichungen mit konstanten Koeffizienten,* am einfachen *Spezialfall* $k = 2$ *und* $\mathbb{K} = \mathbb{R}$: Mit einer $(2, 2)$-Matrix A sei das Differentialgleichungsystem

$$(*) \qquad y' = Ay$$

betrachtet. (erläutern! ∘ ∘ ∘)
Ist A *diagonalisierbar,* d. h.

$$D := S^{-1}AS = \begin{pmatrix} \mu & 0 \\ 0 & \nu \end{pmatrix} \text{ mit } \mu, \nu \in \mathbb{C}$$

für eine geeignete invertierbare $(2,2)$-Matrix S, dann führt die Transformation $z := S^{-1}y$ auf das *‚entkoppelte' Differentialgleichungssystem* $z' = Dz$. Dieses zerfällt für die beiden Komponentenfunktionen z_1 und z_2 in die skalaren Differentialgleichungen $z_1' = \mu z_1$ und $z_2' = \nu z_2$, wovon nicht-triviale Lösungen sofort durch $z_1(x) = \exp(\mu x)$ und $z_2(x) = \exp(\nu x)$ gegeben sind. Die Rücktransformation $y = Sz$ liefert die Komponentenfunktionen y_1 und y_2 als Linearkombinationen von z_1 und z_2.
Die erste Spalte S_1 von S ist ein *Eigenvektor* zum *Eigenwert* μ von A:

$$AS_1 = ASe_1 = SDe_1 = S\mu e_1 = \mu S_1 .$$

Entsprechend ist die zweite Spalte S_2 von S ein Eigenvektor zum Eigenwert ν.

Definition 7.1.1.

Ein $\lambda \in \mathbb{K}$ heißt genau dann *„Eigenwert"* zu T, wenn ein $x \in \mathfrak{V} \setminus \{0\}$ mit $Tx = \lambda x$ existiert.

Ist $\lambda \in \mathbb{K}$ ein Eigenwert zu T, dann heißt

$$E(\lambda) := E(\lambda; T) := \{x \in \mathfrak{V} \mid Tx = \lambda x\} = \ker(\lambda \,\mathrm{id}_{\mathfrak{V}} - T)$$

„Eigenraum" zu λ und T und *jedes*[10] $x \in E(\lambda)$ *„Eigenvektor"*[11] zu λ und T. Für A sind diese Begriffe entsprechend über f_A: $\begin{array}{ccc} \mathbb{K}^n & \longrightarrow & \mathbb{K}^n \\ \cup & & \cup \\ x & \longmapsto & Ax \end{array}$ erklärt. $v(\lambda) := \dim E(\lambda)$ heißt *„geometrische Vielfachheit"* von λ.

Beispiele

(B1) $\mathrm{id}_{\mathfrak{V}}$ hat den Eigenwert 1 mit $E(1; \mathrm{id}_{\mathfrak{V}}) = \mathfrak{V}$.

(B2) Eine Matrix im $\mathbb{R}^2$, die eine Drehung um den Koordinatenursprung mit Winkel $\pi/2$ beschreibt, hat keinen (reellen) Eigenwert. ◁

Wir hatten in Bemerkung 4.5.3 gesehen, wie sich die darstellende Matrix $A = \mathcal{M}_{\mathcal{A}}(T)$ beim Übergang von einer Basis $\mathcal{A}$ zu einer Basis $\mathcal{A}'$ transformiert, nämlich:

$$A' = S^{-1} A S$$

mit $A' := \mathcal{M}_{\mathcal{A}'}(T)$, $S := T^{\mathcal{A}}_{\mathcal{A}'}$ Dies führt zu:

[10] Der Nullvektor gehört also — anders als in vielen Darstellungen — auch dazu!
[11] Ist $\mathfrak{V}$ ein Raum von Funktionen, dann spricht man auch von *Eigenfunktionen.*

Definition 7.1.2.
Zwei Matrizen $N, M \in \mathbb{K}^{n\times n}$ heißen genau dann *„ähnlich“*, in Zeichen $M \approx N$, wenn ein $S \in \mathrm{GL}(n,\mathbb{K})$ mit $N = S^{-1}\,M\,S$ existiert.

Bemerkung 7.1.3.

a) *Zwei Matrizen aus $\mathbb{K}^{n\times n}$ sind genau dann ähnlich, wenn sie bezüglich geeigneter Basen den gleichen Endomorphismus darstellen.*

b) *$\approx$ ist eine Äquivalenzrelation.*

Beweis:

a) : Bemerkungen 4.3.4 und 4.5.3

b) : ∘ ∘ ∘ □

Natürlich stellt sich die *Frage,* wie findet man zu einer gegebenen Matrix eine ähnliche Matrix, die ‚möglichst einfach‘ ist. Anders formuliert: Wie findet man zu dem gegebenen Endomorphismus T eine (geordnete) Basis von $\mathfrak{V}$ derart, daß die darstellende Matrix eine ‚möglichst einfache‘ Gestalt hat? Dieses *Normalformenproblem* ist nicht einfach! Man erhält eine Klassifizierung durch Repäsentanten in Jordan-*Normalform*. Darauf gehen wir in dieser Vorlesung in Abschnitt 7.4 nur ganz kurz ein. Wir beschränken uns weitgehend auf die ungleich einfachere Untersuchung, wann sogar *Diagonalgestalt* erreichbar ist.

Dazu überlegen wir vorweg:
Falls ein $S \in \mathrm{GL}(n,\mathbb{K})$ so existiert, daß $D := S^{-1}\,A\,S$ *Diagonalmatrix* ist, also alle Elemente außerhalb der Hauptdiagonale Null sind, d. h.

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

mit $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$; dann hat eine solche Matrix S *als ν-te Spalte S_ν gerade einen Eigenvektor zum Eigenwert λ_ν*:

Beweis: $D = S^{-1}AS \iff AS = SD$
$\iff \forall \nu \in \mathbb{N}_n \;\; AS_\nu = ASe_\nu = SDe_\nu = S\lambda_\nu e_\nu = \lambda_\nu S_\nu$ □

Wir wissen daher schon, wie man — in diesem Fall — eine solche Transformationsmatrix S erhält: Man muß ‚nur‘ die Eigenwerte und zugehörige Eigenvektoren von A bestimmen.

Von nun an setzen wir wieder — wie in den Abschnitten 6.3, 6.4 und 6.5 — voraus, daß $1 + 1 \neq 0$ gilt.

Definition 7.1.4.

Wir betrachten als wichtiges Hilfsmittel das durch

$$\chi_T(x) := \det(x\,\mathrm{id}_{\mathfrak{V}} - T)$$

(für $x \in \mathbb{K}$) definierte *„charakteristische Polynom“* $\chi_T\colon \mathbb{K} \longrightarrow \mathbb{K}$.

Bemerkung 7.1.5.

Ein $\lambda \in \mathbb{K}$ ist genau dann ein Eigenwert zu T, wenn $\chi_T(\lambda) = 0$ gilt.

Beweis: λ Eigenwert zu $T \iff \exists x \in \mathfrak{V} \setminus \{0\}\ (\lambda\,\mathrm{id}_{\mathfrak{V}} - T)\,x = 0$
$\iff \lambda\,\mathrm{id}_{\mathfrak{V}} - T$ ist nicht injektiv $\iff \det(\lambda\,\mathrm{id}_{\mathfrak{V}} - T) = 0$ □

Satz 7.1.6.

Ist $\mathcal{B}$ eine geordnete Basis von $\mathfrak{V}$ mit $\mathcal{M}_{\mathcal{B}}(T) =: A = (\alpha^\mu_\nu)$, dann ist für $x \in \mathbb{K}$

$$\chi_T(x) = \det(x\,\mathbb{1}_n - A) = \begin{vmatrix} x-\alpha^1_1 & -\alpha^1_2 & \cdots & -\alpha^1_n \\ -\alpha^2_1 & x-\alpha^2_2 & \cdots & -\alpha^2_n \\ \vdots & \vdots & \ddots & \vdots \\ -\alpha^n_1 & -\alpha^n_2 & \cdots & x-\alpha^n_n \end{vmatrix} =: \varphi_A(x)\,.$$

Es existieren $c_1, \ldots, c_n \in \mathbb{K}$ derart, daß

$$\varphi_A(x) = x^n + c_1 x^{n-1} + \cdots + c_{n-1}x + c_n$$

(normiertes Polynom vom Grade n), wobei

$$-c_1 = \sum_{\nu=1}^{n} \alpha^\nu_\nu := \mathrm{spur}(A) \quad \text{(„Spur von } A\text{“)}$$

und

$$c_n = (-1)^n \det A\,.$$

Beweis:

$$\chi_T(x) = \det(x\,\mathrm{id}_{\mathfrak{V}} - T) \underset{(6.4.6)}{=} \det(\mathcal{M}_{\mathcal{B}}(x\,\mathrm{id}_{\mathfrak{V}} - T))$$
$$\underset{(4.3.4)}{=} \det(x\,\mathcal{M}_{\mathcal{B}}(\mathrm{id}_{\mathfrak{V}}) - \mathcal{M}_{\mathcal{B}}(T)) = \det(x\,\mathbb{1}_n - A)$$

Wir bezeichnen: $x\,\mathbb{1}_n - A =: (\beta^\mu_\nu)$, also $\beta^\nu_\nu = x - \alpha^\nu_\nu$ und $\beta^\mu_\nu = -\alpha^\mu_\nu \quad (\nu \neq \mu)$. Dann hat man:

$$\det(x\,\mathbb{1}_n - A) = \prod_{\nu=1}^{n}(x - \alpha^\nu_\nu) + \sum_{\sigma \in \mathfrak{S}_n \setminus \{e\}} \mathrm{sign}\,\sigma \prod_{\nu=1}^{n} \beta^{\sigma(\nu)}_\nu$$

Für $\sigma \in \mathfrak{S}_n \setminus \{e\}$ beschreibt das Produkt $\prod_{\nu=1}^{n} \beta^{\sigma(\nu)}_\nu$ ein Polynom q vom Grade höchstens $n-2$ $\quad(\circ\ \circ\ \circ)$. Zusammen haben wir

$\det(x\mathbb{1}_n - A) = x^n - \left(\sum_{\nu=1}^{n} \alpha_\nu^\nu\right) x^{n-1} + \widetilde{q}(x)$ mit einem Polynom $\widetilde{q}$ vom Grade höchstens $n-2$ und $c_n = \varphi_A(0) = \det(-A) = (-1)^n \det(A)$. □

Folgerung 7.1.7.

Ähnliche Matrizen haben das gleiche charakteristische Polynom, damit gleiche Determinante und gleiche Spur.

Beweis: ∘ ∘ ∘ □

Satz 7.1.8.

Für paarweise verschiedene Eigenwerte $\lambda_1, \ldots, \lambda_m \in \mathbb{K}$ zu T (mit $m \in \mathbb{N}$) gelten:

a) Die Summe $E(\lambda_1) + \cdots + E(\lambda_m)$ ist direkt, wir notieren:
$\bigoplus_{\mu=1}^{m} E(\lambda_\mu)$ *beziehungsweise* $E(\lambda_1) \oplus \cdots \oplus E(\lambda_m)$ (vgl. Übung (5.4))

b) Vektoren $x_1, \ldots, x_m$ mit $x_\mu \in E(\lambda_\mu) \setminus \{0\}$ $(\mu \in \mathbb{N}_m)$ sind linear unabhängig.

c) $\sum_{\mu=1}^{m} v(\lambda_\mu) \leq \dim(\mathfrak{V})$

d) $\mathfrak{V} = \bigoplus_{\mu=1}^{m} E(\lambda_\mu) \iff \sum_{\mu=1}^{m} v(\lambda_\mu) = \dim(\mathfrak{V})$

Beweis:

a): (induktiv): $\underline{m=1}$: ✓ $\underline{m-1 \rightsquigarrow m}$:
Aus $0 = x_1 + \cdots + x_m$ mit $x_\mu \in E(\lambda_\mu)$ $(\mu \in \mathbb{N}_m)$ folgen:
$0 = T0 = \lambda_1 x_1 + \cdots + \lambda_m x_m$ und
$0 = \lambda_m x_1 + \cdots + \lambda_m x_m$, also
$0 = (\lambda_m - \lambda_1)x_1 + \cdots + (\lambda_m - \lambda_{m-1})x_{m-1}$
Nach Induktionsvoraussetzung hat man $(\lambda_m - \lambda_\mu)x_\mu = 0$, somit $x_\mu = 0$ $(\mu \in \mathbb{N}_{m-1})$ und schließlich $x_m = 0$.

b): $0 = \sum_{\mu=1}^{m} \underbrace{\alpha_\mu x_\mu}_{\in E(\lambda_\mu)} \underset{a)}{\Longrightarrow} \alpha_\mu x_\mu = 0$ $(\mu \in \mathbb{N}_m)$, also $\alpha_\mu = 0$ $(\mu \in \mathbb{N}_m)$.

c): $\sum_{\mu=1}^{m} v(\lambda_\mu) = \sum_{\mu=1}^{m} \dim E(\lambda_\mu) \underset{\text{(Bem. 3.5.3)}}{=} \dim\left(\bigoplus_{\mu=1}^{m} E(\lambda_\mu)\right) \leq \dim \mathfrak{V}$

d): $\sum_{\mu=1}^{m} v(\lambda_\mu) \underset{c)}{=} \dim\left(\bigoplus_{\mu=1}^{m} E(\lambda_\mu)\right) = \dim \mathfrak{V} \underset{(3.3.7)}{\Longleftrightarrow} \bigoplus_{\mu=1}^{m} E(\lambda_\mu) = \mathfrak{V}$

□

7.2. **Diagonalisierbarkeit.**

Es seien wieder $\mathbb{K}$ ein Körper mit $1+1 \neq 0$, $n \in \mathbb{N}$, $\mathfrak{V}$ ein $\mathbb{K}$-Vektorraum mit $\dim \mathfrak{V} = n$, $T \in \mathrm{End}(\mathfrak{V})$ und $A \in \mathbb{K}^{n\times n}$.

Definition 7.2.1.

T heißt genau dann *„diagonalisierbar“*, wenn eine (geordnete) Basis $\mathcal{B}$ von $\mathfrak{V}$ so existiert, daß $\mathcal{M}_{\mathcal{B}}(T)$ Diagonalmatrix ist.

Satz 7.2.2.

T ist genau dann diagonalisierbar, wenn es eine Basis von $\mathfrak{V}$ aus Eigenvektoren von T gibt.

Beweis:

Ist T diagonalisierbar, dann existieren eine (geordnete) Basis $\mathcal{B} = (b_1, \ldots, b_n)$ von $\mathfrak{V}$ und $\lambda_\nu \in \mathbb{K}$ $(\nu \in \mathbb{N}_n)$ derart, daß

$$\mathcal{M}_{\mathcal{B}}(T) = \begin{pmatrix} \lambda_1 & 0 & & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & \cdots & & 0 & \lambda_n \end{pmatrix}$$

gilt. Das bedeutet aber gerade $Tb_\nu = \lambda_\nu b_\nu$ $(\nu \in \mathbb{N}_n)$.

Ist $\mathcal{B} = (b_1, \ldots, b_n)$ eine Basis von $\mathfrak{V}$ aus Eigenvektoren von T, dann existieren $\lambda_\nu \in \mathbb{K}$ mit $Tb_\nu = \lambda_\nu b_\nu$ $(\nu \in \mathbb{N}_n)$; damit ist dann:

$$\mathcal{M}_{\mathcal{B}}(T) = \begin{pmatrix} \lambda_1 & 0 & & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & \cdots & & 0 & \lambda_n \end{pmatrix}$$

□

Beispiele

(B3) $\boxed{\mathbb{K} := \mathbb{R}}$, $A := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$

Hier hat man $x\mathbb{1}_2 - A = \begin{pmatrix} x & 1 \\ -1 & x \end{pmatrix}$, also $\varphi_A(x) = x^2 + 1$. Es existiert kein Eigenwert zu A; daher ist A *nicht diagonalisierbar*.

(B4) $\boxed{\mathbb{K} := \mathbb{C}}$, $A := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$

Hier ist $x\mathbb{1}_2 - A = \begin{pmatrix} x & -1 \\ 0 & x \end{pmatrix}$ und somit $\varphi_A(x) = x^2$. $x = 0$ ist einziger Eigenwert zu A mit

$$E(0) = \ker(A) = \left\{ \begin{pmatrix} \alpha \\ 0 \end{pmatrix} : \alpha \in \mathbb{C} \right\} \neq \mathbb{C}^2$$

Daher ist auch diese Matrix A *nicht diagonalierbar*. ◁

Definition 7.2.3.

Ist $\lambda \in \mathbb{K}$ ein Eigenwert zu T, dann bezeichnen wir die ‚Vielfachheit' von λ als Nullstelle des charakteristischen Polynoms χ_T mit $o(\lambda)$ also:

$$o(\lambda) = \max\{r \in \mathbb{N} \mid \chi_T(x) = (x-\lambda)^r q(x) \text{ mit einem Polynom } q\}$$

Wir sagen auch *„Ordnung von λ"* oder *„algebraische Multiplizität"*.

Bemerkung 7.2.4.

Ist $\lambda \in \mathbb{K}$ ein Eigenwert zu T, dann gilt:

$$v(\lambda) \leq o(\lambda)$$

Beweis: Es gelten:

$T(E(\lambda)) \subset E(\lambda)$, $T_0 := T_{/E(\lambda)} = \lambda \,\mathrm{id}_{E(\lambda)}$, also mit $p := v(\lambda)$:

$$\begin{aligned}\chi_{T_0}(x) &= \det\left(x\,\mathrm{id}_{E(\lambda)} - T_0\right) = \det\left((x-\lambda)\,\mathrm{id}_{E(\lambda)}\right)\\ &= \det\left((x-\lambda)\mathbb{1}_p\right) = (x-\lambda)^p.\end{aligned}$$

Ergänzt man eine Basis von $E(\lambda)$ zu einer Basis $\mathcal{B}$ von $\mathfrak{V}$, dann hat die Matrix $A := \mathcal{M}_{\mathcal{B}}(T)$ die Blockgestalt wie in Satz 6.5.1. Damit gilt:

$$\chi_T(x) = \varphi_A(x) = \det\left(x\,\mathbb{1}_p - A_1^1\right)\det\left(x\,\mathbb{1}_q - A_2^2\right)$$

Das durch $Q(x) := \det(x\,\mathbb{1}_q - A_2^2)$ gegebene Polynom Q liefert also $\chi_T(x) = (x-\lambda)^p Q(x)$, somit $v(\lambda) = p \leq o(\lambda)$. □

In Beispiel (B4) gilt $1 = v(0) < o(0) = 2$.

Definition 7.2.5.

Der Körper $\mathbb{K}$ heißt genau dann *„algebraisch abgeschlossen"*, wenn jedes Polynom P über $\mathbb{K}$ in ein Produkt von ‚Linarfaktoren' ‚zerfällt':

$$P(x) = \alpha_0 \prod_{\nu=1}^{n}(x-\alpha_\nu) \text{ mit } \alpha_\nu \in \mathbb{K}$$

Der klassische *Fundamentalsatz der Algebra*, den man *einfach* mit funktionentheoretischen Hilfsmitteln beweist (vgl. dazu etwa das schöne Buch [4]), besagt gerade, daß $\mathbb{C}$ algebraisch abgeschlossen ist. *Sie* dürfen die Annahme „$\mathbb{K}$ algebraisch abgeschlossen" durch „$\mathbb{K}$= $\mathbb{C}$" ersetzen!

Satz 7.2.6.

Ist $\mathbb{K}$ algebraisch abgeschlossen[12]*, dann gilt:*

T diagonalisierbar $\iff$ Für alle Eigenwerte λ von T ist $v(\lambda) = o(\lambda)$

Beweis: Existiert eine Basis $\mathcal{B} = (b_1, \ldots, b_n)$ von $\mathfrak{V}$ und $\mu_j \in \mathbb{K}$ so, daß

$$\mathcal{M}_{\mathcal{B}}(T) = \begin{pmatrix}\mu_1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \mu_n\end{pmatrix},$$

[12] Diese Voraussetzung läßt sich offenbar abschwächen zu: Das charakteristische Polynom zu T zerfällt in Linearfaktoren.

also $(*)\ Tb_j = \mu_j b_j \quad (j \in \mathbb{N}_n)$, dann ist $\chi_T(x) = \prod_{\nu=1}^{n}(x-\mu_\nu)$. Ist λ ein Eigenwert zu T mit $o(\lambda) =: r$, dann sind genau r der μ_ν gleich λ; nach $(*)$ ist dann auch $v(\lambda) = r$. Es seien (in der anderen Richtung) $\lambda_1, \ldots, \lambda_m$ die paarweise verschiedenen Eigenwerte von T. Da $\mathbb{K}$ algebraisch abgeschlossen ist:

$$\dim \mathfrak{V} \overset{\checkmark}{=} \sum_{\mu=1}^{m} o(\lambda\mu) \underset{\text{(Vor.)}}{=} \sum_{\mu=1}^{m} v(\lambda\mu)$$

Nach Satz 7.2.2 folgt die Diagonalisierbarkeit von T. □

7.3. **Der Spektralsatz für normale Endomorphismen.**

Es sei nun $(\mathfrak{V}, \langle\ ,\ \rangle)$ ein endlich-dimensionaler unitärer Vektorraum, also insbesondere $\boxed{\mathbb{K} = \mathbb{C}}$.

Bemerkung 7.3.1.

Ist λ ein Eigenwert eines normalen *Endomorphismus T von $\mathfrak{V}$ mit zugehörigem Eigenvektor x, dann ist $\overline{\lambda}$ ein Eigenwert von T^* mit gleichem Eigenvektor x.*

Beweis: Nach den Bemerkungen 5.3.8 und 5.3.4 hat man

$$\|(\lambda\,\mathrm{id}_{\mathfrak{V}} - T)x\| = \|(\overline{\lambda}\,\mathrm{id}_{\mathfrak{V}} - T^*)x\|$$

für beliebige $\lambda \in \mathbb{K}$ und $x \in \mathfrak{V}$; denn mit T ist $\lambda\,\mathrm{id}_{\mathfrak{V}} - T$ normal. Daraus liest man die Behauptung unmittelbar ab. □

Bemerkung 7.3.2.

Sind λ_1, λ_2 verschiedene Eigenwerte eines normalen *Endomorphismus T, dann sind die Eigenräume $E(\lambda_1)$ und $E(\lambda_2)$ orthogonal.*

Beweis: Für $x \in E(\lambda_1)$ und $y \in E(\lambda_2)$ hat man:

$$\lambda_1 \langle x, y\rangle = \langle \lambda_1 x, y\rangle = \langle Tx, y\rangle = \langle x, T^*y\rangle \underset{(7.3.1)}{=} \langle x, \overline{\lambda_2} y\rangle = \lambda_2 \langle x, y\rangle$$

Mit $\lambda_1 \neq \lambda_2$ folgt so: $x \perp y$. □

Definition 7.3.3.

Sind $\mathfrak{V}_1, \ldots, \mathfrak{V}_m$ (für ein $m \in \mathbb{N}$) Unterräume von $\mathfrak{V}$, dann heißt $\mathfrak{V}$ *„orthogonale Summe“* dieser Unterräume, in Zeichen

$$\mathfrak{V} = \bigoplus_{\mu=1}^{m} \mathfrak{V}_\mu \quad \text{oder} \quad \mathfrak{V} = \mathfrak{V}_1 \oplus \cdots \oplus \mathfrak{V}_m,$$

genau dann, wenn $\mathfrak{V} = \mathfrak{V}_1 + \cdots + \mathfrak{V}_m$ und $\mathfrak{V}_\mu \perp \mathfrak{V}_\nu$ für alle $\mu, \nu \in \mathbb{N}_m$ mit $\nu \neq \mu$ gelten.

Bemerkung 7.3.4.

Jede orthogonale Summe ist direkt.

Beweis: ∘ ∘ ∘ □

Definition 7.3.5.

Eine Projektion[13] $P\colon \mathfrak{V} \longrightarrow \mathfrak{V}$ heißt genau dann *„orthogonal“*, wenn $\operatorname{im} P \perp \ker P$ gilt.

Beispiel

(B5) In jeder orthogonalen Summe sind die zugehörigen Projektionen orthogonal. ◁

Der folgende Satz ist wohl der wichtigste und leistungsfähigste der Linearen Algebra! Der Beweis ist etwas aufwendig; deshalb werden sehr oft nur Spezialfälle behandelt. Ich gehe — aus Zeitgründen — auf den Beweis nicht mehr ein, behandle stattdessen lieber einige Folgerungen. Wer es unbedingt genauer wissen möchte, findet die Ideen beispielsweise in den Büchern [6] und [9].

Satz 7.3.6. *Hauptsatz (Spektralsatz für normale Endomorphismen)*

Es seien — für ein $m \in \mathbb{N}$ — $\lambda_1, \ldots, \lambda_m \in \mathbb{C}$ *paarweise verschiedene Eigenwerte eines Endomorphismus* T *von* $\mathfrak{V}$. *Dann sind folgende Aussagen* äquivalent:

(a) T *ist normal.*

(b) $o(\lambda_\mu) = v(\lambda_\mu) \quad (\mu \in \mathbb{N}_m)$ *und* $E(\lambda_\mu) \perp E(\lambda_\nu)$ *für alle* $\mu, \nu \in \mathbb{N}_m$ *mit* $\nu \neq \mu$

(c) $\mathfrak{V} = \overset{m}{\underset{\mu=1}{\perp\!\!\!\bigoplus}} E(\lambda_\mu)$

(d) *Es existiert eine Basis* $\mathcal{B} = (b_1, \ldots, b_n)$ *von* $\mathfrak{V}$ *mit* $\langle b_\mu, b_\nu \rangle = \delta^\mu_\nu$ *(...) derart, daß* $\mathcal{M}_\mathcal{B}(T)$ *Diagonalmatrix ist.*

(e) *Es existieren* Orthogonalprojektionen $P_\mu\colon \mathfrak{V} \longrightarrow \mathfrak{V}$ *für* $\mu \in \mathbb{N}_m$ *so, daß*

(1) $\mathrm{id}_\mathfrak{V} = \sum_{\mu=1}^{m} P_\mu$

(2) $P_\mu \circ P_\nu = 0$ *für* $\nu \neq \mu \quad (\ldots)$

(3) $T = \sum_{\mu=1}^{m} \lambda_\mu P_\mu$ *(„Spektraldarstellung“)*

[13] Es gilt $P^2 = P$.

In (d) gilt: *Die Diagonalelemente von* $\mathcal{M}_{\mathcal{B}}(T)$ *sind die Eigenwerte von* T*, die Basisvektoren* b_ν *sind zugehörige Eigenvektoren.*

Folgerung 7.3.7.

Ist T *ein* hermitescher *oder* unitärer *Endomorphismus von* $\mathfrak{V}$ *mit paarweise verschiedenen Eigenwerten* $\lambda_1, \ldots, \lambda_m \in \mathbb{C}$*, dann gelten die Aussagen* $(b), (c), (d)$ *und* (e) *aus Satz 7.3.6.*

Beweis der Folgerung: T ist jeweils normal. □

Die Eigenwerte eines *normalen* Endomorphismus T geben viel Information über T:

Satz 7.3.8.

Für die paarweise verschiedenen Eigenwerte $\lambda_1, \ldots, \lambda_m \in \mathbb{C}$ *(mit einem* $m \in \mathbb{N}$*) eines* normalen *Endomorphismus* T *von* $\mathfrak{V}$ *gelten:*

a) T *ist hermitesch* $\iff \forall\, \mu \in \mathbb{N}_m \;\; \lambda_\mu \in \mathbb{R}$

b) T *ist unitär* $\iff \forall\, \mu \in \mathbb{N}_m \;\; |\lambda_\mu| = 1$

c) T *ist invertierbar* $\iff \forall\, \mu \in \mathbb{N}_m \;\; \lambda_\mu \neq 0$

d) T *ist idempotent (d. h.* $T^2 = T$*)* $\iff \forall\, \mu \in \mathbb{N}_m \;\; \lambda_\mu \in \{0, 1\}$

e) T *ist positiv semidefinit* $\iff \forall\, \mu \in \mathbb{N}_m \;\; \lambda_\mu \in [0, \infty)$,

f) T *ist positiv definit* $\iff \forall\, \mu \in \mathbb{N}_m \;\; \lambda_\mu \in (0, \infty)$,

Definition 7.3.9.

Natürlich heißt $T \in \mathrm{End}(\mathfrak{V})$ genau dann *„positiv semidefinit“*, wenn $\langle Tx, x\rangle \in [0, \infty)$ für alle $x \in \mathfrak{V}$ gilt, und *„positiv definit“*, wenn $\langle Tx, x\rangle \in (0, \infty)$ für alle $x \in \mathfrak{V} \setminus \{0\}$ gilt.

Beweis:

Nach Satz 7.3.6 existieren — mit $n := \dim \mathfrak{V}$ — eine orthonormierte Basis $\mathcal{B} = (b_1, \ldots, b_n)$ von $\mathfrak{V}$ und $\mu_\nu \in \mathbb{K}$ $(\nu \in \mathbb{N}_n)$ derart, daß

$$A := \mathcal{M}_{\mathcal{B}}(T) = \begin{pmatrix} \mu_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \mu_n \end{pmatrix}$$

Dabei sind die μ_ν gerade die Eigenwerte von T.

$a)$: T hermitesch $\underset{(5.3.13)}{\iff}$ A hermitesch $\iff A = A^*$

$\iff \mu_\nu = \overline{\mu_\nu} \quad (\nu \in \mathbb{N}_n) \iff \lambda_\mu = \overline{\lambda_\mu} \quad (\mu \in \mathbb{N}_m)$

$b)$: T unitär $\underset{(5.3.13)}{\iff}$ A unitär $\overset{\checkmark}{\iff} |\mu_\nu| = 1$ $(\nu \in \mathbb{N}_n) \iff |\lambda_\mu| = 1$ $(\mu \in \mathbb{N}_m)$

$c)$: T invertierbar $\underset{\text{(4.3.7) und ...}}{\iff}$ A invertierbar $\iff \det A \neq 0$
$\iff \mu_\nu \neq 0\ (\nu \in \mathbb{N}_n) \iff \lambda_\mu \neq 0\ (\mu \in \mathbb{N}_m)$

$d)$: T idempotent $\iff A^2 = A \iff \mu_\nu^2 = \mu_\nu\ (\nu \in \mathbb{N}_n)$
$\iff \lambda_\mu^2 = \lambda_\mu\ (\mu \in \mathbb{N}_m) \iff \lambda_\mu^2 \in \{0,1\}\ (\mu \in \mathbb{N}_m)$

$e)$: T positiv semidefinit $\iff \forall\,(\alpha_1,\dots,\alpha_n) \in \mathbb{C}^n$

$$0 \leq \left\langle T\left(\sum_{\nu=1}^n \alpha_\nu b_\nu\right), \sum_{k=1}^n \alpha_k b_k \right\rangle = \left\langle \sum_{\nu=1}^n \alpha_\nu \mu_\nu b_\nu, \sum_{k=1}^n \alpha_k b_k \right\rangle$$

$$= \sum_{\nu,k} \alpha_\nu \overline{\alpha_k} \mu_\nu \langle b_\nu, b_k \rangle = \sum_{\nu=1}^n |\alpha_\nu|^2 \mu_\nu .$$ Dies bedeutet aber gerade $\mu_\nu \geq 0\ (\dots)$ beziehungsweise $\lambda_\mu \geq 0\ (\dots)$

$f)$: ‚ebenso' wie $e)$ □

Definition 7.3.10.

Eine Matrix $A \in \mathbb{C}^{n\times n}$ heißt genau dann *„normal"*, wenn gilt:

$$AA^* = A^*A$$

Insbesondere sind also hermitesche und unitäre Matrizen normal.

Nach den Bemerkungen 4.3.6 und 5.3.5 ist T genau dann normal, wenn eine zugehörige darstellende Matrix normal ist. Für Matrizen liefert der Spektralsatz:

Folgerung 7.3.11.

Es seien $n \in \mathbb{N}$ und $A \in \mathbb{C}^{n\times n}$ normal. Dann existiert ein $U \in U(n)$, derart, daß

$$U^{-1}AU = U^*AU \quad \textit{Diagonalmatrix ist.}$$

Sprechweise: *„A ist unitär diagonalisierbar."*

Beweis: Auf dem $\mathbb{C}^n$ betrachten wir das Standard-Skalarprodukt und wenden den Spektralsatz auf: f_A:

$$\begin{array}{ccc} \mathbb{C}^n & \longrightarrow & \mathbb{C}^n \\ \cup & & \cup \\ x & \longmapsto & Ax \end{array}$$

∘ ∘ ∘ □

Folgerung 7.3.12.

Ist eine Matrix $A \in \mathbb{K}^{n\times n}$ unitär diagonalisierbar, dann ist sie normal.

Beweis: Ist $U^*AU =: D$ eine Diagonalmatrix mit einem $U \in U(n)$, dann ist D normal und damit:

$$AA^* = (UDU^*)(UD^*U^*) = UDD^*U^* = UD^*DU^* = \dots = A^*A$$

□

Auf wichtige Folgerungen für den Fall $\boxed{\mathbb{K} = \mathbb{R}}$ gehe ich — aus Zeitgründen — nicht mehr ein.

7.4. **Anmerkungen zur JORDAN-Normalform.**

Marie Ennemond Camille JORDAN (1838–1922) war sehr vielseitig in seinem wissenschaftlichen Schaffen. Er arbeitete anfangs auf dem Gebiet der algebraischen Gleichungen. Ab etwa 1867 stand bei ihm die reelle Analysis im Vordergrund. Er führte den Begriff „Funktionen von beschränkter Schwankung“ ein und zeigte dessen Beziehung zu monotonen Funktionen. Neben maßtheoretischen Überlegungen beschäftigte er sich in der Topologie mit den heute nach ihm benannten Kurven, bewies seinen berühmten Kurvensatz und führte den Begriff der Homotopie ein. Zudem beschäftigte er sich mit Kristallographie, Wahrscheinlichkeitsrechnung und Anwendungen der Mathematik in der Technik. Viele Studenten der Mathematik ‚leiden‘ zu Beginn ihres Studiums beim Bemühen, seine Normalform von Matrizen richtig zu verstehen.

Es seien k eine natürliche Zahl, $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $A \in \mathbb{M}_k := \mathbb{K}^{k\times k}$. Wir notieren auch $E := \mathbb{1}_k$.

Wir erläutern wieder vorbereitend einige der folgenden Überlegungen am einfachen *Spezialfall* $k = 2$: Wir wissen schon: Ist eine $(2,2)$-Matrix A *diagonalisierbar,* also

$$D := S^{-1}AS = \begin{pmatrix} \mu & 0 \\ 0 & \nu \end{pmatrix} \quad \text{mit} \quad \mu, \nu \in \mathbb{C}$$

für eine geeignete invertierbare $(2,2)$-Matrix S, dann ist die erste Spalte S_1 von S ein *Eigenvektor* zum *Eigenwert* μ von A:

$$AS_1 = ASe_1 = SDe_1 = S\mu e_1 = \mu S_1 .$$

Entsprechend ist die zweite Spalte S_2 von S ein Eigenvektor zum Eigenwert ν von A.

Ist eine Matrix *nicht* diagonalisierbar, dann wird man versuchen, wenigstens eine „fast“ so einfache Form zu finden. Dieses sichert der Satz über die JORDAN-Normalform (siehe Seite 113).

Ist A *nicht diagonalisierbar,* dann existiert nach diesem Satz eine invertierbare $(2,2)$-Matrix S so, daß

$$J := S^{-1}AS = \begin{pmatrix} \mu & 1 \\ 0 & \mu \end{pmatrix} \quad \text{mit einem } \mu \in \mathbb{C} .$$

Die erste Spalte S_1 von S ist wieder ein *Eigenvektor* zum *Eigenwert* μ von A. Die zweite Spalte S_2 ist ein zugehöriger *Hauptvektor der Stufe 2:*

$$(A-\mu E)\,S_2 = ASe_2 - \mu S_2 = SJe_2 - \mu S_2 = S(e_1+\mu e_2) - \mu S_2 = S_1 .$$

Wir betrachten wieder das Differentialgleichungssystem

$$(*) \qquad y' = Ay$$

Die Transformation $z := S^{-1}y$ führt hier auf das Differentialgleichungssystem $z' = Jz$ und mit den beiden Komponentenfunktionen z_1 und z_2 zu den einfachen skalaren Differentialgleichungen $z_2' = \mu z_2$ und $z_1' = \mu z_1 + z_2$. Mit komplexen Zahlen c und d ist die Lösung z_2 der ersten Differentialgleichung durch $z_2(x) = d\exp(\mu x)$

und dann die der zweiten durch $z_1(x) = (c + dx)\exp(\mu x)$ gegeben. Damit führt

$$y(x) = \begin{pmatrix} y_1(x) \\ y_2(x) \end{pmatrix} = S\,z(x) = S \begin{pmatrix} z_1(x) \\ z_2(x) \end{pmatrix}$$

zur Lösung von $(*)$.

Man sieht schon hier, daß für den allgemeinen Fall von solchen (speziellen) Differentialgleichungssystemen für die vereinfachende Transformation auf *Normalformen Eigenwerte, Eigenvektoren* und gegebenenfalls *Hauptvektoren* eine zentrale Rolle spielen werden.

☺ Obwohl es vielleicht überflüssig ist, sehen wir uns noch den Fall $k = 3$ an: Es seien $A \in \mathbb{M}_3$ und $S \in \mathrm{GL}(3, \mathbb{K})$ mit

$$S^{-1}AS = J := \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}.$$

Mit den Spalten S_1, S_2, S_3 von S und $E := \mathbb{1}_3$ gelten dann:

$$AS = SJ, \text{ somit } AS_1 = S\lambda e_1 = \lambda S_1\,.$$

λ ist also ein *Eigenwert* mit zugehörigem Eigenvektor S_1:

$$(A - \lambda E)\,S_1 = 0$$

$$AS_2 = SJe_2 = S(e_1 + \lambda e_2) = S_1 + \lambda S_2, \text{ also}$$

$(A - \lambda E)\,S_2 = S_1$, somit $(A - \lambda E)^2\,S_2 = 0$. S_2 ist ein *Hauptvektor der Stufe 2.*

$$AS_3 = SJe_3 = S(e_2 + \lambda e_3) = S_2 + \lambda S_3, \text{ also}$$

$(A - \lambda E)\,S_3 = S_2$, somit $(A - \lambda E)^3\,S_3 = 0$. S_3 ist ein *Hauptvektor der Stufe 3.*

Gesucht ist somit ein $S_3 \in \mathbb{K}^3$ mit $(A - \lambda E)^3\,S_3 = 0$, aber $(A - \lambda E)^2\,S_3 \neq 0$. Mit $S_2 := (A - \lambda E)\,S_3$ und $S_1 := (A - \lambda E)\,S_2 = (A - \lambda E)^2\,S_3$ erhält man eine *Hauptvektorkette.*

Allgemein: Sind $\varrho \in \mathbb{N}$, $\lambda \in \mathbb{C}$ und $b \in \mathbb{C}^k$ mit $(A - \lambda E)^\varrho b = 0$ und $(A - \lambda E)^{\varrho-1} b \neq 0$, also λ ein *Eigenwert* zu A und b ein *Hauptvektor* zu λ der *Ordnung (Stufe)* ϱ, dann ist $b_{\varrho-\nu} := (A - \lambda E)^\nu b$ für $\nu = 0, \ldots, \varrho - 1$ ein Hauptvektor der Ordnung $\varrho - \nu$, b_1 also ein *Eigenvektor.*

Die Menge aller Hauptvektoren zu λ bezeichnet man als *Hauptraum* (zu λ).

Zu $r \in \mathbb{N}$ und $\lambda \in \mathbb{C}$ bezeichnen wir die (r, r)-Matrix

$$J(\lambda) := J_r(\lambda) := \begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix} = \lambda\mathbb{1}_r + N_r$$

$$\text{mit } N_r := \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix} = J_r(0)$$

als JORDAN-*Elementarmatrix*, JORDAN-*Block* oder JORDAN-*Kästchen.*

Alle Einträge auf der Hauptdiagonalen sind gleich, in der Nebendiagonalen oberhalb der Hauptdiagonalen stehen lauter Einsen, alle anderen Einträge sind Null.
Die Ordnung (algebraische Multiplizität) ist hier r, die geometrische Vielfachheit 1.

Der Satz über die JORDAN-Normalform läßt sich damit wie folgt formulieren:

Zu $A \in \mathbb{M}_k$ existiert eine invertierbare Matrix $S \in \mathbb{M}_k$ derart, daß

$$S^{-1}AS = \begin{pmatrix} J(\lambda_1) & & & & & & \mathbb{O} \\ & \ddots & & & & & \\ & & J(\lambda_1) & & & & \\ & & & \ddots & & & \\ & & & & \ddots & & \\ & & & & & J(\lambda_s) & \\ & & & & & & \ddots \\ \mathbb{O} & & & & & & J(\lambda_s) \end{pmatrix}$$

mit JORDAN-*Kästchen $J(\lambda_\sigma)$ geeigneter ‚Länge' zu den paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_s$.* Diese spezielle Gestalt einer Matrix heißt JORDAN-*Normalform*.

Es ist also die spezielle Form einer quadratischen Matrix, bei der längs der Hauptdiagonalen lauter JORDAN-Kästchen angeordnet sind und die ansonsten nur Nullen als Einträge aufweist. Bis auf die Anordnung der JORDAN-Kästchen ist die JORDAN-Normalform eindeutig bestimmt.

Ein *Beweis* dieses Satzes ist aufwendig und nicht einfach. Er ist wohl der schwierigste der elementaren Linearen Algebra! Wir gehen darauf nicht mehr ein, bringen aber zumindest ein ausführlich durchgerechnetes einfaches Beispiel dazu.

Beispiel

$$A := \begin{pmatrix} 0 & 1 & 0 \\ 4 & 3 & -4 \\ 1 & 2 & -1 \end{pmatrix}$$

Zunächst sind die *Eigenwerte von* A zu bestimmen:

$$\det(\lambda E - A) = \begin{vmatrix} \lambda & -1 & 0 \\ -4 & \lambda - 3 & 4 \\ -1 & -2 & \lambda + 1 \end{vmatrix}$$
$$\overset{14}{=} -4(-2\lambda - 1) + (\lambda + 1)(\lambda^2 - 3\lambda - 4)$$

$$= \lambda^3 - 2\lambda^2 + \lambda \;=\; \lambda(\lambda^2 - 2\lambda + 1) \;=\; \lambda(\lambda - 1)^2$$

Die Eigenwerte sind 0 mit $o(0) = 1$ und 1 mit $o(1) = 2$. Jetzt sind zu diesen beiden Eigenwerten *Eigenvektoren und ‚Hauptvektorketten'* zu bestimmen:

Ohne Rechnung erkennt man sofort $A \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = 0$:

Somit ist $b_1^{(0)} := \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ ein Eigenvektor zu 0.

Die Matrix $A - 1E = A - E = \begin{pmatrix} -1 & 1 & 0 \\ 4 & 2 & -4 \\ 1 & 2 & -2 \end{pmatrix}$ hat offenbar den Rang 2. Folglich ist ihr Kern, also der Eigenraum zum Eigenwert 1, eindimensional. Es gibt also keine zwei linear unabhängigen Eigenvektoren zu 1. Damit kennen wir schon die Gestalt der zugehörigen JORDAN-Normalform

$$\left(\begin{array}{c|cc} 0 & 0 & 0 \\ \hline 0 & 1 & 1 \\ 0 & 0 & 1 \end{array}\right)$$

— bis auf mögliche Vertauschung der beiden Kästchen.

Es ist $(A - 1E)^2 = \begin{pmatrix} 5 & 1 & -4 \\ 0 & 0 & 0 \\ 5 & 1 & -4 \end{pmatrix}$ mit $\operatorname{rang}(A - 1E)^2 = 1$. Ein Hauptvektor zum Eigenwert 1 ist $b_2^{(1)} := \begin{pmatrix} 0 \\ 4 \\ 1 \end{pmatrix}$ mit zugehörigem Eigenvektor $b_1^{(1)} := (A - 1E)b_2^{(1)} = \begin{pmatrix} 4 \\ 4 \\ 6 \end{pmatrix}$. ◁

Für die praktische Rechnung halten wir noch allgemein fest:

Zu jedem Eigenwert gibt es seiner geometrischen Vielfachheit entsprechend viele Jordanblöcke. (Anzahl der Kästchen)

Die Ordnung (algebraische Multiplizität) gibt die Länge des zugehörigen Jordanblocks.

[14] Hier entwickelt man z. B. nach der dritten Spalte.

Alles hat ein

Ende

— nur die Wurst hat zwei . . .

So hat auch unsere Vorlesung nun ein Ende — bis auf den kleinen Termin am 27. Februar. ☺

Literatur

1. Martin Barner, Friedrich Flohr *Analysis I*, 5. Aufl., de Gruyter, Berlin, 2000.
2. Gerd Fischer, *Lineare Algebra*, 17. Aufl., Grundkurs Mathematik, vol. 17, Vieweg & Teubner, Wiesbaden, 2010.
3. Fischer, Kaul *Mathematik für Physiker*, Teubner, Stuttgart, 1988.
4. Wilhelm Forst, Dieter Hoffmann *Funktionentheorie erkunden mit Maple*, 2. (überarbeitete und aktualisierte) Aufl., Springer-Verlag, Heidelberg, 2012.
5. W. H. Greub, *Lineare Algebra*, Springer-Verlag, Heidelberg, 1967.
6. Paul R. Halmos, *Finite-dimensional vector spaces*, Springer-Verlag, New York, Heidelberg, 1974.
7. Jürgen Hausen, *Lineare Algebra*, Shaker-Verlag, Aachen, 2007.
8. Dieter Hoffmann, *Analysis für Wirtschaftswissenschaftler und Ingenieure*, Springer-Verlag, Heidelberg, 1995.
9. Harald Holmann, *Lineare und multilineare Algebra*, Bibliographisches Institut, Mannheim, 1970.
10. Max Koecher, *Lineare Algebra und analytische Geometrie*, 4. Aufl., Springer-Verlag, Heidelberg, 2003.
11. Serge Lang, *Linear algebra*, third ed., Undergraduate Texts in Mathematics, Springer-Verlag, New York, 1989.
12. Falko Lorenz, *Lineare Algebra. I*, 3. Aufl., Bibliographisches Institut, Mannheim, 1992.
13. Urs Stammbach, *Lineare Algebra*, Teubner Studienskripten, vol. 82, B. G. Teubner, Stuttgart, 1980.

Dieter Hoffmann, Fachbereich Mathematik und Statistik, Universität Konstanz, 78457 Konstanz, Germany
E-mail address: Dieter.Hoffmann@uni.kn