

Lineare Algebra

Hubert Grassmann

HU Berlin, Inst. f. Mathematik

1. September 2011

$$\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & 1 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & 0 & 1 & 0 & 0 \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & 0 & 0 & 1 & 0 \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} & 0 & 0 & 0 & 1 \end{pmatrix}$$

↓

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 16 & -120 & 240 & -140 \\ 0 & 1 & 0 & 0 & -120 & 1200 & -2700 & 1680 \\ 0 & 0 & 1 & 0 & 240 & -2700 & 6480 & -4200 \\ 0 & 0 & 0 & 1 & -140 & 1680 & -4200 & 2800 \end{pmatrix}$$

Inhaltsverzeichnis

-1	Vorwort	5
0	Einführung	7
0.1	Mengen	7
0.2	Abbildungen	10
0.3	Äquivalenzrelationen	10
0.4	Zahlen	11
1	Lineare Gleichungssysteme	13
1.1	Grundlagen	13
1.2	Eigenschaften von Gleichungssystemen	15
1.3	Elementare Operationen	16
1.4	Gaußscher Algorithmus	19
1.5	Computerlösungen	22
2	Grundbegriffe der Theorie der Vektorräume	25
2.1	Vektorräume, Unterräume, lineare Hüllen	25
2.2	Lineare Unabhängigkeit, Basen, Dimension	28
2.3	Anwendung auf lineare Gleichungssysteme	35
3	Lineare Abbildungen und Matrizen	39
3.1	Grundlegende Eigenschaften	39
3.2	Darstellungsmatrizen	44
3.3	Matrixmultiplikation, Inverse von Matrizen	46
3.4	Basiswechsel	51
3.5	Idempotente Abbildungen und direkte Summen	54
4	Affine Geometrie	57
4.1	Affine Räume und Unterräume	57
4.2	Affine Abbildungen	62
4.3	Zweidimensionale Geometrie I	66
5	Linearformen	69

6	Bilinearformen	73
6.1	Darstellungsmatrizen und Basiswechsel, Diagonalisierung	73
6.2	Jacobi-Diagonalisierung	78
6.3	Strassens schnelle Matrixmultiplikation	80
6.4	Klassifikation der Quadriken	82
6.5	Bilinearformen in der Analysis	86
7	Determinanten	89
7.1	Existenz und Eindeutigkeit	89
7.2	Eigenschaften und Anwendungen	93
7.3	Zweidimensionale Geometrie II	100
7.4	Abgeschnittene Pyramiden	102
8	Eigenwerte und Eigenvektoren	105
9	Komplexe Zahlen, Quaternionen usw.	115
10	Grundlegende algebraische Strukturen	123
10.1	Der Ring \mathbb{Z} der ganzen Zahlen	123
10.2	Lineare diophantische Gleichungen	127
10.3	Gruppen, Untergruppen, Homomorphismen	127
10.4	Die symmetrischen Gruppen	136
10.5	Gruppenoperationen	138
10.6	Endlich erzeugte abelsche Gruppen	142
10.7	Lineare Codes	148
10.8	Ringe und Moduln	152
10.9	Polynome	158
10.10	Gleichungen dritten und vierten Grades	163
	Index	166

Kapitel -1

Vorwort

Dies ist eine Ausarbeitung einer Anfängervorlesung zur linearen Algebra, die ich seit 1985 mehrfach an der Humboldt-Universität gehalten habe.

Ich habe Ende der 60er Jahre an der Humboldt-Universität Mathematik studiert und war danach lange Zeit als Assistent beschäftigt. Dadurch hatte ich das Glück, eine ganze Reihe von Berliner Algebraikern bei ihren Vorlesungen zur linearen Algebra beobachten zu können, wenn ich als Übungsleiter in den entsprechenden Übungen eingesetzt war. Ich konnte so bei ganz verschiedenartigen Lesenden Erfahrungen sammeln und gleichzeitig in der Arbeit mit den Studenten feststellen, welche Art und Weise der Anordnung des Stoffs und seiner Darstellung es den Studenten leichter oder schwerer macht, sich die notwendigen Kenntnisse anzueignen.

In der linearen Algebra gibt es zunächst drei Schwerpunkte, die zu bedienen sind:

- lineare Gleichungssysteme,
- Vektorräume und lineare Abbildungen,
- analytische Geometrie.

Alle drei sind gleichwertig, genauer gesagt: Jeder wesentliche Satz in einer der drei Komponenten ist auch in jeder der restlichen ausdrückbar. Es ist also schon eine Frage, von wo aus man das Knäuel aufwickeln soll.

Ein zentraler und schwieriger Begriff ist der der linearen Unabhängigkeit. Nachdem man sich diesen Begriff angeeignet hat, sind gegebene Mengen von Vektoren auf lineare Unabhängigkeit hin zu überprüfen. Dazu ist meist ein lineares Gleichungssystem zu lösen. Also ist es sicher nicht abwegig, die Theorie der linearen Gleichungssysteme an den Anfang zu stellen. Dieser Weg ist von den meisten meiner Lehrer nicht beschritten worden, ich selbst habe sogar auf Veranlassung eines dieser Herren während meines Studiums einen Beitrag zu einem Skript verbrochen, worin die Einführung in die lineare Algebra mit der Behandlung der Kategorie der Matrizen begann.

Wir beginnen also mit der Behandlung linearer Gleichungssysteme und dem Gaußschen Algorithmus (Kapitel 1). Um die Struktur der Lösungsmenge eines homogenen Gleichungssystems beschreiben zu können, werden anschließend die Grundlagen der Theorie

der Vektorräume gelegt (Kapitel 2). Die neuen Begriffe werden in die Sprache der Gleichungssysteme übertragen (Kapitel 3). Im Kapitel 4 werden lineare Abbildungen und Matrizen im Zusammenhang studiert. Im Kapitel 5 wird in die affine Geometrie eingeführt (Beschreibung von Unterräumen durch Gleichungssysteme, affine Abbildungen und ihre Matrixdarstellungen). Das kurze Kapitel 6 behandelt den Begriff des dualen Vektorraums. Im Kapitel 7 werden Bilinearformen behandelt: Matrixdarstellung, Lagrange-Diagonalisierung, Trägheitssatz. Ferner wird die Jacobi-Diagonalisierung und Strassens schnelle Matrixmultiplikation eingeführt, als Anwendung der Diagonalisierungssätze werden Quadriken klassifiziert. Die Einführung des Begriffs der Determinante (Kapitel 8) folgt der Weierstraßschen Definition, der Laplacesche Entwicklungssatz beweist die Existenz und die „Leibnizsche Definition“ die Einzigkeit der Determinantenfunktion. Das Kapitel 9 führt über die Quaternionen zum Skalar- und Vektorprodukt. Im Kapitel 10 werden Eigenwerte und -vektoren von Matrizen behandelt. Zum Ende des ersten Semesters werden „zur Erholung“ Polynome behandelt (Kapitel 11): größter gemeinsamer Teiler, Newtonsche Formeln für symmetrische Polynome und als Anwendung eine Rekursionsformel zur Berechnung der Koeffizienten des charakteristischen Polynoms einer Matrix.

Der Beginn des zweiten Semesters wird durch eine Folge von langen Beweisen geprägt, als deren Ergebnis die Jordansche Normalform erscheint (Kapitel 12). Zu bemerken ist, daß konsequent auf den Begriff des Faktorraums verzichtet wird, der in der Vektorraumtheorie ja eigentlich auch überflüssig ist. Als Anwendung werden rekursive Folgen behandelt. Es folgt ein umfangreiches Kapitel 13 über Euklidische Vektorräume. Hier wird neben dem Üblichen auf einige für numerische Anwendungen relevante Verfahren eingegangen. Kapitel 14 behandelt einige Fragen der Euklidischen Geometrie und führt in die projektive Geometrie ein. Danach werden Polynommatrizen und deren Normalformen behandelt, ein Thema, das nicht zum Standardumfang der linearen Algebra gehört, aber einen abrundenden Rückblick gestattet (Kapitel 15).

Kapitel 0

Einführung

0.1 Mengen

¹ Bei jedem mathematischen Teilgebiet steht (mehr oder weniger deutlich) am Beginn der Begriff der „Menge“: Man betrachtet Mengen von „Elementen“, wobei diese Elemente irgendetwas sein können. Als erste Beispiele seien nur genannt: Die Menge

- \mathbb{N} der natürlichen Zahlen, d.h. der positiven ganzen Zahlen,
- \mathbb{Z} der ganzen Zahlen,
- \mathbb{Q} der rationalen Zahlen,
- \mathbb{R} der reellen Zahlen,
- der stetigen Funktionen im offenen Intervall von 0 bis 1,
- der in diesem Augenblick lebenden Menschen,
- der Atome des Universums, usw.

Für viele mathematische Untersuchungen genügt es, wenn man die Mengen „naiv“ auffaßt: Danach ist eine Menge eine Zusammenfassung von wohl unterscheidbaren Objekten unseres Denkens oder unserer Anschauung zu einem Ganzen. Darin soll eingeschlossen sein, daß von jedem „Ding“ an sich feststeht, ob es zur Menge gehört oder nicht. Man verwendet die folgenden Bezeichnungen und Redeweisen:

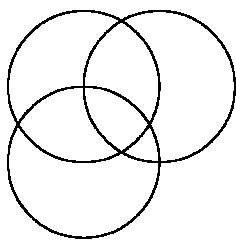
¹Diesen Abschnitt habe ich dem Skript „Lineare Algebra“ von Max Koecher, München 1968 entnommen

Bezeichnung:	Bedeutung:	Redeweise
$x \in A$	x ist Element von A	x aus A
$x \notin A$	x ist keine Element von A	x nicht aus A
$B \subset A$	für jedes $x \in B$ gilt $x \in A$	B Teilmenge von A
$B = A$	$B \subset A$ und $A \subset B$	B gleich A
$B \subsetneq A$	$B \subset A$ und $B \neq A$	B echte Teilmenge von A
$A = \{a, b, \dots\}$	A besteht aus den nicht notwendig verschiedenen Elementen a, b, \dots	die Menge der a, b, \dots
$\{x \mid x \in A, E(x)\}$	Die Menge der $x \in A$, welche die Eigenschaft $E(x)$ haben	
\emptyset	die Menge, die kein Element enthält	leere Menge
$A \cup B$	$\{x \mid x \in A \text{ oder } x \in B\}$	Vereinigung von A und B
$A \cap B$	$\{x \mid x \in A \text{ und } x \in B\}$	Durchschnitt von A und B
$A \cap B = \emptyset$		A und B sind disjunkt
$\bigcup(A_i \mid i \in I) = \bigcup_{i \in I} A_i$	$\{x \mid \text{es gibt } i \in I \text{ mit } x \in A_i\}$	Vereinigung der A_i
$\bigcap(A_i \mid i \in I) = \bigcap_{i \in I} A_i$	$\{x \mid \text{für alle } i \in I \text{ gilt } x \in A_i\}$	Durchschnitt der A_i
$P(A)$	$\{B \mid B \subset A\}$	Potenzmenge von A

Rechenregeln:

$$\begin{aligned}
 A \subset A, \quad \emptyset \subset A, \quad x \in A &\Leftrightarrow \{x\} \subset A, \\
 \text{aus } A \subset B \text{ und } B \subset C \text{ folgt } A &\subset C, \\
 A \cup A = A, \quad A \cap A = A, \quad A \cap B &\subset A \subset A \cup B, \\
 A \cup \emptyset = A, \quad A \cap \emptyset = \emptyset, \\
 A \cup (B \cup C) = (A \cup B) \cup C = A \cup B \cup C, \\
 A \cap (B \cap C) = (A \cap B) \cap C = A \cap B \cap C, \\
 A \cup (B \cap C) = (A \cup B) \cap (A \cup C), \\
 A \cap (B \cup C) = (A \cap B) \cup (A \cap C).
 \end{aligned}$$

Die beiden letzten Regeln kann man sich anschaulich klarmachen: Färben Sie die entsprechenden Bereiche!



Sie bedürfen jedoch eines Beweises, der z.B. wie folgt abläuft:

Behauptung: $A \cup (B \cap C) \subset (A \cup B) \cap (A \cup C)$

Beweis: $x \in A \cup (B \cap C) \Rightarrow (x \in A) \text{ oder } (x \in B \text{ und } x \in C) \Rightarrow x \in A \cup B \text{ und } x \in A \cup C \Rightarrow x \in (A \cup B) \cap (A \cup C)$. \square

Behauptung: $(A \cup B) \cap (A \cup C) \subset A \cup (B \cap C)$

Beweis: $x \in (A \cup B) \cap (A \cup C) \Rightarrow x \in A \cup B \text{ und } x \in A \cup C \Rightarrow (x \in A \text{ oder } x \in B) \text{ und } (x \in A \text{ oder } x \in C) \Rightarrow (x \in A \text{ oder } (x \in B \text{ und } x \in C)) \Rightarrow x \in A \cup (B \cap C)$. \square

Die andere Regel beweist man durch „Dualisierung“: Man vertauscht \cap und \cup und „und“ und „oder“.

Sind A_1, \dots, A_n endlich viele Mengen, dann bildet man die (geordneten) n -Tupel

$$(a_1, \dots, a_n) \text{ mit } a_i \in A_i \text{ für } i = 1, \dots, n$$

und definiert $(a_1, \dots, a_n) = (b_1, \dots, b_n)$ dann und nur dann, wenn $a_i = b_i$ für $i = 1, \dots, n$. Die Menge dieser n -Tupel nennt man das direkte Produkt der Mengen A_1, \dots, A_n und schreibt

$$A_1 \times A_2 \times \dots \times A_n = \{(a_1, \dots, a_n) \mid a_i \in A_i\}.$$

Rechenregeln:

$$\begin{aligned}
 A \times B = \emptyset &\Rightarrow A = \emptyset \text{ oder } B = \emptyset, \quad (A \times C) \cup (B \times C) = (A \cup B) \times C, \\
 (A \times C) \cap (B \times D) &= (A \cap B) \times (C \cap D).
 \end{aligned}$$

Man beachte, das nicht alles, was hingeschrieben werden kann, eine Menge ist. Z.B. hat es keinen Sinn, von der „Menge aller Mengen“ zu sprechen. In der Mathematik kommt

man mit den folgenden „mengenerzeugenden“ Prinzipien aus: $\{a\}$ ist eine Menge; wenn A eine Menge ist und $B \subset A$, so ist B eine Menge; wenn A und B Mengen sind, so ist $A \times B$ eine Menge; wenn $(A_i \mid i \in I)$ eine Familie von Mengen ist, so ist $\bigcup_{i \in I} A_i$ eine Menge; wenn A eine Menge ist, so ist $P(A)$ eine Menge; \mathbb{Z} ist eine Menge.

0.2 Abbildungen

X und Y seien zwei nichtleere Mengen, $F \subset X \times Y$ heißt Graph von X in Y , wenn es zu jedem $x \in X$ genau ein $y \in Y$ mit $(x, y) \in F$ gibt. Man definiert dann $f(x) = y$, falls $(x, y) \in F$ ist und nennt f die zum Graphen F gehörige Abbildung (oder Funktion) auf X mit Werten in Y . Hierfür werden die folgenden Schreibweisen verwendet:

$$f : X \longrightarrow Y; \quad x \mapsto f(x)$$

Man nennt X die Quelle (den Definitionsbereich) und Y das Ziel (den Wertevorrat) von f . Für $A \subset X, B \subset Y$ definiert man $f(A) = \{f(x) \mid x \in A\}$ als Bild von A und $f^{-1}(B) = \{x \mid x \in X \text{ und } f(x) \in B\} \subset X$ als Urbild von B . Ferner nennt man die Abbildung $f : X \longrightarrow Y$ surjektiv, wenn $f(X) = Y$, injektiv, wenn $f(x) = f(y)$ nur für $x = y$ gilt, und bijektiv, wenn f surjektiv und injektiv ist.

Die Abbildung $f : X \longrightarrow Y$ induziert auf jeder Teilmenge A von X eine Abbildung $f \mid A : A \longrightarrow Y$ vermöge $(f \mid A)(x) = f(x)$ für $x \in A$, sie heißt die Einschränkung von f auf A .

Die identische Abbildung $x \mapsto x$ von X auf sich wird mit id_X bezeichnet.

Für zwei Abbildungen $f : X \longrightarrow Y, g : Y \longrightarrow Z$ kann man die komponierte Abbildung $g \circ f : X \longrightarrow Z, x \mapsto g(f(x))$ erklären.

Es gilt $h \circ (g \circ f) = (h \circ g) \circ f$.

Ist $f : X \longrightarrow Y$ bijektiv, so existiert die Umkehrabbildung $f^{-1} : Y \longrightarrow X$, die durch $f^{-1}(y) = x$ definiert ist, falls $y = f(x)$ ist. Dann ist f^{-1} ebenfalls bijektiv und es gilt $f \circ f^{-1} = id_Y$ und $f^{-1} \circ f = id_X$.

Zwei Mengen heißen gleichmächtig, wenn eine bijektive Abbildung zwischen ihnen gibt. Eine Menge heißt endlich (bzw. unendlich), wenn sie nur endlich viele (bzw. unendlich viele) Elemente besitzt. Eine zur Menge \mathbb{N} gleichmächtige Menge X heißt abzählbar; ihre Elemente kann man durch die natürlichen Zahlen indizieren: $X = \{a_i \mid i = 1, 2, \dots\}$.

0.3 Äquivalenzrelationen

Sei A eine Menge. Eine Menge $R \subset A \times A$ heißt Äquivalenzrelation auf A , wenn gilt

1. Für $a \in A$ gilt $(a, a) \in R$ (Reflexivität).
2. Aus $(a, b) \in R$ folgt $(b, a) \in R$ (Symmetrie).
3. Aus $(a, b) \in R, (b, c) \in R$ folgt $(a, c) \in R$ (Transitivität).

Für $(a, b) \in R$ schreibt man auch $a \sim b$, wenn klar ist, um welche Relation es sich handelt.

Die Menge

$$[a]_R = \{b \in A \mid b \sim a\}$$

heißt die Äquivalenzklasse von a . Es gilt

$$A = \bigcup ([a]_R \mid a \in A)$$

und es ist $[a]_R \cap [b]_R \neq \emptyset$ genau dann, wenn $a \sim b$, also wenn die Äquivalenzklassen übereinstimmen.

Die durch

$$A/R = \{[a]_R \mid a \in A\}$$

definierte Teilmenge von $P(A)$ heißt die Faktormenge von A nach R . Offenbar ist die „kanonische“ Abbildung $p: A \rightarrow A/R, a \mapsto [a]_R$ surjektiv.

0.4 Zahlen

Das Lösen von Gleichungen ist eine grundlegende mathematische Aufgabenstellung. Eine Gleichung kann man in der Form $AX = B$ schreiben, dabei seien A und B gegeben und X gesucht. In jedem konkreten Sachverhalt muß man sich aber darüber im klaren sein, was A, B, X für Objekte sein sollen, wie das Zeichen „=“ zu interpretieren ist und wie aus A und X das Objekt AX entstehen soll. Wir werden sehen, daß sich sehr allgemeine Gleichungen in der beschriebenen Weise darstellen lassen, wenn diese Interpretation in geeigneter Weise gewählt wird.

Beispiele für Gleichungen sind:

$$3x = 9; \quad x^2 + ax + b = 0; \quad x_1 + 2x_2 = 5; \quad \sin(x) = 0,5.$$

Meist kommen in Gleichungen Zahlenkoeffizienten vor und die Unbekannten sind Zahlen aus einem bestimmten Zahlbereich. Sie kennen aus der Schule die folgenden Zahlbereiche:

\mathbb{N} , die Menge der natürlichen Zahlen,

\mathbb{Z} , die Menge der ganzen Zahlen,

\mathbb{Q} , die Menge der rationalen Zahlen,

\mathbb{R} , die Menge der reellen Zahlen und

\mathbb{C} , die Menge der komplexen Zahlen.

Die letzten drei dieser Bereiche haben gemeinsame Eigenschaften, die man in den folgenden Axiomen zusammenfaßt:

Definition: Eine Menge R heißt Körper, wenn zu je zwei Elementen $r, s \in R$ eine „Summe“ $r + s$ und ein „Produkt“ rs gegeben ist (dies sollen wieder Elemente aus R sein), so daß folgendes gilt:

1. $(r + b) + c = r + (b + c)$,
(Assoziativgesetz der Addition)

2. es gibt ein Element 0 mit $r + 0 = r$ für alle r ,
(Existenz eines neutralen Elements)
3. zu jedem $r \in R$ gibt es ein r' mit $r + r' = 0$,
(Existenz eines zu r inversen Elements, man schreibt für r' gewöhnlich $-r$)
4. $r + s = s + r$ für alle r, s
(Kommutativgesetz der Addition)
(Wenn die Eigenschaften 1...4 erfüllt sind, sagt man: R bildet bezüglich der Addition eine kommutative Gruppe.)
5. $(rs)t = r(st)$
(Assoziativgesetz der Multiplikation)
6. $(r + s)t = rt + st$
(Distributivgesetz)
7. $rs = sr$
(Kommutativgesetz der Multiplikation)
8. es gibt ein Element 1 in R mit $1r = r$ für alle r ,
(Existenz eines neutralen Elements)

(Wenn die Eigenschaften 1...7 erfüllt sind, so sagt man: R ist ein kommutativer Ring mit Einselement.)
9. zu jedem $r \neq 0$ aus R gibt es ein r'' mit $rr'' = 1$.
(Existenz eines zu r inversen Elements; man schreibt für r'' gewöhnlich r^{-1}).

Ohne dafür Beweise anzugeben, werden wir im folgenden stets benutzen, daß \mathbb{Z} ein Ring ist und daß $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ Körper sind. Wir werden Körperelemente kurz als „Zahlen“ bezeichnen.

Im folgenden werden wir stets einen fixierten Zahlbereich R zugrundelegen; Sie können ohne weiteres annehmen, daß dies der Körper \mathbb{R} der reellen Zahlen ist.

Wir werden einige Abkürzungen verwenden, die hier aufgezählt werden sollen:

Mit $f : A \rightarrow B$ bezeichnen wir eine Abbildung f einer Menge A in eine Menge B , und wenn $C \subseteq A$ eine Teilmenge ist, so bezeichnet $f|_C$ die Einschränkung der Abbildung f auf die Teilmenge C .

Das Ende eines Beweises wird so angezeigt: □

Kapitel 1

Lineare Gleichungssysteme

1.1 Grundlagen

Lineare Gleichungssysteme sind Ihnen aus der Schule bekannt. Wir betrachten ein Beispiel: Das folgende Gleichungssystem sei gegeben:

$$\begin{aligned}ax + by &= c \\dx + ey &= f\end{aligned}$$

(a, \dots, f sind gegebene Zahlen, x, y sind gesucht).

Als Lösung dieses Gleichungssystems bezeichnen wir jedes Paar (x, y) von Zahlen, das beide Gleichungen erfüllt. Wir nehmen an, daß eine Lösung existiert und nehmen mit den vier Zahlen $ax + by$, c , $dx + ey$, f , von denen je zwei gleich sind, folgende Umformungen vor: Wir multiplizieren die Zahlen der ersten Zeile mit e , die der zweiten mit b , subtrahieren beides und erhalten:

$$\begin{aligned}eax + eby &= ec \\bdx + eby &= bf\end{aligned}$$

und

$$eax - bdx = ec - bf,$$

also ist, falls $ea - bd \neq 0$ ist,

$$x = \frac{ec - bf}{ea - bd} \quad y = \frac{af - cd}{ea - bd}.$$

Wir machen noch die Probe:

$$(aec - abf + baf - bcd) : (ea - bd) = c$$

(usw.) Hier haben wir also eine eindeutig bestimmte Lösung gefunden.

Im folgenden werden wir versuchen, beliebige lineare Gleichungssysteme zu lösen.

Eine Lösung eines Gleichungssystems ist ein Paar, ein Tripel, ... von Zahlen (je nach dem, wieviele Unbekannte das System hat).

Definition: Für $i = 1, \dots, m$ und $j = 1, \dots, n$ seien Zahlen a_{ij} und b_i gegeben, dann nennt man die folgenden Bedingungen

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ &\dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned}$$

ein lineares Gleichungssystem mit m Gleichungen und den n Unbekannten x_1, \dots, x_n .

Die Menge aller n -tupel

$$x = (x_1, \dots, x_n)$$

bezeichnen wir mit R^n . Ein n -tupel (x_1, \dots, x_n) , dessen Komponenten die Gleichungen erfüllen, heißt eine Lösung des Systems S ; die Menge aller Lösungen von S bezeichnen wir mit $LM(S)$. Ein Gleichungssystem, wo alle b_i gleich Null sind, heißt homogen, wenn dies nicht der Fall ist, heißt es inhomogen. Zum gegebenen (inhomogenen) Gleichungssystem

$$\sum_{j=1}^n a_{ij}x_j = b_i, i = 1, \dots, m \quad (S)$$

nennen wir das homogene Gleichungssystem

$$\sum_{j=1}^n a_{ij}x_j = 0, i = 1, \dots, m \quad (H)$$

das zu S gehörige homogene System.

Bemerkung zur Verwendung des Summenzeichens:

Aus schreibtechnischen Gründen werden wir oft auf die Angabe des Summationsindex und seiner Grenzen verzichten. Meist ist aus dem Zusammenhang klar, welche Werte dieser Index zu durchlaufen hat. Außerdem ist der Summationsindex von anderen Indizes leicht zu unterscheiden: er tritt in dem dem \sum -Symbol folgenden Term doppelt auf!

1.2 Eigenschaften homogener und inhomogener Gleichungssysteme

Wir führen zunächst in R^n die folgenden Operationen ein: Seien x und y n -tupel und r eine Zahl, dann setzen wir

$$x + y = (x_1 + y_1, \dots, x_n + y_n)$$

und

$$rx = (rx_1, \dots, rx_n).$$

Sei

$$\sum a_{ij}x_j = 0, i = 1, \dots, m \quad (H)$$

ein homogenes Gleichungssystem. Dann gilt:

1. Es existiert stets eine Lösung von H , nämlich die triviale Lösung $(0, \dots, 0)$.
2. Wenn $x = (x_1, \dots, x_n)$ eine Lösung von H und r eine Zahl ist, so ist auch das Vielfache $rx = (rx_1, \dots, rx_n)$ eine Lösung von H .
3. Wenn $x = (x_1, \dots, x_n)$ und $y = (y_1, \dots, y_n)$ Lösungen von H sind, so ist auch die Summe $x + y = (x_1 + y_1, \dots, x_n + y_n)$ eine Lösung von H .

Wenn $x, y, z, \dots \in R^n$ und r, s, t, \dots Zahlen sind, so nennen wir das n -tupel $rx + sy + tz + \dots$ eine Linearkombination von x, y, z, \dots

Dann erhalten wir sofort

4. Jede Linearkombination von Lösungen des Systems H ist eine Lösung von H .

Sei nun wieder

$$\sum a_{ij}x_j = b_i, i = 1, \dots, m \quad (S)$$

ein inhomogenes System und

$$\sum a_{ij}x_j = 0, i = 1, \dots, m \quad (H)$$

das zugehörige homogene System.

5. Wenn y eine Lösung von S und x eine Lösung von H ist, so ist $x + y$ eine Lösung von S .

Beweis: $\sum a_{ij}(x_j + y_j) = \sum a_{ij}x_j + \sum a_{ij}y_j = b_i + 0$.

6. Sei y eine Lösung von S ; dann hat jede Lösung von S die Form $y + x$, wo x eine geeignete Lösung von H ist.

Beweis: Seien y und y' Lösungen von S , d.h. es gilt

$$\sum a_{ij}y_j = b_i, i = 1, \dots, m$$

und

$$\sum a_{ij}y'_j = b_i, i = 1, \dots, m.$$

Durch Subtraktion dieser Zahlen erhalten wir

$$\sum a_{ij}(y'_j - y_j) = 0, i = 1, \dots, m,$$

d.h. das n -tupel $x = y' - y$ ist eine Lösung von H und es gilt $y' = y + x$. \square

In Bezug auf lineare Gleichungssysteme werden wir die folgenden drei Fragen behandeln:

1. Wann existieren Lösungen ?
2. Wie kann man alle Lösungen berechnen ?
3. Welche Struktur hat die Lösungsmenge ?

1.3 Elementare Operationen

Wir werden nun Operationen mit den Gleichungen eines gegebenen Gleichungssystems einführen, die uns bei der Bestimmung der Lösungsmenge nützlich sein werden.

Sei das folgende lineare Gleichungssystem gegeben:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ &\dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned}$$

Typ 1. Sei $c \neq 0$ eine Zahl, $1 \leq k \leq m$, dann sei S_1 das folgende System:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ &\dots \\ ca_{k1}x_1 + ca_{k2}x_2 + \dots + ca_{kn}x_n &= cb_k \\ &\dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned}$$

(Die k -te Gleichung wird mit c multipliziert.)

Typ 2. Sei $1 \leq i, k \leq m$; dann sei S_2 das folgende System:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ &\dots \\ (a_{i1} + a_{k1})x_1 + (a_{i2} + a_{k2})x_2 + \dots + (a_{in} + a_{kn})x_n &= b_i + b_k \\ &\dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned}$$

(Die i -te Gleichung wird zur k -ten addiert.)

Typ 3. Sei $1 \leq i, k \leq m$; dann sei S_3 das folgende System:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ &\dots \\ a_{k1}x_1 + a_{k2}x_2 + \dots + a_{kn}x_n &= b_k \\ &\dots \\ a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n &= b_i \\ &\dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned}$$

(Die i -te und k -te Gleichung werden vertauscht.)

Dann gilt der folgende

Satz 1.3.1 Die Operationen vom Typ 1, 2, 3 verändern die Lösungsmenge des Gleichungssystems nicht, d.h. es gilt

$$LM(S) = LM(S_1) = LM(S_2) = LM(S_3).$$

Beweis: 1. Sei $x = (x_1, \dots, x_n) \in LM(S)$, dann gilt

$$\sum a_{ij}x_j = b_i \text{ für } i = 1, \dots, m.$$

Wir betrachten die k -te Gleichung:

$$\sum a_{kj}x_j = b_k.$$

Dann ist auch

$$c \sum a_{kj}x_j = cb_k,$$

die anderen Gleichungen sind auch erfüllt, also ist $x \in LM(S_1)$.

Folglich ist $LM(S)$ bei beliebigen Operationen vom Typ 1 in $LM(S_1)$ enthalten; umgekehrt läßt sich S_1 durch eine Operation vom Typ 1 (nämlich durch Multiplikation der k -ten Gleichung mit $\frac{1}{c}$) in S überführen, also müssen beide Lösungsmengen gleich sein.

2. Sei wieder $x = (x_1, \dots, x_n) \in LM(S)$, also

$$\sum a_{ij}x_j = b_i \text{ für } i = 1, \dots, m.$$

Wir betrachten die i -te und die k -te Gleichung:

$$\sum a_{ij}x_j = b_i$$

$$\sum a_{kj}x_j = b_k.$$

Dann ist

$$\sum a_{ij}x_j + \sum a_{kj}x_j = b_i + b_k = \sum (a_{ij} + a_{kj})x_j$$

also $x \in LM(S_2)$ für beliebige Operationen vom Typ 2.

Umgekehrt läßt sich S_2 durch Operationen der Typen 1 und 2 wieder in S überführen, also stimmen beide Lösungsmengen überein.

3. Eine Operation vom Typ 3 läßt sich aus Operationen der Typen 1 und 2 zusammensetzen, jedesmal bleibt die Lösungsmenge ungeändert. \square

Folgerung 1.3.1 Sei $c \neq 0 \in R, i, j \leq m$; wenn das c -fache der i -ten Gleichung von S zur k -ten Gleichung addiert wird, so ändert sich die Lösungsmenge nicht. \square

Mit diesen elementaren Operationen können wir Gleichungssysteme in eine übersichtliche Form bringen, wo die Lösungsmenge leicht abzulesen ist.

Es erhebt sich die Platzfrage: Wie schreibt man ein Gleichungssystem rationell auf? Zum Beispiel:

$$\begin{aligned}x_1 - 2x_2 - 3x_3 &= 4 \\-4x_1 + x_2 - 2x_3 &= 5 \\-3x_1 + 5x_2 + x_3 &= 6.\end{aligned}$$

Alle Information steckt im folgenden Schema (einer sogenannten Matrix):

$$\begin{pmatrix} 1 & -2 & -3 & 4 \\ -4 & 1 & -2 & 5 \\ -3 & 5 & 1 & 6 \end{pmatrix}$$

Wir streben an, die Matrix durch elementare Operationen mit ihren Zeilen, die den obigen Operationen mit Gleichungen entsprechen, in die Form

$$\begin{pmatrix} 1 & 0 & 0 & a \\ 0 & 1 & 0 & b \\ 0 & 0 & 1 & c \end{pmatrix},$$

in eine „reduzierte Form“ zu überführen; dem entspricht dann das Gleichungssystem

$$\begin{aligned}x_1 &= a \\x_2 &= b \\x_3 &= c,\end{aligned}$$

dessen Lösungsmenge man sofort ablesen kann (das wird nicht in jedem Fall möglich sein). Überlegen Sie sich, welche Operationen bei der folgenden Rechnung angewandt wurden:

$$\begin{pmatrix} 1 & -2 & -3 & 4 \\ 0 & -7 & -14 & 21 \\ 0 & -1 & -8 & 18 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -2 & -3 & 4 \\ 0 & 1 & 2 & -3 \\ 0 & 0 & -6 & 15 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 1 & -2 \\ 0 & 1 & 2 & -3 \\ 0 & 0 & 1 & -\frac{5}{2} \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & \frac{1}{2} \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & -\frac{5}{2} \end{pmatrix}$$

also erhalten wir die einzige Lösung $(\frac{1}{2}, 2, -\frac{5}{2})$.

1.4 Gaußscher Algorithmus

Wir wollen dieses Verfahren nun für ein beliebiges Gleichungssystem durchführen; das folgende Verfahren wird als Gaußscher Algorithmus bezeichnet.

Sei also ein Gleichungssystem

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ &\dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned}$$

gegeben, dazu gehört die Matrix

$$\begin{pmatrix} a_{11} & \dots & a_{1n} & b_1 \\ & \dots & & \\ a_{i1} & \dots & a_{in} & b_i \\ & \dots & & \\ a_{m1} & \dots & a_{mn} & b_m \end{pmatrix}.$$

Wir setzen zuerst $k = 1$ (wir beginnen mit der ersten Zeile). Wir suchen den kleinsten Spaltenindex $j \geq k$, so daß die j -te Spalte ein von Null verschiedenes Element a_{ij} , $i \geq k$ enthält, und bringen die i -te Zeile durch Zeilenvertauschung in die k -te Zeile (falls nicht schon $a_{kj} \neq 0$ war). Nun multiplizieren wir die k -te Zeile mit $(a_{kj})^{-1}$, dann steht an der Stelle (k, j) eine 1. Unter- und überhalb der 1 werden in der j -ten Spalte Nullen erzeugt, indem wir das a_{ij} -fache der k -ten Zeile von der i -ten subtrahieren ($1 \leq i < k, k < i \leq m$).

Schließlich erhöhen wir, falls $k < m$ ist, den Index k um 1 und beginnen von vorn, bis wir keine von Null verschiedene Zahl mehr finden können. Die entstandenen Spalten, die eine 1 in der 1., 2., ... Zeile und sonst nur Nullen enthalten, heißen **ausgezeichnete** Spalten.

Als Ergebnis erhalten wir eine Matrix, die im allgemeinen folgende Gestalt haben kann (in konkreten Fällen werden einige [nichtausgezeichnete] Spalten fehlen; die ausgezeichneten Spalten haben die Nummern $k_1 \dots k_r$):

$$\begin{pmatrix} 0 & \dots & 1 & a_{1,k_1+1} & \dots & a_{1,k_2-1} & 0 & a_{1,k_2+1} & \dots & a_{1,k_r-1} & 0 & \dots & b_1 \\ 0 & \dots & 0 & & \dots & 0 & 1 & a_{2,k_2+1} & \dots & a_{2,k_r-1} & 0 & a_{2,k_r+1} & \dots & b_2 \\ & \dots & & & & & & & & & & & & \\ 0 & \dots & 0 & & \dots & & & & & & 1 & a_{r,k_r+1} & \dots & b_r \\ & \dots & & & & & & & & & & & & \\ 0 & & & & & & \dots & & & & & & 0 & b_{r+1} \\ & \dots & & & & & & & & & & & & \\ 0 & & & & & & \dots & & & & & & 0 & b_m \end{pmatrix}$$

Das dieser Matrix entsprechende Gleichungssystem, das dieselbe Lösungsmenge wie das gegebene besitzt, hat dann die folgende Gestalt S' :

Ein formaler Algorithmus

Eine Basis des Lösungsraums eines homogenen Gleichungssystems läßt sich leicht ermitteln, wenn die „Anfangseinsen“ in der reduzierten Zeilenstufenform „am Anfang“ stehen. Der allgemeine Fall läßt sich wie folgt erledigen:

1. Fülle die reduzierte Form so durch Nullzeilen auf, daß die Anfangseinsen in der Diagonalen stehen.

2. Subtrahiere die „Einheitsmatrix“
$$\begin{pmatrix} 1 & 0 & \cdots & & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ & & \cdots & & \\ 0 & \cdots & 1 & 0 & \\ 0 & \cdots & 0 & 1 & \end{pmatrix}.$$

3. Die von Null verschiedenen Spalten bilden eine „Basis“ des Lösungsraums.

Wir führen das am obigen Beispiel durch:

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

Abschließend beweisen wir den folgenden

Satz 1.4.1 Sei $\sum a_{ij}x_j = 0, i = 1, \dots, m$, ein homogenes Gleichungssystem, für das $n > m$ gilt (es gibt also mehr Unbekannte als Gleichungen) dann existiert eine Lösung $(x_1, \dots, x_n) \neq (0, \dots, 0)$.

Beweis: Die reduzierte Form der Koeffizientenmatrix sieht etwa folgendermaßen aus:

$$\begin{pmatrix} 1 & & & 0 \\ & 1 & & 0 \\ & & \cdots & 0 \\ & & & 1 & 0 \end{pmatrix}$$

Sie habe m Zeilen und n Spalten, davon r ausgezeichnete. (Wir haben nur die ausgezeichneten Spalten angedeutet.)

Da die Einsen der ausgezeichneten Spalten in verschiedenen Zeilen stehen, sind es derer höchstens m , also weniger als n , es gibt also mindestens eine nichtausgezeichnete Unbekannte, deren Wert von Null verschieden gewählt werden kann. \square

1.5 Computerlösungen

Man kann einen Computer zur Lösung linearer Gleichungssysteme nutzen. Ich habe einige Java-Programme geschrieben, die im Java-Paket `HUMath.Algebra` zusammengefaßt sind, deren Dokumentation findet man unter

<http://www.mathematik.hu-berlin.de/~lamour/software/JAVA/HUMath/>

Es gibt dort eine Klasse `DM.java`, wo mit Matrizen gerechnet wird, deren Komponenten Fließkommazahlen (`double`) sind, und eine Klasse `QM`, wo die Matrixkomponenten rationale Zahlen sind, dort wird also ohne Rundungsfehler gerechnet.

Wir schreiben ein kurzes Java-Programm:

```
import HUMath.Algebra.*;
public class lg
{
public static void main(String[] arg)
{
    QM a = QM.fromFile("ein");
    QM.write(a);
    QM b = QM.GAUSS(a);
    QM.write(b);
    QM e = QM.loesung(b);
    QM.write(e);
}
}
```

In der Datei `ein` steht die Koeffizientenmatrix unseres Gleichungssystems, zuerst die Zeilen-, dann die Spaltenzahl, dann die Komponenten, jeweils in einer neuen Zeile. Diese Datei wird gelesen und die Matrix ausgegeben. Dann wird der Gaußsche Algorithmus angewandt und das Ergebnis interpretiert: Die letzte Spalte ist eine spezielle Lösung des (inhomogenen) Gleichungssystems, die Spalten davor sind Lösungen des zugehörigen homogenen Gleichungssystems (später werden wir sehen, daß sie eine „Basis“ bilden). Die Ausgabe des Computers sieht so aus:

```
2 3 4 1 0
4 6 5 2 3
6 9 3 4 5
```

```
1 3/2 0 0 4
0 0 1 0 -1
0 0 0 1 -4
```

```
-3/2 4
1 0
0 -1
0 -4
```

Die allgemeine Lösung ist also

$$x = \begin{pmatrix} 4 \\ 0 \\ -1 \\ -4 \end{pmatrix} + t \begin{pmatrix} -\frac{3}{2} \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

Wenn das Programm MATLAB verwendet wird, wäre einzugeben:

```
a = [2 3 4 1 0; 4 6 5 2 3; 6 9 3 4 5]
```

```
b = [0; 3; 5]
```

```
x = a\b
```

Das Kommando `rref(a b)` ergibt die reduzierte Zeilenstufenform; mittels `rrefmovie(a)` kann man die einzelnen Schritte der Rechnung verfolgen.

Kapitel 2

Grundbegriffe der Theorie der Vektorräume

2.1 Vektorräume, Unterräume, lineare Hüllen

Sei R ein Körper. Eine Menge V heißt R -Vektorraum, wenn zu je zwei Elementen $v, w \in V$ ein Element von V existiert, das mit $v + w$ bezeichnet wird und Summe von v und w heißt, und wenn zu $v \in V$ und jeder Zahl $r \in R$ ein Element $rv \in V$ existiert (dies wird als Produkt von r und v bezeichnet), so daß für alle $u, v, w \in V$ und alle $r, s \in R$ folgende Eigenschaften erfüllt sind:

1. $(u + v) + w = u + (v + w)$
(Assoziativgesetz),
2. es gibt ein Element $o \in V$, so daß für alle $v \in V$ gilt $v + o = v$
(Existenz eines neutralen Elements),
3. zu jedem $v \in V$ gibt es ein $v' \in V$ mit $v + v' = o$
(Existenz des zu v inversen Elements)
4. $v + w = w + v$
(Kommutativgesetz),
5. $r(sv) = (rs)v$
(Assoziativgesetz),
6. $r(v + w) = rv + rw$
(1. Distributivgesetz),
7. $(r + s)v = rv + sv$
(2. Distributivgesetz),
8. $1v = v$.

Die Elemente eines Vektorraums werden Vektoren genannt. Das neutrale Element o wird der Nullvektor von V genannt, wir werden das Symbol „ o “ hierfür reservieren;

anstelle von v' schreiben wir $-v$ und anstelle von $v + (-w)$ einfach $v - w$.¹

Beispiele:

- a) $V =$ Menge der Verschiebungen der Ebene (eine Verschiebung kann man durch einen Pfeil kennzeichnen), die Summe zweier Verschiebungen ist die Nacheinanderausführung beider Verschiebungen, das Produkt einer Verschiebung mit einer reellen Zahl ist die entsprechend „verlängerte“ Verschiebung.
 b) $V = R^n =$ Menge aller n -tupel (r_1, \dots, r_n) , Addition und Multiplikation sind (wie im Kapitel 1) komponentenweise definiert.
 c) $V =$ Menge aller Lösungen des homogenen Gleichungssystems

$$\sum a_{ij}x_j = 0, \quad i = 1, \dots, m,$$

die Addition und Multiplikation sind wie in R^n definiert.

d) Sei V ein Vektorraum. Wenn $v_1, \dots, v_n \in V$ und $r_1, \dots, r_n \in R$ sind, so heißt der Vektor $r_1v_1 + \dots + r_nv_n \in V$ eine Linearkombination der Vektoren v_1, \dots, v_n .

Sei $\mathcal{L}(v_1, \dots, v_n)$ die Menge aller Linearkombinationen von v_1, \dots, v_n , also

$$\mathcal{L}(v_1, \dots, v_n) = \{v \in V \mid \text{es gibt } r_1, \dots, r_n \in R \text{ mit } v = \sum r_i v_i\}.$$

Diese Menge heißt die lineare Hülle von v_1, \dots, v_n .

Lemma 2.1.1 $\mathcal{L}(v_1, \dots, v_n)$ ist ein Vektorraum (Summe und Produkt sind wie in V definiert).

Beweis: Wir überprüfen die Axiome. Die Summe zweier Linearkombinationen von v_1, \dots, v_n ist ebenfalls eine Linearkombination von v_1, \dots, v_n

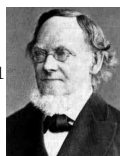
$$\sum r_i v_i + \sum s_i v_i = \sum (r_i + s_i) v_i,$$

das Vielfache einer Linearkombination von v_1, \dots, v_n ist ebenfalls eine Linearkombination von v_1, \dots, v_n :

$$r \sum r_i v_i = \sum (rr_i) v_i.$$

Der Nullvektor o ist eine Linearkombination von v_1, \dots, v_n :

$$o = \sum 0v_i$$



Hermann Günther Graßmann (1809 – 1877) Gymnasialprofessor in Stettin, Verfasser der „Linealen Ausdehnungslehre“ (1844), hier wurde die Theorie der Vektorräume dargestellt. In der 2., überarbeiteten Auflage von 1878 heißt es in der Vorrede:

Das Werk, dessen zweite Auflage ich hiermit der Öffentlichkeit übergebe, hat in den ersten 25 Jahren nach seinem ersten Erscheinen nur eine geringe und meist nur gelegentliche Beachtung gefunden. Diesen Mangel an Erfolg konnte ich nicht der behandelten Wissenschaft als solcher zur Last legen; denn ich kannte deren fundamentale Wichtigkeit, ja deren Nothwendigkeit vollkommen; sondern ich konnte die Ursache davon nur in der streng wissenschaftlichen, auf die ursprünglichen Begriffe zurückgehenden Behandlungsweise finden. Eine solche Behandlungsweise erforderte aber ein nicht bloss gelegentliches Auffassen dieser oder jener Resultate, sondern ein sich verfenken in die zu Grunde liegenden Ideen und eine zusammenhängende Auffassung des ganzen auf dies Fundament aufgeführten Baues, dessen einzelne Theile erst durch das Uebersehen des Ganzen ihr volles Verständnis erhalten konnte. ... Meine Hoffnung, einen akademischen Lehrstuhl zu gewinnen, und dadurch jüngere Kräfte in die Wissenschaft einzuführen und sie zum weiteren Ausbau derselben anzuregen, schlug fehl.

und der zu $\sum r_i v_i$ inverse Vektor auch:

$$-\sum r_i v_i = \sum (-r_i) v_i.$$

Die Gültigkeit der Axiome 1,4,...,8 versteht sich von selbst, da dies ja für alle Elemente von V gilt. \square

Definition: Sei V ein Vektorraum und $U \neq \emptyset$ eine Teilmenge von V , so daß mit $u, u' \in U$ und $r \in R$ auch $u + u'$ sowie ru Elemente von U sind. Dann heißt U ein Unterraum von V .

Also haben wir gezeigt, daß $\mathcal{L}(v_1, \dots, v_n)$ ein Unterraum von V ist.

Allgemeiner: Sei V ein Vektorraum und M eine (nicht notwendigerweise endliche) Teilmenge von V , dann setzen wir

$$\mathcal{L}(M) = \{v \in V \mid \text{es gibt } v_1, \dots, v_n \in M \text{ und } r_1, \dots, r_n \in R \text{ mit } v = r_1 v_1 + \dots + r_n v_n\}.$$

$\mathcal{L}(M)$ heißt die Menge der Linearkombinationen über M . Es ist wieder klar, daß $\mathcal{L}(M)$ ein Unterraum von V ist. Wir sagen, daß M den Unterraum $\mathcal{L}(M)$ erzeugt.

Satz 2.1.1 Sei V ein Vektorraum und $M \subseteq V$ eine Teilmenge. Dann ist $\mathcal{L}(M)$ der kleinste Unterraum von V , der M enthält, d.h. wenn U ein Unterraum von V ist, der M enthält, so ist $\mathcal{L}(M)$ in U enthalten.

Beweis: Trivialerweise ist M in $\mathcal{L}(M)$ enthalten. Wenn andererseits M eine Teilmenge von U ist, so sind alle Linearkombinationen von Elementen von M , also alle Elemente von $\mathcal{L}(M)$ in U enthalten, d.h. $\mathcal{L}(M) \subseteq U$. \square

Definition: Sei V ein Vektorraum und $M \subseteq V$ eine Teilmenge, so daß $\mathcal{L}(M) = V$ ist. Dann heißt M ein Erzeugendensystem von V .

Beispiele:

1. v sei eine Verschiebung der Ebene, dann ist $\mathcal{L}(\{v\})$ die Menge aller Vielfachen von v , also die Menge aller Verschiebungen in der Richtung von v . Wenn v und w zwei Verschiebungen mit verschiedenen Richtungen sind, so ist $\mathcal{L}\{v, w\}$ die Menge aller Verschiebungen der Ebene.

2. $V = R^3, v = (1, 2, 0), w = (2, 1, 0)$, dann ist

$$\mathcal{L}(\{v\}) = \{v \in R^3 \mid v = (r, 2r, 0) \text{ mit beliebigem } r \in R\},$$

$$\mathcal{L}(\{v, w\}) = \{v = (r, s, 0) \mid r, s \text{ beliebig}\}.$$

(Den Beweis der letzten Aussage überlassen wir dem Leser.)

3. Wir betrachten den Lösungsraum des folgenden homogenen Gleichungssystems, den wir natürlich erst einmal bestimmen müssen:

$$\begin{aligned} x_1 + 3x_2 + 2x_3 + x_4 &= 0 \\ 2x_1 - x_2 + 3x_3 - 4x_4 &= 0 \\ 3x_1 - 5x_2 + 4x_3 - 9x_4 &= 0 \\ x_1 + 17x_2 + 4x_3 + 13x_4 &= 0 \end{aligned}$$

Dazu gehört die folgende Matrix, die wir dem Gaußschen Algorithmus unterwerfen:

$$\begin{pmatrix} 1 & 3 & 2 & 1 & 0 \\ 2 & -1 & 3 & -4 & 0 \\ 3 & -5 & 4 & -9 & 0 \\ 1 & 17 & 4 & 13 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 3 & 2 & 1 & 0 \\ 0 & -7 & -1 & -6 & 0 \\ 0 & -14 & -2 & -12 & 0 \\ 0 & 14 & 2 & 12 & 0 \end{pmatrix} \rightarrow$$

$$\begin{pmatrix} 1 & 3 & 2 & 1 & 0 \\ 0 & 1 & \frac{1}{7} & \frac{6}{7} & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & \frac{11}{7} & -\frac{11}{7} & 0 \\ 0 & 1 & \frac{1}{7} & \frac{6}{7} & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Dazu gehört wiederum das Gleichungssystem

$$\begin{aligned} x_1 + \frac{11}{7}x_3 - \frac{11}{7}x_4 &= 0 \\ x_2 + \frac{1}{7}x_3 + \frac{6}{7}x_4 &= 0, \end{aligned}$$

wo wir $x_3 = s$ und $x_4 = t$ als Parameter wählen können; die Lösungsmenge hat dann die Form

$$L(S) = \left\{ \begin{pmatrix} -\frac{11}{7} \\ \frac{1}{7} \\ 1 \\ 0 \end{pmatrix} s + \begin{pmatrix} \frac{11}{7} \\ -\frac{6}{7} \\ 0 \\ 1 \end{pmatrix} t \mid s, t \text{ aus } R \text{ beliebig} \right\}.$$

Wie Sie sehen, finden wir so ein Erzeugendensystem des Lösungsraums.

2.2 Lineare Unabhängigkeit, Basen, Dimension

Sei nun $M = \{v_1, \dots, v_k\}$ ein Erzeugendensystem des Vektorraums V , also $\mathcal{L}(M) = V$. Dann ist auch $\mathcal{L}(M \cup N) = V$ für jede Teilmenge $N \subseteq V$. Es erhebt sich daher die Frage, ob man aus einem gegebenen Erzeugendensystem den einen oder anderen Vektor weglassen und den Vektorraum mit den restlichen erzeugen kann. Dies führt auf die

Definition: Ein Erzeugendensystem M von V heißt minimal, wenn für jeden Vektor $w \in M$ gilt $\mathcal{L}(M \setminus \{w\}) \neq \mathcal{L}(M) = V$.

Welche Erzeugende kann man denn nun weglassen?

Es sei $M = \{v_1, \dots, v_k\}$. Der Vektor v_k ist überflüssig, wenn $\mathcal{L}(M \setminus \{v_k\}) = \mathcal{L}(M)$ ist, also wenn $v_k \in \mathcal{L}(v_1, \dots, v_{k-1})$ ist. Dann gibt es also Zahlen r_1, \dots, r_{k-1} mit

$$v_k = r_1 v_1 + \dots + r_{k-1} v_{k-1}$$

bzw.

$$0 = r_1 v_1 + \dots + r_k v_k$$

mit $r_k \neq 0$ (nämlich $r_k = -1$). Anders ausgedrückt: Der Nullvektor läßt sich als Linearkombination der v_i darstellen, wobei nicht alle Koeffizienten gleich Null sind.

Definition: Die Menge $\{v_1, \dots, v_k\} \subseteq V$ heißt linear unabhängig, wenn aus

$$r_1v_1 + \dots + r_kv_k = o \quad (r_i \in R)$$

folgt, daß $r_1 = r_2 = \dots = r_k = 0$ ist. Nicht linear unabhängige Mengen heißen linear abhängig, für diese gilt: Es gibt eine Darstellung $r_1v_1 + \dots + r_kv_k = o$ und mindestens ein r_i ist nicht Null.

Minimale Erzeugendensysteme werden wie folgt charakterisiert:

Satz 2.2.1 *Ein Erzeugendensystem M von V ist genau dann minimal, wenn M linear unabhängig ist.*

Beweis: Sei $M = \{v_1, \dots, v_k\}$ ein minimales Erzeugendensystem von V . Wir nehmen zuerst an, M wäre linear abhängig. Dann gibt es Zahlen r_1, \dots, r_k , von denen etwa r_i ungleich Null ist, so daß

$$r_1v_1 + \dots + r_kv_k = o$$

gilt. Es folgt

$$v_i = -\frac{r_1}{r_i}v_1 - \dots - \frac{r_k}{r_i}v_k,$$

also wäre v_i in M überflüssig, was der Voraussetzung widerspricht.

Nun sei M linear unabhängig. Wäre M nicht minimal, so wäre etwa

$$v_k = r_1v_1 + \dots + r_{k-1}v_{k-1}$$

und damit

$$o = r_1v_1 + \dots + r_{k-1}v_{k-1} - 1v_k.$$

In dieser Linearkombination ist ersichtlich ein Koeffizient von Null verschieden, was der vorausgesetzten linearen Unabhängigkeit widerspricht. \square

Satz 2.2.2 *Jede Teilmenge M_1 einer linear unabhängigen Menge M_2 von Vektoren ist linear unabhängig.*

Den Beweis führen wir indirekt: Sei $M_1 = \{v_1, \dots, v_n\}$ linear abhängig, d.h. es gibt Zahlen r_1, \dots, r_n , unter denen etwa $r_i \neq 0$ ist, so daß $o = r_1v_1 + \dots + r_nv_n$.

Wir nehmen weitere Vektoren v_{n+1}, \dots, v_m hinzu (die Gesamtmenge sei M_2) und erhalten die folgende nichttriviale Linearkombination

$$o = r_1v_1 + \dots + r_nv_n + 0v_{n+1} + \dots + 0v_m,$$

damit ist auch die größere Menge linear abhängig, ein Widerspruch. \square

Sei nun M eine linear unabhängige Teilmenge von V . Wir stellen die Frage, ob man weitere Vektoren aus V zu M hinzunehmen kann, so daß auch die größere Menge linear unabhängig bleibt. Wenn dies nicht möglich ist, nennen wir die Menge M eine maximale linear unabhängige Teilmenge:

Definition: Eine linear unabhängige Teilmenge $M \subseteq V$ heißt maximal, wenn $M \cup \{w\}$ für jeden Vektor w aus V linear abhängig ist.

Der folgende Satz charakterisiert maximale linear unabhängige Teilmengen:

Satz 2.2.3 Sei $M \subseteq V$ linear unabhängig. M ist genau dann eine maximale linear unabhängige Teilmenge, wenn $\mathcal{L}(M) = V$, also wenn M ein minimales Erzeugendensystem ist.

Beweis: $M = \{v_1, \dots, v_k\}$ sei eine maximale linear unabhängige Teilmenge. Sei $v \in V$ ein beliebiger Vektor. Wir wissen, daß $M \cup \{v\}$ linear abhängig ist, also läßt sich der Nullvektor wie folgt kombinieren:

$$0 = r_1 v_1 + \dots + r_k v_k + r v,$$

mindestens ein Koeffizient ist von Null verschieden. Wäre $r = 0$, so bliebe

$$0 = r_1 v_1 + \dots + r_k v_k,$$

worin noch ein von Null verschiedener Koeffizient vorkommen soll, was der linearen Unabhängigkeit von M widerspricht. Also muß r von Null verschieden sein, dann läßt sich aber v als Linearkombination aus den v_i darstellen, d.h. M ist ein Erzeugendensystem von V .

Sei umgekehrt M linear unabhängig und $\mathcal{L}(M) = V$. Sei $w \in V$ beliebig, dann liegt w in $\mathcal{L}(M)$, also ist $M \cup \{w\}$ linear abhängig, d.h. M ist eine maximale linear unabhängige Teilmenge. \square

Wir kommen damit zu einem zentralen Begriff:

Definition: Eine Teilmenge $B \subseteq V$ heißt Basis von V , wenn B eine maximale linear unabhängige Teilmenge von V ist.

Es ist äquivalent:

1. B ist eine Basis von V ,
2. B ist eine maximale linear unabhängige Teilmenge von V ,
3. B ist ein linear unabhängiges Erzeugendensystem von V ,
4. B ist ein minimales Erzeugendensystem von V .

Satz 2.2.4 Sei $B = \{v_1, \dots, v_k\}$ eine Basis von V und $v \in V$, dann gibt es eindeutig bestimmte Zahlen r_1, \dots, r_k , so daß $v = r_1 v_1 + \dots + r_k v_k$ ist.

Beweis: Die Existenz folgt aus $\mathcal{L}(B) = V$. Sei etwa

$$v = r_1 v_1 + \dots + r_n v_n = s_1 v_1 + \dots + s_n v_n,$$

dann ist

$$0 = (r_1 - s_1)v_1 + \dots + (r_n - s_n)v_n,$$

wegen der linearen Unabhängigkeit von B folgt $r_i - s_i = 0$ für $i = 1, \dots, k$. \square

Die Zahlen r_1, \dots, r_k heißen die Koordinaten von v bezüglich der Basis B .

Im obigen Beispiel 3 (Lösungsraum eines homogenen Gleichungssystems) sind die erzeugenden Vektoren linear unabhängig, die Zahlen s, t sind also die Koordinaten der Lösung (x_1, \dots, x_4) .

Im Vektorraum R^n der n -tupel gibt es eine sehr einfache Basis, die aus den „Einheitsvektoren“

$$e_i = (0, \dots, 0, 1, 0, \dots, 0)$$

↑ i -te Stelle

besteht. Die Bezeichnung „ e_i “ wollen wir für diese „kanonische“ Basis des R^n reservieren.

Als nächstes beweisen wir den

Satz 2.2.5 (Beschränkungssatz) *Seien $v_1, \dots, v_k \in V$ und $w_1, \dots, w_m \in \mathcal{L}(v_1, \dots, v_k)$. Wenn $\{w_1, \dots, w_m\}$ linear unabhängig ist, so ist $m \leq k$.*

Beweis: Wir nehmen an, es gelte $m > k$. Dann betrachten wir eine Linearkombination $w = r_1 w_1 + \dots + r_m w_m$. Wir fragen uns, ob denn die Zahlen r_1, \dots, r_m so gewählt werden können, daß nicht alle gleich Null sind, aber dennoch $w = o$ ist. Es sei $w_i = \sum a_{ij} v_j$, dann ist

$$w = r_1 \sum a_{1j} v_j + \dots + r_m \sum a_{mj} v_j = \sum (r_1 a_{1j} + \dots + r_m a_{mj}) v_j.$$

Nun ist sicher $w = o$, wenn die Koeffizienten der v_j null sind, also wenn

$$r_1 a_{11} + \dots + r_m a_{m1} = 0$$

...

$$r_1 a_{1k} + \dots + r_m a_{mk} = 0$$

gilt. Dies ist aber ein homogenes Gleichungssystem für die r_j mit k Gleichungen und m Unbekannten, wegen $m > k$ besitzt gibt es ein m -tupel $(r_1, \dots, r_m) \neq (0, \dots, 0)$, das diese Gleichungen erfüllt, für diese Zahlen gilt also

$$w = r_1 w_1 + \dots + r_m w_m = o,$$

d.h. $\{w_1, \dots, w_m\}$ wäre linear abhängig, was der Voraussetzung widerspricht. Folglich ist $m \leq k$. □

Folgerung 2.2.1 *Die Maximalzahl linear unabhängiger Vektoren im R^n ist gleich n .*

Beweis: Wir haben ein Erzeugendensystem aus n Elementen. □

Wir benötigen das einfache

Lemma 2.2.1 *Wenn $\{u_1, \dots, u_k\}$ linear unabhängig ist und u_{k+1} kein Element von $\mathcal{L}\{u_1, \dots, u_k\}$ ist, so ist $\{u_1, \dots, u_{k+1}\}$ linear unabhängig.*

Beweis: Es sei $r_1 u_1 + \dots + r_{k+1} u_{k+1} = o$. Wenn $r_{k+1} \neq 0$ wäre, so könnte man durch r_{k+1} dividieren und hätte u_{k+1} als Linearkombination von u_1, \dots, u_k dargestellt, was nicht möglich ist. Folglich ist $r_{k+1} = 0$ und es bleibt $r_1 u_1 + \dots + r_k u_k = o$. Wegen der linearen Unabhängigkeit von $\{u_1, \dots, u_k\}$ ist auch $r_1 = \dots = r_k = 0$. □

Satz 2.2.6 *Sei V ein Vektorraum, der ein endliches Erzeugendensystem besitzt und $U \subseteq V$ ein Unterraum. Dann besitzt U eine (endliche) Basis.*

Beweis: Wir konstruieren eine maximale linear unabhängige Teilmenge B . Falls $U = \{o\}$ ist, so sei B die leere Menge. Andernfalls wählen wir ein $u_1 \neq o$ aus U . Die Menge $\{u_1\}$ ist natürlich linear unabhängig. Falls $U = \mathcal{L}(u_1)$ ist, so sei $B = \{u_1\}$. Andernfalls wählen wir ein $u_2 \in U$, das nicht in $\mathcal{L}(u_1)$ liegt. Nach dem Lemma ist $\{u_1, u_2\}$ linear unabhängig. Und so weiter: Sei eine linear unabhängige Teilmenge $\{u_1, \dots, u_k\}$ schon gefunden. Wenn $U = \mathcal{L}\{u_1, \dots, u_k\}$ ist, so sei $B = \{u_1, \dots, u_k\}$. Andernfalls wählen wir ein u_{k+1} , das nicht in $\mathcal{L}\{u_1, \dots, u_k\}$ liegt, dann ist wieder $\{u_1, \dots, u_{k+1}\}$ linear unabhängig.

Nach höchstens so vielen Schritten, wie das Erzeugendensystem von V Elemente hat, muß das Verfahren abbrechen, d.h. es tritt der Fall $U = \mathcal{L}(u_1, \dots, u_k)$ ein und wir haben eine Basis konstruiert. \square

Satz 2.2.7 *Je zwei endliche Basen eines Vektorraums V besitzen gleichviele Elemente.*

Beweis: Seien $\{u_1, \dots, u_l\}$ und $\{v_1, \dots, v_k\}$ Basen von V , dann gilt einerseits $v_1, \dots, v_k \in \mathcal{L}(u_1, \dots, u_l)$, diese Vektoren sind linear unabhängig, also ist $k \leq l$. Analog zeigt man $l \leq k$. \square

Definition: Die Zahl der Elemente einer Basis von V heißt die Dimension $\dim V$ von V .

Wir setzen im folgenden stets voraus, daß alle betrachteten Vektorräume eine endliche Basis besitzen.

Nun beweisen wir den

Satz 2.2.8 (Austauschsatz) *Sei $E \subseteq V$ ein Erzeugendensystem des Vektorraums V und $M \subseteq V$ eine linear unabhängige Teilmenge. Dann gibt es eine Teilmenge $F \subseteq E$, so daß $F \cup M$ eine Basis von V ist.*

Beweis: Sei etwa $M = \{u_1, \dots, u_m\}$, $E = \{v_1, \dots, v_k\}$. Die Menge $E \cup M$ erzeugt V . Wir lassen nun schrittweise Elemente aus E weg, solange dies möglich ist, wobei wir stets sichern, daß die verbleibende Menge noch den Vektorraum V erzeugt. Sei nun $F = \{v_1, \dots, v_p\}$ und $F \cup M$ sei ein Erzeugendensystem von V , aus dem kein Element von F weggelassen werden darf, ohne das Erzeugnis zu verkleinern. Wir zeigen, daß $F \cup M$ linear unabhängig ist. Sei also

$$\sum r_i v_i + \sum s_j u_j = o$$

und wir nehmen an, das nicht alle Koeffizienten verschwinden. Nun können nicht alle r_i gleich Null sein, da $\{u_1, \dots, u_m\}$ linear unabhängig ist. D.h. $r_i \neq 0$ für ein i , dann läßt sich also v_i durch die restlichen Vektoren linear kombinieren, kann also aus F weggelassen werden, was der Konstruktion von F widerspricht. Also ist $F \cup M$ eine Basis von V . \square

Als Folgerung erhalten wir den

Satz 2.2.9 (Ergänzungssatz) *Jede linear unabhängige Teilmenge $M \subseteq V$ kann zu einer Basis von V ergänzt werden.*

Beweis: Wir wenden den Austauschsatz für $E = V$ an. \square

Satz 2.2.10 Sei $U \subseteq V$ ein Unterraum. Dann gilt $\dim U \leq \dim V$ und wenn $\dim U = \dim V$ ist, so gilt $U = V$.

Beweis: In U kann es nicht mehr linear unabhängige Vektoren als in V geben, also ist $\dim U \leq \dim V$.

Sei nun $\dim U = \dim V$. Sei $B = \{u_1, \dots, u_m\}$ eine Basis von U . Wir betrachten B als Teilmenge von V ; sie ist linear unabhängig, kann also zu einer Basis B' von V ergänzt werden. Da B' ebenfalls $m = \dim V$ Elemente haben muß, ist $B = B'$ und damit $V = \mathcal{L}(B) = U$. \square

Seien U und W Unterräume des Vektorraums V . Wir überlassen es dem Leser zu zeigen, daß auch der Durchschnitt $U \cap W$ ein Unterraum von V ist.

Wir überlassen es ebenfalls dem Leser, sich davon zu überzeugen, daß die Vereinigung $U \cup W$ im allgemeinen kein Unterraum ist (die Summe eines Vektors aus U und eines Vektors aus W liegt nicht notwendigerweise in $U \cup W$).

Definition: Seien U und W Unterräume des Vektorraums V . Dann heißt $U + W = \mathcal{L}(U \cup W)$ die Summe von U und W . $U + W$ ist also der kleinste Unterraum von V , der U und W enthält.

Lemma 2.2.2 $U + W = \{v \mid \text{es gibt } u \in U \text{ und } w \in W \text{ mit } v = u + w\}$.

Beweis: Jedes Element von $U + W$ ist eine Linearkombination von Vektoren aus U oder W . \square

Nun folgt der

Satz 2.2.11 (Dimensionsatz) $\dim(U + W) = \dim U + \dim W - \dim(U \cap W)$

Beweis: $U \cap W$ ist ein Unterraum von U und von W , diese sind Unterräume von $U + W$. Wir wählen nun eine Basis $B = \{v_1, \dots, v_k\}$ von $U \cap W$, ergänzen sie mit Hilfe von $B_1 = \{u_1, \dots, u_l\}$ zu einer Basis $B \cup B_1$ von U sowie durch $B_2 = \{w_1, \dots, w_m\}$ zu einer Basis $B \cup B_2$ von W . Dann ist

$$U + W = \mathcal{L}(U \cup W) = \mathcal{L}(B \cup B_1, B \cup B_2) = \mathcal{L}(B \cup B_1 \cup B_2).$$

Wir zeigen, daß $B \cup B_1 \cup B_2$ linear unabhängig ist. Es sei also

$$\sum r_i v_i + \sum s_j u_j + \sum t_k w_k = o, \quad (r_i, s_j, t_k \in R),$$

also ist der Vektor

$$\sum r_i v_i + \sum s_j u_j = - \sum t_k w_k$$

sowohl in U wie in W enthalten, also in $U \cap W$. Er ist also durch die Basis B darstellbar:

$$- \sum t_k w_k = \sum p_i v_i$$

oder

$$\sum p_i v_i + \sum t_k w_k = o,$$

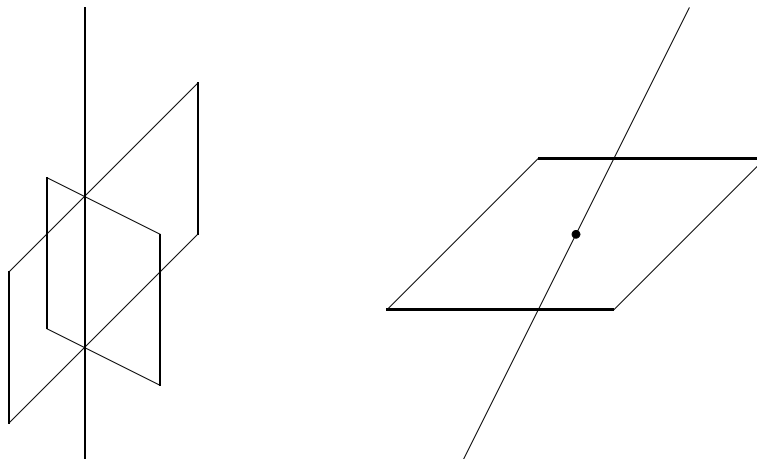
da $B \cup B_2$ linear unabhängig ist, sind alle Koeffizienten gleich Null, also $-\sum t_l w_l = o$, d.h.

$$\sum r_i v_i + \sum s_j u_j = o$$

und aus der linearen Unabhängigkeit von $B \cup B_1$ folgt, daß auch hier alle Koeffizienten verschwinden. Also ist $B \cup B_1 \cup B_2$ eine Basis von $U + W$ und es gilt

$$\dim(U + W) = k + l + m = \dim U + \dim W - k. \square$$

Veranschaulichen Sie sich den Sachverhalt an folgenden Skizzen:



Definition: Wenn $U \cap W = \{o\}$ gilt, so heißt die Summe $U + W$ direkt; man schreibt dann $U \oplus W$.

Es folgt $\dim U \oplus W = \dim U + \dim W$.

Lemma 2.2.3 Die Summe von U und W sei direkt. Dann ist die Darstellung von $v \in U \oplus W$ als $v = u + w$ mit $u \in U$, $w \in W$ eindeutig bestimmt.

Beweis: Sei $v = u + w = u' + w'$ mit $u, u' \in U, w, w' \in W$. Dann ist $u - u' = w' - w$ sowohl in U als auch in W gelegen, also

$$u - u' = w - w' = o. \square$$

Diese Eigenschaft wollen wir zur Definition einer direkten Summe mehrerer Summanden verwenden:

Definition: Die Summe der Unterräume U_1, \dots, U_k von V heißt direkt, wenn die Darstellung jedes Vektors $v = \sum u_i$ mit $u_i \in U_i$ eindeutig bestimmt ist.

Satz 2.2.12 Die Summe der Unterräume U_1, \dots, U_n ist genau dann direkt, wenn für alle i gilt

$$U_i \cap \sum_{k \neq i} U_k = \{o\}.$$

Beweis: Sei die Bedingung erfüllt und $v = \sum u_i = \sum u'_i$ mit $u_i, u'_i \in U_i$. Dann ist

$$u_i - u'_i = \sum_{k \neq i} (u'_k - u_k),$$

dies ist ein Vektor aus $U_i \cap \sum_{k \neq i} U_k = \{o\}$.

Die Umkehrung ist genauso leicht zu zeigen. \square

Beispiele für Vektorräume

Eine Abbildung von einer Menge M in eine Menge N ist eine Vorschrift f , wie jedem Element aus M ein Element aus N zuzuordnen ist.

Die aus der Schule bekannten „Funktionen“ sind Abbildungen, hier sind der Definitionsbereich und der Wertevorrat Teilmengen von \mathbb{R} :

$$f_1(x) = 5x + 7, \quad f_2(x) = x^2, \quad f_3(x) = \sin(x).$$

Wenn V ein Vektorraum ist, so haben wir Abbildungen

$$\text{add} : V \times V \longrightarrow V, \quad \text{add}(v, w) = v + w,$$

$$\text{mult} : R \times V \longrightarrow V, \quad \text{mult}(r, w) = rw.$$

Sei M eine Menge, sei $\text{Abb}(M, R) = \{f : M \longrightarrow R\}$ die Menge aller Abbildungen von M in R . Wir führen folgende Operationen ein: Wenn $f_1, f_2 : M \longrightarrow R$, so setzen wir $(f_1 + f_2)(m) = f_1(m) + f_2(m)$, dann gilt

$$((f_1 + f_2) + f_3)(m) = (f_1 + f_2)(m) + f_3(m) = (f_1(m) + f_2(m)) + f_3(m) = (f_1 + (f_2 + f_3))(m)$$

für alle $m \in M$, also $(f_1 + f_2) + f_3 = f_1 + (f_2 + f_3)$. Für die Nullabbildung $o(m) = 0$ für alle $m \in M$ gilt $f + o = o + f = f$, die Abbildung $-f$ mit $(-f)(m) = -f(m)$ ist zu f additiv invers zu f und es gilt $f_1 + f_2 = f_2 + f_1$. Für $r \in R$ definieren wir $(rf)(m) = rf(m)$ und die restlichen Axiome sind genauso leicht nachzuweisen.

2.3 Anwendung auf lineare Gleichungssysteme

Sei

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ & \dots & \\ a_{i1} & \dots & a_{in} \\ & \dots & \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

eine Matrix. Wir bezeichnen ihre Zeilen mit

$$z_1 = (a_{11} \dots a_{1n}), \dots, z_m = (a_{m1} \dots a_{mn}).$$

Die Vektoren z_1, \dots, z_m erzeugen den Unterraum $\mathcal{L}(z_1, \dots, z_m) = Z(A)$ von R^n , den sogenannten Zeilenraum von A . Die Dimension von $Z(A)$ heißt der Zeilenrang $zr(A)$ von A :

$$zr(A) = \dim \mathcal{L}(z_1, \dots, z_m).$$

Der Zeilenrang ist die Maximalzahl linear unabhängiger Zeilen der Matrix A .

Satz 2.3.1 Wenn A' durch elementare Zeilenoperationen aus A hervorgeht, so ist $zr(A) = zr(A')$.

Beweis: Es ist $\mathcal{L}(z_1, \dots, z_m) = \mathcal{L}(z_1, \dots, cz_i, \dots, z_m) = \mathcal{L}(z_1, \dots, z_k + z_i, \dots, z_m)$, also stimmen sogar die Zeilenräume überein ($c \neq 0$). \square

Mittels des Gaußschen Algorithmus können wir A in eine reduzierte Form bringen, dabei ändert sich der Zeilenraum und damit der Zeilenrang nicht. Wenn die Anzahl der ausgezeichneten Spalten gleich r ist, so sind die ersten r Zeilen linear unabhängig, also ist $zr(A) = r$.

Sei nun

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ &\dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned}$$

ein lineares Gleichungssystem. Wir wollen annehmen, das die reduzierte Form seiner Koeffizientenmatrix die einfache Form

$$\begin{pmatrix} 1 & & a_{1,r+1} & \dots & a_{1n} & b_1 \\ 0 & 1 & a_{2,r+1} & \dots & a_{2n} & b_2 \\ & \dots & & & & \\ 0 & \dots & 1 & a_{r,r+r} & \dots & a_{rn} & b_r \\ 0 & \dots & & & & & 0 \\ & \dots & & & & & \\ 0 & \dots & & & & & 0 \end{pmatrix}$$

besitzt (nach Spaltenvertauschen wäre das zu erreichen). Dann kann man die Lösungsmenge folgendermaßen beschreiben: Jede Lösung hat die Form

$$\begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_r \\ x_{r+1} \\ \dots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_r \\ 0 \\ \dots \\ 0 \end{pmatrix} - t_1 \begin{pmatrix} a_{1,r+1} \\ a_{2,r+1} \\ \dots \\ a_{r,r+1} \\ 1 \\ \dots \\ 0 \end{pmatrix} - \dots - t_{n-r} \begin{pmatrix} a_{1,n} \\ a_{2,n} \\ \dots \\ a_{r,n} \\ 0 \\ \dots \\ 1 \end{pmatrix}$$

Wir sehen also, daß die Zahl der Parameter nicht vom Lösungsweg abhängt.

Folgerung 2.3.1 Sei H ein homogenes Gleichungssystem mit n Unbekannten und der Koeffizientenmatrix A . Dann ist $\dim LM(H) = n - zr(A)$. \square

Sei wieder

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ & \dots & \\ a_{i1} & \dots & a_{in} \\ & \dots & \\ a_{m1} & \dots & a_{mn} \end{pmatrix}.$$

Wir bezeichnen die Spalten von A mit

$$s_1 = \begin{pmatrix} a_{1,1} \\ \dots \\ a_{i,1} \\ \dots \\ a_{m,1} \end{pmatrix}, \dots, s_n = \begin{pmatrix} a_{1,n} \\ \dots \\ a_{i,n} \\ \dots \\ a_{m,n} \end{pmatrix}$$

Diese erzeugen $\mathcal{L}(s_1, \dots, s_n) = S(A)$, den Spaltenraum von A . Die Dimension von $S(A)$ heißt Spaltenrang von A und wird mit $sr(A)$ bezeichnet, dies ist die Maximalzahl linear unabhängiger Spalten.

Es gilt der wichtige

Satz 2.3.2 *Wenn die Matrix A' durch elementare Zeilenoperationen aus der Matrix A hervorgegangen ist, so gilt $sr(A) = sr(A')$.*

Beweis: Ohne Beschränkung der Allgemeinheit können wir annehmen, daß die Spalten

$$s_1 = \begin{pmatrix} a_{1,1} \\ \dots \\ a_{i,1} \\ \dots \\ a_{m,1} \end{pmatrix}, \dots, s_l = \begin{pmatrix} a_{1,l} \\ \dots \\ a_{i,l} \\ \dots \\ a_{m,l} \end{pmatrix}$$

linear unabhängig sind. Bei einer Zeilenoperation (vom Typ 2) gehen sie über in Spalten

$$t_1 = \begin{pmatrix} a_{1,1} \\ \dots \\ a_{i,1} + a_{k,1} \\ \dots \\ a_{m,1} \end{pmatrix}, \dots, t_l = \begin{pmatrix} a_{1,l} \\ \dots \\ a_{i,l} + a_{k,l} \\ \dots \\ a_{m,l} \end{pmatrix}$$

Wir zeigen, daß $\{t_1, \dots, t_l\}$ linear unabhängig ist. Sei nämlich

$$r_1 t_1 + \dots + r_l t_l = o,$$

d.h.

$$\begin{aligned} r_1 a_{11} + \dots + r_l a_{1l} &= 0 \\ &\dots \\ r_1(a_{i1} + a_{k1}) + \dots + r_l(a_{il} + a_{kl}) &= 0 \\ &\dots \\ r_1 a_{m1} + \dots + r_l a_{ml} &= 0. \end{aligned}$$

Aus diesen Gleichungen folgt aber sofort

$$\begin{aligned} r_1 a_{11} + \dots + r_l a_{1l} &= 0 \\ &\dots \\ r_1 a_{i1} + \dots + r_l a_{il} &= 0 \\ &\dots \\ r_1 a_{m1} + \dots + r_l a_{ml} &= 0. \end{aligned}$$

Dieses Gleichungssystem hat aber nur die triviale Lösung, weil s_1, \dots, s_l linear unabhängig sind. Also gilt $sr(A') \geq sr(A)$ und die Gleichheit folgt aus Symmetriegründen.

□

Satz 2.3.3 *Für jede Matrix A gilt $sr(A) = zr(A)$. Diese Zahl wird als Rang $rg(A)$ von A bezeichnet.*

Beweis: Wir überführen A in die reduzierte Form

$$\begin{pmatrix} 0 & \dots & 1 & a_{1,k_1+1} & \dots & a_{1,k_1-1} & 0 & a_{1,k_2+1} & \dots & a_{1,k_r-1} & 0 & \dots & a_{1n} \\ 0 & \dots & 0 & & \dots & 0 & 1 & a_{2,k_2+1} & \dots & a_{2,k_r-1} & 0 & a_{2,k_r+1} & \dots & a_{2n} \\ 0 & \dots & 0 & & \dots & & & & & & 1 & a_{r,k_r+1} & \dots & a_{rn} \\ \dots & & & & & & & & & & & & & \\ 0 & & & & & & \dots & & & & & & & 0 \\ \dots & & & & & & & & & & & & & \\ 0 & & & & & & \dots & & & & & & & 0 \end{pmatrix}$$

Es ist $zr(A) = r$, denn die ersten r Zeilen sind linear unabhängig. Und es ist $sr(A) = r$, da die r ausgezeichneten Spalten linear unabhängig sind, die übrigen aber Linearkombinationen davon sind. □

Satz 2.3.4 (Kronecker/Capelli) *Das Gleichungssystem $\sum a_{ij}x_j = b_i$ ist genau dann lösbar, wenn*

$$rg \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} = rg \begin{pmatrix} a_{11} & \dots & a_{1n} & b_1 \\ \dots & \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} & b_m \end{pmatrix}$$

ist.

Den Beweis überlassen wir dem Leser. □

Abschließend erwähnen wir, daß, wie aus dem oben Gesagten folgt, der Gaußsche Algorithmus geeignet ist, die Dimension und eine Basis eines Unterraums des R^n zu bestimmen, der durch ein Erzeugendensystem gegeben ist. Dazu werden die erzeugenden Vektoren zeilenweise in eine Matrix A geschrieben, auf die Matrix werden entsprechende Zeilenoperationen angewandt, die ja neue Zeilen produzieren, die im selben Vektorraum liegen. Die Dimension des Unterraums ist gleich $rg(A)$ und die ersten $rg(A)$ Zeilen bilden eine Basis des Unterraums.

Kapitel 3

Lineare Abbildungen und Matrizen

3.1 Grundlegende Eigenschaften

Wir beginnen mit einem Beispiel.

Sei V ein zweidimensionaler Vektorraum und $B = \{u, v\}$ eine Basis von V . Dann kann man einen beliebigen Vektor $w \in V$ in eindeutiger Weise als $w = ru + sv$ ($r, s \in R$) darstellen, dabei sind die Zahlen r und s die Koordinaten von w bezüglich der gewählten Basis. Wir ordnen dem Vektor w dieses Zahlenpaar zu: Sei

$$k_B : V \rightarrow R^2 \text{ mit } k(w) = (r, s)$$

die „Koordinatenabbildung“, die jedem Vektor aus V sein Koordinatenpaar zuordnet. Diese Abbildung k hat folgende Eigenschaften:

Sei w' ein weiterer Vektor aus V mit den Koordinaten (r', s') , also $k_B(w') = (r', s')$. Wegen $w' = r'u + s'v$ gilt

$$w + w' = (r + r')u + (s + s')v,$$

also

$$k_B(w + w') = (r + r', s + s') = k_B(w) + k_B(w').$$

Sei t eine Zahl, dann hat der Vektor tw die Koordinaten (tr, ts) , also gilt

$$k_B(tw) = (tr, ts) = tk_B(w).$$

Es ist sicher verständlich, daß Abbildungen, die sich derart gut gegenüber den Operationen in Vektorräumen verhalten, in der Theorie der Vektorräume eine gewisse Rolle spielen werden.

Definition: Seien V und W R -Vektorräume und $f : V \rightarrow W$ eine Abbildung von V in W , für die für beliebige $u, v \in V$ und $r \in R$

$$f(u + v) = f(u) + f(v) \quad (f \text{ ist „additiv“}) \text{ sowie}$$

$$f(rv) = rf(v) \quad (f \text{ ist „homogen“})$$

gilt, dann heißt f „lineare Abbildung“.

Beispiele für lineare Abbildungen:

1. Wenn $\dim V = n$ und eine Basis B von V gewählt ist, so erhalten wir in Verallgemeinerung des obigen Beispiels die Koordinatenabbildung $k_B : V \rightarrow R^n$, die jedem Vektor sein Koordinaten- n -tupel bezüglich B zuordnet. Dieses Beispiel wird uns später noch beschäftigen.
2. Sei $i \leq n$, wir betrachten die „Projektionsabbildung“

$$p_i : R^n \rightarrow R, p_i(r_1, \dots, r_n) = r_i,$$

sie ist linear.

3. Allgemeiner: Sei $V = U \oplus W$ direkte Summe von Unterräumen und für $v \in V$ sei $v = u + w$ mit $u \in U, w \in W$; dann ist $p : V \rightarrow U$ mit $p(v) = u$ die Projektion auf U .
4. Die „identische“ Abbildung $id : V \rightarrow V, id(v) = v$ für alle $v \in V$ ist linear.
5. Zwischen beliebigen Vektorräumen V, W gibt es eine Nullabbildung $o : V \rightarrow W, o(v) = o$ für alle $v \in V$, hierbei bezeichnen die beiden ersten o 's die Abbildung, das dritte o ist der Nullvektor von W . Die Bezeichnungskonfusion darf man ausnahmsweise durchgehen lassen, denn wir werden sehen, daß die Nullabbildung das neutrale Element eines gewissen Vektorraums ist, und für derartige Vektoren hatten wir ausdrücklich das Symbol o reserviert.

Für Abbildungen mit bestimmten Eigenschaften haben sich Attribute eingebürgert, die wir nun kurz aufzählen wollen.

Seien A und B Mengen und $f : A \rightarrow B$ eine Abbildung von A in B . Die Abbildung f heißt „injektiv“ (oder „1-1-deutig“), wenn aus $f(a) = f(a')$ stets $a = a'$ folgt, wobei a, a' beliebige Elemente von A sind.

Die Abbildung f heißt „surjektiv“, wenn für jedes Element $b \in B$ ein Element $a \in A$ existiert, so daß $f(a) = b$ ist (eine surjektive Abbildung von A auf B heißt auch „Abbildung auf B “ [es gibt keine deutsche Übersetzung des Adjektivs „surjektiv“]). Die Abbildung f heißt bijektiv, wenn sie injektiv und surjektiv ist.

Lineare Abbildungen werden gelegentlich auch als „Homomorphismen“ von Vektorräumen bezeichnet. Davon leiten sich die folgenden, häufig anzutreffenden Bezeichnungen ab:

- ein „Monomorphismus“ ist eine injektive lineare Abbildung,
- ein „Epimorphismus“ ist eine surjektive lineare Abbildung,
- ein „Isomorphismus“ ist eine bijektive lineare Abbildung,
- ein „Endomorphismus“ ist eine lineare Abbildung eines Vektorraums V in sich,
- ein „Automorphismus“ ist ein bijektiver Endomorphismus.

Untersuchen Sie, welche Attribute für die in den obigen Beispielen angegebenen linearen Abbildungen zutreffen!

Wir wollen nun Operationen zwischen linearen Abbildungen einführen:

Seien $f, g : V \rightarrow W$ zwei lineare Abbildungen von V in W . Wir konstruieren eine lineare Abbildung $f + g : V \rightarrow W$ von V in W wie folgt:

$(f + g)(v) = f(v) + g(v)$ für alle $v \in V$.

Sei $s \in R$ eine Zahl, wir konstruieren eine lineare Abbildung $sf : V \rightarrow W$ wie folgt:

$(sf)(v) = sf(v)$ für alle $v \in V$.

Lemma 3.1.1 *Die Abbildungen $f + g$ und sf sind linear.*

Beweis: Wir prüfen die Axiome nach: Seien $v, v' \in V$ und $r \in R$, dann gilt

$$\begin{aligned} (f + g)(v + rv') &= f(v + rv') + g(v + rv') \\ &\quad \text{nach Definition von } f + g, \\ &= f(v) + rf(v') + g(v) + rg(v') \\ &\quad \text{wegen der Linearität von } f \text{ und } g, \\ &= (f + g)(v) + r(f + g)(v') \end{aligned}$$

wieder nach Definition von $f + g$. Für $r = 1$ erhalten wir die Additivität von $f + g$, für $v = o$ erhalten wir die Homogenität. Weiter ist

$$\begin{aligned} (sf)(v + rv') &= sf(v + rv') \\ &= s(f(v) + rf(v')) \\ &= sf(v) + (sr)f(v') \\ &= (sf)(v) + r(sf)(v'). \quad \square \end{aligned}$$

Definition: Die Menge aller linearer Abbildungen eines Vektorraums V in einen Vektorraum W wird mit $\text{Hom}(V, W)$ bezeichnet.

Satz 3.1.1 $\text{Hom}(V, W)$ ist ein Vektorraum.

Beweis: Summen und Vielfache linearer Abbildungen sind linear, wie wir eben gesehen haben. Es bleiben die Vektorraumaxiome zu überprüfen. Da wäre etwa die Frage nach der Existenz eines neutralen Elements zu stellen. Wir zeigen, daß die Nullabbildung der Nullvektor von $\text{Hom}(V, W)$ ist:

Sei $f : V \rightarrow W$ eine beliebige lineare Abbildung, dann ist $(f + o)(v) = f(v) + o(v) = f(v) + o = f(v)$ für beliebige Vektoren $v \in V$, also ist $f + o = f$.

Wir wollen lediglich noch ein Distributivgesetz beweisen, der Rest bleibt dem Leser überlassen. Seien $f, g : V \rightarrow W$ lineare Abbildungen von V in W , $v \in V$ und $r \in R$, dann gilt:

$$\begin{aligned} (r(f + g))(v) &= r((f + g)(v)) \\ &= r(f(v) + g(v)) \\ &= rf(v) + rg(v) \\ &= (rf + rg)(v), \end{aligned}$$

und zwar für beliebige $v \in V$. Das heißt, daß die Abbildungen $r(f + g)$ und $rf + rg$ gleich sind. \square

Wir führen noch eine weitere Operation zwischen linearen Abbildungen ein: Seien $g : V \rightarrow W$ und $f : W \rightarrow U$ lineare Abbildungen. Wir konstruieren die Abbildung $f \circ g : V \rightarrow U$ wie folgt:

$$(f \circ g)(v) = f(g(v)) \text{ für } v \in V.$$

Nur in dieser Situation (der Definitionsbereich von f stimmt mit dem Wertevorrat von g überein) ist das „Produkt“ (oder die „Komposition“) von f und g definiert.

Lemma 3.1.2 *Die Abbildung $f \circ g$ ist linear.*

Beweis: Seien $v, v' \in V$ und $r \in R$, dann gilt

$$\begin{aligned} (f \circ g)(v + rv') &= f(g(v + rv')) \\ &\quad \text{nach Definition,} \\ &= f(g(v) + rg(v')) \\ &\quad \text{wegen der Linearität von } g, \\ &= f(g(v)) + rf(g(v')) \\ &\quad \text{wegen der Linearität von } f, \\ &= (f \circ g)(v) + r(f \circ g)(v') \\ &\quad \text{nach Definition von } f \circ g. \end{aligned}$$

□

Bezüglich dieser (nicht uneingeschränkt ausführbaren) Multiplikation verhalten sich die verschiedenen identischen Abbildungen wie „Einselemente“:

Lemma 3.1.3 *Sei $f : V \rightarrow W$ eine lineare Abbildung und seien $id_V : V \rightarrow V$ sowie $id_W : W \rightarrow W$ die jeweiligen identischen Abbildungen, dann gilt $f \circ id_V = f = id_W \circ f$.*

Beweis: $(f \circ id_V)(v) = f(id_V(v)) = f(v) = id_W(f(v)) = (id_W \circ f)(v)$ für alle $v \in V$, also folgt die Behauptung. □

Wenn die lineare Abbildung $f : V \rightarrow W$ bijektiv ist, so existiert eine Abbildung $g : W \rightarrow V$ mit $f \circ g = id_W$ und $g \circ f = id_V$, wir konstruieren nämlich g wie folgt:

Sei $w \in W$ gewählt, da f surjektiv ist, gibt es ein $v \in V$ mit $f(v) = w$. Dieser Vektor v ist eindeutig bestimmt, denn wenn noch $f(v') = w$ wäre, so folgt $v = v'$ aus der Injektivität von f . Wir setzen dann $g(w) = v$.

Lemma 3.1.4 *Die Abbildung g ist linear.*

Beweis: Sei $g(w) = v, g(w') = v'$ sowie $r \in R$. Dies ist genau dann der Fall, wenn $f(v) = w$ und $f(v') = w'$ ist. Aus der Linearität von f folgt $f(v + rv') = w + rw'$, d.h. $g(w + rw') = g(w) + rg(w')$. □

Definition: Die soeben konstruierte Abbildung g heißt die zu f inverse Abbildung, sie wird mit f^{-1} bezeichnet.

Zu einer linearen Abbildung $f : V \rightarrow W$ gehören zwei Unterräume von V bzw. von W :

Definition: Sei $f : V \rightarrow W$ eine lineare Abbildung.

$$\text{Ker}(f) = \{v \in V \mid f(v) = o\}$$

heißt der Kern von f .

$$\text{Im}(f) = \{w \in W \mid \text{es gibt ein } v \in V \text{ mit } f(v) = w\} = f(V)$$

heißt das Bild von f .

Lemma 3.1.5 $\text{Ker}(f) \subseteq V$ und $\text{Im}(f) \subseteq W$ sind Unterräume.

Beweis: Seien $v, v' \in \text{Ker}(f)$ und $r \in R$, d.h. es ist $f(v) = f(v') = o$. Dann ist $f(v + rv') = f(v) + rf(v') = o + o = o$. Seien $w, w' \in \text{Im}(f)$ und $r \in R$, d.h. es gibt $v, v' \in V$ mit $f(v) = w$ und $f(v') = w'$. Dann ist $w + rw' = f(v) + rf(v') = f(v + rv') \in \text{Im}(f)$. \square

Nützlich, wenn auch trivial ist das folgende

Lemma 3.1.6 Die lineare Abbildung $f : V \rightarrow W$ ist genau dann surjektiv, wenn $\text{Im}(f) = W$. Die lineare Abbildung $f : V \rightarrow W$ ist genau dann injektiv, wenn $\text{Ker}(f) = \{o\}$.

Beweis: Die erste Aussage ergibt sich aus der Definition, ist also wirklich trivial.

Sei nun f injektiv und $v \in \text{Ker}(f)$, also $f(v) = o$. Nun gibt es aber einen Vektor $u \in V$, der auf alle Fälle im Kern von f liegt, nämlich $u = o$ (es ist $f(o) = o$). Wegen der Injektivität von f muß also $v = u = o$ sein, also ist $\text{Ker}(f) = \{o\}$.

Sei umgekehrt $\text{Ker}(f) = \{o\}$ und sei $f(v) = f(v')$, dann ist $f(v - v') = f(v) - f(v') = o$, also liegt $v - v'$ im Kern von f , also $v - v' = o$, d.h. $v = v'$, folglich ist f injektiv. \square

Wir wollen im folgenden untersuchen, wie lineare Abbildungen auf linear abhängige bzw. unabhängige sowie erzeugenden Teilmengen wirken.

Mit $f(M)$ bezeichnen wir die Menge

$$f(M) = \{w \in W \mid \text{es gibt } v \in M \text{ mit } f(v) = w\}.$$

In diesem Sinne ist $\text{Im}(f) = f(V)$.

Satz 3.1.2 Sei $f : V \rightarrow W$ eine lineare Abbildung und $M \subseteq V$ ein Erzeugendensystem von V . Dann ist $f(M)$ ein Erzeugendensystem von $\text{Im}(f)$.

Beweis: Sei $w \in \text{Im}(f)$, dann gibt es ein $v \in V$ mit $w = f(v)$. Es sei

$$v = \sum r_i v_i \text{ mit } v_i \in M,$$

dann ist

$$w = \sum r_i f(v_i) \in \mathcal{L}(f(M)). \square$$

Sei nun $f : V \rightarrow W$ eine lineare Abbildung und $\{v_1, \dots, v_k\}$ eine linear abhängige Teilmenge von V . Dann gibt es Zahlen r_1, \dots, r_k , die nicht alle null sind, so daß $r_1 v_1 + \dots + r_k v_k = o$.

Durch Anwendung von f und Ausnutzung der Linearität von f erhalten wir

$$\begin{aligned} o &= f(r_1v_1 + \dots + r_kv_k) \\ &= f(r_1v_1) + \dots + f(r_kv_k) \\ &= r_1f(v_1) + \dots + r_kf(v_k), \end{aligned}$$

also ist auch $\{f(v_1), \dots, f(v_k)\}$ linear abhängig.

Wir erhalten den

Satz 3.1.3 Sei $f : V \rightarrow W$ eine lineare Abbildung und v_1, \dots, v_k Vektoren aus V . Wenn $\{f(v_1), \dots, f(v_k)\}$ linear unabhängig ist, so ist $\{v_1, \dots, v_k\}$ auch linear unabhängig. \square

Satz 3.1.4 Sei $f : V \rightarrow W$ eine lineare Abbildung, weiter sei $U \subseteq V$ ein Teilraum von V , so daß der Durchschnitt von U und $\text{Ker}(f)$ nur den Nullvektor enthält. Wenn nun $\{v_1, \dots, v_k\}$ eine linear unabhängige Teilmenge von U ist, so ist auch $\{f(v_1), \dots, f(v_k)\}$ linear unabhängig.

Beweis: Sei $\sum r_i f(v_i) = o = f(\sum r_i v_i)$, also liegt $\sum r_i v_i$ im Durchschnitt von $\text{Ker}(f)$ und U , also gilt $\sum r_i v_i = o$ und aus der linearen Unabhängigkeit von $\{v_1, \dots, v_k\}$ folgt $r_1 = \dots = r_k = o$. \square

Den folgenden Satz werden wir oft anwenden:

Satz 3.1.5 Sei $f : V \rightarrow W$ eine lineare Abbildung, dann gibt es einen Unterraum $U \subseteq V$ mit $U \oplus \text{Ker}(f) = V$ und es gilt $\dim V = \dim \text{Ker}(f) + \dim \text{Im}(f)$.

Beweis: Wir wählen eine Basis $\{v_1, \dots, v_k\}$ von $\text{Ker}(f)$ und ergänzen sie zur Basis $\{v_1, \dots, v_n\}$ von V . Wir setzen $U = \mathcal{L}(\{v_{k+1}, \dots, v_n\})$, dann ist $\text{Ker}(f) \oplus U = V$. Da $\mathcal{L}(\{v_1, \dots, v_n\}) = V$ und $f(v_1) = \dots = f(v_k) = o$ ist, gilt $\mathcal{L}(\{f(v_1), \dots, f(v_n)\}) = \mathcal{L}(\{f(v_{k+1}), \dots, f(v_n)\}) = \text{Im}(f)$. Nach dem vorigen Satz ist $\{f(v_{k+1}), \dots, f(v_n)\}$ linear unabhängig, also eine Basis von $\text{Im}(f)$ und es folgt

$$\dim V = n = k + (n - k) = \dim \text{Ker}(f) + \dim \text{Im}(f). \quad \square$$

Folgerung 3.1.1 Wenn $f : V \rightarrow W$ ein Isomorphismus ist (also eine bijektive lineare Abbildung), dann ist $\dim V = \dim W$.

Beweis: Es ist $\text{Ker}(f) = \{o\}$ und $\text{Im}(f) = W$, nach der obigen Dimensionsformel ist $\dim V = \dim W$. \square

3.2 Darstellungsmatrizen

Der folgende Satz zeigt, daß eine lineare Abbildung schon durch die Bildvektoren einer Basis bestimmt ist.

Satz 3.2.1 (Prinzip der linearen Fortsetzung) Sei $B = \{v_1, \dots, v_n\}$ eine Basis von V und $w_1, \dots, w_n \in W$ beliebig gewählte Vektoren. Dann gibt es genau eine lineare Abbildung

$$f : V \rightarrow W \text{ mit } f(v_i) = w_i \text{ für } i = 1, \dots, n.$$

Beweis: Wir zeigen zunächst die Einzigkeit: Sei f eine derartige Abbildung und $v \in V$, es sei $v = \sum r_i v_i$, dann folgt aus der Linearität von f , daß

$$f(v) = \sum r_i f(v_i) = \sum r_i w_i$$

ist. Zur Existenz: Wir setzen für $v = \sum r_i v_i \in V$ fest:

$$f(v) = \sum r_i w_i.$$

Diese Abbildung ist linear: Sei noch $v' = \sum r'_i v_i$ und $r \in R$. Dann ist

$$\begin{aligned} f(v + rv') &= \sum (r_i + rr'_i) w_i \\ &= \sum r_i w_i + r \sum r'_i w_i \\ &= f(v) + rf(v'). \quad \square \end{aligned}$$

Lemma 3.2.1 Sei $B = \{v_1, \dots, v_n\}$ eine Basis des Vektorraums V , dann ist die durch $k_B(v_i) = e_i = (0, \dots, 1, \dots, 0)$ gegebene Koordinatenabbildung $k_B : V \rightarrow R^n$ ein Isomorphismus.

Beweis: Die Abbildung ist surjektiv, denn ein gegebenes n -tupel (r_1, \dots, r_n) ist Bild von $\sum r_i v_i$. Sie ist injektiv, denn falls $k_B(v) = (0, \dots, 0)$ ist, ist $v = 0$. \square

Wir wenden das Prinzip der linearen Fortsetzung an, um lineare Abbildungen zahlenmäßig beschreiben zu können:

Sei $f : V \rightarrow W$ eine lineare Abbildung. Wir wählen Basen $B = \{v_1, \dots, v_n\}$ von V und $C = \{w_1, \dots, w_m\}$ von W . Dann können wir jeden Vektor $f(v_i)$ durch die Basis C ausdrücken:

$$f(v_i) = \sum f_{ji} w_j, \quad i = 1, \dots, n.$$

In der i -ten Spalte stehen die Koordinaten des Bildes des i -ten Basisvektors.

Die Matrix (f_{ji}) (mit m Zeilen und n Spalten) bezeichnen wir mit

$$A_{BC}(f) = (f_{ji})$$

und nennen sie die f bezüglich B und C zugeordnete Darstellungsmatrix.

Beispiel:

$f : R^4 \rightarrow R^2$ sei die folgende (lineare) Abbildung:

$$f(w, x, y, z) = (w + x + y, z - w - x),$$

$$B = \{(1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0), (0, 0, 0, 1)\},$$

$C = \{(1, 0), (0, 1)\}$ seien die „kanonischen“ Basen, dann ist

$$A_{BC}(f) = \begin{pmatrix} 1 & 1 & 1 & 0 \\ -1 & -1 & 0 & 1 \end{pmatrix}.$$

Es ist klar, daß sich (bei gegebenen Basen) die lineare Abbildung f und die ihr zugeordnete Matrix $A_{BC}(f)$ gegenseitig eindeutig bestimmen, wir haben also eine bijektive Abbildung

$$A_{BC} : \text{Hom}(V, W) \rightarrow M_{mn},$$

dabei bezeichnet M_{mn} den Vektorraum der Matrizen mit m Zeilen und n Spalten. Wir zeigen, daß die Abbildung A_{BC} linear ist: Seien also $f, f' : V \rightarrow W$ lineare Abbildungen und $r \in R$, $B = \{v_1, \dots, v_n\}$ sei eine Basis von V , $C = \{w_1, \dots, w_m\}$ eine Basis von W und

$$f(v_i) = \sum f_{ji} w_j, \quad f'(v_i) = \sum f'_{ji} w_j,$$

also

$$A_{BC}(f) = (f_{ji}), \quad A_{BC}(f') = (f'_{ji}).$$

Dann ist

$$\begin{aligned} (f + rf')(v_i) &= f(v_i) + rf'(v_i) \\ &= \sum f_{ji} w_j + r \sum f'_{ji} w_j \\ &= \sum (f_{ji} + rf'_{ji}) w_j \end{aligned}$$

Also ist

$$A_{BC}(f + rf') = A_{BC}(f) + rA_{BC}(f')$$

Damit erhalten wir die

Folgerung 3.2.1 Sei $\dim V = n$ und $\dim W = m$, dann sind die Vektorräume $\text{Hom}(V, W)$ und M_{mn} isomorph, sie haben die Dimension $m \cdot n$. \square

3.3 Matrixmultiplikation, Inverse von Matrizen

Seien nun lineare Abbildungen $f : V \rightarrow W$ und $g : W \rightarrow U$ gegeben, dann ist $g \circ f$ eine lineare Abbildung von V in U . Wir bestimmen nun die $g \circ f$ zugeordnete Darstellungsmatrix. Dazu wählen wir Basen $B = \{v_1, \dots, v_n\}$ von V , $C = \{w_1, \dots, w_m\}$ von W und $D = \{u_1, \dots, u_l\}$ von U . Es sei

$$f(v_i) = \sum f_{ji} w_j, \quad g(w_j) = \sum g_{kj} u_k,$$

dann ist

$$\begin{aligned} g \circ f(v_i) &= g\left(\sum f_{ji} w_j\right) \\ &= \sum f_{ji} g(w_j) \\ &= \sum_j f_{ji} \sum_k g_{kj} u_k \\ &= \sum_{kj} \left(\sum f_{ji} g_{kj}\right) u_k, \end{aligned}$$

also ist

$$A_{BD}(g \circ f) = \left(\sum_j g_{kj} f_{ji} \right).$$

Wir kommen damit zur

Definition: Die Matrix $(h_{ki}) \in M_{ln}$ mit $h_{ki} = \sum g_{kj} f_{ji}$ heißt das Produkt der Matrizen $(g_{kj}) \in M_{lm}$ und $(f_{ji}) \in M_{mn}$.

Damit gilt

$$A_{BD}(g \circ f) = A_{CD}(g)A_{BC}(f).$$

Es ist nützlich, sich die Art und Weise, wie zwei Matrizen multipliziert werden, genau zu merken: um die (k, i) -Komponente des Produkts GF der Matrizen G und F zu erhalten, werden die Komponenten der k -ten Zeile von G mit denen der i -ten Spalte von F multipliziert und alles addiert. Dazu müssen natürlich die Anzahl der Komponenten in den Zeilen von G (also die Spaltenzahl von G) und die Zahl der Komponenten in den Spalten von F (also die Zeilenzahl von F) übereinstimmen, dies ist durch die Voraussetzungen gesichert.

Beispiel:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} = \begin{pmatrix} 22 & 28 \\ 49 & 64 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 9 & 12 & 15 \\ 19 & 26 & 33 \\ 29 & 40 & 51 \end{pmatrix}$$

Der Leser möge sich bitte selbst überlegen, daß für die Matrixmultiplikation die folgenden Eigenschaften gelten

$$H(GF) = (HG)F,$$

$$H(G + F) = HG + HF,$$

$$(H + G)F = HF + GF.$$

Man kann diese Identitäten entweder durch Nachrechnen verifizieren, oder man überlegt, daß die Matrixmultiplikation so definiert wurde, daß bei dem obigen Isomorphismus zwischen dem Raum der linearen Abbildung und dem Raum der Matrizen das Produkt von Abbildungen dem Matrixprodukt entspricht, und daß für Abbildungen analoge Identitäten gelten.

Betrachten wir die identische Abbildung $id : V \rightarrow V$. Wir wählen eine Basis $B = \{v_1, \dots, v_n\}$ von V , dann ist $id(v_i) = v_i$, also

$$A_{BB}(id) = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ & \dots & & \\ 0 & \dots & 1 & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}.$$

Diese Matrix heißt Einheitsmatrix, wir reservieren hierfür die Bezeichnung E_n oder auch einfach E . Dann gilt $E_m F = F = F E_n$.

Wenn die lineare Abbildung $f : V \rightarrow W$ ein Isomorphismus ist, so existiert eine zu f inverse Abbildung $f^{-1} : W \rightarrow V$ und für die zugeordneten Matrizen gilt

$$A_{BC}(f)A_{CB}(f^{-1}) = A_{CC}(id_W) = E.$$

Dies motiviert die folgende

Definition: Wenn für zwei quadratische Matrizen F, G gilt $FG = GF = E$, so heißt G die zu F inverse Matrix, wir schreiben $G = F^{-1}$. Wenn F eine Inverse besitzt, so nennen wir F regulär, andernfalls singulär.

Also gilt

$$A_{CB}(f^{-1}) = A_{BC}(f)^{-1}.$$

Mit der oben eingeführten Matrixmultiplikation kann man ein lineares Gleichungssystem

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ &\dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned}$$

als Matrixprodukt schreiben:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & & & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{pmatrix},$$

oder kurz $AX = B$, wo $A \in M_{mn}$ die Koeffizientenmatrix, $X \in M_{n1}$ der Spaltenvektor der Unbekannten und $B \in M_{m1}$ die rechte Seite des Gleichungssystems ist.

Wenn nun A eine reguläre Matrix ist (das kommt vor), so kann man die eindeutig bestimmte Lösung des Gleichungssystems $AX = B$ sehr leicht bestimmen, wenn A^{-1} bekannt ist: $X = A^{-1}B$.

Es stellen sich also wieder zwei Fragen:

Wann existiert eine Inverse einer Matrix?

Wie kann man eine Inverse einer Matrix berechnen?

Zunächst beweisen wir den

Satz 3.3.1 Sei $f : V \rightarrow W$ eine lineare Abbildung, B eine Basis von V und C eine Basis von W , dann ist $rg(A_{BC}(f)) = \dim \text{Im}(f)$.

Beweis: Sei $B = \{v_1, \dots, v_n\}$, dann ist $\{f(v_1), \dots, f(v_n)\}$ ein Erzeugendensystem von $\text{Im}(f)$, sei oBdA. $\{f(v_1), \dots, f(v_r)\}$ eine maximale linear unabhängige Teilmenge. Die Spalten von $A_{BC}(f)$ sind nun die Bilder $k_C(f(v_i))$ der $f(v_i)$ unter der Koordinatenabbildung k_C . Da diese ein Isomorphismus ist, sind die ersten r Spalten linear unabhängig und die restlichen sind Linearkombinationen der ersten r Spalten. Also ist $rg(A_{BC}(f)) = r$. \square

Wir fassen eine gegebene Matrix $F \in M_{lk}$ wie folgt als Abbildung von R^k in R^l auf: Das Bild des Spaltenvektors $X \in R^k$ sei einfach das Matrixprodukt FX . Die Abbildung $F : R^k \rightarrow R^l$ ist offenbar linear.

Sei nun wieder $f : V \rightarrow W$ eine lineare Abbildung, B, C Basen von V bzw. W und $F = A_{BC}(f)$. Dann setzen wir aus den Abbildungen k_B, k_C, f und F das folgende „Diagramm“ zusammen:

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ k_B \downarrow & & \downarrow k_C \\ R^n & \xrightarrow{F} & R^m \end{array}$$

Wir zeigen, daß $k_C \circ f = F \circ k_B$ gilt (ein derartiges Diagramm heißt „kommutativ“):

Es sei also $f(v_i) = \sum f_{ji}w_j$ und $v = \sum r_i v_i$. Dann gilt

$$k_C(f(v)) = k_C(f(\sum r_i v_i)) = k_C(\sum r_i f(v_i)) = k_C(\sum_{ij} r_i f_{ji} w_j) = \begin{pmatrix} \sum r_i f_{1i} \\ \vdots \\ \sum r_i f_{mi} \end{pmatrix}$$

und

$$F \cdot k_B(v) = F \cdot \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} = \begin{pmatrix} \sum f_{1i} r_i \\ \vdots \\ \sum f_{mi} r_i \end{pmatrix},$$

das heißt, die Koordinaten von $f(v)$ erhält man, wenn man die Darstellungsmatrix mit dem Koordinatentupel von v multipliziert.

Beispiel: Wir betrachten $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3x - y \\ -2x + 5y \end{pmatrix}$.

Dann ist $F = \begin{pmatrix} 3 & -1 \\ -2 & 5 \end{pmatrix}$ und $f \begin{pmatrix} -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 & -1 \\ -2 & 5 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 2 \end{pmatrix} = \begin{pmatrix} -5 \\ 12 \end{pmatrix}$.

Nun können wir sagen, wann eine zu F inverse Matrix existiert:

Satz 3.3.2 F^{-1} existiert genau dann, wenn f^{-1} existiert.

Beweis: Wenn f^{-1} existiert, so ist die zugehörige Matrix zu F invers. Wenn F^{-1} existiert, so setzen wir $f' = k_B^{-1} \circ F^{-1} \circ k_C$, dabei haben wir wie oben die Matrix F^{-1} als Abbildung aufgefaßt. Man rechnet schnell nach, daß $f' \circ f = id$ und $f \circ f' = id$ ist. \square

Wir haben auch gleich gesehen, daß F^{-1} eindeutig bestimmt ist.

Folgerung 3.3.1 Sei $F \in M_{nn}$, F ist genau dann regulär, wenn $rg(F) = n$ ist.

Beweis: Die Abbildung $f : V \rightarrow V$ ist genau dann ein Isomorphismus, wenn $\text{Ker}(f) = \{o\}$ und $\text{Im}(f) = V$ ist. Wir haben also $n = \dim V = \dim \text{Im}(f) + \dim \text{Ker}(f) = rg(F)$. \square

Folgerung 3.3.2 Seien G und F multiplizierbare Matrizen, dann ist $rg(GF) \leq rg(G)$ und $rg(GF) \leq rg(F)$. Wenn G regulär ist, so gilt $rg(GF) = rg(F)$.

Beweis: Anstelle von Matrizen betrachten wir lineare Abbildungen $f : V \rightarrow W$ und $g : W \rightarrow U$. Sei $\{v_1, \dots, v_n\}$ eine Basis von V , dann ist $\{f(v_1), \dots, f(v_n)\}$ ein Erzeugendensystem von $\text{Im}(f)$, also $\dim \text{Im}(f) \leq \dim V$ und ebenso folgt $\dim g(f(V)) \leq \dim f(V)$, also $\text{rg}(GF) \leq \text{rg}(F)$. Weiter ist $\text{Im}(g \circ f)$ in $\text{Im}(g)$ enthalten, also ist $\dim \text{Im}(g \circ f) \leq \dim \text{Im}(g)$, also $\text{rg}(GF) \leq \text{rg}(G)$.

Wenn g ein Isomorphismus ist, so ist $\dim T = \dim g(T)$ für jeden Unterraum $T \subseteq W$, also ist $\dim \text{Im}(g \circ f) = \dim \text{Im}(f)$. \square

Als nächstes wollen wir eine Beziehung zu den elementaren Zeilenoperationen des Gaußschen Algorithmus herstellen. Wir betrachten die folgenden sogenannten Elementarmatrizen aus M_{nn} :

$$M(i, r) = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ & & \dots & \\ 0 & \dots & r & 0 \\ 0 & \dots & & 0 & 1 \end{pmatrix}, \quad A(i, j) = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ & & \ddots & \\ 0 & 1 & \dots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}.$$

Dabei ist $r \in R$, $r \neq 0$, in $M(i, r)$ steht diese Zahl in der i -ten Zeile und i -ten Spalte, in $A(i, j)$ steht die Eins außerhalb der Diagonalen in der i -ten und j -ten Zeile und Spalte. Der Rest sind Nullen.

Sei nun F eine Matrix aus M_{nn} , dann stimmt, wie man durch Nachrechnen findet, die Matrix $M(i, r)F$ bis auf die i -te Zeile mit F überein, die i -te Zeile aber ist das r -fache der i -ten Zeile von F . Auch die Matrix $A(i, j)F$ unterscheidet sich von F nur in der i -ten Zeile, hier steht die Summe der i -ten und der j -ten Zeile von F .

Wir sehen also, daß die elementaren Zeilenoperationen als gewisse Matrixmultiplikationen aufgefaßt werden können.

Lemma 3.3.1 *Die Elementarmatrizen sind regulär.*

Beweis: Es ist $M(i, r)^{-1} = M(i, r^{-1})$ und $A(i, j)^{-1} = 2E - A(i, j)$. \square

Wir erhalten damit etwas bereits bekanntes:

Folgerung 3.3.3 *Die elementaren Zeilenoperationen ändern den Rang der Matrix nicht.* \square

Es sei nun F eine reguläre Matrix aus M_{nn} . Wir werden ein Berechnungsverfahren für F^{-1} vorstellen:

Es ist $\text{rg}(F) = n$, also ist die reduzierte Form von F die Einheitsmatrix, d.h. F kann durch elementare Zeilenoperationen z_1, \dots, z_k in E überführt werden. Jeder dieser Zeilenoperation ordnen wir die entsprechende Elementarmatrix Z_i zu, dann ist

$$Z_k \cdots Z_1 F = E.$$

Folglich ist die Matrix $Z_k \cdots Z_1$ zu F invers. Nun können wir das Produkt $Z_k \cdots Z_1$ aber auch als Anwendung elementarer Zeilenoperationen auf die Einheitsmatrix interpretieren:

$$Z_k \cdots Z_1 = Z_k \cdots Z_1 E.$$

Also: Wenn dieselben Zeilenoperationen, die F in die Einheitsmatrix überführen, auf die Einheitsmatrix angewandt werden, erhält man F^{-1} .

Damit man nicht vergißt, welche Operation man auf F angewandt hat, schreibt man am Besten die Einheitsmatrix gleich neben F und wendet den Gaußschen Algorithmus auf die „große“ Matrix an.

Als Beispiel wollen wir die Inverse der allgemeinen 2×2 -Matrix berechnen:

$$\begin{pmatrix} a & b & 1 & 0 \\ c & d & 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & \frac{b}{a} & \frac{1}{a} & 0 \\ 0 & \frac{ad-bc}{a} & -\frac{c}{a} & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & \frac{b}{a} & \frac{1}{a} & 0 \\ 0 & 1 & \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{pmatrix} \rightarrow \\ \begin{pmatrix} 1 & 0 & \frac{d}{D} & -\frac{b}{D} \\ 0 & 1 & -\frac{c}{D} & \frac{a}{D} \end{pmatrix},$$

wobei $D = ad - bc$ ist. Die Inverse existiert also, wenn $D \neq 0$ ist.

3.4 Basiswechsel

Die Zuordnungen

Vektor \rightarrow Koordinaten und

Abbildung \rightarrow Matrix

hängen natürlich von der Wahl der Basen ab. Wir fragen uns also, wie sich die Koordinaten eines Vektors bezüglich verschiedener Basen zueinander verhalten.

Seien also $B = \{v_1, \dots, v_n\}$ und $C = \{w_1, \dots, w_n\}$ Basen von V . Dann existieren Zahlen $r_{ji} \in R$ mit

$$v_i = \sum r_{ji} w_j, \quad i = 1, \dots, n.$$

Wir können dies auch anders interpretieren:

$$id_V(v_i) = v_i = \sum r_{ji} w_j,$$

d.h. die Matrix $A = (r_{ji})$ ist die Darstellungsmatrix der identischen Abbildung bezüglich der Basen B, C .

Wie oben betrachten wir das Diagramm

$$\begin{array}{ccc} V & \xrightarrow{id} & V \\ k_B \downarrow & & \downarrow k_C \\ R^n & \xrightarrow{A} & R^n \end{array}$$

und sehen: Das Koordinatentupel $k_C(v)$ des Vektors v bezüglich der Basis C erhalten wir als

$$k_C(v) = k_C(id(v)) = Ak_B(v),$$

also als Produkt der Matrix A mit dem Koordinatentupel von v bezüglich B .

Beispiel:

Sei $V = \mathbb{R}^3$, die Basis B bestehe aus den Vektoren $b_1 = (1, 1, 1)$, $b_2 = (1, -1, -1)$, $b_3 = (1, 1, -1)$ und C aus den Vektoren $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$. Dann hat die Übergangsmatrix von B zu C die Form

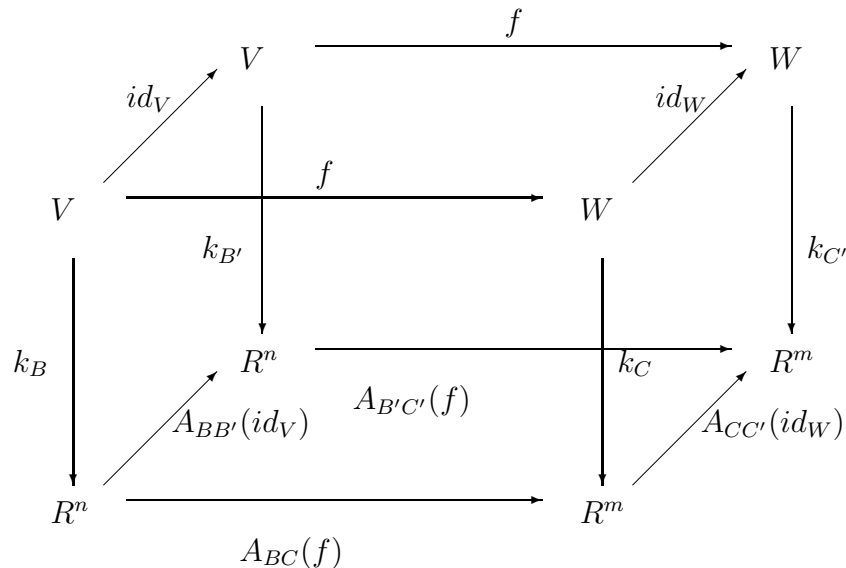
$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & -1 & -1 \end{pmatrix}$$

und das Koordinatentupel von $v = 5b_1 + 7b_2 + 2b_3 = (14, 0, -4)$ bezüglich B ist

$$k_B(v) = \begin{pmatrix} 5 \\ 7 \\ 2 \end{pmatrix}, \text{ w\u00e4hrend das Koordinatentupel von } v \text{ bezüglich } C \text{ gleich } A \begin{pmatrix} 5 \\ 7 \\ 2 \end{pmatrix} = \begin{pmatrix} 14 \\ 0 \\ -4 \end{pmatrix}$$

ist.

Seien nun eine lineare Abbildung $f : V \rightarrow W$ und Basen B, B' von V und Basen C, C' von W gegeben. Um den Zusammenhang von $A_{BC}(f)$ und $A_{B'C'}(f)$ zu erkennen, betrachten wir das folgende Diagramm:



Alle Diagramme auf den Seitenfl\u00e4chen und der Deckfl\u00e4che sind kommutativ, damit ist auch das Diagramm auf der unteren Fl\u00e4che kommutativ und wir erhalten

$$A_{B'C'}(f) = A_{CC'}(id_W)A_{BC}(f)A_{BB'}(id_V)^{-1}.$$

Wir wissen, da\u00df eine beliebige Matrix mit Hilfe von Zeilen- und Spaltenoperationen in die Form

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 1 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 \end{pmatrix}.$$

gebracht werden kann. Daraus erhalten wir das

Lemma 3.4.1 *Sei $f : V \rightarrow W$ eine lineare Abbildung; dann gibt es Basen $\{v_1, \dots, v_n\}$ von V und $\{w_1, \dots, w_m\}$ von W , so daß $f(v_i) = w_i$ für $i = 1, \dots, r$ und $f(v_i) = 0$ für $i > r$ gilt.* \square

Wir wollen nun die sogenannte LU -Zerlegung einer Matrix herleiten. Die Matrix A habe den Rang r . Wir setzen voraus, daß die ersten r Spalten von A linear unabhängig sind, dann hat die reduzierte Form von A die Gestalt

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ & \dots & & \\ 0 & \dots & 1 & 0 \\ & \dots & & \\ 0 & \dots & 0 & 0 \end{pmatrix}.$$

Genauer gesagt: Mit Zeilenoperationen, die nur Vielfache der „oberen“ Zeilen zu unteren addieren, kann A in die Form

$$\begin{pmatrix} a_1 & \star & \dots & \star \\ 0 & a_2 & \dots & \star \\ & \dots & & \\ 0 & \dots & a_r & \star \\ & \dots & & \\ 0 & \dots & 0 & 0 \end{pmatrix}$$

überführt werden (der Stern bedeutet, daß dort irgendeine Zahl steht), also gilt

$$U = M_k \dots M_1 A = \begin{pmatrix} a_1 & \star & \dots & \star \\ 0 & a_2 & \dots & \star \\ & \dots & & \\ 0 & \dots & a_r & \star \\ & \dots & & \\ 0 & \dots & 0 & 0 \end{pmatrix},$$

dies ist eine obere Dreiecksmatrix und die M_i haben die Form

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ & \dots & & \\ & r & \dots & 1 & 0 \\ & \dots & & & \\ 0 & \dots & 0 & 0 \end{pmatrix},$$

dies sind also untere Dreiecksmatrizen, die auf der Diagonalen nur Einsen zu stehen haben. Dann ist auch $L = (M_k \dots M_1)^{-1}$ eine untere Dreiecksmatrix mit Einsen auf der Diagonalen und wir erhalten den

Satz 3.4.1 (LU-Zerlegung) *Unter der genannten Voraussetzung gibt es eine obere Dreiecksmatrix U und eine untere Dreiecksmatrix L , die auf der Diagonalen nur Einsen besitzt, so daß $A = LU$ gilt.* \square

3.5 Idempotente Abbildungen und direkte Summen

Wir betrachten noch eine Reihe spezieller Matrizen und Endomorphismen.

Definition: Sei $f : V \rightarrow V$ eine lineare Abbildung von V in sich, also ein Endomorphismus von V . Die Abbildung f heißt idempotent, wenn $f \circ f = f^2 = f$ gilt, sie heißt involutiv, wenn $f^2 = id$ gilt, und nilpotent, wenn eine natürliche Zahl n existiert, so daß $f^n = o$ ist. Matrizen mit entsprechenden Eigenschaften werden entsprechend benannt.

Zum Beispiel sind die Matrizen $\begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix}$ und $\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$ idempotent, die Matrix $\begin{pmatrix} -1 & 0 \\ -4 & 1 \end{pmatrix}$ ist involutiv und die Matrix $\begin{pmatrix} -2 & 1 \\ -4 & 2 \end{pmatrix}$ ist nilpotent.

Wir betrachten zuerst nilpotente Abbildungen:

Satz 3.5.1 *Wenn $f : V \rightarrow V$ nilpotent ist, so gibt es ein $m \leq \dim V$ mit $f^m = o$.*

Beweis: Zunächst ist $\text{Im}(f) \subset V$ ein echter Unterraum, denn bei $\text{Im}(f) = V$ hätten wir $\dim \text{Ker}(f) = 0$, also wäre f injektiv und niemals nilpotent. Ganz genauso sieht man, daß $\text{Im}(f^2)$ ein echter Unterraum von $\text{Im}(f)$ ist. Insgesamt erhalten wir eine echt absteigende Folge

$$V \supset \text{Im}(f) \supset \text{Im}(f^2) \supset \dots \supset \text{Im}(f^{m-1}) \supset \text{Im}(f^m) = \{o\}$$

von Unterräumen von V , die beim Nullraum endet. Da die Dimension dieser Unteräume sich bei jedem Schritt verkleinert, muß $m \leq \dim V$ sein. \square

Satz 3.5.2 *Wenn f ein nilpotenter Endomorphismus ist, so ist $g = id + f$ ein Isomorphismus.*

Beweis: Sei $f^n = o$, wir setzen $h = id - f + f^2 - \dots + (-1)^{n-1} f^{n-1}$. Dann ist

$$\begin{aligned} gh &= (id + f)(id - f + f^2 - \dots + (-1)^{n-1} f^{n-1}) \\ &= id - f + f^2 - \dots + (-1)^{n-1} f^{n-1} + f - f^2 + \dots + (-1)^{n-2} f^{n-1} \\ &= id. \square \end{aligned}$$

Beispiel:

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 0 & 2 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 8 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -2 & 5 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{pmatrix}$$

Wir betrachten nun idempotente Abbildungen. Typische Beispiele sind Projektionen: Sei $V = U \oplus W$ eine direkte Summe der Unterräume U und W . Wir konstruieren folgendermaßen einen Endomorphismus von V : für $v = u + w$ ($u \in U, w \in W$) setzen wir $p(v) = u$. Dann ist $\text{Im}(p) = U$ und für $u \in U$ ist $u = u + o$ die einzige Zerlegung von u in Summanden aus U und W , also ist $p(u) = u$, d.h. $p^2 = p$. Wir nennen p die Projektion von V auf U (in Richtung von W). Es gilt $\text{Ker}(p) = W$, wir sehen, daß das kein Zufall ist:

Satz 3.5.3 Wenn $f : V \rightarrow V$ idempotent ist, so gilt $V = \text{Ker}(f) \oplus \text{Im}(f)$.

Beweis: Sei $v \in V$, dann liegt $f(v)$ in $\text{Im}(f)$ und $v - f(v)$ in $\text{Ker}(f)$, da $f(v - f(v)) = f(v) - f(v) = o$ ist. Also ist V die Summe der Unterräume $\text{Ker}(f)$ und $\text{Im}(f)$. Sei nun ein Vektor v sowohl in $\text{Ker}(f)$ als auch in $\text{Im}(f)$ enthalten, dann ist $f(v) = o$ und es gibt einen Vektor w mit $v = f(w)$. Dann gilt aber $o = f(v) = f(f(w)) = f(w) = v$. \square

Satz 3.5.4 Wenn $f : V \rightarrow V$ idempotent ist, so ist $g = id - 2f$ involutiv. Wenn g involutiv ist, so ist $f = \frac{1}{2}(id - g)$ idempotent. Wenn f idempotent ist, so ist auch $(id - f)$ idempotent und es gilt $(id - f)f = o$.

Den Beweis möge der Leser durch einfaches Nachrechnen führen. \square

Satz 3.5.5 Seien $f, g : V \rightarrow V$ idempotente Abbildungen mit $f + g = id$. Dann ist $V = \text{Im}(f) \oplus \text{Im}(g)$.

Beweis: Wir zeigen $\text{Im}(g) = \text{Ker}(f)$. Es ist $g = id - f$, also $gf = fg = o$. Sei $g(v) \in \text{Im}(g)$, dann ist $f(g(v)) = o$, also ist $g(v) \in \text{Ker}(f)$. Sei umgekehrt $v \in \text{Ker}(f)$, dann ist $f(v) = o$, also $g(v) = v - f(v) = v$, d.h. v liegt in $\text{Im}(g)$. \square

Wenn umgekehrt V die direkte Summe von Unterräumen U und W ist, so haben wir zwei Projektionen f, g von V mit $\text{Im}(f) = U$ und $\text{Im}(g) = W$ und für $v = u + w$ mit $u \in U$, $w \in W$ gilt $f(v) = u$, $g(v) = w$, also $(f + g)(v) = u + w = v$, d.h. $f + g = id$.

Satz 3.5.6 Seien $f_1, \dots, f_k : V \rightarrow V$ idempotente Abbildungen, für die $f_i \circ f_j = o$ für $i \neq j$ sowie $f_1 + \dots + f_k = id$ gilt. Dann ist

$$V = \text{Im}(f_1) \oplus \dots \oplus \text{Im}(f_k).$$

Beweis: Sei v ein beliebiger Vektor aus V , dann ist $v = id(v) = (f_1 + \dots + f_k)(v) = f_1(v) + \dots + f_k(v)$, also $\text{Im}(f_1) + \dots + \text{Im}(f_k) = V$. Sei weiter v ein Vektor, der in $\text{Im}(f_i)$ und in der Summe der $\text{Im}(f_j)$ ($j \neq i$) liegt. Dann gibt es w_j , so daß

$$v = f_i(w_i) = \sum_{j \neq i} f_j(w_j)$$

gilt. Dann ist $f_i(v) = f_i^2(w_i) = f_i(w_i) = v$ und $f_i(v) = \sum f_i(f_j(w_j)) = o$, also $v = o$. \square

Kapitel 4

Affine Geometrie

4.1 Affine Räume und Unterräume

In diesem Abschnitt wollen wir uns mit einfachen geometrischen Objekten, wie Punkten, Geraden, Ebenen beschäftigen.

Wenn in der Ebene ein Koordinatensystem gegeben ist, so kann man Punkte durch ihre Koordinaten und Geraden z.B. durch eine Gleichung $y = mx + n$ beschreiben. Wir wollen diese Begriffe im folgenden präzisieren.

Definition: Sei A eine Menge und V ein R -Vektorraum. Das Paar (A, V) heißt affiner Raum, wenn eine Operation $+ : A \times V \rightarrow A$ gegeben ist, die dem Paar (P, v) mit $P \in A$, $v \in V$ das Element $P + v$ zuordnet, so daß

1. $(P + v) + w = P + (v + w)$ für alle $P \in A$, $v, w \in V$ gilt und
2. zu beliebigen $P, Q \in A$ ein eindeutig bestimmter Vektor v existiert, so daß $Q = P + v$ ist (dieser Vektor heißt der Verbindungsvektor von P und Q und wird mit \overrightarrow{PQ} bezeichnet).

Die Elemente von A nennen wir dann Punkte.

Manchmal sagen wir auch, daß A ein affiner Raum ist, wenn klar ist, welches der zugehörige Vektorraum sein soll. Dieser Vektorraum ist durch A eindeutig bestimmt: Er besteht aus der Menge aller Verbindungsvektoren der Punkte aus A .

Beispiele:

1. Sei A die Menge der „Punkte“ einer Ebene und V der Vektorraum aller Verschiebungen der Ebene in sich. Wenn P ein Punkt und v eine Verschiebung ist, so sei $P + v$ das Ergebnis der Verschiebung v auf P . Dann ist die obige Bedingung 1 erfüllt und zu zwei Punkten gibt es genau eine Verschiebung der Ebene, die den ersten in den zweiten überführt.
2. Sei V ein Vektorraum, wir setzen $A = V$, die Addition von Punkt und Vektor definieren wir durch die Addition in V . Dann ist (V, V) ein affiner Raum.
3. Sei S ein beliebiges Gleichungssystem und H das zugehörige homogene Gleichungssystem, dann ist $(LM(S), LM(H))$ ein affiner Raum.

Wir wissen nun, was Punkte sind, nämlich Elemente eines affinen Raums. Wir präzisieren nun solche Begriffe wie „Gerade“, „Ebene“, ...

Definition: Sei (A, V) ein affiner Raum. Eine nichtleere Teilmenge H von A heißt affiner Unterraum von A , wenn es einen Punkt $P \in H$ und einen Unterraum U von V gibt, daß

$$H = P + U = \{Q \mid \text{es gibt ein } u \in U \text{ mit } Q = P + u\}$$

ist.

Lemma 4.1.1 Sei $H = P + U$ ein affiner Unterraum von (A, V) . Dann ist $H = Q + U$ für alle $Q \in H$. Weiter gilt: aus $H = P + U = Q + W$ folgt $U = W$.

Beweis: Sei $Q \in H$, also $Q = P + u$ für ein $u \in U$, dann ist $Q + U = P + u + U = P + U = H$. Wenn $P + U = Q + W$ ist, so liegt Q auch in $P + U$, also ist $P + U = Q + U = Q + W$. Sei nun $u \in U$, dann gibt es ein $w \in W$ mit $Q + u = Q + w$, da der Verbindungsvektor von Q und $Q + u$ eindeutig bestimmt ist, gilt $u = w$, d.h. u liegt in W , also gilt $U \subseteq W$, analog folgt $W \subseteq U$, also $U = W$. \square

Definition: Sei $H = P + U$ ein affiner Unterraum, wir setzen $\dim H = \dim U$.

Nulldimensionale Unterräume bestehen also aus einem einzigen Punkt, eindimensionale Unterräume nennen wir „Geraden“, zweidimensionale „Ebenen“ usw.

Wir sagen, daß die Punkte $P_0, P_1, \dots, P_k \in A$ sich in allgemeiner Lage befinden, wenn es keinen $(k - 1)$ -dimensionalen Unterraum von A gibt, der sie enthält.

Zum Beispiel sind zwei verschiedene Punkte in allgemeiner Lage, drei Punkte sind in allgemeiner Lage, wenn sie nicht auf einer Geraden liegen usw.

Satz 4.1.1 Die Punkte P_0, \dots, P_k sind genau dann in allgemeiner Lage, wenn die Vektoren $v_1 = \overrightarrow{P_0P_1}, \dots, v_k = \overrightarrow{P_0P_k}$ linear unabhängig sind.

Beweis: Seien die Punkte P_0, \dots, P_k in allgemeiner Lage. Wir setzen $H = P_0 + L\{v_1, \dots, v_k\}$, dann ist $P_0 \in H$, die $P_i = P_0 + v_i$ liegen auch in H und es ist $\dim H \leq k$. Wenn $\dim H < k$ wäre, so wären die Punkte P_0, \dots, P_k nicht in allgemeiner Lage, folglich ist $\dim H = k$, d.h. $\{v_1, \dots, v_k\}$ ist eine linear unabhängige Menge.

Sei $\{v_1, \dots, v_k\}$ linear unabhängig. Wir nehmen an, daß die Punkte P_0, \dots, P_k in einem Unterraum $H = Q + U$ mit $\dim U \leq k - 1$ liegen. Es ist $P_0 \in H$, also $H = P_0 + U$ und damit liegen die v_i in U , also ist $\dim U \geq k$. \square

Lemma 4.1.2 Seien $P_0, \dots, P_k \in A$ Punkte in allgemeiner Lage, dann gibt es einen eindeutig bestimmten k -dimensionalen Unterraum H von A , der P_0, \dots, P_k enthält.

Beweis: Die Existenz ist klar: $H = P_0 + L\{\overrightarrow{P_0P_1}, \dots, \overrightarrow{P_0P_k}\}$ hat die Dimension k . Sei umgekehrt $H = P + U = P_0 + U$ irgendein Unterraum, der die P_i enthält, dann liegen die Vektoren $\overrightarrow{P_0P_i}$ in U , also ist $L\{\overrightarrow{P_0P_1}, \dots, \overrightarrow{P_0P_k}\}$ in U enthalten und beide Räume haben dieselbe Dimension, sind also gleich. \square

Definition: Sei $H = P + U$ ein affiner Unterraum von A und $\{b_1, \dots, b_k\}$ eine Basis von U , dann heißt $\{P, b_1, \dots, b_k\}$ ein Koordinatensystem von H .

Wenn ein Koordinatensystem $\{P, b_1, \dots, b_k\}$ von $H = P + U$ gegeben ist, so gibt es für jeden Punkt Q von H eindeutig bestimmte Zahlen r_1, \dots, r_k mit $Q = P + \sum r_i b_i$, diese „Punktkoordinaten“ fassen wir in einem $(k + 1)$ -tupel $(1, r_1, \dots, r_k)$ zusammen (die führende 1 soll anzeigen, daß es sich um Koordinaten eines Punkts handelt).

Zum Vektor $u \in U$ haben wir Zahlen s_1, \dots, s_k mit $u = \sum s_i b_i$, diese „Vektorkoordinaten“ fassen wir im $(k + 1)$ -tupel $(0, s_1, \dots, s_k)$ zusammen.

Die Operationen im affinen Raum spiegeln sich wie folgt in den Koordinaten wider:

Lemma 4.1.3 *Sei $\{P, b_1, \dots, b_k\}$ ein Koordinatensystem von $H = P + U$, das Koordinatentupel des Punkts $Q \in H$ sei $(1, q_1, \dots, q_k)$, das von S sei $(1, s_1, \dots, s_k)$ und das des Vektors $v \in U$ sei $(0, r_1, \dots, r_k)$. Dann ist das Koordinatentupel von $Q + v$ gleich $(1, q_1 + r_1, \dots, q_k + r_k)$ und der Verbindungsvektor von Q nach S hat die Koordinaten $(0, s_1 - q_1, \dots, s_k - q_k)$.*

Den Beweis überlassen wir dem Leser. □

Sei nun $\{P, e_1, \dots, e_n\}$ ein Koordinatensystem des affinen Raums A selbst. Seien Matrizen $(a_{ij}) \in M_{mn}$ und $(b_i) \in M_{m1}$ gegeben, dann ist die Menge H der Punkte X mit dem Koordinatentupel $(1, x_1, \dots, x_n)$, für die

$$\sum a_{ij} x_j = b_i, \quad i = 1, \dots, m$$

gilt, ein affiner Unterraum von A (oder leer). In der Tat: Sei

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix} + \sum_{i=1}^{n-r} \begin{pmatrix} c_{1i} \\ \vdots \\ c_{ni} \end{pmatrix} t_i$$

die Lösungsmenge des Gleichungssystems, dann gilt

$$(1, x_1, \dots, x_n) = (1, p_1, \dots, p_n) + \sum (0, c_{1i}, \dots, c_{ni}) t_i$$

oder

$$H = (P + \sum p_i e_i) + L\left\{\sum_j c_{j1} e_j, \dots, \sum_j c_{jn} e_j\right\}.$$

Beispiel:

Wir betrachten den R^3 mit dem Koordinatensystem $\{(0, 0, 0), e_1, e_2, e_3\}$ und das Gleichungssystem $x_1 + x_2 + x_3 = 1$. Der Lösungsraum des zugehörigen homogenen Systems ist

$$U = LM(x_1 + x_2 + x_3 = 0) = L\{(-1, 0, 1), (0, -1, 1)\}$$

und eine spezielle Lösung des inhomogenen Systems ist $(1, 0, 0)$, also

$$H = (1, 0, 0) + L\{e_3 - e_1, e_3 - e_2\}.$$

Wir werden nun sehen, daß jeder affine Unterraum durch ein lineares Gleichungssystem beschrieben werden kann:

Satz 4.1.2 Sei H ein affiner Unterraum von A , $\{P, e_1, \dots, e_n\}$ ein Koordinatensystem von A . Dann existiert ein lineares Gleichungssystem $\sum a_{ij}x_j = b_i$, $i = 1, \dots, m$, so daß der Punkt X genau dann in H liegt, wenn sein Koordinatentupel $(1, x_1, \dots, x_n)$ das Gleichungssystem erfüllt.

Beweis: Wir wählen ein Koordinatensystem $\{Q, b_1, \dots, b_k\}$ von H . Dann gilt für $X \in A$, daß X genau dann in H liegt, wenn es Zahlen r_1, \dots, r_k gibt, so daß $X = Q + \sum r_i b_i$ ist. Wir stellen dies im Koordinatensystem $\{P, e_1, \dots, e_n\}$ dar: Die Koordinatentupel von b_i , X und Q seien $(0, b_{1i}, \dots, b_{ni})$, $(1, x_1, \dots, x_n)$ bzw. $(1, q_1, \dots, q_n)$. Dann bedeutet die obige Relation, das

$$\begin{aligned} b_{11}r_1 + \dots + b_{1k}r_k &= x_1 - q_1 \\ &\dots \\ b_{n1}r_1 + \dots + b_{nk}r_k &= x_n - q_n \end{aligned}$$

genau dann eine eindeutig bestimmte Lösung (r_1, \dots, r_k) besitzt, wenn X in H liegt. Dies ist genau dann der Fall, wenn der Rang der Koeffizientenmatrix gleich k ist. Das heißt, daß die reduzierte Form der Koeffizientenmatrix folgendermaßen aussieht:

$$\begin{pmatrix} 1 & 0 & \dots & f_1(x) \\ & 1 & 0 & \dots & f_2(x) \\ & & \dots & & \\ & & & 1 & \dots & f_k(x) \\ & & & 0 & & f_{k+1}(x) \\ & & & & \dots & \\ & & & & & f_n(x) \end{pmatrix}$$

Der Rang dieser Matrix ist genau dann gleich k , wenn die $n - k$ Gleichungen

$$\begin{aligned} f_{k+1}(x) &= 0 \\ &\dots \\ f_n(x) &= 0 \end{aligned}$$

erfüllt sind. Dies ist unser gesuchtes (inhomogenes) Gleichungssystem.

Beispiel:

Sei $H = (1, 1, 1) + \mathcal{L}(e_1 + e_2, e_2 + e_3)$ im affinen Raum R^3 . Der Punkt (x_1, x_2, x_3) liegt genau dann in H , wenn $(x_1 - 1, x_2 - 1, x_3 - 1)$ in $\mathcal{L}((1, 1, 0), (0, 1, 1))$ liegt, also wenn

$$\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} r_1 + \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} r_2 = \begin{pmatrix} x_1 - 1 \\ x_2 - 1 \\ x_3 - 1 \end{pmatrix}.$$

Wir wenden den Gaußschen Algorithmus auf die Koeffizientenmatrix an:

$$\begin{pmatrix} 1 & 0 & x_1 - 1 \\ 1 & 1 & x_2 - 1 \\ 0 & 1 & x_3 - 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & x_1 - 1 \\ 0 & 1 & x_2 - x_1 \\ 0 & 0 & x_3 - x_2 + x_1 - 1 \end{pmatrix}$$

also liegt der Punkt genau dann in H , wenn $x_3 - x_2 + x_1 = 1$ ist.

Als nächstes wollen wir uns mit dem Durchschnitt affiner Unterräume befassen. Seien $H_1 = P_1 + U_1$ und $H_2 = P_2 + U_2$ zwei affine Unterräume eines affinen Raums A .

Lemma 4.1.4 *Der Durchschnitt $H_1 \cap H_2$ ist leer oder gleich $P + U_1 \cap U_2$, wobei P ein beliebiger Punkt von $H_1 \cap H_2$ ist.*

Beweis: Sei $P \in H_1 \cap H_2$, dann ist $H_1 = P + U_1$ und $H_2 = P + U_2$. Ein Punkt X liegt genau dann im Durchschnitt, wenn es Vektoren $u_1 \in U_1, u_2 \in U_2$ gibt, so daß $X = P + u_1 = P + u_2$ ist, d.h. es ist $u_1 = u_2 \in U_1 \cap U_2$. \square

Wenn die Koordinaten des Punktes $X \in H_1$ bzw. H_2 bezüglich eines in A gewählten Koordinatensystems durch die Gleichungssysteme

$$MX = B \text{ bzw. } NX = C$$

beschrieben werden, so sind die Koordinaten von Punkten aus $H_1 \cap H_2$ gerade die Lösungen von

$$\begin{pmatrix} M \\ N \end{pmatrix} X = \begin{pmatrix} B \\ C \end{pmatrix}$$

denn X liegt genau dann in $H_1 \cap H_2$, wenn $MX = B$ und $NX = C$ ist.

Lemma 4.1.5 *Ein k -dimensionaler Unterraum H eines n -dimensionalen affinen Raums ist als Durchschnitt von $n - k$ ($n - 1$)-dimensionalen Unterräumen (sog. Hyperebenen) darstellbar.*

Beweis: Wir wählen ein Gleichungssystem mit $n - k$ Gleichungen, das die Koordinaten der Punkte von H beschreibt. Jede einzelne Gleichung hat als Lösungsmenge die Koordinaten der Punkte einer Hyperebene, der Durchschnitt dieser Hyperebenen ist gerade H . \square

Definition: Zwei affine Unterräume $H_1 = P_1 + U_1$ und $H_2 = P_2 + U_2$ heißen parallel, wenn $U_1 \subseteq U_2$ oder $U_2 \subseteq U_1$ gilt. Wenn sie nicht parallel sind und ihr Durchschnitt leer ist, so heißen sie windschief.

Satz 4.1.3 *Sei $\dim A = 3$ und H_1, H_2 zwei Geraden in A . Dann sind die Geraden entweder parallel oder windschief oder sie schneiden sich.*

Beweis: Wir wählen ein Koordinatensystem und stellen H_1 und H_2 durch zwei Gleichungssysteme mit je zwei Gleichungen dar:

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1$$

$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2$$

und

$$a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3$$

$$a_{41}x_1 + a_{42}x_2 + a_{43}x_3 = b_4$$

Alle vier Gleichungen zusammen beschreiben den Durchschnitt beider Geraden. Es sei r der Rang der „kleinen“ Koeffizientenmatrix und R der Rang der „großen“ Koeffizientenmatrix. Es ist $2 \leq r \leq R \leq 4$.

1. Fall: $r = R = 2$, dann ist $H_1 = H_2$.
2. Fall: $r = 2, R = 3$, dann ist $U_1 = U_2$ und der Durchschnitt der Geraden ist leer, also sind sie parallel.
3. Fall: $r = R = 3$, dann hat das Gleichungssystem eine eindeutig bestimmte Lösung, also schneiden sich die Geraden in einem Punkt.
4. Fall: $r = 3, R = 4$, dann ist $U_1 \neq U_2$ und der Durchschnitt ist leer, also sind die Geraden windschief. \square

Satz 4.1.4 H_1 und H_2 seien Hyperebenen in einem n -dimensionalen affinen Raum, dann tritt einer der folgenden Fälle auf:

1. $H_1 = H_2$,
2. H_1 und H_2 sind parallel,
3. $H_1 \cap H_2$ hat die Dimension $n - 2$. \square

Definition: Seien $H_1, H_2 \subseteq A$ Unterräume, $H = H_1 \vee H_2$ sei der kleinste Unterraum, der H_1 und H_2 umfaßt, er heißt der Verbindungsraum von H_1 und H_2 .

Lemma 4.1.6 Sei $H_1 = P_1 + U_1, H_2 = P_2 + U_2$. Wenn der Durchschnitt von H_1 und H_2 nichtleer ist, so liegt der Verbindungsvektor von P_1 und P_2 in $U_1 + U_2$.

Beweis: Es sei P ein Punkt von $H_1 \cap H_2$, dann ist $P_1 = P + u_1, P_2 = P + u_2$ mit $u_1 \in U_1, u_2 \in U_2$, also ist $\overrightarrow{P_1 P_2} = \overrightarrow{P_1 P} + \overrightarrow{P P_2} = -u_1 + u_2 \in U_1 + U_2$. \square

Satz 4.1.5 $H_1 \vee H_2 = P_1 + \mathcal{L}(\overrightarrow{P_1 P_2}) + (U_1 + U_2)$.

Beweis: Sowohl H_1 als auch H_2 sind in $P_1 + \mathcal{L}(\overrightarrow{P_1 P_2}) + (U_1 + U_2)$ enthalten. Sei $H = H_1 \vee H_2 = P_1 + U = P_2 + U$. Dann ist $U_1 \subseteq U, U_2 \subseteq U, \overrightarrow{P_1 P_2}$ liegt in U , also ist $P_1 + \mathcal{L}(\overrightarrow{P_1 P_2}) + (U_1 + U_2) \subseteq H$. \square

Folgerung 4.1.1 Wenn $H_1 \cap H_2$ nichtleer ist, so ist

$$\dim H_1 \vee H_2 = \dim H_1 + \dim H_2 - \dim H_1 \cap H_2.$$

Wenn $H_1 \cap H_2$ leer ist, so ist

$$\dim H_1 \vee H_2 = \dim H_1 + \dim H_2 - \dim U_1 \cap U_2 + 1. \square$$

4.2 Affine Abbildungen

Definition: Seien $(A, V), (A', V')$ affine Räume, dann heißt $F : A \rightarrow A'$ eine affine Abbildung, wenn eine lineare Abbildung $f : V \rightarrow V'$ existiert, so daß $F(P + v) = F(P) + f(v)$ ist.

Beispiele:

1. Parallelprojektion: Sei (A, V) ein affiner Raum und $H = P + U \subseteq A$ ein affiner Unterraum, U' ein Komplement von U in V , d.h. $V = U \oplus U'$. Sei Q ein Punkt von A , $Q = P + u + u'$, wo $u \in U$ und $u' \in U'$ ist. Wir setzen $F(Q) = P + u$ und $f(u + u') = u$, dann ist f linear und es gilt $F(Q + w) = F(Q) + f(w)$, wie man sofort sieht.

2. Translation: Seien Punkte P, Q gegeben, wir setzen $T(P + v) = Q + v$, $t = id$, dies ist die Verschiebung des Raums um den Verbindungsvektor von P nach Q .

Satz 4.2.1 *Sei $H \subseteq A$ ein Unterraum und $F : A \rightarrow A'$ eine affine Abbildung, dann ist $F(H) \subseteq A'$ ein affiner Unterraum und $\dim F(H) \leq \dim H$. Wenn H und H' parallel sind, so sind auch $F(H)$ und $F(H')$ parallel.*

Beweis: Sei $H = P + U$, dann ist $F(H) = F(P) + f(U)$ ein affiner Unterraum und $\dim f(U) \leq \dim U$. Sei noch $H' = P' + U'$ parallel zu H , etwa $U \subseteq U'$, dann ist auch $f(U) \subseteq f(U')$, also sind die Bilder auch parallel. \square

Dem Leser überlassen wir die folgende Aussage zum Beweis:

Folgerung 4.2.1 *Das Bild einer Geraden ist eine Gerade oder ein Punkt. Wenn H und H' parallele Geraden und $F(H)$ ein Punkt ist, so ist $F(H')$ auch ein Punkt.* \square

Seien nun (A, V) und (A', V') affine Räume und $F : A \rightarrow A'$ eine affine Abbildung. Wir wählen Koordinatensysteme:

$$A = P + \mathcal{L}(b_1, \dots, b_n) \text{ und } A' = P' + \mathcal{L}(b'_1, \dots, b'_m).$$

Wir wollen der Abbildung F eine Matrix zuordnen, die F eindeutig bestimmt.

Sei Q ein beliebiger Punkt von A mit dem Koordinatentupel $(1, r_1, \dots, r_n)$, d.h. $Q = P + \sum r_i b_i$, dann ist

$$F(Q) = F(P) + f\left(\sum r_i b_i\right) = F(P) + \sum r_i f(b_i),$$

also ist F durch $F(P)$ und $f(b_1), \dots, f(b_n)$ eindeutig bestimmt, also durch die Darstellungsmatrix (f_{ji}) der Abbildung f und das Koordinatentupel $(1, s_1, \dots, s_m)$ des Punktes $F(P)$. Wir schreiben dies alles in die folgende Matrix

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ s_1 & f_{11} & \dots & f_{1n} \\ & & \dots & \\ s_m & f_{m1} & \dots & f_{mn} \end{pmatrix}$$

in deren Spalten die Koordinaten von $F(P), f(b_1), \dots, f(b_n)$ stehen.

Wir wollen nun nachzuweisen, daß bei einer derartigen Matrixzuordnung das Produkt affiner Abbildungen dem Matrixprodukt der Darstellungsmatrizen der affinen Abbildungen entspricht.

Wir betrachten drei affine Räume A , B , C und zwei affine Abbildungen $F : A \rightarrow B$, $G : B \rightarrow C$. Wir wählen Koordinatensysteme (P, b_1, \dots, b_n) von A , (Q, c_1, \dots, c_m) von B und (R, d_1, \dots, d_l) von C . Es sei

$$\begin{aligned} F(P) &= Q + \sum f_j c_j, & f(b_i) &= \sum f_{ji} c_j, \\ G(Q) &= R + \sum g_k d_k, & g(c_j) &= \sum g_{kj} d_k. \end{aligned}$$

Dann ist

$$\begin{aligned} G \circ F(P) &= G(Q) + g\left(\sum f_j c_j\right) \\ &= R + \sum g_k d_k + \sum f_j \sum g_{kj} d_k \\ &= R + \sum_k \left(g_k + \sum_j g_{kj} f_j\right) d_k \end{aligned}$$

Die Ausdrücke in den Klammern sind gerade die Komponenten der ersten Spalte der Produktmatrix

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ g_1 & g_{11} & \dots & g_{1m} \\ & & \dots & \\ g_l & g_{l1} & \dots & g_{lm} \end{pmatrix} \begin{pmatrix} 1 & 0 & \dots & 0 \\ f_1 & f_{11} & \dots & f_{1n} \\ & & \dots & \\ f_m & f_{m1} & \dots & f_{mn} \end{pmatrix}.$$

Die Koordinaten y_j des Punkts $F(X)$ kann man aus den Koordinaten von X wie folgt berechnen: Sei der Einfachheit halber $F : A \rightarrow A$, $X = P + \sum x_i b_i$, dann ist

$$F(X) = F(P) + \sum x_i f(b_i) = Q + \sum (f_j + \sum f_{ji} x_i) b_j = P + \sum y_j b_j$$

also

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ f_1 & f_{11} & \dots & f_{1n} \\ & & \dots & \\ f_m & f_{m1} & \dots & f_{mn} \end{pmatrix} \begin{pmatrix} 1 \\ x_1 \\ \dots \\ x_n \end{pmatrix} = \begin{pmatrix} 1 \\ y_1 \\ \dots \\ y_n \end{pmatrix}.$$

Sei zum Beispiel d die Drehung um den Ursprung um 90 Grad und v die Verschiebung um den Vektor $(1,1)$. Dazu gehören die Matrizen

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \text{ bzw. } V = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

Zum Produkt $d \circ v$ gehört die Matrix

$$DV = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 0 & -1 \\ 1 & 1 & 0 \end{pmatrix}$$

also ist

$$dv(X) = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 0 & -1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ -1 - y \\ 1 + x \end{pmatrix}.$$

Wir fragen abschließend, ob dv einen Fixpunkt besitzt (d.h. $dv(X) = X$). Wir finden, daß der Punkt $(-1, 0)$ fest bleibt. (Machen Sie eine Skizze!)

Seien A, B, T drei Punkte auf einer Geraden Und es sei $T = A + k \overrightarrow{AB}$, dann heißt die Zahl k das Teilverhältnis der Punkte A, B, T , dies wird mit dem Symbol (ATB) bezeichnet.

Sei nun F eine affine Abbildung und f die zugehörige lineare Abbildung. Dann gilt

$$F(T) = F(A + k \overrightarrow{AB}) = F(A) + kf(\overrightarrow{AB}) = F(A) + k F(A)\overrightarrow{F(B)},$$

das heißt $k = (F(A), F(T)F(B))$, das Teilverhältnis dreier Punkte bleibt bei affinen Abbildungen erhalten.

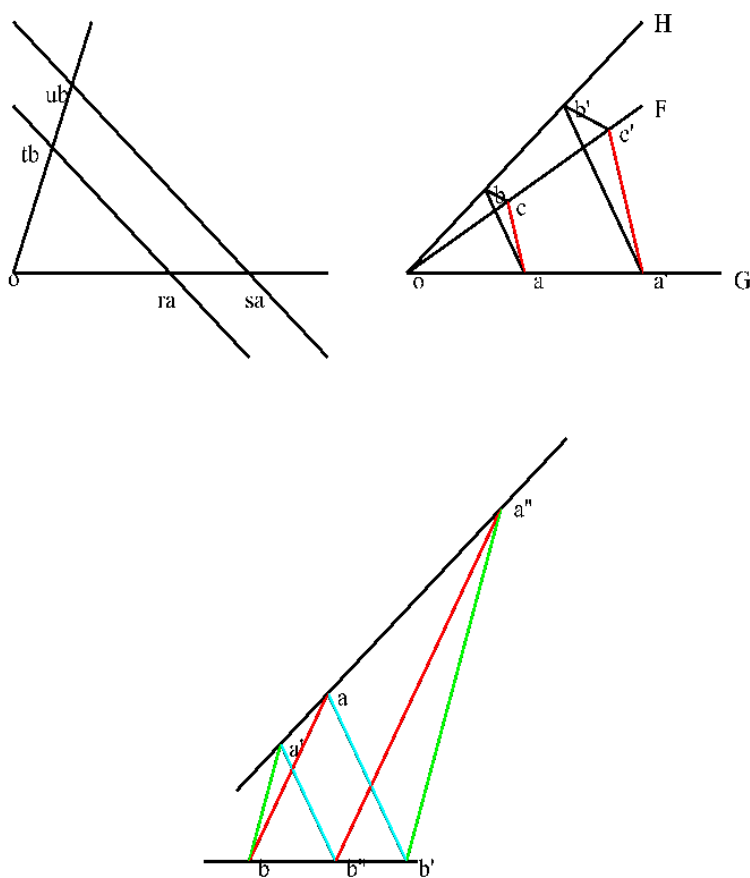
4.3 Zweidimensionale Geometrie I

Die folgenden Ausführungen lehnen sich an Koecher/Krieg, Ebene Geometrie an. Wir betrachten \mathbb{R}^2 als affinen Raum, in diesem Abschnitt bezeichnen wir mit r, s, t, u Elemente von $\mathbb{R} \setminus \{0\}$, a, b, c, d bezeichnen Punkte, große Buchstaben Geraden und mit $a \vee b$ wird die Verbindungsgerade von a und b bezeichnet.

Satz 4.3.1 Wir führen die Abkürzung $[a, b] = a_1b_2 - a_2b_1$ ein, dann ist der Schnittpunkt der Geraden $G = \{a + rb \mid r \in \mathbb{R}\}$ und $H = \{c + sd \mid s \in \mathbb{R}\}$ durch

$$\frac{1}{[b, d]} ([c, d]a - [a, b]d)$$

gegeben.



Satz 4.3.2 (Strahlensatz) Die beiden folgenden Aussagen sind äquivalent:

1. Die Verbindungsgeraden $ra \vee tb$ und $sa \vee ub$ sind parallel.
2. $ru = st$

Beweis: Die Verbindungsgeraden sind genau dann parallel, wenn $ra - tb$ und $sa - ub$ linear abhängig sind, wenn also $\frac{r}{s} = \frac{t}{u}$ gilt. \square

Satz 4.3.3 (Desargues) *Seien F, G, H paarweise verschiedene Geraden durch den Punkt o und $a, a' \in F, b, b' \in G$ und $c, c' \in H$ mit $a \vee b \parallel a' \vee b'$ und $b \vee c \parallel b' \vee c'$, dann gilt auch $a \vee c \parallel a' \vee c'$.*

Beweis: Nach Voraussetzung ist $a' = ra, b' = sb, c' = tc$, die Voraussetzungen und die Behauptung sind nach dem Strahlensatz gleichwertig mit $s = r, t = s$ sowie $t = r$ \square

Satz 4.3.4 (Pappus) *Seien F, G verschiedene Geraden, die sich in o schneiden, und seien $a, a', a'' \in F \setminus G$ und $b, b', b'' \in G \setminus F$ jeweils paarweise verschieden. Wenn $a \vee b' \parallel a' \vee b$ und $a' \vee b'' \parallel a'' \vee b'$, so gilt $a \vee b \parallel a'' \vee b''$.*

Beweis: Nach Voraussetzung sind a, b linear unabhängig, wir setzen $a' = ra, a'' = sa, b' = tb, b'' = ub$. Aus dem Strahlensatz folgt $u = rt$ und $s = rt$, also $u = s$. Dann gilt aber auch $a \vee b \parallel a'' \vee b''$. \square

Kapitel 5

Linearformen

Die Menge aller linearer Abbildung $\text{Hom}(V, W)$ eines Vektorraums V in einen Vektorraum W ist selbst ein Vektorraum, speziell ist $V^* = \text{Hom}(V, R)$ ein Vektorraum, der dieselbe Dimension wie V besitzt. Wir nennen die Elemente von V^* Linearformen auf V .

Wir wiederholen:

Wenn l und l' Linearformen auf V sind, so ist für $v \in V$ und $r \in R$ stets $(l + l')(v) = l(v) + l'(v)$ und $(rl)(v) = rl(v)$. Die Linearformen $\{l_1, \dots, l_k\}$ auf V sind genau dann linear unabhängig, wenn aus $\sum r_i l_i = 0$ folgt, daß alle r_i Null sind, also: Wenn für alle $v \in V$ die Relation $\sum r_i l_i(v) = 0$ gilt, so ist $r_i = 0$ für $i = 1, \dots, k$.

Zur Abkürzung hat sich die folgende Funktion δ eingebürgert:

$$\delta_{ij} = \begin{cases} 0 & \text{für } i \neq j, \\ 1 & \text{für } i = j, \end{cases}$$

sie wird als „Kroneckersymbol“ bezeichnet.

Satz 5.0.5 *Seien $\{v_1, \dots, v_k\}$ linear unabhängige Vektoren aus V , dann gibt es linear unabhängige Linearformen $l_1, \dots, l_k \in V^*$, so daß $l_i(v_j) = \delta_{ij}$.*

Beweis: Sei $V = \mathcal{L}(v_1, \dots, v_k) \oplus U$ eine direkte Summe, dann hat jeder Vektor $v \in V$ eine eindeutig bestimmte Darstellung $v = \sum r_i v_i + w$, wobei w in U liegt. Wir setzen $l_i(v) = r_i$, dann ist $l_i(v_j) = 0$ für $i \neq j$ und $l_i(v_i) = 1$.

Wir zeigen noch die lineare Unabhängigkeit: Sei $\sum s_i l_i(v) = 0$ für alle $v \in V$. Wir setzen speziell $v = v_j$, dann ist $0 = \sum s_i l_i(v_j) = \sum s_i \delta_{ij} = s_j$, also sind alle s_j Null. \square

Folgerung 5.0.1 *Zu einer gegebenen Basis $B = \{b_1, \dots, b_n\}$ von V existiert eine Basis $\{l_1, \dots, l_n\}$ von V^* mit $l_i(b_j) = \delta_{ij}$, sie heißt die zu B duale Basis.* \square

Wenn die zur Basis B von V duale Basis von V^* bekannt ist, kann man die Koordinaten eines Vektors leicht berechnen:

Sei $v = \sum r_i b_i$, dann ist $l_j(v) = \sum r_i l_j(b_i) = \sum r_i \delta_{ij} = r_j$, also ist $v = \sum l_i(v) b_i$.

Wir betrachten als Spezialfall den Vektorraum R^n der Zeilenvektoren und es sei $X \in R^n$. Sei weiter

$$L = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

ein Spaltenvektor, dann ist $XL \in M_{11}$, also eine Zahl, und die Zuordnung $l : X \rightarrow XL$ ist eine Linearform.

Es gibt Mengen aus n linear unabhängigen Spaltenvektoren, dazu gehören n linear unabhängige Linearformen, die Basen von R^{n*} bilden, also kann der Raum der Spaltenvektoren als der zum Raum der Zeilenvektoren duale angesehen werden.

Definition: Sei $M \subseteq V$ eine Teilmenge, wir setzen

$$\text{Ann}(M) = \{l \in V^* \mid l(m) = 0 \text{ für alle } m \in M\},$$

und für $L \subseteq V^*$ setzen wir

$$\text{Ann}(L) = \{v \in V \mid l(v) = 0 \text{ für alle } l \in L\}.$$

Lemma 5.0.1 $\text{Ann}(M)$ ist ein Unterraum von V^* , $\text{Ann}(L)$ ist ein Unterraum von V .

Beweis: Wenn $l_1(m) = l_2(m) = 0$ für alle m gilt, so ist auch $(rl_1 + l_2)(m) = 0$, also $rl_1 + l_2 \in \text{Ann}(M)$. Aus $l(v_1) = l(v_2) = 0$ für alle l folgt $l(rv_1 + v_2) = 0$, also $rv_1 + v_2 \in \text{Ann}(L)$. \square

Lemma 5.0.2 $\dim \text{Ann}(M) = \dim V - \dim \mathcal{L}(M)$.

Beweis: Wir wählen eine Basis $\{v_1, \dots, v_m\}$ von $\mathcal{L}(M)$ und ergänzen sie zu einer Basis $\{v_1, \dots, v_m, v_{m+1}, \dots, v_n\}$ von V , die dazu duale Basis von V^* sei $\{v_1^*, \dots, v_m^*, v_{m+1}^*, \dots, v_n^*\}$. Dann ist $l = \sum r_i v_i^*$ genau dann aus $\text{Ann}(M)$, wenn $l(v_1) = \dots = l(v_m) = 0$. Es ist $l(v_i) = \sum r_k v_k^*(v_i) = r_i$, also ist l genau dann aus $\text{Ann}(M)$, wenn $l \in \mathcal{L}(v_{m+1}^*, \dots, v_n^*)$ ist. \square

Lemma 5.0.3 $\text{Ann}(\text{Ann}(M)) = \mathcal{L}(M)$.

Beweis: Wenn m in M liegt, so ist $l(m) = 0$ für alle $l \in \text{Ann}(M)$, also ist $m \in \text{Ann}(\text{Ann}(M))$ und damit ist $\mathcal{L}(M) \subseteq \text{Ann}(\text{Ann}(M))$ und aus Dimensionsgründen folgt die Gleichheit. \square

Satz 5.0.6 $\text{Ann}(U + W) = \text{Ann}(U) \cap \text{Ann}(W)$, $\text{Ann}(U \cap W) = \text{Ann}(U) + \text{Ann}(W)$, wobei $U, W \subseteq V$ Teilräume sind.

Beweis: Sei $l \in \text{Ann}(U + W)$, d.h. $l(u + w) = 0$ für alle $u \in U, w \in W$. Dann ist speziell $l(u) = 0$ für alle $u \in U$ und $l(w) = 0$ für alle $w \in W$, also $l \in \text{Ann}(U)$ und $l \in \text{Ann}(W)$, also $l \in \text{Ann}(U) \cap \text{Ann}(W)$, also $\text{Ann}(U + W) \subseteq \text{Ann}(U) \cap \text{Ann}(W)$.

Sei nun $l \in \text{Ann}(U) \cap \text{Ann}(W)$, also $l(u) = l(w) = 0$ für alle u, w , dann ist $l(u + w) = 0$, also $l \in \text{Ann}(U + W)$. \square

Sei $f : V \rightarrow W$ eine lineare Abbildung und sei $l \in W^*$ beliebig, d.h. $l : W \rightarrow R$ ist auch linear. Dann ist $l \circ f : V \rightarrow R$ linear, also liegt $l \circ f$ in V^* .

Definition: Die Abbildung $f^* : W^* \rightarrow V^*$ mit $f^*(l) = l \circ f$ heißt die zu f duale Abbildung.

Lemma 5.0.4 Seien $f : V \rightarrow W, g : W \rightarrow U$ lineare Abbildungen und $f^* : W^* \rightarrow V^*, g^* : U^* \rightarrow W^*$ die dualen Abbildungen. Dann gilt $(g \circ f)^* = f^* \circ g^*$.

Beweis: Sei $l \in U^*$, dann ist $(g \circ f)^*(l) = l(g \circ f) = (l \circ (g \circ f)) = (l \circ g) \circ f = g^*(l) \circ f = f^*(g^*(l)) = f^* \circ g^*(l)$. \square

Satz 5.0.7 Sei $f : V \rightarrow W$ eine lineare Abbildung und $f^* : W^* \rightarrow V^*$ die dazu duale. Die Abbildung f ist genau dann injektiv, wenn f^* surjektiv ist, und f ist genau dann surjektiv, wenn f^* injektiv ist.

Beweis: Wir berechnen $\text{Ker}(f^*)$ und $\text{Im}(f^*)$: Genau dann ist $l \in \text{Ker}(f^*)$, wenn $f^*(l) = lf = 0$, also wenn $l(f(v)) = 0$ für alle $v \in V$ gilt, also ist $\text{Ker}(f^*) = \text{Ann}(\text{Im}(f))$. Speziell: Wenn $\text{Ker}(f^*) = \{0\}$ ist, so gilt $\text{Im}(f) = W$.

Sei weiter $k \in \text{Im}(f^*)$, dann gibt es ein $l \in W^*$ mit $k = f^*(l) = lf$, also gilt $k(v) = l(f(v))$ für alle $v \in V$. Wenn nun v in $\text{Ker}(f)$ liegt, so ist $k(v) = 0$, folglich ist $\text{Ker}(f)$ in $\text{Ann}(\text{Im}(f^*))$ enthalten. Schließlich ist

$$\begin{aligned} \dim \text{Ann}(\text{Im}(f^*)) &= \dim V^* - \dim \text{Im}(f^*) = \dim V^* - (\dim W^* - \dim \text{Ker}(f^*)) \\ &= \dim V - \dim W + \dim \text{Ann}(\text{Im}(f)) = \dim V - \dim \text{Im}(f) = \dim \text{Ker}(f) \end{aligned}$$

und wegen $\text{Ker}(f) \subseteq \text{Ann}(\text{Im}(f^*))$ folgt die Gleichheit der Vektorräume. Wenn also $\text{Ker}(f) = \{0\}$ ist, so ist $\text{Im}(f^*) = V^*$ und umgekehrt. \square

Seien nun in V und W Basen B und C gewählt, $f : V \rightarrow W$ sei eine lineare Abbildung und $f^* : W^* \rightarrow V^*$ sei die dazu duale Abbildung. Wir wollen die Darstellungsmatrix $A_{C^*B^*}(f^*)$ bestimmen.

Sei $B = \{b_1, \dots, b_n\}$, $C = \{c_1, \dots, c_m\}$, $C^* = \{c_1^*, \dots, c_m^*\}$ und $B^* = \{b_1^*, \dots, b_n^*\}$. Schließlich sei $f(b_i) = \sum f_{ji}c_j$, also $F = (f_{ji}) = A_{BC}(f)$.

Wir betrachten nun $f^*(c_j^*) : V \rightarrow R$, es ist

$$f^*(c_j^*)(b_i) = c_j^*(f(b_i)) = c_j^*\left(\sum_k f_{ki}c_k\right) = f_{ji} = \sum_k f_{jk}b_k^*(b_i),$$

also ist

$$f^*(c_j^*) = \sum_k f_{jk}b_k^*.$$

Die Matrix $F^T = (f'_{ij})$ mit $f'_{ij} = f_{ji}$ heißt die zu F transponierte Matrix. So erhalten wir den

Satz 5.0.8 $A_{C^*B^*}(f^*) = (A_{BC}(f))^T$, $(AB)^T = B^T A^T$, $(A+B)^T = A^T + B^T$. \square

Zum Abschluß betrachten wir den Vektorraum $V^{**} = \text{Hom}(V^*, R)$. Wir haben hier eine kanonische Abbildung

$$i : V \rightarrow V^{**},$$

die folgendermaßen gegeben ist: Für $v \in V$ legen wir die Linearform $i(v)$ auf V^* durch $i(v)(l) = l(v)$ ($l \in V^*$) fest. Die Abbildung i ist linear:

$$i(v + rv')(l) = l(v + rv') = l(v) + rl(v') = i(v)(l) + ri(v')(l) = (i(v) + ri(v'))(l)$$

für alle $l \in V^*$.

Die Abbildung i ist injektiv: Andernfalls gibt es ein $v \neq 0$ mit $i(v) = 0$, d.h. $i(v)(l) = l(v) = 0$ für alle $l \in V^*$. Wir ergänzen $v = v_1$ zu einer Basis von V , die duale Basis sei $\{v_1^*, \dots, v_n^*\}$, nun wählen wir $l = v_1^*$ und erhalten den Widerspruch $v_1^*(v_1) = 0$. Da V ein endlichdimensionaler Vektorraum ist, ist i ein Isomorphismus.

Dies sollten Sie sich merken.

Kapitel 6

Bilinearformen

6.1 Darstellungsmatrizen und Basiswechsel, Diagonalisierung

Sei V ein R -Vektorraum. Eine Bilinearform b auf V ist eine Abbildung $b : V \times V \rightarrow R$, die in jeder Komponente linear ist, d.h.

$$b(v + rv', w) = b(v, w) + rb(v', w),$$

$$b(v, w + rw') = b(v, w) + rb(v, w')$$

für alle $v, v', w, w' \in V$ und $r \in R$.

Beispiele:

1. $V = R$, $b : R \times R \rightarrow R$ sei die Multiplikation. Die Bilinearität ist durch das Distributivgesetz gesichert.
2. $V = R^n$, $b((x_1, \dots, x_n), (y_1, \dots, y_n)) = \sum x_i y_i$.

Mit $B(V)$ bezeichnen wir die Menge aller Bilinearformen auf V . Durch die Festlegungen

$$(b + b')(v, w) = b(v, w) + b'(v, w) \text{ und } (rb)(v, w) = rb(v, w)$$

wird $B(V)$ ein Vektorraum.

Lemma 6.1.1 $B(V)$ ist isomorph zu $\text{Hom}(V, V^*)$.

Beweis: Sei $f : V \rightarrow V^*$ linear, dann setzen wir $b(v, w) = f(v)(w)$, dies ist sicher eine Bilinearform. Sei umgekehrt $b \in B(V)$, dann legen wir die Abbildung $f : V \rightarrow V^*$ durch $f(v)(w) = b(v, w)$ fest, f ist natürlich linear.

Wir setzen nun $H(b) = f$, dann ist H eine bijektive Abbildung von $B(V)$ in $\text{Hom}(V, V^*)$ und aus der obigen Definition der Operationen mit Bilinearformen ergibt sich die Linearität von H . \square

Es ist

$$\begin{aligned} \text{Ker}(H(b)) &= \text{Ker}(f) = \{v \in V \mid f(v) = 0\} \\ &= \{v \in V \mid f(v)(w) = 0 \text{ für alle } w \in W\} \end{aligned}$$

$$= \{v \in V \mid b(v, w) = 0 \text{ für alle } w \in V\}.$$

Definition: Die Bilinearform b heißt nichtausgeartet, wenn $\text{Ker}(H(b)) = \{o\}$ ist.

Lemma 6.1.2 Die Bilinearform b ist genau dann nichtausgeartet, wenn zu jedem $v \in V, v \neq o$, ein $w \in V$ existiert, so daß $b(v, w) \neq 0$ ist.

Beweis: Andernfalls gibt es ein $v \neq o$, so daß für alle Vektoren w gilt $b(v, w) = 0$, d.h. v liegt in $\text{Ker}(H(b))$. \square

Beispiele:

1. $V = \mathbb{R}^2$, $b((x_1, x_2), (y_1, y_2)) = x_1 y_1 + x_2 y_2$, dies ist genau dann für alle y_1, y_2 gleich Null, wenn $x_1 = x_2 = 0$ ist, d.h. b ist nicht ausgeartet.

2. $V = \mathbb{R}^2$, $b((x_1, x_2), (y_1, y_2)) = (x_1 + x_2)(y_1 + y_2)$, dies ist eine ausgeartete Bilinearform, denn $b((1, -1), (y_1, y_2)) = 0$ für alle (y_1, y_2) .

Wir wollen nun Bilinearformen durch Matrizen beschreiben. Sei dazu $C = \{v_1, \dots, v_n\}$ eine Basis von V und $b : V \times V \rightarrow \mathbb{R}$ eine Bilinearform. Es sei $b(v_i, v_j) = b_{ij}$ und wir setzen

$$M_C(b) = B = \begin{pmatrix} b_{11} & \dots & b_{1n} \\ & \dots & \\ b_{n1} & \dots & b_{nn} \end{pmatrix},$$

dies sei die bezüglich C zu b gehörige Darstellungsmatrix. Wenn nun $v = \sum r_i v_i, w = \sum s_j v_j$, dann ist $b(v, w) = b(\sum r_i v_i, \sum s_j v_j) = \sum \sum r_i s_j b(v_i, v_j) = \sum \sum r_i b_{ij} s_j$ oder in Matrixschreibweise

$$b(v, w) = (r_1, \dots, r_n) \begin{pmatrix} b_{11} & \dots & b_{1n} \\ & \dots & \\ b_{n1} & \dots & b_{nn} \end{pmatrix} \begin{pmatrix} s_1 \\ \dots \\ s_n \end{pmatrix} = k_C(v)^T M_C(b) k_C(w).$$

Der Summe von Bilinearformen entspricht die Summe der Darstellungsmatrizen, ebenso ist es mit der Vervielfachung, also ist $B(V)$ isomorph zu M_{nn} ($\dim V = n$).

Lemma 6.1.3 $M_C(b) = A_{C, C^*}(H(b))^T$.

Beweis: Sei $C = \{v_1, \dots, v_n\}$ und $C^* = \{v_1^*, \dots, v_n^*\}$ die zu C duale Basis, weiter sei $H(b)(v_i) = \sum f_{ki} v_k^*$. Dann ist $b_{ij} = b(v_i, v_j) = H(b)(v_i)(v_j) = \sum f_{ki} v_k^*(v_j) = \sum f_{ki} \delta_{kj} = f_{ji}$. \square

Folgerung 6.1.1 Die Bilinearform b ist genau dann nichtausgeartet, wenn $M_C(b)$ regulär ist.

Beweis: Genau in diesem Fall ist $H(b)$ injektiv. \square

Wir zeigen nun, wie sich die Darstellungsmatrizen von Bilinearformen beim Basiswechsel verhalten.

Satz 6.1.1 Seien $C = \{v_1, \dots, v_n\}$ und $D = \{w_1, \dots, w_n\}$ Basen von V , $A = A_{CD}(id)$, $B = M_C(b)$, $B' = M_D(b)$, dann gilt $B = A^T B' A$.

Beweis: Wegen $(id_V)^* = id_{V^*}$ ist das folgende Diagramm kommutativ:

$$\begin{array}{ccc} V & \xrightarrow{H(b)} & V^* \\ id_V \downarrow & & \downarrow id_V^* \\ V & \xrightarrow{H(b)} & V^* \end{array}$$

also $H(b) = id_{V^*} \circ H(b) \circ id_V$, d.h. für die Darstellungsmatrizen gilt

$$M_C(b) = A_{D^*C^*}(id_V^*)M_D(b)A_{CD}(id_V)$$

und die Darstellungsmatrix der dualen Abbildung ist die Transponierte der originalen Darstellungsmatrix, daraus ergibt sich die Behauptung. \square

Definition: Die Bilinearform b heißt symmetrisch, wenn für alle $v, w \in V$ gilt $b(v, w) = b(w, v)$, und alternierend (oder antisymmetrisch), wenn $b(v, w) = -b(w, v)$ ist.

Lemma 6.1.4 *Zu einer symmetrischen Bilinearform b gehört bezüglich jeder Basis von V eine symmetrische Matrix.* \square

Der folgende wichtige Satz besagt, daß symmetrische Matrizen „diagonalisierbar“ sind. Wir geben zwei äquivalente Formulierungen an, beweisen aber nur die erste.

Satz 6.1.2 *1. Sei b eine symmetrische Bilinearform auf V , dann existiert eine Basis B von V , so daß*

$$M_B(b) = \begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & & \dots & \\ 0 & & \dots & d_n \end{pmatrix}$$

eine Diagonalmatrix ist.

2. Wenn A eine symmetrische Matrix ist, so existiert eine reguläre Matrix C , so daß

$$C^T A C = \begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & & \dots & \\ 0 & & \dots & d_n \end{pmatrix}$$

eine Diagonalmatrix ist.

Beweis: Wir führen die Induktion über $\dim V$. Der Induktionsanfang ist trivial. Die Behauptung sei für alle Vektorräume mit einer Dimension, die kleiner als die von V ist, bereits bewiesen. Wenn $b(v, w) = 0$ für alle $v, w \in V$ gilt, so ist nichts zu zeigen. Seien also $u, w \in V$ so gewählt, daß $b(u, w) \neq 0$ ist. Wir suchen einen Vektor v , für den $b(v, v) \neq 0$ ist.

Falls $b(u, u) \neq 0$ ist, so wählen wir $v = u$, wenn $b(w, w) \neq 0$ ist, so wählen wir $v = w$. Wenn aber $b(u, u) = b(w, w) = 0$ ist, so setzen wir $v = u + w$, in der Tat ist

$$b(v, v) = b(u + w, u + w) = b(u, u) + b(u, w) + b(w, u) + b(w, w) = 2b(u, w) \neq 0.$$

Nun wollen wir den Vektor v so zu einer Basis $\{v_1 = v, v_2, \dots, v_n\}$ von V ergänzen, daß $b(v_1, v_i) = 0$ für $i > 1$ ist. Bezüglich dieser Basis gehört dann zu b die Matrix

$$\begin{pmatrix} b(v, v) & 0 & \dots & 0 \\ 0 & ? & \dots & ? \\ 0 & & \dots & \\ 0 & & & \end{pmatrix}$$

und das Problem wäre auf ein kleineres reduziert.

Sei also $\{v_1, w_2, \dots, w_n\}$ (mit $v_1 = v$) irgendeine Basis und $w = \sum r_i w_i$, der Bilinearform b entspreche die Matrix M , genau dann ist $b(v, w) = 0$, wenn

$$(1 \ 0 \ \dots \ 0)M \begin{pmatrix} r_1 \\ \dots \\ r_n \end{pmatrix} = 0$$

ist. Dies ist eine Gleichung mit n Unbekannten, der Lösungsraum hat also die Dimension $n - 1$. Sei $\{v_2, \dots, v_n\}$ eine Basis des Lösungsraums. Dann ist $b(v_1, v_i) = 0$ für $i > 1$. Wir zeigen, daß $\{v_1, \dots, v_n\}$ linear unabhängig ist.

Sei also $\sum r_i v_i = o$, dann gilt

$$0 = b(v_1, \sum r_i v_i) = r_1 b(v_1, v_1) + r_2 b(v_1, v_2) + \dots + r_n b(v_1, v_n),$$

die letzten $n - 1$ Summanden sind null und $b(v_1, v_1)$ ist nicht null, es folgt $r_1 = 0$. Da $\{v_2, \dots, v_n\}$ bereits linear unabhängig waren, sind auch die übrigen r_i null. Bezüglich der Basis $\{v_1, \dots, v_n\}$ hat also b eine Darstellungsmatrix der Form

$$\begin{pmatrix} * & 0 & \dots & 0 \\ 0 & ? & \dots & ? \\ 0 & & \dots & \\ 0 & & & \end{pmatrix}$$

und diese Form hat sie auch bezüglich jeder Basis $\{v_1, w_2, \dots, w_n\}$, wenn nur die $w_i \in \mathcal{L}(v_2, \dots, v_n)$ sind, denn es ist $b(v_1, w_i) = 0$. Wir schränken nun die Bilinearform b auf den Unterraum $\mathcal{L}(v_2, \dots, v_n)$ zu b' ein, in diesem Vektorraum gibt es nach Induktionsvoraussetzung eine Basis (oBdA sei es bereits $\{v_2, \dots, v_n\}$), so daß die Darstellungsmatrix von b' Diagonalgestalt hat. Bezüglich der Basis $\{v_1, \dots, v_n\}$ hat dann die Bilinearform b eine Diagonalmatrix als Darstellungsmatrix. \square

Beispiel:

Wir betrachten die Bilinearform $b((x_1, x_2), (y_1, y_2)) = y_1 x_2 + y_2 x_1$ auf dem R^2 . Bezüglich der kanonischen Basis $\{e_1, e_2\}$ ist ihre Darstellungsmatrix $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Es ist $b(e_1, e_1) =$

$b(e_2, e_2) = 0$ und $b(e_1 + e_2, e_1 + e_2) = 2$. Also setzen wir $v = (1, 1)$. Wir suchen a, b mit $(1 \ 1) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = (1 \ 1) \begin{pmatrix} a \\ b \end{pmatrix} = 0$ und finden $w = (1, -1)$. Bezüglich der Basis

$\{v, w\}$ hat b die Darstellungsmatrix $\begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$.

Wir können die Aussage des obigen Satzes noch verschärfen:

also

$$b\left(\sum r_i v_i, \sum r_i v_i\right) = \sum r_i^2 = b\left(\sum s_j w_j, \sum s_j w_j\right) = -\sum s_j^2,$$

diese Zahl ist sowohl nichtnegativ als auch nichtpositiv, also ist $r_1 = \dots = s_n = 0$. \square

Wenn man eine symmetrische Matrix A nur in eine Diagonalgestalt überführen will und an der Transformationsmatrix (bzw. an der neuen Basis) gar nicht interessiert ist, so kann man den „symmetrischen“ Gaußschen Algorithmus anwenden, d.h. zu jeder Zeilenoperation hat man „dieselbe“ Spaltenoperation auf A anzuwenden. Denn sei

$$E^{ij} = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & \ddots \\ & & & & 1 \end{pmatrix}$$

eine Elementarmatrix, wo an der Stelle (i, j) eine Eins steht. Dann entsteht $E^{ij}A$ aus A durch Addition der j -ten Zeile zur i -ten, während AE^{ji} aus A durch Addition der j -ten Spalte zur i -ten entsteht. Wenn wir also eine Zeilenoperation auf A anwenden, die die Komponente a_{ij} zu Null macht, so verschwindet wegen der Symmetrie von A nach der entsprechenden Spaltenoperation die Komponente a_{ji} .

6.2 Jacobi-Diagonalisierung

Von Jacobi (also aus dem 19. Jahrhundert) stammt das folgende Iterationsverfahren zur Überführung einer symmetrischen Matrix in eine Diagonalform.

Es sei eine symmetrische Matrix A gegeben, wir betrachten Matrizen

$$J_{ij}(w) = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & c & s \\ & & -s & c \\ & & & \ddots \\ & & & & 1 \end{pmatrix}$$

die sich nur an vier Stellen von der Einheitsmatrix unterscheiden und wo $s^2 + c^2 = 1$ ist. Die Zahlen c und s können wir als Cosinus bzw. Sinus eines Winkels w auffassen, dann ist die Matrix $J_{ij}(w)$ die Darstellungsmatrix einer Drehung in der i, j -Ebene um den Winkel w .

Wir werden die Matrix A mit einer Folge derartiger Drehmatrizen transformieren, also Operationen der Form

$$A \rightarrow B = J_{ij}(w)^T A J_{ij}(w)$$

durchführen, und zwar suchen wir die Drehmatrizen so aus, daß in jedem Schritt die Zahl

$$\text{off}(A) = \sum_{i \neq j} a_{ij}^2$$

kleiner wird. Die Matrix A nähert sich also immer weiter an eine Diagonalmatrix an. Wir wählen die Drehmatrix so, daß nacheinander an den Stellen $(1,2), (1,3), \dots, (1,n), (2,3), \dots, (n-1,n)$ Nullen entstehen. Daß eine solche Wahl gelingt, wollen wir uns im Fall von 2×2 -Matrizen verdeutlichen.

Wir berechnen

$$\begin{aligned} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} &= \begin{pmatrix} c & -s \\ s & c \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} c & s \\ -s & c \end{pmatrix} \\ &= \begin{pmatrix} c^2 a_{11} - cs(a_{21} + a_{12}) + s^2 a_{22} & c^2 a_{21} + cs(a_{11} - a_{22}) - s^2 a_{12} \\ sc(a_{11} - a_{22}) - s^2 a_{21} + c^2 a_{12} & s^2 a_{11} + sc(a_{12} + a_{21}) + c^2 s_{22} \end{pmatrix} \end{aligned}$$

Es soll nun

$$b_{21} = sc(a_{11} - a_{22}) + (c^2 - s^2)a_{21} = 0$$

sein, d.h. es muß gelten

$$\frac{c^2 - s^2}{2cs} = \frac{a_{22} - a_{11}}{2a_{21}} = x,$$

die Zahl x ist bekannt, c bzw. s sind gesucht.

Wir denken daran, daß $c = \cos(w)$ und $s = \sin(w)$ ist, dann ist $x = \cot(2w) = \frac{1}{2} \cot(w) - \frac{1}{2} \tan w$. Wir setzen $\tan(w) = t$ und erhalten $2x - \frac{1}{t} + t = 0$ oder $t^2 + 2xt - 1 = 0$ und damit $t = -x \pm \sqrt{x^2 + 1}$, also $c = \frac{1}{\sqrt{1+t^2}}$, $s = tc$.

Genauso geht das natürlich auch für $n \times n$ -Matrizen.

Wir bezeichnen die Zahl $\sum a_{ij}^2$ mit $F(A)$, dies ist die sogenannte Frobenius-Norm der Matrix A .

Lemma 6.2.1 Sei J eine Drehmatrix wie oben, dann ist $F(J^T A J) = F(A)$.

Beweis: Wir setzen der Einfachheit halber $i = 1, j = 2$ und berechnen

$$J^T A = \begin{pmatrix} c & -s & & & \\ s & c & & & \\ & & 1 & & \\ & & & \dots & \\ & & & & 1 \end{pmatrix} \begin{pmatrix} a_{11} & & & & \\ a_{21} & & & & \\ & \dots & & & \\ & & 1 & & \\ & & & \dots & \end{pmatrix} = \begin{pmatrix} ca_{11} - sa_{21} & & & & \\ sa_{11} + ca_{21} & & & & \\ & \dots & & & \\ & & \dots & & \\ & & & \dots & \end{pmatrix}$$

und sehen

$$(ca_{11} - sa_{21})^2 + (sa_{11} + ca_{21})^2 = a_{11}^2 + a_{21}^2,$$

d.h. bereits für je zwei benachbarte Stellen bleibt die Quadratsumme konstant, und dies gilt auch beim Übergang von A zu AJ . \square

Wie unterscheiden sich nun $\text{off}(A)$ und $\text{off}(B)$? Wir bemerken zunächst, daß sich A und B überhaupt nur in der 1. und 2. Zeile bzw. Spalte voneinander unterscheiden. Es ist weiter $\text{off}(A) = F(A) - \sum a_{ii}^2$ und es gilt

$$\text{off}(B) = F(B) - \sum b_{ii}^2 = F(A) - b_{11}^2 - b_{22}^2 - \sum_{i>2} a_{ii}^2,$$

da $a_{ii} = b_{ii}$ für $i > 2$. Weiter gilt

$$b_{11}^2 + 2b_{12}^2 + b_{22}^2 = a_{11}^2 + 2a_{12}^2 + a_{22}^2$$

also

$$\text{off}(B) = \text{off}(A) - 2a_{12}^2 + 2b_{12}^2 = \text{off}(A) - 2a_{12}^2.$$

Es sei also a_{pq} das betragsgrößte Element von A außerhalb der Diagonalen, wir transformieren A mit $J_{pq}(w)$ mit geeignetem w , dabei wird $\text{off}(A)$ um $2a_{pq}^2$ verkleinert.

Es sei $N = \frac{n(n-1)}{2}$ die Zahl der Elemente von A oberhalb der Diagonalen. Dann ist $2a_{pq}^2 \geq \frac{\text{off}(A)}{N}$, also, wenn A_k das nach k Schritten erhaltene Ergebnis ist, gilt

$$\text{off}(A_k) = \text{off}(A_{k-1}) - 2a_{pq}^2 \leq \left(1 - \frac{1}{N}\right)\text{off}(A_{k-1}) \leq \left(1 - \frac{1}{N}\right)^k \text{off}(A),$$

also konvergiert das Verfahren.

6.3 Strassens schnelle Matrixmultiplikation

Normalerweise benötigt man zur Berechnung des Produkts zweier $n \times n$ -Matrizen n^3 Multiplikationen. Im folgenden stellen wir ein Verfahren vor, das es gestattet, große Matrizen mit nur $n^{2,8}$ Multiplikationen zu multiplizieren.

Wir nehmen an, wir hätten schon ein schnelles Verfahren zur Multiplikation von $\frac{n}{2} \times \frac{n}{2}$ -Matrizen. Dann teilen wir die $n \times n$ -Matrix in vier $\frac{n}{2} \times \frac{n}{2}$ -Matrizen und wenden auf diese Blockmatrizen das schnelle Multiplikationsverfahren für 2×2 -Matrizen an. Es bleibt also nur eine schnelle Multiplikation für 2×2 -Matrizen zu finden. Wir werden sehen, daß man 2×2 -Matrizen mit 7 (anstelle von 8) Multiplikationen multiplizieren kann.

Wir betrachten das Produkt zweier 2×2 -Matrizen:

$$\begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} \begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix} = \begin{pmatrix} c_1 & c_2 \\ c_3 & c_4 \end{pmatrix}$$

und fassen die Matrizen als Elemente des R^4 auf. Die c_i sind dann Bilinearformen auf R^4 :

$$c_1 = a_1b_1 + a_2b_3$$

$$c_2 = a_1b_2 + a_2b_4$$

$$c_3 = a_3b_1 + a_4b_3$$

$$c_4 = a_3b_2 + a_4b_4.$$

Bezüglich der kanonischen Basis des R^4 entsprechen den c_i die folgenden Matrizen:

$$c_1 : \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad c_2 : \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad c_3 : \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad c_4 : \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Jede der obigen Bilinearformen „enthält“ zwei Produkte, ihre Darstellungsmatrizen haben den Rang 2.

Eine Bilinearform, die nur ein Produkt enthält, hat die Form

$$(r_1 a_1 + r_2 a_2 + r_3 a_3 + r_4 a_4)(s_1 b_1 + s_2 b_2 + s_3 b_3 + s_4 b_4),$$

ihre Darstellungsmatrix

$$\begin{pmatrix} r_1 s_1 & r_2 s_1 & r_3 s_1 & r_4 s_1 \\ r_1 s_2 & r_2 s_2 & r_3 s_2 & r_4 s_2 \\ r_1 s_3 & r_2 s_3 & r_3 s_3 & r_4 s_3 \\ r_1 s_4 & r_2 s_4 & r_3 s_4 & r_4 s_4 \end{pmatrix}$$

hat den Rang 1. Das Problem besteht nun darin, die den c_i entsprechenden Matrizen als Summe möglichst weniger Matrizen vom Rang 1 darzustellen. Strassen zeigte 1969, das hierfür die folgenden 7 Matrizen ausreichen:

$$m_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 \end{pmatrix} \quad m_2 = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \quad m_3 = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$m_4 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad m_5 = \begin{pmatrix} 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad m_6 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 \end{pmatrix}$$

$$m_7 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

Diesen Matrizen entsprechen die folgenden Bilinearformen:

$$(a_2 - a_4)(b_3 + b_4), (a_1 + a_4)(b_1 + b_4), (a_1 - a_3)(b_1 + b_2),$$

$$(a_1 + a_2)b_4, a_1(b_2 - b_4), a_4(b_3 - b_1), (a_3 + a_4)b_1.$$

Die Bilinearformen c_1, \dots, c_4 lassen sich aus diesen linear kombinieren:

$$c_1 = m_1 + m_2 - m_4 + m_6, \quad c_2 = m_4 + m_5, \quad c_3 = m_6 + m_7, \quad c_4 = m_2 - m_3 + m_5 - m_7.$$

Vom Numerik-System LAPACK wird Strassens Methode genutzt.

Zum Abschluß dieses Abschnitts wollen wir unsere Kenntnisse in der Geometrie anwenden.

6.4 Klassifikation der Quadriken

Sei A ein affiner Raum und $\{P, v_1, \dots, v_n\}$ ein Koordinatensystem, ein Punkt $X \in A$ habe die Koordinaten $(1, x_1, \dots, x_n)$. Wie wir wissen, läßt sich die Zugehörigkeit des Punkts X zu einem gegebenen Unterraum H von A daran erkennen, ob sein Koordinatentupel eine Lösung eines gewissen linearen Gleichungssystems ist. Also: Eine lineare Gleichung beschreibt eine Hyperebene.

Wir wollen nun „quadratische“ Gleichungen betrachten und feststellen, was für ein geometrisches Gebilde die Lösungstupel solcher Gleichungen darstellen.

Definition: Sei $\{P, v_1, \dots, v_n\}$ ein Koordinatensystem des affinen Raums A . Die Menge aller Punkte X mit den Koordinaten $(1, x_1, \dots, x_n)$ mit

$$Q : \sum_{i,j=1}^n a_{ij}x_i x_j + 2 \sum_{i=1}^n a_i x_i + a_0 = 0$$

heißt eine Quadrik (oder quadratische Hyperfläche).

Wir können die linke Seite der Gleichung auch als Matrixprodukt schreiben:

$$Q : \begin{pmatrix} 1 & x_1 & \dots & x_n \end{pmatrix} \begin{pmatrix} a_0 & a_1 & \dots & a_n \\ a_1 & a_{11} & \dots & a_{1n} \\ & & \dots & \\ a_n & a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Wir können oBdA festlegen, daß $a_{ij} = a_{ji}$ ist, daß also die Matrix $A = (a_{ij})$ symmetrisch ist. Mit dieser Abkürzung können wir die quadratische Gleichung einfach in der Form $X^T A X = 0$ schreiben (wir haben hier den Punkt X mit seinem Koordinatentupel identifiziert).

Bei einem anderen Koordinatensystem möge der Punkt das Koordinatentupel X' haben, dann gilt $X = B X'$ für eine gewisse reguläre Matrix B . Dann ist aber

$$X^T A X = X'^T (B^T A B) X',$$

also ist $B^T A B$ die Matrix, die die gegebene Quadrik bezüglich des neuen Koordinatensystems beschreibt.

Wir fragen uns nun, ob es ein Koordinatensystem gibt, bezüglich dessen die die Quadrik beschreibende Matrix „möglichst einfach“ aussieht.

Dazu betrachten wir zunächst die symmetrische Bilinearform

$$b(v, w) = \sum a_{ij} v_i w_j.$$

Wir wissen, daß es eine Basis $\{u_1, \dots, u_n\}$ gibt, so daß $b(u_i, u_j) = d_i \delta_{ij}$ gilt, d.h. die Darstellungsmatrix hat eine Diagonalgestalt. Also können wir für A die Gestalt

$$\begin{pmatrix} a_0 & a_1 & \dots & a_r & \dots & a_n \\ a_1 & a_{11} & 0 & & \dots & 0 \\ & & \dots & & & \\ a_r & 0 & \dots & a_{rr} & \dots & 0 \\ a_n & 0 & & & \dots & 0 \end{pmatrix}$$

annehmen. Dann wird die Quadrik Q durch die Gleichung

$$\sum_{i=1}^r a_{ii}x_i^2 + 2 \sum a_i x_i + a_0 = 0$$

beschrieben. Für $i = 1, \dots, r$ führen wir eine quadratische Ergänzung durch, wir setzen

$$x'_i = x_i + \frac{a_i}{a_{ii}},$$

die Gleichung für Q hat dann in den gestrichenen Koordinaten die Form

$$\sum_{i=1}^r a_{ii}x_i'^2 + \sum_{i=r+1}^n a_i x_i' + a'_0 = 0.$$

Wenn alle a_i null sind oder $r = n$ ist, so hat die Gleichung die einfache Gestalt

$$\sum a_{ii}x_i^2 + a'_0 = 0.$$

Nun betrachten wir den Fall, daß $r < n$ ist, es sei mindestens ein $a_i \neq 0$, etwa $a_n \neq 0$. Wir setzen für $i = 1, \dots, n-1$ $x''_i = x'_i$ und

$$x''_n = x'_n + \frac{a_{r+1}}{a_n}x'_{r+1} + \dots + \frac{a_{n-1}}{a_n}x'_{n-1} + \frac{a'_0}{a_n}$$

dann erhält die Gleichung in den zweigestrichenen Koordinaten die Form

$$\sum a_{ii}x_i''^2 + 2a_n x_n'' = 0.$$

Unter den Koeffizienten a_{ii} seien oBdA die ersten p positiv und die restlichen $r-p$ negativ, wir ersetzen sie durch $\pm d_i$, wobei $d_i > 0$ sein soll. Wir ersetzen die gestrichenen Koordinaten wieder durch die ursprünglichen und dividieren noch durch a_0 (falls von Null verschieden) bzw. $2a_n$.

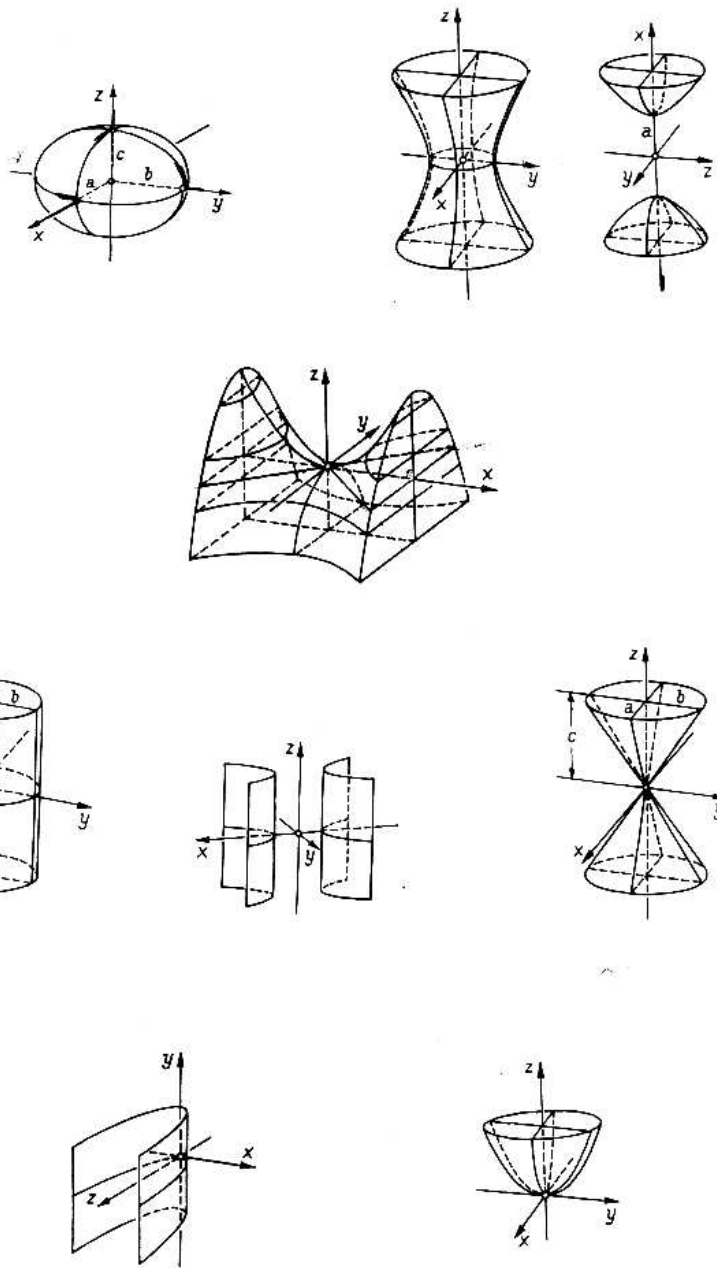
Insgesamt können folgende drei Fälle auftreten:

$$\sum_{i=1}^p d_i x_i^2 - \sum_{i=p+1}^r d_i x_i^2 = \begin{cases} 0 & \text{(Fall 1)} \\ 1 & \text{(Fall 2)} \\ x_{r+1} & \text{(Fall 3)}. \end{cases}$$

In den folgenden Tabellen geben wir eine Übersicht über alle Quadriken für $n = 2$ (quadratische Kurven) und $n = 3$ (quadratische Flächen), dabei sind d_1, \dots durch a, b, c und x_1, \dots durch x, y, z ersetzt:

$n = 2$			
(p, r)	Fall 1	Fall 2	Fall 3
(2,2)	$ax^2 + by^2 = 0$ Punkt	$ax^2 + by^2 = 1$ Ellipse	
(1,2)	$ax^2 - by^2 = 0$ Geradenpaar	$ax^2 - by^2 = 1$ Hyperbel	
(0,2)	$-ax^2 - by^2 = 0$ Geradenpaar	$-ax^2 - by^2 = 1$ leer	
(1,1)	$ax^2 = 0$ Gerade	$ax^2 = 1$ parallele Geraden	$ax^2 = y$ Parabel
(0,1)	$-ax^2 = 0$ Punkt	$-ax^2 = 1$ leer	$-ax^2 = y$ Parabel

$n = 3$			
(p, r)	Fall 1	Fall 2	Fall 3
(3,3)	$ax^2 + by^2 + cz^2 = 0$ Punkt	$ax^2 + by^2 + cz^2 = 1$ Ellipsoid	
(2,3)	$ax^2 + by^2 - cz^2 = 0$ Doppelkegel	$ax^2 + by^2 - cz^2 = 1$ einschaliges Hyperboloid	
(1,3)	$ax^2 - by^2 - cz^2 = 0$ Doppelkegel	$ax^2 - by^2 - cz^2 = 1$ zweischaliges Hyperboloid	
(2,2)	$ax^2 + by^2 = 0$ Gerade	$ax^2 + by^2 = 1$ elliptischer Zylinder	$ax^2 + by^2 = z$ Paraboloid
(1,2)	$ax^2 - by^2 = 0$ schneidende Flächen	$ax^2 - by^2 = 1$ hyperbolischer Zylinder	$ax^2 - by^2 = z$ hyperbolisches Paraboloid
(1,1)	$ax^2 = 0$ Ebene	$ax^2 = 1$ parallele Ebenen	$ax^2 = y$ parabolischer Zylinder



6.5 Bilinearformen in der Analysis

Sei $C(a, b)$ die Menge der im Intervall (a, b) stetigen Funktionen, dies ist ein unendlichdimensionaler Vektorraum. Für Funktionen $f, g \in C(a, b)$ ist durch

$$I(f, g) = \int_a^b f(x)g(x)dx$$

eine symmetrische Bilinearform gegeben.

Wir betrachten den endlichdimensionalen Unterraum $\mathcal{L}(\sin(nx) \mid n = 0, 1, \dots, k)$ und das Intervall $(-\pi, \pi)$, bezüglich dieser Basis ist die Bilinearform I diagonalisiert: Es gilt

$$\sin(mx)\sin(nx) = \frac{1}{2}(\cos((m-n)x) - \cos((m+n)x))$$

und damit für $m \neq n$

$$\begin{aligned} \int_{-\pi}^{\pi} \sin(mx) \cdot \sin(nx) dx &= \frac{1}{2} \int_{-\pi}^{\pi} (\cos((m-n)x) - \cos((m+n)x)) dx \\ &= \frac{1}{2} \left[\frac{1}{m-n} \sin((m-n)x) \right]_{-\pi}^{\pi} - \frac{1}{2} \left[\frac{1}{m+n} \sin((m+n)x) \right]_{-\pi}^{\pi} = 0 \end{aligned}$$

und für $m = n$ haben wir

$$\frac{1}{2} \int_{-\pi}^{\pi} 1 dx - \frac{1}{2} \int_{-\pi}^{\pi} \cos(2mx) dx = \pi + 0.$$

Im Vektorraum $P_n = \{\sum_{i=0}^n a_i x^i\}$ der Polynome vom Grad $\leq n$ betrachten wir die Bilinearform

$$I(f, g) = \int_{-1}^1 f(x)g(x)dx.$$

Für die Legendre-Polynome

$$L_0(x) = 1, \quad L_1(x) = x, \quad L_2(x) = \frac{1}{2}(3x^2 - 1), \quad L_3(x) = \frac{1}{2}(5x^3 - 3x),$$

allgemein

$$(k+1)L_{k+1}(x) = (2k+1)x \cdot L_k(x) - k \cdot L_{k-1}(x)$$

gilt

$$\int_{-1}^1 L_n(x)L_m(x)dx = \frac{2}{2m+1}\delta_{mn}.$$

Wenn wir die Bilinearform

$$I(f, g) = \int_{-1}^1 f(x)g(x) \frac{1}{\sqrt{1-x^2}} dx$$

wählen, so leisten die Tschebyscheff-Polynome

$$T_0(x) = 1, \quad T_1(x) = x, \quad T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$$

gutes:

$$I(f, g) = \int_{-1}^1 T_n(x)T_m(x) \frac{1}{\sqrt{1-x^2}} dx = \begin{cases} 0 & m \neq n \\ \frac{\pi}{2} & m = n \neq 0 \\ \pi & m = n = 0 \end{cases}$$

Kapitel 7

Determinanten

7.1 Existenz und Eindeutigkeit

Es sei (A, V) ein affiner Raum; wir wollen den Begriff des Flächeninhalts fassen.

Sei dazu O ein Punkt und seien v, w Vektoren, diese bestimmen ein Parallelogramm mit den Eckpunkten $O, O+v, O+w, O+v+w$, dessen „Flächeninhalt“ wir mit $F(v, w)$ bezeichnen wollen. Der Flächeninhalt soll die folgenden Eigenschaften haben:

1. $F(rv, w) = rF(v, w) = F(v, rw)$ ($r \in R$),
2. $F(v+v', w) = F(v, w) + F(v', w)$,
3. $F(v, w+w') = F(v, w) + F(v, w')$,
4. $F(v, v) = 0$.

Diese Forderungen haben zu Folge, daß gilt

$$0 = F(v+w, v+w) = F(v, v) + F(v, w) + F(w, v) + F(w, w) = F(v, w) + F(w, v),$$

d.h. der Flächeninhalt, falls es so eine Funktion überhaupt gibt, ist „orientiert“.

Sei $\{e_1, e_2\}$ eine Basis von V und

$$v = r_1e_1 + r_2e_2, \quad w = s_1e_1 + s_2e_2,$$

dann ist

$$\begin{aligned} F(v, w) &= F(r_1e_1 + r_2e_2, s_1e_1 + s_2e_2) \\ &= r_1s_1F(e_1, e_1) + r_1s_2F(e_1, e_2) + r_2s_1F(e_2, e_1) + r_2s_2F(e_2, e_2) \\ &= (r_1s_2 - r_2s_1)F(e_1, e_2), \end{aligned}$$

d.h. wir brauchen nur den Flächeninhalt des Parallelogramms, das von e_1, e_2 aufgespannt wird, festzulegen und können $F(v, w)$ aus den Koordinaten der Vektoren berechnen.

Der Term $r_1s_2 - r_2s_1$ wird als Determinante der aus den Koordinatentupeln gebildeten Matrix

$$\begin{pmatrix} r_1 & s_1 \\ r_2 & s_2 \end{pmatrix}$$

bezeichnet.

Wir wollen dies verallgemeinern:

Definition: Eine Funktion $F : (M_{n1})^n \rightarrow R$ heißt Multilinearform, wenn für fixierte $a_j \in M_{n1}$ jede Abbildung

$$F_i(v) = F(a_1, a_2, \dots, a_{i-1}, v, a_{i+1}, \dots, a_n) : M_{n1} \rightarrow R$$

linear ist. Eine Multilinearform heißt alternierend, wenn $F(a_1, \dots, a_n) = 0$ ist, falls $\{a_1, \dots, a_n\}$ linear abhängig ist.

Wir fassen Multilinearformen meist als Abbildungen von M_{nn} in R auf und sagen dann, daß sie linear in den Spalten der Matrix sind.

Definition: Eine alternierende Multilinearform $D : M_{nn} \rightarrow R$, deren Wert auf der Einheitsmatrix E gleich 1 ist, heißt Determinante vom Grad n .

Wie oben beim Flächeninhalt erhalten wir nun aus der Definition folgende Eigenschaften alternierender Multiplinearformen:

1. $F(\dots, a, \dots, a, \dots) = 0$,
2. Beim Vertauschen von Spalten ändert sich das Vorzeichen:
 $0 = F(\dots, a + b, \dots, a + b, \dots) =$
 $F(\dots, a, \dots, a, \dots) + F(\dots, a, \dots, b, \dots) + F(\dots, b, \dots, a, \dots) + F(\dots, b, \dots, b, \dots),$
 also gilt
 $F(a_1, \dots, a_i, \dots, a_j, \dots, a_n) = -F(a_1, \dots, a_j, \dots, a_i, \dots, a_n).$
3. Elementare Spaltenoperationen ändern den Wert nicht:
 $F(a_1 + ra_2, a_2, \dots, a_n) = F(a_1, a_2, \dots, a_n) + rF(a_2, a_2, \dots, a_n)$
 und der zweite Summand ist null.
4. Wenn f eine Multilinearform mit der Eigenschaft $f(\dots, a \dots, a \dots) = 0$ ist, so ist f alternierend, denn wenn von den Vektoren v_1, \dots, v_n einer eine Linearkombination der übrigen ist, so kann man durch elementare Operationen zwei gleich Vektoren herstellen.

Satz 7.1.1 *Es gibt eine Funktion $D : M_{nn} \rightarrow R$, die eine Determinante vom Grad n ist.*

Beweis: Wir führen die Induktion über n .

Für $n = 1$ können wir $M_{11} = R$ annehmen, dann setzen wir $D = id$, diese Funktion erfüllt die Bedingungen. Sei D eine Determinante vom Grad $n - 1$, wir konstruieren eine Determinante D' vom Grad n wie folgt:

Sei $A = (a_{ij})$ eine $n \times n$ -Matrix; die $(n - 1) \times (n - 1)$ -Matrix, die aus A entsteht, wenn die i -te Zeile und die j -te Spalte gestrichen wird, bezeichnen wir mit A_{ij} . Sei nun i eine beliebige Zahl zwischen 1 und n , dann setzen wir

$$D'(A) = (-1)^{i+1}a_{i1}D(A_{i1}) + (-1)^{i+2}a_{i2}D(A_{i2}) + \dots + (-1)^{i+n}a_{in}D(A_{in})$$

(diese Formel heißt Laplacescher Entwicklungssatz für die i -te Zeile).

Wir zeigen nun die Linearität der Abbildung D' in den Spalten. Betrachten wir die erste Spalte a_1 von A und halten a_2, \dots, a_n fest:

In A_{i1} kommt die erste Spalte von A gar nicht vor, also ist $D(A_{i1})$ konstant und die Abbildung

$$A \rightarrow (-1)^{i+1}a_{i1}D(A_{i1})$$

ist offenbar linear. Weiter sind $D(A_{i2}), \dots, D(A_{in})$ nach Induktionsvoraussetzung linear in der ersten Spalte und die Faktoren a_{i2}, \dots, a_{in} hängen von der ersten Spalte von A nicht ab, also sind auch die Abbildungen

$$A \rightarrow (-1)^{i+j}a_{ij}D(A_{ij})$$

für $j > 1$ linear in der ersten Spalte von A . Folglich ist $D'(A)$ als Summe linearer Abbildungen in der ersten Spalte von A linear. Die Linearität in den anderen Spalten zeigt man analog. Wir prüfen noch, ob D' alternierend ist. Wir haben oben gezeigt, daß dies dann erfüllt ist, wenn der Funktionswert einer Multilinearform für solche Matrizen verschwindet, bei denen zwei Spalten übereinstimmen. Sei oBdA $a_1 = a_2$, dann ist

$$D'(A) = (-1)^{i+1}a_{i1}D(A_{i1}) + (-1)^{i+2}a_{i2}D(A_{i2}) + 0,$$

da die restlichen A_{ij} zwei gleiche Spalten besitzen. Nun ist $a_{i1} = a_{i2}$ und $A_{i1} = A_{i2}$ und beide Summanden haben unterschiedliche Vorzeichen, also ist $D'(A) = 0$.

Schließlich ist $D'(E_n) = 1$, wie man leicht sieht: Beim Streichen der i -ten Zeile und j -ten Spalte verschwindet die 1 an der Stelle (i, i) und es entsteht eine Nullspalte. Wenn $j = i$ ist, so wird diese Nullspalte gestrichen. \square

Zum Beispiel wäre

$$D' \begin{pmatrix} a & b \\ c & d \end{pmatrix} = a \cdot d - b \cdot c$$

und

$$D' \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = 1 \cdot D \begin{pmatrix} 5 & 6 \\ 8 & 9 \end{pmatrix} - 2 \cdot D \begin{pmatrix} 4 & 6 \\ 7 & 8 \end{pmatrix} + 3 \cdot D \begin{pmatrix} 4 & 5 \\ 7 & 8 \end{pmatrix}.$$

Für die folgenden Betrachtungen brauchen wir einen neuen Begriff:

Definition: Die Menge aller bijektiven Abbildungen der Menge $\{1, \dots, n\}$ in sich wird mit S_n bezeichnet, ihre Elemente heißen Permutationen.

Permutationen kann man multiplizieren, genauer gesagt: Das Produkt (die Nacheinanderausführung) zweier bijektiver Abbildungen von $\{1, \dots, n\}$ in sich ist wieder eine derartige Abbildung, die identische Abbildung ist bei dieser Multiplikation ein neutrales Element und zu jeder Abbildung $f : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ ist die Abbildung f^{-1} invers, d.h. $f \circ f^{-1} = f^{-1} \circ f = id$. Darüberhinaus gilt für die Multiplikation das Assoziativgesetz. Wir sagen, die Menge S_n ist eine „multiplikative Gruppe“, sie wird auch als die Symmetrische Gruppe vom Grade n bezeichnet.

Nun beweisen wir den

Satz 7.1.2 *Es gibt genau eine Determinante vom Grad n .*

Beweis: Sei $D : M_{nn} \rightarrow R$ eine Determinante, also ist $D(A)$ für eine Matrix $A \in M_n$ eine in den in den Spalten von A lineare Funktion. Wir bezeichnen die Spalten von A mit $a_i = \sum a_{ji}e_j$, dann ist

$$D(A) = D\left(\sum a_{j(1),1}e_{j(1)}, \dots, \sum a_{j(n),n}e_{j(n)}\right) = \sum a_{j(1),1} \dots a_{j(n),n} D(e_{j(1)}, \dots, e_{j(n)}),$$

wobei die Summation über alle Indexsysteme $(j(1), \dots, j(n))$ zu erstrecken ist. Nun ist aber $D(e_{j(1)}, \dots, e_{j(n)}) = 0$, wenn nicht alle Indizes voneinander verschieden sind, also sind nur die Summanden von Interesse, wo $\{j(1), \dots, j(n)\} = \{1, \dots, n\}$ ist, d.h. wo die Zuordnung $k \rightarrow j(k)$ eine Permutation ist, also ist

$$D(A) = \sum a_{p(1),1} \dots a_{p(n),n} D(e_{p(1)}, \dots, e_{p(n)}),$$

wo über alle Permutationen $p \in S_n$ zu summieren ist. Der Faktor

$$D(e_{p(1)}, \dots, e_{p(n)})$$

ist die Determinante einer Matrix, die aus der Einheitsmatrix durch gewisse Vertauschungen der Spalten hervorgeht, wegen der Festlegung $D(E) = 1$ ist er also gleich 1 oder -1 , diese Zahl wird als Signum $\text{sgn}(p)$ der Permutation p bezeichnet.

Folglich ist

$$D(A) = \sum_{p \in S_n} a_{p(1),1} \dots a_{p(n),n} \text{sgn}(p),$$

diese Formel heißt „Leibnizsche Definition“ der Determinante. □

Wir haben also eine explizite Formel für die Funktion D gefunden, also gibt es genau eine Determinantenfunktion vom Grade n . Die somit eindeutig bestimmte Determinante einer Matrix A wird mit $\det(A)$ oder auch kurz mit $|A|$ bezeichnet.

Wir erhalten noch die

Folgerung 7.1.1 *Die obige Laplacesche Formel ergibt für jeden Zeilenindex i denselben Wert $D(A)$.*

Satz 7.1.3 *Sei $F : M_{nn} \rightarrow R$ eine alternierende Multilinearform, dann gilt*

$$F(A) = \det(A)F(E).$$

Der Beweis verläuft analog. □

Obwohl die Leibnizsche Formel die explizite Berechnung von $D(A)$ gestattet, ist sie doch nur für kleine Werte von n (etwa $n = 2$ oder 3) zu gebrauchen, für $n = 2$ ergibt sich der anfangs angegebene Wert, für $n = 3$ gibt es eine leicht zu merkende Formel (die „Sarrussche Regel“) zur Determinantenberechnung, die wir hier nicht angeben wollen (Schreiben Sie doch einfach alle sechs Summanden einer Determinante vom Grade 3 auf).

Für größere Werte von n ist die Leibnizsche Formel zu unhandlich, es wären ja $(n-1)n!$ Multiplikationen und $n! - 1$ Additionen nötig. Besser ist die Formel von Laplace geeignet, wenn sie geschickt verwendet wird; wird sie aber nur stur (etwa durch einen Computer) angewandt, werden allerdings ebensoviele Rechenoperationen ausgeführt. Wir wissen allerdings, daß sich der Wert einer Determinante bei elementaren Zeilen- und Spaltenoperationen nicht oder (bei Vertauschungen) nur um das Vorzeichen ändert. Mit Hilfe von etwa n^3 Spaltenoperationen können wir eine Matrix A in eine Dreiecksform überführen:

$$\det(A) = \det\left(\begin{pmatrix} a_{11} & 0 & \dots & 0 \\ & & \dots & \\ a_{n1} & & \dots & a_{nn} \end{pmatrix}\right),$$

und wenn wir jetzt einen Blick auf die Leibnizsche Formel werfen, sehen wir, daß die Summanden für fast alle Permutationen gleich Null sind, da ein Faktor $a_{p(i),i}$ Null ist. Nur die identische Permutation $p = id$ liefert einen (evtl.) von Null verschiedenen Wert, also gilt für eine Dreiecksmatrix A

$$\det(A) = a_{11} \dots a_{nn}.$$

Dies ist das geeignete Verfahren zur Determinantenberechnung:

n	Add.	Mult.	Add.	Mult.
2	1	2	1	3
3	5	9	5	10
4	23	40	14	23
5	119	205	30	45
10	3 628 799	6 235 300	285	339

7.2 Eigenschaften und Anwendungen

Wir beweisen zuerst den

Satz 7.2.1 (Multiplikationssatz) *Seien A, B zwei $n \times n$ -Matrizen, dann gilt*

$$\det(AB) = \det(A) \det(B).$$

Beweis: Sei $B = (b_1, \dots, b_n)$, die Spalten von AB sind dann Ab_1, \dots, Ab_n , also ist $\det(AB) = \det(Ab_1, \dots, Ab_n)$. Wir setzen $F(b_1, \dots, b_n) = \det(Ab_1, \dots, Ab_n)$.

Die Abbildung $F : M_n \rightarrow R$ ist multilinear:

$$\begin{aligned} F(b_1 + rb'_1, b_2, \dots, b_n) &= \det(A(b_1 + rb'_1), Ab_2, \dots) \\ &= \det(Ab_1 + rAb'_1, \dots) \\ &= \det(Ab_1, \dots) + r \det(Ab'_1, \dots) \\ &= F(b_1, \dots, b_n) + rF(b'_1, \dots, b_n). \end{aligned}$$

Die Abbildung F ist auch alternierend: Sei $\{b_1, \dots, b_n\}$ linear abhängig, d.h. $rg(B) < n$, dann ist $rg(AB) \leq rg(B) < n$, also sind die Spalten von AB linear anhängig, d.h. $\det(AB) = F(B) = 0$, also nach der obigen Verallgemeinerung

$$\begin{aligned} \det(AB) &= F(B) \\ &= \det(B)F(E) \\ &= \det(B) \det(Ae_1, \dots, Ae_n) \\ &= \det(B) \det(A) \\ &= \det(A) \det(B). \quad \square \end{aligned}$$

Wir betrachten folgenden Spezialfall: Seien p und q Permutationen aus S_n und $A = (e_{p(1)}, \dots, e_{p(n)})$ sowie $B = (e_{q(1)}, \dots, e_{q(n)})$ Matrizen, die aus der Einheitsmatrix durch Vertauschen der Spalten entstanden sind. Wie sieht dann AB aus? Wir fassen dazu A und B als Darstellungsmatrizen linearer Abbildungen des R^n bezüglich der kanonischen Basis auf: Die zu B gehörige Abbildung bildet e_i in $e_{q(i)}$ ab, zu A gehört die Abbildung, die e_j in $e_{p(j)}$ abbildet. Der Matrix AB entspricht das Produkt dieser Abbildungen, wobei e_i in $e_{p(q(i))}$ überführt wird. Also ist $AB = (e_{pq(1)}, \dots, e_{pq(n)})$ und wir erhalten die Folgerung

Folgerung 7.2.1 $\operatorname{sgn}(pq) = \operatorname{sgn}(p) \operatorname{sgn}(q)$, $\operatorname{sgn}(p) = \operatorname{sgn}(p^{-1})$.

Beweis: Dies folgt aus dem Multiplikationssatz. □

Satz 7.2.2 Die Determinantenfunktion ist auch eine multilineare alternierende Funktion der Zeilen.

Beweis: Wir zeigen, daß $\det(A) = \det(A^T)$ gilt. Sei also $A = (a_{ij})$, $B = A^T = (b_{ij})$ mit $b_{ij} = a_{ji}$. Wenn P eine Permutation ist und $p(i) = j$ gilt, so ist $a_{i,p(i)} = a_{p^{-1}(j),j}$. Dann ist

$$\begin{aligned} \det(B) &= \sum_{p \in S_n} \operatorname{sgn}(p) b_{p(1),1} \dots b_{p(n),n} \\ &= \sum_{p \in S_n} \operatorname{sgn}(p) a_{1,p(1)} \dots a_{n,p(n)} \\ &= \sum_{p \in S_n} \operatorname{sgn}(p) a_{p^{-1}(1),1} \dots a_{p^{-1}(n),n} \\ &= \sum_{p^{-1} \in S_n} \operatorname{sgn}(p^{-1}) a_{p^{-1}(1),1} \dots a_{p^{-1}(n),n} \\ &= \det(A). \quad \square \end{aligned}$$

Wir wissen, daß $\det(A) = 0$ gilt, wenn die Spalten von A linear abhängig sind. Gilt aber auch die Umkehrung?

Satz 7.2.3 Sei $A \in M_{nn}$, genau dann ist $\det(A) \neq 0$, wenn der Rang von A gleich n ist.

Beweis: (\Rightarrow) Klar nach Definition.

(\Leftarrow) Sei $rg(A) = n$, dann läßt sich A durch Zeilen und Spaltenoperationen, die die Determinante ja nicht verändern, in Diagonalform

$$\begin{pmatrix} r_1 & 0 & \dots & 0 \\ & \dots & & \\ 0 & \dots & & r_n \end{pmatrix}$$

mit $r_i \neq 0$ bringen, dann ist $\det(A) = r_1 \dots r_n \neq 0$. □

Satz 7.2.4 (Cramersche Regel) Das Gleichungssystem $Ax = b$, genauer

$$\sum_{j=1}^n a_{ij}x_j = b_i, \quad i = 1, \dots, n$$

mit der quadratischen Koeffizientenmatrix A hat genau dann eine eindeutig bestimmte Lösung, wenn $\det(A) \neq 0$ ist. Diese Lösung ist durch

$$x_k = \frac{\det(A_k)}{\det(A)}, \quad k = 1, \dots, n$$

gegeben, dabei entsteht die Matrix A_k aus A dadurch, daß das n -tupel $\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ anstelle der k -ten Spalte in A eingetragen wird.

Beweis: Eine eindeutig bestimmte Lösung existiert genau dann, wenn $rg(A) = rg(A, b) = n$ ist, d.h. es muß $\det(A) \neq 0$ sein. Sei nun (x_1, \dots, x_n) diese Lösung. Dann gilt $\sum a_{ij}x_j = b_i$. Wir betrachten

$$\begin{aligned} \det(A_k) &= \det(a_1, \dots, a_{k-1}, b, a_{k+1}, \dots, a_n) \\ &= \det(a_1, \dots, a_{k-1}, \sum a_j x_j, a_{k+1}, \dots, a_n) \\ &= \sum \det(a_1, \dots, a_{k-1}, a_j, a_{k+1}, \dots, a_n) x_j \\ &= \det(A) x_k, \end{aligned}$$

da $\det(a_1, \dots, a_{k-1}, a_j, a_{k+1}, \dots, a_n) = 0$ für $j \neq k$ ist. Damit ist die obige Formel bewiesen. □

Wir wenden uns noch einmal dem Laplaceschen Entwicklungssatz zu:

$$\det(A) = \sum_j (-1)^{i+j} a_{ij} \det(A_{ij}), \quad (1)$$

dabei ist i eine (beliebige) Zahl zwischen 1 und n und A_{ij} entsteht aus A durch Streichen der i -ten Zeile und der j -ten Spalte.

Nun ändern wir diese Formel nur an einer Stelle und fragen, „was dann herauskommt“:

$$? = \sum_j (-1)^{k+j} a_{ij} \det(A_{kj}) \text{ mit } k \neq i. \quad (2)$$

Wir können den Wert durch Anwendung der Laplaceschen Formel bestimmen, dies ist doch die Determinante der Matrix, deren k -te Zeile gleich (a_{i1}, \dots, a_{in}) ist, die nach der k -ten Zeile zu entwickeln ist. Diese Determinante hat aber den Wert 0, da zwei Zeilen der Matrix übereinstimmen.

Nun interpretieren wir die Formeln (1) und (2) als ein Matrixprodukt, sie lauten zusammengefaßt

$$\sum (-1)^{k+j} a_{ij} \det(A_{kj}) = \delta_{ik} \det(A)$$

und besagen dann, daß

$$\frac{1}{\det(A)} \left((-1)^{k+j} \det(A_{kj}) \right)^T = A^{-1},$$

wir haben damit eine explizite Formel für die Inverse einer regulären Matrix gefunden.

Wir wenden uns noch dem „klassischen“ Rangbegriff zu.

Definition: Ein s -Minor einer beliebigen (rechteckigen) Matrix A ist die Determinante einer $s \times s$ -Untermatrix von A , die durch Streichen gewisser Spalten und Zeilen von A entstanden ist.

Satz 7.2.5 Die größte Zahl s , für die es einen von Null verschiedenen s -Minor von A gibt, ist gleich dem Rang von A .

Beweis: Sei oBdA $A = \begin{pmatrix} B & \star \\ \star & \star \end{pmatrix}$ in Blockmatrizen zerlegt, die linke obere Untermatrix B sei eine $s \times s$ -Matrix mit $\det(B) \neq 0$. Dann sind die Spalten von B linear unabhängig, also sind auch die ersten s Spalten von A linear unabhängig, demnach ist $rg(A) \geq s$. Wir zeigen nun: Wenn $rg(A) = r$ ist, so existiert ein von Null verschiedener r -Minor von A . Sei $A = (a_1, \dots, a_n)$, oBdA sei $\{a_1, \dots, a_r\}$ linear unabhängig. Wir setzen $B = (a_1, \dots, a_r)$, dann ist natürlich $rg(B) = r$, also besitzt B auch r linear unabhängige Zeilen, diese Zeilen aus B bilden zusammen eine $r \times r$ -Untermatrix vom (Zeilen-)Rang r , also mit von Null verschiedener Determinante. \square

Es folgen einige Resultate über die Determinanten spezieller Matrizen.

Wir unterteilen eine Matrix A wie folgt in Teilmatrizen auf:

$$\begin{pmatrix} a & z \\ s & B \end{pmatrix},$$

wobei $B \in M_{n-1, n-1}$, $a \in R$, z eine Zeile und s eine Spalte ist. Wenn dann $a \neq 0$ ist, so gilt

$$\begin{pmatrix} a & z \\ s & B \end{pmatrix} \begin{pmatrix} 1 & -\frac{1}{a}z \\ 0 & E \end{pmatrix} = \begin{pmatrix} a & 0 \\ s & -\frac{1}{a}sz + B \end{pmatrix},$$

also

Satz 7.2.6

$$\det(A) = a \cdot \det\left(-\frac{1}{a}sz + B\right) = \frac{1}{a^{n-2}} \det(aB - sz). \square$$

Satz 7.2.7 Sei $A \in M_{nn}$ eine schiefsymmetrische Matrix (d.h. $A^T = -A$) und n eine ungerade Zahl, dann gilt $\det(A) = 0$.

Beweis: $\det(A) = \det(A^T) = \det(-A) = (-1)^n \det(A) = -\det(A)$. □

Satz 7.2.8 Sei $A \in M_{nn}$ eine schiefsymmetrische Matrix und n eine gerade Zahl, dann gilt $\det(A) \geq 0$.

Beweis: Die Diagonaleinträge einer schiefsymmetrischen Matrix sind Nullen. Wenn an der Stelle (1,2) etwas von Null verschiedenes steht, so überspringen wir die folgenden Operationen. Sei in A an der Stelle (i, j) ein von Null verschiedener Eintrag a vorhanden, und es sei P die Permutationsmatrix, die die Stellen j und 2 miteinander vertauscht. Dann gilt $\det(PAP) = \det(A)$ und PAP hat folgende Gestalt:

$$\begin{pmatrix} 0 & a & B \\ -a & 0 & \\ -B^T & & C \end{pmatrix}.$$

Wenn wir $S = \begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix}$ setzen, so gilt

$$\begin{aligned} & \begin{pmatrix} E & 0 \\ B^T S^{-1} & E \end{pmatrix} \begin{pmatrix} S & B \\ -B^T & C \end{pmatrix} \begin{pmatrix} E & -S^{-1}B \\ 0 & E \end{pmatrix} \\ &= \begin{pmatrix} S & B \\ 0 & B^T S^{-1}B + C \end{pmatrix} \begin{pmatrix} E & -S^{-1}B \\ 0 & E \end{pmatrix} \\ &= \begin{pmatrix} S & 0 \\ 0 & B^T S^{-1}B + C \end{pmatrix} \end{aligned}$$

und deren Determinante ist gleich $a^2 \cdot \det(B^T S^{-1}B + C)$, die Matrix C ist schiefsymmetrisch und es ist

$$(B^T S^{-1}B)^T = B^T S^{-1T}B = -B^T S^{-1}B,$$

also ist die „Restmatrix“ schiefsymmetrisch und wir erhalten das Resultat durch Induktion.

Satz 7.2.9 (Vandermondsche Determinante)

$$\det \begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ & & \dots & & \\ & & & & \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{pmatrix} = \prod_{i>j} (x_i - x_j).$$

Beweis: Wir subtrahieren das x_1 -fache der i -ten Spalte von der $(i+1)$ -ten und erhalten

$$\det \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & x_2 - x_1 & x_2(x_2 - x_1) & \dots & x_2^{n-2}(x_2 - x_1) \\ & & \dots & & \\ 1 & x_n - x_1 & x_n(x_n - x_1) & \dots & x_n^{n-2}(x_n - x_1) \end{pmatrix},$$

deren Determinante hat den Wert

$$(x_2 - x_1) \cdots (x_n - x_1) \det \begin{pmatrix} 1 & x_2 & x_2^2 & \dots & x_2^{n-2} \\ & & \dots & & \\ 1 & x_n & x_n^2 & \dots & x_n^{n-2} \end{pmatrix}$$

und wir erhalten wieder durch Induktion das Resultat. \square

Zum Schluß wollen wir noch einem Endomorphismus $f : V \rightarrow V$ eines Vektorraums V eine Determinante zuordnen. Dazu wählen wir irgendeine Basis B von V , sei $M = A_{BB}(f)$ die Darstellungsmatrix von f ; wir können nun $\det(M)$ bilden, aber hängt das nicht sehr von der Wahl der Basis B ab? Sei also M' die Darstellungsmatrix von f bezüglich einer anderen Basis von V , dann „unterscheiden“ sich M und M' um eine reguläre Matrix X :

$$M = X^{-1}M'X$$

und damit ist $\det(M) = \det(X)^{-1} \det(M') \det(X) = \det(M')$ von der Wahl der Basis unabhängig. Wir setzen also $\det(f) = \det(M)$.

Interpolation

Wir wenden die Cramersche Regel an, um explizite Formeln zu gewinnen.

Es sei

$$s_n = 1 + 2 + 3 + \dots + n$$

die Summe der ersten n natürlichen Zahlen. Es ist $s_0 = 0, s_1 = 1, s_2 = 3$; das genügt schon, wenn wir uns erinnern, daß es eine Formel der Form

$$s_n = an^2 + bn + c$$

gab, d.h. s_n ist ein Polynom 2-ten Grades in n .

Wir setzen die ersten Werte ein:

$$a \cdot 0 + b \cdot 0 + c = 0$$

$$a \cdot 1 + b \cdot 1 + c = 1$$

$$a \cdot 4 + b \cdot 2 + c = 3$$

Cramersche Regel:

$$D_N = \begin{vmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \end{vmatrix} = 2 - 4 = -2,$$

$$D_a = \begin{vmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 3 & 2 & 1 \end{vmatrix} = 2 - 3 = -1, \quad D_b = \begin{vmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 3 & 1 \end{vmatrix} = 3 - 4 = -1, \quad D_c = \begin{vmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 4 & 2 & 3 \end{vmatrix} = 0.$$

Damit ergibt sich $a = 1/2, b = 1/2, c = 0$, also $s_n = n^2/2 + n/2 = \frac{n(n+1)}{2}$.

Beispiel: Wir bestimmen $t_n = \sum_{i=0}^n i^2$, dies ist ein Polynom in n vom Grad 3:

$$\sum_{i=1}^n i^2 = a_3 n^3 + a_2 n^2 + a_1 n + a_0;$$

wir setzen $n = 0, 1, 2, 3$:

$$\begin{array}{rccccrc} 0a_3 & +0a_2 & +0a_1 & +a_0 & = & 0 \\ a_3 & +a_2 & +a_1 & +a_0 & = & 1 \\ 8a_3 & +4a_2 & +2a_1 & +a_0 & = & 5 \\ 27a_3 & +9a_2 & +3a_1 & +a_0 & = & 14 \end{array}$$

Bei MATLAB schreiben wir

```
v = [0 1 2 3]
a = vander(v)
b = [0; 1; 5; 14]
x = a \ b
rats(x)
```

und erhalten

$$\frac{n(n+1)(2n+1)}{6}.$$

7.3 Zweidimensionale Geometrie II

Wir betrachten wieder \mathbb{R}^2 als affinen Raum. Mit $|a, b|$ bezeichnen wir die Determinante mit den Spalten $a, b \in \mathbb{R}^2$. Wenn $b \neq o$, so ist $|a, b| = 0$ gdw. $a = rb$ für ein $r \in \mathbb{R}$.

Wir folgen wieder Koecher/Krieg:

Lemma 7.3.1 *Drei Punkte a, b, c liegen genau dann auf einer Geraden, wenn es $r, s, t \in \mathbb{R}$ gibt mit*

$$ra + sb + tc = o \text{ und } r + s + t = 0.$$

Beweis: Wenn c auf der Geraden durch a und b liegt, so gibt es ein $s \in \mathbb{R}$ mit $c = a + s(b - a)$. \square

Die Punkte der Geraden $G_{a,v} = \{x|a + rv, r \in \mathbb{R}\}$ sind dann genau die x mit $|x, v| = |a, v|$, denn dies ist äquivalent zu $|x - a, v| = 0$, also $x - a = rv$, d.h. $x = a + rv$.

Lemma 7.3.2 *Zwei nichtparallele Geraden $G_{a,u}, G_{b,v}$ schneiden sich im Punkt*

$$s = \frac{1}{|u, v|} (|b, v|u - |a, u|v).$$

Beweis: u, v sind linear unabhängig, wir machen den Ansatz $s = xu + yv$ mit $x, y \in \mathbb{R}$. Nach dem Obigen ist

$$s \in G_{a,u} \Leftrightarrow |a, u| = |s, u| = y|v, u|,$$

$$s \in G_{b,v} \Leftrightarrow |b, v| = |s, v| = x|u, v|$$

und daraus folgt die Behauptung. \square

Mit $u = b - a, v = d - c$ erhalten wir den symmetrischen Ausdruck

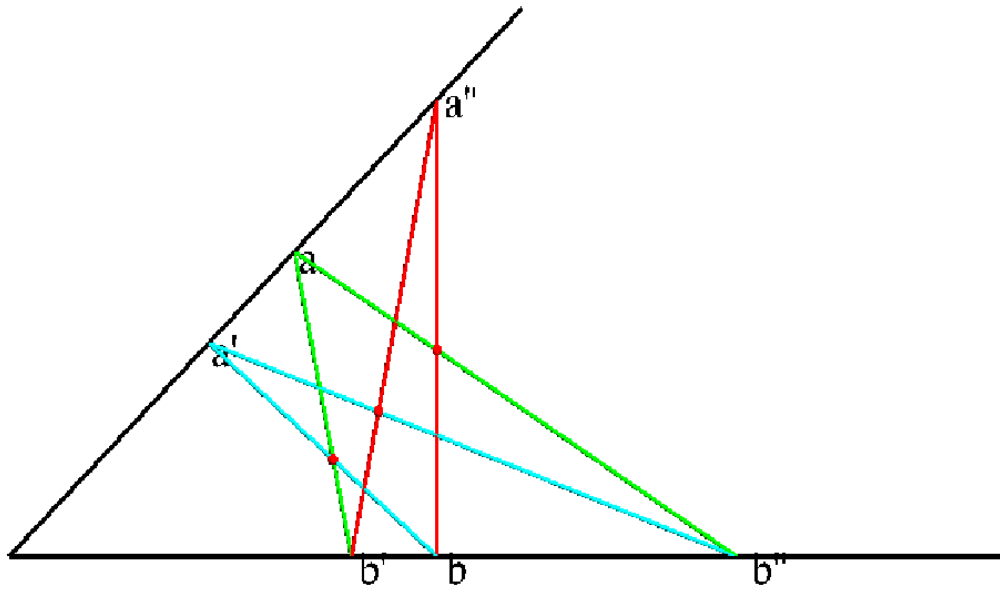
$$(a \vee b) \cap (c \vee d) = \frac{1}{|b - a, d - c|} (|c, d|(b - a) - |a, b|(d - c))$$

und speziell

$$(ra \vee sb) \cap (ta \vee ub) = \frac{1}{|sb - ra, ub - ta|} (|ta, ub|(sb - ra) - |ra, sb|(ub - ta)),$$

der Nenner ist gleich $|sb, -ta| + |-ra, ub| = st|b, -a| + ru|a, b| = (ru - st)|a, b|$, der Vektor ist gleich $(-ru(ta + sb) + st(ra + ub))|a, b|$, also erhalten wir den Schnittpunkt

$$\frac{1}{ru - st} (ru(ta + sb) - st(ra + ub))$$



Satz 7.3.1 (Pascal) Seien F, G nichtparallele Geraden mit Schnittpunkt o sowie $a, a', a'' \in F \setminus G$ und $b, b', b'' \in G \setminus F$ paarweise verschiedene Punkte, so daß die Schnittpunkte

$$P = (a \vee b') \cap (a' \vee b),$$

$$Q = (a' \vee b'') \cap (a'' \vee b'),$$

$$R = (a \vee b'') \cap (a'' \vee b)$$

existieren. Dann liegen P, Q, R auf einer Geraden.

Beweis: Es seien $a' = ra$, $a'' = sa$, $b' = tb$, $b'' = ub$. Wir setzen $k = a + b$, $l = ra + tb$, $m = sa + ub$, dann gilt

$$P = (ra \vee b) \cap (a \vee tb) = \frac{1}{rt-1}(rt(a+b) - (ra+tb)) = \frac{1}{rt-1}(rtk - l),$$

$$Q = (sa \vee tb) \cap (ra \vee ub) = \frac{1}{su-tr}(su(ra+tb) - tr(sa+ub)) = \frac{1}{su-tr}(sul - trm),$$

$$R = (a \vee ub) \cap (sa \vee b) = \frac{1}{1-us}((sa+ub) - us(a+b)) = \frac{1}{1-us}(m - usl),$$

Es folgt

$$su(rt-1)P + (su-tr)Q + rt(1-us)R = o$$

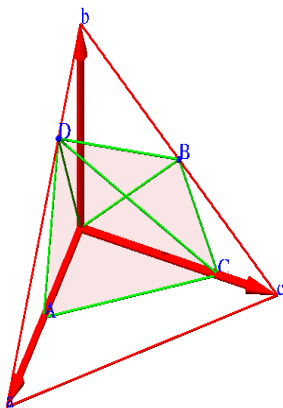
und die Koeffizientensumme ist Null, also liegen P, Q, R auf einer Geraden. \square

7.4 Abgeschnittene Pyramiden

Ein Spat ist ein durch drei linear unabhängige Vektoren a, b, c im \mathbb{R}^3 aufgespannter Körper, sei Volumen ist gleich $\det(a, b, c)$. Wir interessieren uns für einen halben Spat, also eine dreisetige Pyramide:



Die Seitenhalbierenden zweier Dreiecksflächen bestimmen zusammen mit der gegenüberliegenden Kante eine Ebene, diese zerlegt die Pyramide in zwei volumengleiche Pyramiden. Allgemeiner gilt:



Satz 7.4.1 *Seien A und B die Mittelpunkte zweier gegenüberliegender Seiten der Pyramide, dann zerlegt jede Ebene durch diese Punkte die Pyramide in zwei volumengleiche Teile.*

Beweis: Es sei also $A = \frac{1}{2}a$, $B = \frac{1}{2}(b + c)$ und $C = z \cdot c$ ein Punkt auf der dritten Kante. Wir bestimmen den Schnittpunkt D der durch A, B, C bestimmten Ebene mit

der vierten Kante: Es ist

$$D = xa + yb = C + r(A - C) + s(B - C) = (z - rz + \frac{1}{2}s - zs)c + \frac{1}{2}ra + \frac{1}{2}b,$$

sowie $x + y = 1$, daraus ergibt sich

$$\begin{array}{rcl} zr & + & (z - \frac{1}{2})s = z \\ r & + & s = 2 \end{array}$$

also $r = 2 - 2z$, $s = 2z$, $x = 1 - z$, $y = z$.

Das durch das Viereck abgeschnittene Pyramidenstück setzt sich aus den durch DAC , DBC und bBD gegebenen Pyramiden zusammen, die entsprechenden Determinanten sind

$$\begin{vmatrix} 1-z & z & 0 \\ 0 & 0 & z \\ \frac{1}{2} & 0 & 0 \end{vmatrix} = \frac{1}{2}z^2, \quad \begin{vmatrix} 1-z & z & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & z \end{vmatrix} = \frac{1}{2}z - \frac{1}{2}z^2 \quad \begin{vmatrix} 0 & 1 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ i-z & z & 0 \end{vmatrix} = \frac{1}{2} - \frac{1}{2}z,$$

deren Summe ist gleich $\frac{1}{2}$. □

Kapitel 8

Eigenwerte und Eigenvektoren

Sei $f : V \rightarrow V$ ein Endomorphismus des Vektorraums V . Wir fragen uns, ob es einen Vektor $v \in V$ gibt, der unter der Wirkung von f seine Richtung nicht ändert, für den es also eine Zahl z gibt, so daß $f(v) = zv$ gilt. Solch einen Vektor v nennen wir einen Eigenvektor von f , die Zahl z heißt der zugehörige Eigenwert. (Trivialerweise hat der Nullvektor die oben genannte Eigenschaft, ihn wollen wir aber ausdrücklich nicht als Eigenvektor ansehen.)

Sei nun z ein Eigenwert von f , d.h. es gibt ein $v \neq o$ aus V mit $f(v) = zv$. Dann sei V_z die Menge aller $v \in V$ mit $f(v) = zv$ (einschließlich des Nullvektors), V_z heißt der Eigenraum von f zum Eigenwert z .

Dies wird durch das folgende Lemma gerechtfertigt:

Lemma 8.0.1 V_z ist ein Unterraum von V .

Beweis: Seien v_1, v_2 Eigenvektoren von f (oder Null), d.h. $f(v_i) = zv_i$, dann gilt $f(v_1 + rv_2) = f(v_1) + rf(v_2) = zv_1 + rzv_2 = z(v_1 + rv_2)$ für beliebige $r \in R$. \square

Satz 8.0.2 Seien z_1, \dots, z_m paarweise verschiedene Eigenwerte von f und v_1, \dots, v_m zugehörige Eigenvektoren, dann ist $\{v_1, \dots, v_m\}$ linear unabhängig.

Beweis: Induktion über m : $\{v_1\}$ ist linear unabhängig, da $v_1 \neq o$ ist.

Sei der Satz also für $m - 1$ verschiedene Eigenvektoren bewiesen. Wir nehmen an, daß $v_m = r_1v_1 + \dots + r_{m-1}v_{m-1}$ ist und wenden f an:

$$\begin{aligned} f(v_m) &= z_m v_m \\ &= z_m r_1 v_1 + \dots + z_m r_{m-1} v_{m-1} \\ &= f(r_1 v_1 + \dots + r_{m-1} v_{m-1}) \\ &= z_1 r_1 v_1 + \dots + z_{m-1} r_{m-1} v_{m-1}. \end{aligned}$$

Aus der linearen Unabhängigkeit von $\{v_1, \dots, v_{m-1}\}$ folgt $z_i = z_m$ für $i = 1, \dots, m - 1$, ein Widerspruch. \square

Nach diesen abstrakten Betrachtungen wollen wir uns der Frage stellen, ob denn Eigenvektoren und -werte wirklich existieren (das sollte man eigentlich zuerst tun). Dazu übertragen wir die gesamte Problematik in die Sprache der Matrizen.

Definition: Sei A eine Matrix aus M_{nn} und $v = (v_1, \dots, v_n)^T \neq o$ ein Spaltenvektor aus R^n , dann heißt v Eigenvektor von A , wenn eine Zahl z existiert, so daß $Av = zv$ gilt. Die Zahl z heißt der zu v gehörige Eigenwert.

Die Bedingung $Av = zv$ ist äquivalent zu

$$(A - zE)v = o,$$

dies ist ein homogenes Gleichungssystem mit der Koeffizientenmatrix $A - zE$ und den Unbekannten v_1, \dots, v_n , wie wir wissen, existiert genau dann eine nichttriviale Lösung, wenn $rg(A - zE) < n$ ist. Dies ist wiederum genau dann der Fall, wenn

$$\det(A - zE) = 0$$

gilt.

Wenn wir z als Variable auffassen, so ist $\det(A - zE)$ ein Polynom in z vom Grade n , es wird als das charakteristische Polynom $c_A(z)$ von A bezeichnet.

Schauen wir uns das charakteristische Polynom einer Matrix genauer an, wir bezeichnen die Koeffizienten (bis aufs Vorzeichen) mit c_i :

$$c_A(z) = (-1)^n z^n + (-1)^{n-1} c_1 z^{n-1} + \dots + c_n.$$

Man sieht sofort, daß $c_n = \det(A)$ ist, daraus folgt, daß die Zahl 0 genau dann ein Eigenwert der Matrix A ist, wenn $\det(A) = 0$ ist. Weiter gilt $c_1 = a_{11} + a_{22} + \dots + a_{nn}$. Die Summe der Diagonalelemente von A , also c_1 , heißt die Spur $Sp(A)$ von A .

Sei nun $f : V \rightarrow V$ ein Endomorphismus, B eine Basis von V und $F = A_{BB}(f)$ die Darstellungsmatrix von f . Dann setzen wir $c_f(z) = c_F(z)$ und nennen dies das charakteristische Polynom von f . Zur Rechtfertigung beweisen wir das

Lemma 8.0.2 *Die Koeffizienten von $c_f(z)$ sind unabhängig von der Wahl der Basis B .*

Beweis: Sei C eine andere Basis von V und F' die entsprechende Darstellungsmatrix, dann gilt $F' = X^{-1}FX$ für eine gewisse reguläre Matrix X . Es gilt

$$\begin{aligned} c_{F'}(z) &= \det(X^{-1}FX - zE) \\ &= \det(X^{-1}(F - zE)X) \\ &= \det(X^{-1}) \det(F - zE) \det(X) \\ &= \det(X)^{-1} \det(X) c_F(z) \\ &= c_F(z). \quad \square \end{aligned}$$

Folgerung 8.0.1 $Sp(X^{-1}AX) = Sp(A)$. □

Das folgende Lemma ist leicht zu beweisen, folgt aber nicht aus der obigen Folgerung.

Lemma 8.0.3 *Für beliebige (nicht notwendig reguläre) Matrizen A, B gilt*

$$Sp(AB) = Sp(BA). \quad \square$$

Definition: Sei $A \in M_{nn}$. Die $(n-1)$ -reihige Matrix A_{ik} möge aus A durch Streichen der i -ten Zeile und der k -ten Spalte entstehen. Die Determinante $\det(A_{ik})$ heißt dann ein $(n-1)$ -Minor von A . Seien weiter $I = \{i_1, \dots, i_{n-t}\}$ und $K = \{k_1, \dots, k_{n-t}\}$ zwei $(n-t)$ -elementige Mengen natürlicher Zahlen zwischen 1 und n . Die t -reihige Matrix A_{IK} möge aus A durch Streichen der Zeilen aus I und der Spalten aus K hervorgehen. Dann heißt die Determinante $\det(A_{IK})$ ein t -Minor von A . Ein t -Hauptminor ist ein Minor der Form $\det(A_{II})$, wo in A „dieselben“ Zeilen und Spalten gestrichen sind.

Satz 8.0.3 Sei $c_A(z) = (-1)^n z^n + (-1)^{n-1} c_1 z^{n-1} + \dots + c_n$. Dann ist c_i die Summe der i -Hauptminoren von A .

Beweis: Wir halten uns an die Leibnizsche Determinantendefinition: Zur Berechnung einer Determinante ist eine alternierende Summe zu bilden, deren Summanden Produkte sind, deren Faktoren jeweils aus verschiedenen Zeilen und aus verschiedenen Spalten der Matrix zu wählen sind. Den Term $(-1)^i z^i$ erhalten wir, wenn wir i Elemente $(a_{jj} - z)$, $j = k_1, \dots, k_i$ auf der Diagonalen wählen, für die restlichen Faktoren dürfen wir dann die Zeilen und die Spalten mit den Nummern k_1, \dots, k_i nicht mehr verwenden, wir können sie also auch streichen. Wenn wir alles zusammenfassen, was mit dem Produkt unserer festgehaltenen $(a_{jj} - z)$ multipliziert wird, erhalten wir einen i -Hauptminor von A . Wenn wir nun die Faktoren auf der Diagonalen variieren lassen, erhalten wir als Koeffizienten von $(-1)^i z^i$ gerade die Summe aller i -Hauptminoren. \square

Wenn wir davon ausgehen, daß die betrachteten Matrizen reelle Komponenten haben, dann sind die Koeffizienten des entsprechenden charakteristischen Polynoms auch reell, jedoch kann es durchaus vorkommen, daß nicht alle Eigenwerte (oder auch gar keiner) reell sind. Betrachten wir zum Beispiel eine Drehung um den Winkel w :

$$A = \begin{pmatrix} \cos w & \sin w \\ -\sin w & \cos w \end{pmatrix}.$$

Wenn w nicht gerade ein Vielfaches von 180° ist, gibt es keinen vom Nullvektor verschiedenen Vektor, der seine Richtung behält, wie es ein Eigenvektor tun müßte. Die beiden Eigenwerte von A sind ja gleich $\exp(\pm iw)$, also nicht reell.

Wenn wir auf die Existenz von Eigenwerten nicht verzichten wollen, müssen wir eventuell unseren Grundkörper erweitern, wir halten nicht am Körper \mathbb{R} der reellen Zahlen fest, sondern verwenden den Körper \mathbb{C} der komplexen Zahlen.

In besonderen Fällen können wir aber die Realität der Eigenwerte garantieren:

Satz 8.0.4 Wenn A eine symmetrische Matrix ist, so sind alle Eigenwerte von A reell.

Beweis: Sei $a + bi$ eine Nullstelle von $c_A(z) = \det(A - zE)$, dann gibt es einen Vektor (z_1, \dots, z_n) mit komplexen Komponenten, die nicht alle gleich Null sind, so daß

$$(A - (a + bi)E) \begin{pmatrix} z_1 \\ \dots \\ z_n \end{pmatrix} = \begin{pmatrix} 0 \\ \dots \\ 0 \end{pmatrix}$$

ist. Sei $z_k = x_k + iy_k$, x_k, y_k reell, dann gilt

$$\begin{aligned} \sum a_{kl}z_l - (a + bi)z_k &= 0 \\ &= \sum a_{kl}(x_l + iy_l) - (ax_k - by_k) - (bx_k + ay_k)i. \end{aligned}$$

Wir betrachten den Realteil:

$$\sum a_{kl}x_l - ax_k + by_k = 0,$$

wir multiplizieren dies mit y_k und addieren (über k). Den Imaginärteil

$$\sum a_{kl}y_l - bx_k - ay_k = 0$$

multiplizieren wir mit x_k und addieren ebenfalls. Wir erhalten

$$\sum (\sum a_{kl}x_ly_k - ax_ky_k + by_k^2) = 0$$

und

$$\sum (\sum a_{kl}y_lx_k - bx_k^2 - ax_ky_k) = 0.$$

Wir subtrahieren beide Terme und erhalten unter Beachtung von $a_{kl} = a_{lk}$

$$b \sum (x_k^2 + y_k^2) = 0,$$

nach Voraussetzung ist der zweite Faktor von Null verschieden, also muß $b = 0$ sein, d.h. unser Eigenwert ist reell. \square

Die Eigenwerte symmetrischer Matrizen sind nicht nur reell, sondern auch recht einfach zu berechnen. Wir erinnern uns daran, daß man eine symmetrische Matrix durch eine Transformation der Form

$$A \rightarrow X^TAX$$

(X ist eine reguläre Matrix) in eine Diagonalmatrix überführen kann, leider bleiben dabei die Eigenwerte im allgemeinen nicht erhalten.

Jedoch haben wir die Matrix A beim Jacobischen Diagonalisierungsverfahren mit Drehmatrizen der Form

$$J = \begin{pmatrix} 1 & & & \\ & \dots & & \\ & c & & s \\ & -s & & c \\ & & \dots & \\ & & & 1 \end{pmatrix}$$

transformiert, und die Matrix J hat die angenehme Eigenschaft, daß $J^T = J^{-1}$ ist, d.h. die Eigenwerte von A und von J^TAJ stimmen überein. Somit haben wir mit dem Jacobischen Verfahren ein Näherungsverfahren zur Berechnung der Eigenwerte symmetrischer Matrizen gefunden.

Sei $A \in M_{nn}$ eine Matrix mit den Eigenwerten z_1, \dots, z_n und zugehörigen Eigenvektoren v_1, \dots, v_n , also

$$Av_i = z_i v_i.$$

Wir wissen, daß $\{v_1, \dots, v_n\}$ linear unabhängig sind, wenn die z_i paarweise verschieden sind, also:

Satz 8.0.5 *Wenn $A \in M_{nn}$ lauter verschiedene Eigenwerte hat, so besitzt \mathbb{R}^n eine Basis aus Eigenvektoren von A .*

Diese Bedingung ist aber nicht notwendig, wie wir an folgendem Beispiel sehen: Sei

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

ihr charakteristisches Polynom

$$c_A(z) = -z^3 + 5z^2 - 8z + 4 = (z - 1)(z - 2)^2$$

hat die Zahl $z = 2$ als doppelte Nullstelle, dennoch bilden die Eigenvektoren

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

eine Basis des \mathbb{R}^3 .

Es gibt aber nicht zu jeder Matrix eine Basis aus Eigenvektoren.

Sei $\begin{pmatrix} -3 & 2 \\ -2 & 1 \end{pmatrix}$, es ist $c_A(z) = z^2 + 2z + 1 = (z + 1)^2$, aber $A - 1E = \begin{pmatrix} -2 & 2 \\ -2 & 2 \end{pmatrix}$ hat den Rang 1, also hat A nur einen eindimensionalen Eigenraum.

Wir können solche Matrizen, für die eine Basis aus Eigenvektoren existiert, genau beschreiben:

Satz 8.0.6 *Zur $n \times n$ -Matrix A existiert genau dann eine Basis des \mathbb{R}^n aus Eigenvektoren, wenn es eine invertierbare Matrix V gibt, so daß $V^{-1}AV = D$ eine Diagonalmatrix ist.*

Beweis: Die Matrix V habe die Spalten (v_1, \dots, v_n) , dann gilt

$$AV = A(v_1, \dots, v_n) = (Av_1, \dots, Av_n) = (v_1, \dots, v_n) \begin{pmatrix} z_1 & & 0 \\ & \dots & \\ 0 & & z_n \end{pmatrix} = (z_1 v_1, \dots, z_n v_n),$$

also $Av_i = z_i v_i$, also sind die Vektoren v_1, \dots, v_n Eigenvektoren von A , und als Spalten einer invertierbaren Matrix sind sie linear unabhängig. \square

Allgemeiner gilt der folgende

Satz 8.0.7 *Das charakteristische Polynom der Matrix $A \in M_{nn}$ habe in R n Nullstellen (d.h. $c_A(z) = \prod_{i=1}^n (z - z_i)$, dies ist insbesondere für $R = \mathbf{C}$ stets erfüllt). Dann gibt*

es eine reguläre Matrix X , so daß $X^{-1}AX = \begin{pmatrix} r_1 & \dots & \star \\ 0 & \dots & \star \\ 0 & \dots & r_n \end{pmatrix}$ eine Dreiecksmatrix ist.

Beweis: Wir führen die Induktion über n ; sei für $(n-1)$ -reihige Matrizen schon alles bewiesen.

Sei z_1 ein Eigenwert von A und $v_1 \in \mathbf{C}^n$ ein zugehöriger Eigenvektor ($v_1 \neq 0$). Wir ergänzen v_1 zu einer Basis $\{v_1, \dots, v_n\}$ des \mathbf{C}^n , nun sei X die Matrix mit den Spalten v_1, \dots, v_n . Dann gilt

$$AX = A(v_1, \dots, v_n) = (Av_1, \dots, Av_n) = (z_1v_1, Av_2, \dots, Av_n),$$

also ist

$$X^{-1}AX = \begin{pmatrix} z_1 & \dots \\ 0 & B \end{pmatrix}$$

wobei B eine $(n-1)$ -reihige Matrix ist. Nach Voraussetzung gibt es eine reguläre Matrix Y , so daß $Y^{-1}BY$ eine Dreiecksmatrix ist. Wir setzen

$$X' = X \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \dots & & Y & \\ 0 & & & \end{pmatrix},$$

dann ist $X'^{-1}AX'$ eine Dreiecksmatrix. □

Wir rechnen ein nichttriviales Beispiel durch:

$$A = \begin{pmatrix} -1 & 2 & 3 \\ -2 & 3 & 7 \\ 0 & 0 & 1 \end{pmatrix}, \quad c_A(z) = (1-z)(z^2 - 2z + 1) = -(z-1)^3,$$

wir erhalten einen eindimensionalen Eigenraum, der z.B. von Vektor $v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ aufgespannt wird. Wir ergänzen v_1 (willkürlich) durch $v_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ und $v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ zu einer

Basis von R^3 schreiben diese Vektoren in die Matrix

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

deren Inverse ist

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

Dann ist

$$B^{-1}AB = \begin{pmatrix} 1 & 5 & 7 \\ 0 & -3 & -4 \\ 0 & 4 & 5 \end{pmatrix}$$

schon „fast“ eine Dreiecksmatrix.

Nun befassen wir uns mit der Untermatrix $A' = \begin{pmatrix} -3 & -4 \\ 4 & 5 \end{pmatrix}$, die wir als im Raum $U = \mathcal{L}(v_2, v_3)$ operierend auffassen, d.h. wir suchen *dort* eine Basis, so daß diese Matrix in Dreiecksgestalt transformiert wird. Zum Eigenwert $z = 1$ finden wir einen Eigenvektor $w_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} = v_2 - v_3$, den wir durch $w_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = v_2 + v_3$ zu einer Basis von U ergänzen. Wir bilden wieder eine Matrix $B' = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$, deren Inverse ist $B'^{-1} = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ und wir erhalten die Dreiecksmatrix $B'^{-1}A'B' = \begin{pmatrix} 1 & -8 \\ 0 & 1 \end{pmatrix}$. Schließlich bilden wir

$$C = B \begin{pmatrix} 1 & 0 \\ 0 & B' \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Am Ende erhalten wir

$$C^{-1}AC = \begin{pmatrix} 1 & -2 & 12 \\ 0 & 1 & -8 \\ 0 & 0 & 1 \end{pmatrix},$$

die zugehörige Basis ist $\{v_1, v_2 - v_3, v_2 + v_3\}$.

Satz 8.0.8 1. Sei $X^{-1}AX = \begin{pmatrix} r_1 & & * \\ & \dots & \\ 0 & & r_n \end{pmatrix}$ eine Dreiecksmatrix. Dann sind r_1, \dots, r_n

die Eigenwerte von A .

2. Wenn r_1, \dots, r_n die Eigenwerte von A sind, so sind die Eigenwerte von A^k gerade die Zahlen r_1^k, \dots, r_n^k . (Dies gilt, falls es einen Sinn hat, auch für negatives k .)

Beweis: 1. Die Determinante von $X^{-1}AX - zE$ hat den Wert $(r_1 - z) \dots (r_n - z)$.

2. Bei der Multiplikation von Dreiecksmatrizen multiplizieren sich die Diagonalelemente.

□

Der folgende Satz ist eigentlich zuunrecht nach Cayley benannt, denn von diesem wurde er nur für 2- oder 3-reihige Matrizen bewiesen, das war aber der Stil der Zeit:

Satz 8.0.9 (Hamilton-Cayley) Sei A eine n -reihige Matrix und $c_A(z) = \sum b_{n-i}z^i$ ihr charakteristisches Polynom, dann ist $\sum b_{n-i}A^i = 0$ die Nullmatrix aus M_{nn} .

(Wenn man eine Matrix in ihr charakteristisches Polynom einsetzt, kommt null heraus.)

Wir bemerken, daß Cayley de Satz in der naheliegenden, wenn auch unsinnigen Form „ $|A - A| = 0$ “ formulierte.

Beweis: Seien z_1, \dots, z_n die Eigenwerte von A und z eine von den z_k verschiedene Zahl. Dann ist $B = A - zE$ eine reguläre Matrix, sie besitzt also eine Inverse und diese hat, wie wir früher gesehen haben, die Gestalt

$$(A - zE)^{-1} = \frac{1}{\det(A - zE)} \begin{pmatrix} b_{11} & \cdots & b_{n1} \\ \vdots & \ddots & \vdots \\ b_{1n} & \cdots & b_{nn} \end{pmatrix},$$

die b_{ij} sind Minoren von $A - zE$. Wir setzen

$$B = \begin{pmatrix} b_{11} & \cdots & b_{n1} \\ \vdots & \ddots & \vdots \\ b_{1n} & \cdots & b_{nn} \end{pmatrix} = B_{n-1}z^{n-1} + \cdots + B_1z + B_0,$$

dabei sollen die B_i von z unabhängige Matrizen sein. Es gilt also

$$\det(A - zE)E = (A - zE)B$$

oder ausführlicher

$$(z^n + b_1z^{n-1} + b_2z^{n-2} + \cdots + b_n)E = (A - zE)(B_{n-1}z^{n-1} + \cdots + B_1z + B_0).$$

Wir vergleichen die Koeffizienten von z^i und erhalten

$$\begin{aligned} b_n E &= AB_0 \\ b_{n-1} E &= AB_1 - B_0 \\ b_{n-2} E &= AB_2 - B_1 \\ &\dots \\ b_1 E &= AB_{n-1} - B_{n-2} \\ E &= -B_{n-1}. \end{aligned}$$

Wir multiplizieren die Gleichungen von links mit $E, A, A^2, \dots, A^{n-1}, A^n$ und addieren alles:

$$A^n + b_1 A^{n-1} + \cdots + b_{n-1} A + b_n E = 0E. \square$$

Man kann den Satz von Hamilton-Cayley zur Berechnen der Matrixinversen nutzen:

Sei $A = \begin{pmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{pmatrix}$, dann ist $c_A(z) = z^3 - 5z^2 + 8z - 4 = (z - 1)(z - 2)^2$. Aus

$$A^3 - 5A^2 + 8A - 4E = 0 \text{ folgt } A^2 - 5A + 8E = 4A^{-1}. \text{ Es ist } A^2 = \begin{pmatrix} -2 & 0 & -6 \\ 3 & 4 & 3 \\ 3 & 0 & 7 \end{pmatrix},$$

$$\text{also } A^{-1} = \frac{1}{4} \begin{pmatrix} 6 & 0 & 4 \\ -2 & 2 & -2 \\ -2 & 0 & 0 \end{pmatrix}.$$

Schließlich wollen wir ein Verfahren behandeln, daß es, wenn man Glück hat, gestattet, Eigenvektoren ohne Kenntnis von Eigenwerten zu berechnen:

Sei die Matrix $A - xE$ regulär, also x kein Eigenwert, und sei $w_0 \in \mathbb{R}^n$ beliebig. Wir lösen das Gleichungssystem

$$(A - xE)v_i = w_{i-1}$$

und setzen $w_i = \frac{1}{a_i}v_i$, wo a_i die größte Komponente von v_i ist. Unter bestimmten Voraussetzungen konvergiert v_i gegen einen Eigenvektor von A :

\mathbb{R}^n besitze eine Basis $\{b_1, \dots, b_n\}$ aus Eigenvektoren von A , die zugehörigen Eigenwerte seien z_1, \dots, z_n und es sei $w_0 = \sum r_i b_i$. Dann hat $A - xE$ die Eigenwerte $z_1 - x, \dots, z_n - x$ und $(A - xE)^{-1}$ hat die Eigenwerte $\frac{1}{z_1 - x}, \dots, \frac{1}{z_n - x}$, also ist

$$(A - xE)^{-1}b_i = \frac{1}{z_i - x}b_i$$

und damit ist

$$v_1 = (A - xE)^{-1} \sum r_i b_i = \sum \frac{r_i}{z_i - x} b_i,$$

also

$$w_k = \frac{1}{a_1 \cdots a_k} \sum \frac{r_i}{(z_i - x)^k} b_i.$$

Wenn nun x dichter an z_i als an den anderen Eigenwerten liegt, so überwiegt dieser Summand, also konvergiert w_k gegen b_i .

Rekursive Folgen

Wir betrachten ein Beispiel:

Seien Zahlen u_0, u_1 gegeben, die nächsten Folgenglieder seien durch

$$u_{n+1} = u_n + 6u_{n-1}$$

gegeben. Wir suchen nach einer expliziten Formel für u_n .

Es sei

$$A = \begin{pmatrix} 1 & 6 \\ 1 & 0 \end{pmatrix},$$

dann gilt

$$\begin{pmatrix} u_{n+1} \\ u_n \end{pmatrix} = \begin{pmatrix} 1 & 6 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} u_n \\ u_{n-1} \end{pmatrix} = \begin{pmatrix} 1 & 6 \\ 1 & 0 \end{pmatrix}^n \begin{pmatrix} u_1 \\ u_0 \end{pmatrix}.$$

Wir bestimmen die Eigenwerte von A : $z^2 - z - 6 = 0$ ergibt $z_1 = 3, z_2 = -2$.

Als Eigenvektoren erhalten wir $\begin{pmatrix} 3 \\ 1 \end{pmatrix}$ für $z = 3$ und $\begin{pmatrix} -2 \\ 1 \end{pmatrix}$ für $z = -2$. Wir bilden die Matrix

$$M = \begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix}, \text{ es ist } M^{-1} = \frac{1}{5} \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix},$$

also

$$M^{-1}AM = \begin{pmatrix} 3 & 0 \\ 0 & -2 \end{pmatrix} = D$$

und damit $A = MDM^{-1}$, also

$$A^n = MD^nM^{-1} = \begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 3^n & 0 \\ 0 & (-2)^n \end{pmatrix} \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} \frac{1}{5}.$$

Wir müssen nicht alle Terme wirklich ausrechnen (es reicht die zweite Zeile); nach dem Zusammenfassen erhalten wir

$$u_n = \frac{1}{5}(3^n(u_1 + 2u_0) + (-2)^n(-u_1 + 3u_0)).$$

Aufgabe:

Für die Fibonacci-Folge $f_{n+1} = f_n + f_{n-1}$ mit den Anfangswerten 0, 1 verifizieren Sie

$$f_n = \frac{1}{\sqrt{5}}\left(z^n - \frac{(-1)^n}{z^n}\right), \quad z = 1/2 + 1/2\sqrt{5} = 1.616\dots$$

Kapitel 9

Komplexe Zahlen, Quaternionen usw.

Zum Beginn wollen wir die Eigenschaften einer speziellen Sorte von 2×2 -Matrizen über dem Körper \mathbb{R} der reellen Zahlen untersuchen.

Es sei

$$\mathcal{C} = \left\{ \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \mid a, b \in \mathbb{R} \right\},$$

diese Menge bildet offenbar einen zweidimensionalen \mathbb{R} -Vektorraum. Wir stellen fest, daß auch das Produkt zweier Matrizen aus \mathcal{C} ein Element von \mathcal{C} ist:

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \begin{pmatrix} c & d \\ -d & c \end{pmatrix} = \begin{pmatrix} ac - bd & ad + bc \\ -bc - ad & ac - bd \end{pmatrix}.$$

Für $A, B \in \mathcal{C}$ gilt $AB = BA$, es ist $\det \begin{pmatrix} a & b \\ -b & a \end{pmatrix} = a^2 + b^2$, also ist jede von der Nullmatrix verschiedene Matrix aus \mathcal{C} invertierbar und die Inverse $\frac{1}{a^2+b^2} \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ ist wieder ein Element aus \mathcal{C} . Also ist \mathcal{C} ein Körper. Die Matrizen

$$E = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ und } I = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

bilden eine Basis des Vektorraums \mathcal{C} , es gilt $E^2 = E$ und $I^2 = -E$, also ist die Zuordnung $k : \mathbb{C} \rightarrow \mathcal{C}$ mit $k(a + bi) = aE + bI$ ein Isomorphismus.

Die komplexen Zahlen vom Betrag 1 sind von der Form $\cos(\alpha) + i \sin(\alpha)$, ihnen entsprechen die Drehmatrizen $\begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix}$. Seien nun $\begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix}$ und $\begin{pmatrix} \cos(\beta) & \sin(\beta) \\ -\sin(\beta) & \cos(\beta) \end{pmatrix}$ zwei Drehmatrizen, dann gehört zu ihrem Produkt die Drehung um den Winkel $\alpha + \beta$, aus dieser Produktmatrix liest man die Additionstheoreme für die Winkelfunktionen ab:

$$\begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix} \begin{pmatrix} \cos(\beta) & \sin(\beta) \\ -\sin(\beta) & \cos(\beta) \end{pmatrix}$$

$$\begin{aligned}
&= \begin{pmatrix} \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta) & \cos(\alpha)\sin(\beta) + \sin(\alpha)\cos(\beta) \\ -\cos(\alpha)\sin(\beta) + \sin(\alpha)\cos(\beta) & \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta) \end{pmatrix} \\
&= \begin{pmatrix} \cos(\alpha + \beta) & \sin(\alpha + \beta) \\ -\sin(\alpha + \beta) & \cos(\alpha + \beta) \end{pmatrix}.
\end{aligned}$$

Die Konstruktion des Körpers \mathbb{C} der komplexen Zahlen als Körper von Matrizen kann man wie folgt verallgemeinern:

Es sei

$$\mathcal{H} = \left\{ \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \mid a, b \in \mathbb{C} \right\}.$$

Satz 9.0.10 Für $h, g \in \mathcal{H}, r \in \mathbb{R}$ gilt

1. $h + g \in \mathcal{H}$,
2. $-h \in \mathcal{H}$ (also ist \mathcal{H} eine additive Untergruppe von $M_{22}(\mathbb{C})$),
3. $rh \in \mathcal{H}$ (also ist \mathcal{H} ein \mathbb{R} -Vektorraum, es ist $\dim_{\mathbb{R}}(\mathcal{H}) = 4$),
4. $hg \in \mathcal{H}$,
5. $h^{-1} \in \mathcal{H}$ (man könnte meinen, daß \mathcal{H} ein Körper ist; vergleichen Sie die Körperaxiome auf S. 1, aber:)
6. das Kommutativgesetz der Multiplikation gilt nicht.

Beweis: Wir zeigen nur 4):

$$\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \begin{pmatrix} c & d \\ -\bar{d} & \bar{c} \end{pmatrix} = \begin{pmatrix} ac - b\bar{d} & ad + b\bar{c} \\ -\bar{b}c - \bar{a}\bar{d} & \bar{a}\bar{c} - \bar{b}\bar{d} \end{pmatrix}. \quad \square$$

Eine Menge, in der eine Addition und eine Multiplikation definiert ist, so daß außer dem Kommutativgesetz der Multiplikation alle Körperaxiome gelten, nennt man einen Schiefkörper oder eine Divisionsalgebra.

Satz 9.0.11 Die Matrizen

$$E = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, I = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, J = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, K = \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix}$$

bilden eine \mathbb{R} -Basis von \mathcal{H} und es gilt

$$E^2 = E, \quad I^2 = J^2 = K^2 = -E,$$

$$\begin{aligned}
IJ &= K, & JK &= I, & KI &= J, \\
JI &= -K, & KJ &= -I, & IK &= -J.
\end{aligned}$$

Den Beweis führt man durch Nachrechnen. \square

Wir wollen die Schreibweise vereinfachen: Wir setzen $E = 1, I = i$ (also $L(E, I) = L(1, i) = \mathbb{C}$) und weiter $J = j, K = k$ und bezeichnen den von $1, i, j, k$ erzeugten Vektorraum mit

$$\mathbb{H} = \{a + bi + cj + dk \mid a, b, c, d \in \mathbb{R}\},$$

die Elemente von \mathbb{H} heißen Quaternionen. ¹ Dieser Schiefkörper wurde von Hamilton ² im Jahre 1843 entdeckt, nachdem er jahrelang vergeblich versucht hatte, eine umkehrbare Multiplikation in einem dreidimensionalen Vektorraum zu definieren. Da es sich bei der Quaternionenmultiplikation um die Multiplikation spezieller Matrizen handelt, ist die Gültigkeit des Assoziativgesetzes völlig klar. Das konnte Hamilton aber nicht wissen, denn die Matrixmultiplikation wurde erst 1858 eingeführt. An seinen Sohn schreibt er 1865 kurz vor seinem Tode (Math. Papers 3, p. XV): „Every morning, as my coming down to breakfast, you used to ask me: 'Well, Papa, can you multiply triplets?' Whereto I was always obliged to reply, with a sad shake of the head: 'No, I can only add and subtract them'.“ ³

Der Durchbruch gelang HAMILTON am 16. Oktober 1843 auf dem Wege zur Sitzung der Royal Irish Academy; noch auf jener Sitzung kündigt er seine Erfindung der Quaternionen an. Sein weiteres Leben widmet er ausschließlich der Erforschung der Quaternionen. Den Augenblick der Entdeckung hat er selbst 1858 wie folgt beschrieben (North. British Review 14, 1858): „... Tomorrow will be the fifteenth birthday of the Quaternions. They started into life, or light, full grown, on the 16th of October, 1843, as I was walking with Lady Hamilton to Dublin, and came up to Brougham Bridge. That is to say, I then and there felt the galvanic circuit of thought closed, and the sparks which fell from it were the fundamental equations between i, j, k exactly such as I have used them ever since. I pulled out, on the spot, a pocketbook, which still exists, and made an entry, on which, at the very moment, I felt that it might be worth my while to expend the labour of at least ten (or it might be fifteen) years to come. But then it is fair to say that this was because I felt a problem to have been at that moment solved, an intellectual want relieved, which had haunted me for at least fifteen years before . . .“ ⁴

¹Der Name entstammt der Bibel, Apostelgeschichte 12.4: es ist die Bezeichnung für 4 Rotten von je 4 Kriegsknechten des Herodes, die Petrus im Gefängnis bewachten. Um diese Zeit legte König Herodes die Hände an etliche von der Gemeinde, sie zu peinigen. Er tötete aber Jakobus, den Bruder des Jonas, mit dem Schwert. Und da er sah, daß es den Juden gefiel, fuhr er fort und fing Petrus auch. Da er ihn nun ergriff, legte er ihn ins Gefängnis und überantwortete ihn 4 Rotten, je von 4 Kriegsknechten, ihn zu bewachen.



²William Rowan Hamilton (1821–1895) Dublin, königlicher Astronom

³Google-Übersetzung: „Jeden Morgen, und mein herabkommen Frühstück, benutzt du mich fragen: 'Na, Papa, können Sie vermehren Drillinge?' Wohin war ich immer gezwungen, zu antworten, mit einem traurigen Kopfschütteln den Kopf: 'Nein, ich kann nur addieren und subtrahieren sie'.“

⁴Morgen wird der fünfzehnte Geburtstag der Quaternionen werden. Sie gestartet ins Leben, oder Licht, voll gewachsen, am 16. Oktober 1843, wie ich war Gehen mit Lady Hamilton in Dublin, und kam bis zu Brougham Bridge. Das heißt, ich dann und dort fühlte den galvanischen Kette von Denken

(zitiert aus „Zahlen“, Ebbinghaus et al., Springer 1983; der Beitrag stammt von M. Koecher und R. Remmert.)

Sei $a = a_1 + a_2i + a_3j + a_4k$ ein Quaternion, dann nennen wir a_1 seinen skalaren Anteil und $a_2i + a_3j + a_4k$ seinen vektoriellen Anteil, wir stellen uns den vektoriellen Anteil als einen „richtigen“ Vektor (einen Pfeil) im von i, j, k aufgespannten dreidimensionalen Raum vor, dabei möge (O, i, j, k) ein rechtwinkliges (kartesisches) Koordinatensystem sein.

Wir betrachten nun das Produkt zweier vektorieller Quaternionen $a = a_2i + a_3j + a_4k$ und $b = b_2i + b_3j + b_4k$:

$$(a_2i + a_3j + a_4k)(b_2i + b_3j + b_4k) = \\ -(a_2b_2 + a_3b_3 + a_4b_4) + (a_3b_4 - a_4b_3)i + (a_4b_2 - a_2b_4)j + (a_2b_3 - a_3b_2)k.$$

Den Ausdruck

$$\langle a, b \rangle = a_2b_2 + a_3b_3 + a_4b_4$$

nennt man das Skalarprodukt der Vektoren a und b , den Ausdruck

$$a \times b = (a_3b_4 - a_4b_3)i + (a_4b_2 - a_2b_4)j + (a_2b_3 - a_3b_2)k$$

nennt man das Vektorprodukt von a und b . Also gilt

$$ab = -\langle a, b \rangle + a \times b.$$

Wir bemerken

$$a \times b = \frac{1}{2}(ab - ba).$$

Wir werden sofort den Zusammenhang mit den Produkt-Konstruktionen herstellen, die Sie in der Schule kennengelernt haben.

Wenn wir ein Skalarprodukt durch

$$\langle a, b \rangle = |a| |b| \cos(\alpha)$$

eingeführen, wobei $|a|$ die „Länge“ des Vektors a ist und α den zwischen a und b eingeschlossenen Winkel bezeichnet, so haben wir die Übereinstimmung beider Definitionen zu zeigen.

Sei $A = O + a_2i + a_3j + a_4k$ und $B = O + b_2i + b_3j + b_4k$, wir betrachten das Dreieck OAB . Dessen Seiten haben folgende Längen:

$$|OA| = \sqrt{a_2^2 + a_3^2 + a_4^2},$$

geschlossen, und die Funken, fiel von ihr waren die Grundgleichungen zwischen i, j, k genau wie ich verwendet habe sie seitdem. Ich zog an Ort und Stelle, eine Brieftasche, die noch vorhanden ist, und einen Eintrag, auf dem, in dem Augenblick, ich fühlte, dass es sich lohnen könnte meinen, während der Arbeit an aufwenden mindestens zehn (oder es vielleicht fünfzehn) Jahre. Aber dann ist es fair zu sagen, dass dies, weil ich ein Problem empfunden worden zu haben war diesem Augenblick gelöst, ein Intellektueller, dass entlastet, die hatte verfolgte mich für mindestens fünfzehn Jahre vor.

$$|OB| = \sqrt{b_2^2 + b_3^2 + b_4^2},$$

$$|AB| = \sqrt{(b_2 - a_2)^2 + (b_3 - a_3)^2 + (b_4 - a_4)^2}.$$

Nach dem Cosinussatz gilt

$$|b - a|^2 = |a|^2 + |b|^2 - 2|a||b|\cos(\alpha),$$

also

$$\begin{aligned} (b_2 - a_2)^2 + (b_3 - a_3)^2 + (b_4 - a_4)^2 = \\ a_2^2 + a_3^2 + a_4^2 + b_2^2 + b_3^2 + b_4^2 - 2|a||b|\cos(\alpha) \end{aligned}$$

und daraus folgt

$$a_2b_2 + a_3b_3 + a_4b_4 = |a||b|\cos(\alpha).$$

Wie man leicht nachrechnet, hat das Skalarprodukt folgende Eigenschaften:

1. $\langle a + b, c \rangle = \langle a, c \rangle + \langle b, c \rangle,$
2. $\langle ra, b \rangle = r \langle a, b \rangle \quad (r \in \mathbb{R}),$
3. $\langle a, b \rangle = \langle b, a \rangle,$
4. $|a| = \sqrt{\langle a, a \rangle},$
5. $\langle a, b \rangle = 0$ gdw. $a \perp b.$

Das Vektorprodukt

$$a \times b = (a_3b_4 - a_4b_3)i + (a_4b_2 - a_2b_4)j + (a_2b_3 - a_3b_2)k = \frac{1}{2}(ab - ba)$$

kann formal als Determinante geschrieben werden, wenn man nämlich die Determinante

$$\det \begin{pmatrix} i & j & k \\ a_2 & a_3 & a_4 \\ b_2 & b_3 & b_4 \end{pmatrix}$$

nach der ersten Zeile entwickelt, erhält man gerade $a \times b$. Aus den Determinanteneigenschaften erkennen wir sofort

1. $(a + rb) \times c = a \times c + rb \times c \quad (r \in \mathbb{R}),$
2. $a \times b = -b \times a,$
3. $a \times b = 0$ gdw. $\{a, b\}$ ist linear abhängig,

4. Der Vektor $a \times b$ steht senkrecht auf a und auf b .

Beweis: Wegen $ab = -\langle a, b \rangle + a \times b$ folgt $a \times b = ab + \langle a, b \rangle$ und speziell $a^2 = -|a|^2$, also

$$a(a \times b) = a(ab + \langle a, b \rangle) = a^2b + \langle a, b \rangle a = -|a|^2 b + \langle a, b \rangle a,$$

dies ist ein vektorielles Quaternion(!), folglich ist das Skalarprodukt (der skalare Anteil des Produkts) von a und $a \times b$ gleich Null:

$$\langle a, a \times b \rangle = 0.$$

□

5. Der Betrag des Vektors $a \times b$ ist gleich dem Flächeninhalt des Parallelogramms, das durch a und b aufgespannt wird.

Beweis: Es ist

$$\begin{aligned} |a \times b|^2 &= (a_3b_4 - a_4b_3)^2 + (a_4b_2 - a_2b_4)^2 + (a_2b_3 - a_3b_2)^2 \\ &= (a_2^2 + a_3^2 + a_4^2)(b_2^2 + b_3^2 + b_4^2) - (a_2b_2 + a_3b_3 + a_4b_4)^2 \\ &= |a|^2 |b|^2 - \langle a, b \rangle^2 \\ &= |a|^2 |b|^2 - |a|^2 |b|^2 \cos^2(\alpha) \\ &= |a|^2 |b|^2 \sin^2(\alpha). \end{aligned}$$

□

Diese Konstruktionen erlauben interessante geometrische Anwendungen.

Die Menge der Punkte $P = (x, y, z)$, deren Koordinaten eine lineare Gleichung

$$ax + by + cz + d = 0$$

erfüllen, ist eine Ebene E . Sei $P_1 = (x_1, y_1, z_1)$ ein fixierter Punkt von E , also

$$ax_1 + by_1 + cz_1 + d = 0,$$

es folgt

$$a(x - x_1) + b(y - y_1) + c(z - z_1) = 0.$$

Wenn wir den Vektor

$$n = (a, b, c)$$

und den Verbindungsvektor $\overrightarrow{PP_1}$ verwenden, so gilt

$$\langle n, \overrightarrow{PP_1} \rangle = 0 \text{ für alle } P \in E,$$

d.h. der Vektor $n = (a, b, c)$ steht senkrecht auf der durch die Gleichung $ax + by + cz + d = 0$ gegebenen Ebene, man nennt ihn einen Normalenvektor.

Wenn zwei Ebenen E_1 und E_2 einen gemeinsamen Punkt P_0 besitzen, so lauten ihre Gleichungen

$$\langle n_1, \overrightarrow{PP_0} \rangle = 0 \text{ bzw. } \langle n_2, \overrightarrow{PP_0} \rangle = 0,$$

wobei n_1, n_2 jeweils Normalenvektoren der Ebenen sind. Wir suchen die Schnittgerade $E_1 \cap E_2$. Ihre Richtung ist senkrecht zu n_1 und zu n_2 , also lautet ihre Parameterdarstellung

$$P = P_0 + n_1 \times n_2 \cdot t, \quad t \in \mathbb{R}.$$

Der Abstand eines Punkts P_1 von einer Geraden, die durch eine Parameterdarstellung

$$P = P_0 + a \cdot t, \quad t \in \mathbb{R}$$

gegeben ist, ist gleich der Höhe im von den Vektoren a und $b = P_0P_1$ aufgespannten Parallelogramms, also gleich $|b| \sin(\alpha)$ oder gleich

$$|a \times b| / |a|.$$

Für die Multiplikation von Quaternionen gilt das Assoziativgesetz. Nun seien speziell a, b, c vektorielle Quaternionen, dann gilt

$$a(bc) = -a \langle b, c \rangle + a(b \times c) = -a \langle b, c \rangle - \langle a, b \times c \rangle + a \times (b \times c),$$

$$(ab)c = -\langle a, b \rangle c + (a \times b)c = -\langle a, b \rangle c - \langle a \times b, c \rangle + (a \times b) \times c.$$

Die skalaren Anteile müssen übereinstimmen, dies nennt man das Spatprodukt der Vektoren a, b, c ; es ist gleich

$$\det \begin{pmatrix} a_2 & a_3 & a_4 \\ b_2 & b_3 & b_4 \\ c_2 & c_3 & c_4 \end{pmatrix},$$

wie man durch Entwicklung sieht, also gleich dem Volumen des „Spats“, der von den Vektoren a, b, c aufgespannt wird.

Lemma 9.0.4 *Die Vektoren a, b, c liegen genau dann in einer Ebene, wenn $\langle a, b \times c \rangle = 0$ ist.* □

Wenn wir die vektoriellen Teile der Produkte betrachten, erkennen wir, daß das Vektorprodukt nicht assoziativ ist. Vielmehr gilt die sogenannte Jacobi-Identität

$$a \times (b \times c) + b \times (c \times a) + c \times (a \times b) = o.$$

Wir kommen nochmals zu den Quaternionen zurück. Obwohl diese eine doch sehr spezielle Teilmenge von $M_{44}(\mathbb{R})$ bilden, ist ihre Multiplikation „sehr“ nichtkommutativ:

Lemma 9.0.5 *Seien $a, b \in \mathbb{H}$; da gilt $ab = ba$ genau dann, wenn $\{1, a, b\}$ linear abhängig ist.*

Beweis: Sei $\{1, a, b\}$ linear unabhängig und $a = a_1 + a_2, b = b_1 + b_2$ deren Zerlegung in den skalaren und den vektoriellen Anteil. Dann ist auch $\{1, a_2, b_2\}$ linear unabhängig ($a_2 = a - a_1 \cdot 1$). Nun gilt

$$ab = (a_1 + a_2)(b_1 + b_2) = a_1b_1 + a_2b_1 + b_2a_1 + a_2b_2$$

$$ba = (b_1 + b_2)(a_1 + a_2) = b_1a_1 + b_2a_1 + a_2b_1 + b_2a_2$$

Die ersten drei Summanden beider Gleichungen stimmen überein, also ist $a_2b_2 = b_2a_2$, d.h. $a_2 \times b_2 = \frac{1}{2}(a_2b_2 - b_2a_2) = 0$, also ist $a_2 \in L(b_2)$. \square

Kapitel 10

Grundlegende algebraische Strukturen

10.1 Der Ring \mathbb{Z} der ganzen Zahlen

In diesem Abschnitt verstehen wir unter „Zahlen“ stets ganze Zahlen.

Die Division mit Rest ist ein nützliches Hilfsmittel: Seien $a, b \in \mathbb{Z}$, dann gibt es Zahlen q und r , so daß

$$a = bq + r \text{ und } 0 \leq r < |b|.$$

Diese Darstellung ist nicht eindeutig: $100 = 17 \cdot 6 - 2 = 17 \cdot 5 + 15$.

Seien a, b ganze Zahlen, dann nennen wir a einen Teiler von b und schreiben $a \mid b$, wenn eine ganze Zahl c mit $ac = b$ existiert.

Teilbarkeitsregeln für 2,3,4,5,6 sind bekannt. Originell ist die folgende Regel für die 7:

1. Streiche letzte Ziffer.
2. Subtrahiere das Doppelte der gestrichenen Ziffer von der neuen Zahl.
3. Wiederhole 1 und 2

Wenn das Ergebnis durch 7 teilbar ist, so ist es auch die Ausgangszahl.

Beispiele:

$$\begin{array}{r} 3 \quad 9 \quad 8 \quad 2 \quad 3 \\ -6 \\ \hline 3 \quad 9 \quad 7 \quad 6 \\ -1 \quad 2 \\ \hline 3 \quad 8 \quad 5 \\ -1 \quad 0 \\ \hline 2 \quad 8 \end{array}$$

$$\begin{array}{r} 5 \quad 5 \quad 2 \quad 7 \quad 1 \quad 7 \\ -1 \quad 4 \\ \hline 5 \quad 5 \quad 2 \quad 5 \quad 7 \\ -1 \quad 4 \\ \hline 5 \quad 5 \quad 1 \quad 1 \\ -2 \\ \hline 5 \quad 4 \quad 9 \\ -1 \quad 8 \\ \hline 3 \quad 6 \end{array}$$

dies ist durch 7 teilbar

dies nicht

Was ist passiert? Wir haben nicht 6, 5, 10 subtrahiert, sondern 63, 126, 105, das sind durch 7 teilbare Zahlen (bei letzter Ziffer x war das $20x + x = 21x$); es entsteht eine

durch 10 teilbare Zahl. Die Division durch 10 ändert nichts am Teilbarkeitsverhalten durch 7.

Die (positive) Zahl d heißt größter gemeinsamer Teiler der Zahlen a und b , wenn $d \mid a$ und $d \mid b$ gilt (wenn d also ein gemeinsamer Teiler von a und b ist) und wenn für jeden gemeinsamen Teiler t von a und b gilt, daß $t \mid d$ (d ist bezüglich der Teilbarkeitsrelation der GröÙte). Wir schreiben $d = \text{ggT}(a, b)$.

Zur Berechnung des größten gemeinsamen Teilers zweier Zahlen benutzen wir den Euklidischen Algorithmus:

Seien f_1, f_2 gegeben, wir dividieren fortlaufend mit Rest, bis die Division aufgeht:

$$\begin{aligned} f_1 &= q_1 f_2 + f_3 \\ f_2 &= q_2 f_3 + f_4 \\ f_3 &= q_3 f_4 + f_5 \\ &\dots \\ f_{m-3} &= q_{m-3} f_{m-2} + f_{m-1} \\ f_{m-2} &= q_{m-2} f_{m-1} \end{aligned}$$

Wegen $f_2 > f_3 > f_4 > \dots$ muß nach endlich vielen Schritten ein Rest gleich Null sein, hier ist es f_m .

Behauptung: $\text{ggT}(f_1, f_2) = f_{m-1}$.

Beweis:

1. Klar ist, daß f_{m-2} von f_{m-1} geteilt wird. Weiter ist

$$f_{m-3} = (q_{m-3} q_{m-2} + 1) f_{m-1}$$

durch f_{m-1} teilbar. Jetzt haben wir den Anfang in der Hand: Schauen Sie sich die obigen Gleichungen von der letzten bis zur ersten an! Die Zahl f_{m-1} teilt die beiden f 's auf der rechten Seite, damit aber auch das f mit kleinerem Index auf der linken Seite. Am Ende sehen wir, daß f_{m-1} sowohl f_1 als auch f_2 teilt.

2. Sei h ein gemeinsamer Teiler von f_1 und f_2 . Es ist $f_3 = f_1 - q_1 f_2$, also ist h ein Teiler von f_3 . So schrauben wir uns an den Indizes nach oben und erhalten zum Schluß, daß h die Zahl f_{m-1} teilt. \square

Lemma 10.1.1 Sei $d = \text{ggT}(f_1, f_2)$, dann gibt es Zahlen g_1, g_2 mit

$$f_1 g_1 + f_2 g_2 = d.$$

Beweis: Wir lesen die obigen Gleichungen von rechts nach links und von unten nach oben und sehen: Die Zahl f_i läßt sich aus f_{i-1} und f_{i-2} kombinieren. Also läßt sich f_{m-1} aus f_1 und f_2 mit gewissen Faktoren kombinieren. \square

Beispiel: Wir berechnen den $\text{ggT}(1027, 499)$

$$1027 = 2 \cdot 499 + 29$$

$$499 = 17 \cdot 29 + 6$$

$$29 = 4 \cdot 6 + 5$$

$$6 = 1 \cdot 5 + 1$$

also $\text{ggT}(1027, 499) = 1$. Wenn wir die Zahlen rückwärts in die Gleichungen einsetzen, erhalten wir $1 = 177 \cdot 499 - 86 \cdot 1027$.

Interessanter ist das

Lemma 10.1.2 *Der größte gemeinsame Teiler von f_1 und f_2 ist die kleinste positive Zahl d , so daß $f_1g_1 + f_2g_2 = d$ ist.*

Beweis: Sei $d = f_1g_1 + f_2g_2$ und d minimal.

1. Wir dividieren mit Rest:

$$f_1 = q_1d + r_1 = q_1g_1f_1 + q_1g_2f_2 + r_1,$$

hier ist $0 \leq r_1 < d$; wir nehmen an, daß $r_1 \neq 0$ wäre und erhalten

$$r_1 = f_1(1 - q_1g_1) - f_2q_1g_2,$$

aber wegen $r_1 < d$ ist dies ein Widerspruch zur Minimalität von d , also ist $r_1 = 0$, d.h. $d \mid f_1$.

2. Sei h ein gemeinsamer Teiler von f_1 und f_2 , dann ist h auch ein Teiler von $f_1g_1 + f_2g_2 = d$. \square

Seien $a, b, m \in \mathbb{Z}$, wir sagen, daß a und b kongruent modulo m sind, wenn a und b bei der Division durch m denselben Rest lassen, also wenn

$$a - b = km \text{ für ein } k \in \mathbb{Z}.$$

Wir schreiben dann

$$a \equiv b \pmod{m}.$$

Die Menge aller zu einer Zahl a kongruenten Zahlen nennen wir eine Restklasse modulo m , dies ist die Menge $a + m\mathbb{Z}$, manchmal bezeichnen wir diese mit \bar{a} , hier erkennt man aber nicht mehr den „Modul“.

Die Menge aller Restklassen modulo m bezeichnet man mit $\mathbb{Z}/m\mathbb{Z}$. In dieser Menge kann man Rechenoperationen einführen:

$$\bar{a} + \bar{b} = \overline{a + b}, \quad \bar{a} \cdot \bar{b} = \overline{a \cdot b}.$$

Für diese Operationen gelten Assoziativ-, Kommutativ- und Distributivgesetz, es gibt neutrale Elemente 0 und 1 und die Addition ist umkehrbar. Bei der Division ist es schwieriger. Wenn aber a und m zueinander teilerfremd sind, so besitzt \bar{a} ein multiplikatives Inverses modulo m : Es ist $\text{ggT}(a, m) = 1$, also gibt es u, v mit

$$1 = ua + vm,$$

d.h. $1 \equiv ua \pmod{m}$, also ist $\bar{a}^{-1} = \bar{u}$.

Eine Zahl p heißt Primzahl, wenn aus $a \mid p$ folgt, daß $a = \pm 1$ oder $a = \pm p$ gilt.

Ist nun p eine Primzahl, so besitzt jedes von Null verschiedene Element von $\mathbb{Z}/p\mathbb{Z}$ ein multiplikatives Inverses, also ist $\mathbb{Z}/p\mathbb{Z}$ ein Körper.

Zum Schluß wollen wir uns davon überzeugen, daß sich jede positive ganze Zahl als Produkt von Primzahlen darstellen kann.

Lemma 10.1.3 Sei $1 < a \in \mathbb{Z}$, dann gibt es eine Primzahl p mit $p \mid a$.

Beweis: Sei T die Menge aller Teiler von a , die größer als 1 sind. Diese Menge ist nicht leer, besitzt also ein kleinstes Element p . Angenommen, die Zahl p hat einen echten Teiler q , dann gälte $q \in T$ und $q < p$ im Widerspruch zur Auswahl von p . \square

Folgerung 10.1.1 Jede ganze Zahl a ist Produkt von Primzahlen.

Beweis: Die Zahl a besitzt einen Primteiler p_1 , also $a = p_1 a_1$, wenn $a_1 \neq \pm 1$ ist, so gilt $a_1 = a_2 p_2$ und so weiter. Irgendwann wird $a_{n+1} = \pm 1$, also $a = p_1 \dots p_n$. \square

Lemma 10.1.4 Seien $a, b \in \mathbb{Z}$ und p eine Primzahl. Wenn $p \mid ab$ gilt, so gilt $p \mid a$ oder $p \mid b$.

Beweis: Wenn p kein Teiler von a ist, so ist $\text{ggT}(p, a) = 1 = up + va$ für gewisse $u, v \in \mathbb{Z}$. Dann folgt $b = upb + vab$, die rechte Seite wird von p geteilt, also gilt $p \mid b$. \square

Satz 10.1.1 Die Primzahlzerlegung ist (bis auf die Reihenfolge der Faktoren) eindeutig.

Beweis: Es sei $p_1 \dots p_r = q_1 \dots q_s$ für gewisse Primzahlen p_i, q_j . Wir führen die Induktion über die Zahl r . Wenn $r = 1$ ist, so gilt $p_1 = q_1 \dots q_s$, also muß $p_1 = q_1$ und $s = 1$ gelten.

Sei die Behauptung für $r - 1$ Faktoren (links) bewiesen. Die rechte Seite von $p_1 \dots p_r = q_1 \dots q_s$ ist durch p_1 teilbar, also ist ein Faktor, etwa q_1 , durch p_1 teilbar, d.h. $p_1 = q_1$. Dann bleibt $p_2 \dots p_r = q_2 \dots q_s$ und nach Induktionsvoraussetzung ist $r = s$ und $p_i = q_i$ (bei geeigneter Numerierung der Faktoren). \square

Beispiele: Die Fermatschen „Primzahlen“ haben die Form $2^{2^m} + 1$, dies sind $2^1 + 1 = 3$, $2^2 + 1 = 5$, $2^4 + 1 = 17$, $2^8 + 1 = 257$, aber $2^{16} + 1 = 4294967297$ hat den Teiler 641.

Die Mersenneschen Primzahlen haben die Form $M_p = 2^p - 1$, dabei muß p eine Primzahl sein, denn

$$2^{n \cdot m} = (2^n - 1)(2^{m(n-1)} + 2^{m(n-2)} + \dots + 1).$$

Nicht alle sind Primzahlen: $2^2 - 1 = 3$, $2^3 - 1 = 7$, $2^5 - 1 = 31$, $2^{11} - 1 = 2047 = 23 \cdot 89$. Im Jahre 1998 war die größte bekannte Mersenne-Primzahl von der Größenordnung $2^{3Mio} = 10^{900000}$, gedruckt nahm sie 173 A4-Seiten ein; 2006 wurde die 44ste Mersennezahl gefunden: 2^{32Mio} , sie hat 9 Millionen Stellen.

Mein Computer berechnete M_{110503} in 0 Sekunden; zur Bestätigung der Primalität brauchte er 1433 Minuten.

Zur Bestimmung aller Primzahlen ≤ 50 Mio brauchte er 9 Minuten.

Primzahldichte:

Zahlen bis	10	100	1000	10000	100000	1000000
Primzahlen (%)	40	25	16.8	12.3	9.6	7.8

10.2 Lineare diophantische Gleichungen

Wir betrachten Gleichungen der Form

$$ax + by = c, \quad a, b, c \in \mathbb{Z} \text{ gegeben, } x, y \in \mathbb{Z} \text{ gesucht.} \quad (1)$$

Die Gleichung

$$ax + by = 0 \quad (2)$$

nennen wir die zugehörige homogene Gleichung.

Wenn nun $(x_1, y_1), (x_2, y_2)$ Lösungen von (1) sind, also

$$ax_1 + by_1 = c$$

$$ax_2 + by_2 = c$$

so ist $a(x_1 - x_2) + b(y_1 - y_2) = 0$. Wir erhalten also die allgemeine Lösung von (1) als Summe einer speziellen Lösung von (1) und der allgemeinen Lösung von (2).

Die homogene Gleichungen $ax - by = 0$ ist immer lösbar: Wir haben $ax = -by$; sei $t = \text{ggT}(a, b)$, $a' = a/t$, $b' = b/t$, dann folgt $a'x = b'y$ und a', b' sind teilerfremd, also müssen die Teiler von a' in y und die von b' in x aufgehen.

Also ist $x = b', y = -a'$ eine Lösung, aber für jedes ganze Zahl k ist auch $x = kb', y = -ka'$ eine Lösung, und das sind wohl alle.

Zurück zur inhomogenen Gleichung:

Satz 10.2.1 $ax + by = c$ ist genau dann lösbar, wenn $\text{ggT}(a, b) \mid c$.

Beweis: 1. Sei $\text{ggT}(a, b) = t \mid c$, also gibt es $u, v, d \in \mathbb{Z}$ mit $c = dt$, $t = ua + vb$. Dann ist $x = du, y = dv$ eine Lösung.

2. Wenn x, y Lösung ist, also $ax + by = c$ und $\text{ggT}(a, b) = t$ ist, so gilt $t \mid a, t \mid b$, also auch $t \mid c$ □

Beispiel: $7x + 3y = 20$

$\text{ggT}(7, 3) = 1 = -2 \cdot 7 + 5 \cdot 3$, also $7 \cdot (-40) + 3 \cdot 100 = 20$.

10.3 Gruppen, Untergruppen, Homomorphismen

Definition: Sei G eine Menge und $\cdot: G \times G \rightarrow G$ eine Abbildung, die dem Paar (g_1, g_2) das Element $\cdot(g_1, g_2) = g_1 \cdot g_2$ zuordnet. Wir nennen diese Abbildung eine Multiplikation. Wenn folgende Eigenschaften erfüllt sind, so heißt G eine Gruppe:

- 1) $g_1 \cdot (g_2 \cdot g_3) = (g_1 \cdot g_2) \cdot g_3$ für alle $g_1, g_2, g_3 \in G$ (Assoziativgesetz),
- 2) es gibt ein $e \in G$, so daß $g \cdot e = e \cdot g = g$ für alle $g \in G$ gilt,
- 3) zu jedem $g \in G$ gibt es ein Element $g^{-1} \in G$ mit $g \cdot g^{-1} = g^{-1} \cdot g = e$.

Das ausgezeichnete Element e heißt neutrales Element und das Element g^{-1} heißt das zu g inverse Element. Das Multiplikationszeichen werden wir künftig weglassen. Wenn

besonders hervorgehoben werden soll, um welche Operation es sich in der Menge G handelt, so bezeichnen wir die Gruppe mit (G, \cdot) .

Falls die Gruppe G eine endliche Menge ist, so bezeichnen wir mit $|G|$ die Zahl ihrer Elemente, diese Zahl heißt die Ordnung von G .

Falls für alle $g_1, g_2 \in G$ die Gleichung $g_1g_2 = g_2g_1$ gilt, so heißt die Gruppe G kommutativ.

Für kommutative (und nur solche) Gruppen ist auch eine additive Schreibweise üblich:

$$+ : G \times G \rightarrow G, (g_1, g_2) \mapsto g_1 + g_2,$$

das neutrale Element wird Nullelement genannt und mit 0 bezeichnet, also

$$g + 0 = 0 + g = g \text{ für alle } g \in G,$$

das zu g inverse Element wird mit $-g$ bezeichnet, also

$$g + (-g) = 0.$$

Anstelle von $g_1 + (-g_2)$ schreibt man dann einfach $g_1 - g_2$.

Sie kennen folgende Beispiele von Gruppen:

$$(\mathbb{Z}, +), (\mathbb{Q}, +), (\mathbb{R}, +), (\mathbb{C}, +), (\mathbb{Q} \setminus \{0\}, \cdot), (\mathbb{R} \setminus \{0\}, \cdot), (\mathbb{C} \setminus \{0\}, \cdot),$$

$$(\mathbb{R}^n, +), (\text{Hom}(V, W), +), (M_{mn}, +),$$

all diese Gruppen sind kommutativ. Die Menge GL_n aller invertierbarer Matrizen ist eine nichtkommutative Gruppe, ebenso die Menge S_n aller Permutationen von n Ziffern.

Die Gruppenaxiome können wir folgt abgeschwächt werden:

Sei G eine Menge mit einer assoziativen Multiplikation; wir fordern

(A) Es gibt ein $e \in G$ mit $ge = g$ für alle $g \in G$ (d.h. es gibt ein universelles rechtes Einselement, es kann mehrere derartige Elemente geben).

(B) Zu jedem derartigen e und jedem $g \in G$ existiert ein $g' \in G$ mit $gg' = e$ (also ein privates rechtinverses Element).

Wir zeigen die Gültigkeit der obigen Gruppenaxiome:

Sei $gg' = e$, dann gilt $g' = g'e = g'gg'$; wir multiplizieren von rechts mit dem zu g' Inversen:

$$e = g'(g')' = g'gg'(g')' = g'ge = g'g,$$

also ist jedes Rechtsinverse auch linksinvers, d.h. g' ist zu g invers. Weiter ist

$$eg = (gg')g = g(g'g) = g,$$

also ist e auch ein linksseitiges neutrales Element und wenn f ein weiteres neutrales Element ist, so gilt $e = ef = f$.

Die Bedingung (A) wird zum Beispiel bei folgenden Multiplikationen erfüllt:

	a	b	c
a	a	a	a
b	a	b	b
c	a	b	c

	a	b	c
a	a	a	a
b	b	b	b
c	c	c	c

Im ersten Fall spielt c die Rolle von e , im zweiten Fall gibt es mehrere Rechtseinsen; die Forderung (B), daß e in jeder Zeile auftreten soll, ist nicht erfüllt.

Anmerkung: Es gibt $3^9 = 19689$ Möglichkeiten, eine 3×3 -Multiplikationstafel auszufüllen, den Assoziativtest überstehen 113; wenn noch „isomorphe“ und „antiisomorphe“ ausgesondert werden, verbleiben 18 „Halbgruppen“, unter denen befindet sich eine Gruppe. Der Computer schafft dies alles in weniger als einer Sekunde. Wenn dasselbe mit 4 Elementen probiert wird, dauert das 46 Minuten, bei 5 Elementen würde es 1000 Jahre dauern.

Definition: Eine nichtleere Teilmenge $U \subseteq G$ einer Gruppe G heißt eine Untergruppe von G , wenn für alle $u, v \in U$ auch $uv \in U$ und $u^{-1} \in U$ gilt.

Wir sehen sofort, daß jede Untergruppe $U \subseteq G$ das neutrale Element e von G enthalten muß: Da $U \neq \emptyset$ gilt, gibt es ein $u \in U$. Dann muß auch $u^{-1} \in U$ sein und folglich ist auch $e = uu^{-1} \in U$.

Lemma 10.3.1 *Wenn U und V Untergruppen von G sind, so ist auch $U \cap V$ eine Untergruppe von G .* □

Wir werfen einen Blick auf die obigen Beispiele: Unter den additiven Gruppen $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ ist jeweils die kleinere eine Untergruppe der größeren, ebenso gilt dies für die multiplikativen Gruppen $\mathbb{Q} \setminus \{0\}, \mathbb{R} \setminus \{0\}, \mathbb{C} \setminus \{0\}$.

Wir betrachten als Beispiel die einfachste nichtkommutative Gruppe

$$S_3 = \left\{ e = \begin{pmatrix} 123 \\ 123 \end{pmatrix}, a = \begin{pmatrix} 123 \\ 132 \end{pmatrix}, b = \begin{pmatrix} 123 \\ 321 \end{pmatrix}, c = \begin{pmatrix} 123 \\ 213 \end{pmatrix}, d = \begin{pmatrix} 123 \\ 231 \end{pmatrix}, f = \begin{pmatrix} 123 \\ 312 \end{pmatrix} \right\}.$$

Die Multiplikation in S_3 , die Nacheinanderausführung der Permutationen, kann man in einer Multiplikationstafel beschreiben:

	e	a	b	c	d	f
e	e	a	b	c	d	f
a	a	e	d	f	b	c
b	b	f	e	d	c	a
c	c	d	f	e	a	b
d	d	c	a	b	f	e
f	f	b	c	a	e	d

e ist das neutrale Element, $a^{-1} = a$, $b^{-1} = b$, $c^{-1} = c$, $d^{-1} = f$.

Die Gruppe S_3 hat folgende Untergruppen:

$$\{e\}, \{e, a\}, \{e, b\}, \{e, c\}, \{e, d, f\}, S_3.$$

Wir suchen die Untergruppen der additiven Gruppe \mathbb{Z} :

Zunächst ist $m \cdot \mathbb{Z} = \{mk \mid k \in \mathbb{Z}\}$ eine Untergruppe.

Behauptung: Jede Untergruppe von \mathbb{Z} hat diese Form.

Beweis: Wenn $U = \{0\}$ ist, so wählen wir $m = 0$. Sonst sei m das kleinste positive Element in U , dann ist $m\mathbb{Z} \subseteq U$. Sei nun $u \in U$ beliebig; wir dividieren mit Rest: $u = qm + r$, $0 \leq r < m$. Wegen $u, m \in U$ folgt $r \in U$, also $r = 0$ und $u \in m\mathbb{Z}$.

Wenn $E \subseteq G$ eine Teilmenge ist, so bezeichnen wir mit $\langle E \rangle$ die kleinste Untergruppe von G , die E enthält, sie besteht aus allen Produkten von Elementen aus E und von Inversen von Elementen aus E . Wir sagen, die Untergruppe $\langle E \rangle$ wird von der Menge E erzeugt.

Zum Beispiel:

$$\begin{aligned} (\mathbb{Z}, +) &= \langle 1 \rangle, (\mathbb{Q} \setminus \{0\}, \cdot) = \langle \mathbb{Z} \setminus \{0\} \rangle, \\ \{e, a\} &= \langle a \rangle, \{e, b\} = \langle b \rangle, \{e, c\} = \langle c \rangle, \\ \{e, d, f\} &= \langle d \rangle = \langle f \rangle, S_3 = \langle a, b \rangle = \langle a, d \rangle \text{ usw.} \end{aligned}$$

Eine Gruppe G , die von einem Element g erzeugt wird, heißt zyklische Gruppe, es gilt also $G = \{e = g^0, g = g^1, g^2, \dots\}$, die Gruppe kann endlich oder unendlich sein.

Wir überlegen, wie eine endliche zyklische Gruppe $G = \langle g \rangle$ aussehen könnte. Die Potenzen g, g^2, g^3, \dots von g können nicht alle verschieden sein, denn es gibt unendlich viele. Also gilt für gewisse Exponenten m und k , daß $g^m = g^{m+k}$ ist. Wir multiplizieren mit $(g^m)^{-1}$ und erhalten $e = g^0 = g^k$, also besteht G genau aus den k verschiedenen Elementen $e = g^0, g, g^2, \dots, g^{k-1}$. Die Gruppe werden wir mit C_k bezeichnen.

Die additive Gruppe \mathbb{Z} ist eine unendliche zyklische Gruppe, die Menge der Drehungen um Vielfache von 120° ist eine endliche zyklische Gruppe, sie hat die Ordnung 3.

Wenn $M, N \subseteq G$ Teilmengen einer Gruppe sind, so bezeichnen wir mit $M \cdot N$ die Menge $\{mn \mid m \in M, n \in N\}$ und mit M^{-1} die Menge $\{m^{-1} \mid m \in M\}$. Dann ist $U \subseteq G$ also eine Untergruppe, wenn $UU \subseteq U$ und $U^{-1} \subseteq U$ gilt. Überlegen Sie sich, daß in beiden Fällen sogar Gleichheit gilt.

Sei $U \subseteq G$ eine Untergruppe. Wir führen in der Menge G eine Relation \sim ein: für $g, h \in G$ gelte $g \sim h$ genau dann, wenn $gh^{-1} \in U$ ist. Wir sagen: g und h sind äquivalent *modulo* U . Äquivalent dazu ist $Uh = Ug$.

Lemma 10.3.2 Die Relation \sim ist eine Äquivalenzrelation auf G , die Menge aller zu $g \in G$ äquivalenten Elemente ist $Ug = \{ug \mid u \in U\}$.

Beweis: Für alle $g \in G$ gilt $g \sim g$, da $gg^{-1} = e \in U$ ist. Sei $g \sim h$, also $gh^{-1} \in U$, dann ist auch $(gh^{-1})^{-1} = hg^{-1} \in U$, also gilt $h \sim g$.

Sei schließlich $g \sim h$ und $h \sim k$, also $gh^{-1} \in U$ und $hk^{-1} \in U$, dann ist auch $(gh^{-1})(hk^{-1}) = gk^{-1} \in U$, also $g \sim k$.

Schließlich ist $g \sim ug$ für alle $u \in U$, denn $g(ug)^{-1} = gg^{-1}u^{-1} = u^{-1} \in U$. \square

Wenn G eine additiv geschriebene Gruppe und U eine Untergruppe ist, so gilt $g \sim h$, wenn $g - h \in U$ ist, und die Äquivalenzklasse von g wird mit $g + U$ bezeichnet.

Lemma 10.3.3 Für alle $g \in G$ gilt $|Ug| = |U|$, d.h. alle Äquivalenzklassen sind gleichmächtig.

Beweis: Sei $g \in G$, wir betrachten die Abbildung $f : U \rightarrow Ug$ mit $f(u) = ug$. Diese Abbildung ist surjektiv (klar), wir zeigen, daß sie injektiv ist: Sei $u_1g = u_2g$, dann gilt $u_1gg^{-1} = u_2gg^{-1} = u_1 = u_2$. Also ist f bijektiv und damit gilt $|Ug| = |U|$. \square

Beispiel:

Die Menge aller durch 5 teilbaren ganzen Zahlen (wir bezeichnen sie mit $5\mathbb{Z}$) ist eine Untergruppe der additiven Gruppe \mathbb{Z} . Die Menge \mathbb{Z} ist die Vereinigung aller Äquivalenzklassen modulo $5\mathbb{Z}$:

$$\begin{aligned} 5\mathbb{Z} &= \{0, \pm 5, \pm 10, \pm 15, \dots\}, \\ 1 + 5\mathbb{Z} &= \{1, 6, 11, \dots, -4, -9, \dots\}, \\ 2 + 5\mathbb{Z} &= \{2, 7, 12, \dots, -3, -8, \dots\}, \\ 3 + 5\mathbb{Z} &= \{3, 8, 13, \dots, -2, -7, \dots\}, \\ 4 + 5\mathbb{Z} &= \{4, 9, 14, \dots, -1, -6, \dots\}. \end{aligned}$$

Wenn $U \subseteq G$ eine Untergruppe ist, so bezeichnet man die Menge aller Äquivalenzklassen modulo U mit G/U .

Satz 10.3.1 (Lagrange) Sei G eine endliche Gruppe und $U \subseteq G$ eine Untergruppe, dann ist die Zahl $|U|$ ein Teiler von $|G|$.

Beweis: Es ist $G = U \cup Ug_2 \cup Ug_3 \cup \dots \cup Ug_t$ für gewisse $g_i \in G$, denn G ist die disjunkte Vereinigung seiner Äquivalenzklassen modulo U , also gilt $|G| = t|U|$. \square

Folgerung 10.3.1 Jede Gruppe von Primzahlordnung ist zyklisch.

Beweis: Sei $|G| = p$ eine Primzahl und $e \neq g \in G$, dann ist $\langle g \rangle$ eine Untergruppe mit mehr als einem Element, da die Ordnung von $\langle g \rangle$ ein Teiler von p ist, folgt $|\langle g \rangle| = p$, also $G = \langle g \rangle$. \square

Definition: Sei G eine Gruppe und $g \in G$, dann heißt die kleinste Zahl $n > 0$ mit $g^n = e$ die Ordnung von g .

Die Ordnung von $g \in G$ ist also gleich der Ordnung der von g erzeugten zyklischen Untergruppe $\langle g \rangle$, also ein Teiler von $|G|$. Also gilt das

Lemma 10.3.4 Sei $|G| = n$ und $g \in G$, dann ist $g^n = e$. \square

Folgerung 10.3.2 (Kleiner Satz von Fermat) Sei p eine Primzahl. Wenn $ggT(a, p) = 1$ ist, so gilt $a^{p-1} \equiv 1 \pmod{p}$.

Wenn zwei (nicht notwendigerweise verschiedene) Gruppen G und H gegeben sind, so kann man in der Menge $G \times H$ eine Multiplikation einführen, so daß $G \times H$ wieder eine Gruppe wird:

$$(g_1, h_1)(g_2, h_2) = (g_1g_2, h_1h_2).$$

Wenn e das neutrale Element von G und f das neutrale Element von H ist, so ist (e, f) das neutrale Element von $G \times H$ und $(g, h)^{-1} = (g^{-1}, h^{-1})$. Das Assoziativgesetz ist leicht nachzuprüfen.

Von nun ab wollen wir das neutrale Element einer multiplikativ geschriebenen Gruppe mit „1“ bezeichnen, wenn es nicht zu Verwechslungen führt.

Wir wollen uns einen Überblick über die Gruppen mit „wenigen“ Elementen verschaffen. Wir stellen uns die Multiplikationstafel vor, dort müssen in jeder Zeile und in jeder Spalte alle Gruppenelemente auftreten.

1. $\{1\} = C_1$

2. $\{1, g\}$

Es kann nicht $g^2 = g$ gelten, also ist $g^2 = 1$, dies ist also C_2 .

3. $\{1, g, h\}$

Wenn $g^2 = 1$ wäre, müßte $gh = h$ sein, das geht aber nicht. Also ist $g^2 = h$. Dann muß aber auch $gh = 1$ sein, also $g^3 = 1$, die Gruppe ist also C_3 .

4. Eine Möglichkeit wäre C_4 .

Eine nichtzyklische Gruppe mit vier Elementen müßte wie folgt aussehen: $\{1, g, h, k\}$. Wenn $g^2 = h$ wäre, müßte $g^3 = 1$ oder $g^3 = k$ sein, das erste geht nicht, weil dann $\{1, g, g^2\}$ eine Untergruppe mit drei Elementen wäre (3 ist kein Teiler von 4), das zweite geht nicht, weil dann $g^4 = 1$ wäre, die Gruppe wäre also zyklisch. Folglich ist $g^2 = 1$, analog $h^2 = k^2 = 1$ und schließlich $gh = k$. Diese Gruppe ist „isomorph“ zu $C_2 \times C_2$. Diese Gruppe heißt „Kleinsche Vierergruppe“.

5. Die Gruppenordnung ist eine Primzahl, die einzige Möglichkeit ist C_5 .

6. Wie immer haben wir eine zyklische Gruppe C_6 , eine andere Gruppe mit sechs Elementen ist S_3 , dies sind „bis auf Isomorphie“ alle. Frage: Was ist mit $C_2 \times C_3$?

Definition: Seien (H, \cdot) und $(G, *)$ Gruppen. Eine Abbildung $f : H \rightarrow G$ heißt Gruppenhomomorphismus, wenn $f(h_1 \cdot h_2) = f(h_1) * f(h_2)$ für alle $h_1, h_2 \in H$ gilt.

Sei $f : H \rightarrow G$ ein Homomorphismus, dann gilt $f(1) = 1$ und $f(h^{-1}) = f(h)^{-1}$, denn $f(1) = f(1 \cdot 1) = f(1) * f(1)$ und $1 = f(1) = f(hh^{-1}) = f(h)f(h^{-1})$.

Beispiele: Die Inklusionsabbildungen $\mathbb{Z} \rightarrow \mathbb{Q} \rightarrow \mathbb{R} \rightarrow \mathbb{C}$ sind Homomorphismen der additiven Gruppen, für die Logarithmusfunktion gilt $\ln(ab) = \ln(a) + \ln(b)$, also ist $\ln : (\mathbb{R}_+, \cdot) \rightarrow (\mathbb{R}, +)$ ein Homomorphismus. Die Funktion $\text{sgn} : S_n \rightarrow \{\pm 1\}$, die jeder Permutation ihr Signum zuordnet, ist ein Homomorphismus. Für jeden Homomorphismus $f : G \rightarrow H$ und jede Untergruppe $U \subseteq G$ ist die Einschränkung $f|_U : U \rightarrow H$ ebenfalls ein Homomorphismus. Schließlich ist für jedes $x \in \mathbb{Z}$ die Abbildung $l_x : \mathbb{Z} \rightarrow \mathbb{Z}$ mit $l_x(a) = xa$ ein Homomorphismus der additiven Gruppen. Für $G = \langle a \rangle$, $a^3 = e$ ist durch $f(a) = a^2$ ein nichttrivialer Homomorphismus von G in sich gegeben; für $G = \langle a \rangle$, $a^4 = e$ haben wir $f_1 = \text{id}$, $f_2(a) = a^2$, $f_3(a) = a^3$; bei der Kleinschen Vierergruppe ist jede Permutation der von e verschiedenen Elemente ein Homomorphismus.

Wenn $U \subset G$ eine Untergruppe ist, so ist $f(U) \subset H$ eine Untergruppe und wenn $V \subset H$ eine Untergruppe ist, so ist $f^{-1}(V) \subset G$ eine Untergruppe.

Definition: Sei $f : H \rightarrow G$ ein Homomorphismus, dann ist der Kern von f die Teilmenge $\text{Ker}(f) = \{h \in H \mid f(h) = 1\}$.

Lemma 10.3.5 Wenn $f : H \rightarrow G$ ein Homomorphismus ist, so ist $\text{Ker}(f)$ eine Untergruppe von G .

Beweis: Seien $h_1, h_2 \in \text{Ker}(f)$, also $f(h_1) = 1 = f(h_2)$, dann ist $f(h_1 h_2) = f(h_1) f(h_2) = 1 \cdot 1 = 1$ und $f(h_1^{-1}) = f(h_1)^{-1} = 1$. \square

Wir bemerken, daß der Kern eines Homomorphismus eine weitere Eigenschaft hat: Wenn $h \in \text{Ker}(f)$ ist, so gilt für beliebige $g \in G$ folgendes: $f(g^{-1} h g) = f(g^{-1}) f(h) f(g) = f(g)^{-1} f(g) = 1$, also $g^{-1} h g \in \text{Ker}(f)$.

Definition: Sei $N \subseteq G$ eine Untergruppe, sie heißt normale Untergruppe (oder Normalteiler), wenn $g^{-1} N g = N$ für alle $g \in G$ gilt.

In einer kommutativen Gruppe ist jede Untergruppe normal, der Kern eines Homomorphismus ist eine normale Untergruppe.

Wir erinnern daran, daß $G/N = \{Ng\}$ die Menge aller Äquivalenzklassen modulo der Untergruppe N bezeichnete.

Satz 10.3.2 Sei $N \subseteq G$ eine normale Untergruppe, dann ist die Menge G/N mit folgender Multiplikation eine Gruppe: $(Ng)(Nh) = Ngh$. Die Ordnung von G/N ist $\frac{|G|}{|N|}$.

Beweis: Wegen $g^{-1} N g = N$ gilt $Ng = gN$, also gilt für das Produkt der Teilmengen Ng und Nh wirklich $NgNh = NNgh = Ngh$. Der Rest ist klar: $(Ng_1 Ng_2) Ng_3 = N(g_1 g_2) g_3 = Ng_1 (g_2 g_3) = Ng_1 (Ng_2 N_3)$, das neutrale Element ist N , da $NNg = Ng = NgN$ gilt, invers zu Ng ist Ng^{-1} . \square

Den im folgenden Lemma auftretenden Homomorphismus nennt man einen „kanonischen“ Homomorphismus.

Lemma 10.3.6 Sei $N \subseteq G$ eine normale Untergruppe, dann ist die Abbildung $k : G \rightarrow G/N$ mit $k(g) = Ng$ ein Homomorphismus und es gilt $\text{Ker}(k) = N$.

Beweis: $k(g_1 g_2) = Ng_1 g_2 = Ng_1 N g_2 = k(g_1) k(g_2)$ und $k(g) = N$ gilt genau dann, wenn $g \in N$ ist. \square

Definition: Sei $f : H \rightarrow G$ ein Homomorphismus, dann ist das Bild von f die folgende Menge $\text{Im}(f) = \{g \in G \mid \text{es gibt ein } h \in H \text{ mit } g = f(h)\}$.

Lemma 10.3.7 $\text{Im}(f)$ ist eine Untergruppe von G . \square

Satz 10.3.3 Sei $f : H \rightarrow G$ ein Homomorphismus. Dann gilt:

f ist genau dann injektiv, wenn $\text{Ker}(f) = \{1\}$ ist,

f ist genau dann surjektiv, wenn $\text{Im}(f) = G$ ist.

Beweis: Sei f injektiv und $g \in \text{Ker}(f)$, also $f(g) = 1 = f(1)$, dann muß $g = 1$ sein. Sei umgekehrt $\text{Ker}(f) = \{1\}$ und $f(h) = f(g)$, dann gilt $1 = f(g) f(h)^{-1} = f(gh^{-1})$, also $gh^{-1} \in \text{Ker}(f) = \{1\}$, d.h. $gh^{-1} = 1$, also $g = h$.

Die zweite Aussage ist trivial. \square

Ein injektiver und surjektiver Homomorphismus heißt Isomorphismus. Wenn zwischen zwei Gruppen H und G ein Isomorphismus existiert $f : H \rightarrow G$ existiert, so heißen sie isomorph, man schreibt dann $H \simeq G$.

Es folgen einige Sätze, die die Isomorphie gewisser Gruppen sichern.

Satz 10.3.4 (Homomorphiesatz) Sei $f : H \rightarrow G$ ein Homomorphismus, dann ist die durch $F(h \cdot \text{Ker}(f)) = f(h)$ gegebene Abbildung $F : H/\text{Ker}(f) \rightarrow \text{Im}(f)$ ein Isomorphismus.

Beweis: Wir zeigen zuerst, daß F wohldefiniert ist: Sei $h_1 \text{Ker}(f) = h_2 \text{Ker}(f)$, also $h_1 h_2^{-1} \in \text{Ker}(f)$, d.h. $1 = f(h_1 h_2^{-1}) = f(h_1) f(h_2)^{-1}$, also $F(h_1 \text{Ker}(f)) = f(h_1) = f(h_2) = F(h_2 \text{Ker}(f))$. Weiter gilt

$$\begin{aligned} F(h_1 \text{Ker}(f) \cdot h_2 \text{Ker}(f)) &= F(h_1 h_2 \text{Ker}(f)) = f(h_1 h_2) = f(h_1) f(h_2) \\ &= F(h_1 \text{Ker}(f)) \cdot F(h_2 \text{Ker}(f)), \end{aligned}$$

also ist F ein Homomorphismus. Die Surjektivität von F ist klar und die Injektivität folgt sofort: Sei $F(h \text{Ker}(f)) = 1 = f(h)$, dann ist $h \in \text{Ker}(f)$, also ist $h \text{Ker}(f) = \text{Ker}(f)$ das neutrale Element in $H/\text{Ker}(f)$. \square

Lemma 10.3.8 H sei eine Untergruppe von G , N sei eine normale Untergruppe von G , dann gilt:

1. $H \cap N$ ist eine normale Untergruppe von H ,
2. wenn $N \subseteq H$ ist, so ist N eine normale Untergruppe von H ,
3. HN ist eine Untergruppe von G und N ist eine normale Untergruppe von HN ,
4. wenn $N \subseteq H$ und $H \subseteq G$ ebenfalls eine normale Untergruppe ist, so ist H/N eine normale Untergruppe von G/N .

Beweis: 1. Sei $f : G \rightarrow G/N$ der kanonische Homomorphismus, dann ist die Einschränkung $f|_H : H \rightarrow G/N$ ebenfalls ein Homomorphismus, dessen Kern gerade $H \cap N$ ist.

2. Trivial.

3. Sei $h_i \in H, n_i \in N$, dann sind $h_1 n_1, h_2 n_2 \in HN$, weiter ist $h_2^{-1} n_1 h_2 = n \in N$ wegen der Normalteilereigenschaft, also $n_1 h_2 = h_2 n$. Nun folgt $h_1 n_1 \cdot h_2 n_2 = h_1 h_2 n n_2 \in HN$ und $(h_1 n_1)^{-1} = n_1^{-1} h_1^{-1} = h_1^{-1} n'$ mit $n' = h_1 n_1^{-1} h_1^{-1} \in N$.

4. Es gilt $H/N = \{Nh \mid h \in H\} \subseteq \{Ng \mid g \in G\} = G/N$, weiter $Nh_1 \cdot Nh_2 = Nh_1 h_2 \in H/N$ und $(Nh)^{-1} = Nh^{-1} \in H/N$, also ist H/N eine Untergruppe von G/N . Diese ist normal: $(Ng)^{-1} Nh Ng = Ng^{-1} h g$ und $g^{-1} h g \in H$, also liegt $(Ng)^{-1} Nh Ng$ in H/N . \square

Satz 10.3.5 (1. Isomorphiesatz) Seien $H, N \subseteq G$ Untergruppen, N sei normal, dann gilt

$$H/(N \cap H) \simeq HN/N.$$

Beweis: Sei $f : H \rightarrow HN/N$ die durch $f(h) = hN$ gegebene Abbildung, dies ist ein Homomorphismus. Die Abbildung f ist surjektiv, denn sei $hnN \in HN/N$, wegen $nN = N$ ist dies gleich $hN = f(h) \in \text{Im}(f)$. Sei $h \in \text{Ker}(f)$, also $f(h) = hN = N$, d.h. $h \in N$, also $h \in N \cap H$. Die Behauptung folgt nun aus dem Homomorphiesatz. \square

Satz 10.3.6 (2. Isomorphiesatz) Seien $N \subseteq H \subseteq G$ normale Untergruppen, dann gilt

$$G/H \simeq (G/N)/(H/N).$$

Beweis: Wir betrachten die Abbildung $f : G/N \rightarrow G/H$, die durch $f(gN) = gH$ gegeben ist (sie ist wegen $gN \subseteq gH$ wohldefiniert), offenbar surjektiv und ein Homomorphismus. Es ist genau dann $gN \in \text{Ker}(f)$, wenn $gH = H$, also $g \in H$ gilt. Der Kern ist somit gleich H/N und die Behauptung folgt aus dem Homomorphiesatz. \square

Beispiele:

1. $G = S_3$, $H = \{e, a\}$, $N = \{e, d, f\}$, dann ist $HN = S_3$ und $H \cap N = \{e\}$, also $S_3/N \cong H$.
2. $G = \mathbb{Z}$, $H = m\mathbb{Z}$, $N = km\mathbb{Z}$. Dann ist $G/H = \mathbb{Z}/m\mathbb{Z}$ eine zyklische Gruppe der Ordnung m und $(\mathbb{Z}/km\mathbb{Z})/(m\mathbb{Z}/km\mathbb{Z}) \cong C_k$.

Satz 10.3.7 Sei $U \subset V$ ein Unterraum des Vektorraums V , dann ist $\dim V/U = \dim V - \dim U$.

Beweis: Sei $\{v_1, \dots, v_k\}$ eine Basis von U , wir ergänzen sie zu einer Basis $\{v_1, \dots, v_n\}$ von V . Die Elemente von V/U haben die Gestalt $v + U$, $v \in V$. Sei $v = \sum r_i v_i$, dann ist

$$v + U = \underbrace{(r_1 v_1 + U)}_{=U} + \dots + \underbrace{(r_k v_k + U)}_{=U} + (r_{k+1} v_{k+1} + U) + \dots + (r_n v_n + U),$$

also erzeugen $v_{k+1} + U, \dots, v_n + U$ den Vektorraum V/U . Sind sie linear unabhängig? Wir betrachten

$$(r_{k+1} v_{k+1} + U) + \dots + (r_n v_n + U) = (r_{k+1} v_{k+1} + \dots + r_n v_n) + U,$$

dies ist genau dann gleich U , wenn $r_{k+1} v_{k+1} + \dots + r_n v_n \in U$ gilt, also

$$r_{k+1} v_{k+1} + \dots + r_n v_n = r_1 v_1 + \dots + r_k v_k,$$

und hieraus folgt $(r_1 = \dots = r_k =) r_{k+1} = \dots = r_n = 0$. \square

Aus dem ersten Isomorphiesatz folgt nun $\dim(U_1 + U_2) - \dim U_1 = \dim U_2 - \dim U_1 \cap U_2$, das wußten wir schon.

Bemerkungen zur Ordnung

Wir bezeichnen das neutrale Element einer (multiplikativen) Gruppe von nun ab mit 1.

Wenn $a \in G$ ist, so nennen wir $\text{ord}(a) = |\langle a \rangle|$, die kleinste positive Zahl k mit $a^k = 1$, die Ordnung von a .

1. Wenn alle $a \in G$ die Ordnung 2 haben, so ist G kommutativ.

Beweis: Aus $a^2 = 1$ folgt $a^{-1} = a$, also ist $(ab)(ab) = 1$, also $ab = (ab)^{-1} = b^{-1} a^{-1} = ba$.

2.

Lemma 10.3.9 Wenn G kommutativ und $N_1, N_2 \subset G$ Untergruppen mit $N_1 N_2 = G$, $N_1 \cap N_2 = \{1\}$ sind, dann ist $f : N_1 \times N_2 \rightarrow G$, $f(a_1, a_2) = a_1 a_2$ ein Isomorphismus.

Beweis: f ist offensichtlich surjektiv. Wenn $F(a_1, a_2) = a_1 a_2 = 1$ ist, so ist $a_2 = a_1^{-1} \in N_1 \cap N_2 = \{1\}$, also $a_1 = a_2 = 1$, d.h. f ist injektiv. \square

3. Wenn $a, b \in G$ und $ab = ba$, dann gilt $ord(ab) = ord(a) \cdot ord(b)$ genau dann, wenn $ggT(ord(a), ord(b)) = 1$ ist.

Beweis: Wir betrachten $\langle a \rangle, ord(a) = m, \langle b \rangle, ord(b) = n, ord(ab) = r$. Sei d ein gemeinsamer Teiler von m, n , etwa $m = dm', n = dn'$. Dann ist $(ab)^{dm'n'} = a^{mn'} \cdot b^{nm'} = 1$, also $r < mn$. Sei nun $ggT(m, n) = 1$. Es ist $\langle a \rangle \cap \langle b \rangle$ eine Untergruppe sowohl von $\langle a \rangle$ als auch von $\langle b \rangle$, also teilt die Ordnung von $\langle a \rangle \cap \langle b \rangle$ sowohl m als auch n , also ist $\langle a \rangle \cap \langle b \rangle = \{1\}$. Also ist $\langle a \rangle \cdot \langle b \rangle = \langle a \rangle \times \langle b \rangle$ und da die Ordnungen teilerfremd sind, ist $\langle a \rangle \times \langle b \rangle$ zyklisch.

4. Seien $a_1, \dots, a_n \in G, a_i a_j = a_j a_i$ für alle i, j und $ord(a_i) = k_i$, dann ist $ord(a_1 \cdots a_n)$ ein Teiler von $kgV(k_1, \dots, k_n)$.

Beweis: Es ist $(a_1 \cdots a_n)^k = a_1^k \cdots a_n^k$; wenn k Vielfaches aller k_i ist, gilt $a_i^k = 1$, also gilt für $k = kgV(k_1, \dots, k_n)$ auch $(a_1 \cdots a_n)^k = 1$.

5. Für nichtkommutative Gruppen ist es nicht so einfach. Wir betrachten die multiplikative Gruppe $SL(2, \mathbb{Z}) = \{A \in M_{22} \mid det(a) = 1\}$. Wir geben einige Element A_i an, deren Ordnung i ist:

$$A_2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, A_3 = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}, A_3^2 = \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix}, A_4 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

$$A_6 = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}, A_\infty = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, A_4 \cdot A_3 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = A_\infty.$$

6. Die einzigen möglichen endlichen Ordnungen von Elementen aus $SL(2, \mathbb{Z})$ sind 1, 2, 3, 4, 6.

Beweis: Sei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ und $a^n = E_2$. Sei z ein Eigenwert von A , dann ist $z^n = 1$, also ist z eine Einheitswurzel. Es ist $c_A(z) = z^2 - (a+d)z + |A|, z + \bar{z} \in \mathbb{Z}, |z| = |\bar{z}| = 1$, also gilt $z + \bar{z} \in \{-2, -1, 0, 1, 2\}$. Die Nullstellen von $z^2 + z + 1$ sind $-\frac{1}{2} \pm \frac{1}{2}\sqrt{3}i$, dies sind dritte Einheitswurzeln.

10.4 Die symmetrischen Gruppen

Wir wollen nun die Gruppen $S_n = \{p : \{1, \dots, n\} \rightarrow \{1, \dots, n\}\}$ der bijektiven Abbildungen der Menge $\{1, \dots, n\}$ in sich betrachten.

Zunächst wollen wir mit gruppentheoretischen Mitteln deren Ordnung bestimmen.

Es sei eine Ziffer $m, 1 \leq m \leq n$, fixiert. Die Menge

$$S_n^{(m)} = \{p \in S_n \mid p(m) = m\}$$

ist eine Untergruppe von S_n , denn aus $p(m) = q(m) = m$ folgt $pq(m) = p(m) = m$ und $p^{-1}(m) = m$.

Lemma 10.4.1 Für $p, q \in S_n$ gilt $pS_n^{(m)} = qS_n^{(m)}$ genau dann, wenn $p(m) = q(m)$.

Beweis: Sei $p(m) = j = q(m)$, also $q^{-1}(j) = m$. Dann ist $q^{-1}(p(m)) = q^{-1}(j) = m$, also $q^{-1}p \in S_n^{(m)}$. Umgekehrt folgt aus $q^{-1}p(m) = m$ sofort $p(m) = q(m)$. \square

Folgerung 10.4.1 Die Anzahl der Nebenklassen von S_n nach $S_n^{(m)}$ ist gleich n .

Beweis: Jede Nebenklasse ist durch das Bild der Ziffer m eindeutig bestimmt. \square

Da nun $S_n^{(m)} \cong S_{n-1}$ ist, folgt aus dem Satz von Lagrange und einer induktiven Argumentation, daß $|S_n| = n!$ ist.

Definition: Eine Permutation p heißt Zyklus, wenn es eine Teilmenge

$$\{i_1, \dots, i_m\} \subset \{1, \dots, n\}$$

gibt, so daß

$$p(i_k) = i_{k+1}, \quad k = 1, \dots, m-1,$$

$$p(i_m) = i_1,$$

$$p(j) = j \text{ sonst}$$

gilt. Wir schreiben dann $p = (i_1, \dots, i_m)$, z.B. $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = (1 \ 2 \ 3)$.

Zyklen, die keine Ziffern gemeinsam bewegen, heißen disjunkt.

Man kann zeigen, daß jede Permutation ein Produkt disjunkter Zyklen ist. Wir begnügen uns damit, dies an einem Beispiel zu demonstrieren:

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

Ein Zweierzyklus $(i \ j)$ heißt Transposition, Transpositionen haben das Signum -1 . Ein Zyklus der Länge k ist ein Produkt von $k-1$ Transpositionen, hat also das Signum $(-1)^{k-1}$, denn

$$(i_1, i_2, \dots, i_k) = (i_1, i_k) \cdots (i_1, i_3)(i_1, i_2).$$

Jede Transposition läßt sich als Produkt von „Nachbartranspositionen“ $(i, i+1)$ darstellen: Sei $i < j$, dann gilt

$$(i, j) = \underbrace{(j-1, j)(j-2, j-1) \dots (i+1, i+2)}_{\text{links vergrößern}} (i, i+1) \overbrace{(i+1, i+2) \dots (j-1, j)}_{\text{rechts verkleinern}},$$

denn j sinkt rechts bis zu i , i wächst links bis zu j , und für $i < k < j$ wird k rechts 1mal verkleinert und links 1mal vergrößert, also $k \mapsto k$.

Beispiel: $(5 \ 8) = (7 \ 8)(6 \ 7)(5 \ 6)(6 \ 7)(7 \ 8)$

Jede Transposition ist als Produkt einer ungeraden Zahl von Nachbartranspositionen darstellbar: In der Mitte steht $(i, i+1)$ und rechts und links stehen dieselben Transpositionen.

10.5 Gruppenoperationen

Definition: Sei X eine Menge und G eine Gruppe; wir nennen X eine G -Menge, wenn eine Abbildung $\cdot : G \times X \longrightarrow X$, $(g, x) \mapsto g \cdot x$ gegeben ist, so daß $g(hx) = (gh)x$ sowie $1x = x$ für alle $g, h \in G, x \in X$ gilt.

Beispiele:

1. $X = \{1, \dots, n\}$, $G = S_n$, $p \cdot i = p(i)$.
2. V sei ein \mathbb{R} -Vektorraum, hier operiert die multiplikative Gruppe von \mathbb{R} durch Multiplikation.
3. (A, V) sei ein affiner Raum, dann operiert die Vektor-Gruppe auf der Punkt-Menge durch Translation.
4. Wir wählen $X = G$, $g \cdot x$ sei das Produkt. Die Gruppe G operiert durch Linksmultiplikation auf sich selbst. Aber: Die Menge G mit der Rechtsmultiplikation ist keine G -Menge, das Assoziativgesetz ist verletzt, wenn G nicht kommutativ ist.
5. Wir betrachten wieder $X = G$ mit der Operation $g \cdot x = xg^{-1}$. Dann ist das Assoziativgesetz erfüllt.
6. Wieder $X = G$ mit der Konjugation als Operation: $g \cdot x = gxg^{-1}$.
7. Sei $H \subset G$ eine Untergruppe und $X = G/H = \{xH \mid x \in G\}$ die Menge der rechten Nebenklassen. Hier operiert G auf natürliche Weise.

Definition: X sei eine G -Menge und $x \in X$. Dann heißt $G_x = \{g \in G \mid gx = x\}$ der Stabilisator von x und $O_x = \{gx \mid g \in G\}$ heißt die Bahn (der Orbit) von x .

Bestimmen Sie die Stabilisatoren und Bahnen in den obigen Beispielen. Im Fall der Konjugation ist der Stabilisator von x die Menge der mit x vertauschbaren Elemente, die Bahn von x ist die Klasse der zu x konjugierten Elemente.

Satz 10.5.1 Sei X eine G -Menge und $x \in X$. Dann ist die Abbildung $f : G/G_x \longrightarrow O_x$, $f(gG_x) = gx$ bijektiv und mit der G -Operation verträglich, d.h. $f(g \cdot hG_x) = g \cdot f(hG_x)$.

Beweis: Die Abbildung f ist wohldefiniert, denn wenn $gG_x = hG_x$ ist, so ist $g^{-1}h \in G_x$, also $g^{-1}hx = x$, also $gx = hx$.

Injektivität: Sei $f(gG_x) = f(hG_x) = gx = hx$, dann ist $g^{-1}hx = x$, d.h. $g^{-1}h \in G_x$, also $gG_x = hG_x$.

Surjektivität: Sei $gx \in O_x$ beliebig, dann ist $f(gG_x) = gx$. □

Folgerung 10.5.1 Wenn G und X endlich sind, so gilt $|O_x| = \frac{|G|}{|G_x|}$.

Beweis: $|O_x| = |G/G_x| = \frac{|G|}{|G_x|}$. □

Folgerung 10.5.2 Die Anzahl der zu $g \in G$ konjugierten Elemente ist ein Teiler der Gruppenordnung.

Satz 10.5.2 (Cauchy) Sei G eine endliche Gruppe und p eine Primzahl mit $p \mid |G|$, dann gibt es ein $g \in G$ mit der Ordnung p , also $g^p = 1$.

Beweis: Es sei $X = \{(g_0, \dots, g_{p-1}) \in G \times \dots \times G \mid g_0 \cdots g_{p-1} = 1\}$, diese Menge ist nicht leer, denn sie enthält $(1, \dots, 1)$. Zu $g_0, \dots, g_{p-2} \in G$ gibt es ein eindeutig bestimmtes g_{p-1} , so daß $(g_0, \dots, g_{p-1}) \in X$ ist. Also ist $|X| = |G|^{p-1}$, d.h. $|X|$ ist ein Vielfaches von p . Wir interpretieren die Indizes von g_0, \dots, g_{p-1} als Elemente von $\mathbb{Z}/p\mathbb{Z} = H$. Die (additive) Gruppe H operiert auf X :

$$h \cdot (g_0, \dots, g_{p-1}) = (g_h, g_{h+1}, \dots, g_{h-1})$$

und X ist wirklich eine H -Menge:

$$0 \cdot x = x, (h+k) \cdot x = h \cdot (k \cdot x).$$

Nach der obigen Folgerung ist also

$$|O_x| = \frac{|\mathbb{Z}/p\mathbb{Z}|}{|G_x|} = \frac{p}{|G_x|},$$

also ist $|O_x| = 1$ oder $|O_x| = p$.

Die Bahn von $(1, \dots, 1)$ enthält nur dieses eine Element.

Wenn alle anderen Bahnen p Elemente hätten (es seien etwa k Stück), so wäre $|X| = 1 + kp$, also nicht durch p teilbar. Es muß also ein weiteres $x \in X$ geben, so daß $O_x = \{(g_0, \dots, g_{p-1})\}$ einelementig ist. Dann ist aber $g_0 = \dots = g_{p-1} \neq 1$, also $g_0^p = 1$. \square

Wir wollen nun noch einmal systematischer die Gruppen kleiner Ordnung untersuchen.

Lemma 10.5.1 Wenn $|G| = p$ eine Primzahl ist, so ist G zyklisch.

Beweis: Es sei $1 \neq g \in G$ beliebig, dann ist $\langle g \rangle$ eine nichttriviale Untergruppe von G , deren Ordnung ein Teiler von p , also gleich p ist. Somit ist $G = \langle g \rangle$. \square

Diedergruppen

Sei D_n die Menge der Kongruenzabbildungen, die ein regelmäßiges n -Eck in sich überführen. Sei $a \in D_n$ eine Drehung um $\alpha = 360/n$ Grad, dann ist a^k eine Drehung um $k \frac{360}{n}$ Grad und $a^n = 1$.

Sei b die Spiegelung an einer Geraden durch den Mittelpunkt und einen Eckpunkt des n -Ecks. Dann ist $b^2 = 1$.

Die Transformation ba^k können wir auf zwei Weisen realisieren: zuerst eine Drehung um $k \cdot \alpha$, dann spiegeln, oder zuerst spiegeln, dann eine Drehung um $-k \cdot \alpha$, also $ba^k = a^{-k}b$. Somit ist

$$D_n = \{1, a, a^2, \dots, a^{n-1}, b, ba, ba^2, \dots, ba^{n-1}\},$$

diese Gruppe hat $2n$ Elemente und wird durch die Relationen

$$a^n = 1, b^2 = 1, ab = ba^{-1}$$

charakterisiert.

Satz 10.5.3 Sei $p > 2$ eine Primzahl und $|G| = 2p$, dann ist $G = C_{2p}$ zyklisch oder $G = D_p$.

Beweis: Nach Cauchy existieren $x, y \in G$ mit $x^p = 1$, $y^2 = 1$. Wegen $2 \nmid p$ ist $y \notin \langle x \rangle$, also $x^k y \neq x^l$ für alle k, l . Das heißt

$$\langle x \rangle \cap \langle x \rangle \cdot y = \emptyset,$$

also

$$G = \langle x \rangle \cup \langle x \rangle \cdot y,$$

und analog folgt

$$G = \langle x \rangle \cup y \cdot \langle x \rangle,$$

also

$$\langle x \rangle y = y \langle x \rangle,$$

d.h. $\langle x \rangle$ ist eine normale Untergruppe.

Nun betrachten wir das Element xy , es hat die Ordnung 1, 2, p oder $2p$. Nun, der Fall 1 scheidet aus ($x \neq y$).

Wenn die Ordnung gleich $2p$ ist, so ist die Gruppe zyklisch.

Wenn die Ordnung gleich 2 ist, also $(xy)(xy) = 1$, so ist $yx = x^{-1}y$, also ist $G = D_p$.

Wenn schließlich die Ordnung gleich p sein sollte, so gälte

$$\langle x \rangle = \langle x \rangle (xy)^p = (\langle x \rangle xy)^p,$$

da $\langle x \rangle$ normal ist. Dann wäre aber

$$\langle x \rangle = \langle x \rangle y \langle x \rangle y \cdots \langle x \rangle y = \langle x \rangle y^p = \langle x \rangle y,$$

da p ungerade ist, ein Widerspruch. \square

Satz 10.5.4 Sei p eine Primzahl und $|G| = p^2$, dann ist $G = C_{p^2}$ oder $G = C_p \times C_p$.

Beweis: Es genügt zu zeigen, daß G kommutativ (abelsch) ist.

Sei O_g die Klasse der zu g konjugierten Elemente, sie enthält 1, p oder p^2 Elemente.

Die Menge $O_1 = \{1\}$ hat ein Element. Wenn für alle $g \neq 1$ die Bahn O_g mehr als ein Element hätte, also p oder p^2 Elemente, so wäre $|G| = 1 + kp \neq p^2$, ein Widerspruch.

Es gibt also ein $x \neq 1$ mit $|O_x| = 1$, d.h. $g^{-1}xg = x$ oder $xg = gx$ für alle $g \in G$.

Wenn die Ordnung von x gleich p^2 ist, so ist G zyklisch. Andernfalls ist die Ordnung von x gleich p , es gibt also ein $y \notin \langle x \rangle$. Dann sind die Elemente von G genau die $x^i y^k$, $0 \leq i, k \leq p-1$, also ist G abelsch. \square

Damit kennen wir alle Gruppen mit bis zu 15 Elementen, mit einer Ausnahme: $|G| = 8$. Die kommutativen Gruppen der Ordnung 8 sind $C_8, C_2 \times C_4, C_2 \times C_2 \times C_2$, außerdem kennen wir die Diedergruppe D_4 .

Es gibt noch eine weitere, die Quaternionengruppe

$$H = \{\pm 1, \pm i, \pm j, \pm k\}$$

mit

$$i^2 = j^2 = k^2 = -1, \quad ij = -ji = k.$$

Hier ist jede Untergruppe normal, denn die Untergruppen können nur die Ordnung 1, 2, 4 oder 8 haben; die einzige Untergruppe der Ordnung 2 ist $\{1, -1\}$, diese ist normal.

Satz 10.5.5 *Das sind alle Gruppen der Ordnung 8.*

Beweis: Sei G nichtkommutativ. Dann hat G kein Element der Ordnung 8 (sonst wäre sie zyklisch) und nicht alle Elemente haben die Ordnung 2 (sonst wäre sie kommutativ). Sei also $y \in G$ ein Element der Ordnung 4 und $x \notin \langle y \rangle$. Die Untergruppe $N = \langle y \rangle$ ist normal, denn sie hat den Index 2. Es ist $|G/N| = 2$, also $(xN)^2 = N = x^2N$, also $x^2 \in N$.

Fall $x^2 = y$ oder $x^2 = y^{-1}$ wäre, so hätte x die Ordnung 8. Folglich gilt $x^2 = 1$ oder $x^2 = y^2$. Wir behaupten: $xyx^{-1} = y^{-1}$.

Nun, es gilt $xNx^{-1} = N$, also $xyx^{-1} = y^k$ und bestimmen k :

Wegen $x^2 \in N$ gilt

$$y = x^2yx^{-2} = x(xyx^{-1})x^{-1} = xy^kx^{-1} = (xyx^{-1})^k = (y^k)^k = y^{k^2},$$

also $y^{k^2-1} = 1$. Demnach ist $k^2 - 1$ ein Vielfaches von 4, also ist k ungerade.

Wenn $k = 1$ wäre, also $xyx^{-1} = y$, d.h. $xy = yx$, so wäre G kommutativ. Es bleibt also nur $k = 3$, und das hatten wir behauptet.

Wir kommen nun zu den beiden Fällen zurück:

$x^2 = 1, y^4 = 1, xyx^{-1} = y^{-1}$, dies ist die Diedergruppe.

$x^2 = y^2, y^4 = 1, xyx^{-1} = y^{-1}$. Wir setzen $x = i, y = j$ und bezeichnen x^2 mit -1 und $xy = k$. Die i, j, k erfüllen nun die Relationen der Quaternionengruppe. \square

Wie gesagt haben wir damit alle Gruppen bis zur Ordnung 15 in der Hand. Bei der Ordnung 15 gibt es noch eine Besonderheit. Wir wissen, daß eine Gruppe von Primzahlordnung zyklisch ist, also: Wenn p eine Primzahl ist, so existiert nur eine Gruppe der Ordnung p . Die Umkehrung gilt nicht.

Satz 10.5.6 *Sei $|G| = 15$, dann ist $G = C_{15}$.*

Beweis: Es gibt $x, y \in G$ mit $x^5 = y^3 = 1$. Wir zeigen, daß $H = \langle x \rangle$ eine normale Untergruppe ist:

H operiert durch Linksmultiplikation auf G/H : $h \cdot gH = hg \cdot H$. Die Zahl der Elemente einer Bahn ist ein Teiler von 5. Wegen $|G/H| = 3$ hat also jede Bahn nur ein Element. Das heißt

$$hgH = gH \text{ oder } g^{-1}hg \in H \text{ für alle } h \in H, g \in G,$$

also ist H normal.

Nun betrachten wir den durch $f(h) = yhy^{-1}$ gegebenen Homomorphismus $f: H \rightarrow H$; dessen Kern ist offenbar gleich $\{1\}$, er ist also bijektiv. Es gilt $f^3 = id$. Wir zeigen $f^4 = id$:

Es gilt $f(x) = x^k$ und f ist durch k eindeutig bestimmt, mögliche Werte sind $k = 1, 2, 3, 4$. Wegen

$$f^l(x) = x^{k^l}$$

und

$$k^4 \equiv 1 \pmod{5}$$

gilt $f^4 = id$, also

$$f = f^4 \circ f^{-3} = id,$$

also $yhy^{-1} = h$ oder $yh = hy$. Wir setzen $K = \langle y \rangle$, dann hat $H \cdot K$ 15 Elemente, also

$$G = H \cdot K = C_5 \times C_3 = C_{15}.$$

□

10.6 Endlich erzeugte abelsche Gruppen

Eine abelsche Gruppe ist nichts anderes als eine kommutative Gruppe. Wir verwenden hier die additive Schreibweise. Das direkte Produkt $A_1 \times \dots \times A_n$ nennen wir hier die direkte Summe und bezeichnen sie mit $A_1 \oplus \dots \oplus A_n$.

Sei also A eine abelsche Gruppe und $a \in A$, dann schreiben wir als Abkürzung für $a + a + \dots + a$ (m Summanden) einfach $m \cdot a$. Umgekehrt, wenn $m \in \mathbb{Z}$ ist, so soll $ma = a + \dots + a$ (m Summanden, wenn $m \geq 0$) bzw. $ma = -a - \dots - a$ ($-m$ Summanden, wenn $m < 0$) gelten. (Später werden wir sehen, daß eine abelsche Gruppe auf diese Weise als \mathbb{Z} -Modul aufgefaßt werden kann.)

Wenn $|A| = n$ ist so gilt $na = 0$ für alle $a \in A$.

Wenn A eine abelsche Gruppe ist und $A_1, A_2 \subseteq A$ Untergruppen sind, so nennen wir in Analogie zur multiplikativen Schreibweise die Menge $A_1 + A_2 = \{a_1 + a_2 \mid a_i \in A_i\}$ als die Summe von A_1 und A_2 , dies ist die kleinste A_1 und A_2 umfassende Untergruppe von A .

Es gelte nun $A_1 + A_2 = A$, wenn zusätzlich $A_1 \cap A_2 = \{0\}$ ist, so schreiben wir $A = A_1 \oplus A_2$ und nennen dies eine direkte Summe.

Lemma 10.6.1 *Sei $A = A_1 \oplus A_2$, dann ist jedes Element $a \in A$ in eindeutiger Weise als $a = a_1 + a_2$, $a_i \in A_i$ darstellbar. Dies ist genau dann der Fall, wenn sich das Nullelement von A nur auf die triviale Weise als Summe von Elementen aus A_1 und A_2 darstellen läßt.*

Beweis: Da $A = A_1 + A_2$ gilt, gibt es für jedes $a \in A$ eine derartige Darstellung. Wir nehmen an, es gäbe zwei:

$$a = a_1 + a_2 = b_1 + b_2, \quad a_i, b_i \in A_i.$$

Dann ist $a_1 - b_1 = b_2 - a_2$, der linke Summand liegt in A_1 , der rechte in A_2 und wegen $A_1 \cap A_2 = \{0\}$ folgt $a_i = b_i$. □

Sie haben sicher bemerkt, daß wir den Begriff der direkten Summe auf zwei verschiedene Weisen verwenden (vgl. ganz oben). Nach dem soeben bewiesenen Lemma ist aber $A \times B \simeq (A \times \{0\}) \oplus (\{0\} \times B)$, was uns diese Schludrigkeit verzeiht.

Beispiel:

$\mathbb{Z}/6\mathbb{Z} = \{\bar{0}, \bar{2}, \bar{4}\} \oplus \{\bar{0}, \bar{3}\}$, wobei $\bar{m} = m + 6\mathbb{Z}$.

$\mathbb{Z}/4\mathbb{Z}$ läßt sich nicht in eine direkte Summe von Untergruppen zerlegen, da es nur eine einzige nichttriviale Untergruppe besitzt.

Definition: Sei A eine abelsche Gruppe, dann ist

$$t(A) = \{a \in A \mid \text{es gibt ein } m \in \mathbb{Z} \text{ mit } ma = 0\}$$

die Torsionsuntergruppe von A .

Lemma 10.6.2 $t(A)$ ist eine Untergruppe.

Beweis: Wenn $ma = 0 = nb$, so ist $mn(a + b) = 0$. □

Falls $|A| < \infty$ ist, so gilt $t(A) = A$.

Falls $t(A) = \{0\}$ ist, so heißt A torsionsfrei.

Definition: Sei p eine Primzahl und A eine abelsche Gruppe, dann heißt

$$A_p = \{a \in A \mid p^i a = 0 \text{ für ein } i > 0\}$$

die p -Torsionsuntergruppe von A .

Lemma 10.6.3 A_p ist eine Untergruppe von A .

Beweis: Wenn $p^i a = 0 = p^j b$ ist, so gilt $p^k(a + b) = 0$ für $k = \max(i, j)$. □

Wir teilen hier mit, daß es zu jedem Primteiler p der Ordnung n einer Gruppe für ein gewisses k eine Untergruppe der Ordnung p^k gibt. Folglich ist die Ordnung von A_p eine Potenz von p (derartige Gruppen heißen p -Gruppen).

Wir erhalten einen ersten Struktursatz:

Satz 10.6.1 Sei $|A| = n = p_1^{i_1} \dots p_k^{i_k}$ mit verschiedenen Primzahlen p_i , dann ist $A = A_{p_1} \oplus \dots \oplus A_{p_k}$.

Beweis: Wir führen die Induktion über die Anzahl der Primfaktoren von n .

Wenn $n = p^i$ ist, so gilt $A = A_p$, denn $p^i A = \{0\}$.

Sei $n = uv$ mit $\text{ggT}(u, v) = 1 = ru + sv$, dann ist

$$A = 1 \cdot A = ruA + svA \subseteq uA + vA \subseteq A,$$

also $A = uA + vA$. Sei $a \in uA \cap vA$, also $a = ub$ mit $b \in A$, dann gilt $va = vub = nb = 0$ und analog $ua = 0$, also $a = 0$. Also ist $A = uA \oplus vA$ eine direkte Summe und für diese Untergruppen kann die behauptete Zerlegung als bewiesen angenommen werden. □

Satz 10.6.2 Jede endliche abelsche p -Gruppe ist eine direkte Summe zyklischer Untergruppen.

Beweis: Sei $p^n A = \{0\}$ und $p^{n-1} A \neq \{0\}$, wir führen die Induktion über n (die Zahl p^n heißt die Periode von A).

Sei $n = 1$, also $pA = \{0\}$, dann ist A ein Vektorraum über dem Körper $K = \mathbb{Z}/p\mathbb{Z}$ (überprüfen Sie einfach die Vektorraumaxiome, die Multiplikation $\cdot : K \times A \rightarrow A$ ist durch $\bar{m} \cdot a = ma$ gegeben, wegen $\bar{0}a = pa = 0$ ist dies wohldefiniert). Aus der linearen Algebra ist bekannt, daß ein K -Vektorraum der Dimension n isomorph zu K^n ist, also gilt $A \simeq K^n \simeq \mathbb{Z}/p\mathbb{Z} \oplus \dots \oplus \mathbb{Z}/p\mathbb{Z}$.

Sei nun der Satz für Gruppen der Periode p^{n-1} bereits bewiesen. Es gilt $pA \subset A$, die Gruppe pA hat die Periode p^{n-1} , also gilt

$$pA = \langle a_1 \rangle \oplus \dots \oplus \langle a_k \rangle$$

und es gibt $h_1, \dots, h_k \in A$ mit $ph_i = a_i$. Wir setzen

$$H = \langle h_1 \rangle + \dots + \langle h_k \rangle$$

und behaupten, daß die Summe der $\langle h_i \rangle$ direkt ist. In der Tat, sei

$$0 = m_1 h_1 + \dots + m_k h_k$$

dann ist

$$0 = p0 = m_1 ph_1 + \dots + m_k ph_k = m_1 a_1 + \dots + m_k a_k,$$

also $m_i = 0$.

Sei $B \subset A$ die maximale Untergruppe mit $B \cap H = \{0\}$, wir nehmen an, daß $B + H \neq A$ wäre.

Sei also $a \notin B + H$, dann ist $pa = \sum m_i a_i = \sum m_i ph_i = ph$ für ein $h \in H$. Wir setzen $a' = a - h$, dann ist $a' \notin B + H$ und $pa' = 0$. Wir setzen $B' = \langle a', B \rangle$. Nach Konstruktion von B gilt $B' \cap H \neq \{0\}$, also gibt es ein $h' \in H$ mit

$$h' = ka' + b, \quad b \in B, \quad 0 < k < p.$$

Sei $sk \equiv 1 \pmod{p}$, dann ist $a' = ska' = sh' - sb \in H + B$, ein Widerspruch zur Konstruktion von a' . Somit gilt $A = B \oplus H$ und nach Induktionsvoraussetzung ist B eine direkte Summe zyklischer Untergruppen, somit gilt die Behauptung auch für A . \square

Wir werfen den Blick noch von einer anderen Richtung auf unsere abelschen Gruppen werfen:

Sei $A = \langle a_1, \dots, a_n \rangle$; der durch $f(e_i) = a_i$ gegebene Homomorphismus $F : \mathbb{Z}^n \rightarrow A$ ist surjektiv, deren Kern ist eine freie Untergruppe: $\text{Ker } f = \langle m_1, \dots, m_n \rangle, m_i \in \mathbb{Z}^n$. Sei M die Matrix mit den Spalten m_1, \dots, m_n . Solchen Matrizen gilt unser Augenmerk. Wenn M nämlich eine Diagonalmatrix mit den Einträgen d_1, \dots, d_n ist, so folgt $A \simeq \mathbb{Z}/d_1\mathbb{Z} \oplus \dots \oplus \mathbb{Z}/d_n\mathbb{Z}$.

Wir wollen ganzzahligen Matrizen Operationen folgenden Typs unterwerfen:

1. Vertauschen von Zeilen bzw. Spalten,
2. Addition des a -fachen einer Zeile zu einer anderen, dasselbe auch für Spalten.

Definition: Zwei Matrizen heißen äquivalent, wenn sie durch eine Folge von elementaren Operationen auseinander hervorgehen.

Zum Beispiel gehen die folgenden Matrizen durch elementare Operationen auseinander hervor:

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{pmatrix} \begin{pmatrix} 1 & & \\ & 1 & \\ & & 2 \end{pmatrix}$$

Satz 10.6.3 *Jede Matrix ist zu einer Matrix der Form*

$$\begin{pmatrix} i_1 & & & 0 \\ & \dots & & \\ & & i_r & \\ 0 & & & 0 \end{pmatrix}$$

äquivalent, wobei jeweils i_k ein Teiler von i_{k+1} ist.

Beweis: Durch Zeilen- und Spaltenvertauschungen wird erreicht, daß $|a_{11}|$ minimal ist. Die Zahl a_{1k} aus der ersten Zeile wird mit Rest durch a_{11} dividiert:

$$a_{1k} = qa_{11} + r, |r| < |a_{11}|.$$

Nun subtrahieren wir das q -fache der ersten Spalte von der k -ten Spalte, dann bleibt an der Stelle $(1, k)$ das r stehen. Wenn $r = 0$ ist, ist es gut, sonst bringen wir es an die Stelle $(1,1)$ und beginnen von vorn. Nach endlich vielen Schritten sind alle Elemente der ersten Zeile (außer dem ersten) gleich Null. Dasselbe veranstalten wir mit der ersten Spalte. Also ist A äquivalent zur Matrix

$$\begin{pmatrix} a_{11} & 0 & \dots & 0 \\ \vdots & & & \\ 0 & & A_1 & \end{pmatrix}$$

Wenn a_{11} alle Komponenten von A_1 teilt, so bleibt das auch bei allen Operationen, die wir künftig mit A_1 ausführen, erhalten. Wenn etwa a_{ij} nicht von a_{11} geteilt wird, so addieren wir die i -te Zeile zur ersten und beginnen von vorn. Dabei wird sich der Betrag von a_{11} weiter verkleinern. Wenn wir erreicht haben, daß a_{11} alle Komponenten von A_1 teilt, widmen wir uns A_1 und bringen es in Diagonalgestalt. Irgendwann sind wir fertig. \square

Wir fragen uns nun, ob die Zahlen i_1, i_2, \dots von den gewählten elementaren Operationen oder nur von der Matrix A abhängen. Die Antwort können wir aber erst etwas später geben. Zuerst überlegen wir uns, daß die Wirkung dieser Operationen durch Multiplikation mit Matrizen folgender Art realisiert werden kann:

$$\begin{pmatrix} 1 & & 0 \\ & \dots & \\ & r & \\ & \dots & \\ & & 1 \end{pmatrix}, \begin{pmatrix} 1 & & 0 \\ & \dots 1 & \\ & 1 \dots & \\ & & 1 \end{pmatrix}.$$

Dies sind Matrizen, deren Determinante gleich 1 ist, die also im Bereich der ganzzahligen Matrizen eine Inverse besitzen.

Definition: Sei $A = (a_{ij})$ eine ganzzahlige Matrix.

Sei d_1 der größte gemeinsame Teiler aller a_{ij} ,

d_2 der größte gemeinsame Teiler aller 2-Minoren von A ,

...

d_i der größte gemeinsame Teiler aller i -Minoren von A ,

...

$d_n = \det A$. Die d_i heißen die Determinantenteiler von A .

Lemma 10.6.4 Für alle i gilt: d_i teilt d_{i+1} .

Beweis: Nach dem Entwicklungssatz ist jeder $(i+1)$ -Minor von A eine Linearkombination von i -Minoren, also teilt d_i jeden $(i+1)$ -Minor und damit auch d_{i+1} . \square

Definition: Wir setzen $i_0 = 1$, $i_k = \frac{d_k}{d_{k-1}}$, die i_k heißen die Invariantenteiler von A .

Satz 10.6.4 Die Determinantenteiler einer Matrix ändern sich bei elementaren Operationen nicht. Äquivalente Matrizen haben dieselben Determinantenteiler.

Beweis: Wir betrachten die äquivalenten Matrizen A und PAQ , wo P und Q Produkte von Elementarmatrizen sind, ihre Inversen sind also auch ganzzahlige Matrizen. Sei b_j ein l -Minor von PAQ , nach dem verallgemeinerten Determinantenmultiplikationssatz gilt

$$b_j = \sum p_i a_i q_i,$$

wo die p_i, a_i, q_i jeweils gewisse l -Minoren von P, A bzw. Q sind. Nun sei d_l der l -te Determinantenteiler von A . Dann teilt d_l jedes a_i , also teilt es auch jeden l -Minor von PAQ und damit auch den l -ten Determinantenteiler von PAQ . Da durch Multiplikation von PAQ mit P^{-1} und Q^{-1} wieder A erhalten wird, stimmen die Determinantenteiler überein. \square

Satz 10.6.5 Sei A zu $\begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix}$ äquivalent, weiter möge jedes a_k ein Teiler von a_{k+1} sein, dann sind die a_k die Invariantenteiler von A .

Beweis: Beide Matrizen haben dieselben Determinantenteiler d_k , da sie äquivalent sind. Das Polynom a_1 teilt alle Elemente der zweiten Matrix, also ist $d_1 = a_1$. Die 2-Minoren haben die Form $a_i a_j$, sie werden alle von $a_1 a_2$ geteilt, also ist $d_2 = a_1 a_2$. Analog sieht man $d_k = a_1 \dots a_k$.

Nun ist $i_1 = d_1 = a_1$, allgemeiner

$$i_k = \frac{d_k}{d_{k-1}} = \frac{a_1 \dots a_k}{a_1 \dots a_{k-1}} = a_k. \square$$

Damit können wir unsere obige Frage beantworten: Die oben verbliebenen Diagonalelemente sind die Invariantenteiler der Matrix.

Folgerung 10.6.1 *Zwei Matrizen sind genau dann äquivalent, wenn sie dieselben Invariantenteiler besitzen.* \square

Zum Abschluß wollen wir noch torsionsfreie abelsche Gruppen untersuchen. Dazu benötigen wir ein Lemma über Matrizen, deren Komponenten ganzzahlig sind. Wir bemerken zuvor, daß die Inverse einer ganzzahligen Matrix, deren Determinante gleich 1 ist (solche Matrizen heißen unimodular), ebenfalls ganzzahlig ist .

Lemma 10.6.5 *Seien $x, y \in \mathbb{Z}$, dann gibt es eine unimodulare Matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ mit*

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} t \\ 0 \end{pmatrix}.$$

Beweis: Sei $t = \text{ggT}(x, y) = ax + by$, wir setzen $c = -\frac{y}{t}$, $d = \frac{x}{t}$, dann ist

$$\begin{pmatrix} a & b \\ -\frac{y}{t} & \frac{x}{t} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ -\frac{xy}{t} + \frac{xy}{t} \end{pmatrix} = \begin{pmatrix} t \\ 0 \end{pmatrix}$$

und $\det \begin{pmatrix} a & b \\ -\frac{y}{t} & \frac{x}{t} \end{pmatrix} = \frac{ax}{t} + \frac{by}{t} = \frac{t}{t} = 1$. \square

Lemma 10.6.6 *Sei A eine endlich erzeugte abelsche Gruppe und $a_1, \dots, a_n \in A$, $x_1, \dots, x_n \in \mathbb{Z}$ vorgegeben. Dann gibt es $b_1, \dots, b_n \in A$ mit*

$$\langle a_1, \dots, a_n \rangle = \langle b_1, \dots, b_n \rangle$$

und $x_1 a_1 + \dots + x_n a_n = t b_1$, wobei $t = \text{ggT}(x_1, \dots, x_n)$ ist.

Beweis: Wir beginnen mit $n = 2$. Wir wählen eine unimodulare Matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ mit

$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} t \\ 0 \end{pmatrix}$ und setzen $(b_1, b_2) = (a_1, a_2) \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1}$, dann ist jede Linearkombination von b_1, b_2 auch eine von a_1, a_2 und umgekehrt. Weiter ist

$$\begin{aligned} a_1 x_1 + a_2 x_2 &= (a_1, a_2) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (b_1, b_2) \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= (b_1, b_2) \begin{pmatrix} t \\ 0 \end{pmatrix} = t b_1. \end{aligned}$$

Nun sei $n = 3$. Dann gibt es b_2, b_3 mit $\langle a_2, a_3 \rangle = \langle b_2, b_3 \rangle$ und $x_2 b_2 + x_3 b_3 = d b_2$, wie wir soeben sahen. Ebenso gibt es b_1, b'_2 , so daß $\langle a_1, d b_2 \rangle = \langle b_1, b'_2 \rangle$ und $x_1 a_1 + 1 \cdot d b_2 = t b_1$. Dann gilt $\langle a_1, a_2, a_3 \rangle = \langle b_1, b'_2, b_3 \rangle$ und $x_1 a_1 + x_2 a_2 + x_3 a_3 = t b_1$. Und so weiter. \square

Definition: Sei A eine abelsche Gruppe, dann heißen die Elemente a_1, \dots, a_n linear unabhängig, wenn aus $\sum x_i a_i = 0$ ($x_i \in \mathbb{Z}$) folgt, daß $x_i = 0$ für alle i gilt.

Eine abelsche Gruppe heißt frei, wenn sie ein linear unabhängiges Erzeugendensystem besitzt. Eine endlich erzeugte freie abelsche Gruppe ist isomorph zu $\mathbb{Z} \times \dots \times \mathbb{Z}$.

Satz 10.6.6 Sei A eine endlich erzeugte torsionsfreie abelsche Gruppe und $r(A)$ die Minimalzahl von Erzeugenden von A . Dann gilt:

1. Jedes Erzeugendensystem von A mit $r(A)$ Elementen ist linear unabhängig.
2. A ist frei.

Beweis: Sei $\{a_1, \dots, a_n\}$ ein Erzeugendensystem von A mit $n = r(A)$ und es sei $x_1 a_1 + \dots + x_n a_n = 0$. Dann gibt es ein Erzeugendensystem b_1, \dots, b_n mit $t b_1 = x_1 a_1 + \dots + x_n a_n = 0$, hier muß $b_1 \neq 0$ sein, denn $\{b_2, \dots, b_n\}$ enthält zuwenig Elemente, um A zu erzeugen. Folglich muß $t = 0$ sein, wenn eines der x_i von Null verschieden wäre, so wäre auch $t \neq 0$, also sind die a_i linear unabhängig. \square

10.7 Lineare Codes

Eine „Nachricht“ ist eine Folge von Nullen und Einsen, wir werden diese Ziffern kurz „Bits“ nennen. Ein „Wort“ der Länge n ist also eine Bitfolge der Länge n , mit Folgen der Länge 10 kann man also 1024 verschiedene Worte darstellen.

Bei der Benutzung eines Computers denkt man sofort, daß ein Wort, z.B. 0 0 0 1 1 1 0 als binäre Darstellung einer natürlichen Zahl aufgefaßt werden kann, hier wäre es $2^3 + 2^2 + 2^1 = 14$. Wir wollen ein Wort aber lieber als Element von $\{0, 1\}^n$ auffassen, weil diese Menge eine algebraische Struktur besitzt:

$$\{0, 1\}^n \leftrightarrow (\mathbb{Z}/2\mathbb{Z})^n,$$

dies können wir als direkte Summe abelscher Gruppen auffassen, besser aber noch als n -dimensionalen Vektorraum über dem Körper $\mathbb{Z}/2\mathbb{Z}$, den wir mit $\mathbf{2}$ bezeichnen wollen. Das Rechnen in $\mathbf{2}^n = \mathbb{W}$ geschieht also komponentenweise modulo 2. Diese Rechenart kann auch ein Computer ausführen: Die komponentenweise Addition zweier `int`-Zahlen bei Java wird mit dem `xor`-Operator (exclusive or) `^` ausgeführt: `14 ^ 15 = 1`.

Durch physikalische Einflüsse auf den Übertragungskanälen können einzelne Bits verändert werden, nicht aber die Wortlänge.

Der Hamming-Abstand zweier Worte ist die Anzahl

$$d(v, w) = |\{i \mid v_i \neq w_i\}|.$$

Eigenschaften:

1. $d(v, w) \geq 0$
2. $d(v, v) = 0$
3. $d(v, w) = d(w, v)$
4. $d(u, m) \leq d(u, v) + d(v, w)$

Dies heißt, daß d eine „Metrik“ ist; die letzte Eigenschaft heißt „Dreiecksungleichung“, zum Beweis machen Sie sich einfach eine Skizze.

Wenn $v \in \mathbb{W}$ und $r \in \mathbb{N}$ ist, so sei

$$S_r(v) = \{w \in \mathbb{W} \mid d(v, w) \leq r\},$$

dies ist die Kugel um v mit dem Radius r .

Ein Code C ist eine Teilmenge von \mathbb{W} ; C heißt t -fehlerkorrigierend, wenn für $v, w \in C$ gilt $d(v, w) \geq 2t + 1$. Verschiedene Codeworte von C liegen dann in disjunkten Kugeln vom Radius t .

Als Beispiel betrachten wir den folgenden Code $C \subset \mathbf{2}^7$, der 1-fehlerkorrigierend ist.
¹ In der zweiten Spalte ist die dezimale Darstellung der Interpretation der Codeworte als Binärzahlen angegeben; die letzten 8 Worte sind komplementär zu den ersten 8.

0000000	0
0001110	14
0010111	23
0011001	25
0100101	37
0101011	47
0110010	50
0111100	60
1111111	127
1110001	123
1101000	104
1100110	102
1011010	90
1010100	84
1001101	77
1000011	67

Um eine Nachricht dekodieren zu können, muß jedes Codewort gespeichert werden, zum Vergleich Nachrichtenwort/Codewort sind je $|C|$ Vergleiche nötig; zur Bestimmung des Hammingabstands zweier Codeworte sind $|C|^2$ Vergleiche nötig.

Ein Code $C \subset \mathbb{W}$ heißt linear, wenn C ein Unterraum des Vektorraums \mathbb{W} ist.

Sei C ein linearer Code der Dimension k und $\{c_1, \dots, c_k\}$ eine Basis von C . Mit $G \in M_{kn}(\mathbf{2})$ bezeichnen wir die Matrix mit den Zeilen c_1, \dots, c_k , sie heißt die Generatormatrix von C . In unserem Beispiel kann

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

gewählt werden, die Zeilen entsprechen den Dezimalzahlen 67, 37, 23, 14.

Das Gewicht $w(x)$ eines Wortes sei die Zahl der Bits $\neq 0$: $w(x) = d(x, 0)$. Mit $w(C)$ bezeichnen wir das minimale Gewicht von Worten aus C .

¹Das Beispiel stammt von Beutelspacher, Lineare Algebra, Vieweg 2003

Satz 10.7.1 *Der minimale Hamming-Abstand $d(C)$ zweier Worte eines linearen Codes C ist $d(C) = w(C)$.*

Beweis: 1. $d(C) = \min\{d(c, c')\} \leq \min\{d(c, 0)\} = w(C)$.

2. Wir zeigen, daß ein Wort $c_0 \in C$ existiert, wo das Minimalgewicht angenommen wird und seien $c, c' \in C$ mit $d(c, c') = d(C)$. Dann ist $w(c - c') = d(c - c', 0) = d(c, c') = d(C)$, wir setzen $c_0 = c - c'$, dieses Wort liegt in C . \square

Sei $C \subset \mathbb{W}$ ein linearer Code; wir nennen

$$C^\perp = \{v \in \mathbb{W} \mid \sum c_i v_i = 0 \pmod{2} \text{ für alle } c \in C\}$$

den zu C orthogonalen Code ².

Satz 10.7.2 *C^\perp ist ein Unterraum von \mathbb{W} der Dimension $n - k$.*

Beweis: Es ist $v \in C^\perp$ genau dann, wenn v eine Lösung des homogenen Gleichungssystems $Gv = 0$ ist; der Rang von G ist gleich k . \square

Folgerung 10.7.1 $(C^\perp)^\perp = C$.

Beweis: Es ist $C \subseteq (C^\perp)^\perp$, denn für $c \in C$ gilt $cv = 0$ für alle $v \in C^\perp$, d.h. $c \in (C^\perp)^\perp$; die Behauptung folgt aus der Dimensionsformel. \square

Sei nun $H \in M_{n, n-k}(\mathbf{2})$ eine Matrix, deren Spalten eine Basis von C^\perp bilden, sie heißt Kontrollmatrix von C . Für $v \in \mathbb{W}$ heißt der Zeilenvektor $s(v) = vH \in \mathbf{2}^{n-k}$ das Syndrom von v .

Satz 10.7.3 $C = \{v \in \mathbb{W} \mid s(v) = 0\}$.

Der Beweis ist trivial: $s(v) = 0 = vH$ heißt $v \perp C^\perp$, also $v \in (C^\perp)^\perp = C$. \square

Folgerung 10.7.2 *Es ist $s(v) = s(w)$ genau dann, wenn die Nebenklassen $v + C$ und $w + C$ übereinstimmen.*

Beweis: $s(v) = s(w)$ gdw. $vH = wH$ gdw. $(v - w)H = 0$ gdw. $v - w \in C$ gdw. $v + C = w + C$. \square

Ein Wort $u \in v + C$ heißt Anführer der Nebenklasse, wenn sein Gewicht $w(u)$ minimal in $v + C$ ist.

Satz 10.7.4 *Sei $C \subset \mathbb{W}$ ein linearer t -korrigierender Code. Dann ist jeder Vektor v vom Gewicht $\leq t$ Anführer einer Nebenklasse. Die Anführer derjenigen Nebenklassen, die ein Wort vom Gewicht $\leq t$ enthalten, sind eindeutig bestimmt.*

²Beutelspacher nennt ihn den zu C dualen Code, distanziert sich aber von dieser Benennung.

Beweis: Sei $w(v) \leq t$ und $v \neq v' \in v + C$; wir zeigen $w(v') > t$: Wegen $v + C = v' + C$ gilt $v - v' \in C$, also $w(v - v') \geq 2t + 1$, also

$$2t + 1 \leq w(v - v') = d(v - v', 0) = d(v, v') \leq d(v, 0) + d(0, v') =$$

$$w(v) + w(v') \leq t + w(v'), \text{ also } w(v') \geq t + 1.$$

□

Beim Senden des Worts x geht dies evtl. in das Wort $v = x + f$ über, wir nennen f den Fehlervektor, dieser hat das gleiche Syndrom wie v , denn wegen $x \in C$ ist $v + C = f + C$, und sein Gewicht ist $\leq t$, also ist f der Anführer der Nebenklasse $x + C$ und $x + f + f = x$ ist das korrigierte Codewort.

Die Dekodierung kann also wie folgt verlaufen: Sei v das empfangene Wort; wir berechnen $s(v)$ und suchen diesen Vektor in der Liste aller möglichen Syndrome, bestimmen den zugehörigen Nebenklassenanführer f und dekodieren v zu $x = v + f$.

Im obigen Beispiel sei H die folgende Matrix mit $G \cdot H = 0$:

```
100
010
001
110
101
011
111
```

Es ist $\dim C = 4$, also $|C| = 2^4 = 16$, die Anzahl der Nebenklassen von \mathbb{W} modulo C ist gleich

$$|\mathbb{W}/C| = \frac{|\mathbb{W}|}{|C|} = \frac{128}{16} = 8.$$

Die Anführer und ihre Syndrome sind

```
0000000 000
0000001 111
0000010 011
0000100 101
0001000 110
0010000 001
0100000 010
1000000 100
```

Wenn z.B. $v = 0010001$ empfangen wird, so ist $s(v) = 110$, $f = 0001000$, $c = v + f = 0011001$.

10.8 Ringe und Moduln

Definition: Sei R eine Menge, in der zwei Operationen

$$+ : R \times R \rightarrow R \quad (r, s) \mapsto r + s$$

und

$$\cdot : R \times R \rightarrow R \quad (r, s) \mapsto r \cdot s$$

gegeben sind; R heißt (zusammen mit den gegebenen Operationen) ein Ring, wenn die „üblichen“ Rechenregeln gelten:

1. $(r + s) + t = r + (s + t)$ für alle $r, s, t \in R$, (Assoziativgesetz der Addition)
2. es gibt ein Element 0 mit $r + 0 = r$ für alle $r \in R$, (Existenz eines neutralen Elements)
3. zu jedem $r \in R$ gibt es ein r' mit $r + r' = 0$, (Existenz eines zu r inversen Elements, man schreibt für r' gewöhnlich $-r$)
4. $r + s = s + r$ für alle $r, s \in R$ (Kommutativgesetz der Addition)
5. $(rs)t = r(st)$ für alle $r, s, t \in R$ (Assoziativgesetz der Multiplikation)
6. $(r + s)t = rt + st$ für alle $r, s, t \in R$ (1. Distributivgesetz)
7. $t(r + s) = tr + ts$ für alle $r, s, t \in R$ (2. Distributivgesetz)
8. es gibt ein Element $1 \in R$ mit $1r = r$ für alle $r \in R$, (Existenz eines neutralen Elements)

Wenn zusätzlich

9. $rs = sr$ für alle $r, s \in R$ (Kommutativgesetz der Multiplikation) erfüllt ist, so heißt R ein kommutativer Ring.

Beispiele für kommutative Ringe sind \mathbb{Z} , \mathbb{Q} , \mathbb{R} , \mathbb{C} , $K[x]$, $\mathbb{Z}/m\mathbb{Z}$, während die Menge M_{nn} der quadratischen n -reihigen Matrizen ein nichtkommutativer Ring ist. Wir werden uns vorwiegend mit nichtkommutativen Ringen beschäftigen.

Eine additive kommutative Gruppe M heißt linker R -Modul, wenn eine Operation

$$\cdot : R \times M \rightarrow M \quad ((r, m) \mapsto r \cdot m)$$

gegeben ist, so daß wiederum die üblichen Rechenregeln gelten:

1. $r(sm) = (rs)m$,
2. $r(m + n) = rm + rn$,
3. $(r + s)m = rm + sm$,

$$4. 1m = m.$$

Eine additive kommutative Gruppe M heißt rechter R -Modul, wenn eine Operation

$$\cdot : M \times R \rightarrow M \quad ((m, r) \mapsto m \cdot r)$$

gegeben ist, so daß wiederum die üblichen Rechenregeln gelten:

1. $(mr)s = m(rs)$,
2. $(m + n)r = mr + nr$,
3. $m(r + s) = mr + ms$,
4. $m1 = m$.

Wenn wir den Begriff „Modul“ ohne Attribut verwenden, so meinen wir linke Moduln.

Beispiele:

Ein Vektorraum über einem Körper K ist ein K -Modul. Eine abelsche Gruppe ist ein \mathbb{Z} -Modul. Die Menge M_{n1} aller Spaltenvektoren ist ein linker M_{nn} -Modul. Die Menge M_{1n} aller Zeilenvektoren ist ein rechter M_{nn} -Modul. Jeder Ring R ist sowohl ein linker als auch ein rechter R -Modul.

Sei $U \subseteq M$ eine additive Untergruppe des R -Moduls M . Wenn für alle $r \in R$ und $u \in U$ gilt $ru \in U$, so nennen wir U einen Untermodul von M .

Seien M und N linke R -Moduln und $f : M \rightarrow N$ ein Homomorphismus der additiven Gruppen. Wir nennen f eine R -lineare Abbildung (oder einen R -Modulhomomorphismus), wenn $f(rm) = rf(m)$ für alle $r \in R$ und $m \in M$ gilt.

Lemma 10.8.1 Sei $f : M \rightarrow N$ ein R -Homomorphismus und $U \subseteq M$ sowie $V \subseteq N$ Untermoduln. Dann sind auch

$$f(U) = \{n \in N \mid \text{es gibt ein } u \in U \text{ mit } n = f(u)\} \subseteq N$$

und

$$f^{-1}(V) = \{m \in M \mid f(m) \in V\} \subseteq M$$

Untermoduln. □

Speziell sind $f(M) = \text{Im}(f)$ und $f^{-1}(\{0\}) = \text{Ker}(f)$ Untermoduln. Ein R -Homomorphismus $f : M \rightarrow N$ ist genau dann surjektiv, wenn $\text{Im}(f) = N$ ist und genau dann injektiv, wenn $\text{Ker}(f) = \{0\}$ ist. Ein injektiver und surjektiver R -Homomorphismus heißt Isomorphismus.

Sei M ein R -Modul und $U \subseteq M$ ein Untermodul. Die Relation \sim auf M , die durch

$$m \sim m' \text{ gdw. } m - m' \in U$$

gegeben ist, ist eine Äquivalenzrelation und aus $m \sim m'$ folgt $rm \sim rm'$ für alle $r \in R$. Die Menge der Äquivalenzklassen wird mit M/U bezeichnet, die Elemente von M/U haben die Form $m + U$ mit $m \in M$.

Die Faktorgruppe M/U wird ein R -Modul, wenn wir eine Multiplikation wie folgt einführen:

$$r(m + U) = rm + U.$$

(Die Mengen auf der linken und der rechten Seite stimmen überein, was die Repräsentantenunabhängigkeit der Definition zeigt. Das Überprüfen der Modulaxiome wollen wir uns ersparen.)

Wenn $U, V \subseteq M$ Untermoduln sind, so ist die Summe der Untergruppen $U+V$ ebenfalls ein Untermodul, und wenn $U \cap V = \{0\}$ gilt, so nennen wir die Summe direkt und schreiben $U \oplus V$. In diesem Fall läßt sich jedes Element $m \in M$ in genau einer Weise in der Form $m = u+v$ mit $u \in U$ und $v \in V$ schreiben. Wir können also zwei Abbildungen $p_U : M \rightarrow U$ und $p_V : M \rightarrow V$ definieren:

$$p_U(m) = u, \quad p_V(m) = v.$$

Diese Abbildungen sind R -linear und es gilt

$$p_U \circ p_U = p_U, \quad p_V \circ p_V = p_V, \quad p_U \circ p_V = p_V \circ p_U = 0, \quad p_U + p_V = id_M.$$

Wir nennen diese die Projektionen auf die Summanden U, V .

Da jeder R -Homomorphismus $f : M \rightarrow N$ auch ein Gruppenhomomorphismus ist, haben wir nach dem Homomorphiesatz einen Isomorphismus

$$F : M/\text{Ker}(f) \rightarrow \text{Im}(f),$$

der durch $F(m + \text{Ker}(f)) = f(m)$ gegeben ist. Dies ist sogar ein R -Isomorphismus, denn

$$F(r(m + \text{Ker}(f))) = f(rm) = rf(m) = rF(m + \text{Ker}(f)).$$

Nun können wir die beiden Isomorphiesätze, die ja direkte Folgerungen aus dem Homomorphiesatz waren, analog herleiten:

$$(U + V)/U \simeq V/U \cap V$$

$$(M/U)/(V/U) \simeq M/V \text{ für } U \subseteq V \subseteq M$$

Seien $m_1, \dots, m_k \in M$ und $r_1, \dots, r_k \in R$, dann heißt $\sum r_i m_i$ eine Linearkombination der m_i . Wenn $N \subseteq M$ eine Teilmenge ist, so bezeichnen wir mit RN die Menge aller Linearkombinationen von Elementen aus N . Falls $RN = M$ gilt, so heißt N ein Erzeugendensystem des Moduls M .

Die Elemente m_1, \dots, m_k heißen linear unabhängig, wenn aus

$$\sum r_i m_i = 0 \quad (r_i \in R)$$

folgt, daß $r_1 = \dots = r_k = 0$ gilt. Ein linear unabhängiges Erzeugendensystem von M heißt eine Basis, ein R -Modul, der eine Basis besitzt, heißt frei.

Lemma 10.8.2 Sei M ein freier R -Modul und $\{m_1, \dots, m_n\}$ eine Basis von M , dann ist $M \simeq R \times \dots \times R = R^n$.

Beweis: Jedes $m \in M$ läßt sich in eindeutiger Weise als Linearkombination $\sum r_i m_i$ darstellen, wir ordnen m das n -tupel $(r_1, \dots, r_n) \in R^n$ zu. \square

Lemma 10.8.3 *Jeder endlich erzeugte R -Modul ist isomorph zu einem Faktormodul eines freien R -Moduls.*

Beweis: Sei $M = R\{m_1, \dots, m_n\}$ und $m \in M$ beliebig, also $m = \sum r_i m_i$, dann ist die Abbildung $f : R^n \rightarrow M$ mit $f(r_1, \dots, r_n) \mapsto \sum r_i m_i$ surjektiv. Wir setzen $U = \text{Ker}(f)$, dann gilt nach dem Homomorphiesatz $M \simeq R^n/U$. \square

Definition: Eine additive Untergruppe $L \subseteq R$ eines Rings R heißt Linksideal, wenn $rL \subseteq L$ für alle $r \in R$ gilt. Ein Linksideal ist also ein Untermodul des linken R -Moduls R .

Eine additive Untergruppe $D \subseteq R$ eines Rings R heißt Rechtsideal, wenn $Dr \subseteq D$ für alle $r \in R$ gilt. Ein Rechtsideal ist also ein Untermodul des rechten R -Moduls R .

Eine Teilmenge $I \subseteq R$, die sowohl ein Links- als auch ein Rechtsideal ist, heißt (zweiseitiges) Ideal.

Seien R und S zwei Ringe. Ein Homomorphismus $f : R \rightarrow S$ der additiven Gruppen von R und S heißt Ringhomomorphismus, wenn $f(r_1 r_2) = f(r_1) f(r_2)$ und $f(1) = 1$ gilt. Als Kern von f bezeichnen wir wieder die Menge

$$\text{Ker}(f) = \{r \in R \mid f(r) = 0\}.$$

Ein Ringhomomorphismus ist genau dann injektiv, wenn $\text{Ker}(f) = \{0\}$ ist.

Lemma 10.8.4 *Sei $f : R \rightarrow S$ ein Ringhomomorphismus, dann ist $\text{Ker}(f)$ ein Ideal von R .*

Beweis: Die Abbildung f ist ein Gruppenhomomorphismus, also ist $\text{Ker}(f) \subseteq R$ eine Untergruppe. Sei $a \in \text{Ker}(f)$ und $r \in R$, dann gilt

$$f(ra) = f(r)f(a) = f(r) \cdot 0 = 0,$$

$$f(ar) = f(a)f(r) = 0 \cdot f(r) = 0,$$

also $ra \in \text{Ker}(f)$ und $ar \in \text{Ker}(f)$. \square

Sei $I \subseteq R$ ein Ideal, dies ist insbesondere ein linker R -Untermodul, also können wir den Faktormodul R/I bilden. Wir führen in R/I eine Multiplikation ein:

$$(r + I)(s + I) = rs + I.$$

Wir zeigen, daß dies eine repräsentantenunabhängige Definition ist: Sei $r + I = r' + I$ und $s + I = s' + I$, also $a = r - r' \in I$ und $b = s - s' \in I$. Dann ist

$$(r' + I)(s' + I) = (r + a + I)(s + b + I) = (r + a)(s + b) + I =$$

$$rs + as + rb + ab + I = rs + I,$$

da die übrigen Summanden in I liegen.

Zum Ideal $I \subseteq R$ haben wir den kanonischen Homomorphismus

$$f : R \rightarrow R/I, f(r) = r + I,$$

dessen Kern gleich I ist.

Definition: Ein Ideal $I \subseteq R$ eines kommutativen Rings R heißt Hauptideal, wenn es aus allen Vielfachen eines Elements a besteht: $I = aR$.

Wir betrachten als Beispiel den Ring \mathbb{Z} .

Sei $I \subseteq \mathbb{Z}$ ein Ideal. Wir wollen zeigen, daß I ein Hauptideal ist. Sei $0 \neq a \in I$ das Element von minimalem Betrag. Wir zeigen, daß I von a erzeugt wird: Sei $b \in I$ ein beliebiges Element, wir dividieren mit Rest:

$$b = qa + r, 0 \leq r < a.$$

Wenn $r \neq 0$ wäre, so wäre $r = b - qa \in I$ im Widerspruch zur Minimalität von a , also ist $r = 0$ und $a \mid b$. \square

Wir wollen uns nun etwas genauer mit Polynomen mit rationalen bzw. ganzzahligen Koeffizienten beschäftigen.

Wenn K ein Körper ist, so ist der Polynomring $K[x]$ ein Hauptidealring. Man beweist dies wie oben mithilfe der Restdivision.

Definition: Ein Polynom $p(x) \in \mathbb{Q}[x]$ heißt irreduzibel (oder Primpolynom), wenn in jeder Zerlegung $p(x) = f(x)g(x)$ mit $f, g \in \mathbb{Q}[x]$ einer der Faktoren ein konstantes Polynom ist.

Bei einer Zerlegung eines Polynoms in ein Produkt von Polynomen kommt es uns auf konstante Faktoren nicht an. Wenn wir ein Polynom $f(x) \in \mathbb{Q}[x]$ mit einer geeigneten ganzen Zahl multiplizieren, so daß sich alle Nenner der Koeffizienten wegheben, erhalten wir ein Polynom mit ganzzahligen Koeffizienten. Wir werden sehen, daß bei diesem Übergang die Irreduzibilität erhalten bleibt.

Wir beweisen dazu zwei Resultate, die in der Literatur häufig unter den unten verwendeten Namen zu finden sind.

Definition: Sei $f(x) = a_0x^n + a_1x^{n-1} + \dots + a_n \in \mathbb{Z}[x]$, dann heißt die Zahl $\text{cont}(f) = \text{ggT}(a_0, \dots, a_n)$ der Inhalt von $f(x)$. Ein Polynom $f(x) \in \mathbb{Z}[x]$ heißt primitiv, wenn sein Inhalt gleich 1 ist.

Lemma 10.8.5 (Hilfssatz von Gauß) *Das Produkt primitiver Polynome ist primitiv.*

Beweis: Seien $f, g \in \mathbb{Z}[x]$ primitiv und $h = f \cdot g$ sei nicht primitiv. Dann besitzen die Koeffizienten von $h(x)$ einen gemeinsamen Primfaktor p . Wenn wir jeden Koeffizienten der Polynome durch seine Restklasse modulo p ersetzen, erhalten wir Polynome $f_p, g_p, h_p \in \mathbb{Z}/p\mathbb{Z}[x]$, für die gilt $h_p = f_p g_p$. Nun ist aber h_p das Nullpolynom und f_p und g_p sind keine Nullpolynome. Dieser Widerspruch beweist die Primitivität von h . \square

Satz 10.8.1 (Satz von Gauß) *Wenn $f(x) \in \mathbb{Z}[x]$ in $\mathbb{Q}[x]$ zerlegbar ist, so ist $f(x)$ bereits in $\mathbb{Z}[x]$ zerlegbar.*

Beweis: Für jedes Polynom $g(x) \in \mathbb{Q}[x]$ gibt es ein Polynom $g^\#(x) \in \mathbb{Z}[x]$ mit $g(x) = \frac{1}{b} \cdot g^\#(x)$ und $b \in \mathbb{Z}$. Sei noch $a = \text{cont}(g^\#(x))$, dann gibt es ein primitives Polynom $g^*(x)$ und

$$g(x) = \frac{a}{b} \cdot g^*(x).$$

Sei nun

$$f(x) = g_1(x)g_2(x) = \frac{a}{b}g_1^*(x)g_2^*(x)$$

mit primitiven Polynomen g_1^*, g_2^* . Links steht ein Polynom mit ganzzahligen Koeffizienten und das Produkt $g_1^*(x)g_2^*(x)$ ist primitiv, also kann sich der Nenner b gegen keinen Koeffizienten der rechten Polynome wegekürzen, also muß $\frac{a}{b}$ eine ganze Zahl sein. \square

Ein Kriterium, ob ein Polynom mit ganzzahligen Koeffizienten irreduzibel ist, ist durch den folgenden Satz gegeben.

Satz 10.8.2 (Satz von Eisenstein) Sei $f(x) = a_n x^n + \dots + a_0 \in \mathbb{Z}[x]$ und p eine Primzahl, so daß p kein Teiler von a_n ist, $p \mid a_{n-1}, \dots, p \mid a_0$, aber p^2 kein Teiler von a_0 ist. Dann ist $f(x)$ irreduzibel.

Beweis: Sonst wäre

$$f(x) = (b_m x^m + \dots + b_0)(c_l x^l + \dots + c_0)$$

und oBdA $p \mid b_0$, p teilt nicht c_0 , da ja $a_0 = b_0 c_0$ gilt. Nun sind nicht alle b_i durch p teilbar, denn sonst wäre p ein Teiler von a_n . Sei also

$$b_0 \equiv \dots \equiv b_{k-1} \equiv 0 \pmod{p},$$

und p ist kein Teiler von b_k für ein $k \leq m < n$. Dann ist

$$a_k = \sum_{i=0}^k b_i c_{k-i} \equiv b_k c_0 \not\equiv 0 \pmod{p},$$

ein Widerspruch. \square

Wir wenden dieses Kriterium auf das Polynom $f(x) = \frac{x^p - 1}{x - 1} = x^{p-1} + x^{p-2} + \dots + x + 1$ an, wo p eine Primzahl ist.

Wir setzen $x = y + 1$:

$$\frac{(y+1)^p - 1}{y} = \frac{1}{y} \sum_{i=1}^p \binom{p}{i} y^i = y^{p-1} + \binom{p}{1} y^{p-2} + \dots + \binom{p}{p-1}.$$

Die Binomialkoeffizienten

$$\binom{p}{i} = \frac{p \cdot (p-1) \cdot \dots \cdot (p-i+1)}{1 \cdot 2 \cdot \dots \cdot i} \in \mathbb{Z}$$

sind ganzzahlig, aber der Faktor p kann sich nicht gegen die kleineren Zahlen im Nenner kürzen, also sind sie durch p teilbar. Der erste Summand ist nicht durch p teilbar und der letzte nicht durch p^2 , also ist $f(x)$ nach dem Satz von Eisenstein irreduzibel.

10.9 Polynome

Satz 10.9.1 (*Division mit Rest*): Seien $f(x), g(x) \in \mathbb{R}[x]$ gegeben, dann gibt es eindeutig bestimmte Polynome $q(x), r(x)$, so daß $f(x) = q(x) \cdot g(x) + r(x)$ gilt, wobei der Grad von $r(x)$ kleiner als der Grad von $g(x)$ ist.

Beweis: Wenn $\deg(g) > \deg(f)$ ist, so setzen wir $q = 0$ und $r = f$. Weiterhin sei $\deg(f) \geq \deg(g)$. Wir führen die Induktion über $n = \deg(f)$.

Ohne Beschränkung der Allgemeinheit können wir annehmen, daß die höchsten Koeffizienten von f und g gleich 1 sind (solche Polynome heißen „normiert“).

Sei also $n = 1$, d.h. $f(z) = z + a$. Dann ist $g(z) = z + b$ oder $g(z) = 1$, im ersten Fall wählen wir $q(z) = 1, r(z) = a - b$ und im zweiten Fall $q(z) = f(z), r(z) = 0$.

Wir setzen nun voraus, daß den Satz für Polynome von einem Grad, der kleiner als n ist, bewiesen ist. Sei also

$$f(z) = z^n + a_1 z^{n-1} + \dots, \quad g(z) = z^m + b_1 z^{m-1} + \dots,$$

dann setzen wir $q_1(z) = z^{n-m}$, es ist

$$q_1(z)g(z) = z^n + b_1 z^{n-1} + \dots$$

und das Polynom

$$f_1(z) = f(z) - q_1(z)g(z) = (a_1 - b_1)z^{n-1} + \dots$$

hat einen Grad, der kleiner als n ist, also gibt es Polynome $q_2(z)$ und $r(z)$ mit $f = q_2 g + r$ und wir wissen, daß $r = 0$ oder $\deg(r) < \deg(g)$ gilt. Dann haben wir mit

$$f = f_1 + q_1 g = (q_2 + q_1)g + r$$

die gewünschte Zerlegung gefunden.

Wir zeigen noch die Einzigkeit von $q(x)$ und $r(x)$. Sei etwa noch

$$f(x) = s(x) \cdot g(x) + t(x), \quad \deg(t) < \deg(g),$$

dann ist

$$(q(x) - s(x)) \cdot g(x) = t(x) - r(x).$$

Wenn $q(x) \neq s(x)$ wäre, stünde links ein Polynom, dessen Grad mindetens gleich $\deg(g)$ ist, und rechts ein Polynom, dessen Grad kleiner als $\deg(g)$ ist.

Folgerung 10.9.1 Sei $a \in \mathbb{R}$ eine Nullstelle von $f(x)$, dann ist das lineare Polynom $x - a$ ein Teiler von $f(x)$.

Beweis: Wir teilen $f(x)$ durch $x - a$:

$$f(x) = q(x) \cdot (x - a) + r,$$

der Grad von r ist kleiner als 1, also ist r eine Konstante. Wir setzen $x = a$ ein und finden $f(a) = r$, also ist $r = 0$. Das heißt, $f(x)$ ist durch $x - a$ teilbar.

Lemma 10.9.1 *Ein Polynom $f(x)$ besitzt genau dann eine mehrfache Nullstelle r , wenn $f'(r) = 0$ ist.*

Beweis: Sei $f(x) = (x - r)^k \cdot g(x)$, $k > 1$. Dann ist

$$f'(x) = k \cdot (x - r)^{k-1} \cdot g(x) + (x - r)^k \cdot g'(x),$$

also ist r eine gemeinsame Nullstelle von f und f' .

Folgerung 10.9.2 *Sei $g = ggT(f, f')$; dann besitzt f/g dieselben Nullstellen wie f , aber jede Nullstelle ist einfach.*

Wir wollen uns nun mit einem klassischen Verfahren der näherungsweise Nullstellenberechnung befassen.

Das einfachste Näherungsverfahren ist die Newton-Interpolation: Wenn für eine reelle Zahl a der Wert von $f(a)$ beinahe Null ist, so ist $a - f(a)/f'(a)$ eine bessere Näherung. Eine Voraussetzung dafür ist aber, daß f'' in der Nähe des Startpunkts der Iteration sein Vorzeichen nicht ändert. Man muß also erst einmal in die Nähe einer Nullstelle geraten.

Als erstes kann man ein Intervall angeben, in dem alle Nullstellen liegen müssen.

Satz 10.9.2 (Cauchysche Ungleichung)

Sei $f(x) = \sum a_i x^{n-i} \in \mathbb{C}[x]$, dann gilt für jede Nullstelle z von $f(x)$ die Ungleichung

$$|z| < 1 + \frac{\max(|a_1|, \dots, |a_n|)}{|a_0|}.$$

Beweis: Sei $h = \max(|a_1|, \dots, |a_n|)$ und $f(z) = 0$, dann ist

$$a_0 z^n = -a_1 z^{n-1} - \dots - a_n,$$

$$|a_0| |z|^n \leq h \cdot (|z|^{n-1} + \dots + 1) < \frac{h \cdot |z|^n}{|z| - 1},$$

also $|a_0| \cdot (|z| - 1) < h$.

Als nächstes behandeln wir ein Verfahren, das es gestattet, die Anzahl der Nullstellen eines Polynoms in einem gegebenen Intervall zu berechnen. Durch fortgesetzte Intervallteilung kann man jede einzelne Nullstelle dann beliebig genau einschachteln.

Der Sturmsche Lehrsatz

Sei $f(x) \in \mathbb{R}[x]$ und $a \in \mathbb{R}$. Wie oben haben wir

$$f(x) = (x - a)q(x) + f(a),$$

also

$$q(x) = \frac{f(x) - f(a)}{x - a}.$$

Wenn a eine einfache Nullstelle von $f(x)$ ist, so ist

$$q(a) = f'(a) \neq 0$$

und dies gilt in einer Umgebung von a . Es gibt zwei Fälle:

- 1) $f'(a) > 0$, dann ist $f(x)$ bei a monoton wachsend; wenn x von links durch a hindurch läuft, ist $f(x)$ zuerst negativ, dann positiv.
- 2) $f'(a) < 0$, dann ist $f(x)$ bei a monoton fallend; wenn x von links durch a hindurch läuft, ist $f(x)$ zuerst positiv, dann negativ.

Beiden Fällen ist folgendes gemeinsam: Wenn x wachsend durch a läuft, gehen die Funktionen $f(x)$ und $f'(x)$ von verschiedenen zu gleichen Vorzeichen über.

Wir konstruieren nun eine „Sturmsche Kette“, wir setzen $f_1(x) = f'(x)$ und dividieren fortlaufend mit Rest:

$$\begin{aligned} f &= q_1 \cdot f_1 - f_2 \\ f_1 &= q_2 \cdot f_2 - f_3 \\ &\dots \\ f_{i-1} &= q_i \cdot f_i - f_{i+1} \\ &\dots \\ f_m &= c \text{ konstant.} \end{aligned}$$

Beachten Sie das Minuszeichen.

Wenn $f(x)$ nur einfache Nullstellen besitzt, ist $\text{ggT}(f, f_1) = 1$, also $f_m \neq 0$.

Nun gilt: Für keine Zahl $r \in \mathbb{R}$ gibt es ein i mit $f_i(r) = 0 = f_{i+1}(r)$, denn sonst wäre $f(r) = 0 = f'(r)$, also r eine mehrfache Nullstelle. Sei r eine reelle Zahl, sei $w(r)$ die Zahl der Indizes i mit $\text{sgn}(f_i(r)) \neq \text{sgn}(f_{i+1}(r))$, dies ist die Zahl der Vorzeichenwechsel in der Sturmschen Kette $f(r), f_1(r), \dots, f_m(r)$.

Satz 10.9.3 Die Zahl der im Intervall (a, b) gelegenen Nullstellen von $f(x)$ ist gleich $w(a) - w(b)$.

Beweis: Wenn r die x -Achse entlangläuft, so kann sich $w(r)$ nur dann ändern, wenn für ein i galt $f_i(r) = 0$. Dann ist aber

$$f_{i-1}(r) = -f_{i+1}(r).$$

Schauen wir uns die Werte von f_{i-1}, f_i und f_{i+1} in der Nähe von r an:

	f_{i-1}	f_i	f_{i+1}
$r - \epsilon$	$+t$	p	$-t$
r	$+t$	0	$-t$
$r + \epsilon$	$+t$	$-p$	$-t$

Dabei sei $p, t \in \{\pm 1\}$; egal, welchen Wert p hat, p oder $-p$ stimmt mit einem ihrer Nachbarn überein. Beim Durchgang durch r findet also hier gar keine Änderung der Wechselzahl statt. Eine Änderung der Wechselzahl sich kann also nur beim Durchgang durch eine Nullstelle von f ereignen, und oben haben wir gesehen, das dort ein Vorzeichenwechsel zwischen f und f' verlorenght.

Schauen wir uns das in einem Beispiel an:

$$\begin{aligned} f(x) &= x^5 - 4x - 2 \\ f_1(x) &= 5x^4 - 4 \\ x^5 - 4x - 2 &= (5x^4 - 4) \cdot x/5 - 16/5x + 2, \\ \text{wir setzen } f_2(x) &= 8x + 5 \\ f_3(x) &> 0 \end{aligned}$$

x	f	f_1	f_2	f_3	$w(x)$
-2	-	+	-	+	3
-1	+	+	-	+	2
0	-	-	+	+	1
1	-	+	+	+	1
2	+	+	+	+	0

Also liegen zwischen -2 und -1, zwischen -1 und 0 sowie zwischen 1 und 2 je eine Nullstelle.

Die folgende Konstruktion erlaubt es festzustellen, ob zwei Polynome gemeinsame Nullstellen besitzen.

Resultante und Diskriminante

Seien zwei Polynome

$$\begin{aligned} f(x) &= a_0x^m + a_1x^{m-1} + \dots + a_m \\ g(x) &= b_0x^n + b_1x^{n-1} + \dots + b_n \end{aligned}$$

gegeben.

Satz 10.9.4 *Die Polynome $f(x)$ und $g(x)$ haben genau dann einen nicht konstanten größten gemeinsamen Teiler, wenn es von Null verschiedene Polynome*

$$\begin{aligned} h(x) &= c_0x^{m-1} + c_1x^{m-2} + \dots + c_{m-1} \\ k(x) &= d_0x^{n-1} + d_1x^{n-2} + \dots + d_{n-1} \end{aligned}$$

so daß

$$k(x) \cdot f(x) = h(x) \cdot g(x).$$

(Es ist $\deg(h) < \deg(f)$ und $\deg(k) < \deg(g)$).

Beweis: Sei $k \cdot f = h \cdot g$, dann können nicht alle Teiler von f in h aufgehen, da der Grad von h zu klein ist, also muß f einen gemeinsamen Teiler mit g besitzen.

Sei umgekehrt $t(x)$ ein gemeinsamer Teiler von $f(x)$ und $g(x)$. Dann gibt es Polynome $h(x)$ und $k(x)$ mit $f = t \cdot h$, $g = t \cdot k$, also

$$k \cdot f = k \cdot t \cdot h = g \cdot h.$$

□

10.10 Gleichungen dritten und vierten Grades

Nun sollen explizite Formeln für die Nullstellen von Polynomen mit reellen Koeffizienten entwickelt werde, soweit dies möglich ist.

Sei

$$f(y) = y^3 + ay^2 + by + c$$

ein Polynom dritten Grades, dessen Nullstellen wir suchen. Wenn wir $y = x + m$ setzen, so hat der Koeffizient von x^2 in $f(y)$ den Wert $3m + a$, er verschwindet für $m = -\frac{a}{3}$. Somit können wir annehmen, daß die Gleichung in der Form

$$x^3 + px + q = 0$$

zu lösen ist. Wir leiten zunächst die nach Geronimo Cardano (1501-1576) benannten und von Niccolo Targaglia (1500-1557) gefundenen Cardanische Formel her.

Wir machen den Ansatz $x = u + v$ und erhalten durch Einsetzen von x^3 in die obige Gleichung

$$u^3 + v^3 + 3uv(u + v) + p(u + v) + q = 0.$$

Wir suchen solche Werte u, v , daß

$$u^3 + v^3 + q = 0$$

und

$$3uv + p = 0$$

gilt, dann wäre die Gleichung gelöst. Dies ist der Fall, wenn

$$u^3 + v^3 = -1$$

und

$$u^3 v^3 = -\frac{p^3}{27}$$

gelten. Nach der Formel von Viètà heißt das, daß die Werte u^3 und v^3 die Lösungen der Gleichung

$$z^2 + qz - \frac{p^3}{27} = 0$$

sind. Für quadratische Gleichungen kennen wir schon eine Lösungsformel, also gilt

$$u^3 = -\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}$$

$$v^3 = -\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}$$

Sei $\epsilon \in \mathbf{C}$ eine primitive dritte Einheitswurzel, dann haben u^3 und v^3 drei verschiedene dritte Wurzeln, die sich jeweils um einen Faktor ϵ unterscheiden, dies ergäbe neun

mögliche Werte für x . Beachten wir aber (s.o.), daß uv reell sein soll, so erhalten wir als folgende drei Lösungen der gegebenen Gleichung:

$$\begin{aligned}x_1 &= {}^3\sqrt{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + {}^3\sqrt{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} \\x_2 &= \epsilon {}^3\sqrt{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \epsilon^2 {}^3\sqrt{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} \\x_3 &= \epsilon^2 {}^3\sqrt{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \epsilon {}^3\sqrt{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}\end{aligned}$$

Dies also sind die Cardanischen Formeln. Sie haben allerdings ihre Tücken, wie den Mathematikern des 16. Jahrhunderts schmerzlich bewußt wurde. Betrachten wir ein Beispiel:

$$x^3 - 981x - 11340 = 0,$$

wir erhalten

$$x_1 = {}^3\sqrt{-5670 + \sqrt{32148900 - 34965783}} + \dots$$

hier wären also dritte Wurzeln aus (echt) komplexen Zahlen zu ziehen. Die „alten Herren“ haben vielleicht weniger die imaginären Terme gestört, seit jeher haben Mathematiker mit Dingen operiert, die sie nicht verstanden haben. Aber: die drei Lösungen dieser Gleichung sind reell, wie wir noch sehen werden, sie lassen sich aber nicht durch „normale“ Rechenoperationen (Addition, Multiplikation, Wurzelziehen) aus den Koeffizienten der Gleichung berechnen. Eine dritte Wurzel aus einer beliebigen komplexen Zahl durch „Radikale“ darstellen zu können, würde bedeuten, daß man einen Winkel mit Zirkel und Lineal in drei gleiche Teile teilen könnte; die Galois-Theorie lehrt, daß dies unmöglich ist. Um diesem Problem wenigstens einen Namen zu geben, nannte man diesen Fall „casus irreducibilis“.

Mit algebraischen Methoden ist dieses Problem prinzipiell nicht lösbar, hier zeigen sich auch Grenzen der Computeralgebra.

Jedoch hat Viètà um 1600 eine „transzendente“ Lösung gefunden:

Wir haben

$$x^3 + px + q = 0$$

und die Zahl p ist notwendigerweise negativ. Sei

$$u^3 = -\frac{q}{2} + i\sqrt{-\frac{q^2}{4} - \frac{p^3}{27}} = r(\cos \alpha + i \sin \alpha),$$

dann errechnet man $r = \sqrt{-\frac{p^3}{27}}$ und $\cos \alpha = -\frac{q}{2r}$ also

$$\begin{aligned}x_1 &= r^{1/3} \left(\cos \frac{\alpha}{3} + i \sin \frac{\alpha}{3} \right) + r^{1/3} \left(\cos \frac{\alpha}{3} - i \sin \frac{\alpha}{3} \right) \\&= 2\sqrt{-\frac{p}{3}} \cos \frac{\alpha}{3}\end{aligned}$$

$$x_2 = 2\sqrt{\frac{-p}{3}} \cos \frac{\alpha - \pi}{3}$$

$$x_3 = 2\sqrt{\frac{-p}{3}} \cos \frac{\alpha + \pi}{3}$$

Im oben angeführten Beispiel gilt

$$r \approx 5913,1872$$

$$\cos \alpha \approx 0,9588$$

$$\alpha \approx 16,5$$

$$\cos(5,5) \approx 0,9954$$

$$x_1 \approx 35,99$$

$$x_1 \approx -21$$

$$x_1 \approx -15$$

Wer nachrechnen möchte, wird sehen, daß die Lösungen ganzzahlig sind, sich aber nicht „einfacher“ aus den Koeffizienten berechnen lassen.

Als nächstes betrachten wir Gleichungen 4. Grades, diese möge bereits in die Form

$$x^4 + px^2 + qx + r = 0$$

gebracht worden sein.

Wir machen wieder einen Ansatz $x = u + v + w$ und führen folgende Rechnungen durch:

$$\begin{aligned} x^2 &= u^2 + v^2 + w^2 + 2(uv + uw + vw), \\ x^2 - u^2 + v^2 + w^2 &= 2(uv + uw + vw), \\ x^4 - 2x^2(u^2 + v^2 + w^2) + (u^2 + v^2 + w^2)^2 \\ &= 4(u^2v^2 + u^2w^2 + v^2w^2) + 8(u^2vw + v^2uw + w^2uv)^2 \\ &= 4(u^2v^2 + u^2w^2 + v^2w^2) + 8uvw, \end{aligned}$$

nun setzen wir x^4 in die obige Gleichung ein:

$$2x^2(u^2 + v^2 + w^2) - (u^2 + v^2 + w^2)^2 + 4(u^2v^2 + u^2w^2 + v^2w^2) + 8(u^2vw + v^2uw + w^2uv)^2 + 4(u^2v^2 + u^2w^2 + v^2w^2) + 8uvw + px^2 + qx + r = 0.$$

Nun wählen wir u, v, w so, daß die Koeffizienten von x^2, x und 1 Null werden:

$$u^2 + v^2 + w^2 = -\frac{p}{2},$$

$$8uvw = -q,$$

daraus folgt

$$u^2v^2w^2 = \frac{q^2}{64},$$

und

$$u^2v^2 + u^2w^2 + v^2w^2 = \frac{p^2}{16} - \frac{r}{4},$$

d.h. die Zahlen u^2 , v^2 , w^2 sind die Nullstellen des Polynoms

$$x^3 + \frac{p}{2}x^2 + \left(\frac{p^2}{64} - \frac{r}{4}\right)x - \frac{q^2}{64}.$$

Diese seien gleich z_1 , z_2 , z_3 , dann erhalten wir

$$u = \pm\sqrt{z_1}, v = \pm\sqrt{z_2}, w = \pm\sqrt{z_3},$$

wegen $8uvw = -q$ legen die Vorzeichen von u und v bereits das von w fest, es gibt also vier Lösungen.

Man könnte versuchen, nach demselben Verfahren Gleichungen fünften Grades zu lösen, ich habe es nicht probiert, weil ich weiß, das es nicht geht, dies ist ein Resultat der Galois-Theorie. Hier bei dem konkreten Ansatz $x = t + u + v + w$ wird man wahrscheinlich wieder auf eine Gleichung für Terme, die aus t, u, v, w gebildet werden, geführt, und diese wird mindestens den Grad 5 haben.

Index

- äquivalente Matrizen, 144
- affine Abbildung, 62
- affiner Raum, 57
- affiner Unterraum, 58
- allgemeine Lage, 58
- Annulator, 70
- ausgezeichnete Spalten, 19

- Basis, 30
- Bild, 133
- Bilinearformen, 73

- charakteristisches Polynom, 106

- Determinante, 90
- Determinantenteiler, 146
- Dimension, 32
- Dimensionssatz, 33
- direkte Summe, 34
- duale Abbildung, 70

- Eigenvektor, 106
- Eigenwert, 106
- elementare Operationen, 16
- Erzeugendensystem, 27

- Gaußscher Algorithmus, 19
- Gleichungssysteme, 13
- Gruppe, 127
- Gruppenhomomorphismus, 132

- Hauptideal, 156
- Hom, 41
- Homomorphiesatz, 133

- Ideal, 156
- idempotent, 54
- Invariantenteiler, 146
- inverse Abbildung, 42

- inverse Matrix, 48
- involutiv, 54
- irreduzibles Polynom, 156

- Körper, 11
- Kern, 43, 132
- Koordinaten, 30
- Koordinatensystem, 58

- Laplacescher Entwicklungssatz, 91
- Leibnizsche Determinantendefinition, 92
- linear unabhängig, 29
- lineare Abbildung, 39
- lineare Hülle, 26
- lineares Gleichungssystem, 14
- Linearkombination, 15
- LU-Zerlegung, 53

- Matrix, 18
- Matrixprodukt, 47
- maximale linear unabhängige Menge, 29
- minimales Erzeugendensystem, 28
- Minor, 96, 107
- Multilinearform, 90

- nichtausgeartete Bilinearform, 74
- nilpotent, 54
- normale Untergruppe, 133
- Normalteiler, 133

- Ordnung einer Gruppe, 131

- parallel, 61
- Permutation, 91
- primitives Polynom, 156
- Primzahlzerlegung, 126

- Quadrik, 82

- Rang, 38

Ring, 152

Satz von Hamilton-Cayley, 111

Satz von Kronecker/Capelli, 38

Spaltenrang, 37

Spaltenraum, 37

Summe von Unterräumen, 33

symmetrische Bilinearform, 75

Torsionsuntergruppe, 143

Transposition, 137

Untergruppe, 129

Unterraum, 27

Vektoren, 25

Vektorraum, 25

Verbindungsraum, 62

windschief, 61

Zeilenrang, 35

Zeilenraum, 35

Zyklus, 137