Algebra und Geometrie

Vorlesungsskript,
http://www-irm.mathematik.hu-berlin/∼ hgrass/

Hubert Grassmann

28. Juni 2000

Inhaltsverzeichnis

0 Einführung **7**
0.1 Aufgaben 8

1 Lineare Gleichungssysteme **11**
1.1 Grundlagen 11
1.2 Eigenschaften von Gleichungssystemen 13
1.3 Elementare Operationen 14
1.4 Gaußscher Algorithmus 17
1.5 Aufgaben 19

2 Grundbegriffe der Theorie der Vektorräume **23**
2.1 Vektorräume, Unterräume, lineare Hüllen 23
2.2 Lineare Unabhängigkeit, Basen, Dimension 26
2.3 Anwendung auf lineare Gleichungssysteme 33
2.4 Aufgaben 36

3 Lineare Abbildungen und Matrizen **41**
3.1 Grundlegende Eigenschaften 41
3.2 Darstellungsmatrizen 46
3.3 Matrixmultiplikation, Inverse von Matrizen 48
3.4 Basiswechsel 52
3.5 Idempotente Abbildungen und direkte Summen 55
3.6 Aufgaben 57

4 Affine Geometrie **63**
4.1 Affine Räume und Unterräume 63
4.2 Affine Abbildungen 68
4.3 Aufgaben 71

5 Linearformen **77**

6 Bilinearformen **81**
6.1 Darstellungsmatrizen und Basiswechsel, Diagonalisierung 81
6.2 Jacobi-Diagonalisierung 86
6.3 Strassens schnelle Matrixmultiplikation 88
6.4 Klassifikation der Quadriken 89

7 Determinanten **93**
7.1 Existenz und Eindeutigkeit 93
7.2 Eigenschaften und Anwendungen 97
7.3 Aufgaben 102

8 Dreidimensionale Geometrie **103**

9 Eigenwerte und Eigenvektoren **109**
9.1 Aufgaben 117

10 Polynome **119**

11 Normalformen von Matrizen **127**
11.1 Invariante Unterräume 127
11.2 Nilpotente Endomorphismen 129
11.3 Jordansche Normalform 132
11.4 Rekursive Folgen 134
11.5 Lineare Differentialgleichungssysteme 136

12 Euklidische Vektorräume **139**
12.1 Skalarprodukt, Orthonormalbases 139
12.2 Orthogonale Abbildungen und Matrizen 141
12.3 Die adjungierte Abbildung 144
12.4 Pseudoinverse Matrizen 148
12.5 Unlösbare und unterbestimmte Gleichungssysteme 151
12.6 Householder-Transformationen 152
12.7 QR-Zerlegung 153
12.8 Hessenberg-Zerlegung 154
12.9 Singularwertzerlegung 157
12.10 Vektor- und Matrixnormen 158
12.11 Positiv definite Matrizen 160
12.12 Aufgaben 162

13 Euklidische und projektive Geometrie **167**
13.1 Euklidische Geometrie 167
13.2 Sphärische Geometrie 172
13.3 Konvexe Mengen und lineare Ungleichungssysteme 174
13.4 Projektive Geometrie 178

14 Polynommatrizen **185**
14.1 Smithsche Normalform 187
14.2 Die rationale Normalform 192
14.3 Lokale Minimalpolynome eines Endomorphismus 193

15 Elementare Gruppentheorie **197**
15.1 Der Ring $\mathbb{Z}$ der ganzen Zahlen . . . 197
15.2 Gruppen, Untergruppen, Homomorphismen . . . 199
15.3 Die symmetrischen Gruppen . . . 206
15.4 Endlich erzeugte abelsche Gruppen . . . 207
15.5 Gruppenoperationen . . . 210
15.6 Aufgaben . . . 215

16 Ringe und Moduln **217**
16.1 Grundbegriffe . . . 217
16.2 Universelle Konstruktionen; abstract nonsens . . . 223
16.3 Tensorprodukte . . . 228
16.4 Das Jacobson-Radikal . . . 232

17 Halbeinfache Algebren und Moduln **239**
17.1 Grundlagen . . . 239
17.2 Darstellungen endlicher Gruppen . . . 246
17.3 Charaktere . . . 251
17.4 Die diskrete Fourier-Transformation . . . 256

18 Zerlegung endlichdimensionaler Algebren **259**

19 Boolesche Algebren und Boolesche Funktionen **267**

Index **271**

Vorwort

Dies ist eine Ausarbeitung einer Anfängervorlesung zur linearen Algebra, zur analytischen Geometrie sowie die Ring- und Modultheorie, 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 füehrt 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 fhrt 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).

Der zweite Teil führt in die Elemente der Algebra ein. Das Kapitel 16 behandelt Elemente der Gruppentheorie, im Kapitel 17 wird auf Ringe, Ideale und Moduln eingegangen. Es werden Tensorprodukte eingeführt und das Jacobson-Radikal eines Rings wird mit ringtheoretisch charakterisiert. Etwas tiefergehende Betrachtungen über halbeinfache Algebren über Körpern und deren Moduln, eine Einführung in die Darstellungstheorie endlicher Gruppen (Kapitel 18) sowie über die Zerlegung beliebiger (endlichdimensionaler) Algebren (Kapitel 19) fügen sich an, dazu gehört die modultheoretische Charakterisierung des Jacobson-Radikals.

Kapitel 0

Einführung

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:

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

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 $1 \ldots 4$ 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 $1 \ldots 7$ 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:
A und B seien Mengen, dann bezeichnet $A \cup B$ die Vereinigung und $A \cap B$ den Durchschnitt von A und B, die Relation $A \subseteq B$ bedeutet, daß A eine Teilmenge von B ist. Mit $f : A \to B$ bezeichnen wir eine Abbildung f einer Menge A in eine Menge B, und wenn $C \subseteq A$ eine Teilmenge ist, so bezeichnet $f \mid C$ die Einschränkung der Abbildung f auf die Teilmenge C. Das Zeichen $\square$ zeigt das Ende eines Beweises an.

0.1 Aufgaben

1. Man gebe in $\mathbf{C} \cong \mathbf{R}^2$ die Eckpunkte eines gleichseitigen Dreiecks an welches $(1 + i)$ als Mittelpunkt, sowie $(3 + 2i)$ als einen Eckpunkt besitzt!

2. Zeigen Sie, daß man in der 4-elementigen Menge $K = \{0, 1, a, b\}$ zwei Operationen $+$ und $\cdot$ so einführen kann, daß K ein Körper der Charakteristik 2 wird. Sind die Operationen eindeutig bestimmt?

3. In der Menge der Paare rationaler Zahlen seien Operationen $\oplus$ und $\odot$ definiert durch: $(a, b) \oplus (c, d) := (a + c, b + d)$, $(a, b) \odot (c, d) := (a \cdot c + 2b \cdot d, a \cdot d + b \cdot c)$ Erhält man einen Körper? Verändert sich die Antwort, wenn man statt dessen Paare reeller Zahlen betrachtet?

4. Beweisen Sie: Im Körper der komplexen Zahlen kann keine Ordnungsrelation $\leq$ eingeführt werden, die gleichzeitig folgende Bedingungen erfüllt:

   1. Für alle $z_1, z_2 \in \mathbf{C}$ gilt $z_1 \leq z_2$ oder $z_2 \leq z_1$;
   2. Aus $z_1 \leq z_2$ folgt $z_1 + z_3 \leq z_2 + z_3$ für beliebige $z_3 \in \mathbf{C}$)
   3. Falls $z_1 \leq z_2$ und $0 \leq z_3 für z_1, z_2, z_3 \in \mathbf{C}$ gilt , so folgt $z_1 \cdot z_3 \leq z_2 \cdot z_3$.

5. Man bestimme alle Körperautomorphismen $\varphi : \mathbf{K} \to \mathbf{K}$ für folgende Körper $\mathbf{K}$:

   a) $\mathbf{K} = \mathbf{Q}$,

   b) $\mathbf{K} = \mathbf{R}$,

   c) $\mathbf{K} = \mathbf{C}$, wobei $\varphi(\mathbf{R}) \subseteq \mathbf{R}$ gelten möge.

Kapitel 1

Lineare Gleichungssysteme

1.1 Grundlagen

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

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

($a, \ldots, f$ 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,

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

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). Die Menge aller n-tupel

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

bezeichnen wir mit R^n.

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

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

ein lineares Gleichungssystem mit m Gleichungen und den n Unbekannten $x_1, \ldots, x_n$. Ein n-tupel $(x_1, \ldots, x_n)$, 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

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

nennen wir das homogene Gleichungssystem

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

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

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

und

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

Sei

$$\sum a_{ij} x_j = 0, i = 1, \ \ldots, m \tag{H}$$

ein homogenes Gleichungssystem. Dann gilt:

1. Es existiert stets eine Lösung von H, nämlich die triviale Lösung $(0, \ldots, 0)$.

2. Wenn $x = (x_1, \ldots, x_n)$ eine Lösung von H und r eine Zahl ist, so ist auch das Vielfache $rx = (rx_1, \ldots, rx_n)$ eine Lösung von H.

3. Wenn $x = (x_1, \ldots, x_n)$ und $y = (y_1, \ldots, y_n)$ Lösungen von H sind, so ist auch die Summe $x + y = (x_1 + y_1, \ldots, x_n + y_n)$ eine Lösung von H.

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

   Dann erhalten wir sofort

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

   Sei nun wieder

   $$\sum a_{ij} x_j = b_i, i = 1, \ldots, m \tag{S}$$

   ein inhomogenes System und

   $$\sum a_{ij} x_j = 0, i = 1, \ldots, m \tag{H}$$

   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.

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

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

   und

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

Durch Subtraktion dieser Zahlen erhalten wir

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

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

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:

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

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

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

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

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

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

(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:

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

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

Dann gilt der folgende

Satz 1.1 *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 $x = (x_1, \ldots, x_n) \in LM(S)$, dann gilt

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

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 $x = (x_1, \ldots, x_n) \in LM(S)$, also

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

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. □

Folgerung 1.2 *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.* □

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):

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

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

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

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

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

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:

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

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

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

gegeben, dazu gehört die Matrix

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

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} enthält, und bringen die i-te Zeile durch Zeilenoperationen 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 $1., 2., \ldots$ 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 $k_1 \ldots k_r$):

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

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

$$\begin{array}{rcl} x_{k_1} + a_{1,k_1+1}x_{k_1+1} + \quad \ldots \quad + a_{1n}x_n & = & b_1 \\ & \ldots & \\ x_{k_r} + a_{r,k_r+1}x_{k_r+1} + \ldots + a_{rn}x_n & = & b_r \\ 0 & = & b_{r+1} \\ & \ldots & \\ 0 & = & b_m \end{array}$$

Nun können wir die Lösungsmengen $LM(S) = LM(S')$ ablesen:
Wenn die Zahlen $b_{r+1}, \ldots, b_m$ von Null verschieden sind, so existiert keine Lösung, andernfalls existiert eine Lösung, was wir wie folgt einsehen:
Die ausgezeichneten Spalten entsprechen ausgezeichneten Unbekannten $x_{k_1}, \ldots, x_{k_r}$, für die restlichen (nicht-ausgezeichneten) Unbekannten wählen wir beliebige Werte

$$x_i = t_i,\ i = 1, \ldots, n,\ i \neq k_l,\ l = 1, \ldots, r.$$

Dann ist jedes n-tupel $(x_1, \ldots, x_n)$ mit

$$x_i = b_i - \sum_{j \neq k_l} a_{ij} t_j$$

eine Lösung von S.

Wir führen dies an einem Beispiel aus: Wir haben ein Gleichungssystem

$$\begin{array}{rcrcrcrcr} x_1 & + & x_2 & & & + & 3x_4 & = & 2 \\ 2x_1 & + & 2x_2 & + & x_3 & + & 7x_4 & = & 6 \\ -3x_1 & - & 3x_2 & - & x_3 & - & 10x_4 & = & -8 \end{array}$$

Dazu gehört die Koeefizentenmatrix

$$\begin{pmatrix} 1 & 1 & 0 & 3 & 2 \\ 2 & 2 & 1 & 7 & 6 \\ -3 & -3 & -1 & -10 & -8 \end{pmatrix},$$

deren reduzierte Form ist

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

also sind x_1 und x_3 ausgezeichnete Unbekannte, $x_2 = t_1$ und $x_4 = t_2$ können beliebig gewählt werden, also

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 2 - 3t_2 - t_1 \\ t_1 \\ 2 - t_2 \\ t_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 2 \\ 0 \end{pmatrix} + t_1 \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t_2 \begin{pmatrix} 3 \\ 0 \\ -1 \\ 1 \end{pmatrix}.$$

Nun beweisen wir den folgenden

Satz 1.3 *Sei* $\sum a_{ij}x_j = 0, i = 1, \ldots, m$, *ein homogenes Gleichungssystem, für das* $n > m$ *gilt (es gibt also mehr Unbekannte als Gleichungen) dann existiert eine Lösung* $(x_1, \ldots, x_n) \neq (0, \ldots, 0)$.

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

$$\begin{bmatrix} 1 & & & \\ & 1 & & \\ & \cdots & & \\ & & 1 & \end{bmatrix}$$

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. □

1.5 Aufgaben

Lösen Sie folgende Gleichungsysteme! (Aufg. 1 - 7)

1. $x_1 + x_2 - x_3 = 0$
   $3x_1 + 2x_2 + x_3 = 5$
   $4x_1 - x_2 + 5x_3 = 3$

2. $x_1 - 2x_2 + 2x_3 = -5$
   $2x_1 + x_2 - x_3 = 5$
   $7x_1 + x_2 - x_3 = 10$

3. $2x_1 + 3x_2 + 2x_3 = 9$
   $x_1 + 2x_2 - 3x_3 = 14$
   $3x_1 + 4x_2 + x_3 = 16$

4. $x_1 + 2x_2 + 3x_3 - x_4 = 0$
   $x_1 - x_2 + x_3 + 2x_4 = 4$
   $x_1 + 5x_2 + 5x_3 - 4x_4 = -4$
   $x_1 + 8x_2 + 7x_3 - 7x_4 = -8$

5. $x_1 - x_2 + x_3 - x_4 = -2$
   $x_1 + 2x_2 - 2x_3 - x_4 = -5$
   $2x_1 - x_2 - 3x_3 + 2x_4 = -1$
   $x_1 + 2x_2 + 3x_3 - 6x_4 = -10$

6. $x_1 - 2x_2 + 3x_3 = 4$

   $3x_1 + x_2 - 5x_3 = 5$

   $2x_1 - 3x_2 + 4x_3 = 7$

7. $2x_1 - x_2 - x_3 + 2x_4 = 3$

   $6x_1 - 2x_2 + 3x_3 - x_4 = -3$

   $-4x_1 + 2x_2 + 3x_3 - 2x_4 = -2$

   $2x_1 \qquad + 4x_3 - 3x_4 = -1$

8. Für welche reellen Zahlen c besitzt das folgende Gleichungssystem nicht-triviale Lösungen?

   $3x + 2y + z = 0$

   $x + y + z = 0$

   $2x + y + cz = 0$

9. Für welche Zahlen λ ist das folgende Gleichungssystem nichttrivial lösbar ?

   $x_2 + x_5 = \lambda x_1$

   $x_1 + x_3 = \lambda x_2$

   $x_2 + x_4 = \lambda x_3$

   $x_3 + x_5 = \lambda x_4$

   $x_1 + x_4 = \lambda x_5$ (vgl. auch Aufg. 25 aus Kap. 3)

10. a) Nennen Sie ein Kriterium für die Lösbarkeit eines linearen Gleichungssystems $Ax = b$ mit $A_{m,n} = \begin{bmatrix} a_{11} & \dots & a_{m1} \\ & \dots & \\ a_{1n} & \dots & a_{mn} \end{bmatrix}, x = [x_1 \dots x_n]^T, b = [b_1 \dots b_m]^T$.

    b) Entscheiden Sie, ob das Gleichungssystem lösbar ist! Begründen Sie Ihre Antwort!

    $$\begin{aligned} x + 3y &= 2 \\ x + z &= 0 \\ 3x + 9z &= 4 \end{aligned}$$

    c) Lösen Sie das Gleichungssystem!

    $$\begin{aligned} 2x_1 - x_2 - x_3 + 3x_4 &= 1 \\ 4x_1 - 2x_2 - x_3 + x_4 &= 5 \\ 6x_1 - 3x_2 - x_3 - x_4 &= 9 \\ 2x_1 - x_2 + 2x_3 - 12x_4 &= 10 \end{aligned}$$

    d) Ist es möglich, daß ein homogenes lineares Gleichungssystem genau eine nichttriviale Lösung hat ? (Begründung)

11. a) Lösen Sie folgendes Gleichungssystem!

$$\begin{aligned} 2r_1 \quad\quad + r_3 + 3r_4 &= 0 \\ r_1 \quad\quad + r_3 \quad\quad &= 0 \\ 2r_2 \quad\quad + 4r_4 &= 0 \end{aligned}$$

b) Was können Sie über die Struktur der Lösungsmenge aussagen? Begründen Sie Ihre Antwort!

c) Sei (G): $Ax = b$ ein inhomogenes lineares Gleichungssystem mit Sei $x \in R^n$ eine Lösung von (G) und $y \in R^n$ eine Lösung des zugehörigen homogenen linearen Gleichungssystems. Zeigen Sie, daß $x + y$ eine Lösung von (G) ist.

12. Mit Hilfe des Kriteriums von Kronecker-Capelli untersuche man, für welche Werte von $c \in \mathbf{R}$ das folgende Gleichungs- system lösbar ist; man gebe eine Parameterdarstellung der Lösungsmenge an!

$$\begin{aligned} x_1 - 2x_2 + 3x_3 &= 5 \\ 2x_1 + 3x_2 + x_3 &= 2 \\ 4x_1 + 13x_2 - 3x_3 &= -4 \\ 5x_1 + cx_2 \quad\quad &= 1 \end{aligned}$$

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 $v - w$.

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 $(r_1, \ldots, r_n)$, Addition und Multiplikation sind (wie im Kapitel 1) komponentenweise definiert.
c) V = Menge aller Lösungen des homogenen Gleichungssystems

$$\sum a_{ij}x_j = 0,\ i = 1, \ldots, m, \tag{S}$$

die Addition und Multiplikation sind wie in R^n definiert.
d) Sei V ein Vektorraum. Wenn $v_1, \ldots, v_n \in V$ und $r_1, \ldots, r_n \in R$ sind, so heißt der Vektor $r_1v_1 + \ldots + r_nv_n \in V$ eine Linearkombination der Vektoren $v_1, \ldots, v_n$.

Sei $L(v_1, \ldots, v_n)$ die Menge aller Linearkombinationen von $v_1, \ldots, v_n$, also

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

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

Lemma 2.1 *$L(v_1, \ldots, v_n)$ ist ein Vektorraum (Summe und Produkt sind wie in V definiert).*

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

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

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

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

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

$$o = \sum 0v_i$$

und der zu $\sum r_iv_i$ inverse Vektor auch:

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

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

Definition: Sei V ein Vektorraum und U 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 $L(v_1, \ldots, v_n)$ ein Unterraum von V ist.

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

$$L(M) = \{v \in V \mid \text{ es gibt } v_1, \ldots, v_n \in M \text{ und } r_1, \ldots, r_n \in R \text{ mit } v = r_1v_1 + \ldots + r_nv_n\}.$$

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

Satz 2.2 *Sei V ein Vektorraum und $M \subseteq V$ eine Teilmenge. Dann ist $L(M)$ der kleinste Unterraum von V, der M enthält, d.h. wenn U ein Unterraum von V ist, der M enthält, so ist $L(M)$ in U enthalten.*

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

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

Beispiele:
1. v sei eine Verschiebung der Ebene, dann ist $L(\{v\})$ 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 $L(\{v, w\})$ die Menge aller Verschiebungen der Ebene.
2. $V = R^3, v = (1, 2, 0), w = (2, 1, 0)$, dann ist

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

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

(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:

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

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

Dazu gehört wiederum das Gleichungssystem

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

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

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

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

2.2 Lineare Unabhängigkeit, Basen, Dimension

Sei nun $M = \{v_1, \ldots, v_k\}$ ein Erzeugendensystem des Vektorraums V, also $L(M) = V$. Dann ist auch $L(M \cup N) = V$ 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 $L(M \backslash \{w\}) \neq L(M) = V$.

Welche Erzeugende kann man denn nun weglassen?
Es sei $M = \{v_1, \ldots, v_k\}$. Der Vektor v_k ist überflüssig, wenn $L(M \backslash \{v_k\}) = L(M)$ ist, also wenn $v_k \in L(v_1, \ldots, v_{k-1})$ ist. Dann gibt es also Zahlen $r_1, \ldots, r_{k-1}$ mit

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

bzw.

$$o = r_1 v_1 + \ldots + r_k v_k$$

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 $\{v_1, \ldots, v_k\} \subseteq V$ heißt linear unabhängig, wenn aus

$$r_1 v_1 + \ldots + r_k v_k = o \; (r_i \in R)$$

folgt, daß $r_1 = r_2 = \ldots = r_k = 0$ ist. Nicht linear unabhängige Mengen heißen linear abhängig.

Minimale Erzeugendensysteme werden wie folgt charakterisiert:

Satz 2.3 *Ein Erzeugendensystem M von V ist genau dann minimal, wenn M linear unabhängig ist.*

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

$$r_1 v_1 + \ldots + r_k v_k = o$$

gilt. Es folgt

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

also wäre v_i in M überflüssig, was der Voraussetzung widerspricht.
Nun sei M linear unabhängig. Wäre M nicht minimal, so wäre etwa

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

und damit

$$o = r_1 v_1 + \ldots + r_{k-1} v_{k-1} - 1 v_k.$$

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

Satz 2.4 *Jede Teilmenge einer linear unabhängigen Menge von Vektoren ist linear unabhängig.*

Den Beweis führen wir indirekt: Sei $\{v_1, \ldots v_n\}$ linear abhängig, d.h. es gibt Zahlen $r_1, \ldots, r_n$, unter denen etwa $r_i \neq 0$ ist, so daß $o = r_1 v_1 + \ldots + r_n v_n$.
Wir nehmen weitere Vektoren $v_{n+1}, \ldots, v_m$ hinzu und erhalten die folgende nichttriviale Linearkombination

$$o = r_1 v_1 + \ldots + r_n v_n + 0 v_{n+1} + \ldots + 0 v_m,$$

damit ist auch die größere Menge linear abhängig. □

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 2.5 *Sei $M \subseteq V$ linear unabhängig. M ist genau dann eine maximale linear unabhängige Teilmenge, wenn $L(M) = V$, also wenn M ein minimales Erzeugendensystem ist.*

Beweis: $M = \{v_1, \ldots, v_k\}$ 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:

$$o = r_1 v_1 + \ldots + r_k v_k + r v,$$

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

$$o = r_1 v_1 + \ldots + r_k v_k,$$

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 $L(M) = V$. Sei $w \in V$ beliebig, dann liegt w in $L(M)$, also ist $M \cup \{w\}$ linear abhängig, d.h. M ist maximal. □

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

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

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

dann ist

$$o = (r_1 - s_1) v_1 + \ldots + (r_n - s_n) v_n,$$

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

Die Zahlen $r_1, \ldots, r_k$ 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 $(x_1, \ldots, x_4)$.

Im Vektorraum R^n der n-tupel gibt es eine sehr einfache Basis, die aus den „Einheitsvektoren“
$e_i = (0, \ldots 0, 1, 0, \ldots 0)$
$\uparrow$ i-te Stelle
besteht. Die Bezeichnung „e_i“ wollen wir für diese „kanonische“ Basis des R^n reservieren.

Als nächstes beweisen wir den

Satz 2.7 (Beschränkungssatz) *Seien $v_1, \ldots, v_k \in V$ und $w_1, \ldots, w_m \in L(v_1, \ldots, v_k)$. Wenn $\{w_1, \ldots, w_m\}$ linear unabhängig ist, so ist $m \leq k$.*

Beweis: Wir nehmen an, es gelte $m > k$. Dann betrachten wir eine Linearkombination $w = r_1 w_1 + \ldots + r_m w_m$. Wir fragen uns, ob denn die Zahlen $r_1, \ldots, r_m$ 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

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

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

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

$$\ldots$$

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

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 $(r_1, \ldots, r_m) \neq (0, \ldots, 0)$, das diese Gleichungen erfüllt, für diese Zahlen gilt also

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

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

Folgerung 2.8 *Die Maximalzahl linear unabhängiger Vektoren im R^n ist gleich n.*

Beweis: Wir haben ein Erzeugendensystem aus n Elementen. □

Wir benötigen das einfache

Lemma 2.9 *Wenn $\{u_1, \ldots, u_k\}$ linear unabhängig ist und u_{k+1} kein Element von $L(\{u_1, \ldots, u_k\})$ ist, so ist $\{u_1, \ldots, u_{k+1}\}$ linear unabhängig.*

Beweis: Es sei $r_1 u_1 + \ldots + r_{k+1} u_{k+1} = o$. Wenn $r_{k+1} \neq 0$ wäre, so könnte man durch r_{k+1} dividieren und hätte v_{k+1} als Linearkombination von $v_1, \ldots, v_k$ dargestellt, was nicht möglich ist. Folglich ist $r_{k+1} = 0$ und es bleibt $r_1 u_1 + \ldots + r_k u_k = o$. Wegen der linearen Unabhängigkeit von $\{u_1, \ldots, u_k\}$ ist auch $r_1 = \ldots = r_k = 0$. □

Satz 2.10 *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 $U = L(u_1)$ ist, so sei $B = \{u_1\}$. Andernfalls wählen wir ein $u_2 \in U$, das nicht in $L(u_1)$ liegt. Nach dem Lemma ist $\{u_1, u_2\}$ linear unabhängig.Und so weiter: Sei eine linear unabhängige Teilmenge $\{u_1, \ldots, u_k\}$ schon gefunden. Wenn $U = L\{u_1, \ldots, u_k\}$ ist, so sei $B = \{u_1, \ldots, u_k\}$. Andernfalls wählen wir ein u_{k+1}, das nicht in $L(\{u_1, \ldots, u_k\})$ liegt, dann ist wieder $\{u_1, \ldots, u_{k+1}\}$ linear unabhängig.
Nach höchtens so vielen Schritten, wie das Erzeugendensystem von V Elemente hat, muß das Verfahren abbrechen, d.h. es tritt der Fall $U = L(u_1, \ldots, u_{k+1})$ ein und wir haben eine Basis konstruiert. □

Satz 2.11 *Je zwei endliche Basen eines Vektorraums V besitzen gleichviele Elemente.*

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

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 2.12 (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 $M = \{u_1, \ldots, u_m\}, E = \{v_1, \ldots, v_k\}$. 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 $F = \{v_1, \ldots, v_p\}$ und $M \cup F$ 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 $\{u_1, \ldots, u_m\}$ 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 $M \cup F$ eine Basis von V. □

Als Folgerung erhalten wir den

Satz 2.13 (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. □

Satz 2.14 *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 $B = \{u_1, \ldots, u_m\}$ 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 $V = L(B) = U$. □

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 V liegt nicht notwendigerweise in $U \cup W$).

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

Lemma 2.15 $U + W = \{v \mid$ *es gibt* $u \in U$ *und* $w \in W$ *mit* $v = u + w\}$.

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

Nun folgt der

Satz 2.16 (Dimensionssatz) $\dim(U + W) = \dim U + \dim W - \dim(U \cap W)$

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

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

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

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

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:

$$-\sum t_l w_l = \sum p_i v_i$$

oder

$$\sum p_i v_i + \sum t_l w_l = o,$$

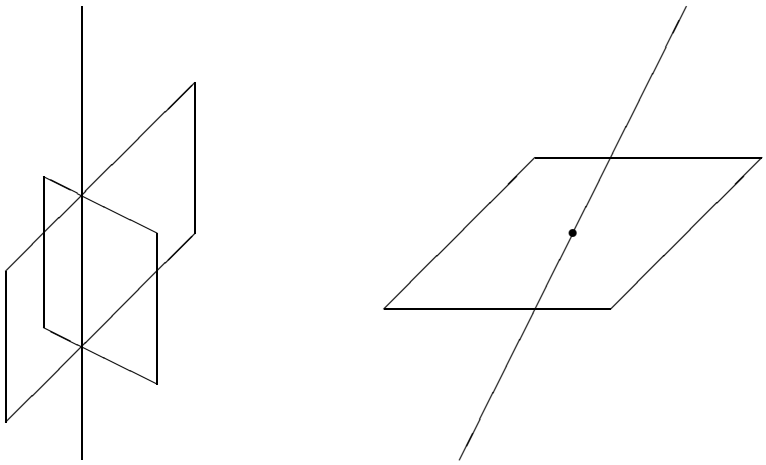
da $B \cup B_2$ linear unabhängig ist, sind alle Koeffizienten gleich Null, also $-\sum t_l w_l = o$, d.h.

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

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 2.17 *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 $U_1, \ldots, U_k$ von V heißt direkt, wenn die Darstellung jedes Vektors $v = \sum u_i$ mit $u_i \in U_i$ eindeutig bestimmt ist.

Satz 2.18 *Die Summe der Unterräume $U_1, \ldots, U_n$ 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 genausoleicht zu zeigen. $\square$

2.3 Anwendung auf lineare Gleichungssysteme

Sei

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

eine Matrix. Wir bezeichnen ihre Zeilen mit

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

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

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

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

Satz 2.19 *Wenn A' durch elementare Zeilenoperationen aus A hervorgeht, so ist $zr(A) = zr(A')$.*

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

Mittels des Gaußschen Algorithmus können wir A in eine reduzierte Form bringen, dabei ändert sich 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

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

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

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

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

Folgerung 2.20 *Sei H ein homogenes Gleichungssystem mit n Unbekannten und der Koeffizientenmatrix A. Dann ist* $\dim LM(H) = n - zr(A)$. □

Sei wieder

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

Wir bezeichnen die Spalten von A mit

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

Diese erzeugen $L(s_1, \dots, s_n) = S(A)$, 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 2.21 *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

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

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

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

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

Beweis: Wir überführen A in die reduzierte Form

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

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

$$rg\left(\begin{bmatrix} a_{11} & \dots & a_{1n} \\ & \dots & \\ a_{m1} & \dots & a_{mn} \end{bmatrix}\right) = rg\left(\begin{bmatrix} a_{11} & \dots & a_{1n} & b_1 \\ & \dots & & \\ a_{m1} & \dots & a_{mn} & b_m \end{bmatrix}\right)$$

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.

2.4 Aufgaben

1. Man zeige, daß in jedem Vektorraum folgendes gilt:

   a) es existiert genau ein Nullvektor, d. h. es gibt genau einen Vektor $o \in V$ mit $a + o = o + a = a (a \in V)$

   b) wenn ein Vektor $a \in V$ die Beziehung $a + a = a$ erfüllt, so gilt $a =$ o.

2. Beweisen Sie, daß in jedem Vektorraum folgendes gilt: $c \cdot a = o$ gdw. $c = 0$ oder $a = o$ $(a \in V, c \in K)$

3. Für einen beliebigen Vektorraum V zeige man:

   a) für jedes $x \in V$ existiert genau ein Vektor $x' \in V$ mit $x + x' = x' + x = o$; dabei gilt $x' = (-1) \cdot x$ (Bezeichnung $-x := (-1) \cdot x$)

   b) für alle $x \in V$ gilt $-(-x) = x$

   c) für alle $x, y \in V$ und aller $\in R : r(x - y) = rx - ry$

4. Man zeige, daß die Kommutativität der Addition aus den übrigen Axiomen eines Vektorraums folgt!

5. Seien U_i $(i = 1, 2, 3, 4)$ Teilmengen des R^4.

   $U_1 = \{(x_1, 0, 0, x_4) \mid x_1, x_4 \in R\}$

   $U_2 = \{(x_1, x_2, 0, 0) \mid x_1, x_2 \in R\}$

   $U_3 = U_1 \cap U_2$

   $U_4 = U_1 + U_2$

   a) Weisen Sie nach, daß U_i $(i = 1, \dots, 4)$ ein Unterraum des R^4 ist.

   b) Geben Sie für jeden dieser Unterräume ein Erzeugendensystem an!

6. Sei V ein reeller Vektorraum. Beweisen Sie:

   a) Für Teilmengen $M \subseteq N \subseteq V$ gilt : $L(M) \subseteq L(N)$.

   b) Sei $M \subseteq V$ ein Erzeugendensystem von V und $N \subseteq V$ mit $M \subseteq L(N)$, dann ist auch N ein Erzeugendensystem von V.

7. Man zeige, daß $M = \{f_1, f_2, f_3\}$ mit $f_1(x) = 2, f_2(x) = x - 1$ und $f_3(x) = x^2 - 2x$ ein Erzeugendensystem von P_2 (Vektorraum der Polynome höchstens 2. Grades) ist!

8. a) Bilden die Vektoren (4, 7, 0, 0); (2, 0, 1, 0); (3, 0, 0, 9); (-27, 0, 0, 0) eine Basis des R^4 ?

   b) Beweisen Sie: Die Vektoren $a_1 = (a_{11}, 0, \ldots, 0)$ $a_2 = (0, a_{22}, 0, \ldots, 0) \ldots a_n = (0, 0, 0, \ldots, a_{nn})$ des R^n sind linear unabhängig gdw. $a_{ii} \neq 0$ für alle i.

9. Im R^5 erzeugen die Vektoren (1, 1, 0, 1, 1); (0, 0, 1, 1, 0); (0, 1, 0, 0, 0); (1, 0, 0, 1, 1); (1, 0, 1, 0, 1) einen Unterraum. Bestimmen Sie die Dimension dieses Unterraums!

10. Für welche $t \in \mathbf{R}$ gilt $w \in L\{v_1, v_2\}$, mit $w = (0, -1, -1)$ und $v_1 = (1{+}2t, 3, 1), v_2 = ({}^1/_2, t, t) \in \mathbf{R}^3$?

11. $$M_1 = \{(1, 0, 0, 0),\ (1, 1, 0, 0),\ (1, 2, 1, 0),\ (-1, 1, 1, 0)\}$$

    und

    $$M_2 = \{(3, 0, 0, 0), (0, -1, 0, 0), (2, 2, 1, 1), (0, 0, 2, 2)\}$$

    sind Vektormengen des R^4. Es sei $U_i = L(M_i)$ $(i = 1, 2)$. Geben Sie die Dimension von $U_1, U_2, U_1 \cap U_2$ und $U_1 + U_2$ an!

12. Man zeige, daß die reellen Zahlen $1, \sqrt{2}, \sqrt{3}$ linear unabhängig über $\mathbf{Q}$ sind!

13. Sei a_n die n-te Wurzel aus $2 (n = 0, 1, 2, \ldots)$, sowie $A = \{a_0, a_1, a_2, a_3, \ldots\}$. Man zeige, daß die lineare Hülle von A über $\mathbf{Q}$ nicht endlich erzeugt ist!

14. Beweisen Sie, daß der Vektorraum aller reeller Zahlenfolgen unendlichdimensional ist!

15. Beweisen Sie, daß $B = (a_1, a_2, a_3)$ mit $a_1 = (2, 2, -1), a_2 = (2, -1, 2)$ und $a_3 = (-1, 2, 2)$ eine Basisdes R^3 ist und berechnen Sie die Koordinaten von $x = (1, 1, 1)$ bzgl. B!

16. Sei $B = (b_1, b_2, b_3)$ eine Basis des R^3 und c ein Vektor mit dem Koordinatentripel (0, -1, 2) bezüglich B.

    a) Beweisen Sie, daß $B^* = (c_1, c_2, c_3)$ mit $c_1 = b_1 + b_2 + b_3, c_2 = b_1 - b_2 - b_3$ und $c_3 = b_3$ ebenfalls eine Basis des R^3 ist!

    b) Berechnen Sie die Koordinaten von c bezüglich B^*!

17. Seien Unterräume $U, W \subseteq \mathbf{R}^4$ wie folgt definiert:

    $W := \{(x_1, x_2, x_3, x_4) : x_1 + x_2 + x_3 + x_4 = 0\}$

    $U := \{x = (x_1, x_2, x_3, x_4) : x$ erfüllt das Gleichungssystem $\}$ mit

    $x_1 + 2x_2 + 3x_3 - x_4 = 0$

    $x_1 - x_2 + x_3 + 2x_4 = 0$

    $x_1 + 5x_2 + 5x_3 - 4x_4 = 0;$

    Man gebe jeweils eine Basis von U, W sowie $U \cap W$ an, und zeige, daß $U + W = \mathbf{R}^4$ gilt!

18. Sei $C(I)$ der Vektorraum der stetigen Funktionen auf dem Intervall $I = [0, 1]$. Man zeige: Die Funktionen $f_1 \equiv 1$, $f_2 = \sin x$, $f_3 = \cos x$ sind linear unabhängige Vektoren.

19. Welche der angegebenen Mengen sind Unterräume von $\mathbf{R}[x]$?

    $\{f \in \mathbf{R}[x] : f(2) = 0\}$

    $\{f \in \mathbf{R}[x] : f(1) \cdot f(0) = 0\}$

    $\{f \in \mathbf{R}[x] :$ grad $(f) = 3\}$(bzw. ≤ 3)

    $\{f \in \mathbf{R}[x] : (x^2 + 1)$ teilt $f\}$.

20. Sei folgende Teilmenge des $\mathbf{R}$ gegeben: $M = \{(1, 0, 0), (2, 1, 1), (0, -1, -1)\}$.

    a) Bilden Sie $L(M)$. Geben Sie eine Basis von $L(M)$ an und begründen Sie, daß es sich um eine Basis handelt.

    b) Sei $U = \{(x, 0, z) \mid x, z \in \mathbf{R}\}$. Bilden Sie $L(M) \cap U$ und geben Sie die Dimension dieses Unterraumes an! Bestimmen Sie die Dimension von $L(M) + U$!

21. a) Beweisen Sie: Der Durchschnitt zweier Unterräume eines Vektorraums V ist ein Unterraum von V.

    b) Gilt eine analoge Aussage auch für die Vereinigung zweier Unterräume? (Begründung!)

22. Ist $B = \{a_1, a_2, a_3, a_4\}$ mit $a_1 = (1, 2, 3, 4), a_2 = (2, 0, 1, 3), a_3 = (2, 0, -1, 4)$ und $a_4 = (0, 0, -2, 1)$ eine Basis des R^4? Begründen Sie Ihre Antwort!

    b) Falls das nicht der Fall ist, geben Sie eine Basis des R^4 an, die möglichst viele der Vektoren von B enthält!

    c) Geben Sie für den Vektor (2,0,1,3) die Koordinatendarstellung bezüglich der von Ihnen konstruierten Basis an!

23. Sei $F = \{(a_n)_{n \in \mathbf{N}} \mid \lim a_n$ existiert $\}$ die Menge der konvergenten, reellen Zahlenfolgen.

    a) Man zeige: Bezüglich der komponentenweisen Addition von Zahlenfolgen und der komponentenweisen Multiplikation einer solchen Folge mit einer reellen Zahl ist F ein reeller Vektorraum.

b) Die Menge $F_o := \{(a_n) \in F; \lim a_n = 0\}$ ist ein Unterraum von F!

c) Geben Sie einen Unterraum $F_1 \subseteq F$ mit $F = F_0 \oplus F_1$ an!

24. Es seien W_1, W_2, W_3 Unterräume eines Vektorraumes V. Finden Sie eine Formel zur Berechnung von $\dim(W_1 + W_2 + W_3) = ?$!

25. Seien $W_1, \ldots, W_k$ die Lösungsräume zu den in Aufgabe 9, Kap.2 ermittelten Zahlen $\lambda_1, \ldots, \lambda_k$. Man bestimme Basen $B_1, \ldots, B_k$ von $W_1, \ldots, W_k$ (als Unterräume des Vektorraumes $\mathbf{R}^5$); weiter zeige man, daß $W_1 \oplus \ldots \oplus W_k = \mathbf{R}^5$ gilt!

26. Sei V ein Vektorraum über K sowie V_1, V_2, V_3 Unterräume von V.

    Beweisen Sie folgende Inklusionen:

    a) $(V_1 \cap V_2) + (V_2 \cap V_3) \subseteq V_2 \cap (V_1 + V_3)$

    b) $V_1 + (V_2 \cap V_3) \subseteq (V_1 + V_2) \cap (V_1 + V_3)$

    Unter der zusätzlichen Voraussetzung $V_1 \subseteq V_2$ zeige man die Gleichheit in a) und b).

27. Sei $C[0,1]$ die Menge der reellwertigen stetigen Funktionen auf [0,1], und $a, b, c \in [0,1]$ seien drei verschiedene Punkte.

    a) Man zeige: Bezüglich der Operationen $(f+g)(x) := f(x) + g(x), x \in [0,1]$,

28. $(f \cdot \lambda)(x) := f(x) \cdot \lambda, x \in [0,1], \lambda \in \mathbf{R}$, ist $C[0,1]$ ein reeller Vektorraum.

    b) Beweisen Sie, daß $V = \{f \in C[0,1] : f(a) = f(b) = f(c) = 0\}$ ein Unterraum von $C[0,1]$ ist!

    c) Man finde Funktionen $u, v, w \in C[0,1]$ (also stetige Funktionen!) derart, daß für $W := L(\{u, v, w\})$ die Beziehung $C[0,1] = V \oplus W$ erfüllt ist!

29. Seien V und W Vektorräume, $W_1 \subseteq W$ ein Unterraum sowie $f : V \to W$ eine lineare Abbildung. Man zeige, daß $f^{-1}(W_1)$ ein Unteraum von V ist!

30. Sei $\mathbf{R}[x]$ die Menge der polynomialen Funktionen über $\mathbf{R}$, und $a \in \mathbf{R}$ ein fester Punkt. Beweisen Sie: $V := \{F \in \mathbf{R}[x] : (x-a) \mid F(x)\}$ ist ein linearer Unterraum von $\mathbf{R}[x]$.

31. Sei $\mathbf{K}$ ein Körper und $\mathbf{K}[x]$ der Vektorraum der Polynome über $\mathbf{K}$.

    a) Man beweise, daß sein Dual $(\mathbf{K}[x])^*$ isomorph ist zum Vektorraum der formalen Potenzreihen $\mathbf{K}[[x]]$, wobei $\mathbf{K}[[x]] := \{\sum_{i=0}^{\infty} a_i x_i; a_i \in \mathbf{K}$ für $i = 0, 1, 2, 3, \ldots\}$ ist.

    b) Für den kleinsten Körper $\mathbf{K} = \{0, 1\}$ beweise man, daß $\mathbf{K}[x]$ nicht isomorph zu $(\mathbf{K}[x])^*$ ist!

32. Eine quadratische Matrix n-ter Ordnung mit Elementen aus dem Körper $\mathbf{K}$ heißt magisches Quadrat der Ordnung n, falls die Summen der Elemente jeder Zeile, jeder Spalte und der beiden Diagonalen gleich sind.

a) zeigen Sie: $\mathrm{Mag}(n, \mathbf{K})$, die Menge der magischen Quadrate n-ter Ordnung, bildet einen linearen Unterraum des Raumes aller quadratischen Matrizen $M(n, \mathbf{K})$ der Ordnung n.

b) Für $n = 1, 2, 3$ und $\mathbf{K} = \mathbf{R}$ berechne man dim $\mathrm{Mag}(n, \mathbf{K})$ und gebe jeweils eine Basis an!

æ

Kapitel 3

Lineare Abbildungen und Matrizen

3.1 Grundlegende Eigenschaften

Wir beginnen mit einem Beispiel.
Sei V ein zweidimensionaler Vektorraum und $\{u, v\}$ 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

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

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 $k(w') = (r', s')$. Wegen $w' = r'u + s'v$ gilt

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

also

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

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

$$k(tw) = (tr, ts) = tk(w).$$

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 $f : V \to W$ eine Abbildung von V in W, für die für beliebige $u, v \in V$ und $r \in R$
$f(u + v) = f(u) + f(v)$ (f ist „additiv“) sowie
$f(rv) = rf(v)$ (f 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 $k_B : V \to R^n$, 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“
$$p_i : R^n \to R,\ p_i(r_1, \ldots, r_n) = r_i,$$
sie ist linear.

3. Die „identische“ Abbildung $id : V \to V,\ id(v) = v$ für alle $v \in V$ ist linear.

4. Zwischen beliebigen Vektorräumen V, W gibt es eine Nullabbildung $o : V \to W$, $o(v) = o$ 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 $f : A \to B$ 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 angegeben linearen Abbildungen zutreffen!

Wir wollen nun Operationen zwischen linearen Abbildungen einführen:
Seien $f, g : V \to W$ zwei lineare Abbildungen von V in W. Wir konstruieren eine lineare Abbildung $f + g : V \to W$ 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 $sf : V \to W$ wie folgt:
$(sf)(v) = sf(v)$ für alle $v \in V$.

Lemma 3.1 *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

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

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

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

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

Satz 3.2 $\mathrm{Hom}(V,W)$ *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, das die Nullabbildung der Nullvektor von $\mathrm{Hom}(V,W)$ ist:
Sei $f: V \to W$ 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 $f, g: V \to W$ 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 $g: V \to W$ und $f: W \to U$ lineare Abbildungen. Wir konstruieren die Abbildung $f \circ g: V \to U$ 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 3.3 *Die Abbildung $f \circ g$ ist linear.*

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

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

□

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

Lemma 3.4 *Sei $f : V \to W$ eine lineare Abbildung und seien $id_V : V \to V$ sowie $id_W : W \to W$ 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 $f : V \to W$ bijektiv ist, so existiert eine Abbildung $g : W \to V$ 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 3.5 *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 $f : V \to W$ gehören zwei Unterräume von V bzw. von W:

Definition: Sei $f : V \to W$ eine lineare Abbildung. $\mathrm{Ker}(f) = \{v \in V \mid f(v) = o\}$ heißt der Kern von f. $\mathrm{Im}(f) = \{w \in W \mid \text{es gibt ein } v \in V \text{ mit } f(v) = w\}$ heißt das Bild von f.

Lemma 3.6 $\mathrm{Ker}(f) \subseteq V$ *und* $\mathrm{Im}(f) \subseteq W$ *sind Unterräume.*

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

Nützlich, wenn auch trivial ist das folgende

Lemma 3.7 *Die lineare Abbildung $f : V \to W$ ist genau dann surjektiv, wenn* $\mathrm{Im}(f) = W$. *Die lineare Abbildung $f : V \to W$ ist genau dann injektiv, wenn* $\mathrm{Ker}(f) = \{o\}$.

Beweis: Die erste Aussage ergibt sich aus der Definition, ist also wirklich trivial.
Sei nun f injektiv und $v \in \mathrm{Ker}(f)$, 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 $\mathrm{Ker}(f) = \{o\}$.
Sei umgekehrt $\mathrm{Ker}(f) = \{o\}$ 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. □

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 $\mathrm{Im}(f) = f(V)$.

Satz 3.8 *Sei $f : V \to W$ eine lineare Abbildung und $M \subseteq V$ ein Erzeugendensystem von V. Dann ist $f(M)$ ein Erzeugendensystem von* $\mathrm{Im}(f)$.

Beweis: Sei $w \in \mathrm{Im}(f)$, 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

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

Sei nun $f : V \to W$ eine lineare Abbildung und $\{v_1, \ldots, v_k\}$ eine linear abhängige Teilmenge von V. Dann gibt es Zahlen $r_1, \ldots, r_k$, die nicht alle null sind, so daß $r_1 v_1 + \ldots + r_k v_k = o$.
Durch Anwendung von f und Ausnutzung der Linearität von f erhalten wir

$$\begin{aligned} o &= f(r_1 v_1 + \ldots + r_k v_k) \\ &= f(r_1 v_1) + \ldots + f(r_k v_k) \\ &= r_1 f(v_1) + \ldots + r_k f(v_k), \end{aligned}$$

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

Wir erhalten den

Satz 3.9 *Sei $f : V \to W$ eine lineare Abbildung und $v_1, \ldots, v_k$ Vektoren aus V. Wenn $\{f(v_1), \ldots, f(v_k)\}$ linear unabhängig ist, so ist $\{v_1, \ldots, v_k\}$ auch linear unabhängig.* □

Satz 3.10 *Sei $f : V \to W$ eine lineare Abbildung, weiter sei $U \subseteq V$ ein Teilraum von V, so daß der Durchschnitt von U und* $\mathrm{Ker}(f)$ *nur den Nullvektor enthält. Wenn nun $\{v_1, \ldots, v_k\}$ eine linear unabhängige Teilmenge von U ist, so ist auch $\{f(v_1), \ldots, f(v_k)\}$ 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 $\mathrm{Ker}(f)$ und U, also gilt $\sum r_i v_i = o$ und aus der linearen Unabhängigkeit von $\{v_1, \ldots, v_k\}$ folgt $r_1 = \ldots = r_k = o$. □

Den folgenden Satz werden wir oft anwenden:

Satz 3.11 *Sei $f : V \to W$ eine lineare Abbildung, dann gibt es einen Unterraum $U \subseteq V$ mit $U \oplus \mathrm{Ker}(f) = V$ und es gilt* $\dim V = \dim \mathrm{Ker}(f) + \dim \mathrm{Im}(f)$.

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

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

Folgerung 3.12 *Wenn $f : V \to W$ ein Isomorphismus ist (also eine bijektive lineare Abbildung), dann ist* $\dim V = \dim W$.

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

3.2 Darstellungsmatrizen

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

Satz 3.13 (Prinzip der linearen Fortsetzung) *Sei $B = \{v_1, \ldots, v_n\}$ eine Basis von V und $w_1, \ldots, w_n \in W$ beliebig gewählte Vektoren. Dann gibt es genau eine lineare Abbildung*

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

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

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

Lemma 3.14 *Sei $B = \{v_1, \ldots, v_n\}$ eine Basis des Vektorraums V, dann ist die durch $k_B(v_i) = e_i = (0, \ldots, 1, \ldots, 0)$ gegebene Koordinatenabbildung $k_B : V \to R^n$ ein Isomorphismus.*

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

Wir wenden das Prinzip der linearen Fortsetzung an, um lineare Abbildungen zahlenmäßig beschreiben zu können:
Sei $f : V \to W$ eine lineare Abbildung. Wir wählen Basen $B = \{v_1, \ldots, v_n\}$ von V und $C = \{w_1, \ldots, w_m\}$ von W. Dann können wir jeden Vektor $f(v_i)$ durch die Basis C ausdrücken:

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

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:
$f : R^4 \to R^2$ 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

$$A_{BC} : \mathrm{Hom}(V, W) \to M_{mn},$$

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 $f, f' : V \to W$ lineare Abbildungen und $r \in R$, $B = \{v_1, \ldots, v_n\}$ sei eine Basis von V, $C = \{w_1, \ldots, w_m\}$ eine Basis von W und

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

also

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

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 3.15 *Sei* $\dim V = n$ *und* $\dim W = m$, *dann sind die Vektorräume* $\mathrm{Hom}(V, W)$ *und* M_{mn} *isomorph, sie haben die Dimension* mn. □

3.3 Matrixmultiplikation, Inverse von Matrizen

Seien nun lineare Abbildungen $f : V \to W$ und $g : W \to U$ 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 $B = \{v_1, \ldots, v_n\}$ von V, $C = \{w_1, \ldots, w_m\}$ von W und $D = \{u_1, \ldots, u_l\}$ von U. Es sei

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

dann ist

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

also ist

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

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.

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 $id : V \to V$. Wir wählen eine Basis $B = \{v_1, \ldots, v_n\}$ von V, dann ist $id(v_i) = v_i$, also

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

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 $f : V \to W$ ein Isomorphismus ist, so existiert eine zu f inverse Abbildung $f^{-1} : W \to V$ und für die zugeordneteten 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

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

als Matrixprodukt schreiben:

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

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 3.16 *Sei $f : V \to W$ eine lineare Abbildung, B eine Basis von V und C eine Basis von W, dann ist $rg(A_{BC}(f)) = \dim \mathrm{Im}(f)$.*

Beweis: Sei $B = \{v_1, \ldots, v_n\}$, dann ist $\{f(v_1), \ldots, f(v_n)\}$ ein Erzeugendensystem von $\mathrm{Im}(f)$, sei oBdA. $\{f(v_1), \ldots, f(v_r)\}$ 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 $rg(A_{BC}(f)) = r$. □

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 $F : R^k \to R^l$ ist offenbar linear.

Sei nun wieder $f : V \to W$ 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:

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

Wir überlassen es dem Leser zu zeigen, daß $k_C \circ f = F \circ k_B$ gilt (ein derartiges Diagramm heißt „kommutativ“).

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

Satz 3.17 *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. □

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

Folgerung 3.18 *Sei $F \in M_{nn}, F$ ist genau dann regulär, wenn $rg(F) = n$ ist.*

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

Folgerung 3.19 *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 $f : V \to W$ und $g : W \to U$. Sei $\{v_1, \ldots, v_n\}$ eine Basis von V, dann ist $\{f(v_1), \ldots, f(v_n)\}$ ein Erzeugendensystem von $\mathrm{Im}(f)$, also $\dim \mathrm{Im}(f) \leq \dim V$ und ebenso folgt $\dim g(f(V)) \leq \dim f(V)$, also $rg(GF) \leq rg(F)$. Weiter ist $\mathrm{Im}(g \circ f)$ in $\mathrm{Im}(g)$ enthalten, also ist $\dim \mathrm{Im}(g \circ f) \leq \dim \mathrm{Im}(g)$, also $rg(GF) \leq rg(G)$.
Wenn g ein Isomorphismus ist, so ist $\dim T = \dim g(T)$ für jeden Unterraum $T \subseteq W$, also ist $\dim \mathrm{Im}(g \circ f) = \dim \mathrm{Im}(f)$. □

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

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

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 3.20 *Die Elementarmatrizen sind regulär.*

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

Wir erhalten damit etwas bereits bekanntes:

Folgerung 3.21 *Die elementaren Zeilenoperationen ändern den Rang der Matrix nicht.*
□

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

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

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

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

$$Z_k \dots Z_1 = Z_k \dots Z_1 E.$$

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 Einheitsmatix 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:

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

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

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

3.4 Basiswechsel

Die Zuordnungen
Vektor $\to$ Koordinaten und
Abbildung $\to$ 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

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

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

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

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 $V = R^3$, 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 $k_B(v) = \begin{pmatrix} 5 \\ 7 \\ 2 \end{pmatrix}$, während das Koordinatentupel von v bezüglich C gleich $A \begin{pmatrix} 5 \\ 7 \\ 2 \end{pmatrix} = \begin{pmatrix} 14 \\ 0 \\ -4 \end{pmatrix}$ ist.
Seien nun eine lineare Abbildung $f : V \to W$ 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:

$$\begin{tikzcd}
 & V \arrow[rr, "f"] \arrow[dd, "k_{B'}"] & & W \arrow[dd, "k_{C'}"] \\
V \arrow[ur, "id_V"] \arrow[rr, "f"] \arrow[dd, "k_B"'] & & W \arrow[ur, "id_W"'] \arrow[dd, "k_C"] & \\
 & R^n \arrow[rr, "A_{B'C'}(f)"] & & R^m \\
R^n \arrow[ur, "A_{BB'}(id_V)"] \arrow[rr, "A_{BC}(f)"'] & & R^m \arrow[ur, "A_{CC'}(id_W)"'] &
\end{tikzcd}$$

Alle Diagramme auf den Seitenflächen und der Deckfläche sind kommutativ, damit ist auch das Diagramm auf der unteren Fläche kommutativ und wir erhalten

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

Wir wissen, daß eine beliebige Matrix mit Hilfe von Zeilen- und Spaltenoperationen in

die Form

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

gebracht werden kann. Daraus erhalten wir das

Lemma 3.22 *Sei $f : V \to W$ 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 $f(v_i) = o$ für $i > r$ gilt.* □

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 3.23 (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.* □

3.5 Idempotente Abbildungen und direkte Summen

Wir betrachten noch eine Reihe spezieller Matrizen und Endomorphismen.

Definition: Sei $f : V \to V$ 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 3.24 *Wenn $f : V \to V$ nilpotent ist, so gibt es ein $m \leq \dim V$ mit $f^m = o$.*

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

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

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

Satz 3.25 *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 $\mathrm{Im}(p) = U$ 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 $\mathrm{Ker}(p) = W$, wir sehen, daß das kein Zufall ist:

Satz 3.26 *Wenn $f : V \to V$ idempotent ist, so gilt $V = \mathrm{Ker}(f) \oplus \mathrm{Im}(f)$.*

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

Satz 3.27 *Wenn $f : V \to V$ 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. □

Satz 3.28 *Seien $f, g : V \to V$ idempotente Abbildungen mit $f + g = id$. Dann ist $V = \mathrm{Im}(f) \oplus \mathrm{Im}(g)$.*

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

Wenn umgekehrt V die direkte Summe von Unterräumen U und W ist, so haben wir zwei Projektionen f, g von V mit $\mathrm{Im}(f) = U$ und $\mathrm{Im}(g) = W$ 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 3.29 *Seien $f_1, \ldots, f_k : V \to V$ idempotente Abbildungen, für die $f_i \circ f_j = o$ für $i \neq j$ sowie $f_1 + \ldots + f_k = id$ gilt. Dann ist*

$$V = \mathrm{Im}(f_1) \oplus \ldots \oplus \mathrm{Im}(f_k).$$

Beweis: Sei v ein beliebiger Vektor aus V, dann ist $v = id(v) = (f_1 + \ldots + f_k)(v) = f_1(v) + \ldots + f_k(v)$, also $\mathrm{Im}(f_1) + \ldots + \mathrm{Im}(f_k) = V$. Sei weiter v ein Vektor, der in $\mathrm{Im}(f_i)$ und in der Summe der $\mathrm{Im}(f_j)$ $(j \neq i)$ 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 $v = o$. □

3.6 Aufgaben

1. Entscheiden Sie, ob die folgenden beiden Matrizen invertierbar sind und berechnen Sie gegebenenfalls die inverse Matrix!

$$A = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} 3 & -2 & 1 \\ 2 & 1/2 & -1 \\ 1/3 & 3 & 7 \end{pmatrix}$$

2. Lösen Sie folgende Matrizengleichungen!

   a) $\begin{pmatrix} 1 & 1/2 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x & y \\ u & v \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 6 & 5 \end{pmatrix}$

   b) $\begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix} \begin{pmatrix} x & y \\ u & v \end{pmatrix} = \begin{pmatrix} 1 & 1/2 \\ -1 & 0 \end{pmatrix}$

3. $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ sei eine fixierte Matrix. Man zeige: Wenn für jede beliebige 2×2-Matrix $\begin{pmatrix} x & y \\ u & v \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x & y \\ u & v \end{pmatrix}$ erfüllt ist, so existiert eine reelle Zahl r, so daß $x = v = r$ und $y = u = 0$ gilt.

4. Gibt es quadratische Matrizen $A_{n,n} (n \geq 2)$ mit folgenden Eigenschaften: $A * A = E$ und $A \neq E$?

5. Entscheiden Sie, bei welchen der folgenden Abbildungen es sich um lineare Abbildungen handelt. Begründen Sie Ihre Entscheidung.

   a) $f : \mathbf{R} \to \mathbf{R}$ mit $f(x) = x^2$

   b) $f : \mathbf{R}^2 \to \mathbf{R}$ mit $f(x, y) = \frac{x+y}{2}$

   c) $f : \mathbf{R}^2 \to \mathbf{R}$ mit $f(x, y) = \sqrt{xy}$

   d) $f : \mathbf{R} \to \mathbf{R}$ mit $f(x) = |\, x \,|$

6. Weisen Sie nach, daß $f : \mathbf{R}^3 \to \mathbf{R}^3$ mit $f(x, y, z) = (3x - y, 4x + 2z, x + y + 2z)$ eine lineare Abbildung ist. Bestimmen Sie Ker f und Im f.

7. Sei die lineare Abbildung $f : \mathbf{R}^2 \to \mathbf{R}^2$ gegeben durch $f((1, 1)) = (1, 2), f((1, -1)) = (1, 1)$. Geben Sie die Matrix von f bezüglich der kanonischen Basis von $\mathbf{R}^2$ an!

8. Sei $f : \mathbf{R}^3 \to \mathbf{R}^3$ eine lineare Abbildung mit $f(x, y, z) = (3x + y - z, x + y + z, -x - y - z)$. Es seien folgende Basen gegeben: B kanonische Basis, $B' = \{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}, B'' = \{(2, 1, 1), (1, 1, 0), (2, 0, 0)\}$.

   a) Ermitteln Sie $A_{BB'}(f)$ und $A_{BB''}(f)$!

   b) Sei $g : \mathbf{R}^3 \to \mathbf{R}^3$ mit der beschreibenden Matrix $A_{BB''}(g) = \begin{pmatrix} 0 & 0 & 1 \\ 2 & 1 & 0 \\ 3 & 2 & -1 \end{pmatrix}$

   Geben Sie eine Abbildungsgleichung für g an!

9. Seien V und V' reelle Vektorräume mit den Basen $B = \{b_1, b_2, b_3\}$ und $B' = \{a_1, a_2\}$. Eine lineare Abbildung $f : V \to V'$ werde folgendermaßen festgelegt: $f(b_1) = a_1, f(b_2) = a_2$ und $f(b_3) = a_1 + a_2$.

   a) Bestimmen Sie $f(b_1 + 3b_2 - 5b_3)$, $\mathrm{Im} f$, $\mathrm{Ker} f$ und $rg f$! b) Geben Sie die Matrixdarstellung von f bzgl. B und B' an!

10. a) Weisen Sie nach, daß

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

   und

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

   Basen des $\mathbf{R}^3$ sind!

   b) Berechnen Sie die Koordinatentripel von $\vec{x} = (3,4,5)_B, \vec{x}_1 = (1,2,3)_B, \vec{x}_2 = (4,5,6)_B$ und $\vec{x}_3 = (7,8,9)_B$ bzgl. B^*!

   c) Welche Vektoren des $\mathbf{R}^3$ sind $\vec{x}, \vec{x}_1, \vec{x}_2$ und $\vec{x}_3$?

11. Im Vektorraum $\mathbf{R}^2$ seinen folgende Basen gegeben: $B_0 = \{(1,0), (0,1)\}, B_1 = \{(1,1), (0,1)\}, B_2 = \{-1,0), (-1,-1)\}$. Ermitteln Sie die Matrizen A, B und C für den Basiswechsel

   (A): $B_0 \to B_1, (B) : B_1 \to B_2, (C) : B_0 \to B_2$.

   Gilt $A * B = C$?

12. a) Sei $\mathbf{V}$ ein reeller, endlich-dimensionaler Vektorraum und f, g lineare Abbildungen von $\mathbf{V}$ in sich mit $f \circ g = id_{\mathbf{V}}$. Beweisen Sie, daß f bijektiv ist!

   b) Zeigen Sie anhand eines Beispiels, daß diese Behauptung für unendlich-dimensionale Vektorräume nicht stimmt.

13. Zeigen Sie, daß folgendes gilt: Ist $A_{n,n}$ eine reguläre Matrix und $B_{n,n}$ eine beliebige quadratische Matrix, so sind AB und BA ähnlich.

14. Sei $f : \mathbf{R}^3 \to \mathbf{R}^3$ eine lineare Abbildung, die durch $f(x,y,z) = (3x - y, 4x + 2z, x + y + 2z)$ gegeben ist; als Basen in $\mathbf{R}^3$ betrachten wir folgende Mengen: es sei $B = \{(1,1,1), (1,1,0), 1,0,0)\}$, und B' sei die geordnete kanonische Basis des $\mathbf{R}^3$.

   a) Bestimmen Sie Kern und Bildraum von f sowie deren Dimensionen!

   b) Ermitteln Sie die zu f gehörende Matrix bzgl. B und B'!

   c) Ist f ein Isomorphismus? (Begründung!)

15. Weisen Sie nach, daß für jede lineare Abbildung f gilt: f ist injektiv genau dann, wenn Ker $f = \{\vec{0}\}$.

16. a) Weisen Sie nach, daß $f : R^3 \to R^2$ mit $f(x, y, z) = (2x + z, y - x)$ eine lineare Abbildung ist!

    b) Bestimmen Sie Ker f und Im f.

    c) Ist f ein Isomorphismus?

17. Seien B(kanonische Basis) und $B' = \{(2, 1, 1), (1, 1, 0), (2, 0, 0)\}$ Basen des R^3. Eine lineare Abbildung $g : R^3 \to R^3$ sei bezüglich dieser Basen durch folgende Matrix gegeben: $\begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 0 \\ 3 & 2 & -1 \end{pmatrix}$. Geben Sie eine Abbildungsgleichung für g an! $(g(x, y, z) = \ldots)$

18. Ist $A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & -1 \\ -1 & 2 & 3 \end{pmatrix}$ regulär? Wenn ja, berechnen Sie A^{-1}!

19. Sei $f : V \to W$ ein Isomorphismus zweier K-Vektorräume V und W; ferner sei $a_1, \ldots, a_n$ eine Basis von V. Zeigen Sie, daß $f(a_1), \ldots, f(a_n)$ eine Basis von W ist!

20. Sei $f : V \to W$ ein Isomorphismus zwischen den Vektorräumen V und W. Zeigen Sie, daß dann auch $f^{-1} : V \to W$ ein Isomorphismus ist!

21. Sei V ein dreidimensionaler Vektorraum über $\mathbf{R}$ mit der Basis $\{a_1, a_2, a_3\}; f : V \to V$ sei eine lineare Abbildung mit $f(a_1) = \alpha a_1 + a_3, f(a_2) = a_1 + \beta a_2, f(a_3) = a_2 + \gamma a_3$. Geben Sie eine notwendige und hinreichende Bedingung in Abhängigkeit von $\alpha, \beta, \gamma \in \mathbf{R}$ dafür an, daß f ein linearer Isomorphismus ist!

22. Sei V ein reeller, zweidimensionaler Vektorraum, und $f : V \to V$ eine lineare Abbildung. Beweisen Sie, daß die linearen Abbildungen $id = f^o, f = f^1$ und $f^2 = f \circ f$ linear abhängige Vektoren in $\mathrm{Hom}(V, V)$ sind.

23. a) Sei $f : \mathbf{R}^4 \to \mathbf{R}^3$ die durch $f(x_1, x_2, x_3, x_4) = (x_1 - x_2 + x_3 + x_4, x_1 + 2x_3 - x_4, x_1 + x_2 + 3x_3 - 3x_4)$ definierte lineare Abbildung. Bestimmen Sie die Dimension von Ker f und Im f, und geben Sie jeweils eine Basis dieser Räume an!

    b) Man gebe eine lineare Abbildung $f : \mathbf{R}^3 \to \mathbf{R}^3$ an, deren Bild erzeugt wird von (1,2,3) und $(4, 5, 6)$, während der Kern von f duch (-1,2,0) erzeugt werden soll.

24. In $\mathbf{R}^4$ seien folgende Vektoren gegeben: $a_1 = (1, 1, 2, 1), a_2 = (-2, 3, -2, 1), a_3 = (-1, t, 2, 5), b_1 = (0, -2, 4, 3), b_2 = (3, 1, 0, 1)$. Für welche Werte von $t \in \mathbf{R}$ existiert eine lineare Abbildung $\phi : \mathbf{R}^4 - \to \mathbf{R}^4$ mit Ker $\phi = L(\{a_1, a_2, a_3\})$ und Im $\phi = L(\{b_1, b_2\})$? Für alle diese Werte von t gebe man eine solche lineare Abbildung ϕ und ihre Matrixdarstellung in der Standardbasis an!

25. Sei $I = [0, 1] \subseteq \mathbf{R}$, sowie $C(I) = \{f : I \to \mathbf{R}; f \text{ stetig}\}$. Für beliebiges $c \in I$ sei $\varphi_c \in C(I)'$ definiert durch $\varphi_c(f) := f(c)$, mit $f \in C(I)$.

a) Man zeige, daß die Menge $\{\varphi_c\}_{c\in I}$ linear unabhängig ist!

b) Zeigen Sie: Der durch $\psi(f) := \int f(x)dx$, mit $f \in C(I)$, definierte Kovektor $\psi \in C(I)^*$ liegt nicht in der linearen Hülle der Menge $\{\varphi_c\}_{c\in I}$!

26. Sei $I = (a, b)$ ein offenes Intervall in $\mathbf{R}$, und weiterhin $C(I; \mathbf{C}) := \{f : I \to \mathbf{C}; f$ ist stetige Abbildung$\}$.

a) Zeigen Sie, daß die Funktionen $\{e^{i\lambda x}\}_{\lambda\in\mathbf{R}}$ eine linear unabhängige Teilmenge des reellen Vektorraums $C(I; \mathbf{C})$ bilden! (Hinweis: Man betrachte die durch $Df := \frac{d}{d}\frac{f}{x}$ gegebene lineare Abbildung $D : C(I; \mathbf{C}) \to C(I; \mathbf{C})$!)

b) Man schlußfolgere daraus, daß auch die Funktionen $\{\cos nx, \sin nx\}_{n\in\mathbf{N}}$ linear unabhängig in $C(I; \mathbf{R})$ sind!

c) Sei $U := L(\{1, \sin x, \cos x, \sin 2x, \cos 2x\}) \subseteq C(I; \mathbf{R})$. Man gebe die Matrix von $D = \frac{d}{d}\bar{x} \in L(U)$ bezüglich dieser Basis von U an, und bestimme $\mathrm{Ker}(D)$ und $\mathrm{Im}(D)$.

27. Sei V ein Vektorraum, und $f : V \to V$ eine lineare Abbildung mit $f \circ f = f$. Beweisen Sie, daß unter diesen Voraussetzungen $V = \mathrm{Ker}(f) \oplus \mathrm{Im}(f)$ gilt!

28. Sei $U \subseteq V$ ein Unterraum des Vektorraumes V, und $\dim V < \infty$. Ferner sei $f : V \to V$ eine lineare Abbildung.

a) Man zeige: $\dim U - \dim\ \mathrm{Ker}(f) \leq \dim f(U) \leq \dim U$,

$\dim U \leq \dim f^{-1}(U) \leq \dim U + \dim\ \mathrm{Ker}(f)$

b) Man finde notwendige und/oder hinreichende Bedingungen dafür, daß bei den einzelnen Teilen von Punkt a) das Gleichheitszeichen gilt!

29. Sei V ein endlich-dimensionaler Vektorraum, und $f, g \in \mathrm{Hom}(V)$. Beweisen Sie: $\dim \mathrm{Ker}(f \circ g) \leq \dim \mathrm{Ker(f)} + \dim \mathrm{Ker(g)}$. Wann gilt die Gleichheit ?

30. Sei V ein Vektorraum, und $f \in \mathrm{Hom}(V)$ ein Endomorphismus, der $f^2 + f - id_V = 0$ erfüllen möge.

a) Beweisen Sie: f ist ein Automorphismus.

b) Geben Sie eine Formel für f^{-1} an!

31. V sei ein Vektorraum über einem Körper K, und $f_0, f_1, \ldots, f_n$ seien lineare Abbildungen von V nach K. Man beweise: Es existieren $\alpha_1, \ldots, \alpha_n \in K$ mit $f_0 = \alpha_1 f_1 + \ldots + \alpha_n f_n$ genau dann, wenn gilt: $\bigcap_{i=1}^n \mathrm{Ker}(f_i) \subseteq \mathrm{Ker}(f_0)$.

32. Sei V ein reeller, endlich-dimensionaler Vektorraum. Zeigen Sie, daß man jeden Endomorphismus $f \in \mathrm{Hom}(V, V)$ als Summe zweier Automorphismen von V darstellen kann!

33. Sei V ein $\mathbf{K}$-Vektorraum, und $f \in \mathrm{Hom}(V, V)$. Die Abbildung f heißt *nilpotent* , wenn es eine natürliche Zahl $m \geq 1$ derart gibt, daß $f^m = 0$ gilt.

a) Beweisen Sie folgende Aussagen:

(i) Wenn f nilpotent und $\dim V = n < \infty$, so ist $f^n = 0$.

(ii) Sei f nilpotent und g ein Automorphismus von V

mit $g \circ f = f \circ g$. Dann ist $(g + f)$ ein Automorphismus.

b) Sei f nilpotent und $g = id_V$(d.h. $f + id_V$ ist ein Automorphismus). Man bestimme $(f + id_V)^{-1}$!

c) Man gebe ein Beispiel für ein nilpotentes f und einen Automorphismus g derart an, daß $f + g$ kein Automorphismus ist!

34. Wir betrachten folgende Unterräume von $\mathbf{R}^n$: $M := \{x = (x_1, \ldots, x_n); x_1 + \ldots + x_n = 0\}$, $U := \{x = (x_1, \ldots, x_n); x_1 = \ldots = x_n\}$.

a) Zeigen Sie, daß $\mathbf{R}^n = M \oplus U$ gilt! Sei $x = m(x) + v(x)$ die entsprechende eindeutige Zerlegung eines Vektors $x \in \mathbf{R}^n$ in Komponenten aus M bzw. V.

b) Zeigen Sie, daß die Abbildung $f_m : \mathbf{R}^n \to \mathbf{R}^n$, die durch $f_m(x) := m(x)$ definiert wird, linear ist!

c) Man gebe eine Basis von $\mathbf{R}^n$ an, bzgl. welcher die Abbildung f_m die folgende Matrix besitzt: $\begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & \\ 0 & 0 & 1 & 0 & \\ & & \ldots & & \\ 0 & \ldots & 0 & 1 & 0 \\ 0 & & \ldots & 0 & 0 \end{pmatrix}$

35. Man bestimme eine lineare Abbildung $\varphi : \mathbf{R}^4 \to \mathbf{R}^4$, die

$$\mathrm{Ker}(\varphi) = L\{(1, 1, -1, 0), (1, 1, 0, 1)\}$$

und

$$\mathrm{Im}(\varphi) = L\{(1, 1, 1, 1), (1, 0, 1, 0)\}$$

erfüllt, und gebe die Matrix von φ bezüglich der Standardbasis von $\mathbf{R}^4$ an!

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 $+ : A \times V \to A$ 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 *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 $Q+U = P+v+U = P+U = H$. 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 $U = W$. □

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 $P_0, P_1, \ldots P_k \in A$ 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 4.2 *Die Punkte $P_0, \ldots, P_k$ sind genau dann in allgemeiner Lage, wenn die Vektoren $v_1 = \overrightarrow{P_0P_1}, \ldots, v_k = \overrightarrow{P_0P_k}$ linear unabhängig sind.*

Beweis: Seien die Punkte $P_0, \ldots, P_k$ in allgemeiner Lage. Wir setzen $H = P_0 + L\{v_1, \ldots v_k\}$, 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 $P_0, \ldots, P_k$ nicht in allgemeiner Lage, folglich ist $\dim H = k$, d.h. $\{v_1, \ldots, v_k\}$ ist eine linear unabhängige Menge.
Sei $\{v_1, \ldots, v_k\}$ linear unabhängig. Wir nehmen an, daß die Punkte $P_0, \ldots, P_k$ 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 $\dim U \geq k$. □

Lemma 4.3 *Seien $P_0, \ldots, P_k \in A$ Punkte in allgemeiner Lage, dann gibt es einen eindeutig bestimmten k-dimensionalen Unterraum H von A, der $P_0, \ldots, P_k$ enthält.*

Beweis: Die Existenz ist klar: $H = P_0 + L\{\overrightarrow{P_0P_1}, \ldots, \overrightarrow{P_0P_k}\}$ 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 $L\{\overrightarrow{P_0P_1}, \ldots, \overrightarrow{P_0P_k}\}$ in U enthalten und beide Räume haben dieselbe Dimension, sind also gleich. □

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

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

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

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

Lemma 4.4 *Sei $\{P, b_1, \ldots, b_k\}$ ein Koordinatensystem von $H = P+U$, das Koordinatentupel des Punkts $Q \in H$ sei $(1, q_1, \ldots, q_k)$, das von S sei $(1, s_1, \ldots, s_k)$ und das des Vektors $v \in U$ sei $(0, r_1, \ldots, r_k)$. Dann ist das Koordinatentupel von $Q + v$ gleich $(1, q_1 + r_1, \ldots, q_k + r_k)$ und der Verbindungsvektor von Q nach S hat die Koordinaten $(0, s_1 - q_1, \ldots, s_k - q_k)$.*

Den Beweis überlassen wir dem Leser. □

Sei nun $\{P, e_1, \ldots, e_n\}$ 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 $(1, x_1, \ldots, x_n)$, für die

$$\sum a_{ij} x_j = b_i, \; i = 1, \ldots, m$$

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

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

die Lösungsmenge des Gleichungssystems, dann gilt

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

oder

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

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

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

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

genau dann eine eindeutig bestimmt Lösung $(r_1, \ldots, r_k)$ 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:

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

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

$$\begin{array}{rcl} f_{k+1}(x) & = & 0 \\ & \ldots & \\ f_n(x) & = & 0 \end{array}$$

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

Beispiel:
Sei $H = (1, 1, 1) + L(e_1 + e_2, e_2 + e_3)$ 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 $L((1, 1, 0), (0, 1, 1))$ 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:

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

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 4.6 *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 $u_1 = u_2 \in U_1 \cap U_2$. □

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 4.7 *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. □

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 4.8 *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. □

Satz 4.9 *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 $n - 2$.* □

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 4.10 *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 $\overrightarrow{P_1P_2} = \overrightarrow{P_1P} + \overrightarrow{PP_2} = -u_1 + u_2 \in U_1 + U_2$. □

Satz 4.11 $H_1 \vee H_2 = P_1 + L(\overrightarrow{P_1P_2}) + (U_1 + U_2)$.

Beweis: Sowohl H_1 als auch H_2 sind in $P_1 + L(\overrightarrow{P_1P_2}) + (U_1 + U_2)$ enthalten. Sei $H = H_1 \vee H_2 = P_1 + U = P_2 + U$. Dann ist $U_1 \subseteq U$, $U_2 \subseteq U$, $\overrightarrow{P_1P_2}$ liegt in U, also ist $P_1 + L(\overrightarrow{P_1P_2}) + (U_1 + U_2) \subseteq H$. □

Folgerung 4.12 *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 $f : A \to A'$ eine affine Abbildung, wenn eine lineare Abbildung $f' : V \to V'$ 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 4.13 *Sei $H \subseteq A$ ein Unterraum und $f : A \to A'$ 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. □

Dem Leser überlassen wir die folgende Aussage zum Beweis:

Folgerung 4.14 *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.* □

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

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

Wir wollen der Abbildung f eine Matrix zuordnen, die f eindeutig bestimmt.
Sei Q ein beliebiger Punkt von A mit dem Koordinatentupel $(1, r_1, \ldots, r_n)$, d.h. $Q = P + \sum r_i b_i$, dann ist

$$f(Q) = f(P) + f'(\sum r_i b_i) = f(P) + \sum r_i f'(b_i),$$

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

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

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

Wir wollen nun nachzuweisen, daß bei einer derartigen Matrixzuordnung das Produkt affiner Abbildungen dem Matrixprodukt der Darstellungsmatrizen der affinen Abbildungen entspricht.
Wir betracheten drei affine Räume A, B, C und zwei affine Abbildungen $f : A \longrightarrow B$, $g : B \longrightarrow C$. Wir wählen Koordinatensysteme $(P, b_1, \ldots, b_n)$ von A, $(Q, c_!, \ldots, c_m)$ von B und $(R, d_1, \ldots, d_l)$ von C. Es sei

$$f(P) = Q + \sum f_j c_j, \qquad \bar{f}(b_i) = \sum f_{ji} c_j,$$

$$g(Q) = R + \sum g_k d_k, \qquad \bar{g}(c_j) = \sum g_{kj} d_k.$$

Dann ist

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

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 $f : A \longrightarrow A,\ X = P + \sum x_i b_i$, dann ist

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

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 Beipiel 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!)

4.3 Aufgaben

1. Sei A ein affiner Raum. Beweisen Sie: Für beliebige Punkte $P_1, \ldots, P_{n+1}$ aus A gilt: Aus $P_{n+1} = P_0$ folgt $\sum_{i=1}^{n} \overrightarrow{P_iP_{i+1}} = \vec{o}$.

2. Seien $P_0, \ldots, P_n$ Punkte eines affinen Raumes und es gelte, daß die Vektoren $\overrightarrow{P_0P_i}$ $(i = 1, \ldots, n)$ linear unabhängig sind. Beweisen Sie, daß dann auch die Vektoren $\overrightarrow{P_kP_j}, 0 \leq j \leq n, j \neq k\}$ mit $1 \leq k \leq n$ (k beliebig, aber fest), linear unabhängig sind!

3. Seien P, Q, X, Y Punkte eines affinen Raumes A und (O, B) ein Koordinatensystem von A; seien $(p_1, \ldots, p_n), (q_1, \ldots, q_n), (x_1, \ldots, x_n)$, und $(y_1, \ldots, y_n)$ die Koordinaten-n-tupel dieser Punkte bzgl. (O, B). Beweisen Sie: Es gilt $\overrightarrow{PQ} = \overrightarrow{XY}$ gdw. $q_i - p_i = y_i - x_i$ für alle $i = 1, \ldots, n$.

4. Beweisen Sie: Drei Punkte A, B, C eines affinen Raumes, die durch ihre Ortsvektoren $\vec{a}, \vec{b}, \vec{c}$ gegeben sind, liegen genau dann auf einer Geraden, wenn es reelle Zahlen r, s, t gibt, die nicht alle Null sind, so daß gilt: $r\vec{a} + s\vec{b} + t\vec{c} = o$ und $r + s + t = 0$.

5. Untersuchen Sie die gegenseitige Lage der Geraden g_1 und g_2.

   a) $g_1 : X = (2, 3, 1) + r(2, 1, 1) g_2 : X = (-2, 1, 0) + s(1, 1, 0)$

   b) $g_1 : X = (3, 0, 4) + r(-4, 2, -6) g_2 : X = (-1, 2, -2) + s(2, -1, 3)$

   c) $g_1 : X = (1, 0, 0) + r(0, 1, 1) g_2 : X = (0, 1, 0) + s(1, 1, 2)$

6. Gegeben seien die Punkte A, B, C, D eines dreidimensionalen affinen Raumes durch folgende Koordinaten: $A : (2, 3, 5); B : (0, 5, 10); C : (3, 4, 6); D : (5, 0, 2)$.

   a) Bilden die vier Punkte ein ebenes Viereck? Begründen Sie Ihre Antwort!

   b) Zeigen Sie, daß die Geraden durch die Mittelpunkte benachbarter Seiten des Vierecks paarweise zueinander parallel sind!

   c) Untersuchen Sie, ob die Aussage b) für jedes nichtebene Viereck gilt!

7. Ein Parallelogramm werde von den Vektoren $a = \overrightarrow{AB}$ und $b = \overrightarrow{BC}$ aufgespannt. M_a und M_b seien die Mittelpunkte der Seiten AB bzw. BC.

   a) Ermitteln Sie Parametergleichungen der Geraden $g_1(A, M_b), g_2(C, M_a)$ und $g_3(B, D)$!

   b) Zeigen Sie, daß diese drei Geraden einen gemeinsamen Schnittpunkt S haben!

   c) Welche Rolle spielt S im Dreieck ABC ?

8. (P, Q, R, S, T, U, P) sei ein Sechseck in einem affinen Raum und $M_i (i = 1, \ldots, 6)$ seien die Mittelpunkte der Seiten dieses Sechsecks. Beweisen Sie: Die Schwerpunkte der Dreiecke $M_1M_3M_5$ und $M_2M_4M_6$ stimmen überein.

9. Sei $\mathcal{A}$ ein affiner Raum und Z, A, B, C, D paarweise verschiedene Punkte mit $B \in g_1(Z, A)$ und $D \in g_2(Z, C)$, sowie $g_1 \neq g_2$. Zeigen Sie, daß folgende Bedingungen äquivalent sind:

   (A) Es existiert eine reelle Zahl $s \in \mathbf{R}$ mit $\overrightarrow{ZB} = s\ \overrightarrow{ZA}$ und $\overrightarrow{ZD} = s\ \overrightarrow{ZC}$.

   (B) Es existiert eine reelle Zahl $t \in \mathbf{R}$ mit $\overrightarrow{ZA} = t\ \overrightarrow{AB}$ und $\overrightarrow{ZC} = t\ \overrightarrow{CD}$.

10. Seien P, Q, R paarweise verschiedene, kollineare Punkte eines affinen Raumes und $TV(P, Q, R) = \lambda$. Berechnen Sie $TV(Q, R, P), TV(R, P, Q)$ und $TV(R, Q, P)$ und untersuchen Sie, welche Beziehung jeweils zu $TV(P, Q, R)$ besteht!

11. Im $\mathbf{R}^3$ seien zwei Ebenen ϵ_1 und ϵ_2 jeweils durch drei Punkte gegeben: $\epsilon_1 : (0, 0, 1); (1, 0, 0); (0, 1, 0),\ \epsilon_2 : (1, 1, -4); (-3, 4, -3); (-4, 3, -4)$. Untersuchen Sie, welche Lage ϵ_1 und ϵ_2 zueinander haben!

12. Die Punkte A, B, C, D eines affinen Raumes liegen in einer Ebene genau dann, wenn es reelle Zahlen $\alpha, \beta, \gamma, \delta$ gibt, die nicht alle gleich Null sind, so daß für die Ortsvektoren $\vec{a}, \vec{b}, \vec{c}, \vec{d}$ dieser Punkte gilt: $\alpha\vec{a}+\beta\vec{b}+\gamma\vec{c}+\delta\vec{d} = \vec{o}$ und $\alpha+\beta+\gamma+\delta = 0$.

13. Sei A ein affiner Raum, $Z \in A$ und $k \in \mathbf{R}, k \neq 0$ und $\phi : A \longrightarrow A$ eine zentrische Streckung von A auf sich mit dem Zentrum Z und dem Streckfaktor k. Beweisen Sie, daß für beliebige Punkte $P, Q, R, S \in A$ folgende Beziehung gilt: Wenn $\overrightarrow{RS} = c\ \overrightarrow{PQ}$, so ist $\overrightarrow{\phi(r)\phi(S)} = c\ \overrightarrow{\phi(P)\phi(Q)}$.

14. Zeigen Sie, daß bei injektiven affinen Abbildungen die Dimension von Teilräumen eine Invariante ist.

15. Gegeben sei ein dreidimensionaler affiner Raum A mit einem Koordinatensystem (O, B) und drei Punkte X, Y, Z aus A mit den Koordinaten (-2, -3, 0); (-1, -1, 1); (-1, 1, 3) bzgl. (O, B).

    a) Zeigen Sie, daß X, Y, Z ein Dreieck bilden.

    b) ϕ sei eine zentrische Streckung mit dem Zentrum $P = (2, 1, 0)_{(O,B)}$ und dem Streckungsfaktor $k = -3$. In welches Dreieck wird das Dreieck XYZ überführt?

    c) Geben Sie eine Koordinatendarstellung von ϕ an!

16. Im A^3 seien die folgenden Punkte gegeben. $P_1 = (0, 0, 0),\ Q_1 = (-1, 1, 2), P_2 = (1, 2, -1),\ Q_2 = (-2, 4, -1), P_3 = (3, 2, 0),\ Q_3 = (6, 16, -3), P_4 = (1, 1, 1),\ Q_4 = (3, 3, 0)$.

    a) Gibt es eine affine Abbildung von A^3 in sich, die P_i auf Q_i abbildet? Wieviele solcher affinen Abbildungen gibt es?

    b) Im Falle der Existenz einer solchen Abbildung gebe man ein Beispiel an!

17. Seien A, B, C, D, E Punkte des A^3, die bezüglich des kanonischen Koordinatensystems folgende Koordinaten haben: $A = (3, -1, 0); B = (-1, 5, 1); C = (-1, -1, 2); D = (0, 0, 0); E = (5, 5, 5)$

a) Beweisen Sie, daß durch A, B, C genau eine Ebene geht!

b) Geben Sie für diese Ebene eine Parameterdarstellung und eine parameterfreie Darstellung an!

c) Ermitteln Sie die gegenseitige Lage der Ebene ABC und der Geraden DE (Bestimmen Sie gegebenenfalls den Durchschnitt!)!

18. Seien A, B, C paarweise verschiedene, nicht kollineare Punkte eines affinen Raumes. M_1 bzw. M_2 seien die Mittelpunkte der Dreiecksseiten $\overrightarrow{BC}$ bzw. $\overrightarrow{CA}$. Beweisen Sie, daß $\overrightarrow{M_1M_2}$ parallel zu $\overrightarrow{AB}$ ist!

19. Sei A ein zweidimensionaler affiner Raum mit einem Koordinatensystem (O, B), $P, Q \in A$ mit $P = (1, 0)_{0,B}, Q = (0, 1)_{0,B}$ und $\phi : A \to A$ sei gegeben durch: $\phi(0) = 0'; \phi(P) = P'; \phi(Q) = Q'$, mit $0' = (-4, -2)_{0,B}, P' = (-3, -1)_{0,B}$ und $Q' = (-5, -1)_{0,B}$.

    a) Begründen Sie, daß dadurch eine affine Abbildung von A in sich gegeben ist!

    b) Geben Sie die durch ϕ induzierte lineare Abbildung an!

    c) Geben Sie das Bild des Dreiecks X, Y, Z mit $X = (4, 1)_{0,B}, Y = (2, 2)_{0,B}$ und $Z = (3, 4)_{0,B}$ an!

    d) Geben Sie ϕ durch Transformationsgleichungen an!

20. a) Wann sind Teilräume eines affinen Raums zueinander parallel?

    b) Durch das Gleichungssystem

    $3x + 4y + 5z + 6u = 1$

    $3x + 5y + 6z + 4u = 2$

21. ist ein Teilraum des R^4 gegeben. Begründen Sie dies und geben Sie eine Parameterdarstellung dieses Teilraums an.

    c) Geben Sie einen weiteren Teilraum an, der zu dem unter b) gegebenen parallel ist.

22. a) Beweisen Sie, daß die Diagonalen in einem Parallelogramm einander halbieren.

    b) Beweisen Sie: Die Seitenhalbierenden eines Dreiecks in $\mathbf{R}^2$ teilen einander im Verhältnis 2:1.

23. a) Sind die Punkte P, Q, R, S mit $P = (1, 1, 1), Q = (2, 2, 2), R = (3, 2, 1), S = (5, 4, 3)$ (bzgl. des kanonischen Koordinatensystems) affin unabhängig?

    b) Beweisen Sie, daß es in einem n-dimensionalen affinen Raum höchstens $n + 1$ Punkte in allgemeiner Lage gibt.

24. Sei $f : \mathbf{R}^3 - \to \mathbf{R}^3$ gegeben durch $f(x_1, x_2, x_3) = (2x_1 + x_3 + 1, -x_1 + x_2 + 3x_3 + 5, x_2 + 7x_3 - 1)$;

    a) man zeige: f ist eine affine Abbildung.

b) Stellen Sie f als Produkt (=Hintereinanderausführung) einer Translation τ und einer linearen Abbildung ϕ des $\mathbf{R}^3$ dar!

c) Man bestimme die Fixpunktmenge C_f von f; dies ist ein affiner Unterraum von $\mathbf{R}^3$. Charakterisieren Sie den zugehörigen Vektorraum $V(C_f)$ durch die Abbildung ϕ!

d) Für die Gerade $H(p,b)$ mit $p = (-1,0,1), b = (2,1,0)$ bestimme man die Urbildmenge $f^{-1}(H)$!

e) Durch $P_1 = (2,1,0), P_2 = (3,1,1), P_3 = (1,0,1)$ und $P_4 = (0,0,0)$ wird ein Parallelogramm $(P_1P_2P_3P_4)$ in $\mathbf{R}^3$ definiert; man bestimme $Q_j = f(P_j)$ $(j = 1,\ldots,4)$ und zeige, daß $(Q_1Q_2Q_3Q_4)$ wiederum ein Parallelogramm ist!

25. Sei $f : \mathbf{R}^n \to \mathbf{R}^n$ eine affine Abbildung mit f^2=id. Zeigen Sie, daß f mindestens einen Fixpunkt besitzt! (Zusatz: Im Falle eines Körpers $\mathbf{K}$ mit char $(\mathbf{K}) = 2$ gebe man eine affine Abbildung $f : \mathbf{K}^2 \to \mathbf{K}^2$ mit $f^2 = id$ an, die keinen Fixpunkt besitzt.)

26. Zeigen Sie: für $n \geq 2$ gibt es eine affine Abbildung $f : \mathbf{R}^n \to \mathbf{R}^n$, die keinen Fixpunkt hat und trotzdem keine Translation ist!

27. Sei (A,V) ein affiner Raum. Ferner seien Punkte $p_1,\ldots,p_n \in A$ und Skalare $a_1,\ldots,a_n \in \mathbf{K}$ mit $a_1 + \ldots + a_n = 1$ gegeben.

    a) man zeige, daß der Ausdruck $S(x) = x + \sum_{k=1}^{n} \overrightarrow{xp_k} \cdot a_k$ nicht von der Wahl des Punktes $x \in A$ abhängt. (Der Punkt $S \equiv S(x)$ wird Schwerpunkt der Punkte $p_1,\ldots,p_n$ bzl. der Gewichte $a_1,\ldots,a_n$ genannt.)

    b) Finden Sie in $\mathbf{R}^3$ den Schwerpunkt der Punkte $p_1 = (0,1,0)$, $p_2 = (1,1,1)$ und $p_3 = (2,0,1)$ bezüglich der Gewichte $a_1 = 2$, $a_2 = -2$ und $a_3 = 1$.

28. Beweisen Sie: Eine Abbildung $f : A \to B$ zwischen affinen Räumen ist genau dann affin, wenn sie den Schwerpunkt eines beliebigen Tripels (p_1,p_2,p_3) von Punkten aus A erhält, d. h. mit $S = S(p_1,p_2,p_3)$ und $q_j = f(p_j)$ gilt $f(S) = S(q_1,q_2,q_3)$. (analog für den Schwerpunkt von Punktepaaren)

29. Man finde eine Homothetie $f : \mathbf{R}^3 \to \mathbf{R}^3$ mit $f((2,0,1)) = (4,1,2)$ und $f((4,2,2)) = (0,-3,0)$; gesucht sind Zentrum und Dehnungsfaktor dieser Homothetie!

30. Im affinen Punktraum $\mathbf{R}^3$ gebe man eine nichtleere Menge M an, die keine Ebene ist und für die $V(M) := \{\overrightarrow{pq}\,; p,q \in M\}$ ein zweidimensionaler Unterraum von $\mathbf{R}^3$ ist!

31. Sei $[A^n, V^n]$ eine affiner Raum, $o \in A$ ein fixierter Punkt. Beweisen Sie: Eine Menge $M \subseteq A$ ist eine k-Ebene genau dann, wenn zu je zwei Punkten $p_1, p_2 \in M$ und je zwei Zahlen $s,t \in \mathbf{R}$ mit $s+t = 1$ auch der Punkt $P := o + s\cdot \overrightarrow{p_1} + t\cdot \overrightarrow{p_2}$ zu M gehört.

32. Im affinen Raum $\mathbf{R}^3$ sei eine Menge M definiert durch $M := \{(x_1, x_2, x_3) \in \mathbf{R}^3; 2x_1 - x_3 = 2\}$. Ferner sei der Vektor $b = (1, 1, 3)$ gegeben.

    a) Man zeige: für alle Punkte $p \in \mathbf{R}^3$ schneidet die Gerade $H(p, b)$ die Ebene M in genau einem Punkt $p' \in M$.

    b) Sei $\phi : \mathbf{R}^3 \to M$ definiert durch $\phi(p) := p'$. Man gebe die Koordinatendarstellung von ϕ an und zeige, daß ϕ eine affine Abbildung ist, die $\mathbf{R}^3$ surjektiv auf M abbildet. (ϕ heißt Parallelprojektion auf M in Richtung von b).

    c) Sei $N \subseteq \mathbf{R}^3$ eine 2-dimensionale Ebene. Beweisen Sie: $\Phi\mid_N$ ist injektiv gdw. $b \notin V(N) = \{\overrightarrow{pq}; p, q \in N\}$.

33. (Zentralprojektion) Im reellen affinen Raum $\mathbf{R}^3$ seien ein Punkt $z = (1, 3, 1)$ sowie die beiden parallelen 2-Ebenen $H_1 := \{(x_1, x_2, x_3) \in \mathbf{R}^3; 2x_1 - x_2 + x_3 = 1\}'$ $H_2 := \{(x_1, x_2, x_3) \in \mathbf{R}^3; 2x_1 - x_2 + x_3 = 4\}$.

    a) Zeigen Sie: Für jedes $x \in H_1$ schneidet die Gerade $H(x, z)$ die Ebene H_2 in genau einem Punkt $x' \in H_2$.

    b) Durch die Zuordnung $\phi : x \mapsto x' =: \phi(x)$ wird eine affine Abbildung $\phi : H_1 \to H_2$ definiert.

    c) Geben Sie die Koordinatendarstellung von ϕ an!

34. Seien A, B affine Räume, $f : A \to B$ eine affine Abbildung sowie $H_1, H_2 \subseteq B$ affine Teilräume in B mit $f^{-1}(H_1) \neq \emptyset, f^{-1}(H_2) \neq \emptyset$. Zeigen Sie:

    a) $f^{-1}(H_i)$ sind affine Teilräume von A. $(i = 1, 2)$

    b) Wenn $H_1 \parallel H_2$ gilt, so folgt auch $f^{-1}(H_1) \parallel f^{-1}(H_2)$.

35. In $\mathbf{R}^2$ sei ein kartesisches Koordinatensystem mit Ursprung O fixiert. Sei weiterhin $F : \mathbf{R}^2 \to \mathbf{R}^2$ die Spiegelung des $\mathbf{R}^2$ an der Geraden $2x - 4y = 6$.

    a) Geben Sie einen Punkt Q und eine lineare Abbildung ϕ an mit $F(P) = Q + \Phi(\overrightarrow{OP})$ für alle $P \in \mathbf{R}$!

    b) Man beschreibe F^{-1} in Koordinaten!

36. Sei $W := L(\{(4, 3, 1, 0, 1), (1, -1, -2, 2, 1)\}) \subseteq \mathbf{R}^5$ gegeben, ferner sei der Punkt $P = (1, 1, 0, 1, 2) \in \mathbf{R}^5$ fixiert.

    a) Man bestimme eine Basis von $W^\perp$!

    b) Man gebe ein lineares Gleichungssystem an, das die Menge $M := P + W \subseteq \mathbf{R}^5$ als Lösungsraum besitzt!

37. Sei $\Delta = (A, B, C)$ ein Dreieck in $\mathbf{R}^2$.

    a) Man drücke die Seitenhalbierenden $\vec{s_1}, \vec{s_2}, \vec{s_3}$ von Δ durch die Seiten $\vec{a}, \vec{b}, \vec{c}$ von Δ aus!

    b) Man zeige, daß die Seitenhalbierenden von Δ ebenfalls ein Dreieck bilden, d.h. es gilt $\vec{s_1} + \vec{s_2} + \vec{s_3} = 0$.

38. Sei $W := L\{(1,-2,0,3,1),(1,-2,1,2,-2)\} \subseteq \mathbf{R}^5$ gegeben.

    a) Geben Sie ein homogenes lineares Gleichungssystem an, das W als Lösungsmenge besitzt!

    b) Finden Sie ein inhomogenes lineares Gleichungssystem, das die Ebene $H^2 = \xi + W$ als Lösungsmenge besitzt!

39. Man gebe die Gleichung einer Ebene H in $\mathbf{R}^3$ an, die die Punkte $A = (5,1,3)$ und $B = (1,6,2)$ enthält, und parallel zur Geraden $G(P,Q)$ mit $P = (5,0,4), Q = (4,0,6)$ verläuft!

æ

Kapitel 5

Linearformen

Die Menge aller linearer Abbildung $\mathrm{Hom}(V, W)$ eines Vektorraums V in einen Vektorraum W ist selbst ein Vektorraum, speziell ist $V^* = \mathrm{Hom}(V, R)$ 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 $\{l_1, \ldots, l_k\}$ 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 $\sum r_i l_i(v) = o$ gilt, so ist $r_i = 0$ für $i = 1, \ldots, k$.

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

Beweis: Sei $V = L(v_1, \ldots, v_k) \oplus U$ 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. □

Folgerung 5.2 *Zu einer gegebenen Basis $B = \{b_1, \ldots, b_n\}$ von V existiert eine Basis $\{l_1, \ldots, l_n\}$ von V^* mit $l_i(b_j) = \delta_{ij}$, sie heißt die zu B duale Basis.* □

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 $v = \sum l_i(v) v_i$.

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

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

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

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

und für $L \subseteq V^*$ setzen wir

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

Lemma 5.3 Ann*(M) ist ein Unterraum von V^*,* Ann*(L) ist ein Unterraum von V.*

Lemma 5.4 $\dim \mathrm{Ann}(M) = \dim V - \dim L(M)$.

Beweis: Wir wählen eine Basis $\{v_1, \dots, v_m\}$ von $L(M)$ 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 $\mathrm{Ann}(M)$, 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 $\mathrm{Ann}(M)$, wenn $l \in L(v_{m+1}^*, \dots, v_n^*)$ ist. □

Lemma 5.5 Ann*(*Ann*(M)) = L(M).*

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

Satz 5.6 $\mathrm{Ann}(U + W) = \mathrm{Ann}\ (U) \cap \mathrm{Ann}\ (W)$, $\mathrm{Ann}(U \cap W) = \mathrm{Ann}\ (U) + \mathrm{Ann}\ (W)$, *wobei $U, W \subseteq V$ Teilräume sind.* □

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

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

Lemma 5.7 *Seien $f : V \to W, g : W \to U$ lineare Abbildungen und $f^* : W \to V^*$, $g^* : U^* \to W^*$ die dualen Abbildungen. Dann gilt $(g \circ f)^* = f^* \circ g^*$.*

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

Satz 5.8 *Sei $f : V \to W$ eine lineare Abbildung und f^* 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 $\mathrm{Ker}(f^*)$ und $\mathrm{Im}(f^*)$: Genau dann ist $l \in \mathrm{Ker}(f^*)$, wenn $f^*(l) = lf = 0$, also wenn $l(f(v)) = 0$ für alle $v \in V$ gilt, also ist $\mathrm{Ker}(f^*) = \mathrm{Ann}(\mathrm{Im}(f))$. Speziell: Wenn $\mathrm{Ker}(f^*) = \{0\}$ ist, so gilt $\mathrm{Im}(f) = W$.
Sei weiter $k \in \mathrm{Im}(f^*)$, 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 $\mathrm{Ker}(f)$ liegt, so ist $k(v) = 0$, folglich ist $\mathrm{Ker}(f)$ in $\mathrm{Ann}(\mathrm{Im}(f^*))$ enthalten. Schließlich ist

$$\dim\ \mathrm{Ann}(\mathrm{Im}(f^*)) = \dim V^* - \dim\ \mathrm{Im}(f^*) = \dim V^* - (\dim W^* - \dim\ \mathrm{Ker}(f^*))$$

$$= \dim V - \dim W + \dim\ \mathrm{Ann}(\mathrm{Im}(f)) = \dim V - \dim\ \mathrm{Im}(f) = \dim\ \mathrm{Ker}(f)$$

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

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

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

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

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

also ist

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

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

Satz 5.9 $A_{C^*B^*}(f^*) = (A_{BC}(f))^T,\ (AB)^T = B^T A^T,\ (A+B)^T = A^T + B^T.$ □

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

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

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 $v \neq o$ mit $i(v) = o$, 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 $\{v_1^*, \ldots, v_n^*\}$, 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 $b : V \times V \to R$, 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. $V = R$, $b : R \times R \to R$ sei die Multiplikation. Die Bilinearität ist durch das Distributivgesetz gesichert.
2. $V = R^n$, $b((x_1, \ldots, x_n), (y_1, \ldots, y_n)) = \sum x_i y_i$.

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

Beweis: Sei $f : V \to V^*$ 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 $f : V \to V^*$ 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 $\mathrm{Hom}(V, V^*)$ und aus der obigen Definition der Operationen mit Bilinearformen ergibt sich die Linearität von H. □

Es ist

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

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

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

Lemma 6.2 *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 $\mathrm{Ker}(H(b))$. □

Beispiele:
1. $V = R^2$, $b((x_1,x_2),\ (y_1,y_2)) = x_1y_1 + x_2y_2$, 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. $V = R^2$, $b((x_1,x_2),(y_1,y_2)) = (x_1+x_2)(y_1+y_2)$, dies ist eine ausgeartete Bilinearform.

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

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

dies sei die bezüglich C zu b gehörige Darstellungsmatrix. Wenn nun $v = \sum r_i v_i, w = \sum s_i v_i$, 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

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

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 6.3 $M_C(b) = A_{C,C^*}(H(b))^T$.

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

Folgerung 6.4 *Die Bilinearform b ist genau dann nichtausgeartet, wenn $M_C(b)$ regulär ist.*

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

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

Satz 6.5 *Seien $C = \{v_1, ..., v_n\}$ und $B = \{w_1, ..., w_n\}$ 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:

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

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

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

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

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

Lemma 6.6 *Zu einer symmetrischen Bilinearform b gehört bezüglich jeder Basis von V eine symmetrische Matrix.* □

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

Satz 6.7 *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 $b(u, u) = b(v, v) = 0$ 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 $\{v_1 = v, v_2, \ldots, v_n\}$ von V ergänzen, daß $b(v_1, v_i) = 0$ für $i > 1$ ist. Bezüglich dieser Basis gehört dann zu b die Matrix

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

und das Problem wäre auf ein kleineres reduziert.
Sei also $\{v_1, w_2, \ldots, w_n\}$ (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

$$[1 \quad 0 \quad \ldots \quad 0] M \begin{pmatrix} r_1 \\ \cdots \\ r_n \end{pmatrix} = 0$$

ist. Dies ist eine Gleichung mit n Unbekannten, der Lösungsraum hat also die Dimension $n - 1$. Sei $\{v_2, \ldots, v_n\}$ eine Basis des Lösungsraums. Dann ist $b(v_1, v_i) = 0$ für $i > 1$. Wir zeigen, daß $\{v_1, \ldots, v_n\}$ linear unabhängig ist.
Sei also $\sum r_i v_i = o$, dann gilt

$$0 = b(v_1, \sum r_i v_i) = r_1 b(v_1, v_1) + r_2 b(v_1, v_2) + \ldots + r_n b(v_1, v_n),$$

die letzten $n - 1$ Summanden sind null und $b(v_1, v_1)$ ist nicht null, es folgt $r_1 = 0$. Da $\{v_2, \ldots, v_n\}$ bereits linear unabhängig waren, sind auch die übrigen r_i null. Bezüglich der Basis $\{v_1, \ldots, v_n\}$ hat also b eine Darstellungsmatrix der Form

$$\begin{pmatrix} * & 0 & \ldots & 0 \\ 0 & ? & \ldots & ? \\ 0 & & \ldots & \\ 0 & & & \end{pmatrix}$$

und diese Form hat sie auch bezüglich jeder Basis $\{v_1, w_2, \ldots, w_n\}$, wenn nur die $w_i \in L(v_2, \ldots, v_n)$ sind, denn es ist $b(v_1, w_i) = 0$. Wir schränken nun die Bilinearform b auf den Unterraum $L(v_2, \ldots, v_n)$ zu b' ein, in diesem Vektorraum gibt es nach Induktionsvoraussetzung eine Basis (oBdA sei es bereits $\{v_2, \ldots v_n\}$), so daß die Darstellungsmatrix von b' Diagonalgestalt hat. Bezüglich der Basis $\{v_1, \ldots v_n\}$ hat dann die Bilinearform b eine Diagonalmatrix als Darstellungsmatrix. □

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 $b(e_1, e_1) = b(e_2, b_2) = 0$ und $b(e_1 + e_2, e_1 + e_2) = 2$. Also setzen wir $v = [1, 1]$. Wir suchen a, b mit $[1 \quad 1] \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = [1 \quad 1] \begin{pmatrix} a \\ b \end{pmatrix} = 0$ 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:

Satz 6.8 *Zur Bilinearform b auf V gibt es eine Basis, bezüglich derer die Darstellungsmatrix von b die Form*

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

hat.

Beweis: Sei $\{v_1, \dots, v_n\}$ eine Basis von V, bezüglich derer die Darstellungsmatrix von b Diagonalform hat. Nach einer eventuellen Änderung der Reihenfolge der Basisvektoren sind etwa die ersten Diagonalelemente positiv, die folgenden negativ und die letzten null, also

$$b(v_i, v_i) = \begin{cases} d_i^2 & \text{für} \quad 1 \le i \le s, \\ -d_i^2 & \text{für} \quad s+1 \le i \le r \\ 0 & \text{für} \quad i > r. \end{cases}$$

Nun gehen wir zur Basis $\{\frac{1}{d_1}v_1, \dots, \frac{1}{d_r}v_r, v_{r+1}, \dots, v_n\}$über und haben hier die gewünschte Darstellungsmatrix. □

Satz 6.9 (Trägheitssatz von Sylvester) *Sei*

$$b(v_i, v_i) = \begin{cases} 1 & \textit{für} \quad 1 \le i \le s, \\ -1 & \textit{für} \quad s+1 \le i \le r \\ 0 & \textit{für} \quad i > r. \end{cases}$$

Dann sind die Zahlen s und r von der Wahl der Basis unabhängig.

Beweis: Sei $\{w_1, \dots, w_n\}$ eine Basis mit $b(w_i, w_j) = 0$ für $i \neq j$ und $b(w_i, w_i) = 1$ für $1 \le i \le t$, $b(w_i, w_i) = -1$ für $t+1 \le i \le r$ sowie $b(w_i, w_i) = 0$ für $i > r$.
(In beiden Fällen steht dasselbe r, dies ist der Rang der Darstellungsmatrix, der natürlich von der Basis unabhängig ist.) Es genügt zu zeigen, daß $s \le t$ ist. Wir zeigen, daß die Menge $\{v_1, \dots, v_s, w_{t+1}, \dots, w_n\}$ linear unabhängig ist, dann folgt $s+n-t \le n$, also $s \le t$. Sei also

$$\sum_{i=1}^{s} r_i v_i - \sum_{j=t}^{n} s_j w_j = 0$$

Dann ist

$$\sum r_i v_i = \sum s_j w_j,$$

also

$$b(\sum r_i v_i, \sum r_i v_i) = \sum r_i^2 = b(\sum s_j w_j, \sum s_j w_j) = -\sum s_j^2,$$

diese Zahl ist sowohl nichtnegativ als auch nichtpositiv, also ist $r_1 = \ldots = s_n = 0$. $\square$

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

$$E^{ij} = \begin{pmatrix} 1 & & & \\ & \ldots & & \\ & & 1 & \\ & & & 1 \end{pmatrix}$$

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 vergangenen Jahrhundert) stammt das folgende Iterationsverfahren zur Überführung einer symmetrischen Matrix in eine Diagonalform.
Es sei eine symmetrische Matrix A gegeben, wir betrachten Matrizen

$$J_{ij}(w) = \begin{pmatrix} 1 & & & & & \\ & & \ldots & & & \\ & c & & s & & \\ & & \ldots & & & \\ & -s & & c & & \\ & & \ldots & & & \\ & & & & & 1 \end{pmatrix}$$

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

$$A \to B = J_{ij}(w)^T A J_{ij}(w)$$

durchführen, und zwar suchen wir die Drehmatrizen so aus, daß in jedem Schritt die Zahl

$$\operatorname{off}(A) = \sum_{i \neq j} a_{ij}^2$$

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), ..., (1,n),

(2,3), ..., $(n-1,n)$ Nullen entstehen. Daß eine solche Wahl gelingt, wollen wir uns im Fall von 2×2-Matrizen verdeutlichen.

Wir berechnen

$$\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} c & -s \\ s & c \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} c & s \\ -s & c \end{pmatrix}$$

$$= \begin{pmatrix} c^2 a_{11} - cs(a_{21} + a_{12}) + s^2 a_{22} & c^2 a_{21} + cs(a_{11} - a_{22}) - s^2 a_{12} \\ sc(a_{11} - a_{22}) - s^2 a_{21} + c^2 a_{12} & s^2 a_{11} + sc(a_{12} + a_{21}) + c^2 s_{22} \end{pmatrix}$$

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 $2x - \dfrac{1}{t} + t = 0$ oder $t^2 + 2xt - 1 = 0$ und damit $t = -x \pm \sqrt{x^2+1}$, also $c = \dfrac{1}{\sqrt{1+t^2}}$, $s = tc$.
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 6.10 *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

$$J^T A = \begin{pmatrix} c & -s & & & \\ s & c & & & \\ & & 1 & & \\ & & & \dots & \\ & & & & 1 \end{pmatrix} \begin{pmatrix} a_{11} & & & & \\ a_{21} & & & & \\ & \dots & & & \\ & & 1 & & \\ & & & & \dots \end{pmatrix} = \begin{pmatrix} ca_{11} - sa_{21} & & & \\ sa_{11} + ca_{21} & & & \\ & & \dots & \\ & & & \dots \end{pmatrix}$$

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. □

Wie unterscheiden sich nun off(A) und 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 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

$$\operatorname{off}(B) = \operatorname{off}(A) - 2a_{12}^2 + 2b_{12}^2 = \operatorname{off}(A) - 2a_{12}^2.$$

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 $\operatorname{off}(A)$ um $2a_{pq}^2$ verkleinert.
Es sei $N = \frac{n(n-1}{2}$ die Zahl der Elemente von A oberhalb der Diagonalen. Dann ist $2a_{pq}^2 \geq \frac{\operatorname{off}(A)}{N}$, also, wenn A_k das nach k Schritten erhaltene Ergebnis ist, gilt

$$\operatorname{off}(A_k) = \operatorname{off}(A_{k-1}) - 2a_{pq}^2 \leq (1 - \frac{1}{N})\operatorname{off}(A_{k-1}) \leq (1 - \frac{1}{N})^k \operatorname{off}(A),$$

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:

$$\begin{aligned} c_1 &= a_1b_1 + a_2b_3 \\ c_2 &= a_1b_2 + a_2b_4 \\ c_3 &= a_3b_1 + a_4b_3 \\ c_4 &= a_3b_2 + a_4b_4. \end{aligned}$$

Bezüglich der kanonischen Basis des R^4 entsprechen den c_i die folgenden Matrizen:

$$c_1 : \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} c_2 : \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} c_3 : \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} c_4 : \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Jede der obigen Bilinearformen „enthält“ zwei Produkte, ihre Darstellungsmatrizen haben den Rang 2.

Eine Bilinearform, die nur ein Produkt enthält, hat die Form

$$(r_1a_1 + r_2a_2 + r_3a_3 + r_4a_4)(s_1b_1 + s_2b_2 + s_3b_3 + s_4b_4),$$

ihre Darstellungsmatrix

$$\begin{pmatrix} r_1s_1 & r_2s_1 & r_3s_1 & r_4s_1 \\ r_1s_2 & r_2s_2 & r_3s_2 & r_4s_2 \\ r_1s_3 & r_2s_3 & r_3s_3 & r_4s_3 \\ r_1s_4 & r_2s_4 & r_3s_4 & r_4s_4 \end{pmatrix}$$

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:

$$(a_2 - a_4)(b_3 + b_4),\ (a_1 + a_4)(b_1 + b_4),\ (a_1 - a_3)(b_1 + b_2),$$

$$(a_1 + a_2)b_4,\ a_1(b_2 - b_4),\ a_4(b_3 - b_1),\ (a_3 + a_4)b_1.$$

Es soll dem Leser überlassen bleiben nachzuweisen, wie sich die Bilinearformen $c_1, \ldots, c_4$ aus diesen linear kombinieren lassen.

Zum Abschluß dieses Abschnitts wollen wir unsere Kenntnisse in der Geometrie anwenden.

6.4 Klassifikation der Quadriken

Sei A ein affiner Raum und $\{P, v_1, \ldots, v_n\}$ ein Koordinatensystem, ein Punkt $X \in A$ habe die Koordinaten $[1, x_1, \ldots, x_n]$. 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 $\{P, v_1, \ldots, v_n\}$ ein Koordinatensystem des affinen Raums A. Die Menge aller Punkte X mit den Koordinaten $[1, x_1, \ldots, x_n]$ 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:

$$Q : \begin{pmatrix} 1 & x_1 & \ldots & x_n \end{pmatrix} \begin{pmatrix} a_0 & a_1 & \ldots & a_n \\ a_1 & a_{11} & \ldots & a_{1n} \\ & & \ldots & \\ a_n & a_{n1} & \ldots & a_{nn} \end{pmatrix} \begin{pmatrix} 1 \\ x_1 \\ \ldots \\ x_n \end{pmatrix}$$

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 $X = BX'$ 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, das es eine Basis $\{u_1, \ldots, u_n\}$ 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

$$\begin{pmatrix} a_0 & a_1 & \ldots & a_r & \ldots & a_n \\ a_1 & a_{11} & 0 & & \ldots & a_{1n} \\ & & \ldots & & & \\ a_r & 0 & \ldots & a_{rr} & \ldots & 0 \\ a_n & 0 & & & \ldots & 0 \end{pmatrix}$$

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 $i = 1, \ldots, r$ 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 $i = 1, \ldots, n-1 x_i'' = x_i'$ und

$$x_n'' = x_n' + \frac{a_{r+1}}{a_n} x_{r+1}' + \ldots + \frac{a_{n-1}}{a_n} x_{n-1}' + \frac{a_0'}{a_n}$$

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 $d_1, \ldots$ durch a, b, c und $x_1, \ldots$ 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

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. $F(rv, w) = rF(v, w) = F(v, rw) \qquad (r \in R)$,
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

$$v = r_1e_1 + r_2e_2,\ w = s_1e_1 + s_2e_2,$$

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 $r_1 s_2 - r_2 s_1$ 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 $F : (M_{n1})^n \to R$ heißt Multilinearform, wenn für fixierte $a_j \in M_{n1}$ jede Abbildung

$$F_i(v) = F(a_1, a_2, \ldots, a_{i-1}, v, a_{i+1}, \ldots, a_n) : M_{n1} \to R$$

linear ist. Eine Multilinearform heißt alternierend, wenn $F(a_1, \ldots, a_n) = 0$ ist, falls $\{a_1, \ldots, a_n\}$ 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 $D : M_{nn} \to R$, 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. $F(\ldots, a, \ldots, a, \ldots) = 0$,
2. Beim Vertauschen von Spalten ändert sich das Vorzeichen:
   $0 = F(\ldots, a+b, \ldots, a+b, \ldots) =$
   $F(\ldots, a, \ldots, a, \ldots) + F(\ldots, a, \ldots, b, \ldots) + F(\ldots, b, \ldots, a, \ldots) + F(\ldots, b, \ldots, b, \ldots)$,
   also gilt
   $F(a_1, \ldots, a_i, \ldots, a_j, \ldots, a_n) = -F(a_1, \ldots, a_j, \ldots, a_i, \ldots, a_n)$.
3. Elementare Spaltenoperationen ändern den Wert nicht:
   $F(a_1 + ra_2, a_2, \ldots, a_n) = F(a_1, a_2, \ldots, a_n) + rF(a_2, a_2, \ldots, a_n)$
   und der zweite Summand ist null.
4. Wenn f eine Multilinearform mit der Eigenschaft $f(\ldots, a \ldots, a \ldots) = 0$ ist, so ist f altenierend, denn wenn von den Vektoren $v_1, \ldots, v_n$ einer eine Linearkombination der übrigen ist, so kann man durch elementare Operationen zwei gleich Vektoren herstellen.

Satz 7.1 *Es gibt eine Funktion $D : M_{nn} \to R$, 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

$$D'(A) = (-1)^{i+1} a_{i1} D(A_{i1}) + (-1)^{i+2} a_{i2} D(A_{i2}) + \ldots + (-1)^{i+n} a_{in} D(A_{in})$$

(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 $a_2, \ldots, a_n$ fest:
In A_{i1} kommt die erste Spalte von A gar nicht vor, also ist $D(A_{i1})$ konstant und die Abbildung

$$A \to (-1)^{i+1} a_{i1} D(A_{i1})$$

ist offenbar linear. Weiter sind $D(A_{i2}), \ldots, D(A_{in})$ nach Induktionsvoraussetzung linear in der ersten Spalte und die Faktoren $a_{i2}, \ldots, a_{in}$ hängen von der ersten Spalte von A nicht ab, also sind auch die Abbildungen

$$A \to (-1)^{i+j} a_{ij} D(A_{ij})$$

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 $D'(E_n) = 0$, wie man leicht sieht. □

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 Begiff:

Definition: Die Menge aller bijektiven Abbildungen der Menge $\{1, \ldots, n\}$ 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 $\{1, \ldots, n\}$ in sich ist wieder eine derartige Abbildung, die identische Abbildung ist bei dieser Multiplikation ein neutrales Element und zu jeder Abbildung $f : \{1, \ldots, n\} \to \{1, \ldots, n\}$ 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 7.2 *Es gibt genau eine Determinante vom Grad n.*

Beweis: Sei $D : M_{nn} \to R$ 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

$$D(A) = D(\sum a_{j(1),1} e_{j(1)}, \ldots, \sum a_{j(n),n} e_{j(n)}) = \sum a_{j(1),1} \ldots a_{j(n),n} D(e_{j(1)}, \ldots, e_{j(n)}),$$

wobei die Summation über alle Indexsysteme $(j(1), \ldots, j(n))$ zu erstrecken ist. Nun ist aber $D(e_{j(1)}, \ldots, e_{j(n)}) = 0$, wenn nicht alle Indizes voneinander verschieden sind, also sind nur die Summanden von Interesse, wo $\{j(1), \ldots, j(n)\} = \{1, \ldots, n\}$ ist, d.h. wo die Zuordnung $k \to j(k)$ eine Permutation ist, also ist

$$D(A) = \sum a_{p(1),1} \ldots a_{p(n),n} D(e_{p(1)}, \ldots, e_{p(n)}),$$

wo über alle Permutationen $p \in S_n$ zu summieren ist. Der Faktor

$$D(e_{p(1)}, \ldots, e_{p(n)})$$

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 $\mathrm{sgn}(p)$ der Permutation p bezeichnet.
Folglich ist

$$D(A) = \sum_{p \in S_n} a_{p(1),1} \ldots a_{p(n),n} \mathrm{sgn}(p),$$

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 7.3 *Die obige Laplacesche Formel ergibt für jeden Zeilenindex i denselben Wert $D(A)$.*

Satz 7.4 *Sei $F : M_{nn} \to R$ eine alternierende Multilinearform, dann gilt*

$$F(A) = \det(A) F(E).$$

Der Beweis verläuft analog. □

Obwohl die Leibnitzsche 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 Leibnitzsche 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 Spaltenoperationen nicht oder (bei Spaltenvertauschungen) nur um das Vorzeichen ändert. Mit Hilfe von etwa n^3 Spaltenoperationen können wir eine Matrix A in eine Dreiecksform überführen:

$$\det(A) = \det(\begin{pmatrix} a_{11} & 0 & \dots & 0 \\ & & \dots & \\ a_{n1} & & \dots & a_{nn} \end{pmatrix}),$$

und wenn wir jetzt einen Blick auf die Leibnitzsche 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 Dreieckmatrix A

$$\det(A) = a_{11} \dots a_{nn}.$$

7.2 Eigenschaften und Anwendungen

Wir beweisen zuerst den

Satz 7.5 (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 $F : M_n \to R$ 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

$$\begin{aligned} \det(AB) &= F(B) \\ &= \det(B)F(E) \\ &= \det(B)\det(Ae_1, \dots, Ae_n) \\ &= \det(B)\det(A) \\ &= \det(A)\det(B). \ \square \end{aligned}$$

Wir betrachten folgenden Spezialfall: Seien p und q Permutationen aus S_n und $A = (e_{p(1)}, \ldots, e_{p(n)})$ sowie $B = (e_{q(1)}, \ldots, e_{q(n)})$ 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 $AB = (e_{pq(1)}, \ldots, e_{pq(n)})$ und wir erhalten die Folgerung

Folgerung 7.6 $\operatorname{sgn}(pq) = \operatorname{sgn}(p)\operatorname{sgn}(q)$, $\operatorname{sgn}(p) = \operatorname{sgn}(p^{-1})$.

Beweis: Dies folgt aus dem Multiplikationssatz. □

Satz 7.7 *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

$$\begin{aligned}
\det(B) &= \sum_{p \in S_n} \operatorname{sgn}(p)\, b_{p(1),1} \ldots b_{p(n),n} \\
&= \sum \operatorname{sgn}(p)\, a_{1,p(1)} \ldots a_{n,p(n)} \\
&= \sum \operatorname{sgn}(p)\, a_{p^{-1}(1),1} \ldots a_{p^{-1}(n),n} \\
&= \sum -p^{-1} \in S_n \operatorname{sgn}(p^{-1})\, a_{p^{-1}(1),1} \ldots a_{p^{-1}(n),n} \\
&= \det(A). \square
\end{aligned}$$

Wir wissen, daß $\det(A) = 0$ gilt, wenn die Spalten von A linear abhängig sind. Gilt aber auch die Umkehrung ?

Satz 7.8 *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

$$\begin{pmatrix} r_1 & 0 \ldots & 0 \\ & \ldots & \\ 0 & \ldots & r_n \end{pmatrix}$$

mit $r_i \neq 0$ bringen, dann ist $\det(A) = r_1 \ldots r_n \neq 0$. □

Satz 7.9 (Cramersche Regel) *Das Gleichungssystem $Ax = b$, genauer*

$$\sum_{j=1}^{n} a_{ij} x_j = b_i, \; i = 1, \ldots, n$$

mit der quadratischen Koeffizientenmatrix A hat genau dann eine eindeutig bestimmte Lösung, wenn $\det(A) \neq 0$ *ist. Diese Lösung ist durch*

$$x_k = \frac{\det(A_k)}{\det(A)},\ k = 1, \ldots, n$$

gegeben, dabei entsteht die Matrix A_k aus A dadurch, daß das n-tupel $\begin{pmatrix} b_1 \\ \ldots \\ b_n \end{pmatrix}$ *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 $(x_1, \ldots, x_n)$ diese Lösung. Dann gilt $\sum a_{ij} x_j = b_i$. Wir betrachten

$$\begin{aligned} \det(A_k) &= \det(a_1, \ldots, a_{k-1}, b, a_{k+1}, \ldots, a_n) \\ &= \det(a_1, \ldots, a_{k-1}, \sum a_j x_j, a_{k+1}, \ldots, a_n) \\ &= \sum \det(a_1, \ldots, a_{k-1}, a_j, a_{k+1}, \ldots, a_n) x_j \\ &= \det(A) x_k, \end{aligned}$$

da $\det(a_1, \ldots, a_{k-1}, a_j, a_{k+1}, \ldots, a_n) = 0$ für $j \neq k$ ist. Damit ist die obige Formel bewiesen. □

Wir wenden uns noch einmal dem Laplaceschen Entwicklungssatz zu:

$$\det(A) = \sum_j (-1)^{i+j} a_{ij} \det(A_{ij}), \tag{1}$$

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“:

$$? = \sum_j (-1)^{k+j} a_{ij} \det(A_{kj}) \text{ mit } k \neq i. \tag{2}$$

Wir können den Wert durch Anwendung der Laplaceschen Formel bestimmen, dies ist doch die Determinante der Matrix, deren k-te Zeile gleich $[a_{i1}, \ldots, a_{in}]$ ist, die nach der k-ten Zeile zu entwicklen 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 7.10 *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 $A = (a_1, \ldots, a_n)$, oBdA sei $\{a_1, \ldots, a_r\}$ linear unabhängig. Wir setzen $B = (a_1, \ldots, a_r)$, 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. □

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 7.11

$$\det(A) = a \cdot \det(-\frac{1}{a}sz + B) = \frac{1}{a^{n-2}} \det(B - sz). \square$$

Satz 7.12 *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 7.13 *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:

$$\begin{pmatrix} 0 & a & & \\ -a & 0 & B & \\ & & & \\ -B^T & & & C \end{pmatrix}.$$

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

$$(B^T S^{-1} B)^T = B^T S^{-1\,T} B = -B^T S^{-1} B,$$

also ist die „Restmatrix“ schiefsymmetrisch und wir erhalten das Resultat durch Induktion.

Satz 7.14 (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

$$(x_2 - x_1) \cdots (x_n - x_1) \det \begin{pmatrix} 1 & x_2 & x_2^2 & \dots & x_2^{n-2} \\ & & \dots & & \\ 1 & x_n & x_n^2 & \dots & x_n^{n-12} \end{pmatrix}$$

und wir erhalten wieder durch Induktion das Resultat. □

Zum Schluß wollen wir noch einem Endomorphismus $f : V \to V$ 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 $\det(f) = \det(M)$.

7.3 Aufgaben

1. Berechnen Sie:

   a) $\det \begin{pmatrix} 1 & 3 & 2 & 1 \\ 3 & 1 & 3 & 0 \\ 1 & 1 & 0 & 1 \\ 2 & 0 & 4 & 2 \end{pmatrix}$

   b) $\det \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \end{pmatrix}$

2. A und B seien 2×2-Matrizen.

   a) Geben Sie ein Paar A, B an, für das gilt: $\det(A + B) = \det(A) + \det(B)$.

   b) Geben Sie ein Paar A, B an, für das gilt: $\det(A + B) \neq \det(A) + \det(B)$.

3. Lösen Sie das Gleichungssystem $A\vec{x} = \vec{b}$ mit Hilfe der Cramerschen Regel:

   $A = \begin{pmatrix} 3 & 2 & 4 \\ 2 & -1 & 1 \\ 1 & 2 & 3 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$

4. * Aus dem Laplaceschen Entwicklungssatz leite man her, daß die Inverse einer regulären 3×3-Matrix A folgendermaßen berechnet wird:

   $A^{-1} = \frac{1}{\det A} \begin{pmatrix} detA_{11} & -detA_{21} & detA_{31} \\ -detA_{12} & detA_{22} & -detA_{32} \\ detA_{13} & -detA_{23} & detA_{33} \end{pmatrix}$

Kapitel 8

Dreidimensionale Geometrie

Zum Beginn wollen wir die Eigenschaften einer speziellen Sorte von 2×2-Matrizen über dem Körper $\mathbf{R}$ der reellen Zahlen untersuchen.
Es sei

$$\mathcal{C} = \{\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \mid a, b \in \mathbb{R}\},$$

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 $\frac{1}{a^2+b^2} \begin{pmatrix} a & -b \\ ba & \end{pmatrix}$ 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 $k : \mathbb{C} \to \mathcal{C}$ 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:

$$\begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix} \begin{pmatrix} \cos(\beta) & \sin(\beta) \\ -\sin(\beta) & \cos(\beta) \end{pmatrix}$$

$$= \begin{pmatrix} \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta) & \cos(\alpha)\sin(\beta) + \sin(\alpha)\cos(\beta) \\ -\cos(\alpha)\sin(\beta) + \sin(\alpha)\cos(\beta) & \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta) \end{pmatrix}$$

$$= \begin{pmatrix} \cos(\alpha+\beta) & \sin(\alpha+\beta) \\ -\sin(\alpha+\beta) & \cos(\alpha+\beta) \end{pmatrix}.$$

Die Konstruktion des Körpers **C** der komplexen Zahlen als Körper von Matrizen kann man wie folgt verallgemeinern:
Es sei

$$\mathcal{H} = \{\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \mid a, b \in \mathbb{C}\}.$$

Satz 8.1 *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 3):

$$\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \begin{pmatrix} c & d \\ -\bar{d} & \bar{c} \end{pmatrix} = \begin{pmatrix} ac - b\bar{d} & ad + b\bar{c} \\ -\bar{b}c - \bar{a}\bar{d} & \bar{a}\bar{c} - \bar{b}\bar{d} \end{pmatrix}. \square$$

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 8.2 *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

$$E^2 = E, \quad I^2 = J^2 = K^2 = -E,$$

$$\begin{array}{rclrclrcl} IJ & = & K, & JK & = & I, & KI & = & J, \\ JI & = & -K, & KJ & = & -I, & IK & = & -J. \end{array}$$

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 bezeichen den von $1, i, j, k$ erzeugten Verktorraum mit

$$\mathbb{H} = \{a + bi + cj + dk \mid a, b, c, d \in \mathbb{R}\},$$

die Elemente von $\mathbb{H}$ heißen Quaternionen. Dieser Schiefkörper wurde von Hamilton 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.

Sei $a = a_1 + a_2 i + a_3 j + a_4 k$ ein Quaternion, dann nennen wir a_1 seinen skalaren Anteil und $a_2 i + a_3 j + a_4 k$ 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 (karthesisches) Koordinatensystem sein.
Wir betrachten nun das Produkt zweier vektorieller Quaternionen $a = a_2 i + a_3 j + a_4 k$ und $b = b_2 i + b_3 j + b_4 k$:

$$(a_2 i + a_3 j + a_4 k)(b_2 i + b_3 j + b_4 k) =$$

$$-(a_2 b_2 + a_3 b_3 + a_4 b_4)$$

$$+(a_3 b_4 - a_4 b_3)i + (a_4 b_2 - a_2 b_4)j + (a_2 b_3 - a_3 b_2)k.$$

Den Ausdruck

$$< a, b >= a_2 b_2 + a_3 b_3 + a_4 b_4$$

nennt man das Skalarprodukt der Vektoren a und b, den Ausdruck

$$a \times b = (a_3 b_4 - a_4 b_3)i + (a_4 b_2 - a_2 b_4)j + (a_2 b_3 - a_3 b_2)k$$

nennt man das Vektorprodukt von a und b . Also gilt

$$ab = - < a, b > + a \times b.$$

Wir werden sofort den Zusammenhang mit den Produkt-Konstruktionen herstellen, die Sie in der Schule kennengelernt haben.
Wenn wir ein Skalarprodukt durch

$$< a, b >=| a || b | \cos(\alpha)$$

einführen, 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 $A = O + a_2 i + a_3 j + a_4 k$ und $B = O + b_2 i + b_3 j + b_4 k$, wir betrachten das Dreieck OAB . Dessen Seiten haben folgende Längen:

$$| OA |= \sqrt{a_2^2 + a_3^2 + a_4^2},$$

$$| OB |= \sqrt{b_2^2 + b_3^2 + b_4^2},$$

$$| AB |= \sqrt{(b_2 - a_2)^2 + (b_3 - a_3)^2 + (b_4 - a_4)^2}.$$

Nach dem Cosinussatz gilt

$$\mid b-a \mid^2 = \mid a \mid^2 + \mid b \mid^2 + 2 \mid a \mid\mid b \mid \cos(\alpha),$$

also

$$(b_2-a_2)^2+(b_3-a_3)^2+(b_4-a_4)^2 =$$
$$a_2^2+a_3^2+a_4^2+b_2^2+b_3^2+b_4^2+2 \mid a \mid\mid b \mid \cos(\alpha)$$

und daraus folgt

$$a_2b_2+a_3b_3+a_4b_4 = \mid a \mid\mid b \mid \cos(\alpha).$$

Wie man leicht nachrechnet, hat das Skalarprodukt folgende Eigenschaften:

1. $< a+b, c > = < a, c > + < b, c >$,
2. $< ra, b > = r < a, b >$ $(r \in \mathbb{R})$,
3. $< a, b > = < b, a >$,
4. $\mid a \mid = \sqrt{< a, a >}$,
5. $(a, b > = 0$ gdw. $a \perp b$.

Das Vektorprodukt

$$a \times b = (a_3b_4 - a_4b_3)i + (a_4b_2 - a_2b_4)j + (a_2b_3 - a_3b_2)k$$

kann formal als Deteminante 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. $(a+rb) \times c = a \times c + rb \times c$ $(r \in \mathbb{R})$,
2. $a \times b = -b \times a$,
3. $a \times b = o$ gdw. $\{a, b\}$ ist linear abhängig,
4. Der Vektor $a \times b$ steht senkrecht auf a und auf b.

   Beweis: Wegen ab $= - < a, b > + a \times b$ folgt $a \times b =$ ab $+ < a, b >$ und speziell $a^2 = - \mid a \mid^2$, also

   $$a(a \times b) = a(\text{ab} + < a, b >) = a^2b + < a, b > a = - \mid a \mid^2 b + < a, b > a,$$

   dies ist ein vektorielles Quaternion, folglich ist das Skalarprodukt (der skalare Anteil des Produkts) von a und $a \times b$ gleich Null:

   $$< a, a \times b > = 0.$$

5. Der Betrag des Vektors $a \times b$ ist gleich dem Flächeninhalt des Parallelogramms, das durch a und b aufgespannt wird.

   Beweis: Es ist

$$\begin{aligned} \mid a \times b \mid^2 &= (a_3b_4 - a_4b_3)^2 + (a_4b_2 - a_2b_4)^2 + (a_2b_3 - a_3b_2)^2 \\ &= (a_2^2 + a_3^2 + a_4^2)(b_2^2 + b_3^2 + b_4^2) - (a_2b_2 + a_3b_3 + a_4b_4)^2 \\ &= \mid a \mid^2 \mid b \mid^2 - < a, b >^2 \\ &= \mid a \mid^2 \mid b \mid^2 - \mid a \mid^2 \mid b \mid^2 \cos^2(\alpha) \\ &= \mid a \mid^2 \mid b \mid^2 \sin^2(\alpha). \end{aligned}$$

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 Verbindugsvektor $\overrightarrow{PP_1}$ verwenden, so gilt

$$< n, \overrightarrow{PP_1} >= 0 f\ddot{u}r \text{ alle } P \in E,$$

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

$$< n_1, \overrightarrow{PP_0} >= 0 \text{ bzw. } < n_2, PP_0 >= 0,$$

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

$$P = P_0 + n_1 \times n_2 \cdot t, \ t \in \mathbb{R}.$$

Der Abstand eines Punkts P_1 von einer Geraden, die durch eine Parameterdarstellung

$$P = P_0 + a \cdot t, \ t \in \mathbb{R}$$

gegeben ist, ist gleich der Höhe im von den Vektoren a und $b = P_0P_1$ aufgespannten Parallelogramms, also gleich $\mid b \mid \sin(\alpha)$ oder gleich

$$\mid a \times b \mid / \mid a \mid .$$

Für die Multiplikation von Quaternionen gilt das Assoziativgesetz. Nun seien speziell a, b, c vektorielle Quaternionen, dann gilt

$$a(bc) = -a < b, c > +a(b \times c) = -a < b, c > - < a, b \times c > +a \times (b \times c),$$

$$(ab)c = - < a, b > c + (a \times b)c = - < a, b > c - < a \times b, c > +(a \times b) \times c.$$

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 8.3 *Die Vektoren a, b, c liegen genau dann in einer Ebene, wenn $< a, b \times c >= 0$ ist.* □

Wenn wir die vektoriellen Teile der Produkte betrachten, erkennen wir, daß das Vektorprodukt nicht assoziativ ist. Vielmehr gilt die sogenannte Jacobi-Identität

$$a \times (b \times c) + b \times (c \times a) + c \times (a \times b) = o.$$

æ

Kapitel 9

Eigenwerte und Eigenvektoren

Sei $f : V \to V$ 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 9.1 *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 $r \in R$. □

Satz 9.2 *Seien $z_1, \ldots z_m$ paarweise verschiedene Eigenwerte von f und $v_1, \ldots, v_m$ zugehörige Eigenvektoren, dann ist $\{v_1, \ldots, v_m\}$ 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 $v_m = r_1 v_1 + \ldots + r_{m-1} v_{m-1}$ ist und wenden f an:

$$\begin{aligned} f(v_m) &= z_m v_m \\ &= z_m r_1 v_1 + \ldots + z_m r_{m-1} v_{m-1} \\ &= f(r_1 v_1 + \ldots + r_{m-1} v_{m-1}) \\ &= z_1 r_1 v_1 + \ldots + z_{m-1} r_{m-1} v_{m-1}. \end{aligned}$$

Aus der linearen Unabhängigkeit von $\{v_1, \ldots, v_{m-1}\}$ folgt $z_i = z_m$ für $i = 1, \ldots, m-1$, ein Widerspruch. □

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 $v = (v_1, \ldots, v_n)^T \neq o$ 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 $v_1, \ldots, v_n$, 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:

$$c_A(z) = (-1)^n z^n + (-1)^{n-1} c_1 z^{n-1} + \ldots + c_n.$$

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 $c_1 = a_{11} + a_{22} + \ldots + a_{nn}$. Die Summe der Diagonalelemente von A, also c_1, heißt die Spur $\mathrm{Sp}(A)$ von A.

Sei nun $f : V \to V$ 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 9.3 *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

$$\begin{aligned} c_{F'}(z) &= \det(X^{-1}FX - zE) \\ &= \det(X^{-1}(F - zE)X) \\ &= \det(X^{-1})\det(t - zE)\det(X) \\ &= \det(X)^{-1}\det(X)c_F(z) \\ &= c_F(z). \ \square \end{aligned}$$

Folgerung 9.4 *$Sp(X^{-1}AX) = Sp(A)$.* $\square$

Das folgende Lemma ist leicht zu beweisen, folgt aber nicht aus der obigen Folgerung.

Lemma 9.5 *Für beliebige (nicht notwendig reguläre) Matrizen A, B gilt*

$$\mathrm{Sp}(AB) = \mathrm{Sp}(BA). \square$$

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 $I = \{i_1, \ldots, i_{n-t}\}$ und $K = \{k_1, \ldots, k_{n-t}\}$ 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 9.6 *Sei $c_A(z) = (-1)^n z^n + (-1)^{n-1} c_1 z^{n-1} + \ldots + c_n$. Dann ist c_i die Summe der i-Hauptminoren von A.*

Beweis: Wir halten uns an die Leibnitzsche 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 $(a_{jj} - z)$, $j = k_1, \ldots, k_i$ auf der Diagonalen wählen, für die restlichen Faktoren dürfen wir dann die Zeilen und die Spalten mit den Nummern $k_1, \ldots, k_i$ 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. □

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 9.7 *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 $(z_1, \ldots z_n)$ mit komplexen Komponenten, die nicht alle gleich Null sind, so daß

$$(A - (a+bi)E) \begin{pmatrix} z_1 \\ \ldots \\ z_n \end{pmatrix} = \begin{pmatrix} 0 \\ \ldots \\ 0 \end{pmatrix}$$

ist. Sei $z_k = x_k + iy_k,\ x_k, y_k$ reell, dann gilt

$$\sum a_{kl} z_l - (a + bi) z_k = 0$$

$$= \sum a_{kl}(x_l + iy_l) - (ax_k - by_k) - (bx_k + ay_k)i.$$

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

$$\sum (\sum a_{kl} x_l y_k - ax_k y_k + by_k^2) = 0$$

und

$$\sum (\sum a_{kl} y_l x_k - bx_k^2 - ax_k y_k) = 0.$$

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. □

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

$$A \to X^T A X$$

(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

$$J = \begin{pmatrix} 1 & & & & & \\ & & \cdots & & & \\ & c & & s & & \\ & & \cdots & & & \\ & -s & & c & & \\ & & \cdots & & & \\ & & & & & 1 \end{pmatrix}$$

transformiert, und die Matrix J hat die angenehme Eigenschaft, daß $J^T = J^{-1}$ ist, d.h. die Eigenwerte von A und von $J^T A J$ stimmen überein. Somit haben wir mit dem Jacobischen Verfahren ein Näherungsverfahren zur Berechnung der Eigenwerte symmetrischer Matrizen gefunden.

Sei also A eine $n \times n$-Matrix mit den Eigenwerten z_i und zugehörigen Eigenvektoren v_i. Wir wissen: Wenn die z_i alle voneinander verschieden sind, so ist $\{v_1, \ldots, v_n\}$ eine linear unabhängige Menge, also eine Basis des R^n. Sei $A \in M_{nn}$ eine Matrix mit den Eigenwerten $z_1, \ldots, z_n$ und zugehörigen Eigenvektoren $v_1, \ldots, v_n$, also

$$Av_i = z_i v_i.$$

Wir wissen, daß $\{v_1, \ldots, v_n\}$ linear unabhängig sind, wenn die z_i paarweise verschieden sind, also:

Satz 9.8 *Wenn $A \in M_{nn}$ lauter verschiedene Eigenwerte hat, so besitzt R^n eine Basis aus Eigenvektoren von A.*

Diese Bedingung ist aber nicht notwendig, wie wir an folgendem Beipiel sehen: Sei

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

ihr charkteristisches 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 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 9.9 *Zur $n \times n$-Matrix A existiert genau dann eine Basis des 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 $(v_1, \ldots, v_n)$, dann gilt

$$AV = A(v_1, \ldots, v_n) = (Av_1, \ldots, Av_n) = (v_1, \ldots, v_n) \begin{pmatrix} z_1 & & 0 \\ & \ldots & \\ 0 & & z_n \end{pmatrix} = (z_1 v_1, \ldots, z_n v_n),$$

also $Av_i = z_i v_i$, also sind die Vektoren $v_i, \ldots, v_n$ Eigenvektoren von A, und als Spalten einer invertierbaren Matrix sind sie linear unabhängig. □
Allgemeiner gilt der folgende

Satz 9.10 *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 ($v_1 \neq o$). 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

$$AX = A(v_1, \dots, v_n) = (Av_1, \dots, Av_n) = (z_1 v_1, Av_2, \dots, Av_n),$$

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 $U = L(v_2, v_3)$ operierend auffassen, d.h. wir suchen *dort* eine Basis, so daß diese Matrix in Dreicksgestalt 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 9.11 *1. Sei* $X^{-1}AX = \begin{pmatrix} r_1 & & * \\ & \dots & \\ 0 & & r_n \end{pmatrix}$ *eine Dreieckmatrix. 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 Dreieckmatrizen 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 9.12 (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 „$\mid A - A \mid = 0$“ formulierte.

Beweis: Seien $z_1, \ldots, z_n$ 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

$$(A - zE)^{-1} = \frac{1}{\det(A - zE)} \begin{pmatrix} b_{11} & \ldots & b_{n1} \\ & \ldots & \\ b_{1n} & \ldots & b_{nn} \end{pmatrix},$$

die b_i sind Minoren von $A - zE$. Wir setzen

$$B = \begin{pmatrix} b_{11} & \ldots & b_{n1} \\ & \ldots & \\ b_{1n} & \ldots & b_{nn} \end{pmatrix} = B_{n-1}z^{n-1} + \ldots + B_1 z + B_0,$$

dabei sollen die B_i von z unabhängige Matrizen sein. Es gilt also

$$\det(A - zE)E = (A - zE)B$$

oder ausführlicher

$$(z^n + b_1 z^{n-1} + b_2 z^{n-2} + \ldots + b_n)E = (A - zE)(B_{n-1}z^{n-1} + \ldots + B_1 z + B_0).$$

Wir vergleichen die Koeffizienten von z^i und erhalten

$$\begin{aligned} b_n E &= AB_0 \\ b_{n-1}E &= AB_1 - B_0 \\ b_{n-2}E &= AB_2 - B_1 \\ &\ldots \\ b_1 E &= AB_{n-1} - B_{n-2} \\ E &= -B_{n-1}. \end{aligned}$$

Wir multiplizieren die Gleichungen von links mit $E, A, A^2, \ldots A^{n-1}, A^n$ und addieren alles:

$$A^n + b_1 A^{n-1} + \ldots + b_{n-1}A + b_n E = 0E. \square$$

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 $w_0 \in \mathbf{R}^n$ 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:

$\mathbf{R}^n$ besitze eine Basis $\{b_1, \ldots, b_n\}$ aus Eigenvektoren von A, die zugehörigen Eigenwerte seien $z_1, \ldots, z_n$ und es sei $w_0 = \sum r_i b_i$. Dann hat $A - xE$ die Eigenwerte $z_1 - x, \ldots, z_n - x$ und $(A - xE)^{-1}$ hat die Eigenwerte $\frac{1}{z_1 - x}, \ldots, \frac{1}{z_n - x}$, 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.

9.1 Aufgaben

1. $f : \mathbf{R}^2 \to \mathbf{R}^2$ mit $f(x, y) = (3x + 3y, x + 5y)$ ist eine lineare Abbildung. Geben Sie alle Eigenwerte von f an und zu jedem Eigenraum eine Basis. Entscheiden Sie, ob f diagonalisierbar ist!

2. Zeigen Sie, daß die folgende Matrix zu einer Diagonalmatrix ähnlich ist und geben Sie eine solche Diagonalmatrix an!
$$A = \begin{pmatrix} 1 & 2 & 2 \\ 1 & 2 & -1 \\ 0 & -1 & 4 \end{pmatrix}$$

3. Geben Sie das charakteristische Polynom der folgenden Matrix an und berechnen Sie die Eigenwerte von A! Ist A diagonalisierbar?
$$A = \begin{pmatrix} 5 & 2 & 2 \\ 2 & 2 & 1 \\ 2 & 1 & 2 \end{pmatrix}$$

4. $^*f$ und g seien lineare Abbildungen von V in V. Beweisen Sie: $f \circ g$ und $g \circ f$ besitzen dieselben Eigenwerte.

5. Ist $A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & -1 \\ -1 & 2 & 3 \end{pmatrix}$ regulär? Wenn ja, berechnen Sie A^{-1}!

6. a) Berechnen Sie die Eigenwerte und die zugehörigen Eigenräume der Matrix A!

   b) Ist die Matrix A diagonalisierbar? Begründen Sie Ihre Antwort.
$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

7. Sei $f : V \longrightarrow V$ eine lineare Abbildung derart, daß ein Vektor $\vec{x} \in V, \vec{x} \neq \vec{o}$ mit $f(-\vec{x}) = r\vec{x}$ existiert. Welche der folgenden Aussagen ist/sind dann wahr?

   (1) -$\vec{x}$ ist Eigenvektor von f zum Eigenwert r

   (2) $\vec{x}$ ist Eigenvektor zum Eigenwert $-r$

   (3) - $\vec{x}$ ist Eigenvektor von f zum Eigenwert $-r$

   Begründen Sie Ihre Antwort!

8. Ist die Matrix $A = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix}$ diagonalisierbar? Begründen Sie Ihre Antwort und geben Sie gegebenenfalls eine zu A ähnliche Diagonalmatrix an!

9. Seien V und W endlich-dimensionale Vektorräume über einem Körper K, und $\varphi \in \mathrm{Hom}(V), \psi \in \mathrm{Hom}(W)$ diagonalisierbare Endomorphismen. Man zeige, daß die durch $f(a) := \psi \circ a \circ \varphi$ definierte lineare Abbildung $f : \mathrm{Hom}(V, W) \to \mathrm{Hom}(V, W)$ ebenfalls diagonalisierbar ist!

10. Sei $f \in \mathrm{Hom}(V)$, mit $\dim V < \infty$. Sei λ eine k - fache Nullstelle des charakteristischen Polynoms von f, und sei E_λ der zu λ gehörende Eigenunterraum. Man zeige, daß $\dim E_\lambda \leq k$ gilt!

11. Sei V der Vektorraum der reellen Polynome vom Grad ≤ 3, und $\varphi : V \to V$ gegeben durch $\varphi(f) = \frac{d}{d}\bar{x}[(x+3) \cdot f], f \in V$. Zeigen Sie, daß φ linear ist, und bestimmen Sie die Eigenwerte und Eigenvektoren von φ!

12. Sei $a \in \mathrm{End}(\mathbf{R}^4)$ bezüglich der Standardbasis von $\mathbf{R}^4$ durch die Matrix

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

gegeben. Ist a diagonalisierbar ? Wenn ja, so gebe man in $\mathbf{R}^4$ eine Basis an, die aus Eigenvektoren von a besteht, und bestimme die zugehörigen Eigenwerte!

13. Man überprüfe, ob die Matrizen $A = \begin{pmatrix} -1 & 3 & -1 \\ -3 & 5 & -1 \\ -3 & 3 & 1 \end{pmatrix}$ und $B = \frac{1}{2}\begin{pmatrix} 3 & -1 & 1 \\ -1 & 3 & 1 \\ 0 & 0 & 4 \end{pmatrix}$ ähnlich sind. Falls dies zutrifft, so gebe man eine Matrix $T \in GL(3, \mathbf{R})$ mit $A = T \circ B \circ T^{-1}$ an!

Kapitel 10

Polynome

Wir verlassen für ein paar Augenblicke die lineare Algebra und stellen uns einige Hilfsmittel zusammen, die wir später verwenden werden.

Ein Term der Form $f(z) = a_0 z^n + a_1 z^{n-1} + \ldots + a_n$ heißt ein Polynom in z, die Zahlen $a_0, \ldots, a_n$ heißen die Koeffizienten des Polynoms. Die Summe $f + g$ zweier Polynome f und g ist wieder ein Polynom in z, ebenso ihr Produkt fg. Wenn in der obigen Darstellung der Koeffizient a_0 von Null verschieden ist, so sagt man, das Polynom f habe den Grad n, als Abkürzung schreiben wir $\deg(f) = n$. Die Menge alle Polynome in z mit Koeffizienten aus R bezeichnet man mit $R[z]$.

Sei $\deg(f) = n, \deg(g) = m$. Bitte überlegen Sie sich, daß $\deg(f + g) \leq \max(m, n)$ und $\deg(fg) = n + m$ ist.

Das Polynom g heißt Teiler des Polynoms f, wenn ein Polynom h existiert, so daß $f = gh$ ist.

Wenn $r \neq 0$ eine Zahl (also ein Polynom vom Grade 0) ist, so gibt es immer ein Polynom h mit $f = rh$. Gerade diesen trivialen Fall wollen wir immer ausschließen, wenn wir von Teilbarkeit sprechen.

Zum Beispiel hat das Polynom $f(z) = z^2 + pz + q$ die (nichtkonstanten) Teiler

$$z + p/2 \pm (p^2/4 - q)^{1/2}$$

.

Nicht zu unterschätzen ist das folgende Lemma über die Division mit Rest:

Lemma 10.1 *Seien f, g Polynome, $g \neq 0$. Dann gibt es Polynome q und r, so daß $f = gq + r$ gilt und entweder $r = 0$ oder $\deg(r) < \deg(g)$ ist.*

Beweis: Wenn $\deg(g) > \deg(f)$ ist, so setzen wir $q = 0$ und $r = f$. Weiterhin sei $\deg(f) >= \deg(g)$. 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 das Lemma für Polynome von einem Grad, der kleiner als n ist, bewiesen ist. Sei also

$$f(z) = z^n + a_1 z^{n-1} + \dots, \; g(z) = z^m + b_1 z^{m-1} + \dots,$$

dann setzen wir $q_1(z) = z^{n-m}$, es ist

$$q_1(z)g(z) = x^n + b_1 x^{n-1} + \dots$$

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 wenden das Resultat für ein spezielles Polynom an: Sei

$$g(z) = z - a$$

ein Polynom vom Grad 1 und $f(z)$ ein beliebiges Polynom. Wir dividieren mit Rest:

$$f(z) = q(z)(z - a) + r,$$

dabei muß der Grad von r kleiner als 1 sein, d.h. r ist ein konstantes Polynom, also eine Zahl. Wenn wir in dieser Gleichung $z = a$ „einsetzen“, erhalten wir $f(a) = r$. Also: Wenn $f(a) = 0$ ist, so ist $z - a$ ein Teiler von $f(z)$.

Definition: Seien f_1 und f_2 Polynome. Ein Polynom d heißt größter gemeinsamer Teiler von f_1 und f_2, wenn gilt:
1. d ist ein Teiler von f_1 und von f_2 (also ein gemeinsamer Teiler),
2. wenn h irgendein gemeinsamer Teiler von f_1 und f_2 ist, so ist h ein Teiler von d.

Als Abkürzung schreiben wir $d = \mathrm{ggT}(f_1, f_2)$.

Trivial ist das

Lemma 10.2 *Der größte gemeinsame Teiler zweier Polynome ist bis auf einen konstanten Faktor eindeutig bestimmt.*

Beweis: Die Polynome d_1 und d_2 mögen die Bedingungen der Definition erfüllen. Aus 2. folgt, daß Polynome p und q existieren, so daß $d_1 = pd_2$ und $d_2 = qd_1$. Daraus folgt $\deg(p) = \deg(q) = 0$. □

Zur Berechnung des größten gemeinsamen Teilers zweier Polynome benutzen wir den Euklidischen Algorithmus:

Seien f_1, f_2 gegeben, wir dividieren fortlaufend mit Rest, bis die Division aufgeht:

$$f_1 = q_1 f_2 + f_3$$

$$f_2 = q_2 f_3 + f_4$$
$$f_3 = q_3 f_4 + f_5$$
$$\dots$$
$$f_{m-3} = q_{m-3} f_{m-2} + f_{m-1}$$
$$f_{m-2} = q_{m-2} f_{m-1}$$

Wegen $\mathrm{def}(f_2) > \deg(f_3) > \deg(f_4) > \dots$ muß nach endlich vielen Schritten ein Rest gleich Null sein, hier ist es f_m.
Behauptung: $\mathrm{ggT}(f_1, f_2) = f_{m-1}$.
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! Das Polynom 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 das Polynom f_{m-1} teilt. □

Lemma 10.3 *Sei $d = ggT(f_1, f_2)$, dann gibt es Polynome 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: Das Polynom 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. □

Interessanter ist das

Lemma 10.4 *Der größte gemeinsame Teiler von f_1 und f_2 ist das (normierte) Polynom d von minimalem Grad, für das Polynome g_1 und g_2 existieren, so daß $f_1 g_1 + f_2 g_2 = d$ ist.*

Beweis: Sei $d = f_1 g_1 + f_2 g_2$ und $\deg(d)$ minimal.
1. Wir dividieren mit Rest:

$$f_1 = q_1 d + r_1 = q_1 g_1 f_1 + q_1 g_2 f_2 + r_1,$$

also

$$r_1 = f_1(1 - q_1 g_1) - f_2 q_1 q_2,$$

aber wegen $\deg(r_1) < \deg(d)$ ist dies ein Widerspruch zur Minimalität des Grades von d.

2. Sei h ein gemeinsamer Teiler von f_1 und f_2, dann ist h auch ein Teiler von $f_1g_1 + f_2g_2 = d$. $\square$

Wir wollen uns nun genauer mit dem Zusammenhang zwischen den Koeffizienten und den Nullstellen eines Polynoms befassen. Sei

$$f(z) = z^n + b_1z^{n-1} + \ldots + b_{n-1}z + b_n = (z - z_1)\ldots(z - z_n)$$

ein Polynom mit den Koeffizienten b_i und den Nullstellen z_i. Wir wissen, daß $f(z)$ durch die Linearfaktoren $z - z_i$ teilbar ist. Wie sehen die Koeffizienten von $\dfrac{f(z)}{z - z_i}$ aus?

Satz 10.5 $\dfrac{\sum_{i=0}^n b_{n-i}z^i}{z - z_1} = \displaystyle\sum_{i=0}^{n-1} z^{n-1-i} \sum_{j=0}^{i} b_j z_1^{i-j}$.

Beweis: Wir multiplizieren $\sum\ z^{n-1-i} \sum b_j z_1^{i-j}$ und $z - z_1$ miteinander und stellen Sie fest, ob $f(z)$ herauskommt.

$$\begin{aligned}
&(\sum_{i=0}^{n-1} z^{n-1-i} \sum_{j=0}^{i} b_j z_1^{i-j}) \cdot (z - z_1) \\
&\quad = \sum_{i=0}^{n-1} z^{n-i} \sum_{j=0}^{i} b_j z_1^{i-j} - \sum_{i=0}^{n-1} z^{n-1-i} \sum_{j=0}^{i} b_j z_1^{i-j+1} \\
&\quad = \sum_{i=0}^{n-1} z^{n-i} \sum_{j=0}^{i} b_j z_1^{i-j} - \sum_{k=0}^{n} z^{n-k} \sum_{j=0}^{k-1} b_j z_1^{k-j} \qquad (k := i+1) \\
&\quad = \sum_{i=1}^{n-1} z^{n-i} \underbrace{(\sum_{j=0}^{i} b_j z_1^{i-j} - \sum_{j=0}^{i-1} b_j z_1^{i-j})}_{= b_i} + z^n + b_n - \sum_{j=0}^{n-1} b_j z_1^{n-j} - b_n \\
&\quad = \sum_{i=0}^{n} z^{n-i} b_i + \sum_{i=0}^{n} z_1^{n-i} b_i \\
&\quad = f(z) - f(z_1) = f(z). \ \square
\end{aligned}$$

Abkürzung: Sei $f(z)$ ein Polynom vom Grade n mit den Nullstellen $z_1, \ldots, z_n$. Wir setzen

$$\begin{aligned}
s_0 &= n, \\
s_1 &= z_1 + \ldots + z_n \\
s_2 &= z_1^2 + \ldots + z_n^2, \\
&\ldots \\
s_i &= z_1^i + \ldots + z_n^i.
\end{aligned}$$

Die Zahl s_i heißt die $i - te$ Potenzsumme der x_j.

Wir stellen nun eine Beziehung zwischen den Potenzsummen der Wurzeln und den Koeffizienten des Polynoms auf, dies sind die sogenannten Newtonschen Formeln.

Satz 10.6 $b_i = -\frac{1}{i}\sum_{j=0}^{i-1} b_j s_{i-j}$

Beweis: Es ist $f(z) = \prod(z - z_i)$. Wir betrachten die Ableitung $f'(z)$ von $f(z)$:

$$f'(z) = \sum_k \prod_{i \neq k}(z - z_i) = \sum \frac{f(z)}{z - z_k}$$

$$= \sum_{k=1}^{n}\sum_{i=0}^{n-1} z^{n-i-1}\sum_{j=0}^{i} b_j z_k^{i-j}$$

$$= \sum_{i=0}^{n-1}\sum_{j=0}^{i} z^{n-i-1} b_j \sum_k z_k^{i-j}$$

$$= \sum_{i=0}^{n-1}\sum_{j=0}^{i} z^{n-i-1} b_j s_{i-j}.$$

Andererseits gilt

$$f'(z) = \sum z^{n-i-1}(n-i)b_i$$

und durch Koeffizientenvergleich erhalten wir

$$(n-i)b_i = \sum_{j=0}^{i} b_j s_{i-j} = \sum_{j=0}^{i-1} b_j s_{i-j} + nb_i,$$

und daraus ergibt sich die Behauptung. □

Wir kehren nun doch nach diesem Seitensprung wieder zu unseren lieben Matrizen zurück. Aber wir wenden das Gelernte an:

Lemma 10.7 *Seien $z_1, \ldots z_n$ die Eigenwerte der Matrix A, dann ist $s_i = Sp(A^i)$.*

Beweis: A^i hat die Eigenwerte $z_1^i, \ldots, z_n^i$ und die Spur einer Matrix ist die Summe ihrer Eigenwerte. □

Nun können wir die Newtonschen Formeln verwenden, um die Koeffizienten des charakteristischen Polynoms einer Matrix zu bestimmen, ohne irgendwelche Determinanten ausrechnen zu müssen.

Folgerung 10.8 *Sei A eine Matrix und $c_A(z) = \sum b_i z^{n-i}$ ihr charakteristisches Polynom. Dann ist $b_i = -\frac{1}{i}\sum_{j=0}^{i-1} b_j Sp(A^{i-j})$.* □

Sei $f(z) = \sum a_i z^{n-i}$ ein normiertes Polynom (also $a_0 = 1$) und sei A eine Matrix, dann setzen wir $f(A) = \sum a_i A^{n-i}$, dies ist wieder eine Matrix.

Wenn $f(A) = 0$ die Nullmatrix ist, so heißt f ein die Matrix A annullierendes Polynom. Wie wir im Satz von Hamilton-Cayley gesehen haben, ist das charakteristische Polynom ein annullierendes Polynom.

Definition: Ein (normiertes) Polynom f mit $f(A) = 0$, das den kleinstmöglichen Grad hat, heißt Minimalpolynom von A.

Lemma 10.9 *Sei $m(z)$ ein Minimalpolynom der Matrix A und $f(z)$ irgendein annullierendes Polynom. Dann ist m ein Teiler von f.*

Beweis: Wir dividieren mit Rest:

$$f(z) = q(z)m(z) + r(z),$$

es ist $r = 0$ oder $\deg(r) < \deg(m)$. Wenn $r = 0$ ist, so folgt die Teilbarkeit. Sonst setzen wir A ein:

$$0 = f(A) = q(A)m(A) + r(A),$$

wegen $m(A) = 0$ folgt $r(A) = 0$, d.h. $r(z)$ wäre ein A annullierendes Polynom mit einem Grad, der kleiner als der von m ist, ein Widerspruch. □

Folgerung 10.10 *Das Minimalpolynom von A ist eindeutig bestimmt.*

Beweis: Wir nehmen an, wir hätten zwei Minimalpolynome. Dann teilen sie sich gegenseitig, und da sie normiert sein sollten, sind sie gleich. □

Folgerung 10.11 *Die Nullstellen des Minimalpolynoms der Matrix A sind Eigenwerte von A.*

Beweis: Das Minimalpolynom teilt das charakteristische Polynom. □

Wir bemerken, daß auch die Umkehrung gilt: Jeder Eigenwert von A ist Nullstelle des Minimalpolynoms, dazu benötigen wir noch ein paar Hilfmittel.

Satz 10.12 *Sei $f(x) = \sum a_i x^i \in R[x]$ ein Polynom P eine invertierbare Matrix und A eine beliebige Matrix. Dann gilt $f(P^{-1}AP) = P^{-1}f(A)P$.*

Beweis: Die Idee ist $(P^{-1}AP)^2 = P^{-1}APP^{-1}AP = P^{-1}A^2P$, also $f(P^{-1}AP) = \sum a_i(P^{-1}AP)^i = \sum a_i P^{-1}A^iP = P^{-1}(\sum a_i A^i)P$. □

Folgerung 10.13 *Wenn $f(A) = 0$ ist, so ist auch $f(P^{-1}AP) = 0$.* □

Satz 10.14 *Wenn $f(A) = 0$ ist, so gilt $f(z_i) = 0$ für alle Eigenwerte z_i von A.*

Beweis: Wir wissen, daß es eine invertierbare Matrix P gibt, so daß $P^{-1}AP = D$ eine obere Dreiecksmatrix ist und daß auf der Diagonalen von D die Eigenwerte z_i von A stehen. Wie wir eben gesehen haben, gilt $f(D) = 0$, man rechnet nun leicht aus, daß auf der Diagonalen von $f(D)$ gerde die Ausdrücke $f(z_i)$ stehen. □

Folgerung 10.15 *Jeder Eigenwert von A ist Nullstelle des Minimalpolynoms von A.*

Wir können dies anwenden:
Wenn A idempotent ist, so ist sein Minimalpolynom ein Teiler von $z^2 - z$, also hat A nur die Eigenwerte 1 und 0.
Wenn A nilpotent ist, so hat sein Minimalpolynom die Form z^k, also ist 0 der einzige Eigenwert.

Wenn A involutiv ist, so ist sein Minimalpolynom von der Form $z^2 - 1$, also kommen nur die Eigenwerte 1 und -1 in Frage.

Wir haben bisher recht naiv mit Polynomen $f(x)$ gerechnet; was ist eigentlich das x? Manchmal nennt man es eine Unbestimmte, aber kann man mit unbestimmten Objekten rechnen?

Wir machen folgende Konstruktion:

$$F = \{(a_0, a_1, \ldots, a_n, 0, \ldots) \mid a_i \in R\}$$

sei die Menge aller endlichen Folgen, d.h. nur endlich viele Glieder dürfen von Null verschieden sein. Wir führen eine Addition und eine Multiplikation in F ein:

$$(a_i) + (b_i) = (a_i + b_i),$$

$$(a_i) \cdot (b_i) = (c_k) = (\sum_{i+j=k} a_i b_j),$$

man sieht leicht, daß die rechten Seiten ebenfalls endliche Folgen sind.
Bezüglich der Addition bildet F eine kommutative Gruppe, auch die Multiplikation ist kommutativ. Wir zeigen die Gültigkiet des Assoziativgesetzes:

$$\begin{aligned} ((a_i)(b_j))(c_l) &= (\sum_{i+j=k} a_i b_j)(c_l) \\ &= \sum_{k+l=p} \sum_{i+j=k} a_i b_j c_l \\ &= \sum_{i+j+l=p} a_i b_j c_l \end{aligned}$$

und diesen symmetrischen Ausdruck können wir in die gewünschte Form überführen.
Bei der Multiplikation ist die Folge $(1, 0, 0, \ldots)$ das neutrale Element und die Elemente der Form $(a, 0, 0, \ldots)$ verhalten sich bei Addition und Multiplikation wie Elemente von R, wir werden also $(a, 0, 0, \ldots)$ mit $a \in R$ identifizieren.
Wir setzen nun

$$x = (0, 1, 0, \ldots),$$

dann ist

$$x \cdot x = (0, 0, 1, 0, \ldots)$$

und

$$x^i = (0, \ldots, 0, 1, 0, \ldots),$$

wo die 1 an der $(i+1)$-ten Position steht. Dann hat jedes Element von F die Form

$$(a_0, a_1, \ldots, a_n, 0, \ldots) = \sum_{i=0}^{n} a_i x^i.$$

Wir können also sagen: Die Unbestimmte x ist die obengenannte Folge und ein Polynom ist ein Element von F.

Für die Festlegung der Rechenoperationen in F war die Forderung nach der Endlichkeit der Folgen eigentlich nicht wesentlich. Wir setzen

$$P = \{(a_0, a_1, \ldots) \mid a_i \in R\}$$

und vereinbaren eine Addition und eine Multiplikation wie soeben. Dann hat jedes Element von P die Gestalt

$$f = (a_i) = \sum_{i=0}^{\infty} a_i x^i,$$

dies nennen wir eine formale Potenzreihe (wir kümmern uns nicht um Konvergenzfragen).

Aufgabe: Betrachten Sie die (formalen) Potenzreihen, die die Sinus- und die Cosinusfunktion darstellen, und zeigen Sie mithilfe der obigen Rechenregeln, daß $\sin^2(x) + \cos^2(x) = 1$ gilt.

Sei $f(x) = \sum a_i x^i$ eine Potenzreihe mit $a_0 \neq 0$; wir zeigen, daß eine Potenzreihe $g(x)$ mit $f(x) \cdot g(x) = 1$ existiert, d.h. $f(x)$ ist invertierbar.
Wir machen einen Ansatz $g(x) = \sum b_j x^j$, dann soll

$$\sum_{i+j=k} a_i b_j = \begin{cases} 0 & \text{für } k > 0 \\ 1 & \text{für } k = 0 \end{cases}$$

gelten. Also muß der Reihe nach $a_0 b_0 = 1,\ a_0 b_1 + a_1 b_0 = 0,\ \ldots$ gelten, und diese Gleichungen sind offenbar lösbar.

Aufgabe: Berechnen Sie $\frac{1}{\cos(x)}$ und daraus $\tan(x) = \frac{\sin(x)}{\cos(x)}$.

Zum Schluß behandeln wir noch das Problem der Interpolation: Zu n gegebenen, verschiedenen Zahlen $x_1, \ldots, x_n$ und gegebenen $y_1, \ldots, y_n$ ist ein Polynom $f(x)$ vom Grad $n-1$ gesucht, so daß $f(x_i) = y_i$ ist.
Dazu sind die folgenden, von Lagrange gefundenen Polynome hilfreich:

$$L_i = \frac{(x - x_1) \cdots (x - x_{i-1})(x - x_{i+1}) \cdots (x - x_n)}{(x_i - x_1) \cdots (x_i - x_{i-1})(x_i - x_{i+1}) \cdots (x_i - x_n)}.$$

Offensichtlich gilt $L_i(x_j) = \delta_{ij}$ und damit ist $f(x) = \sum y_i L_i(x)$ das gesuchte Polynom.

Satz 10.16 $\sum L_i(x) = 1$.

Beweis: Wir wählen $y_i = 1,\ i = 1, \ldots n$. □

Kapitel 11

Normalformen von Matrizen

11.1 Invariante Unterräume

Während dieses gesamten Kapitels sei V ein fixierter Vektorraum und $f : V \to V$ ein Endomorphismus. Alle einzuführenden Begriffe beziehen sich auf diese Situation, auch wenn es später nicht ausdrücklich erwähnt wird.

Definition: Ein Unterraum $U \subseteq V$ heißt invariant (bezüglich f), wenn $f(U) \subseteq U$ gilt.

Sei U_1 ein invarianter Unterraum von V. Wir wählen eine Basis B_1 von U_1 und ergänzen sie durch eine Menge B_2 zu einer Basis B von V. Wie sieht die Darstellungsmatrix $A_{BB}(f)$ aus?

Nun, das Einzige, was wir wissen, ist, das für $b_i \in B_1$ das Bild $f(b_i)$ in $L(B_1)$ liegt, also

$$A_{BB}(f) = \begin{pmatrix} \star & \dots & \star & ? & \dots & ? \\ & \dots & & & & \\ \star & \dots & \star & ? & \dots & ? \\ & \dots & & & & \\ 0 & \dots & 0 & ? & \dots & ? \\ & \dots & & & & \\ 0 & \dots & 0 & ? & \dots & ? \end{pmatrix}$$

Die Darstellungsmatrix besitzt links unten einen Null-Block.
In besonderen Fällen kann es vorkommen, daß die Darstellungsmatrix auch noch rechts oben einen Null-Block besitzt, etwa so:

$$\begin{pmatrix} \star & \dots & \star & 0 & \dots & 0 \\ & \dots & & & & \\ \star & \dots & \star & 0 & \dots & 0 \\ & \dots & & & & \\ 0 & \dots & 0 & ? & \dots & ? \\ & \dots & & & & \\ 0 & \dots & 0 & ? & \dots & ? \end{pmatrix}$$

Das ist dann der Fall, wenn $f(B_2) \subseteq L(B_2)$ ist, d.h. der Unterraum $U_2 = L(B_2)$ ist ebenfalls invariant. Wir bemerken nebenbei, daß V die direkte Summe von U_1 und U_2 ist: $V = U_1 \oplus U_1$.

Definition: Der Vektorraum V heißt (bezüglich f) zerlegbar, wenn invariante Unterräume U_1, U_2 von V existieren, so daß $V = U_1 \oplus U_2$ ist, anderenfalls heißt V unzerlegbar.

Lemma 11.1 *Es gibt unzerlegbare invariante Unterräume $U_1, \ldots, U_r$ mit*

$$V = U_1 \oplus \ldots \oplus U_r.$$

Beweis: Wenn V unzerlegbar ist, so setzen wir $r = 1$ und $U_1 = V$. Wenn V zerlegbar ist, so zerlegen wir V in invariante Unterräume: $V = U_1 \oplus U_2$. Wenn U_1 und U_2 unzerlegbar sind, sind wir fertig, wenn nicht, so zerlegen wir diese Unterräume weiter, bis wir lauter unzerlegbare invariante Summanden erhalten. □

Der folgende Satz hat einen etwas längeren Beweis, ist aber sehr nützlich.

Satz 11.2 *Sei $f : V \to V$ ein Endomorphismus und V unzerlegbar, dann ist $U =$*Im(f) *ein unzerlegbarer invarianter Unterraum.*

Beweis: Die Invarianz ist leicht einzusehen:

$$f(U) = f(f(V)) \subseteq f(V) = U.$$

Wir nehmen an, U wäre in invariante Unterräume zerlegbar:

$$U = U_1 \oplus U_2,$$

sei B_1 eine Basis von U_1 und B_2 eine Basis von U_2. Weiter sei $B_1 = \{b_1, \ldots, b_k\}$, wir wählen Urbilder c_i der b_i:

$$f(c_i) = b_i.$$

Sei $C_1 = B_1 \cup D_1$ eine maximale linear unabhängige Teilmenge von $\{b_1, \ldots, b_k, c_1, \ldots, c_k\}$, die B_1 enthält, sei oBdA $D_1 = \{c_1, \ldots, c_l\}$, also $U_1 = \mathrm{Im}(U_1) \oplus L(b_1, \ldots, b_l)$. Dann ist $V_1 = L(C_1)$ ein invarianter Unterraum, denn $f(c_i) = b_i$ liegt in V_1 und $f(b_i)$ liegt in $U_1 \subseteq V_1$.
Nach Konstruktion gibt es zu jedem Vektor $u \in U_1$ einen Vektor $v \in V_1$ mit $u = f(v)$ und die Menge aller Urbilder von u ist gleich $\{v + x \mid x \in \mathrm{Ker}(f)\}$. Damit die die Menge $f^{-1}(U_1)$ aller Urbilder von Vektoren aus U_1 gleich V_1+Ker(f).
Analog bilden wir zur Basis B_2 die Teilmengen D_2 und C_2 sowie den Unterraum V_2. Es ist wieder $f^{-1}(U_2) = V_2$+Ker(f). Wir überlegen uns, daß die Summe von V_1 und V_2 direkt ist. Dazu ist zu zeigen, daß die Menge $B_1 \cup D_1 \cup B_2 \cup D_2$ linear unabhängig ist. Sei $u_i \in L(B_i) = U_i$ und $d_i \in L(D_i)$, $i = 1, 2$. Wir nehmen an, es gelte

$$u_1 + d_1 + u_2 + d_2 = o,$$

dann liegt

$$f(u_1) + f(d_1) = -f(u_2) - f(d_2)$$

im Durchschnitt von U_1 und U_2, ist also gleich o. Nun ist aber $f(u_1) \in \mathrm{Im}(U_1)$ und $f(d_1) \in L(b_1, \ldots, b_l)$ entsprechend der obigen Zerlegung von U_1, folglich sind beide Summanden null. Nach Wahl von D_1 sind aber die Bilder von $c_1, \ldots, c_l$ linear unabhängig, also folgt aus $f(d_1) = o$ schon $d_1 = o$. Ganz analog zeigt man $d_2 = o$.
Aus der obigen Gleichung bleibt dann

$$u_1 + u_2 = o$$

und daraus folgt $u_1 = u_2 = o$, da die Summe von U_1 und U_2 direkt ist.
Damit ist die Direktheit der Summe von V_1 und V_2 bewiesen. Weiter haben wir

$$V = f^{-1}(U) = V_1 + V_2 + \ \mathrm{Ker}(f).$$

Wir setzen nun $V_1 \oplus V_2 = W$.
Der Durchschnitt von W und $\mathrm{Ker}(f)$ ist ein Unterraum von $\mathrm{Ker}(f)$, sei T ein komplementärer Unterraum:

$$(W \cap \mathrm{Ker}(f)) \oplus T = \mathrm{Ker}(f).$$

Da T in $\mathrm{Ker}(f)$ liegt, ist T ein invarianter Unterraum. Wir zeigen nun

$$V = W \oplus T = V_1 \oplus V_2 \oplus T,$$

was der Unzerlegbarkeit von V widerspricht.
Wir bestimmen zuerst den Durchschnitt von W und T: Der Vektor u liege in W und in T. Wegen $T \subseteq \mathrm{Ker}(f)$ liegt u in $W \cap \mathrm{Ker}(f)$ und in T, ist also gleich o. Also ist die Summe von W und T direkt. Wir berechnen nun die Dimension von $W \oplus T$:

$$\begin{aligned}
\dim T &= \dim\ \mathrm{Ker}(f) - \dim(W \cap\ \mathrm{Ker}(f)), \\
\dim V &= \dim(W +\ \mathrm{Ker}(f)) \\
&= \dim W + \dim\ \mathrm{Ker}(f) - \dim(W \cap\ \mathrm{Ker}(f)) \\
&= \dim W + \dim T.
\end{aligned}$$

Damit ist der Satz bewiesen. □

11.2 Nilpotente Endomorphismen

Definition: Ein Endomorphismus $f : V \to V$ heißt nilpotent vom Grade n, wenn $f^{n-1} \neq o$ und $f^n = o$ ist.

Das folgende Lemma ist trivial.

Lemma 11.3 *Sei $f : V \to V$ nilpotent vom Grade n und $V = V_1 \oplus \ldots \oplus V_r$ eine direkte Summe invarianter Unterräume. Es sei $f_i = f \mid V_i$ die Einschränkung von f auf den Unterraum V_i. Dann ist f_i nilpotent vom Grade $\leq n$.* □

Satz 11.4 *Sei $f : V \to V$ nilpotent vom Grade n, dann ist*

$$\{o\} \subset \ Ker(f) \subset \ Ker(f^2) \subset \ldots \subset \ Ker(f^{n-1}) \subset V$$

und alle Inklusionen sind echt.

Beweis: Wenn $f^i(v) = o$ ist, so ist auch $f^{i+1}(v) = o$, also ist $\mathrm{Ker}(f^i)$ in $\mathrm{Ker}(f^{i+1})$ enthalten.
Wir nehmen an, daß $\mathrm{Ker}(f^i) = \mathrm{Ker}(f^{i+1})$ für ein i gelte. Das heißt, wenn ein Vektor in $\mathrm{Ker}(f^{i+1})$ liegt, so liegt er auch in $\mathrm{Ker}(f^i)$. Nun existiert ein v mit $f^{n-1}(v) \neq o$, dann ist

$$o = f^n(v) = f^{i+1}(f^{n-i-1}(v))$$

und damit

$$f^i(f^{n-i-1}(v)) = f^{n-1}(v) = o,$$

ein Widerspruch. □

Satz 11.5 *Sei f nilpotent vom Grade n und $f^{n-1}(v) \neq o$, dann ist $\{v, f(v), \ldots, f^{n-1}(v)\}$ linear unabhängig.*

Beweis: Es ist $f^i(v) \in \mathrm{Ker}(f^{n-i})$, denn $f^{n-i}(f^i(v)) = f^n(v) = o$, aber $f^i(v)$ liegt nicht in $\mathrm{Ker}(f^{n-i-1})$, wie man schnell nachrechnet. Sei nun

$$r_0 v + r_1 f(v) + \ldots + r_{n-1} f^{n-1}(v) = o$$

und es sei i die kleinste Zahl, so daß $r_i \neq 0$ ist (also $r_0 = \ldots = r_{i-1} = 0$). Dann ist

$$-r_i f^i(v) = r_{i-1} f^{i+1}(v) + \ldots + r_{n-1} f^{n-1}(v),$$

die Summanden auf der rechten Seite liegen aber alle in $\mathrm{Ker}(f^{n-i-1})$, ein Widerspruch.
□

Satz 11.6 *Sei $f : V \to V$ nilpotent vom Grade n und V unzerlegbar. Dann gibt es einen Vektor $v \in V$, so daß $\{v, f(v), \ldots, f^{n-1}(v)\}$ eine Basis von V ist (insbesondere ist $\dim V = n$).*

Beweis: Wir führen die Induktion über n.
Wenn $n = 1$ ist, so heißt das $f = o$. Bezüglich der Nullabbildung ist jeder Unterraum von V invariant, da V unzerlegbar sein sollte, muß $\dim V = 1$ sein und die Behauptung ist gezeigt.
Sei der Satz nun für beliebige Abbildungen vom Nilpotenzgrad höchstens $n-1$ bewiesen. Wir betrachten den invarianten Unterraum $U = \mathrm{Im}(f)$, dieser ist ebenfalls unzerlegbar, wie wir gesehen haben.
Sei $u \in U$, wir wählen ein v mit $u = f(v)$, dann ist $f^{n-1}(u) = f^n(v) = o$, also hat die Einschränkung von f auf U höchstens den Nilpotenzgrad $n-1$. Weiter gibt es ein $v \in V$ mit $f^{n-1}(v) \neq o$, also ist $f^{n-1}(f(v)) \neq o$, demnach hat $f \mid U$ genau den Nilpotenzgrad $n-1$. Nach Induktionsvoraussetzung besitzt U eine Basis

$$\{u, f(u), \ldots, f^{n-2}(u)\}$$

für ein gewisses $u \in U$. Wir wählen wieder ein Urbild v von u, dann ist $f^i(u) = f^{i+1}(v)$. Wir betrachten nun die Menge

$$\{v, f(v), \ldots, f^{n-1}(v)\}$$

und zeigen ihre lineare Unabhängigkeit.
Da die letzten $n-1$ Vektoren nach Voraussetzung linear unabhängig sind, untersuchen wir, ob v eine Linearkombination dieser Vektoren ist:

$$v = \sum r_i f^i(v).$$

Dann ist aber

$$u = f(v) = \sum r_i f^i(u)$$

im Widerspruch zur Voraussetzung.
Es bleibt zu zeigen, daß $\dim V = n$ ist, oder daß $\dim \operatorname{Ker}(f) = 1$ ist.
Wir wissen, daß $f^{n-1}(v)$ im Kern von f liegt, wir nehmen an, es gäbe dort noch einen zweiten, von diesem linear unabhängigen Vektor w. Es sind zwei Fälle möglich:
1. w liegt in $\operatorname{Im}(f)$, dann ist

$$w = r_1 f(v) + \ldots + r_{n-1} f^{n-1}(v)$$

für gewisse r_i. Daraus folgt

$$o = f(w) = r_1 f^2(v) + \ldots + r_{n-2} f^{n-2}(v) + o,$$

also $r_1 = \ldots = r_{n-2} = 0$, also $w = r_{n-1} f^{n-1}(v)$ im Widerspruch zur Voraussetzung.
2. w liegt nicht in $\operatorname{Im}(f)$, dann finden wir eine Basis von V von der Form

$$\{w, w_2, \ldots, w_m, v, f(v), \ldots, f^{n-1}(v)\}.$$

Sei

$$U = L\{w_2, \ldots, w_m, v, f(v), \ldots, f^{n-1}(v)\}.$$

Für alle w_i liegt $f(w_i)$ in $\operatorname{Im}(f)$, also in U, damit ist U ein invarianter Unterraum und $L(w) \oplus U = V$ im Widerspruch zur Unzerlegbarkeit von V. Damit ist der Satz bewiesen. □

Folgerung 11.7 *Sei $f : V \to V$ nilpotent, dann gibt es Vektoren $v_1, \ldots, v_k \in V$, so daß*

$$\{v_1, f(v_1), \ldots, f^{n(1)-1}(v_1), \ldots, v_k, \ldots, f^{n(k)-1}(v_k)\}$$

eine Basis von V ist.

Beweis: Wir zerlegen V in eine direkte Summe unzerlegbarar invarianter Unterräume und wenden den obigen Satz an. □

Wenn man die gewünschte Basis tatsächlich bestimmen muß, ist das angegebene Verfahren wenig hilfreich, denn wie soll man die unzerlegbaren Unterräume bestimmen.
Man kann wie folgt vorgehen: Wir wählen eine Basis von $\operatorname{Ker}(f)$ und bestimmen Urbilder der Basisvektoren (in $\operatorname{Ker}(f^2)$), diese Urbilder sind linear unabhängig, wir ergänzen sie zu einer Basis von $\operatorname{Ker}(f^2)$ und verfahren mit dieser Basis analog.

Nun übertragen wir dies auf Matrizen, indem wir sie als Darstellungsmatrizen von Endomorphismen auffassen:

Folgerung 11.8 *Sei A eine nilpotente Matrix, dann gibt es eine reguläre Matrix X, so daß $X^{-1}AX$ eine Block-Diagonalmatrix ist:*

$$X^{-1}AX = \begin{pmatrix} A_1 & & 0 \\ & \dots & \\ 0 & & A_k \end{pmatrix},$$

und die A_i sind n_i-reihige Matrizen der Form

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

Beispiel: $A = \begin{pmatrix} -4 & 2 & 3 \\ -6 & 3 & 5 \\ -2 & 1 & 1 \end{pmatrix}$, $A^2 = \begin{pmatrix} -2 & 1 & 1 \\ -4 & 2 & 2 \\ 0 & 0 & 0 \end{pmatrix}$, $A^3 = 0$. Der Vektor $v = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ liegt nicht in $\mathrm{Ker}(f^2)$, es ist $Av = \begin{pmatrix} -4 \\ -6 \\ -2 \end{pmatrix}$, $A^2v = \begin{pmatrix} -2 \\ -4 \\ 0 \end{pmatrix}$, und mit $X = \begin{pmatrix} 1 & -4 & -2 \\ 0 & -6 & -4 \\ 0 & -2 & 0 \end{pmatrix}$ haben wir $X^{-1}AX = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$.

11.3 Jordansche Normalform

Definition: Eine Matrix der Form

$$J(z) = \begin{pmatrix} z & & & \dots & 0 \\ 1 & z & & \dots & 0 \\ 0 & 1 & z & \dots & 0 \\ & & & \dots & \\ 0 & \dots & 0 & 1 & z \end{pmatrix}$$

heißt Jordankästchen. Wir sagen, eine Matrix liegt in Jordanscher Normalform vor, wenn sie eine Blockdiagonalmatrix ist, deren Diagonalblöcke Jordankästchen sind:

$$\begin{pmatrix} J(z_1) & \dots & 0 \\ & \dots & \\ 0 & \dots & J(z_k) \end{pmatrix}$$

Die Eigenwerte einer Matrix in Jordanscher Normalform sind offenbar die $z_1, \dots, z_k$, die Eigenwerte in verschiedenen Jordankästchen müssen nicht voneinander verschieden sein, z.B. sind für dreireihige Matrizen folgende Jordanschen Normalformen möglich:

$$\begin{pmatrix} x & & \\ & y & \\ & & z \end{pmatrix}, \begin{pmatrix} x & & \\ & x & \\ & & y \end{pmatrix}, \begin{pmatrix} x & & \\ 1 & x & \\ & & y \end{pmatrix}, \begin{pmatrix} x & & \\ & x & \\ & & x \end{pmatrix}, \begin{pmatrix} x & & \\ 1 & x & \\ & & x \end{pmatrix}, \begin{pmatrix} x & & \\ 1 & x & \\ & 1 & x \end{pmatrix}.$$

Wir werden sehen, daß zu jeder Matrix A eine reguläre Matrix X existiert, so daß $X^{-1}AX$ Jordansche Normalform besitzt. Dies folgt aus dem

Satz 11.9 *Sei $f : V \to V$ ein Endomorphismus, dann gibt es eine Basis B von V, so daß $A_{BB}(f)$ Jordansche Normalform besitzt.*

Beweis: Sei z ein Eigenwert von f, dann ist $(f - z\,id)$ nicht injektiv, wir setzen $g = f - z\,id$. Es gilt $\mathrm{Ker}(g) \neq \{o\}$ und es sei

$$\{o\} \subset \mathrm{Ker}(g) \subset \mathrm{Ker}(g^2) \subset \mathrm{Ker}(g^3) \subset \ldots \subset \mathrm{Ker}(g^m) = \mathrm{Ker}(g^{m+1}),$$

die ersten Inklusionen seien alle echt, wir überlegen uns, daß die Kerne der noch höheren Potenzen von g alle übereinstimmen: Sei $g^{m+2}(v) = o$, dann ist $g^{m+1}(g(v)) = o$, also auch $g^m(g(v)) = g^{m+1}(v) = o$, usw.
Wir setzen nun $U_1 = \mathrm{Ker}(g^m)$ und $U_2 = \mathrm{Im}(g^m)$.
Behauptung: $V = U_1 \oplus U_2$. In der Tat: Sei $v \in U_1 \cap U_2$, also $v = g^m(w)$, dann ist $o = g^m(v) = g^{2m}(w)$, also liegt w in $\mathrm{Ker}(g^{2m}) = \mathrm{Ker}(g^m)$, also ist $v = g^m(w) = o$.
Andererseits gilt $\dim V = \dim \mathrm{Im}(g^m) + \dim \mathrm{Ker}(g^m)$, also ist V die direkte Summe von U_1 und U_2.
Man sieht leicht, daß U_1 und U_2 invariante Unterräume sind und daß die Einschränkung von g auf U_1 nilpotent vom Grade m ist.
Wir wenden nun unsere Kenntnisse über nilpotente Abbildung an: Wir zerlegen U_1 in eine direkte Summe unzerlegbarer invarianter Unterräume, oBdA können wir annehmen, daß U_1 selbst schon unzerlegbar ist. Also gibt es eine Basis B von U_1, die folgende Gestalt hat

$$B = \{v, g(v), \ldots, g^{m-1}(v)\}.$$

Wie wirkt nun f auf diese Basis? Es ist $f = g + z\,id$.

$$\begin{aligned}
f(v) &= g(v) + zv \\
f(g(v)) &= g^2(v) + zg(v) \\
&\ldots \\
f(g^{m-2}(v)) &= g^{m-1}(v) + zg^{m-2}(v) \\
f(g^{m-1}(v)) &= o + zg^{m-1}(v).
\end{aligned}$$

Die Darstellungsmatrix der Einschränkung von f auf U_1 ist also ein Jordankästchen.
Nun schränken wir f auf U_2 ein und beginnen von vorn. Wir suchen einen Eigenwert, bilden ein neues g, stellen fest, wo sich die aufsteigende Folge der Kerne der Potenzen von g stabilisiert usw. Damit ist der Satz bewiesen. □

Sei nun A eine Matrix und $J = X^{-1}AX$ ihre Jordansche Normalform. Sei $m(z)$ das Minimalpolynom von A, dann ist $m(z)$ auch das Minimalpolynom von J, wie man sich schnell überlegt.

Wir betrachten ein Beispiel:

$$J = \begin{pmatrix} J(z_1) & \\ & J(z_2) \end{pmatrix}$$

Das Kästchen zum Eigenwert z_1 habe p Reihen, das zum Eigenwert z_2 habe q Reihen. Das Minimalpolynom $m(z)$ kann nur die Nullstellen z_1 und z_2 haben, wie wir gesehen haben, wir müssen noch ihre Vielfachkeit erraten. Wir wollen es nicht zu spannend machen, das Minimalpolynom ist in unserem Beispiel

$$m(z) = (z - z_1)^p (z - z_2)^q,$$

rechnen Sie es nach! Dann haben Sie die Beweisidee der

Folgerung 11.10 *Die Matrix A habe die verschiedenen Eigenwerte $z_1, \ldots, z_l$, das größte Jordankästchen zum Eigenwert z_i habe p_i Reihen. Dann ist $\prod (z - z_i)^{p_i}$ das Minimalpolynom von A.* □

Folgerung 11.11 *Jeder Eigenwert von A ist Nullstelle des Minimalpolynoms von A.*

Beweis: Die Eigenwerte von A und $X^{-1}AX$ stimmen überein, also muß in der Jordanschen Normalform von A zu jedem Eigenwert mindestens ein Kästchen vorhanden sein. □

Wir können nun die Frage beantworten, wann es zu einer Matrix A eine reguläre Matrix X gibt, so daß $X^{-1}AX$ eine Diagonalmatrix ist, oder, was dasselbe heißt, ob der Vektorraum R^n eine Basis aus Eigenvektoren von A besitzt.

Satz 11.12 *Die Matrix A ist genau dann diagonalisierbar, wenn ihr Minimalpolynom nur einfache Nullstellen besitzt, d.h.*

$$m(z) = \prod (z - z_i), \; z_i \neq z_j \text{ für } i \neq j.$$

Beweis: In diesem Fall haben alle Jordankästchen der Normalform von A die Größe 1, d.h. die Jordansche Normalform ist eine Diagonalmatrix. □

Beispiel:
Sei A eine idempotente Matrix, also $A^2 = A$, das Minimalpolynom von A ist $m(z) = (z - 1)z$, hat also einfache Nullstellen, also ist A diagonalisierbar.

Folgerung 11.13 *Sei A eine idempotente Matrix, dann ist $rg(A) = Sp(A)$.*

Beweis: Der Rang einer diagonalisierbaren Matrix ist gleich der Differenz der Reihenzahl und der Vielfachheit des Eigenwerts 0. Da alle Eigenwerte von A gleich 0 oder 1 sind, ist der Rang gleich der Spur. □

11.4 Rekursive Folgen

Wir betrachten eine Folge $(x_n)_{n \geq 0}$, deren Glieder einer Rekursionsformel genügt:

$$x_n = a_1 x_{n-1} + \ldots + a_k x_{n-k}.$$

Alle Glieder sind eindeutig bestimmt, wenn die Anfangsglieder $x_0, x_1, \ldots, x_{k-1}$ gegeben sind. Das Problem besteht darin, eine explizite Formel für x_n zu finden, so daß man

etwa x_{1000} sofort ausrechnen kann und nicht vorher x_{999}, x_{998} usw. kennen muß. Wir schreiben die Formel als Matrixprodukt:

$$\begin{pmatrix} x_n \\ x_{n-1} \\ \dots \\ \dots \\ x_{n-k+1} \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & \dots & & a_k \\ 1 & 0 & & & 0 \\ & 1 & 0 & & \\ & & & \dots & \\ & & & 1 & 0 \end{pmatrix} \begin{pmatrix} x_{n-1} \\ x_{n-2} \\ \dots \\ \dots \\ x_{n-k} \end{pmatrix}$$

Den Vektor auf der linken Seite nennen wir X_n, auf der rechten Seite steht dann X_{n-1}, multipliziert mit der Matrix A. Der Vektor X_{k-1} enthält die Anfangswerte $x_0, \dots, x_{k-1}$. Dann gilt also

$$X_n = AX_{n-1} = A^i X_{n-i} = A^{n-k+1} X_{k-1}.$$

Die erste Zeile dieser Formel berechnet unser x_n, jedoch ist dies eigentlich keine explizite Formel, da für A^i keine explizite Formel bekannt ist, man muß eine Potenz nach der anderen ausrechnen (wenn man genauer hinsieht, stellt man fest, daß man zur Berechnung der i-ten Potenz nicht i Rechenoperationen, sondern nur $\log(i)$ durchführen muß). Nichtsdestoweniger ist es überhaupt nicht trivial, z.B. für $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^n$ eine explizite Formel zu finden.

Bei Jordankästchen jedoch ist das Potenzieren ein Kinderspiel (Mit $\binom{n}{i}$ bezeichnen wir den Binomialkoeffizienten „n über i“):

Lemma 11.14

$$\begin{pmatrix} z & & & \dots & 0 \\ 1 & z & & \dots & 0 \\ 0 & 1 & z & \dots & 0 \\ & & & \dots & \\ 0 & \dots & 0 & 1 & z \end{pmatrix}^n = \begin{pmatrix} z^n & & & \dots & 0 \\ \binom{n}{1} z^{n-1} & z^n & & \dots & 0 \\ \binom{n}{2} z^{n-2} & \binom{n}{1} z^{n-1} & z^n & \dots & 0 \\ & & & \dots & \end{pmatrix}$$

Den Induktionsbeweis überlassen wir dem Leser. □

Wir kehren zu unserer Folge zurück. Sei J die Jordansche Normalform von A und $A = YJY^{-1}$, dann ist

$$X_n = A^{n-k+1} X_{k-1} = (YJY^{-1})^{n-k+1} X_{k-1} = YJ^{n-k+1} Y^{-1} X_{k-1},$$

also ist X_n ein Vektor, in dem Linearkombinationen der Potenzen der Eigenwerte von A stehen, damit ist eine explizite Formel für x_n gegeben.

Wir müssen uns aber nicht die Mühe machen, die Jordansche Normalform von A zu bestimmen, sondern wir können für x_n einen Ansatz $x_n = \sum \binom{n}{j} b_{ij} z_i^j$ machen, dabei sind die z_i die Eigenwerte von A und die b_{ij} bestimmt man aus den Anfangswerten der Folge. Wenn alle Eigenwerte von A paarweise verschieden sind, so ist J eine Diagonalmatrix und es reicht der Ansatz $x^n = \sum b_{ij} z_i^n$.
Nun ist aber A nicht irgendeine Matrix, es ist nicht schwierig, ihr charakteristisches Polynom zu bestimmen:

Lemma 11.15

$$c_A(z) = (-1)^k \sum (-a_i) z^{k-i} \quad (a_0 = 1).$$

Also einfacher gehts wirklich nicht. Den Beweis überlassen wir allerdings wieder dem Leser. □

Beipiele:

$$x_n = 2x_{n-1} - x_{n-2}$$

$$A = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}, \; c_A(z) = z^2 - 2z + 1, \text{ Eigenwerte 1,1;}$$

Transformation in Jordansche Normalform:

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

$$\begin{pmatrix} x_n \\ x_{n-1} \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_{n-1} \\ x_{n-2} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^n \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_0 \end{pmatrix}$$

$$= \begin{pmatrix} n(x_1 - x_0) + x_0 \\ (n-1)(x_1 - x_0) + x_0 \end{pmatrix}$$

also ist $x_n = n(x_1 - x_0) + x_0$.

Fibonacci-Zahlen

$x_n = x_{n-1} + x_{n-2}$, Anfangswerte $x_0 = 0, x_1 = 1$. Das charakteristische Polynom ist $z^2 - z - 1$ und hat die einfachen Nullstellen $z_i = \frac{1 \pm \sqrt{5}}{2}$.
Wir machen den Ansatz $x_n = az_1^n + bz_2^n$ und bestimmen a und b aus den Anfangswerten zu $a = \frac{1}{\sqrt{5}} = -b$.

11.5 Lineare Differentialgleichungssysteme

Wir setzen hier ein paar Vorkenntnisse aus der Analysis voraus. Diesen Abschnitt behandeln wir nicht in der Vorlesung, er ist für späteres Nachschlagen gedacht.

Sei $y(x)$ eine differenzierbare Funktion und $a(y)$ eine gegebene Funtion, dann heißt eine Gleichung der Form $y'(x) = a(y(x))$ eine Differentialgleichung.

Wenn n^2 Funktionen $a_{ij}(y)$ gegeben und n Funktionen $y_1(x), \ldots, y_n(x)$ gesucht sind, so daß

$$y_1' = a_{11}(y_1) + \ldots + a_{1n}(y_n)$$

$$\ldots$$

$$y_n' = a_{n1}(y_1) + \ldots + a_{nn}(y_n)$$

gilt, so nennen wir diese Bedingungen ein „lineares homogenes Differentialgleichungssystem 1. Ordnung“. Wir schreiben es auch kurz in der Form $y' = Ay$.

Lemma 11.16 *Die Menge aller Lösungen von $y' = Ay$ bildet einen Vektorraum.* □

Ein „Anfangswertproblem“ besteht darin, eine Lösung zu finden, so daß $y_i(0) = c_i,\ (i = 1, \ldots, n)$ für ein gegebenes n-tupel $c \in \mathbf{R^n}$ gilt.

Im allereinfachsten Fall ist $n = 1,\ a(y) = y$, die Differentialgleichung $y' = y$ hat die Lösung $y(x) = ce^x$.
Der nächsteinfachste Spezialfall ist $y' = ay,\ \ a \in \mathbf{R}$, hier haben wir die Lösung $y = e^{ax}$. Wir werden im folgenden versuchen, eine Exponentialfunktion für Matrizen einzuführen. Dazu überlegen wir zuerst, wie man die Matrix einer Differentialgleichung transformieren kann. Wir beschränken uns auf den Spezialfall, daß A eine konstante Matrix ist.

Sei $y = (y_1, \ldots, y_n)$, dann setzen wir $y' = (y_1', \ldots, y_n')$. Weiter sei M eine reguläre Matrix, dann besteht $w = My$ aus Linearkombinationen der y_i, und da die Ableitung eine lineare Abbildung ist, gilt $w_i' = (\sum m_{ij} y_j)' = \sum m_{ij} y_j'$, also $(My)' = My'$. Also gilt

$$y' = Ay \quad \text{gdw.} \quad M^{-1}w' = AM^{-1}w \quad \text{gdw.} \quad w' = MAM^{-1}w,$$

Wir können also ohne Beschränkung der Allgemeinheit annehmen, daß die Koeffizientenmatrix A in Jordanschr Normalform vorliegt.
Das Differntialgleichungssystem

$$y' = \begin{pmatrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_k \end{pmatrix}$$

zerfällt nun in k voneinander unabhängigen Differentialgleichungen, so daß wir nur noch den Fall zu untersuchen haben, wo A ein Jordan-Block ist.

Die gewöhnliche Exponentialfunktion ist durch

$$e^x = \sum \frac{1}{i!} x^i$$

gegeben, diese Reihe konvergiert für alle $x \in \mathbf{R}$. Wir definieren: Eine Matrixfolge $C^{(k)} = (c_{ij}^{(k)}$ konvergiert, wenn alle Folgen $c_{ij}^{(k)}$ konvergieren, und der Grenzwert der Matrixfolge sei die Matrix der Grenzwerte.
Wir definieren nun

$$e^C = \sum \frac{1}{i!} C^i$$

und setzen zunächst voraus, daß diese Reihe konvergiert.
Dann gilt

$$e^{M^{-1}CM} = \sum \frac{1}{i!} (M^{-1}CM)^i = M^{-1}(\sum \frac{1}{i!} C^i)M = M^{-1}e^C M,$$

es genügt also, die Exponentialfunktion für Jordan-Blöcke zu berechnen. Sei also nun

$$C = \begin{pmatrix} z & & & \\ 1 & z & & \\ & & \ddots & \\ & & 1 & z \end{pmatrix}$$

Lemma 11.17 *Die Matrixreihe e^C konvergiert.*

Beweis: In der Matrixreihe steht an der Stelle $(k+l+1, l)$ die Summe

$$\sum \frac{1}{i!}\binom{i}{k} z^{i-k} = \sum \frac{1}{i!}\frac{i!}{k!(i-k)!} z^{i-k} = \frac{1}{k!}\sum \frac{1}{(i-k)!} z^{i-k} = \frac{1}{k!} e^z,$$

und diese Summe existiert. □

In unserem Fall ist $C = Ax$ und C^i hat die Komponenten $\binom{i}{k} z^{i-k} x^i$, also hat e^{Ax} die Komponenten

$$\sum \frac{1}{i!}\binom{i}{k} z^{i-k} x^i = \frac{1}{k!}\sum \frac{1}{(i-k)!} z^{i-k} x^i = \frac{1}{k!} x^k \sum \frac{1}{(i-k)!} (zx)^{i-k} = \frac{1}{k!} x^k e^{zx}.$$

Satz 11.18 *Für jedes n-tupel $c \in \mathbf{R^n}$ ist $y(x) = e^{Ax}c$ eine Lösung von $y' = Ay$.*

Beweis: Wir berechnen $(e^{Ax}c)'$:

$$(\sum \frac{1}{k!} x^k e^{zx} c_l)' = \sum \frac{1}{k!}(kx^{k-1}e^{zx}c_l + zx^k e^{zx} c_l) = \sum (\frac{1}{(k-1)!} x^{k-1} e^{zx} + \frac{1}{k!} z x^k e^{zx}) c_l,$$

dies sind die Komponenten von

$$\begin{pmatrix} 0 & & & \\ 1 & 0 & & \\ & \ddots & & \\ & & 1 & 0 \end{pmatrix} e^A xc + \begin{pmatrix} z & & & \\ & z & & \\ & & \ddots & \\ & & & z \end{pmatrix} e^{Ax} c = Ae^{Ax}c.$$

□

æ °

Kapitel 12

Euklidische Vektorräume

12.1 Skalarprodukt, Orthonormalbases

Definition: Sei $b : V \times V \to R$ eine Bilinearform, b heißt symmetrisch, wenn $b(v, w) = b(w, v)$ gilt, b heißt positiv definit, wenn $b(v, v) \geq 0$ für alle $v \in V$ und $b(v, v) = 0$ nur für $v = o$ gilt. Eine positiv definite symmetrische Bilinearform heißt Skalarprodukt. Zur Abkürzung schreiben wir bei Skalarprodukten $b(v, w) = \langle v, w\rangle$. Ein Vektorraum, in dem ein Skalarprodukt ausgezeichnet ist, heißt Euklidischer Vektorraum. Die Zahl $|v| = \sqrt{\langle v, v\rangle}$ heißt der Betrag des Vektors v.

Wenn wir zum Beispiel im Vektorraum R^2 eine Basis $\{v_1, v_2\}$ wählen, und zwei Vektoren $v = r_1v_1 + r_2v_2,\ w = s_1v_1 + s_2v_2$ gegeben sind, so ist durch $\langle v, w\rangle = r_1s_1 + r_2s_2$ ein Skalarprodukt gegeben.

Eigenschaften des Betrags:
1. $|v| \geq 0$, wenn $|v| = 0$ ist, so ist $v = o$.
2. $|rv| = |r|\,|v|$ für $r \in R$.
3. $|\langle v, w\rangle| \leq |v|\,|w|$ (Cauchy-Schwarzsche Ungleichung)

Beweis: Sei $r \in R$, wir betrachten $u = v + rw$. Es gilt

$$0 \leq |u|^2 = \langle v + rw, v + rw\rangle = \langle v, v\rangle + 2r\langle v, w\rangle + r^2\langle w, w\rangle.$$

Wenn $w = o$ ist, so ist die Behauptung richtig. Nun sei $w \neq o$, wir setzen $r = -\dfrac{\langle v, w\rangle}{|w|^2}$ ein:

$$0 \leq \langle v, v\rangle - 2\frac{\langle v, w\rangle^2}{|w|^2} + \frac{\langle v, w\rangle^2}{|w|^2},$$

also

$$0 \leq |v|^2\,|w|^2 - \langle v, w\rangle^2.$$

4. $|v + w| \leq |v| + |w|$ (Dreiecksungleichung)

Beweis: $|v + w|^2 = \langle v + w, v + w\rangle = \langle v, v\rangle + 2\langle v, w\rangle + \langle w, w\rangle \leq |v|^2 + |w|^2 + 2\,|v|\,|w| = (|v| + |w|)^2$

5. Die Zahl $c = \dfrac{\langle v, w\rangle}{|v|\,|w|}$ liegt zwischen -1 und 1, wir setzen $\cos x = c$ und definieren x als den „Winkel“ zwischen den Vektoren v und w. Dann bedeutet $\langle v, w\rangle = 0$, daß v und w senkrecht aufeinander stehen.

Definition: Eine Menge $\{v_1, \ldots, v_n\}$ heißt ein Orthonormalsystem, wenn $\langle v_i, v_j\rangle = \delta_{ij}$ gilt.

Lemma 12.1 *Orthonormalsysteme sind linear unabhängig.*

Beweis: Sei $\sum s_i v_i = o$, wir multiplizieren skalar mit v_j und erhalten $0 = \sum s_i \langle v_i, v_j\rangle = s_j$. □

Definition: Eine Basis von V, die ein Orthonormalsystem ist, heißt Orthonormalbasis.

Satz 12.2 (Orthonormierungsverfahren von E. Schmidt) *Sei $\{v_1, \ldots v_n\}$ eine Basis von V, dann gibt es eine Orthonormalbasis $\{e_1, \ldots, e_n\}$ von V, so daß*

$$L(e_1) = L(v_1),$$

$$L(e_1, e_2) = L(v_1, v_2),$$

$$\ldots$$

$$L(e_1, \ldots, e_i) = L(v_1, \ldots, v_i),\ (i = 1, \ldots, n).$$

Beweis: Wir setzen $e_1 = \dfrac{v_1}{|v_1|}$, dann ist $|e_1| = 1$ und $L(e_1) = L(v_1)$. Sei $e_1, \ldots, e_{i-1}$ schon konstruiert. Wir machen den Ansatz

$$e_i = r_1 e_1 + \ldots + r_{i-1} e_{i-1} + v_i.$$

Die Bedingungen $\langle e_j, e_i\rangle = 0$ für $j = 1, \ldots, i-1$ dienen zur Berechnung der r_j, indem wir e_i skalar mit e_j multiplizieren:

$$\begin{aligned} 0 &= \langle e_j, e_i\rangle \\ &= \langle e_j, r_1 e_1\rangle + \ldots + \langle e_j, r_{i-1} e_{i-1}\rangle + \langle e_j, v_i\rangle \\ &= r_j + \langle e_j, v_i\rangle, \end{aligned}$$

da e_j schon senkrecht auf $e_1, \ldots, e_{i-1}$ steht. Damit kann r_j berechnet werden. Falls nun $|e_i| \neq 1$ ist, so ersetzen wir e_i durch $\dfrac{e_i}{|e_i|}$ (e_i ist keinesfalls der Nullvektor, denn sonst läge v_i in $L(e_1, \ldots, e_{i-1}) = L(v_1, \ldots, v_{i-1})$, was unmöglich ist).
Schließlich ist $L(e_1, \ldots e_i) \subseteq L(v_1, \ldots, v_i) = L(e_1, \ldots, e_{i-1}, v_i)$ und umgekehrt $v_i \in L(e_1, \ldots, e_i)$, also stimmen beide Mengen überein. □

Wenn in einem Euklidischen Vektorraum eine Orthonormalbasis gewählt wird, vereinfachen sich die Rechnungen:

Lemma 12.3 *Sei $\{e_1, \ldots, e_n\}$ eine Orthonormalbasis von V und $v = \sum v_i e_i$, $w = \sum w_i e_i$, dann gilt $v_i = \langle v, e_i\rangle$, also $v = \sum \langle v, e_i\rangle e_i$, $\langle v, w\rangle = \sum v_i w_i$, $|v|^2 = \sum v_i^2$.* □

Lemma 12.4 (Besselsche Ungleichung) *Sei $\{e_1, \ldots, e_k\}$ ein Orthonormalsystem und $v \in V$, dann gilt $|v|^2 \geq \sum \langle v, e_i \rangle^2$.*

Beweis: Wir ergänzen $\{e_1, \ldots, e_k\}$ zu einer Orthonormalbasis (das geht!) und haben

$$|v|^2 = \sum_{i=1}^{k} \langle v, e_i \rangle^2 + \sum_{i=k+1}^{n} \langle v, e_i \rangle^2,$$

die zweite Summe ist nichtnegativ. □

Definition: Sei $U \subseteq V$ ein Unterraum, dann sei

$$U^\perp = \{v \in V \mid \langle v, u \rangle = 0 \text{ für alle } u \in U\}.$$

$U^\perp$ heißt das orthogonale Komplement von U.

Lemma 12.5 *$U^\perp$ ist ein Unterraum von V, es gilt $(U_1 + U_2)^\perp = U_1^\perp \cap U_2^\perp$, $U^{\perp\perp} = U$ und $U \cap U^\perp = \{o\}$.* □

12.2 Orthogonale Abbildungen und Matrizen

V und W seien Euklidische Vektorräume und $f : V \to W$ sei eine lineare Abbildung, f heißt orthogonale Abbildung, wenn

$$\langle f(v), f(w) \rangle = \langle v, w \rangle$$

für alle $v, w \in V$ gilt.

Lemma 12.6 *Sei $f : V \to W$ eine orthogonale Abbildung, dann gilt $|\, f(v) \,| = |v|$. Wenn v auf w senkrecht steht, so steht $f(v)$ auf $f(w)$ senkrecht und der Winkel zwischen $f(v)$ und $f(w)$ ist gleich dem Winkel zwischen v und w.* □

Lemma 12.7 *Wenn $|\, f(v) \,| = |v|$ für alle $v \in V$ gilt, so ist f eine orthogonale Abbildung.*

Beweis: $\langle v + w, v + w \rangle = \langle v, v \rangle + 2\langle v, w \rangle + \langle w, w \rangle$, also ist

$$\langle v, w \rangle = \frac{1}{2}(|v + w|^2 - |v|^2 - |w|^2)$$

und damit

$$\langle f(v), f(w) \rangle = \frac{1}{2}(|f(v + w)|^2 - |\, f(v) \,|^2 - |\, f(w) \,|^2). \ \square$$

Lemma 12.8 *Orthogonale Abbildungen sind injektiv.*

Beweis: Sei $f(v) = o$, dann ist $0 = \langle f(v), f(v) \rangle = \langle v, v \rangle$, also ist $v = o$. □

Satz 12.9 *Sei $f : V \to V$ ein orthogonaler Endomorphismus und $B = \{e_1, \ldots, e_n\}$ eine Orthonormalbasis von V. Dann ist $A_{BB}(f)^T = A_{BB}(f)^{-1}$.*

Beweis: Sei $f(e_i) = \sum f_{ji} e_j$, dann ist $A_{BB}(f) = (f_{ji}) = F$ und $\langle f(e_i), f(e_k)\rangle = \sum f_{ji} f_{jk} = \langle e_i, e_k\rangle = \delta_{ik}$, d.h. $F^T F = E$. □

Definition: Eine Matrix A mit $A^T = A^{-1}$ heißt orthogonal.

Also gehören zu orthogonalen Abbildungen bezüglich Orthonormalbasen orthogonale Darstellungsmatrizen.

Man kann die Orthogonalität einer Matrix so veranschaulichen: Bezüglich des Skalarprodukts $\langle v, w\rangle = \sum v_i w_i$ im R^n sind die Beträge der Zeilen und der Spalten gleich 1, das Skalarprodukt verschiedener Zeilen oder Spalten ist null.

Lemma 12.10 *Das Produkt und die Inverse von orthogonalen Matrizen sind orthogonal. Die n-reihigen orthogonalen Matrizen bilden eine Gruppe $O(n)$, die „orthogonale“ Gruppe.*

Beweis: Sei $A^T A = E, B^T B = E$, dann ist $(AB)^T AB = B^T A^T AB = B^T B = E$. □

Die Determinante einer orthogonalen Matrix hat den Wert -1 oder 1, eine Matrix mit Determinante 1 heißt speziell und die speziellen orthogonalen Matrizen bilden die „spezielle orthogonale“ Gruppe $SO(n)$.

Wir wollen uns eine Übersicht über die orthogonalen Matrizen verschaffen. Für kleine Matrizen ist dies trivial: $O(1) = \{1, -1\}$.

Sei $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in O(2)$, dann ist $E = A^T A = \begin{bmatrix} a^2 + b^2 & ac + bd \\ ac + bd & c^2 + d^2 \end{bmatrix}$, wir setzen $a = \cos w,\ b = \sin w$ und aus den anderen Relationen folgt $c = \pm b, d = \mp a$, also treten zwei Fälle auf:

$$A = \begin{bmatrix} \cos w & \sin w \\ -\sin w & \cos w \end{bmatrix},$$

dies ist eine Drehung, oder

$$A = \begin{bmatrix} \cos w & \sin w \\ \sin w & -\cos w \end{bmatrix},$$

dies ist das Produkt einer Drehung und einer Spiegelung.

Satz 12.11 *Jede Matrix $A \in SO(n)$ läßt sich als Produkt von $\dfrac{n(n-1)}{2}$ Drehmatrizen darstellen.*

Beweis: Die Matrix, die in der Ebene der i-ten und der j-ten Koordinatenachse eine Drehung um den Winkel w veranstaltet, bezeichnen wir mit $D_{ij}(w)$.
Wir multiplizieren die Matrix A der Reihe nach mit $D_{12}, D_{13}, \ldots$ und bestimmen w so, daß im Produkt an den Stellen $(1,2), (1,3), \ldots$ Nullen stehen und an der Stelle (1,1) eine positive Zahl steht. Das diese beiden Forderungen erfüllbar sind, möge sich der Leser bitte klarmachen (eine analoge Rechnung haben wir im Zusammenhang mit der Jacobi-Diagonalisierung durchgefürt). Nach $n-1$ Schritten haben wir

$$AD = \begin{bmatrix} a & 0 & \ldots & 0 \\ & & \ldots & \\ * & & \ldots & * \end{bmatrix},$$

diese Matrix ist orthogonal, also ist der Betrag der ersten Zeile gleich 1, also ist $a = 1$. Auch der Betrag der ersten Spalte ist gleich 1, also stehen in der ersten Spalte unter der 1 lauter Nullen. So fahren wir fort, bis wir als Produkt die Einheitsmatrix erhalten. □

Als nächstes wollen wir eine besondere Klasse orthogonaler Abbildungen untersuchen: Sei V ein Vektorraum und $x \neq o$ ein Vektor aus V. Wir definieren eine Abbildung $s_x : V \to V$ durch

$$s_x(w) = w - 2\frac{\langle x, w\rangle}{|x|^2}x.$$

Diese Abbildung ist orthogonal, denn

$$\langle s_x(v), s_x(w)\rangle = \langle v - 2\frac{\langle x, w\rangle}{|x|^2}x,\ w - 2\frac{\langle x, w\rangle}{|x|^2}x\rangle = \langle v, w\rangle.$$

Es gilt $s_{rx} = s_x$ für $r \in R$ und $s_x(x) = -x$. Wenn aber $\langle x, w\rangle = 0$ ist, so gilt $s_x(w) = w$, wie man leicht nachrechnet. Also: Die Elemente von $L(x)^\perp$ werden bei s_x nicht verändert, d.h. s_x ist die Spiegelung an $L(x)^\perp$.
Daß s_x eine lineare Abbildung ist, folgt aus dem

Satz 12.12 *Sei $f : V \to V$ eine Abbildung mit $\langle f(v), f(w)\rangle = \langle v, w\rangle$ für alle $v, w \in V$, dann ist f linear.*

Beweis:

$$\begin{aligned}&\langle f(v+w) - f(v) - f(w), f(v+w) - f(v) - f(w)\rangle\\ &= \langle f(v+w), f(v+w)\rangle - \langle f(v+w), f(v)\rangle - \dots\\ &= \langle v+w, v+w\rangle - \langle v+w, v\rangle - \dots = 0.\square\end{aligned}$$

Sei nun $U \subseteq V$ ein Unterraum und $x \in U$, dann ist $s_x : U \to U$ eine Spiegelung in U. Wenn aber v ein beliebiger Vektor aus V ist, so ist auch $s_x(v)$ definiert, wir haben also eine Spiegelung des gesamten Raums V, die die Spiegelung von U fortsetzt und $U^\perp$ festhält.

Satz 12.13 *Sei $f : V \to V$ eine orthogonale Abbildung, $f \neq id$, $\dim V = n$, dann ist f ein Produkt von höchstens n Spielungen.*

Beweis: Wir führen die Induktion über n.
Wenn $n = 1$ und $f \neq id$ ist, so ist $f(v) = -v$ und dies ist eine Spiegelung.
Sei der Satz für die Dimension $n - 1$ bewiesen.
Wir treffen noch eine zusätzliche Voraussetzung: Es soll ein Vektor $v \neq o$ existieren, für den $f(v) = v$ gilt.
Wenn nun $\langle w, v\rangle = 0$ ist, so ist auch $\langle f(w), f(v)\rangle = \langle f(w), v\rangle = 0$, also ist $H = L(v)^\perp$ ein invarianter Unterraum. Also ist die Einschränkung $f \mid H$ von f auf H ein Produkt von höchstens $n - 1$ Spiegelungen. Dies sind auch Spiegelungen von V, die $H^\perp = L(v)$ festlassen, also ist ihr Produkt gleich f.
Wir betrachten nun den allgemeinen Fall. Da $f \neq id$ ist, gibt es einen Vektor v mit $f(v) \neq v$. Wir setzen $w = f(v) - v, H = L(w)^\perp$, sei s_w die Spiegelung an H. Wir zeigen $s_w(f(v)) = v$:

Es ist $s_w(w) = -w$, also

$$s_w(f(v) - v) = -f(v) + v = s_w(f(v)) - s_w(v) \tag{1}$$

und es gilt

$$\langle f(v) + v, f(v) - v\rangle = \langle f(v), f(v)\rangle - \langle v, v\rangle = 0,$$

also liegt $f(v) + v$ in H und damit ist

$$s_w(f(v) + v) = f(v) + v = s_w(f(v)) + s_w(v) \tag{2}$$

Die Gleichungen (1) und (2) addiert ergeben $2s_w(f(v)) = 2v$. Damit erfüllt die Abbildung $s_w \circ f$ die obige spezielle Voraussetzung, ist also Produkt von höchstens $n - 1$ Spiegelungen. Damit ist f ein Produkt von höchstens n Spiegelungen. □

12.3 Die adjungierte Abbildung

Das Skalarprodukt auf V ist eine bilineare Abbildung, dazu gehört also eine Abbildung $t : V \to V^*$, $t(v)(w) = \langle v, w\rangle$.

Lemma 12.14 *Die Abbildung t ist bijektiv.*

Beweis: Sei $t(v) = o$, d.h. $t(v)(w) = \langle v, w\rangle = 0$ für alle $w \in V$, speziell ist $t(v)(v) = \langle v, v\rangle = 0$, also $v = o$. Also ist t injektiv und wegen der Dimensionsgleichheit von V und V^* ist t bijektiv. □

Folgerung 12.15 *Sei $l : V \to R$ eine Linearform, dann gibt es einen eindeutig bestimmen Vektor $w \in V$, so daß $l(v) = \langle v, w\rangle$ für alle $v \in V$ gilt.*

Beweis: Sei $w = t^{-1}(l)$, dann ist $t(w) = l$, also $t(w, v) = l(v) = \langle v, w\rangle$. □

Satz 12.16 *Seien V, W Euklidische Vektorräume und $f : V \to W$ eine lineare Abbildung. Dann gibt es eine eindeutig bestimmte lineare Abbildung $f^* : W \to V$ mit $\langle f(v), w\rangle = \langle v, f^*(w)\rangle$.*

Beweis: Für festes $w \in W$ ist die Zuordnung $v \to \langle f(v), w\rangle$ eine Linearform, also gibt es einen Vektor $u \in V$ mit $\langle f(v), w\rangle = \langle v, u\rangle$. Wir setzen dann $f^*(w) = u$. Die Linearität von f^* ergibt sich aus der Linearität von $\langle v, \quad \rangle$. □

Definition: f^* heißt die zu f adjungierte Abbildung.

Lemma 12.17 $(f + g)^* = f^* + g^*$, $id^* = id$, $(f \circ g)^* = g^* \circ f^*$, $f^{**} = f$. □

Lemma 12.18 *Sei $B = \{e_1, \ldots, e_n\}$ eine Orthonormalbasis von V und $f : V \to V$ eine lineare Abbildung. Dann ist $A_{BB}(f^*) = A_{BB}(f)^T$.*

Beweis: Sei $f(e_i) = \sum f_{ji} e_j$, dann gilt

$$\langle f(e_i), e_k\rangle = \langle e_i, f^*(e_k)\rangle = \langle \sum f_{ji} e_j, e_k\rangle = \sum f_{ji} \langle e_j, e_k\rangle = f_{ki},$$

also

$$f^*(e_k) = \sum f_{ki} e_i. \square$$

Lemma 12.19 *Wenn f eine orthogonale Abbildung ist, so ist $f^* = f^{-1}$.* □

Definition: Wenn $f^* = f$ ist, so heißt f selbstadjungiert.

Lemma 12.20 *Sei B eine Orthonormalbasis, f ist genau dann selbstadjungiert, wenn $A_{BB}(f)$ eine symmetrische Matrix ist.* □

Wir wollen für kurze Zeit als Grundkörper den Körper $\mathbb{C}$ der komplexen Zahlen wählen. Hier ist die durch $\langle \sum v_i e_i, \sum w_j e_j \rangle = \sum v_i w_i$ gegebene Bilinearform nicht mehr positiv definit. Abhilfe schaffen hier Hermitesche Formen:
Eine Abbildung $\langle \quad , \quad \rangle : V \times V \to \mathbb{C}$ heißt Hermitesch, wenn sie linear im ersten Faktor ist und $\langle v, w \rangle = \overline{\langle w, v \rangle}$ gilt (der Strich bedeutet die konjugiert komplexe Zahl). Für Hermitesche Formen gelten die bisher bewiesenen Resultate ebenfalls, blättern Sie zurück und beweisen Sie es.

Eine Abbildung, für die $\langle f(v), f(w) \rangle = \langle v, w \rangle$ gilt, heißt hier unitär. Die Darstellungsmatrix B zur adjungierten Abbildung ist die komplex konjugierte der Transponierten der Darstellungsmatrix A der gegebenen Abbildung, die Matrix B wird dann als die zu A adjungierte Matrix bezeichnet: $B = A^*$.

Satz 12.21 *Die Eigenwerte einer selbstadjungierten Abbildung sind reell.*

Beweis: Sei $f(v) = zv$, dann ist

$$\langle f(v), v \rangle = \langle zv, v \rangle = z\langle v, v \rangle = \langle v, f(v) \rangle = \langle v, zv \rangle = \bar{z}\langle v, v \rangle,$$

also ist $z = \bar{z}$ reell. □

Wir wollen nun die folgende Frage beantworten:
V sei ein Euklidischer Vektorraum und $f : V \to V$ ein Endomorphismus. Gibt es dann eine Orthonormalbasis von V, die aus Eigenvektoren von f besteht?
Oder anders ausgedrückt: Sei eine n-reihige Matrix A gegeben. Gibt es dann eine orthogonale (bzw. unitäre) Matrix X, so daß X^*AX eine Diagonalmatrix ist?

Wir werden sehen, daß dies für eine spezielle Klasse von Endomorphismen bzw. Matrizen der Fall ist.

Der folgende Satz stammt von I. Schur.

Satz 12.22 (Schur-Zerlegung) *Zu jeder Matrix A gibt es eine unitäre Matrix U, so daß U^*AU eine obere Dreiecksmatrix ist.*

Beweis: Wir führen die Induktion über n. Für $n = 1$ ist nichts zu zeigen. Sei der Satz für $(n-1)$-reihige Matrizen bewiesen.
Sei z ein Eigenwert und u_1 ein entsprechender Eigenvektor von A, also $Au_1 = zu_1$. Wir ergänzen u_1 zu einer Orthonormalbasis $\{u_1, \dots u_n\}$ von $\mathbb{C}^n$, also gilt $u_i^* u_j = \delta_{ij}$, d.h. die Matrix U_1 mit den Spalten $u_1, \dots, u_n$ ist unitär. Es gilt

$$AU_1 = U_1 \begin{bmatrix} z & \star \\ 0 & \\ \dots & \\ 0 & A_1 \end{bmatrix}$$

also

$$U_1^* A U_1 = \begin{bmatrix} z & \star \\ 0 & \\ \dots & \\ 0 & A_1 \end{bmatrix}$$

Nun sei U_2 eine unitäre Matrix, so daß $U_2^* A_1 U_2 = D_1$ eine Dreiecksmatrix ist. Dann ist

$$\begin{bmatrix} 1 & \dots & 0 \\ 0 & & U_2 \end{bmatrix}^* U_1^* A U_1 \begin{bmatrix} 1 & \dots & 0 \\ 0 & & U_2 \end{bmatrix} = \begin{bmatrix} z \dots & \star & \\ 0 & & D_1 \end{bmatrix}$$

eine Dreiecksmatrix. □

Satz 12.23 (reelle Schur-Zerlegung) *Zu einer reellen Matrix A gibt es eine orthogonale Matrix Q, so daß $Q^T AQ$ eine Block-Dreiecksgestalt hat, deren Diagonalblöcke R_k die folgende Form haben: $R_k = (d_k) \in M_{11}$, wenn d_k ein reeller Eigenwert von A ist, und $R_k \in M_{22}$, wenn $c \pm is$ ein Paar komplexer Eigenwerte von A ist.*

Beweis: Relle Eigenwerte werden wie oben bearbeitet.
Sei $z = c + is$ ein komplexer Eigenwert von A, dann gibt es $x, y \in R^n$ mit

$$A(x + iy) = (c + is)(x + iy),$$

also

$$Ax = cx - sy, \; Ay = sx + cy.$$

Die Zahl $c - is$ ist ebenfalls ein Eigenwert von A, der zugehörige Eigenvektor ist $x - iy$, denn

$$A(x - iy) = cx - sy - isx - icy = (c - is)(x - iy).$$

Da $c + is \neq c - is$ ist, ist $\{x + iy, x - iy\}$ eine linear unabhängige Menge. Wir zeigen, daß auch $\{x, y\}$ linear unabhängig ist.
Sei $rx + ty = o$, wir betrachten $(r - it)(x + iy)$ und dem dazu komplex konjugieren Vektor:

$$(r - it)(x + iy) + (r + it)(x - iy) = 2(rx + ty) = o,$$

wegen der linearen Unabhängigkeit von $\{x + iy, x - iy\}$ muß $r + it = 0$ sein, d.h. $r = t = 0$.
Aus den obigen Beziehungen sehen wir, daß $L(x, y)$ ein A-invarianter Unterraum ist, in dem wir eine Orthonormalbaisis $\{u, v\}$ wählen. Nun schreiben wir die Komponenten von u und v in die ersten beiden Spalten von Q und verfahren weiter wie bei der komplexen Schur-Zerlegung. □

Definition: Ein Endomorphismus $f : V \to V$ heißt normal, wenn $f \circ f^* = f^* \circ f$ gilt. Eine Matrix A heißt normal, wenn $AA^* = A^*A$ gilt.

Lemma 12.24 *Sei A normal und U unitär, dann ist U^*AU normal.*

Beweis: $(U^*AU)(U^*AU)^* = U^*AUU^*A^*U = U^*AA^*U = U^*A^*AU = U^*A^*UU^*AU = (U^*AU)^*(U^*AU)$. □

Lemma 12.25 *Eine normale Dreiecksmatrix hat Diagonalgestalt.*

Beweis: Sei $A = (a_{ij}),\ a_{ij} = 0$ für $i > j$, dann ist das j-te Diagonalelement der Matrix AA^* gleich

$$\sum_{i=1}^{n} a_{ji}\bar{a}_{ji}$$

und das j-te Diagonalelement von A^*A ist gleich

$$\sum_{i=1}^{n} \bar{a}_{ji}a_{ji},$$

aus der Gleichheit folgt der Reihe nach $a_{12} = \ldots = a_{1n} = 0,\ a_{23} = \ldots = a_{2n} = 0$ usw. □

Folgerung 12.26 (Spektralsatz) *1. Wenn A eine normale Matrix ist, so gibt es eine unitäre Matrix U, so daß U^*AU eine Diagonalmatrix ist.*
2. Wenn $f : V \to V$ ein normaler Endomorphismus ist, so besitzt V eine Orthonormalbasis $\{e_1, \ldots, e_n\}$ aus Eigenvektoren von f.
3. Wenn die e_i eine Orthonormalbasis von Eigenvektoren von f bilden, dann sind die e_i sind auch Eigenvektoren von f^*.

Beweis: 1. Es gibt eine unitäre Matrix U, so daß U^*AU eine normale Dreiecksmatrix ist.
2. ist äquivalent zu 1.
3. Sei $f(e_i) = z_ie_i$, dann ist

$$\langle f(e_i), e_j\rangle = \langle z_ie_i, e_j\rangle = \langle e_i, \bar{z}_ie_j\rangle = z_i\delta_{ij} = \langle e_i, f^*(e_j)\rangle,$$

dies ist gleich Null für $i \neq j$, also liegt $f^*(e_j)$ in $L(e_j)$, für $j = i$ erhalten wir $\langle e_i, f^*(e_i)\rangle = \bar{z}_i$, also $f^*(e_i) = \bar{z}_ie_i$. □

Es gilt aber auch die Umkehrung:

Satz 12.27 *Der Vektorraum V besitze eine Orthonormalbasis aus Eigenvektoren von $f : V \to V$. Dann ist f normal.*

Beweis: Wie oben folgt $f^*(e_i) = \bar{z}_ie_i$, dann ist $f \circ f^*(e_i) = z_i\bar{z}_ie_i = f^* \circ f(e_i)$ für alle i, also $ff^* = f^*f$. □

Analog beweist man den

Satz 12.28 *Sei $f : V \to V$ normal. Wenn alle Eigenwerte von f reell sind, so ist f selbstadjungiert. Wenn alle Eigenwerte von f den Betrag 1 haben, so ist f unitär.* □

Den folgenden Zusammenhang werden wir später benötigen.

Satz 12.29 *Sei $f : V \to W$ eine lineare Abbildung, dann ist* $(\mathrm{Ker}(f^*))^\perp = \mathrm{Im}(f)$.

Beweis: Sei $f^*(v) = o$, dann ist für alle Vektoren w die Gleichung

$$\langle f^*(v), w\rangle = 0 = \langle v, f(w)\rangle,$$

erfüllt, also liegt v in $(\text{Im } (f))^\perp$, also ist $\text{Im } (f) \subseteq (\text{Ker } (f^*))^\perp$. Sei F die Darstellungsmatrix von f bezüglich irgendwelcher Basen, sei $\dim V = n$, $\dim W = m$, dann ist

$$\dim \text{Im}(f) = rg(F) = rg(F^*) = \dim \text{Im}(f^*) = r$$

und

$$\dim \text{Ker}(f^*) = m - r,$$

also

$$\dim(\text{Ker}(f^*))^\perp = r,$$

daraus folgt die Behauptung. □

12.4 Pseudoinverse Matrizen

Wir wollen den Begriff der inversen Matrix auf nichtreguläre und nichtquadratische Matrizen verallgemeinern.
Wir betrachten Matrizen von fixierter Größe, und zwar seinen sie stets entweder aus M_{nm} oder aus M_{mn}, wenn Produkte $A_1A_2A_3\ldots$ gebildet werden, so wollen wir immer voraussetzen, daß diese Matrizen sich auch wirklich multiplizieren lassen, also abwechselnd aus M_{mn} und aus M_{nm} sind.

Definition: Eine Matrix A heißt pseudo-regulär, wenn eine Matrix X mit $AXA = A$ existiert. Die Matrizen A und X heißen zueinander pseudoinvers, wenn $AXA = A$ und $XAX = X$ gilt.

Lemma 12.30 *Wenn die Matrix A regulär ist, so ist sie auch pseudoregulär und die einzige zu A pseudo-inverse Matrix ist A^{-1}.* □

Lemma 12.31 *Wenn A pseudo-regulär ist, so besitzt es eine pseudo-inverse Matrix.*

Beweis: Sei $AXA = A$ und $Y = XAX$, dann ist $AYA = AXAXA = AXA = A$, $YAY = XAXAXAX = XAXAX = XAX = Y$, also sind A und Y zueinander pseudo-invers. □

Satz 12.32 *Sei $M = \{X \mid AXA = A\}$, dann ist jede Matrix $Y = X_1AX_2$, wo $X_1, X \in M$ sind, zu A pseudo-invers und jede zu A pseudoinverse Matrix hat diese Form. Seien A und X zueinander pseudoinvers, dann sind AX und XA idempotente Matrizen.*

Beweis: 1. $AYA = AX_1AX_2A = AX_2A = A$, $YAY = X_1AX_2AX_1AX_2A = X_1AX_1AX_2A = X_1AX_2 = Y$.
2. Sei A zu Y pseudo-inverse, dann liegt $Y = YAY$ in M.
3. $(AX)^2 = AXAX = AX$, analog für XA. □

Folgerung 12.33 *Seien A und B zueinander pseudoinvers, dann gilt rg(A) =rg(B) und Sp(AB) =Sp(BA) =rg(A).*

Beweis: Aus $ABA = A$ folgt rg(A) $\leq$rg(B), aus Symmetriegründen folgt die Gleichheit. die zweite Aussage folgt aus der Idempotenz von AB. □

Satz 12.34 *Sei B pseudoinvers zu A. Das Gleichungssystem $Ax = b$ ist genau dann lösbar, wenn $ABb = b$ ist und die allgemeine Lösung hat die Form $x = y - BAy + Bb$, wobei $y \in \mathbf{R}^n$ beliebig ist.*

Beweis: Sei u eine Lösung, also $Au = b$, dann ist $ABb = ABAu = Au = b$. Wenn umgekehrt $ABb = b$ gilt, so ist Bb eine Lösung.
Wegen $Ay - ABAy = 0$ ist $x = y - BAAy$ eine Lösung des homogenen Systems $Ax = 0$; wenn umgekehrt $Ax = 0$ ist, so hat x die Form $x = x - BAAx$. □

Wir betrachten ein Beispiel:
Sei $A = (1,\ 1)$, dann ist $B = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$ zu A pseudoinvers. Wir betrachten das Gleichungssystem $Ax = 5$. Es ist $AB = (1)$, also $ABb = b$ und wegen $BA = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$ ist die allgemeine Lösung gleich

$$\begin{pmatrix} a \\ b \end{pmatrix} - \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} \frac{5}{2} \\ \frac{5}{2} \end{pmatrix} = \begin{pmatrix} \frac{1}{2}a - \frac{1}{2}b + \frac{5}{2} \\ -\frac{1}{2}a + \frac{1}{2}b + \frac{5}{2} \end{pmatrix}$$

Definition: Die Matrix X heißt Moore-Penrose-Inverse der Matrix A, wenn
1. $AXA = A$, 2. $XAX = X$, 3. $(AX)^* = AX$, 4. $(XA)^* = XA$
gilt, d.h. A und X sind zueinander pseudo-invers und AX und XA sind selbstadjungiert.

Lemma 12.35 *Eine Matrix A besitzt höchstens eine Moore-Penrose-Inverse.*

Beweis: Seien X und Y pseudo-invers zu A, dann gilt

$$X = XAX = XX^*A^* = XX^*A^*Y^*A^* = XX^*A^*AY = XAXAY = XAY,$$

analog zeigt man $Y = XAY$, damit ist $X = Y$. □

Wir wollen nun zeigen, das jede Matrix eine Moore-Penrose-Inverse besitzt. Dazu brauchen wir ein paar Hilfsbetrachtungen.

Lemma 12.36 *Wenn $AA^* = 0$ ist, so ist $A = 0$. Wenn $BA^*A = 0$ ist, so ist $BA^* = 0$.*

Beweis: Sei $AA^* = 0$, dann ist $\mathrm{Sp}(AA^*) = 0$, diese Spur ist aber gleich $\sum a_{ij}\bar{a}_{ij}$, folglich sind alle $a_{ij} = 0$.
Wenn $BA^*A = 0$ ist, so ist auch $BA^*AB^* = (BA^*)(BA^*)^* = 0$ und daraus folgt die Behauptung. □

Satz 12.37 *Sei A eine beliebige (rechteckige) Matrix. Das Minimalpolynom $m(z)$ von A^*A hat die Null höchstens als einfache Nullstelle.*

Beweis: Wir nehmen an, $m(z)$ hätte die Null als k-fache Nullstelle, $k > 1$, also $m(z) = g(z)z^k$. Dann gilt

$$g(A^*A)(A^*A)^k = 0,$$

daraus folgt

$$g(A^*A)(A^*A)^{k-1}A^* = 0$$

und

$$g(A^*A)(A^*A)^{k-1} = 0$$

im Widerspruch dazu, daß $m(z)$ das Minimalpolynom ist. □

Der folgende Satz wurde von Penrose 1956 veröffentlicht.

Satz 12.38 *Jede Matrix A besitzt eine Moore-Penrose-Inverse A^-.*

Beweis: Wenn 0 kein Eigenwert von A^*A ist, so ist A regulär und A^{-1} exisiert. Andernfalls hat das Minimalpolynom $m(z)$ von A^*A (bis auf einen konstanten Faktor) die Form

$$m(z) = g(z)z^2 - z,$$

wir setzen

$$X = g(A^*A)A^*.$$

Es gilt

$$XAA^*A = g(A^*A)A^*AA^*A = g(A^*A)(A^*)^2 = A^*A,$$

also $(XA - E)A^*A = 0$, daraus folgt $(XA - E)A^* = 0$, d.h.

$$XAA^* = A^*.$$

Nun folgt

$$(XA)^* = A^*X^* = XAA^*X^* = X(XAA^*)^* = XA,$$
$$AXA = A(XA)^* = AA^*X^* = A.$$

Die Matrix A^*A ist selbstadjungiert, damit ist auch die Matrix $g(A^*A)$ selbstadjungiert, daraus folgt

$$AX = Ag(A^*A)A^* = Ag(A^*A)^*A^* = (Ag(A^*A)A^*)^* = (AX)^*,$$
$$XAX = X(AX)^* = XX^*A^* = g(A^*A)A^*X^*A^* = g(A^*A)(AXA)^* =$$
$$g(A^*A)A^* = X.$$

Damit sie die vier Eigenschaften bewiesen. □

Der Beweis liefert auch gleich ein Konstruktionsverfahren für die Moore-Penrose-Inverse, das aber in der Praxis nicht angewandt wird.

Wir wollen aber ein Beispiel betrachten. Sei $A = \begin{pmatrix} 3 & 2 & 1 \end{pmatrix}$, wir wollen die Moore-Penrose-Inverse von A bestimmen. Es ist

$$A^*A = \begin{pmatrix} 9 & 6 & 3 \\ 6 & 4 & 2 \\ 3 & 2 & 1 \end{pmatrix},$$

$rgA = rg(A^*A) = 1$, also hat das charakteristische Polynom die Form $z^3 - 14z^2$ und das Minimalpolynom ist $\frac{1}{14}z^2 - z$, also ist $X = \frac{1}{14}A^*$.

12.5 Unlösbare und unterbestimmte Gleichungssysteme

In praktischen Beispielen kommen lineare Gleichungssysteme vor, die zu wenige Gleichungen enthalten, um eine eindeutig bestimmte Lösung zu besitzen. Manchmal ist es sinnvoll, aus der Lösungsmannigfaltigkeit eine Lösung mit minimalem Betrag auszuwählen. Wie das zu geschehen hat, wollen wir uns ansehen.

Sei also $Ax = b$ ein lineares Gleichungssystem, $x \in R^m, b \in R^n$. Wir wissen vom Ende des letzten Kapitels, daß

$$R^m = \mathrm{Im}(A^*) \oplus \mathrm{Ker}(A)$$

gilt. Sei x eine Lösung des Systems, wir zerlegen $x = x_1 + x_2$, wo x_1 in $\mathrm{Im}(A^*)$ und x_2 in $\mathrm{Ker}(A)$ liegt, dann ist

$$Ax = Ax_1 + Ax_2 = Ax_1,$$

also ist x_1 auch eine Lösung, weiter ist

$$|x|^2 = |x_1|^2 + |x_2|^2 \geq |x_1|^2$$

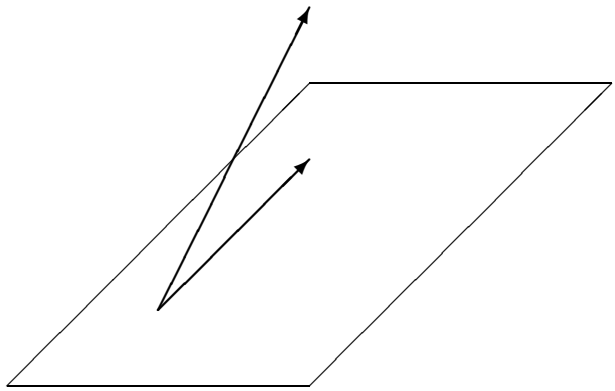
und die untere Grenze wird für $x \in \mathrm{Im}(A^*)$ angenommen.

Weiter kann es vorkommen, daß ein Gleichungssystem nicht lösbar ist, die Ursache kann in kleinen (vernachlässigbaren) Ungenauigkeiten der Eingangsdaten liegen, die sich aber fatal auswirken: Der Gaußsche Algorithmus liefert nicht etwa eine Näherungslösung, sondern gar keine. Wir wollen eine „Lösung“ x suchen, wo $|Ax - b|$ minimal ist, sicher ist das ein vernünftiger Kompromiß.

Der Wert von $|Ax - b|$ ist dann minimal, wenn Ax die orthogonale Projektion von b auf Im (A) ist. Es ist

$$R^n = \mathrm{Im}(A) \oplus \mathrm{Ker}(A^*)$$

wir zerlegen b entsprechend: $b = b_1 + b_2$. Da b_1 in $\mathrm{Im}(A)$ liegt, ist $Ax = b_1$ lösbar und $|Ax - b|$ ist minimal. Dadurch ist x noch nicht notwendigerweise endeutig bestimmt, evtl. ist noch eine „Lösung“ minimalen Betrags auszuwählen.



Der folgende Satz wurde von Penrose 1957 veröffentlicht.

Satz 12.39 *Für $x = A^- b$ sind $|\,Ax - b\,|$ und $|x|$ minimal.*

Beweis: Zu zeigen ist, daß x in $\mathrm{Im}(A^*)$ liegt und daß $AA^- b = b_1$ ist. Die erste Aussage folgt aus

$$A^- b = g(A^* A)A^* b$$

(mit den Bezeichnungen des ersten Satzes von Penrose).
Weiter ist $AA^- b$ die Orthogonalprojetion von b auf Im (A), denn $AA^- b \in$ Im (A) und

$$\langle AA^- b - b, AA^- b\rangle = \langle AA^- b, AA^- b\rangle - \langle b, AA^- b\rangle,$$

AA^- ist selbstadjungiert, also

$$= \langle b, AA^- AA^- b\rangle - \langle b, AA^- b\rangle$$

$$= \langle b, AA^- b\rangle - \langle b, AA^- b\rangle = 0. \ \square$$

Beispiel:
Wir betrachten das unlösbare Gleichungssystem

$$\begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Wir erhalten $\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{50} \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \frac{1}{50} \begin{bmatrix} -2 \\ -4 \end{bmatrix}$, die Probe ergibt $\begin{bmatrix} -\frac{1}{5} \\ \frac{3}{5} \end{bmatrix}$, doch eine gehörige Abweichung.

Moore-Penrose-Inverse von normalen Matrizen lassen sich leicht berechnen. Zunächst rechnen Sie bitte nach, daß man die Moore-Penrose-Inverse einer Diagonalmatrix dadurch erhält, daß man die von Null verschiedenen Diagonalelemente durch ihre Inversen ersetzt.

Satz 12.40 *Sei A eine (quadratische) normale Matrix, U eine unitäre Matrix und D eine Diagonalmatrix, so daß $A = UDU^*$ ist. Dann ist $A^- = UD^- U^*$.*

Beweis: Man rechnet es eben aus:
$AA^- A = UDU^* UD^- U^* UDU^* = UDD^- DU^* = UDU^* = A$ usw. $\square$

12.6 Householder-Transformationen

Seien $v = [v_1, \ldots, v_n]^T,\ x = [x_1, \ldots, x_n]^T,\ e = [1, 0, \ldots, 0]^T$ Spaltenvektoren aus M_{n1}, wir setzen

$$P = E - 2\frac{vv^T}{v^T v},$$

Es ist $Px = x - 2\frac{v^T x}{v^T v} v$, also $Pv = -v$ und wenn $\langle x, v\rangle = 0$ ist, so folgt $Px = x$. Demnach ist P die Spiegelung am Unterraum $L(v)^\perp$. Derartige Abbildungen werden als Householder-Transformationen bezeichnet.

Lemma 12.41 *Sei x gegeben, dann kann v so gewählt werden, daß Px in $L(e)$ liegt.*

Beweis: Wenn Px in $L(e)$ liegen soll, so muß v in $L(e,x)$ liegen. Wir setzen $v = x + re$, dann ist

$$Px = x - 2\frac{x^Tx + rx_1}{x^Tx + 2rx_1 + r^2}x - 2r\frac{x^Tx + rx_1}{v^Tv}e,$$

der Koeffizient von x ist gleich Null, wenn $r = |x|$ ist, also ist $v = x \pm |x|\, e$ zu wählen. (Das ist auch anschaulich klar: v ist die Winkelhalbierende oder die Normale der Winkelhalbierenden zwischen x und e.) □

12.7 QR-Zerlegung

Sei eine Matrix A gegeben, gesucht wird eine orthogonale Matrix Q, so daß $Q^TA = R$ eine obere Dreiecksmatrix ist. Dann gilt also $A = QR$, man nennt dies die QR-Zerlegung von A, sie hat viele Anwendungen in der numerischen Mathematik.
Wir werden diese Zerlegung mit Hilfe von Householder-Transformationen herstellen.
Die Spalten von A seien $a_1, \ldots, a_n$. Es sei P_1 eine Householder-Matrix, so daß

$$P_1a_1 = \begin{bmatrix} \star \\ 0 \\ \ldots \\ 0 \end{bmatrix}$$

ist, wir bilden

$$P_1A = \begin{bmatrix} \star & \star & \ldots & \star \\ 0 & \clubsuit & & \\ \ldots & & & \\ 0 & \clubsuit & \ldots & \star \end{bmatrix}$$

(Beachten Sie, daß wir die Elemente in der zweiten Spalte hervorgehoben haben.)
Nun sei P_2' die Householder-Matrix (mit einer Reihe weniger als P_1), für die

$$P_2' \begin{bmatrix} \clubsuit \\ \\ \ldots \\ \clubsuit \end{bmatrix} = \begin{bmatrix} \star \\ 0 \\ \ldots \\ 0 \end{bmatrix}$$

ist, wir setzen

$$P_2 = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & & & \\ \ldots & & & \\ 0 & & P_2' & \end{bmatrix}$$

Dann sind in P_2P_1A schon die ersten beiden Spalten richtig eingerichtet. So fahren wir fort, nach n Schritten haben wir $Q = P_1 \ldots P_n$ gefunden. □

Aufgabe: Sei die Matrix A gegeben, sei

$$A = Q_0 R_0$$

ihre QR-Zerlegung. Wir setzen

$$A_1 = R_0 Q_0$$

und bilden wieder die QR-Zerlegung:

$$A_1 = Q_1 R_1.$$

Nun sei wieder

$$A_2 = R_1 Q_1.$$

So fahren wir fort. Die Folge der A_i konvergiert gegen eine Matrix B (das kann man zeigen), probieren Sie aus, welche Matrix das ist! Ein Computerprogramm dafür ist schnell geschrieben. In der Numerik-Vorlesung lernen Sie, daß dieses Verfahren hervorragend zur Eigenwertberechnung geeignet ist.

12.8 Hessenberg-Zerlegung

Sei A eine Matrix, dann gibt es eine orthogonale Matrix Q, so daß

$$Q^T A Q = \begin{bmatrix} \star & & & \dots & & \star \\ \star & \star & & \dots & & \star \\ 0 & \star & \star & & & \star \\ 0 & 0 & \star & & & \star \\ & & & \dots & & \\ 0 & \dots & & 0 & \star & \star \end{bmatrix}$$

eine „Hessenberg“-Matrix ist. (Dies ist eine Dreiecksmatrix, wo unter der Diagonalen noch eine Reihe besetzt ist.)

Die Existenz ist klar: Die reelle Schur-Zerlegung hat eine solche Gestalt. Wir werden sehen, daß man hier Householder-Transformationen nutzen kann.

Sei $A = (a_{ij})$, es sei P_1' eine Householder-Matrix, die den Vektor $\begin{bmatrix} a_{21} \\ a_{31} \\ \dots \\ a_{n1} \end{bmatrix}$ in $\begin{bmatrix} \star \\ 0 \\ \dots \\ 0 \end{bmatrix}$ überführt. Wir rändern die Matrix P_1' mit Nullen in der ersten Zeile und Spalte und einer Eins links oben:

$$P_1 = \begin{bmatrix} 1 & 0 & \dots & 0 \\ & & \dots & \\ 0 & & & P_1' \end{bmatrix}$$

dann ist

$$P_1 A = \begin{bmatrix} \star & & & \dots & & \star \\ \star & \star & & \dots & & \star \\ 0 & \star & \star & & & \star \\ & & & \dots & & \\ 0 & \star & & \dots & \star & \star \end{bmatrix}$$

wenn wir nun noch, wie gefordert, von rechts mit P_1^T heranmultiplizieren, bleiben die Nullen in der ersten Spalte erhalten. Nun suchen wir eine $(n-2)$-reihige Householdermatrix, die einen Teil der zweiten Spalte von $P_1 A P_1^T$ in Nullen überführt, rändern wieder zu einer n-reihigen Matrix usw. □

Lemma 12.42 *Wenn die Matrix A symmetrisch und $H = Q^T AQ$ eine Hessenbergmatrix ist, so ist H tridiagonal, d.h. in H sind nur die Diagonale und die beiden Reihen über und unter der Diagonalen besetzt.*

Beweis: $H^T = (Q^T AQ)^T = Q^T A^T Q = Q^T AQ = H$. □

Wir wissen zwar, daß eine symmetrische Matrix sogar diagonalisierbar ist, haben dafür z. B. Iterationsverfahren zur Verfügung. Die tridiagonale Form kann man jedoch in wenigen Schritten erzeugen.

Satz 12.43 *Sei*

$$T_r = \begin{bmatrix} a_1 & b_2 & & & & \\ b_2 & a_2 & b_3 & & & \\ & b_3 & \dots & & & \\ & & & \dots & & \\ & & & & & b_r \\ & & & & b_r & a_r \end{bmatrix}$$

eine tridiagonale symmetrische r-reihige Matrix und $c_r(z)$ ihr charakteristisches Polynom, dann gilt

$$c_r(z) = (a_r - z)c_{r-1}(z) - b_r^2 c_{r-2}(z).$$

Beweis: Es ist

$$\det \begin{bmatrix} a_1 - z & b_2 & & & & \\ b_2 & a_2 - z & b_3 & & & \\ & b_3 & \dots & & & \\ & & & \dots & & \\ & & & & & b_r \\ & & & & b_r & a_r - z \end{bmatrix}$$

$$= (a_r - z)c_{r-1}(z) - b_r \; det \begin{bmatrix} a_1 - z & b_2 & & & & \\ b_2 & a_2 - z & b_3 & & & \\ & b_3 & \dots & & & \\ & & & \dots & & \\ & & & & a_{r-2} - z & 0 \\ & & & & b_{r-1} & b_r \end{bmatrix}$$

und durch Entwicklung des zweiten Summanden nach der letzten Spalte erhalten wir die Behauptung. □

Weil wir ihn als ein Hilfsmittel für die anschließende Herleitung der Singulärwertzerlegung verwenden können, beweisen wir hier den verallgemeinerten Determinantenmultiplikationssatz:

Satz 12.44 *Seien $A \in M_{mn}$ und $B \in M_{nm}$ Matrizen und es sei $m \leq n$. Dann gilt $\det(AB) = \sum_K \det A_{IK} \det B_{KI}$, wobei $I = \{1, \ldots, m\}$ ist und K alle Indexsysteme $\{k_1, \ldots, k_m\}$ mit $1 \leq k_1 < k_2 < \ldots < k_m \leq n$ durchläuft, A_{IK} ist die Untermatrix von A, die nur die Zeilen aus I und die Spalten aus K enthält.*

Beweis: Es gilt

$$AB = \begin{bmatrix} \sum_{i_1} a_{1i_1} b_{i_1 1} & \ldots & \sum_{i_m} a_{1i_m} b_{i_m m} \\ & \ldots & \\ \sum_{i_1} a_{mi_1} b_{i_1 1} & \ldots & \sum_{i_m} a_{mi_m} b_{i_m m} \end{bmatrix}$$

Die Determinantenfunktion ist linear in jeder Spalte, also gilt

$$\det(AB) = \sum_{i_1} \ldots \sum_{i_m} \det \begin{bmatrix} a_{1i_1} & \ldots & a_{1i_m} \\ \ldots & & \\ a_{mi_1} & \ldots & a_{mi_m} \end{bmatrix} b_{i_1 1} \ldots b_{i_m m}$$

Die Determinante auf der rechten Seite ist null, wenn einige Indizes übereinstimmen, wir betrachten also nur Summanden, wo alle i_l paarweise verschieden sind. Sei $\{i_1, \ldots, i_m\} = \{k_1, \ldots, k_m\}$, wo $k_1 < \ldots < k_m$ ist. Sei p die Permutation mit $p(i_l) = k_l$. Dann ist die in Frage stehende Determinante gleich

$$\operatorname{sgn}(p) \det A_{IK}$$

und damit

$$\det(AB) = \sum_K \det A_{IK} \sum_p \operatorname{sgn}(p) b_{i_1 1} \ldots b_{i_m m},$$

wegen $i_l = p^{-1}(k_l)$ ist die rechte Summe gleich $\det(B_{KI})$. □

Folgerung 12.45 *Sei $C = AB$, sei C_{JK} eine Untermatrix von C, dann ist*

$$C_{JK} = A_{JI} B_{IK} \quad (I = \{1, \ldots, n\})$$

und damit

$$\det C_{JK} = \sum \det A_{JL} \det B_{LK}. \square$$

Wir wenden dies für den Fall $B = A^T$ an und betrachten den Hauptminor $\det C_{JJ}$ von AA^T: $\quad \det C_{JJ} = \sum \det A_{JK} \det A_{JK}$.

Folgerung 12.46 *Die charakteristischen Polynome von AA^T und $A^T A$ unterscheiden sich um den Faktor z^{n-m}, d.h. die von Null verschiedenen Eigenwerte von AA^T und $A^T A$ stimmen überein.*

Beweis: Der Koeffizient von z^{n-k} im charakteristischen Polynom von AA^T ist die Summe der k-Hauptminoren, also gleich

$$\sum_I \sum_J \det A_{IJ} \det A_{IJ},$$

bei $A^T A$ ist der Koeffizient von z^{m-k} gleich

$$\sum_I \sum_J \det A_{JI} \det A_{JI},$$

also stimmen sie überein. □

12.9 Singularwertzerlegung

Sei $f : V \to W$ eine lineare Abbildung, dabei sei $\dim V = n,\ \dim W = m,\ \dim \mathrm{Im}(f) = r$. Wir wissen, daß die von Null verschiedenen Eigenwerte von f^*f und von ff^* übereinstimmen und daß sie reell sind. Wir überlegen, daß die Eigenwert nicht negativ sein können: Sei $A^TAv = zv$, dann ist $v^TA^TAv = |Av|^2 = zv^Tv = z\,|v|^2 \geq 0$. Wir können also die gemeinsamen Eigenwerte von ff^* und f^*f mit $a_1^2, \ldots, a_s^2$ bezeichnen.
Nun gibt es eine Orthonormalbasis $B = \{v_1, \ldots, v_n\}$ von V aus Eigenvektoren von f^*f und eine Orthonormalbasis $C = \{w_1, \ldots, w_m\}$ aus Eigenvektoren von ff^*. Also ist

$$f^*f(v_i) = a_i^2 v_i,$$

also

$$f(f^*f(v_i)) = (ff^*)(f(v_i)) = a_i^2 f(v_i),$$

d.h $f(v_i)$ ist ein Eigenvektor von ff^* zum Eigenwert a_i^2, also gilt

$$f(v_i) = rw_i.$$

Analog erhalten wir

$$f^*(w_i) = pv_i.$$

Nun gilt

$$\langle f(v_i), w_i\rangle = \langle rw_i, w_i\rangle = r = \langle v_i, f^*(w_i\rangle = \langle v_i, pv_i\rangle = p.$$

Wir setzen oben $v_i = f^*(\frac{1}{r}w_i)$ ein:

$$ff^*ff^*(\frac{1}{r}w_i) = \frac{1}{r}a_i^4 w_i = a_i^2 f(v_i) = a_i^2 rw_i,$$

folglich ist $|r| = a_i$ und wir können (evtl. nach Ersetzung von w_i durch $-w_i$) $r = a_i$ annehmen, also

$$f(v_i) = a_i w_i.$$

Die Zahlen a_i heißen die Singulärwerte von f und die Basen B und C ein Paar singulärer Basen.

Folgerung 12.47 *1. Es gibt Orthogonalbasen B, C von V bzw. W mit*

$$A_{BC}(f) = \begin{bmatrix} a_1 & & 0 \\ & \ldots & \\ 0 & & a_s \end{bmatrix}.$$

2. Zur Matrix A gibt es orthogonale Matrizen U, V, so daß

$$A = U^T \begin{bmatrix} a_1 & & 0 \\ & \ldots & \\ 0 & & a_s \end{bmatrix} V. \square$$

12.10 Vektor- und Matrixnormen

Wir wählen in diesem Abschnitt den Körper $\mathbf{R}$ der rellen Zahlen als Grundkörper.
Eine Zuordnung $\mathbf{R} \longrightarrow \mathbf{R}$, $x \mapsto \|x\|$ mit den Eigenschaften

1. $\|x\| \geq 0$, $\|x\| = 0$ genau für $x = 0$,
2. $\|rx\| = |r|\,\|x\|$ $(r \in \mathbf{R})$,
3. $\|x + y\| \leq \|x\| + \|y\|$

heißt eine Vektornorm.
Zum Beispiel ist der Betrag $\|x\|_2 = \sqrt{\langle x, x\rangle}$ eine Vektornorm, sie heißt die euklidische Norm. Weitere Beispiele sind die
Summennorm $\|x\|_1 = \sum |x_i|$
und die Maximumnorm $\|x\|_\infty = \max |x_i|$.
Man rechnet die obigen Eigenschaften leicht nach. Veranschaulichen Sie sich durch eine Skizze, wie die jeweiligen „Einheitskreise“ aussehen.

Lemma 12.48 $\|x\|_\infty \leq \|x\| \leq n\,\|x\|_\infty$,
$\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\,\|x\|_\infty$,
$\|x\|_2 \leq \|x\|_1 \leq \sqrt{n}\,\|x\|_2$.

Definition: Sei $\|\ \|_p$ eine Vektornorm, $p \in \{1, 2, \infty\}$; wir setzen für eine Matrix A

$$\|A\|_p = \max\left\{\frac{\|Ax\|_p}{\|x\|_p} \mid x \neq 0\right\} = \max_{\|x\|_p} \|Ax\|_p\,,$$

dies nennen wir die durch $\|\ \|_p$ induzierte Matrixnorm.
Also gilt $\|Ax\|_p \leq \|A\|_p \cdot \|x\|_p$.

Weitere Matrixnormen sind die folgenden:

$$\|A\|_z = \max_k \sum_j |a_{kj}| \qquad \text{Zeilensummennorm}$$

$$\|A\|_s = \max_j \sum_k |a_{kj}| \qquad \text{Spaltensummennorm}$$

$$\|A\|_F = \sqrt{\sum a_{ij}^2} \qquad \text{Frobeniusnorm}$$

Lemma 12.49 *Die Spaltensummennorm wird durch die Maximumnorm, die Zeilensummennorm wird durch die Summennorm induziert.*

Beweis: Zunächst gilt

$$\|A\|_\infty = \max_i \left|\sum_j a_{ij}x_j\right| \leq \max_i \sum_j |a_{ij}| \cdot |x_j| \leq \|x\|_\infty \max_i \sum_j |a_{ij}| = \|x\|_\infty \cdot \|A\|_z\,,$$

wir zeigen, daß die obere Schranke tatsächlich angenommen wird. Sei die k-te die Zeilensumme, wo das Maximum angenommen wird:

$$\max_i \sum_j |a_{ij}| = \sum_j |a_{kj}| \,,$$

dann setzen wir

$$x_i = \begin{cases} \frac{|a_{ki}|}{a_{ki}} & \text{für } a_{ki} \neq 0 \\ 0 & \text{sonst} \end{cases} \,,$$

dann ist $\|x\|_\infty = 1$ und $\max |\sum a_{ij} x_j| = \sum_j |a_{kj}|$.
Die andere Aussage wird analog bewiesen. □

Die verschiedenen Matrixnormen haben folgende Eigenschaften:

1. Für jede orthogonale Matrix Q gilt $\|x\|_2^2 = x^T x = x^T Q^T Q x = \|Qx\|_2$.

2. Wenn Q und R orthogonale Matrizen sind, so gilt $\|QAR\|_2 = \|A\|_2$, denn $\max \|QARx\|_2 = \max \|ARx\|_2 = \max \|Ay\|_2$.

3. $\|ABx\|_p = \|A(Bx)\|_p \leq \|A\|_p \|Bx\|_p \leq \|A\|_p \|B\|_p \|x\|_p$, also $\|AB\|_p \leq \|A\|_p \|B\|_p$.

4. Sei
$$D = \begin{pmatrix} d_1 & & \\ & \dots & \\ & & d_n \end{pmatrix}$$
eine Diagonalmatrix, dann ist $\|D\|_1 = \|D\|_2 = \|D\|_\infty = \max |d_i|$.

5. Sei
$$QAR = \begin{pmatrix} a_1 & & \\ & \dots & \\ & & a_n \end{pmatrix}$$
die Singulärwertzerlegung von A und $a_1 = \max a_i$, dann ist $\|A\|_2 = a_1$.

6. Sei z ein Eigenwert von A, dann $|z| \leq \|A\|_p$, denn für einen zugehörigen Eigenvektor v gilt $|z| \cdot \|v\| = \|zv\| = \|Av\| \leq \|A\| \cdot \|v\|$.

7. Sei A eine quadratische Matrix, für ihre Eigenwerte gelte $|z_1| \geq \ldots \geq |z_n|$ und ihre Singulärwerte seien $a_1 \geq \ldots \geq a_n$. Dann gilt $\prod a_i = |\det(A)|$, da orthogonale Matrizen die Determinante ± 1 haben, und und $a_n \leq |z_i| \leq a_1$. Die rechte Ungleichung folgt aus dem oben Gesagten, die linke Ungleichung ist für $a_n = 0$ trivial, wenn aber $a_n \neq 0$ ist, so ist A regulär und hat die Eigenwerte $\frac{1}{z_i}$ und die Singulärwerte $\frac{1}{a_i}$, die linke Ungleichung folgt dann aus der rechten.

8. Wenn A eine symmetrische Matrix ist, so gilt $A^T A v_i = a_i^2 v_i = A^2 v_i = z_i^2 v_i$, also $a_i = |z_i|$.

Durch einige Rechnung kann man die folgende Hölder-Ungleichung beweisen:

$$\sum u_k v_k \leq (\sum u_k^p)^{\frac{1}{p}} (\sum u_k^q)^{\frac{1}{q}}, \qquad \text{wobei } \frac{1}{p} + \frac{1}{q} = 1.$$

Hieraus folgt für $p > 1$ die Minkowski-Ungleichung

$$(\sum |x_k + y_k|^p)^{\frac{1}{p}} \leq (\sum |x_k|^p)^{\frac{1}{p}} + (\sum |y_k|^p)^{\frac{1}{p}},$$

dies ist für $p = 2$ die Dreiecksungleichung und allgemein bedeutet es, daß für $\|x\|_p = (\sum |x_k|^p)^{\frac{1}{p}}$ die oben geforderte dritte Normeigenschaft gilt (die beiden anderen sind ebenfalls erfüllt), damit haben wir eine ganze Serie von Vektornormen und durch sie induzierter Matrixnormen erhalten.

12.11 Positiv definite Matrizen

Wir beginnen mit ein paar Vorüberlegungen. Quadratische Gleichungen löst man durch quadratische Ergänzung:
Wegen

$$ax^2 + 2bxy + cy^2 = a(x + \frac{b}{a}y)^2 + (c - \frac{b^2}{a})y^2 = 0$$

gilt

$$x = \left(-\frac{b}{a} \pm \sqrt{\frac{1}{a}(c - \frac{b^2}{a})} \right) y.$$

Man kann dies in Matrixform schreiben:

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

wobei $a \neq 0$ und $d = c - \frac{b^2}{a}$ ist.
Dies kann in zwei verschiedenen Weisen auf symmetrische $n \times n$-Matrizen verallgemeinert werden. Sei

$$S = \begin{pmatrix} T & w \\ w^T & a \end{pmatrix},$$

wo T eine $(n-1)$-reihige Matrix und $a \neq 0$ ist. Dann gilt

$$S = \begin{pmatrix} E & \frac{1}{a}w \\ 0 & 1 \end{pmatrix} \begin{pmatrix} T - \frac{1}{a}ww^T & 0 \\ 0 & a \end{pmatrix} \begin{pmatrix} E & 0 \\ \frac{1}{a}w^T 1 & \end{pmatrix}.$$

Wenn T eine invertierbare Matrix ist, so gilt

$$S = \begin{pmatrix} E & 0 \\ w^T(T^{-1})^T & 1 \end{pmatrix} \begin{pmatrix} T & 0 \\ 0 & a - w^T T^{-1} w \end{pmatrix} \begin{pmatrix} E & T^{-1}w \\ 0 & 1 \end{pmatrix}.$$

Wir setzen $d = a - w^T T^{-1} w$, es ist $d = \dfrac{\det(S)}{\det(T)}$. Die jeweiligen rechten und linken Faktoren sind zueinander transponierte Dreiecksmatrizen, auf deren Diagonalen Einsen stehen, solche Matrizen heißen unipotent .

Definition: Die Determinanten der Untermatrix einer Matrix A, die durch streichen der letzten $n-i$ Zeilen und Spalten entstehen, nennen wir die i-Anfangsminoren von A.
In der Literatur werden diese oft als „Hauptminoren“ bezeichnet, wir haben diesen Begriff aber schon vergeben.

Satz 12.50 (Jacobi) *Die Anfangsminoren $d_1, \ldots, d_n$ der symmetrischen Matrix S seien von Null verschieden, dann gibt es eine unipotente Matrix W mit*

$$S = W^T \begin{pmatrix} d_1 & & & & \\ & \frac{d_2}{d_1} & & & \\ & & \frac{d_3}{d_2} & & \\ & & & \cdots & \\ & & & & \frac{d_n}{d_{n-1}} \end{pmatrix}.$$

Beweis: Den Anfang haben wir soeben gemacht: Die oben eingeführte Zahl d ist gleich $\frac{d_n}{d_{n-1}}$ und da $\det(T) = d_{n-2} \neq 0$ ist, kann das ganze Verfahren auf die Untermatrix T angewandt werden, usw. □

Definition: Eine symmetrische Matrix S heißt positiv definit, falls die Bilinearform $b_S(y,x) = x^T S y$ positiv definit ist.

Zum Beispiel sei $S = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$, dann ist

$$a \cdot x^T S x = a^2 x_1^2 + 2abx_1x_2 + acx_2^2 = (ax_1 + bx_2)^2 + (ac - b^2)x_2^2 > 0$$

genau dann, wenn $a > 0$ und $ac - b^2 > 0$.

Wir bemerken, daß Diagonalmatrizen mit positiven Diagonalelementen positiv definit sind.

Wenn W invertierbar und S positiv definit sind, so ist $W^T S W$ auch positiv definit, denn $(Wx)^T S W x = x^T W^T S W x = 0$ genau dann, wenn $Wx = 0$, also wenn $x = 0$.

Schließlich ist mit

$$S = \begin{pmatrix} T & \star \\ \star & \star \end{pmatrix},$$

auch die $m \times m$-Matrix T positiv definit, dies sehen wir, wenn wir b_T auf Vektoren anwenden, deren letzte $n-m$ Komponenten gleich Null sind.

Satz 12.51 *Sei S eine symmetrische Matrix, dann sind folgende Bedingungen äquivalent:*
1. S ist positiv definit,
2. es gibt eine invertierbare Matrix W mit $S = W^T W$,
3. alle Anfangsminoren von S sind positiv.

Beweis: ($2 \Rightarrow 1$) Es ist $S = W^T E W$ und E ist positiv definit.
($1 \Rightarrow 3$) Sei T eine Anfangs-Teilmatrix von S:

$$S = \begin{pmatrix} T & \star \\ \star & \star \end{pmatrix},$$

dann ist T positiv definit, also ist b_T eine nichtausgeartete Bilinearform, also ist $\det(T) \neq 0$, die Anfangsminoren sind von Null verschieden. Nach dem Satz von Jacobi gibt es eine unipotente (also invertierbare) Matrix U, so daß $S = U^T DU$ ist, wobei D eine Diagonalmatrix mit den Diagonaleinträgen $\frac{d_i}{d_{i-1}}$ ist. Mit S ist auch die Diagonalmatrix D positiv definit, also gilt $d_i > 0$.
($3 \Rightarrow 2$) Wenn die Anfangsminoren positiv sind, so ist $S = U^T DU$ und es gibt reelle Zahlen a_i mit $\frac{d_i}{d_{i-1}} = a_i^2$. Wir setzen

$$A = \begin{pmatrix} a_1 & & \\ & \dots & \\ & & a_n \end{pmatrix},$$

dann ist $D = A^2$; wenn wir $W = AU$ setzen, erhalten wir $S = U^T DU = U^T a^2 U = W^T W$. □

12.12 Aufgaben

1. Durch welche der folgenden Funktionen ist im R^2 ein Skalarprodukt für jedes Vektorpaar definiert?

   $f_1(\vec{x}, \vec{y}) = x_1 y_1 + x_2 y_2 - x_1 y_2 - x_2 y_1$

   $f_2(\vec{x}, \vec{y}) = x_1 y_1 + 2x_2 y_2$

   $f_3(\vec{x}, \vec{y}) = 3x_2 y_2 + x_1(y_1 + 2y_2)$

   $f_4(\vec{x}, \vec{y}) = x_1 + x_2 y_2 + y_2$

2. Beweisen Sie, daß in einem euklidischen Vektorraum gilt: Wenn $< \vec{a}, \vec{x} >=< \vec{b}, \vec{x} >$ für alle $\vec{x} \in V$, so $\vec{a} = \vec{b}$.

3. Im A^3 sei das kartesische Koordinatensystem $(O, \mathbf{B})$ sowie die Punkte $A = (2, -5, 1)_{OB}, B = (6, -3, 5)_{OB}, C = (6, -4, 9)_{OB}$ gegeben. Berechnen Sie die Längen der Dreiecksseiten, wenn das kanonische Skalarprodukt im A^3 gegeben ist.

4. Zeigen Sie, daß in einem euklidischen Vektorraum das sogenannte Parallelogrammgesetz gilt: $\| \vec{x} + \vec{y} \|^2 + \| \vec{x} - \vec{y} \|^2 = 2(\| \vec{x} \|^2 + \| \vec{y} \|^2)$. (Fertigen Sie sich für den R^2 eine Skizze an!)

5. Berechnen Sie für einen Würfel die Größe folgender Winkel:

   a) Winkel zwischen einer Flächendiagonale und den anstoßenden Kanten,

   b) Winkel zwischen einer Raumdiagonalen und den anstoßenden Kanten,

   c) Winkel zwischen zwei Raumdiagonalen,

   d) Winkel zwischen einer Raumdiagonalen und einer anstoßenden Flächendigonalen.

6. Seien $\vec{a}$ und $\vec{b}$ Elemente eines euklidischen Vektorraums. Beweisen Sie, daß dann folgendes gilt:

   a) Wenn r und s positive reelle Zahlen sind, dann gilt: $\angle(r\vec{a}, s\vec{b}) = \angle(\vec{a}, \vec{b})$

   b) $\vec{a}$ und $\vec{b}$ sind linear abhängig genau dann, wenn entweder $\angle(\vec{a}, \vec{b}) = 0$ oder $\angle(\vec{a}, \vec{b}) = \pi$ gilt.

7. Sei V ein euklidischer Vektorraum; für einen Vektor $\vec{x} \in V$ mit $\vec{x} \neq \vec{o}$ sei $\vec{x}^o$ folgendermaßen definiert: $\vec{x}^o := \frac{1}{\|\vec{x}\|}\vec{x}$. Beweisen Sie, daß folgende Aussagen äquivalent sind :

   a) $\vec{x}$ und $\vec{y}$ sind linear abhängig;

   b) $\vec{x}^o = \vec{y}^o$ oder $\vec{x}^o =$ - $\vec{y}^o$;

   c) $< \vec{x}^o, \vec{y}^o >= 1$.

8. Beweisen Sie, daß für beliebige Vektoren $\vec{x}$ und $\vec{y}$ eines euklidischen Vektorraums gilt: $\Big| \parallel \vec{x} \parallel - \parallel \vec{y} \parallel \Big| \leq \parallel \vec{x} + \vec{y} \parallel$

9. Beweisen Sie, daß zu jeder Geraden g und zu jedem Punkt P mit $P \notin g$ genau ein Lot von P auf g existiert.

10. Beweisen Sie vektoriell:

    a) den Satz des Thales,

    b) Ein Viereck, in dem sich die Diagonalen halbieren, ist ein Parallelogramm.

    c) In jedem regelmäßigen Tetraeder sind die Vektoren, die zu zwei windschiefen Kanten gehören, zueiander orthogonal.

11. Sei der $\mathbf{R}^3$ mit folgendem Skalarprodukt gegeben: $< \vec{x}, \vec{y} >= x_1y_1 + 2x_2y_2 + x_3y_3 - x_2y_3 - x_3y_2$ Bestimmen Sie ausgehend von der Basis $B = \{(1,0,0), (0,1,0), (0,0,1)\}$ eine ONB für diesen euklidischen Vektorraum.

12. Bestimmen Sie für die folgende Matrix $A = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix}$ eine obere Dreiecksmatrix R und eine orthogonale Matrix Q, so daß $A = QR$ gilt.

13. Sei $\mathbf{R}^3$ mit kanonischen Skalarprodukt sowie $U_1 = \text{lin}\ \{\vec{a}_1, \vec{a}_2\}$ mit $\vec{a}_1 = (1,1,0)$ und $\vec{a}_2 = (0,1,1)$ gegeben.

    a) Geben Sie das orthogonale Komplement zu U_1 an!

    b) Geben Sie einen Unterraum U_2 von $\mathbf{R}^3$ an mit $U_2 \neq U_1^\perp$ und $\mathbf{R}^3$ ist die direkte Summe von U_1 und U_2.

14. Im $\mathbf{R}^3$ seien der Punkt $P = (1, -3, 1)$ und eine Ebene ϵ durch die Gleichung $2x - 5y + z = -2$ gegeben. Bestimmen Sie den Abstand von P und ϵ!

15. a) Ermitteln Sie die gegenseitige Lage der durch die folgenden Parametergleichungen gegebenen Geraden des $\mathbf{R}^3$! $g_1 : \vec{x} = (1, -4, 2) + t(1, -3, -2)$; $g_2 : \vec{x} = (9, 0, 0) + s(4, 2, -1)$.

    b) Ermitteln Sie den Schnittwinkel der beiden Geraden!

    c) Ermitteln Sie eine Parametergleichung und eine Normalengleichung für die durch g_1 und g_2 bestimmte Ebene ϵ!

16. Gegeben Sei eine Ebene ϵ und eine Geradenschar $g(t)$ durch: $\epsilon : x + 2y = 3$ und $g(t) : \vec{x} = (1 + t, 2 + t, 1) + s(t, 1 + t, 1 - t)$

    a) Für welche reellen Zahlen t ist $g(t)$ parallel zu ϵ, für welche t liegt $g(t)$ in ϵ ?

    b) Für welches t ist $g(t)$ orthogonal zu ϵ ?

17. Beweisen Sie, daß für alle Vektoren a, b, c des $\mathbf{R}^3$ gilt:

    a) $\vec{a}x\vec{b} = -\vec{b}x\vec{a}$,

    b) $\vec{a}x(\vec{b}x\vec{c}) + \vec{b}x(\vec{c}x\vec{a}) + \vec{c}x(\vec{a}x\vec{b}) = \vec{o}$.

18. a) Zeigen Sie, daß es sich bei den folgenden Abbildungen des $\mathbf{R}^2$ auf sich um orthogonale Abbildungen handelt; dabei sei $B = \{\vec{e}_1, \vec{e}_2\}$ eine Orthonormalbasis von $\mathbf{R}^2$, und $\vec{x} \in \mathbf{R}^2$ ein beliebiger Vektor aus $\mathbf{R}^2$ mit den Koordinaten (x_1, x_2) bzgl. B.: $\varphi_1(\vec{x}) = -x_1\vec{e}_1 + x_2\vec{e}_2$, $\varphi_2(\vec{x}) = -x_1\vec{e}_1 - x_2\vec{e}_2$

    b) Bestimmen Sie zu φ_1 und φ_2 eine Spiegelung ϕ, so daß gilt: $\phi \circ \varphi_1 = \varphi_2$

    c) Finden Sie zu φ_1 und ϕ eine ONB $\{\vec{a}_1, \vec{a}_2\}$ bzw. $\{\vec{b}_1, \vec{b}_2\}$ von $\mathbf{R}^2$, bezüglich derer φ_1 bzw. ϕ durch $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ dargestellt werden.

19. Zeigen Sie: Ist φ eine eigentlich orthogonale Abbildung des $\mathbf{R}^3$, so gilt: $\varphi(\vec{a} \times \vec{b}) = \varphi(\vec{a}) \times \varphi(\vec{a})$.

20. Bestimmen Sie alle orthogonalen Abbildungen des $\mathbf{R}^3$, die sich durch eine Diagonalmatrix darstellen lassen und interpretieren Sie diese Abbildungen geometrisch!

21. Seien a und b Vektoren eines euklidischen Vektorraumes mit $|\,a\,| = 5, |\,b\,| = 8$, und $\mathrm{arc}(a, b) = {}^{\pi}/_{3}$. Man finde $|\,a + b\,|$ und $|\,a - b\,|$!

22. Sei in $\mathbf{R}^3$ ein Skalarprodukt gegeben durch $< (x_1, x_2, x_3), (y_1, y_2, y_3) >= x_1y_1 + x_1y_2 + x_2y_1 + 2x_2y_2 + 3x_3y_3$

    a) Wird dadurch $\mathbf{R}^3$ ein euklidischer Raum ?

    b) Man finde in $[\mathbf{R}^3, <, >]$ die Kosinuswerte der Winkel des Dreiecks (p_1, p_2, p_3) mit $p_1 = (1, 0, 0), p_2 = (0, 1, 0)$ und $p_3 = (0, 0, 1)$.

23. Wir betrachten $\mathbf{R}^3$ mit dem gewöhnlichen euklidischen Skalarprodukt. Berechnen Sie den Abstand der Punkte $P = (1, 2, 1)$ sowie $Q = (0, 2, 4)$ von der durch $H := \{(x_1, x_2, x_3) \in \mathbf{R}^3; 3x_1 - x_2 + 2x_3 = 1\}$ definierten Ebene an!

24. Sei $V := \{f : \mathbf{R} \to \mathbf{R}; f$ ist Polynom vom grad $\leq 3\}$.

    a) Zeigen Sie, daß durch $b(f,g) := \int_{-1}^{1} f(t) \cdot g(t)dt$ ein euklidisches Skalarprodukt in V definiert wird.

    b) Mittels des Orthogonalisierungsverfahrens von E.Schmidt überführe man die Standardbasis $\{1, x, x^2, x^3\}$ von V in eine ON-Basis bezüglich b!

25. Seien $v_1, v_2, v_3 \in \mathbf{R}^3$ Vektoren in $\mathbf{R}^3$ derart, daß die Summe $c(L) :=\mid pr_L v_1 \mid^2 +\mid pr_L v_2 \mid^2 + \mid pr_L v_3 \mid^2$ unabhängig vom 2-dimensionalen Unterraum $L \subseteq \mathbf{R}^3$ ist. Man zeige, daß dann die Vektoren v_1, v_2, v_3 paarweise orthogonal sind, und überdies die gleiche Länge haben.

Kapitel 13

Euklidische und projektive Geometrie

13.1 Euklidische Geometrie

Sei V ein euklidischer Vektorraum mit dem Skalarprodukt $\langle\ ,\ \rangle$. Sei $\{v_1, \ldots, v_n\} \subset V$ eine Basis und $b_1, \ldots, b_n \in \mathbf{R}$ gegeben. Wir suchen den Vektor x mit

$$\langle v_i, x\rangle = b_i, \quad i = 1, \ldots, n, \tag{1}$$

wir setzen dazu $x = \sum y_j v_j$ ein:

$$\langle v_i, \sum y_j v_j\rangle = \sum_j \langle v_i, v_j\rangle y_j = b_i, \tag{2}$$

d.h. die Bedingung (1) ist äquivalent zum Gleichungssystem (2) für die Koordinaten y_i von x; dessen Koeffizientenmatrix wird als Gram-Matrix bezeichnet:

$$G(v_1, \ldots, v_n) = (\,\langle v_i, v_j\rangle\,)\,.$$

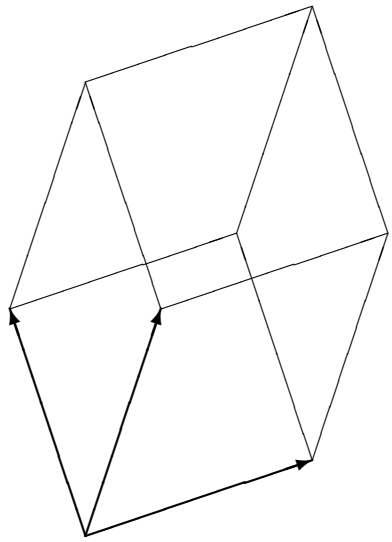
Es ist $\det(G(v_1, \ldots, v_n)) \neq 0$. Solch eine Matrix kann auch für eine beliebige Anzahl auch linear abhängiger Vektoren eingeführt werden:

$$G(v_1, \ldots, v_m) = (\,\langle v_i, v_j\rangle\,)\,.$$

Lemma 13.1 *Wenn $\{v_1, \ldots, v_m\} \subset V$ linear unabhängig ist, so ist $G(v_1, \ldots, v_m)$ positiv definit.*

Seien $r_1, \ldots, r_m \in \mathbf{R}$ nicht alle null, dann ist $x = \sum r_i v_i \neq 0$ und $0 < \langle x, x\rangle = \sum r_i r_j \langle v_i, v_j\rangle$, d.h. die Matrix $G(v_1, \ldots, v_m)$ ist positiv definit. □

Wir können die Gram-Matrix zur Volumenberechnung verwenden:



Wir stellen uns vor, daß die Vektoren $a_1, \ldots, a_m$ einen von Parallelogrammen begrenzten Körper aufspannen, so etwas nennt man ein Parallelepiped. Dann setzen wir

$$\mathrm{vol}(a_1, \ldots, a_m) = \sqrt{\det(G(a_1, \ldots, a_m)}$$

(wir wissen, daß der Radikand nicht negativ ist).

Zur Rechtfertigung dieser Definition betrachten wir Spezialfälle:
m = 1: $\mathrm{vol}(a) = \sqrt{\det(\langle a, a\rangle)} = |a|$.
m = 2: $\mathrm{vol}(a, b) = \sqrt{|a|^2 |b|^2 - \langle a, b\rangle} = |a|\,|b| \sin(\alpha)$.

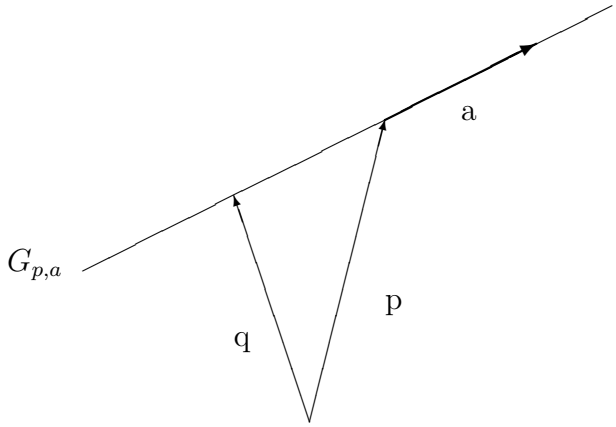
Wir überlegen, wie sich das Volumen ändert, wenn die Vektoren einer linearen Abbildung unterworfen werden.

Lemma 13.2 *Seien* $b_1, \ldots, b_m \in L(a_1, \ldots, a_m)$ *und* $a_j = \sum a_{jk} b_k$, *dann ist* $\mathrm{vol}(a_1, \ldots, a_m) = |\det(A)| \cdot \mathrm{vol}(b_1, \ldots, b_m)$ *mit* $A = (a_{ij})$.

Beweis: Es ist $\langle a_i, a_j\rangle = \langle \sum a_{il} b_l, \sum a_{jk} b_k\rangle = \sum a_{il} a_{jk} \langle b_l, b_k\rangle$, also $G(a_1, \ldots, a_m) = A \cdot G(b_1, \ldots, b_m) \cdot A^T$. □

Wir wollen nun Abstände und Winkel zwischen Geraden und Hyperebenen berechnen. Wir identifizieren dabei Punkte mit ihren „Ortsvektoren“, d.h. wir fassen den Vektorraum V als affinen Raum auf.

Eine Gerade ist durch einen ihrer Punkte p und einen Richtungsvektor a bestimmt.



Wenn $M \subset V$ eine Teilmenge und $q \in V$ ein Punkt ist, so definieren wir deren Abstand als

$$d(q, M) = \min\{|q - m| \mid m \in M\}.$$

Lemma 13.3 *Der Abstand des Punkts q von der Geraden $G_{p,a}$ ist*

$$d(q, G_{p,a}) = \frac{1}{|a|}\sqrt{|a|^2 |p-q|^2 - \langle p-q, a\rangle^2},$$

der Fußpunkt des Lotes von q auf $G_{p,a}$ ist

$$f = p - \frac{\langle p-q, a\rangle}{|a|^2}\, a.$$

Beweis: Für $r \in \mathbf{R}$ gilt

$$\begin{aligned} |(p+ra)-q|^2 &= |p-q|^2 + 2r\langle p-q, a\rangle + r^2 |a|^2 \\ &= (r|a| + \frac{\langle p-q, a\rangle}{|a|})^2 + |p-q|^2 - \frac{\langle p-q, a\rangle^2}{|a|^2} \end{aligned}$$

und das Minimum wird für $\quad r|a|^2 = -\langle p-q, a\rangle$ angenommen. □

Sei $U \subset V$ ein Unterraum mit $\dim U = \dim V - 1$ und $p \in V$, dann heißt $H = p + U$ eine Hyperebene. Es gilt $x \in H$ genau dann, wenn die Koordinaten von x *eine* Gleichung erfüllen, etwa

$$H = H_{c,r} = \{v \in V \mid \langle c, v\rangle = r\},$$

dann ist $U = L(c)^\perp$, denn es ist $u \in U$ gdw. $u = v_1 - v_2$ mit $v_i \in H$, also $\langle c, u\rangle = \langle c, v_1\rangle - \langle c, v_2\rangle = r - r = 0$.

Die Unterräume $p_1 + U_1$ und $p_2 + U_2$ sind parallel, wenn $U_1 \subset U_2$ oder $U_2 \subset U_1$ gilt, also sind zwei Geraden $G_{p,a}$ und $G_{q,b}$ genau dann parallel, wenn a und b linear abhängig sind, und zwei Hyperebenen $H_{c,r}$ und $H_{d,s}$ sind genau dann parallel, wenn c und d linear abhängig sind. Schließlich ist die Gerade $G_{p,a}$ genau dann parallel zur Hyperebene $H_{c,r}$, wenn $a \perp c$ ist.

Satz 13.4 *Sei $G = G_{p,a}$ eine Gerade und $H = H_{c,r}$ eine Hyperebene. Dann sind folgende Bedingungen äquivalent:*
1. G und H schneiden sich in einem Punkt.
2. G und H sind nicht parallel.
In diesem Fall ist der Schnittpunkt gleich $p + \frac{r-\langle p,c\rangle}{\langle a,c\rangle}\, a$.

Beweis: Wir suchen ein $x \in \mathbf{R}$ mit $\langle c, p + xa\rangle = r$, also $x\langle c, a\rangle = r - \langle c, p\rangle$, solch ein x existiert, wenn $\langle c, a\rangle \neq 0$ ist, und damit ergibt sich der obige Parameter für den Schnittpunkt. Wenn $\langle c, a\rangle = 0$ ist, so sind G und H parallel. □

Folgerung 13.5 *Das Lot vom Punkt p auf die Hyperebene $H_{c,r}$ ist gerade $G_{p,c}$, der Fußpunkt des Lots ist $f = p + \frac{r-\langle p,c\rangle}{|c|^2}\, c$ und $d(p, H) = \frac{|r-\langle p,c\rangle|}{|c|}$.* □

Satz 13.6 *Sei $U \subset V$ ein Unterraum und $\{b_1, \ldots, b_m\}$ eine Orthonormalbasis von U, dann ist für $x \in V$*

$$p_U(x) = \sum_{i=1}^{m} \langle b_i, x\rangle b_i$$

die orthogonale Projektion von x auf U und

$$d(x, U) = \sqrt{|x|^2 - |p_U(x)|^2}.$$

Beweis: Man ergänzt die Basis von U zu einer Orthonormalbasis von V. □

Damit können wir den Abstand beliebiger affiner Teilräume bestimmen. Sei $X_1 = p_1 + U_1,\ X_2 = p_2 + U_2$, dann ist

$$\begin{aligned} d(X_1, X_2) &= \min(|x_1 - x_2|) \\ &= \min(|p_1 + u_1 - p_2 - u_2|) \\ &= \min(|p_1 - p_2 - (u_2 - u_1)|) \\ &= d(p_1 - p_2, U_1 + U_2). \end{aligned}$$

Als nächstes betrachten wir Kreise (oder Sphären): Sei M ein Punkt und r eine reelle Zahl,

$$\begin{aligned} S_{M,r} &= \{x \mid d(x, m) = r\} \\ &= \{x \mid |x - M| = r\} \\ &= \{x \mid \textstyle\sum (x_i - m_i)^2 = r^2\} \end{aligned}$$

ist die Sphäre mit dem Mittelpunkt M und dem Radius r. Wir betrachten eine Gearde und einen speziellen Kreis:

$$x = p + ta, \quad |a| = 1,$$
$$|x|^2 = r^2.$$

Ein Punkt x ist ein Schnittpunkt, wenn

$$\langle p + ta, p + ta\rangle = r^2$$

d.h.

$$|p|^2 + 2\langle p, a\rangle t + t^2 - r^2 = 0, \tag{1}$$

also

$$(t + \langle p, a\rangle)^2 = \langle p, a\rangle^2 - (|p|^2 - r^2).$$

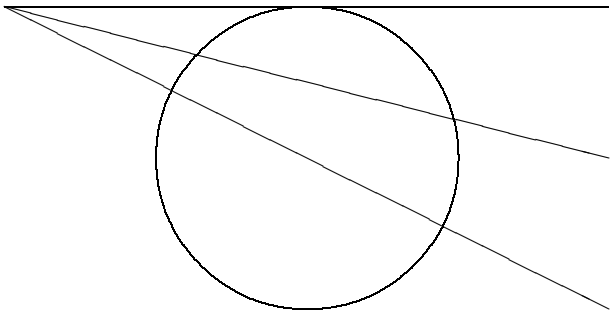
In Abhängigkeit vom Vorzeichen der rechten Seite gibt es zwei, einen oder keinen Schnittpunkt. Die Abstände von p zu den Schnittpunkten sind gleich den Lösungen der quadratischen Gleichung (1):

$$d(p, x_1) = |t_1 a| = |t_1|\,,$$

$$d(p, x_2) = |t_2 a| = |t_2|\,,$$

ihr Produkt ist gleich dem konstanten Term in (1), also

$$d(p, x_1)d(p, x_2) = |t_1 t_2| = |p|^2 - r^2.$$



Diese Konstante hängt nicht von der Richtung der Geraden ab, sondern nur vom Punkt p. Diese Aussage ist als Sekanten-Tangentensatz bekannt.

Für die folgenden Untersuchungen benötigen wir als Hilfsmittel das Vektorprodukt von Vektoren im $\mathbf{R}^3$. Sei $\{i, j, k\}$ eine Orthonormalbasis des $\mathbf{R}^3$, dann war fr $a, b \in \mathbf{R}^3$ als

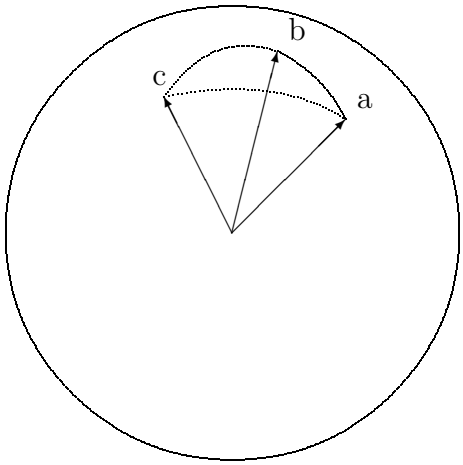
$$a \times b = \det \begin{pmatrix} i & j & k \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix}$$

definiert. Wir werden nun einige Identitäten für Vektorprodukte herleiten.
Wir erinnern daran, daß das Produkt $a \times b$ linear in a und in b ist, daher genügt es, wenn die zu beweisenden Gleichungen nachgewiesen werden, wenn die einzelnen Faktoren Basiselemente sind. Somit erhalten wir

1. $\langle a \times b, c\rangle = \det(a\ b\ c)$
2. $a \times (b \times c) = \langle a, c\rangle b - \langle a, b\rangle c$ Graßmann-Identität
3. $a \times (b \times c) + b \times (c \times a) + c \times (a \times b) = 0$ Jacobi-Identität
4. $(a \times b) \times (c \times d) = \langle a \times b, d\rangle c - \langle a \times b, c\rangle d = \det(a\ b\ d)c - \det(a\ b\ c)d$
5. $\langle a \times b, c \times d\rangle = \langle a, c\rangle\langle b, d\rangle - \langle b, c\rangle\langle a, d\rangle$

13.2 Sphärische Geometrie

Als nächstes wollen wir uns mit der Geometrie auf der Oberfläche der Einheitskugel beschäftigen.



Ein sphärisches Dreieck ist durch drei Einheitsvektoren a, b, c gegeben. Wir wählen die Reihenfolge so, daß $\det(a\ b\ c) > 0$ ist. Die „Seiten“ des Dreiecks sind die Winkel A, B, C zwischen den Vektoren a, b, c, die „Winkel“ α, β, γ des Dreiecks sind die Winkel zwischen den Ebenen $L(a, b),\ L(b, c),\ L(a, c)$, also die Winkel zwischen deren Normalenvektoren. Da die Vektoren den Betrag 1 haben, gelten folgende Beziehungen:

$$\cos A = \langle b, c\rangle \quad \cos B = \langle a, c\rangle \quad \cos C = \langle b, a\rangle$$

$$\sin A = |a \times c| \quad \sin B = |a \times c| \quad \sin C = |a \times b|$$

$$\cos\alpha = \frac{\langle a \times c, a \times b\rangle}{|a \times c|\,|a \times b|} \quad \cos\beta = \frac{\langle b \times a, b \times c\rangle}{|b \times a|\,|b \times c|} \quad \cos\gamma = \frac{\langle c \times b, c \times a\rangle}{|c \times b|\,|c \times a|}$$

Aus den obigen Formeln für das Vektorprodukt folgt

$$\begin{aligned} |(a \times b) \times (a \times c)| &= |\det(a\ b\ c)|\,|a| \\ &= \det(a\ b\ c) \\ &= |a \times b|\,|a \times c| \sin\alpha \\ &= \sin C \cdot \sin B \cdot \sin\alpha. \end{aligned}$$

Daraus folgt der spärische Sinussatz:

Satz 13.7

$$\frac{\sin\alpha}{\sin A} = \frac{\det(a\ b\ c)}{\sin A \sin B \sin C} = \frac{\sin\beta}{\sin B} = \frac{\sin\gamma}{\sin C}. \square$$

Wenn die Seiten klein sind, so können wir sie die Sinuswerte durch die Argumente ersetzen und erhalten den ebenen Sinussatz.

Wir erhalten zwei Cosinussätze:

Satz 13.8 *1.* $\cos A = \cos B \cos C + \sin B \sin C \cos\alpha$,
2. $\sin C \cos B = \sin B \cos C \cos\alpha + \sin A \cos\beta$.

Aus der ersten Formel geht hervor, daß man die Winkel aus den Seiten berechnen kann.
Beweis: 1. $\sin B \sin C \cos\alpha = |a \times c|\,|a \times b| \cos\alpha = \langle a \times c, a \times b\rangle = \langle a, a\rangle\langle b, c\rangle - \langle a, b\rangle\langle a, c\rangle = \cos A - \cos B \cos C$.
2. Die Behauptung ist genau dann richtig, wenn

$$|a \times b|\,\langle a, c\rangle = |a \times c|\,\langle a, c\rangle \frac{\langle a \times c, a \times b\rangle}{|a \times c|\,|a \times b|} + |b \times c|\,\frac{\langle b \times a, b \times c\rangle}{|b \times a|\,|b \times c|}$$

gdw.

$$|a \times b|^2\,\langle a, c\rangle = \langle a, b\rangle\langle a \times c, a \times b\rangle + \langle b \times a, b \times c\rangle,$$

die linke Seite ist gleich $(1 - \langle a, b\rangle^2)\langle a, c\rangle$, die rechte Seite ist gleich

$$\langle a, b\rangle(\langle a, a\rangle\langle c, b\rangle - \langle c, a\rangle\langle a, b\rangle) + \langle b, b\rangle\langle a, c\rangle - \langle a, b\rangle\langle b, c\rangle,$$

dies stimmt mit der linken Seite überein. □

Das Dreieck, dessen definierenden Vektoren senkrecht auf den Seiten des durch a, b, c gegebenen Dreiecks stehen, wird das Polardreieck genannt, es hat die Ecken

$$a' = \frac{b \times c}{|b \times c|}, \quad b' = \frac{c \times a}{|c \times a|}, \quad c' = \frac{a \times b}{|a \times b|}$$

die Seiten A', B', C' und die Winkel α', β', γ'.

Satz 13.9 (Viètá-Formeln)

$$\cos A' = -\cos\alpha,\ \cos\alpha' = -\cos A.$$

Beweis:

$$\cos A' = \langle b', c'\rangle = \frac{\langle c \times a, a \times b\rangle}{|c \times a|\,|a \times b|}.\square$$

Als Folgerung erhalten wir den polaren Cosinussatz:

Folgerung 13.10

$$-\cos\alpha = \cos\beta\cos\gamma + \sin\beta\sin\gamma\cos A.\square$$

Das bedeutet, daß die Seiten des Dreiecks aus den Winkeln berechnet werden können. Es gibt also keine ähnlichen Dreiecke, die nicht kongruent sind.

Wenn man die geografischen Längen L_i und Breiten B_i zweier Orte kennt, so kann man mit dem ersten Cosinussatz deren Entfernung berechnen. Man betrachtet das Dreieck, das von beiden Orten und dem Nordpol gebildet wird, dessen zwei Seiten sind gleich $\pi/2 - B_i$ und der der gesuchten Seite gegenüberliegende Winkel ist gleich $L_1 - L_2$.
Beispiel: Paris: $(2,3°; 48,8°)$, Berlin: $(13,4°; 52,5°)$, damit berechnet man $\cos A = 0,99...$, also $A = 7,939°$, bei einem Erdradius von 6378 km ergibt dies eine Entfernung von 884 km.

13.3 Konvexe Mengen und lineare Ungleichungssysteme

Sei V ein Vektorraum und $U \subset V$ ein Unterraum, dann ist die Menge $v + U$ ein affiner Unterraum des affinen Raums (V, V), wir nennen sie eine affine Teilmenge. Wenn $M \subset V$ eine beliebige Teilmenge ist, so sei I_M eine derartige Indexmenge, daß

$$\{A_i \mid i \in I_M\} = \{A \supseteq M \mid A \text{ affiner Unterraum}$$

die Menge aller affinen Unterräume von V ist, die M umfassen. Den Durchschnitt all dieser Mengen

$$A(M) = \bigcap_{i \in I_M} A_i$$

bezeichnen wir als die affine Hülle von M.

Lemma 13.11 *$A(M) = m + \bigcap_{i \in I_M} U_i$ mit $m \in M$.*

Beweis: Sei $m \in M$ beliebig, dann ist $m + \bigcap U_i \subset m + U_j$ für alle $j \in I_M$, also $m + \bigcap U_i \subset A(M)$.
Sei andererseits $x \in A(M)$, dann ist $x \in m + U_i$, also $x - m \in U_i$ für alle i, folglich ist $x - m \in \bigcap U_i$, also $x \in m + \bigcap U_i$ und damit $A(M) \subset m + \bigcap U_i$. □

Wir bezeichnen den Unterraum $\bigcap U_i$ mit U_M.

Folgerung 13.12 *Sei $M = \{m_1, \ldots, m_k\}$.Es ist $U_M = L(\{m_i - m_1\}) = \{\sum r_i m_i \mid \sum r_i = 0\}$.*

Beweis: Der zu einem affinen Unterraum gehörige Vektorraum besteht aus den Verbindungsvektoren der Punkte. Jeder Vektor $u \in U_M$ hat die Gestalt $u = \sum r_i(m_i - m_1) = \sum r_i m_i - (\sum r_i)m$ und die Summe aller Koeffizienten ist Null. □

Folgerung 13.13 *$A(M) = \{\sum r_i m_i \mid \sum r_i = 1\}$.* □

Definition: Eine Teilmenge $K \subset V$ heißt konvex, wenn mit $u, v \in K,\ r, s \in \mathbf{R},\ r, s \geq 0,\ r + s = 1$ auch $ru + sv \in K$ ist.

Eine konvexe Teimenge enthält also mit je zwei Punkten auch deren Verbindungsstrecke.

Beispiele:
1. Jeder Unterraum ist konvex.
2. Die Menge $\{v \mid |v| \leq 1\}$ ist konvex, denn sei $|u| \leq 1, |v| \leq 1$, dann ist $|ru + sv| \leq |r|\,|u| + |s|\,|v| \leq r + s = 1$.

Definition: Seien $v_1, \ldots, v_n \in V,\ r_1, \ldots, r_n \in \mathbf{R}$, dann heißt $v = \sum r_i v_i$ mit $r_i \leq 0, \sum r_i = 1$ eine konvexe Linearkombination der v_i.

Lemma 13.14 *Eine konvexe Menge enthält alle konvexen Linearkombinationen seiner Elemente.*

Beweis: Sei K eine konvexe Menge. Nach Definition enthält K jede konvexe Linearkombination zweier seiner Elemente. Es sei schon gezeigt, daß K jede konvexe Linearkombination von $m-1$ seiner Elemente enthält, und seien $v_1, \ldots, v_m \in K$. Sei $r_i \geq 0, \sum r_i = 1$, wir betrachten den Vektor $v = \sum r_i v_i$. Falls $r_m = 1$ ist, so ist $v = v_m \in K$. Andernfalls ist $r_m < 1$, wir setzen $r = \sum_{i=1}^{m-1} r_i = 1 - r_m$, dann ist $v = r \sum \frac{r_i}{r} v_i + r_m v_m$, wobei $\sum \frac{r_i}{r} = \frac{\sum r_i}{\sum r_i} = 1$ ist, der erste Summand liegt nach Voraussetzung in K, also ist v als konvexe Linearkombination zweier Elemente von K auch in K enthalten. □

Definition: $K(M) = \bigcap\{K \mid M \subset K, K \text{ konvex}\}$ heißt die konvexe Hülle von K.

Satz 13.15 *$K(M)$ ist die Menge aller (endlichen) konvexen Linearkombinationen von Elementen aus M.*

Beweis: Die Menge

$$K_0 = \{v = \sum r_i v_i \mid v_i \in M, \sum r_i = 1\}$$

umfaßt M und ist konvex, denn seien $v = \sum r_i v_i,\ w = \sum s_i w_i \in M, r + s = 1$, dann ist $rv + sw = \sum r r_i v_i + \sum s s_i w_i$ und $\sum r r_i + \sum s s_i = r \sum r_i + s \sum s_i = r + s = 1$. □

Definition: Eine Linearkombination $\sum r_i v_i$ heißt positiv, wenn $r_i \geq 0$ ist.
Wenn $M = \{v_1, \ldots, v_m\}$ eine endliche Menge ist, so heißt $K(M)$ ein konvexes Polyeder.

Ein Element eines konvexen Polyeders $K(M)$ heißt Ecke, wenn es nicht als konvexe Linearkombination anderer Elemente von $K(M)$ dargestellt werden kann.

Satz 13.16 *Die Ecken von $K(v_1, \ldots, v_m)$ sind genau diejenigen v_i, die sich nicht als konvexe Linearkombination der restlichen v_j darstellen lassen.*

Beweis: Es sei v_m keine konvexe Linearkombination von $v_1, \ldots, v_{m-1}$, wir zeigen, daß v_m auch keine konvexe Linearkombination anderer Elemente ist.
Angenommen, es gilt $v_m = \sum_{i=1}^{k} a_i w_i,\ \sum a_i = 1, a_i > 0$, es sei $w_i = \sum_{j=1}^{m} a_{ij} v_j,\ \sum a_{ij} = 1$. Dann gilt $v_m = \sum_j \sum_i a_i a_{ij} v_j$, wir setzen $r_j = \sum_i a_i a_{ij}$, es gilt $\sum r_j = 1$. Falls $r_m = \sum a_i a_{im} = 1$ gilt, so muß $a_{im} = 1$ für ein i und $a_{ij} = 0$ für $j < m$ gelten, also wäre $v_m = w_i$ nicht als konvexe Linearkombination dargestellt.
Also gilt $r_m < 1$ und wir haben v_m als Linearkombination von $v_1, \ldots, v_{m-1}$ dargestellt:

$$(1 - r_m) v_m = \sum_{i=1}^{m-1} r_i v_i.$$

Dies ist eine konvexe Linearkombination, wie man leicht nachrechnet. Damit erhalten wir einen Widerspruch. □

Definition: Sei $M = \{m_0, \ldots, m_k\}$; wenn $\dim A(M) = s$ ist, so heißt $K(M)$ ein s-dimensionales Polyeder. Wenn $\dim K(M) = k$ ist, so heißt es ein k-Simplex. Sei $S = K(x_0, \ldots, x_k)$ ein k-Simplex, dann ist $K(x_{i_0}, \ldots, x_{i_r})$ ein r-Simplex, es heißt Seite von S.

Satz 13.17 *Der Durchschnitt eines Simplex und eines Untervektorraums ist ein konvexes Polyeder oder leer.*

Beweis: Sei $P = K(x_0, \ldots, x_m)$ ein Simplex und $U \subset V$ ein Unterraum. Jeder Punkt $x \in \cap U$ liegt auf gewissen Seiten von P, sei S_x die Seite kleinster Dimension, die x enthält. Wir nennen einen Punkt $x \in P \cap U$ markant, wenn $S_x \cap U = \{x\}$ ist, also wenn $S_x \not\subset U$ ist. Die Zahl der markanten Punkte ist endlich, da die Zahl der Seiten endlich ist. Wir zeigen

$$P \cap U = K(\{x \mid x \text{ markant})\}.$$

Wenn $x \in P \cap U$ nicht markant ist, so enthält S_x noch einen Punkt z aus $P \cap U$. Dann gilt $x, z \in U$, also ist $y = z - x \in U$ nicht der Nullvektor. Es sei $S_x = K(x_0, \ldots, x_r)$ und $x = \sum a_k x_k$ mit $\sum a_k = 1$, dann ist $a_k > 0$ für alle k wegen der Minimalität von S_x. Sei $z = \sum b_k x_k,\ b_k \geq 0,\ \sum b_k = 1$. Dann ist $y = \sum (b_k - a_k) x_k$, wir setzen $c_k = b_k - a_k$, dann gilt $\sum c_k = 0$, aber nicht alle c_k sind null. Also ist mindestens eins der c_k positiv und mindestens eins ist negativ. Sei

$$a = \min_{c_k < 0} \left(-\frac{a_k}{c_k} \right),$$

diese Zahl ist positiv. Dann gilt

$$a_k + a c_k \geq 0,\ k = 0, \ldots, r$$

und das Gleichheitszeichen kommt für ein k vor. Analog gibt es ein $b > 0$ mit

$$a_k - b c_k \geq 0,\ k = 0, \ldots, r$$

und auch hier kommt das Gleichheitszeichen vor. Wir betrachten

$$u = x + ay = \sum (a_k + a c_k) x_k,$$

hier sind die Koeffizienten nicht negativ und die Koeffizientensumme ist gleich 1. Analog sei

$$v = x - by = \sum (a_k - b c_k) x_k,$$

auch dies ist eine konvexe Linearkombination der x_k. Nun liegen aber u und v auf einer echten Seite von S_x, denn jeweils ein Koeffizient ist null. Folglich ist

$$x = \frac{1}{a+b}(bu + av)$$

eine konvexe Linearkombination von Elementen, die auf niedrigerdimensionalen Seiten als x liegen. Wenn $u,\ v$ markant sind, so sind wir fertig. Wenn nicht, so zerlegen wir, wie soeben x, nun u und v. □

Definition: Eine Menge der Form $P = P(v_0, \ldots, v_m) = \{\sum r_i v_i \mid r_i \geq 0\}$ heißt konvexe Pyramide. Zwei Vektoren u, v heißen positiv parallel, wenn ein $r > 0$ existiert, so daß $ru = v$ ist. Eine Pyramide P heißt spitz, wenn $P \cap (-P) = \{0\}$ ist.
OBdA seien unter den v_i keine positiv parallelen Vektoren; dann heißen diese die Kanten von P. Keine Kante ist eine positive Linearkombination der restlichen.

Seien $H_1, \ldots, H_k \subset \mathbf{R}^n$ Hyperebenen, die jeweils durch eine homogene Gleichung $f_i(x) = 0$ beschrieben seien. Dann ist $H = \bigcap H_i$ ein Unterraum, wir betrachten $H^+ = \{x \in H \mid x_i \geq 0\}$. Die Elemente von H^+ können also durch Gleichungen und Ungleichungen beschreiben werden.

Satz 13.18 *H^+ ist eine spitze konvexe Pyramide.*

Beweis: Sei $\{e_1, \ldots, e_n\}$ die kanonische Basis, sie erzeugen das $(n-1)$-Simplex $P = K(e_1, \ldots, e_n)$ und der Durchschnitt $P \cap H$ ist ein konvexes Polyeder, also von der Form $K(x_0, \ldots, x_m)$. Nun ist aber H^+ die positive Hülle der x_i, H^+ ist spitz, da in H^+ nur Vektoren mit nichtnegativen Komponenten, in $-H^+$ aber nur Vektoren mit nichtpositiven Komponenten vorkommen. □

Wir betrachten jetzt ein lineares Ungleichungssystem, dies ist eine Folge von Bedingungen der Form

$$f(x) \geq a,\ g(x) \leq b,\ h(x) = c,$$

wo $f,\ g,\ h$ Linearformen sind.

Durch Einfürung neuer Unbekannter und neuer Bedingungen können wir das Bild vereinheitlichen:
$f(x) \geq a$ ist äquivalent zu $f(x) - y = a,\ y \geq 0$,
$g(x) \leq b$ ist äquivalent zu $g(x) + z = a,\ z \geq 0$,
wir können die Ungleichungen also durch Gleichungen und Positivitätsforderungen an die Unbekannten ersetzen. Wenn an eine Unbekannte z keine Positivitäsforderung gestellt ist, so ergänzen wir $z = z' - z'',\ z' \geq 0,\ z'' \geq 0$.

Somit können wir eine einfache Form des Ungleichungssystems annehmen:

$$Ax = b,\ x_i \geq 0,\ (i = 1, \ldots, n).$$

Satz 13.19 *Die Menge der Lösungen des homogenen Ungleichungssystems $Ax = 0,\ x_i \geq 0$ ist eine spitze konvexe Pyramide.*

Beweis: Die Lösungsmenge hat die Form H^+ (s.o.). □

Zum inhomogene Ungleichungssystem

$$Ax = b,\ x_i \geq 0 \tag{1}$$

betrachten wir das Ungleichungssystem

$$AZ - bz_0 = 0,\ z_i \geq 0, \tag{2}$$

hierbei sei $Z = (z_1, \ldots, z_n)^T$, wir betrachten die Lösungen von (2) mit $z_0 > 0$, aus diesen erhalten wir Lösungen von (1): $x_i = \frac{z_i}{z_0}$.

Satz 13.20 *Die Lösungsmenge eines inhomogenen linearen Ungleichungssystems ist eine direkte Summe eines konvexen Polyeders und einer konvexen Pyramide.*

Beweis: Die Menge aller Lösungen $Z = (z_1, \ldots, z_n, z_0)$ von (2) bilden eine konvexe Pyramide, diese habe die Kanten

$$Z_k = (z_{1k}, \ldots, z_{nk}, z_{0k}),\ k = 1, \ldots, s,$$

also hat jede Lösung die Form $Z = \sum r_i Z_i,\ r_i \geq 0$.

Falls für $k = 1, \ldots, s$ stets $z_{0k} = 0$ ist, so ist $z_0 = 0$ für jede Lösung von (2), d.h. (1) besitzt keine Lösung.
Sei also oBdA $z_{01} > 0, \ldots, z_{0r} > 0,\ z_{0,r+1} = \ldots = 0$, es ist $z_j = \sum r_i z_{ji}$ und damit

$$x_j = \frac{z_j}{z_0} = \sum_{k=1}^{r} r_k \frac{z_{0k}}{z_0} \frac{z_{jk}}{z_{0k}} + \sum_{k=r+1}^{s} \frac{r_k}{z_0} z_{jk},$$

die zweite Summe beschreibt ein Element einer Pyramide; wir betrachten die Koeffizientensumme der ersten Summe:

$$\sum_{j=1}^{r} r_k \frac{z_{0k}}{z_0} = \frac{1}{z_0} \sum r_k z_{0k} = \frac{z_0}{z_0} = 1,$$

also beschreibt die erste Summe ein Element eines konvexen Polyeders. □

13.4 Projektive Geometrie

Definition: Sei V ein Vektorraum, $P(V)$ die Menge aller 1-dimensionalen Unterräume von V, $G(V)$ sei die Menge aller 2-dimensionalen Unterräume von V und ϵ sei die Inklusionsrelation auf $P(V) \times G(V)$, also $p \,\epsilon\, g \Longleftrightarrow p \subset g$. Dann heißt $(P(V), G(V), \epsilon)$ der projektive Raum über V. Die Elemente von $P(V)$ heißen Punkte, die Elemente von $G(V)$ heißen Geraden. Der Punkt p liegt auf der Geraden g, wenn $p \,\epsilon\, g$, d.h. $p \subset g$ gilt.
Sei $U \subset V$ ein Unterraum, dann gilt $P(U) \subset P(V)$ und $G(U) \subset G(V)$ und $(P(U), G(U), \epsilon\,|_{P(U)\times G(U)})$ heißt projektiver Unterraum. Wir setzen $\dim(P(V)) = \dim V - 1$.

Lemma 13.21 *$P(U)$ und $P(W)$ seien Unteräume von $P(V)$. Wenn $P(U) = P(W)$ ist, dann gilt $U = W$.*

Beweis: Sei $P(U) \subset P(W)$ und $o \neq u \in U$, dann hat $p = \mathbf{R}u$ die Dimension 1, also gilt $p \in P(U)$, also $p \in P(W)$ und damit $U \subset W$. □

Satz 13.22 *$P(V)$ sei ein projektiver Raum, dann gilt:*
(1) Zu zwei Punkten p, q gibt es genau eine Gerade $g = (p, q)$ mit $p \,\epsilon\, g,\ q \,\epsilon\, g$.
(2) Auf jeder Geraden liegen mindestens drei Punkte.
(3) Seien p, q, r, s verschiedene Punkte. Wenn die Geraden (p, q) und (r, s) einen Schnittpunkt haben, so schneiden sich auch (p, r) und (q, s).

Beweis: 1. Seien $p, q \in P(V),\ p \neq q$, dann ist der Unterraum $p + q \subset V$ 2-dimensional und es gilt $p \subset p + q,\ q \subset p + q$, also ist $p + q \in G(V)$ und $p \,\epsilon\, p + q,\ q \,\epsilon\, p + q$ und $(p, q) = p + q$ ist der einzige zweidimensionale Unterraum von V mit dieser Eigenschaft.
2. Sei $g = \mathbf{R}x + \mathbf{R}y$ eine Gerade, dann sind x, y linear unabhängig, also sind $\mathbf{R}x,\ \mathbf{R}y,\ \mathbf{R}(x + y)$ drei verschiedene Punkte auf g.
3. Sei $p = \mathbf{R}u,\ q = \mathbf{R}v,\ r = \mathbf{R}x,\ s = \mathbf{R}y$, nach Voraussetzung sind jeweils zwei der Vektoren u, v, x, y linear unabhängig. Es ist $(p, q) = \mathbf{R}u + \mathbf{R}v,\ (r, s) = \mathbf{R}x + \mathbf{R}y$. Wenn

(p,q) und (r,s) sich schneiden, so ist $(\mathbf{R}u+\mathbf{R}v)\cap(\mathbf{R}x+\mathbf{R}y)$ 1-dimensional, enthält also einen vom Nullvektor verschiedenen Vektor

$$z = au+bv = cx+dy,\ a,b,c,d\in\mathbf{R}.$$

Dann ist

$$z' = au-cx = -bv+dy\ \epsilon\ (p,r)\cap(q,s),$$

denn falls $z'=o$ wäre, so wäre wegen $a=b=c=d=0$ auch $z=o$. □

Alle im Folgenden zu beweisenden Aussagen können aus den drei Aussagen dieses Satzes hergeleitet werden, diese Aussagen könnten also als Axiome der projektiven Geometrie dienen.

Wir beweisen einige elementare geometrische Eigenschaften.

Satz 13.23 *(4) Zwei verschiedene Geraden g,h schneiden sich höchstens in einem Punkt.*
(5) Sei $\dim(P(V))=2$ *(d.h. $P(V)$ ist eine projektive Ebene), dann schneiden sich zwei verschiedene Geraden genau in einem Punkt.*

Beweis: 4. Seien $p,q\in P(V)$ mit $p\ \epsilon\ h,\ q\ \epsilon\ h,\ p\ \epsilon\ g,\ q\ \epsilon\ g,\ p\neq q$, also $h=(p,q),\ g=(p,h)$, wegen (1) folgt $g=h$.
5. Seien $g,h\in G(V)$. Wenn $g\cap h=\{o\}$, dann ist

$$\dim g+\dim h = 4 > dimV = 3,$$

aus dem Widerspruch folgt $\dim(g\cap h)=1$ und der Schnittpunkt ist wegen (4) eindeutig bestimmt. □

Satz 13.24 *Seien p,q,r drei Punkte, dann ist $p\ \epsilon\ (q,r)$ genau dann, wenn $q\ \epsilon\ (p,r)$.*

Beweis: Sei $p\ \epsilon\ (q,r)$, dann ist auch $p\ \epsilon\ (p,r)$, beide Geraden enthalten p und r, also ist $(q,r)=(p,r)$ und damit $q\ \epsilon\ (p,r)$. □

Lemma 13.25 *$P(U)$ sei ein Unterraum von $P(V)$, dann gilt: (6) Wenn $p,q\in P(U),\ p\neq q$, dann ist $(p,q)\in G(U)$.*
(7) Sei $g\in G(U),\ p\in P(V)$, wenn $p\ \epsilon\ g$, so ist $p\in p(U)$.

Beweis: 6. Wir haben $p,q\subset U$, also $p+q\subset U$, also ist $(p,q)=p+q\in G(U)$.
7. Wegen $p\ \epsilon\ g$ ist $p\subset g\subset U$, also $p\subset U$. □

Wir können Unterräume eines projektiven Raums wie folgt charakterisieren:

Satz 13.26 *Seien $P\subset P(V),\ G\subset G(V)$ Teilmengen mit folgenden Eigenschaften:*
(a) wenn $p,q\in P$ und $p\neq q$ ist, so ist $(p,q)\in G$,
(b) sei $g\in G$ und $p\in P(V)$, wenn $p\ \epsilon\ g$ ist, so ist $p\in P$.
Dann gibt es einen Unterraum $U\subset V$ mit $P=P(U),\ G=G(U)$.

Beweis: Wir setzen $U = \{u \in V \mid u = o \text{ oder } \mathbf{R}u \in P\}$ und zeigen, daß dies ein Unterraum ist.
Seien $u, v \in U$, also $\mathbf{R}u, \mathbf{R}v \in P$. Wenn u, v linear abhängig sind, $\mathbf{R}(u+v) = \mathbf{R}u \in P$, also $u + v \in U$. Seien nun u, v linear unabhängig, dann ist $\mathbf{R}u + \mathbf{R}v \in G$ nach (a). Weiter ist $\mathbf{R}(u+v) \,\epsilon\, \mathbf{R}u + \mathbf{R}v$, wegen (b) ist also $\mathbf{R}(u+v) \in P$, d.h. $u + v \in U$.
Für $r \in \mathbf{R}$ ist $ru \in U$, da $\mathbf{R}ru = \mathbf{R}u$.
Nach Konstruktion ist $P(U) = P$ und aus (a) und (7) gilt $G(U) \subset G$; aus (6) und (b) folgt $G \subset G(U)$. □

Definition: Seien $X, Y \in P(V)$ Teilmengen, wir setzen

$$X + Y = \{p \in P(V) \mid \text{ es gibt } x \in X,\ y \in Y \text{ mit } p \,\epsilon\, (x, y)\} \cup X \cup Y,$$

$2X = X + X, \ldots, nX = (n-1)X + X$,
$H(X) = \bigcup_n nX$ heißt die lineare Hülle von X.

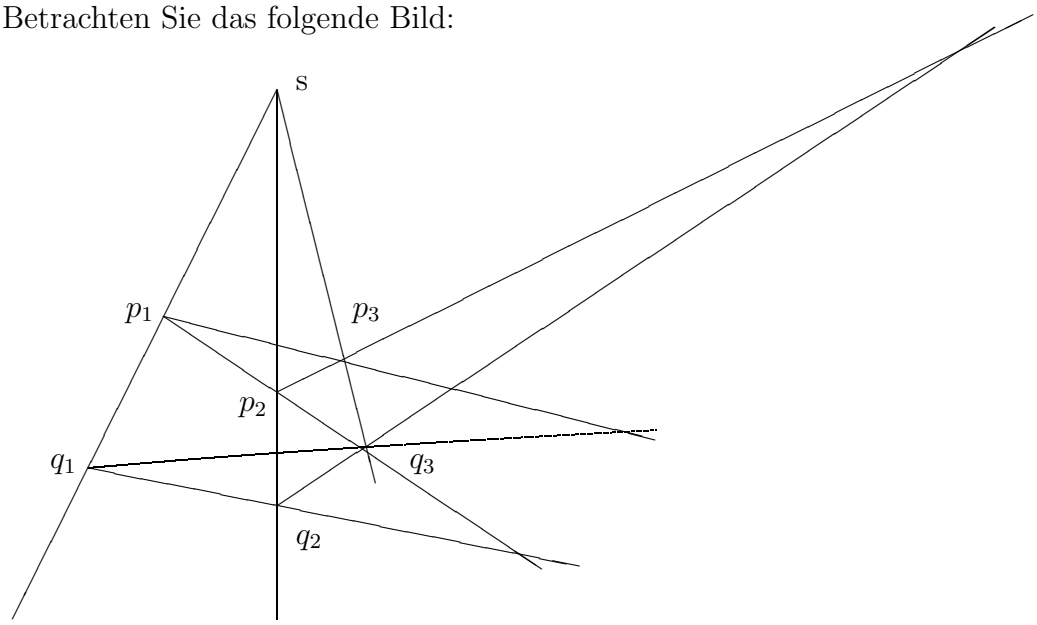
Lemma 13.27 $(X + Y) + Z = X + (Y + Z)$ □

Satz 13.28 *Seien* $p_0, \ldots, p_n \in P(V)$, $p_i = \mathbf{R}x_i$, $L(x_0, \ldots, x_n) = U$, *dann gilt* $H(p_0, \ldots, p_n) = P(U)$, $\dim(H(p_0, \ldots, p_n) \leq n$. □

Folgerung 13.29 $\dim P + \dim Q = \dim(p * q) + \dim(P \cap Q)$. □

Sei nun $H \subset P(V)$ eine Hyperebene und g eine Gerade, dann ist $\dim(g \cap H) = \dim H + \dim g - \dim(H + g) \geq n - 1 + 1 - n = 0$, also schneiden sie sich in jedem Fall. In einem projektiven Raum gibt es keine Parallelität.

Betrachten Sie das folgende Bild:



Satz 13.30 (Desargues(1591-1661)) *Seien* $p_1, p_2, p_3, q_1, q_2, q_3, s$ *paarweise verschiedene Punkte und* $s \,\epsilon\, (p_i, q_i)$, $i = 1, 2, 3$. *Dann gibt es kollineare (d.h. auf einer Geraden gelegene Punkte)* b_{12}, b_{23} b_{13} *mit* $b_{ij} \,\epsilon\, (p_i, p_j)$ *und* $b_{ij} \,\epsilon\, (q_i, q_j)$.

Beweis: Sei $p_i = \mathbf{R}x_i,\ q_i = \mathbf{R}y_i,\ s = \mathbf{R}z$, dann sind jeweils x_i, y_i linear unabhängig. Es ist aber $z \in L(x_i, y_i)$, oBdA können wir $z = x_i - y_i,\ (i = 1, 2, 3)$ annehmen, also

$$x_i - x_j = y_j - y_i.$$

Die gesuchten Schnittpunkte sind dann

$$b_{ij} = \mathbf{R}(x_i - x_j)\ \epsilon\ \mathbf{R}x_i + \mathbf{R}x_j = (p_i, p_j)$$

und

$$\mathbf{R}(y_i - y_j)\ \epsilon\ \mathbf{R}y_i + \mathbf{R}y_j = (q_i, q_j).$$

Weiter ist $x_1 - x_3 = (x_1 - x_2) + (x_2 - x_3)$, also

$$\mathbf{R}(x_1 - x_3) \subset \mathbf{R}(x_1 - x_2) + \mathbf{R}(x_2 - x_3),$$

also $b_{13}\ \epsilon\ (b_{12}, b_{23}$. □

Wir wollen nun einen Zusammenhang mit den von früher bekannten affinen Räumen herstellen.

Satz 13.31 *Sei $V = W \oplus \mathbf{R}a$ und $A = P(V)\backslash P(W)$, dann gibt es zu jedem Punkt $p \in A$ genau einen Vektor $f(p) \in W$ mit $p = \mathbf{R}(f(p) + a)$.*

Beweis: Sei $p = \mathbf{R}x \in A$, dann ist $\mathbf{R}x \not\subset W$, also gibt es ein eindeutig bestimmtes $w \in W$ und $t \in \mathbf{R}$ mit $x = w + ta$, wir setzen $f(p) = \frac{1}{t}w$. Dann gilt $\frac{1}{t}x = f(p) + a$. Wenn $y = sw,\ (s \neq 0)$ ein anderer Repräsentant von p ist, so gilt $y = sw + sta$, also ist $f(p) = \frac{1}{sr}sw = \frac{1}{t}w$ wohldefiniert. □

Satz 13.32 *Sei $P(W)$ eine projektive Hyperebene in $P(V)$. Dann ist $A = P(V)\backslash P(W)$ ein affiner RAum mit dem Translationsraum W.*

Wir nennen $P(W)$ die bezüglich A uneigentliche Hyperebene.
Beweis: Wegen $\dim W = \dim V - 1$ ist $V = W \oplus \mathbf{R}a$ für ein $a \in V\backslash W$, hierzu haben wir die obige Abbildung $f : A \longrightarrow W$. Sei $p = \mathbf{R}x,\ w \in W$, wir setzen $q = \mathbf{R}(w + tx)$ mit $tx = f(p) + a$, dann ist$w + tx = (w + f(p)) + a = f(q) + a$ (nach Konstruktion von f), also setzen wir $\overrightarrow{pq} = w = f(q) - f(p)$. Man rechnet leicht nach, daß $\overrightarrow{pq} = \overrightarrow{pr} + \overrightarrow{rq}$ gilt. □

Folgerung 13.33 *Sei $g \in G(V),\ g \not\subset P(W)$, dann ist $g\backslash P(W) = g \cap A$ eine affine Gerade und $g \cap P(W)$ besteht aus genau einem Punkt.*

Beweis: Es ist $\dim W = \dim V - 1,\ \dim g = 2$, also $\dim(g \cap W) = 1$. Sei $g = \mathbf{R}a + \mathbf{R}w$ mit $a \notin W$, wir können $w \in W$ wählen. Seei $p_0 = \mathbf{R}a \in g \cap A$, dann ist ein beliebiger Punkt $p \in g \cap A$ von der Form $p + \mathbf{R}(tw + a)$ (0BdA kann man den FAktor von a gleich 1 wählen), d.h. $f(p) = tw$ ist der Verbindungsvektor von $p_)$ zu p, also gilt (im affinen Sinn) $p = p_0 + tw$, diese Punkte bilden eine affine Gerade. Weiter ist $g \cap P(W) \neq \emptyset$, und wenn im Durchschnitt zwei Punkte lägen, so läge g in $P(W)$. □

Satz 13.34 *Wenn $h \subset A$ eine affine Gerade ist, dann gibt es genau eine projektive Gerade $\bar{h}$ mit $h = \bar{h} \cap A$.*

Beweis: Sei $h = \mathbf{R}w + p_0 \subset A$, $p_0 = \mathbf{R}a \in A$, dann ist $\bar{h} = \mathbf{R}a + \mathbf{R}w$ eine projektive Gerade. Jeder Punkt $\mathbf{R})tw+a)$ von h liegt in $\bar{h} \cap A$, also ist $h \subset \bar{h} \cap A$. Wenn umgekehrt $\epsilon\ \bar{h} \cap A$ ist, so ist $p = \mathbf{R}(tw + a)\ \epsilon\ h$. □

Folgerung 13.35 *Die Zuordnung $h \longrightarrow \bar{h}$ ist eine Bijektion zwischen den Geraden von A und den projektiven Geraden, die nicht in $P(W)$ enthalten sind.*

Folgerung 13.36 *Die affinen Geraden h_1, h_2 sind genau dann parallel, wenn $\bar{h}_1 \cap P(W) = \bar{h}_2 \cap P(W)$.*

Beweis: Seien (affin) $h_i = p_i + \mathbf{R}w_i$ mit $p_i = \mathbf{R}a_i$, dann sind die zugenörigen projektiven Geraden gerade $\bar{h}_i = \mathbf{R}w_i + \mathbf{R}a_i$, deren Schnittpunkt mit $P(W$ sind die Punkte $\mathbf{R}w_i$. Affin gilt aber $h_1 \parallel h_2$ genau dann, wenn $\mathbf{R}w_1 = \mathbf{R}w_2$. □

Der affine Raum (A, W) wird also zu einem projektiven Raum „vervollständigt“, indem zu jeder Parallelenschar in A ein „uneigentlicher“ Punkt hinzugefügt wird. Alle Punkte bilden eine „uneigentliche Hyperebene“. Jede Gerade wird zu einer projektiven Geraden verlängert.

Wir wollen nun die Punkte eines projektiven Raums durch Koordinaten beschreeiben.

Definition: Sei $\dim p(V) = n$. Die Punkte $p_0, \ldots, p_n$ bilden eine projetive Basis, wenn sie nicht in einer Hyperebene enthalten sind.

Satz 13.37 *Die Punkte $p_0 = \mathbf{R}x_0, \ldots, p_n = \mathbf{R}x_n$ bilden genau dann eine projektive Basis von $P(V)$, wenn $\{x_0, \ldots x_n\}$ eine Basis von V bilden.*

Beweis: $\{x_0, \ldots, x_n\}$ sind genau dann linear abhängig, wenn sie in einem n-dimensionalen Unterraum W von V enthalten sind, dann liegen die entsprechenden Punkte aber in der Hyperebene $P(W)$. □

Definition: Die Punkte $p_0, \ldots, p_{n+1}$ bilden eine projektives Koordinatensystem, wenn je $n+1$ dieser Punkte eine projektive Basis bilden.
Sei $\{p_0, \ldots, p_{n+1}\}$ ein projektives Koordinatensystem und $p_i = \mathbf{R}x_i$. Dann sind die Vektoren $x_0, \ldots x_{n+1}$ linear abhängig, aber je $n+1$ von ihnen sind linear unabhängig. Es gibt also eine Linearkombination

$$t_0 x_0 + \cdots + t_{n+1} x_{n+1} = o$$

wo alle Koeffizienten T_i ungleich Null sind. Wir können also ohne Beschrnkung der Allgemeinheit annehmen, daß

$$x_0 + \cdots + x_{n+1} = o$$

ist. Die Punkte $p_0, \ldots, p_n$ heißen dann Grundpunkte, p_{n+1} heißt Einheitspunkt.
Wenn nun $p \in P(V)$, $p = \mathbf{R}x$ irgendein Punkt ist, so ist $x = a_0 x_0 + \cdots + a_n x_n$ und die Zahlen $a_0, \ldots, a_n$ heißen die homogenen Koordinaten von p. Wegen des folgenden Satzes wird das Koordinatentupel mit $(a_0 : \ldots : a_n)$ bezeichnet.

Satz 13.38 *Die homogenen Koordinaten sind bis auf ein Vielfaches mit einem Faktor* $\neq 0$ *eindeutig bestimmt. Zu jedem Tupel* $(a_0 : \ldots : a_n) \neq (0 : \ldots : 0)$ *gibt es einen Punkt mit diesen homogenen Koordinaten.*

Beweis: Seien $p_i = \mathbf{R}x_i = \mathbf{R}x'_i$ mit $x_0 + \cdots + x_{n+1} = 0 = x'_0 + \cdots + x'_{n+1}$, also $x'_i = t_i x_i$. Dann ist $\{x - 1, \ldots, x_n\}$ eine Basis von V und es gilt

$$-t_{n+1}x_{n+1} = t_0 x_0 + \cdots + t_n x_n,$$

aber auch

$$-t_{n+1}x_{n+1} = t_{n+1}x_0 + \cdots + t_{n+1}x_n,$$

also $t_0 = \ldots = t_{n+1} = t$. Wenn nun $p = \mathbf{R}x$ ist, so haben wir $x = a_0 x'_0 + \cdots + a_n x'_n = t a_0 x_0 + \cdots + t a_n x_n$, also unterscheiden sich die homogenen Koordinaten bei verschiedenen Repräsentanten der Elemente des projektiven Koordinatensystems nur um die Konstante t.
Zum Tupel $= (a_0 : \ldots : a_n)$ haben wir $p = \mathbf{R}x$ mit $x = a_0 x'_0 + \cdots + a_n x'_n$. □

Folgerung 13.39 *Die homogenen Koordinaten der Grundpunkte sind* $(0 : \ldots : 1 : \ldots : 0)$ *und die des Einheitspunkts sind* $(1 : \ldots : 1)$. □

Folgerung 13.40 *Sei* $P(W)$ *die Hyperebene durch* $p_1, \ldots, p_n$ *und* p *habe die homogenen Koordinaten* $(a_0 : \ldots : a_n)$*, dann ist* p *ein eigentlicher Punkt, wenn* $a_0 \neq 0$ *ist, sonst ein uneigentlicher Punkt. Die affinen Koordinaten eines eigentlichen Punkts sind* $(1 : \frac{a_1}{a_0} : \ldots : \frac{a_n}{a_0})$

Beweis: Es ist $x = a_0 x_0 + \cdots + a_n x_n \in W$ genau dann, wenn $a_0 = 0$ gilt. □

Folgerung 13.41 *Bewzüglich des affinen Koordinatensystems* $\{a = \mathbf{R}x_0, x_1, \ldots x_n\}$ *möge* p *doe Koodinaten* $(p_1, \ldots, p_n)$ *haben, dann sind seine homogenen Koordinaten* $(1 : p_1 : \ldots : p_n)$. □

Sei eine affine Gerade durch eine Gleichung

$$a x_1 + \cdots + b x_2 = c$$

gegeben, wir wollen ihre uneigentlichen Punkte finden. Wir homogenisieren die Gleichung, indem wir x_i durch $\frac{x_i}{x_0}$ ersetzen und die Nenner „hochmultiplizieren“:

$$a x_1 + \cdots + b x_2 = c x_0. \qquad (\star)$$

Ihre eigentlichen Punkte haben die homogenen Koordinaten $(1 : x_1, x_2)$, die $(\star)$ erfüllen. Es gibt genau einen uneigentlichen Punkt auf der Geraden, er hat die homogenen Koordinaten $(0 : -b : a)$, man sieht die Verwandtschft zum Richtungsvektor der Geraden.

Analog besitzt eine Ebene genau eine uneigentliche Gerade; suche Sie sie!

Ein Kreis ist durch eine Gleichung

$$(x_1 - a)^2 + (x_2 - b)^2 = r^2$$

gegeben, wir homogenisieren:

$$x_1^2 - 2ax_0 + x_0^2 + x_2^2 - 2bx_0 + x_0^2 = r^2 x_o^2.$$

Um uneigentliche Punkte zu finden, setzen wir $x_0 = 0$:

$$x_1^2 + x_2^2 = 0,$$

die Lösungen sind $(0 : x_1 : x_2) = (0 : 1 : x_2)$ mit $x_2^2 = -1$, also $(0 : 1 : i)$ und $(0 : 1 : -i)$. In diesen beiden unendlich fernen imaginären Punkten schneiden sich also alle Kreise, denn die Parameter a, b, r sind herausgefallen. Diese Punkte werden die „Kreispunkte genannt.

Überprüfen Sie, daß sich alle zueinander ähnlichen Ellipsen ebenfalls jeweils in zwei imaginären uneigentlichen Punkten schneiden.

Betrachten wir nun die durch

$$ax_1^2 + x_2 = 0$$

gegebene Parabel. Ihre homogene Gleichung lautet

$$ax_1^2 + x_0 x_2 = 0$$

und ihr uneigentlicher Punkt ist $(0 : 0 : 1)$.

Betrachten wir schließlich eine Hyperbel, ihre homogene Gleichung lautet

$$ax_1^2 - bx_2^2 = cx_0^2,$$

sie hat die uneigentlichen Punkte $(0 : 1 : \pm\sqrt{\frac{a}{b}})$, sie entsprechen den Richtungen der Asymptoten.
æ

Kapitel 14

Polynommatrizen

Definition: $M_n(R[x])$ sei die Menge aller $n \times n$-Matrizen $A(x) = (a_{ij}(x))$, wo die $a_{ij}(x)$ Polynome sind. Solche Matrizen heißen Polynommatrizen.

Sei $A(x)$ eine Polynommatrix, k sei das Maximum der Grade der $a_{ij}(x)$, dann heißt k der Grad von $A(x)$, $k = \deg(A(x))$. Dann können wir jedes $a_i(x)$ als

$$a_{ij}(x) = a_{ij}^{(0)} x^k + a_{ij}^{(1)} x^{k-1} + .. + a_{ij}^{(k)}$$

schreiben und mindestens ein $a_{ij}^{(0)}$ ist von Null verschieden.
Wir betrachten nun die Matrizen

$$A_l = (a_{ij}^{(l)}) \in M_{nn},$$

dann ist

$$A(x) = A_0 x^k + A_1 x^{k-1} + \ldots + A_k$$

und A_0 ist nicht die Nullmatrix.

Zum Beispiel:

$$\begin{bmatrix} x^2 + x + 1 & x^3 - x + 2 \\ 2x & x - 3x - 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x^3 + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} x^2 + \begin{bmatrix} 1 & -1 \\ 2 & -3 \end{bmatrix} x + \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}$$

Definition: Die Polynommatrix $a(x)$ heißt regulär, wenn A_0 regulär ist.

Polynommatrizen werden wie üblich addiert und multipliziert.

Lemma 14.1 $\deg(A(x) + B(x)) \leq \max(\deg(A(x), \deg(B(x))$,
$\deg(A(x)B(x)) \leq \deg(A(x)) + \deg(B(x))$, *wenn $A(x)$ oder $B(x)$ regulär ist, so gilt im zweiten Fall Gleichheit.*

Beweis: Sei $A(x) = A_0 x^n + \ldots,\ B(x) = B_0 x^n + \ldots$, wo $A_0 \neq 0$ oder $B_0 \neq 0$ ist, dann ist

$$A(x) + B(x) = (A_0 + B_0) x^k + \ldots,$$

also ist der Grad höchstens gleich n. Weiter sei $A_0 \neq 0$ und $B_0 \neq 0$, dann ist

$$A(x)B(x) = A_0 B_0 x^{n+m} + \ldots,$$

also ist der Grad höchstens gleich $n + m$. Wenn z.B. A_0 regulär ist, so ist $A_0 B_0 \neq 0$, der Grad also gleich $n + m$. □

Satz 14.2 (Divsion mit Rest) *Seien $A(x)$, $B(x)$ Polynommatrizen, $B(X)$ sei regulär. Dann gibt es eindeutig bestimmte Polynommatrizen $Q(x)$, $R(x)$ mit $A(x) = Q(x)B(x) + R(x)$, wobei $R(x) = 0$ oder $\deg R(x) < \deg B(X)$ gilt. $Q(x)$ heißt der rechte Quotient, $R(x)$ der rechte Rest von $A(x)$ bzgl. $B(x)$.*

Beweis: Sei $\deg(A(x)) = l$, $\deg(B(x)) = m$. Falls $l < m$ ist, setzen wir $Q(x) = 0$, $R(x) = A(x)$.
Sei also $l \geq m$, wir führen die Induktion über l. Es gilt

$$B_0^{-1}B(x)x^{l-m} = Ex^l + B_0^{-1}B_1x^{l-1} + \ldots + B_0^{-1}B_mx^{l-m},$$

die Matrix

$$A_0B_0^{-1}B(x)x^{l-m} = A_0x^l + \ldots$$

hat denselben höchsten Koeffizienten wie $A(x)$, also hat

$$A(x) - A_0B_0^{-1}B(x)x^{l-m}$$

höchstens den Grad $l-1$. Nach Induktionsvoraussetzung gibt es also Matrizen $P(x)$ und $R(x)$, wo $\deg(R) < \deg(B)$ oder $R = 0$ gilt, so daß

$$A(x) - A_0B_0^{-1}B(x)x^{l-m} = P(x)B(x) + R(x),$$

d.h.

$$A(x) = (P(x) + A_0B_0^{-1}x^{l-m})B(x) + R(x).$$

Den Faktor vor $B(x)$ nennen wir $Q(x)$.
Die Matrizen Q und R sind eindeutig bestimmt, denn sonst gäbe es P und S mit

$$A = QB + R = RB + S$$

also

$$(Q - P)B = R - R,$$

da $Q - P \neq 0$ sein sollte, steht links eine Matrix vom Grad $\geq m$, rechts aber eine Matrix vom Grad $< m$, also ist $P = Q$ und $R = S$. □

Folgerung 14.3 *Dasselbe gilt mit vertauschten Faktoren: $A = BP + S$, $S = 0$ oder $\deg(S) < \deg(B)$, P heißt linker Quotient und S linker Rest.* □

Es ist sicher nicht verwunderlich, das bei linker und rechter Division unterschiedliche Quotienten und Reste auftreten, es kann sogar vorkommen, daß eine Matrix A bei rechter Division durch B einen von Null verschiedenen Rest läßt, aber von links durch B teilbar ist. (Suchen Sie ein Beispiel!)

Früher haben wir in ein gegebenes Polynom eine (skalare) Matrix eingesetzt. Nun wollen wir in eine Polynommmatrix $A(x) \in M_n(R[x])$ eine Matrix $B \in M_{nn}$ einsetzen, dabei müssen wir aber aufpassen: Sei

$$A(x) = A_0x^k + A_1x^{k-1} + \ldots + A_k,$$

dann definieren wir

$$A(B) = A_0B^k + A_1B^{k-1} + \ldots + A_k.$$

Satz 14.4 *Es sei $B \in M_{nn}$ und $A(x) = Q(x)(Ex - B) + R$, dann hat R den Grad 0, ist also eine skalare Matrix und es gilt $A(B) = R$.*

Beweis: Wie Sie selbst bitte überprüfen, gilt

$$Ex^i - B^i = (Ex^{i-1} + Bx^{i-2} + \ldots + B^{i-2}x + B^{i-1})(Ex - B),$$

wir multiplizieren von links mit A_{k-i} und summieren:

$$A_0Ex^k - A_0B^k + A_1Ex^{k-1} - A_1B^{k-1} + \ldots + A_k - A_k$$

$$= A(x) - A(B) = \sum A_{k-i}(Ex^{i-1} + \ldots + B^{i-1})(Ex - B),$$

den Faktor vor $(Ex - B)$ bezeichnen wir mit $Q(x)$ und erhalten

$$A(x) = Q(x)(Ex - B) + A(B),$$

also $A(B) = R$. □

14.1 Smithsche Normalform

Wir wollen Polynommatrizen Operationen folgenden Typs unterwerfen:
1. Vertauschen von Zeilen bzw. Spalten,
2. Multiplikation einer Reihe mit einer Zahl $r \neq 0$,
3. Addition des $f(x)$-fachen einer Zeile zu einer anderen, dasselbe auch für Spalten, dabei sei $f(x)$ ein Polynom.

Definition: Zwei Polynommatrizen heißen äquivalent, wenn sie durch eine Folge von elementaren Operationen auseinander hervorgehen.

Zum Beispiel gehen die folgenden Matrizen durch elementare Operationen auseinander hervor:

$$\begin{bmatrix} x & x+1 \\ x^2 - x & x^2 - 1 \end{bmatrix} \begin{bmatrix} x & x+1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

Satz 14.5 *Jede Polynommatrix ist zu einer Polynommmatrix der Form*

$$\begin{bmatrix} i_1 & & & & 0 \\ & \ldots & & & \\ & & i_r & & \\ & & & 0 & \\ 0 & & & & 0 \end{bmatrix}$$

äquivalent, wobei jeweils i_k ein Teiler von $i_{k+1}(x)$ ist.

Beweis: Durch Zeilen- und Spaltenvertauschungen wird erreicht, daß $\deg(a_{11}(x))$ minimal ist. Das Polynom a_{1k} aus der ersten Zeile wird mit Rest durch a_{11} dividiert:

$$a_{1k} = qa_{11} + r, \deg(r) < \deg(a_{11}) \text{ oder } r = 0.$$

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 $A(x)$ äquivalent zur Matrix

$$\begin{bmatrix} a_{11}(x) & 0 & \dots & 0 \\ \dots & & & \\ 0 & & A_1(x) & \end{bmatrix}$$

Wenn a_{11} alle Komponenten von $A_1(x)$ teilt, so bleibt das auch bei allen Operationen, die wir künftig mit $A_1(x)$ ausführen, erhalten. Wenn etwa $a_{ij}(x)$ nicht von a_{11} geteilt wird, so addieren wir die i-te Zeile zur ersten und beginnen von vorn. Dabei wird sich der Grad von a_{11} weiter verkleinern. Wenn wir erreicht haben, daß a_{11} alle Komponenten von $A_1(x)$ teilt, widmen wir uns $A_1(x)$ und bringen es in Diagonalgestalt. Irgendwann sind wir fertig. □

Wir fragen uns nun, ob die Polynome $i_1, i_2, \dots$ von den gewählten elementaren Operationen oder nur von der Matrix $A(x)$ abhängen. Die Antwort können wir aber erst etwas später geben. Zuerst überlegen wir uns, daß die Wirkung dieser Operationen durch Multiplikation mit Matrizen folgernder Art realisiert werden kann:

$$\begin{bmatrix} 1 & & & 0 \\ & \dots & & \\ & r & & \\ & \dots & & \\ & & 1 & \end{bmatrix}, \begin{bmatrix} 1 & & & 0 \\ & \dots & & \\ & & f(x) & \\ & \dots & & \\ & & & 1 \end{bmatrix}, \begin{bmatrix} 1 & & & 0 \\ & \dots 1 & & \\ & & & \\ & 1 \dots & & \\ & & & 1 \end{bmatrix}.$$

Dies sind Polynommatrizen, deren Determinante nicht von x abhängt, die also im Bereich der Polynommatrizen eine Inverse besitzen.

Definition: Sei $A(x) = (a_{ij}(x))$ eine Polynommatrix.
Sei $d_1(x)$ der größte gemeinsame Teiler aller $a_{ij}(x)$,
$d_2(x)$ der größte gemeinsame Teiler aller 2-Minoren von $A(x)$,
...
$d_i(x)$ der größte gemeinsame Teiler aller i-Minoren von $A(x)$,
...
$d_n(x) = \det A(x)$. Alle d_i seien normiert. Die d_i heißen die Determinantenteiler von $A(x)$.

Lemma 14.6 *Für alle i gilt: $d_i(x)$ teilt $d_{i+1}(x)$.*

Beweis: Nach dem Entwicklungssatz ist jeder $(i+1)$-Minor von $A(x)$ eine Linearkombination von i-Minoren, also teilt d_i jeden $(i+1)$-Minor und damit auch d_{i+1}. □

Definition: Wir setzen $i_0(x) = 1,\ i_k(x) = \dfrac{d_k(x)}{d_{k-1}(x)}$, die $i_k(x)$ heißen die Invariantenteiler von $A(x)$.

Satz 14.7 *Die Determinantenteiler einer Matrix ändern sich bei elementaren Operationen nicht. Äquivalente Matrizen haben dieselben Determinantenteiler.*

Beweis: Wir betrachten die äquivalenten Matrizen $A(x)$ und $P(x)A(x)Q(x)$, wo $P(x)$ und $Q(x)$ Produkte von Elementarmatrizen sind, ihre Inversen sind also auch Polynommatrizen. Sei $b_j(x)$ ein l-Minor von $P(x)A(x)Q(x)$, nach dem verallgemeinerten Determinantenmultiplikationssatz gilt

$$b_j = \sum p_i a_i q_i,$$

wo die Polynome p_i, a_i, q_i jeweils gewisse l-Minoren von $P(x), A(x)$ bzw. $Q(x)$ sind. Nun sei d_l der l-te Determinantenteiler von $A(x)$. 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. □

Satz 14.8 *Sei $A(x)$ zu* $\begin{bmatrix} a_1 & & \\ & \dots & \\ & & a_n(x) \end{bmatrix}$ *äquivalent, weiter möge jedes a_k ein Teiler von a_{k+1} sein, dann sind die $a_k(x)$ die Invariantenteiler von $A(x)$.*

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 14.9 *Zwei Polynommatrizen sind genau dann äquivalent, wenn sie dieselben Invariantenteiler besitzen.* □

Definition: Zwei Matrizen $A, B \in M_n$ heißen ähnlich, wenn eine reguläre Matrix $X \in M_{nn}$ existiert, so daß $X^{-1}AX = B$ ist.

Im Kapitel über Normalformen haben wir uns ständig mit ähnlichen Matrizen befaßt (ohne es zu wissen).

Satz 14.10 *Die Matrizen A und B sind genau dann ähnlich, wenn die Polynommatrizen $A - Ex$ und $B - Ex$ äquivalent sind, also dieselben Invariantenteiler besitzen.*

Beweis: Sei $X^{-1}AX = B$, dann ist

$$X^{-1}(A - Ex)X = X^{-1}AX - Ex = B - Ex,$$

also sind $A - Ex$ und $B - Ex$ äquivalent.

Seien umgekehrt $A - Ex$ und $B - Ex$ äquivalent, dann gibt es invertierbare Polynommatrizen $P(x), Q(x)$, so daß

$$P(x)(A - Ex)Q(x) = B - Ex$$

gilt. Wir setzen

$$R(x) = P(x)^{-1},$$

dann gilt

$$(A - Ex)Q(x) = R(x)(B - Ex).$$

Wir dividieren nun $Q(x)$ von rechts mit Rest durch $B - Ex$ und $R(x)$ von links durch $A - Ex$:

$$Q(x) = T(x)(B - Ex) + Q_0,$$

$$R(x) = (A - Ex)S(x) + R_0,$$

dabei sind Q_0 und R_0 skalare Matrizen (sie haben den Grad 0). Also ist

$$(A - Ex)(T(x)(B - Ex) + Q_0) = ((A - Ex)S(x) + R_0)(B - Ex)$$

$$(A - Ex)(T(x) - S(x))(B - Ex) = -(A - Ex)Q_0 + R_0(B - Ex)$$

Falls $S \neq T$ ist, hat die linke Matrix einen Grad ≥ 2, die rechte Matrix hat aber höchstens den Grad 1, also ist

$$S(x) = T(x)$$

und damit

$$(A - Ex)Q_0 = R_0(B - Ex) = AQ_0 - Q_0x = R_0B - R_0x,$$

also

$$R_0 = Q_0 \text{ und } AR_0 = R_0B.$$

Um die Ähnlichkeit von A und B zu beweisen, müssen wir noch zeigen, daß R_0 regulär ist. Dazu dividieren wir $P(x) = R(x)^{-1}$ mit Rest durch $(B - Ex)$:

$$P(x) = (B - Ex)U(x) + P_0,$$

dann ist

$$E = R(x)P(x) = ((A - Ex)S(x) + R_0)((B - Ex)U(x) + P_0)$$

$$= (A - Ex)S(x)(B - Ex)U(x) + R_0(B - Ex)U(x) + (A - Ex)S(x)P_0 + R_0P_0,$$

Es ist

$$R_0(B - Ex) = (A - Ex)Q_0,$$

also ist

$$E = (A - Ex)(Q(x)U(x) + S(x)P_0) + R_0P_0,$$

dies ist eine Darstellung der Restdivision von E durch $(A - Ex)$, die sieht aber so aus:

$$E = (A - Ex)0 + E,$$

also ist $R_0P_0 = E$ und R_0 eine reguläre Matrix. □

Wir wenden uns nun dem Minimalpolynom der Matrix $A \in M_{nn}$ zu. Dazu betrachten wir die Polynommatrix $A - Ex$. Die Matrix

$$B(x) = (b_{ij}(x))$$

sei die aus den Adjunkten (den $(n-1)$-Minoren) von $A - Ex$ gebildete Matrix, sie hat den Grad $n-1$, es sei $d_1(x)$ der erste Determinantenteiler von $B(x)$, also der größte gemeinsame Teiler der $b_{ij}(x)$. Wir teilen alle $b_{ij}(x)$ durch $d_1(x)$, es gibt also eine Polynommatrix $C(x)$, deren erster Determinantenteiler gleich 1 ist, mit

$$B(x) = d_1(x)C(x).$$

Aus der Formel für die Inverse einer Matrix erkennen wir

$$(A - Ex)B(x) = \det(A - Ex)E = c_A(x)E,$$

dabei ist $c_A(x)$ das charakteristische Polynom von A. Also gilt

$$c_A(x)E = d_1(x)(A - Ex)C(x),$$

also ist $c_A(x)$ durch $d_1(x)$ teilbar:

$$c_A(x) = d_1(x)m(x),$$

$$m(x)E = (A - Ex)C(x),$$

d.h. die Polynommatrix $m(x)E$ ist ohne Rest durch $A - Ex$ teilbar, also gilt

$$m(A)E = m(A) = 0,$$

also ist $m(x)$ ein annulierendes Polynom für A.

Satz 14.11 *$m(x)$ ist das Minimalpolynom von A, es gilt $m(x)d_1(x) = c_A(x)$.*

Beweis: Sei $n(x)$ das Minimalpolynom von A, dann ist $m(x) = f(x)n(x)$ und $n(x)E$ ist durch $A - Ex$ teilbar:

$$n(x)E = (A - Ex)D(x),$$

also

$$m(x)E = (A - Ex)D(x)f(x) = (A - Ex)C(x),$$

folglich ist $C(x) = D(x)f(x)$, d.h. $f(x)$ ist ein gemeinsamer Teiler der Komponenten von $C(x)$, also ist $f(x) = 1$ und $m(x) = n(x)$. □

Folgerung 14.12 (Hamilton-Cayley) $c_A(A) = 0$. □

Folgerung 14.13 *Das Minimalpolynom von A ist gleich dem höchsten Invariantenteiler von A - Ex.*

Beweis: Sei

$$P(x)(A - Ex)Q(x) = \begin{bmatrix} i_1 & & \\ & \dots & \\ & & i_n \end{bmatrix},$$

Wir wissen, daß $c_A(x) = i_1 \dots i_n$ ist. Sei wieder $B(x)$ die Matrix der Adjunkten von $A - Ex$, dann ist

$$\begin{aligned} (A - Ex)B(x) &= c_A(x)E \\ &= P(x)c_A(x)P(x)^{-1} \\ &= P(x)(A - Ex)Q(x)Q(x)^{-1}B(x)P^{-1} \\ &= \begin{bmatrix} i_1 & & \\ & \dots & \\ & & i_n \end{bmatrix} \begin{bmatrix} b_n & ? \\ \dots & \\ ? & b_1 \end{bmatrix} = \begin{bmatrix} i_1 \dots i_n & & \\ & \dots & \\ & & i_1 \dots i_n \end{bmatrix} \end{aligned}$$

da die $i_k \neq 0$ sind, ist auch die zweite Matrix eine Diagonalmatrix und es gilt

$$\begin{aligned} b_n &= i_2 \dots i_n, \\ b_{n-1} &= i_1 i_3 \dots i_n, \\ &\dots \\ b_2 &= i_1 \dots i_{n-2} i_n, \\ b_1 &= i_1 \dots i_{n-1}. \end{aligned}$$

Nun teilt b_1 das Polynom b_2, b_2 teilt b_3 usw., also sind die b_k die Invariantenteiler von $B(x)$, es ist $c_A(x) = b_1 m(x)$, also ist $m(x) = i_n(x)$. □

14.2 Die rationale Normalform

Zum Schluß wollen wir noch eine weitere Normalform einer skalaren Matrix finden.

Lemma 14.14 *Sei $f(x) = x^n + a_1 x^{n-1} + \dots + a_n$ und*

$$M(f) = \begin{bmatrix} 0 & & & & -a_n \\ 1 & 0 & & & -a_{n-1} \\ 0 & 1 & 0 & & \\ & & & \dots & \\ & & & 1 & -a_1 \end{bmatrix}$$

dann ist $det(M(f) - xE) = f(x)$ das Minimalpolynom von $M(f)$.

Beweis: Das das charakteristische Polynom von $M(f)$ gleich $f(x)$ ist, sollten Sie schon wissen. Die zweite Behauptung folgt aus der ersten. Wir bestimmen die Determinantenteiler von $M(f) - xE$: Es gibt Minoren der Ordnung $1, \dots, n-1$, die gleich 1 sind, damit ist $d_1 = \dots = d_{n-1} = 1$ und $d_n = f$. □

Satz 14.15 (rationale Normalform) *Sei A eine skalare Matrix und $i_r, \ldots, i_n$ seien die nichtkonstanten Invariantenteiler von $A - xE$. Dann gibt es eine regulre Matrix X, so daß*

$$X^{-1}AX = \begin{bmatrix} M(i_r) & & \\ & \ldots & \\ & & M(i_n) \end{bmatrix}$$

eine Blockdiagonalmatrix ist.

Beweis: Nach dem Lemma stimmen die Invariantenteiler der zugehörigen Polynommatrizen überein. □

14.3 Lokale Minimalpolynome eines Endomorphismus

Wir hatten früher gesehen, daß man am Minimalpoynom einer Matrix erkennen kann, ob es eine Basis aus Eigenvektoren gibt oder nicht: Dies ist genau dann der Fall, wenn alle Nullstellen des Minimalpolynoms einfach sind.

Ob mehrfache Nullstellen vorhanden sind, kann man erkennen, ohne diese berechnen zu müssen:

Lemma 14.16 *Wenn $f(x) = (x - x_0)^k g(x),\ g(x_0) \neq 0,\ k > 1$ eine mehrfache Nullstelle x_0 besitzt, so ist x_0 auch eine Nullstelle von $f'(x)$, und umgekehrt.*

Beweis: Es ist $f'(x) = k(x-x_0)^{k-1}g(x) + (x-x_0)^k g'(x)$ und wegen $k > 1$ ist $f'(x_0) = 0$; wenn $k = 1$ gilt, so ist $f'(x_0) = g(x_0) \neq 0$. □

Folgerung 14.17 *Das Polynom $f(x)$ hat genau dann mehrfache Nullstellen, wenn $ggT(f, f') \neq 1$ ist.* □

Wenn A eine „zufällige“ Matrix ist, so sind deren Eigenwerte auch zufällig, also „oft“ voneinander verschieden. Demnach ist „fast jede“ Matrix diagonalisierbar. Schwieriger zu behandeln, aber mathematisch interessant sind die Sonderfälle.

Wir wollen uns nun näher mit Minimalpolynomen beschäftigen.

Satz 14.18 *Sei $m_f(x) = g_1(x)g_2(x)$ mit teilerfremden Polynomen $g_1,\ g_2$, dann gibt es invariante Unterräume $U_1, U_2 \subset V$ und das Minimalpolynom der Einschränkung $f\mid_{U_i}$ ist gleich $g_i(x)$.*

Beweis: Wir setzen $U_i = \{v \in V \mid g_i(f)(v) = o\}$. Wegen der Teilerfremdheit gibt es Polynome $h_i(x)$ mit

$$g_1h_1 + g_2h_2 = 1,$$

also

$$g_1(f) \circ h_1(f) + g_2(f) \circ h_2(f) = id_V.$$

Sei $v \in V$ beliebig, dann gilt

$$v = id_V(v) = g_1(f) \circ h_1(f)(v) + g_2(f) \circ h_2(f)(v)$$

und der erste Summand liegt in U_2 und der zweite in U_1, denn (z.B. $i = 1$)

$$g_1(f) \circ g_2(f) \circ h_2(f)(v) = m_f(f) \circ h_2(f)(v) = o.$$

Somit ist $V = U_1 + U_2$. Sei nun $v \in U_1 \cap U_2$, also

$$g_1(f)(v) = g_2(f)(v) = o,$$

dann ist

$$v = id_V(v) = g_1(f) \circ h_1(f)(v) + g_2(f) \circ h_2(f)(v) = o,$$

also $U_1 \cap U_2 = \{o\}$.
Nach Konstruktion ist g_i ein annullierenden Polynom für $f\mid_{U_i}$. Wenn ein echter Teiler $h(x)$ von g_1 schon ein annullierendes Polynom für $f\mid_{U_1}$ wäre, so wäre $h \circ g_2$ ein annullierendes Polynom für f im Widerspruch zur Minimalität von $m_f(x)$. □

Beispiel: $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{pmatrix}$, $A^2 = \begin{pmatrix} 14 & 28 & 42 \\ 28 & 56 & 84 \\ 42 & 84 & 126 \end{pmatrix}$, $m_A(x) = x^2 - 14x$, denn A hat den Rang 1. Wir setzen $g_1(x) = x - 14$, $g_2(x) = x$, $U_1 = \{v \in R^3 \mid (A - 14E)v = o\} = L\begin{pmatrix} 16 \\ -2 \\ 3 \end{pmatrix}$, $U_2 = \{v \mid Av = o\} = L\left\{\begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}\right\}$.

Definition: Sei $f : V \longrightarrow V$ ein Endomorphismus und $v \in V$. Das normierte Polynom $m_{f,v}(x) \in R[x]$ heißt Minimalpolynom von f für v, wenn es das Polynom kleinsten Grades ist, so daß $m_{f,v}(f)(v) = o$ ist.

Beispiel: Sei $A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$, dann ist $A^2 = \begin{pmatrix} 4 & 4 \\ 0 & 4 \end{pmatrix}$ also $Ae_1 = 2e_1$, $(A-2E)e_1 = o$, also $m_{A,e_1}(x) = x-2$; $Ae_2 = e_1+2e_2$, $A^2e_2 = 4e_1+4e_2 = 4Ae_2-4Ee_2$, da $4Ae_2 = 4e_1+8e_2$ ist. Also gilt $(A^2 - 4A + 4E)e_2 =$, demnach ist $m_{A,e_2} = x^2 - 4x + 4 = (x - 2)^2$.

Wenn $m_f(x)$ das Minimalpolynom von f ist, also $m_f(f) = 0$ ist, dann ist $m_f(f)(v) = 0$ für alle $v \in V$, also ist $m_{f,v}(x)$ ein Teiler von $m_f(x)$.

Welche Beziehungen bestehen nun zwischen verschiedenen Minimalpolynomen? Wir halten den Endomorphismus f fest und schreiben m_v anstelle von $m_{f,v}$.

Satz 14.19 *Seien $v, w \in V$; wenn m_v und m_w teilerfremd sind, so ist $m_{v+w} = m_v m_w$.*

Beweis: Sei $h(x)$ ein annullierendes Polynom für $v + w$, d.h. $h(f)(v + w) = o$. Dann ist

$$m_w(f)h(f)(v) = m_w(f)h(f)(v + w) - h(f)\underbrace{m_w(f)(w)}_{=o} = o$$

(Polynome in f kommutieren), also gilt $m_v \mid m_w h$, wegen der Teilerfremdheit von m_v und m_w folgt $m_v \mid h$ und analog $m_w \mid h$. Also wird jedes $v + w$ annullierende Polynom von $m_v m_w$ geteilt, also ist dies das Minimalpolynom. □

Lemma 14.20 *Sei $\{v_1, \ldots, v_n\} \subset V$ eine Basis, dann ist $m_f(x)$ das kleinste gemeinsame Vielfache der m_{v_i}.*

Beweis: Sei $g(x)$ ein gemeinsames Vielfaches der m_{v_i}, also $g(x) = h_i(x)m_{v_i}(x)$ und sei $v = \sum r_i v_i \in V$, dann gilt

$$g(f)(v) = \sum r_i h_i(f) m_{v_i}(f)(v_i) = o.$$

Wenn umgekehrt $g(f)$ alle Vektoren in V annulliert, so annulliert es auch die v_i, also ist $g(x)$ durch m_{v_i} teilbar, also ein gemeinsames Vielfaches der m_{v_i}. Das Polynom minimalen Grades mit dieser Eigenschaft ist das kleinste gemeinsame Vielfache. □

Satz 14.21 *Es gibt einen Vektor $v \in V$ mit $m_{f,v} = m_v$.*

Beweis: Wir betrachten zunächst einen Spezialfall:
Sei $m_f(x) = g(x)^k$ die Potenz eines irreduziblen Polynoms und sei $\{v_1, \ldots, v_n\}$ eine Basis von V. Die Minimalpolynome der Basisvektoren sind dann Teiler von $g(x)^k$, also $m_{v_i}(x) = g(x)^{k_i}$. Sei nun $m = \max k_i$, dann ist $g(x)^m = kgV(g(x)^{k_i}) = m_f(x) = g(x)^k$, also $l = k$ und ein Basisvektor v_i, wo das Maximum angenommen wir, leistet das Verlangte.
Sei nun

$$m_f(x) = \prod_{i=1}^{m} g_i(x)^{k_i},\ ggT(g_i, g_j) = 1 \text{ für } i \neq j,$$

dann ist $V = U_1 \oplus \cdots \oplus U_m$ mit zu den $g(x)^{k_i}$ gehörigen invarianten Unterräumen, diese Polynome sind paarweise teilerfremd. Wir wählen Vektoren $u_i \in U_i$ mit den „richtigen“ Minimalpolynomen, das Minimalpolynom von $u_1, \ldots, u_m$ ist dann gleich $\prod_{i=1}^{m} g_i(x)^{k_i} = m_f(x)$. □

Also

Folgerung 14.22 (rationale Normalform) *Sei V bezüglich f unzerlegbar, dann gibt es einen Vektor v, so daß $m_{f,v}(x) = m_f(x) = x^n + a_1 x^{n-1} + \cdots a_n$ ist, die Vektoren $f^i(v)$, $i = 0, \ldots, n-1$ sind linear unabhängig, die Darstellungsmatrix von f hat dann die Form*

$$M(f) = \begin{pmatrix} 0 & & & & -a_n \\ 1 & 0 & & & -a_{n-1} \\ 0 & 1 & 0 & & \\ & & & \cdots & \\ & & & 1 & -a_1 \end{pmatrix}.$$

□

Kapitel 15

Elementare Gruppentheorie

15.1 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

$$a = bq + r \text{ und } 0 \leq r < |\, b \,|\,.$$

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.
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 Größte). Wir schreiben $d = \mathrm{ggT}(a, b)$.

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: $\mathrm{ggT}(f_1, f_2) = f_{m-1}$.
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. □

Lemma 15.1 *Sei $d = ggT(f_1, f_2)$, 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. □
Interessanter ist das

Lemma 15.2 *Der größte gemeinsame Teiler von f_1 und f_2 ist die kleinste positive Zahl d, so daß $f_1 g_1 + f_2 g_2 = d$ ist.*

Beweis: Sei $d = f_1 g_1 + f_2 g_2$ und d minimal.
1. Wir dividieren mit Rest:

$$f_1 = q_1 d + r_1 = q_1 g_1 f_1 + q_1 g_2 f_2 + r_1,$$

also

$$r_1 = f_1(1 - q_1 g_1) - f_2 q_1 q_2,$$

aber wegen $r_1 < d$ ist dies ein Widerspruch zur Minimalität von d.
2. Sei h ein gemeinsamer Teiler von f_1 und f_2, dann ist h auch ein Teiler von $f_1 g_1 + f_2 g_2$ = d. □

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:

$$\bar{a} + \bar{b} = \overline{a + b}, \ \bar{a} \cdot \bar{b} = \overline{a \cdot b}.$$

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 $\mathrm{ggT}(a, m) = 1$, also gibt es u, v mit

$$1 = ua + vm,$$

d.h. $1 \equiv ua \pmod{m}$, also ist $\bar{a}^{-1} = \bar{u}$.
Ist insbesondere 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.

Eine Zahl p heißt Primzahl, wenn aus $a \mid p$ folgt, daß $a = \pm 1$ oder $a = \pm p$ gilt.

Lemma 15.3 *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. □

Folgerung 15.4 *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 $a = p_1 \dots p_n$. □

Lemma 15.5 *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 $\mathrm{ggT}(p, a) = 1 = up + va$ für gewisse $u, v \in \mathbb{Z}$. Dann folgt $b = upb + vab$, die rechte Seite wird von p geteilt, also gilt $p \mid b$. □

Satz 15.6 *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). □

15.2 Gruppen, Untergruppen, Homomorphismen

Definition: Sei G eine Menge und $\cdot : G \times G \to G$ 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 $g_1 g_2 = g_2 g_1$ gilt, so heißt die Gruppe G kommutativ.
Für kommutative (und nur solche) Gruppen ist auch eine additive Schreibweise üblich:

$$+ : G \times G \to G,\ (g_1, g_2) \mapsto g_1 + g_2,$$

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:

$$(\mathbb{Z}, +),\ (\mathbb{Q}, +),\ (\mathbb{R}, +),\ (\mathbb{C}, +),\ (\mathbb{Q}\backslash\{0\}, \cdot),\ (\mathbb{R}\backslash\{0\}, \cdot),\ (\mathbb{C}\backslash\{0\}, \cdot),$$

$$(\mathbb{R}^n, +),\ (\mathrm{Hom}(V, W), +),\ (M_{mn}, +),$$

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.

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 15.7 *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 $\mathbb{Q}\backslash\{0\}, \mathbb{R}\backslash\{0\}, \mathbb{C}\backslash\{0\}$.

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:

$$\{e\},\ \{e,a\},\ \{e,b\},\ \{e,c\},\ \{e,d,f\},\ S_3.$$

Wenn $E \subseteq G$ eine Teilmenge ist, so bezeichnen wir mit $< E >$ 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 $< E >$ wird von der Menge E erzeugt.
Zum Beispiel:

$$(\mathbb{Z},+) =< 1 >,\ (\mathbb{Q}\backslash\{0\},\cdot) =< \mathbb{Z}\backslash\{0\} >,$$

$$\{e,a\} =< a >,\ \{e,b\} =< b >,\ \{e,c\} =< c >,$$

$$\{e,d,f\} =< d >=< f >,\ S_3 =< a,b >=< a,d > \text{ usw.}$$

Eine Gruppe G, die von einem Element g erzeugt wird, heißt zyklische Gruppe, es gilt also $G = \{e = g^0, g = g^1, g^2, \ldots\}$, die Gruppe kann endlich oder unendlich sein.

Wir überlegen, wie eine endliche zyklische Gruppe $G =< g >$ aussehen könnte. Die Potenzen $g, g^2, g^3, \ldots$ 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 $e = g^0, g, g^2, \ldots, g^{k-1}$. 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.

Lemma 15.8 *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 $g(ug)^{-1} = gg^{-1}u^{-1} = u^{-1} \in U$. □

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 15.9 *Für alle $g \in G$ gilt $|Ug| = |U|$, d.h. alle Äquivalenzklassen sind gleichmächtig.*

Beweis: Sei $g \in G$, wir betrachten die Abbildung $f : U \to Ug$ mit $f(u) = ug$. Diese Abbildung ist surjektiv (klar), wir zeigen, daß sie injektiv ist: Sei $u_1 g = u_2 g$, dann gilt $u_1 gg^{-1} = u_2 gg^{-1} = u_1 = u_2$. Also ist f bijektiv und damit gilt $|Ug| = |U|$. □

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 $5\mathbb{Z}$:

$$\begin{aligned} 5\mathbb{Z} &= \{0, \pm 5, \pm 10, \pm 15, \ldots\}, \\ 1 + 5\mathbb{Z} &= \{1, 6, 11, \ldots, -4, -9, \ldots\}, \\ 2 + 5\mathbb{Z} &= \{2, 7, 12, \ldots, -3, -8, \ldots\}, \\ 3 + 5\mathbb{Z} &= \{3, 8, 13, \ldots, -2, -7, \ldots\}, \\ 4 + 5\mathbb{Z} &= \{4, 9, 14, \ldots, -1, -6, \ldots\}. \end{aligned}$$

Wenn $U \subseteq G$ eine Untergruppe ist, so bezeichnet man die Menge aller Äquivalenzklassen modulo U mit G/U.

Satz 15.10 (Lagrange) *Sei G eine endliche Gruppe und $U \subseteq G$ eine Untergruppe, dann ist die Zahl $|U|$ ein Teiler von $|G|$.*

Beweis: Es ist $G = U \cup Ug_2 \cup Ug_3 \cup \ldots \cup Ug_t$ für gewisse $g_i \in G$, denn G ist die disjunkte Vereinigung seiner Äquivalenzklassen modulo U, also gilt $|G| = t\,|U|$. □

Folgerung 15.11 *Jede Gruppe von Primzahlordnung ist zyklisch.*

Beweis: Sei $|G| = p$ eine Primzahl und $e \neq g \in G$, dann ist $< g >$ eine Untergruppe mit mehr als einem Element, da die Ordnung von $< g >$ ein Teiler von p ist, folgt $|< g >| = p$, also $G =< g >$. □

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 $< g >$, also ein Teiler von $|G|$. Also gilt das

Lemma 15.12 *Sei $|G| = n$ und $g \in G$, dann ist $g^n = e$.* □

Folgerung 15.13 (Kleiner Satz von Fermat) *Sei p eine Primzahl. Wenn $ggT(a,p) = 1$ ist, so gilt $a^{p-1} \equiv 1 (mod p)$.*

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:

$$(g_1, h_1)(g_2, h_2) = (g_1 g_2, h_1 h_2).$$

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 $C_2 \times C_2$.
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 $f : H \to G$ 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 $f : H :\to G$ 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 $\mathbb{Z} \to \mathbb{Q} \to \mathbb{R} \to \mathbb{C}$ sind Homomorphismen der additiven Gruppen, für die Logarithmusfunktion gilt $\ln(ab) = \ln(a) + \ln(b)$, also ist $\ln : (\mathbb{R}_+, \cdot) \to (\mathbb{R}, +)$ ein Homomorphismus. Die Funktion sgn:$S_n \to \{\pm 1\}$, die jeder Permutation ihr Signum zuordnet, ist ein Homomorphismus. Für jeden Homomorphismus $f : G \to H$ und jede Untergruppe $U \subseteq G$ ist die Einschränkung $f \mid U : U \to H$ ebenfalls ein Homomorphismus. Schließlich ist für jedes $x \in \mathbb{Z}$ die Abbildung $l_x : \mathbb{Z} \longrightarrow \mathbb{Z}$ mit $l_x(a) = xa$ ein Homomorphismus der additiven Gruppen.

Definition: Sei $f : H \to G$ ein Homomorphismus, dann ist der Kern von f die Teilmenge $\mathrm{Ker}(f) = \{h \in H \mid f(h) = 1\}$.

Lemma 15.14 *Wenn $f : H \to G$ ein Homomorphismus ist, so ist* Ker*(f) eine Untergruppe von G.*

Beweis: Seien $h_1, h_2 \in \mathrm{Ker(f)}$, also $f(h_1) = 1 = f(h_2)$, dann ist $f(h_1h_2) = f(h_1)f(h_2) = 1 \cdot 1 = 1$ und $f(h_1^{-1}) = f(h_1)^{-1} = 1$. □

Wir bemerken, daß der Kern eines Homomorphismus eine weitere Eigenschaft hat: Wenn $h \in \mathrm{Ker}(f)$ ist, so gilt für beliebige $g \in G$ folgendes: $f(g^{-1}hg) = f(g^{-1})f(h)f(g) = f(g)^{-1}f(g) = 1$, also $g^{-1}hg \in \mathrm{Ker}f(f)$.

Wenn $A, B \subseteq G$ Teilmengen einer Gruppe sind, so bezeichneten wir mit AB die Menge aller Produkte ab, wo $a \in A$ und $b \in B$ ist. Wenn $B = \{b\}$ ist so schreiben wir für $A\{b\}$ einfach Ab.

Definition: Sei $N \subseteq G$ eine Untergruppe, sie heißt normale Untergruppe (oder Normalteiler), wenn $g^{-1}Ng = N$ 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 15.15 *Sei $N \subseteq G$ eine normale Untergruppe, dann ist die Menge G/N mit folgender Multiplikation eine Gruppe: $(Ng)(Nh) = Ngh$.*

Beweis: Wegen $g^{-1}Ng = N$ gilt $Ng = gN$, also gilt für das Produkt der Teilmengen Ng und Nh wirklich $NgNh = NNgh = Ngh$. Der Rest ist klar: $(Ng_1Ng_2)Ng_3 = N(g_1g_2)g_3 = Ng_1(g_2g_3) = Ng_1(Ng_2N_3)$, das neutrale Element ist N, da $NNg = Ng = NgN$ gilt, Invers zu Ng ist Ng^{-1}. □

Den im folgenden Lemma auftretenden Homomorphismus nennt man einen „kanonischen“ Homomorphismus.

Lemma 15.16 *Sei $N \subseteq G$ eine normale Untergruppe, dann ist die Abbildung $k : G \to G/N$ mit $k(g) = Ng$ ein Homomorphismus und es gilt* Ker*(k) = N.*

Beweis: $k(g_1g_2) = Ng_1g_2 = Ng_1Ng_2 = k(g_1)k(g_2)$ und $k(g) = N$ gilt genau dann, wenn $g \in N$ ist. □

Definition: Sei $f : H \to G$ ein Homomorphismus, dann ist das Bild von f die folgende Menge $\mathrm{Im(f)} = \{g \in G \mid \text{es gibt ein } h \in H \text{ mit } g = f(h)\}$.

Lemma 15.17 Im*(f) ist eine Untergruppe von G.* □

Satz 15.18 *Sei $f : H \to G$ ein Homomorphismus. Dann gilt:*
f ist genau dann injektiv, wenn Ker*(f) = {1} ist,*
f ist genau dann surjektiv, wenn Im*(f) = G ist.*

Beweis: Sei f injektiv und $g \in \mathrm{Ker}(f)$, also $f(g) = 1 = f(1)$, dann muß $g = 1$ sein. Sei umgekehrt $\mathrm{Ker}(f) = \{1\}$ und $f(h) = f(g)$, dann gilt $1 = f(g)f(h)^{-1} = f(gh^{-1})$, also $gh^{-1} \in \mathrm{Ker}(f) = \{1\}$, d.h. $gh^{-1} = 1$, also $g =$ h.
Die zweite Aussage ist trivial. □

Ein injektiver und surjektiver Homomorphismus heißt Isomorphismus. Wenn zwischen zwei Gruppen H und G ein Isomorphismus existiert $f : H \to G$ existiert, so heißen sie isomorph, man schreibt dann $H \simeq G$.

Es folgen einige Sätze, die die Isomorphie gewisser Gruppen sichern.

Satz 15.19 (Homomorphiesatz) *Sei $f : H \to G$ ein Homomorphismus, dann ist die durch $F(h{\cdot}\mathrm{Ker}(f)) = f(h)$ gegebene Abbildung $F : H/\mathrm{Ker}(f) \to \mathrm{Im}(f)$ ein Isomorphismus.*

Beweis: Wir zeigen zuerst, daß F wohldefiniert ist: Sei $h_1\mathrm{Ker}(f) = h_2\mathrm{Ker}(f)$, also $h_1h_2^{-1} \in \mathrm{Ker}(f)$, d.h. $1 = f(h_1h_2^{-1}) = f(h_1)f(h_2)^{-1}$, also $F(h_1\mathrm{Ker}(f)) = f(h_1) = f(h_2) = F(h_2\mathrm{Ker}(f))$. Weiter gilt $F(h_1\mathrm{Ker}(f) \cdot h_2\mathrm{Ker}(f)) = F(h_1h_2\mathrm{Ker}(f)) = f(h_1h_2) = f(h_1)f(h_2) = F(h_1\mathrm{Ker}(f)) \cdot F(h_2\mathrm{Ker}(f))$, also ist F ein Homomorphismus. Die Surjektivität von F ist klar und die Injektivität folgt sofort: Sei $F(h\mathrm{Ker}(f)) = 1 = f(h)$, dann ist $h \in \mathrm{Ker}(f)$, also ist $h\mathrm{Ker}(f) = \mathrm{Ker}(f)$ das neutrale Element in $H/\mathrm{Ker}(f)$. □

Lemma 15.20 *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 ebenfalls eine normale Untergruppe ist, so ist H/N eine normale Untergruppe von G/N.

Beweis: 1. Sei $f : G \to G/N$ der kanonische Homomorphismus, dann ist die Einschränkung $f \mid H : H \to G/N$ ebenfalls ein Homomorphismus, dessen Kern gerade $H \cap N$ ist.
2. Trivial.
3. Sei $h_i \in H, n_i \in N$, dann sind $h_1n_1, h_2n_2 \in HN$, weiter ist $h_2^{-1}n_1h_2 = n \in N$ wegen der Normalteilereigenschaft, also $n_1h_2 = h_2n$. Nun folgt $h_1n_1 \cdot h_2n_2 = h_1h_2nn_2 \in HN$ und $(h_1n_1)^{-1} = n_1^{-1}h_1^{-1} = h_1^{-1}n'$ mit $n' = h_1n_1^{-1}h_1^{-1} \in N$.
4. Es gilt $H/N = \{Nh \mid h \in H\} \subseteq \{Ng \mid g \in G\} = G/N$, weiter $Nh_1 \cdot Nh_2 = Nh_1h_2 \in H/N$ und $(Nh)^{-1} = Nh^{-1} \in H/N$, also ist H/N eine Untergruppe von G/N. Diese ist normal: $(Ng)^{-1}NhNg = Ng^{-1}hg$ und $g^{-1}hg \in H$, also liegt $(Ng)^{-1}NhNg$ in H/N. □

Satz 15.21 (1. Isomorphiesatz) *Seien $H, N \subseteq G$ Untergruppen, N sei normal, dann gilt*

$$H/(N \cap H) \simeq HN/N.$$

Beweis: Sei $f : H \to HN/N$ 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 $hN = f(h) \in \mathrm{Im}(f)$. Sei $h \in \mathrm{Ker}(f)$, also $f(h) = hN = N$, d.h. $h \in N$, also $h \in N \cap H$. Die Behauptung folgt nun aus dem Homomorphiesatz. □

Satz 15.22 (2. Isomorphiesatz) *Seien $N \subseteq H \subseteq G$ normale Untergruppen, dann gilt*

$$G/H \simeq (G/N)/(H/N).$$

Beweis: Wir betrachten die Abbildung $f : G/N \to G/H$, die durch $f(gN) = gH$ gegeben ist (sie ist wegen $gN \subseteq gH$ wohldefiniert), offenbar surjektiv und ein Homomorphismus. Es ist genau dann $gN \in \mathrm{Ker}(f)$, wenn $gH = H$, also $g \in H$ gilt. Der Kern ist somit gleich H/N und die Behauptung folgt aus dem Homomorphiesatz. □

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 $(\mathbb{Z}/km\mathbb{Z})/(m\mathbb{Z}/km\mathbb{Z}) \cong C_k$.

15.3 Die symmetrischen Gruppen

Wir wollen nun die Gruppen $S_n = \{p : \{1, \ldots, n\} \longrightarrow \{1, \ldots, n\}\}$ der bijektiven Abbildungen der Menge $\{1, \ldots, n\}$ 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 15.23 *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 $p(m) = q(m)$. □

Folgerung 15.24 *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. □

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 $\{i_1, \ldots, i_m\} \subset \{1, \ldots, n\}$ gibt, so daß

$$p(i_k) = p_{k+1},\ k = 1, \ldots, m-1,$$
$$p(i_m) = i_1,$$
$$p(j) = j \text{ sonst}$$

gilt. Wir schreiben dann $p = (i_1, \ldots, i_m)$, 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, Transpositonen haben das Signum -1. Ein Zyklus der Länge k ist ein Produkt von $k-1$ Transpositionen, hat also das Signum $(-1)^{k-1}$, denn

$$(i_1, i_2, \ldots, i_k) = (i_1, i_k) \cdots (i_1, i_3)(i_1, i_2).$$

15.4 Endlich erzeugte abelsche Gruppen

Eine abelsche Gruppe ist nichts anderes als eine kommutative Gruppe. Wir verwenden hier die additive Schreibweise. Das direkte Produkt $A_1 \times \ldots \times A_n$ nennen wir hier die direkte Summe und bezeichnen sie mit $A_1 \oplus \ldots \oplus A_n$.

Sei also A eine abelsche Gruppe und $a \in A$, dann schreiben wir als Abkürzung für $a + a + \ldots + a$ (m Summanden) einfach $m{\cdot}$a. Umgekehrt, wenn $m \in \mathbb{Z}$ ist, so soll $ma = a + \ldots + a$ (m Summanden, wenn $m \geq 0$) bzw. $ma = -a - \ldots - a$ ($-m$ 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 15.25 *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:

$$a = a_1 + a_2 = b_1 + b_2,\ a_i, b_i \in A_i.$$

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

$$t(A) = \{a \in A \mid \text{ es gibt ein } m \in Z \text{ mit } ma = 0\}$$

die Torsionsuntergruppe von A.

Lemma 15.26 *$t(A)$ 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 15.27 *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 15.28 *Sei $|\ A\ | = n = p_1^{i_1} \dots p_k^{i_k}$, p_i verschiedene Primzahlen, 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 $\mathrm{ggT}(u, v) = 1 = ru + sv$, 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 $ua = o$, 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 15.29 *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 $\cdot : K \times A \to A$ 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 $A \simeq K^n \simeq \mathbb{Z}/p\mathbb{Z} \oplus \ldots \oplus \mathbb{Z}/p\mathbb{Z}$.
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

$$pA =< a_1 > \oplus \ldots \oplus < a_k >$$

und es gibt $h_1, \ldots, h_k \in A$ mit $ph_i = a_i$. Wir setzen

$$H =< h_1 > + \ldots + < h_k >$$

und behaupten, daß die Summe der $< h_i >$ direkt ist. In der Tat, sei

$$0 = m_1 h_1 + \ldots + m_k h_k$$

dann ist

$$0 = p0 = m_1 ph_1 + \ldots + m_k ph_k = m_1 a_1 + \ldots + m_k a_k,$$

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 $pa = \sum m_i a_i = \sum m_i ph_i = ph f\ddot{u}r$ ein $h \in H$. Wir setzen $a' = a - h$, dann ist $a' \notin B + H$ und $pa' = 0$. Wir setzen $B' =< a', B >$. Nach Konstruktion von B gilt $B' \cap H \neq \{0\}$, also gibt es ein $h' \in H$ mit

$$h' = ka' + b,\ b \in B,\ 0 < k < p.$$

Sei $sk \equiv 1 (\text{mod } p)$, 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 Beahuptung auch für A. □

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 15.30 *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 $t = ggT(x, y) = ax + by$, 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 $\det \begin{pmatrix} a & b \\ -\frac{y}{t} & \frac{x}{t} \end{pmatrix} = \frac{ax}{t} + \frac{by}{t} = \frac{t}{t} = 1.$ □

Lemma 15.31 *Sei A eine endlich erzeugte abelsche Gruppe und $a_1, \ldots, a_n \in A$, $x_1, \ldots, x_n \in \mathbb{Z}$ vorgegeben. Dann gibt es $b_1, \ldots, b_n \in A$ mit $< a_1, \ldots, a_n >=< b_1, \ldots, b_n >$ und $x_1a_1 + \ldots + x_na_n = tb_1$, wobei $t = ggT(x_1, \ldots, x_n)$ 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

$$a_1x_1 + a_2x_2 = (a_1, a_2)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (b_1, b_2)\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$= (b_1, b_2)\begin{pmatrix} t \\ 0 \end{pmatrix} = tb_1.$$

Nun sei $n = 3$. Dann gibt es b_2, b_3 mit $< a_2, a_3 >=< b_2, b_3 >$ und $x_2b_2 + x_3b_3 = db_2$, wie wir soeben sahen. Ebenso gibt es b_1, b_2', so daß $< a_1, db_2 >=< b_1, b_2' >$ und $x_1a_1 + 1 \cdot db_2 = tb_1$. Dann gilt $< a_1, a_2, a_3 >=< b_1, b_2', b_3 >$ und $x_1a_1 + x_2a_2 + x_3a_3 = tb_1$. Und so weiter. □

Definition: Sei A eine abelsche Gruppe, dann heißen die Elemente $a_1, \ldots, a_n$ linear unabhängig, wenn aus $\sum x_ia_i = 0$ ($x_i \in \mathbb{Z}$) folgt, daß $x_i = 0$ für alle i gilt.
Eine abelsche Gruppe heißt frei, wenn sie ein linear unabhängiges Erzeugendensystem besitzt. Eine endliche erzeugte freie abelsche Gruppe ist isomorph zu $\mathbb{Z} \times \ldots \times \mathbb{Z}$.

Satz 15.32 *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 $r(A)$ Elemente ist linear unabhängig.
2. A ist frei.

Beweis: Sei $\{a_1, \ldots, a_n\}$ ein Erzeugendensystem von A mit $n = r(A)$ und es sei $x_1a_1 + \ldots + x_na_n = 0$. Dann gibt es ein Erzeugendensystem $b_1, \ldots, b_n$ mit $tb_1 = x_1a_1 + \ldots + x_na_n = 0$, hier muß $b_1 \neq 0$ sein, denn $\{b_2, \ldots, b_n\}$ 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. □

15.5 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. $X = \{1, \ldots, n\}, G = S_n, p \cdot i = p(i)$.
2. V sei ein R-Vektorraum, hier operiert die multiplikative Gruppe von 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 15.33 *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 verträglch, 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 15.34 *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 15.35 *Die Anzahl der zu $g \in G$ konjugierten Elemente ist ein Teiler der Gruppenordnung.*

Satz 15.36 (Cauchy) *Sei G eine endliche Gruppe und p eine Primzahl mit $p \mid |G|$, dann gibt es eine gibt es ein $g \in G$ mit der Ordnung p, also $g^p = 1$.*

Beweis: Es sei $X = \{(g_0, \dots g_{p-1}) \in G \times \cdots \times G \mid g_0 \cdots g_{p-1} = 1\}$, 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 $\mathrm{Z}/p\mathrm{Z} = H$. 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 $(1, \ldots, 1)$ 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 $O_x = \{(g_o, \ldots, g_{p-1})\}$ einelementig ist. Dann ist aber $g_0 = \ldots = g_{p-1} \neq 1$, also $g_0^p = 1$.
□

Wir wollen nun noch einmal systematischer die Gruppen kleiner Ordnung untersuchen.

Lemma 15.37 *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 $G = \langle g \rangle$. □

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

$$D_n = \{1, a, a^2, \ldots a^{n-1}, b, ba, ba^2, \ldots ba^{n-1}\},$$

diese Gruppe hat $2n$ Elemente und wird durch die Relationen

$$a^n = 1, b^2 = 1, ab = b^{-1}a$$

charakterisiert.

Satz 15.38 *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 $2 \not| p$ 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 $G = D_n$.
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. □

Satz 15.39 *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 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. □

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

$$i^2 = j^2 = k^2 = -1, \; ij = -ji = k.$$

Satz 15.40 *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 $/xN)^2 = N = x^2 N$, 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

$$y = x^2yx^{-2} = x(xyx^{-1})x^{-1} = xy^kx^{-1} = (xyx^{-1})^k = (xyx^{-1})^k = (y^k)^k = y^{k^2},$$

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. □

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 15.41 *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 $f : H \longrightarrow H$; 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

$$k^4 \equiv 1 (\mathrm{mod} 5)$$

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

$$G = H \cdot K = C_5 \times C_3 = C_{15}.$$

□

15.6 Aufgaben

1. Man beweise card $S_n = n!$ für die symmetrische Gruppe der Ordnung n ($n \in \mathbb{N}$); ferner zeige man, daß die Gruppen S_n für $n \geq 3$ nicht abelsch sind!

2. Sei $[G, \circ]$ eine Gruppe mit neutralem Element e. Man beweise:

   a) wenn $g^2 = e$ für alle $g \in G$ gilt, so ist G abelsch.

   b) wenn es genau ein Element $g \in G, g \neq e$, mit der Eigenschaft $g^2 = e$ gibt, so ist $x \circ g = g \circ x$ für alle Elemente $x \in$ G.

3. Es sei $A := \{x \in \mathbb{R}; x > 1\}$. Auf A sei eine Operation $\circ$ definiert durch: $x \circ y := xy - x - y + 2; x, y \in A$. Es ist zu beweisen, daß A eine Gruppe ist. Ist diese Gruppe abelsch ?

4. Man finde alle Untergruppen der symmetrischen Gruppe S_3!

5. a) Sei $g = (a_1, a_2, \ldots, a_k) \in S_n$ ein Zyklus der Länge k; beweisen Sie, daß die Ordnung von g gleich k ist!

   b) Sei $g = g_1 \cdot g_2 \cdot \ldots \cdot g_r$ die Zerlegung eines Elements$g \in S_n$ in elementfremde Zyklen der Längen $k_1, k_2, \ldots, k_r$. Finden Sie eine Formel für die Ordnung von g in Abhängigkeit von $k_1, \ldots, k_r$! (Man benutze, daß elementfremde Zyklen miteinander kommutieren!)

   c) Sei p eine Primzahl, k eine beliebige natürliche Zahl und sei $n := p^k$. Beweisen Sie: wenn $g \in S_n$ ein Element der Ordnung ord(g) $= n$ ist, so muß g ein Zyklus der Länge n sein.

   d) Geben Sie ein Gegenbeispiel für die unter c) getroffene Aussage für den Fall an, daß n keine Primzahlpotenz ist!

6. Man zeige: die Menge der Matrizen der Form $\left[\begin{pmatrix} \epsilon \cdot \cos\varphi & -\sin\varphi \\ \epsilon \cdot \sin\varphi & \cos\varphi \end{pmatrix}\right]$, $(\varphi \in \mathbb{R}, \epsilon = \mp 1)$ ist eine Untergruppe von $GL(2, \mathbb{R})$. Ist sie abelsch ?

7. Zeigen Sie, daß die Menge der Matrizen $H := \{A \in GL(3, \mathbb{R}); A = \begin{bmatrix} 1 & x & y \\ 0 & 1 & x \\ 0 & 0 & 1 \end{bmatrix}; x, y, z \in \mathbb{R}\}$ eine nichtabelsche Gruppe bildet! Für ein beliebiges Element $A \in H$ gebe man A^{-1} an!

8. Sei K ein Körper. Beweisen Sie: Das Zentrum der Gruppe $GL(n, K)$ ist gleich $Z(GL(n, K)) = \{\mu \cdot E; \mu \in K\}$, wobei$E = (\delta_{ij})$ die Einheitsmatrix bezeichnet.

Kapitel 16

Ringe und Moduln

16.1 Grundbegriffe

Definition: Sei R eine Menge, in der zwei Operationen

$$+ : R \times R \to R \; (r, s) \; \mapsto r + s)$$

und

$$\cdot : R \times R \to R \; (r, s) \mapsto r \cdot s)$$

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

$$\cdot : R \times M \to M \ ((r, m) \mapsto r \cdot m)$$

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

$$\cdot : M \times R \to M \ ((m, r) \mapsto m \cdot r)$$

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 $f : M \to N$ 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 16.1 *Sei $f : M \to N$ 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 $f(M) = \mathrm{Im}(f)$ und $f^{-1}(\{0\}) = \mathrm{Ker}(f)$ Untermoduln. Ein R-Homomorphismus $f : M \to N$ ist genau dann surjektiv, wenn $\mathrm{Im}(f) = M$ ist und genau dann injektiv, wenn $\mathrm{Ker}(f) = \{0\}$ 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 Reprasentantenunabhängigkeit 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 $p_U : M \to U$ und $p_V : M \to V$ definieren:

$$p_U(M) = m,\ p_V(m) = v.$$

Diese Abbildungen sind R-linear und es gilt

$$p_U \circ p_U = p_U,\ p_V \circ p_V = p_V,\ p_U \circ p_V = p_V \circ p_U = 0,\ p_U + p_V = id_M.$$

Wir nennen diese die Projektionen auf die Summanden U, V.

Da jeder R-Homomorphismus $f : M \to N$ auch ein Gruppenhomomorphismus ist, haben wir nach dem Homomorphiesatz einen Isomorphismus

$$F : M/\mathrm{Ker}(f) \to \mathrm{Im}(f),$$

der durch $F(m + \mathrm{Ker}(f)) = f(m)$ gegeben ist. Dies ist sogar ein R-Isomorphismus, denn

$$F(r(m + \mathrm{Ker}(f))) = f(rm) = rf(m) = rF(m + \mathrm{Ker}(f)).$$

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 $m_1, \ldots, m_k \in M$ und $r_1, \ldots, r_k \in R$, 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 $m_1, \ldots, m_k$ heißen linear unabhängig, wenn aus

$$\sum r_i m_i = 0 \ (r_i \in R)$$

folgt, daß $r_1 = \ldots = r_k = 0$ gilt. Ein linear unabhängiges Erzeugendensystem von M heißt eine Basis, ein R-Modul, der eine Basis besitzt, heißt frei.

Lemma 16.2 *Sei M ein freier R-Modul und $\{m_1, \ldots, m_n\}$ eine Basis von M, dann ist $M \simeq R \times \ldots \times R = R^n$.*

Beweis: Jedes $m \in M$ läßt sich in eindeutiger Weise als Linearkombination $\sum r_i m_i$ darstellen, wir ordnen m das n-tupel $(r_1, \ldots, r_n) \in R^n$ zu. □

Lemma 16.3 *Jeder endlich erzeugte R-Modul ist isomorph zu einem Faktormodul eines freien R-Moduls.*

Beweis: Sei $M = R\{m_1, \ldots, m_n\}$ und $m \in M$ beliebig, also $m = \sum r_i m_i$, dann ist die Abbildung $f : R^n \to M$ mit $f(r_1, \ldots, r_n) \mapsto \sum r_i m_i$ surjektiv. Wir setzen $U = \mathrm{Ker}(F)$, dann gilt nach dem Homomorphiesatz $M \simeq R^n/U$. □

Definition: Eine additive Untergruppe $L \in R$ 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 $D \in R$ 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 $I \in R$, die sowohl ein Links- als auch ein Rechtsideal ist, heißt (zweiseitiges) Ideal.

Seien R und S zwei Ringe. Ein Homomorphismus $f : R \to S$ 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

$$\mathrm{Ker}(f) = \{r \in R \mid f(r) = 0\}.$$

Ein Ringhomomorphismus ist genau dann injektiv, wenn $\mathrm{Ker}(f) = \{0\}$ ist.

Lemma 16.4 *Sei $f : R \to S$ ein Ringhomomorphismus, dann ist* $\mathrm{Ker}(f)$ *ein Ideal von R.*

Beweis: Die Abbildung f ist ein Gruppenhomomorphismus, also ist $\mathrm{Ker}(f) \subseteq R$ eine Untergruppe. Sei $a \in \mathrm{Ker}(f)$ 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 $ra \in \mathrm{Ker}(f)$ und $ar \in \mathrm{Ker}(f)$. □

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

$$f : R \to R/I, f(r) = r + I,$$

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 $a \mid b$. □

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 geegneten 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 $f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_n \in \mathbb{Z}$, dann heißt die Zahl $\mathrm{cont}(f) = \mathrm{ggT}(a_0, \ldots, a_n)$ der Inhalt von $f(x)$. Ein Polynom $f(x) \in \mathbb{Z}[x]$ heißt primitiv, wenn sein Inhalt gleich 1 ist.

Lemma 16.5 (Hilfssatz von Gauß) *Das Produkt primitiver Polynome ist primitiv.*

Beweis: Seien $f, g \in \mathbb{Z}[x]$ primitv 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. □

Satz 16.6 (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 $a = \mathrm{cont}(g^{\#}(x))$, 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 wegkürzen, also muß $\frac{a}{b}$ eine ganze Zahl sein. □

Ein Kriterium, ob ein Polynom mit ganzzahligen Koeffizienten irreduzibel ist, ist durch den folgenden Satz gegeben.

Satz 16.7 (Satz von Eisenstein) *Sei $f(x) = a_n x^n + \ldots + a_0 \in \mathbb{Z}[x]$ und p eine Primzahl, so daß p kein Teiler von a_n ist, $p \mid a_{n-1}, \ldots, p \mid a_0$, aber p^2 kein Teiler von a_0 ist. Dann ist $f(x)$ irreduzibel.*

Beweis: Sonst wäre

$$f(x) = (b_m x^m + \ldots + b_0)(c_l x^l + \ldots + c_0)$$

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

$$b_0 \equiv \ldots \equiv b_{k-1} \equiv 0 \pmod{p},$$

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. □

Wir wenden dieses Kriterium auf das Polynom $f(x) = \dfrac{x^p - 1}{x - 1} = x^{p-1} + x^{p-2} + \cdots + x + 1$ an, wo p eine Primzahl ist.

Wir setzen $x = y + 1$:

$$\frac{(y+1)^p - 1}{y} = \frac{1}{y}\sum_{i=1}^{p}\binom{p}{i}y^i = y^{p-1} + \binom{p}{1}y^{p-2} + \cdots + \binom{p}{p-1}.$$

Die Binomialkoeffizienten

$$\binom{p}{i} = \frac{p\cdot(p-1)\cdots(p-i+1)}{1\cdot 2\cdots i} \in \mathbb{Z}$$

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.

æ

16.2 Universelle Konstruktionen; abstract nonsens

Der Begriff der *Klasse* wird nicht definiert; stellen Sie sich eine Ansammlung von *Objekten* vor. Wenn M ein Objekt der Klasse C ist, so schreiben wir $M \in C$.

Definition
Eine Kategorie C besteht aus einer Klasse $Ob(C)$ von Objekten, wobei zu jedem Paar $A, B \in Ob(C)$ eine Menge $Mor(A, B)$ von Morphismen existiert. Weiter soll zu je drei Objekten A, B, C eine Kompositionsabbildung

$$Mor(B, C) \times Mor(A, B) \longrightarrow Mor(A, C)$$

$$(f, g) \mapsto f \circ g$$

existieren, so daß folgende Eigenschaften erfüllt sind:

1. $Mor(A, B) \cap Mor(A', B') = \emptyset$, wenn $A \neq A'$ oder $B \neq B'$ ist.
2. Zu jedem $A \in Ob(C)$ gibt es einen Morphismus $id_A \in Mor(A, A)$, so daß für alle $f \in Mor(A, B),\ g \in Mor(B, A)$ gilt
$$f \circ id_A = f,\ id_A \circ g = g.$$
3. Wenn $f \in Mor(A, B),\ g \in Mor(B, C),\ h \in Mor(C, D)$ gilt, so folgt
$$(h \circ g) \circ f = h \circ (g \circ f).$$

Anstelle von $f \in Mor(A, B)$ schreiben wir auch einfach $f : A \longrightarrow B$.

Beispiele für Kategorien sind die Kategorie der Vektorrräume über einem Körper K mit den linearen Abbildungen als Morphismen, die Kategorie der Gruppen (der endlichen Gruppen, der abelschen Gruppen) mit den Gruppenhomomorphismen als Morphismen, die Kategorie der linken Moduln über einem Ring R mit den R-Modulhomomorphismen

als Morphismen sowie die Kategorie der Mengen mit den Abbildungen als Morphismen. In all diesen Beispielen ist der Morphismus id_A die identische Abbildung.

Definition
Ein Morphismus $f : A \longrightarrow B$ heißt Isomorphismus, wenn ein Morphismus $g : B \longrightarrow A$ mit $f \circ g = id_B,\ g \circ f = id_A$ existiert. Ein Morphismus f heißt Monomorphismus, wenn aus $f \circ g_1 = f \circ g_2$ folgt $g_1 = g_2$. Ein Morphismus f heißt Epimorphismus, wenn aus $g_1 \circ f = g_2 \circ f$ folgt $g_1 = g_2$.

Die Begriffe „Monomorphismus“ und „Epimorphismus“ sind *duale* Begriffe; sie vertauschen sich, wenn man die Pfeilrichtungen vertauscht.

Satz 16.8 *In den obengenannten Kategorien sind die Monomorphismen genau die injektiven Abbildungen und Epimorphismen sind genau die surjektiven Abbildungen.*

Beweis: (Wir behandeln der Einfachheit halber die Kategorie der Mengen.) Es sei $f : Y \longrightarrow Z$ eine injektive Abbildung, $g_1, g_2 : X \longrightarrow Y$ und $fg_1 = fg_2$, dann ist $f(g_1(x)) = f(g_2(x))$ für alle x; wegen der Injektivität folgt $g_1(x) = g_2(x)$ für alle x, also $g_1 = g_2$.
Sei f nicht injektiv, also $f(y_1) = f(y_2)$ mit $y_1 \neq y_2$). Wir wählen $X = \{x\}$ und setzen $g_1(x) = y_1,\ g_2(x) = y_2$. Dann ist $g_1 \neq g_2$, aber $fg_1 = fg_2$.
Nun sei $f : X \longrightarrow Y$ surjektiv und $g_1, g_2 : Y \longrightarrow Z$ mit $g_1 f = g_2 f$ gegeben. Dann gilt $g_1(f(x)) = g_2(f(x))$ für alle x. Jedes $y \in Y$ hat wegen der Surjektivität die Form $y = f(x)$, also gilt $g_1(y) = g_2(y)$ für alle y, also $g_1 = g_1$.
Sei f nicht injektiv, dann gibt es ein $y \notin Im(f)$, wir wählen ein $y_1 \neq y$ und setzen $Z = Y,\ g_1 = id_Y,\ g_2(y') = y'$ für $y' \neq y,\ g_2(y) = y_1$. Dann gilt $g_1 f = g_2 f$. □

Satz 16.9 *Wenn f, g Monomorphismen sind, so ist $f \circ g$ auch ein Monomorphismus. Wenn f, g Epimorphismen sind, so ist $f \circ g$ auch ein Epimorphismus. Wenn $f \circ g$ ein Monomorphismus ist, so ist g ein Monomorphismus. Wenn $f \circ g$ ein Epimorphismus ist, so ist f ein Monomorphismus.*

Beweis: Wir beweisen jeweils die Aussagen für Monomorphismen, der Rest geschieht durch Dualisierung.
Sei $f \circ g \circ h_1 = f \circ g \circ h_2$, dann ist wegen der Monomorphie von f auch $g \circ h_1 = g \circ h_2$ und aus der Monomorphie von g folgt $h_1 = h_2$.
Sei $g \circ h_1 = g \circ h_2$, dann gilt $(f \circ g) \circ h_1 = (f \circ g) \circ h_2$, also $h_1 = h_2$. □

Definition

Seien A, B Objekte; ein Produkt von A, B ist ein Tripel $(p, f : P \longrightarrow A, g : P \longrightarrow B)$, so daß für alle $\phi : C \longrightarrow A,\ \psi : C \longrightarrow B$ ein eindeutig bestimmter Morphismus $h : C \longrightarrow P$ existiert, so daß $\phi = f \circ h,\ \psi = g \circ h$ gilt, d.h. das folgende Diagramm ist kommutativ:

$$\begin{tikzcd}
 & C \arrow[d, "h"] \arrow[ddl, "\phi"'] \arrow[ddr, "\psi"] & \\
 & P \arrow[dl, "f"] \arrow[dr, "g"'] & \\
A & & B
\end{tikzcd}$$

Satz 16.10 *Wenn ein Produkt von A, B existiert, dann ist es bis auf Isomorphie eindeutig bestimmt; man bezeichnet es mit $A \times B$ oder $A \sqcup B$.*

Beweis: Seien (P_1, f_1, g_1) und (P_2, f_2, g_2) zwei Produkte von A und B. Dann existieren eindeutig bestimmte h_1, h_2 mit $f_1 = f_2 \circ h_1,\ f_2 = f_1 \circ h_2$ und $g_1 = g_2 \circ h_1,\ g_2 = g_1 \circ h_2$.

$$\begin{tikzcd}
 & P_1 \arrow[dl, "f_1"'] \arrow[d, shift right, "h_1"'] \arrow[dr, "g_1"] & \\
A & P_2 \arrow[l, "f_2"'] \arrow[r, "g_2"] \arrow[u, shift right, "h_2"'] & B
\end{tikzcd}$$

Es folgt $f_1 = f_1 \circ h_2 \circ h_1$ und $g_1 = g_1 \circ h_2 \circ h_1$. Wegen der Einzigkeit von h folgt $h_2 \circ h_1 = id_{P_1}$ und analog $h_1 \circ h_2 = id_{P_2}$, also sind P_1 und P_2 isomorph. □

Satz 16.11 *In der Kategorie der Mengen (Gruppen, R-Moduln,...) existiert das Produkt je zweier Objekte, und zwar ist es das kartesische Produkt zusammen mit den beiden Projektionen auf die Faktoren.*

Beweis:

$$\begin{tikzcd}
 & C \arrow[d, "h"] \arrow[ddl, "\phi"'] \arrow[ddr, "\psi"] & \\
 & A \times B \arrow[dl, "f"] \arrow[dr, "g"'] & \\
A & & B
\end{tikzcd}$$

Die Abbildung h ist durch $h(c) = (\phi(c), \psi(c))$ gegeben. □

Wir dualisieren den Begriff des Produkts.

Definition

Seien A, B Objekte; ein Koprodukt von A, B ist ein Tripel $(p, f : A \longrightarrow S, g : B \longrightarrow S)$, so daß für alle $\phi : A \longrightarrow C,\ \psi : B \longrightarrow C$ ein eindeutig bestimmter Morphismus $h : S \longrightarrow C$ existiert, so daß $\phi = h \circ f,\ \psi = h \circ g$ gilt.

Satz 16.12 *Wenn ein Koprodukt von A, B existiert, dann ist es bis auf Isomorphie eindeutig bestimmt; man bezeichnet es mit $A \oplus B$ oder $A \sqcup B$.*

Der Beweis ist dual zum obigen. □

Satz 16.13 *In der Kategorie der Mengen existiert das Koprodukt beliebiger Mengen A, B, es besteht aus der disjunkten Vereinigung $A \sqcup B$ zusammen mit den beiden Einbettungen von A udn B in $A \sqcup B$.*

Beweis: Die gesuchte Abbildung $h : A \sqcup B \longrightarrow C$ ist durch $h(a) = \phi(a),\ h(b) = \psi(b)$ gegeben. □

Satz 16.14 *In der Kategorie der R-Moduln existiert das Koprodukt beliebiger Moduln A, B, es besteht aus der direkten Summe $A \oplus B$ zusammen mit den beiden Einbettungen von A und B in $A \oplus B$.*

Beweis: Die gesuchte Abbildung $h : A \oplus B \longrightarrow C$ ist durch $h(a + b) = \phi(a) + \psi(b)$ gegeben. □

Definition
Ein Objekt U heißt universell, wenn es zu jedem Objekt A genau einen Morphismus $f_A : U \longrightarrow A$ gibt.

Wir fragen uns nach der Existenz. Offenbar sind alle universellen Objekte einer Kategorie isomorph, falls welche existieren.

In der Kategorie der Mengen ist die leere Menge universell, in der Kategorie der Gruppen leistet dies eine einelementige Gruppe, in der Kategorie der R-Moduln ist der Nullmodul universell.

Wir betrachten nun eine Kategorie, deren Objekte keine Mengen (mit einer algebraischen Struktur) sind:

Sei S eine Menge und R ein Ring. Die Objekte der Kategorie R_S seien Abbildungen $\phi : S \longrightarrow M$, wobei M ein R-Modul ist. Ein Morphismus von $\phi : S \longrightarrow M$ in $\psi : S \longrightarrow N$ ist ein R-Modulhomomorphismus $f : M \longrightarrow N$, so daß das Diagramm

$$\begin{tikzcd} & M \arrow[dd, "f"] \\ S \arrow[ur, "\phi"] \arrow[dr, "\psi"'] & \\ & N \end{tikzcd}$$

kommutativ wird, d.h. $f \circ \phi = \psi$.

Ein universelles Objekt der Kategorie R_S ist also eine Abbildung $\phi : S \longrightarrow F$ in einen R-Modul F, so daß zu jeder Abbildung $\psi : S \longrightarrow M$ in einen R-Modul M ein eindeutig bestimmter R-Modulhomomorphismus $f : F \longrightarrow M$ mit kommutativem Diagramm

$$\begin{tikzcd} & \mathrm{F} \arrow[dd, "f"] \\ \mathrm{S} \arrow[ur, "\phi"] \arrow[dr, "\psi"'] & \\ & \mathrm{M} \end{tikzcd}$$

existiert. Wir werden gleich sehen, daß ein derartiges universelles Objekt existiert; der Modul F heißt der von S erzeugte freie R-Modul.

Satz 16.15 *Sei $F = \{\sum r_i s_i \in r_i \in R,\ s_i \in S\}$ die Menge aller formaler Linearkombinationen und $\phi : S \hookrightarrow F$ die Einbettung; dies ist ein universelles Objekt in R_S.*

Beweis: Offenbar ist F ein R-Modul. Wenn $\psi : S \longrightarrow M$ eine beliebige Abbildung ist, so ist durch $f(\sum r_i s_i) = \sum r_i \psi(s_i)$ ein R-Modulhomomorphismus gegeben und es gilt $f \circ \phi = \psi$. □

Definition
Ein Objekt O heißt Nullobjekt, wenn es sowohl universell als auch kouniversell ist, d.h. zu jedem Objekt A gibt es genau einen Morphismus $O \longrightarrow A$ und genau einen Morphismus $A \longrightarrow O$.
Ein Morphismus $o : A \longrightarrow B$ heißt Nullmorphismus, wenn er gleich einem Produkt $A \longrightarrow O \longrightarrow B$ ist.
Sei $f : A \longrightarrow B$ ein Morphismus; ein Morphismus $k : K \longrightarrow A$ heißt Kern von f, wenn $f \circ k = o$ ist und wenn es zu jedem $g : G \longrightarrow A$ mit $f \circ g = o$ einen eindeutig bestimmten Morphismus $h : G \longrightarrow A$ mit $g = k \circ h$ gibt.
$K \to A \to B$
$\uparrow \nearrow$
G

Lemma 16.16 *Jeder Kern ist ein Monomorphismus.*

Seien $g_1, g_2 : X \longrightarrow K$ Morphismen mit $kg_1 = kg_2$. Wir setzen $g = kg_1 : X \longrightarrow A$, dann ist $fg = fkg_1 = o$, also existiert ein eindeutig bestimmtes $h : X \longrightarrow K$ mit $g = kh = kg_1 = kg_2$, also ist $h = g_1 = g_2$. □

Lemma 16.17 *Der Kern eines Monomorphismus ist der Nullmorphismus.*

Beweis: Aus $fk = o = fo$ und der Monomorphie von f folgt $k = o$. □

Nun nehmen Sie ein Buch über Kategorientheorie und beweisen Sie als Übungsaufgaben alle Sätze.

16.3 Tensorprodukte

Wir machen eine Vorbemerkung: Sei im folgenden sei stets R ein kommutativer Ring und M und N seien R-Moduln.

Wir machen eine Vorbemerkung:
Sei $f : M \longrightarrow N$ eine R-lineare Abbildung und $U \subset M$ ein Untermodul mit $f(U) = \{0\}$. Dann wird durch f eine Abbildung $f' : M/U \longrightarrow N$ induziert, wir setzen $f'(m + U) = f(m)$, dies ist wohldefiniert, wie man schnell nachrechnet.

Wir wollen einen neuen R-Modul $M \otimes_R N$ konstruieren, der von Elementen der Form $m \otimes n,\ m \in M, n \in N$ erzeugt wird, wobei folgende Rechenregeln gelten sollen:

$$(m + m') \otimes n = m \otimes n + m' \otimes n$$

$$m \otimes (n + n') = m \otimes n + m \otimes n'$$

$$(rm) \otimes n = m \otimes (rn), \qquad (r \in R)$$

Wir konstruieren diesen Modul folgendermaßen: Sei $F(M \times N)$ der von der *Menge* $M \times N$ erzeugte freie R-Modul, sei S der Untermodul von $F(M \times N)$, der von allen Elementen der Form

$$(m + m', n) - (m, n) - (m', n)$$

$$(m, n + n') - (m, n) - (m, n')$$

$$(rm, n) - (m, rn)$$

$$(m, m' \in M,\ n, n' \in N,\ a \in R)$$

erzeugt wird.
Wir setzen $M \otimes_R N = F(M \times N)/S$, die Äquivalenzklasse von (m, n) bezeichnen wir mit $m \otimes n$, dann sind die obigen Rechenregeln erfüllt. Der Modul $M \otimes_R N$ wird das Tensorprodukt von M und N genannt. Die Elemente von $M \otimes N$ haben die Gestalt $\sum_{i \in I} m_i \otimes n_i,\ m_i \in M,\ n_i \in N$.

Ein Element der Form $m \otimes n$ heißt zerfallender Tensor, die zerfallenden Tensoren bilden ein Erzeugendensystem. Zum Beispiel ist

$$m_1 \otimes n_1 + 2m_1 \otimes n_2 + 2m_2 \otimes n_1 + 4m_2 \otimes n_2$$

$$= m_1 \otimes (n_1 + 2n_2) + 2m_2(n_1 + 2n_2) = (m_1 + 2m_2) \otimes (n_1 + 2n_2)$$

ein zerfallender Tensor.

Definition: Seien M, N, P R-Moduln; eine Abbildung $f : M \times N \longrightarrow P$ mit

$$f(m + m', n) = f(m, n) + f(m, n')$$

$$f(m, n + n') = f(m, n) + f(m, n')$$

$$f(rm, n) = rf(m, n) = f(m, rn)$$

heißt R-bilinear.

Beispiel: Die kanonische Abbildung $k : M \times N \longrightarrow M \otimes_R N,\ (m, n) \mapsto m \otimes n$, ist bilinear.

Satz 16.18 *Sei $f : M \times N \longrightarrow P$ eine bilineare Abbildung. Dann gibt es eine eindeutig bestimmte R-lineare Abbildung $g : M \otimes_R N \longrightarrow R$, so daß das Diagramm*

$$\begin{tikzcd} M \times N \arrow[r] \arrow[d, "k"'] & P \\ M \otimes_R N \arrow[ur, "g"'] & \end{tikzcd}$$

kommutativ ist. („Die bilineare Abbildung f induziert die lineare Abbildung g".)

Beweis: Wir setzen zunächst f linear auf $F(M \times N)$ fort:

$$f' : F(M \times N) \longrightarrow P, \quad f'(\sum r_i(m_i, n_i)) = \sum r_i f(m_i, n_i)).$$

Wir überlegen, daß $f'(s) = 0$ für $s \in S$ gilt. Sei etwa $s = (m + m') - (m, n) - (m', n)$, dann ist

$$\begin{aligned} f'(s) &= f'((m + m', n) - (m, n) - (m', n)) \\ &= f'(m + m') - f'(m, n) - f'(m'n) \\ &= f(m + m') - f(m, n) - f(m'n) = 0 \end{aligned}$$

wegen der Bilinearität von f. Also Induziert f' eine lineare Abbildung $g : M \otimes_R N \longrightarrow P$ mit $g(m \otimes n) = g((m, n) + S) = f'(m, n) = f(m, n)$.
Das die Elemente der Form $m \otimes n$ den R-Modul $M \otimes_R N$ erzeugen, ist die Abbildung g eindeutig bestimmt. □

Satz 16.19 (Isomorphieeigenschaften) *1. $M \otimes_R N \cong N \otimes_R M$,*
2. $(M \otimes_R N) \otimes_R P \cong M \otimes_R (N \otimes_R P)$,
3. $R \otimes_R M \cong M \cong M \otimes_R R$.

Beweis: 1. Die Abbildung

$$f : N \times M \longrightarrow M \otimes_R N, \; f(n, m) = m \otimes n,$$

ist bilinear und induziert eine lineare Abbildung $g : N \otimes_R M \longrightarrow M \otimes_R N$, diese besitzt offensichtlich eine Inverse.
2. Wir fixieren ein $p \in P$ und betrachten

$$h_p : M \times N \longrightarrow M \otimes_R (N \otimes_R P),$$

$$h_p(m, n) = m \otimes (n \otimes p),$$

sie ist bilinear und induziert eine lineare Abbildung

$$g_p : M \otimes_R N \longrightarrow M \otimes_R (N \otimes_R P).$$

Wir betrachten nun

$$g : (M \otimes_R N) \times P \longrightarrow M \otimes_R (N \otimes_R P),$$

$$(m \otimes n, p) \mapsto m \otimes (n \otimes p),$$

diese Abbildung ist bilinear und induziert eine lineare Abbildung

$$(M \otimes_R N) \otimes_R P \longrightarrow M \otimes_R (N \otimes_R P),$$

die wiederum offensichtlich eine Inverse besitzt.
3. Die Abbildung $f : A \times M,\ f(a, m) = am$, ist bilinear, usw. □

Seien $f : M \longrightarrow P,\ g : N \longrightarrow Q$ R-lineare Abbildungen, dann ist die Abbildung

$$h : M \times N \longrightarrow P \otimes_R Q$$

$$(m, n) \mapsto f(m) \otimes g(n)$$

bilinear, also wird eine lineare Abbildung

$$f \otimes g : M \otimes_R N \longrightarrow P \otimes_R Q$$

induziert, sie heißt das Tensorprodukt der Abbildungen f und g.

Lemma 16.20

$$\begin{aligned}
f \otimes g(m \otimes n) &= f(m) \otimes g(n) \\
(f + f') \otimes g &= f \otimes g + f' \otimes g \\
f \otimes (g + g') &= f \otimes g + f \otimes g' \\
(af) \otimes g &= f \otimes (ag) = a\, f \otimes g \\
(f_2 \circ f_1) \otimes (g_2 \circ g_1) &= (f_2 \otimes g_2) \circ (f_1 \otimes g_1)
\end{aligned}$$ □

Lemma 16.21 $(M \oplus M') \otimes_R N \cong M \otimes_R N \oplus M' \otimes_R N$.

Beweis: Sei $P = M \oplus N$, wir können die direkte Summe durch idempotente Endomorphismen charakterisieren: Wir haben die Projektionen

$$e_1, e_2 : P \longrightarrow P,\ e_1 + e_2 = id, e_i \circ e_j = \delta_{ij}\, e_i,\ M = \operatorname{Im} e_1,\ M' = \operatorname{Im} e_2.$$

Wir tensorieren mit $id : N \longrightarrow N$, also $f_i = e_i \otimes id : P \otimes_R N \longrightarrow P \otimes_R N$ und erhalten für die f_i analoge Relationen. □

Sei nun $M = R^m$ ein freier R-Modul, dann folgt

$$M \otimes_R N = \left(\bigoplus_{i=1}^{m} R \right) \otimes_R N = \bigoplus_{i=1}^{m} (R \otimes_R N) \cong \bigoplus_{i=1}^{m} N = N^m.$$

Wenn auch noch $N = R^n$ frei ist, so folgt

$$M \otimes_R N = (R^m)^n = R^{mn}.$$

Folgerung 16.22 *Sei K ein Körper und seien V, W K-Vektorräume, dann gilt* $\dim V \otimes_K W = \dim V \cdot \dim W$. *Seien $\{v_1, \ldots, v_n\}$, $\{w_1, \ldots, w_m\}$ Basen von V bzw. W, dann ist $\{v_i \otimes w_j \mid i = 1, \ldots, n;\ j = 1 \ldots, m\}$ eine Basis von $V \otimes_K W$ und jedes Element $t \in V \otimes_K W$ läßt sich mit $t = \sum t_{ij}\, v_i \otimes w_j$ durch einen „Tensor“ (t_{ij}) beschreiben.*

Seien nun $f_1 : V_1 \longrightarrow W_1$, $f_2 : V_2 \longrightarrow W_2$ lineare Abbildungen, dazu erhalten wir die lineare Abbildung

$$f_1 \otimes f_2 : V_1 \otimes V_2 \longrightarrow W_1 \otimes W_2,$$

wir wollen deren Darstellungsmatrix bestimmen. Dazu wählen wir Basen $\{b_i\}$ von V_1, $\{c_j\}$ von V_2, $\{d_k\}$ von W_1, $\{e_l\}$ von W_2 und ordnen die Basiselemente von $V_1 \otimes V_2$ bzw. $W_1 \otimes W_2$ folgendermaßen an:

$$G = \{b_1 \otimes c_1, \ldots, b_n \otimes c_1, b_1 \otimes c_2, \ldots, b_n \otimes c_2, \ldots, b_1 \otimes c_m, \ldots, b_n \otimes c_m\},$$

$$H = \{d_1 \otimes e_1, \ldots, d_p \otimes e_1, d_1 \otimes e_2, \ldots, d_p \otimes c_2, \ldots, d_1 \otimes e_q, \ldots, d_p \otimes e_q\}.$$

Sei $X = A_{BD}(f_1)$, $Y = A_{CE}(f_2)$, dann hat A_{GH} die folgende Block-Gestalt:

$$\begin{pmatrix} Xy_{11} & \ldots & Xy_{1m} \\ & \ldots & \\ Xy_{p1} & \ldots & Xy_{pq} \end{pmatrix}.$$

Diese Matrix heißt das Kroneckerprodukt (oder auch Tensorprodukt) der Matrizen X und Y.

Wir betrachten die folgende (offenbar lineare) Abbildung

$$f : V \otimes W^* \longrightarrow \mathrm{Hom}(W, V), f(v \otimes l) = f_{v \otimes l},$$

$$f_{v \otimes l}(w) = l(w)\, v.$$

(W^* ist der zu W duale Vektorraum.
Wir zeigen, daß f ein Isomorphismus ist. Da

$$\dim(V \otimes W^*) = \dim V \cdot \dim W^* = \dim V \cdot \dim W = \dim \mathrm{Hom}(W, V)$$

ist, genügt es, die Surjektivität nachzuweisen.
Seien $B = \{b_i\}$, $\{c_j\}$ Basen von W bzw. V. Dann bilden die Abbildungen $G_{ij} : W \longrightarrow V$ mit $g_{ij}(b_k) = \delta_{ik}\, c_j$ eine Basis von $\mathrm{Hom}(W, V)$. Sei $\{b_i^*\}$ die B duale Basis von W^*, dann ist $f(c_j \otimes b_i^*) = g_{ij}$, folglich ist f surjektiv.

Folgerung 16.23

$$\mathrm{Hom}(V, W) \cong V^* \otimes W. \square$$

Wir kehren wieder zu beliebigen Grundringen zurück.

Satz 16.24 *Seien $U \subset M$, $T \subset N$ Untermoduln, dann ist*

$$(M/U) \otimes_R (N/T) \cong M \otimes_R N/(U \otimes_R N + M \otimes_R T).$$

Beweis: Seien $f : M \longrightarrow M/U$, $g : N \longrightarrow N/T$ die kanonischen Abbildungen, wir betrachten die Abbildung

$$f \otimes g : M \otimes_R N \longrightarrow M/U \otimes_R N/T$$

$$m \otimes n \mapsto \bar{m} \otimes \bar{n} = f(m) \otimes g(n).$$

Sei $m \in U$ oder $t \in T$, dann ist $f \otimes g(m \otimes n) = 0$, also

$$f \otimes g|_{U \otimes_R N + M \otimes_R T} = 0,$$

also wird eine Abbildung

$$h : M \otimes N/(U \otimes_R N + M \otimes_R T) \longrightarrow M/U \otimes_R N/T$$

induziert, wobei $\overline{m \otimes n} \mapsto \bar{m} \otimes \bar{n}$. Wir zeigen, daß die Umkehrabbildung

$$k :\longrightarrow M/U \otimes_R N/T \longrightarrow M \otimes N/(U \otimes_R N + M \otimes_R T)$$

$$\bar{m} \otimes \bar{n} \mapsto \overline{m \otimes n}$$

wohldefiniert ist:

$$(\overline{m+u} \otimes \overline{n+t}) = \bar{m} \otimes \bar{n} + \bar{u} \otimes \bar{n} + \bar{m} \otimes t + \bar{u} \otimes \bar{t},$$

der erste Summand wir auf $\overline{m \otimes n}$, die restlichen auf Null abgebildet.

Folgerung 16.25 *Seien I und J Ideale von R, dann gilt*

$$R/I \otimes_R R/J \cong A/(I+J). \square$$

Beispiel: $\mathbf{Z}/2\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}/3\mathbf{Z} = 0, \mathbf{Z}/2\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}/2\mathbf{Z} = \mathbf{Z}/2\mathbf{Z}$.

16.4 Das Jacobson-Radikal

In diesem Anschnitt bezeichne R stets einen Ring, der nicht notwendigerweise eine 1 besitzen muß. Dann können wir eine Eins „adjungieren“: In der additiven Gruppe $R^1 = \mathbf{Z} \oplus R$ führen wir auf natürliche Weise eine Multiplikation ein:

$$(n+r)(m+q) = nm + mr + nq + rq,$$

der erste Summand liegt in $\mathbf{Z}$, die anderen in R.

Für $x \in R$ betrachten wir folgende Abbildung

$$L(x) : R \longrightarrow R,\ L(x)(y) = xy.$$

Aus dem Assoziativgesetz $x(yz) = (xy)z$ folgt

$$L(x)(L((y)(z)) = L(xy)(z),$$

also

$$L(x) \circ L(y) = L(xy),$$

und aus dem Distributivgesetz $(x+y)z = xy + xz$ folgt

$$L(x+y)(z) = L(x)(z) + L(y)(z) = (L(x) + L(y))(z),$$

also

$$L(x+y) = L(x) + L(z),$$

somit ist $L : R \longrightarrow \mathrm{End}_R(R)$ ein Ringhomomorphismus.
Wenn R eine 1 besitzt, so ist $\mathrm{End}(R) \cong R$, dann entspricht L der identischen Abbildung.

Definition: Ein Element $x \in R$ heißt quasi-invertierbar mit dem Quasi-Inversen y, wenn $1 - x \in R^1$ invertierbar mit dem Inversen $1 + y$ ist.

Bemerkung: Sei $1-x \in R^1$ invertierbar mit einem Inversen $m+z$, dann gilt $(1-x)(m+z) = m - mx + z - xz = 1$, also muß $m = 1$ sein, d.h. das Inverse hat stets die Form $1 + y$.

Lemma 16.26 *Wenn R eine Eins besitzt, dann sind x und $L(x)$ gleichzeig invertierbar.*

Satz 16.27 *Für $x \in R$ ist äquivalent:*
a) x ist quasi-invertierbar,
b) es gibt ein $y \in R$ mit $y - x = xy = yx$,
c) $id - L(x) \in \mathrm{End}R$ ist invertierbar.
Das Quasi-Inverse von x ist in diesem Fall gleich $y = (id - L(x))^{-1}(x)$.

Beweis: a) $\Leftrightarrow$ b) folgt aus $(1-x)(1+y) = 1 - x + y - xy = 1$.
a) $\Leftrightarrow$ c): $1 - x \in R^1$ ist genau dann invertierbar, wenn $L(1-x) \in R^1$ invertierbar ist. Es gilt $L(1-x)(n+r) = n - nx + r - xr$, dies ist gleich 0 genau dann, wenn $n = 0$ und $r - xy = 0$ ist, also wenn $r \in \mathrm{Ker}(id - L(x)$, also ist $L(1-x)$ genau dann bijektiv, wenn $id - L(x) \in \mathrm{End}(R)$ bijektiv ist. □

Beispiele: 1. Wann ist $x \in \mathbf{Z}$ invertierbar? Es muß also $y - x = xy$ sein, d.h. $y(1-x) = x$. Dies ist einerseits für $x = y = 0$ der Fall, sonst ist $x-1$ ein Teiler von x, also $x - 1 \leq \sqrt{x}$, hieraus folgt $x^2 - 1 \leq 3x$ und damit $x = 2,\ y = -2$.
2. Wenn x nilpotent vom Grade n ist, dann gilt

$$(1-x)(1 + x + x^2 + \cdots + x^{n-1}) = 1,$$

also ist x quasi-invertierbar.

Wir fixieren ein Element $u \in R$ und führen in R eine neue Multiplikation ein:

$$x \circ y = x\,u\,y.$$

Wenn wir die Abhängigkeit von u hervorheben wollen, schreiben wir $\circ_u$.
Die Ringaxiome für die neue Multiplikation berprüft man schnell, z.B.

$$(x \circ y) \circ z = xuyuz = x \circ (y + z),$$

$$x \circ (y+z) = xu(y+z) = xuy + xuz = x \circ y + x \circ z.$$

(Das neutrale Element wäre u^{-1}, falls dies existiert.)
Wir beweichnen diesen Ring mit R_u, er heißt das u-Homotop von R. Die Linksmultiplikation in R_u bezeichnen wir mit $L_u : L_u(x) = x \circ y = xuy$, also

$$L_u(x) = L(x)\, L(u) = L(xu).$$

Folgerung 16.28 *Für $x \in R$ sind äquivalent:*
a) x ist in R_u quasi-invertierbar,
es gibt ein y mit $y - x = xuy = yux$,
c) $id - L(xu)$ ist invertierbar.
Das Quasi-Inverse von x ist in diesem Fall gleich $(id - L(xu))^{-1}(x)$. □

Wir bezeichnen in folgenden das Quasi-Inverse von x in R_u mit $q(x,u)$, falls es existiert, und setzen $B(x,u) = id - L(xu) = id - L_u(x)$.

Satz 16.29 (Symmetriesatz) *Für $x, y \in R$ sind äquivalent:*
a) $q(x,u)$ existiert,
b) $q(xu,1)$ existiert,
c) $q(u,x)$ existiert (d.h. u ist quasi-invertierbar in R_x),
d) $q(ux,1)$ existiert,
e) $B(x,u)$ ist invertierbar,
f) $B(u,x)$ ist invertierbar.
Dann gilt $q(x,u) = B(x,u)^{-1}(x)$ und $q(u,x) = uq(x,u)u + u$.

Beweis: a) $\Leftrightarrow$ e) folgt aus der Folgerung, b) $\Leftrightarrow$ e) folgt aus dem obigen Satz.
a) $\Leftrightarrow$ c): Sei $w = q(x,u)$, also $w - x = x \circ_u w = xuw = wux$, dann setzen wir $z = uwu + u$ und erhalten

$$z - u = (uwu + u) - u = uwu = uxu + uxuwu = u \circ_x (u + uwu) = u \circ_x z,$$

also $z - u = u \circ_x z$, d.h. $z = q(u,x)$.
Der Rest folgt aus Symmetriegründen. □

Bemerkung: Seien $a, b \in R^1, x \in R$, dann ist

$$axb = (m + a')x(n + b') = \cdots \in R \ (!).$$

Satz 16.30 (Verschiebungssatz) *Seien $a, b \in R^1,\ x, u \in R$, dann existiert $q(axb,u)$ genau dann, wenn $q(x,bua)$ existiert, und dann gilt $q(axb,u) = aq(x,bua)b$.*

Beweis: Sei $w = q(x,bua)$, also $w - x = wbuax = xbuaw$, also $awb - axb = (awb)u(axb) = (axb)u(awb)$, d.h. es existiert $q(axb,u) = awb$.
Wenn umgekehrt $q(axb,u)$ existiert, dann existiert wegen der Symmetrie auch $q(u,axb)$, wegen des soeben Bewiesenen folgt die Existenz von $q(bua,x)$ und wegen der Symmetrie folgt wieder die Existenz von $q(x,bua)$. □

Satz 16.31 (Additionssatz) *Es existiere $q(x,u)$, dann gilt*

$$B(x,u)B(q(x,u),z) = B(x,u+z)$$

und $q(q(x,u),z)$ existiert genau dann, wenn $q(x,u+z)$ existiert; in diesem Fall gilt $q(q(x,u),z) = q(x,u+z)$.

Beweis: Sei $w = q(x,u)$, also $w - x = wux = xuw$, daraus folgt $xuwz = wz - xz$. Weiter gilt

$$\begin{aligned} B(x,u)B(q(x,u),z) &= (id - L(xu))(id - L(xz)) \\ &= id - L(xu) - L(wz) + L(xuwz) \\ &= id - L(x(u+z)) \\ &= B(y,u+z). \end{aligned}$$

$B(x,u)$ ist invertierbar, also ist $B(q(x,u),z)$ genau dann invertierbar, wenn $B(x,u+z)$ invertierbar ist, und dann gilt

$$q(x,u+z) = B(x,u+z)^{-1}(x) = B(w,z)^{-1}B(x,u)^{-1}(x) = B(w,z)^{-1}(w) = q(w,z). \square$$

Satz 16.32 *$J = \{x \in R \mid x$ ist quasi-invertierbar allen Ringen $R_u\}$ ist ein Ideal von R.*

Beweis: Seien $a, b \in R^1$, $x \in J$. Dann existiert $q(x,u)$, also existiert auch $q(x,aub)$ für jedes u. Durch Verschiebung ehalten wir: $q(bxa,u)$ existiert für alle u, d.h. $bxa \in J$. Seien $x, y \in J$, wegen der Symmetrie existieren dann $q(u,x)$ und $q(v,y)$ für alle u, v. Wir setzen speziell $v = q(u,x)$, dann existiert $q(q(u,x),y) = q(u,x+y)$ und damit existiert $q(x+y,u)$ für alle u, also ist $x+y \in J$. $\square$

Definition: $J = \mathrm{Rad}(R)$ heißt das Jacobson-Radikal von R. Wenn $\mathrm{Rad}(R) = 0$ ist, so heißt R semi-primitiv.

Satz 16.33

$$\mathrm{Rad}(R/\mathrm{Rad}(R)) = \{0\}$$

.

Beweis: Wir setzen $\bar{R} = R/\mathrm{Rad}(R)$, sei $\bar{x} \in \mathrm{Rad}(\bar{R})$, also gibt es zu jedem $\bar{u} \in \bar{R}$ ein $\bar{w} \in \bar{R}$ mit $\bar{w} - \bar{x} = \overline{wux} = \overline{xuw}$, also folgt für die Reprräsentanten

$$w - x - wux \in \mathrm{Rad}(R).$$

Damit ist $B(w - x - xuw, -u)$ invertierbar und

$$\begin{aligned} B(w - x - xuw, -u) &= id - L(-wu + xu - xuwu) \\ &= (id - L(xu))(id + L(wu)) \\ &= B(x,u)B(w,u), \end{aligned}$$

also ist auch $B(x,u)$ invertierbar und damit $x \in \mathrm{Rad}(R)$, folglich $\bar{x} = \bar{0}$. $\square$

Definition: Ein (Links-, Rechts-) Ideal von R heißt quasi-invertierbar, wenn jedes Element quasi-invertierbar ist. Es heißt nil, wenn jedes Element nilpotent ist.

Satz 16.34 *L sei ein quasi-invertierbares Linksisdeal, dann ist $L \subset \mathrm{Rad}(R)$.*

Beweis: Seien $x \in L,\ u \in R$, dann ist $ux \in L$, also existiert $q(ux, 1)$, folglich exisiert $q(x, u)$ für alle u, also ist $x \in \mathrm{Rad}(R)$. □

Folgerung 16.35 *Jedes nil-Linksideal ist in* $\mathrm{Rad}(R)$ *enthalten.*

Folgerung 16.36 $\mathrm{Rad}(R)$ *ist das maximale quasi-invertierbare Linksideal von R.*

Beweis: Sei $x \in \mathrm{Rad}(R)$, wir zeigen, daß x quasi-invertierbar ist.
Zunächst existiert $q(x, x)$, also auch $q(x^2, 1)$, d.h.$1 - x^2 \in R^1$ ist invertierbar. Wegen $1 - x^2 = (1 - x)(1 + x)$ ist auch $(1 - x) \in R^1$ invertierbar, also ist x quasi-invertierbar. □

Folgerung 16.37 $\mathrm{Rad}(R) = \{x \in R \mid xu$ *ist quasi-invertierbar für alle* $u\}$.

Der Beweis folgt aus dem Symmetriesatz. □

Bemerkung: Sei I ein Links- oder Rechtsideal und $x \in I$ quasi-invertierbar in R_u, d.h. $q(x, u)$ und $q(u, x)$ existieren. Dann liegt $q(x, u) = xq(u, x)x + x$ ebenfalls in I.

Satz 16.38 *Sei $I \subset R$ ein Ideal, dann ist* $\mathrm{Rad}(I) = I \cap \mathrm{Rad}(R)$.

Beweis: Sei $x \in I \cap \mathrm{Rad}(R)$, dann ist $w = q(x, u) \in I$, dann gilt $w - x = wux = xuw$ speziell für alle $u \in I$, also ist $x \in \mathrm{Rad}(I)$.
Sei umgekehrt $x \in \mathrm{Rad}(I)$, d.h. zu jedem $y \in I$ gibt es ein $v \in I$ mit $v - x = vyx = xyv$, d.h. $q(x, y)$ existiert für alle $y \in I$. Folglich existiert $q(x, uxu)$ für alle $u \in R$, also ist die folgende Abbildung invertierbar:

$$\begin{aligned} B(x, uxu) &= id - L(xuxu) \\ &= (id - L(xu))(id + L(xu)) \\ &= B(x, u)B(x, -u), \end{aligned}$$

damit ist auch $B(x, u)$ invertierbar, d.h. $q(x, u)$ existiert, also ist $x \in \mathrm{Rad}(R)$. Es folgt $\mathrm{Rad}(U) \subset I \cap \mathrm{Rad}(R)$. □

Folgerung 16.39 *Jedes Ideal eines semi-primitiven Rings ist semi-primitiv.* □

Satz 16.40

$$\mathrm{Rad}(R_u) = \{x \in R \mid uxu \in \mathrm{Rad}(R)\}$$

$$\mathrm{Rad}(R) = \bigcap_{u \in R} \mathrm{Rad}(R_u)$$

Beweis: 1. Für die Homotope von R_u gilt $(R_u)_v = R_{uvu}$, denn $x \circ_{uvu} y = xuvuy = x \circ_u v \circ_u y$.
2.

$$\begin{aligned} x \in \mathrm{Rad}(R_u) \quad &\Leftrightarrow \quad x \text{ ist quasi-invertierbar in } (R_u)_v \text{ für alle } v \\ &\Leftrightarrow \quad x \text{ ist quasi-invertierbar in } R_{uvu} \text{ für alle } v \\ &\Leftrightarrow \quad q(x, uvu) \text{ existiert für alle } v \\ &\Leftrightarrow \quad q(uxu, v) \text{ existiert für alle } v \text{ (Verschiebung)} \\ &\Leftrightarrow \quad uxu \in \mathrm{Rad}(R). \end{aligned}$$

3. $\mathrm{Rad}(R)$ ist ein Ideal, also gilt $u\mathrm{Rad}(R)u \subset \mathrm{Rad}(R)$ für alle u, also gilt $\mathrm{Rad}(R) \subset \mathrm{Rad}(R_u)$ und damit $\mathrm{Rad}(R) \subset \bigcap \mathrm{Rad}(R_u)$. Sei umgekehrt $x \in \mathrm{Rad}(R_u)$ für alle u, dann existiert $q(uxu, y)$ für alle $u,\ y \in R$, also ist die Abbildung $B(uxu, x) = B(u, x)B(u, -x)$ invertierbar, damit ist auch $B(u, x)$ invertierbar, also ist $x \in \mathrm{Rad}(R)$. □

Wir berechnen zum Abschluß das Radikal eines Matrixrings.

Satz 16.41

$$\mathrm{Rad}(M_{nn}(R)) = M_{nn}(\mathrm{Rad}(R)).$$

Beweis: Sei $x \in \mathrm{Rad}(R)$ und $M = \begin{array}{cc} \left(\begin{array}{ccc} & & \\ & u & \leftarrow \\ & \uparrow & \\ & i & \end{array}\right) & j \end{array}$. Es existiert ein w mit $w - x = wux = xuw$. Dann gilt

$$wE_{ij} - xE_{ij} = (wE_{ij})\, M\, (xE_{ij}) = (xE_{ij})\, M\, (wE_{ij}),$$

also $xE_{ij} \in \mathrm{Rad}(M_{nn}(R))$.
Sei umgekehrt $X \in \mathrm{Rad}(M_{nn})$ und u beliebig, dann gibt es zu uE_{ij} ein W mit $W - X = WuE_{ij}X = XuE_{ij}W$. Wir betrachten die Stelle (i, j):

$$w_{ij} - x_{ij} = w_{ij}ux_{ij} = x_{ij}uw_{ij},$$

also ist $x_{ij} \in \mathrm{Rad}(R)$. □

Beispiele: 1. Wenn K ein Körper ist, dann ist $\mathrm{Rad}(K) = \{0\}$, also ist auch $\mathrm{Rad}(M_{nn}(K)) = \{0\}$.
2. $\mathrm{Rad}(K[[x]]) = \{\sum a_i x^i \mid a_0 = 0\}$.

Kapitel 17

Halbeinfache Algebren und Moduln

17.1 Grundlagen

Definition: Sei K ein Körper, A ein K-Vektorraum und gleichzeitig ein Ring; vermöge der Abbildung $K \to A$ mit $k \mapsto k \cdot 1$ wird K in A eingebettet, wir identifizieren K mit $K \cdot 1$. Es gelte

$$k \cdot a = a \cdot k \text{ für alle } k \in K, a \in A.$$

Dann heißt A eine K-Algebra.

Sei M ein linker A-Modul, wegen $K \subseteq A$ operiert auch K auf M, d.h. M ist auch ein K-Vektorraum.

Beispiele:
$K[x]$, $K[x_1, \ldots, x_n], K[x_1, \ldots, x_n]/I$ für ein Ideal I von $K[x_1, \ldots, x_n]$, $M_{nn}(K)$, die Menge $T_n(K)$ der oberen Dreiecksmatrizen, die Menge D_n der Diaginalmatrizen, $K \times \ldots \times K$ mit komponentenweiser Addition und Multiplikation.

Wir vereinbaren, daß alle in diesem Abschnitt betrachteten Vektorräume endlichdimensional sind (dann fallen zwei der obigen Beipiele aus dem Rahmen).

Sei $\dim A = n$, dann wählen wir eine Basis $\{e_1, \ldots, e_n\}$ des Vektorraums A, dann lassen sich die Produkte $e_i e_j$ als Linearkombination der e_k darstellen:

$$e_i e_j = \sum a_{ijk} e_k \text{ mit } a_{ijk} \in K.$$

Die n^3 Zahlen a_{ijk} heißen die Strukturkonstanten der Algebra (bezüglich der gewählten Basis), durch sie ist die Multiplikation in A eindeutig bestimmt:

$$\sum x_i e_i \cdot \sum y_j e_j = \sum x_i y_j a_{ijk} e_k.$$

Die Strukturkonstanten sind zur Konstruktion einer Algebra nicht willkürlich wählbar, denn die Multiplikation soll assoziativ sein, also muß gelten:

$$(e_i e_j) e_l = \sum a_{ijk} e_k e_l = \sum a_{ijk} a_{klm} e_m,$$

$$e_i(e_je_l) = e\sum_i a_{jlk}e_k = \sum a_{jlk}a_{ikm}e_m,$$

dafür ist notwendig und hinreichend, daß

$$\sum a_{ijk}a_{klm} = \sum a_{jlk}a_{ikm} \text{ für alle } i, j, l, m.$$

Eine Menge S heißt Halbgruppe, wenn eine Multiplikation $\cdot : S \times S \to S$ gegeben ist, die das Assoziativgesetz erfüllt und für die es ein neutrales Element 1 gibt.

Sei S eine endliche Halbgruppe, wir ordnen ihr die folgende „Halbgruppenalgebra“ K^S zu. Wir setzen

$$K^S = \{f : S \to K\}$$

und definieren Addition, K-Multiplikation und Ringmultiplikation wie folgt:

$$(f_1 + f_2)(s) = f_1(s) + f_2(s),$$

$$(k \cdot f)(s) = k \cdot f(s),$$

$$(f_1 \cdot f_2)(s) = \sum_{rt=s} f_1(r) \cdot f_2(t),$$

Es ist nicht schwierig nachzurechnen, daß K^S eine K-Algebra der Dimension $|\,S\,|$ ist.

Definition: Sei A ein Ring und M ein linker A-Modul, M heißt einfach, wenn M keine echten Untermoduln besitzt. Ein Linksideal $L \subseteq A$ heißt minimal, wenn $\{0\}$ das einzige echt in L enthaltene Linksideal ist. Ein Linksideal L heißt maximal, wenn A das einzige L echt enthaltende Linksideal ist.

Satz 17.1 *1. Jedes minimale Linksideal ist ein einfacher A-Modul.*
2. Für jedes maximale Linksideal $L \subseteq A$ ist A/L ein einfacher A-Modul.
3. Jeder einfache A-Modul ist isomorph zu A/L für ein geeignetes maximales Linksideal $L \subseteq A$.

Beweis: Die beiden ersten Aussagen sind trivial, wir beweisen nur die dritte: Sei M einfach und $0 \neq m \in M$, dann ist $\{0\} \neq Am \subseteq M$ ein Untermodul, also $Am = M$. Wir betrachten den folgenden Modulhomomorphismus

$$f : A \to M,\ a \mapsto am,$$

dieser ist offenbar surjektiv und $L = \mathrm{Ker}(f)$ ist ein Linksideal von A. Nach dem Homomorphiesatz gilt $A/L \cong M$ und wegen der Einfachheit von M muß L maximal sein. □

Beispiele:
Sei $A = K$, die einfachen K-Moduln sind 1-dimensionale Vektorräume, also isomorph zu K, K besitzt das minimale Ideal K und das maximale Ideal $\{0\}$.

Wir betrachten die Matrixalgebra $A = M_{22}(K)$ und deren Unterräume, wir bezeichen mit $\begin{pmatrix} * & 0 \\ * & 0 \end{pmatrix}$ die Menge aller Matrizen, deren zweite Spalte Null ist. Die Menge $\begin{pmatrix} * & 0 \\ * & 0 \end{pmatrix}$ ist ein minimales und ein maximales Linksideal von A, der Vektorraum $K^2 = \begin{pmatrix} * \\ * \end{pmatrix}$ der Spaltenvektoren ist ein einfacher A-Modul.

Der Original-Bewies des folgenden Satzes ist etwa eine Druckseite lang.

Lemma 17.2 (Schur) *1. Sei M ein einfacher A-Modul und $f : M \to M$ eine A-lineare Abbildung, dann ist f ein Isomorphismus oder $f = 0$.*
2. Sei A eine $\mathbb{C}$-Algebra, M ein einfacher A-Modul und $f : M \to M$ eine A-lineare Abbildung, dann gilt $f = z \cdot id_M$ für ein $z \in \mathbb{C}$.

Beweis: 1. $\mathrm{Ker}(f)$, $\mathrm{Im}(f) \subseteq M$ sind Untermoduln, also müssen sie gleich $\{0\}$ oder gleich M sein.
2. Sei $z \in \mathbb{C}$ ein Eigenwert der linearen Abbildung f, dann ist $f - z \cdot id_M$ kein Isomorphismus, also $f - z \cdot id = 0$. □

Eine A-lineare Abbildung eines A-Moduls M in sich nennt man einen A-Endomorphismus, die Menge aller A-Endomorphismen von M wird mit $\mathrm{End}_A(M)$ bezeichnet. Mit der Nacheinanderausführung von Endomorphismen als Multiplikation wird $\mathrm{End}_A(M)$ eine K-Algebra.

Folgerung 17.3 *Sei A eine $\mathbb{C}$-Algebra und M ein einfacher A-Modul. Dann gilt $End_A(M) \cong \mathbb{C}$.*

Beweis: Wir ordnen dem Endomorphismus f seinen Eigenwert zu. □

Definition: Eine K-Algebra A heißt einfache Algebra, wenn A genau zwei Ideale besitzt, nämlich $\{0\}$ und A.

Beispiele:
1. Ein Körper K ist eine einfache K-Algebra.
2. Sei R ein kommutativer Ring ohne echte Ideale. Sei $0 \neq r \in R$, dann ist das von r erzeugte Ideal $rR \neq \{0\}$, also $rR = R \ni 1$, also gibt es ein $s \in R$ mit $rs = 1$. Jedes von Null verschiedene Element von R besitzt ein Inverses, also ist R ein Körper.
3. Wir wollen nachweisen, daß die Matrixalgebra $A = M_{nn}(K)$ eine einfache K-Algebra ist. Sei also $\{0\} \neq I \subseteq A$ ein Ideal. Dann muß für jede Matrix $M = (m_{ij}) \in I$ und alle Matrizen $X, Y \in A$ auch das Produkt XMY in I liegen. Mit E_{ij} bezeichnen wir die Matrix, die nur an der Stelle (i, j) eine 1 und sonst Nullen enthält. Sei $m_{ij} \neq 0$, dann gilt

$$\frac{1}{m_{ij}}(E_{1i}ME_{j1} + E_{2i}ME_{j2} + \ldots + E_{ni}ME_{jn}) = E \in I,$$

also $I = A$.

Satz 17.4 (Wedderburn) *Jede (endlichdimensionale) einfache $\mathbb{C}$-Algebra A ist isomorph zu einer Matrixalgebra $M_{nn}(\mathbb{C})$.*

Beweis: Sei $L \subseteq A$ ein minimales Linksideal. Zum Element $x \in L$ betrachten wir die Abbildung

$$f_x : L \to L,\ f_x(l) = l \cdot x,$$

sie ist A-linear, denn

$$f_x(al_1 + l_2) = (al_1 + l_2)x = al_1x + l_2x = af_x(l_1) + f_x(l_2).$$

Also gilt $f_x \in \mathrm{End}_A(L) \cong \mathbb{C}$, also $f_x = z_x \cdot id$ mit $z_x \in \mathbb{C}$. Sei nun $a \in A$ beliebig, wir betrachten die Abbildung

$$g_a : L \to L,\ g_a(l) = al.$$

Die Abbildung g_a ist offenbar $\mathbb{C}$-linear. Es gilt

$$g_1 = id_A,\ g_{a+b} = g_a + g_b,\ g_{ab} = g_a \cdot g_b,$$

also ist

$$g : A \to \ \mathrm{End}_{\mathbb{C}}(L),\ g(a) = g_a,$$

ein Ringhomomorphismus, es ist $g \neq 0$, somit ist $\mathrm{Ker}(g) \neq A$ ein Ideal von A, also $\mathrm{Ker}(g) = \{0\}$, somit ist g injektiv. Wir wollen noch beweisen, daß g surjektiv ist.
Die Menge $L \cdot A$ ist ein Ideal von A, also gilt $L \cdot A = A$, also

$$g(A) = g(LA) = g(L) \cdot g(A).$$

Wir zeigen nun, daß $g(L) \subseteq \mathrm{End}_{\mathbb{C}}(L)$ ein Linksideal ist.
Sei also $h \in \mathrm{End}_{\mathbb{C}}(L)$ und $x, l \in L$. Dann gilt $h \circ g_l(x) = h(lx) = h(f_x(l)) = h(z_x \cdot l) = z_x \cdot h(l) = f_x \circ h(l) = h(l) \cdot x = g_{h(l)}(x)$, also $h \circ g_l = g_{h(l)} \in g(L)$. Folglich ist $\mathrm{End}_{\mathbb{C}}(L) \cdot g(L) = g(L)$ und wegen $1 \in g(A)$ gilt $\mathrm{End}_{\mathbb{C}}(L) \cdot g(A) = \mathrm{End}_{\mathbb{C}}(L)$.
Nun folgt $g(A) = g(L) \cdot g(A) = \mathrm{End}_{\mathbb{C}}(L) \cdot g(L) \cdot g(A) = \mathrm{End}_{\mathbb{C}}(L) \cdot g(A) = \mathrm{End}_{\mathbb{C}}(L)$, also ist g surjektiv, d.h. $A \cong \mathrm{End}_{\mathbb{C}}(L)$.
Wenn nun $\dim_{\mathbb{C}}(L) = n$ ist, so ist $\mathrm{End}_{\mathbb{C}}(L) \cong M_{nn}(\mathbb{C})$. □

Definition: Sei A eine beliebige K-Algebra, ein A-Modul M heißt halbeinfach, wenn $M = M_1 \oplus \ldots \oplus M_k$ eine direkte Summe einfacher A-Moduln ist.

Satz 17.5 *Sei M ein halbeinfacher A-Modul und $U \subseteq M$ ein Untermodul. Dann gibt es einen halbeinfachen Untermodul $V \subseteq M$ mit $U \oplus V = M$.*

Beweis: Sei $M = \bigoplus_{i \in I} M_i$ mit einfachen Moduln M_i. Wir wählen eine maximale Teilmenge $J \subseteq I$, so daß die Summe $U + \bigoplus_{j \in J} M_j$ direkt ist. Dann ist

$$M_i \subseteq U + \bigoplus_{j \in J} M_j \text{ für alle } i \in I,$$

andernfalls wäre

$$M_i \cap (U + \bigoplus_{j \in J} M_j) \subset M_i, \text{ also}$$

$$M_i \cap (U + \bigoplus_{j \in J} M_j) = \{0\},$$

daß heißt, die Menge J wäre nicht maximal. Damit gilt

$$M = \bigoplus M_i = U + \bigoplus_{j \in J} M_j$$

und $V = \bigoplus_{j \in J} M_j$ ist ein halbeinfacher A-Modul. □

Wir können auch die Umkehrung beweisen:

Satz 17.6 *Sei M ein A-Modul, $\dim_K(M) < \infty$, jeder Untermodul von M sei ein direkter Summand. Dann ist M halbeinfach.*

Beweis: Wir wählen einen minimalen Untermodul $M_1 \subseteq M$, dieser ist ein einfacher Modul und es gilt

$$M = M_1 \oplus U_1$$

für einen Untermodul U_1. Sei $M_2 \subseteq U_1$ ein minimaler (also einfacher) Untermodul, damm ist $M_1 \cap M_2 = \{0\}$, also ist deren Summe direkt und es gibt einen Untermodul U_2 mit $M = M_1 \oplus M_2 \oplus U_2$, usw.
Nach endlich vielen Schritten haben wir M in einfache Untermoduln zerlegt. □

Satz 17.7 *Untermoduln und Faktormoduln halbeinfacher Moduln sind halbeinfach.*

Beweis: Sei M halbeinfach und $U \subseteq M$ ein Untermodul. Dann existiert ein halbeinfacher Modul $V \subseteq M$ mit $U \oplus V = M$, und zu V existiert ein halbeinfacher Modul $W \subseteq M$ mit $V \oplus W = M$. Nun ist

$$M/V = (U \oplus V)/V \cong U \cong (V \oplus W)/V \cong W,$$

also ist U halbeinfach, und

$$M/U \cong V$$

ist ebenfalls halbeinfach. □

Satz 17.8 *Sei $M = M_1 \oplus \ldots \oplus M_n$ mit einfachen Untermoduln M_i und $U \subseteq M$ sei ein weiterer einfacher Untermodul. Es gelte $U \cong M_i$ für $i = 1, \ldots, r$ und $M_j \ncong U$ für $j > r$. Dann gilt $U \subseteq M_1 \oplus \ldots \oplus M_r$.*

Beweis: Es sei $p_i : M \to M_i$ die Projektion; für $u \in U$ gilt dann

$$u = \sum p_i(u)$$

mit $p_i(u) \in M_i$. Die Einschränkung $p_i \mid U : U \to M_i$ ist A-linear, wenn also $p_i(u) \neq 0$ ist, so ist $p_i \mid U$ ein Isomorphismus zwischen U und M_i, also $p_j \mid U = 0 für j >$ r. □

Wir betrachten nun eine spezielle Klasse von K-Algebren, die sich als direkte Summen von Linksidealen darstellen lassen. Dazu erweitern wir unsere Kenntnisse über Modulhomomorphismen. Die Menge aller A-linearen Abbildungen $f : M \to N$ zwischen zwei R-Moduln bezeichnen wir mit $\mathrm{Hom}_A(M, N)$.

Lemma 17.9 $\mathrm{Hom}_A(A, M) \cong M$.

Beweis: Sei $f : A \to M$ eine A-lineare Abbildung, dann gilt

$$f(r) = f(r \cdot 1) = rf(1) \text{ für alle } r \in A,$$

also ist f durch $f(1)$ eindeutig bestimmt. □

Folgerung 17.10 $\mathrm{End}_A(A) \cong A$. □

Wir hatten früher gesehen, daß zu direkten Zerlegungen

$$M = M_1 \oplus \ldots M_n$$

Endomorphismen $p_1, \ldots, p_n \in \mathrm{End}(M)$ gehören, für die $p_i \circ p_j = p_i \delta_{ij}$ galt (man nennt solche Elemente „orthogonale Idempotente“) und es galt $M_i = p_i(M)$.
Wenn wir nun eine Zerlegung einer Algebra $A = L_1 \oplus \ldots \oplus L_n$ in Linksideale vornehmen, so entspricht dem die Existenz orthogonaler Idempotenter e_i in $\mathrm{End}_R(R) = R$ und es gilt $L_i = Ae_i$.

Definition: Eine K-Algebra A heißt halbeinfach, wenn $A = L_1 \oplus \ldots \oplus L_n$ eine direkte Summe minimaler Linksideale ist.

Eine Algebra ist also genau dann halbeinfach, wenn sie als linker Modul über sich selbst halbeinfach ist. Wir zeigen, daß dies keine besondere Bevorzugung der linken Seite bedeutet.

Satz 17.11 *Eine Algebra A ist genau dann eine direkte Summe minimaler Linksideale, wenn sie eine direkte Summe minimaler Rechtsideale ist.*

Beweis: Sei $A = \bigoplus L_i = \bigoplus Ae_i$ mit $e_i e_j = e_i \delta_{ij}$ und $\sum e_i = 1$ eine Zerlegung von A in eine direkte Summe minimaler Linksideale. Dann ist

$$A = \bigoplus e_i A$$

eine direkte Zerlegung in Rechtsideale. Wir zeigen, daß diese minimal sind.
Sei $a \in e_i A e_i \subseteq e_i A$, also $a = e_i x e_i$ für ein x, dann ist $e_i a = e_i^2 x e_i = e_i x e_i = a$, also $aA \subseteq e_i A$. Analog gilt $Aa \subseteq Ae_i$, also $Aa = Ae_i$. Wir betrachten die Abbildung $f : Ae_i \to Aa$ mit $f(xe_i) = xe_i a = xa$, dies ist ein Homomorphismus linker A-Moduln, also ein Isomorphismus. Weiter sei $A = Aa \oplus U$, wir definieren $g : A \to A$ durch $g(xa + u) = f^{-1}(xa) = xe_i$, dies ist ein Homomorphismus linker A-Moduln, also $g(y) = y \cdot g(1)$, wir setzen $g(1) = b$. Dann ist $e_i = g(b) = a \cdot b$, also $e_i \in aA \subseteq e_i A$, also $aA = e_i A$. □

Die Moduln halbeinfacher Algebren sind besonders einfach:

Satz 17.12 *Sei M ein (endlich erzeugter) Modul über der halbeinfachen Algebra A. Dann ist M halbeinfach.*

Beweis: Es gibt einen freien A-Modul F, so daß $M \cong F/U$ mit $U \subseteq F$ gilt. Mit A ist auch F halbeinfach, also auch dessen Faktormudul M. □

Satz 17.13 *Jeder einfache Modul einer halbeinfachen Algebra A ist isomorph zu einem minimalen Linksideal von A.*

Beweis: Sei M ein einfacher A-Modul, dann gilt $M \cong A/L$ für ein maximales Linksideal L. Es gibt ein Linksideal $H \subseteq A$ mit $A = L \oplus H$ und da L maximal ist, kann H nur minimal sein. Nun gilt aber $M \cong A/L \cong H$. □

Wie sehen eigentlich die zweiseitigen Ideale einer halbeinfachen Algebra aus?

Satz 17.14 *Sei $A = \oplus L_i$ eine halbeinfache Algebra, die L_i seien minimale Linksideale und $L_1 \cong \ldots \cong L_r$ und $L_1 \ncong L_i$ für $i > r$. Dann ist $I = L_1 \oplus \ldots \oplus L_r$ ein minimales zweiseitiges Ideal von A.*

Beweis: Sei $a \in A$; zu $i \leq r$ betrachten wir die Abbildung $f_a : L_i \to A$ mit $f_a(l) = la$, die Abbildung f_a ist A-linear, also ist $\mathrm{Im}(f_a) = \{0\}$ oder $\mathrm{Im}(f_a) \cong L_i$. Folglich ist $\mathrm{Im}(f_a) = L_i a \subseteq L_1 \oplus \ldots \oplus L_r$, also ist I ein zweiseitiges Ideal.
Wir zeigen noch: jedes in I enthaltene minimale Linksideal (etwa L_1) erzeugt I als Ideal.
Sei $p_1 : A \to L_1$ die Projektion und $f : L_1 \to L_j$ ein A-Isomorphismus. Wir betrachten $f \circ p_1 : A \to L_j$, es ist $f \circ p_1(a) = af \circ p_1(1)$ für $a \in A$. Sei $l \in L_1$, dann gilt $p_1(l) = l$, also $f \circ p_1(l) = f(l) = lf \circ p_1(1)$, also ist $L_j = f(L_1) = L_1 \cdot f \circ p_1(1) \subseteq L_1 A$, also ist I ein minimales Ideal. □

Folgerung 17.15 *Sei A eine halbeinfache Algebra, dann ist A eine direkte Summe minimaler Ideale: $A = I_1 \oplus \ldots \oplus I_s$, jedes I_i ist eine direkte Summe paarweise isomorpher minimaler Linksideale.* □

Satz 17.16 *Sei $A = I_1 \oplus \ldots \oplus I_s$ mit minimalen Idealen I_i. Dann gilt $I_i \cdot I_j = \{0\}$ für $i \neq j$ und jedes I_i ist eine einfache Algebra.*

Beweis: *Für*$i \neq j$ gilt $I_i \cdot I_j \subseteq I_i \cap I_j = \{0\}$. Sei $1 = e_1 + \ldots + e_s$ mit $e_i \in I_i$, dann ist e_i das neutrale Element von I_i. Sei $J \subseteq I_1$ ein Ideal von I_1, also $I_1 J I_1 = J$, dann ist $AJA = AI_1JI_1A = I_1JI_1 = J$, da $AI_1 = I_1A = I_1$ ist. Also ist J ein Ideal von A, also $J = \{0\}$ oder $J = I_1$. □

Sei nun $G = \{g_1, \ldots, g_n\}$ eine endliche Gruppe. Mit

$$KG = \{\sum r_i g_i \mid r_i \in K\}$$

bezeichnen wir die Menge aller formaler Linearkombinationen der Elemente der Gruppe G, dies ist ein n-dimensionaler Vektorraum. Wir führen eine Multiplikation ein, die durch die Multiplikation in G induziert wird:

$$(\sum r_i g_i)(\sum s_j g_j) = \sum r_i s_j (g_i g_j),$$

diese erfüllt offenbar das Assoziativgesetz, die Körperelemente kommutieren mit allen Elementen und das Einselement von G ist das neutrale Element. Die Menge KG ist also eine K-Algebra, sie heißt die Gruppenalgebra der Gruppe G.
Wir bemerken, daß KG isomorph zu Halbgruppenalgebra K^G ist.

Satz 17.17 (Maschke) *Wenn die Zahl $|G|$ in K invertierbar ist, so ist KG eine halbeinfache Algebra.*

Beweis: Wir zeigen viel mehr: Jeder Untermodul eines KG-Moduls ist ein direkter Summand.
Sei M ein KG-Modul und $U \subseteq M$ ein Untermodul. Speziell ist U ein Unterraum von M und es gibt einen Unterrraum V von M mit $M = U \oplus V$, also gibt es eine K-lineare

Abbildung $p : M \to M$ mit $p^2 = p$ und $p(M) = U$, nämlich die Projektion auf U. Wir konstruieren nun eine KG-lineare Abbildung:

$$q(x) = \frac{1}{|\,G\,|} \sum g_i^{-1} p(g_i x) \text{ für } x \in M.$$

Zunächst gilt $\mathrm{Im}(q) \subseteq \mathrm{Im}(p)$, wenn $u \in U$ ist, so gilt

$$q(u) = \frac{1}{|\,G\,|} \sum g_i^{-1} p(g_i u) = q = \frac{1}{|\,G\,|} \sum g_i^{-1} g_i u = u,$$

denn wegen $g_i u \in U$ ist $p(g_i u) = g_i u$, also gilt $\mathrm{Im}(q) = \mathrm{Im}(p)$. Sei nun $h \in G$, dann ist

$$h^{-1} q(hx) = \frac{1}{|\,G\,|} \sum h^{-1} g_i^{-1} p(g_i h x) = q(x),$$

denn mit g_i durchläuft auch $g_i h$ die Gruppe G, d.h.

$$q(hx) = hq(x),$$

also ist q ein KG-Homomorphismus. Schließlich ist

$$q \circ q(x) = q(\frac{1}{|\,G\,|} \sum g_i^{-1} p(g_i x)) = \frac{1}{|\,G\,|} \sum g_i^{-1} q \circ p(g_i x),$$

wegen $p(g_i x) \in U$ ist dies gleich $\frac{1}{|\,G\,|} \sum g_i^{-1} p(g_i x) = q(x)$, also ist q idempotent und damit $M = \mathrm{Im}(q) \oplus \mathrm{Ker}(q)$, somit ist U ein direkter Summand von M. □

17.2 Darstellungen endlicher Gruppen

In diesem Abschnitt werden wir uns mit den Elementen der Darstellungstheorie beschäftigen. Das Ziel kann grob so umrissen werden, daß „abstrakte“ Gruppen als „konkrete“ Gruppen von Matrizen beschrieben werden sollen.

Definition: Sei G eine Gruppe und V ein Vektorraum über dem Körper K, dann heißt ein Gruppenhomomorphismus

$$\rho : G \longrightarrow GL(V)$$

eine Darstellung von G in V; ein Gruppenhomomorphismus

$$R : G \longrightarrow GL(n, K)$$

heißt Matrixdarstellung von G über K vom Grade n.

Für eine Darstellung ρ von G gilt also für $g, h \in G$: $\rho(gh) = \rho(g) \circ \rho(h)$ sowie $\rho(g^{-1}) = \rho(g)^{-1}$, $\rho(1) = id$.

Sei nun $t : V \xrightarrow{\sim} W$ ein Isomorphismus von Vektorräumen und ρ eine Darstellung von G, dann haben wir also Automorphismen $\rho(g) : V \longrightarrow V$. Wir konstruieren nun eine neue Darstellung $\rho' : G \longrightarrow GL(W)$ wie folgt: $\rho'(g) : W \longrightarrow W$ sei durch

$$\rho'(g) = t \circ \rho(g) \circ t^{-1}$$

gegeben, d.h. wir haben ein kommutatives Diagramm

$$\begin{array}{ccccc} & V & \xrightarrow{\rho(g)} & V & \\ t & \downarrow & & \downarrow & t \\ & W & \xrightarrow{\rho'(g)} & W & \end{array}$$

und es gilt $\rho'(gh) = t\rho(g)t^{-1}t\rho(h)t^{-1} = \rho'(g)\rho'(h)$, also ist ρ' eine Darstellung von G in W.

Definition: Zwei Darstellungen $\rho : G \longrightarrow GL(V),\ \rho' : G \longrightarrow GL(W)$ heißen äquivalent, wenn ein Isomorphismus $t : V \xrightarrow{\sim} W$ existiert, so daß

$$\rho'(g) = t \circ \rho(g) \circ t^{-1} \text{ für alle } g \in G$$

gilt.

Zu einer Darstellung ρ von G erhalten wir eine Matrixdarstellung, indem wir eine Basis B in V wählen und jedem Element $g \in G$ die Darstellungsmatrix des Automorphismus $\rho(g)$ zuordnen:

$$R(g) = A_{BB}(\rho(g)).$$

Wenn B' eine andere Basis und X die Basiswechselmatrix ist, so erhalten wir eine neue Matrixdarstellung R', die aber wegen $R'(g) = X^{-1}R(g)X$ für alle $g \in G$ zu R äquivalent ist.

Beispiele:
1. $\rho_1 : G \longrightarrow GL(K) = K^*,\ \rho_1(g) = 1$ heißt 1-Darstellung von G.
2. X sei eine Menge, auf der die Gruppe G operiere, d.h. es gibt eine Abbildung $G \times X \longrightarrow X,\ (g, x) \mapsto g \cdot x$ mit $(gh)x = g(hx),\ 1x = x$. Dann ist $V = \{f : X \longrightarrow K\}$ ein Vektorraum und die Abbildung $\rho : G \longrightarrow GL(V)$ mit

$$(\rho(g)(f))(x) = f(g^{-1}x)$$

ist eine Darstellung, denn

$$(\rho(gh)(f))(x) = f(h^{-1}g^{-1}x) = (\rho(h)(f))(g^{-1}x) = \rho(g)(r(h)(f))(x).$$

3. Sei $G = \{1 = g_1, g_2, \ldots, g_n\},\ V = L(b_1, \ldots, b_n)$ ein n-dimensionaler Vektorraum, wir setzen

$$\rho(g_i)(b_j) = b_m, \text{ falls } g_i g_j = g_m$$

ist. Wegen $g_i G = G$ ist $\rho(g_i)$ invertierbar, daß ρ ein Homomorphismus ist, rechnet man leicht nach. Diese Darstellung heißt die reguläre Darstellung von G .

4. Sei $K_4 = \{1, g, h, gh\}$ die Kleinsche Vierergruppe. Sei R ihre reguläre Matrixdarstellung, dann haben wir

$$\begin{aligned} R(g)(b_1) &= b_2, \text{ da } g \cdot 1 = g \\ R(g)(b_2) &= b_1, \text{ da } g^2 = 1 \\ R(g)(b_3) &= b_4, \\ R(g)(b_4) &= b_3. \end{aligned}$$

also

$$R(g) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

und analog für die anderen Elemente.
5. Sei $G = C - n = \langle a \rangle,\ a^n = 1$ die zyklische Gruppe der Ordnung n, für deren reguläre Darstellung gilt

$$\begin{aligned} \rho(a)(b_i) &= b_{i+1},\ i = 1, \ldots, n-1 \\ \rho(a)(b_n) &= b_1 \end{aligned}$$

also

$$R(a) = \begin{pmatrix} 0 & & \ldots & 0 & 1 \\ 1 & 0 & \ldots & & \\ & & \ldots & & \\ & & & & \\ 0 & & \ldots & 1 & 0 \end{pmatrix}.$$

6. Sei wieder $G = C_n$ und $A \in GL(V)$ mit $A^n = E$, wir setzen $\rho(a^i) = A^i$, dies ist eine Darstellung. Wenn speziell $z \in \mathbb{C}$ eine n-te Einheitswurzel ist, so erhalten wir mit $\rho(a^i) = z^i$ eine Darstellung vom Grad 1. Wenn $z_i, \ldots, z_m$ n-te Einheitswurzeln sind, so ist durch

$$T(a) = \begin{pmatrix} z_1, & 0 & \ldots & 0 \\ & & \ldots & \\ 0 & & \ldots & z_m \end{pmatrix}$$

eine Darstellung vom Grad m gegeeben.
7. Wenn $m = n$ und die z_i paarweise verschieden sind, so sind die Darstellungen unter 5. und 6. äquivalent: Sei

$$S = \begin{pmatrix} z_1 & \ldots & z_n \\ z_1^2 & \ldots & z_n^2 \\ & \ldots & \\ 1 & \ldots & 1 \end{pmatrix}$$

die Vandermondsche Determinante, dann gilt

$$R(A)^T S = S\, T(A).$$

8. Sei $G \subset S_n$ eine Untergruppe, $p \in G$ eine Permutation von $\{1, \ldots, n\}$ und $\{x_1, \ldots, x_n\}$ eine Basis von V. Wir definieren $\rho(p)(x_i) = x_{p(i)}$, dies ist eine Darstellung von G.

Zum Beispiel kann die Kleinsche Vierergruppe als Gruppe von Permutationen dargestellt werden:

$$K_4 = \{1 = (1),\, g = (1\,2)(3\,4),\, h = (\,3)(2\,4),\, gh = (1\,4)(2\,3)\}.$$

Dann ist

$$R(1),\ R(g) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix},\ R(h) = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix},\ R(gh) = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

eine zu K_4 isomorphe Matrixgruppe.

Als nÄchstes wollen wir uns einen Überblick über alle 1-dimensionalen komplexen Darstellung einer Gruppe verschaffen. Dies sind also Homomorphismen $\rho : G \longrightarrow \mathbb{C}^*$ in die multiplikative Gruppe des Körpers $\mathbb{C}$.
Ein Gruppenelement der Form $g^{-1}h^{-1}gh$ wird als Kommutator bezeichnet. In einer kommutativen Gruppe sind alle Kommutatoren gleich 1. Die von allen Kommutatoren erzeugte sogenannte Kommutatorgruppe

$$G' = \langle g^{-1}h^{-1}gh \mid g, h \in G \rangle$$

stellt also ein Maß für die Abweichung von der Kommutativität dar.
Da $\mathbb{C}^*$ kommuativ ist, gilt $G' \subset \mathrm{Ker}(\rho)$, also wird ein Homomorphismus

$$\bar{\rho} : G/G' \longrightarrow \mathbb{C}^*$$

induziert; ρ ist durch $\bar{\rho}$ eindeutig bestimmt: $\rho(g) = \bar{\rho}(gG')$.
Da die Gruppe G/G' abelsch ist, müssen nur die eindimensionalen Darstellungen abelscher Gruppen behandelt werden. Nach dem Hauptsatz über endliche abelsche Gruppen ist

$$G \cong \mathbb{Z}/p_1^{n_1}\mathbb{Z} \times \cdots \mathbb{Z}/p_m^{n_m}\mathbb{Z},$$

wir whlen Erzeugende $a_1, \ldots, a_m$ mit $a_i^{p_i^{n_i}} = 1$ und $p_i^{n_i}$-te Einheitswurzeln z_i, dann liefert $\rho(a_i) = z_i$ eine 1-dimensionale Darstellung.

Satz 17.18 *G sei eine abelsche Gruppe der Ordnung $p_1^{n_1} \cdots p_m^{n_m}$, dann gibt es genau $|G|$ verschiedene 1-dimensionale Darstellungen von G über $\mathbb{C}$.*

Beweis: Jede Darstellung ist durch die Bilder der Erzeugenden eindeutig bestimmt, es gibt $p_i^{n_i}$ verschiedene $p_i^{n_i}$-te Einheitswurzeln, insgesamt also $p_1^{n_1} \cdots p_m^{n_m} = |G|$ Wahlmglichkeiten. □

Folgerung 17.19 *Eine endliche Gruppe G besitzt genau $|G/G'|$ verschiedene 1-dimensionale Darstellung über $\mathbb{C}$.*

Seien nun $\rho_1,\ \rho_2$ Darstellungen von G in $V_1,\ V_2$, dann können wir die Darstellung

$$\rho_1 \oplus \rho_2 = \rho : G \longrightarrow \mathrm{GL}(V_1 \oplus V_2)$$

mit

$$\rho(g)(v_1 + v_2) = \rho_1(g)(v_1) + \rho_2(g)(v_2)$$

betrachten, sie heißt die direkte Summe der Darstellungen $\rho_1,\ \rho_2$.
Analog definiert man die direkte Summe von Matrixdarstellungen:

$$R_1 \oplus R_2(g) = \begin{pmatrix} R_1(g) & 0 \\ 0 & R_2(g) \end{pmatrix}.$$

Das Tensorprodukt

$$\rho_1 \otimes \rho_2 = \rho : G \longrightarrow \mathrm{GL}(V_1 \otimes V_2)$$

mit

$$\rho_1 \otimes \rho_2(g) = \rho_1(g) \otimes \rho_2(g)$$

ist ebenfalls eine Darstellung, die entsprechenden Darstellungsmatrizen sind die Kroneckerprodukte der gehabten Darstellungsmatrizen.

Wenn wieder $\rho : G \longrightarrow GL(V)$ eine Darstellung und KG die Gruppenalgebra ist, so wird V durch $(\sum a_i g_i) \cdot v := \sum a_i\, \rho(g_i)(v)$ ein linker KG-Modul. Umgekehrt: Wenn V ein KG-Modul ist, so definiert $\rho(g)(v) = g \cdot v$ eine Darstellung von G in V. Wenn wir KG als linken Modul über sich selbst auffassen, so entspricht dem die reguläre Darstellung von G.

Definition: Sei $\rho : G \longrightarrow GL(V)$ eine Darstellung. Ein Unterraum $U \subset V$ heißt ρ-invariant, wenn $\rho(g)(U) \subset U$ für alle $g \in G$ gilt. (In diesem Fall ist U ein KG-Untermodul von V.) Die Darstellung ρ heißt irreduzibel, wenn es außer $\{0\}$ und V keine invarianten Unterräume gibt, sonst reduzibel. (Der KG-Modul V ist im ersten Fall einfach, sonst nicht.)

Aus dem Satz von Maschke folgt:

Folgerung 17.20 *Jede endlichdimensionale Darstellung ist eine direkte Summe irreduzibler Darstellungen.* □

Der folgende Satz erlaubt schon einmal eine Zerlegung eine Darstellung in eine direkte Summe von Unterdarstellungen, denn idempotenten Endomorphismen entsprechen direkte Summanden:

Satz 17.21 *Für jede Darstellung $\rho : G \longrightarrow GL(V)$ ist*

$$p = \frac{1}{|G|} \sum_{g \in G} \rho(g) : V \longrightarrow V$$

ein idempotenter Endomorphismus.

Beweis: Für $h \in G$ ist

$$\rho(h) \circ p = \frac{1}{|G|} \sum_{g \in G} \rho(h) \circ \rho(g) = \frac{1}{|G|} \sum_{g \in G} \rho(hg) = p,$$

also

$$p \circ p = \frac{1}{|G|} \sum_{g \in G} \rho(h) \circ p = \frac{1}{|G|} \sum_{g \in G} p = \frac{|G|}{|G|} = p.$$

Die Verträglichkeit mit der KG-Operation zeigt man analog. □

Das Lemma von Schur liest sich im Darstellungszusammenhang folgendermaßen:

Satz 17.22 *1. Seien $\rho_i : G \longrightarrow GL(V_i)$, $1 = 1, 2$ irreduzible Darstellungen und $f : V_1 \longrightarrow V_2$ eine lineare Abbildung mit $f \circ \rho_1(g) = \rho_2(g) \circ f$ für alle $g \in G$, dann ist $f = 0$ oder f ist ein Isomorphismus, d.h. ρ_1 und ρ_2 sind äquivalent.*
2. Wenn $V_1 = V_2$ und zusätzlich $K = \mathbb{C}$ ist, so ist $f = z \cdot id$. □

Mit demselben Trick wie beim Satz von Maschke oder im obigen Satz erhält man Vertauschungsrelationen, für die man das Schursche Lemma anwenden kann:
Sei $h : V_1 \longrightarrow V_2$ beliebig und $a \in G$, wir setzen

$$h_a = \sum_g \rho_2(g) h \rho_1(ag^{-1}), \qquad (*)$$

dies ist eine lineare Abbildung und es gilt

$$\rho_2(u) h \rho_1(u^{-1}) = \sum_g \rho_2(ug) h \rho_1(a(ug)^{-1}) = h_a,$$

also

$$\rho_1(u) h_a = h_a \rho_1(u).$$

Folgerung 17.23 *Wenn ρ_1, ρ_2 irreduzibel und nicht äquivalent sind, so ist $h_a = 0$. Wenn ρ_1, ρ_2 äquivalent sind, so gibt es ein $z_a \in \mathbb{C}$ mit $h_a = z_a id$. Dabei gilt $z_e = \frac{|G|}{\dim V} Sp(h)$.*

Beweis: Wir bilden in (*) die Spur. □

17.3 Charaktere

Definition: Sei $\rho : G \longrightarrow GL(V)$ eine Darstellung, wir konstruieren dazu die Funktion $\chi : G \longrightarrow K$ mit $\chi(g) = Sp(\rho(g))$, sie heißt der zu ρ gehörige Charakter der Gruppe G. Wenn ρ eine irreduzible Darstellung ist, so nennt man χ einen irreduziblen Charakter.

Beispiel: Sei ρ die reguläre Darstellung, der entspricht als Modul die Gruppenalgebra. Wie operiert $\rho(g_i)$ auf $\{g_1 = 1, g_2, \ldots, g_n\}$?

$$\rho(g_i)(g_j) = \begin{cases} g_i g_j \neq g_j & \text{für} \quad i \neq 1 \\ g_i & \text{für} \quad i = 1 \end{cases},$$

d.h. $\rho(g_1)$ hat als Darstellungsmatrix die Einheitsmatrix, also $\chi(1) = Sp(\rho(1)) = n$; für $i \neq 1$ hat die Darstellungsmatrix von $\rho(g_i)$ nur Nullen auf der Diagonalen, also $\chi(g_i) = Sp(\rho(g_1)) = 0$.

Satz 17.24 *1. Die Charaktere zu äquivalenten Darstellungen sind gleich.*
2. $\chi(gh) = \chi(hg)$ *für* $g, h \in G$.
3. Ein Charakter ist konstant auf den Klassen konjugierter Elemente.
4. Seien χ_i *die zu* ρ_i *gehörigen Charaktere, dann gehört zur Darstellung* $\rho_1 \oplus \rho_2$ *der Charakter* $\chi_1 + \chi_2$.

Beweis: 1. Es gilt $\rho_1(g) = t^{-1}\rho_2(g)t$ für einen Isomorphismus t, also sind die Spuren von $\rho_1(g)$ und $\rho_2(g)$ gleich.
2. gilt wegen $Sp(AB) = SP(BA)$.
3. $\chi(h^{-1}gh) = Sp(\rho(h^{-1}gh) = Sp(\rho(h)^{-1}\rho(g)\rho(h)) = Sp(\rho(g)) = \chi(g)$.
4. Die zugehörigen Darstellungsmatrizen sind $\begin{pmatrix} R_1(g) & 0 \\ 0 & R_2(g) \end{pmatrix}$, deren Spur ist gleich $Sp(R_1(g)) + Sp(R_2(g))$. □

Satz 17.25 *Jeder Charakter von* G *ist eine Summe irreduzibler Charaktere.*

Beweis: Jede Darstellung ist eine direkte Summe irreduzibler Darstellungen. □

Definition: Seien $\phi, \psi : G \longrightarrow K$ beliebige Abbildungen; wir definieren ein Skalarprodukt

$$\langle \phi, \psi \rangle = \frac{1}{|G|} \sum_{g \in G} \phi(g)\psi(g^{-1}).$$

Satz 17.26 *Das Skalarprodukt* $\langle\ ,\ \rangle$ *ist symmetrisch und nicht ausgeartet.*

Beweis: $\langle \phi, \psi \rangle = \frac{1}{|G|} \sum_{g \in G} \phi(g)\psi(g^{-1}) = \frac{1}{|G|} \sum_{h^{-1} \in G} \phi(h^{-1})\psi(h) = \langle \psi, \phi \rangle$.
Sei $\langle \phi, \psi \rangle = 0$ für alle Abbildungen $\psi : G \longrightarrow K$. Für ein beliebiges $h \in G$ definieren wir

$$\psi_h(g) = \begin{cases} 1, & g = h^{-1} \\ 0 & \text{sonst} \end{cases},$$

dann gilt

$$0 = \langle \phi, \psi_h \rangle = \frac{1}{|G|}\phi(h),$$

also $\phi = 0$. □

Beispiel: Es sei χ_{reg} der Charakter zur regulären Darstellung von G, also

$$\chi_{reg}(g) = \begin{cases} |G|, & g = 1 \\ 0 & \text{sonst} \end{cases},$$

χ sei der Charakter zu $\rho : G \longrightarrow GL(V)$, dann ist

$$\langle \chi, \chi_{reg} \rangle = \frac{1}{|G|} \sum_{g \in G} \chi_{reg}(g)\chi(g^{-1}) = \frac{1}{|G|}\, |G|\, \chi(1) = \dim(V).$$

Da die Spur des Kroneckerprodukts $A \otimes B$ zweier Matrizen gleich $Sp(A)Sp(B)$ ist, gilt entsprechendes für den Charakter zum Tensorprodukt: Seien $\rho_i : G \longrightarrow GL(V_i)$ Darstellungen und χ_i die zugehörigen Charaktere, weiter sei χ der Charakter zu $\rho_1 \otimes \rho_2$, dann gilt

$$\chi(g) = \chi_1(g)\chi_2(g).$$

Wir wollen nun eine Darstellung von G im $\mathrm{Hom}(V_1, V_2)$ konstruieren:

$$\rho(g) : V_1^* \otimes V_2 \longrightarrow V_1^* \otimes V_2,$$

$$\rho(g) = \rho_1(g)^{-1} \otimes \rho_2(g),$$

führt zu

$$\rho(g) : \mathrm{Hom}(V_1, V_2) \longrightarrow \mathrm{Hom}(V_1, V_2),$$

$$\rho(g)(f) = \rho_2(g) \circ f \circ \rho_1(g).$$

Wegen $Sp(A^T) = Sp(A)$ gehört zu ρ der Charakter χ mit

$$\chi(g) = \chi_2(g)\chi_1(g^{-1}).$$

Zu dieser Darstellung bilden wir den idempotenten Endomorphismus

$$p = \frac{1}{|G|} \sum_g \rho(g) = \frac{1}{|G|} \sum \rho_2(g) \otimes \rho_1(g^{-1})^*,$$

es gilt

$$rg(p) = Sp(p) = \frac{1}{|G|} \sum \chi_2(g)\chi_1(g^{-1}) = \langle \chi_2, \chi_1 \rangle.$$

Nun können wir die nützlichen Orthogonalitätsrelationen für irreduzible Charaktere beweisen:

Satz 17.27 *Seien $\rho_i : G \longrightarrow GL(V)$ irreduzible Darstellungen über C und χ_1, χ_2 die zugehörigen Charaktere, dann gilt*

$$\langle \chi_1, \chi_2 \rangle = \begin{cases} 0, & \textit{wenn } \rho_1,\ \rho_2 \textit{äquivalent,} \\ 1, & \textit{wenn } \rho_1,\ \rho_2 \textit{nicht äquivalent.} \end{cases}$$

Beweis: Wir betrachten die obige Abbildung p, es gilt $\langle \chi_1, \chi_2 \rangle = rg(p) = \dim(\mathrm{Im}(p))$. Sei $h \in \mathrm{Hom}(v_1, V_2)$, dann ist (vergleiche die obige Konstrution (*))

$$p(h) = \frac{1}{G\,|G|} \sum_g \rho_2(g) h \rho_1(g^{-1}) = \frac{1}{|G|} h_e = 0,$$

falls ρ_1 und ρ_2 nicht äquivalent sind. Wenn aber ρ_1, ρ_2 äquivalent sind, so sei oBdA $\rho_1 = \rho_2,\ V_1 = V_2$, dann ist

$$p(h) = \frac{1}{|G|} h_e = z_h id$$

für alle h, d.h. $Im(p) = L(id)$ ist eindimensional, also $\langle \chi_1, \chi_1 \rangle = 1$. □

Satz 17.28 *Sei $\rho : G \longrightarrow GL(V)$ eine Darstellung mit dem Charakter χ, $\phi : G \longrightarrow GL(U)$ sei eine irreduzible Darstellung mit dem Charakter ψ, dann ist $\langle \chi, \psi \rangle$ gleich der Anzahl der irreduziblen Summanden von ρ, die äquivalent zu ϕ sind.*

Beweis: Wir zerlegen den KG-Modul V in einfache Untermoduln und fassen zueinander isomorphe zusammen:

$$V = \bigoplus n_i V_i,$$

die zugehörige Darstellung ist $\rho = \sum n_i \rho_i$ mit dem Charakter $\chi = \sum n_i \chi_i$. OBdA sei ψ äquivalent zu ρ_1, dann folgt aus den Orthogonalitätsrelationen

$$\langle \chi, \psi \rangle = \sum n_i \langle \chi_i, \psi \rangle = n_1. \square$$

Wesentlich einfacher als im Satz von Krull-Schmidt erhalten wir die

Folgerung 17.29 *Eine Zerlegung einer Darstellung in irreduzible Darstellungen ist bis auf die Reihenfolge der Summanden eindeutig bestimmt.*

Beweis: Die Vielfachheiten ergeben sich als Skalarprodukte. $\square$

Folgerung 17.30 *Seien $\chi_1, \ldots, \chi_k$ alle irreduziblen Charaktere von G und $n_1, \ldots, n_k$ die Dimensionen der zugehörigen Darstellungsräume V_i, dann gilt $\chi_{reg} = \sum n_i \chi_i$ und $|G| = \sum n_i^2$.*

Beweis: $n_i = \langle \chi_{reg}, \chi_i \rangle = \dim(V_i)$, $\chi_{reg}(1) = |G| = \sum n_i \chi_i(1) = \sum n_i^2$, da $\chi(1) = \dim(V)$. $\square$

Satz 17.31 *Seien $\rho, \rho' : G \longrightarrow GL(V_i)$ Darstellungen mit Charakteren χ, χ'. Die Darstellungen ρ, ρ' sind genau dann äquivalent, wenn $\chi = \chi'$.*

Beweis: Sei $\rho = \sum n_i \rho_i$, $\rho' = \sum l_i \rho_i$ mit irreduziblen ρ_i. Dann folgt aus $\chi = \chi'$ sofort $n_i = \langle \chi, \chi_i \rangle = \langle \chi', \chi_i \rangle = l_i$. $\square$

Satz 17.32 *Eine Darstellung $\rho, \rho' : G \longrightarrow GL(V_i)$ mit dem Charakter χ ist genau dann irreduzibel, wenn $\langle \chi, \chi \rangle = 1$ ist.*

Beweis: Wenn ρ irreduzibel ist, so folgt $\langle \chi, \chi \rangle$ aus den Orthogonalitätsrelationen. Sei umgekehrt $\rho = \sum m_i \rho_i$, dann gilt $\langle \chi, \chi \rangle = \sum m_i^2 = 1$ genau dann, wenn ein m_i gleich 1 und die restlichen gleich Null sind, also wenn ρ irreduzibel ist. $\square$

Definition: Eine Abbildung $f : G \longrightarrow K$ heißt Klassenfunktion, wenn $f(gh) = f(hg)$ für alle $g, h \in G$ gilt, d.h. f ist auf den Klassen konjugierter Elemente konstant.

Charaktere sind Klassenfunktionen.

Sei $f : G \longrightarrow K$ eine Klassenfunktion und $\rho : G \longrightarrow GL(V)$ eine irreduzible Darstellung, wir betrachten

$$T(f, \rho) = \sum_g f(g^{-1}) \rho(g) \in \mathrm{End}(V).$$

Es gilt

$$\begin{aligned} \rho(h) T(f, \rho) \rho(h^{-1}) &= \sum f(g^{-1}) \rho(hgh^{-1}) \\ &= \sum f(h^{-1} g^{-1} h) \rho(hgh^{-1}) \\ &= \sum f(u^{-1}) \rho(u) \\ &= T(f, \rho), \end{aligned}$$

also ist $T(f,\rho)$ mit allen $\rho(h)$ vertauschbar, aus dem Lemma von Schur folgt also

$$T(f,\rho) = z_{f,\rho} id,\ z_{f,\rho} \in \mathbb{C}.$$

Wir bilden die Spur:

$$Sp(T(f,\rho)) = \dim(V) z_{f,\rho}$$

oder

$$\frac{\dim(V)}{|G|} z_{f,\rho} = \frac{1}{|G|} Sp(T(f,\rho)) = \frac{1}{|G|} \sum f(g^{-1})\chi(g) = \langle f, \chi \rangle,$$

also

$$z_{f,\rho} = \frac{|G|}{\dim(V)} \langle f, \chi \rangle.$$

Satz 17.33 *Jede Klassenfunktion ist eine Linearkombination irreduzibler Charaktere.*

Beweis: Andernfalls gibt es eine Klassenfunktio f mit $\langle f, \chi_i \rangle$ für alle irreduziblen Charaktere χ_i, dann ist die obige Konstante z_{f,ρ_i} für alle i gleich Null. Also ist $T(f,\rho_i) = 0$. Sei $\rho_{reg} = \sum n_i \rho_i$, dann ist

$$T(f,\rho_{reg}) = \sum \imath n_i \sum_g f(g^{-1}\rho_i(g) = \sum n_i T(f,\rho_i) = 0.$$

Wir wenden dies auf 1 an und beachten $\rho_{reg}(g)(1) = g$:

$$\sum f(g^{-1})\rho_{reg}(1) = \sum f(g^{-1})g = 0,$$

also $f(g^{-1}) = 0$, d.h. $f = 0$. □

Satz 17.34 *Die irreduziblen Charaktere bilden eine Orthonormalbasis des Vektorraums der Klassenfunktionen.* □

Folgerung 17.35 *Die Anzahl der irreduziblen Charaktere ist gleich der Anzahl der Konjugationsklassen von G.*

Beweis: Die Dimension des Raums der Klassenfunktionen ist gleich der Zahl der Konjugationsklassen. □

Satz 17.36 *Sei G eine abelsche Gruppe, dann ist jede irreduzible Darstellung eindimensional, d.h. jeder irreduzible Charakter $\chi_i : G \longrightarrow \mathbb{C}^*$ ist ein Homomorphismus.*

Beweis: Sei $|G| = k$, alle Konjugationsklassen sind einelementig, also gibt es k Stück. Weiter ist $k = |G| = \sum_{i=1}^k n_i^2$, also $n_i = 1$ für $i = 1, \ldots, k$. □

Ohne Beweis teilen wir abschließend mit, daß die Dimensionen der irreduziblen Darstellungen Teiler der Gruppenordnung sind.

17.4 Die diskrete Fourier-Transformation

Die Gruppe G sei kommutativ, dann ist KG kommutativ. Wir betrachten den Spezialfall $K = \mathbb{C}$. Nach dem Satz von Maschke ist $\mathbb{C}G$ eine halbeinfache Algebra, wir haben gesehen, das halbeinfache Algebren sich als eine direkte Summe einfacher Algebren darstellen lassen:

$$\mathbb{C}G = A_1 \oplus \ldots \oplus A_m.$$

Nach dem Satz von Wedderburn ist jede einfache $\mathbb{C}$-Algebra isomorph zu einer Matrixalgebra:

$$A_i \cong M_{n_i n_i}(\mathbb{C}).$$

Für $n_i > 1$ ist diese Algebra nichtkommutativ, also müssen alle $n_i = 1$ sein, also $A_i \cong \mathbb{C}$. Insgesamt erhalten wir

$$\mathbb{C}G \cong \mathbb{C} \times \ldots \times \mathbb{C}.$$

Es sei $C_n = \{1, g, g^2, \ldots, g^{n-1}\}$ mit $g^n = 1$ die zyklische Gruppe mit n Elementen. Dann ist $\mathbb{C}C_n$ isomorph zur Faktoralgebra $\mathbb{C}[x]/(x^n - 1)$ des Polynomrings $\mathbb{C}[x]$. Die Multiplikation in $\mathbb{C}[x]/(x^n - 1)$ ist relativ komplex, wenn das Produkt zweier Polynome zu berechnen ist, so sind etwa n^2 Multiplikationen von Körperelementen durchzuführen. Nach den obigen Resultaten ist aber

$$\mathbb{C}C_n \cong \mathbb{C} \times \ldots \times \mathbb{C}$$

und die Multiplikation in dieser Algebra geschieht komponentenweise, für eine Multiplikation von Algebra-Elementen benötigt man also nur n Multiplikationen von Körperelementen. Es wäre schön, wenn wir den obigen Isomorphismus explizit kennen würden.

Dies ist möglich. Wir bestimmen nämlich die Idempotenten e_i mit $A_i = \mathbb{C}C_n e_i$.

Lemma 17.37 *Sei G eine kommutative Gruppe und $f : G \to \mathbb{C}\backslash\{0\}$ ein Homomorphismus von multiplikativen Gruppen, dann ist*

$$\frac{1}{\mid G \mid} \sum f(g_i) g_i \in \mathbb{C}G$$

idempotent.

Beweis:

$$\frac{1}{\mid G \mid} \sum f(g_i) g_i \cdot \frac{1}{\mid G \mid} \sum f(g_j) g_j = \frac{1}{\mid G \mid \mid G \mid} \sum f(g_i g_j) g_i g_j$$
$$= \frac{1}{\mid G \mid \mid G \mid} \sum f(g_i) g_i \cdot \mid G \mid = \frac{1}{\mid G \mid} \sum f(g_i) g_i. \square$$

Wenn speziell $G = C_n =< g >$ ist, so ist jeder Homomorphismus $f : C_n \to \mathbb{C}\backslash\{0\}$ durch $f(g) \in \mathbb{C}$ bestimmt, wegen $g^n = 1$ muß $f(g)^n = 1$ sein, d.h. $f(g)$ ist eine n-te Einheitwurzel. Sei also ω eine primitive n-te Einheitwurzel, dann gibt es die folgenden n Homomorphismen $f_i : C_n \to \mathbb{C}\backslash\{0\}$ mit

$$f_i(g) = \omega^i.$$

Also haben wir n Idempotente in der Gruppenalgebra:

$$e_i = \frac{1}{n}(1 + \omega^i g + \omega^{2i} g^2 + \ldots + \omega^{(n-1)i} g^{n-1}),\ i = 0, \ldots, n-1.$$

Also hat der Isomorphismus

$$F^{-1} : \bigoplus \mathbb{C}C_n e_i \to \mathbb{C}C_n$$

bezüglich der Basen $\{e_1, \ldots, e_n\}$ und $\{1, g, \ldots, g^{n-1}\}$ die Darstellungsmatrix

$$\frac{1}{n}\begin{bmatrix} 1 & 1 & \ldots & 1 \\ 1 & \omega & \ldots & \omega^{n-1} \\ & & \ldots & \\ 1 & \omega^{n-1} & \ldots & \omega^{(n-1)^2} \end{bmatrix}$$

der (eigentlich interessante) inverse Isomorphismus F hat die zu dieser inverse Darstellungsmatrix, diese hat die Gestalt

$$\begin{bmatrix} 1 & 1 & \ldots & 1 \\ 1 & \xi & \ldots & \xi^{n-1} \\ & & \ldots & \\ 1 & \xi^{n-1} & \ldots & \xi^{(n-1)^2} \end{bmatrix}$$

mit $\xi = \omega^{-1}$. æ

Kapitel 18

Zerlegung endlichdimensionaler Algebren

Von nun an sei A eine beliebige endlichdimensionale K-Algebra.

Definition: Ein Element $a \in A$ heißt nilpotent, wenn ein $n \in \mathbb{N}$ mit $a^n = 0$ existiert. Ein Linksideal $L \subseteq A$ heißt nilpotent, wenn $L^n = \{0\}$ für ein $n \in \mathbb{N}$, d.h. wenn alle Produkte $l_1 \dots l_n$ von Elementen aus L Null sind.

Beispiele: In $K[x]/(x^n)$ ist das von $\bar{x}$ erzeugte Ideal nilpotent.
In der Algebra $\begin{pmatrix} * & * \\ 0 & * \end{pmatrix}$ der Dreiecksmatrizen ist das Linksideal $\begin{pmatrix} 0 & * \\ 0 & 0 \end{pmatrix}$ nilpotent.

Lemma 18.1 *In einer halbeinfachen Algebra A gibt es keine von $\{0\}$ verschiedenen nilpotenten Linksideale.*

Beweis: Ein Linksideal $L \subseteq A$ ist ein direkter Summand, wird also von einem Idempotent e erzeugt, von dem keine Potenz verschwindet. □

Deshalb traten bisher keine nilpotenten Linksideale auf.

Lemma 18.2 *Die Summe zweier nilpotenter Linksideale ist nilpotent.*

Beweis: Seinen $N, L \subseteq A$ Linksideale und $N^p = L^q = \{0\}$. Wir betrachten ein Produkt von $p + q + 1$ Faktoren aus $N + L$:

$$(n_1 + l_1) \dots (n_{p+q+1} + l_{p+q+1}),$$

wenn dies ausmultipliziert wird, so enthalte ein Summand r Faktoren aus N und $p + q + 1 - r$ Faktoren aus L. Wenn $r \geq p + 1$ ist, so liegt er in $N^p(N + L) = \{0\}$, wenn $r < p + 1$ ist, so ist $p + q + 1 - r > q$, also liegt der Summand in $L^q(N + L) = \{0\}$. □

In einer endlichdimensionalen Algebra ist die Summe unendlich vieler Linksideale stets gleich der Summe von nur endlich vielen dieser Linksideale, denn sonst könnte man eine unendliche echt aufsteigende Kette von Linksidealen konstruieren. Somit erhalten wir den

Satz 18.3 *Die Summe aller nilpotenter Linksideale von A ist nilpotent.* □

Satz 18.4 *Sei $a \in A$, dann sind die folgenden Bedingungen äquivalent:*
1. Es ist $aM = \{0\}$ für alle einfachen linken A-Moduln M.
2. Das Element a liegt im Durchschnitt aller maximalen Linksideale von A.
3. Für alle $b \in A$ besitzt $1 - ba$ ein Inverses.
4. Für alle $b \in A$ ist ba nilpotent.
5. Das Element a liegt in einem nilpotenten Linksideal.
6. Es ist $Ma = \{0\}$ für alle einfachen rechten A-Moduln M.
7. Das Element a liegt im Durchschnitt aller maximalen Rechtsideale von A.
8. Für alle $b \in A$ besitzt $1 - ab$ ein Inverses.
9. Für alle $b \in A$ ist ab nilpotent.
10. Das Element a liegt in einem nilpotenten Rechtsideal.

Beweis: Wir zeigen $1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 4 \Rightarrow 5 \Rightarrow 1$ und $4 \Leftrightarrow 9$, daraus folgt der Rest.
$1 \Rightarrow 2$: Sei L ein maximales Linksideal und $a \notin L$, dann ist der A-Modul A/L einfach und $a \cdot (1 + L) = a + L \neq 0 + L$, also ist $a \cdot (A/L) \neq \{0\}$, ein Widerspruch.
$2 \Rightarrow 3$: Wir zeigen zuerst, daß $1 - ba$ ein Linksinverses besitzt. Das Element ba liegt in jedem maximalen Linksideal, also liegt $1 - ba$ in keinem maximalen Linksideal, denn aus $ba \in L$ und $1 - ba \in L$ folgt $1 \in L$, also $L = A$. Folglich ist $A(1 - ba) = A$, also gibt es ein $c \in A$ mit $c(1 - ba) = 1$. Aus dem soeben bewiesenen folgt, daß auch $1 + cba = c$ ein Linksinverses d besitzt: $dc = 1$, also hat c ein Links- und ein Rechtsinverses, somit stimmen beide überein: $d = 1 - ba$.
$3 \Rightarrow 4$: Sei $L = Aa$, es ist

$$L \supset L^2 \supset L^3 \supset \ldots \supset L^n = L^{n+1}$$

und wir nehmen an, das $L^n \neq \{0\}$ wäre. Sei dann $N \subseteq L$ ein Linksideal, das minimal mit der Eigenschaft $L^n \cdot N \neq \{0\}$ ist, also gibt es ein $x \in N$ mit $L^n x \neq \{0\}$. Wir betrachten das Linksideal $L^n x$: es ist

$$L^n L^n x = L^n x \neq \{0\},$$

also $L^n x = N$. Also gibt es ein $y \in L^n$ mit $yx = x$ und es gilt $y \in L = Aa$, also $y = ba$ für ein $b \in A$. Folglich besitzt $1 - y$ ein Inverses c. Nun folgt $x = 1 \cdot x = c(1 - y)x = c(x - yx) = 0$, ein Widerspruch, also ist das Linksideal Aa nilpotent.
$4 \Rightarrow 5$: Nach Vorausesetzung besteht $L = Aa$ aus nilpotenten Elementen, sei wieder $L^n = L^{n+1} \neq \{0\}$, wie oben erhalten wir Elemente $0 \neq x, y \in L$ mit $yx = x$, daraus folgt $y^n x = x$ für alle n, aber für großes n ist $y^n = 0$, ein Widerspruch.
$5 \Rightarrow 1$: Sei M ein einfacher A-Modul, dann gilt entweder $AaM = \{0\}$ oder $AaM = M$, denn AaM ist ein Untermodul von M. Im ersten Fall folgt $aM = \{0\}$, im zweiten $(Aa)^n M = M$ für alle n, wegen der Nilpotenz von Aa ist dies ein Widerspruch.
$4 \Leftrightarrow 9$: Wenn $(ba)^n = 0$ ist, so ist ebenfalls $(ab)^{n+1} = a(ba)^n b = 0$. □

Satz 18.5 *Die Menge J aller Elemente $a \in A$, die den Bedingungen des vorigen Satzes genügen, ist ein Ideal von A.*

Beweis: Nach 2. ist J der Durchschnitt aller maximaler Linksideale, also selbst ein Linksideal, nach 7. ist J auch ein Rechtsideal. □

Dieses Ideal J wird als Jacobson-Radikal bezeichnet.

Folgerung 18.6 *J ist das eindeutig bestimmte maximale nilpotente (Links-, Rechts-) Ideal von A. Wenn A halbeinfach ist, so ist $J = \{0\}$.* □

Satz 18.7 *Das Radikal von A/J ist Null.*

Beweis: Sei $N/J \subseteq A/J$ ein nilpotentes Linksideal, dabei ist $J \subseteq N \subseteq A$. Aus $(N/J)^n = J/J = \{\bar{0}\}$ folgt $N^n \subseteq J$ und aus $J^m = \{0\}$ folgt $N^{nm} = \{0\}$, also ist N nilpotent, also $N \subseteq J$, d.h. N/J ist Null. □

Die Umkehrung des folgenden Satzes haben wir oben gesehen.

Satz 18.8 *Wenn $J = \{0\}$ ist, so ist A halbeinfach.*

Beweis: Wir zeigen zuerst, daß jedes minimale Linksideal $L \subseteq A$ ein idempotentes Element enthält.
Es ist $L^2 \subseteq L$, also $L^2 = \{0\}$ oder $L^2 = L$, wobei der erste Fall nicht eintreten kann, weil es wegen $J = \{0\}$ keine nilpotenten Linksideale gibt. Also gibt es ein $a \in L$ mit $La \neq \{0\}$, also $La = L$. Wir betrachten $N = \{b \in L \mid ba = 0\}$, dies ist ein Linksideal, das in L enthalten und von L verschieden ist, also muß $N = \{0\}$ sein, d.h. aus $ba = 0$ folgt $b = 0$. Wegen $La = L$ gibt es ein $e \in L$ mit $ea = a$, dann ist auch $e^2a = a$, d.h. $(e^2 - e)a = 0$, also $e^2 - e = 0$, also ist e idempotent.
Nun zerlegen wir A schrittweise in eine direkte Summe minimaler Linksideale.
Das Linksideal Ae ist nicht Null und in L enthalten, also $L = Ae$. Sei nun $a \in A$ beliebig, dann ist $a = ae + (a - ae)$, dabei ist der erste Summand ein Element von L und die Elemente der Form $a - ae$ bilden ein Linksideal N, folglich ist $A = L + N$.
Sei $b \in L \cap N$, dann ist einerseits $b = b_1e \in L = Ae$, also $be = b_1e^2 = b_1e = b$, andererseits ist $b = b_2 - b_2e \in N$, also $be = b_2e - b_2e^2 = b_2e - b_2e = 0$, also ist $L \cap N = \{0\}$, d.h. $A = L \oplus N$. Nun wählen wir ein minimales Linksideal, das in N enthalten ist und spalten es als direkten Summanden ab. Nach endlich vielen Schritten sind wir fertig. □

Folgerung 18.9 *A/J ist halbeinfach.* □

Satz 18.10 *Ein A-Modul M ist genau dann halbeinfach, wenn $J \cdot M = \{0\}$ ist.*

Beweis: Sei $M = \oplus M_i$ eine direkte Summe einfacher Moduln, dann gilt $J \cdot M_i = \{0\}$ nach Definition des Radikals.
Sei umgekehrt $J \cdot M = \{0\}$, dann wird M wie folgt ein A/J-Modul:

$$(a + J) \cdot m = am,$$

dies ist wegen $Jm = \{0\}$ wohldefiniert. Also ist M als A/J-Modul halbeinfach, es gibt einfache A/J-Moduln M_i mit $M = \oplus M_i$. Jeder A/J-Modul ist aber auch ein A-Modul, also ist M halbeinfach als A-Modul. □

Da wir über A keine weiteren Voraussetzungen machen, können wir nicht erwarten, daß sich jeder Modul in eine direkte Summe einfacher Untermoduln zerlegen läßt. Aber Zerlegungen wird es schon geben.

Definition: Sei M ein A-Modul. Wenn nichttriviale Untermoduln $U, V \subseteq M$ existieren, so daß $M = U \oplus V$ gilt, so heißt M zerlegbar, andernsfalls heißt M unzerlegbar.

Satz 18.11 *Jeder (endlichdimensionale) A-Modul M ist eine direkte Summe unzerlegbarer Untermoduln.*

Beweis: Entweder ist M unzerlegbar oder eine direkte Summe von Untermoduln. Diese Summanden sind entweder unzerlegbar oder lassen sich in Summen zerlegen. Und so weiter. □

Definition: Ein idempotentes Element $e \in A$ heißt primitiv, wenn es keine orthogonalen Idempotenten s, t mit $e = s + t$ gibt.

Da Idempotente zu direkten Summen führen, gilt der

Satz 18.12 *Sei $M = \oplus M_i$ und $e_i : M \to M$ sei die Projektion auf M_i. Die Moduln M_i sind genau dann unzerlegbar, wenn die e_i primitive orthogonale Idempotente in $End_A(M)$ sind und $\sum e_i = id_M$ ist.* □

Folgerung 18.13 *Seien $e_1, \ldots, e_k \in A$ idempotente Elemente, dann ist äquivalent:*
1. $A = Ae_1 \oplus \ldots \oplus Ae_k$ und die Ae_i sind unzerlegbare Linksideale.
2. $A = e_1A \oplus \ldots \oplus e_kA$ und die e_iA sind unzerlegbare Rechtsideale.
3. $\sum e_i = 1$ und die e_i sind primitive orthogonale Idempotente. □

Beispiel: $A = \begin{pmatrix} * & * \\ 0 & * \end{pmatrix}$, primitive Idempotente sind $e_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ und $e_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, sie erzeugen die Linksideale $Ae_1 = \begin{pmatrix} * & 0 \\ 0 & 0 \end{pmatrix}$ und $Ae_2 = \begin{pmatrix} 0 & * \\ 0 & * \end{pmatrix}$, letzeres ist nicht minimal, denn es enthält das Linksideal $\begin{pmatrix} 0 & * \\ 0 & 0 \end{pmatrix}$, dieses wiederum ist kein direkter Summand von A.

Wir werden nun einen Zusammenhang zwischen den Zerlegungen von A/J in eine direkte Summe minimaler Linksideale und von A in eine Summe unzerlegbarer Linksideale herstellen.

Satz 18.14 (Liften von Idempotenten) *Seien $f_i + J \in A/J$ orthogonale Idempotente, dann gibt es orthogonale Idempotente $e_i \in A$ mit $e_i + J = f_i + J$.*

Beweis: Wir setzen $f_1 = a$, dann gilt $a^2 - a \in J$. Wir machen einen Ansatz

$$e = a + x(1 - 2a),$$

wobei $x \in J$ sein wird, also $a + J = e + J$. Dann ist

$$e^2 = a^2 + x^2(1 + 4a^2 - 4a) + 2ax(1 - 2a),$$

$$\begin{aligned} e^2 - e &= a^2 + x^2 + 4a^2x^2 - 4ax^2 + 2ax - 4a^2x - a - x + 2ax \\ &= (x^2 - x)(1 + 4(a^2 - a)) + a^2 - a \end{aligned}$$

und dies ist jedenfalls dann gleich Null, wenn

$$x = \frac{1}{2}(1 - \frac{1}{\sqrt{1 + 4n}}) = \frac{1}{2}(2n - \tbinom{4}{2}n^2 + \tbinom{6}{3}n^3 - \ldots),$$

dabei haben wir zur Abkürzung $n = a^2 - a$ gesetzt. Nebenrechnung: $\sqrt{\frac{1}{4} - \frac{n}{1+4n}} = \frac{1}{2\sqrt{1+4m}}$. Wegen $n = a^2 - a \in J$ ist n nilpotent und die Potenzreihe bricht ab. Also haben wir ein Idempotent e_1 in der Klasse von f_1 gefunden.
Analog sei e_2 eine Liftung von f_2. Diese Idempotente sind eventuell nicht orthogonal. Aber es ist

$$e_1 e_2 \equiv f_1 f_2 \equiv 0 \pmod{J},$$

also $e_1 e_2 \in J$, folglich besitzt $1 - e_1 e_2$ ein Inverses (nämlich $1 + e_1 e_2 + (e_1 e_2)^2 + \ldots$), wir betrachten nun

$$e_2^* = (1 - e_1 e_2) e_2 (1 - e_1 e_2)^{-1},$$

dann ist $e_2^{*2} = e_2^*$ und $e_2^* \equiv f_2 \pmod{J}$, denn $e_2^*(1 - e_1 e_2) = e_2 - e_1 e_2 = e_s^* - e_2^* e_1 e_2$, also ist $e_2^* - e_2 = e_2^* e_1 e_2 - e_1 e_2 \in J$. Weiterhin gilt

$$e_1 e_2^* = (e_1 e_2 - e_1 e_2)(1 - e_1 e_2)^{-1} = 0.$$

Wir setzen nun $e_2^\# = e_2^*(1 - e_1)$, dann ist $e_1 e_2^\# = e_2^\# e_1 = 0$ und

$$e_2^{\#2} = e_2^*(1 - e_1) e_2^*(1 - e_1) = e_2^*(e_2^* - e_1 e_2^*)(1 - e_1) = e_2^*(1 - e_1) = e_2^\#.$$

Seien nun $e_1, e_2, e_3 \in A$ idempotente Elemente und $e_1 e_2 = e_2 e_1 = 0$ und $e_1 e_3 = e_3 e_1 = 0$. Nun wird e_3 zu $e_3^\#$ geändert, so daß $e_2 e_3^\# = e_3^\# e_2 = 0$ gilt. Wir rechnen nach, ob dann noch $e_1 e_3^\# = e_3^\# e_1 = 0$ gilt: Es ist

$$e_3^\# = (1 - e_2 e_3) e_3 (1 + e_2 e_3 + (e_2 e_3)^2 + \ldots + (e_2 e_3)^k)(1 - e_2),$$

also gilt

$$e_1 e_3^\# = e_3^\# e_1 = 0.$$

So kann jede Menge orthogonaler Idempotente von A/J zu A geliftet werden. □

Folgerung 18.15 *Jede direkte Zerlegung von A/J in unzerlegbare (d.h. minimale) Linksideale läßt sich zu einer Zerlegung von A in unzerlegbare Linksideale liften.*

Beweis: Sei $A/J = \bigoplus (A/J)(f_i + J)$, wobei die $f_i + J$ primitive othogonale Idempotente mit $\sum f_i + J = 1 + J$ sind. Seien e_i orthogonale Liftungen dieser Idempotenten, wir zeigen, daß die e_i primitiv sind und daß $\sum e_i = 1$ gilt.
Andernfalls wäre $e_i = p_i + q_i$ mit $p_i q_i = q_i p_i = 0$, dann ist aber auch $e_i + J = p_i + J + q_i + J$ und $(p_i + J) \cdot (q_i + J) = J$ im Widerspruch zur Primitivität der $f_i + J$.
Falls $\sum e_i \neq 1$ wäre, so ist $1 - \sum e_i$ ein weiteres Idempotent, aber $(1 - \sum e_i) + J = (1 + J) - (1 + J) = J$, also $1 - \sum e_i \in J$, ein Widerspruch. □

Wir wollen den Zusammenhang zwischen den minimalen Linksidealen von A/J und deren Liftungen noch genauer untersuchen. Dazu sind einige Hilfsmittel nötig.

Satz 18.16 *Jedes nichtnilpotente Linksideal $L \subseteq A$ enthält ein Idempotent.*

Beweis: Sei $N \subseteq L$ ein Linksideal, das in der Menge der in L enthaltenen nichtnilpotenten Linksidealen minimal ist. Dann ist $N^2 = N$, also gibt es ein $a \in N$ mit $Na = N$, denn falls $Na \neq N$ für alle $a \in N$ gälte, so wäre N nilpotent). Folglich ist

$$Q = \{n \in N \mid na = 0\} \subset N$$

ein echtes Unterideal, also nilpotent. Nun gibt es ein $c \in N$ mit $ca = a$, damit $c^2a = a$, also $(c^2 - c)a = 0$, d.h. $c^2 - c \in Q$ und wie vorhin können wir ein idempotentes Element $e = c + x(1 - 2c)$ finden. □

Satz 18.17 *Ein unzerlegbares Linksideal $L \subseteq A$ ist genau dann ein direkter Summand, wenn L nicht nilpotent ist.*

Beweis: Wenn L ein direkter Summand ist, so ist es von der Form $L = Ae$ mit idempotentem e, also ist L nicht nilpotent.
Sei umgekehrt L nicht nilpotent, dann enthält es ein Idempotent e. Nun ist Le das Bild der Multiplikation von L mit e, deren Kern ist $Q = \{l \in L \mid le = 0\}$ und aus der Idempotenz von e folgt

$$L = Le \oplus Q,$$

wegen der Unzerlegbarkeit von L muß also $Q = 0$ und $L = Le$ sein. Schließlich ist $Le \subseteq Ae \subseteq L$, also ist $L = Ae$ ein direkter Summand.

Satz 18.18 *Ein nichtnilpotentes Linksideal L ist genau dann unzerlegbar, wenn alle echt in L enthaltenen Linksideale nilpotent sind.*

Beweis: Seien Alle $N \subset L$ nilpotent, dann kann L nicht die Summe von Untermoduln sei, es wäre sonst selbst nilpotent.
Sei umgekehrt L unzerlegbar und $N \subseteq L$ minimal unter den nichtnilpotenten Linksidealen. N kann nicht Summe von Untermoduln sei (sonst wäre es nilpotent), also ist N unzerlegbar, folglichein direkter Summand von A und damit auch ein direkter Summand von L. Also muß $N = L$ gelten, also jeder Untermodul von L ist nilpotent.□

Satz 18.19 *Sei $L = Ae, e^2 = e$, ein unzerlegbares Linksideal, dann ist $Je \subseteq Ae$ das eindeutig bestimmte maximale Unterlinksideal, somit ist Ae/Je ein einfacher A-Modul.*

Beweis: Jeder Linksmodul $N \subset L$ ist nilpotent, also $N \subseteq J \cap L = Je$, also ist Je maximal in L. □

Wir können nun die einfachen Moduln einer Algebra genauer beschreiben:

Satz 18.20 *Sei $A = \oplus Ae_i$ mit orthogonalen primitiven Idempotenten e_i. Sei M ein einfacher A-Modul, dann gibt es ein i mit $M \cong Ae_i/Je_i$.*

1. Beweis: Es iat $AM \neq \{0\}$, also $Ae_iM \neq \{0\}$ für ein i, also $Ae_im = M$ für ein $m \in M$. Damit haben wir einen surjektiven Modulhomomorphismus $f : Ae_i \to M$ mit $f(ae_i) = ae_im$, dessen Kern ist maximal in Ae_i, also gleich Je_i.
2. Beweis: M ist einfach, also $JM = 0$, d.h. M ist ein A/J-Modul und als einfacher Modul isomorph zu einem minimalen Ideal der halbeinfachen Algebra A/J, etwa zu Ae_i/Je_i. □

Lemma 18.21 (Fitting) *Sei M ein unzerlegbarer A-Modul und $f : M \to M$ ein A-Endomorphismus, dann ist f ein Isomorphismus oder nilpotent.*

Beweis: Wir haben eine absteigende Folge

$$M \supset f(M) \supset f^2(M) \supset \ldots \supset f^n(M) = f^{n+1}(M),$$

die sich stabilisiert, also ist die Einschränkung

$$f^n \mid f^n(M) : f^n(M) \to f^n(M)$$

surjektiv und folglich auch injektiv (alle Vektorräume sind endlichdimensional). Also gilt $f^n(M) \cap \operatorname{Ker}(f^n) = \{0\}$. Sei nun $m \in M$, dann gilt $f^n(m) = f^{2n}(b)$ für ein $b \in M$ und

$$m = f^n(b) + (m - f^n(b)),$$

der erste Summand liegt in $f^n(M)$, der zweite in $\operatorname{Ker}(f^n)$, also ist $M = f^n(M) \oplus \operatorname{Ker}(f^n)$ eine direkte Summe, daraus folgt $M = f^n(M)$ oder $f^n = 0$. □

Folgerung 18.22 *Wenn M ein unzerlegbarer A-Modul ist, so ist dessen Radikal $J(\operatorname{End}_A(M))$ das eindeutig bestimmte maximale Ideal (d.h. $End_A(M)$) ist ein lokaler Ring).* □

Satz 18.23 *Seien $e_1, e_2 \in A$ primitive Idempotente, dann gilt $Ae_1/Je_1 \cong Ae_2/Je_2$ genau dann, wenn $Ae_1 \cong Ae_2$.*

Beweis: Sei $f : Ae_1/Je_1 \to Ae_2/Je_2$ ein Isomorphismus und $f(e_1 + J) = r + J$ mit $r \in Ae_2$. Dann ist $r + J = f(e_1 + J) = f(e_1^2 + J) = e_1 f(e_1 + J) = e_1 r + J$. Es gilt $Ae_1 r \subseteq Ae_2$ und $e_1 r$ ist nicht nilpotent, also ist $Ae_2 = Ae_1 r$, d.h. $g : Ae_1 \to Ae_2$ mit $g(x) = xr$ ist ein surjektiver Homomorphismus linker A-Moduln. Analog erhalten wir einen Homomorphismus $h : Ae_2 \to Ae_1$ mit $h(y) = ys$, also ist $g \circ h : Ae_1 \to Ae_1$ ein surjektiver Endomorhismus, also nicht nilpotent. Nach dem Lemma von Fitting ist $g \circ h$ ein Isomorphismus, also ist h auch ein Isomorphismus.
Wenn umgekehrt $h : Ae_1 \to Ae_2$ ein Isomorphismus ist, so betrachten wir die Komposition

$$Ae_1 \to Ae_2 \to Ae_2/Je_2,$$

deren Kern ist gleich Je_1. Nach dem Homomorphiesatz folgt $Ae_1/Je_1 \cong Ae_2/Je_2$. □

Wir hatten gesehen, daß sich jeder Modul in eine direkte Summe unzerlegbarer Untermoduln zerlegen läßt. Wir stellen uns die Frage, ob dies auf mehrere verschiedene Weisen möglich sein kann.

Lemma 18.24 *Sei M ein unzerlegbarar A-Modul und für $i = 1, \ldots, n$ seien $f_i : M \to M$ A-lineare Abbildungen, so daß $\sum f_i$ ein Isomorphismus ist. Dann ist zumindest eines der f_i ein Isomorphismus.*

Beweis: Andernfalls sind alle f_i nilpotent, liegen also im Radikal von $\operatorname{End}_A(M)$, dort liegt auch deren Summe, kann also kein Isomorphismus sein. □

Lemma 18.25 *Sei $f : M_1 \oplus M_2 \to N_1 \oplus N_2$ ein Isomorphismus, es sei $f(m_1, 0) = (h(m_1), g(m_1))$ und $h : M_1 \to N_1$ sei ein Isomorphismus, dann ist $M_2 \cong N_2$.*

Beweis: Sei $f(m_1, m_2) = (n_1, n_2)$, wir setzen

$$p(m_1, m_2) = (n_1, n_2 - gh^{-1}(n_1)).$$

Die Abbildung p ist injektiv, denn aus $(n_1, n_2 - gh^{-1}(n_1)) = (0, 0)$ folgt $n_1 = 0 = n_2$ und aus der Injektivität von f folgt $m_1 = m_2 = 0$. Wegen $\dim(M_1 \oplus M_2) = \dim(N_1 \oplus N_2)$ ist p auch ein Isomorphismus. Weiter gilt

$$p(m_1, 0) = (h(m_1), g(m_1) - gh^{-1}(h(m_1))) = (h(m_1), 0),$$

also $M_2 \cong M_1 \oplus M_2/M_1 \oplus 0 \xrightarrow{p} (M_1 \oplus M_2)/h(M_1) \oplus 0 = N_1 \oplus N_2/N_1 \oplus 0 \cong N_2$. □

Nun können wir zeigen, daß es keine wesentlich verschiedenen Zerlegungsmöglichkeiten eines Moduls gibt:

Satz 18.26 (Krull/Schmidt/Remak/Wedderburn) *Wenn $M_1 \oplus \ldots \oplus M_m \cong N_1 \oplus \ldots \oplus N_n$ und die M_i und N_j unzerlegbare A-Moduln sind, so ist $m = n$ und bei geeigneter Numerierung $M_i \cong N_i$.*

Beweis: Sei $f : M_1 \oplus \ldots \oplus M_m \to N_1 \oplus \ldots \oplus N_n$ ein Isomorphismus, wir verknüpfen ihn und sein Inverses mit Einbettungen und Projektionen:

$$g_k : M_k \xrightarrow{i_k} M_1 \oplus \ldots \oplus M_m \xrightarrow{f} N_1 \oplus \ldots \oplus N_n \xrightarrow{p_1} N_1,$$

$$h_k : N_1 \xrightarrow{j_1} N_1 \oplus \ldots \oplus N_n \xrightarrow{f^{-1}} M_1 \oplus \ldots \oplus M_n \xrightarrow{q_k} M_k,$$

dann ist

$$\sum g_k h_k = \sum p_1 f i_k q_k f^{-1} j_1 = p_1 f \sum i_k q_k f^{-1} j_1 = p_1 j_1 = id_{N_1},$$

also ist einer der Summanden, etwa $g_1 h_1$, ein Isomorphismus. Also ist g_1 surjektiv, weiter ist $h_1 g_1$ nicht nilpotent, also ein Isomorphismus und damit ist g_1 injektiv, also $M_1 \cong N_1$.
Nun ist $f(m_1, 0, \ldots, 0) = (g_1(m_1), \ldots)$, also folgt aus dem obigen Lemma, daß $M_2 \oplus \ldots \oplus M_m \cong N_2 \oplus \ldots \oplus N_n$ gilt. □
æ

Kapitel 19

Boolesche Algebren und Boolesche Funktionen

Definition:
Eine Menge B mit drei Operationen $+ : B \times B \longrightarrow B$, $\cdot : B \times B \longrightarrow B$ und $^{-}: B \longrightarrow B$ sowie zwei ausgezeichnenten Elementen $0, 1 \in B$ heißt Boolesche Algebra, wenn für $a, b, c \in B$ folgende Rechenregeln erfüllt sind:
$a + (b + c) = (a + b) + c, \qquad a \cdot (b \cdot c) = (a \cdot b) \cdot c \qquad$ (Assoziativität)
$a + b = b + a, \qquad a \cdot b = b \cdot a \qquad$ (Kommutativität)
$a + a = a, \qquad a \cdot a = a \qquad$ (Idempotenz)
$a + (b \cdot c) = (a + b) \cdot (a + c), \qquad a \cdot (b + c) = a \cdot b + a \cdot c \qquad$ (Distibutuvität)
$a + (a \cdot b) = a, \qquad a \cdot (a + b) = a \qquad$ (Absorbtion)
$0 + a = a, \qquad 0 \cdot a = 0$
$1 + a = 1, \qquad 1 \cdot a = a$
$a + \bar{a} = 1, \qquad a \cdot \bar{a} = 0.$

Manchmal schreibt man anstelle von $+$ auch $\vee$ oder $\cup$ und nennt diese Operation Disjunktion, Vereinigung oder Supremum; für $\cdot$ schreibt man dann $\wedge$ oder $\cap$ und nennt es Konjunktion, Durchschnitt oder Infimum. Die Operation $^{-}$ heißt Komplementierung oder Negation.

Das einfachste Beipiel einer Booleschen Algebra ist die Algebra $\mathbb{B} = \{0, 1\}$, wo sich die Definition der Rechenoperationen schon aus den obigen Regeln ergibt.
Ein weiteres Beispiel ist die Potenzmenge $P(M) = \{U \mid U \subseteq M\}$ einer Menge M mit Durchschnitt und Vereinigung sowie Komplemmentärtmengenbildung als Rechenoperationen, da, wie man oben sieht, für die Addition und die Multiplikation genau dieselben Rechenregeln gelten, ist es egal, ob wir den Durchschnitt als Addition oder als Multiplikation auffassen.

Wenn B und C Boolesche Algebren sind, so ist $B \times C$ mit komponentenweise definierten Rechenoperationen ebenfalls eine Boolesche Algebra, insbesondere also auch jede kartesische Potenz B^n von B.
Von nun an bezeichne B stets eine Boolesche Algebra.

Für die Komplementierung gelten die folgenden DeMorganschen Regeln:

Satz 19.1 $\overline{x \cdot y} = \bar{x} + \bar{y}$, $\overline{x + y} = \bar{x} \cdot \bar{y}$.

Beweis: Wenn a das Komplement von $x \cdot y$ bezeichnet, so haben wir $a + (x \cdot y) = 1$ und $a \cdot (x \cdot y) = 0$ nachzuweisen:
$(x \cdot y) + (\bar{x} + \bar{y}) = (x + \bar{x} + \bar{y}) \cdot (y + \bar{x} + \bar{y}) = 1 \cdot 1 = 1$,
$(x \cdot y) \cdot (\bar{x} + \bar{x}) = (x \cdot y \cdot \bar{x}) + (x \cdot y \cdot \bar{y}) = 0 + 0 = 0$. Der Beweis der anderen Regel verläuft analog. □

Die soeben benutze Beweismethode („analog“) ist typisch für die Arbeit mit Booleschen Algebren: Man vertauscht die Rechenoperationen miteinander und wendet die analogen Regeln an; dies nennt man „Dualisierung“.

Lemma 19.2 (Kürzungsregel) *Für $x, y, z \in B$ gelte (1) $x \cdot y = x \cdot z$ und (2) $x + y = x + z$. Dann folgt $y = z$.*

Beweis: Zur ersten Gleichung wird sowohl y als auch z addiert:

$$\begin{aligned}(x \cdot y) + y &= (x + y) \cdot (y + y) = (x + y) \cdot y = y \\ &= (x \cdot z) + y = (x + y) \cdot (z + y), \\ (x \cdot z) + z &= (x + z) \cdot (z + z) = (x + y) \cdot z = z \\ &= (x \cdot y) + z = (x + z) \cdot (y + z)\end{aligned}$$

und die beiden letzten Terme jeder Gleichung stimmen wegen (2) überein. □

Wir können in B wie folgt eine Ordnung einführen: $a \leq b$ genau dann, wenn $a \cdot b = a$ gilt.

Lemma 19.3 *$a \leq b$ gdw. $a + b = b$.*

Beweis: $b = (a + b) \cdot b = a \cdot b + b \cdot b = a + b$. □

Definition:
Seien $a \leq b \in B$, dann heißt die Menge $\{x \mid a \leq x \leq b\} = [a, b]$ das durch a und b bestimmte Intervall von B.

Wir bemerken, daß $[a, b]$ bezüglich der Addition und Multiplikation abgeschlossen sind. Wenn wir a als Nullelement und b als Einselement auffassen und die Komplementierung in $[a, b]$ relativ zu diesen durchführt (was auch immer das heißen mag), so wird $[a, b]$ wieder eine Boolesche Algebra.

Eine Abbildung zwischen Booleschen Algebren, die mit den jeweiligen drei Rechenoperationen verträglich ist, heißt Homomorphismus Boolescher Algebren. Ein bijektiver Homomorphismus heißt Isomorphismus.

Nun beweisen wir einen Struktursatz, der eine Übersicht über alle endlichen Boolschen Algebren ergibt.

Satz 19.4 *Wenn B eine endliche Boolesche Algebra ist, so gilt $B \cong \mathbb{B}^n$ für eine natürliche Zahl n.*

Beweis: Wir führen die Induktion über $\mid B \mid$. Wenn $\mid B \mid = 2$ ist, so ist nichts zu zeigen. Sei also die Behauptung für „kleine“ Boolesche Algebren schon bewiesen.
Wir wählen ein Element $a \in B,\ a \neq 0, 1$. Wir setzen

$$X_a = \{(a \cdot b, a + b) \mid b \in B\},$$

dies ist eine Teilmenge von $[0, a] \times [a, 1]$.
Weiter sei $f : B \longrightarrow X_a$ folgende Abbildung: $f(b) = (a \cdot b, a + b)$. Nach der obigen Kürzungsregel ist f injektiv. Wir zeigen die Verträglichkeit mit den Rechenoperationen:

$$f(b \cdot c) = (a \cdot (b \cdot c), a + (b \cdot c)),$$

$$\begin{aligned} f(b) \cdot f(c) &= (a \cdot b, a + b) \cdot (a \cdot c, a + c) \\ &= (a \cdot a \cdot b \cdot c, (a + b) \cdot (a + c)) \\ &= (a \cdot b \cdot c, a + (b \cdot c)), \\ f(b + c) &= f(b) + f(c) \end{aligned}$$

analog. Beim Komplement müssen wir aufpassen: Wir zeigen zunächst, daß $a \cdot \bar{b}$ das Komplement von $a \cdot b$ in $[0, a]$ ist.
$a \cdot \bar{b} + a \cdot b = a \cdot (b + \bar{b}) = a \cdot 1 = a$ ist das größte Element und $(a \cdot \bar{b}) \cdot (a \cdot b) = a \cdot 0 = 0$ ist das kleinste.
Analog: $a + \bar{b}$ ist das Komplement von $a + b$ in $[a, 1]$, da $a + \bar{b} + a + b = 1$ und $(a + \bar{b}) \cdot (a + b) = a + (\bar{b} \cdot b) = a + 0 = a$ ist das kleinste Element.
Nun folgt $f(\bar{b}) = (a \cdot \bar{b}, a + \bar{b}) = \overline{(a \cdot b, a + b)} = \overline{f(b)}$.
Nun ist f auch noch surjektiv, denn $f\ddot{u}r(x, y) \in [0, a]x[a, 1]$ setzen wir $b = y \cdot (\bar{a} + x)$, dann ist $f(b) = (a \cdot y \cdot (\bar{a} + x), a + y \cdot (\bar{a} + x)) = (a \cdot y \cdot \bar{a} + a \cdot y \cdot x, (a + y)(a + \bar{a} + x))$; der erste Term ist Null, der zweite wegen $x \leq a \leq y$ gleich x, der dritte Term ist gleich $(a + y) \cdot (a + \bar{a} + x) = (a + y) \cdot 1 = y$.
Also ist f ein Isomorphismus Boolescher Algebren.
Da nun sowohl $[0, a]$ als auch $[a, 1]$ weniger Elemente als B haben, gilt für sie die Induktionsvoraussetzung: $[0, a] \cong \mathbb{B}^k,\ [a, 1] \cong \mathbb{B}^m$, also $B \cong \mathbb{B}^{k+m}$. □

Die Menge $\mathbb{B}^n$ ist isomorph zur Potenzmenge der Menge $\{1, \ldots, n\}$, wir ordnen dem Tupel $(i_1, \ldots, i_n)$ die Menge der k mit $i_k \neq 0$ zu. Dies ist mit den Operationen verträglich.

Folgerung 19.5 (Stonescher Darstellungssatz) *$B \cong P(M)$ für eine endliche Menge M.* □

Folgerung 19.6 *Zwei gleichmächtige endliche Boolsche Algebren (mit 2^n Elementen) sind isomoprh (zu $\mathbb{B}^n$).* □

Wir betrachten nun n-stellige Abbildungen der Form $f : B^n \longrightarrow B$. Wenn f, g zwei solcher Abbildungen sind, so können wir $(f \cdot g)(x) = f(x) \cdot g(x),\ (f + g)(x) = f(x) + g(x)$ und $(\bar{f})(x) = \overline{f(x)}$ setzen und es ist nicht schwer nachzuweisen, daß die Menge $F_n(B) = \{f : B^n \longrightarrow B\}$ so eine Boolsche Algebra wird.

Definition:
Ein Boolsches Polynom in $x_1, \ldots, x_n$ ist folgendes:
(1) $x_1, \ldots, x_n, 0, 1$ sind Boolesche Polynome,
(2) wenn p und q Boolesche Polynome sind, so sind auch $(p) + (q), (p) \cdot (q)$ und $\overline{(p)}$ Boolesche Polynome.

Ein Boolesches Polynom ist also einfach eine Zeichenkette, es gilt $x_1 + x_2 \neq x_2 + x_1$. Wenn aber $f(x_1, \ldots, x_n)$ ein Boolesches Polynom und B eine Boolesche Algebra ist, so können wir eine Funtion $f^* : B^n \longrightarrow B$ durch $f^*(b_1, \ldots, b_n) = f(b_1, \ldots, b_n)$ konstruieren, indem wir die b_i einfach in f einsetzen und den Wert ausrechnen. Dann gilt natürlich $(x_1 + x_2)^* = (x_2 + x_1)^*$.

Definition:
Zwei Boolsche Polynome f, g heißen äquivalent ($f \sim g$), wenn die zugehörigen Funktionen auf der Algebra $\mathbb{B}$ gleich sind.

Zur Vereinfachung führen wir folgende Schreibweisen ein: $x^1 = x, x^{-1} = \bar{x}$.

Satz 19.7 *Jedes Boolesche Polynom ist äquivalent zu einer „disjunktiven Normalform“*

$$f_d(x_1, \ldots, x_n) = \sum_{i_1, \ldots i_n \in \{0,1\}} d_{i_1 \ldots i_n} \cdot x_1^{i_1} \cdot \ldots \cdot x_n^{i_n},\ d_{i_1 \ldots i_n} \in \{0, 1\}.$$

Jedes Boolesche Polynom ist äquivalent zu einer „konjunktiven Normalform“

$$f_k(x_1, \ldots, x_n) = \prod (k_{i_1 \ldots i_n} + x_1^{i_1} + \ldots + x_n^{i_n}),\ k_{i_1 \ldots i_n} \in \{0, 1\}.$$

Beweis: Es ist $f^*(1^{j_1}, \ldots, 1^{j_n}) = \sum d_{i \ldots i} 1^{i_1 j_1} \ldots 1^{i_n j_n}$ und ein Summand ist genau dann gleich 1, wenn $i_1 = j_1, \ldots, i_n = j_n$ und $d_{i_1 \ldots i_n} = 1$ ist, das heißt, die f_d mit verschiedenen d sind jeweils inäquivalent. Nun ist aber die Anzahl der disjunktiven Normalformen gleich 2^{2^n}, also gleich der Zahl aller Funktionen $\mathbb{B}^n \longrightarrow \mathbb{B}$.
Die zweite Aussage ergibt sich durch Dualisierung. □

Folgerung 19.8 *In der obigen Darstellung ist* $d_{i_1 \ldots i_n} = f^*(1^{i_1}, \ldots, 1^{i_n})$.

Beispiel: $f = ((x_1 + x_2) \cdot \bar{x}_1) + (x_2 \cdot (x_1 + \bar{x}_2))$, dann ist $f(0, 0) = f(1, 0) = 0$ und $f(0, 1) = f(1, 1) = 1$, die disjunktive Normalform von f erhalten wir, indem wir in der Wertetabelle die Stellen aufsuchen, wo der Wert 1 angenommen wird. Wenn hier ein Argument gleich 0 ist, so ist die entsprechende Variable zu komplementieren, sonst nicht. Also $f \sim \bar{x}_1 x_2 + x_1 x_2$. Dies kann weiter vereinfacht werden: $f \sim (\bar{x}_1 + x_2) \cdot x_2 = 1 \cdot x_2 = x_2$.

Wir überlegen nun, wie man eine Darstellung von Polynomen vereinfachen kann.

Definition Es seien p und q Boolesche Polynome; wir sagen, daß p das Polynom q impliziert, wenn aus $p^*(b_1, \ldots, b_n) = 1$ folgt, daß auch $q^*(b_1, \ldots, b_n) = 1$ gilt (dabei ist $b_i \in \{0, 1\}$.
Wir bezeichnen ein Polynom als „Produkt“, wenn es kein +-Zeichen enthält.
Das Polynom p heißt Primimplikant von q, wenn gilt
1) p ist ein Produkt,

2) p impliziert q,
3) kein Teilprodukt von p impliziert q.

Sei zum Beispiel $q = x_1x_2x_3 + x_1\bar{x}_2x_3 + \bar{x}_1\bar{x}_2\bar{x}_3$ und $p = x_1x_2$, dann wird q von p impliziert, denn $p^* = 1$ gilt nur für $x_1 = x_3 = 1$ und es ist $q^*(1, x_2, 1) = (x_2+\bar{x}_2+\bar{x}_2)^* = 1$, aber z.B. x_1 impliziert q nicht, da $q^*(1, x_2, x_3) = (x_2x_3 + \bar{x}_2x_3 + 0)^* = (x_2 + \bar{x}_2)x_3 = x_3 \neq 1$ ist.

Wir bemerken, daß ein Produkt genau dann 1 ist, wenn alle nichtkomplementierten Variablen gleich 1 und alle komplementierten Variablen gleich 0 gesetzt werden.
Alle Summanden einer disjunktiven Normalform sind Implikanten.

Satz 19.9 *Jedes Polynom ist äquivalent zur Summe seiner Primimplikanten.*

Beweis: Seien $p_1, \ldots, p_m$ die Primimplikanten von q, wir setzen $p = p_1 + \ldots p_m$. Sei nun $p^*(b_1, \ldots, b_n) = 1$, dann gibt es ein p_i mit $p_i(b_1, \ldots, b_n) = 1$ und da p_i das Polynom q impliziert, gilt auch $q^*(b_1, \ldots, b_n) = 1$.
Sei umgekehrt $q^*(b_1, \ldots, b_n) = 1$, wir setzen $s = x_1^{i_1} \cdots x_n^{i_n}$ mit $i_k = 1$, falls $b_k = 1$ und $i_k = -1$ für $b_k = 0$, dann ist s ein Implikant von q. Wir lassen nun aus dem Term s alle die x_i weg, für die $q^*(b_1, \ldots, b_{i-1}, \bar{b}_i, \ldots) = 1$ ist; das Ergebnis sei r. Dann gilt: r impliziert q, aber kein Teilwort von r impliziert q, folglich ist r als Primimplikant gleich eienm der p_j, also folgt $p^*(b_1, \ldots, b_n) = 1$, d.h. $p \sim q$. □

Von der disjunktiver Normalform eines Polynoms ausgehend kann man eine Darstellung als Summe von Primimplikanten erhalten, indem man für alle Paare von Summanden, wo dies möglich ist, die Regel $px + p\bar{x} \sim p$ anwendet.

Index

ähnliche Matrizen, 189
äquivalente Darstellungen, 247
äquivalente Polynommatrizen, 187

adjungierte Abbildung, 144
affine Abbildung, 68
affine Hülle, 174
affiner Raum, 63
affiner Unterraum, 64
Algebra, 239
allgemeine Lage, 64
Anfangsminoren, 161
Annulator, 78
ausgezeichnete Spalten, 17

Basis, 28
Bild, 204
Bilinearformen, 81

Charakter, 251
charakteristisches Polynom, 110

Darstellung, 246
Determinante, 94
Determinantenteiler, 188
Dimension, 30
Dimensionssatz, 31
direkte Summe, 32
duale Abbildung, 78

Eigenvektor, 110
Eigenwert, 110
einfache Algebra, 241
einfacher Modul, 240
Einsdarstellung, 247
elementare Operationen, 14
Erzeugendensystem, 25

Gaußscher Algorithmus, 17
Gleichungssysteme, 11
größter gemeinsamer Teiler, 120
Gram-Matrix, 167
Gruppe, 199
Gruppenhomomorphismus, 203

halbeinfache Algebra, 244
halbeinfacher Modul, 242
Hauptideal, 221
Hom, 43
Homomorphiesatz, 205

Ideal, 221
idempotent, 55
Invariantenteiler, 188
invarianter Unterraum, 127, 250
inverse Abbildung, 44
inverse Matrix, 49
involutiv, 55
irreduzibel, 250
irreduzibles Polynom, 221

Jacobson-Radikal, 235
Jordankästchen, 132
Jordansche Normalform, 132

Körper, 7
Kern, 44, 203
Klassenfunktion, 254
Kommutator, 249
konvex, 174
konvexe Hülle, 175
konvexe Pyramide, 176
konvexes Polyeder, 175
Koordinaten, 28
Koordinatensystem, 64

Laplacescher Entwicklungssatz, 95
Leibnizsche Deteminantendefinition, 96
linear unabhängig, 26

lineare Abbildung, 41
lineare Hülle, 24
lineares Gleichungssystem, 12
Linearkombination, 13
LU-Zerlegung, 55

Matrix, 16
Matrixdarstellung, 246
Matrixnorm, 158
Matrixprodukt, 48
maximale linear unabhängige Menge, 27
minimales Erzeugendensystem, 26
minimales Linksideal, 240
Minimalpolynom, 123
Minor, 100, 111
Moore-Penrose-Inverse, 149
Multilinearform, 94

Newtonsche Formeln, 122
nichtausgeartete Bilinearform, 82
nilpotent, 55, 129, 259
normale Matrix, 146
normale Untergruppe, 204
normaler Endomorphismus, 146
Normalteiler, 204

Ordnung einer Gruppe, 202
orthogonale Matrix, 142
orthogonales Komplement, 141
Orthogonalitätsrelationen, 253
Orthonormalbasis, 140
Orthonormalsystem, 140

parallel, 67
Permutation, 95
positiv definit, 161
positive Linearkombination, 175
Potenzsummen, 122
primitiv, 262
primitives Polynom, 221
Primzahlzerlegung, 199
pseudoreguläre Matrix, 148

Quadrik, 90
quasi-invertierbar, 233

Radikal, 260
Rang, 35
reduzibel, 250
reguläre Darstellung, 247
reguläre Polynommatrix, 185
Ring, 217

Satz von Hamilton-Cayley, 115, 191
Satz von Kronecker/Capelli, 36
Seite eines Simplex, 175
selbstadjungierte Abbildung, 145
semi-primitiv, 235
Simplex, 175
Skalarprodukt, 139
Smithsche Normalform, 187
Spaltenrang, 34
Spaltenraum, 34
sphärische Geometrie, 172
Summe von Unterraumen, 31
symmetrische Bilinearform, 83

Tensorprodukt von linearen Abbildungen, 230
Tensorprodukt von Moduln, 228
Torsionsuntergruppe, 208
Tranposition, 207

unipotent, 160
Untergruppe, 200
Unterraum, 24
unzerlegbarar Vektorraum, 128
unzerlegbarer Modul, 261

Vektoren, 23
Vektornorm, 158
Vektorprodukt, 171
Vektorraum, 23
Verbindungsraum, 68

windschief, 67

Zeilenrang, 33
Zeilenraum, 33
zerlegbarer Modul, 261
zerlegbarer Vektorraum, 128
Zyklus, 206