Kapitel 1

Einführung

1.1 Literatur

G. Fischer: *Lineare Algebra*, Vieweg Studium Verlag

Weitere Bücher:

F. Lorentz: *Lineare Algebra I und II*, BI Wissenschaftsverlag

K. Jänich: *Lineare Algebra*, Springer Verlag

Stammbach: *Lineare Algebra*, Teubner Verlag

W. Klingenberg: *Lineare Algebra und Geometrie*, Springer Verlag

S. Lang: *Linear Algebra*, Addison-Wesley

van der Waerden: *Mathematik für Naturwissenschaftler*, BI Wissenschaftsverlag

Courant und Hilbert Methoden der Mathematischen Physik I+II

1.2 Etwas über die Formale Struktur der Mathematik

- **Definitionen:** Begriffe oder Schreibweisen werden festgelegt.

  Zum Beispiel kommt die folgende Schreibweise überall in der linearen Algebra vor:

  Definition 1. $\sum_{i=1}^{n} a_i \equiv a_1 + a_2 + \cdots + a_n$.

  (d.h. etwa: $\sum_{i=1}^{3} i = 1 + 2 + 3$)

  Bemerkung. es gilt $\sum_{i=1}^{0} a_i \equiv 0$.

- **Behauptungen:** (Lemmata, Sätze, Korollare) sind Aussagen, die—mehr oder weniger—mit Hilfe der vorher festgelegten Definitionen bewiesen werden müssen.

- **Beweise**: Es gibt im allgemeinen drei Beweismethoden:

  - direkte Beweise,
  - indirekte Beweise (durch Widerspruch), und
  - induktive Beweise.

- Beispiel für einen direkten Beweis:

  Satz 1.1 (Gauß als Schulkind. ca. 1787). *Es gilt* $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$.

  Beweis. Es gilt:

$$\sum_{i=1}^{n} i = 1 + 2 + \cdots + n,$$

$$\sum_{i=1}^{n} i = n + (n-1) + \cdots + 1.$$

Daher

$$\left(\sum_{i=1}^{n} i\right) + \left(\sum_{i=1}^{n} i\right) = (1+n) + (2+(n-1)) + \cdots + (n+1)$$
$$= \underbrace{(n+1) + (n+1) + \cdots + (n+1)}_{n-\text{mal}}$$
$$= n(n+1)$$

$$\Rightarrow \sum_{i=1}^{n} i = \frac{n(n+1)}{2}.$$

□

- Beispiel für einen indirekten Beweis:

Satz 1.2. *$\sqrt{2}$ ist irrational.*

Beweis. Angenommen, $\sqrt{2}$ ist doch rational, so daß $\sqrt{2} = a/b$, mit a und $b > 0$ Integerzahlen (wir können natürlich annehmen, daß $b \neq 0$). Seien a, b die *kleinsten* Integerzahlen mit $\sqrt{2} = a/b$. Es gilt:

$$(\sqrt{2})^2 = 2 = \left(\frac{a}{b}\right)^2 = \frac{a^2}{b^2}$$
$$\Rightarrow 2b^2 = a^2$$
$$\Rightarrow a^2 \text{ ist eine gerade Zahl}$$
$$\Rightarrow a \text{ ist eine gerade Zahl}$$

Wir können daher schreiben $a = 2c$, wobei $c \neq 0$ auch eine Integerzahl ist. D.h.:

$$2b^2 = (2c)^2 = 4c^2$$
$$\Rightarrow b^2 = 2c^2$$

Das heißt, b ist eine gerade Zahl, etwa $b = 2d$, und sicherlich ist $d < b$. Es gilt

$$\sqrt{2} = \frac{a}{b} = \frac{2c}{2d} = \frac{c}{d},$$

wobei c und d Integerzahlen sind, die kleiner sind als a und b. Widerspruch. Folglich ist unsere Annahme falsch: d.h. es existieren *keine* Integerzahlen a und b mit $\sqrt{2} = a/b$. □

- Beispiel für einen Beweis mittels (vollständiger) Induktion [1]:

Satz 1.3. *Es gilt $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$ für alle $n \geq 0$.*

Beweis. (durch Induktion nach n).

1. **Induktionsanfang** ($n = 0$).

$$\sum_{i=1}^{0} i = 0 = \frac{0 \times (0+1)}{2}$$

2. **Induktionsschritt** ($n \to n+1$).

Sei $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$. Wir müssen nun zeigen, daß

$$\sum_{i=1}^{n+1} i = \frac{(n+1) \times (n+2)}{2}$$

[1] Früher hat man 'Beweise' auch durch *nicht*-vollständige Induktion geführt, indem man einige Beispiele einfach ausgerechnet hat. Beispiel: $\sum_{i=1}^{1} i = 1$, $\sum_{i=1}^{2} i = 3$, $\sum_{i=1}^{3} i = 6$. Ah-Ha! Es scheint zu stimmen! QED(?)

Aber hierfür gilt

$$\sum_{i=1}^{n+1} i = \left(\sum_{i=1}^{n} i\right) + (n+1) = \frac{1}{2}n(n+1) + (n+1)$$
$$= \frac{1}{2}(n+1)n + \frac{1}{2}(n+1)2$$
$$= \frac{1}{2}(n+1)(n+2).$$

□

Im allgemeinen geht es hier darum, eine Aussage $A(n)$ zu beweisen, wobei $n = 1, 2, 3, \ldots$. (D.h. es gibt eigentlich unendlich viele verschiedene Aussagen: $A(1)$, $A(2)$, u.s.w.) Um $A(n)$ ganz allgemein zu beweisen (etwa für alle $n \geq n_0$, wobei n_0 eine vorgegebene Integerzahl ist), genügt es, nur zwei bestimmte Aussagen zu beweisen, nämlich:

(I) $A(n_0)$ ist richtig. (Induktionsanfang)

(II) Für beliebige $n \geq n_0$ gilt: *Falls* $A(n)$ richtig ist, dann ist auch $A(n+1)$ richtig.

Die Beweismethode der vollständigen Induktion mag zunächst neu erscheinen, aber alle, die Erfahrung mit der Computerprogrammierung haben, werden feststellen, daß diese Methode nichts anderes ist, als die Idee der rekursiven Programmierung:

```
function summiere (n : integer) : integer;
   begin
      if n <= 0 then summiere := 0
         else summiere := summiere(n-1) + n
   end;
```

Manchmal ist es zweckmäßig, einen Beweis in kleinere Abschnitte zu zerlegen. Zum Beispiel:

Satz 1.4. *Jede positive Integerzahl besitzt eine eindeutige Primfaktorzerlegung.*

Wir benutzen hierbei die übliche Definition

Definition 2. Sei $N = \{1, 2, 3, \ldots\}$ die Menge der 'natürlichen' Zahlen. Eine positive Integerzahl p heißt *prim*, falls kein $q \in N$ existiert mit $1 < q < p$ und $p = qr$, wobei auch $r \in N$.

Um den Satz zu beweisen, werden wir zunächst einige Hilfssätze beweisen. Wir brauchen auch eine weitere Definition.[2]

Definition 3. Seien $a, b \in N$. a heißt *Teiler* von b, geschrieben $a|b$, falls $\exists c \in N$ mit $ac = b$. Seien $n, m \in N$. Die Zahl $a \in N$ heißt größter gemeinsamer Teiler ($ggt(n,m)$), falls $a|n$ und $a|m$, und $\exists$ kein $b \in N$ mit $a < b$, aber b teilt sowohl n als auch m.

Zwei besondere Fälle:

- Falls $n = m$, dann ist $ggt(n,m) = n = m$.
- Es gilt immer $1|a$, für alle $a \in N$. Falls $ggt(n,m) = 1$, dann heißen m und n *relativ prim* zueinander.

Insbesondere ist eine Primzahl p zu allen Zahlen, die kleiner als p sind, relativ prim.

Lemma. *Seien m, n relativ prim. $\Rightarrow \exists$ Integerzahlen a, b mit*

$$am + bn = 1.$$

Beweis. Falls $m = 1$ oder $n = 1$, dann ist das Ergebnis trivial. Falls $m = n > 1$, dann sind m und n nicht zueinander relativ prim. Sei daher etwa $1 < m < n$. Wir benutzen Induktion über n.

Für $m = 2$, $n = 3$ gilt $1 = (-1) \times 2 + 1 \times 3$. Sei daher $1 < m < n$ gegeben mit $n > 3$, und sei das Lemma wahr für alle relativ prim Paare, die kleiner als (m, n) sind. Wir schreiben

$$n = xm + y$$

wobei $x, y \in N \cup \{0\}$, $y < m$. Es gilt $0 < y$, da sonst $m|n$ wäre, was der Voraussetzung m, n relativ prim widersprechen würde.

[2]Bemerken Sie, daß $\exists$ heißt 'es existiert' und $\forall$ heißt 'für alle'.

Aber y, m sind zueinander auch relativ prim. Warum? Sonst würde ein $z \in N$ existieren, $z > 1$, mit $z|y$ und $z|m$.

$$\Rightarrow z|n, \text{ wobei } n = xm + y.$$

Das heißt, z würde sowohl m als auch n teilen. Widerspruch, da m und n doch zueinander relativ prim sind.

Nun, das Paar (y, m) ist kleiner als das Paar (m, n). Folglich können wir die induktive Hypothese in Anspruch nehmen, um auf die Existenz von zwei weiteren Zahlen c und d zu schließen mit der Eigenschaft, daß $1 = cy+dm$. Aber dann gilt

$$\begin{aligned} 1 &= cy + dm \\ &= c(n - xm) + dm \\ &= (d - cx)m + cn \end{aligned}$$

Das heißt, wir können einfach $a \equiv d - cx$ und $b \equiv c$ nehmen. □

Korollar 1.4.1. *Angenommen $p, a, b \in N$, mit p eine Primzahl und $p|ab$. Dann gilt entweder $p|a$ oder $p|b$.*

Beweis. Falls *nicht* $p|a$, dann sind p und a relativ prim. $\Rightarrow \exists x, y$ mit

$$1 = xp + ya.$$

$$\Rightarrow b = xpb + yab.$$

Aber da $p|ab \Rightarrow p|b$. □

Beweis des Satzes über Primfaktorzerlegung.

Offensichtlich existieren immer Primfaktorzerlegungen für beliebige positive Integerzahlen. Wir müssen nur beweisen, daß diese Zerlegungen eindeutig sind. Um einen Widerspruch zu finden, sei angenommen, daß der Satz doch nicht richtig sei. Das würde heißen, daß eine Zahl $n \in N$ existieren würde mit zwei verschiedenen Primfaktorzerlegungen, etwa

$$n = p_1 p_2 \ldots p_n = q_1 q_2 \ldots q_m.$$

Sei n die kleinste Zahl in N, die zwei verschiedene Primfaktorzerlegungen besitzt. Kann es sein, daß $p_i = q_j$ für irgendwelche i, j? Nein, sonst wäre

$$p_1 \ldots p_{i-1} p_{i+1} \ldots p_n = q_1 \ldots q_{j-1} q_{j+1} \ldots q_m$$

ein kleineres Beispiel. Es gilt nun

$$p_1 | n = q_1 \ldots q_m = q_1(q_2 \ldots q_m)$$

Aber p_1 ist sicherlich kein Teiler von q_1, da q_1 doch prim ist. Folglich (nach unserem Korollar) muß $p_1 | q_2 \ldots q_m$. Und so weiter...! Letzten Endes müssen wir schließen, daß $p_1 | q_m$, was wiederum unmöglich ist. Insgesamt haben wir einen Widerspruch.

1.3 Etwas über die Logik

Ich habe das gewöhnliche mathematische Zeichen '$\Rightarrow$' in diesen Beweisen benutzt. $A \Rightarrow B$ heißt 'aus A folgt B' oder 'A impliziert B'. Logiker versuchen, diese Idee noch genauer zu beschreiben, indem sie eine 'Wahrheitstabelle' aufstellen.

'$\Rightarrow$'	wahr	falsch
wahr	wahr	falsch
falsch	wahr	wahr

Das heißt z.B., daß die gesamte Aussage '$A \Rightarrow B$' wahr ist, falls etwa A und B beide wahr sind. Aber nach dieser Vereinbarung gilt auch '$A \Rightarrow B$' ist wahr, falls A falsch und B wahr! Ich will nicht viel über diese logische Beziehungen reden. In Wirklichkeit läuft auch alles in der Mathematik über den 'gesunden Menschenverstand'. Viel einfacher ist die Beziehung $A \Leftrightarrow B$. Das heißt, A ist wahr genau dann, wenn B auch wahr ist.

'$\Leftrightarrow$'	wahr	falsch
wahr	wahr	falsch
falsch	falsch	wahr

Sie sollten sich auf jeden Fall (entweder mittels Wahrheitstabellen oder durch 'Verstehen') die folgenden logischen Beziehungen durch den Kopf gehen lassen.

$$(A \Leftrightarrow B) \quad \Leftrightarrow \quad (A \Rightarrow B \text{ und } B \Rightarrow A)$$

$$(A \Rightarrow B) \quad \Leftrightarrow \quad (\text{nicht } B \Rightarrow \text{nicht } A)$$

$$(\text{nicht (nicht } A)) \quad \Leftrightarrow \quad A$$

$$(A \Rightarrow (\text{nicht } B)) \quad \Leftrightarrow \quad (B \Rightarrow (\text{nicht } A))$$

und so weiter... Zum Beispiel:

$$\forall n \geq 0, \quad \exists m > n$$

ist sicherlich wahr für geeignete Integerzahlen n und m.

Hier ist eine kleine zusätzliche Übung zur allgemeinen Überlegung: Sei 'A' irgendeine logische Aussage, die die '$\exists$' und '$\forall$' Zeichen enthält (und vielleicht auch '$\Rightarrow$', '$\Leftarrow$', '$\Leftrightarrow$' und 'nicht'). Was ist dann die logische Aussage 'nicht A'? Zum Beispiel gilt nach dem 'gesunden Menschenverstand' sicherlich:

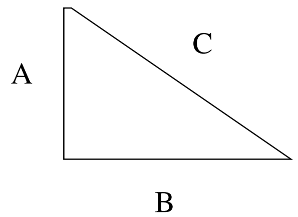
$$\begin{aligned}&(\text{nicht } (\forall n \geq 0, \exists m \text{ so daß } m > n)) \\ &\quad \Leftrightarrow \quad (\exists n \geq 0 \text{ so daß nicht } m > n, \forall m).\end{aligned}$$

Gibt es eine formale Regel für diese logische Umkehrung?

1.4 Was ist lineare Algebra?

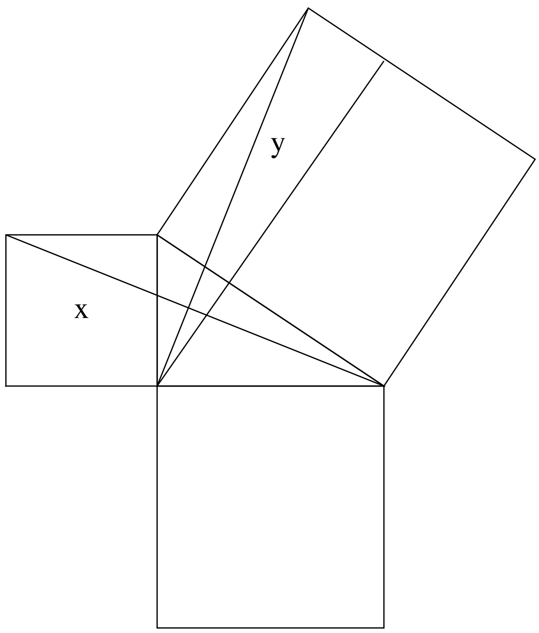
Antwort: Lineare Algebra ist das Studium von Algebra und Geometrie. (Zumindest heißt es so in der Prüfung!) In der linearen Algebra werden wir uns jedenfalls mit der 'traditionellen' Hälfte der Mathematik auseinandersetzen. (Die Analysis hingegen ist erst in den letzten 300 Jahren entstanden.) Dabei geht es zuerst darum, zu überlegen, was die Mathematik überhaupt ist.

Beispiel (aus der Schulmathematik):



Satz 1.5 (Pythagoras). *Für das rechtwinklige Dreieck im Bild gilt*

$$A^2 + B^2 = C^2.$$



Beweis.

Wie man sieht, gilt: Flächeninhalt des Quadrats x gleich Flächeninhalt des Rechtecks y. u.s.w. $\square$

Frage: ist dies ein Beweis? Was sind die Annahmen? Was wird überhaupt bewiesen? Laut Plato's 'Meno' Dialog kennt jeder schon im voraus alle Wahrheiten. Nach dieser Auffassung wäre es meine Aufgabe, Sie an diese Wahrheiten einfach zu erinnern, so daß Sie die mathematischen Beweise 'wieder' sehen. Aber wir in der modernen Welt kennen viele bizarre Ideen, die zu Zweifel an Plato's einfacher Auffassung von 'Wahrheit' führen. Zum Beispiel:

Definition 4. (Nach Plato) Zwei geometrische Objekte sind gleich groß, falls sie 'zerlegungsgleich' sind, wobei X und Y zerlegungsgleich sind, falls disjunkte Zerlegungen von X und Y in gleich viele Teile existieren ($\{X_1, \ldots, X_n\}$ und $\{Y_1, \ldots, Y_n\}$), mit der Eigenschaft, daß für jedes i die Teile X_i und Y_i sich so bewegen lassen, daß sie miteinander übereinstimmen.

Für Plato war dies praktisch das elementarste Prinzip überhaupt, und es diente als Grundlage für die Erforschung vieler anderer Wahrheiten. Aber es gilt:

Satz 1.6 (Banach, Tarski - 1924). *Sei $B \subset \mathbf{R}^3$ die Einheitskugel. Dann gibt es eine disjunkte Zerlegung in zwei Teile $B = B_1 \cup B_2$, so daß alle drei Mengen B_1, B_2 und B zerlegungsgleich sind!*

Das heißt, nach dieser Auffassung wäre die 3-Kugel 2-mal so groß wie sie selbst! So etwas ist offensichtlich absurd, und es ist klar, daß wir eine einfachere, 'vernünftigere' Definition als die von Plato suchen müssen. Insgesamt müssen wir viel vorsichtiger mit 'Geometrie' umgehen als Sie es von der Schule her gewöhnt sind! Deswegen ist es nötig, daß wir uns schon jetzt mit einigen scheinbar langen, abstrakten und unmotivierten Definitionen auseinandersetzen.

Kapitel 2

Zahlensysteme, Körper, Vektorräume

2.1 Die Zahlen

Die 'üblichen' Zahlensysteme sind:

- die natürlichen Zahlen: $\mathbf{N} = \{1, 2, 3, \dots\}$
- die Integerzahlen: $\mathbf{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$
- die rationalen Zahlen: $\mathbf{Q} = \{\frac{a}{b} : a \in \mathbf{Z}, b \in \mathbf{N}\}$
- die reellen Zahlen: $\mathbf{R}$, Z.B $1/2$, π, $\sin(e^{\sqrt{2}})$, u.s.w.
- die komplexen Zahlen: $\mathbf{C} = \{x + yi : x, y \in \mathbf{R}, i = \sqrt{-1}\}$

Die Zahlensysteme $\mathbf{N}$, $\mathbf{Z}$ und $\mathbf{Q}$ sind natürlich allgemein bekannt; die reellen Zahlen sind komplizierter. Es gilt, daß $\mathbf{N} \subset \mathbf{Z} \subset \mathbf{Q} \subset \mathbf{R} \subset \mathbf{C}$). Alle reellen Zahlen lassen sich als Dezimalbruchzahlen ausdrücken. Zum Beispiel

$$\sqrt{2} = 1,41421\dots = 1 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{2 + \dots}}}}$$

(Der dritte Ausdruck hier ist eine Kettenbruchzahl. Möglicherweise (nach Plato's 'Parmedides' Dialog) war dies der eigentliche Gegenstand der griechischen Mathematik.) Eine weitere gängige reelle Zahl ist

$$\pi = 3,14159\dots$$

Die folgende interessante Beziehung (mit Kettenbruchzahlen) hat Brouncker (1654) entdeckt:

$$\frac{\pi}{4} = \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{3^2}{2 + \cfrac{5^2}{2 + \cfrac{7^2}{2 + \dots}}}}}$$

Auf jeden Fall gilt: eine reelle Zahl $x \in \mathbf{R}$ ist auch rational $\Leftrightarrow$ die entsprechende Dezimalbruchzahl periodisch ist. (Auch $\Leftrightarrow$ die entsprechende Kettenbruchentwicklung endlich ist.) Zum Beispiel:

$$\frac{4}{3} = 1,33333333\dots$$

2.2 Die moderne Auffassung von Geometrie

Statt wie die alten Griechen mit Zirkel und Lineal zu arbeiten, wird die Geometrie in der modernen Mathematik nach der Methode von Descartes (1596-1650) definiert.

Definition 5. Die kartesische Ebene ist die Menge $\mathbf{R}^2$, bestehend aus Zahlenpaaren (x, y), wobei sowohl x als auch y reelle Zahlen sind. Das heißt:

$$\mathbf{R}^2 \equiv \{(x, y) : x, y \in \mathbf{R}\}$$

Allgemeiner:

Definition 6. Seien $X_1, X_2, \dots, X_n$ beliebige Mengen. Das kartesische Produkt ist die Menge

$$X_1 \times X_2 \times \cdots \times X_n \equiv \{(x_1, x_2, \dots, x_n) : x_1 \in X_1, \dots, x_n \in X_n\}.$$

Falls $X_1 = X_2 = \cdots = X_n = X$ etwa, dann schreibt man auch

$$X_1 \times X_2 \times \cdots \times X_n \equiv X^n$$

Definition 7. Der n-dimensionale Euklidische Raum ($n \in \mathbf{N}$) ist die Menge

$$\mathbf{R}^n \equiv \{(x_1, x_2, \ldots, x_n) : x_i \in \mathbf{R},\ \forall\, i = 1, \ldots, n\}$$

Mit Hilfe dieser Definition ist es möglich, Räume mit beliebig hohen Dimensionen zu untersuchen. Aber für uns ist im Augenblick wichtig, die Tatsache festzuhalten, daß wir dadurch das Studium der Geometrie auf das Studium von Zahlen reduziert haben. Es ist daher vernünftig, eine genaue Definition des Zahlenbegriffs zu haben. Zuerst aber noch einige Bemerkungen über die Vorteile dieser kartesischen Vorstellung gegenüber der klassischen griechischen Vorstellung.

- Die Kartesische Geometrie führt (über die moderne Algebra) zu Lösungen der klassischen Fragen (die Unmöglichkeit der Quadratur des Kreises, u.s.w.).
- Kartesische Geometrie läßt eine Beschreibung durch mathematische Formeln zu. Z.B. beschreiben Formeln in der Physik die Bewegungen von Objekten im Raum.
- Andererseits können wir jetzt geometrische Vorstellungen benutzen, um algebraische Fragen (etwa bezüglich Systeme von linearer Gleichungen) zu lösen.

2.3 Zahlen im allgemeinen: Körper

Was ist ein Zahlensystem? Zuerst braucht man eine Menge (nennen wir diese Menge M) von 'Zahlen': dann müssen auf dieser Zahlenmenge die Operationen $(+, -, \times, /)$ definiert werden. Zum Beispiel $2 + 3 = 5$ oder $2 \times 3 = 6$. Abstrakt gesehen ist etwa '+' eine Operation, die jedem Zahlenpaar (x, y) eine weitere Zahl (nämlich $x + y$) zuordnet. Das heißt, '+' ist eine *Abbildung* (wir sagen auch 'Funktion')

$$+ : M \times M \to M.$$

Die Vorstellung von Abbildungen zwischen Mengen ist sicherlich die wichtigste Idee der ganzen modernen Mathematik!

Definition 8. Seien X und Y zwei vorgegebene Mengen. Sei f eine Vorschrift, die jedem Element $x \in X$ genau ein Element aus Y zuordnet. Dann heißt f eine '*Abbildung*' von X nach Y. Man schreibt

$$f : X \to Y.$$

Sei $x \in X$, dann ist $f(x) \in Y$ das entsprechende Element in Y. Falls die Menge Y ein 'Zahlensystem' ist, dann heißt f auch *Funktion.*

Beispiele:

(i) $\sin : \mathbf{R} \to \mathbf{R}$ — die Sinusfunktion — z.B. $\sin(\pi/2) = 1$ u.s.w.

(ii) $|\cdot| : \mathbf{R} \to \mathbf{R}$ — die Absolutbetragfunktion — z.B.
$|2| = 2$, während
$|-2| = 2$.

(iii) $+ : \mathbf{R} \times \mathbf{R} \to \mathbf{R}$ — die Plusfunktion — $2 + 3 = 5$ u.s.w.

(iv) $\mathrm{trunc} : \mathbf{R} \to \mathbf{Z}$ — 'Truncate' (diese Funktion wird in den meisten Computersprachen angeboten) — $\mathrm{trunc}(x)$ ist die größte Integerzahl n, mit $n \leq x$.

Jetzt kommt unsere Hauptdefinition: nämlich die von 'Körpern' (d.h. 'Zahlensystemen'). Die üblichen Zahlensysteme $\mathbf{Q}$, $\mathbf{R}$ und $\mathbf{C}$ sind Körper. (Aber die Systeme $\mathbf{N}$ und $\mathbf{Z}$ sind keine Körper. – Mehr davon später!)

Definition 9. Sei F eine Menge, versehen mit zwei Abbildungen:

$$`\,+\,' : F \times F \to F \quad \text{und}$$

$$`\,\cdot\,' : F \times F \to F,$$

genannt *Addition* und *Multiplikation.* F heißt *Körper*, falls die folgenden Axiome erfüllt sind.

(F1) $a + b = b + a$, $\forall\, a, b \in F$ (Kommutativität der Addition)

(F2) $a + (b + c) = (a + b) + c$, $\forall\, a, b, c \in F$ (Assoziativität der Addition)

(F3) $\exists$ ein eindeutig bestimmtes Element, genannt $0 \in F$, mit $a + 0 = a$, $\forall\, a \in F$ (Identitätselement unter Addition)

(F4) $\forall\, a \in F$, $\exists$ ein eindeutig bestimmtes Element, genannt $-a \in F$, mit $a + (-a) = 0$ (additives Inverses: man schreibt auch statt $a + (-b)$ einfach $a - b$.)

(F5) $a \cdot b = b \cdot a$, $\forall\, a, b \in F$ (man schreibt auch statt $a \cdot b$ einfach ab)

(F6) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$, $\forall\, a, b, c \in F$

(F7) $\exists$ ein eindeutig bestimmtes Element $1 \in F$, mit $1 \neq 0$ und $a \cdot 1 = a$, $\forall\, a \in F$. (Identitätselement unter Multiplikation)

(F8) $\forall a \in F$ mit $a \neq 0$, $\exists$ ein eindeutig bestimmtes Element $a^{-1} \in F$ mit $a \cdot (a^{-1}) = 1$ (multiplikatives Inverses)

(F9) $a \cdot (b + c) = a \cdot b + a \cdot c$, $\forall\, a, b, c \in F$ (Distributivitätsgesetz)

Es ist leicht nachzuprüfen, daß **Q** und **R** Körper sind, aber nicht **N** und **Z**. (Bei **Z** fehlt nur (F8), bei **N** werden sowohl (F3), (F4), als auch (F8) verletzt.) **Z** ist kein Körper, da (F8) fehlt. Zahlensysteme wie **Z**, wo nur Axiom (F8), fehlt heißen (kommutative) 'Ringe' (mit eins).

Definition 10. Sei G eine Menge mit einer Operation '$\cdot$' : $G \times G \to G$. Die Menge G, versehen mit dieser Operation, heißt *Gruppe*, falls die Axiome (F6), (F7) – (aber ohne Erwähnung der '0') – und (F8) gelten für das Paar $(G, `\cdot\text{'})$. Falls auch (F5) gilt, dann heißt die Gruppe G *abelsch*.

Man könnte dann sagen, daß ein Körper F eigentlich (fast) in zweifacher Weise eine Gruppe ist: Das Paar $(F, `+\text{'})$ ist eine (abelsche) Gruppe. Aber auch das Paar $(F \setminus \{0\}, `\cdot\text{'})$ ist eine (abelsche) Gruppe.

Beispiele:

(i) (Der einfachste mögliche Körper)

$F = \{0, 1\}$. Man kann sich jetzt überlegen, daß die Operationen '+' und '·' schon gezwungenermaßen festgelegt sind, nämlich:

'+'	0	1
0	0	1
1	1	0

'·'	0	1
0	0	0
1	0	1

(ii) $\mathbf{Q}(\sqrt{2})$ — nämlich die rationalen Zahlen, 'erweitert' durch die irrationale Zahl $\sqrt{2}$. Hier ist

$$\mathbf{Q}(\sqrt{2}) = \{a + \sqrt{2}b : a, b \in \mathbf{Q}\}.$$

Seien $a_1 + \sqrt{2}b_1$ und $a_2 + \sqrt{2}b_2$ zwei Elemente von $\mathbf{Q}(\sqrt{2})$. Dann sind die Additions- und Multiplikationsoperationen durch die folgende Regel definiert:

$$(a_1 + \sqrt{2}b_1) + (a_2 + \sqrt{2}b_2) = (a_1 + a_2) + \sqrt{2}(b_1 + b_2)$$

$$(a_1 + \sqrt{2}b_1) \cdot (a_2 + \sqrt{2}b_2) = (a_1 \cdot a_2 + 2b_1 \cdot b_2) + \sqrt{2}(a_1 \cdot b_2 + a_2 \cdot b_1)$$

Es ist eine einfache Übung, die Axiome nachzuprüfen.

Z.B. (F8):

Sei $a + \sqrt{2}b \in \mathbf{Q}(\sqrt{2})$ mit $a + \sqrt{2}b \neq 0$ $(\equiv 0 + \sqrt{2} \cdot 0)$. Dann ist

$$(a + \sqrt{2}b)^{-1} = \left(\frac{a}{a^2 - 2b^2}\right) + \sqrt{2}\left(\frac{-b}{a^2 - 2b^2}\right)$$

(Bemerkung: Für $a, b \in \mathbf{Q}$ gilt $a^2 - 2b^2 = 0 \;\Rightarrow\; a = b = 0$, da sonst $\sqrt{2} = a/b$ wäre!)

Wir wollen jetzt, einige Sätze über Körper beweisen.

Satz 2.1. *Sei F ein Körper. Für alle $a, b \in F$ gilt:*

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

(ii) $a \cdot (-b) = -(a \cdot b) = (-a) \cdot b$

(iii) $-(-a) = a$

(iv) $(a^{-1})^{-1} = a$, *falls* $a \neq 0$

(v) $(-1) \cdot a = -a$

(vi) $(-a) \cdot (-b) = a \cdot b$

(vii) $a \cdot b = 0 \;\Rightarrow\; a = 0$ *oder* $b = 0$

Beweis. (i)

$$\begin{array}{rrcll} & 0 \cdot a & = & (0+0) \cdot a & (F3) \\ & & = & 0 \cdot a + 0 \cdot a & (F9) \\ \\ \Rightarrow & 0 & = & 0 \cdot a + (-(0 \cdot a)) & (F4) \\ & & = & (0+0) \cdot a + (-(0 \cdot a)) & \text{(oben)} \\ & & = & 0 \cdot a + ((0 \cdot a) + (-(0 \cdot a)) & (F2) \\ & & = & 0 \cdot a & (F4) + (F3) \\ \text{Aber auch} & 0 \cdot a & = & a \cdot 0 & (F5) \end{array}$$

(ii)

$$\begin{array}{rrcll} & (a \cdot (-b)) + a \cdot b & = & a \cdot ((-b) + b) & (F9) \\ & & = & a \cdot 0 & (F4) \\ & & = & 0 & \text{(oben)} \\ \\ \Rightarrow & a \cdot (-b) & = & -(a \cdot b) & \end{array}$$

Die andere Hälfte von (ii) ist ähnlich. u.s.w. . . Ich werde auch (vii) beweisen:

(vii) Sei $a \cdot b = 0$. Angenommen, $a \neq 0$.

$$\Rightarrow \quad \exists\, a^{-1} \in F \text{ mit } a^{-1} \cdot a \;=\; 1 \qquad (F7)$$

$$\begin{array}{rrcll} \Rightarrow & 0 & = & (a^{-1}) \cdot 0 & \text{(oben)} \\ & & = & (a^{-1}) \cdot (a \cdot b) & \text{(Voraussetzung)} \\ & & = & ((a^{-1}) \cdot a) \cdot b & (F6) \\ & & = & 1 \cdot b & (F8) \\ & & = & b & (F7) \end{array}$$

□

Wir haben gesehen, daß $\mathbf{Z}$ ein Ring ist. Noch ein weiteres System, das alle Körperaxiome außer (F8) erfüllt, ist das System der Polynome, nämlich:

Sei T ein beliebiger Ring.

$$T[x] = \left\{ \sum_{i=1}^{n} a_i x^i : a_i \in T,\, n \in \mathbf{N}_0 \right\}$$

ist die Menge der Polynome über *einer* Unbekannten x mit Koeffizienten in T. (Hierbei ist $\mathbf{N}_0 \equiv \mathbf{N} \cup \{0\}$.) Dann ist $T[x]$ auch ein Ring. Zum Beispiel sind $\mathbf{Z}[x]$ oder $\mathbf{Q}[x]$ gängige Ringe, die eine wichtige Rolle in der Computeralgebra spielen.

Warum haben Polynome keine multiplikativen Inversen? Man könnte etwa versuchen,

$$(1-x)^{-1} = \frac{1}{1-x} = 1 + x + x^2 + x^3 + \ldots$$

zu schreiben.[1] Aber wir wollen uns lieber damit abfinden, daß Polynome im allgemeinen *keine* multiplikativen Inversen besitzen.

[1]Für jedes $n \in \mathbf{N}$ gilt:

$$\frac{1 - x^{n+1}}{1-x} = 1 + x + x^2 + \cdots + x^n.$$

Euler hat sich in dieser Richtung viele Gedanken im 18.ten Jahrhundert gemacht. Solche Gedanken führen in der Tat zum Begriff der 'p-adischen' Zahlen.

Polynome haben in der historischen Entwicklung der Mathematik eine sehr große Rolle gespielt. Die moderne Algebra beschäftigt sich auch mit Fragen, die mit der Arithmetik von Polynomen zu tun haben. Für uns sind Polynome wichtig, da wir später (gegen Ende des Semesters) die charakteristischen Polynome von Matrizen untersuchen werden. Eine entscheidende Tatsache ist, daß Polynome mit Koeffizienten aus den reellen Zahlen im allgemeinen nicht faktorisierbar sind in lineare Faktoren mit reellen Koeffizienten. Zum Beispiel das folgende Polynom ist faktorisierbar:

$$2x^2 + 7x + 3 = (2x + 1)(x + 3).$$

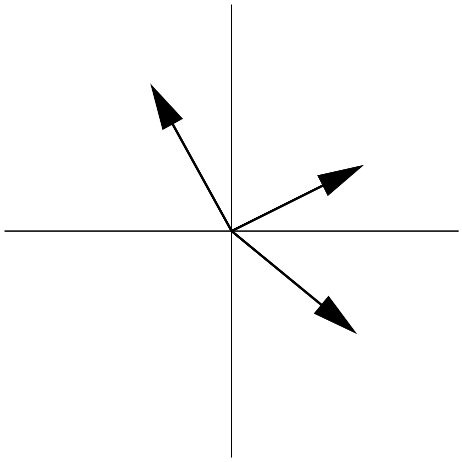
Das heißt, $2x^2 + 7x + 3$ ist faktorisierbar *innerhalb* des Polynomrings $\mathbf{R}[x]$. Andererseits gibt es *keine* reellen Zahlen a, b, c und d mit der Eigenschaft, daß

$$x^2 + 1 = (ax + b)(cx + d).$$

Man sagt deshalb, daß das Zahlensystem $\mathbf{R}$ nicht 'algebraisch abgeschlossen' ist.

Wenn wir aber Polynome mit komplexen Zahlen als Koeffizienten nehmen (d.h. den Ring $\mathbf{C}[x]$), dann ist es doch immer möglich, eine Faktorisierung in lineare Faktoren zu erhalten. Diese Feststellung ist der 'Hauptsatz der Algebra', und wie wir sehen werden, spielt sie eine grosse Rolle in der linearen Algebra.

2.4 Vektorräume



Einige Vektoren in $\mathbf{R}^2$

Nachdem wir jetzt eine allgemeine Definition von 'Zahlen' haben, geht es nun darum, auch eine allgemeine Definition von 'Geometrie' zu finden. Die alte Schulgeometrie führt zu verschiedenen mehr oder weniger obskuren Gedanken — zum Beispiel die nicht-Euklidische Geometrie der hyperbolischen Ebene. Aber für uns ist die Idee des *Vektorraumes* doch die einfachste Art, eine Geometrie zu beschreiben. $\mathbf{R}^n$ heißt in der modernen Mathematik der n-dimensionale *Euklidische* Raum. Die Definition ist vielleicht zunächst etwas abstrakt, aber sie ist gewählt worden, um die spätere Arbeit so einfach wie möglich zu gestalten.

Das Wort 'Geometrie' hat etwas mit *Abständen* zu tun. Man sieht daher, daß es wichtig ist in der Geometrie, verschiedenen Punkten Abstände zuzuordnen.

Definition 11. Sei $a = (a_1, \ldots, a_n)$, $b = (b_1, \ldots, b_n) \in \mathbf{R}^n$. Der Abstand zwischen a und b ist die (nicht-negative) reelle Zahl

$$d(a, b) = \sqrt{(a_1 - b_1)^2 + \cdots + (a_n - b_n)^2}.$$

Bemerkung. Andere Definitionen sind auch möglich. Zum Beispiel die Abstandsfunktion gegeben durch

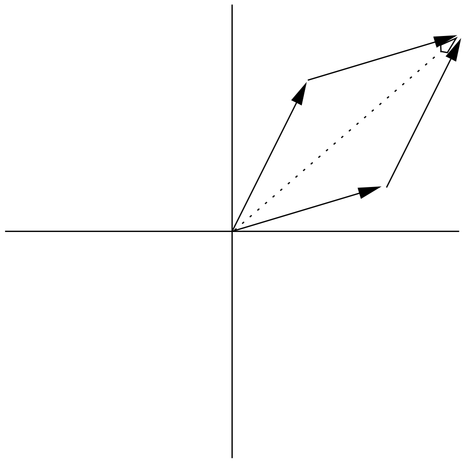
$$d_*(a, b) = \sum_{i=1}^{n} |a_i - b_i|$$

gibt eine etwas andersartige Geometrie. Solche Gedanken führen zum Studium der 'Topologie'. Aber unser Interesse soll auf Vektorräume beschränkt bleiben.

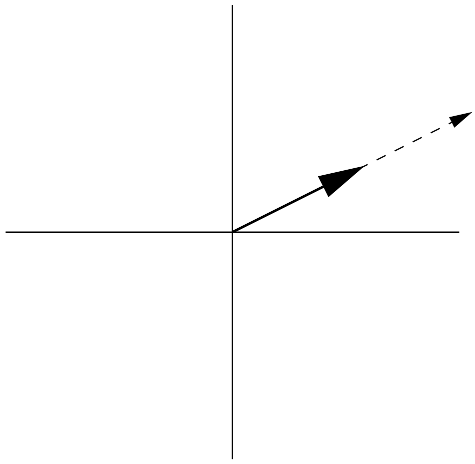
Was sind Vektoren?

- 'Vektoren' sind wie Pfeile — mit 'Schwanz' im Nullpunkt.
- Vektoren haben eine *Richtung* und eine *Länge*.

- Zwei Vektoren können 'addiert' werden.



- Sei $k \in \mathbf{R}$ ein Element des Körpers $\mathbf{R}$ und $\alpha \in \mathbf{R}^n$ ein Vektor. Dann ist $k \cdot \alpha$ auch ein Vektor; nämlich der Vektor mit derselben Richtung wie α, aber die Länge geändert durch den Faktor k.



- Die 'Vektoren' entsprechen auch den 'Punkten' in $\mathbf{R}^n$; indem man nämlich einen Vektor mit dem Punkt auf seiner Spitze identifiziert.

Definition 12. (Vektorräume) Sei F ein Körper und V eine nicht-leere Menge ($V \neq \emptyset$). Die Menge V ist ein Vektorraum über F, falls folgende Axiome erfüllt sind.

(i) Es gibt eine Addition auf V (d.h. '+':$V \times V \to V$) mit

  (V1) $\alpha + \beta = \beta + \alpha$, $\forall\, \alpha, \beta \in V$

  (V2) $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$

  (V3) $\exists$ ein eindeutig bestimmtes Element (genannt der *Nullvektor*), $\mathbf{0} \in V$ mit $\alpha + \mathbf{0} = \alpha$, $\forall\, \alpha \in V$

  (V4) $\forall \alpha \in V$, $\exists$ ein eindeutig bestimmtes Element $(-\alpha) \in V$ mit $\alpha + (-\alpha) = \mathbf{0}$

(ii) Es gibt eine skalare Multiplikation von Elementen (Skalaren) aus F mit Elementen (Vektoren) aus V ('$\cdot$':$F \times V \to V$), so daß

  (V5) $a \cdot (\alpha + \beta) = (a \cdot \alpha) + (a \cdot \beta)$, $\forall\, a \in F$, $\alpha, \beta \in V$

  (V6) $(a + b) \cdot \alpha = (a \cdot \alpha) + (b \cdot \alpha)$

  (V7) $(a \cdot b) \cdot \alpha = a \cdot (b \cdot \alpha)$

  (V8) $1 \cdot \alpha = \alpha$, $\forall\, \alpha \in V$

Beispiele für Vektorräume

(i) $\mathbf{R}^n$ ist ein Vektorraum über $\mathbf{R}$, für alle $n \in \mathbf{N}$. (Der 'normale' n-dimensionale Vektorraum.)

(ii) Jeder Körper ist ein Vektorraum über sich selbst. Man nehme einfach die Additions- und Multiplikationsoperationen des Körpers!

(iii) Lösungen von homogenen linearen Gleichungssystemen.

Zuerst ein einfaches Beispiel: Seien $a, b \in \mathbf{R}$. Wir betrachten die Gleichung

$$ax + by = 0.$$

Sei

$$L \equiv \{(x, y) \in \mathbf{R}^2 : ax + by = 0\}$$

die Menge der Lösungen. Dann ist L ein Vektorraum über $\mathbf{R}$. Z.B. seien (x_1, y_1), $(x_2, y_2) \in L$ zwei Lösungen. Dann gilt

$$ax_1 + by_1 = 0$$

$$ax_2 + by_2 = 0.$$

Aber auch $(x_1, y_1) + (x_2, y_2) \equiv (x_1 + x_2, y_1 + y_2) \in L$. Dies gilt, da

$$a(x_1 + x_2) + b(y_1 + y_2) = (ax_1 + by_1) + (ax_2 + by_2) = 0 + 0 = 0.$$

Skalare Multiplikation: Sei $r \in \mathbf{R}$ beliebig. Dann ist $r(x, y) \equiv (rx, ry) \in L$, da $arx + bry = r(ax + by) = r \cdot 0 = 0$.

Allgemeiner, seien m Gleichungen in n Unbekannten gegeben, etwa:

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & 0 \\ & \vdots & \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = & 0 \end{array}$$

oder kürzer geschrieben

$$\sum_{i=1}^{n} a_{ji}x_i = 0, \quad j = 1, \ldots, m.$$

Sei jetzt

$$L = \left\{ (x_1, \ldots, x_n) \in \mathbf{R}^n : \sum_{i=1}^{n} a_{ji}x_i = 0,\ j = 1, \ldots, m \right\}$$

die Lösungsmenge. Dann ist L wieder ein Vektorraum über $\mathbf{R}$.

Bemerkung. $\sum_{i=1}^{n} a_{ji}x_i = 0$, $j = 1, \ldots, m$ heißt *homogenes* reelles lineares Gleichungssystem. Allgemeiner kann man etwa das folgende System nehmen.

$$\sum_{i=1}^{n} a_{ji}x_i = b_j, \quad j = 1, \ldots, m,$$

mit $b_j \in \mathbf{R}$, $\forall j$, wobei mindestens ein b_j nicht null ist. Dies ist dann ein *inhomogenes* System. Es ist wichtig zu wissen, daß die Lösungsmenge eines inhomogenen Systems *keinen* Vektorraum bildet!

(iv) Sei F ein Körper und $F[x]$ die Menge der Polynome mit Koeffizienten in F. Ein Polynom

$$P(x) = \sum_{i=1}^{n} a_i x^i$$

mit $a_n \neq 0$ hat den *Grad* n. Sei $F_n[x] \subset F[x]$ die Menge der Polynome mit Grad $\leq n$. Dann ist $F_n[x]$ ein Vektorraum über F.

– Addition: Seien $P(x) = \sum_{i=1}^{n} a_i x^i$, $Q(x) = \sum_{i=1}^{n} b_i x^i$ gegeben. Dann ist

$$(P + Q)(x) = \sum_{i=1}^{n} (a_i + b_i) x^i$$

– Skalare Multiplikation: Sei $a \in F$. Dann ist

$$(a \cdot P)(x) = \sum_{i=1}^{n} (a \cdot a_i) x^i.$$

(v) Allgemeiner als nur Polynome: Sei X eine beliebige Menge und F ein Körper. Sei V die Menge aller Abbildungen $X \to F$. Dann ist auch V ein Vektorraum über F.

- Addition: Seien $f, g : X \to F$ gegeben. Dann ist

$$(f + g)(x) \equiv f(x) + g(x),$$

für alle $x \in X$.

- Skalare Multiplikation: Sei $a \in F$. Dann ist

$$(a \cdot f)(x) \equiv a \cdot f(x),$$

für alle $x \in X$.

Dieses letzte Beispiel ist sehr allgemein. Der Vektorraum V enthält *alle* Abbildungen von X nach F, ob stetig oder nicht. (Der Stetigkeitsbegriff aus der Analysis ist natürlich nur sinnvoll bzgl. speziellen Körpern wie etwa **R**.) Welche ‘Dimensionen’ haben diese Vektorräume? Wir werden die formale Definition von Dimension in den nächsten Stunden besprechen. Zunächst aber will ich nur sagen, daß die Beispiele (i) und (iv) Vektorräume von Dimension n sind. ((ii) hat natürlich die Dimension 1.) Wir werden später viel über die Dimension des Raumes (iii) sprechen; sie ist jedenfalls auch endlich. Beispiel (v) hingegen ist ein Vektorraum, der im allgemeinen *keine* endliche Dimension hat. Man könnte sagen, daß V ‘unendlich-dimensional’ sei. Auf jeden Fall sollten Sie daran denken, daß mehrere Sätze, die wir gleich besprechen werden (die Sätze, die durch Induktion über n bewiesen werden), *ungültig* sind für unendlich-dimensionale Vektorräume wie V.

Der folgende Satz ist aber für *alle* Vektorräume gültig.

Satz 2.2. *Sei V ein Vektorraum über dem Körper F, und seien $\alpha, \beta \in V$, $a \in F$. Dann gilt:*

(i) $0 \cdot \alpha = \mathbf{0} = a \cdot \mathbf{0}$

(ii) $a \cdot \alpha = \mathbf{0} \quad \Rightarrow \quad a = 0$ *oder* $\alpha = \mathbf{0}$

(iii) $(-a) \cdot \alpha = a \cdot (-\alpha) = -(a \cdot \alpha)$

Beweis. Eine leichte Übung in der Art von Satz 2.1. □

2.5 Unterräume, Unterkörper, Untergruppen

Zum Beispiel $\mathbf{Q} \subset \mathbf{R} \subset \mathbf{C}$ sind jeweils Unterkörper.

Definition 13. Sei F ein Körper und $E \subset F$ eine Teilmenge. Falls E auch ein Körper ist (mit Addition und Multiplikation aus F), dann heißt E *Unterkörper* von F.

— ähnlich für Gruppen, Ringe, u.s.w.

Definition 14. Sei V ein Vektorraum über einem Körper F und $W \subset V$ eine Teilmenge. Falls W auch ein Vektorraum über F ist (mit Addition und Multiplikation aus V), dann heißt W *Unterraum* von V.

Satz 2.3. *Seien $W \subset V$ und F wie in der Definition. Dann gilt*

$$a \cdot \alpha + b \cdot \beta \in W,$$

$\forall \alpha, \beta \in W$ und $a, b \in F \Leftrightarrow W$ ist Unterraum von V.

Beweis. ‘$\Leftarrow$’ ist klar; es ist daher nur nötig ‘$\Rightarrow$’ zu beweisen. D.h. falls $a\alpha + b\beta \in W$, $\forall \alpha, \beta \in W$ und $a, b \in F$, dann müssen wir zeigen, daß die Axiome für Vektorräume gelten bzgl. W und F.

Ist ‘+’:$W \times W \to W$ und ‘$\cdot$’:$F \times W \to W$? Seien $\alpha, \beta \in W$ beliebige Elemente aus W. Es gilt:

$$\begin{array}{rcll} 1 \cdot \alpha + 1 \cdot \beta & = & \alpha + \beta & (V8) \\ & \in & W. & \text{(Voraussetzung des Satzes)} \\ \text{auch } a \cdot \alpha + \mathbf{0} \cdot \beta & = & a \cdot \alpha + \mathbf{0} & \text{(Satz 2.2)} \\ & = & a \cdot \alpha & (V3) \\ & \in & W. & \text{(Voraussetzung)} \end{array}$$

Das heißt, W ist *abgeschlossen* unter Addition und skalarer Multiplikation. Wir müssen jetzt die Axiome (V1) – (V8) nachprüfen für W. (V1) und (V2) sind offensichtlich klar.

(V3)

$$\begin{aligned} 0\cdot\alpha+0\cdot\beta &= \mathbf{0}+\mathbf{0} && \text{(Satz 2.2)}\\ &= \mathbf{0} && \text{(V3 im Raum V)}\\ &\in W && \text{(Voraussetzung)} \end{aligned}$$

(V4) Sei $\alpha \in W$ gegeben. Dann ist

$$\begin{aligned} \alpha+(-1)\cdot\alpha &= 1\cdot\alpha+(-1)\cdot\alpha \\ &= (1+(-1))\cdot\alpha && (V6),(V8)\\ &= 0\cdot\alpha && (F4)\\ &= \mathbf{0} && \text{(Satz 2.2)}\\ &\in W \end{aligned}$$

$$\Rightarrow \quad (-1)\cdot\alpha = -\alpha$$

Schließlich sind die Axiome (V5) $\rightarrow$ (V8) klar. $\square$

Wie ist die Situation für Gruppen? Zur Erinnerung: eine Gruppe G ist eine Menge mit einer binären Operation '$\cdot$':$G\times G\rightarrow G$, wobei:

(i) $a\cdot b\in G, \quad \forall\, a,b\in G$

(ii) $a\cdot(b\cdot c)=(a\cdot b)\cdot c$

(iii) $\exists\, e\in G$ mit $e\cdot a=a\cdot e=a, \quad \forall\, a\in G$

(iv) $\forall\, a\in G, \exists\, a^{-1}\in G$ mit $a\cdot a^{-1}=e$.

Korollar 2.3.1. Sei $H\subset G$ eine Teilmenge einer Gruppe. H ist eine Untergruppe $\Leftrightarrow a\cdot b^{-1}\in H, \forall\, a,b\in H$.

Beweis. '$\Rightarrow$': Klar.

'$\Leftarrow$': Sei $a\in H\Rightarrow a\cdot a^{-1}=e\in H$. Daher ist $e\cdot a^{-1}=a^{-1}\in H$. Seien $a,b\in H\Rightarrow (b^{-1})^{-1}=b\in H\Rightarrow a\cdot b\in H$. $\square$

Beispiele für Unterräume:

(a) $\mathbf{R}\subset\mathbf{R}^2$ als Vektorräume über $\mathbf{R}$.

(b) Sei wieder $V=\mathbf{R}^2$ (wir identifizieren Punkte in $\mathbf{R}^2$ mit Vektoren in V).

$$W=\{(s,t)\in V : s=k\cdot t, \text{ für ein festes } k\in\mathbf{R}\}$$

ist auch ein Unterraum.

(Dieses zweite Beispiel ist eigentlich identisch mit dem ersten. Bemerken Sie, daß der Nullvektor enthalten sein *muß* in jedem Unterraum.)

(c) Sei V die Menge aller Abbildungen $f:\mathbf{R}\rightarrow\mathbf{R}$ und sei $V_0\subset V$ die Menge der Polynomabbildungen. Dann ist V_0 ein Unterraum von V.

Satz 2.4. *Sei V ein Vektorraum über F und seien X, Y Unterräume von V. Dann ist $X\cap Y$ auch ein Unterraum von V.*

Beweis. Seien $\alpha,\beta\in X\cap Y,\ a,b\in F$.
$\Rightarrow$ $a\alpha+b\beta\in X$ (Satz 2.3), da X Unterraum ist; auch $a\alpha+b\beta\in Y$
$\Rightarrow$ $a\alpha+b\beta\in X\cap Y$
$\Rightarrow$ $X\cap Y$ ist Unterraum (Satz 2.3). $\square$

Kapitel 3

Lineare Unabhängigkeit, Basen, Dimension

3.1 Lineare Unabhängigkeit

Der Begriff 'lineare Unabhängigkeit' spielt eine sehr entscheidende Rolle in der Theorie. Warum? Dadurch werden *Basen* für Vektorräume charakterisiert. Sie geben uns eine einfache und systematische Methode, um (endlich dimensionale) Vektorräume zu beschreiben.

Sei V ein Vektorraum über F, und sei $U = \{\alpha_1, \alpha_2, \ldots, \alpha_n\} \subset V$ eine endliche Teilmenge von V. Die Frage ist nun: Gibt es n Zahlen ('Skalaren') $a_1, a_2, \ldots, a_n \in F$, so daß

$$\sum_{i=1}^{n} a_i \alpha_i = 0 \quad ?$$

Eine triviale Lösung: Man nehme einfach $a_1 = a_2 = \cdots = a_n = 0$. Gibt es andere, *nicht triviale* Lösungen? Falls ja, dann ist die Menge $U \subset V$ linear *abhängig.*

Definition 15. Seien $U \subset V$, und F wie oben. Die Menge U heißt *linear unabhängig*, falls

$$\sum_{i=1}^{n} a_i \alpha_i = 0 \quad \Rightarrow \quad a_i = 0,\ \forall\, i = 1, \ldots, n.$$

Beispiele

- Sei $U = \{\alpha_1, \alpha_2, \alpha_3\}$, wobei $\alpha_1 = (1, 0)$, $\alpha_2 = (0, 1)$ und $\alpha_3 = (1, 1)$. Die Menge $U \subset \mathbf{R}^2$ ist *nicht* linear unabhängig, da $\alpha_3 = \alpha_1 + \alpha_2$. D.h.

$$\alpha_1 + \alpha_2 - \alpha_3 = 1 \cdot \alpha_1 + 1 \cdot \alpha_2 + (-1) \cdot \alpha_3 = 0.$$

- Sei $\mathbf{R}_2[x]$ die Menge der Polynome vom Grad höchstens zwei mit reellen Koeffizienten. Sei $U = \{x, x^2\} \subset \mathbf{R}_2[x]$. Dann ist U linear unabhängig, da

$$ax + bx^2 = 0 \quad \text{(als Polynome!)}$$

$$\Rightarrow a = b = 0.$$

Definition 16. Sei $S = \{\alpha_1, \alpha_2, \ldots, \alpha_n\} \subset V$ eine endliche Teilmenge. Jeden Vektor der Gestalt

$$\alpha = \sum_{i=1}^{n} a_i \alpha_i, \quad \text{wobei } a_i \in F,\ \forall\, i$$

nennt man eine *Linearkombination* der Vektoren in S. Sei $[S] \subset V$ die Menge aller möglichen Linearkombinationen der Vektoren in S. $[S]$ heißt die *lineare Hülle* von S.

Satz 3.1. *$[S]$ ist ein Unterraum von V.*

Beweis. Seien $a, b \in F$, $\alpha, \beta \in [S]$ beliebig. Dann gibt es Skalare c_i, d_i, so daß

$$\alpha = \sum_{i=1}^{n} c_i \alpha_i,$$

$$\beta = \sum_{i=1}^{n} d_i \alpha_i.$$

Daher ist

$$\begin{aligned}
a \cdot \alpha + b \cdot \beta &= a \cdot \left(\sum_{i=1}^{n} c_i \alpha_i \right) + b \cdot \left(\sum_{i=1}^{n} d_i \alpha_i \right) \\
&= \left(\sum_{i=1}^{n} (a \cdot c_i) \alpha_i \right) + \left(\sum_{i=1}^{n} (b \cdot d_i) \alpha_i \right) && \text{(V5), (V7)} \\
&= \sum_{i=1}^{n} (a \cdot c_i + b \cdot d_i) \alpha_i \quad \in \quad [S]
\end{aligned}$$

$\square$

Allgemeiner, sei $T \subset V$ eine beliebige Teilmenge (nicht notwendig endlich!) Wir bezeichnen mit $T_e \subset T$ irgendeine beliebige *endliche* Teilmenge von T. Sei

$$[T] \equiv \bigcup_{\text{alle } T_e} [T_e] \quad : \quad \text{die lineare Hülle von } T$$

Korollar 3.1.1. *Auch wenn T unendlich ist, ist $[T] \subset V$ immer noch ein Unterraum.*

Beispiele

- Sei $\mathbf{Q}[x]$ die Menge aller Polynome über x mit rationalen Zahlen als Koeffizienten; betrachtet als Vektorraum über $\mathbf{Q}$. Ein typischer Vektor in $\mathbf{Q}[x]$ ist dann einfach ein Polynom

$$\sum_{i=0}^{n} a_i x^i = a_0 + a_1 x + \cdots + a_n x^n, \quad a_i \in \mathbf{Q}, \forall i.$$

  $T \subset V$ sei die Menge aller solcher Polynome mit Grad 2 oder weniger. (D.h. $n \leq 2$ hier.) Dann ist auch T ein Untervektorraum von V.

- Andererseits ist die Menge $\mathbf{Z}[x]$ aller Polynome über x mit *Integer* Koeffizienten kein Unterraum von $\mathbf{Q}[x]$. Z.B. $1 + x \in \mathbf{Z}[x]$ und $\frac{1}{2} \in \mathbf{Q}$, aber $\frac{1}{2} \cdot (1 + x) \notin \mathbf{Z}[x]$.

Definition 17. $T \subset V$ wie oben. T heißt Erzeugendensystem von $[T]$. Falls $[T] = V$, dann heißt T ein Erzeugendensystem des Raumes V. V heißt *endlich erzeugt*, falls ein endliches Erzeugendensystem existiert.

Beispiel

- $\mathbf{R}^n$ ist endlich erzeugt, und zwar sei $S = \{\epsilon_1, \ldots, \epsilon_n\}$, wobei

$$\begin{aligned}
\epsilon_1 &= (1, 0, 0, \ldots, 0) \\
\epsilon_2 &= (0, 1, 0, \ldots, 0) \\
&\vdots \\
\epsilon_n &= (0, 0, 0, \ldots, 1)
\end{aligned}$$

  Es ist jetzt offensichtlich, daß $\mathbf{R}^n = [S]$.

Der folgende Satz beschreibt die Beziehung zwischen ‘linearer Unabhängigkeit’ und ‘Erzeugendensystemen’.

Satz 3.2. *Sei $\alpha = \sum_{i=1}^{n} a_i \alpha_i$ eine Linearkombination der Vektoren in $S = \{\alpha_1, \ldots, \alpha_n\}$. Diese Darstellung von α in V durch Vektoren aus S ist eindeutig $\Leftrightarrow$ S ist linear unabhängig.*

Beweis. **‘$\Rightarrow$’:** Sei $\alpha = \sum_{i=1}^{n} a_i \alpha_i$. Dann ist

$$\begin{aligned}
\alpha = \alpha + 0 &= \sum_{i=1}^{n} a_i \alpha_i + \sum_{i=1}^{n} b_i \alpha_i \\
&= \sum_{i=1}^{n} (a_i + b_i) \alpha_i
\end{aligned}$$

Aber falls irgendein $b_i \neq 0$, dann wäre die Darstellung von α nicht eindeutig.[1]

[1] Falls $a_i + b_i = a_i \;\Rightarrow\; (-a_i) + a_i + b_i = (-a_i) + a_i \;\Rightarrow\; b_i = 0$.

'⇐': Um einen Widerspruch zu erzeugen, sei

$$\alpha = \sum_{i=1}^{n} a_i\alpha_i = \sum_{i=1}^{n} c_i\alpha_i,$$

mit $a_i \neq c_i$ für mindestens ein i.

$$\begin{aligned} \Rightarrow \quad 0 &= \alpha - \alpha \\ &= \sum_{i=1}^{n}(a_i - c_i)\alpha_i \end{aligned}$$

$\Rightarrow S$ ist nicht linear unabhängig; ein Widerspruch!

□

Definition 18. Sei $S \subset V$ linear unabhängig mit $[S] = V$. Dann heißt S eine *Basis* für V.

Beispiel

- Der Körper
$$\mathbf{Q}(\sqrt{2}) = \{a + b\sqrt{2} : a, b \in \mathbf{Q}\}$$
kann als *Vektorraum* über $\mathbf{Q}$ aufgefaßt werden. Eine Basis ist
$$S = \{1, \sqrt{2}\}.$$
Ist S tatsächlich linear unabhängig? Sei $a \cdot 1 + b \cdot \sqrt{2} = 0$. Falls $b \neq 0$ dann wäre $-\frac{a}{b} = \sqrt{2}$, mit $a, b \in \mathbf{Q}$. Unmöglich. $\Rightarrow b = 0 \Rightarrow a = 0$ auch.

Lemma. *Sei $S = \{\alpha_1, \ldots, \alpha_n\}$ linear* ab*hängig.* $\Rightarrow \exists\, j \in \{1, \ldots, n\}$ *und* $a_i \in F$, $i = 1, \ldots, j-1, j+1, \ldots, n$ *mit*

$$\alpha_j = \sum_{i \neq j} a_i\alpha_i.$$

Beweis. S linear abhängig $\Rightarrow \exists\, b_i \in F$, $i = 1, \ldots, n$ mit

$$\sum_{i=1}^{n} b_i\alpha_i = 0$$

und $b_j \neq 0$ für mindestens ein $j \in \{1, \ldots, n\}$.

$$\begin{aligned} \Rightarrow \quad b_j\alpha_j &= -\sum_{i \neq j} b_i\alpha_i \\ \Rightarrow \quad \alpha_j &= \sum_{i \neq j} \left(-\frac{b_i}{b_j}\right)\alpha_i \end{aligned}$$

□

Korollar 3.2.1. *(S und V wie oben): Sei*

$$S' = \{\alpha_1, \ldots, \alpha_{j-1}, \alpha_{j+1}, \ldots, \alpha_n, \}.$$

Dann ist $[S] = [S']$.

Satz 3.3. *Sei V endlich erzeugt.* $\Rightarrow \exists$ *eine endliche Basis für V.*

Beweis. Endlich erzeugt $\Rightarrow \exists$ eine endliche Teilmenge $S \subset V$ mit $[S] = V$. Unter allen solchen S gibt es sicherlich welche mit einer minimalen Anzahl von Elementen. Wir nehmen an, daß S minimal ist in diesem Sinne. Falls S linear unabhängig ist, dann sind wir fertig. Andernfalls ist S linear abhängig und nach unserem Korollar muß eine kleinere Menge $S' \subset S$ existieren mit $[S'] = [S] = V$; ein Widerspruch. □

Bemerkungen:

- Sei $S \subset V$ eine endliche Teilmenge. Um die bisherigen Ergebnisse zusammenzufassen: S ist eine Basis für $V \Leftrightarrow S$ ist eine *maximale* linear unabhängige Teilmenge $\Leftrightarrow S$ ist ein *minimales* Erzeugendensystem für V.
- $[S]$ wird oft auch 'Span(S)' genannt.

Lemma. *(Steinitz Austauschlemma): Sei $S = \{\alpha_1, \ldots, \alpha_n\}$ eine Basis für V, und sei $\beta = \sum_{i=1}^n a_i\alpha_i$ mit $\beta \neq 0$. Dann existiert ein $j \in \{1, \ldots, n\}$ so daß*

$$S' = \{\alpha_1, \ldots, \alpha_{j-1}, \beta, \alpha_{j+1}, \ldots, \alpha_n\}$$

auch eine Basis für V ist.

Beweis. Da $\beta \neq 0$, $\exists\ 1 \leq j \leq n$ mit $a_j \neq 0$.

$$\Rightarrow \quad \alpha_j = a_j^{-1} \cdot \beta + \sum_{i \neq j} \left(-\frac{a_i}{a_j}\right) \cdot \alpha_i.$$

Sei nun $\gamma = \sum_{i=1}^n b_i\alpha_i \in [S] = V$ ein beliebiger Vektor.

$$\begin{aligned} \Rightarrow \quad \gamma &= \sum_{i \neq j} b_i\alpha_i + b_j \cdot \left(a_j^{-1}\beta + \sum_{i \neq j} \left(-\frac{a_i}{a_j}\right) \cdot \alpha_i\right) \\ &= \frac{b_j}{a_j} \cdot \beta + \sum_{i \neq j} \left(b_i - \frac{b_j a_i}{a_j}\right) \cdot \alpha_i \end{aligned}$$

$\Rightarrow \gamma \in [S']$. Da aber γ beliebig war, gilt $V \subset [S']$. D.h. $V = [S']$.

Ist S' linear unabhängig? Sei

$$\begin{aligned} 0 &= b\beta + \sum_{i \neq j} b_i\alpha_i \\ &= b \cdot \left(\sum_{i=1}^n a_i\alpha_i\right) + \sum_{i \neq j} b_i\alpha_i \\ &= \sum_{i=1}^n (ba_i + b_i) \cdot \alpha_i \qquad \text{(wobei } b_j \equiv 0) \end{aligned}$$

Aber S ist linear unabhängig. $\Rightarrow ba_i + b_i = 0$, $\forall i = 1, \ldots, n$. Insbesondere gilt $ba_j + b_j = 0$. Aber $a_j \neq 0$. $\Rightarrow$ $b = 0 \Rightarrow b_i = 0$, $\forall i$. □

Satz 3.4 (Austauschsatz). *Sei $S = \{\alpha_1, \ldots, \alpha_n\}$ eine Basis für V. Sei weiter $T = \{\beta_1, \ldots, \beta_m\} \subset V$ linear unabhängig. Dann ist $m \leq n$. Es ist möglich, durch Umnumerierung der Elemente aus S zu erreichen, daß*

$$U = \{\beta_1, \ldots, \beta_m, \alpha_{m+1}, \ldots, \alpha_n\}$$

eine Basis für V ist.

Beweis. Induktion nach m.

1. Fall: $m = 0$: Dabei ist $T = \emptyset$ und $U = S$; daher ist nichts zu beweisen.

2. Fall: Angenommen, der Satz sei wahr für den Fall $m - 1$ (mit $m \geq 1$). Wir zeigen, daß er dann auch wahr ist für den Fall m. Nun, $T' = \{\beta_1, \ldots, \beta_{m-1}\}$ ist linear unabhängig, da T linear unabhängig ist. Daher ist (laut induktiver Hypothese und nach möglicher Umnumerierung)

$$U' = \{\beta_1, \ldots, \beta_{m-1}, \alpha_m, \ldots, \alpha_n\}$$

eine Basis für V. Kann es sein, daß $n = m - 1$? D.h. daß T' auch eine Basis für V ist? Falls ja, dann wäre insbesondere

$$\beta_m = \sum_{i=1}^{m-1} b_i\beta_i.$$

D.h. T wäre nicht linear unabhängig. Widerspruch. Daher muß $n \geq m$.

U' ist aber eine Basis für V. Folglich gibt es

$$b_1, \ldots, b_{m-1}, a_m, \ldots, a_n \in F$$

mit

$$\beta_m = b_1\beta_1 + \cdots + b_{m-1}\beta_{m-1} + a_m\alpha_m + \cdots + a_n\alpha_n.$$

Kann es sein, daß $a_m = \cdots = a_n = 0$? Nein, sonst wäre T linear abhängig. Durch mögliche Umnumerierung können wir annehmen, daß $a_m \neq 0$. Jetzt brauchen wir nur Steinitz' Austauschlemma zu berücksichtigen.

□

Korollar 3.4.1 (Basisergänzungssatz). *Sei V endlich erzeugt und $S \subset V$ eine linear unabhängige Teilmenge. $\Rightarrow \exists$ eine Basis $B \subset V$ mit $S \subset B$.*

Beweis. Sei $[S] \neq V$. Da V endlich erzeugt ist, existiert eine Basis bestehend (etwa) aus n Vektoren. Dann $\exists\, \alpha \in V - [S]$ und $S' = S \cup \{\alpha\}$ ist linear unabhängig. Ist $[S'] \neq V$? Falls nein, dann $\exists \alpha' \in V - S'$, und S' kann ebenfalls ergänzt werden. Nach höchstens n solchen Ergänzungen ist man fertig. □

Korollar 3.4.2. *Sei V endlich erzeugt und $U \subset V$ ein Unterraum. Dann ist U auch endlich erzeugt, und zwar durch höchstens so viele Vektoren wie nötig sind, um V zu erzeugen.*

Satz 3.5. *Angenommen, V besitzt eine Basis mit n Elementen. Dann hat jede Basis genau n Elemente.*

Beweis. Seien $\{\alpha_1, \ldots, \alpha_n\}$, $\{\beta_1, \ldots, \beta_m\}$ Basen. Nach Satz 3.4 gilt sowohl $m \leq n$ als auch $m \geq n$. □

Definition 19. n heißt die *Dimension* von V, geschrieben $\dim(V)$. Falls V endlich erzeugt ist, dann heißt V endlich dimensional.

Definition 20. Sei V ein Vektorraum und X, Y Unterräume. $X + Y \equiv [X \cup Y]$ heißt die *Summe* von X und Y. Falls $X \cap Y = \{0\}$ und $X + Y = V$ dann schreibt man auch $X \oplus Y = V$, die *direkte Summe*.

Satz 3.6 (Dimensionsformel). [2] *Seien $X, Y \subset V$ Unterräume, wobei V endlich dimensional ist. $\Rightarrow$ $dim(X + Y) = dim(X) + dim(Y) - dim(X \cap Y)$.*

Beweis. Sei $S = \{\alpha_1, \ldots, \alpha_n\}$ eine Basis für $X \cap Y$. (Korollar zu Satz 3.4) Nach dem Basisergänzungssatz existieren $T = \{\beta_1, \ldots, \beta_m\}$ und $U = \{\gamma_1, \ldots, \gamma_p\}$ mit $S \cup T$ eine Basis für X und $S \cup U$ eine Basis für Y. Behauptung: $S \cup T \cup U$ ist eine Basis für $X + Y$.
Nun, $X + Y = [S \cup T \cup U]$ ist klar. Wie ist es mit der linearen Unabhängigkeit? Sei

$$\begin{aligned} 0 &= \sum_{i=1}^{n} a_i\alpha_i + \sum_{j=1}^{m} b_j\beta_j + \sum_{k=1}^{p} c_k\gamma_k \\ &= \alpha + \beta + \gamma \quad \text{etwa, wobei } \alpha_i \in S, \beta_j \in T, \gamma_k \in U. \\ \Rightarrow \quad \gamma &= -\alpha - \beta = \sum_{i=1}^{n}(-a_i)\alpha_i + \sum_{j=1}^{m}(-b_j)\beta_j \in X \end{aligned}$$

Aber auch $\gamma \in Y$. D.h. $\gamma \in X \cap Y$. Da aber $S \cup T$ eine Basis für X ist, ist die Darstellung von γ eindeutig. $\Rightarrow$ $b_j = 0$, $\forall j$. Durch symmetrische Argumentation gilt auch $c_k = 0$, $\forall k$. Aber S ist linear unabhängig. $\Rightarrow$ $a_i = 0$, $\forall i$. D.h. $S \cup T \cup U$ ist insgesamt linear unabhängig.
Nun, $\dim(X) = n + m$, $\dim(Y) = n + p$, $\dim(X \cap Y) = n$.

$$\Rightarrow \quad \dim(X + Y) = n + m + p = \dim(X) + \dim(Y) - \dim(X \cap Y).$$

□

Korollar 3.6.1. *$dim(X \oplus Y) = dim(X) + dim(Y)$.*

Satz 3.7. *Sei $X \subset V$ Unterraum eines endlich dimensionalen Vektorraums. $\Rightarrow \exists$ ein Unterraum $Y \subset V$ mit $X \oplus Y = V$.*

Beweis. triviale Folgerung aus dem Basisergänzungssatz und der Definition von '$\oplus$'. □

3.2 Eine andere mögliche Definitionen von 'Dimension'

Sei M eine Menge mit einer 'Topologie'. D.h. insbesondere, es ist sinnvoll zu sagen, wann Gebiete in M 'zusammenhängend' sind. Um die Idee der Dimension hier zu definieren, gibt es die Möglichkeit, wie folgt vorzugehen.

(0) Ein isolierter Punkt P ist 0-dimensional. etwa: $M_0 \equiv \{P\}$

(1) Eine 'zusammenhängende' Menge M_1 ist 1-dimensional, falls $M_1 - M_0$ nicht mehr zusammenhängend ist, für eine 0-dimensionale Teilmenge $M_0 \subset M_1$.

(n) Eine 'zusammenhängende' Menge M_n ist n-dimensional, falls $M_n - M_{n-1}$ nicht mehr zusammenhängend ist, für eine $(n-1)$-dimensionale Teilmenge $M_n \subset M_{n-1}$. u.s.w...

[2] Es gibt noch eine zweite Dimensionsformel, die etwas später vorkommen wird.

Kapitel 4

Mengen, Abbildungen, Linearität

4.1 Etwas über die Mengentheorie

In der modernen Mathematik wird oft über 'Abbildungen' gesprochen. Z.B. 'Matrizen'—d.h. Sammlungen von Zahlen, angeordnet in Zeilen und Spalten—spielen eine große Rolle in der linearen Algebra. Aber es wäre sicherlich falsch zu behaupten, daß solche rechteckige Zahlenmuster im Mittelpunkt der Mathematik stehen. Für uns sind Matrizen nur interessant, weil sie eine besonders einfache Beschreibung einer bestimmten Klasse von Abbildungen (lineare Abbildungen) ermöglichen.

Es ist daher angebracht, zu fragen, warum Abbildungen wichtig sind; wäre es nicht sinnvoller, die mathematischen Objekte an sich zu studieren, statt nur die Abbildungen von Objekten untereinander zu betrachten? Aber diese Frage ist ganz leicht zu beantworten! Sie haben beispielsweise in der Grundschule schon gelernt, daß

$$2 + 2 = 4.$$

Was bedeutet das? Heißt es etwa, daß zwei Äpfel plus zwei Äpfel gleich vier Äpfel sind? oder Birnen? oder was? Was sind zwei Äpfel plus zwei Birnen? Kann man sagen, daß vier von irgendetwas entsteht? u.s.w. Wie Sie sehen, hat die Mathematik eher etwas mit *allgemeinen Eigenschaften* von Objekten zu tun. Wenn man $2 + 2 = 4$ sagt, dann wird stillschweigend angenommen, daß es sich um *wesentlich gleiche* Objekte handelt. Oder anders gesagt, es ist uns egal, ob Äpfel oder Birnen gerade betrachtet werden, da für die Mathematik, die zwei Fälle *im wesentlichlichen gleich* sind; es gibt nämlich eine Abbildung dazwischen, die die mathematischen Eigenschaften, die wir gerade betrachten, unverändert läßt. Es ist daher vernünftig, zu fragen: wann sind zwei Objekte eigentlich 'wesentlich gleich'?

Beispiele:

- Nehmen wir als erstes Beispiel den Körper mit zwei Elementen: $\{0, 1\}$. Ich kann mir aber auch einen weiteren Körper mit zwei Elementen vorstellen, nämlich $\{a, b\}$, wobei das, was mit 'a' bezeichnet wird, das Identitätselement der Addition ist, und 'b' ist das Identitätselement der Multiplikation. Sind diese zwei Körper dann wirklich verschieden? Eigentlich nicht. Ich habe nur verschiedene Zeichen benutzt, um eine *einzige Sache* zu beschreiben. Es gibt daher eine *Struktur-erhaltende* Abbildung zwischen den Körpern $\{0, 1\} \leftrightarrow \{a, b\}$.
- Der Vektorraum $\mathbf{R}^2$ mit Basis $\epsilon_1 = (1, 0)$, $\epsilon_2 = (0, 1)$. Wir haben gesehen, daß ein beliebiger Vektor in $\mathbf{R}^2$ eindeutig darstellbar ist als $x\epsilon_1 + y\epsilon_2$. Wir können noch eine andere Basis für $\mathbf{R}^2$ nehmen, etwa $\nu_1 = (1, 0)$, $\nu_2 = (1, 1)$. Dann ist für jedes Paar von reellen Zahlen (x, y) auch $x\nu_1 + y\nu_2$ ein Vektor in $\mathbf{R}^2$. Im Gedanken bedeutet '$(x, y) \in \mathbf{R}^2$' zuerst ein bestimmter Vektor im 'ϵ'-System, danach ist vielleicht '(x, y)' in der Vorstellung doch ein Vektor im 'ν'-System. Aber *im wesentlichen* handelt es sich immer um einen einzigen Vektorraum!

Definition 21. Seien X, Y zwei Mengen und sei $f : X \to Y$ eine Abbildung.
(i) Falls $f(x) \neq f(x')$ für alle $x \neq x'$ in X, dann heißt f eine *Injektion.*
(ii) Falls $f(X) = Y$; d.h. $\forall y \in Y$, $\exists x \in X$ mit $f(x) = y$, dann heißt f eine *Surjektion.*
(iii) Falls f injektiv und surjektiv ist, dann heißt f eine *Bijektion.*

Bemerkung. Gegeben $f : X \to Y$, und $U \subset X$, $V \subset Y$ jeweils Teilmengen, dann heißt die Menge

$$f(U) = \{y \in Y : \exists x \in U \text{ mit } f(x) = y\}$$

das *Bild* von U unter f. Weiter heißt

$$f^{-1}(V) = \{x \in X : f(x) \in V\}$$

das *Urbild* von V.

Satz 4.1. *Sei $f : X \to Y$ eine Bijektion. Dann existiert eine Abbildung $g : Y \to X$ mit $g(f(x)) = x$ und $f(g(y)) = y$ für alle $x \in X$ und $y \in Y$.*

Beweis. g wird durch die folgende Regel gegeben. Sei $y \in Y$ ein beliebiges Element. Da $f : X \to Y$ surjektiv ist, $\exists x \in X$ mit $f(x) = y$. Da f auch injektiv ist, gibt es nur ein einziges solches x. Die Regel lautet daher: $g(y) = x$. □

Bemerkung: Man schreibt eigentlich f^{-1} statt g.

'Gleichheit' Seien X und Y zwei Mengen. Im Rahmen der Mengentheorie sind X und Y *im wesentlichen gleich*, falls eine Bijektion $f : X \to Y$ existiert.

Nach dieser Auffassung sind dann etwa die zwei Mengen

$$\{\text{Hund}, \text{Katze}\} \quad \text{und} \quad \{\text{Salz}, \text{Pfeffer}\}$$

im wesentlichen gleich. Nun, im Gegensatz zu Körpern oder Vektorräumen, haben Mengen praktisch *keine* Struktur. Die Gleichheit unter Mengen wird nur durch die *Anzahl* der Elemente bestimmt. Diese Idee ist vor etwa 100 Jahren von Georg Cantor weiter entwickelt worden.

Nach seiner Auffassung ist es auch sinnvoll zu sagen, daß die Menge X *mindestens* soviele Elemente wie die Menge Y hat, falls eine *Surjektion* $f : X \to Y$ existiert.

Zum Beispiel ist unser $f : \{0,1\} \to \{a,b\}$ auch eine Surjektion. Daher hat (nach der Cantor'schen Philosophie) die Menge $\{0,1\}$ mindestens soviele Elemente wie die Menge $\{a,b\}$. Andererseits gibt es sicherlich *keine* Surjektion $f : \{0\} \to \{0,1\}$. D.h. die Menge $\{0\}$ ist *echt kleiner* als die Menge $\{0,1\}$.

Falls wir zwei Mengen X und Y haben, und wir wissen, daß *keine* Surjektion $f : X \to Y$ existiert, dann werden wir sagen, daß X *weniger* Elemente als Y hat. Man sagt auch, '$\text{card}(X) < \text{card}(Y)$'. D.h. die *Kardinalität* von X ist kleiner als die Kardinalität von Y. Falls eine Bijektion $f : X \to Y$ existiert, dann ist $\text{card}(X) = \text{card}(Y)$.

Satz 4.2. *Seien X, Y Mengen. Es existiert eine Surjektion $f : X \to Y$ dann und nur dann wenn eine Injektion $g : Y \to X$ existiert.*

Beweis. '$\Rightarrow$' Sei $f : X \to Y$ surjektiv. $\forall y \in Y$, $\exists x \in X$ mit $f(x) = y$. Nehme irgendein bestimmtes solches x. Definiere $g(y) = x$. Offensichtlich ist $g : Y \to X$ dann ein Injektion.

'$\Leftarrow$' Sei $g : Y \to X$ eine Injektion. Sei $y_0 \in Y$ ein beliebiges — aber dann festgelegtes — Element. Für alle $x \in X$ sei

$$f(x) = \begin{cases} \text{das Element } y \in Y, \text{ mit } g(y) = x, \text{ falls } x \in g(Y); \\ y_0 \text{ sonst.} \end{cases}$$

□

Satz 4.3 (Bernstein). *Seien X, Y zwei Mengen. Falls eine* Injektion $f : X \to Y$, *und eine* Surjektion $g : X \to Y$ *existieren, dann existiert auch eine* Bijektion $h : X \to Y$.

Beispiele:

(i) Für endliche Mengen ist die Sache klar; alles hängt von der Anzahl der Elemente ab.

(ii) Die Mengen $\mathbf{N}$ und $\mathbf{Z}$ (die natürlichen und die Integerzahlen). Es gilt $\text{card}(\mathbf{N}) = \text{card}(\mathbf{Z})$. Warum? Weil

$$f(n) = \begin{cases} 2n, & \text{falls } n \in \mathbf{N}; \\ -2n+1 & \text{sonst.} \end{cases}$$

eine Bijektion $f : \mathbf{Z} \to \mathbf{N}$ ist.

(iii) Wie ist es mit $\mathbf{N}$ und $\mathbf{Q}$?

- Natürlich gibt es eine Injektion $f : \mathbf{N} \to \mathbf{Q}$. Man nehme einfach $f(n) = n$, $\forall n \in \mathbf{N}$. (die Identitätsabbildung)
- Gibt es eine *Surjektion* $g : \mathbf{N} \to \mathbf{Q}$? Ja. Sie brauchen nur den Pfeilen in der Tabelle zu folgen:

$$
\begin{array}{ccccccccccccccc}
\frac{-4}{1} & \leftarrow & \frac{-3}{1} & & \frac{-2}{1} & \leftarrow & \frac{-1}{1} & & \frac{0}{1} & \rightarrow & \frac{1}{1} & & \frac{2}{1} & \rightarrow & \frac{3}{1} \\
 & \searrow & & \nwarrow & & \searrow & & \nwarrow & & \swarrow & & \nearrow & & \swarrow & \\
\frac{-4}{2} & & \frac{-3}{2} & & \frac{-2}{2} & & \frac{-1}{2} & & \frac{0}{2} & & \frac{1}{2} & & \frac{2}{2} & & \frac{3}{2} \\
 & & & \searrow & & \nwarrow & & \searrow & & \nearrow & & \swarrow & & & \\
\frac{-4}{3} & & \frac{-3}{3} & & \frac{-2}{3} & & \frac{-1}{3} & & \frac{0}{3} & & \frac{1}{3} & & \frac{2}{3} & & \frac{3}{3} \\
 & & & & & \searrow & & \nwarrow & & \swarrow & & & & & \\
\frac{-4}{4} & & \frac{-3}{4} & & \frac{-2}{4} & & \frac{-1}{4} & & \frac{0}{4} & & \frac{1}{4} & & \frac{2}{4} & & \frac{3}{4}
\end{array}
$$

Das heißt auch card($\mathbf{N}$)=card($\mathbf{Q}$).

Kann es sein, daß alle unendlich großen Mengen gleich groß sind?
Nein!

(iv) Es gilt card($\mathbf{N}$) $<$ card($\mathbf{R}$).

Sie haben dieses letzte Ergebnis bestimmt in der Schule gesehen. Hier nur eine Skizze: Angenommen es gilt doch card($\mathbf{N}$) $\geq$ card($\mathbf{R}$). Das würde heißen, daß eine Surjektion $\mathbf{N} \to \mathbf{R}$ existieren würde. Mit anderen Worten, es wäre möglich, die reellen Zahlen vollständig 'aufzuzählen':

$$\mathbf{R} = \{r_1, r_2, r_3, \dots\}.$$

Für jedes $i \in \mathbf{N}$ sei

$$r_i = d_{in} d_{i\,n-1} \dots d_{i1}, x_{i1} x_{i2} x_{i3} \dots,$$

geschrieben als Dezimalbruchzahl. (Zum Beispiel $\pi = 3,14159\dots$ D.h. wenn π etwa die fünfte reelle Zahl wäre, dann wäre $n = 1$, $d_{51} = 3$, $x_{51} = 1$, $x_{52} = 4$, $x_{53} = 1$, $x_{54} = 5$, $x_{55} = 9$, u.s.w.) Aber es gibt viele Zahlen, die nicht in dieses Schema passen. Z.B. die Zahl $0, y_1 y_2 y_3 \dots$, wobei $y_j \neq x_{jj}$, $\forall\, j \in \mathbf{N}$. Dies zeigt, daß eine solche Surjektion $\mathbf{N} \to \mathbf{R}$ unmöglich ist.

Definition 22. Sei X irgendeine Menge. $\wp(X)$ ist die Menge aller Teilmengen von X (genannt *Potenzmenge*).

Satz 4.4. *Sei $X \neq \emptyset$. Dann ist card($\wp(X)$) $>$ card(X).*

Beweis. Andernfalls wäre card($\wp(X)$) $\leq$ card(X). D.h. es würde eine Surjektion $f : X \to \wp(X)$ existieren. Was würde dies bedeuten?

Für jedes Element $x \in X$, wäre $f(x)$ dann eine Teilmenge von X. Wir wollen jetzt eine besonders boshafte Teilmenge $B \subset X$ betrachten.

$$B = \{x \in X : x \notin f(x)\}.$$

Nun, da f eine Surjektion ist, $\exists\ y \in X$ mit $f(y) = B$. Ist $y \in B$? Nein! Denn das stünde in Widerspruch zur Definition von B. Ist $y \notin B$? Auch nein! Widerspruch. $\Rightarrow$ card($\wp(X)$) $>$ card(X). □

Dieser Beweis ist eigentlich sehr heikel, aber er ist doch (immer noch) in Ordnung. Wir bewegen uns hier am Rande des logischen Chaos, wie das folgende Ergebnis zeigt.

Definition 23. Eine Menge X heißt *gut*, falls $X \notin X$. Sonst ist X *schlecht.*

Beispiel: Sei $U \equiv$ die Menge aller Mengen. Dann ist U offensichtlich schlecht!

Russel's Paradoxon *Seien G die Menge aller guten Mengen und S die Menge aller schlechten Mengen. Daher ist $U = G \cup S$. Frage: ist G gut oder schlecht?*

Antwort:

Angenommen G sei gut. $\Rightarrow G \in G \Rightarrow G$ ist schlecht. Widerspruch.

Angenommen G sei schlecht. $\Rightarrow G \notin G \Rightarrow G$ ist gut. Widerspruch.

Das heißt, *alles* führt zum Widerspruch!

Schlußfolgerung *Vermeiden Sie die Redewendungen: 'die Menge aller Mengen mit der Eigenschaft... u.s.w.'. Als allgemeine Regel könnte man Folgendes sagen. Die 'konkreten' mathematischen Objekte, wie* $\mathbf{R}$, $\mathbf{N}$, $\mathbf{R}^n$, *u.s.w. sind Mengen. Sei M eine Menge. Dann sind sicherlich alle Teilmengen von M auch Mengen. Auch die Menge aller Teilmengen von M ist eine Menge. Es ist andererseits heikel zu sagen: 'Sei M die 'Menge' aller Mengen mit der Eigenschaft $\wp$', für eine bestimmte allgemeine Eigenschaft $\wp$.*

4.2 Mengen mit Struktur

(Z.B. Körper, Vektorräume, u.s.w.)

Definition 24. Sei M eine Menge. Eine *Relation* auf M ist eine Teilmenge $K \subset M \times M$. Sei etwa $(a, b) \in K$. Man schreibt dazu aKb.

Beispiele:

- Die Relation '=' (Gleichheit) ist: $\{(a, a) : a \in M\} \subset M \times M$.
- '$\sim$' $\subset \mathbf{R} \times \mathbf{R}$, wobei $x \sim y \Leftrightarrow x - y \in \mathbf{Z}$ ist auch eine Relation.

Eine Relation heißt

(i) reflexiv, falls aKa, $\forall a \in M$

(ii) symmetrisch, falls $aKb \Rightarrow bKa$

(iii) transitiv, falls aKb und $bKc \Rightarrow aKc$

Falls eine Relation K auf M die Bedingungen (i) — (iii) erfüllt, dann heißt K eine *Äquivalenzrelation* auf M. Sei nun

$$[a] = \{b \in M : (a, b) \in K\}$$

Die Menge $[a]$ heißt die *Äquivalenzklasse* von a in M.

Satz 4.5. *Sei K eine Äquivalenzrelation auf M und seien $a, b \in M$. Dann ist entweder $[a] = [b]$ oder $[a] \cap [b] = \emptyset$.*

Beweis. Angenommen $[a] \cap [b] \neq \emptyset$. Sei $c \in [a] \cap [b]$ und sei x ein beliebiges Element von $[a]$. $\Rightarrow xKa$ und aKc und $cKb \Rightarrow xKc$ und $cKb \Rightarrow xKb \Rightarrow x \in [b] \Rightarrow [a] \subset [b]$. Durch symmetrische Argumentation können wir auch zeigen, daß $[b] \subset [a]$. □

Seien nun X und Y Mengen mit ähnlichen Strukturen. X und Y sind 'gleich', falls eine Bijektion $f : X \to Y$ existiert, die die *Struktur* der Mengen erhält.

Nehmen wir als Beispiel die Gruppen. Seien X, Y Gruppen.

Definition 25. Eine Abbildung $f : X \to Y$ heißt *Homomorphismus*, falls $f(a \cdot b) = f(a) \cdot f(b)$, $\forall\, a, b \in X$.

Sei jetzt f ein Homomorphismus. Man sagt f sei:

$$\begin{array}{rcl} \text{injektiv} & \leftrightarrow & \textit{Monomorphismus} \\ \text{surjektiv} & \leftrightarrow & \textit{Epimorphismus} \\ \text{bijektiv} & \leftrightarrow & \textit{Isomorphismus} \end{array}$$

Falls $X = Y$, dann heißt f ein *Endomorphismus*. Falls f gleichzeitig Isomorphismus und Endomorphismus ist, dann heißt f ein *Automorphismus*.

Zwei Gruppen sind 'gleich', falls ein Isomorphismus zwischen den Gruppen existiert. Man sagt auch, daß sie zueinander *isomorph* sind.

Satz 4.6. *Seien X und Y Gruppen und $f : X \to Y$ ein Homomorphismus. Dann ist $f(1) = 1$ und $f(a^{-1}) = (f(a))^{-1}$, für alle $a \in X$.*

Beweis. Sei $a \in X \Rightarrow 1 \cdot a = a$

$$\begin{array}{rrcl} \Rightarrow & f(a) & = & f(1) \cdot f(a) \\ & \text{Sei}\ \ y & = & (f(a))^{-1} \in Y \\ \Rightarrow & 1 = f(a) \cdot y & = & (f(1) \cdot f(a)) \cdot y \\ & & = & f(1) \cdot \left(f(a) \cdot f(a)^{-1}\right) \\ & & = & f(1) \cdot 1 = f(1) \end{array}$$

Weiter gilt $1 = a \cdot a^{-1}$ in X.

$$\Rightarrow \quad 1 = f(1) = f(a \cdot a^{-1}) = f(a) \cdot f(a^{-1}).$$

□

Diese Begriffe lassen sich auch leicht auf Körper und Vektorräume übertragen. (Man braucht nur daran zu denken, daß solche Mengen sich (fast) wie Gruppen verhalten unter '+' und '·'.)

Definition 26. Seien X und Y Körper. Eine Abbildung $f : X \to Y$ heißt *Homomorphismus*, falls für alle $a, b \in X$ gilt:

$$f(a+b) = f(a) + f(b)$$
$$\text{und} \quad f(a \cdot b) = f(a) \cdot f(b).$$

Seien U und V Vektorräume über einem Körper F. $f : U \to V$ heißt Homomorphismus, falls gilt

$$f(\alpha + \beta) = f(\alpha) + f(\beta) \ \text{ und } \ f(a \cdot \alpha) = a \cdot f(\alpha)$$

für alle $\alpha, \beta \in X$, $a \in F$. Eine solche Abbildung $f : U \to V$ zwischen Vektorräumen heißt F-*linear* (oder einfach eine *lineare Abbildung*).

Korollar 4.6.1. *Für lineare Abbildungen gilt*

$$f(0) = 0 \quad \textit{und} \quad f(-\alpha) = -f(\alpha).$$

Satz 4.7. *Seien X und Y Vektorräume über einem Körper F und sei $f : X \to Y$ eine lineare Abbildung. Sei ferner eine endliche Teilmenge $S \subset X$ gegeben. Es gilt: $f(S)$ ist linear unabhängig $\Rightarrow$ S ist linear unabhängig in X. Seien weiter $U \subset X$, $V \subset Y$ jeweils Unterräume. Dann sind $f(U) \subset Y$ und $f^{-1}(V) \subset X$ auch Unterräume.*[1]

Beweis. Sei $0 = \sum_{i=1}^n a_i \alpha_i$, wobei $a_i \in F$, $\alpha_i \in S$, $\forall\, i$.

$$\begin{aligned}
\Rightarrow \quad 0 &= f(0) && \text{(aus Satz 4.6)} \\
&= f\Big(\sum_{i=1}^n a_i \alpha_i\Big) \\
&= \sum_{i=1}^n f(a_i \alpha_i) \\
&= \sum_{i=1}^n a_i f(\alpha_i), && \text{wobei } f(\alpha_i) \in f(S),\ \forall\, i. \\
\Rightarrow \quad a_i &= 0,\ \forall\, i,\ \text{ da } f(S) \text{ linear unabhängig.}
\end{aligned}$$

Sei jetzt $a, b \in F$, $\alpha, \beta \in U$.

Dann ist $a\alpha + b\beta \in U$. (Satz 2.3.) Folglich

$$af(\alpha) + bf(\beta) = f(a\alpha + b\beta) \in f(U).$$

$\Rightarrow$ $f(U)$ ist ein Unterraum von Y. (Wieder Satz 2.3.)

Umgekehrt, seien $\gamma, \delta \in f^{-1}(V)$.

$$\Rightarrow \quad f(a\gamma + b\delta) = af(\gamma) + bf(\delta) \in V,$$

da $f(\gamma)$, $f(\delta) \in V$ und V ein Unterraum von Y ist.

$$\Rightarrow \quad a\gamma + b\delta \in f^{-1}(V)$$
$$\Rightarrow \quad f^{-1}(V) \quad \text{ist ein Unterraum. (Satz 2.3.)}$$

□

Satz 4.8. *Sei $f : X \to Y$ ein Isomorphismus zwischen Vektorräumen. Dann ist $f^{-1} : Y \to X$ auch ein Homomorphismus.*

Beweis. Seien $\alpha, \beta \in Y$. Dann sind $f^{-1}(\alpha)$, $f^{-1}(\beta)$ eindeutig vorgegebene Elemente aus X. Auch $f^{-1}(\alpha) + f^{-1}(\beta)$ ist ein bestimmtes Element in X. Es gilt

$$\begin{aligned}
f(f^{-1}(\alpha) + f^{-1}(\beta)) &= f(f^{-1}(\alpha)) + f(f^{-1}(\beta)) \quad \text{(da } f \text{ Homomorphismus)} \\
&= \alpha + \beta
\end{aligned}$$

D.h. $f^{-1}(\alpha + \beta) = f^{-1}(\alpha) + f^{-1}(\beta)$. Auch

$$\begin{aligned}
f(a \cdot f^{-1}(\alpha)) &= a \cdot f(f^{-1}(\alpha)) \\
&= a \cdot \alpha \\
\Rightarrow \quad f^{-1}(a \cdot \alpha) &= a \cdot f^{-1}(\alpha)
\end{aligned}$$

□

[1] Gegeben eine Abbildung $g : A \to B$ und $C \subset B$ eine Teilmenge, dann ist

$$g^{-1}(C) \equiv \{a \in A : g(a) \in C\}.$$

Bemerkung f^{-1} *heißt die zu* f inverse Abbildung.

Definition 27. Seien X und Y Vektorräume über einem Körper F. X und Y heißen zueinander *isomorph*, falls ein Isomorphismus

$$f : X \to Y$$

existiert. Man schreibt dazu $X \cong Y$.

(D.h. X und Y sind *im wesentlichen* gleich.)

Kapitel 5

Lineare Abbildungen

Wir haben gesehen, daß es möglich ist, die endlichen Mengen vollständig zu 'klassifizieren': zwei endliche Mengen sind 'gleich' genau dann, wenn sie dieselbe Anzahl von Elementen enthalten. Der folgende Satz gibt eine ähnlich einfache Klassifizierung von endlich dimensionalen Vektorräumen.

Satz 5.1. *Seien X und Y endlich erzeugte Vektorräume über F. Es gilt $X \cong Y \Leftrightarrow dim(X) = dim(Y)$.*

Beweis. "$\Rightarrow$" Sei $f : X \to Y$ ein Isomorphismus und sei $B = \{\alpha_1, \dots, \alpha_n\}$ eine Basis für Y. Dann ist

$$f^{-1}(B) = \{f^{-1}(\alpha_1), \dots, f^{-1}(\alpha_n)\}$$

eine Basis für X, da

1. $f^{-1}(B)$ ist linear unabhängig, wegen Satz 4.8.
2. $[f^{-1}(B)] = X$, da

$$\begin{aligned} [f^{-1}(B)] &= f^{-1}([B]) \qquad (f^{-1} \text{ ist ein Homomorphismus}) \\ &= f^{-1}(Y) \qquad (B \text{ ist eine Basis von } Y) \\ &= X \qquad (f \text{ ist eine Bijektion.}) \end{aligned}$$

Folglich ist $\dim(X) = \dim(Y)$. (Satz 3.5)

"$\Leftarrow$" Sei $\dim(X) = \dim(Y)$. Daher $\exists$ Basen

$$A = \{\alpha_1, \dots, \alpha_n\} \subset X,$$
$$B = \{\beta_1, \dots, \beta_n\} \subset Y.$$

Definiere $f : A \to B$ durch die Regel $f(\alpha_i) = \beta_i$, $\forall\, i$. Die Abbildung f kann erweitert werden zu einer Abbildung $f : X \to Y$ wie folgt. Sei $\alpha \in X$ beliebig, mit der Darstellung $\alpha = \sum_{i=1}^n a_i\alpha_i$. Wir *definieren* $f(\alpha)$ dann als den Vektor $f(\alpha) = \sum_{i=1}^n a_i \cdot f(\alpha_i)$. Ist die dadurch definierte Abbildung $f : X \to Y$ ein Isomorphismus?

1. f ist ein Homomorphismus. (Dies ist eine triviale Folgerung der Definition von f.)
2. f ist eine Bijektion. (Nach Satz 3.2 (Eindeutigkeit der Darstellung) ist f eine Injektion. Aus $f(A) = B$ und $[B] = Y$ folgt, daß f eine Surjektion ist.)

□

Insbesondere ist für jeden Körper F das Kartesische-Produkt F^n ein n-dimensionaler Vektorraum über F. Eine Basis für F^n ist $\{\epsilon_1, \dots, \epsilon_n\}$, wobei

$$\epsilon_i = (0, \dots, 0, 1, 0, \dots, 0).$$

D.h. die i-te Komponente von ϵ_i ist 1, sonst sind alle Komponenten Null.

Seien $\alpha = (\alpha_1, \dots, \alpha_n)$ und $\beta = (\beta_1, \dots, \beta_n) \in F^n$, und $a \in F$. Vektoraddition und skalare Multiplikation werden auf folgende Weise festgelegt.

$$\alpha + \beta = (\alpha_1 + \beta_1, \dots, \alpha_n + \beta_n),$$
$$a \cdot \alpha = (a \cdot \alpha_1, \dots, a \cdot \alpha_n)$$

Korollar 5.1.1. *Alle n-dimensionalen Vektorräume über F sind zu F^n isomorph.*

Das heißt, F^n ist im wesentlichen der *einzige* n-dimensionale F-Vektorraum. Man sagt, daß F^n der *kanonische* n-dimensionale F-Vektorraum ist. Die Basis

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

ist die *kanonische Basis* für diesen Raum.

5.1 Allgemeine Eigenschaften von linearen Abbildungen

Definition 28. Sei $f : V \to W$ eine lineare Abbildung.

$$ker(f) \equiv \{\alpha \in V : f(\alpha) = 0\}$$

heißt der *Kern* von f.

$$im(f) \equiv \{\beta \in W : \exists\, \alpha \in V \text{ mit } f(\alpha) = \beta\}$$

heißt das *Bild* von f.

D.h. $ker(f) = f^{-1}(\{0\})$ und $im(f) = f(V)$.

Satz 5.2. *$ker(f) \subset V$ und $im(f) \subset W$ sind beide Unterräume.*

Beweis. Seien $a, b \in F$, $\alpha, \beta \in ker(f)$ beliebig.

$$\begin{aligned} \Rightarrow\ f(a\alpha + b\beta) &= af(\alpha) + bf(\beta) \\ &= a \cdot 0 + b \cdot 0 \\ &= 0 + 0 = 0 \end{aligned}$$

$\Rightarrow\ a\alpha + b\beta \in ker(f)\ \Rightarrow\ ker(f)$ ist Unterraum.

Seien jetzt $\gamma, \delta \in im(f)$ beliebig.
$\Rightarrow\ \exists\, \alpha, \beta \in V$ mit $f(\alpha) = \gamma$, $f(\beta) = \delta$.

$$\begin{aligned} \Rightarrow\ f(a\alpha + b\beta) &= af(\alpha) + bf(\beta) \\ &= a\gamma + b\delta \in im(f) \end{aligned}$$

$\Rightarrow\ im(f)$ ist auch ein Unterraum. □

Satz 5.3. *$f : V \to W$ ist ein Monomorphismus $\Leftrightarrow$ $ker(f) = \{0\}$.*

Beweis. "$\Rightarrow$" Sei $\alpha \in ker(f)$. D.h. $f(\alpha) = 0$. Aber nach Proposition 4 ist $f(0) = 0$. D.h. $f(\alpha) = f(0)$. Da f eine Injektion ist $\Rightarrow\ \alpha = 0$.

"$\Leftarrow$" Angenommen $f(\alpha) = f(\beta)$ in V. Dann gilt

$$\begin{aligned} f(\alpha) = f(\beta) &\Rightarrow f(\alpha) - f(\beta) = 0 \\ &\Rightarrow f(\alpha - \beta) = 0 \\ &\Rightarrow \alpha - \beta \in ker(f) = \{0\} \\ &\Rightarrow \alpha - \beta = 0 \\ &\Rightarrow \alpha = \beta \end{aligned}$$

Folglich ist f ein Monomorphismus. □

Korollar 5.3.1. *Der Satz gilt allgemeiner für Monomorphismen bei Gruppen.*

Wann sind Abbildungen zwischen Vektorräumen linear? Der folgende Satz gibt eine Bedingung ähnlich wie die Bedingung in Satz 2.3.

Satz 5.4. *Sei $f : V \to W$ eine Abbildung. Dann gilt: f ist linear $\Leftrightarrow$ $f(a\alpha + b\beta) = af(\alpha) + bf(\beta)$, für alle möglichen $a, b \in F$ und $\alpha, \beta \in V$.*

Beweis. "$\Rightarrow$" $f(a\alpha + b\beta) = f(a\alpha) + f(b\beta) = af(\alpha) + bf(\beta)$.

"$\Leftarrow$" Zu zeigen:

- $f(\alpha + \beta) = f(\alpha) + f(\beta)$.
- $f(a\alpha) = af(\alpha)$.

Aber es gilt (i):

$$\begin{aligned} f(\alpha + \beta) &= f(1 \cdot \alpha + 1 \cdot \beta) \\ &= 1 \cdot f(\alpha) + 1 \cdot f(\beta) = f(\alpha) + f(\beta) \end{aligned}$$

und auch (ii):

$$\begin{aligned} f(a \cdot \alpha) &= f(a\alpha + 0) \\ &= f(a\alpha + 0 \cdot 0) \\ &= af(\alpha) + 0 \cdot f(0) \\ &= af(\alpha) + 0 \\ &= af(\alpha) \end{aligned}$$

□

Satz 5.5. *Seien V, W Vektorräume über F und $B = \{\alpha_1, \ldots \alpha_n\}$ eine Basis für V. Sei weiter $S = \{\beta_1, \ldots \beta_n\}$ eine beliebige Menge von n Vektoren ($n = dim(V)$) in W. Dann $\exists$ eine eindeutig bestimmte lineare Abbildung $f : V \to W$ mit der Eigenschaft, daß $f(\alpha_i) = \beta_i$, $\forall i$.*

Beweis. f wird wie folgt definiert. Sei $\alpha = \sum_{i=1}^n a_i\alpha_i \in V$ ein beliebiger Vektor. Dann ist $f(\alpha) \equiv \sum_{i=1}^n a_i\beta_i$. Ist $f(\alpha_i) = \beta_i$, $\forall i$ mit dieser Definition? Ja, da die Darstellung von Vektoren durch eine Basis eindeutig ist.

Ist f linear? Seien $\gamma, \delta \in V$ und $a, b \in F$ beliebig.

$$\Rightarrow \; f(a\gamma + b\delta) = f\left(a \cdot \sum_{i=1}^n c_i\alpha_i + b \cdot \sum_{i=1}^n d_i\alpha_i\right)$$

für geeignete $c_i, d_i \in F$.

$$\begin{aligned} &= f\left(\sum_{i=1}^n (ac_i + bd_i) \cdot \alpha_i\right) \\ &= \sum_{i=1}^n (ac_i + bd_i) \cdot \beta_i \\ &= a\left(\sum_{i=1}^n c_i\beta_i\right) + b\left(\sum_{i=1}^n d_i\beta_i\right) \\ &= a \cdot f(\gamma) + b \cdot f(\delta). \end{aligned}$$

Ist f eindeutig bestimmt?

Sei auch $g : V \to W$ eine lineare Abbildung mit $g(\alpha_i) = \beta_i$, $\forall i$. Sei $\alpha = \sum_{i=1}^n a_i\alpha_i \in V$ beliebig.

$$\begin{aligned} \Rightarrow \; g(\alpha) &= g\left(\sum_{i=1}^n a_i\alpha_i\right) && \text{(Definition von } \alpha\text{)} \\ &= \sum_{i=1}^n a_i g(\alpha_i) && \text{(Linearität von } g\text{)} \\ &= \sum_{i=1}^n a_i\beta_i && \text{(Voraussetzung)} \\ &= f(\alpha) && \text{(Definition von } f\text{)} \end{aligned}$$

□

Beispiele: Automorphismen $f : \mathbf{R}^2 \to \mathbf{R}^2$.

Sei $B = \{(1,0),(0,1)\} \equiv \{\epsilon_1, \epsilon_2\}$ die *kanonische* Basis für $\mathbf{R}^2$.

1. (Spiegelungen): Sei $f : \mathbf{R}^2 \to \mathbf{R}^2$ gegeben durch: $f(\epsilon_1) = \epsilon_1$, $f(\epsilon_2) = -\epsilon_2$. Die Abbildung f ist dann eine 'Spiegelung' der Euklidischen Ebene an die x-Achse.

2. (Drehungen): Jetzt sei $\theta \in [0, 2\pi)$ und f gegeben durch[1]

$$\begin{aligned} f(\epsilon_1) &= (\cos\theta, \sin\theta), \\ f(\epsilon_2) &= (\cos(\theta + \pi/2), \sin(\theta + \pi/2)) \\ &= (-\sin\theta, \cos\theta) \end{aligned}$$

[1] Es gilt:

$$\cos(\theta + \phi) = \cos\theta\cos\phi - \sin\theta\sin\phi$$
$$\sin(\theta + \phi) = \sin\theta\cos\phi + \cos\theta\sin\phi$$

Sei jetzt $\alpha = (a_1, a_2)$ ein beliebiger Vektor in $\mathbf{R}^2$. Behauptung: $f(\alpha)$ ist ein Vektor mit derselben Länge wie α, dessen Richtung um einen Winkel θ vom ursprünglichen Vektor α gedreht ist.

Wir betrachten die zwei Fälle $a_1 = 0$ und $a_1 \neq 0$ getrennt.

$a_1 = 0$: Dann ist $\alpha = (0, a_2)$ und $f(\alpha) = (-a_2 \sin\theta, a_2 \cos\theta)$. Die Länge von $f(\alpha)$ ist

$$\sqrt{(-a_2 \sin\theta)^2 + (a_2 \cos\theta)^2} = |a_2|$$

und dies ist die Länge von α. Daß es sich um eine Drehung handelt, folgt aus der üblichen geometrischen Definition der trigonometrischen Funktionen.

$a_1 \neq 0$: Das heißt $\alpha \neq 0$. Seien $\phi = \arctan(a_2/a_1)$ und $r = \left(a_1^2 + a_2^2\right)^{1/2}$. Dann ist $\alpha = (r \cdot \cos\phi, r \cdot \sin\phi)$. Warum?

$$a_1 = r \cdot \cos\phi \text{ und } a_2 = r \cdot \sin\phi \;\Leftrightarrow\; \frac{a_2}{a_1} = \tan\phi \;\Leftrightarrow\; \phi = \arctan\left(\frac{a_2}{a_1}\right).$$

Nun, es gilt $\sin^2\vartheta + \cos^2\vartheta = 1$, $\forall\,\vartheta$.

$$\begin{aligned} \Rightarrow\; a_1^2 + a_2^2 &= r^2\cos^2\phi + r^2\sin^2\phi \\ &= r^2\left(\cos^2\phi + \sin^2\phi\right) \\ &= r^2 \\ \Rightarrow\; r &= \left(a_1^2 + a_2^2\right)^{\frac{1}{2}} \end{aligned}$$

Was ist eine Drehung von α um den Winkel θ?

$$\alpha = (a_1, a_2) = a_1\epsilon_1 + a_2\epsilon_2 = (r\cos\phi, r\sin\phi).$$

Eine Drehung um θ ist dann:

$$\begin{aligned} f(\alpha) &= (r\cos(\phi+\theta), r\sin(\phi+\theta)) \\ &= (r(\cos\phi\cos\theta - \sin\phi\sin\theta), r(\sin\phi\cos\theta + \cos\phi\sin\theta)) \\ &= ((r\cos\phi)\cos\theta - (r\sin\phi)\sin\theta, (r\cos\phi)\sin\theta + (r\sin\phi)\cos\theta) \\ &= (a_1\cos\theta - a_2\sin\theta, a_1\sin\theta + a_2\cos\theta) \\ &= a_1(\cos\theta, \sin\theta) + a_2(-\sin\theta, \cos\theta) \\ &= a_1 f(\epsilon_1) + a_2 f(\epsilon_2) \end{aligned}$$

Das heißt, die lineare Abbildung f gibt tatsächlich eine Drehung[2] von α um den Winkel θ.

3. (Dehnungen) Seien $c, d \in \mathbf{R}$ vorgegeben. $f : R^2 \to R^2$ wird jetzt gegeben durch $f(\epsilon_1) = c \cdot \epsilon_1$, $f(\epsilon_2) = d \cdot \epsilon_2$. D.h. wenn $\alpha = (a_1, a_2)$, dann ist $f(\alpha) = (ca_1, da_2)$.

Bemerkungen:

1. Dehnungen ändern die Längen von Vektoren; Spiegelungen und Drehungen ändern diese Längen *nicht*.
2. Bei Dehnungen und Spiegelungen gibt es Vektoren $\alpha \neq 0$, die einfach in sich selbst abgebildet werden, vielleicht mit geänderter Länge ($f(\alpha) = k \cdot \alpha$.) Bei Drehungen gibt es im allgemeinen solche "Eigenvektoren" nicht.
3. Alle diese linearen Abbildungen können auch analog im n-dimensionalen Raum $\mathbf{R}^n$ definiert werden.

5.2 Die Darstellung von linearen Abbildungen durch Matrizen

Seien V und W zwei F-Vektorräume. Angenommen $\dim(V) = n$, $\dim(W) = m$ und $f : V \to W$ sei linear. Sei $\{\alpha_1, \ldots, \alpha_n\}$ eine vorgegebene Basis für V, und $\{\gamma_1, \ldots, \gamma_m\}$ sei eine vorgegebene Basis für W.

('Normalerweise' ist $V = F^n$ und $W = F^m$, und die vorgegebenen Basen sind einfach die kanonischen Basen.)

[2] Die Menge der (2-dimensionalen) Drehungen bildet eine Gruppe—die (spezielle) orthogonale Gruppe, $SO(2)$. (Man schreibt auch $SO(2;\mathbf{R})$, um zu zeigen, daß R das zugrundeliegende Zahlensystem ist.) Seien $f, g \in SO(2)$; sei etwa f eine Drehung um den Winkel θ und g um den Winkel ϕ. Dann ist das Produkt fg die Drehung um den Winkel $\theta + \phi$. Es ist wichtig zu wissen, daß $SO(2)$ eine *abelsche* Gruppe ist, aber $SO(n)$ (die Drehungen des n-dimensionalen Raumes $\mathbf{R}^n$) ist *nicht* abelsch, für $n > 2$.

Schreiben wir jetzt $f(\alpha_i) = \beta_i$, $i = 1, \ldots, n$. Dadurch sind n bestimmte Vektoren in dem m-dimensionalen Raum W gegeben. Nach Satz 5.5 ist f durch Kenntnis der n Vektoren $\{\beta_1, \ldots, \beta_n\}$ eindeutig bestimmt. Für jedes i sei

$$\beta_i = \sum_{j=1}^{m} a_{ji}\gamma_j.$$

Dies ergibt eine $m \times n$ Matrix von Elementen aus dem Körper F:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Korollar 5.5.1. *Die Matrix A wird eindeutig bestimmt durch die Abbildung $f : V \to W$ (bei vorgegebenen Basen) und umgekehrt.*

Sei jetzt $\alpha = \sum_{i=1}^{n} b_i\alpha_i \in V$ ein beliebiger Vektor. Dann ist

$$\begin{aligned} f(\alpha) &= \sum_{i=1}^{n} b_i f(\alpha_i) \\ &= \sum_{i=1}^{n} b_i\beta_i = \sum_{i=1}^{n} b_i \left(\sum_{j=1}^{m} a_{ji}\gamma_j\right) \\ &= \sum_{i=1}^{n} \left(\sum_{j=1}^{m} b_i a_{ji}\right) \cdot \gamma_j. \end{aligned}$$

Nun, V ist 'eigentlich' nichts anderes als F^n, und W ist im wesentlichen F^m. Es ist daher vernünftig, zu sagen, daß die vorgegebenen Basen $\{\alpha_1, \ldots, \alpha_n\}$ und $\{\gamma_1, \ldots, \gamma_m\}$ eigentlich auch als die kanonischen Basen für F^n und F^m dargestellt werden können. Diese kanonischen Basisvektorn können als 'Spaltenvektoren' dargestellt werden.

$$\alpha_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad \alpha_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \quad \alpha_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Mit dieser Art der Darstellung folgt, daß

$$\alpha = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

Was ist $f(\alpha)$? Nichts anderes als

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix}$$

wobei $f(\alpha) = \sum_{j=1}^{m} c_j\beta_j$ und

$$c_j = \sum_{i=1}^{n} a_{ji}b_i, \quad j = 1, \ldots, m.$$

So wird die Matrizenmultiplikation erklärt.

Definition 29. Seien $A = (a_{ij})$, $B = (b_{jk})$ und $C = (c_{st})$ Matrizen, wobei $1 \leq i \leq n$, $1 \leq j \leq m$, $1 \leq k \leq p$, $1 \leq s \leq n$ und $1 \leq t \leq p$. Es gilt $A \cdot B = C$ mit $c_{st} \equiv \sum_{u=1}^{m} a_{su}b_{ut}$ für alle s und t.

Um zu sehen, daß alles richtig stimmt, wollen wir jetzt eine Verknüpfung von zwei Abbildungen betrachten. Und zwar seien U, V und W Vektorräume über einem Körper F mit Basen $\{\alpha_1, \ldots, \alpha_n\}$, $\{\beta_1, \ldots, \beta_m\}$ und $\{\gamma_1, \ldots, \gamma_p\}$. Gegeben seien lineare Abbildungen

$$U \xrightarrow{f} V \xrightarrow{g} W.$$

Diese Abbildungen können einzeln dargestellt werden: Seien $A = (a_{ij})$ und $B = (b_{kl})$ die Matrizen, die die Abbildung f bzw. g darstellen bzgl. der vorgegebenen Basen. Sei nun

$$\xi = \sum_{i=1}^{n} x_i \alpha_i = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

ein beliebiger Vektor in U. Dann ist

$$g(f(\xi)) = g\left(f\left(\sum_{i=1}^{n} x_i \alpha_i\right)\right) = g\left(\sum_{i=1}^{n}\sum_{j=1}^{m} a_{ji} x_i \beta_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{p} b_{kj} a_{ji} x_i \gamma_k.$$

Dies läßt sich alles viel einfacher schreiben, wenn man die Matrizenschreibweise benutzt. Es gilt nämlich: $f(g(\xi)) = B \cdot A \cdot \xi$. D.h.

$$BA\xi = \begin{pmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & & \vdots \\ b_{p1} & \cdots & b_{pm} \end{pmatrix} \cdot \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Kapitel 6

Matrizenumformungen

Sei

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

eine $m \times n$ Matrix mit Elementen aus einem Körper F. Wir haben gesehen, daß A als eine lineare Abbildung $f : F^n \to F^m$ aufgefaßt werden kann. Es ist auch möglich, A auf eine ganz andere—einfachere—Weise zu sehen. Man könnte sagen, daß A eigentlich nichts anderes ist, als eine Sammlung von m Zeilen, die jeweils n Elemente enthalten. Oder man könnte auch sagen, daß A nichts anderes ist, als eine Sammlung von n Spalten, die jeweils m Elemente enthalten. D.h.

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = \begin{matrix} (a_{11}, \ldots, a_{1n}) \\ \vdots \\ (a_{m1}, \ldots, a_{mn}) \end{matrix} = \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} \cdots \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

Nun, wir können die einzelnen Zeilen und Spalten für sich alleine betrachten. Zum Beispiel ist die i-te Zeile $(a_{i1}, \ldots, a_{in})$. D.h. die i-te Zeile ist eigentlich ein Punkt (d.h. ein Vektor) in F^n. So betrachtet ist die ganze Matrix A nichts anderes als eine Sammlung von m *Zeilenvektoren* aus F^n. Und wenn wir uns auf die Spalten konzentrieren, dann sehen wir, daß A eigentlich auch eine Sammlung von n *Spaltenvektoren* ist. Jeder solcher Spaltenvektor ist ein Vektor des Raumes F^m. Wir werden zunächst die Zeilenvektoren betrachten und dann später die Spaltenvektoren heranziehen.

6.1 Zeilenvektoren

Für jedes i zwischen 1 und m sei $\alpha_i = (a_{i1}, \ldots, a_{in})$. D.h. die Matrix A ist die (geordnete) Menge $S = \{\alpha_1, \ldots, \alpha_m\} \subset F^n$. Die lineare Hülle $[S] \subset F^n$ heißt der *Zeilenraum* von A.

Es ist wichtig, immer daran zu denken, daß die Menge S *geordnet* ist. Seien $i \neq j$ zwei verschiedene Integerzahlen zwischen 1 und m. Sei etwa $i < j$. Dann wäre

$$S = \{\alpha_1, \ldots, \alpha_i, \ldots, \alpha_j \ldots, \alpha_m\}.$$

Durch einfaches Vertauschen der Elemente α_i und α_j erhalten wir eine neue geordnete Menge

$$S_{ij} = \{\alpha_1, \ldots, \alpha_j, \ldots, \alpha_i \ldots, \alpha_m\}.$$

Das heißt, die Matrix A wird durch eine *elementare Zeilenoperation* umgeformt:

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jn} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \longrightarrow \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jn} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

Es gibt zwei weitere elementare Zeilenoperationen:

$$S_i(a) = \{\alpha_1, \ldots, \alpha_{i-1}, a \cdot \alpha_i, \alpha_{i+1}, \ldots, \alpha_m\}$$

D.h.

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \longrightarrow \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a \cdot a_{i1} & \cdots & a \cdot a_{in} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

und

$$S_{ij}(c) = \{\alpha_1, \ldots, \alpha_{i-1}, \alpha_i + c \cdot \alpha_j, \alpha_{i+1}, \ldots, \alpha_m\}$$

D.h.

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \longrightarrow \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} + c \cdot a_{j1} & \cdots & a_{in} + c \cdot a_{jn} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

Definition 30. Sei A eine $m \times n$ Matrix. Die drei elementaren Zeilenoperationen sind S_{ij}, $S_i(a)$ und $S_{ij}(c)$. S_{ij} ist das Vertauschen der Zeilen $i \neq j$ von S. $S_i(a)$ ist die Multiplikation der i-ten Zeile mit dem Skalar $(a \neq 0)$. Schließlich ist $S_{ij}(c)$ die Addition von c-mal der j-ten Zeile zur i-ten Zeile ($i \neq j$ und $c \neq 0$).

Satz 6.1. *(I)* $[S] = [S_{ij}] = [S_i(a)] = [S_{ij}(c)]$.

(II) Die vier Mengen S, S_{ij}, $S_i(a)$ und $S_{ij}(c)$ sind entweder alle linear unabhängig, oder alle linear abhängig.

Beweis. (I) 1. $[S] = [S_{ij}]$ ist trivial.

2. $[S] = [S_i(a)]$: Sei $\beta = \sum_{j=1}^n b_j \alpha_j \in [S]$. Aber

$$\beta = \left(\sum_{j \neq i} b_j \alpha_j \right) + \left(\frac{b_i}{a} \right) \cdot (a \cdot \alpha_i) \in [S_i(a)]$$

$\Rightarrow [S] \subset [S_i(a)]$. Die andere Richtung ist auch klar.

3. $[S] = [S_{ij}(c)]$: Sei β wie oben.

$$\beta = \left(\sum_{\substack{k \neq i \\ k \neq j}} b_k \beta_k \right) + (b_j - b_i c)\beta_j + b_i(\beta_i + c\beta_j) \ \in \ [S_{ij}(c)]$$

$\Rightarrow \quad [S] \subset [S_{ij}(c)]$. Auch $[S] = [S_{ij}(c)]$.

(II) Sei $\dim([S]) = d = \dim([S_{ij}]) = \dim([S_i(a)]) = \dim([S_{ij}(c)])$. Aber jede Menge hat genau n Elemente. Falls $d = n$, dann ist jede Menge linear unabhängig (sonst gäbe es eine Basis mit weniger als d Elementen). Falls $d < n$, dann ist jede Menge linear abhängig.

□

Definition 31. Die $m \times n$ Matrix $A = (a_{ij})$ ist in *Zeilenstufenform*, falls $\exists\ 0 \leq r \leq m$ und $1 \leq j_1 < j_2 < \cdots < j_r \leq m$ mit $a_{ij_i} = 1$, $\forall\, i = 1, \ldots, r$, und $a_{st} = 0$, $\forall\, s, t$ mit $t < j_s$ oder $s > j_r$.

$$A = \begin{pmatrix} \cdots 1 & a_{1\,j_1+1} & & \cdots & & \cdots & a_{1n} \\ 0 & 1 & a_{2\,j_2+1} & & & \cdots & a_{2n} \\ 0 & 0 & 1 & a_{3\,j_3+1} & & \cdots & a_{3n} \\ \vdots & & & \ddots & & & \vdots \\ 0 & \cdots & \cdots & 0 & 1 & a_{r\,j_r+1} \ \cdots & a_{rn} \\ 0 & & & 0 & & & 0 \\ \vdots & & & \vdots & & & \vdots \end{pmatrix}$$

Zum Beispiel die folgende Matrix A_1 ist in Zeilenstufenform, während A_2 *nicht* in Zeilenstufenform ist.

$$A_1 = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

Satz 6.2. *Sei $A = (a_{ij})$ eine $m \times n$ Matrix. Dann gibt es eine $m \times n$ Matrix B in Zeilenstufenform, die denselben Zeilenraum hat wie A.*

Beweis. Induktion nach m.

(i) $m = 1$. Sei $\beta_1 = (a_{11}, \ldots, a_{1n})$ der (einzige) Zeilenvekter. Falls alle a_{1j} null sind, dann ist A sicherlich schon in Zeilenstufenform. Sei daher j_1 die erste nicht-null Stelle. Was ist a_{1j_1}? Falls $a_{1j_1} = 1$, dann sind wir fertig, sonst führe die Zeilenoperation $S_1(a_{1j_1}^{-1})$ aus.

(ii) Falls $m > 1$, dann können wir annehmen, daß die ersten $m-1$ Zeilen schon in Zeilenstufenform sind, während die letzte Zeile $\beta_m = (a_{m1}, \ldots, a_{mn})$ beliebig ist. Falls $a_{m1} \neq 0$ und $j_1 > 1$, dann führe die Zeilenoperationen $S_m(a_{mj_1}^{-1})$ und S_{1m} aus. Nachher ist A sicherlich in eine Matrix in Zeilenstufenform transformiert worden. Falls $a_{m1} \neq 0$ und $j_1 = 1$, dann führe die Zeilenoperation $S_{m1}(a_{1m}^{-1})$ aus. Jetzt haben wir auf jeden Fall erreicht, daß das erste Element des letzten Zeilenvektors Null ist. Versuche nun so weit wie möglich nach diesem Verfahren—von links nach rechts—die Elemente des letzten Zeilenvektors auf Null zu bringen durch Zeilenoperationen der Art $S_{ms}(a_{ms}^{-1})$. Zum Schluß wird Zeilenstufenform erreicht durch eine $S_m(a_{ms}^{-1})$ Zeilenoperation, möglicherweise gefolgt von einer S_{ms} Zeilenoperation.

$\square$

Satz 6.3. *Sei A in Zeilenstufenform. Dann bilden die nicht-null Zeilen eine Basis des Zeilenraums.*

Beweis. Wir brauchen nur die lineare Unabhängigkeit nach zu prüfen. Seien daher $\{\beta_1, \ldots, \beta_p\}$ die nicht-null Vektoren, wobei

$$\beta_i = (0, \ldots, 0, 1, \ldots, a_{in}) \in F^n, \quad \forall\, i,$$

wobei die ersten j_i-ten Elemente Null sind, und $j_s < j_t$, für $s < t$. Sei

$$0 = \sum_{i=1}^{p} b_i \beta_i,$$

wobei mindestens ein $b_i \neq 0$. Sei etwa $b_{i(1)}$ der erste nicht-null Koeffizient. Dann ist

$$b_{i(1)} \beta_{i(1)} = - \sum_{j > i(1)} b_j \beta_j.$$

Aber dies ist unmöglich, da $\beta_{i(1)} = (0, \ldots, 0, 1, \ldots, a_{i(1)\,n})$, wobei das erste nicht-null Element an der $i(1)$-ten Stelle kommt, während für alle $j > i(1)$ alle Stellen in $\beta_j = (0, \ldots, 0, 1, \ldots, a_{jn})$ Null bis hinter der $i(1)$-ten Stelle sind. $\square$

6.2 Spaltenvektoren

Wir wollen uns nun die alternative Beschreibung einer Matrix A durch Spaltenvektoren anschauen.

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} \cdots \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

Sei $T = \{\gamma_1, \ldots, \gamma_n\}$ mit $\gamma_i = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix} \in F^m$ für $i = 1, \ldots, n$.

Definition 32. $[T] \subset F^n$ heißt der *Spaltenraum* der Matrix A.

Satz 6.4. *Sei A eine $m \times n$ Matrix und $T = \{\gamma_1, \ldots, \gamma_n\}$ die Menge der Spaltenvektoren. Sei A durch eine elementare Zeilenoperation in eine neue Matrix $A \longrightarrow A'$ transformiert. Dadurch werden die Spaltenvektoren auch geändert: $\gamma_i \longrightarrow \gamma_i'$. Insgesamt ändert sich die Menge der Spaltenvektoren: $Y \longrightarrow Y'$. Es gilt: Y ist linear unabhängig $\Leftrightarrow$ Y' ist linear unabhängig.*

Beweis. (für die Zeilenoperation S_{ij}) Wir haben $\gamma_k = \begin{pmatrix} a_{1k} \\ \vdots \\ a_{mk} \end{pmatrix}$ für alle k. Es gilt daher:

$$\begin{aligned} 0 = \sum_{k=1}^{n} c_k \gamma_k \quad &\Leftrightarrow \quad \sum_{k=1}^{n} c_k a_{pk} = 0, \quad \forall\, p = 1, \ldots, m \\ &\Leftrightarrow \quad \sum_{k=1}^{n} c_k \gamma_k' = 0 \end{aligned}$$

Die Beweise für die anderen elementaren Zeilenoperationen $S_i(a)$ und $S_{ij}(c)$ sind ähnlich. □

Definition 33. (Mit S als Menge der Zeilenvektoren) $\dim([S])$ heißt der *Zeilenrang* der Matrix A. $\dim([T])$ heißt der *Spaltenrang* von A.

Satz 6.5. *Sei A eine $m \times n$ Matrix. Es gilt: Zeilenrang = Spaltenrang. Diese gemeinsame Zahl heißt der* Rang *von A: Rang(A).*

Beweis. Nach Satz 6.1 wird der Zeilenrang nicht geändert durch elementare Zeilenoperationen. Nach Satz 6.5 gilt dasselbe für den Spaltenrang. Wir können daher annehmen, daß die Matrix A in Zeilenstufenform ist. Nach Satz 6.3 ist der Zeilenrang gleich der Anzahl der Stufen. Dies ist offensichtlich auch der Spaltenrang. □

Definition 34. Sei A eine quadratische Matrix (etwa $n \times n$). A heißt *regulär*, falls $\text{Rang}(A) = n$, sonst *singulär*.

Satz 6.6. *A ist regulär $\Leftrightarrow \exists$ ein Isomorphismus $f : F^n \to F^n$ mit zugehöriger Matrix A.*

Beweis. '$\Rightarrow$' Sei A regulär. $\Rightarrow$ Spaltenrang $= n$. Sei $\{\epsilon_1, \dots, \epsilon_n\}$ die kanonische Basis für F^n. Eine Abbildung $f : F^n \to F^n$ wird definiert durch die Regel $f(\epsilon_i) = \gamma_i$, $\forall\, i = 1, \dots, n$, wobei γ_i die i-te Spalte von A ist. Da $[\{\gamma_1, \dots, \gamma_n\}] = F^n$, ist f eine Surjektion. Aber die Menge $\{\gamma_1, \dots, \gamma_n\}$ muß auch linear unabhängig sein; folglich ist f eine Injektion.

'$\Leftarrow$' Sei $f : F^n \to F^n$ ein Isomorphismus. Dann ist die Menge von Vektoren $T = \{f(\epsilon_i) : i = 1, \dots, n\}$ linear unabhängig. D.h. $\dim(T) = n$. Sei $\gamma_i \equiv f(\epsilon_i)$ für alle i, geschrieben als Spaltenvektoren. Dann gibt $\{\gamma_1, \dots, \gamma_n\}$ eine Matrix, die wir A nennen können, und A ist regulär.

□

Kapitel 7

Lineare Gleichungssysteme

Sei F ein Körper und

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

ein System von m linearen Gleichungen in n Unbekannten $x_1, \ldots, x_n$, wobei $a_{ij}, b_i \in F$, für alle relevanten i und j. Wir haben schon gesehen, daß es möglich ist, dieses System als eine lineare Abbildung $f : F^n \to F^m$ zu betrachten: $f(\vec{x}) = \vec{b}$. D.h. in der Matrizenschreibweise: $A \cdot x = b$.

Wir wollen jetzt etwas praktischer sein und fragen, gibt es Lösungen zu diesem Gleichungssystem? Falls ja, wie können wir *alle* Lösungen finden?

Definition 35. Sei $Ax = b$ das lineare Gleichungssystem oben. A ist die $m \times n$ Matrix $A = (a_{ij})$. Die $m \times (n+1)$ Matrix

$$A' = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{pmatrix}$$

heißt die *erweiterte* Matrix.

Wir betrachten jetzt die drei elementaren Zeilenoperationen, angewandt auf A'. Z.B. wenn wir zwei Zeilen miteinander vertauschen (S_{ij}), dann ist das System von linearen Gleichungen 'eigentlich' dasselbe. Genauer gesagt, die Lösungsmenge

$$L = \{(x_1, \ldots, x_n) \in F^n : A \cdot x = b\}$$

ist gleich für die erweiterte Matrix A' und die transformierte Matrix $S_{ij}(A')$. Es ist ähnlich leicht einzusehen, daß diese Lösungsmenge auch unter den elementaren Zeilenoperationen $S_i(a)$ und $S_{ij}(c)$ unverändert bleibt. Daher brauchen wir nur die Sätze 21 und 23 zu berücksichtigen, um zu sehen, daß wir annehmen können, daß A' in Zeilenstufenform ist.

Durch mögliche Umnumerierung der Variablen $\{x_1, \ldots, x_n\}$ können wir weiter annehmen, daß die erweiterte Matrix die *einfache* Zeilenstufenform

$$A' = \begin{pmatrix} 1 & a_{12} & \cdots & \cdot & a_{1\,r+1} & \cdots & a_{1n} & b_1 \\ 0 & 1 & a_{23} & \cdot & \vdots & & \vdots & \vdots \\ \vdots & & \ddots & \cdot & \cdot & & \cdot & \cdot \\ 0 & \ldots & 0 & 1 & a_{r\,r+1} & \cdots & a_{rn} & b_r \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & b_{r+1} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & b_m \end{pmatrix}$$

hat. (Das heißt, das obere linke Quadrat ist eine Dreiecksmatrix.)

Aber jetzt haben wir die Möglichkeit, eine systematische Beschreibung der Lösungsmenge zu geben, nämlich:

1. Falls $\exists$ ein j (zwischen $r+1$ und n) mit $b_j \neq 0$, dann hat das Gleichungssystem $Ax = b$ offensichtlich *keine* Lösungen.

2. Falls $r = n$ (und Fall 1 nicht zutrifft), dann gibt es genau eine Lösung, die wir erhalten können durch sukzessives Einsetzen der Lösungen für x_n, dann x_{n-1}, u.s.w.

3. Falls $r < n$, dann ist für gewisse x_1^*, ..., x_r^* der Vektor

$$x = \begin{pmatrix} x_1^* \\ \vdots \\ x_r^* \\ d_{r+1} \\ \vdots \\ d_n \end{pmatrix},$$

   wobei $d_j \in F$, $j = r+1, \dots, n$ beliebig sind, eine Lösung des Systems $Ax = b$.

   Warum? Die r-te Gleichung lautet dann

$$1 \cdot x_r + a_{r\,r+1} d_{r+1} + \cdots + a_{rn} d_n = b_r.$$

   D.h. setze

$$x_r^* = b_r - \sum_{i=r+1}^{n} a_{ri} d_i$$

   und die r-te Gleichung ist erfüllt. Man findet ebenso Zahlen $x_{r-1}^*,\ x_{r-2}^*, \dots, x_1^*$ die den ersten $r-1$ Gleichungen genügen.

Zusammengefaßt können wir sagen:

Satz 7.1. *Sei $Ax = b$ und A' die dazugehörige erweiterte Matrix. Sei L die Menge der Lösungen. Dann gilt $L \neq \emptyset \Leftrightarrow Rang(A) = Rang(A')$.*

Im Falle $Ax = 0$, (d.h. im Falle eines *homogenen* linearen Gleichungssystems), folgt:

Satz 7.2. *Sei L die Menge der Lösungen von $Ax = 0$. Dann ist L ein Unterraum von F^n und $dim(L) + Rang(A') = n$. (Wobei $Ax = 0$ eigentlich m Gleichungen in n Unbekannten sind.)*

Beweis. Fall (1) kann nicht zutreffen. (2) ist klar. Nach Fall (3) ist $\dim(L) = n - r$. Aber $\mathrm{Rang}(A) = r$ (Satz 6.3). (Wir haben schon gesehen, daß die Lösungsmenge eines Systems von homogenen linearen Gleichungen, ein Unterraum von F^n sein muß.) □

7.1 Ein Beispiel

$$\begin{aligned} x + 2y + 3z &= 1 \\ 4x + 5y + 6z &= 2 \\ 7x + 8y + 9z &= 3 \end{aligned}$$

Die erweiterte Matrix ist daher

$$\begin{pmatrix} 1 & 2 & 3 & 1 \\ 4 & 5 & 6 & 2 \\ 7 & 8 & 9 & 3 \end{pmatrix}.$$

Diese Matrix muß jetzt in Zeilenstufenform gebracht werden:

$$\begin{pmatrix} 1 & 2 & 3 & 1 \\ 4 & 5 & 6 & 2 \\ 7 & 8 & 9 & 3 \end{pmatrix} \xrightarrow{S_{21}(-4)} \begin{pmatrix} 1 & 2 & 3 & 1 \\ 0 & -3 & -6 & -2 \\ 7 & 8 & 9 & 3 \end{pmatrix} \xrightarrow{S_{31}(-7)} \begin{pmatrix} 1 & 2 & 3 & 1 \\ 0 & -3 & -6 & -2 \\ 0 & -6 & -12 & -4 \end{pmatrix}$$

$$\xrightarrow{S_{32}(-2)} \begin{pmatrix} 1 & 2 & 3 & 1 \\ 0 & -3 & -6 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{S_2(-\frac{1}{3})} \begin{pmatrix} 1 & 2 & 3 & 1 \\ 0 & 1 & 2 & \frac{2}{3} \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Offensichtlich haben wir Fall (3). D.h. hier ist $r = 2$, $n = 3$. Nach der Theorie, kann $z_0 \in \mathbf{R}$ beliebig gewählt werden, dann ist $(x^*, y^*, z_0) \in \mathbf{R}^3$ eine Lösung, wobei

$$y^* = \frac{2}{3} - 2z_0$$

und daher

$$x^* = 1 - 2y^* - 3z_0 = -\frac{1}{3} - 7z_0.$$

7.2 Ein entsprechendes Computerprogramm

```
program LineareGleichungen (input, output);
const
   Anzahl_der_Gleichungen = 3;  { Wieviele lineare Gleichungen? }
   Anzahl_der_Unbekannten = 4;  { Koeffizienten und Konstanten }
type
   matrix =
   array[1..Anzahl_der_Gleichungen,1..Anzahl_der_Unbekannten]
   of real;
var
   matrix1, matrix2 : matrix;

procedure Lese_Gleichungen (var m : matrix);
var
   i, j : integer;
begin
   for i := 1 to Anzahl_der_Gleichungen do
      for j := 1 to Anzahl_der_Unbekannten do read (m[i,j]);
end; { LeseGleichungen }

procedure Schreibe_Gleichungen (m : matrix);
var
   i, j : integer;
begin
   for i := 1 to Anzahl_der_Gleichungen do begin
      for j := 1 to Anzahl_der_Unbekannten-1 do
      begin
         write (m[i,j]:3); write(' x_');
         write(j:1); write('   ');
      end;
      write(' = '); writeln(m[i,Anzahl_der_Unbekannten]:3);
   end;
end; { SchreibeGleichungen }

{ ELEMENTARE ZEILENOPERATIONEN }
procedure tausch_ij (i, j : integer; var m : matrix);
{ Die Zeilen i und j werden vertauscht }
var
   s : integer; temp : real;
begin
   for s := 1 to Anzahl_der_Unbekannten do begin
      temp := m[i,s];
      m[i,s] := m[j,s];
      m[j,s] := temp;
   end;
end; { tausch_ij }

procedure mult_ia (i : integer; a : real; var m : matrix);
{ Zeile i mit a multipliziert }
var
   s : integer;
begin
   for s := 1 to Anzahl_der_Unbekannten do begin
      m[i,s] := a * m[i,s];
   end;
end; { mult_ia }

procedure addzeil (i, j : integer; c : real; var m : matrix);
{ Zeile j multipliziert mit c wird zu Zeile i addiert }
var
   s : integer; temp : real;
```

```
begin
   for s := 1 to Anzahl_der_Unbekannten do begin
      m[i,s] := m[i,s] + c * m[j,s];
   end;
end; { addzeil }


{ HAUPTPROZEDUR }

procedure Gauss_Elimination (eingegebenematrix : matrix;
                                         var m : matrix);
var
   i, j, k, z, min : integer;
begin
   { setzte zunaechst varible m gleich eingegebenematrix }
   for i := 1 to Anzahl_der_Gleichungen do
      for j := 1 to Anzahl_der_Unbekannten do
         m[i,j] := eingegebenematrix [i,j];

   { Zeilenoperationen, um m in die
     obere Dreiecksform zu bringen }

   i := 1;
   if Anzahl_der_Unbekannten-1 < Anzahl_der_Gleichungen then
      min := Anzahl_der_Gleichungen
   else min := Anzahl_der_Unbekannten-1;
   for j := 1 to min do begin
      z := i;
      while (m[z,j] = 0) and (z <= Anzahl_der_Gleichungen) do
         z := z+1;
      if z <= Anzahl_der_Gleichungen then
      begin
         tausch_ij (i, z, m);
         mult_ia (i, 1/m[i,j], m);
         for z := i+1 to Anzahl_der_Gleichungen do
            if m[z,j] <> 0 then addzeil (z, i, -m[z,j], m);
      end;
      i := i+1;
      writeln; Schreibe_Gleichungen(m); writeln;
   end;
end; { Gauss_Elimination }

procedure Finde_Loesung (m : matrix);
var i, j :  integer;
begin
   writeln ('Die Loesung ist :');
   for i := Anzahl_der_Gleichungen downto 2 do
      for j := (i - 1) downto 1 do
         addzeil (j, i, -m[j,i], m);
   for i := 1 to Anzahl_der_Gleichungen do
   begin
      write ('x_'); write (i:1); write (' = ');
      writeln (m[i,Anzahl_der_Unbekannten]);
   end;
end; { Finde_Loesung }

procedure Dimension_der_LoesungsMenge (m : matrix);
var
   i, j                      : integer;
   keineLoesungen, nullzeile : boolean;
begin
```

```
   keineLoesungen := false;
   nullzeile := true;
   i := Anzahl_der_Gleichungen;
   while nullzeile and (i >= 1) do begin
      for j := 1 to Anzahl_der_Unbekannten-1 do
         if m[i,j] <> 0 then nullzeile := false;
      if nullzeile and (m[i,Anzahl_der_Unbekannten] <> 0) then
         keineLoesungen := true;
      if nullzeile then i := i-1;
   end;
   if keineLoesungen then writeln ('keine Loesungen!')
   else
   begin
      write ('Die Dimension des Loesungsraumes ist ');
      write ((Anzahl_der_Unbekannten - 1) - i: 1);
      writeln ('.');
   end;
   if Anzahl_der_Unbekannten - 1 - i = 0
      then Finde_Loesung (m);
end; { Dimension_der_LoesungsMenge }


begin
   Lese_Gleichungen (matrix1);
   Gauss_Elimination (matrix1,matrix2);
   Dimension_der_LoesungsMenge (matrix2);
end.
```

Kapitel 8

Mehr über lineare Abbildungen

Im Kapitel 5 haben wir die folgende Beschreibung von linearen Abbildungen gesehen. Seien X und Y (endlich dimensionale) Vektorräume über einem Körper F und sei $f : X \to Y$ eine lineare Abbildung. Wir wählen die Basen $S = \{\alpha_1, \dots, \alpha_n\}$ und $T = \{\beta_1, \dots, \beta_m\}$ für X bzw. Y. Dann gilt

$$f(\alpha_i) = \sum_{j=1}^{m} a_{ji}\beta_j$$

für gewisse $a_{ji} \in F$.

Definition 36. Die Matrix $A = (a_{ji})$ heißt die zu f gehörige Matrix bezüglich der Basen S und T.

Sei $\alpha = \sum_{i=1}^{n} x_i\alpha_i$ ein beliebiger Vektor in X. Wenn α als Spaltenvektor dargestellt wird, entsteht durch Matrizenmultiplikation ein weiterer Spaltenvektor, nämlich $A \cdot \alpha = f(\alpha)$, der entsprechende Vektor im Vektorraum Y. Nach Satz 5.2 ist $f(X)$ ein Unterraum von Y. Die Dimension dieses Unterraums ist sicherlich nicht größer als die Dimension des Raumes Y.

Definition 37. Sei $f : X \to Y$ eine lineare Abbildung. Der *Rang* von f ist $\text{Rang}(f) = \dim(\text{im(f)})$.

Satz 8.1. *Seien $f : X \to Y$, S und T wie oben und sei $A = (a_{ji})$ die zu f gehörige Matrix bezüglich S und T. Dann ist*

$$Rang(f) = Rang(A).$$

Beweis. $\text{Rang}(A)$ = Spaltenrang von A (Satz 6.5). Aber wie wir gesehen haben, kann die Matrix A als eine Sammlung von n Spaltenvektoren in F^m betrachtet werden. Wenn wir β_j mit dem j-ten kanonischen Spaltenvektor identifizieren: $\beta_j = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$ für $j = 1, \dots, n$, dann ist $f(\alpha_i) = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix}$, $\forall i$. Aber die Vektoren $f(\alpha_i)$ bilden ein Erzeugendensystem für $f(X)$ in Y. D.h. $f(X) = [\{f(\alpha_i), \dots, f(\alpha_n)\}]$. Wenn wir diese Vektoren in Y betrachten, dann erzeugen sie $f(X)$. Betrachtet in F^m, erzeugen sie den Spaltenraum von A. Aber es handelt sich 'eigentlich' um einen einzigen Vektorraum. (D.h. $f(X)$ und der Spaltenraum von A sind zu einander isomorph.) Die Dimension dieses Raums ist $\text{Rang}(f) = \text{Rang}(A)$. □

In diesem Beweis war eine Basis T für den Vektorraum Y vorgegeben. Wir haben dann gesagt, daß es möglich ist, diese Basis mit der kanonischen Basis für F^m zu identifizieren, wobei das Korollar zu Satz 5.1 ($Y \cong F^m$) benutzt wird. Es ist aber auch möglich, andere Basen S' und T' für X bzw. Y zu wählen.

Korollar 8.1.1. *Seien S', T' zwei andere Basen und sei A' die zu f gehörige Matrix bzgl. S' und T'. Dann ist auch $Rang(f) = Rang(A')$. Insbesondere ist $Rang(A) = Rang(A')$.*

Satz 8.2. *Sei $f : X \to Y$ eine lineare Abbildung. Dann ist*

$$dim(ker(f)) + dim(im(f)) = dim(X).$$

Beweis. $ker(f)$ ist ein Unterraum von X. Sei daher $\{\alpha_1, \dots, \alpha_p\} \subset X$ eine Basis für $ker(f)$. Nach dem Basisergänzungssatz können wir eine Basis $\{\alpha_1, \dots, \alpha_p\, \alpha_{p+1}, \dots, \alpha_n\}$ für X finden, wobei $n = dim(X)$. Wir müssen daher zeigen, daß $dim(im(f)) = n - p$. Es genügt zu zeigen, daß

$$U = \{f(\alpha_{p+1}), \dots, f(\alpha_n)\}$$

eine Basis für $im(f)$ ist.

Nun, $\{f(\alpha_1), \ldots, f(\alpha_n)\}$ ist sicherlich ein Erzeugendensystem für $f(X)$. Aber die ersten p Vektoren sind jeweils einfach der Null-Vektor in Y. (Dies ist die Definition von $ker(f)$!) Daher können diese Vektoren weggelassen werden, und es folgt daß auch U allein ein Erzeugendensystem für $im(f)$ ist. Wir brauchen daher nur zu zeigen, daß die Menge U linear unabhängig ist. Sei

$$0 = \sum_{j=p+1}^{n} b_j f(\alpha_j) = f\left(\sum_{j=p+1}^{n} b_j \alpha_j\right)$$

$$\Rightarrow \sum_{j=p+1}^{n} b_j \alpha_j \in ker(f).$$

Das heißt

$$\sum_{j=p+1}^{n} b_j \alpha_j = \sum_{i=1}^{p} a_i \alpha_i \quad \text{etwa.}$$

$$\text{oder} \quad 0 = \sum_{i=1}^{p} a_i \alpha_i - \sum_{j=p+1}^{n} b_j \alpha_j$$

Es folgt, daß $a_i = 0 = b_j$, $\forall\, i, j$, da $\{\alpha_1, \ldots, \alpha_n\}$ eine Basis für X ist (und daher linear unabhängig). Das heißt, die Menge $U = \{f(\alpha_{p+1}), \ldots, f(\alpha_n)\}$ ist linear unabhängig. □

Korollar 8.2.1. *$Rang(f) + dim(ker(f)) = dim(X)$.*

In Satz 6.6 haben wir eine besondere Klasse von Matrizen—die regulären Matrizen—gesehen. Diese Matrizen entsprechen Isomorphismen zwischen Vektorräumen. Nun, eine ganz besondere $n \times n$ Matrix ist die Identitätsmatrix

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

Satz 8.3. *Angenommen die $n \times n$ Matrix A sei regulär. Dann existiert eine $n \times n$ Matrix B mit $A \cdot B = I_n$.*

Beweis. Die Matrix A ist die Darstellung einer linearen Abbildung $f : F^n \to F^n$, wobei $f(\epsilon_i)$ die i-te Zeile von A ist. (ϵ_i ist der i-te kanonische Basisvektor von F^n.) Laut Satz 6.6 ist f ein Isomorphismus. D.h. $f^{-1} : F^n \to F^n$ ist auch eine lineare Abbildung, und natürlich ist $f^{-1}(f(\epsilon_i)) = id(\epsilon_i) = \epsilon_i$, $\forall\, i$. Sei B die zu f^{-1} gehörige Matrix bezüglich der kanonischen Basisvektoren. Dann ist das Produkt BA die zur Identitätsabbildung gehörige Matrix bezüglich der kanonischen Basisvektoren. Aber dies ist offensichtlich I_n. Nun, die Komposition $f \circ f^{-1} = id$ gilt auch. Daher ist $AB = I_n$. □

Definition 38. Sei die Matrix A regulär. Dann heißt A *invertierbar*. Die Matrix A^{-1} mit $A \cdot A^{-1} = I_n$ heißt die zu A inverse Matrix.[1] Die Menge aller invertierbaren $n \times n$ Matrizen mit Elementen aus dem Körper F wird mit $GL(n; F)$ bezeichnet.[2]

Satz 8.4. *$GL(n; F)$ ist eine Gruppe (mit Matrizenmultiplikation).*

Beweis. Trivialerweise ist $A \cdot I_n = I_n \cdot A = A$, $\forall\, A \in GL(n; F)$. Daher ist I_n das Identitätselement in der Gruppe. Nach der Definition existiert ein $A^{-1} \in GL(n; F)$, $\forall\, A \in GL(n; F)$. Die Assoziativität ist eine allgemeingültige Eigenschaft bei Matrizenmultiplikation.

Es bleibt nur zu zeigen, daß $GL(n; F)$ abgeschlossen unter Matrizenmultiplikation ist. Seien daher $A, B \in GL(n; F)$. Es existieren dann weitere $n \times n$ Matrizen A^{-1} und B^{-1}. Aber

$$(AB) \cdot (B^{-1}A^{-1}) = (A \cdot (BB^{-1})) \cdot A^{-1} = (A \cdot I_n) \cdot A^{-1} = AA^{-1} = I_n.$$

Das heißt, $(AB)^{-1} = B^{-1} \cdot A^{-1}$. □

Zusammengefaßt: Sei $f : X \to Y$ eine lineare Abbildung und A die zu f gehörige Matrix bezüglich Basen $\{\alpha_1, \ldots, \alpha_n\}$ für X und $\{\beta_1, \ldots, \beta_m\}$ für Y. Es gilt: A ist invertierbar $\Leftrightarrow$ A ist regulär $\Leftrightarrow$ $A \in GL(n; F)$ $\Leftrightarrow$ f ist ein Isomorphismus. $\Leftrightarrow$ die Menge

$$\{f(\alpha_1), \ldots, f(\alpha_n)\} \subset Y$$

ist eine Basis für Y.

[1] A^{-1} ist eindeutig, da $AB = I_n$ und $AC = I_n \Rightarrow (CA)B = CI_n \Rightarrow I_nB = C \Rightarrow B = C$.

[2] GL heißt die '*general linear*' Gruppe der Ordnung n über F.

Definition 39. Eine $m \times n$ Matrix A ist in oberer Diagonalform, falls $A = (a_{ij})$, wobei $\exists p \leq m$ und n, und

$$a_{ij} = \begin{cases} 1, & \text{falls } i = j \leq p, \\ 0 & \text{sonst.} \end{cases}$$

$$\left(\begin{array}{cccccc|cc} 1 & 0 & 0 & \dots & & 0 & 0 & 0 \\ 0 & 1 & 0 & \dots & & 0 & 0 & 0 \\ 0 & 0 & 1 & \dots & & 0 & 0 & 0 \\ & \vdots & & \ddots & & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & \dots & & 0 & 0 & 0 \end{array}\right)$$

Satz 8.5. *Sei A irgendeine $m \times n$ Matrix. Dann existieren zwei invertierbare Matrizen $E \in GL(m, F)$ und $D \in GL(n, F)$ so, daß das Produkt $A' = EAD^{-1}$ in oberer Diagonalform ist.*

Beweis. A stellt eine lineare Abbildung $f : V \to W$ dar, wobei $dim(V) = n$, $dim(W) = m$. A ist die Darstellung von f bzgl. den kanonischen Basen für F^n, bzw. F^m. (D.h. die vorgegebenen Basen $\{\gamma_1, \dots, \gamma_n\}$ bzw. $\{\delta_1, \dots, \delta_m\}$ für V und W entsprechen hier den kanonischen Basen in F^n und F^m.) Sei nun $p = n - dim(ker(f))$ und sei $\{\alpha_{p+1}, \dots, \alpha_n\} \subset V$ eine Basis für $ker(f)$. Mittels Basisergänzungssatz, erhalten wir eine Basis $\{\alpha_1, \dots, \alpha_p, \alpha_{p+1}, \dots, \alpha_n\}$ für V. Insbesondere ist $\{f(\alpha_1), \dots, f(\alpha_p)\} \subset W$ linear unabhängig.

Dies folgt, da

$$\begin{aligned} 0 &= \sum_{i=1}^{p} a_i f(\alpha_i) \;\Rightarrow\; 0 = f\left(\sum_{i=1}^{p} a_i \alpha_i\right) \\ &\Rightarrow \quad \alpha = \sum_{i=1}^{p} a_i \alpha_i \;\in\; ker(f) \\ &\Rightarrow \quad \exists\, b_{p+1}, \dots, b_n \text{ mit } \alpha = \sum_{j=p+1}^{n} b_j \alpha_j \\ &\Rightarrow \quad 0 = \sum_{i=1}^{p} a_i \alpha_i - \sum_{j=p+1}^{n} b_j \alpha_j \\ &\Rightarrow \quad a_i, b_j = 0, \quad \text{da } \{\alpha_1, \dots, \alpha_n\} \text{ eine Basis ist.} \end{aligned}$$

Setze $\beta_i \equiv f(\alpha_i)$, $i = 1, \dots, p$. Nach dem Basisergänzungssatz $\exists\, \beta_{p+1}, \dots, \beta_m$ so daß $\{\beta_1, \dots, \beta_p, \beta_{p+1}, \dots, \beta_m\}$ eine Basis für W ist. Seien jetzt $g : V \to V$ und $h : W \to W$ Automorphismen, wobei $g(\alpha_i) = \gamma_i$ und $h(\beta_j) = \delta_j$, $\forall\, i, j$. Dann ist $h \cdot f \cdot g^{-1} : V \to W$ die Abbildung

$$h \cdot f \cdot g^{-1}(\gamma_i) = h \cdot f(\alpha_i) = \begin{cases} h(\beta_i) = \delta_i, & \text{falls } \quad i \leq p, \\ h(0) = 0, & \text{sonst.} \quad (i > p) \end{cases}$$

Sei daher $D \in GL(n, F)$ die Matrizendarstellung von g bzgl. der Basis $\{\gamma_1, \dots, \gamma_n\}$ von V und $E \in GL(m, F)$ die Darstellung von h bzg. der Basis $\{\delta_1, \dots, \delta_m\}$ von W. Dann ist die Matrix $A' = EAD^{-1}$ in oberer Diagonalform. $\square$

Kapitel 9

Ähnlichkeit von Matrizen, elementare Matrizen, invariante Unterräume

Sei V ein n-dimensionaler Vektorraum über einem Körper F. Seien $\{\alpha_1, \dots, \alpha_n\}$ und $\{\beta_1, \dots, \beta_n\}$ zwei Basen für V. Es gilt etwa

$$\begin{aligned} \alpha_i &= \sum_{j=1}^{n} a_{ji}\beta_j \quad \text{und} \\ \beta_j &= \sum_{i=1}^{n} b_{ij}\alpha_i. \end{aligned}$$

Das heißt

$$\alpha_i = \sum_{j=1}^{n} a_{ji}\beta_j = \sum_{j=1}^{n} a_{ji}\left(\sum_{k=1}^{n} b_{kj}\alpha_k\right) = \sum_{k=1}^{n}\left(\sum_{j=1}^{n} b_{kj}a_{ji}\right)\alpha_k.$$

Da $\{\alpha_1, \dots, \alpha_n\}$ eine Basis ist $\Rightarrow$

$$\sum_{j=1}^{n} b_{kj}a_{ji} = \begin{cases} 1, & \text{falls i=k,} \\ 0, & \text{sonst.} \end{cases}$$

D.h. $B \cdot A = I$, wobei $A = (a_{ij})$ und $B = (b_{kl})$. Oder $B = A^{-1}$.

Sei jetzt $\alpha = \sum_{i=1}^{n} x_i\alpha_i$ ein beliebiger Vektor in V. Was ist α bzgl. $\{\beta_1, \dots, \beta_n\}$? Es gilt

$$\alpha = \sum_{i=1}^{n} x_i\alpha_i = \sum_{i=1}^{n} x_i\left(\sum_{j=1}^{n} a_{ji}\beta_j\right) = \sum_{j=1}^{n}\left(\sum_{i=1}^{n} a_{ji}x_i\right)\beta_j.$$

Wir können nun—wie in früheren Kapiteln—die Matrix A als eine Darstellung einer linearen Abbildung $V \to V$ betrachten. Dann ist $A{\cdot}\alpha$ auch ein Vektor in V, dargestellt als Linearkombination der Basisvektoren $\{\beta_1, \dots, \beta_n\}$.

Satz 9.1. *Sei $f : V \to V$ eine lineare Abbildung eines Vektorraumes in sich selbst. $\{\alpha_1, \dots, \alpha_n\}$ und $\{\beta_1, \dots, \beta_n\}$ seien zwei Basen für V. Seien C die zu f gehörige Matrix bzgl. der Basis $\{\alpha_1, \dots, \alpha_n\}$ und C' die zu f gehörige Matrix bzgl. der Basis $\{\beta_1, \dots, \beta_n\}$. Dann ist $C' = ACA^{-1}$ ($= ACB$ mit $B = A^{-1}$) wobei $A = (a_{ij})$, $B = (b_{ij})$, und $\alpha_i = \sum_{j=1}^{n} a_{ji}\beta_j$, $\forall i, j$.*

Beweis. Sei $i \in \{1, \dots, n\}$ beliebig. Da C' die zu f gehörige Matrix bzgl. $\{\beta_1, \dots, \beta_n\}$ ist, gilt

$$f(\beta_i) = \sum_{j=1}^{n} c'_{ji}\beta_j.$$

Andererseits gilt

$$
\begin{aligned}
f(\beta_i) &= f\left(\sum_{j=1}^n b_{ji}\alpha_j\right) \qquad \text{(wobei } B = A^{-1}\text{)} \\
&= \sum_{k=1}^n b_{ki} f(\alpha_k) \\
&= \sum_{k=1}^n b_{ki}\left(\sum_{l=1}^n c_{lk}\alpha_l\right) \qquad (f \text{ bzgl. } \{\alpha_1,\dots,\alpha_n\}) \\
&= \sum_{k=1}^n \left(\sum_{l=1}^n c_{lk} b_{ki}\right)\alpha_l \\
&= \sum_{k=1}^n \left(\sum_{l=1}^n c_{lk} b_{ki}\right)\left(\sum_{j=1}^n a_{jl}\beta_j\right) \\
&= \sum_{j=1}^n \left(\sum_{l=1}^n\sum_{k=1}^n a_{jl}c_{lk}b_{ki}\right)\beta_j
\end{aligned}
$$

Da aber $\{\beta_1,\dots,\beta_n\}$ linear unabhängig ist (und daher die Darstellung von $f(\beta_i)$ eindeutig ist bzgl. dieser Basis), gilt

$$c'_{ji} = \sum_{l=1}^n\sum_{k=1}^n a_{jl}c_{lk}b_{ki}, \quad \forall j, i.$$

D.h. $C' = ACA^{-1}$. □

Korollar 9.1.1. *Es gilt die Umkehrung. D.h. gegeben eine $n\times n$ Matrix C, dann existiert eine lineare Abbildung $f : V \to V$, wobei $dim(V) = n$, so daß C die zu f gehörige Matrix ist bzgl. einer Basis $\{\alpha_1,\dots,\alpha_n\}$. Sei $A \in GL(n, F)$ beliebig. Dann ist $C' = ACA^{-1}$ die zu f gehörige Matrix bzgl. einer anderen Basis $\{\beta_1,\dots,\beta_n\}$.*

(Die Basisvektoren werden durch die Regel $\beta_j = \sum_{ij}^n b_{ij}\alpha_i$ gegeben, wobei $(b_{ij}) = B = A^{-1}$.)

Definition 40. Seien C', und C zwei $n\times n$ Matrizen. C' und C sind zueinander *ähnlich*, falls eine $A \in GL(n, F)$ existiert, mit

$$C' = ACA^{-1}.$$

9.1 Die elementaren Matrizen

Sei $S_i(a)$, $a \neq 0$, die folgende $n \times n$ Matrix:

$$
S_i(a) = \begin{pmatrix}
1 & 0 & & \dots & & & 0 \\
0 & \ddots & & & & & \\
& & 1 & & & & \\
\vdots & & & a & & & \vdots \\
& & & & 1 & & \\
& & & & & \ddots & 0 \\
0 & & & \dots & & 0 & 1
\end{pmatrix}.
$$

D.h. alles ist Null außer der Hauptdiagonalen der Matrix. Dort stehen lauter Einsen außer an der i-ten Stelle, wo ein a steht. Es ist eine leichte Übung nachzuprüfen, daß die elementare Matrizenumformung $S_i(a)$ gegeben ist durch Matrizenmultiplikation mit der elementaren Matrix $S_i(a)$. D.h. sei A eine $n\times m$ Matrix. Was ist $S_i(a)\cdot A$?

$$
\begin{pmatrix}
1 & & \dots & & 0 \\
& \ddots & & & \\
\vdots & & a & & \vdots \\
& & & \ddots & \\
0 & & \dots & & 1
\end{pmatrix}
\cdot
\begin{pmatrix}
a_{11} & \dots & a_{1m} \\
\vdots & \ddots & \vdots \\
a_{n1} & \dots & a_{nm}
\end{pmatrix}
=
\begin{pmatrix}
a_{11} & \dots & a_{1m} \\
\vdots & & \vdots \\
a\cdot a_{i1} & \dots & a\cdot a_{im} \\
\vdots & & \vdots \\
a_{n1} & \dots & a_{nm}
\end{pmatrix}
$$

Die elementare Matrix $S_i(a)$ ist invertierbar, und zwar:

$$S_i^{-1}(a) = \begin{pmatrix} 1 & 0 & & \dots & & & 0 \\ 0 & \ddots & & & & & \\ & & 1 & & & & \\ \vdots & & & a^{-1} & & & \vdots \\ & & & & 1 & & \\ & & & & & \ddots & 0 \\ 0 & & & \dots & & 0 & 1 \end{pmatrix} = S_i(a^{-1}).$$

Es gibt auch elementare $n \times n$ Matrizen S_{ij} und $S_{ij}(c)$, die die entsprechenden elementaren Zeilenoperationen durch Matrizenmultiplikation realisieren, nämlich:

$$S_{ij} = \begin{pmatrix} 1 & & & & & & & & & & \\ & \ddots & & & & 0 & & & & 0 & \\ & & 1 & & & & & & & & \\ & & & 0 & & \rightarrow & & 1 & & & \\ & & & & 1 & & & & & & \\ & 0 & & \downarrow & & \ddots & & \uparrow & & 0 & \\ & & & & & & 1 & & & & \\ & & & 1 & & \leftarrow & & 0 & & & \\ & & & & & & & & 1 & & \\ & 0 & & & & 0 & & & & \ddots & \\ & & & & & & & & & & 1 \end{pmatrix},$$

D.h $S_{ij} = (s_{kl})$, wobei $s_{kl} = \begin{cases} 1, & \text{falls } k = l \, (\neq i \text{ oder } j), \\ 1, & \text{falls } k = i, l = j \text{ oder } k = j, l = i, \\ 0, & \text{sonst.} \end{cases}$

$$S_{ij}(c) = \begin{pmatrix} 1 & & & & & & & & & & \\ & \ddots & & & & & & & & 0 & \\ & & 1 & & & & & & & & \\ & & & 1 & & \rightarrow & & c & & & \\ & & & & 1 & & & & & & \\ & & & & & \ddots & & \uparrow & & & \\ & & & & & & 1 & & & & \\ & & & & & & & 1 & & & \\ & & & & & & & & 1 & & \\ & 0 & & & & & & & & \ddots & \\ & & & & & & & & & & 1 \end{pmatrix}.$$

D.h $S_{ij}(c) = (t_{kl})$, wobei $t_{kl} = \begin{cases} 1, & \text{falls } k = l, \\ c, & \text{falls } k = i \text{ und } l = j, \\ 0, & \text{sonst.} \end{cases}$

Es ist nicht schwer zu sehen, daß $(S_{ij})^{-1} = S_{ij}$ und $(S_{ij}(c)))^{-1} = S_{ij}(-c)$. D.h. S_{ij}, $S_i(a)$ und $S_{ij}(c) \in GL(n, F)$, für alle relevante i, j, a, c. Es gilt noch viel mehr:

Satz 9.2. *$GL(n, F)$ wird von den elementaren Matrizen erzeugt.*

wobei

Definition 41. Sei G eine Gruppe und $H \subset G$ eine Teilmenge. G wird von H *erzeugt*, falls $\forall g \in G, \exists h_1, \dots, h_m \in H$ mit $g = h_1 \cdot h_2 \cdots h_m$.

Zunächst aber sei die invertierbare $n \times n$ Matrix $B = (b_{ij})$ in Zeilenstufenform. Da B invertierbar ist, ist die Anzahl der Stufen auch n. Wir werden sagen, daß eine solche Matrix in der 'oberen Dreiecksform' ist.

$$B = \begin{pmatrix} 1 & b_{12} & b_{13} & \dots & b_{1n} \\ 0 & 1 & b_{23} & \dots & b_{2n} \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \dots & 1 & b_{(m-1)n} \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix}.$$

Lemma. *Sei die invertierbare $n \times n$ Matrix $B = (b_{ij})$ in der oberen Dreiecksform. Dann existieren elementare Zeilenoperationen, die B zur Einheitsmatrix transformieren.*

Beweis. Induktion über n. Für $n = 1$ ist $b_{11} \neq 0$. Man multipliziere nun mit b_{11}^{-1}; dies entspricht einer elementaren Zeilenoperation.

Für $n > 1$ gilt, laut Induktionshypothese, daß die Aussage des Lemmas wahr ist für Matrizen mit $n-1 \times n-1$ Elementen. Insbesondere kann B durch elementare Zeilenoperationen in die folgende Form gebracht werden

$$\begin{pmatrix} b_{11} & b_{12} & b_{13} & \dots & b_{1n} \\ \hline 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}.$$

Aber durch Operationen des Types $S_{1j}(-b_{1j})$ können alle b_{1j}, auf 0 gebracht werden ($j > 1$). Zum Schluß wird eine Operation vom Typ $S_1(b_{11}^{-1})$ benutzt, um die Matrix in eine $n \times n$ Einheitsmatrix zu verwandeln. $\square$

Beweis des Satzes: Sei $A \in GL(n, F)$. Behauptung: $\exists$ elementare Matrizen $S_1, \dots, S_p$ mit $S_p \cdots S_1 \cdot A = I$, die $n \times n$ Einheitsmatrix. Warum? Jede elementare Zeilenoperation entspricht einer elementaren Matrix. Nach Satz 6.2 kann A durch endlich viele elementare Zeilenoperationen in die obere Dreiecksform gebracht werden, etwa $S_1, \dots, S_q$. D.h. die Matrix $S_p \cdots S_1 \cdot A$ ist in der oberen Dreiecksform. Aber nach dem Lemma existieren weitere elementare Matrizen $S_{q+1}, \dots, S_p$ so, daß

$$S_p \cdots S_{q+1} \cdot (S_q \cdots S_1 \cdot A) = I = \text{Einheitsmatrix}.$$

D.h. $(S_p \cdots S_{q+1} \cdot S_p \cdots S_1) \cdot A = I$, oder

$$A = (S_p \cdots S_1)^{-1} = S_1^{-1} \cdot S_2^{-1} \cdots S_p^{-1}.$$

Wie wir gesehen haben, ist jedes S_i^{-1} eine elementare Matrix. Da A beliebig in $GL(n, F)$ war, ist der Satz bewiesen. QED

9.2 Die elementaren Spaltenoperationen

Sei S eine elementare $n \times n$ Matrix und sei A eine $n \times m$ Matrix. Dann ist $S \cdot A$ die Matrix A, abgeändert durch die entsprechende Zeilenoperation S. Was aber ist $A \cdot S$? Wenn wir von *rechts* multiplizieren, dann bekommen wir ein völlig anderes Ergebnis, nämlich die Matrix A multipliziert mit einer entsprechenden *Spalten*operation. (Voraussetzung für eine sinnvolle Multiplikation ist, daß A jetzt eine $m \times n$ Matrix ist—d.h. A hat n Spalten.)

1. $A \cdot S_{ij}$ ist das Vertauschen der *Spalten* i und j.
2. Bei $A \cdot S_i(a)$ wird die i-te Spalte mit der Zahl $a/not = 0$ multipliziert.
3. Bei $A \cdot S_{ij}(c)$ wird die j-te Spalte mit der Zahl $c/not = 0$ multipliziert und zur i-ten Spalte addiert.

Wir können nun die ganze Theorie der elementaren *Zeilen*operationen nochmal anschauen und uns vorstellen, daß die Matrizen transponiert werden, so daß Zeilen plötzlich Spalten und Spalten plötzlich Zeilen sind. Es ist daher klar, daß es möglich ist, Matrizen umzuformen durch elementare Spaltenoperationen auf ähnlicher Weise, wie wir dies getan haben durch elementare Zeilenoperationen. Insbesondere:

- Sei die Matrix A durch eine elementare Spaltenoperation umgeformt. Dann bleibt der Spaltenraum invariant unter dieser Operation, aber der Zeilenraum wird im allgemeinen geändert. Der Zeilenrang (= Spaltenrang) bleibt jedoch unverändert.
- Es ist möglich, durch elementare Spaltenoperationen eine Matrix A umzuformen in eine Matrix in Spaltenstufenform. Falls $A \in GL(n, F)$, dann kann A durch elementare Spaltenoperationen in die Identitätsmatrix umgeformt werden.

9.3 Invariante Unterräume, Eigenvektoren, usw.

Definition 42. Sei $f : V \to V$ eine lineare Abbildung, $U \subset V$ ein Unterraum. U heißt *invariant* bezüglich f, falls $f(U) \subset U$.

Satz 9.3. *Sei U invariant bezüglich $f : V \to V$. Sei A die zu f gehörige Matrix bezüglich einer Basis $\{\alpha_1, \ldots, \alpha_n\}$ von V. Dann existiert eine Matrix $C \in GL(n,F)$ mit $B = CAC^{-1}$ ($B = (b_{ij})$), und ein $r \in \mathbf{N}$ mit $0 < r < n$, wobei $b_{ij} = 0$, falls $i > r$ und $j \le r$. D.h.*

$$B = \left(\begin{array}{ccc|ccc} b_{11} & \ldots & b_{1r} & b_{1(r+1)} & \ldots & b_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ b_{r1} & \ldots & b_{rr} & b_{r(r+1)} & \ldots & b_{rn} \\ \hline & & & b_{(r+1)(r+1)} & \ldots & b_{(r+1)n} \\ & 0 & & \vdots & & \vdots \\ & & & b_{n(r+1)} & \ldots & b_{nn} \end{array}\right).$$

Beweis. Sei $\{\beta_1, \ldots, \beta_r\}$ eine Basis für U. Nach dem Basisergänzungssatz, sei

$$\{\beta_1, \ldots, \beta_r, \beta_{r+1}, \ldots, \beta_n\}$$

eine Basis für V. Dann ist die zu f gehörige Matrix bezüglich $\{\beta_1, \ldots, \beta_n\}$ wie oben. (Vektoren sind Spalten!) Die Matrix C ist dann die Matrix, die den Basiswechsel $\{\alpha_i\} \to \{\beta_i\}$ darstellt (nach Satz 9.1). □

Definition 43. Die $n \times n$ Matrix $A = (a_{ij})$ ist in *Kästchenform*, falls p Zahlen $r_1, \ldots, r_p$ existieren, wobei $a_{ij} = 0$, falls kein k existiert, mit $r_k \le i, j < r_{k-1} - 1$. D.h.

$$A = \begin{pmatrix} A_1 & 0 & \ldots & & 0 \\ 0 & A_2 & 0 & & \\ \vdots & 0 & \ddots & 0 & \vdots \\ & & 0 & A_{k-1} & 0 \\ 0 & & \ldots & 0 & A_k \end{pmatrix},$$

und $A_i = \begin{pmatrix} a_{(r_i)(r_i)} & \ldots & a_{(r_i)(r_{i+1}-1)} \\ \vdots & & \vdots \\ a_{(r_{i+1}-1)(r_i)} & \ldots & a_{(r_{i+1}-1)(r_{i+1}-1)} \end{pmatrix}$, $i = 1, \ldots, k$.

Korollar 9.3.1. *Sei $f : V \to V$ und A die zu f gehörige Matrix bezüglich einer Basis $\{\alpha_1, \ldots, \alpha_n\}$. Angenommen, es existieren Unterräume $U_1, \ldots, U_k \subset V$, die invariant sind bezüglich f, und $U_1 \oplus U_2 \oplus \cdots \oplus U_k = V$. Dann existiert eine Matrix $C \in GL(n,F)$ mit CAC^{-1} in Kästchenform.*

Ein Hauptziel in der linearen Algebra ist es, 'gute' Basen für die Darstellung von vorgegebenen linearen Abbildungen $f : F^n \to F^n$ zu finden. Das heißt, gegeben eine beliebige $n \times n$ Matrix A, versuche irgendeine weitere Matrix $C \in GL(n,F)$ zu finden, so daß die äquivalente Matrix CAC^{-1} so 'einfach wie möglich' ist.

9.4 Ein Spezialfall

Definition 44. Sei $f : V \to V$ und sei $\alpha \in V$ ($\alpha \neq 0$) mit $f(\alpha) = \lambda \cdot \alpha$, $\lambda \in F$. Dann heißt α ein *Eigenvektor* von f mit *Eigenwert* λ. Allgemeiner, sei $\lambda \in F$ gegeben.

$$E_\lambda = \{\alpha \in V | f(\alpha) = \lambda\alpha\}$$

heißt der zu f gehörige *Eigenraum.*

Bemerkung. $E_\lambda \subset V$ ist ein Unterraum.

Beweis. Seien $\beta, \gamma \in E_\lambda$, $b, c \in F$. Dann ist

$$\begin{aligned} f(b\beta + c\gamma) &= bf(\beta) + cf(\gamma) \\ &= b(\lambda\beta) + c(\lambda\gamma) \\ &= \lambda(b\beta + c\gamma) \end{aligned}$$

□

Die Dimension des Eigenraums, $dim(E_\lambda)$, heißt die *geometrische Vielfachheit* von λ.

Satz 9.4. *Sei $f : V \to V$ linear mit Eigenvektoren $\xi_1, \ldots, \xi_m$ und entsprechenden Eigenwerten $\lambda_1, \ldots, \lambda_m$. Angenommen, die Eigenwerte sind paarweise verschieden ($\lambda_i = \lambda_j$, für $i = j$), dann ist die Menge $\{\xi_1, \ldots, \xi_m\}$ linear unabhängig.*

Beweis. Falls nicht, dann würden $z_1, \ldots, z_m \in F$ existieren, nicht alle null, mit $0 = \sum_{i=1}^{n} z_i \xi_i$.

Die ξ_i mit $z_i = 0$ sind uninteressant. Entferne daher alle solche ξ_i, so daß wir annehmen können, daß $z_i \neq 0$, $\forall i$. Weiter können wir annehmen, daß die Darstellung des Nullvektors $\sum_{i=1}^{m} z_i \xi_i$ die *kürzeste* mögliche nicht-triviale Darstellung ist. Wir haben sicherlich $m \geq 2$ (sonst wäre $0 \neq z_1 \xi_1$, aber da $z_1 \neq 0 \Rightarrow \xi_1 = 0$; ein Widerspruch). Dann gilt einerseits:

$$0 = \lambda_1 \cdot 0 = \lambda_1 \cdot \left(\sum_{i=1}^{m} z_i \xi_i \right) = \sum_{i=1}^{m} z_i \cdot \lambda_1 \xi_i.$$

Andererseits gilt:

$$0 = f(0) = f \left(\sum_{i=1}^{m} z_i \xi_i \right) = \sum_{i=1}^{m} z_i f(\xi_i) = \sum_{i=1}^{m} z_i \lambda_i \xi_i.$$

$$\Rightarrow \quad 0 = 0 - 0 = \sum_{i=1}^{m} z_i (\lambda_i - \lambda_1) \xi_i = \sum_{i=2}^{m} z_i (\lambda_i - \lambda_1) \xi_i.$$

(da $\lambda_1 - \lambda_1 = 0$) Dies ist jedoch eine *kürzere* nicht-triviale Darstellung des Nullvektors (da $(\lambda_i - \lambda_1 \neq 0, \forall i \geq 2)$; Widerspruch. □

Korollar 9.4.1.

1. *Sei V ein n-dimensionaler Raum, und sei $f : V \to V$ linear. Dann gibt es höchstens n <u>verschiedene</u> Eigenwerte zu f.*
2. *Falls n verschiedene Eigenwerte existieren, dann hat jeder Eigenwert die geometrische Vielfachheit 1, und es existiert eine Basis für V, bestehend aus lauter Eigenvektoren von f. Sei A die Matrix, die f darstellt, bezüglich irgendeiner Basis von V. Dann ist A zu einer diagonalen Matrix ähnlich. Wobei...*

Definition 45. Die $n \times n$ Matrix $A = (a_{ij})$ heißt *diagonal*, falls

$$a_{ij} = 0, \forall i \neq j.$$

Kapitel 10

Determinanten

Sei $f : \mathbf{R}^n \to \mathbf{R}^n$ eine lineare Abbildung des n-dimensionalen Euklidischen Raumes in sich selbst. Wir haben schon gesehen, daß solche Abbildungen im allgemeinen Drehungen und Verzerrungen nach sich ziehen. Inwieweit wird der Raum eigentlich verzerrt unter der Abbildung f? Ein grobes Maß für die Verzerrung kann wie folgt definiert werden.

Betrachten wir zunächst den üblichen Einheitswürfel

$$W_1 \equiv \{(x_1, \ldots, x_n) \in \mathbf{R}^n : x_i \in [0,1], \forall i\}.$$

Nach einer Standardrechnung ist $vol(W_1) = 1$, d.h. $1 \underbrace{\times \cdots \times}_{n\text{-mal}} 1 = 1$. Nun, das Bild von W_1 unter f, nämlich $f(W_1)$, ist im allgemeinen ein (schräges) Parallelogramm. Was ist das (n-dimensionale) Volumen von $f(W_1)$, d.h. $vol(f(W_1))$? Die Antwort:

$$\frac{vol(f(W_1))}{vol(W_1)} = vol(f(W_1)) = det(A),$$

wobei A die $n \times n$ Matrix ist, die f darstellt. D.h. die 'Determinante' der darstellenden Matrix A von f ist ein 'Maß' für die Verzerrung des Volumens von Würfeln.[1]

Bemerkung. Es kann sein, daß f eine orientierungsverkehrende Abbildung ist (z.B. eine Spiegelung). Dann ist $det(A)$ eine negative Zahl—sozusagen, ein 'negatives Volumen'.

10.1 Drei besondere Eigenschaften

Die Idee, daß $det(A)$ etwas mit Volumen von Würfeln zu tun hat gibt uns noch keine vernünftige Definition. Die übliche Definition ist wie folgt. (Und dies ist natürlich unsere 'offizielle' Definition!)

Definition 46. Sei $M(n \times n, F)$ die Menge der $n \times n$ Matrizen mit Elementen aus einem Körper F. Eine Abbildung $det : M(n \times n, F) \to F$ heißt eine *Determinantenfunktion*, falls die folgenden drei Eigenschaften gelten.

1. Sei I_n die Einheitsmatrix. Dann ist $det(I_n) = 1$.

2. Sei $A = \{\gamma_1, \ldots, \gamma_n\} \in M(n \times n, F)$, wobei $\gamma_1, \ldots, \gamma_n$ die Spaltenvektoren von A (in der richtigen Reihenfolge) sind, und sei $i \in \{1, \ldots, n\}$. Sei weiter $a \in F$ beliebig. Dann ist $det(A') = a \cdot det(A)$, wobei

   $$A' = \{\gamma_1, \ldots, \gamma_{i-1}, a \cdot \gamma_i, \gamma_{i+1}, \ldots, \gamma_n\}.$$

   D.h. wenn wir die i-te Spalte mit dem Faktor a multiplizieren, dann ändert sich der Wert der Determinantenfunktion ebenfalls um den Faktor a.

3. Sei wieder $A \in M(n \times n, F)$ und seien $i \neq j$ zwei Zahlen zwischen 1 und n. Nehmen wir diesesmal

   $$A^* = \{\gamma_1, \ldots, \gamma_{i-1}, \gamma_i + \gamma_j, \gamma_{i+1}, \ldots, \gamma_n\}.$$

   Das heißt, A^* entsteht indem man die j-te Spalte von A nimmt und diese zur i-ten Spalte addiert. Dann gilt $det(A^*) = det(A)$.

[1] Wenn Sie später die Maßtheorie studieren, dann werden Sie sehen, daß die Volumen von allen möglichen geometrischen Objekten im Raum durch approximierende Würfel definiert werden. Daher mißt die Determinante einer Abbildung $f : \mathbf{R}^n \to \mathbf{R}^n$ die Verzererung der Volumen von beliebigen geometrischen Objekten in $\mathbf{R}^n$.

Nun, die Einheitsmatrix ist die zur Identitätsabbildung gehörige Abbildungsmatrix. Im n-dimensionalen Euklidischen Raum (d.h. wenn $F = \mathbf{R}$) ist das Bild des Einheitswürfels unter der Identitätsabbildung natürlich der Einheitswürfel selbst, mit Volumen 1. Dies entspricht der ersten Eigenschaft einer Determinantenfunktion..

Betrachten wir jetzt die zweite Eigenschaft (wieder mit $F = \mathbf{R}$). Nach unserer altbekannten Regel entspricht dem i-ten Spaltenvektor in A das Bild des kanonischen Basisvektors $\epsilon_i \in \mathbf{R}^n$ unter der Abbildung f. Nun, diese kanonischen Basisvektoren laufen (ausgehend von der Ecke in $0 \in \mathbf{R}^n$) entlang verschiedenen Kanten von W_1. D.h. jeder Spaltenvektor von A ist eine Kante von $f(W_1)$. Was passiert, wenn die i-te Kante mit dem Faktor $a \in \mathbf{R}$ multipliziert wird? Antwort: der Würfel $f(W_1)$ wird in dieser 'Dimension' durch den Faktor a 'verlängert'. Insgesamt wird das Volumen der Würfel $f(W_1)$ um den Faktor a geändert.

Betrachten wir schließlich die dritte Eigenschaft. Wenn die j-te Spalte zur i-ten Spalte addiert wird, dann ist die geometrische Bedeutung, daß die i-te Kante von $f(W_1)$ 'schräg' gestellt wird in Richtung ϵ_j. Es ist aber ein Standardergebnis der Schulgeometrie, daß solche schräggestellten Würfel dasselbe Volumen haben wie die ursprünglichen nicht-schräggestellten Würfel.

10.2 Einfache Konsequenzen dieser Definition

1. Sei $S_j(a)$ die elementare *Spalten*operation, die die j-te Spalte mit dem Skalaren a multipliziert. D.h. wenn $S_j(a)$ die entsprechende elementare Matrix ist, dann wird diese Spaltenoperation durch Matrizenmultiplikation von rechts $A \to A \cdot S_j(a)$ realisiert. Es gilt dann
$$det(A \cdot S_j(a)) = det(\gamma_1, \ldots, \gamma_{j-1}, a \cdot \gamma_j, \gamma_{j+1}, \ldots, \gamma_n) = a \cdot det(A).$$
Dies ist die Eigenschaft 2 unserer Definition.

2. Als besonderer Fall sei angenommen, daß die Matrix A eine Nullspalte (bestehend aus lauter Nullen) besitzt. Dann ist $det(A) = 0$. Warum? Sei etwa die j_0-te Spalte Null (wobei $1 \le j_0 \le n$). D.h. $A = (a_{ij})$, $i, j = 1, \ldots, n$, und $a_{ij_0} = 0$, $\forall i$. Sei jetzt $a \in F$ irgendein Element des Körpers — z.B. $a = 0$. Dann ist, laut Eigenschaft 2, $det(A) = a \cdot det(A) = 0 \cdot det(A) = 0$.

3. Sei S_{ij} die elementare Spaltenoperation, die die Zeilen i und j miteinander vertauscht. Dann ist $det(A) = -det(A \cdot S_{ij})$. D.h. das Vertauschen von Spalten bewirkt, daß das Vorzeichen der Determinantenfunktion sich ändert. Denn
$$\begin{aligned} det(A) &= det(\gamma_1, \ldots, \gamma_i, \ldots, \gamma_j, \ldots, \gamma_n) \\ &= det(\gamma_1, \ldots, \gamma_i + \gamma_j, \ldots, \gamma_j, \ldots, \gamma_n) \\ &= det(\gamma_1, \ldots, \gamma_i + \gamma_j, \ldots, \gamma_j - (\gamma_i + \gamma_j), \ldots, \gamma_n) \\ &= det(\gamma_1, \ldots, \gamma_i + \gamma_j, \ldots, -\gamma_i, \ldots, \gamma_n) \\ &= det(\gamma_1, \ldots, \gamma_i + \gamma_j - \gamma_i, \ldots, -\gamma_i, \ldots, \gamma_n) \\ &= det(\gamma_1, \ldots, \gamma_j, \ldots, -\gamma_i, \ldots, \gamma_n) \\ &= -det(\gamma_1, \ldots, \gamma_j, \ldots, \gamma_i, \ldots, \gamma_n) \\ &= -det(A \cdot S_{ij}) \end{aligned}$$

4. Angenommen, zwei Spalten von A seien identisch. Dann ist $det(A) = 0$. Dies folgt, weil das Vertauschen von zwei identischen Spalten die Matrix gar nicht ändern kann. Daher ist $det(A) = -det(A) \Rightarrow det(A) = 0$.

5. Sei $S_{ij}(c)$, wobei $c \neq 0$, die elementare Spaltenoperation, die die j-te Spalte, multipliziert mit dem Skalaren c, zur i-ten Spalte addiert. Dann bleibt die Determinantenfunktion unverändert. D.h. es gilt $det(A) = det(A \cdot S_{ij}(c))$. Dies folgt, da
$$\begin{aligned} det(A) &= det(\gamma_1, \ldots, \gamma_i, \ldots, \gamma_j, \ldots, \gamma_n) \\ &= c^{-1} \cdot det(\gamma_1, \ldots, \gamma_i, \ldots, c \cdot \gamma_j, \ldots, \gamma_n) \\ &= c^{-1} \cdot det(\gamma_1, \ldots, \gamma_i + c \cdot \gamma_j, \ldots, c \cdot \gamma_j, \ldots, \gamma_n) \\ &= det(\gamma_1, \ldots, \gamma_i + c \cdot \gamma_j, \ldots, \gamma_j, \ldots, \gamma_n) \end{aligned}$$

6. Beliebige Kombinationen von Spaltenoperationen können auf die Matrix A angewandt werden, aber nach dem Vorangegangenen gibt es für jede solche Operation eine eindeutige Auswirkung auf die Determinantenfunktion. Insbesondere ist es möglich, eine singuläre Matrix A durch elementare Spaltenoperationen so umzuformen, daß eine Spalte von A null wird. Daher gilt: A ist singulär $\Rightarrow det(A) = 0$.

7. Andererseits, falls A regulär ist, dann ist es möglich, A so umzuformen, daß A zur Einheitsmatrix I_n wird. Da $det(I_n) = 1$ und da keine elementare Spaltenoperation eine Matrix mit Determinantenfunktion null in eine Matrix mit Determinantenfunktion ungleich null umwandeln kann, folgt: A ist regulär $\Rightarrow det(A) \neq 0$. Wenn wir dies mit dem Vorherigen kombinieren, dann haben wir sogar: A ist regulär $\Leftrightarrow det(A) \neq 0$.

8. Schließlich können wir noch mehr sagen. Falls A regulär ist, dann ist es möglich, durch elementare Spaltenoperationen A umzuformen in die Einheitsmatrix I_n. Da wir wissen, was $det(I_n)$ ist (nämlich 1), und da die Auswirkungen der elementaren Spaltenoperationen auf die Determinantenfunktion eindeutig sind, folgt, daß $det(A)$ auch eindeutig festgelegt sein muß.

Insgesamt gilt also

Satz 10.1. *Falls eine Determinantenfunktion existiert, dann ist sie eindeutig und wird durch das gerade beschriebene Verfahren gegeben. Die Determinante hat den Wert Null genau dann, wenn die dargestellte lineare Abbildung singulär ist.*

10.3 Noch ein Computerprogram

```
program matrix_invers (input, output);

const
   grosse = 3;
   { D.h. 3x3 Matrizen. Sie konnen auch andere Zahlen nehmen }
type
   matrix = array [1..grosse, 1..grosse] of real;
var
   mat, mat_inv : matrix;
   invertierbar : boolean;

procedure lese_matrix (var m : matrix);
var
   i, j : integer;
begin
   for i := 1 to grosse do
      for j := 1 to grosse do read (m[i,j]);
end;

procedure schreibe_matrix (m : matrix);
var
   i, j : integer;
begin
   for i := 1 to grosse do begin
      for j := 1 to grosse do write (m[i,j]);
      writeln;
   end;
end;

{ ########### ELEMENTARE ZEILENOPERATIONEN ################## }

procedure tausch_ij (i, j : integer; var m : matrix);
{ Die Zeilen i und j werden vertauscht }
var
   s : integer; temp : real;
begin
   for s := 1 to grosse do begin
      temp := m[i,s];
      m[i,s] := m[j,s];
      m[j,s] := temp;
   end;
end;

procedure mult_ia (i : integer; a : real; var m : matrix);
```

```
{ Zeil i mit a multipliziert }
var
   s : integer;
begin
   for s := 1 to grosse do m[i,s] := a * m[i,s];
end;

procedure add_zeil (i, j : integer; c : real; var m : matrix);
{ Zeil j multipliziert mit c wird zum Zeil i addiert }
var
   s : integer; temp : real;
begin
   for s := 1 to grosse do
      m[i,s] := m[i,s] + c * m[j,s];
end;

{ ############## HAUPTPROZEDUR ###################### }

procedure inverse (    m        : matrix;
                   var  m_inv   : matrix;
                   var regulaer : boolean);
label 1;
var
   z, i, j : integer; temp : real;
begin
   regulaer := true;

   { initializiere m_inv als Einheitsmatrix }
   for z := 1 to grosse do
      for i := 1 to grosse do if z = i
         then m_inv[z,i] := 1 else m_inv[z,i] := 0;

   { Zeilenoperationen, um m in die obere Dreiecksform zu bringen }
   for z := 1 to grosse do begin
      i := z;
      while (m[i,z] = 0) and (i <= grosse) do i := i+1;

      { pruefe, ob die Matrix regulaer ist }
      if i > grosse then begin regulaer := false; goto 1; end;

      { bis jetzt regulaer ... daher weitermachen }
      if i > z then begin
         tausch_ij (i, z, m);
         tausch_ij (i, z, m_inv);
      end;
      for i := z+1 to grosse do
         if m[i,z] <> 0 then begin
            temp := m[i,z] / m[z,z];
            add_zeil (i, z, -temp, m);
            add_zeil (i, z, -temp, m_inv);
         end;
   end;

   { m in obere Dreiecksform -> jetzt Diagonalform }
   for i := grosse downto 2 do begin
      for j := i-1 downto 1 do begin
         temp := m[j,i] / m[i,i];
         add_zeil (j, i, -temp, m);
         add_zeil (j, i, -temp, m_inv);
      end;
   end;
```

```
   {  m in Diagonalform -> jetzt normierung zur Einheitsmatrix }
   for z := 1 to grosse do begin
      temp := 1 / m[z,z];
      mult_ia (z, temp, m);
      mult_ia (z, temp, m_inv);
   end;
   1: end;
{ ******************************************************* }

{ EINE EINFACHE DETERMINANTEN FUNKTION KANN NUN WIE FOLGT FORMULIERT WERDEN }

function det (m : matrix) : real;
label 1;
var
   z, i              : integer;
   determinant, temp : real;

begin
   determinant := 1;

   { m WIRD IN DIE OBERE DREIECKSFORM GEBRACHT }
   for z := 1 to grosse do begin
      i := z;
      while (m[i,z] = 0) and (i <= grosse) do i := i+1;

      { PRUEFE OB DIE MATRIX REGULAER IST }
      if i > grosse then begin det := 0; goto 1; end;

      { BIS JETZT REGULAER ... DAHER WEITERMACHEN }
      if i > z then begin
         tausch_ij (i, z, m);
         determinant := - determinant;
      end;
      for i := z+1 to grosse do
         if m[i,z] <> 0 then begin
            temp := m[i,z] / m[z,z];
            add_zeil (i, z, -temp, m);
         end;
   end;

   { m IST IN OBERE DREIECKSFORM ->
    DETERMINANT IST PRODUKT DER DIAGONALEN ELEMENTEN }
   for z := 1 to grosse do
      determinant := determinant * m[z,z];
   det := determinant;
   1: end;
{ ******************************************************** }

begin
   writeln ('Die Matrix?');
   lese_matrix (mat);
   inverse (mat, mat_inv, invertierbar);
   if invertierbar
      then
   begin
      writeln ('Das Inverse ist :');
      schreibe_matrix (mat_inv);
   end
   else writeln ('nicht invertierbar');
   writeln;
```

```
   write ('Die Determinante ist : ');
   writeln (det (mat));
end.
```

Kapitel 11

Die Existenz von Determinantenfunktionen

Unser Computerprogramm 'matrix_invers' beinhaltet die Funktion 'det', die die Eigenschaften (1) - (3) einer Determinantenfunktion erfüllt. Aber wie können wir dies beweisen? Ein Computerprogramm ist doch nicht besonders geeignet als Ausgangspunkt für rationale menschliche Gedanken!

Ich werde daher eine andere Beschreibung dieser Determinantenfunktion geben, in der Hoffnung, daß Sie diese Funktion dadurch gut begreifen werden. Auch diese Beschreibung ist nicht ganz einfach, aber Sie werden diese Beschreibung in allen üblichen Lehrbüchern über die lineare Algebra finden. (Eine mehr computergerechte Beschreibung fehlt meistens.) Zunächst sollten wir daran denken, daß für jedes $n \in \mathbf{N}$ (und für jeden Körper F) eine andere Determinantenfunktion gesucht wird. Es ist daher naheliegend, diese Funktionen rekursiv zu definieren. D.h. zuerst wird der Fall $n = 1$ behandelt. Dann wird für jedes $n > 1$ zunächst angenommen, daß die Determinantenfunktion schon festgelegt ist für alle Zahlen m zwischen 1 und $n-1$. Anhand dieser Festlegungen wird die Determinantenfunktion dann für den Fall n definiert.

Nun, der Fall $n = 1$ ist sehr einfach. Sei $A \in M(1 \times 1, F)$. D.h. $A = (a)$, eine Matrix, bestehend aus einer einzigen Zahl a. Falls $a = 0$, dann muß (nach Eigenschaft (2)) $det(A) = 0$ gelten. Falls $a \neq 0$ ist (wieder nach Eigenschaft (2)) $det(A) = a \cdot det(I_1) = a \cdot 1 = a$, wobei I_1 die 1×1 Einheitsmatrix, bestehend aus der einzigen Zahl 1 ist. Auf jeden Fall gilt dann $det(A) = det((a)) = a$.

Wie ist es, wenn $n > 1$? Sei $A \in M(n \times n; F)$, etwa

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1(n-1)} & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2(n-1)} & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{(n-1)1} & a_{(n-1)2} & \cdots & a_{(n-1)(n-1)} & a_{(n-1)n} \\ a_{n1} & a_{n2} & \cdots & a_{n(n-1)} & a_{nn} \end{pmatrix}.$$

Wir bezeichnen mit A_{ij} dann die $(n-1) \times (n-1)$ Matrix, die entsteht, wenn man die i-te Zeile und die j-te Spalte von A wegstreicht. (Nennen wir dies die i, j-te *Untermatrix* von A.)

$$A = \left(\begin{array}{ccc|c|ccc} a_{11} & \cdots & a_{1(j-1)} & a_{1j} & a_{1(j+1)} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{(i-1)1} & \cdots & a_{(i-1)(j-1)} & a_{(i-1)j} & a_{(i-1)(j+1)} & \cdots & a_{(i-1)n} \\ \hline a_{i1} & \cdots & a_{i(j-1)} & a_{ij} & a_{i(j+1)} & \cdots & a_{in} \\ \hline a_{(i+1)1} & \cdots & a_{(i+1)(j-1)} & a_{(i+1)j} & a_{(i+1)(j+1)} & \cdots & a_{(i+1)n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n(j-1)} & a_{nj} & a_{n(j+1)} & \cdots & a_{nn} \end{array}\right)$$

$$\longrightarrow A_{ij} = \left(\begin{array}{ccc|ccc} a_{11} & \cdots & a_{1(j-1)} & a_{1(j+1)} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{(i-1)1} & \cdots & a_{(i-1)(j-1)} & a_{(i-1)(j+1)} & \cdots & a_{(i-1)n} \\ \hline a_{(i+1)1} & \cdots & a_{(i+1)(j-1)} & a_{(i+1)(j+1)} & \cdots & a_{(i+1)n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n(j-1)} & a_{n(j+1)} & \cdots & a_{nn} \end{array}\right).$$

Unsere Definition für $det : M(n \times n; F) \to F$ lautet jetzt: Sei A eine beliebige Matrix in $M(n \times n; F)$. Dann ist

$$det(A) = \sum_{j=1}^{n} (-1)^{n-j} a_{nj} \, det(A_{nj}).$$

Satz 11.1. *Die so definierte Funktion ist eine Determinantenfunktion.*

Um diesen Satz zu beweisen, brauchen wir aber zunächst zwei andere Ergebnisse.

Lemma. *Sei $A = \gamma_1, \ldots, \gamma_n$ eine $n \times n$ Matrix, ausgedrückt durch Spalten, und sei $det : M(n \times n; F) \to F$ eine Determinantenfunktion. Seien ξ, ζ irgendwelche Spaltenvektoren in F^n. Sei $A' = \gamma_1, \ldots, \gamma_{i-1}, \xi + \zeta, \gamma_{i+1}, \ldots, \gamma_n$. Dann ist $det(A') = det(A_\xi) + det(A_\zeta)$, wobei*

$$A_\xi = \gamma_1, \ldots, \gamma_{i-1}, \xi, \gamma_{i+1}, \ldots, \gamma_n \quad \textit{und}$$

$$A_\zeta = \gamma_1, \ldots, \gamma_{i-1}, \zeta, \gamma_{i+1}, \ldots, \gamma_n.$$

Beweis. Die Menge $\{\xi, \zeta, \gamma_1, \ldots, \gamma_{i-1}, \gamma_{i+1}, \ldots, \gamma_n\}$ enthält $n+1$ Spaltenvektoren aus F^n. Sie ist daher zwangsläufig linear abhängig. Seien $x, z, c_j \in F$, nicht alle null, mit

$$0 = x\xi + z\zeta + \sum_{j=i} c_j \gamma_j.$$

Falls sowohl $x = 0$ als auch $z = 0$, dann ist die Menge

$$\{\gamma_1, \ldots, \gamma_{i-1}, \gamma_{i+1}, \ldots, \gamma_n\}$$

linear abhängig, und folglich (laut Eigenschaft 6 des letzten Kapitels) ist $det(A') = det(A_\xi) = det(A_\zeta) = 0$, und das Lemma ist wahr. Nehmen wir daher an, daß $x \neq 0$ (der Fall $z \neq 0$ ist ähnlich). Es gilt insbesondere, daß

$$\xi = \left(\frac{-z}{x}\right) \zeta + \sum_{j/not=i} \left(\frac{-c_j}{x}\right) \gamma_j.$$

Nach Eigenschaft 5 des letzten Kapitels ist dann

$$\begin{aligned}
det(A_\xi) &= det(\gamma_1, \ldots, \gamma_{i-1}, \xi, \gamma_{i+1}, \ldots, \gamma_n) \\
&= det(\gamma_1, \ldots, \gamma_{i-1}, \left(\frac{-z}{x}\right) \zeta + \sum_{j/not=i} \left(\frac{-c_j}{x}\right) \gamma_j, \gamma_{i+1}, \ldots, \gamma_n) \\
&= det(\gamma_1, \ldots, \gamma_{i-1}, \left(\frac{-z}{x}\right) \zeta, \gamma_{i+1}, \ldots, \gamma_n) \\
&= \left(\frac{-z}{x}\right) det(\gamma_1, \ldots, \gamma_{i-1}, \zeta, \gamma_{i+1}, \ldots, \gamma_n) \\
&= \left(\frac{-z}{x}\right) det(A_\zeta)
\end{aligned}$$

Es gilt aber auch

$$\begin{aligned}
det(A') &= det(\gamma_1, \ldots, \gamma_{i-1}, \xi + \zeta, \gamma_{i+1}, \ldots, \gamma_n) \\
&= det(\gamma_1, \ldots, \gamma_{i-1}, \zeta + \left(\frac{-z}{x}\right) \zeta + \sum_{j/not=i} \left(\frac{-c_j}{x}\right) \gamma_j, \gamma_{i+1}, \ldots, \gamma_n) \\
&= det(\gamma_1, \ldots, \gamma_{i-1}, \zeta + \left(\frac{-z}{x}\right) \zeta, \gamma_{i+1}, \ldots, \gamma_n) \\
&= \left(1 - \frac{z}{x}\right) det(\gamma_1, \ldots, \gamma_{i-1}, \zeta, \gamma_{i+1}, \ldots, \gamma_n) \\
&= \left(1 - \frac{z}{x}\right) det(A_\zeta) \\
&= det(A_\zeta) + det(A_\xi)
\end{aligned}$$

$\square$

Um das zweite Ergebnis zu beschreiben, wird ein bestimmter Begriff—der später öfters auftauchen wird—benötigt.

Definition 47. Sei $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$ eine Bijektion der ersten n Integerzahlen. Dann heißt σ eine *Permutation* dieser Zahlen. Die Menge aller Permutationen der ersten n Integerzahlen ist eine Gruppe, genannt S_n (die symmetrische Gruppe der n-ten Ordnung). Sei σ eine Permutation der Zahlen $\{1, \ldots, n\}$. Sei $s(\sigma)$ die Anzahl der Zahlenpaare (i, j) mit $1 \le i < j \le n$ und $\sigma(i) > \sigma(j)$. Das *Signum* von σ ist dann

$$sign(\sigma) \equiv (-1)^{s(\sigma)}.$$

Eine besondere Klasse von Permutationen ist die Menge der Transpositionen. Eine Transposition $\tau \in S_n$ vertauscht einfach zwei verschiedene Zahlen miteinander. Alle übrigen Zahlen in $\{1, \ldots, n\}$ bleiben bei der Transposition unverändert.

Satz 11.2. *Jede Transposition τ hat $sign(\tau) = -1$.*

Beweis. Sei τ eine Transposition von $\{1, \ldots, n\}$. Das heißt, τ vertauscht etwa die zwei Zahlen i und j (wir nehmen $i < j$). Sei $k = j - i$. Falls $k = 1$ dann ist offensichtlich $s(\tau) = 1$, daher $sign(\tau) = -1$. Für $k > 1$ gibt es $k - 1$ Zahlen zwischen i und j. Für jede dieser Zahlen—etwa die Zahl l—zwischen i und j gilt sowohl $i < l$ mit $\tau(i) > \tau(l) = l$ als auch $l < j$ mit $l = \tau(l) > \tau(j)$. Aber auch $\tau(i) > \tau(j)$. Zusammen haben wir also $2k + 1$ solcher 'Fehlstände'. Folglich ist $sign(\tau) = (-1)^{2k+1} = -1$. □

Satz 11.3. *Sei $\sigma, \tau \in S_n$, wobei τ eine Transposition ist. Dann gilt $sign(\tau \cdot \sigma) = -sign(\sigma)$.*

Beweis. Fast identisch wie der Beweis von Satz 11.2. Angenommen, τ vertauscht die zwei Zahlen i und j, wobei $i < j$. Für die $k = j - i$ Zahlen dazwischen gilt folgendes: Sei $i < l < j$. Falls unter σ die Zahlen an den Stellen i und l einen Fehlstand ergeben (d.h. falls $\sigma^{-1}(i) > \sigma^{-1}(l)$), dann ist nach der Transposition τ *kein* Fehlstand mehr zwischen die Zahlen an den Stellen l und j. Dies ist doch klar, da die Zahl an der Stelle i durch τ an die Stelle j transportiert wird.

Andererseits, falls unter σ kein Fehlstand zwischen den Zahlen an die Stellen i und l steht, dann entsteht ein neuer Fehlstand zwischen l und j, nachdem die Transposition τ ausgeführt wird. Auf jeden Fall *ändert sich* das Vorzeichen für das Signum der Permutation. Die $k - 1$ Zahlen l zwischen i und j ergeben also $k - 1$ Vorzeichenwechseln. Aber auch zwischen l und j gibt es $k - 1$ Vorzeichenwechseln. Schliesslich ändert sich das Vorzeichen noch einmal in dem Vertausch von i und j. Dies gibt insgesamt $2k+1$ Vorzeichenwechseln gegenüber der Ursprunglichen Situation bei σ. Da $2k + 1$ ungerade ist, folgt das Ergebnis. □

Beispiel:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 1 & 2 \end{pmatrix};$$

dann sind die Paare $(4, 3)$, $(4, 1)$, $(4, 2)$, $(3, 1)$ und $(3, 2)$ Fehlstände unter σ. Also ist $sign(\sigma) = -1$. Sei jetzt

$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix}.$$

D.h. τ vertauscht die Zahlen 1 und 3. Dann ist

$$\tau \cdot \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 3 & 2 \end{pmatrix},$$

und die Fehlstände sind jetzt $(4, 1)$, $(4, 3)$, $(4, 2)$ und $(3, 2)$. Also ist $sign(\tau\sigma) = +1$.

Satz 11.4. *Jede Permutation σ kann als Produkt von endlich vielen Transpositionen $\sigma = \tau_1 \cdot \tau_2 \cdots \tau_q$ realisiert werden. Es gilt dann $sign(\sigma) = (-1)^q$.*

Beweis. Sei p die Anzahl der Elemente in der Menge

$$\{i \in \{1, \ldots, n\} : i \neq \sigma(i)\}.$$

Falls $p = 0$, dann brauchen wir keine Transpositionen. Sei daher $p > 0$. Der Fall $p = 1$ kann offensichtlich nicht vorkommen. Falls $p = 2$, dann haben wir es schon mit einer Transposition zu tun. Der interessante Fall ist $p > 2$. Sei dann $i_0 \in \{1, \ldots, n\}$ die kleinste Zahl mit $i_0 \neq \sigma(i_0)$. Sei $j_0 = \sigma^{-1}(i_0)$ — d.h. $\sigma(j_0) = i_0$. Als nächstes führen wir die Transposition τ_0 aus, die die Zahlen i_0 und j_0 miteinander vertauscht. Dann ist $\sigma \cdot \tau_0$ eine neue Permutation, aber die Zahl p ist hier kleiner. (Dies folgt, da jetzt $\sigma \cdot \tau_0(i_0) = i_0$. Ausserdem war schon $\sigma(j_0) \neq j_0$. Für alle andere Zahlen $k \neq i_0$ oder j_0 gilt $\sigma(k) = \sigma \cdot \tau(k)$.) Nach der induktiven Hypothese ist dann $\sigma \cdot \tau_0 = \tau_1 \cdots \tau_m$, für entsprechende Transpositionen $\tau_1, \ldots, \tau_n$. Folglich ist $\sigma = \tau_1 \cdots \tau_m \cdot \tau_0$. Die Aussage über $sign(\sigma)$ folgt aus Satz 11.3. □

Korollar 11.4.1. *Seien $\sigma_1, \sigma_2 \in S_n$. Dann ist*

$$sign(\sigma_1 \cdot \sigma_2) = sign(\sigma_1) \cdot sign(\sigma_2).$$

Definition 48. Eine Permutation σ heißt *gerade* oder *ungerade* je nachdem, ob eine gerade oder eine ungerade Anzahl von Transpositionen nötig sind, um σ darzustellen. Die Menge der geraden Permutationen ist eine Untergruppe $A_n \subset S_n$, genannt die *alternierende* Gruppe der Ordnung n.

Lemma. *Seien $A = (a_{ij})$ und $A' = (a'_{ij})$ zwei $n \times n$ Matrizen, wobei A' sich von A nur durch Umordnung einiger Spalten unterscheidet. D.h. $\exists$ eine Permutation $\sigma \in S_n$ mit $a'_{ij} = a_{\sigma(i)j}$, $\forall i, j$. Dann gilt $det(A') = sign(\sigma)\, det(A)$.*

Beweis. Nach der Konsequenz 3 des letzten Kapitels gilt für beliebige Spaltenvertauschungsoperationen S_{ij} die Gleichung

$$det(A) = -det(A \cdot S_{ij}).$$

σ kann als Produkt von Transpositionen—d.h. Spaltenvertauschungen—dargestellt werden. Nach dem ersten Korollar ist dann $sign(\sigma) = (-1)^m$ wobei m die Anzahl dieser Vertauschungen ist. ☐

Beweis des Satzes 11.1: Wir müssen die drei Eigenschaften nachprüfen.

1. $det(I_n) = 1$. Nun, es gilt $I_n = (\delta_{ij})$, wobei

$$\delta_{ij} = \begin{cases} 1, & \text{falls } i = j, \\ 0, & \text{sonst.} \end{cases}$$

   Da alle Elemente der letzten Zeile Null sind, außer des letzten, und auch die n,n-te Untermatrix von I_n nichts anderes als I_{n-1} ist, gilt

$$\begin{aligned} det(I_n) &= \sum_{j=1}^{n} (-1)^{n-j} \delta_{nj}\, det((I_n)_{nj}) \\ &= (-1)^{n-n} \times 1 \times det(I_{n-1}) \\ &= det(I_{n-1}) = 1. \end{aligned}$$

   Der letzte Schritt ($det(I_{n-1}) = 1$) folgt wegen der induktiven Hypothese.

2. Sei $A \in M(n \times n; F)$ und $a \in F$ mit $a \neq 0$. Angenommen, $A' = (a'_{ij})$ sei die Matrix, die entsteht, wenn die k-te Spalte von A mit dem Faktor a multipliziert wird. Zu zeigen:

$$det(A') = a \cdot det(A).$$

   Nun, es ist sicherlich so, daß dann in jeder Untermatrix A'_{ij} von A', mit $j \neq k$, auch eine Spalte mit dem Faktor a multiplizipiert wird. Daher gilt, jeweils nach der induktiven Hypothese, $det(A'_{ij}) = a \cdot det(A_{ij})$. Aber auch $a'_{ik} = a \cdot a_{ik}$, da a'_{ik} sich doch in der k-ten Spalte befindet. Daher gilt

$$\begin{aligned} det(A') &= \sum_{j=1}^{n} (-1)^{n-j} a'_{nj}\, det(A'_{nj}) \\ &= (-1)^{n-k} a \cdot a_{jk}\, det(A'_{nk}) + \sum_{j \neq k} (-1)^{n-j} a_{nj}\, det(A'_{nj}) \\ &= a \cdot (-1)^{n-k} a_{jk}\, det(A_{nk}) + a \cdot \sum_{j \neq k} (-1)^{n-j} a_{nj}\, det(A_{nj}) \\ &= a \cdot \sum_{j=1}^{n} (-1)^{n-j} a_{nj}\, det(A_{nj}) \\ &= a \cdot det(A). \end{aligned}$$

3. Sei $A \in M(n \times n; F)$. Angenommen, $A' = (a'_{ij})$ sei die Matrix, die entsteht, wenn die k-te Spalte von A zur l-ten Spalte addiert wird ($k \neq l$). Zu zeigen: $det(A') = det(A)$. Nun, es gilt sicherlich, daß $A'_{nl} = A_{nl}$. (Die l-te Spalte—die anders ist bei A' im Vergleich zu A—ist ja gerade, was fehlt bei A'_{nl}.) Für $j \neq k, l$ gilt aber, daß A'_{nj} aus der Untermatrix A_{nj} entsteht, indem eine Spalte von A_{nj} zu einer anderen Spalte addiert wird. Daher ist (nach der induktiven Hypothese und der dritten Eigenschaft in der Definition einer

Determinantenfunktion) $det(A'_{nj}) = det(A_{nj})$ für $j \neq k$. Wir wissen auch, daß die Elemente $a'_{nj} = a_{nj}$, falls $j \neq l$, und $a'_{nl} = a_{nl} + a_{nk}$. Daher gilt

$$\begin{aligned}
det(A) - det(A') &= \\
&= \sum_{j=1}^{n} (-1)^{n-j} a_{nj}\, det(A_{nj}) - \sum_{j=1}^{n} (-1)^{n-j} a'_{nj}\, det(A'_{nj}) \\
&= (-1)^{n-l} \left[a_{nl}\, det(A_{nl}) - (a_{nl} + a_{nk})\, det(A'_{nl})\right] \\
&\quad\quad + (-1)^{n-k} a_{nk} \left[det(A_{nk}) - det(A'_{nk})\right] \\
&= (-1)^{n-l}[-a_{nk}\, det(A_{nl})] + (-1)^{n-k} a_{nk}[det(A_{nk}) - det(A'_{nk})] \\
&= a_{nk}(-1)^{n-l} \left\{-det(A_{nl}) + (-1)^{l-k}(det(A_{nk}) - det(A'_{nk}))\right\}
\end{aligned}$$

Wir brauchen daher nur zu zeigen, daß

$$-det(A_{nl}) + (-1)^{l-k}(det(A_{nk}) - det(A'_{nk})) = 0.$$

D.h. $det(A'_{nk}) = det(A_{nk}) + (-1)^{l-k-1} det(A_{nl})$. Sei B die Matrix, die entsteht, wenn wir die l-te Spalte (in der ursprünglichen Matrix A) entfernen aus der Matrix A_{nk} und diese Spaltee ersetzen mit der entsprechenden k-ten Spaltee. Nach dem ersten Lemma ist dann $det(A'_{nk}) = det(A_{nk}) + det(B)$. Aber die Matrix B hat jetzt genau dieselben Spalten wie A_{nl}, nur in einer anderen Reihenfolge. Welche Reihenfolge? Um von der Matrix B zur Matrix A_{nl} zu gelangen, braucht man nur die l-te Spalte (immer gezählt in der ursprünglichen Matrix A) zu entfernen, dann müssen alle Spalten zwischen der l-ten und der k-ten Spalte um eine Stelle verschoben werden, um Platz an der k-ten Stelle zu schaffen. (Es sind $|l-k|-1$ solche Spalten.) Dort wird dann die herausgenommene Spalte wieder eingefügt. Diese Operation kann als eine Folge von Transpositionen realisiert werden. Jedesmal wird die Spalte, die ursprünglich an der l-ten Stelle war, an einer der dazwischenliegenden Spalten vorbeigeschoben. Nach unserem zweiten Lemma ist dann $det(B) = (-1)^{l-k-1} det(A_{nl})$. QED

Kapitel 12

Weitere Eigenschaften der Determinantenfunktion

12.1 Zwei besondere Fälle

Sei $A = (a_{ij})$ eine $n \times n$ Matrix mit Elementen aus einem Körper F. Wir haben im letzten Kapitel gesehen, daß

$$det(A) = \sum_{j=1}^{n} (-1)^{n-j} a_{nj} det(A_{nj}).$$

Diese Formel ist vielleicht schön anzuschauen, aber sie ist natürlich völlig ungeeignet für den praktischen Gebrauch.[1] Überlegen wir uns, wieviele Schritte nötig sind für die Ausrechnung von $det(A)$. Falls $n = 1$, dann gibt es nur einen Schritt. Falls A eine 2×2 Matrix ist, dann haben wir eine Summe von zwei Termen, die jeweils eine Multiplikation enthalten. Bei einer 3×3 Matrix A, haben wir drei Terme, die jeweils eine Multiplikation und auch die Ausrechnung einer 2×2 Determinante enthalten. Es ist nicht schwer einzusehen, daß die Rechnung für eine $n \times n$ Matrix mindestens $n!$ Schritte benötigt—und dies ist viel mehr als bei unserem Computerprogramm. Trotzdem werden die besonderen Fälle $n = 2$ und $n = 3$ immer getrennt behandelt, da die Rechnung dann doch noch einfach ist.

$$det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = (-1)^{2-1} a_{21} a_{12} + (-1)^{2-2} a_{22} a_{11} = a_{11} a_{22} - a_{12} a_{21}.$$

Sie haben diese Formel schon in einer früheren Übung geprüft (es ging um den Flächeninhalt von 2-dimensionalen Parallelogrammen).

Der Fall $n = 3$ ist etwas komplizierter.

$$det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{31} det(A_{31}) - a_{32} det(A_{32}) + a_{33} det(A_{33}) =$$

$$= a_{31} a_{12} a_{23} - a_{31} a_{13} a_{22} - a_{32} a_{11} a_{23} + a_{32} a_{13} a_{21} + a_{33} a_{11} a_{22} - a_{33} a_{12} a_{21}.$$

Dies heißt die ‘Regel von Sarrus’. Das Rechenschema für 3×3 Matrizen wird oft wie folgt dargestellt:

$$\left(\begin{array}{ccccccccc}
a_{11} & & a_{12} & & a_{13} & & a_{11} & & a_{12} \\
 & & & \searrow\swarrow & & \searrow\swarrow & & & \\
 & \searrow & & \swarrow\searrow & & \swarrow\searrow & & \swarrow & \\
a_{21} & & a_{22} & & a_{23} & & a_{21} & & a_{22} \\
 & & & \searrow\swarrow & & \searrow\swarrow & & & \\
 & \swarrow & & \swarrow\searrow & & \swarrow\searrow & & \searrow & \\
a_{31} & & a_{32} & & a_{33} & & a_{31} & & a_{32}
\end{array}\right)$$

12.2 Andere Ergebnisse

Sei $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$ eine $n \times n$ Matrix. Wir schreiben $det(A)$ für den Wert der Determinantenfunktion, angewandt auf die Matrix A. Traditionellerweise wird manchmal die Matrix zwischen Absolutbetragzeichen

[1] In dieser Formel wird über die Elemente a_{nj} der letzten Zeile summiert. Es ist auch möglich, jede andere Zeile zu nehmen. Sei i eine Zahl zwischen 1 und n. Dann gilt auch $det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} det(A_{ij})$. Der Beweis ist praktisch identisch mit dem für die n-te Zeile.

gesetzt, um diese Determinante anzudeuten. D.h. $det(A) \equiv \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$.

Satz 12.1. *Seien A, B zwei $n \times n$ Matrizen über einem Körper F. Dann ist*

$$det(A \cdot B) = det(A) \cdot det(B).$$

Beweis. Es gibt zwei Fälle: (i) $A \cdot B$ ist singulär, und (ii) $A \cdot B$ ist regulär. Sei $f : V \to V$ die lineare Abbildung, mit zugehöriger Matrix A, und sei $g : V \to V$ die entsprechende lineare Abbildung für die Matrix B. Dann ist $A \cdot B$ die zu $f \cdot g$ gehörige Matrix bzgl. der vorgegebenen Basis für V. Es gilt $det(A \cdot B) = 0 \Leftrightarrow A \cdot B$ ist singulär $\Leftrightarrow dim(f(g(V)) < n \Leftrightarrow dim(f(V)) < n$ oder $dim(g(V)) < n \Leftrightarrow det(A) = 0$ oder $det(B) = 0$. D.h. im Falle (i) ist $det(A \cdot B) = 0$, und wir haben $det(A) = 0$ oder $det(B) = 0$, so daß $det(A \cdot B) = det(A) \cdot det(B)$.

Fall (ii) ist auch relativ einfach. Es gilt $A \cdot B$ ist regulär $\Leftrightarrow dim(f(g(V)) = n \Leftrightarrow dim(f(V)) = n$ und $dim(g(V)) = n \Leftrightarrow A, B \in GL(n; F) \Leftrightarrow \exists S_1, \ldots, S_p, S_{p+1}, \ldots, S_q$, elementare Matrizen, mit $A = S_1 \cdots S_p$ und $B = S_{p+1} \cdots S_q$. D.h. $A \cdot B = S_1 \cdots S_q$, ein Produkt von elementaren Matrizen. Um den Beweis zu Ende zu führen, brauchen wir offensichtlich nur das folgende Lemma zu beweisen. □

Lemma. *Seien S, D zwei $n \times n$ Matrizen, wobei S eine elementare Matrix ist. Dann ist $det(D \cdot S) = det(D) \cdot det(S)$.*

Beweis. Es gibt drei Fälle.

1. $S = S_{ij}$ vertauscht die Spalten i und j von D. Dann ist (nach Konsequenz 3) $det(D \cdot S) = -det(D)$. Aber (wieder Konsequenz 3) auch $det(S_{ij}) = -1$.

2. $S = S_i(a)$, nämlich die i-te Spalte von D wird mit dem Faktor a multipliziert. Dann gilt (Konsequenz 2) $det(D \cdot S) = a \cdot det(D)$. Aber auch $det(S_i(a)) = a$.

3. $S = S_{ij}(c)$. Dann gilt (Konsequenz 5) $det(D \cdot S) = det(D)$. Aber (wieder Konsequenz 5) es gilt auch $det(S_{ij}(c)) = 1$. (Die elementare Matrix S_{ij} entsteht aus der Identitätsmatrix, durch eine entsprechende Spaltenoperation.)

□

Korollar 12.1.1. *Sei $A \in GL(n; F)$. Dann ist $det(A^{-1}) = (det(A))^{-1}$.*

Daraus folgt:

Korollar 12.1.2. *Sei A eine $n \times n$ Matrix, und sei A' ähnlich wie A. (D.h. $\exists C \in GL(n; F)$ mit $A' = C \cdot A \cdot C^{-1}$.) Dann ist $det(A) = det(A')$.*

Definition 49. $SL(n, F) = \{A \in M(n \times n; F) : det(A) = 1\}$ heißt die **spezielle lineare Gruppe** der Ordnung n.

Korollar 12.1.3. *$SL(n; F)$ ist eine Gruppe.*

Wir haben die Determinantenfunktion mittels Spaltenoperationen definiert. Es wäre auch möglich, die entsprechenden Zeilenoperationen zu nehmen, um die Determinantenfunktion zu definieren. Wir würden trotzdem genau dieselbe Funktion $M(n \times n; F) \to F$ erhalten.

Satz 12.2. *Sei $A \in M(n \times n; F)$ transformiert in eine andere Matrix $A \to A'$ durch elementare Spaltenoperationen: $A \to A' = A \cdot S_1 \cdots S_p$. Andererseits können wir auch A transformieren durch die entsprechenden Zeilenoperationen: $A \to A^* = S_p \cdots S_1 \cdot A$. Dann gilt $det(A') = det(A^*)$. Insbesondere sei $det_S : M(n \times n; F) \to F$ die Determinantenfunktion, definiert nach Spaltenoperationen, und $det_Z : M(n \times n; F) \to F$ die Determinantenfunktion, definiert nach Zeilenoperationen. Dann gilt $det_S = det_Z$.*

Beweis. Dies folgt, da $det(S \cdot D) = det(S) \cdot det(D) = det(D) \cdot det(S) = det(D \cdot S)$, für Matrizen S und D. □

Satz 12.3. *Sei $A = (a_{ij})$ eine $n \times n$ Matrix. Dann gilt*

$$det(A) = \sum_{\sigma \in S_n} sign(\sigma) a_{\sigma(1)\,1} \cdot a_{\sigma(2)\,2} \cdots a_{\sigma(n)\,n}.$$

Beweis. Es genügt, zu zeigen, daß die so definierte Funktion unsere drei Eigenschaften erfüllt.

1. $det(I_n) = \sum_{\sigma \in S_n} sign(\sigma) \prod_{i=1}^n \delta_{\sigma(i)\,i} = \sum_{\sigma = id} sign(\sigma) \prod_{i=1}^n 1 = 1$. Die zweite Gleichung folgt, da $\delta_{i\,\sigma(i)} = 0$ für irgendein i, falls $\sigma \neq id$ (die Identitätsabbildung).

2. Sei $a \in F$ und $A' = \{\gamma_1, \ldots, \gamma_{i-1}, a \cdot \gamma_i, \gamma_{i+1}, \ldots, \gamma_n\}$, wobei $A = \{\gamma_1, \ldots, \gamma_n\}$ die Spalten von A sind. Dann ist

$$\begin{aligned} det(A') &= \sum_{\sigma \in S_n} sign(\sigma) a'_{\sigma(1)\,1} \cdots a'_{\sigma(n)\,n} \\ &= \sum_{\sigma \in S_n} sign(\sigma) a_{\sigma(1)\,1} \cdots a_{\sigma(i-1)\,i-1} a \cdot a_{\sigma(i)\,i} a_{\sigma(i+1)\,i+1} \cdots a_{\sigma(n)\,n} \\ &= a \cdot \sum_{\sigma \in S_n} sign(\sigma) a_{\sigma(1)\,1} a_{\sigma(2)\,2} \cdots a_{\sigma(n)\,n} \\ &= a \cdot det(A). \end{aligned}$$

3. Sei $A^* = \{\gamma_1, \ldots, \gamma_{i-1}, \gamma_i + \gamma_j, \gamma_{i+1}, \ldots, \gamma_n\}$. Dann ist

$$\begin{aligned} det(A^*) &= \sum_{\sigma \in S_n} sign(\sigma) a^*_{\sigma(1)\,1} \cdots a^*_{\sigma(n)\,n} \\ &= \sum_{\sigma \in S_n} sign(\sigma) a_{\sigma(1)\,1} \cdots a_{\sigma(i-1)\,i-1} \cdot \\ &\qquad \cdot (a_{\sigma(i)\,i} + a_{\sigma(j)\,j}) a_{\sigma(i+1)\,i+1} \cdots a_{\sigma(n)\,n} \\ &= \sum_{\sigma \in S_n} sign(\sigma) a_{\sigma(1)\,1} \cdots a_{\sigma(i-1)\,i-1} a_{\sigma(i)\,i} a_{\sigma(i+1)\,i+1} \cdots a_{\sigma(n)\,n} \\ &\quad + \sum_{\sigma \in S_n} sign(\sigma) a_{\sigma(1)\,1} \cdots a_{\sigma(i-1)\,i-1} a_{\sigma(j)\,j} a_{\sigma(i+1)\,i+1} \cdots a_{\sigma(n)\,n} \\ &= \sum_{\sigma \in S_n} sign(\sigma) a_{\sigma(1)\,1} \cdots a_{\sigma(i-1)\,i-1} a_{\sigma(i)\,i} a_{\sigma(i+1)\,i+1} \cdots a_{\sigma(n)\,n} \\ &\quad + 0 \\ &= det(A). \end{aligned}$$

Die vorletzte Gleichung folgt, da die Terme $a_{\sigma(j)\,j}$ immer zweimal vorkommen, und es wird über alle Permutationen summiert. Aber die zwei Stellen, wo die Terme vorkommen, unterscheiden sich durch eine Transposition. Daher ist das Vorzeichen einmal plus und einmal minus (siehe Satz 11.3), und sie summieren sich auf Null.

□

Es ist üblich, diese Formel etwas anders zu schreiben.

Lemma. *Sei $\sigma \in S_n$ und sei $P_1 = \{(\sigma(i), i) : 1 \leq i \leq n\}$ die Menge der Paare, die in der Formel von Satz 12.3 vorkommen. Sei $P_2 = \{(i, \sigma^{-1}(i)) : 1 \leq i \leq n\}$. Dann ist $P_1 = P_2$.*

Beweis. Sei $i \in \{1, \ldots, n\}$. Dann ist $(\sigma(i), i)$ ein typisches Element von P_1. Ist $(\sigma(i), i) \in P_2$? D.h. $\exists j \in \{1, \ldots, n\}$ mit $i = \sigma^{-1}(j)$ und $\sigma(i) = j$? Ja, da $j = \sigma(i) \Rightarrow \sigma^{-1}(j) = \sigma^{-1}(\sigma(i)) = i \Rightarrow P_1 \subset P_2$. Aber ein ähnliches Argument zeigt, daß $P_2 \subset P_1$. Daher $P_1 = P_2$. □

Korollar 12.3.1. *(Darstellung der Determinantenfunktion nach Leibniz) Es gilt*

$$det(A) = \sum_{\sigma \in S_n} sign(\sigma) a_{1\,\sigma(1)} a_{2\,\sigma(2)} \ldots a_{n\,\sigma(n)}.$$

Beweis. Nach Satz 12.3 ist

$$\begin{aligned} det(A) &= \sum_{\sigma \in S_n} sign(\sigma) a_{\sigma(1)\,1} \cdot a_{\sigma(2)\,2} \cdots a_{\sigma(n)\,n} \\ &= \sum_{\sigma^{-1} \in S_n} sign(\sigma^{-1}) a_{1\,\sigma^{-1}(1)} \cdot a_{2\,\sigma^{-1}(2)} \cdots a_{n\,\sigma^{-1}(n)} \\ &= \sum_{\sigma \in S_n} sign(\sigma) a_{1\,\sigma(1)} a_{2\,\sigma(2)} \ldots a_{n\,\sigma(n)}. \end{aligned}$$

Die vorletzte Gleichung gilt wegen des Lemmas, und auch wegen $sign(\sigma^{-1}) = sign(\sigma)$. (Dies ist eine triviale Folgerung des Korollars zu Satz 12.1.) Die letzte Gleichung gilt, da die Abbildung $f : S_n \to S_n$, gegeben durch $f(\sigma) = \sigma^{-1}$, $\forall \sigma \in S_n$ eine Bijektion ist. □

Nun, wenn es um die 'praktische' Mathematik geht, dann ist Leibniz' Formel genauso absurd wie unsere rekursive Formel (Satz 11.1). Man sagt, daß ein Rechenverfahren (die von n Eingabeparametern bestimmt wird, für verschiedene $n \in \mathbf{N}$) in 'Polynomzeit' läuft, falls ein Polynom $P(n)$ existiert mit der Eigenschaft, daß die Rechnung für den Fall $n \in \mathbf{N}$ mit Sicherheit beendet ist nach höchstens $P(n)$ Schritten. Unser Computerprogramm läuft tatsächlich in Polynomzeit. Ein 'schlechtes' Rechenverfahren (typischerweise verursacht durch eine Rekursion) braucht normalerweise 'Exponentialzeit'. Aber Leibniz' Verfahren braucht mehr als $n!$ Rechenschritte, um $det(A)$ für eine $n \times n$ Matrix auszurechnen. Dies ist noch viel schlechter als Exponentialzeit, da $\lim_{n\to\infty} e^{kn}/n! = 0$, für alle möglichen Zahlen $k \geq 0$. Trotzdem ist die Formel für theoretische Zwecke oft sehr nützlich. Z.B.

Satz 12.4. *Sei $A = (a_{ij})$ eine $n \times n$ Matrix, und sei $A^t = (a^t_{ij})$ die zu A transponierte Matrix. (D.h. $a^t_{ij} = a_{ji}$ für alle i und j.) Dann ist $det(A) = det(A^t)$.*

Beweis. Wir haben gerade gesehen, daß

$$\begin{aligned}
det(A) &= \sum_{\sigma\in S_n} sign(\sigma) a_{\sigma(1)\,1} \cdot a_{\sigma(2)\,2} \cdots a_{\sigma(n)\,n} \\
&= \sum_{\sigma\in S_n} sign(\sigma) a_{1\,\sigma(1)} a_{2\,\sigma(2)} \ldots a_{n\,\sigma(n)} \\
&= \sum_{\sigma\in S_n} sign(\sigma) a^t_{\sigma(1)\,1} \cdot a^t_{\sigma(2)\,2} \cdots a^t_{\sigma(n)\,n} \\
&= det(A^t).
\end{aligned}$$

□

12.3 Zwei weitere spezielle Fälle

Auch in höheren Dimensionen gibt es zwei Fälle, wo die Berechnung der Determinantenfunktion relativ einfach ist.

- **Obere Dreiecksform:** Angenommen,

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ 0 & a_{22} & \ldots & a_{2n} \\ 0 & 0 & \ddots & \vdots \\ 0 & \ldots & 0 & a_{nn} \end{pmatrix}.$$

  Dan ist $det(A) = a_{11} \cdot a_{22} \cdots a_{nn}$. Der Beweis folgt durch Induktion über n. Der Fall $n = 1$ ist trivial. Sei daher $n > 1$. Es gilt

$$det(A) = \sum_{j=1}^{n} (-1)^{n-j} a_{nj} det(A_{nj}),$$

  aber in diesem Fall ist $a_{nj} = 0$ für $j \neq n$. D.h.

$$det(A) = (-1)^{n-n} a_{nn} det(A_{nn}) = (+1) \cdot a_{nn} \cdot (a_{11} \cdots a_{n-1\,n-1}).$$

  Die letzte Gleichung ist die induktive Hypothese für den Fall $n - 1$.

- **Kästchenform:** Angenommen, die Matrix A sei in Kästchenform (siehe Kapitel 10). D.h.

$$A = \begin{pmatrix} A_1 & 0 & \ldots & & 0 \\ 0 & A_2 & 0 & & \\ \vdots & 0 & \ddots & 0 & \vdots \\ & & 0 & A_{k-1} & 0 \\ 0 & & \ldots & 0 & A_k \end{pmatrix},$$

  wobei $A_i = \begin{pmatrix} a_{r_i\,r_i} & \cdots & a_{r_i\,r_{i+1}-1} \\ \vdots & & \vdots \\ a_{r_{i+1}-1\,r_i} & \cdots & a_{r_{i+1}-1\,r_{i+1}-1} \end{pmatrix}$, $i = 1, \ldots, k$. Dann ist $det(A) = \prod_{i=1}^{k} det(A_i)$. Hierbei ist der Fall $k = 1$ genauso trivial wie vorher. Sei daher $k > 1$ und für jedes $1 \leq i \leq k$ sei die $n \times n$ Matrix A^*_i

gegeben als:

$$A_i^* = \begin{pmatrix} 1 & 0 & & \dots & & & 0 \\ 0 & \ddots & & & & & \\ & & 1 & & & & \\ \vdots & & & A_i & & & \vdots \\ & & & & 1 & & \\ & & & & & \ddots & 0 \\ 0 & & & \dots & & 0 & 1 \end{pmatrix}.$$

D.h. der Block A_i bleibt unverändert, aber um die ursprüngliche Matrix A in A_i^* zu verwandeln, werden alle anderen Blöcke in die entsprechenden Einheitsmatrixblöcke umgewandelt. Es gilt $det(A_i^*) = det(A_i)$, für jedes i. (Eine leichte Übung.) Aber es gilt auch $A = \prod_{i=1}^k A_i$. Folglich (Satz 12.1) $det(A) = \prod_{i=1}^k det(A_i)$.

12.4 Noch eine Formel

Es gibt unzählige, mehr oder weniger elegante Formeln, die man beweisen kann über die Algebra der Determinantenfunktion und die Darstellung einer inversen Matrix. Aber irgendwann ist unsere Geduld zu Ende, und daher möchte ich nur noch die Cramersche Regel beschreiben.

Sei $A \in M(n \times n; F)$ eine Matrix. Unsere ursprüngliche Regel war

$$det(A) = \sum_{j=1}^{n} (-1)^{n-j} a_{nj} det(A_{nj}).$$

Was passiert, wenn die Zahlen a_{nj} in dieser Formel ersetzt werden durch a_{kj}, für irgendeine andere Zahl $k < n$? Dann haben wir $\sum_{j=1}^{n} (-1)^{n-j} a_{kj} det(A_{nj}) = 0$. Warum? Es ist am einfachsten, dies einzusehen, wenn man die Matrix A' betrachtet, die (fast) identisch ist mit A. Um von A nach A' zu gelangen, braucht man nur die letzte Zeile aus A zu entfernen und stattdessen wird eine Kopie der k-ten Zeile anstelle der n-ten Zeile eingefügt. Da A' zwei identische Zeilen besitzt, gilt offensichtlich $det(A') = 0$. Aber es gilt auch, daß $det(A_{nj}) = det(A'_{nj})$, für alle j. (Die letzte Zeile fehlt ja in der Untermatrix A_{nj}.) Folglich ist

$$\sum_{j=1}^{n} (-1)^{n-j} a_{kj} det(A_{nj}) = \sum_{j=1}^{n} (-1)^{n-j} a'_{nj} det(A'_{nj}) = det(A') = 0.$$

Es ist auch möglich, die Determinantenfunktion als Summe entlang jeder anderen Zeile auszudrücken. Sei etwa $i \le n$. Dann gilt[2]

$$det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} det(A_{ij}).$$

Dies folgt, wenn man sich überlegt, daß es möglich ist, zunächst die Zeilen i und n zu vertauschen, dann die Determinante auszurechnen mit unserer bekannten Formel, und schließlich zu überlegen, daß der Wert der Determinantenfunktion für die vertauschte Matrix jeweils das Negative des Werts für die ursprüngliche Matrix ist.

Wir können daher auch sagen, daß

$$\sum_{j=1}^{n} (-1)^{k+j} a_{ij} det(A_{kj}) = 0, \quad \forall i \neq k.$$

Satz 12.5. *Sei $A \in GL(n; F)$ und sei $A' = (a'_{ij}) = A^{-1}$. Dann ist*

$$a'_{ij} = \frac{1}{det(A)} (-1)^{i+j} det(A_{ji}).$$

Beweis. Wir müssen zeigen, daß $A \cdot A' = I_n$. D.h.

$$\sum_{l=1}^{n} a_{il} a'_{lj} = \delta_{ij}, \quad \forall i, j, \quad \text{wobei,} \quad \delta_{ij} = \begin{cases} 1, & \text{falls } i = j, \\ 0, & \text{sonst.} \end{cases}$$

[2] Es gilt $(-1)^{i+j} = (-1)^{i-j}$. Dies folgt, da $2j = 0 \bmod 2 \Rightarrow j = -j \bmod 2 \Rightarrow i + j = i - j \pmod 2$.

Aber für $i \neq j$ gilt

$$\begin{aligned}
\sum_{l=1}^{n} a_{il} a'_{lj} &= \sum_{l=1}^{n} a_{il} \frac{1}{det(A)} (-1)^{l+j} det(A_{jl}) \\
&= \frac{1}{det(A)} \left\{ \sum_{l=1}^{n} a_{il} (-1)^{l+j} det(A_{jl}) \right\} \\
&= 0
\end{aligned}$$

(Die letzte Gleichung ist doch nur die Formel, die wir gerade vorher bewiesen haben.)

Falls $i = j$, dann ist die Formel

$$\begin{aligned}
\sum_{l=1}^{n} a_{il} a'_{li} &= \sum_{l=1}^{n} a_{il} \frac{1}{det(A)} (-1)^{l+i} det(A_{il}) \\
&= \frac{1}{det(A)} \left\{ \sum_{l=1}^{n} a_{il} (-1)^{l+i} det(A_{il}) \right\} \\
&= \frac{1}{det(A)} \{det(A)\} \\
&= 1.
\end{aligned}$$

□

Bemerkung. Die Matrix $\tilde{A} = (\tilde{a}_{ij})$ mit $\tilde{a}_{ij} = (-1)^{i+j} det(A_{ji})$ (so daß $A \cdot \tilde{A} = det(A) \cdot I_n$) heißt die zu A komplementäre Matrix.

Kapitel 13

Komplexe Zahlen

Sei $\mathbf{C} = \{a + bi : a, b \in \mathbf{R}\}$ und $i = \sqrt{-1}$. Die Addition wird durch

$$(a + bi) + (c + di) = (a + c) + (b + d)i$$

und die Multiplikation durch

$$(a + bi) \cdot (c + di) = (ac - bd) + (ad + bc)i$$

festgelegt. Bekanntlich ist $\mathbf{C}$ damit ein Körper. Insbesondere ist $0 + 0i$ das Nullelement, und $1 + 0i$ ist das Einselement.

Die komplexen Zahlen werden oft als Vektoren im 2-dimensionalen reellen Raum dargestellt. Die zwei kanonischen Basisvektoren $\binom{1}{0}$ und $\binom{0}{1}$ werden mit den Zahlen $1 + 0i$ und $0 + 1i$ identifiziert. Seien $u = a + bi$ und $v = c + di$ zwei beliebige komplexe Zahlen. Dann ist die Vektorsumme $u + v$ (in dem *reellen* Vektorraum $\mathbf{R}^2$) gleich die Summe der komplexen Zahlen $u + v \in \mathbf{C}$.

Es ist manchmal nützlich, komplexe Zahlen durch Polarkoordinaten darzustellen. Sei $\alpha = (x, y) \in \mathbf{R}^2$ beliebig. Dann gibt es zwei reelle Zahlen $r \geq 0$ und $\theta \in [0, 2\pi)$ mit

$$x = r \cdot \cos\theta, \quad \text{und} \quad y = r \cdot \sin\theta,$$

und die entsprechende komplexe Zahl $z \in \mathbf{C}$ ist

$$z = x + yi = r\cos\theta + i \cdot r\sin\theta = r \cdot e^{i\theta}.$$

Seien jetzt $z_1 = r_1\cos\theta_1 + i \cdot r_1 \sin\theta_1$ und $z_2 = r_2\cos\theta_2 + i \cdot r_2\sin\theta_2$. Was ist der Vektor $z_1 \cdot z_2 \in \mathbf{C}$? Es gilt

$$\begin{aligned} z_1 \cdot z_2 &= (r_1\cos\theta_1 + i \cdot r_1\sin\theta_1) \cdot (r_2\cos\theta_2 + i \cdot r_2\sin\theta_2) \\ &= r_1 r_2\{(\cos\theta_1 \cdot \cos\theta_2 - \sin\theta_1 \cdot \sin\theta_2) + \\ &\qquad (\cos\theta_1 \cdot \sin\theta_2 + \sin\theta_1 \cdot \cos\theta_2) \cdot i\} \\ &= r_1 r_2(\cos(\theta_1 + \theta_2) + i \cdot \sin(\theta_1 + \theta_2)) \end{aligned}$$

(Die letzte Gleichung ist eine Standardformel der Trigonometrie.)

Es gibt eine besondere Vorlesung über die 'Funktionentheorie', wo die Analysis der komplexen Zahlen behandelt wird. Die klassischen Ergebnisse sind im 19.ten Jahrhundert entstanden, und dieses Gebiet ist ein zentraler Teil der heutigen Mathematik. Auch in der Linearen Algebra spielt das Rechnen mit komplexen Zahlen eine große Rolle. Insbesondere ist die Tatsache für uns interessant, daß $\mathbf{C}$ ein 'algebraisch abgeschlossener Körper' ist. D.h. insbesondere, daß jede Polynomgleichung in $\mathbf{C}$ eine Nullstelle besitzt.

Definition 50. Sei $P \in K[x]$ ein Polynom über einer Unbestimmten x mit Koeffizienten in einem Körper F. Sei $k \in F$. Falls $P(k) = 0$, dann heißt k eine *Nullstelle* (oder Wurzel) des Polynoms P.

Z.B. sei $P \in \mathbf{C}[x]$, wobei $\mathbf{C} \equiv$ die komplexen Zahlen und

$$P(x) \equiv x^3 - 1.$$

Dann gibt es drei Wurzeln: nämlich

$$\begin{aligned} x &= 1, \\ x &= -\frac{1}{2} + \frac{\sqrt{3}}{2}i, \quad \text{und} \\ x &= -\frac{1}{2} - \frac{\sqrt{3}}{2}i. \end{aligned}$$

Der **Fundamentalsatz der Algebra** lautet:

Satz 13.1. *Sei $P \in \mathbf{C}[x]$ ein Polynom mit grad$(P) > 0$. Dann existiert eine Wurzel $z \in \mathbf{C}$ für P.*

Dieser Satz ist von entscheidender Bedeutung in der linearen Algebra. Leider sind aber die Beweise doch etwas kompliziert. Es ist daher üblich, in den Vorlesungen für Anfänger viel Wert auf diesen Satz zu legen, aber keinen Beweis zu geben; Studenten sollten einfach 'glauben', daß alles so stimmt. Diese Vorgehensweise scheint mir aber doch sehr unbefriedigend. Daher werde ich hier einen 'elementaren' Beweis des Satzes beschreiben.

Nun, solche Mathematiker wie Cauchy und Lagrange im 18.ten Jahrhundert haben den Fundamentalsatz — eigentlich nur als 'Vermutung' — gekannt und ihn geglaubt, aber ihre Ideen über die komplexen Zahlen (und überhaupt über die strenge Führung eines mathematischen Beweises) ließen aus moderner Sicht viel zu wünschen übrig. In den meisten Büchern steht, daß Gauß der erste strenge Beweis des Satzes im Jahre 1799 gelang. Dies stimmt aber nicht. In Wahrheit enthielt sein veröffentlichter Beweis eine bestimmte Lücke. Wahrscheinlich aus diesem Grunde versuchte er es nochmal im Jahre 1814, aber sein zweiter Beweis war auch lückenhaft. Sein dritter Versuch im Jahre 1816 war doch glaubwürdiger, und er benutzte im wesentlichen den Cauchy Integralsatz.[1] (Es muß aber gesagt werden, daß Gauß für diesen Beweis die vereinfachende Annahme gemacht hat, daß nur $P \in \mathbf{R}[x]$.) In seinem vierten Beweis am Ende seines Lebens (1848) versuchte er diesen Mangel aufzuheben.

Sie sehen daher, daß auch ein Mann von der Größe eines Gauß sehr viel Mühe mit diesem wirklich fundamentalen Satz gehabt hat. Aber alles schreitet voran, und wir in der modernen Welt können den Satz doch mit Leichtigkeit bewältigen. In dem Beweis, den ich jetzt beschreiben werde, werden einige Ideen aus den Analysis I und II Vorlesungen benutzt — und gerade diese Ideen fehlten den Mathematikern am Anfang des 19.ten Jahrhunderts. (Die Ergebnisse werden hier nur zittiert; die Beweise finden Sie in der 'anderen Hälfte' der Anfängervorlesungen.)

Definition 51. Sei $z \in \mathbf{C}$ und $n \in \mathbf{N}$. Eine n-te Wurzel von z ist eine Nullstelle des Polynoms $x^n - z$.

Bemerkung. Die Zahlen 1, $-\frac{1}{2} + \frac{\sqrt{3}}{2}i$, und $-\frac{1}{2} - \frac{\sqrt{3}}{2}i$ sind alle 3.te Wurzeln von 1.

Lemma (A). *Sei $z \neq 0$. Dann gibt es genau n verschiedene n-te Wurzeln von z in $\mathbf{C}$.*

Beweis. Sei $z = r_0 e^{i\Theta_0}$. Dann sind die n Zahlen

$$w_j = \sqrt[n]{r_0} \cdot e^{\frac{i(\Theta_0 + 2\pi j)}{n}}$$

alle verschieden, und auch sind sie alle Wurzeln von $x^n - z$, für $j = 1, \ldots, n$. Warum?

$$\begin{aligned}
(w_j)^n &= \left(\sqrt[n]{r_0} \cdot e^{\frac{i(\Theta_0+2\pi j)}{n}}\right)^n \\
&= (\sqrt[n]{r_0})^n \cdot \left(e^{\frac{i(\Theta_0+2\pi j)}{n}}\right)^n \\
&= r_0 \cdot e^{i(\Theta_0 + 2\pi j)} \\
&= r_0 \cdot e^{i\Theta_0} \cdot e^{i(2\pi j)} \\
&= r_0 \cdot e^{i\Theta_0} = z.
\end{aligned}$$

(Die vorletzte Gleichung folgt, da stets $e^{i2\pi j} = 1$.) Aber für $j \neq k$ (und beide $\leq n$) ist

$$\frac{2\pi j}{n} \neq \frac{2\pi k}{n} \pmod{2\pi}.$$

Folglich ist

$$\begin{aligned}
w_j - w_k &= \sqrt[n]{r_0}\left(e^{\frac{i(\Theta_0+2\pi j)}{n}} - e^{\frac{i(\Theta_0+2\pi k)}{n}}\right) \\
&= \sqrt[n]{r_0} \cdot e^{\frac{i\Theta_0}{n}} \underbrace{\left(e^{\frac{i(2\pi j)}{n}} - e^{\frac{i(2\pi k)}{n}}\right)}_{\neq 0} \\
&\neq 0.
\end{aligned}$$

Kann es sein, daß das Polynom $x^n - z$ mehr als n verschiedene Nullstellen hat? Nein! Es gilt nämlich

$$x^n - z = (x - w_1) \cdots (x - w_n) = 0.$$

Falls etwa w_0 noch eine andere Wurzel wäre, dann wäre $w_0 - w_j \neq 0$ für alle $j = 1, \ldots, n$ und

$$\underbrace{(w_0 - w_1)}_{\neq 0} \cdots \underbrace{(w_0 - w_n)}_{\neq 0} = 0,$$

was unmöglich ist, da $\mathbf{C}$ ein Körper ist. □

[1] Diejenigen, die später die Vorlesung über 'Funktionentheorie' höhren, werden diesen 'Standardbeweis' kennenlernen.

Definition 52. Die Binomialkoeffizienten sind

$$\binom{n}{m} \equiv \frac{n}{m!(n-m)!}$$

für $n \geq m$ in $\mathbf{N}$. (Man schreibt auch $\binom{n}{0} \equiv 1$, wobei nach der üblichen Vereinbarung $0! \equiv 1$ gilt.)

Satz 13.2 (Binomialsatz). *Es gilt* $(u+v)^n = \sum_{k=0}^{n} \binom{n}{k} u^{n-k} v^k$.

Beweis. Induktion über n. Für $n = 1$ ist die Sache klar. Sei daher $n > 1$, und sei

$$(u+v)^{n-1} = \sum_{k=0}^{n-1} \binom{n-1}{k} u^{n-k-1} v^k.$$

$$\begin{aligned}
\Rightarrow \quad (u+v)^n &= (u+v) \cdot \left(\sum_{k=0}^{n-1} \binom{n-1}{k} u^{n-k-1} v^k \right) \\
&= \sum_{k=0}^{n-1} \binom{n-1}{k} u^{n-k} v^k + \sum_{k=0}^{n-1} \binom{n-1}{k} u^{n-(k+1)} v^{k+1} \\
&= \sum_{k=0}^{n-1} \binom{n-1}{k} u^{n-k} v^k + \sum_{k=1}^{n} \binom{n-1}{k-1} u^{n-k} v^k \\
&= u^n + \sum_{k=1}^{n-1} \binom{n}{k} u^{n-k} v^k + v^n.
\end{aligned}$$

Wir haben hier die Tatsache benutzt, daß

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}.$$

Dies ist offensichtlich wahr nach der Methode des Pascalschen Dreiecks für die Berechnung der Binomialkoeffizienten.

$$\begin{array}{ccccccccccc}
 & & & & & 1 & & & & & \\
 & & & & 1 & & 1 & & & & \\
 & & & 1 & & 2 & & 1 & & & \\
 & & 1 & & 3 & & 3 & & 1 & & \\
 & 1 & & 4 & & 6 & & 4 & & 1 & \\
1 & & 5 & & 10 & & 10 & & 5 & & 1
\end{array}$$

u.s.w. ∎

(Im folgenden werden wir die übliche Absolutbetragskonvention im Rahmen der komplexen Zahlen benutzen: falls etwa $z = a + bi$, dann ist $\|z\| \equiv \sqrt{a^2 + b^2}$.)

Lemma (B). *Seien* $u, v \in \mathbf{C}$. *Dann gilt* $\|u\| - \|v\| \leq \|u + v\|$.

Beweis. Wir haben die Dreiecksungleichung: $\|a + b\| \leq \|a\| + \|b\|$ für alle $a, b \in \mathbf{C}$. Setze jetzt nur $a = u + v$ und $b = -v$ so daß $\|b\| = \|v\|$. ∎

Korollar 13.2.1. *Seien* $w_0, \ldots, w_p \in \mathbf{C}$. *Dann gilt*

$$\|w_0\| - \|w_1\| - \cdots - \|w_p\| \leq \left\| \sum_{j=0}^{p} w_j \right\|.$$

Beweis. Induktion über p. Der Fall $p = 1$ ist Lemma B. Sei daher etwa $u = w_0 + \cdots + w_{p-1}$. Es gilt dann $\|u\| - \|w_p\| \leq \|u + w_p\| \equiv \left\| \sum_{j=0}^{p} w_j \right\|$. Aber nach der induktiven Hypothese ist doch $\|w_0\| - \|w_1\| - \cdots - \|w_{p-1}\| \leq \left\| \sum_{j=0}^{p-1} w_j \right\| \equiv \|u\|$. ∎

Definition 53. Sei $f : \mathbf{C} \to \mathbf{R}$ eine Funktion. Wir schreiben

$$\lim_{\|z\| \to \infty} f(z) = \infty,$$

falls Folgendes gilt. Für alle natürliche Zahlen $n \in \mathbf{N}$ gibt es ein $r_0 > 0$, $r_0 \in \mathbf{R}$, so daß $f(z) > n$, falls $\|z\| > r_0$.

Lemma (C). . *Sei $P \in \mathbf{C}[z]$ von der Gestalt*

$$P(z) = z^m + a_{m-1}z^{m-1} + \cdots + a_1 z + a_0, \quad m \geq 1.$$

Dann ist $\lim_{\|z\| \to \infty} \|P(z)\| = \infty$.

Beweis. Sei $n \in \mathbf{N}$ gegeben. Wir müssen ein $r_0 > 0$ finden, so daß $\|P(z)\| > n$, für alle $\|z\| > r_0$. Wir definieren daher m Zahlen s_j, $j = 0, \ldots, m-1$ wie folgt.

$$s_j = \begin{cases} 2\|a_j\|, & \text{falls } \|a_j\| \neq 0, \\ 0, & \text{sonst.} \end{cases}$$

Sei nun $r_0 = \max\{s_0, \ldots, s_{m-1}, 2n, 1\}$, und wähle ein beliebiges $z_0 \in \mathbf{C}$ mit $\|z_0\| > r_0$. Es gilt

$$\begin{aligned} \|P(z_0)\| &= \|z_0^m\| \cdot \left\| 1 + \frac{a_{m-1}}{z_0} + \cdots + \frac{a_0}{z_0^m} \right\| \\ &\geq \|z_0^m\| \cdot \left(1 - \left\| \frac{a_{m-1}}{z_0} \right\| - \cdots - \left\| \frac{a_0}{z_0^m} \right\| \right) \qquad \text{(Lemma B)} \\ &\geq \|z_0^m\| \cdot \frac{1}{2} \\ &= \frac{\|z_0\|^m}{2} \quad \geq \quad \frac{\|z_0\|}{2} \qquad (\text{da} \quad \|z_0\| > r_0 \geq 1) \\ &\geq \frac{2n}{2} \quad = \quad n. \end{aligned}$$

□

Korollar 13.2.2. *Lemma C gilt auch für beliebige $P \in \mathbf{C}[z]$ mit*

$$grad(P) \geq 1.$$

Beweis. Sei $P(z) = \sum_{j=0}^{m} a_j z^j$ mit $a_m \neq 0$. Dann ist $a_m^{-1} \cdot P(z)$ wie im obigen Beweis. Wähle daher r_0 wie oben, jedoch multipliziert mit dem Faktor $\|a_m^{-1}\|$. □

Nun, alles was wir bisher gemacht haben, war Gauß sicherlich wohl bekannt. Zu seiner Zeit fehlte jedoch unsere moderne Auffassung von der Analysis. Wir brauchen nur einige ganz elementare Ergebnisse aus den Anfängervorlesungen.

Definition 54. Sei M eine Menge und $F : M \to \mathbf{R}$ eine reellwertige Funktion. Eine Zahl $k \in \mathbf{R}$ heißt eine *obere Schranke* für f, falls $f(m) \leq k$, $\forall m \in M$. Ebenso heißt $g \in \mathbf{R}$ *untere Schranke* für f, falls $f(m) \geq g$, $\forall m \in M$. Die Zahl k_0 ist die *kleinste obere Schranke*, falls k_0 eine obere Schranke ist und $k_0 \leq k$ für alle anderen oberen Schranken k. Ähnlich ist die *größte untere Schranke* definiert. Man schreibt

$$\begin{array}{ll} lub(f), & -\text{‘least upper bound’} \\ glb(f), & -\text{‘greatest lower bound’} \end{array}$$

Satz 13.3 (über kompakte Mengen). *Sei M eine* kompakte *Menge und sei $f : M \to \mathbf{R}$ stetig. Dann gibt es $s, t \in M$ mit $f(s) = glb(f)$ und $f(t) = lub(f)$.*

Die folgenden Beispiele sind für uns wichtig.

1. Sei $P \in \mathbf{C}[z]$ ein Polynom. Dann ist $P : \mathbf{C} \to \mathbf{C}$ stetig. Auch die Funktion $\|P\| : \mathbf{C} \to \mathbf{R}$, gegeben durch $z \to \|P(z)\|$, ist stetig.
2. Die folgenden zwei Mengen sind kompakt.
   (a) $D(r, 0) = \{z \in \mathbf{C} : \|z\| \leq r\}$; die Scheibe mit Radius r um den Nullpunkt in $\mathbf{C}$ ist kompakt.
   (b) $[0, 1] \subset \mathbf{R}$, (das Einheitsintervall in $\mathbf{R}$) ist auch kompakt.

Lemma (D). *Sei $f : M \to \mathbf{R}$ eine Abbildung, die eine untere Schranke, etwa $u \in \mathbf{R}$ besitzt, wobei $M \neq \emptyset$. Dann $\exists$ eine glb für f.*

Beweis. Sei $m \in M$. Dann ist $f(m) \in \mathbf{R}$, und alle unteren Schranken sind $\leq f(m)$. Sei $u_0 = u$, $v_0 = f(m)$. Wir definieren zwei Sequenzen (u_k), (v_k), $k = 0, 1, \ldots$ wobei $u_k \leq v_k$, $\forall k$. Dabei ist (u_k) ist eine steigende Sequenz, (v_k) ist eine fallende Sequenz, $\lim_{k \to \infty} |v_k - u_k| = 0$, und zwar so, daß u_k stets eine untere Schranke für f ist, aber $v_k \geq$ alle möglichen unteren Schranken. Diese Aussagen treffen zu zumindestens für die Zahlen u_0 und v_0. Nach einem induktiven Verfahren sei angenommen, daß sie auch zutreffen für u_k und v_k. Wir definieren u_{k+1} und v_{k+1} dann wie folgt. Sei $x = \frac{v_k + u_k}{2}$. Ist x eine untere Schranke für f? Falls ja, dann sei $u_{k+1} = x$ und $v_{k+1} = v_k$. Falls nein, dann sei $u_{k+1} = u_k$ und $v_{k+1} = x$. Dann ist $glb(f) = \lim_{k \to \infty} u_k$. □

Jetzt haben wir genug Informationen, um den Fundamentalsatz der Algebra zu beweisen.

Beweis des Fundamentalsatzes: Sei $P \equiv \sum_{j=0}^{m} c_j z^j \in \mathbf{C}[z]$ unser vorgegebenes Polynom. Da natürlich $\|P(z)\| \geq 0$, $\forall z$, muß die Zahl 0 eine untere Schranke für $\|P\|$ sein. Nach Lemma D existiert dann eine *glb* für $\|P\|$. Sei daher

$$\alpha = glb\{\|P(z)\| : z \in \mathbf{C}\}.$$

Nach Lemma C muß ein $r_0 > 0$ existieren mit $\|P(z)\| \geq \alpha + 1$, für alle $z \in \mathbf{C}$ mit $\|z\| > r_0$. Dann ist sicherlich auch

$$\alpha = glb\{\|P(z)\| : z \in D(r_0, 0)\}.$$

Aber nach Beispiel 2(a) ist die Menge $D(r_0, 0)$ kompakt, und daher muß nach dem Satz über kompakte Mengen ein $z_0 \in D(r_0, 0)$ existieren mit

$$\|P(z_0)\| = \alpha.$$

Behauptung: Tatsächlich ist $\alpha = 0$, so daß $\|P(z_0)\| = 0$; d.h. $P(z_0) = 0$; d.h. z_0 ist die gesuchte Wurzel von P.

Um dies zu zeigen, sei $u \in \mathbf{C}$ irgendeine willkürlich gewählte komplexe Zahl. Dann ist

$$\begin{aligned}
P(z_0 + u) &= \sum_{j=0}^{m} c_j \cdot (z_0 + u)^j \\
&= \sum_{j=0}^{m} c_j \cdot \left(\sum_{k=0}^{j} \binom{j}{k} z_0^{j-k} u^k \right) \qquad \text{(Binomialsatz)} \\
&= \sum_{j=0}^{m} c_j \cdot \left(z_0^j + \sum_{k=1}^{j} \binom{j}{k} z_0^{j-k} u^k \right) \\
&= P(z_0) + \sum_{j=0}^{m} c_j \left(\sum_{k=1}^{j} \binom{j}{k} z_0^{j-k} u^k \right) \\
&= P(z_0) + \sum_{k=1}^{m} b_k u^k, \quad \text{etwa.}
\end{aligned}$$

Kann es sein, daß $b_1 = b_2 = \cdots = b_m = 0$? Falls ja, dann wäre $P(z_0 + u)$ eine konstante Funktion, d.h. $grad(P) < 1$, ein Widerspruch.

Sei daher $t \in \{1, \ldots, m\}$ die kleinste Zahl mit $b_t \neq 0$. Wir nennen diese Zahl einfach $b \equiv b_t$. Dann können wir schreiben

$$P(z_0 + u) = P(z_0) + bu^t + Q(u) \cdot u^{t+1},$$

wobei $Q \in \mathbf{C}[z]$ auch ein Polynom ist.

Unser Ziel ist zu zeigen, daß $P(z_0) = 0$. Um einen Widerspruch zu erzeugen, nehmen wir an, daß doch $P(z_0) \neq 0$. Dann wäre auch $-\frac{P(z_0)}{b} \neq 0$. Sei nun v eine t-te Wurzel von $-\frac{P(z_0)}{b}$. (siehe Lemma A) D.h.

$$v^t = -\frac{P(z_0)}{b}$$

oder $bv^t = -P(z_0)$. Wir haben jetzt zwei feste komplexe Zahlen, nämlich z_0 und v. Wir betrachten

$$P(z_0 + s \cdot v), \quad \text{wobei} \quad s \in [0, 1] \subset \mathbf{R}.$$

Es gilt

$$\begin{aligned}
P(z_0 + s \cdot v) &= P(z_0) + \underbrace{b \cdot (s \cdot v)^t}_{= -P(z_0) \cdot s^t} + Q(s \cdot v) \cdot (s \cdot v)^{t+1} \\
&= P(z_0)(1 - s^t) + (s \cdot v)^{t+1} Q(s \cdot v).
\end{aligned}$$

Bei festem v ist $v^{t+1} Q(s \cdot v)$ ein Polynom in s; daher stetig (Beispiel 1). Aber $s \in [0, 1]$, eine kompakte Menge (Beispiel 2(b)). Daher existiert nach dem Satz über kompakte Mengen eine obere Schranke, etwa die Zahl A (wobei wir auch annehmen werden, daß A die Bedingung $A > \|P(z_0)\| \geq 0$ erfüllt. D.h.

$$\left\| v^{t+1} Q(s \cdot v) \right\| \leq A, \forall s \in [0, 1].$$

Wir nehmen jetzt $s_0 = \frac{\|P(z_0)\|}{A+1}$. Dann gilt

$$\begin{array}{rcl} \|P(z_0 + s_0 v)\| & = & \big\|P(z_0)(1 - s_0^t) + (s_0 \cdot v)^{t+1} Q(s_0 \cdot v)\big\| \\ & \leq & \|P(z_0)\|(1 - s_0^t) + s_0^{t+1} \cdot A \\ & = & \|P(z_0)\| + \underbrace{s_0^t(s_0 - \frac{\|P(z_0)\|}{A})A}_{\text{eine negative Zahl}}. \end{array}$$

Folglich ist $\|P(z_0 + s_0 v)\| < \|P(z_0)\|$; ein Widerspruch, da

$$\|P(z_0)\| = \alpha = glb(\|P\|). \qquad \text{QED}$$

Kapitel 14

Das charakteristische Polynom

14.1 Das charakteristische Polynom

Seien $A, B \in M(n \times n; F)$ zwei beliebige $n \times n$ Matrizen. Die Summe $A + B$ ist die Matrix $A + B = (c_{ij})$ mit $c_{ij} = a_{ij} + b_{ij}$, für alle i und j. D.h. die entsprechenden Elemente werden einfach zusammen addiert. Ein besonderer Fall ist die Matrix $A - x \cdot I_n$. D.h. wenn

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}, \text{ dann ist}$$

$$A - x \cdot I_n = \begin{pmatrix} a_{11} - x & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - x & & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - x \end{pmatrix}.$$

Definition 55. Das charakterische Polynom von A, genannt $\chi_A(x)$, ist $det(A - x \cdot I_n)$.

Beispiel: $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Dann ist $A - x \cdot I_2 = \begin{pmatrix} 1-x & 2 \\ 3 & 4-x \end{pmatrix}$. Folglich ist

$$\begin{aligned} det(A - x \cdot I_2) &= (1-x) \cdot (4-x) - 2 \cdot 3 \\ &= x^2 - 5x - 2 \\ &= \left(x - \frac{5+\sqrt{33}}{2}\right) \cdot \left(x - \frac{5-\sqrt{33}}{2}\right). \end{aligned}$$

Das heißt, das charakterische Polynom in diesem Beispiel ist $x^2 - 5x - 2$, und dieses Polynom hat zwei *verschiedene* reelle Wurzeln, nämlich $\frac{5-\sqrt{33}}{2}$ und $\frac{5+\sqrt{33}}{2}$. Daraus können wir schließen, daß die lineare Abbildung $f : V \to V$, die durch A dargestellt ist, zwei verschiedene Eigenwerte besitzt, daher zwei verschiedene Eigenvektoren. Aber $dim(V) = 2$. Folglich gibt es eine Basis für V, bestehend aus lauter Eigenvektoren von f. D.h. es ist möglich, eine *diagonale* Matrix A' zu finden, die ähnlich ist zu A. Die Theorie, die wir jetzt besprechen werden, erlaubt uns alle diese Schlüsse zu ziehen.

Satz 14.1. *Sei $f : V \to V$ eine lineare Abbildung mit zugehöriger Matrix A bzgl. einer Basis $\{\alpha_1, \dots, \alpha_n\}$ für V. Die Zahl $\lambda \in F$ ist ein Eigenwert für $f \Leftrightarrow det(A - \lambda \cdot I_n) \equiv \chi_A(\lambda) = 0$.*

Beweis. Angenommen, λ ist ein Eigenwert für f. Dann existiert ein $\gamma /not = 0$ in V mit $f(\gamma) = \lambda \cdot \gamma$. Nun, $\lambda \cdot id : V \to V$, wobei $\lambda \cdot id(\xi) = \lambda \cdot \xi$, $\forall \xi \in V$, ist ebenfalls eine lineare Abbildung. Eine Summe von linearen Abbildungen—etwa $f - \lambda \cdot id : V \to V$—ist wieder eine lineare Abbildung. Nun, $f - \lambda \cdot id$ ist auf jeden Fall singulär, da $f - \lambda \cdot id(\gamma) = 0$. Aber die zu $f - \lambda \cdot id$ gehörige Matrix bzgl. der Basis $\{\alpha_1, \dots, \alpha_n\}$ ist $A - \lambda \cdot I_n$. Da diese Abbildung singulär ist, gilt $det(A - \lambda \cdot I_n) = 0$.

Es ist auch möglich, diesen Beweis umzukehren. Sei $det(A - \lambda \cdot I_n) = 0$. $\Rightarrow f - \lambda \cdot id$ ist singulär. $\Rightarrow \exists \gamma \in V$ $(\gamma = 0)$ mit $f - \lambda \cdot id(\gamma) = 0$. $\Rightarrow f(\gamma) = \lambda \cdot \gamma$. $\Rightarrow \lambda$ ist ein Eigenvektor für f. □

Korollar 14.1.1. *Seien A und A' ähnliche Matrizen. Dann sind die Wurzeln der jeweiligen charakteristischen Polynome identisch.*

(Dies folgt, da ähnliche Matrizen ja nur alternative Beschreibungen einer einzigen linearen Abbildung sind.) Es gilt noch mehr:

Satz 14.2. *Seien A und A' ähnliche Matrizen. Dann sind die zugehörigen charakteristischen Polynome identisch.*

Beweis. Sei $A' = C \cdot A \cdot C^{-1}$ mit $C \in GL(n; F)$. Dann ist (nach Satz 12.1 (und dem Korollar dazu))

$$\begin{aligned} det(A') &= det(C \cdot A \cdot C^{-1}) \\ &= det(C) \cdot det(A) \cdot det(C^{-1}) \\ &= det(C) \cdot det(A) \cdot det(C)^{-1} \\ &= det(A). \end{aligned}$$

Daher ist

$$\begin{aligned} \chi_{A'}(x) = det(A' - x \cdot I_n) &= det(C \cdot A \cdot C^{-1} - x \cdot C \cdot I_n \cdot C^{-1}) \\ &= det(C \cdot (A - x \cdot I_n) \cdot C^{-1}) \\ &= det(C) \cdot det(A - x \cdot I_n) \cdot det(C^{-1}) \\ &= det(A - x \cdot I_n) = \chi_A(x). \end{aligned}$$

□

Betrachten wir das characteristische Polynom einer vorgegebenen Matrix A etwas genauer. Wir können es so schreiben:

$$\chi_A(x) = c_n x^n + c_{n-1}x^{n-1} + \cdots + c_1 x + c_0.$$

Was sind die Koeffizienten c_i, für verschiedene i? Schauen wir uns Leibniz' Formel für $det(A - xI_n)$ an. Es handelt sich um eine Summe über alle möglichen Permutationen in S_n. Nun, es gibt eine ganz besondere Permutation, nämlich die Identitätsabbildung $id : \{1, \dots, n\} \to \{1, \dots, n\}$. Um diesen Term besonders zu behandeln, können wir $\chi_A(x)$ wie folgt schreiben.

$$\begin{aligned} \chi_A(x) &= (a_{11} - x)(a_{22} - x) \cdots (a_{nn} - x) \\ &\quad + \sum_{\substack{\sigma \in S_n \\ \sigma \neq id}} sign(\sigma) \prod_{i=1}^{n} \left\{ a_{i\,\sigma(i)} - x \cdot \delta_{i\,\sigma(i)} \right\} \end{aligned}$$

Das Bemerkenswerte an dieser Formel ist, daß alle Terme in der Summe über die Permutationen $\sigma \neq id$ Polynome in x sind vom Grad $\leq n-2$. Warum? Dies folgt, da für jede Permutation $\sigma \neq id$, $\exists i \in \{1, \dots, n\}$ mit $\sigma(i) \neq i$. Sei etwa $\sigma(i) = j$. Dann ist sicherlich auch $\sigma(j) \neq j$. Folglich sind mindestens zwei nicht-diagonale Elemente aus der Matrix $A - x \cdot I_n$ im σ-Term, und der Grad dieses Terms ist $\leq n-2$.

Dies bedeutet; wenn es darum geht, die Koeffizienten c_n und c_{n-1} zu berechnen, dann brauchen wir nur den Term

$$\begin{aligned} (a_{11} - x) \cdot (a_{22} - x) \cdots (a_{nn} - x) &= (-x)^n \\ &\quad - (a_{11} + a_{22} + \cdots + a_{nn})x^{n-1} \\ &\quad + \text{u.s.w.} \end{aligned}$$

zu betrachten. Das Ergebnis ist:

$$c_n = \begin{cases} 1, & \text{falls } n \text{ gerade,} \\ -1 & \text{sonst,} \end{cases}$$

$$c_{n-1} = -\sum_{i=1}^{n} a_{ii}.$$

Nun, die Zahl c_n ist—wie man sieht—trivial zu berechnen. Aber c_{n-1} ist doch interessant.

Definition 56. Sei $A = (a_{ij}) \in M(n \times n; F)$. Die Zahl $\sum_{i=1}^{n} a_{ii}$ heißt die *Spur* von A.

Korollar 14.2.1. *Seinen A und A' ähnlich. Dann ist*

$$Spur(A) = Spur(A').$$

Kapitel 15

Innere Produkte, Normen

Es ist sicherlich so, daß die Idee von ‘Länge’ sehr wichtig in der Beschreibung einer Geometrie ist. Dafür brauchen wir *orthonormale* Basen für unsere Vektorräume. Eine solche Basis $\{\alpha_1, \ldots, \alpha_n\}$ hat die zwei allgemeine Eigenschaften

- Die ‘Länge’ von α_i ist eins, $\forall i = 1, \ldots, n$.
- α_i ist ‘senkrecht zu’ α_j, für $i \neq j$.

Was sind diese Begriffe: ‘Länge’ und ‘senkrecht zu’? Nun, in dem ‘normalen’ reellen n-dimensionalen Raum $\mathbf{R}^n$ — mit der kanonischen Basis — ist ein typischer Vekter etwa $v = (v_1, \ldots, v_n)$. Die ‘Länge’ von v wird hier einfach als die (nicht-negative) reelle Zahl

$$\|v\| = \sqrt{v_1^2 + \cdots + v_n^2}$$

festgelegt.

Wann sind zwei Vektoren in $\mathbf{R}^n$ zueinander ‘senkrecht’?. Einfachheitshalber, betrachten wir zunächst die ebene Geometrie $\mathbf{R}^2$. Seien $\xi = (r \cdot \cos\theta, r \cdot \sin\theta)$ und $\zeta = (s \cdot \cos\phi, s \cdot \sin\phi)$ zwei beliebige Vektoren in $\mathbf{R}^2$, dargestellt durch Polarkoordinaten bzgl. der kanonischen Basis. Nun, die folgende Formel ist in der Trigonometrie wohlbekannt:

$$\cos\theta\cos\phi + \sin\theta\sin\phi = \cos(\theta - \phi).$$

Was bedeutet es, wenn man sagt, daß ξ und ζ zueinander senkrecht sind? Nichts anderes als daß $\theta - \phi = \pm\frac{\pi}{2}$; d.h. $\cos(\theta - \phi) = 0$. Aber dies gilt genau dann, wenn

$$\xi \cdot \zeta \equiv r\cos\theta \cdot s\cos\phi + r\sin\theta \cdot s\sin\phi = 0.$$

D.h. sei $\xi = (x_1, x_2)$ und $\zeta = (z_1, z_2)$. Dann ist ξ zu ζ senkrecht genau dann, wenn

$$\xi \cdot \zeta \equiv x_1 z_1 + x_2 z_2 = 0.$$

Nun, es gilt

$$\|\xi\| = \sqrt{\xi \cdot \xi}, \quad \|\zeta\| = \sqrt{\zeta \cdot \zeta}$$

und auch

$$\cos\Phi = \frac{\xi \cdot \zeta}{\|\xi\| \cdot \|\zeta\|},$$

wobei Φ der Winkel zwischen ξ und ζ ist. Das ‘innere Produkt’ $\xi \cdot \zeta$ gibt uns daher genug Informationen, um sowohl die Länge eines Vekors, als auch den Winkel zwischen beliebig vorgegebenen Paaren von Vektoren zu berechnen.

Satz 15.1. *Seien $v = (v_1, \cdots, v_n)$ und $\omega = (w_1, \cdots, w_n)$ zwei Vektoren in $\mathbf{R}^n$ (bzgl. des kanonischen Koordinatensystems). Die Abbildung*

$$s : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$$

wird wie folgt definiert:

$$s(v, \omega) \equiv v \cdot \omega \equiv \sum_{i=1}^{n} v_i w_i$$

Dann gilt

1. $v \cdot (\omega + \xi) = v \cdot \omega + v \cdot \xi, \quad \xi \in \mathbf{R}^n,$

2. $(\omega + \xi) \cdot v = \omega \cdot v + \xi \cdot v$,

3. $x(v \cdot \omega) = (xv) \cdot \omega = v \cdot (x\omega), \quad \forall x \in \mathbf{R}$.

Beweis. trivial. □

Allgemeiner:

Definition 57. Sei F ein Körper und V, W Vektorräume über F. Eine Abbildung $s : V \times W \to F$ heißt *Bilinearform*, falls die folgenden Eigenschaften für beliebige $\xi, \xi_1, \xi_2 \in V$, $\zeta, \zeta_1, \zeta_2 \in W$ und $a \in F$ gelten:

(BF1:) $s(\xi_1 + \xi_2, \zeta) = s(\xi_1, \zeta) + s(\xi_2, \zeta)$
$s(a\xi, \zeta) = as(\xi, \zeta)$,

(BF2:) $s(\xi, \zeta_1 + \zeta_2) = s(\xi, \zeta_1) + s(\xi, \zeta_2)$
$s(\xi, a\zeta) = as(\xi, \zeta)$,

Die Abbildung s heißt *symmetrisch*, falls $V = W$ und falls gilt:

(SBF:) $s(\xi, \zeta) = s(\zeta, \xi)$, $\forall \xi, \zeta \in V$.

Sei nun $s : V \times W \to F$ eine Bilinearform. s heißt *nicht ausgeartet*, falls gilt:[1]

(DP1:) aus $s(\xi, \zeta) = 0$, $\forall \zeta \in W$ folgt stets $\xi = 0$,

(DP2:) aus $s(\xi, \zeta) = 0$, $\forall \xi \in V$ folgt stets $\zeta = 0$.

Sonst ist s *ausgeartet.*
Falls F der besondere Körper $\mathbf{C}$ (die komplexen Zahlen) ist, dann gibt es weitere Möglichkeiten für die Festlegung von technischen Begriffen. Sei $f : V \to W$ eine Abbildung zwischen zwei Vektorräumen über $\mathbf{C}$. f heißt *semilinear*, falls für beliebige $\xi, \xi_1, \xi_2 \in V$ und $z \in \mathbf{C}$ gilt:

(SL1:) $f(\xi_1 + \xi_2) = f(\xi_1) + f(\xi_2)$,

(SL2:) $f(z\xi) = \overline{z} f(\xi)$.

Sei $f : V \to W$ eine semilineare Bijektion. Dann heißt f ein *semi-Isomorphismus*. Sei $s : V \times W \to \mathbf{C}$ eine Abbildung, wobei V und W Vektorräume über $\mathbf{C}$ sind. s heißt *Sesquilinearform*, falls gilt:

(SF1:) Die Abbildung $s(\cdot, \zeta) : V \to \mathbf{C}$, definiert durch die Regel $\xi \to s(\xi, \zeta)$, ist für jedes ζ semilinear,

(SF2:) Die Abbildung $s(\xi, \cdot) : W \to \mathbf{C}$, definiert durch die Regel $\zeta \to s(\xi, \zeta)$, ist für jedes ξ eine lineare Abbildung.

Sei jetzt $s : V \times V \to \mathbf{C}$ eine Sesquilinearform. s heißt *Hermitesche Form*, falls gilt:

(HF:) $s(\xi, \zeta) = \overline{s(\zeta, \xi)}$, $\forall \xi, \zeta \in V$.

Bemerkung. Sei $s : V \times V \to \mathbf{C}$ eine Hermitesche Form. Dann ist

$$s(\xi, \xi) = \overline{s(\xi, \xi)}$$

(nach Eigenschaft HF). Folglich ist $s(\xi, \xi) \in \mathbf{R}$. Sei daher

$$s : V \times V \to \{\text{entweder } \mathbf{C} (\text{oder} \mathbf{R})\}$$

eine Hermitesche (bzw. symmetrische) Form. s heißt *positiv definit*, falls $s(\xi, \xi) > 0$, $\forall \xi / not = 0$.

Definition 58. Sei $s : V \times V \to F$ eine positiv definite Form (wobei F = entweder $\mathbf{R}$ oder $\mathbf{C}$ und V ein F Vektorraum ist) so, daß s entweder symmetrisch oder Hermitesch ist (je nachdem, ob $F = \mathbf{R}$ oder $\mathbf{C}$). Dann heißt s ein *Skalarprodukt* in V. Man schreibt $\langle , \rangle : V \times V \to F$. D.h. $s(\xi, \zeta) = \langle \xi, \zeta \rangle$.

Beispiele:

1. $v = (v_1, \ldots, v_n)$, $\omega = (w_1, \ldots, w_n) \in \mathbf{R}^n$.

$$\langle v, \omega \rangle = v^t \cdot \omega = (v_1, v_2, \ldots, v_n) \cdot \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \sum_{j=1}^{n} v_j w_j$$

[1]DP = 'dual pairing'

2. Für den Hermiteschen Fall, seien $v = (v_1, \ldots, v_n)$, und $\omega = (w_1, \ldots, w_n)$ in $\mathbf{C}^n$. Dann ist

$$\langle v, \omega \rangle = \overline{v}^t \cdot \omega = (\overline{v_1}, \overline{v_2}, \ldots, \overline{v_n}) \cdot \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \sum_{j=1}^{n} \overline{v_j} w_j$$

Dies sind die *kanonischen* Skalarprodukte in $\mathbf{R}^n$ bzw. $\mathbf{C}^n$.

Definition 59. Sei V ein Vektorraum über $\mathbf{R}$ mit Skalarprodukt $\langle , \rangle$. Dann heißt V ein *Euklidischer* Vektorraum. Falls V ein Vektorraum über $\mathbf{C}$ ist, mit Skalarprodukt $\langle , \rangle$, dann heißt V ein *unitärer* Vektorraum.

Warum nehmen wir diese 'semilineare' Regel für Skalarprodukte in dem komplexen Fall, statt die einfachere 'bilineare' Regel? Die Antwort ist, daß wir Skalarprodukte eigentlich nur definiert haben, um die 'Länge' von Vektoren zu definieren. Die Länge eines Vektors ξ ist nämlich $\|\xi\| = \sqrt{\langle \xi, \xi \rangle}$. In $\mathbf{R}^n$ sei $\xi = (x_1, \ldots, x_n)$. Dann ist

$$\|\xi\| = \sqrt{\langle \xi, \xi \rangle} = \sqrt{\sum_{i=1}^{n} x_i^2} = \sqrt{\sum_{i=1}^{n} |x_i|^2}.$$

Für alle reellen Zahlen $x \in \mathbf{R}$ gilt natürlich stets $x^2 = |x|^2$; d.h. das Absolutbetragzeichen kann einfach ignoriert werden. Diese Bequemlichkeit gibt es bei den komplexen aber Zahlen nicht! Nehmen wir z.B. die 'einfachste' imaginäre Zahl $i \quad (\equiv \sqrt{-1})$. In der Darstellung der komplexen Zahlen in $\mathbf{R}^2$ mittels Polarkoordinaten haben wir $i = (0,1) = (1 \cdot \cos(\pi/2), 1 \cdot \sin(\pi/2))$. D.h. die Länge des 'Vektors' i ist $\|i\| = 1$, und der Winkel zur reellen Achse ist $\pi/2$. Aber $i^2 = (\sqrt{-1})^2 = -1 = -\|i\|^2$. Im allgemeinen gilt $\|z\|^2 = \overline{z} \cdot z$ für alle $z \in \mathbf{C}$. Daher ist die Länge eines Vektors $\zeta = (z_1, \ldots, z_n)$ im komplexen Raum (d.h. $z_i \in \mathbf{C}$, $\forall i = 1, \ldots, n$, oder $\zeta \in \mathbf{C}^n$) einfach

$$\|\zeta\| = \sqrt{\sum_{i=1}^{n} \|z_i\|^2} = \sqrt{\sum_{i=1}^{n} \overline{z_i} \cdot z_i} = \sqrt{\langle \zeta, \zeta \rangle}.$$

Nun, wir haben das Wort 'Länge' immer wieder benutzt. Der mathematische Fachbegriff heißt 'Norm'.

Definition 60. Sei $F = \mathbf{R}$ oder $\mathbf{C}$ und sei $\| \cdot \| : F \to \mathbf{R}$ der Absolutbetragsfunktion in F. Wir betrachten ein Vektorraum V über F. Eine Abbildung $\| \cdot \| : V \to \mathbf{R}$ heißt eine *Norm* auf V, falls gilt:

(N1:) $\|av\| = \|a\| \cdot \|v\|$, für alle $a \in F$ und $v \in V$,

(N2:) $\|v + \omega\| \leq \|v\| + \|\omega\|$ (die Dreiecksungleichung),

(N3:) $\|v\| = 0 \Leftrightarrow v = 0$.

V versehen mit der Norm $\| \cdot \|$ heißt ein normierter Vektorraum.

Satz 15.2 (Cauchy-Schwarz Ungleichung). *Sei V ein Euklidischer (bzw. unitärer) Vektorraum. Sei $\| \cdot \| : V \to \mathbf{R}$ definiert durch die Regel $\|v\| = \sqrt{\langle v, v \rangle}$, für alle $v \in V$. Dann gilt $\|\langle v, \omega \rangle\| \leq \|v\| \cdot \|\omega\|$, für alle $v, \omega \in V$. Insbesondere gilt $\|\langle v, \omega \rangle\| = \|v\| \cdot \|\omega\| \Leftrightarrow v$, ω linear abhängig sind.*

Beweis. Da die Quadratwurzelfunktion monoton ist, brauchen wir nur zu zeigen, daß $\|\langle v, \omega \rangle\|^2 \leq \langle v, v \rangle \cdot \langle \omega, \omega \rangle$. Nun, falls $\omega = 0$, dann ist $\langle v, \omega \rangle = \langle \omega, \omega \rangle = 0$, da $\langle \cdot \rangle$ bilinear (bzw. semilinear) ist. Dann ist sicherlich $0 \leq 0 = \langle v, v \rangle \cdot 0$. Sei daher angenommen, daß $\omega \neq 0$. Folglich ist $\langle \omega, \omega \rangle \neq 0$, da ein Skalarprodukt positiv definit ist. Nehme jetzt

$$a = \frac{\overline{\langle v, \omega \rangle}}{\langle \omega, \omega \rangle} \in \mathbf{R} \quad \text{(bzw. } \mathbf{C}\text{)}.$$

Dann gilt

$$\begin{aligned} 0 &\leq \langle v - a\omega, v - a\omega \rangle \qquad \text{(da positiv definit)} \\ &= \langle v, v - a\omega \rangle + \langle -a\omega, v - a\omega \rangle \\ &= \langle v, v \rangle + \langle v, -a\omega \rangle + \langle -a\omega, v \rangle + \langle -a\omega, -a\omega \rangle \\ &= \langle v, v \rangle - \underbrace{a\langle v, \omega \rangle}_{\frac{\overline{\langle v,\omega \rangle} \cdot \langle v,\omega \rangle}{\langle \omega,\omega \rangle}} \underbrace{- \underbrace{\overline{a}\overline{\langle v, \omega \rangle}}_{\frac{\langle v,\omega \rangle \cdot \overline{\langle v,\omega \rangle}}{\langle \omega,\omega \rangle}} + \underbrace{a\overline{a}\langle \omega, \omega \rangle}_{\frac{\langle v,\omega \rangle \cdot \overline{\langle v,\omega \rangle}}{\langle \omega,\omega \rangle}}}_{=0} \end{aligned}$$

Daher

$$0 \leq \langle v, v\rangle \cdot \langle \omega, \omega\rangle - \underbrace{\langle v, \omega\rangle \cdot \overline{\langle v, \omega\rangle}}_{\|\langle v,\omega\rangle\|^2}.$$

Wann gilt die Gleichheit? Falls $\omega = 0$, dann gilt sie. Aber falls $\omega \neq 0$, dann gilt die Gleichheit, falls

$$\begin{aligned} 0 &= \langle v - a\omega, v - a\omega\rangle \\ &\Leftrightarrow v - a\omega = 0 \qquad \text{(da Skalarprodukte positiv definit sind)} \\ &\Leftrightarrow v = a\omega \\ &\Leftrightarrow v, \omega \qquad \text{linear abhängig} \end{aligned}$$

□

Korollar 15.2.1. *Sei V ein Vektorraum mit Skalarprodukt. Dann besitzt V eine Norm, nämlich $\|\cdot\| : V \to \mathbf{R}$, gegeben durch $\|v\| = \sqrt{\langle v, v\rangle}$, für beliebige $v \in V$.*

Beweis. $\|v + \omega\| \leq \|v\| + \|\omega\|$ folgt aus der Cauchy-Schwarz Ungleichung, und zwar:

$$\begin{aligned} \|v + \omega\|^2 &= \langle v + \omega, v + \omega\rangle \\ &= \langle v, v\rangle + \langle v, \omega\rangle + \langle \omega, v\rangle + \langle \omega, \omega\rangle \\ &\leq \langle v, v\rangle + \|\langle v, \omega\rangle\| + \|\langle \omega, v\rangle\| + \langle \omega, \omega\rangle \\ &= \langle v, v\rangle + 2\|\langle v, \omega\rangle\| + \langle \omega, \omega\rangle \\ &\leq \langle v, v\rangle + 2\|v\| \cdot \|\omega\| + \langle \omega, \omega\rangle \qquad \text{(Cauchy-Schwarz Ungleichung)} \\ &= (\|v\| + \|\omega\|)^2. \end{aligned}$$

Die erste Ungleichung hier folgt, da $\langle v, \omega\rangle + \langle \omega, v\rangle$ eine reelle Zahl sein muß.

Nun, $\|av\| = \|a\| \cdot \|v\|$, für $a \in F$ ist klar. Für die positiv Definitheit gilt:

$$\|v\| = 0 \Leftrightarrow \sqrt{\langle v, v\rangle} = 0 \Leftrightarrow \langle v, v\rangle = 0 \Leftrightarrow v = 0.$$

□

15.1 Beispiele für Normen

Die Standardbeispiele sind natürlich unsere gewöhnlichen Vektorräume $\mathbf{R}^n$ und $\mathbf{C}^n$, versehen mit Normfunktionen, die den Längen von Vektoren entsprechen. Es gibt aber auch einige andere Normen, die in der Mathematik immer wieder vorkommen. Insgesamt spielt dieser Begriff eine außerordentlich wichtige Rolle in der gesamten modernen Analysis.

1. Sei etwa $\mathcal{F}_{[0,1]}$ die Menge der *beschränkten* Funktionen vom Standard-Einheitsintervall in die reellen Zahlen. (Beschränkt heißt, daß die Menge $\{|f(x)| : x \in [0,1]\}$ (für jedes $f \in \mathcal{F}_{[0,1]}$) eine obere Schranke besitzt. Wir definieren
$$\|f\|_0 \equiv lub\{|f(x)| : x \in [0,1]\}.$$
Dann ist $\|\cdot\|_0 : \mathcal{F}_{[0,1]} \to \mathbf{R}$ eine Norm. Denn (N1) und (N3) sind trivialerweise erfüllt. Für (N2) seien $f, g \in \mathcal{F}_{[0,1]}$. Dann ist
$$\|f + g\|_0 = lub\{|f(x) + g(x)| : x \in [0,1]\}.$$
Um einen Widerspruch zu erzeugen, sei angenommen, daß $\|f + g\|_0 > \|f\|_0 + \|g\|_0$. D.h. $\exists\ x_0 \in [0,1]$ mit $|f(x_0) + g(x_0)| > \|f\|_0 + \|g\|_0$. D.h. für alle möglichen Zahlen $a, b \in [0,1]$ wäre $|f(x_0) + g(x_0)| > |f(a)| + |g(b)|$. Dies stimmt aber nicht, da $|f(x_0) + g(x_0)| \leq |f(x_0)| + |g(x_0)|$ nach der Dreiecksungleichung für reelle Zahlen.

2. $\|\cdot\|_m : \mathbf{R}^n \to \mathbf{R}$ wird auf $\mathbf{R}^n$ wie folgt definiert. Für jedes $\xi = (x_1, \ldots, x_n)$ sei $\|\xi\|_m \equiv max\{|x_1|, \ldots, |x_n|\}$. Wieder sind (N1) und (N3) trivialerweise erfüllt. Für (N2) seien $v, \omega \in \mathbf{R}^n$. Es gilt
$$\begin{aligned} \|v + \omega\| &= max\{|v_1 + w_1|, \ldots, |v_n + w_n|\} \\ &\leq max\{|v_1| + |w_1|, \ldots, |v_n| + |w_n|\} \\ &\leq max\{|v_1|, \ldots, |v_n|\} + max\{|w_1|, \ldots, |w_n|\} \\ &= \|v\|_m + \|\omega\|_m \end{aligned}$$

Zum Schluß noch einen weiteren Begriff, der auch sehr wichtig in der Analysis ist, nämlich *metrische Räume*:

Definition 61. Sei M irgendeine Menge. Eine Abbildung

$$d : M \times M \to \mathbf{R}$$

heißt eine *Metrik*, falls gilt:

1. $d(x, y) = 0 \Leftrightarrow x = y, \quad \forall x, y \in M,$
2. $d(x, y) = d(y, x), \quad \forall x, y \in M,$
3. $d(x, z) \leq d(x, y) + d(y, z), \quad \forall x, y, z \in M.$

Eine Menge, die versehen ist mit einer Metrik, heißt ein metrischer Raum.

Satz 15.3. *Jeder normierte Vektorraum $V, \|\cdot\|$ ist ein metrischer Raum mit Metrik $d(x, y) = \|x - y\|$.*

Beweis. (i) und (ii) sind trivial. Zu (iii).

$$\begin{aligned} d(x, z) = \|x - z\| &= \|(x - y) + (y - z)\| \\ &\leq \|x - y\| + \|y - z\| \qquad \text{(Cauchy-Schwarz)} \\ &= d(x, y) + d(y, z) \end{aligned}$$

□

Aber der Begriff 'Metrik' ist sehr viel allgemeiner als der Begriff 'Norm'. Nehmen wir z.B. die sphärische Geometrie. Die n-dimensionale Sphäre ist

$$S^n = \{x \in \mathbf{R}^{n+1} : \|x\| = 1\}.$$

Nun, S^n ist offensichtlich kein Vektorraum. (Z.B. S^n ist kompakt, aber jeder nicht-triviale Vektorraum über $\mathbf{R}$ ist nicht kompakt.) Andererseits ist jede S^n ein metrischer Raum. Der 'Abstand' zwischen zwei Punkten $x, y \in S^n$ ist die Länge (einer) der kürzesten Verbindungsstrecken entlang einem 'Großkreis' von x nach y.

Der 'Abstandsbegriff' wird noch weiter verallgemeinert, wenn es darum geht, 'topologische Räume' zu definieren. Jede Metrik ist eine 'Topologie', aber nicht umgekehrt. Dies ist der übliche Grad der Verallgemeinerungen in der modernen Mathematik, wenn es darum geht, über 'Geometrien' zu sprechen. Diejenigen, die Lust haben, solche Ideen weiter zu verfolgen, können die Vorlesung 'Topologie' in den nächsten Semestern höhren.

Kapitel 16

Symmetrische und Unitäre Matrizen; Orthonormalbasen

Definition 62. Sei $A = (a_{ij}) \in M(n \times n; \mathbf{R})$ eine $n \times n$ Matrix von reellen Zahlen. A heißt *symmetrisch*, falls $A = A^t$; d.h. $a_{ij} = a_{ji}$, $\forall i, j \in \{1, \ldots, n\}$. Andererseits, sei $B = (b_{ij}) \in M(n \times n; \mathbf{C})$ eine $n \times n$ Matrix von komplexen Zahlen. B heißt *Hermitesch*, falls $B = \overline{B}^t$; d.h. $b_{ij} = \overline{b}_{ji}$, $\forall i, j \in \{1, \ldots, n\}$.

Bemerkung. Wegen $\mathbf{R} \subset \mathbf{C}$ gilt $M(n \times n; \mathbf{R}) \subset M(n \times n; \mathbf{C})$. D.h. eine reelle Matrix ist auch eine komplexe Matrix. Daher gilt für $A \in M(n \times n; \mathbf{R}) \Rightarrow$: falls A symmetrisch, dann ist A auch Hermitesch.

Satz 16.1. *Sei $s : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ eine symmetrische Bilinearform auf $\mathbf{R}^n$ und sei $\{\epsilon_1, \ldots, \epsilon_n\}$ die kanonische Basis für $\mathbf{R}^n$. Die Matrix $A \in M(n \times n; \mathbf{R})$ wird wie folgt definiert: $A = (a_{ij})$, wobei $a_{ij} = s(\epsilon_i, \epsilon_j)$, $\forall i, j$. Dann ist A symmetrisch. Umgekehrt sei A irgendeine symmetrische Matrix in $M(n \times n; \mathbf{R})$. Die Bilinearform $s : \mathbf{R}^n \times \mathbf{R}^n \to R$, definiert durch die Regel $s(\epsilon_i, \epsilon_j) = a_{ij}$ $\forall i, j$, ist symmetrisch. (D.h. es existiert eine Bijektion zwischen der Menge der symmetrischen $n \times n$ Matrizen und der Menge der symmetrischen Bilinearformen auf $\mathbf{R}^n$.)*

Beweis. $s : R^n \times \mathbf{R}^n \to \mathbf{R}$ ist symmetrisch

$$\begin{aligned} &\Leftrightarrow \quad a_{ij} = s(\epsilon_i, \epsilon_j) = s(\epsilon_j, \epsilon_i) = a_{ji} \\ &\Leftrightarrow \quad A \quad \text{ist symmetrisch.} \end{aligned}$$

□

Korollar 16.1.1. *(i) Der Satz gilt auch für symmetrische Matrizen über Körpern im allgemeinen. (ii) Für den besonderen Körper $\mathbf{C}$ gilt ein analoger Satz über Hermitesche Matrizen und Hermitesche Formen.*

Bisher haben wir stets die kanonische Basis für unsere Argumente genommen. Es ist aber wichtig, eine Methode zu finden, um auch andere Basen einzubeziehen, um letzten Endes die geometrischen Ideen von 'Längen' und 'Winkeln' festzulegen, unabhängig von irgendeiner bestimmten Basis.

Definition 63. Sei V ein (endlich dimensionaler) Vektorraum über dem Körper F ($F = \mathbf{R}$ oder $\mathbf{C}$). Sei $\mathcal{A} = \{\alpha_1, \ldots, \alpha_n\}$ eine Basis für V und sei $s : V \times V \to F$ eine symmetrische (bzw. Hermitesche) Form. Die *darstellende Matrix* für s bzgl. der Basis $\mathcal{A}$ ist $M_{\mathcal{A}}(s) = (m_{ij})$, $i, j \in \{1, \ldots, n\}$, wobei $m_{ij} = s(\alpha_i, \alpha_j)$.

Seien jetzt $\xi, \zeta \in V$ beliebige Vektoren. Was ist dann $s(\xi, \zeta)$? Sei $\xi = x_1\alpha_1 + \cdots + x_n\alpha_n$ und $\zeta = z_1\alpha_1 +$

$\cdots + z_n\alpha_n$. Es gilt

$$\underbrace{\begin{aligned} s(\xi,\zeta) &= s(\sum_{i=1}^n x_i\alpha_i, \sum_{i=1}^n z_i\alpha_i) \\ &= \sum_{i=1}^n\sum_{j=1}^n s(x_i\alpha_i, z_j\alpha_j) \\ &= \sum_{i=1}^n\sum_{j=1}^n x_i s(\alpha_i,\alpha_j) z_j \\ &= \sum_{i=1}^n\sum_{j=1}^n x_i m_{ij} z_j \\ &= \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \cdot M_{\mathcal{A}}(s) \cdot \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \end{aligned}}_{\text{Matrizenmultiplikation}}$$

Satz 16.2 (Transformationsformel). *Sei V ein endlich-dimensionaler Vektorraum über dem Körper F. Sei s eine symetrische Bilinearform auf V und seien $\mathcal{A}$ und $\mathcal{B}$ zwei Basen für V. Sei weiterhin $S \in GL(n,F)$ die Matrix des Basiswechsels. (d.h. sei $\mathcal{A} = (\alpha_1,\ldots,\alpha_n)$ und $\mathcal{B} = (\beta_1,\ldots,\beta_n)$. Ein Vektor ξ hat die Koordinaten $(x_1^\alpha,\ldots,x_n^\alpha)$ bzgl. des $\mathcal{A}$ Systems und die Koordinaten $(x_1^\beta,\ldots,x_n^\beta)$ bzgl. des $\mathcal{B}$ Systems. Dann ist $S \cdot \begin{pmatrix} x_1^\alpha \\ \vdots \\ x_n^\alpha \end{pmatrix} = \begin{pmatrix} x_1^\beta \\ \vdots \\ x_n^\beta \end{pmatrix}$.) Sei auch $A = M_{\mathcal{A}}(s)$, $B = M_{\mathcal{B}}(s)$. Dann gilt $A = S^t \cdot B \cdot S$.*

Beweis. Seien i, j beliebig in $\{1,\ldots,n\}$ und für jedes i sei ϵ_i die $n \times 1$ Matrix

$$\epsilon_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow \text{i.te Stelle} \enspace .$$

Dann ist

$$\begin{aligned} s(\alpha_i,\alpha_j) &= ij\text{-te Komponente der Matrix } A \\ &= \epsilon_i^t \cdot A \cdot \epsilon_j \\ &= (S\cdot\epsilon_i)^t \cdot B \cdot (S \cdot \epsilon_j) \\ &= \epsilon_i^t(S^t \cdot B \cdot S)\cdot\epsilon_j \\ &= ij\text{-te Komponente der Matrix } S^t \cdot B \cdot S. \end{aligned}$$

□

Korollar 16.2.1. *Die Bedingungen seien wie im Satz 16.2, jedoch $F = \mathbf{C}$ und 'Hermitesch' statt 'symmetrisch'. Dann gilt $A = \overline{S}^t \cdot B \cdot S$.*

Korollar 16.2.2. *Seien V, W Vektorräume über F und $s : V \times W \to F$ eine Bilinearform. Seien $\mathcal{A}_1$ und $\mathcal{A}_2$ zwei Basen für V und $\mathcal{B}_1$ und $\mathcal{B}_2$ zwei Basen für W, mit Matrizen S und T, die die Basiswechsel $\mathcal{A}_1 \to \mathcal{A}_2$, bzw. $\mathcal{B}_1 \to \mathcal{B}_\in$ darstellen. Sei A die darstellende Matrix für s bzgl. $\mathcal{A}_1$ und $\mathcal{B}_1$ und sei B die darstellende Matrix für s bzgl. $\mathcal{A}_2$ und $\mathcal{B}_2$. Dann ist $B = S^t \cdot A \cdot T$.*

Definition 64. Sei $s : V \times V \to F$ eine symmetrische (bzw. Hermitesche) Bilinearform. Sei die Abbildung $q_s : V \to F$ definiert durch die Regel $q_s(\xi) = s(\xi,\xi)$. Diese Abbildung heißt die zu s zugeordnete *quadratische Form.* Die Vektoren $\xi \in V$ mit $q_s(\xi) = 0$ heißen *isotrop.*

Warum die Bezeichnung 'quadratische Form'? Dies ist offensichtlich, nach der Darstellung $s(\xi,\xi) = \sum_{i=1}^n x_i^2$. Warum 'isotrop'? Nun, das Wort kommt eigentlich gar nicht vor im alltäglichen Gebrauch. Auch im Rahmen der Mathematik ist unsere 'gewöhnliche' Geometrie $\mathbf{R}^3$, und hier (wenn wir die übliche Bilinearform nehmen) gibt es keine nicht-trivialen isotropen Vektoren. Aber laut Relativitätstheorie ist die *eigentliche* Geometrie unserer wirklichen Welt nicht $\mathbf{R}^3$, sondern M^4 — der 4-dimensionaler 'Minkowski-Raum. Dies ist ein Vektorraum mit 'kanonischer Basis' $\{\xi,\psi,\zeta,\tau\}$, nämlich die x- y- und z-'Achsen' der 'räumlichen' Dimensionen und die t-'Achse' der 'Zeit'. Seien $v_1 = (x_1,y_1,z_1,t_1)$ und $v_2 = (x_2,y_2,z_2,t_2)$ zwei Punkte in unserer eigentlichen, reellen Welt. Der 'Abstand' zwischen v_1 und v_2 ist dann[1]

$$\begin{aligned}\|v_1 - v_2\| &= \sqrt{q_s(v_1 - v_2)}\\ &= \sqrt{(t_1-t_2)^2 - (x_1-x_2)^2 - (y_1-y_2)^2 - (z_1-z_2)^2}.\end{aligned}$$

Man sieht sofort, daß $\|v_1 - v_2\| = 0 \not\Rightarrow v_1 = v_2$. Im Gegenteil, $\|v_1 - v_2\| = 0$ bedeutet, daß es möglich wäre, v_1 und v_2 durch einen Lichtstrahl zu verbinden.

16.1 Orthonormale Basen

Sei V ein Vektorraum über dem Körper F ($F = \mathbf{R}$ oder $\mathbf{C}$) mit Skalarprodukt $s : V \times V \to F$, und sei $\mathcal{A} = \{\alpha_1,\dots,\alpha_n\}$ eine Basis für V. Seien $\xi = (x_1,\dots,x_n)$ und $\zeta = (z_1,\dots,z_n)$ zwei Vektoren in V mit gegebenen Koordinaten bzgl. der Basis $\mathcal{A}$. Welche Bedingungen muß die Basis $\mathcal{A}$ erfüllen, so daß wir einfach

$$\langle\xi,\zeta\rangle = \sum_{i=1}^n x_i z_i$$

schreiben können? Die Antwort: $\mathcal{A}$ muß eine Orthonormalbasis sein.

Definition 65. Sei V ein Euklidischer (bzw. unitärer) Vektorraum.

1. Seien v und ω in V. Falls $\langle v,\omega\rangle = 0$, dann sind die zwei Vektoren zueinander *orthogonal*. Man schreibt dann $v \perp \omega$.
2. Falls $U, W \subset V$ lineare Unterräume sind mit $v \perp \omega$, für alle $v \in U$ und $\omega \in W$, dann heißen U und W zueinander orthogonal; $U \perp W$.
3. Sei $U \subset V$ ein linearer Unterraum. Die Menge
$$U^\perp \equiv \{\omega \in V : \omega \perp v, \forall v \in U\}$$
heißt das *orthogonale Komplement* zu U. Trivialerweise gilt: $U^\perp \subset V$ ist ein linearer Unterraum und $U^\perp \perp U$.
4. Sei $X \subset V$ irgendeine Teilmenge. X heißt *orthogonal*, falls $x_1 \perp x_2$, $\forall x_1, x_2 \in X$ $(x_1 \neq x_2)$.
5. Sei $X \subset V$ orthogonal. Falls zusätzlich gilt $\|x\| = 1, \forall x \in X$, dann heißt X *orthonormal*.
6. Falls die orthonormale Menge $X \subset V$ eine Basis für V ist, dann heißt X eine *Orthonormalbasis*.
7. Seien $V_1,\dots,V_k \subset V$ lineare Unterräume. V heißt die *orthogonale Summe* von $V_1,\dots,V_k$, falls $V = V_1 \oplus \cdots \oplus V_k$ und $V_i \perp V_j$, $\forall i \neq j$.

Sei nun $\{\alpha_1,\dots,\alpha_n\} \subset V$ eine Orthonormalbasis und seien $\xi = (x_1,\dots,x_n)$ und $\zeta = (z_1,\dots,z_n)$ zwei beliebige Vektoren in V. Dann ist

$$\begin{aligned}\langle\xi,\zeta\rangle &= \langle x_1\alpha_1 + \cdots + x_n\alpha_n, z_1\alpha_1 + \cdots + z_n\alpha_n\rangle\\ &= \sum_{i=1}^n\left(\sum_{j=1}^n x_i z_j \langle\alpha_i,\alpha_j\rangle\right)\\ &= \sum_{i=1}^n x_i z_i \langle\alpha_i,\alpha_i\rangle\\ &= \sum_{i=1}^n x_i z_i\end{aligned}$$

[1]Dies ist eine 'normale' reelle Zahl, falls v_1 in der Zeit 'über' oder 'unter' v_2 steht. Für 'normale' raumartige Abstände ist diese Quadratwurzel aber eine imaginäre Zahl. Diejenigen, die diese Ideen weiter verfolgen wollen, sollten Physik studieren!

Die vorletzte Gleichung folgt, da α_i und α_j zueinander orthogonal sind, für $i \neq j$. Die letzte Gleichung folgt, da stets $\langle \alpha_i, \alpha_i \rangle = 1$. Man sieht dann, daß gerade die Orthonomalbasen geeignet sind zur Berechnung von Skalarprodukten. Gibt es immer solche Basen?

Satz 16.3. *Sei $\{\alpha_1, \ldots, \alpha_m\} \subset V$ orthogonal (mit $\alpha_i \neq 0$, für alle i). Dann ist $\{\alpha_1, \ldots, \alpha_m\}$ linear unabhängig.*

Beweis. Wir können auch annehmen, daß die Menge orthonormal ist. Falls $m = 1$, dann ist der Satz trivial. Sei daher $m > 1$ und $a_1\alpha_1 + \cdots + a_m\alpha_m = 0$. Für jedes $i \in \{1, \ldots, m\}$ gilt dann

$$\begin{aligned} 0 = \langle \alpha_i, 0 \rangle &= \langle \alpha_i, a_1\alpha_1 + \cdots + a_m\alpha_m \rangle \\ &= a_1 \cdot 0 + \cdots + a_{i-1} \cdot 0 + a_i \cdot 1 + a_{i+1} \cdot 0 + \cdots + a_m \cdot 0 \\ &= a_i \end{aligned}$$

□

Satz 16.4 (Schmidt'sches Orthogonalisierungsverfahren). *Sei V ein endlich dimensionaler Euklidischer (bzw. unitärer) Vektorraum und sei $W \subset V$ ein linearer Unterraum. Gegeben eine Orthonormalbasis $\{\omega_1, \ldots, \omega_m\}$ für W, dann existiert auch eine Orthonormalbasis $\{v_1, \ldots, v_n\}$ für V mit $v_i = \omega_i$ für $i = 1, \ldots, m$.*

Beweis. Durch vollständige Induktion über die Zahl $n-m$. Nun, m ist die Dimension von W, n ist die Dimension von V. Falls $n - m = 0$ dann ist $m = n$ und $W = V$; wir sind schon fertig!

Sei daher $m < n$. Folglich ist $W \subset V$ (mit $W \neq V$), und es existiert ein Vektor $v \in V - W$. Der Vektor

$$\tilde{v} = \sum_{i=1}^{m} \overline{\langle v, \omega_i \rangle} \cdot \omega_i$$

heißt die *senkrechte Projektion* von v auf W. Insbesondere ist $\tilde{v} \in W$. Sei jetzt $\omega = v - \tilde{v}$. $\omega \neq 0$, da sonst $v \in W$ wäre. Folglich ist $\|\omega\| > 0$. Sei $\omega_{m+1} = \|\omega\|^{-1} \cdot \omega$. Dann ist

$$\|\omega_{m+1}\| = \|\|\omega\|^{-1} \cdot \omega\| = \|\omega\|^{-1} \cdot \|\omega\| = 1.$$

D.h. ω_{m+1} ist *normiert.* Aber für $i \leq m$ in $\mathbf{N}$ gilt auch

$$\begin{aligned} \langle \|\omega\| \cdot \omega_{m+1}, \omega_i \rangle &= \langle v - \tilde{v}, \omega_i \rangle \\ &= \langle v - \sum_{j=1}^{m} \overline{\langle v, \omega_j \rangle} \cdot \omega_j, \omega_i \rangle \\ &= \langle v, \omega_i \rangle - \sum_{j=1}^{m} \langle v, \omega_j \rangle \cdot \langle \omega_j, \omega_i \rangle \\ &= \langle v, \omega_i \rangle - \langle v, \omega_i \rangle \quad = \quad 0. \end{aligned}$$

Daher ist die Menge $\{\omega_1, \ldots, \omega_m, \omega_{m+1}\}$ auch orthogonal (und folglich linear unabhängig). Sei $W' \subset V$ der lineare Unterraum, aufgespannt durch die Vektoren $\{\omega_1, \ldots, \omega_m, \omega_{m+1}\}$. Wir haben $dim(W') = m + 1$, und die Menge $\{\omega_1, \ldots, \omega_m, \omega_{m+1}\}$ ist eine Orthonormalbasis für W'. D.h. die Bedingungen des Satzes gelten für W' und V, jedoch mit $n-(m+1)$ kleiner als $n-m$. Nach der induktiven Hypothese gibt es dann eine Orthonormalbasis für V, die unsere Menge $\{\omega_1, \ldots, \omega_m, \omega_{m+1}\}$ enthält. Daher auch $\{\omega_1, \ldots, \omega_m\}$. □

Korollar 16.4.1. *Jeder Euklidische (bzw. unitäre) Raum besitzt eine Orthonormalbasis. (Nehme einfach $W = \emptyset$.)*

Korollar 16.4.2. *Sei $W \subset V$ ein linearer Unterraum. (V Euklidisch oder unitär.) Dann gilt $V = W \oplus W^\perp$ (die orthogonale Summe).*

Beweis. Sei $\mathcal{B} = \{\omega_1, \ldots, \omega_m\}$ eine Orthonormalbasis für W. ($m = 0$, falls $W = \emptyset$.) Durch ein Basisergänzungsverfahren sei $\{\omega_1, \ldots, \omega_m, \omega_{m+1}, \ldots, \omega_n\}$ eine Orthonormalbasis für V. Wir müssen zeigen, daß $\mathcal{B}' = \{\omega_{m+1}, \ldots, \omega_n\}$ eine Orthonormalbasis für $W^\perp$ ist. Aber $\mathcal{B}'$ ist schon orthonormal. D.h. es genügt, zu zeigen, daß $\mathcal{B}'$ eine Basis für $W^\perp$ ist.

Dazu sei zunächst $\alpha = \sum_{j=m+1}^{n} s_j\omega_j$ eine beliebige Linearkombination der Vektoren in $\mathcal{B}'$. Wir können auch $\alpha = \sum_{j=1}^{n} s_j\omega_j$ schreiben, mit $s_j = 0$ für $j \leq m$. Sei nun $\beta = \sum_{k=1}^{m} t_k\omega_k \in W$ ein beliebiger Vektor in W. Wir können wiederum $\beta = \sum_{k=1}^{n} t_k\omega_k$ schreiben, mit $t_k = 0$ für $k > m$. Folglich ist

$$\langle \alpha, \beta \rangle = \sum_{j=1}^{n} \sum_{k=1}^{n} \bar{s}_j t_k \langle \omega_j, \omega_k \rangle = \sum_{j=1}^{n} \bar{s}_j t_j = 0.$$

Die vorletzte Gleichung folgt, da $\langle \omega_j, \omega_k \rangle = 0$, falls $j \neq k$, sonst ist $\langle \omega_j, \omega_j \rangle = 1$. Die letzte Gleichung folgt nach der Definition von den s_j und t_k. Folglich ist tatsächlich $\alpha \in W^\perp$.

Andererseits, sei $\alpha \in W^\perp$ beliebig. Da $\{\omega_1, \dots, \omega_n\}$ eine Basis für V ist, können wir α als Linearkombination $\alpha = \sum_{j=1}^n a_j \omega_j$ schreiben. Kann es sein, daß $a_{j_0} \neq 0$ für irgendein $j_0 \leq m$? Falls ja, dann wäre $\omega_{j_0} \in W$ und

$$\langle \omega_{j_0}, \alpha \rangle = \sum_{j=1}^n a_j \langle \omega_{j_0}, \omega_j \rangle = a_{j_0} \cdot 1 \neq 0.$$

D.h. $\alpha \notin W^\perp$; ein Widerspruch. $\square$

Beispiel für eine orthogonale Menge in einem unendlich-dimensionalen Raum

Sei $\mathcal{F}_{[-\pi,\pi)}$ die Menge der stetigen Abbildungen von dem halb-offenen, halb-abgeschlossenen Intervall $[-\pi, \pi)$ in die reellen Zahlen $\mathbf{R}$. Ein Skalarprodukt wird auf $\mathcal{F}_{[-\pi,\pi)}$ durch

$$\langle f, g \rangle \equiv \int_{-\pi}^{\pi} f(t)g(t)dt$$

für $f, g \in \mathcal{F}_{[-\pi,\pi)}$ definiert. Mit dieser Definition (und $f, g, f_1, f_2 \in \mathcal{F}_{[-\pi,\pi)}$, $x \in \mathbf{R}$) gilt:

1. $\langle xf, g \rangle = \int_{-\pi}^{\pi} (x \cdot f(t))g(t)dt = x \int_{-\pi}^{\pi} f(t)g(t)dt = x\langle f, g \rangle$.
   Auch $\langle f, xg \rangle = x\langle f, g \rangle$.
2. $\langle f_1 + f_2, g \rangle = \int_{-\pi}^{\pi} (f_1 + f_2)(t)g(t)dt$
   $= \int_{-\pi}^{\pi} f_1(t)g(t)dt + \int_{-\pi}^{\pi} f_2(t)g(t)dt = \langle f_1, g \rangle + \langle f_2, g \rangle$.
3. $\langle f, f \rangle = 0 \Leftrightarrow \int_{-\pi}^{\pi} f^2(t)dt = 0 \Leftrightarrow f^2(t) = 0, \forall t \in [-\pi, \pi) \Leftrightarrow f(t) = 0, \forall t$.

Sei jetzt $\mathcal{F}_\Theta$ die folgende Menge von 'cos' und 'sin' Funktionen: $\mathcal{F}_\Theta = \{\cos mt, \sin mt : m = 0, 1, 2, \dots\}$. Dann ist die Menge $\mathcal{F}_\Theta$ orthogonal, da $\int_{-\pi}^{\pi} f(t)g(t) = 0$, $\forall f \neq g$ in $\mathcal{F}_\Theta$. Warum? Fangen wir mit dem Fall $f(t) = \sin mt$, $g(t) = \cos nt$, (beliebige m, n) an. Nun, cos ist eine gerade Funktion, während sin ungerade ist. D.h. $\cos(\theta) = \cos(-\theta)$ und $\sin(\theta) = -\sin(-\theta)$. Folglich ist $\cos(mt)\sin(nt) = -\cos(-mt)\sin(-nt)$. D.h.

$$\int_{-\pi}^{0} \cos(mt)\sin(nt)dt = -\int_{0}^{\pi} \cos(mt)\sin(nt)dt$$

$$\Rightarrow \quad \int_{-\pi}^{\pi} \cos(mt)\sin(nt)dt = \int_{-\pi}^{0} \cos(mt)\sin(nt)dt + \int_{0}^{\pi} \cos(mt)\sin(nt)dt = 0.$$

Wie ist der Fall $f(t) = \sin mt$, $g(t) = \sin nt$? Es gilt im allgemeinen:

$$\sin\alpha \cdot \sin\beta = \frac{\cos(\alpha - \beta) - \cos(\alpha + \beta)}{2}.$$

Folglich ist

$$\int_{-\pi}^{\pi} \sin(mt)\sin(nt)dt =$$

$$= \frac{1}{2}\left\{\int_{-\pi}^{\pi} \cos((m-n)t)dt - \int_{-\pi}^{\pi} \cos((m+n)t)dt\right\} = 0 + 0 = 0.$$

(Die vorletzte Gleichung folgt, da cos eine ungerade Funktion ist.) Der Fall $f(t) = \cos mt$, $g(t) = \cos nt$ ist ähnlich.

Kapitel 17

Orthogonale und Unitäre Abbildungen

Definition 66. Sei V ein Euklidischer (bzw. ein unitärer) Vektorraum und sei $f : V \to V$ eine lineare Abbildung. f heißt *orthogonal* (bzw. *unitär*), falls

$$\langle f(\alpha), f(\beta)\rangle = \langle \alpha, \beta\rangle, \quad \forall \alpha, \beta \in V.$$

D.h. unter einer Euklidischen (unitären) Abbildung bleibt das Skalarprodukt—und folglich die 'Geometrie'—unverändert. Was sind Beispiele für solche Abbildungen? Nehmen wir zunächst den Euklidischen Fall. Sei $f : \mathbf{R}^2 \to \mathbf{R}^2$ eine orthogonale Abbildung. Nun, f muß doch eine 'starre' Abbildung sein, die die Längen von Vektoren und die Winkel zwischen Vektorpaaren nicht verändert.[1] Aber dafür gibt es nur die zwei Möglichkeiten:

- f ist eine 'Drehung' und/oder
- f ist eine 'Spiegelung'.

Beide Möglichkeiten haben wir schon behandelt. Wie ist es in höheren Dimensionen? Sei $f : \mathbf{R}^n \to \mathbf{R}^n$ wieder eine orthogonale Abbildung. Wie wir gleich sehen werden, gibt es auch hier eigentlich nur die zwei Möglichkeiten: Drehungen und/oder Spiegelungen. Aber wer kann sich schon eine Drehung oder Spiegelung in mehr als zwei Dimensionen vorstellen? Wir werden beweisen, daß jede mögliche orthogonale Abbildung in $\mathbf{R}^n$ ($n > 2$) als eine einfache Zusammensetzung von 2-dimensionalen Drehungen oder Spiegelungen verstanden werden kann.

Satz 17.1. *Sei V ein endlich-dimensionaler Euklidischer oder unitärer Vektorraum und sei $f : V \to V$ orthogonal, bzw. unitär. Dann ist f ein Automorphismus und f^{-1} ist auch orthogonal (bzw. unitär).*

Beweis. Für den ersten Teil brauchen wir nur zu beweisen, daß $ker(f) = \{0\}$. Sei daher $\alpha \in V$ mit $f(\alpha) = 0$. D.h.

$$0 = \langle 0, 0\rangle = \langle f(\alpha), f(\alpha)\rangle = \langle \alpha, \alpha\rangle \Rightarrow \alpha = 0.$$

Der letzte Schritt folgt, da $\langle \cdot, \cdot\rangle$ positiv definit ist.

Ist nun f^{-1} tatsächlich orthogonal, bzw. unitär? Seien $\alpha, \beta \in V$ beliebig. Dann sind $f^{-1}(\alpha)$ und $f^{-1}(\beta)$ eindeutige Vektoren in V. Da f schon orthogonal ist, gilt:

$$\langle f^{-1}(\alpha), f^{-1}(\beta)\rangle = \langle f(f^{-1}(\alpha)), f(f^{-1}(\beta))\rangle = \langle \alpha, \beta\rangle.$$

$\square$

Um festzustellen, ob eine vorgegebene Abbildung orthogonal ist, müssen wir, nach der Definition, alle möglichen Paare $\langle \alpha, \beta\rangle$ testen. Dies ist aber nicht nötig, wie das folgende Ergebnis zeigt.

Satz 17.2. *Sei die lineare Abbildung $f : V \to V$ so, daß*

$$\langle f(\alpha), f(\alpha)\rangle = \langle \alpha, \alpha\rangle,$$

für alle α in V. Dann ist f orthogonal (oder unitär).

Beweis. Seien β, γ beliebige Vektoren in V. Dann ist

$$\begin{aligned}
\langle \beta + \gamma, \beta + \gamma\rangle &= \langle \beta, \beta\rangle + \langle \beta, \gamma\rangle + \langle \gamma, \beta\rangle + \langle \gamma, \gamma\rangle \\
&= \langle f(\beta + \gamma), f(\beta + \gamma)\rangle \\
&= \langle f(\beta), f(\beta)\rangle + \langle f(\beta), f(\gamma)\rangle + \langle f(\gamma), f(\beta)\rangle + \langle f(\gamma), f(\gamma)\rangle.
\end{aligned}$$

[1] Dies folgt, da das innere Produkt nach der Abbildung f unverändert bleibt.

Aber es gilt doch $\langle f(\beta), f(\beta)\rangle = \langle \beta, \beta\rangle$, u.s.w. Folglich ist

$$\langle f(\beta), f(\gamma)\rangle + \langle f(\gamma), f(\beta)\rangle = \langle \beta, \gamma\rangle + \langle \gamma, \beta\rangle, \quad \text{oder}$$

$$\begin{aligned}
2re(\langle \beta, \gamma\rangle) &= \langle \beta, \gamma\rangle + \overline{\langle \beta, \gamma\rangle} = \langle \beta, \gamma\rangle + \langle \gamma, \beta\rangle \\
&= \langle f(\beta), f(\gamma)\rangle + \langle f(\gamma), f(\beta)\rangle \\
&= \langle f(\beta), f(\gamma)\rangle + \overline{\langle f(\beta), f(\gamma)\rangle} \\
&= 2re(\langle f(\beta), f(\gamma)\rangle).
\end{aligned}$$

Wir hätten aber auch mit $\alpha - \beta$ statt $\alpha + \beta$ anfangen können. Dann hätten wir die Gleichung $2im(\langle \beta, \gamma\rangle) = 2im(\langle f(\beta), f(\gamma)\rangle$ erhalten. D.h. die zwei Zahlen $\langle \beta, \gamma\rangle$ und $\langle f(\beta), f(\gamma)\rangle$ sind identisch, sowohl im reellen als auch im imaginären Teil. Folglich sind sie identisch als komplexe Zahlen. (Wir haben den Beweis im unitären Fall bewiesen. Der reelle Fall folgt natürlich als einfache Konsequenz.) □

Wie sieht es aus, wenn wir die entsprechenden Matrizen untersuchen? D.h. gegeben eine orthogonale (unitäre) Abbildung $f : V \to V$, dann entspricht dieser Abbildung eine Matrix $A \in GL(n; F)$[2] ($F = \mathbf{R}$ oder $\mathbf{C}$).

Definition 67. Angenommen $\overline{A}^t = A^{-1}$, dann heißt A eine *orthogonale* (bzw. im komplexen Fall *unitäre*) Matrix.

Was können wir über solche Matrizen sagen? Es ist einfach zu zeigen, daß $|det(A)| = 1$. Denn $\overline{A}^t = A^{-1} \Rightarrow \overline{A}^t \cdot A = I$, die Identitätsmatrix.

$$\begin{aligned}
\Rightarrow \quad 1 = det(I) &= det(\overline{A}^t \cdot A) \\
&= det(\overline{A}^t) \cdot det(A) \\
&= det(\overline{A}) \cdot det(A) \\
&= \overline{det(A)} \cdot det(A) \\
&= |det(A)|^2 \\
\Rightarrow \quad |det(A)| &= 1
\end{aligned}$$

Definition 68. Die Menge der orthogonalen $n \times n$ Matrizen heißt $O(n)$. Die Menge der unitären $n \times n$ Matrizen heißt $U(n)$. Für orthogonale Matrizen A gibt es nur zwei Möglichkeiten: $det(A) = \pm 1$. Die Teilmenge der Matrizen mit $det(A) = +1$ heißt $SO(n) \subset O(n)$ (die 'speziellen' orthogonalen Matrizen). Analog ist $SU(n)$ die Menge der komplexen Matrizen $A \in U(n)$ mit $det(A) = +1$.

Bemerkung. Alle Mengen ($O(n)$, $SO(n)$, $U(n)$ und $SU(n)$) sind Gruppen unter Matrizenmultiplikation.

Als Beispiel nehmen wir die Matrix $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Dann ist $A^{-1} = A = A^t = \overline{A}^t$. D.h. A ist sicherlich orthogonal. Aber ist $A \in SO(2)$? Nein! da $det(A) = -1$. Wir wissen schon, daß alle orthogonalen Abbildungen eigentlich nur Drehungen oder Spiegelungen sind. Nun, es ist nicht schwer einzusehen, daß Drehungen immer in $SO(n)$ sind. Folglich muß A eine Spiegelung sein.

Satz 17.3. *Sei $f : V \to V$ eine lineare Abbildung; $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$ sei eine Orthonormalbasis für V. Dann ist f orthogonal (unitär) $\Leftrightarrow$ $M_{\mathcal{B}}(f)$ orthogonal (unitär). (wobei $M_{\mathcal{B}}(f)$ die Matrix ist, die f darstellt bzgl. $\mathcal{B}$.)*

Beweis. (Für den unitären Fall) Es gilt: f unitär $\Leftrightarrow \langle f(\beta), f(\kappa)\rangle = \langle \beta, \kappa\rangle$, für alle β, $\kappa \in V$. Sei $\beta = b_1\alpha_1 + \cdots + b_n\alpha_n$ und $\kappa = c_1\alpha_1 + \cdots + c_n\alpha_n$. Dann ist

$$\langle \beta, \kappa\rangle = \sum_{i=1}^{n} \overline{b}_i \cdot c_i.$$

Wenn wir β und κ als Spaltenvektoren darstellen: $\beta = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ und $\kappa = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$, dann ist $\langle \beta, \kappa\rangle = \beta^t \cdot \kappa$. (Dies

[2] A muß invertierbar sein, da f ein Isomorphismus ist.

folgt, da $\mathcal{B}$ eine Orthonormalbasis ist.) Aber auch

$$\begin{aligned}
\langle f(\beta), f(\kappa)\rangle &= \overline{f(\beta)}^t \cdot f(\kappa) \\
&= \left(\overline{M_{\mathcal{B}}(f)\cdot\beta}\right)^t \cdot (M_{\mathcal{B}}(f)\cdot\kappa) \\
&= \left(\overline{M_{\mathcal{B}}(f)}\cdot\overline{\beta}\right)^t \cdot (M_{\mathcal{B}}(f)\cdot\kappa) \\
&= \overline{\beta}^t \cdot \left(\overline{M_{\mathcal{B}}(f)}^t \cdot M_{\mathcal{B}}(f)\right)\cdot\kappa
\end{aligned}$$

□

Um diesen Beweis zu Ende zu führen, brauchen wir nun das folgende Ergebnis.

Lemma. *Sei $M = (m_{ij})$ eine beliebige $n \times n$ Matrix. Angenommen, $\beta^t \cdot \kappa = \beta^t \cdot M \cdot \kappa$, für alle möglichen $n \times 1$ Spaltenmatrizen β und κ, dann ist $M = I$, die Identitätsmatrix.*

Beweis. Sei $i \in \{1, \ldots, n\}$ und wähle $\beta = \kappa = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow i$-te Stelle. Dann ist $\beta^t \cdot \kappa = 1 = m_{ii}$. D.h. die Hauptdiagonale von M besteht aus lauter 1-sen. Wir nehmen jetzt $i \neq j$ und

$\beta = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow i$-te Stelle, während $\kappa = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow j$-te Stelle. Dann ist $\beta^t \cdot \kappa = 0 = m_{ij}$. D.h. alle Matrizenelemente, die nicht auf der Hauptdiagonalen liegen, sind 0. □

Im Beweis des Satzes ist daher $\overline{M_{\mathcal{B}}(f)}^t \cdot M_{\mathcal{B}}(f) = I$ genau dann, wenn $\langle\beta, \kappa\rangle = \langle f(\beta), f(\kappa)\rangle$, und dies ist genau dann der Fall, wenn $M_{\mathcal{B}}(f)$ unitär ist.

Obwohl die komplexen Zahlen komplizierter sind als die reellen Zahlen, sind unitäre Abbildungen und Matrizen doch sehr viel einfacher als die entsprechenden orthogonalen Abbildungen und Matrizen. Es ist nämlich so, daß für jede unitäre Abbildung $f : V \to V$ eine Basis $\mathcal{B}$ für V existiert, bestehend aus lauter Eigenvektoren von f. D.h. (wenn wir uns an die Theorie vom letzten Semester erinnern) jede unitäre Matrix ist diagonalisierbar. Dies ist leider nicht der Fall bei orthogonalen Abbildungen. Z.B. eine Drehung des 2-dimensionalen Raumes $\mathbf{R}^2$ in sich um den Winkel Θ hat *keine* Eigenvektoren, es sei denn, $\Theta = 0 \pmod{\pi}$. Wir werden diese eher komplizierten orthogonalen Abbildungen etwas später behandeln. Zunächst aber den einfachen unitären Fall.

Satz 17.4. *Seien V ein unitärer Vektorraum und $f : V \to V$ eine unitäre Abbildung. Dann existiert eine Basis $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ für V, bestehend aus Eigenvektoren von f. (D.h. $f(\alpha_i) = \lambda_i\alpha_i$, wobei $\lambda_1, \ldots, \lambda_n \in \mathbf{C}$.)*

Beweis. Sei $dim(V) = n$. Wir benutzen vollständige Induktion über die Zahl n. Falls $n = 1$, dann ist alles ganz trivial. Wir können daher annehmen, daß $n > 1$ und daß der Satz stimmt für Vektorräume mit niedrigeren Dimensionen.

Sei $P_f(z) \in \mathbf{C}[z]$ das charakteristische Polynom von f. Nach der Theorie des letzten Semesters ist $\lambda \in \mathbf{C}$ ein Eigenwert für f genau dann, wenn $P_f(\lambda) = 0$. Aber nach dem Fundamentalsatz der Algebra muß doch ein Eigenwert, etwa λ_1 existieren. (Es gilt $grad(P_f) = n > 1$.) Zu diesem Eigenwert gibt es einen entsprechenden Eigenvektor, etwa $\alpha_1 \in V$. D.h. $\alpha_1 = 0$ und $f(\alpha_1) = \lambda_1\alpha_1$. Falls $\|\alpha_1\| \neq 1$, dann können wir $\frac{\alpha_1}{\|\alpha_1\|}$ nehmen, und es gilt

$$f\left(\frac{\alpha_1}{\|\alpha_1\|}\right) = \frac{1}{\|\alpha_1\|} f(\alpha_1) = \lambda_1 \frac{\alpha_1}{\|\alpha_1\|}.$$

D.h. $\frac{\alpha_1}{\|\alpha_1\|}$ ist auch ein Eigenvektor mit Eigenwert λ_1. Wir werden daher einfach annehmen, daß $\|\alpha_1\| = 1$.

Nun, es gilt $\lambda_1 \neq 0$, da f sonst kein Isomorphismus wäre. Aber wir können noch viel mehr sagen.

$$\begin{aligned}
\|\alpha_1\| &= \sqrt{\langle \alpha_1, \alpha_1 \rangle} \\
&= \sqrt{\langle f(\alpha_1), f(\alpha_1) \rangle} \\
&= \sqrt{\langle \lambda_1 \alpha_1, \lambda_1 \alpha_1 \rangle} \\
&= \sqrt{|\lambda_1^2| \cdot \|\alpha_1\|} \quad = \quad |\lambda_1| \cdot \|\alpha_1\| \\
\Rightarrow \quad |\lambda_1| &= 1
\end{aligned}$$

Sei nun $W = \{\alpha_1\}^\perp = \{\gamma \in V : \langle \gamma, \alpha_1 \rangle = 0\}$. Sei $\gamma \in W$ beliebig. Dann ist auch

$$\begin{aligned}
\lambda_1 \langle f(\gamma), \alpha_1 \rangle &= \langle f(\gamma), \lambda_1 \alpha_1 \rangle \quad = \quad \langle f(\gamma), f(\alpha_1) \rangle \\
&= \langle \gamma, \alpha_1 \rangle \quad = \quad 0
\end{aligned}$$

Folglich ist $f(\gamma) \in W$; d.h. $f(W) \subset W$.
Behauptung: $W^\perp$ wird durch den Vektor $\{\alpha_1\}$ erzeugt (d.h. $W^\perp$ ist ein 1-dimensionaler Unterraum von V). Warum? Sei etwa $\mathcal{B}^* = \{\alpha_1, \alpha_2^*, \ldots, \alpha_n^*\}$ eine Orthonormalbasis für V, beginnend mit dem vorhandenen Eigenvektor α_1. (Siehe Satz 16.4: das Schmidt'sche Orthonormierungsverfahren.) Sei $\beta \in W^\perp$ beliebig, etwa $\beta = b_1\alpha_1 + \sum_{i=2}^n b_i \alpha_i^*$. Kann es sein, daß ein $b_i \neq 0$ für $i > 1$? Falls etwa $b_{i_0} \neq 0$, $i_0 > 1$, dann ist sicherlich $\langle \alpha_1, \alpha_{i_0} \rangle = 0$, so daß $\alpha_{i_0} \in W$. Aber $\langle \beta, \alpha_{i_0} \rangle = b_{i_0} \neq 0$, und wir haben $\beta \notin W^\perp$. D.h. unser beliebiger Vektor $\beta \in W^\perp$ ist einfach $\beta = b_1\alpha_1$, und folglich ist $dim(W^\perp) = 1$.

Nach dem zweiten Korollar zu Satz 16.4 ist dann $V = W \oplus W^\perp$ und $dim(W) = n - 1$. Aber $f|_W : W \to W$ (die Einschränkung von f auf die Teilmenge $W \subset V$) ist eine unitäre Abbildung, und nach der induktiven Hypothese muß eine Orthonormalbasis, etwa $\{\alpha_2, \ldots, \alpha_n\}$, für W existieren, bestehend aus lauter Eigenvektoren bzgl. $f|_W$, d.h. bzgl. f. Wir brauchen jetzt nur den ersten Eigenvektor hinzuzunehmen, und wir erhalten die Orthonormalbasis $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ für V. $\square$

Korollar 17.4.1. *Jeder unitäre Matrix A ist zu einer diagonalen Matrix ähnlich. D.h. $\exists$ eine invertierbare Matrix $S \in GL(n; \mathbf{C})$ mit $S^{-1} \cdot A \cdot S$ diagonal. Die Diagonalelemente sind die Eigenwerte zu einer unitären Abbildung, die A darstellt.*

Kapitel 18

Diagonalisierung und Trigonalisierung

Alle unitären Matrizen sind diagonalisierbar. D.h. sie sind ähnlich zu diagonalen Matrizen. Es liegt (zum Teil) daran, daß die komplexen Zahlen algebraisch abgeschlossen sind. In diesem Kapitel wollen wir die Sache etwas genauer betrachten, und einen Satz beweisen über die genauen Bedingungen, die nötig sind, um diese Diagonalisierbarkeit von Matrizen zu garantieren. Für diese Untersuchung brauchen wir nicht vorauszusetzen, daß unsere Vektorräume Skalarprodukte haben. Im Gegenteil, wir nehmen irgendeinen endlich dimensionalen Vektorraum V über einem Körper F. Sei $f : V \to V$ eine lineare Abbildung mit darstellender Matrix $M_{\mathcal{A}} \in M(n \times n; F)$ bzgl. einer Basis $\mathcal{A} = \{\alpha_1, \ldots, \alpha_n\}$. Die Frage ist dann, wann gibt es eine Matrix $S \in GL(n; F)$ mit $S^{-1} \cdot M_{\mathcal{A}} \cdot S$ diagonal?

Nach unserer früheren Theorie können wir auf jeden Fall sagen, daß, falls A diagonalisierbar ist, eine Basis $\mathcal{A}' = \{\beta_1, \ldots, \beta_n\}$ existiert, bestehend aus Eigenvektoren zu f. Die Matrix S ist die Darstellung des Basiswechsels $\mathcal{A} \to \mathcal{A}'$.

Allgemeiner, sei $\lambda \in F$ ein Eigenwert zu f. Dann ist

$$Eig(f; \lambda) = \{\beta \in V : f(\beta) = \lambda\beta\}$$

der entsprechende Eigenraum. Wir wissen, daß $Eig(f; \lambda) \subset V$ ein Unterraum von V ist. Andererseits, da λ ein Eigenwert ist, gilt $P_f(\lambda) = 0$, wobei P_f das charakteristische Polynom ist. λ ist eine Nullstelle von P_f. Sei daher $n_\lambda \in \mathbf{N}$ die größte Zahl mit $P_f(x) = (x - \lambda)^{n_\lambda} \cdot Q(x)$, wobei $Q(x) \in F[x]$.

Definition 69. n_λ heißt die *algebraische Vielfachheit* und $dim(Eig(f; \lambda))$ heißt die *geometrische Vielfachheit* des Eigenwertes λ.

Satz 18.1. $n_\lambda \geq dim(Eig(f; \lambda))$.

Beweis. Sei $Eig(f; \lambda) = W \subset V$ und sei $\{\beta_1, \ldots, \beta_m\}$ eine Basis für W. Jedes β_i ist dann ein Eigenvektor für f mit Eigenwert λ. Nun sei (mittels Basisergänzungssatz) $\mathcal{B} = \{\beta_1, \ldots, \beta_m, \beta_{m+1}, \ldots, \beta_n\}$ eine Basis für V. Die Matrix, die f darstellt bzgl. dieser Basis, ist

$$M_{\mathcal{B}} = \left(\begin{array}{ccc|c} \lambda & & 0 & \\ & \ddots & & X \\ 0 & & \lambda & \\ \hline & 0 & & A \end{array}\right),$$

wobei der linke obere Block eine $n_\lambda \times n_\lambda$ Untermatrix ist. Wir haben schon gesehen, daß das charakteristische Polynom für solche Matrizen die Gestalt

$$P_f(x) = (\lambda - x)^{n_\lambda} \cdot det(A - I_{n-n\lambda})$$

hat. Nun, $M_{\mathcal{B}}$ ist ähnlich wie $M_{\mathcal{A}}$; folglich sind die charakteristischen Polynome identisch. □

Satz 18.2. *Es gilt*

1. $f : V \to V$ *ist diagonalisierbar* $\Leftrightarrow$
2. *Das charakteristische Polynom zerfällt in lineare Faktoren und* $n_\lambda = dim(Eig(f; \lambda))$ *für alle Eigenwerte* $\lambda \Leftrightarrow$
3. $V = Eig(f; \lambda_1) \oplus \cdots Eig(f; \lambda_m)$*, wobei* $\lambda_1, \ldots, \lambda_m$ *die verschiedene Eigenwerte sind.*

Beweis. Wir werden die Kette $(i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (i)$ beweisen.
Sei f diagonalisierbar. $\Rightarrow \exists$ eine Basis $\mathcal{B} = \{\beta_1, \ldots, \beta_n\}$ bestehend aus Eigenvektoren, mit

$$M_{\mathcal{B}}(f) = \begin{pmatrix} \lambda_1 & & 0 & & & & 0 \\ & \ddots & & & & & \\ 0 & & \lambda_1 & & & & \\ & & & \ddots & & & \\ & & & & \lambda_m & & 0 \\ & & & & & \ddots & \\ 0 & & & & 0 & & \lambda_m \end{pmatrix}$$

$\Rightarrow$ das charakteristische Polynom für f hat die Gestalt

$$P_f(x) = (\lambda_1 - x)^{n_{\lambda_1}} \cdots (\lambda_m - x)^{n_{\lambda_m}}$$

wobei die algebraischen und geometrischen Vielfachheiten für λ_i identisch sind, für alle $i \Rightarrow V = Eig(f;\lambda_1) \oplus \cdots Eig(f;\lambda_m)$, da $n_{\lambda_1} + \cdots + n_{\lambda_m} = n = dim(V) \Rightarrow f$ ist diagonalisierbar, und zwar erhält man eine Basis aus Eigenvektoren, wenn man jeweils Basen zu den einzelnen Eigenräumen $Eig(f;\lambda_i)$ wählt. $\square$

Beispiel: Sei $f : \mathbf{R}^2 \to \mathbf{R}^2$ eine Drehung des 2-dimensionalen Raumes mit darstellender Matrix

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

bzgl. der kanonischen Basis, wobei $\theta \neq 0 \bmod \pi$. Dann ist das charakteristische Polynom

$$P_f(x) = (\cos\theta - x)^2 + (\sin\theta)^2 = (\cos^2\theta + \sin^2\theta) - 2x\cos\theta + x^2$$
$$= x^2 - 2x\cos\theta + 1,$$

wobei $\cos^2\theta < 1$. Nach der bekannten Formel

$$x = \cos\theta \pm \sqrt{\cos^2\theta - 1}$$

hat dieses Polynom offensichtlich *keine* Nullstellen in $\mathbf{R}$; folglich ist die Drehung nicht diagonalisierbar.

Es ist jetzt klar, daß viele Matrizen nicht diagonalisierbar sind. Man fragt sich dann, ob wenigstens Trigonalisierbarkeit erreichbar ist. D.h. gegeben eine $n \times n$ Matrix A, gibt es eine invertierbare Matrix $S \in GL(n;F)$ mit $S^{-1} \cdot A \cdot S$ in der oberen Dreiecksform? Aber sogar die Trigonalisierbarkeit ist nur unter speziellen Umständen gewährleistet. (Und orthogonale Matrizen sind im allgemeinen auch *nicht* trigonalisierbar!) Trotzdem ist die Frage nach der Trigonalisierbarkeit eine sehr vernünftige Frage, wegen der folgenden einfachen algebraischen Aussage.

Satz 18.3. *Eine Matrix $A \in M(n \times n; F)$ ist genau dann trigonalisierbar, wenn das entsprechende charakteristische Polynom in lineare Faktoren zerfällt:*

$$det(A - xI) = (\lambda_1 - x) \cdots (\lambda_n - x).$$

Einfache Folgerung: Nach dem Fundamentalsatz der Algebra ist jede quadratische Matrix über $\mathbf{C}$ trigonalisierbar. Es gilt aber viel mehr. Wir werden später sehen, daß jede solche Matrix zu einer Matrix in der besonders einfachen 'Jordan Normalform' ähnlich ist.

Beweis. (des Satzes 18.3) Angenommen, A sei trigonalisierbar. Folglich hat A dasselbe charakteristische Polynom wie eine Dreiecksmatrix, etwa D. Dann ist auch $D - xI$ eine Dreiecksmatrix und $det(D - xI)$ ist das Produkt der Elemente entlang der Hauptdiagonalen. D.h. das charakteristische Polynom ist ein Produkt von linearen Faktoren der Art $(d_{ii} - x)$.

Umgekehrt, sei angenommen, daß das charakterische Polynom $det(A - xI)$ in lineare Faktoren zerfällt, etwa

$$det(A - xI) = (\lambda_1 - x) \cdots (\lambda_n - x).$$

Wir benutzen Induktion über die Zahl n. Falls $n = 1$, dann ist alles trivial. Sei daher $n > 1$. Wir wissen, daß λ_1 ein Eigenwert zur zugrundeliegenden linearen Abbildung $f : V \to V$ ist (wobei V ein n-dimensionaler Vektorraum über F ist). Sei α_1 ein nicht-trivialer Vektor im Eigenraum $Eig(\lambda_1; f)$. Sei $\mathcal{B} = \{\alpha_1, \beta_2, \ldots, \beta_n\}$

eine Basis für V, die α_1 enthält (Basisergänzungssatz). Wie sieht die Matrix aus, die die Abbildung f bzgl. der Basis $\mathcal{B}$ darstellt? Es gilt

$$M_{\mathcal{B}}(f) = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & b_{11} & \cdots & b_{1\,(n-1)} \\ \vdots & & \ddots & \\ 0 & b_{(n-1)\,1} & \cdots & b_{(n-1)\,(n-1)} \end{pmatrix},$$

wobei $B = (b_{ij})$ eine $(n-1)\times(n-1)$ Matrix über F ist, und '$*$' ist vielleicht 0, aber vielleicht auch nicht 0. Warum hat $M_{\mathcal{B}}(f)$ diese Gestalt? Einfach weil die Spaltenvektoren aus den Koordinaten der Vektoren $f(\alpha_1)$ bzw. $f(\beta_i)$ $(i > 1)$ in der Basis $\mathcal{B}$ gebildet sind. Nun, wir wissen eigentlich nichts über die Koordinaten von $f(\beta_i)$—daher sind alle Spaltenvektoren in $M_{\mathcal{B}}(f)$ außer den ersten ziemlich beliebig. Aber zumindest hat der erste Spaltenvektor eine besondere Form; nämlich alle Koordinaten sind Null außer den ersten. Was ist $det(M_{\mathcal{B}}(f)-xI)$ für Matrizen mit solchen ersten Spaltenvektoren? Nach unserer Theorie ist das charakteristische Polynom $(\lambda_1-x)Q(x)$, wobei $Q(x)$ das charakteristische Polynom zur Matrix B ist. Aber $Q(x)$ läßt sich auch in Linearfaktoren zerlegen. Eine einfache Induktion bringt uns dann zum Schluß, und zwar sei $V_* \subset V$ der Unterraum, erzeugt durch die Vektoren $\{\beta_2, \ldots, \beta_n\}$, und sei $f_* : V_* \to V_*$ die folgende lineare Abbildung. Sei $\beta = \sum_{i=2}^n b_i\beta_i$ ein beliebiger Vektor in V_*. Aber auch $\beta \in V$. Was ist $f(\beta)$? Es gibt irgendeine Darstellung in unserer Basis $\mathcal{B}$, etwa $f(\beta) = a_1\alpha_1 + \sum_{i=2}^n c_i\beta_i$. Dann wird $f_*(\beta)$ definiert als $f(\beta) = \sum_{i=2}^n c_i\beta_i$. Offensichtlich ist die darstellende Matrix $M_{\mathcal{B}_*}(f_*) = B$, wobei $\mathcal{B}_* = \{\beta_2, \ldots, \beta_n\}$. Nach der induktiven Hyphothese existiert eine neue Basis $\mathcal{A}_* = \{\alpha_2, \ldots, \alpha_n\}$ mit $M_{\mathcal{A}_*}(f_*)$ in der oberen Dreiecksform. Dann ist sicherlich $M_{\mathcal{A}}(f)$ in der oberen Dreiecksform, wobei $\mathcal{A} = \{\alpha_1, \ldots, \alpha_n\}$. $\square$

Die Basis $\mathcal{A}$ dieses Beweises ist im allgemeinen *keine* Basis von Eigenvektoren. Trotzdem ist die Gestalt von $\mathcal{A}$ von einer ganz besonderen Art, und wir machen daher eine Definition dazu.

Definition 70. Sei $V_0 \subset V_1 \subset \cdots \subset V_{n-1} \subset V_n = V$ eine Kette von Unterräumen mit $dim(V_i) = i$ für jedes i. Jede solche Kette heißt eine *Fahne*. Sei $f : V \to V$ eine lineare Abbildung. Die Fahne heißt f-invariant, falls $f(V_i) \subset V_i$, für alle i.

Offensichtlich ist f trigonalisierbar genau dann, wenn es eine f-invariante Fahne gibt.

Kapitel 19

Mehr über Orthogonale Abbildungen

Im vorletzten Kapitel haben wir gesehen, daß jede unitäre Abbildung $f : V \to V$ 'diagonalisierbar' ist; d.h. es existiert eine Basis $\mathcal{B}$ für V, bestehend aus lauter Eigenvektoren unter f.[1] Dies gilt aber *nicht* für orthogonale Abbildungen. Eine orthogonale Abbildung ist eine Abbildung des reellen Raumes in sich selbst, die die 'Geometrie' (d.h. die Längen und Winkel zwischen Vektoren) unverändert läßt. Unsere normale Anschauung lehrt uns, daß die einfachen Drehungen des 2-dimensionalen Raumes $\mathbf{R}^2$ im allgemeinen keine Eigenvektoren besitzen. Wir können trotzdem sehr viel über die orthogonalen Abbildungen sagen. Als erstes brauchen wir noch ein weiteres Ergebnis über die Algebra der Polynome.

Satz 19.1. *Sei $P(x) \in \mathbf{R}[x]$ ein Polynom mit reellen Koeffizienten. Dann ist $P(x) = \lambda_0 \cdot (\lambda_1 - x) \cdots (\lambda_s - x) \cdot Q_1(x) \cdots Q_t(x)$, wobei $\lambda_i \in \mathbf{R}$ für alle $0 \le i \le s$ und $Q_j(x) = x^2 + a_j x + b_j$ für alle $1 \le j \le t$, wobei a_j und $b_j \in \mathbf{R}$.*

Um diesen Satz zu beweisen, brauchen wir zuerst einige andere Ergebnisse.

Lemma (A). *Sei $\lambda \in \mathbf{C}$ eine komplexe Zahl, die eine Nullstelle des Polynoms $P(x) \in \mathbf{R}[x]$ ist. (D.h. $P(\lambda) = 0$.) Dann ist auch $\overline{\lambda}$ eine Nullstelle des Polynoms.*

Beweis. Sei $P(x) = \sum_{i=0}^{n} a_i x^i$. Da λ eine Nullstelle ist, gilt $\sum_{i=0}^{n} a_i \lambda^i = 0$. Aber die Zahl Null ist eine reelle Zahl. Daher ist

$$0 = \overline{0} = \overline{P(\lambda)} = \sum_{i=0}^{n} \overline{a_i} \cdot \overline{\lambda^i} = \sum_{i=0}^{n} a_i \cdot \overline{\lambda}^i = P(\overline{\lambda}).$$

Dies folgt, da $\overline{a_i} = a_i$, für $a_i \in \mathbf{R}$, und $\overline{\lambda^i} = \overline{\lambda}^i$ gilt für alle komplexen Zahlen. □

Lemma (B). *Sei $\lambda \in \mathbf{C} - \mathbf{R}$ eine Nullstelle von $P(x) \in \mathbf{R}[x]$. (D.h. λ ist definitiv* nicht *reell.) Dann ist $P(x) = Q(x) \cdot T(x)$, wobei $Q(x) = x^2 + ax + b \in \mathbf{R}[x]$ und $T(x) \in \mathbf{R}[x]$ mit grad$(T) =$ grad$(P) - 2$.*

Beweis. Da sowohl λ als auch $\overline{\lambda}$ Nullstellen sind und $\lambda \ne \overline{\lambda}$, gilt $P(x) = (\lambda - x) \cdot (\overline{\lambda} - x) \cdot T(x)$. (siehe Kapitel 13 über den Fundamentalsatz der Algebra) Es genügt daher zu zeigen, daß

$$(\lambda - x) \cdot (\overline{\lambda} - x) = x^2 + ax + b$$

mit a, $b \in \mathbf{R}$. Sei nun etwa $\lambda = c + di$, mit c, $d \in \mathbf{R}$. Es gilt

$$(x - \lambda) \cdot (x - \overline{\lambda}) = (x - c - di)(x - c + di) = x^2 - 2cx + (c^2 + d^2).$$

Setze daher $a = -2c$ und $b = c^2 + d^2$. □

Beweis. (Satz 19.1) Folgt durch Induktion über *grad*(P). Falls *grad*$(P) \le 1$, dann ist die Aussage trivial. Da nun $P \in \mathbf{R}[x] \subset \mathbf{C}[x]$, gilt der Fundamentalsatz der Algebra, und es existiert eine Nullstelle $\lambda \in \mathbf{C}$. Falls $\lambda \notin \mathbf{R}$, dann ist, nach Lemma B, $P(x) = Q(x) \cdot T(x)$, wobei $Q(x) = x^2 + ax + b \in \mathbf{R}[x]$, und wir brauchen daher die induktive Hypothese nur auf $T(x)$ anzuwenden. Falls $\lambda \in \mathbf{R}$, dann ist $P(x) = (\lambda - x) \cdot T(x)$, wobei $T(x) \in \mathbf{R}[x]$ und *grad*$(T) =$ *grad*$(P) - 1$. Eine einfache Induktion bringt uns wieder zum Schluß. □

---

[1] Ich habe dort die Tatsache benutzt, daß $det(A) = det(B) \times det(C)$, falls $A \in M(n \times n; F)$ ist, wobei $A = \begin{pmatrix} B & X \\ 0 & C \end{pmatrix}$, für $B \in M(r \times r; F)$, $C \in M(s \times s; F)$, $n = r + s$ und $X \in M(r \times s; F)$ beliebig. Dieses relativ triviale Ergebnis habe ich im Skript nicht explizit bewiesen. Dafür braucht man aber nur zu bemerken, daß sowohl B als auch C durch elementare Spaltenoperationen in Dreiecksmatrizen umgeformt werden können. Danach werden die Determinanten bestimmt durch Multiplikation der Elemente entlang der Hauptdiagonalen. Aber die besondere Form der $(r+s) \times (r+s)$ Matrix A erlaubt eine Umformung in eine Dreiecksmatrix mittels der genau entsprechenden $(r + s)$-Spaltenoperationen.

19.1 Der Satz von Cayley-Hamilton

Um weiter zu kommen, müssen wir noch ein weiteres Standardergebnis behandeln.

Satz 19.2. *Sei V ein (endlich-dimensionaler) Vektorraum über F, wobei $F = \mathbf{R}$ oder $\mathbf{C}$. Sei $f : V \to V$ irgendeine lineare Abbildung und sei $P_f(x)$ das entsprechende charakteristische Polynom. Dann ist $P_f(f) : V \to V$ die triviale Abbildung, die alle Vektoren auf den Nullvektor abbildet. D.h. (sozusagen) $P_f(f) = 0$.*

Beispiel $f : \mathbf{R}^2 \to \mathbf{R}^2$ gegeben durch $f((x_1, x_2)) = (x_1 + x_2, x_2)$, für alle $(x_1, x_2) \in \mathbf{R}^2$. Dann ist bekannterweise die Matrix $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ die darstellende Matrix für diese Abbildung. Folglich ist $P_f(x) = (1 - x)^2 = x^2 - 2x + 1$. Was ist dann $P_f(f)$? Nichts anderes als die lineare Abbildung $f^2 - 2f + id : V \to V$ (Die lineare Abbildung $id : V \to V$ ist die Identitätsabbildung.) Aber

$$\begin{aligned}
&(f^2 - 2f + id)(x_1, x_2) \\
&\quad = f^2((x_1, x_2)) - 2f((x_1, x_2)) + id(x_1, x_2) \\
&\quad = f((x_1 + x_2, x_2)) - 2(x_1 + x_2, x_2) + (x_1, x_2) \\
&\quad = ((x_1 + x_2) + x_2, x_2) - (2(x_1 + x_2), 2x_2) + (x_1, x_2) \\
&\quad = ((x_1 + x_2 + x_2 - 2(x_1 + x_2) + x_1, x_2 - 2x_2 + x_2) \\
&\quad = (0, 0) = 0 \in \mathbf{R}^2,
\end{aligned}$$

für beliebige $(x_1, x_2) \in \mathbf{R}^2$. Daher stimmt der Satz von Cayley-Hamilton, zumindest in diesem Beispiel.

Beweis. (Satz 19.2) Falls $F = \mathbf{C}$, dann ist $P_f(x) \in \mathbf{C}[x]$, und nach dem Fundamentalsatz der Algebra ist $P_f(x) = (\lambda_1 - x) \cdots (\lambda_n - x)$, wobei $\lambda_1, \ldots, \lambda_n \in \mathbf{C}$. Wir benutzen Induktion über die Zahl n. Falls $n = 1$, dann ist alles klar (jeder Vektor $\alpha \in V$ ist dann ein Eigenvektor ($f(\alpha) = \lambda_1 \alpha$), und der Satz von Cayley-Hamilton ist einfach die Aussage $P_f(f)(\alpha) = (\lambda_1 - f)(\alpha) = \lambda_1 \alpha - f(\alpha) = 0$.) Falls $n > 1$, dann ist die Sache etwas komplizierter. Nach unserem Satz über die Trigonalisierung muß eine Basis $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ für V existieren, mit der Eigenschaft, daß $M_{\mathcal{B}}(f)$ (die Matrix, die f darstellt, mittels $\mathcal{B}$) eine Dreiecksmatrix ist. Sei nun $\beta = \sum_{i=1}^n b_i \alpha_i \in V$ beliebig. Wir müssen zeigen, daß

$$P_f(f)(\beta) = P_f(f)(\sum_{i=1}^n b_i \alpha_i) = \sum_{i=1}^n b_i P_f(f)(\alpha_i) = 0.$$

(Die zweite Gleichung folgt, da $P_f(f)$ eine *lineare* Abbildung ist.) Offensichtlich genügt es, zu zeigen, daß $P_f(f)(\alpha_i) = 0$, für alle i.

Für jedes i sei daher $V_i \subset V$ der Unterraum, der erzeugt wird durch die Vektoren $\{\alpha_1, \ldots, \alpha_i\}$. Wenn wir die Matrix $M_{\mathcal{B}}(f)$ anschauen, dann sehen wir sofort, daß insbesondere $f(V_{n-1}) \subset V_{n-1}$. (Allgemeiner gilt: $V_1 \subset \cdots \subset V_n = V$ ist eine f-invariante Fahne.) D.h. $f_{|V_{n-1}} : V_{n-1} \to V_{n-1}$ ist eine lineare Abbildung, und nach der induktiven Hypothese ist $P_{f_{|V_{n-1}}}(f_{|V_{n-1}})(\gamma) = 0$ und folglich

$$P_f(f)(\gamma) = (\lambda_n - f) \underbrace{\cdots (\lambda_{i+1} - f)(\lambda_i - f) \cdots (\lambda_1 - f)}_{P_{f_{|V_{n-1}}}}(\gamma),$$

für alle $\gamma \in V_{n-1}$. Da alle $\alpha_j \in V_{n-1}$, für $j < n$, brauchen wir nur zu zeigen, daß $P_f(f)(\alpha_n) = 0$. Aber $f(\alpha_n) = \lambda_n \alpha_n + \lambda_* \omega$, wobei $\omega \in V_{n-1}$ und $\lambda_* \in \mathbf{C}$. Daher ist

$$\begin{aligned}
P_f(f)(\alpha_n) &= ((\lambda_1 - f) \cdots (\lambda_n - f))(\alpha_n) \\
&= ((\lambda_1 - f) \cdots (\lambda_{n-1} - f))(\lambda_n \alpha_n - f(\alpha_n)) \\
&= ((\lambda_1 - f) \cdots (\lambda_{n-1} - f))(\lambda_n \alpha_n - \lambda_n \alpha_n - \lambda_* \omega) \\
&= ((\lambda_1 - f) \cdots (\lambda_{n-1} - f))(\lambda_* \omega) \\
&= \lambda_* ((\lambda_1 - f) \cdots (\lambda_{n-1} - f))(\omega) \\
&= \lambda_* P_{f_{|V_{n-1}}}(f_{|V_{n-1}})(\omega) = 0.
\end{aligned}$$

Soviel für den Fall $F = \mathbf{C}$. Wie ist es, wenn $F = \mathbf{R}$? Nun, es gilt $\mathbf{R} \subset \mathbf{C}$; daher können wir sagen, daß der Vektorraum V, der eigentlich ein Vektorraum über $\mathbf{R}$ ist, auch ein Vektorraum über $\mathbf{C}$ ist. Das charakteristische Polynom $P_f(x)$ muß dann als Polynom mit komplexen Koeffizienten betrachtet werden ($P_f(x) \in \mathbf{C}[x]$). Aber glücklicherweise ist $P_f(x) = det(A - xI)$, wobei $A \in M(n \times n; \mathbf{R}) \subset M(n \times n; \mathbf{C})$; folglich ist doch $P_f(x) \in \mathbf{R}(x)$ ein Polynom mit nur reellen Koeffizienten. Aber wir haben gerade bewiesen, daß die lineare Abbildung $P_f(f) : V \to V$ die Nullabbildung ist ($P_f(f) = 0$). Keine echten komplexen Zahlen kommen vor in $P_f(f)$, und daher ist $P_f(f) : V \to V$ auch eine lineare Abbildung für V als *reellen* Raum. $\square$

Satz 19.3. *Sei V ein (endlich-dimensionaler) Vektorraum über $\mathbf{R}$ und sei $f : V \to V$ irgendeine lineare Abbildung. Dann gibt es einen f-invarianten Unterraum $W \subset V$ mit $dim(W) \leq 2$.*

Beweis. Wir können annehmen, daß $dim(V) > 2$, da sonst alles doch trivial wäre. Nach Satz 19.1 ist

$$P_f(x) = \lambda_0 \cdot (\lambda_1 - x) \cdots (\lambda_s - x) \cdot Q_t(x) \cdots Q_1(x),$$

wobei $\lambda_i \in \mathbf{R}$ für alle $0 \leq i \leq s$ und $Q_j(x) = x^2 + a_j x + b_j$ für alle $1 \leq j \leq t$, wobei a_j und $b_j \in \mathbf{R}$. Falls $s \geq 1$, dann ist λ_1 ein Eigenwert, und es existiert ein Eigenvektor α_{λ_1} mit $f(\alpha_{\lambda_1}) = \lambda_1 \alpha_{\lambda_1}$. In diesem Fall sei W der Unterraum, erzeugt durch den Vektor α_{λ_1} alleine.

Falls $s = 0$, dann ist auf jeden Fall $grad(Q_i) = 2$, für alle i. Sei $\alpha \neq 0$ irgendein nicht-trivialer Vektor in V. Dann ist $P_f(f)(\alpha) = Q_t(f) \cdots Q_1(f)(\alpha) = 0$. Wir definieren t Vektoren in V wie folgt. Für jedes i ist $\alpha_i = Q_i(f) \cdots Q_1(f)(\alpha)$. Sei nun i_* die kleinste Zahl mit $\alpha_{i_*}/not = 0$, jedoch $\alpha_{(i_*+1)} = 0$. D.h. $Q_{i_*+1}(\alpha_{i_*}) = 0$, jedoch $\alpha_{i_*} \neq 0$. Sei W der Unterraum von V, erzeugt durch die zwei Vektoren $\{\alpha_{i_*}, f(\alpha_{i_*})\}$. Dann ist W f-invariant. Warum? Sei etwa $Q_{i_*+1}(f) = f^2 + af + b$. Dann ist $Q_{i_*+1}(\alpha_{i_*}) = f(f(\alpha_{i_*})) + af(\alpha_{i_*}) + b\alpha_{i_*} = 0$. D.h. $f(f(\alpha_{i_*})) = +af(\alpha_{i_*}) + b\alpha_{i_*} \in W$, aber auch $f(\alpha_{i_*}) \in W$. □

19.2 Zurück zu orthogonalen Abbildungen

Satz 19.4. *Sei $A \in M(n \times n; \mathbf{R})$ eine orthogonale Matrix. Dann ist A zu einer Matrix A^* ähnlich, wobei*

$$A^* = \begin{pmatrix} +1 &&&&&&&& \\ & \ddots &&&&&&& \\ && +1 &&&&&& \\ &&& -1 &&&&& \\ &&&& \ddots &&&& \\ &&&&& -1 &&& \\ &&&&&& A_1 && \\ &&&&&&& \ddots & \\ &&&&&&&& A_k \end{pmatrix},$$

und $A_i = \begin{pmatrix} \cos\theta_i & \mp\sin\theta_i \\ \sin\theta_i & \pm\cos\theta_i \end{pmatrix}$, $\theta_i \in [-\pi, \pi)$ für $i = 1, \ldots, k$.

Beweis. Sei $f : V \to V$ die lineare Abbildung, die durch die Matrix A dargestellt wird bzgl. einer vorgegebenen Basis $\mathcal{B}$. Angenommen, f besitzt den Eigenvektor α. Dann ist etwa $f(\alpha) = \lambda\alpha$. Was ist λ? Es gilt $\langle \alpha, \alpha \rangle = \langle f(\alpha), f(\alpha) \rangle = \langle \lambda\alpha, \lambda\alpha \rangle = \lambda^2 \langle \alpha, \alpha \rangle$. Folglich ist $\lambda^2 = 1$. D.h. $\lambda = \pm 1$. (λ ist eine *reelle* Zahl!) Nun, mit dieser Erkentnis werden wir unseren Beweis mittels Induktion über n ($= dim(V)$) führen. Falls $n = 1$, dann gibt es auf jeden Fall einen Eigenvektor, und die Matrix A ist eine (1×1) Matrix, nämlich $A = A_* = (\pm 1)$. Sei daher $n > 1$. Nach Satz 19.3 (und einer weiteren Sekunde Nachdenken) ist klar, daß

$$V = X_1 \oplus \cdots \oplus X_s \oplus Y_1 \oplus \cdots \oplus Y_t \oplus Z_1 \oplus \cdots \oplus Z_k,$$

wobei alle X_i und Y_i 1-dimensionale Unterräume sind, erzeugt durch Eigenvektoren mit Eigenwerten $+1$, bzw. -1, und Z_i jeweils 2-dimensionale f-invariante Unterräume ohne Eigenvektoren.[2] Sei $\mathcal{B}_*$ eine Basis für f, bestehend aus den Eigenvektoren (in der gegebenen Reihenfolge), und jeweils zwei Basisvektoren für jede Z_i. Dann ist die darstellende Matrix für f bzgl. $\mathcal{B}_*$ von der Gestalt

$$\begin{pmatrix} +1 &&&&&&&& \\ & \ddots &&&&&&& \\ && +1 &&&&&& \\ &&& -1 &&&&& \\ &&&& \ddots &&&& \\ &&&&& -1 &&& \\ &&&&&& B_1 && \\ &&&&&&& \ddots & \\ &&&&&&&& B_k \end{pmatrix},$$

[2]Sei etwa $Z \subset V$ ein f-invarianter Unterraum. Da orthogonale Abbildungen stets Isomorphismen sind, ist die Einschränkung $f_{|_Z} : Z \to Z$ selbst ein Isomorphismus. Sei $Z^\perp = \{\beta \in V : \langle \alpha, \beta \rangle = 0, \forall \alpha \in Z\}$ der Senkrechtraum zu Z. Dann ist $f_{|_{Z^\perp}} : Z^\perp \to Z^\perp$ so, daß $f_{|_{Z^\perp}}(Z^\perp) = Z^\perp$. Um dies zu sehen, sei $\alpha \in Z$ und $\beta \in Z^\perp$. Es gilt $\langle \alpha, f(\beta) \rangle = \langle f^{-1}(\alpha), \beta \rangle = 0$, da $f^{-1}(\alpha) \in Z$ und $\beta \in Z^\perp$. Folglich ist auch $f(\beta) \in Z^\perp$. Daher muß $f_{|_{Z^\perp}}$ auch eine orthogonale Abbildung sein, aber mit $dim(Z^\perp) < dim(V)$. Die Induktion folgt.

mit jedem B_i eine 2×2 Untermatrix, die auch orthogonal sein muß. Wir brauchen daher nur zu zeigen, daß jede 2×2 orthogonale Matrix zu einer Matrix der Gestalt $\begin{pmatrix} \cos\theta_i & \mp\sin\theta_i \\ \sin\theta_i & \pm\cos\theta_i \end{pmatrix}$ ähnlich ist. □

Lemma. *Sei* $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M(2 \times 2; \mathbf{R})$ *orthogonal. Dann ist* $A = \begin{pmatrix} \cos\theta_i & \mp\sin\theta_i \\ \sin\theta_i & \pm\cos\theta_i \end{pmatrix}$.

Beweis. A orthogonal heißt $A^t = A^{-1}$ oder $A \cdot A^t = I$, die Identitätsmatrix. D.h. $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Diese Matrixgleichung ist eigentlich identisch mit vier kleinen Gleichungen mit reellen Zahlen, nämlich:

$$\begin{aligned} a^2 + b^2 &= 1 \\ ac + bd &= 0 \\ ac + bd &= 0 \\ c^2 + d^2 &= 1 \end{aligned}$$

Nun, die erste Gleichung kann nur gelöst werden, wenn $-1 \leq a \leq +1$. D.h. es existiert irgendeine Zahl $\theta \in [-\pi, \pi)$ mit $\cos\theta = a$. Aber dann ist offensichtlich $b = \pm\sin\theta$. Ähnlich schließt man, daß eine Zahl $\phi \in [-\pi, \pi)$ existieren muß, mit $c = \sin\phi$ und $d = \pm\cos\phi$. Die mittleren Gleichungen sind dann (nach einer Standardformel der Trigonometrie)

$$\cos\theta \sin\phi + \sin\theta \cos\phi = \sin = \sin(\theta + \phi) = 0.$$

D.h. $\theta + \phi = 0 \bmod \pi$. Der Rest ist wieder elementare Trigonometrie. □

Kapitel 20

Selbstadjungierte Abbildungen

Definition 71. Sei $f : V \to V$ eine lineare Abbildung eines endlich-dimensionalen Raumes V mit Skalarprodukt (Euklidisch oder unitär) in sich selbst. Die Abbildung heißt *selbstadjungiert*, falls $\langle \alpha, f(\beta)\rangle = \langle f(\alpha), \beta\rangle$, für alle $\alpha, \beta \in V$.

Warum sind selbstadjungierte Abbildungen interessant? Was ist die geometrische Bedeutung dieses Begriffs? Lassen wir zunächst die zweite Frage beseite liegen. In diesem Kapitel werden wir beweisen, daß alle symmetrische Matrizen in $M(n \times n; \mathbf{R})$ diagonalisierbar sind. Zur Erinnerung: eine $n \times n$ Matrix $A = (a_{ij})$ heißt *symmetrisch*, falls $a_{ij} = a_{ji}$ für alle $i, j \in \{1, \ldots, n\}$. (Für den komplexen Fall gibt es eine entsprechende Definition. Sei $A \in M(n \times n; \mathbf{C})$. Die Matrix $A = (a_{kl})$ heißt *hermitesch*, falls $A = \overline{A}^t$. D.h. $a_{kl} = \overline{a_{lk}}$ für alle $k, l \in \{1, \ldots, n\}$.) Wir werden zeigen, daß für jede reelle symmetrische Matrix, eine weitere Matrix $S \in GL(n; \mathbf{R})$ existiert, mit $S^{-1} \cdot A \cdot S = I$ eine Diagonalmatrix.

Satz 20.1. *Sei V Euklidisch oder unitär (d.h. es existiert ein Skalarprodukt) und sei $\mathcal{B}$ eine Orthonormalbasis für V. Dann gilt: $f : V \to V$ ist selbstadjungiert $\Leftrightarrow$ $M_{\mathcal{B}}(f)$ ist symmetrisch (bzw. hermitesch im unitären Fall).*

Beweis. Sei etwa $M_{\mathcal{B}}(f) = A = (a_{kl})$. Wir schreiben die Basisvektoren als Spaltenvektoren: $\alpha_k = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow k$-te Stelle. Dann ist

$$f(\alpha_k) = A \cdot \alpha_k = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} a_{1k} \\ \vdots \\ a_{nk} \end{pmatrix}.$$

Das heißt, $\langle \alpha_j, f(\alpha_k)\rangle = (0 \quad \cdots \quad \underbrace{1}_{j\text{-te Stelle}} \quad \cdots \quad 0) \cdot \begin{pmatrix} a_{1k} \\ \vdots \\ a_{nk} \end{pmatrix} = a_{jk}$. Aber wir haben vorausgesetzt, daß f selbstadjungiert ist. D.h.

$$a_{jk} = \langle \alpha_j, f(\alpha_k)\rangle = \langle f(\alpha_j), \alpha_k\rangle = \overline{\langle \alpha_k, f(\alpha_j)\rangle} = \overline{a_{kj}},$$

für alle $j, k \in \{1, \ldots, n\}$. $\square$

Satz 20.2. *Sei $f : V \to V$ selbstadjungiert, wobei V ein unitärer Raum sei. Angenommen, $\lambda \in \mathbf{C}$ ist ein Eigenwert zu f. Dann ist λ tatsächlich eine reelle Zahl.*

Beweis. Falls λ ein Eigenwert ist, dann muß ein entsprechender Eigenvektor $\alpha \neq 0$ in V existieren, mit $f(\alpha) = \lambda\alpha$. Dann ist $\lambda\langle \alpha, \alpha\rangle = \langle \alpha, \lambda\alpha\rangle = \langle \alpha, f(\alpha)\rangle = \langle f(\alpha), \alpha\rangle = \langle \lambda\alpha, \alpha\rangle = \overline{\lambda}\langle \alpha, \alpha\rangle$. Da $\langle \cdot, \cdot\rangle$ positiv definit ist, ist $\langle \alpha, \alpha\rangle \neq 0$. Folglich ist $\lambda = \overline{\lambda}$ und daher $\lambda \in \mathbf{R}$. $\square$

Satz 20.3. *Sei $f : V \to V$ selbstadjungiert. Dann zerfällt das charakteristische Polynom $P_f(x) \in F[x]$ ($F = \mathbf{R}$ oder $\mathbf{C}$) in lineare Faktoren.*

Beweis. Der Fall $F = \mathbf{C}$ folgt aus unserem bekannten Hauptsatz der Algebra. Wie ist es im Falle $F = \mathbf{R}$? Nun, bekannterweise ist $\mathbf{R} \subset \mathbf{C}$. D.h. wenn wir die Sache in $\mathbf{C}$ betrachten, dann ist $P_f(x) = (\lambda_a - x) \cdots \lambda_n - x)$, wobei $\lambda_j \in \mathbf{C}$ für $j = 1, \ldots n$. Aber diese λ_j müssen Eigenwerte für f sein; d.h. nach unserem Satz 20.2 gilt $\lambda_j \in \mathbf{R}$ für alle j. An dieser Stelle können wir die komplexen Zahlen wieder verlassen und einfach beobachten, daß die Gleichung $P_f(x) = (\lambda_a - x) \cdots \lambda_n - x)$ eine Faktorisierung in *reelle* Faktoren darstellt. □

Satz 20.4. *Sei $f : V \to V$ selbstadjungiert. Dann existiert eine Orthonormalbasis $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ von V, bestehend aus Eigenvektoren unter f.*

Beweis. Induktion über $n = dim(V)$. Falls $n = 1$, dann ist der Satz, wie immer, trivial. Sei daher $n > 1$ und sei die induktive Hypothese angenommen, für die Fälle $< n$. Nach Satz 20.3 zerfällt das charakteristische Polynom $P_f(x)$ in lineare Faktoren, und daher muß ein Eigenvektor $\alpha_1 \neq 0$ existieren, etwa mit $f(\alpha_1) = \lambda_1 \alpha_1$. (Falls zunächst $\|\alpha_1\| \neq 1$, dann können wir, nach unserem Standardverfahren, α_1 ersetzen mit dem normierten Vektor $\alpha_1/\|\alpha_1\|$. Sei nun $U \subset V$ der 1-dimensionale Unterraum, erzeugt durch den einzigen Vektor α_1. Sei $W = U^\perp = \{\beta \in V : \langle \beta, \alpha_1 \rangle = 0\}$ der Senkrechtraum zu U. Da $V = U \oplus W$ und $dim(U) = 1$, gilt $dim(W) = n - 1$. Aber auch $f(W) \subset W$. Warum? Sei $\beta \in W$. Dann ist $\langle f(\beta), \alpha_1 \rangle = \langle \beta, f(\alpha_1) \rangle = \langle \beta, \lambda_1 \alpha_1 \rangle = \lambda_1 \langle \beta, \alpha_1 \rangle = \lambda_1 \cdot 0 = 0$. D.h. $f(\beta) \in U^\perp = W$. Wir können daher die Einschränkung von f auf den Unterraum W als eine Abbildung $f_{|W} : W \to W$ betrachten. Nun, $f_{|W}$ ist sicherlich selbstadjungiert, da f schon selbstadjungiert ist. Aber dann erfüllt die Abbildung $f_{|W} : W \to W$ die Voraussetzungen des Satzes, mit $dim(W) < n$. Nach Induktion existiert daher eine Basis $\{\alpha_2, \ldots, \alpha_n\}$, bestehend aus Eigenvektoren unter $f_{|W}$. Aber diese Vektoren sind auch Eigenvektoren unter f. □

Korollar 20.4.1. *Sei $A \in M(n \times n; F)$ ($F = \mathbf{R}$ oder $\mathbf{C}$) symmetrisch oder Hermitesch. Dann existiert ein $S \in GL(n, F)$ mit $S^{-1} \cdot A \cdot S \in M(n \times n; \mathbf{R})$ diagonal. (Bemerken Sie: auch im Falle $F = \mathbf{C}$ muß die diagonalisierte Matrix reell sein, wegen Satz 20.2.)*

Lemma. *Sei V ein Vektorraum mit Skalarprodukt und seien $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ und $\mathcal{B}' = \{\alpha'_1, \ldots, \alpha'_n\}$ zwei Orthonormalbasen für V. Sei $S = (s_{kl}) \in GL(n; F)$ die Matrix, die den Basiswechsel $\mathcal{B} \to \mathcal{B}'$ darstellt. Dann ist S unitär (bzw. orthogonal).*

Beweis. Seien $k, l \in \mathbf{C}$ (bzw. $\mathbf{R}$) beliebig. Dann ist

$$\langle \alpha'_k, \alpha'_l \rangle = \langle S \cdot \alpha_k, S \cdot \alpha_l \rangle = \langle \sum_{i=1}^n s_{ik} \alpha_i, \sum_{j=1}^n s_{jl} \alpha_j \rangle = \sum_{i=1}^n \sum_{j=1}^n \overline{s}_{ik} s_{jl} \langle \alpha_i, \alpha_j \rangle$$

$$= \sum_{i=1}^n \overline{s}_{ik} s_{il} = \begin{cases} 1, & \text{falls } k = l, \\ 0 & \text{sonst,} \end{cases}$$

da $\mathcal{B}'$ orthonormal ist. Aber

$$\sum_{i=1}^n \overline{s}_{ik} s_{il} = \sum_{i=1}^n \overline{s}'_{ki} s_{il},$$

für $S' = (s'_{ki}) = S^t$ die zu S transponierte Matrix. Insgesamt haben wir gezeigt, daß $\overline{S}^t \cdot S = I$ oder $\overline{S}^t = S^{-1}$. D.h. S ist unitär (bzw. orthogonal). □

Korollar 20.4.2. *Wir können annehmen, daß die Matrix S orthogonal, bzw. unitär ist, so daß*

$$\overline{S}^t \cdot A \cdot S = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, \quad \lambda_j \in \mathbf{R}, \forall j.$$

Beweis. Wir behandeln den komplexen Fall (f unitär). Der reelle Fall folgt dann als einfache Konsequenz. Nun, die Korrespondenz f unitär $\Leftrightarrow M_{\mathcal{B}}(f)$ unitär setzt voraus, daß die Basis $\mathcal{B}$ orthonormal ist. Es ist daher am einfachsten, wenn wir unsere vorgegebene unitäre Matrix A als die Darstellung einer selbstadjungierten Abbildung $f : \mathbf{C}^n \to \mathbf{C}^n$ bzgl. der kanonischen Basis (die automatisch orthonormal ist) von $\mathbf{C}^n$ betrachten. Aber $\mathcal{B}'$, nämlich die Basis von Eigenvektoren unter f, ist auch orthonormal. Wir brauchen daher nur unser Lemma in Anspruch zu nehmen. □

Kapitel 21

Dualräume

Im letzten Kapitel haben wir 'selbst-adjungierte' lineare Abbildungen gesehen. Die Definition $\langle \alpha, f(\beta) \rangle = \langle f(\alpha), \beta \rangle$ scheint angenehm symmetrisch zu sein. Aber was ist die algebraische oder geometrische Bedeutung von 'symmetrisch' oder 'selbst-adjungiert'? Nun, in Kapitel 16 haben wir gesehen, daß es möglich ist, eine Bilinearform durch eine Matrix darzustellen. Symmetrische Bilinearformen entsprachen symmetrischen Matrizen. Sei $A \in M(n \times n : F)$ eine solche Matrix, berechnet bzgl. einer vorgegebenen Basis $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$. Nach dem Satz über die Existenz von Orthonormalbasen gibt es eine solche Basis $\mathcal{B}_*$. Die Matrix $S \in GL(n : F)$, die den Basiswechsel $\mathcal{B} \to \mathcal{B}_*$ darstellt, gibt dann — durch die Transformationsformel — eine diagonale Matrix, nämlich $S^t \cdot A \cdot S$. Diese Matrix ist die Darstellung der Bilinearform bzgl. der Orthonormalbasis $\mathcal{B}_*$. Andererseits haben wir im letzten Kapitel gesehen, daß eine selbst-adjungierte Matrix A, die eine lineare *Abbildung*—nicht eine *Bilinearform*—darstellt, durch eine Ähnlichkeitstransformation $S^{-1} \cdot A \cdot S$ in eine diagonale Matrix umgeformt werden kann. Diese zwei Tatsachen scheinen zunächst gar nicht zusammen zu passen, da S^{-1} und S^t im allgemeinen zwei ganz verschiedene Matrizen sind. Aber falls S *orthogonal* ist, dann gilt $S^{-1} = S^t$ (bzw. $S^{-1} = \overline{S}^t$ für S unitär). Um diese Zusammenhänge besser zu verstehen, werden wir jetzt Dualräume untersuchen.

Definition 72. Seien V und W Vektorräume über einem Körper F.[1] Die Menge $Hom_F(V, W)$ ist die Menge aller linearen Abbildungen $f : V \to W$.

Es ist ziemlich trivial zu beweisen, daß $Hom_F(V, W)$ ein Vektorraum über F ist. Die Vektoraddition ist $(\phi + \psi)(\alpha) = \phi(\alpha) + \psi(\alpha) \in W$ für alle $\alpha \in V$. Die Skalarmultiplikation ist $(a \cdot \phi)(\alpha) = a \cdot \phi(\alpha)$.

Nun, für jeden Körper F gibt es einen ganz besonderen Vektorraum: nämlich der Körper F selbst! Wir können daher $W = F$ nehmen und wir erhalten die folgende Definition.

Definition 73. $V^* = Hom_F(V, F)$ heißt der Dualraum zu V. Jedes Element $\phi \in V^*$ heißt eine *Linearform* oder lineares *Funktional* auf V.

Beispiele

1. Sei $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$ eine Basis für V. Jeder Vektor β in V kann als Linearkombination $\beta = \sum_{i=1}^n b_i \alpha_i$ geschrieben werden. Eine Linearform $\phi \in V^*$ wird dann gegeben durch die Regel $\phi(\beta) = b_1$, für alle $\beta \in V$.

2. Sei $S_{[0,1]}$ die Menge der stetigen reellen Funktionen $f : [0, 1] \to \mathbf{R}$. Wir haben schon gesehen, daß $S_{[0,1]}$ ein Vektorraum über $\mathbf{R}$ ist. Sei jetzt $\phi \in S^*_{[0,1]}$ die folgende Abbildung. $\phi(f) = f(0)$ für alle $f \in S_{[0,1]}$. Ist $\phi : S_{[0,1]} \to \mathbf{R}$ tatsächlich eine lineare Abbildung? Dazu seien $a, b \in \mathbf{R}$ und $f, g \in S_{[0,1]}$. Dann ist $\phi(af + bg) = af(0) + bg(0) = a\phi(f) + b\phi(g)$.

3. Angenommen, V besitzt ein inneres Produkt $\langle \cdot, \cdot \rangle$. Sei $\alpha' \in V$ ein fest vorgegebener Vektor in V. Dann gibt es ein Element $\phi \in V^*$, definiert durch die Regel $\phi(\beta) = \langle \alpha', \beta \rangle$ für alle $\beta \in V$.

4. Sei V ein beliebiger (endlich dimensionaler) Vektorraum über F mit Basis $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$. Wir können, ähnlich wie in Beispiel 1, n verschiedene Linearformen ϕ_i, $i = 1, \dots, n$ definieren wie folgt. Sei $\beta = \sum_{i=1}^n b_i \alpha_i$ ein beliebiger Vektor in V. Dann ist $\phi_j(\beta) = b_j$ für alle j. Dadurch werden n verschiedene Vektoren $\phi_1, \dots, \phi_n \in V^*$ im Dualraum zu V definiert.

Satz 21.1. *Die Menge $\mathcal{B}^* = \{\phi_1, \dots, \phi_n\}$ von Beispiel 4 ist eine Basis des Dualraumes V^*. $\mathcal{B}^*$ heißt die zu $\mathcal{B}$ duale Basis von V^*.*

[1] In diesem Abschnitt werden wir wieder beliebige Körper—nicht nur $\mathbf{R}$ oder $\mathbf{C}$—zulassen.

Beweis. Ist V^* erzeugt durch $\mathcal{B}^*$? Sei $\phi \in V^*$ beliebig. D.h. $\phi : V \to F$ ist eine lineare Abbildung. Wir können zunächst die Auswirkung von ϕ, angewandt auf die Basis $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$, untersuchen. Für jedes i ist $\phi(\alpha_i) = \lambda_i \in F$ irgendein festes Element von F. Sei jetzt $\beta = \sum_{i=1}^n b_i\alpha_i$ ein beliebiger Vektor in V. Dann ist

$$\phi(\beta) = \phi(\sum_{i=1}^n b_i\alpha_i) = \sum_{i=1}^n b_i\phi(\alpha_i) = \sum_{i=1}^n b_i\lambda_i = \sum_{i=1}^n \lambda_i\phi_i(\beta) = \left(\sum_{i=1}^n \lambda_i\phi_i\right)(\beta).$$

Die vorletzte Gleichung folgt, da $b_i = \phi_i(\beta)$ für jedes i. D.h. wir haben unsere zufällig gewählte Linearform $\phi \in V^*$ als Linearkombination der Vektoren in $\mathcal{B}^*$ dargestellt.

Ist $\mathcal{B}^*$ linear unabhängig? Sei $\phi = \sum_{i=1}^n x_i\phi_i$ irgendeine Linearkombination mit $x_{i_0} \neq 0$ für mindestens ein $i_0 \in \{1, \ldots, n\}$. Dann ist $\phi(\alpha_{i_0}) = \phi_{i_0}(\alpha_{i_0}) = x_i \neq 0$. D.h. $\phi \neq 0$ in V^*. Folglich muß $\mathcal{B}^*$ linear unabhängig sein. □

Korollar 21.1.1. *$dim(V^*) = n = dim(V)$.*

Korollar 21.1.2. *Gegeben V und $\mathcal{B}$, dann gibt es einen eindeutigen Isomorphismus $V \to V^*$, wobei $\mathcal{B} \to \mathcal{B}^*$.*

D.h. wir haben eine Korrespondenz $\mathcal{B} \leftrightarrow \mathcal{B}^*$, gegeben durch $\alpha_i \leftrightarrow \phi_i$. Es ist daher einfacher zu schreiben $\alpha_i^* \equiv \phi_i$. D.h. $\mathcal{B}^* = \{\alpha_1^*, \ldots, \alpha_i^*\}$.

Um die ganze Sache noch interessanter zu machen, können wir den Dualraum von dem Dualraum $(V^*)^*$ nehmen. Was ist dann $(V^*)^*$? Wir haben $V \leftrightarrow V^* \leftrightarrow (V^*)^*$; daher ist sicherlich V zu $(V^*)^*$ isomorph, wobei $\alpha_i \leftrightarrow \alpha_i^* \leftrightarrow (\alpha_i^*)^*$.

Es ist üblich, $(V^*)^*$ *gleichzusetzen* mit dem ursprünglichen Raum V. Warum? Nun, die Korrespondenz $V \leftrightarrow V^*$ hing von der Wahl der Basis $\mathcal{B}$ ab. Aber die Korrespondenz $V \leftrightarrow (V^*)^*$ kann ohne Bezug zu irgendeiner Basis definiert werden. Sei $\alpha \in V$ beliebig. Dann ist $\alpha \leftrightarrow \alpha^{**}$ mit $\alpha^{**}(\phi) = \phi(\alpha)$, für alle $\phi \in V^*$. Die dadurch definierte Abbildung $\alpha^{**} : V^* \to F$ ist linear, da

$$\alpha^{**}(a\phi + b\psi) = (a\phi + b\psi)(\alpha) = a\phi(\alpha) + b\psi(\alpha) = a\alpha^{**}(\phi) + b\alpha^{**}(\psi)$$

für alle a, $b \in F$ und ϕ, $\psi \in V^*$. Daß $\alpha \leftrightarrow \alpha^{**}$, sieht man, indem auch die Basen $\mathcal{B}$ und $\mathcal{B}^*$ herangezogen werden. D.h. wenn $\alpha = \sum_{i=1}^n a_i\alpha_i$, dann ist der entsprechende Vektor in V^* die Linearkombination $\sum_{i=1}^n a_i\alpha_i^*$, und in V^{**} haben wir $\sum_{i=1}^n a_i\alpha_i^{**}$. Dann ist für beliebige $\phi \in V^*$

$$\left(\sum_{i=1}^n a_i\alpha_i^{**}\right)(\phi) = \sum_{i=1}^n a_i\alpha_i^{**}(\phi) = \sum_{i=1}^n a_i\phi(\alpha_i) = \phi\left(\sum_{i=1}^n a_i\alpha_i\right).$$

21.1 Duale Abbildungen

Definition 74. Sei $f : V \to W$ eine lineare Abbildung zwischen endlich dimensionalen Räumen. Die Abbildung $f^* : W^* \to V^*$, gegeben durch $f^*(\phi) = \phi \cdot f$ für alle $\phi \in W^*$, heißt die zu f *duale* Abbildung.

Ist tatsächlich $f^*(\phi) \in V^*$? Sei $\alpha \in V$. Dann ist $f^*(\phi)(\alpha) = \phi(f(\alpha)) \in F$ wohldefiniert, da doch $f(\alpha) \in W$. (Daß $f^*(\phi)$ linear ist, folgt trivialerweise aus den Definitionen.)

Aber nach dem Korollar 2 zum Satz 21.1 sind die Räume V und V^* und W und W^* jeweils zueinander isomorph (d.h. im Rahmen der linearen Algebra sind sie 'gleich'). Diese Überlegung veranlaßt uns, das folgende Diagramm zu zeichnen.

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \tilde{s}\downarrow & & \downarrow\tilde{r} \\ V^* & \xleftarrow{f^*} & W^* \end{array}$$

Beide Abbildungen $\tilde{r}$ und $\tilde{s}$ sind Isomorphismen, so daß wir eigentlich frei sind, die Pfeile entweder in der einen oder in der anderen Richtung (entsprechend etwa $\tilde{s}$ oder $\tilde{s}^{-1}$) zu zeichnen. Nun, die Abbildung f läuft offensichtlich von V nach W. Aber es gibt auch einen erlaubten Weg von W zurück nach V: nämlich die Komposition $f^{ad} \equiv \tilde{s}^{-1} \cdot f^* \cdot \tilde{r} : W \to V$.

Definition 75. f^{ad} heißt die zu f *adjungierte* Abbildung.

Was heißt dann 'selbst-adjungiert'? Wir haben diesen Begriff früher definiert als: $f : V \to V$ ist selbst-adjungiert, falls $\langle\alpha, f(\beta)\rangle = \langle f(\alpha), \beta\rangle$, für alle α, $\beta \in V$, wobei V ein Vektorraum mit Skalarprodukt ist. Es scheint nun sinnvoll, zu erwarten, daß eine selbst-adjungierte Abbildung identisch ist mit ihrer adjungierten Abbildung. D.h. $f = f^{ad}$ für solche Abbildungen. Stimmt diese Gleichung?

Um diese Frage zu beantworten, müssen wir die Definition von adjungierten Abbildungen einschränken auf den speziellen Fall von selbst-Abbildungen von Vektorräumen mit Skalarprodukten. Sei V ein solcher Raum mit Skalarprodukt $\langle \cdot, \cdot \rangle : V \times V \to F$. Wir können dieses Skalarprodukt auch so bezeichnen: $s : V \times V \to F$. D.h. $s(\alpha, \beta) \equiv \langle \alpha, \beta \rangle$, für alle α, $\beta \in V$. Für jedes $\alpha \in V$ ist dann die Abbildung $\tilde{s}(\alpha) : V \to F$, gegeben durch $\tilde{s}(\alpha)(\beta) = s(\alpha, \beta)$ für alle $\beta \in V$, eine lineare Abbildung, wie wir schon gesehen haben. D.h. $\tilde{s}(\alpha) \in V^*$ für alle $\alpha \in V$, und folglich wird eine Abbildung $\tilde{s} : V \to V^*$ definiert.

Satz 21.2. *Seien $s : V \times V \to F$ und $r : W \times W \to F$ nicht ausgeartete Bilinearformen (bzw. Sesquilinearformen), so daß die Abbildungen $\tilde{r}$ und $\tilde{s}$ Isomorphismen sind. Sei $f : V \to W$ linear. Dann ist $s(f^{ad}(\alpha), \beta) = r(\alpha, f(\beta))$ für alle $\alpha \in W$ und $\beta \in V$.*

Beweis. Es gilt

$$\begin{aligned} r(\alpha, f(\beta)) &= \tilde{r}(\alpha)(f(\beta)) \qquad (\text{da } \tilde{r}(\alpha)(\cdot) = r(\alpha, \cdot)) \\ &= f^*(\tilde{r}(\alpha))(\beta) \qquad (\text{da } \phi \cdot f = f^* \cdot \phi,\ \forall \phi \in W^*) \\ &= \tilde{s} \cdot \tilde{s}^{-1} f^*(\tilde{r}(\alpha))(\beta) \\ &= \tilde{s}(\tilde{s}^{-1} f^*(\tilde{r}(\alpha)))(\beta) \\ &= s(f^{ad}(\alpha), \beta) \end{aligned}$$

□

Nun, Skalarprodukte sind auf jeden Fall nicht ausgeartet. Daher gilt, falls eine Selbstabbildung $f : V \to V$ 'selbst-adjungiert' ist (und dies heißt für uns jetzt, daß $f^{ad} = f$), dann gilt $\langle f^{ad}(\alpha), \beta \rangle = \langle f(\alpha), \beta \rangle = \langle \alpha, f(\beta) \rangle$.

Dieses Ergebnis zeigt, daß unsere frühere Diskussion über selbst-adjungierte Abbildungen einen vernünftigen theoretischen Rahmen erhält, wenn wir die Theorie der Dualräume benutzen. Es ging dabei um die Diagonalisierbarkeit von selbst-adjungierten Abbildungen. Was ist diese Bedingung im Rahmen der Dualräume?

Definition 76. Sei $f : V \to V$ surjektiv, wobei V endlich dimensional ist; d.h. f ist ein Isomorphismus. Die Abbildung f heißt *normal*, falls $f \cdot f^{ad} = f^{ad} \cdot f$.

Beispiele

1. Falls f selbst-adjungiert ist, dann ist $f = f^{ad}$, und daher ist sicherlich $f \cdot f^{ad} = f \cdot f = f^{ad} \cdot f$.
2. Falls $f^{ad} = f^{-1}$, dann ist $f \cdot f^{ad} = f \cdot f^{-1} = id = f^{-1} \cdot f = f^{ad} \cdot f$, und f ist wieder normal. Aber was heißt $f^{ad} = f^{-1}$?

Satz 21.3. *Sei f unitär. $\Rightarrow f^{ad} = f^{-1}$.*

Beweis. f unitär $\Rightarrow \langle f(\alpha), f(\beta) \rangle = \langle \alpha, \beta \rangle$, für alle α, $\beta \in V$. Da nun f ein Isomorphismus ist, muß die lineare Abbildung $f^{-1} : V \to V$ existieren. Sei $\gamma \in V$ so, daß $f(\gamma) = \alpha$. D.h. $\gamma = f^{-1}(\alpha)$. Es gilt $\langle \alpha, f(\beta) \rangle = \langle f(\gamma), f(\beta) \rangle = \langle \gamma, \beta \rangle = \langle f^{-1}(\alpha), \beta \rangle = \langle f^{ad}(\alpha), \beta \rangle$. Diese Gleichung gilt für beliebige α, $\beta \in V$. Daher können wir folgern, daß $f^{ad} = f^{-1}$. □

Nun, wir haben bewiesen, daß alle unitären Abbildungen diagonalisierbar sind. Wir können viel mehr beweisen: nämlich daß sogar alle normalen Abbildungen diagonalisierbar sind.

Lemma (1). *Sei $f : V \to V$ normal, wobei V ein Raum mit innerem Produkt ist. Dann ist $ker(f) = ker(f^{ad})$.*

Beweis. $\alpha \in ker(f) \Leftrightarrow f(\alpha) = 0 \Leftrightarrow \langle f(\alpha), f(\alpha) \rangle = 0$. Aber

$$\begin{aligned} 0 = \langle f(\alpha), f(\alpha) \rangle &= \langle f^{ad}(f(\alpha)), \alpha \rangle \qquad (\text{Satz 21.2}) \\ &= \langle f(f^{ad}(\alpha)), \alpha \rangle \qquad (\text{da } f \text{ normal ist.}) \\ &= \overline{\langle f(f^{ad}(\alpha)), \alpha \rangle} \qquad (\text{da } 0 = \overline{0}) \\ &= \langle \alpha, f(f^{ad}(\alpha)) \rangle \\ &= \langle f^{ad}(\alpha), f^{ad}(\alpha) \rangle \end{aligned}$$

Folglich ist $\alpha \in ker(f) \Leftrightarrow \alpha \in ker(f^{ad})$. D.h. $ker(f) = ker(f^{ad})$. □

Lemma (2). *Die Bedingungen wie im Lemma (1). Sei $\alpha \in V$ ein Eigenvektor unter f mit $f(\alpha) = \lambda\alpha$. Dann ist $f^{ad}(\alpha) = \overline{\lambda}\alpha$.*

Beweis. Sei $g = f - \lambda \cdot id$. Dann ist $g^{ad} = f^{ad} - (\lambda \cdot id)^{ad} = f^{ad} - \overline{\lambda} \cdot id$.[2] Die Abbildung g ist auch normal, da $g \cdot g^{ad} = (f - \lambda \cdot id) \cdot (f^{ad} - \overline{\lambda} \cdot id) = f \cdot f^{ad} - \overline{\lambda} \cdot f - \lambda \cdot f^{ad} + \lambda\overline{\lambda} \cdot id = g^{ad} \cdot g$. Aber nach Lemma 1 ist $\alpha \in ker(g) \Leftrightarrow \alpha \in ker(g^{ad}) \Leftrightarrow f(\alpha) = \lambda\alpha \Leftrightarrow f^{ad}(\alpha) = \overline{\lambda}\alpha$. ☐

Satz 21.4. *Sei $f : V \to V$, wobei V ein unitärer Raum ist. (Insbesondere ist V ein Vektorraum über den komplexen Zahlen $\mathbf{C}$.) Dann gilt: f ist normal $\Leftrightarrow \exists$ eine Orthonormalbasis für V, bestehend aus lauter Eigenvektoren unter f.*

Beweis. "$\Leftarrow$" Sei $\{\alpha_1, \dots, \alpha_n\}$ die Orthonormalbasis mit $f(\alpha_j) = \lambda_j\alpha_j$; $\lambda_j \in \mathbf{C}$ für alle j. Sei jetzt j beliebig und $f^{ad}(\alpha_j) = \sum_{k=1}^n a_k\alpha_k$ mit jeweils $a_k \in \mathbf{C}$. Es gilt

$$\begin{aligned} \langle \alpha_k, f^{ad}(\alpha_j)\rangle &= \overline{\langle f^{ad}(\alpha_j), \alpha_k\rangle} \\ &= \overline{\langle \alpha_j, f(\alpha_k)\rangle} \\ &= \overline{\lambda_k\langle \alpha_j, \alpha_k\rangle} = \begin{cases} 0, & \text{falls } j \neq k \\ \overline{\lambda_k}, & \text{falls } j = k. \end{cases} \end{aligned}$$

Was sind die a_k? Wir haben

$$a_k = \langle \alpha_k, \sum_{l=1}^n a_l\alpha_l\rangle = \langle \alpha_k, f^{ad}(\alpha_j)\rangle = \begin{cases} 0, & \text{falls } j \neq k \\ \overline{\lambda_k}, & \text{falls } j = k. \end{cases}$$

$\Rightarrow f^{ad}(\alpha_j) = \overline{\lambda_j}\alpha_j$ für alle j zwischen 1 und n. $\Rightarrow f(f^{ad}(\alpha_j)) = f(\overline{\lambda_j}\alpha_j) = \overline{\lambda_j}f(\alpha_j) = \overline{\lambda_j}\lambda_j\alpha_j = \lambda_j\overline{\lambda_j}\alpha_j = f^{ad}(\lambda_j\alpha_j) = f^{ad}(f(\alpha_j))$. Daher ist die Normalitätsbedingung $f^{ad} \cdot f = f \cdot f^{ad}$ gültig, zumindest auf der Basis. Aber wie wir längst wissen, ist eine lineare Abbildung bestimmt durch das Verhalten auf einer Basis. Folglich muß f normal sein.

"$\Rightarrow$" Sei $n = dim(V)$ mit $n \geq 1$. Nach dem Fundamentalsatz der Algebra existiert ein Eigenwert $\lambda \in \mathbf{C}$ unter f. Sei $\alpha \in V$ ein entsprechender Eigenvektor mit $f(\alpha) = \lambda\alpha$. Sei W jetzt der Senkrechtraum: $W = \{\beta \in V : \langle\alpha, \beta\rangle = 0\}$. Für alle $\beta \in W$ gilt $\langle\alpha, f(\beta)\rangle = \langle f^{ad}(\alpha), \beta\rangle = \langle\overline{\lambda}\alpha, \beta\rangle = \lambda\langle\alpha, \beta\rangle = 0$. Folglich ist $f(\beta) \in W$. D.h. $f(W) \subset W$. Aber es gilt $dim(W) = n - 1$, und $f_{|W} : W \to W$ muß auch normal sein. Nach Induktion gibt es dann eine Orthonormalbasis $\{\alpha_2, \dots, \alpha_n\}$ für W, bestehend aus Eigenvektoren. Aber unser α ist senkrecht zu allen α_j, $j > 1$, und α ist auch ein Eigenvektor. Daher muß $\{\alpha_1 \equiv \alpha, \alpha_2, \dots, \alpha_n\}$ eine Orthonormalbasis sein, bestehend aus Eigenvektoren unter f. ☐

Korollar 21.4.1. *Sei $A \in M(n \times n : \mathbf{C})$. Dann gilt $A \cdot \overline{A}^t = \overline{A}^t \cdot A \Leftrightarrow A$ ist diagonalisierbar.*

Dadurch haben wir eine leicht nachprüfbare algebraische Bedingung, die genau besagt, wann eine (komplexe) Matrix diagonalisierbar ist.

[2] Für alle linearen Abbildungen $f_1, f_2 : V \to V$ gilt stets $(f_1 + f_2)^{ad} = f_1^{ad} + f_2^{ad}$, da

$$\begin{aligned} \langle (f_1 + f_2)^{ad}(\alpha), \beta\rangle &= \langle\alpha, (f_1 + f_2)(\beta)\rangle \\ &= \langle\alpha, f_1(\beta)\rangle + \langle\alpha, f_2(\beta)\rangle \\ &= \langle f_1^{ad}(\alpha), \beta\rangle + \langle f_2^{ad}(\alpha), \beta\rangle \\ &= \langle (f_1^{ad} + f_2^{ad})(\alpha), \beta\rangle \end{aligned}$$

auch

$$\begin{aligned} \langle (\lambda \cdot id)^{ad}(\alpha), \beta\rangle &= \langle\alpha, \lambda\beta\rangle \\ &= \lambda\langle\alpha, \beta\rangle\rangle \\ &= \langle\overline{\lambda}\alpha, \beta\rangle\rangle \end{aligned}$$

Kapitel 22

Jordansche Normalform

In diesem Kapitel werden wir uns hauptsächlich mit quadratischen Matrizen über den *komplexen* Zahlen beschäftigen; d.h. die Menge $M(n \times n : \mathbf{C})$. Wir haben schon gesehen, daß alle Matrizen in $M(n \times n : \mathbf{C})$ trigonalisierbar (aber nicht notwendigerweise diagonalisierbar) sind. Trotzdem ist es möglich, alle quadratischen komplexen Matrizen so umzuformen, daß sie doch eine sehr einfache — 'fast diagonale' — Gestalt haben.

Definition 77. Sei K ein Körper und $\lambda \in K$. Die $r \times r$ Matrix

$$J = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ & & & \lambda & 1 \\ 0 & & \cdots & 0 & \lambda \end{pmatrix} \in M(n \times n : F)$$

heißt die Jordanmatrix (der Ordnung r) zum Eigenwert λ. Sei weiter

$$M = \begin{pmatrix} J_1 & & & 0 \\ & J_2 & & \\ & & \ddots & \\ 0 & & & J_l \end{pmatrix}$$

eine $r \times r$ Matrix, wobei jedes J_i eine Jordanmatrix ist. Dann heißt M eine Matrix in Jordan Normalform.

Beispiele

$$\left(\begin{array}{cc|c} 3 & 1 & 0 \\ 0 & 3 & 0 \\ \hline 0 & 0 & 2 \end{array}\right) \text{ oder } \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ oder } \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Satz 22.1 (Jordan Normalform). *Sei V ein endlich dimensionaler Vektorraum über einem Körper F (nicht notwendigerweise $\mathbf{C}$!) und sei $f : V \to V$ eine lineare Abbildung. Angenommen, das charakteristische Polynom $P_f(x) \in F[x]$ zerfällt in lineare Faktoren. Dann existiert eine Basis $\mathcal{B}$ für V so, daß die darstellende Matrix $M_{\mathcal{B}}(f)$ in Jordan Normalform ist.*

Korollar 22.1.1. *Eine Matrix ist genau dann trigonalisierbar, wenn sie ähnlich ist mit einer Matrix in Jordan Normalform.*

Der Beweis des Satzes ist doch etwas vielfältiger als das, woran wir bis jetzt gewöhnt sind. Es wird nötig sein, zuerst einige Zwischenergebnisse zu betrachten.

Definition 78. Sei $\lambda \in F$ ein Eigenwert zu f. Der *Hauptraum* zum Eigenwert λ ist

$$H(f : \lambda) = \{\alpha \in V : \exists s \in N \text{ mit } (f - \lambda \cdot id)^{(s)}(\alpha) = 0\},$$

wobei $(f - \lambda \cdot id)^{(s)}(\alpha) = \underbrace{(f - \lambda \cdot id) \cdots (f - \lambda \cdot id)}_{s-mal}(\alpha)$.

Es ist klar, daß

$$Eig(f : \lambda) = \text{ der Eigenraum zum Eigenwert } \lambda \subset H(f : \lambda).$$

(Nehme einfach $s = 1$, um dies zu sehen.) Andererseits ist auch klar, daß

$$H(f:\lambda) = \cup_{s=1} ker(f - \lambda \cdot id)^{(s)},$$

wobei

$$ker(f - \lambda \cdot id)^{(s)} = \{\alpha \in V : (f - \lambda \cdot id)^{(s)})(\alpha) = 0\}.$$

Satz 22.2. *Sei $P_f(x) \in F[x]$ das charakteristische Polynom zu f, das in lineare Faktoren zerfällt. Sei $\lambda \in F$ ein Eigenwert mit der algebraischen Vielfachheit $r \in N$. (D.h. $P_f(x) = (x-\lambda)^{(r)} \cdot Q(x)$, wobei $(x-\lambda) \not| Q(x)$). Sei $W = H(f:\lambda)$. Dann gilt*

1. *$dim(W) = r$.*
2. *$f(W) \subset W$.*
3. *$\exists$ ein linearer Unterraum $U \subset V$ mit $V = W \oplus U$ und $f(U) \subset U$.*

Beweis. Da $P_f(x)$ in lineare Faktoren zerfällt, ist f trigonalisierbar. D.h. $\exists$ eine Basis $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ für V mit $M_{\mathcal{B}}(f)$ eine Dreiecksmatrix mit den Eigenwerten zu f als Elemente entlang der Hauptdiagonalen. Wir können die Basis so wählen, daß *zuerst* r Kopien von λ in der Hauptdiagonalen erscheinen. D.h.

$$M_{\mathcal{B}}(f) = \left(\begin{array}{ccc|c} \lambda & \cdots & * & * \\ & \ddots & \vdots & \vdots \\ 0 & & \lambda & * \\ \hline & 0 & & * \end{array}\right) = \left(\begin{array}{c|c} X & C \\ \hline 0 & D \end{array}\right),$$

wobei '$*$' steht für 'irgendetwas' (vielleicht nicht Null) in F, und

$$X = \begin{pmatrix} \lambda & x_{12} & x_{13} & \cdots & x_{1r} \\ & \lambda & x_{23} & & x_{2r} \\ & & \ddots & & \vdots \\ & 0 & & \lambda & x_{r-1\,r} \\ & & & & \lambda \end{pmatrix}.$$

D.h. $X = (x_{ij})$ mit $x_{ij} = \begin{cases} 0, & \text{falls } i > j, \\ \lambda, & \text{falls } i = j, \end{cases}$ und auch die $(n-r) \times (n-r)$ Matrix D ist in Dreiecksform. (Die Form der Matrix C ist egal.)

Wir können auch $X = \lambda \cdot I_r + N$ schreiben, wobei I_r die $r \times r$ Identitätsmatrix ist und

$$N = \begin{pmatrix} 0 & x_{12} & x_{13} & \cdots & x_{1r} \\ & 0 & x_{23} & & x_{2r} \\ & & \ddots & & \vdots \\ & 0 & & 0 & x_{r-1\,r} \\ & & & & 0 \end{pmatrix}.$$

D.h. $N = (n_{ij})$ mit $n_{ij} = \begin{cases} 0, & \text{falls } i \geq j, \\ x_{ij}, & \text{sonst.} \end{cases}$ Wir schreiben jetzt

$$M_{\mathcal{B}}(f) - \lambda I_n = \left(\begin{array}{c|c} N & C \\ \hline 0 & D_1 \end{array}\right) = B,$$

etwa. Nun, die Hauptdiagonale der Matrix D_1 enthält sicherlich keine Nullen, da sonst λ doch eine größere algebraische Vielfachheit als r haben würde. Andererseits ist auch D_1 eine Dreiecksmatrix. D.h. $Rang(D_1) = n - r$.

Behauptung 1: Für alle $s \in N$ gilt $B^s = \left(\begin{array}{c|c} N^s & C_s \\ \hline 0 & D_1^s \end{array}\right)$, wobei C_s irgendeine Matrix ist, aber $N^s = \underbrace{N \times \cdots \times N}_{s-mal}$ und $D_1^s = \underbrace{D_1 \times \cdots \times D_1}_{s-mal}$.

Beweis der Behauptung. Induktion über s. Für $s = 2$ ist die Sache eigentlich klar. Es gilt nämlich:

$$\begin{aligned} B \cdot B &= \left(\begin{array}{c|c} N & C \\ \hline 0 & D_1 \end{array}\right) \cdot \left(\begin{array}{c|c} N & C \\ \hline 0 & D_1 \end{array}\right) \\ &= \left(\begin{array}{c|c} N \cdot N + C \cdot 0 & N \cdot C + C \cdot D_1 \\ \hline 0 \cdot N + D_1 \cdot 0 & 0 \cdot C + D_1 \cdot D_1 \end{array}\right) = \left(\begin{array}{c|c} N^2 & C_1 \\ \hline 0 & D_1^2 \end{array}\right), \end{aligned}$$

wobei $C_1 = N \cdot C + C \cdot D_1$. Der induktive Schritt für $n > 1$ ist jetzt auch klar.
Behauptung 2:

$$N^r = (0) = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix}.$$

Beweis. Wir werden noch mehr beweisen: nämlich für alle $s \in N$ sei $N^s = (n_{ij}^s)$. Dann ist $n_{ij}^s = 0$, falls $j < i + s$. Dazu benutzen wir Induktion über s. Falls $s = 1$, dann ist die Bedingung einfach die Beschreibung des Hauptmerkmals der Matrix N. Sei daher $s > 1$. Dann ist

$$n_{ij}^s = \sum_{k=1}^{r} n_{ik} n_{kj}^{s-1},$$

wobei $\begin{cases} n_{ik} = 0, & \text{für } k \leq i \text{ und} \\ n_{kj}^{s-1} = 0, & \text{für } j < k + s - 1. \end{cases}$

Sei $j < i + s$. Falls $k > i$, dann ist $k \geq i + 1$; d.h. $i \leq k - 1$ oder $j < i + s \leq k + s - 1$. D.h. $n_{kj}^{s-1} = 0$. Andererseits, wenn $k \leq i$, ist $n_{ik} = 0$. Folglich sind alle Terme in der Summe $\sum_{k=1}^{n} n_{ik} n_{kj}^{s-1}$ null.

Nun, $Rang(D_1) = Rang(D_1^s) = n - r$, für alle $s \in N$. Aber für $s > r$ gilt $Rang(B^s) = Rang(N^s) + Rang(D_1^s) = n - r$. Sei jetzt $\alpha_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \longleftarrow i$–te Stelle mit $i \leq r$. Dann ist $B^s \cdot \alpha_i = 0 = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{pmatrix}$. Aber $B = M_{\mathcal{B}}(f) - \lambda \cdot I_n$. D.h. $(f - \lambda \cdot id)^{(s)}(\alpha_i) = 0$. Folglich ist $\alpha_i \in H(f;\lambda)$ für alle $i \leq r$. D.h. $dim(H(f;\lambda)) \geq r$. Andererseits ist

$$dim(Im(f - \lambda \cdot id)^{(s)})) = Rang(B^{(s)}) = n - r.$$

Aber wir wissen, daß $dim(Ker(f - \lambda \cdot id)^{(s)}) \geq r$ und

$$dim(Im(f - \lambda \cdot id)^{(s)}) + dim(ker(f - \lambda \cdot id)^{(s)}) = n.$$

Folglich ist $dim(W) = r$ und (i) ist bewiesen.

Wir müssen auch zeigen, daß $f(W) \subset W$. Aber $W = Ker(f - \lambda \cdot id)^{(s)})$, für $s \geq r$. Insbesondere ist $W = Ker(f - \lambda \cdot id)^{(r+1)}$. D.h. $\forall \omega \in W$ gilt $(f - \lambda \cdot id)^{(r)}(\omega) = 0$.

$$\begin{aligned} \Rightarrow 0 &= (f - \lambda \cdot id)(0) = (f - \lambda \cdot id)(f - \lambda \cdot id)^{(r)}(\omega) \\ &= (f - \lambda \cdot id)^{(r)}(f - \lambda \cdot id)(\omega) \\ \Rightarrow (f - \lambda \cdot id)(\omega) &\in Ker((f - \lambda \cdot id)^{(r)} = W \\ \Rightarrow f(\omega) - \lambda\omega &\in W \end{aligned}$$

Aber $\lambda\omega \in W$, da $\omega \in W$. $\Rightarrow f(\omega) \in W$.

Nun sei

$$U = Im((f - \lambda \cdot id)^{(r)}) = \{\alpha \in V : \exists \beta \in V mit (f - \lambda \cdot id)^{(r)}(\beta) = \alpha\}.$$

Es gilt $dim(U) = Rang(B^r) = n - r$. Aber $U \cap W = \{0\}$, und wir haben $dim(W) = r$. Folglich ist $dim(U + W) = dim(U) + dim(W) = n = dim(V)$. $\Rightarrow V = W \oplus U$.

Schließlich müssen wir zeigen, daß $f(U) \subset U$. Aber sei $\alpha \in V$. Dann ist $\beta \equiv (f - \lambda \cdot id)^{(r)}(\alpha) \in U$, und jedes Element von U kann so beschrieben werden. Daher ist

$$(f - \lambda \cdot id)(\beta) = (f - \lambda \cdot id)^{(r)}((f - \lambda \cdot id)(\alpha)) \in U.$$

Es folgt $f(\beta) - \lambda\beta \in U$ und $f(\beta) \in U$. □

Korollar 22.2.1. *Sei $f_{|W} : W \to W$ die Einschränkung von f auf W. Es gilt $P_{f_{|W}}(x) = (\lambda - x)^{(r)}$ und $(f_{|W} - \lambda \cdot id)^{(r)} = 0$.*

Beweis. Im Beweis des Satzes ist gezeigt worden, daß $(f_{|W} - \lambda \cdot id)^{(r)}(\omega) = 0$ für alle $\omega \in W$. Folglich ist $(f_{|W} - \lambda \cdot id)^{(r)} = 0$. Auch

$$M_{\mathcal{B}}(f_{|W}) = X = \begin{pmatrix} \lambda & * & * \\ & \ddots & * \\ 0 & & \lambda \end{pmatrix} \Rightarrow P_{f_{|W}}(x) = (\lambda - x)^{(r)}.$$

□

Satz 22.3. *Sei $f : V \to V$ mit $P_f(x) = \prod_{i=1}^k (\lambda_i - x)^{(r_i)}$, wobei $r_i > 0$, $\lambda_i \neq \lambda_j$ für $i \neq j$ und $W_i = H(f; \lambda_i)$. Dann ist $V = W_1 \oplus \cdots \oplus W_k$ mit $dim(W_i) = r_i$ und $W_i = ker((f - \lambda_i \cdot id)^{(r_i)})$ und $f(W_i) \subset W_i$, für alle i.*

Beweis. Induktion über die Zahl k. Falls $k = 0$, dann ist $0 = grad(P_f) = dim(V)$. D.h. $V = \{0\}$ ist der triviale Raum. Sei daher $k \geq 1$. Nach unserem Satz 22.2 existieren Unterräume W_1, $U \subset V$ mit $V = W_1 \oplus U$ und $dim(W_1) = r_1$ und auch $f(W_1) \subset W_1$ und $f(U) \subset U$. Aber

$$P_f(x) = (\lambda_1 - x) \cdot \underbrace{Q(x)}_{\prod_{i=2}^n (\lambda_i - x)^{r_i}}$$

und nach der Darstellung von $M_{\mathcal{B}}(f)$ im Beweis des Satzes 22.2 ist klar, daß $P_{f_{|U}}(x) = Q(x)$. Nach der induktiven Hypothese ist dann $U = W_2 \oplus \cdots \oplus W_r$ mit $dim(W_i) = r_i$, $W_i = ker((f - \lambda_i \cdot id)^{(r_i)})$, $f(W_i) \subset W_i$, $\forall\, i \geq 2$. Dies alles gilt auch für W_1, außer vielleicht die Bedingung, daß $W_1 = ker((f - \lambda_1 \cdot id)^{(r_1)})$. Aber nach Satz 22.2 wissen wir, daß $W_1 = ker((f - \lambda_1 \cdot id)^{(s)})$ für alle $s \geq r_1$. Daher gilt auch $W_1 = ker((f - \lambda_1 \cdot id)^{(r_1)})$. □

Korollar 22.3.1. *Sei $A \in M(n \times n; F)$ so, daß P_A in lineare Faktoren zerfällt. ($P_A(x) = (\lambda_1 - x) \cdots (\lambda_n - x)$, wobei $\lambda_i \in F$ für alle i.) Dann existiert eine Matrix $S \in GL(n; F)$ mit $S^{-1} \cdot A \cdot S$ von der Gestalt*

$$S^{-1} \cdot A \cdot S = \begin{pmatrix} K_1 & & & \\ & K_2 & 0 & \\ & 0 & \ddots & \\ & & & K_r \end{pmatrix}$$

wobei jedes K_i eine Dreiecksmatrix von der Gestalt

$$K_i = \begin{pmatrix} \lambda_i & * & * \\ & \lambda_i & * \\ 0 & & \lambda_i \end{pmatrix}$$

ist.

Beweis. A ist die Darstellung einer linearen Abbildung $f : V \to V$ bzgl. einer vorgegebenen Basis $\mathcal{B}$. Aber nach dem Satz existiert eine 'bessere' Basis $\mathcal{B}'$ für V, die zusammengesetzt ist aus Basen für die jeweiligen Räume W_i. Dann ist die gesuchte Matrix S die Matrix des Basiswechsels $\mathcal{B} \to \mathcal{B}'$. □

Damit sind wir einen weiten Schritt vorwärts gekommen in unserem Programm für den Beweis des Satzes über die Jordansche Normalform. Offensichtlich brauchen wir nur noch die Haupträume zu untersuchen. Unser Problem jetzt ist zu zeigen, daß jeder Hauptraum eine Basis besitzt mit der Eigenschaft, daß die Matrix, die die lineare Abbildung darstellt bzgl. dieser Basis, eine Jordanmatrix ist.

Definition 79. Sei $A \in M(n \times n; F)$. Falls eine Zahl $p \in \mathbf{N}$ existiert mit

$$A^p = 0 \equiv \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$$

dann heißt A *nilpotent*. Sei auch $h : V \to V$ eine lineare Abbildung. h heißt *nilpotent*, falls ein $p \in \mathbf{N}$ existiert mit $h^{(p)} = \underbrace{h \cdot h \cdots h}_{p-mal} = 0$, die triviale Abbildung, die alles einfach auf den Nullvektor abbildet. D.h. $h^{(p)}(\alpha) = 0$, für alle $\alpha \in V$.

Offensichtlich gilt: h ist nilpotent $\Leftrightarrow$ $M_{\mathcal{B}}(h)$ ist nilpotent, für jede beliebige Basis $\mathcal{B}$ von V.

Satz 22.4. *Sei $h : W \to W$ eine nilpotente lineare Abbildung mit $dim(W) = r$. Sei $p \in \mathbf{N}$ die kleinste Zahl mit $h^{(p)} = 0$. Für alle $i = 1, \ldots, p$ sei $V_i = ker(h^{(i)}) = \{\omega \in W : h^{(i)}(\omega) = 0\}$. Dann gilt:*

1. *$\{0\} \equiv V_0 \subset V_1 \subset \cdots \subset V_p = W$.*
2. *$V_{i-1} \neq V_i$ und $h^{-1}(V_{i-1}) = V_i$, für alle i.*
3. *Sei $U \subset W$ ein linearer Unterraum mit $U \cap V_i = \{0\}$ für ein $i \geq 1$. Dann ist $h_{|U}$ injektiv.*

Beweis. 1. Die Gleichung $h(0) = 0$ ist wahr für jede lineare Abbildung. Sei nun angenommen, daß $h^{(i-1)}(\omega) = 0$ für ein i zwischen 1 und p und ein $\omega \in W$. Dann ist $h^{(i)}(\omega) = h(h^{(i-1)}(\omega)) = h(0) = 0$. Da h nilpotent ist, folgt daß $V_p = W$.

2. Es gilt

$$\begin{aligned} \alpha \in V_i &\Leftrightarrow \alpha \in ker(h^{(i)}) \\ &\Leftrightarrow h^{(i)}(\alpha) = 0 \\ &\Leftrightarrow h(\alpha) \in V_{i-1} \\ &\Leftrightarrow \alpha \in h^{-1}(V_{i-1}) \\ &\Rightarrow h^{-1}(V_{i-1}) = V_i \end{aligned}$$

   Sei nun angenommen, daß $V_{i-1} = V_i$ und sei $\omega \in W$ mit der Eigenschaft, daß zwar $h^{(p)}(\omega) = 0$, aber $h^{(p-1)}(\omega) \neq 0$. Aber $h^{(i)}(h^{(p-i)}(\omega)) = 0$, und daher $h^{(p-i)}(\omega) \in V_i = V_{i-1}$. D.h. $h^{(i-1)}(h^{(p-i)}(\omega)) = h^{(p-1)}(\omega) = 0$. Ein Widerspruch.

3. Es gilt $h_{|_U}$ ist injektiv $\Leftrightarrow ker(h_{|_U}) = \{0\}$. Sei daher $v \in ker(h_{|_U})$. Dann gilt: $h(v) = 0 \Rightarrow v \in ker(h^{(1)}) = V_1 \subset V_i$ für $i = 1, \ldots, p$. $\Rightarrow v \in U \cap V_i = \{0\} \Rightarrow ker(h_{|_U}) = \{0\}$.

□

Satz 22.5. *Seien $h : W \to W$, $p \in \mathbf{N}$ mit $h^{(p)} = 0$ und $V_0 \subset \cdots \subset V_p$ wie im Satz 22.4. Dann ist $W = U_1 \oplus \cdots \oplus U_p$, wobei U_i jeweils ein Unterraum ist mit $V_i = U_1 \oplus \cdots \oplus U_i$, für alle $i \geq 1$ und $h(U_i) \subset U_{i-1}$ und $h_{|_{U_i}}$ ist injektiv.*

Beweis. Induktion über p. Für $p = 1$ nehme $V_1 = U_1 = W$ und benutze Satz 22.4. Wir können daher annehmen, daß $p \geq 2$, und wir betrachten V_{p-1}. Mittels Basisergänzungssatz finden wir dann einen Unterraum U_p mit $W = U_p \oplus V_{p-1}$. Insbesondere ist $U_p \cap V_{p-1} = \{0\}$. Nach Satz 22.4 ist $h^{-1}(V_{p-2}) = V_{p-1}$. D.h. es existiert kein $u \in U_p$ mit $h(u) \in V_{p-2}$ und $u \neq 0$, oder $h(U_p) \cap V_{p-2} = \{0\}$. Wir benutzen wieder den Basisergänzungssatz, um ein $U_{p-1} \supset h(U_p)$ zu finden, und $V_{p-1} = U_{p-1} \oplus V_{p-2}$. Nach der induktiven Hypothese können wir schreiben $V_{p-1} = U_{p-1} \oplus \cdots \oplus U_1$. Dann ist $W = U_p \oplus U_{p-1} \oplus \cdots \oplus U_1$. Es gilt $h(U_p) \subset U_{p-1}$, und daß $h_{|_{U_{|p}}}$ injektiv ist, folgt aus Satz 22.4, da $U_p \cap V_{p-1} = \{0\}$. □

Satz 22.6. *Sei $h : W \to W$ nilpotent. $\Rightarrow$ es existiert eine Basis $\mathcal{B} = \{\omega_1, \ldots, \omega_r\}$ von W mit entweder $h(\omega_i) = \omega_{i-1}$ oder $h(\omega_i) = 0$ für alle i.*

Beweis. Nach Satz 22.5 ist $W = U_1 \oplus \cdots \oplus U_p$. Sei $i > 1$ und sei etwa $\{\omega_1, \ldots, \omega_q\}$ eine Basis für U_i. Da $h_{|_{U_i}}$ injektiv ist (wobei $h : U_i \to U_{i-1}$), sind die Vektoren $h(\omega_j)$, $j = 1, \ldots, q$ linear unabhängig in U_{i-1}. Wir können jetzt den Basisergänzungssatz benutzen, um weitere Vektoren, etwa $\omega_{q+1}, \ldots, \omega_t$ zu finden, mit $\{h(\omega_1), \ldots, h(\omega_q), \omega_{q+1}, \ldots, \omega_t\}$ eine Basis für U_{i-1}. D.h. es existiert ein System von Basisvektoren für die verschiedenden U_i der Art:

$$\begin{aligned} \{v_1^{(p)}, \ldots, v_{l_p}^{(p)}\} &\to U_p \\ \{h(v_1^{(p)}), \ldots, h(v_{l_p}^{(p)}), v_1^{(p-1)}, \ldots, v_{l_{p-1}}^{(p-1)}\} &\to U_{p-1} \\ &\vdots \\ \{h^{(p-1)}(v_1^{(p)}), \ldots, h^{(p-1)}(v_{l_p}^{(p)}), h^{(p-2)}(v_1^{(p-1)}), \ldots, v_1^{(1)}, \ldots, v_{l_1}^{(1)}\} &\to U_1 \end{aligned}$$

Die Basis $\mathcal{B} = \{\omega_1, \ldots, \omega_r\}$ wird so definiert:

$$\begin{aligned} \omega_1 &= v_1^{(p)} \\ \omega_2 &= h(\omega_1) \\ &\vdots \\ \omega_p &= h^{(p-1)}(\omega_1) \\ \omega_{p+1} &= v_2^{(p)} \\ &\vdots \\ \omega_{2p} &= h^{(p-1)}(\omega_{p+1}) \\ &\vdots \\ &\text{u.s.w.} \end{aligned}$$

□

Satz 22.7. *Sei die Matrix $A \in M(r \times r; F)$ in oberer Dreiecksgestalt mit* $A = \begin{pmatrix} \lambda & * & * \\ & \ddots & * \\ 0 & & \lambda \end{pmatrix}$. *Dann gibt es eine Matrix $S \in GL(r; F)$ mit $S^{-1} \cdot A \cdot S$ in Jordan Normalform.*

Beweis. Sei $A = \lambda \cdot I + N$, wobei N eine nilpotente Matrix ist. Nach Satz 22.6 muß eine Matrix S existieren, mit

$$S^{-1} \cdot N \cdot S = \begin{pmatrix} 0 & \epsilon_1 & 0 & \cdots & 0 \\ & 0 & \epsilon_2 & 0 & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ & & & 0 & \epsilon_{r-1} \\ 0 & & \cdots & & 0 \end{pmatrix}$$

wobei $\epsilon_i = 0$ oder 1 für $i = 1, \ldots, r-1$. Aber

$$\begin{aligned} S^{-1} \cdot A \cdot S &= S^{-1}(\lambda \cdot I + N)S \\ &= S^{-1}(\lambda \cdot I)S + S^{-1}NS \\ &= \lambda \cdot I + S^{-1}NS \quad \longrightarrow \quad \text{Jordan Normalform.} \end{aligned}$$

□

Beweis des Satzes über Jordan Normalformen

folgt, da die Matrix S in Satz 22.7 uns eine geeignete Basis in einem vorgegebenen Hauptraum gibt. Wir brauchen daher nur die entsprechenden Basen in den jeweiligen Haupträumen zusammen zu nehmen, um eine Basis $\mathcal{B}$ für den gesamten Raum V zu erhalten mit der Eigenschaft, daß die darstellende Matrix $M_{\mathcal{B}}(f)$ in Jordan Normalform ist.