Lineare Algebra II

Prof. Dr. Peter Müller

Inhaltsverzeichnis

11 Polynome über Körper

11.1 Definition

Sei K ein Körper. Dann ist ein Polynom in der Variablen x eine formale Summe $f = \sum_{i=0}^{n} a_i x^i$ mit $n \in \mathbb{N}$ und $a_i \in K$ für $1 \leq i \leq n$, wobei $x^0 = 1$ gilt. Die Elemente a_i werden Koeffizienten genannt. Für $a_n \neq 0$ heißt a_n Leitkoeffizient und es ist $\operatorname{grad} f = n$ der Grad des Polynoms. Dabei setzt man $\operatorname{grad} 0 = -\infty$. Das Polynom f heißt normiert, falls $a_n = 1$ gilt.

Die Menge aller Polynome in x über K bezeichnet man mit $K[x]$. Auf dieser Menge $K[x]$ wird dann die Addition koeffizientenweise durch

$$\sum_{i=0}^{n} a_i x^i + \sum_{i=0}^{m} b_i x^i = \sum_{i=0}^{n} (a_i + b_i) x^i + \sum_{i=n+1}^{m} b_i x^i$$

für $n \leq m$ definiert. Die Multiplikation ist durch $ax = xa$ für $a \in K$, $x^i x^j = x^{i+j}$ für $i, j \in \mathbb{N}$ und das Distributivgesetz festgelegt.

Beispiel

Es gilt $(a_0 + a_1 x + a_2 x^2)(b_0 + b_1 x) = a_0(b_0 + b_1 x) + a_1 x(b_0 + b_1 x) + a_2 x^2(b_0 + b_1 x) = a_0 b_0 + a_0 b_1 x + a_1 x b_0 + a_1 x b_1 x + a_2 x^2 b_0 + a_2 x^2 b_1 x = a_0 b_0 + a_0 b_1 x + a_1 b_0 x + a_1 b_1 x x + a_2 b_0 x^2 + a_2 b_1 x^2 x = a_0 b_0 + a_0 b_1 x + a_1 b_0 x + a_1 b_1 x^2 + a_2 b_0 x^2 + a_2 b_1 x^2 = a_0 b_0 + (a_0 b_1 + a_1 b_0) x + (a_1 b_1 + a_2 b_0) x^2 + a_2 b_1 x^3$.

Bemerkung

Sei K ein Körper. Dann bildet $K[x]$ einen kommutativen Ring und einen Vektorraum.

Beweis

Da die Addition auf $K[x]$ koeffizientenweise definiert ist, bildet $K[x]$ eine abelsche Gruppe mit neutralem Element 0.

Aus der Definition der Multiplikation folgt unmittelbar die Existenz des neutralen Elements 1, die Kommutativität und die Distributivität. Schließlich folgt die Assoziativität aus

$$(\sum_{i=1}^{n} a_i x^i)(\sum_{j=1}^{m} b_j x^j)(\sum_{k=1}^{p} c_k x^k) = \sum_{i=1}^{n} \sum_{j=1}^{m} \sum_{k=1}^{p} a_i b_j c_k x^{i+j+k}.$$

Da die Abbildung $a \in K \longmapsto ax^0 \in K[x]$ injektiv ist, wird K als Teilmenge von $K[x]$ betrachtet. Daher fasst man das Produkt $a \cdot f$ für $a = a \cdot x^0 \in K[x]$ und $f \in K[x]$ als

Skalarmultiplikation auf. Die Vektorraumaxiome folgen dann aus den Ringeigenschaften von $K[x]$.

11.2 Satz

Seien $f, g \in K[x]$. Dann gilt $\operatorname{grad}(f+g) \leq \max\{\operatorname{grad} f, \operatorname{grad} g\}$ und $\operatorname{grad} fg = \operatorname{grad} f + \operatorname{grad} g$.

Beweis

Sei o.B.d.A $f = a_0 + \ldots + a_m x^m$ und $g = b_0 + \ldots + b_n x^n$ mit $a_m \neq 0$ und $b_n \neq o$ sowie $m \geq n$.

Für $m > n$ ist dann $f + g = (a_0 + b_0) + \ldots + (a_n + b_n)x^n + a_{n+1}x^{n+1} + \ldots + a_m x^m$, d.h. $\operatorname{grad}(f+g) = m$. Für $m = n$ gilt $f + g = (a_0 + b_0) + \ldots + (a_n + b_n)x^n$, d.h. mit $a_n + b_n \neq 0$ folgt $\operatorname{grad}(f+g) = n$ und mit $a_n + b_n = 0$ folgt $\operatorname{grad}(f+g) < n$.

Weiter ist $fg = a_0b_0 + (a_1b_0 + a_0b_1)x + \ldots + a_nb_m x^{n+m}$ mit $a_nb_m \neq 0$. Daher gilt $\operatorname{grad} fg = n+m$.

Korollar

1. $K[x]$ ist nullteilerfrei.
2. Aus $fg = hg$ für $f, g, h \in K[x]$ und $g \neq 0$ folgt stets $f = h$.

Beweis

1 Sei $f, g \in K[x]$ mit $fg = 0$ und $f \neq 0$. Dann gilt $\operatorname{grad} f \geq 0$ und daher $\operatorname{grad} g \leq \operatorname{grad} g + \operatorname{grad} f = \operatorname{grad} gf < 0 \Longrightarrow g = 0$.

2 Sei nun $fg = hg$ für $f, g, h \in K[x]$ und $g \neq 0$. Dann gilt $(f-h)g = fg - gh = 0 \Longrightarrow f - h = 0 \Longrightarrow f = h$.

Bemerkung

Früher fasste man Polynome $f \in K[x]$ für $K = \mathbb{R}$ als Abbildungen $K \longrightarrow K$ auf. Dies ist halbwegs gerechtfertigt, da $\varphi : f \in K[x] \longmapsto \big(x \longmapsto f(x)\big) \in \operatorname{Abb}(K, K)$ ein Ringhomomorphismus ist. Für $|K| < \infty$ ist diese Abbildung jedoch nicht injektiv. Dazu betrachte man etwa $K = \mathbb{F}_2$. Dann gilt $0 \neq x^2 + x \in \operatorname{Kern}\varphi$, da $\varphi(x^2+x)(0) = 0 + 0 = 0$ und $\varphi(x^2+x)(1) = 1 + 1 = 0$.

11.3 Satz (Division mit Rest)

Seien $f, g \in K[x]$ mit $g \neq 0$. Dann gibt es eindeutige Polynome $q, r \in K[x]$ mit $f = qg + r$ und $\operatorname{grad} r < \operatorname{grad} g$.

Beweis

Zunächst untersuchen wir die Eindeutigkeit. Sei dazu $f = qg + r$ und $f = \widetilde{q}g + \widetilde{r}$ mit $\operatorname{grad} r < \operatorname{grad} g$ und $\operatorname{grad} \widetilde{r} < \operatorname{grad} g$. Es folgt $r - \widetilde{r} = (f - qg) - (f - \widetilde{q}g) = \widetilde{q}g - qg = g(\widetilde{q} - q)$. Damit gilt nun $\operatorname{grad} g > \max\{\operatorname{grad} r, \operatorname{grad} \widetilde{r}\} \geq \operatorname{grad}(r - \widetilde{r}) = \operatorname{grad} g(\widetilde{q} - q) = \operatorname{grad} g + \operatorname{grad}(\widetilde{q} - q) \Longrightarrow \operatorname{grad}(\widetilde{q} - q) < 0 \Longrightarrow \widetilde{q} - q = 0 \Longrightarrow \widetilde{q} = q$. Schließlich ist auch $\widetilde{r} = f - \widetilde{q}g = f - qg = r$.

Nun sei $g = b_m x^m + \ldots + b_0$ mit $b_m \neq 0$, d.h. $\operatorname{grad} g = m$. Man setze $\widetilde{g} = b_m^{-1} g$, d.h. $\widetilde{g}$ ist normiert mit $\operatorname{grad} \widetilde{g} = m = \operatorname{grad} g$. Sei weiter $f = \widetilde{q}\widetilde{g} + r$ für $\operatorname{grad} r < \operatorname{grad} \widetilde{g}$ eine Divsion mit Rest. Dann ist auch $f = \widetilde{q}\widetilde{g} + r = \widetilde{q}(b_m^{-1} g) + r = (\widetilde{q} b_m^{-1})g + r$ wegen $\operatorname{grad} r < \operatorname{grad} \widetilde{g} = \operatorname{grad} g$ eine Division mit Rest. Daher können wir im Folgenden o.E. annehmen, $g \in K[x]$ ist normiert.

Für $\operatorname{grad} f < \operatorname{grad} g$ ist $f = 0 \cdot g + f$ eine Division der verlangten Art. Sei also $\operatorname{grad} f \geq \operatorname{grad} g$. Dann beweisen wir die Aussage durch vollständige Induktion über $n = \operatorname{grad} f$. Für $n = 0$ gilt wegen $g \neq 0$ und $\operatorname{grad} f \geq \operatorname{grad} g$ stets $0 \leq \operatorname{grad} g \leq \operatorname{grad} f = n = 0$, d.h. $\operatorname{grad} g = 0$. Da g normiert ist, folgt $g = 1$. Also ist $f = fg + 0$ wegen $\operatorname{grad} 0 = -\infty < 0 = \operatorname{grad} g$ eine Division mit Rest.

Sei nun a der Leitkoeffizient von f. Wegen $\operatorname{grad} f \geq \operatorname{grad} g$ ist $x^{\operatorname{grad} f - \operatorname{grad} g} \in K[x]$ wohldefiniert. Man setze $\widetilde{f} = f - a x^{\operatorname{grad} f - \operatorname{grad} g} g$. Da g und $x^{\operatorname{grad} f - \operatorname{grad} g}$ normiert sind, ist $-a$ der Leitkoeffizient von $-a x^{\operatorname{grad} f - \operatorname{grad} g} g$. Wegen $\operatorname{grad}(-a x^{\operatorname{grad} f - \operatorname{grad} g} g) = \operatorname{grad}(-a) + \operatorname{grad} x^{\operatorname{grad} f - \operatorname{grad} g} + \operatorname{grad} g = 0 + \operatorname{grad} f - \operatorname{grad} g + \operatorname{grad} g = \operatorname{grad} f$ gilt daher $\operatorname{grad} \widetilde{f} < n$.

Nach Induktionsannahme ist dann $\widetilde{f} = qg + r$ für $\operatorname{grad} r < \operatorname{grad} g$. Es folgt $f = \widetilde{f} + a x^{\operatorname{grad} f - \operatorname{grad} g} g = qg + r + a x^{\operatorname{grad} f - \operatorname{grad} g} g = (q + a x^{\operatorname{grad} f - \operatorname{grad} g})g + r$.

Bemerkung

Der Beweis von Satz 11.3 ist die algorithmische Beschreibung der Polynomdivision. Sei etwa $f = x^6 + 2x^4 + x^3 + x^2 + 3x + 4$ und $g = x^2 + 1$. Dann ist $a = 1$ der Leitkoeffizient von f und man setzt $\widetilde{f} = f - a x^{\operatorname{grad} f - \operatorname{grad} g} g = f - x^4 g = x^4 + x^3 + x^2 + 3x + 4$. Mit $\widetilde{f} = qg + r = g(x^2 + x) + (2x + 4)$ erhält man $f = (q + a x^{\operatorname{grad} f - \operatorname{grad} g})g + r = (x^4 + x^2 + x)g + (2x + 4)$.

11.4 Definition / Satz

Sei R ein Ring und $I \subseteq R$ ein Ideal. Dann heißt I Hauptideal, wenn ein $x \in I$ mit $I = xR = (x)$ existiert. Wenn jedes Ideal $I \subseteq R$ ein Hauptideal ist, wird R Hauptidealring genannt.

$K[x]$ ist ein Hauptidealring.

Beweis

Sei $I \subseteq K[x]$ ein Ideal. Offenbar gilt $I = (0)$ für $I = \{0\}$. Sei also $I \neq \{0\}$. Dann existiert ein normiertes Polynom $g \in I$ mit minimalem Grad. Insbesondere gilt $g \neq 0$. Somit gibt es für $f \in I$ nach Satz 11.3 eine Darstellung $f = qg + r$ mit $q, r \in K[x]$ und $\operatorname{grad} r < \operatorname{grad} g$.

Da nun I ein Ideal ist, gilt $qg \in I$. Es folgt $r = f - qg \in I$. Sei dann a der Leitkoeffizient von r. Für $a \neq 0$ ist $a^{-1}r$ normiert und man erhält wegen $\operatorname{grad} a^{-1}r = \operatorname{grad} r < \operatorname{grad} g$ einen Widerspruch. Insgesamt gilt also $r = 0 \Longrightarrow f = gq \Longrightarrow I = (g)$.

Bemerkung

Sei $I \subseteq K[x]$ ein Ideal für $I = (g)$. Dann ist $g \in I$ bis auf einen Faktor $a \in K \setminus \{0\}$ eindeutig bestimmt.

Beweis

Für $I = \{0\}$ ist das erzeugende Element offenbar eindeutig. Sei also $(h) = I = (g)$ für $I \neq \{0\}$, d.h. $h, g \neq 0$. Dann gilt $h = qg$ und $g = rh$ für $0 \neq q, r \in K[x]$. Es folgt $h = (qr)h$ und daher $\operatorname{grad} h = \operatorname{grad} qr + \operatorname{grad} h \Longrightarrow \operatorname{grad} q + \operatorname{grad} r = \operatorname{grad} qr = 0$. Wegen $0 \leq \operatorname{grad} q$ und $0 \leq \operatorname{grad} r$ folgt damit $\operatorname{grad} q = 0 = \operatorname{grad} r$, also $q, r \in K$.

11.5 Satz

Sei $a \in K$ eine Nullstelle von $f \in K[x]$, d.h. $f(a) = 0$. Dann gibt es ein $g \in K[x]$ mit $f = (x - a)g$.

Beweis

Sei $f = g(x - a) + r$ mit $\operatorname{grad} r < \operatorname{grad}(x - a)$ eine Division mit Rest. Wegen $\operatorname{grad}(x - a) = 1$ folgt $\operatorname{grad} r \leq 0 \Longrightarrow r \in K$. Damit gilt $0 = f(a) = g(a - a) + r = 0 + r = r$.

11.6 Definition / Bemerkung

Sei $0 \neq f = (x-a)^{\widetilde{e}} g \in K[x]$ ein Polynom mit Nullstelle $a \in K$ für $0 < \widetilde{e} \in \mathbb{N}$. Dann gilt insbesondere $g \neq 0$ und es folgt $\operatorname{grad} f = \operatorname{grad}(x-a)^{\widetilde{e}} + \operatorname{grad} g = \widetilde{e} + \operatorname{grad} g \geq \widetilde{e}$. Da die Menge $M = \{\widetilde{e} \in \mathbb{N}^+ \mid f = (x-a)^{\widetilde{e}} g \text{ für ein } g \in K[x]\}$ somit nach oben beschränkt ist, existiert $e = \max M$. Man bezeichnet e als algebraische Vielfachheit der Nullstelle a.

Bemerkung

Sei $0 \neq f = (x-a)^e g \in K[x]$. Dann hat die Nullstelle $a \in K$ genau dann die Vielfachheit e, wenn $g(a) \neq 0$ gilt.

Beweis

Sei $g(a) = 0$. Nach Satz 11.5 ist dann $g = (x-a)h$ für ein $h \in K[x]$. Damit ist e wegen $f = (x-a)^{e+1}h$ nicht die Vielfachheit von a. Sei nun $k > e$ die Vielfachheit von a. Dann ist $(x-a)^e g = f = (x-a)^k h$ für $h \in K[x]$, d.h. $g = (x-a)^{k-e}h$ mit $k - e > 0$. Somit gilt $g(a) = 0$.

11.7 Satz

Sei ein Polynom $f \in K[x]$ mit $\operatorname{grad} f = n \geq 0$ gegeben. Seien weiter $a_1, \ldots, a_m$ verschiedene Nullstellen von f mit Vielfachheiten $e_1, \ldots, e_m$. Dann gibt es ein $\overline{f} \in K[x]$ mit $f = (x - a_1)^{e_1} \cdots (x - a_m)^{e_m} \overline{f}$. Insbesondere gilt $e_1 + \ldots + e_m \leq n$.

Beweis

Sei zunächst $f = (x-a_1)^{e_1} \cdots (x-a_m)^{e_m} \overline{f}$ für $\overline{f} \in K[x]$. Wegen $\operatorname{grad} f \geq 0 \Longrightarrow f \neq 0$ gilt $\overline{f} \neq 0 \Longrightarrow \operatorname{grad} \overline{f} \geq 0$. Damit folgt $n = \operatorname{grad} f = e_1 + \ldots + e_m + \operatorname{grad} \overline{f} \geq e_1 + \ldots + e_m$.

Nun beweisen wir die erste Aussage durch vollständige Induktion über $\operatorname{grad} f = n$. Der Fall $n = 0$ ist trivial, da dann f keine Nullstellen hat. Sei nun $n \geq 1$ und a_1 eine Nullstelle von f der Vielfachheit e_1. Dann ist $f = (x-a_1)^{e_1} g$ für ein $g \in K[x]$.

Sei weiter $k_i \in \mathbb{N}$ die Vielfachheit der Nullstelle a_i von g. Für $g(a_i) \neq 0$ setze man $k_i = 0$. Dann existiert ein $h_i \in K[x]$ mit $g = (x-a_i)^{k_i} h_i$ und $h_i(a_i) \neq 0$. Mit $\overline{h}_i = h_i (x-a_1)^{e_1}$ erhält man $f = (x-a_i)^{k_i} \overline{h}_i$. Dabei gilt offenbar $\overline{h}_i(a_i) \neq 0 \Longrightarrow k_i = e_i$. Insgesamt hat also g die Nullstellen $a_2, \ldots, a_m$ mit Vielfacheiten $e_2, \ldots, e_m$.

Wegen $\operatorname{grad} g < \operatorname{grad} f$ ergibt sich damit aus der Induktionsannahme $g = (x-a_2)^{e_2}\cdots(x-a_m)^{e_m}\overline{f}$. Schließlich folgt $f = (x-a_1)^{e_1}g = (x-a_1)^{e_1}\cdots(x-a_m)^{e_m}\overline{f}$.

Bemerkung

Satz 11.7 gilt i.A. für Polynome über Ringen nicht. Etwa hat mit $R = \mathbb{Z}\,|_{8\mathbb{Z}}$ das Polynom $x^3 = f \in R[x]$ die Nullstellen $a_1 = 0$, $a_2 = 2$, $a_3 = 4$ und $a_4 = 6$. Für die entsprechenden Vielfachheiten folgt dann $e_1 + e_2 + e_3 + e_4 \geq 4 > \operatorname{grad} f$.

11.8 Definition

Sei $f \in K[x]$ definiert durch $f = \sum_{i=0}^{n} a_i x^i$. Dann ist $f' = \sum_{i=1}^{n} i a_i x^{i-1} \in K[x]$ die Ableitung von f.

Bemerkung

Seien $f, g \in K[x]$. Dann gilt $(f+g)' = f' + g'$ und $(fg)' = f'g + fg'$.

Beweis

Sei o.B.d.A. $f = \sum_{i=0}^{n} a_i x^i$ und $g = \sum_{i=0}^{m} b_i x^i$ mit $m \leq n$. Dann gilt $f + g = (a_0 + b_0) + \ldots + (a_m + b_m)x^m + a_{m+1}x^{m+1} + \ldots + a_n x^n$. Es folgt $f' + g' = a_1 + \ldots + a_n n x^{n-1} + b_1 + \ldots + b_m m x^{m-1} = (a_1 + b_1) + \ldots + (a_m + b_m) m x^{m-1} + a_{m+1}(m+1)x^m + \ldots + a_n n x^{n-1} = (f+g)'$.

Nun untersuchen wir $f b_r x^r$ für $b_r \in K$. Es gilt

$$\begin{aligned}
(f b_r x^r)' &= (a_0 b_r x^r + a_1 b_r x^{r+1} + \ldots + a_n b_r x^{n+r})' \\
&= a_0 b_r r x^{r-1} + a_1 b_r (r+1) x^r \ldots + a_n b_r (n+r) x^{n+r-1} \\
&= a_0 b_r r x^{r-1} + a_1 b_r r x^r + \ldots + a_n b_r r x^{n+r-1} + a_1 b_r x^r + \ldots + a_n b_r n x^{n+r-1} \\
&= (a_0 + a_1 x + \ldots + a_n x^n) b_r r x^{r-1} + (a_1 + \ldots + a_n n x^{n-1}) b_r x^r \\
&= f(b_r x^r)' + f'(b_r x^r)
\end{aligned}$$

Damit folgt $(fg)' = (f \sum_{i=0}^{m} b_i x^i)' = (\sum_{i=0}^{m} f b_i x^i)' = \sum_{i=0}^{m} (f b_i x^i)' = \sum_{i=0}^{m} \big(f(b_i x^i)' + f'(b_i x^i)\big) = \sum_{i=0}^{m} f(b_i x^i)' + \sum_{i=0}^{m} f'(b_i x^i) = f \sum_{i=0}^{m} (b_i x^i)' + f' \sum_{i=0}^{m} b_i x^i = fg' + f'g$.

Bemerkung

Sei $f \in K[x]$. Dann folgt aus $f' = 0$ nicht notwendig $f \in K$. Für $K = \mathbb{F}_2$ und $f = x^2$ gilt etwa $f' = 2x = 0x = 0$.

11.9 Satz

Sei $a \in K$ eine Nullstelle des Polynoms $f \in K[x]$. Dann ist a genau dann eine einfache Nullstelle, wenn $f'(a) \neq 0$ gilt.

Beweis

Sei $f = (x - a)g$ für $g \in K[x]$. Dann ist a genau dann eine einfache Nullstelle, wenn $g(a) \neq 0$ gilt. Die Behauptung folgt nun wegen $f' = (x-a)'g+(x-a)g' = g+(x-a)g'$, also $f'(a) = g(a)$.

11.10 Definition

Ein Polynom $f \in K[x]$ mit $\operatorname{grad} f \geq 1$ heißt irreduzibel, wenn es keine Zerlegung $f = gh$ mit $\operatorname{grad} g < \operatorname{grad} f$ und $\operatorname{grad} h < \operatorname{grad} f$ gibt.

Bemerkung

Ein Polynom $f \in K[x]$ mit $\operatorname{grad} f \geq 1$ ist genau dann irreduzibel, wenn es keine Zerlegung $f = gh$ mit $g, h \in K[x] \setminus K$ gibt.

Beweis

Sei f reduzibel. Dann existiert eine Zerlegung $f = gh$ mit $\operatorname{grad} g < \operatorname{grad} f$ und $\operatorname{grad} h < \operatorname{grad} f$. Wegen $\operatorname{grad} f = \operatorname{grad} g + \operatorname{grad} h$ gilt $\operatorname{grad} g = \operatorname{grad} f - \operatorname{grad} h > 0 \Longrightarrow g \notin K$. Analog folgt $h \notin K$.

Sei nun $f = gh$ eine Zerlegung mit $g, h \notin K$. Insbesondere gilt dann $\operatorname{grad} g > 0$ und $\operatorname{grad} h > 0$. Daher folgt $\operatorname{grad} f = \operatorname{grad} g + \operatorname{grad} h > \operatorname{grad} h$ und analog $\operatorname{grad} f > \operatorname{grad} g$, d.h. f ist reduzibel.

Beispiel

1. Lineare Polynome $ax + b$ für $a \neq 0$ sind stets irreduzibel.

2. Das Polynom $x^2 + 1$ ist in $\mathbb{R}[x]$ irreduzibel und in $\mathbb{C}[x]$ wegen $x^2 + 1 = (x - i)(x + i)$ reduzibel.

11.11 Definition

Seien $f, g \in K[x]$ mit $g \neq 0$. Dann heißt g ein Teiler von f, wenn ein $h \in K[x]$ mit $f = gh$ existiert. Man schreibt dafür g/f.

11.12 Satz

Sei $f \in K[x]$ mit $\operatorname{grad} f \geq 1$. Dann ist

1. f ist irreduzibel
2. für alle $a, b \in K[x]$ folgt aus f/ab stets f/a oder f/b

äquivalent.

Beweis

1 $\Longrightarrow$ **2** Wir zeigen die Aussage durch Widerspruch und wählen dazu ein Gegenbeispiel mit minimalem Grad. Sei also f irreduzibel und $a, b \in K[x]$ mit f/ab sowie $f \nmid a$ und $f \nmid b$. Es folgt insbesondere $f/ab \Longrightarrow fg = ab$ für ein $g \in K[x]$ sowie $f \nmid a \Longrightarrow a \neq 0$ und $f \nmid b \Longrightarrow b \neq 0$.

Sei zunächst $(*)$ $f = aq + r$ mit $\operatorname{grad} r < \operatorname{grad} a$ für $\operatorname{grad} f \geq \operatorname{grad} a$ eine Division mit Rest. Wegen $\operatorname{grad} r < \operatorname{grad} f$ gilt dabei $q \neq 0$. Mit $rb = (f - aq)b = fb - abq = fb - fgq = f(b - gq)$ gilt weiter f/rb. Aus $b \neq 0$ und $\operatorname{grad} r < \operatorname{grad} a$ folgt dann $\operatorname{grad} rb < \operatorname{grad} ab$.

Da ab ein Gegenbeispiel minimalen Grades ist, folgt mit f/rb und $f \nmid b$ nun f/r. Wegen $q \neq 0$ gilt weiter $\operatorname{grad} r < \operatorname{grad} a \leq \operatorname{grad} aq = \operatorname{grad}(f - r) \leq \max\{\operatorname{grad} f, \operatorname{grad} r\} \Longrightarrow \max\{\operatorname{grad} f, \operatorname{grad} r\} = \operatorname{grad} f \Longrightarrow \operatorname{grad} r < \operatorname{grad} f$. Aus f/r erhält man daher $r = 0 \Longrightarrow f = aq$.

Da f aber irreduzibel ist, folgt $a \in K \Longrightarrow$ oder $q \in K$. Für $a \in K$ erhält man den Widerspruch $f/ab \Longrightarrow f/b$ und für $q \in K$ erhält man den Widerspruch $f = aq \Longrightarrow fq^{-1} = a \Longrightarrow f/a$.

Sei nun $a = fq + r$ mit $\operatorname{grad} r < \operatorname{grad} f$ für $\operatorname{grad} a > \operatorname{grad} f$. Analog zu $(*)$ erhält man $r = 0 \Longrightarrow a = fq \Longrightarrow f/a$ – ein Widerspruch.

2 $\Longrightarrow$ **1** Sei f reduzibel, also $f = ab$ mit $\operatorname{grad} a < \operatorname{grad} f$ und $\operatorname{grad} b < \operatorname{grad} f$. Wegen $\operatorname{grad} f \geq 1$ gilt $f \neq 0$, also $a \neq 0$ und $b \neq 0$. Insgesamt gilt dann $f \nmid a$ und $f \nmid b$.

11.13 Korollar

Jedes normierte Polynom $f \in K[x]$ mit $\operatorname{grad} f \geq 1$ hat eine eindeutige Darstellung als Produkt irreduzibler, normierter Faktoren.

Beweis

[Existenz] Sei zunächst f irreduzibel. Dann ist f selbst ein Produkt irreduzibler, normierter Polynome. Sei nun f reduzibel, d.h. $f = gh$ für $g, h \in K[x]$. Durch a bzw. b seien die Leitkoeffizienten von g bzw. h gegeben. Wegen $f \neq 0$ gilt dabei $a \neq 0$ und $b \neq 0$. Dann sind $a^{-1}g \in K[x]$ bzw. $b^{-1}h \in K[x]$ normiert. Man beginne für g und h erneut.

Mit jeder Anwendung der Zerlegung erhält man eine Darstellung $f = g_1 \cdots g_n$. Wegen $\operatorname{grad} g_i \geq 1$ für $1 \leq i \leq n$ folgt $n \leq \operatorname{grad} g_1 + \ldots + \operatorname{grad} g_n = \operatorname{grad} f$. Daher terminiert dieser Algorithmus.

[Eindeutigkeit] Wir beweisen durch Widerspruch. Sei dazu $f \in K[x]$ ein Gegenbeispiel mit minimalem Grad, d.h. $f = p_1 \cdots p_r = q_1 \cdots q_s$ für $p_1, \ldots, p_r, q_1, \ldots, q_s$ irreduzibel und normiert. Aus Satz 11.12 folgt wegen $p_1 / p_1 \cdots p_r = q_1 \cdots q_s$ zunächst p_1 / q_i für ein $1 \leq i \leq s$. Daher existiert ein $h \in K[x]$ mit $q_i = p_1 h$. Da nun q_i irreduzibel ist, erhält man $h \in K$. Da weiter q_i normiert ist, gilt $h = 1$. Insgesamt gilt also $p_1 = q_i$.

Es folgt $\overline{f} = p_2 \cdots p_r = q_1 \cdots q_{i-1} q_{i+1} \cdots q_s$. Da nun f eine Gegenbeispiel minimalen Grades war, erhält man wegen $\operatorname{grad} \overline{f} < \operatorname{grad} \overline{f} + \operatorname{grad} p_1 = \operatorname{grad} f$ die Identität $\{p_2, \ldots, p_r\} = \{q_1, \ldots, q_{i-1}, q_{i+1}, \ldots, q_s\}$. Mit $p_1 = q_i$ folgt $\{p_1, \ldots, p_r\} = \{q_1, \ldots, q_s\}$ – ein Widerspruch.

11.14 Bemerkung / Definition

Seien $\alpha, \beta \in K$ die Leitkoeffizienten von $f, g \in K[x] \setminus K$. Dann sind $\overline{f} = \alpha^{-1} f$ und $\overline{g} = \beta^{-1} g$ normiert. Nach Korollar 11.13 ist nun die Menge $\{p_1, \ldots, p_r\}$ aller normierter, irreduzibler Teiler von $\overline{f}$ oder $\overline{g}$ eindeutig bestimmt. Es folgt $f = \alpha\, p_1^{e_1} \cdots p_r^{e_r}$ und $g = \beta\, p_1^{k_1} \cdots p_r^{k_r}$ mit Potenzen $e_i, k_i \in \mathbb{N}$. Dabei setze man $e_i = 0$ bzw. $k_i = 0$, falls p_i kein Teiler von f bzw. g ist. Dann ist

$$\operatorname{ggT}(f, g) = \prod_{i=1}^{r} p_i^{\min(e_i, f_i)}$$

der größte gemeinsame Teiler und

$$\operatorname{kgV}(f, g) = \prod_{i=1}^{r} p_i^{\max(e_i, f_i)}$$

das kleinste gemeinsame Vielfache von f und g. Dabei ist $\operatorname{ggT}(f, g)$ bzw. $\operatorname{kgV}(f, g)$ als Produkt normierter Faktoren selbst normiert. Für $h = 0$ setzt man $\operatorname{ggT}(f, h) = \overline{f}$ und für $h \in K \setminus \{0\}$ ist $\operatorname{ggT}(f, h) = 1$.

Beweis

Offenbar ist $\mathrm{ggT}(f,g)$ ein Teiler von f und g. Sei nun $t \in K[x]$ ein weiterer gemeinsamer Teiler von f und g. Sei dann γ der Leitkoeffizient von t und $t = \gamma\, t_1^{l_1} \cdots t_s^{l_s}$ die eindeutige Darstellung als Produkt irreduzibler, normierter Faktoren. Es folgt $t_i^{l_i}/f$ und $t_i^{l_i}/g$ für $1 \leq i \leq s$, d.h. $t_i = p_j$ mit $l_i \leq e_j$ und $l_i \leq k_j$ für ein $1 \leq j \leq r$. Damit gilt $t/\mathrm{ggT}(f,g)$.

Analog ist $\mathrm{kgV}(f,g)$ ein Vielfaches von f und g mit $f/t \wedge g/t \Longrightarrow \mathrm{kgV}(f,g)/t$.

11.15 Definition

1. Sei $p \in K[x]$ irreduzibel und normiert. Dann nennt man p^e für $e \in \mathbb{N}$ eine Primpotenz.
2. Polynome $f, g \in K[x]$ mit etwa $f \neq 0$ heißen teilerfremd, falls $\mathrm{ggT}(f,g) = 1$ gilt.

Bemerkung

Sei $I = K[x]f + K[x]g = \{uf + vg \mid u, v \in K[x]\}$. Dann ist I ein Ideal und wird von $\mathrm{ggT}(f,g)$ erzeugt.

Beweis

Offenbar gilt $I \neq \emptyset$ und $I \subseteq K[x]$. Sei nun $a, b \in I$. Dann folgt $a = u_1 f + v_1 g$ und $b = u_2 f + v_2 g$ mit $u_1, v_1, u_2, v_2 \in K[x]$, d.h. $a - b = (u_1 - u_2)f + (v_1 - v_2)g \in I$. Daher ist $(I, +)$ eine Untergruppe. Weiter gilt für $h \in K[x]$ stets $ha = h(u_1 f + v_1 g) = (hu_1)f + (hv_1)g \in I$. Insgesamt ist I ein Ideal.

Nach Satz 11.4 gilt dann $J = (d)$ für ein normiertes Polynom $d \in K[x]$. Wegen $f = 1 \cdot f + 0 \cdot g \in J = K[x]d$ gilt d/f. Analog erhält man d/g, also insgesamt $d/\mathrm{ggT}(f,g)$. Andererseits gilt $d = 1 \cdot d \in K[x]d = J = K[x]f + K[x]g$, d.h. $d = rf + sg$ für $r, s \in K[x]$. Mit $\mathrm{ggT}(f,g)/f \Longrightarrow f = \mathrm{ggT}(f,g) \cdot h_1$ und $\mathrm{ggT}(f,g)/g \Longrightarrow g = \mathrm{ggT}(f,g) \cdot h_2$ folgt $d = rf + sg = (rh_1 + sh_2) \cdot \mathrm{ggT}(f,g) \Longrightarrow \mathrm{ggT}(f,g)/d$.

Wegen $d/\mathrm{ggT}(f,g)$ und $\mathrm{ggT}(f,g)/d$ gilt $d = \alpha \cdot \mathrm{ggT}(f,g)$ für $\alpha \in K$. Da d und $\mathrm{ggT}(f,g)$ normiert sind, ergibt sich $\alpha = 1$.

11.16 Korollar (Satz von Bézout)

Für $f, g \in K[x] \setminus \{0\}$ existieren $r, s \in K[x]$ mit $rf + sg = \mathrm{ggT}(f,g)$.

Beweis

Mit $d = \mathrm{ggT}(f, g)$ erhält man $d \in (d) = K[x]f + K[x]g$, also $d = rf + sg$ für $r, s \in K[x]$.

11.17 Satz (euklidischer Algorithmus)

Seien $f, g \in K[x]$ mit $g \neq 0$. Sei weiter $f = qg + r$ eine Division mit Rest. Dann gilt $\mathrm{ggT}(f, g) = \mathrm{ggT}(g, r)$.

Beweis

Sei $\mathrm{ggT}(f, g) = d$. Dann gilt $d/f \Longrightarrow f = dh_1$ und $d/g \Longrightarrow g = dh_2$, d.h. $r = f - qg = d(h_1 - qh_2) \Longrightarrow d/r$. Wegen d/r und d/g gilt dann $d/\,\mathrm{ggT}(g, r)$. Analog folgt $\mathrm{ggT}(g, r)/d$. Da nun $\mathrm{ggT}(g, r)$ und d normiert sind, folgt $\mathrm{ggT}(g, r) = d$.

Beispiel

Seien $f, g \in \mathbb{Q}[x]$ durch $f = x^3 + x^2 + 2x$ und $g = x^2 + x + 1$ gegeben. Dann folgt aus

$$\begin{aligned} f &= x \cdot g + x \\ g &= (x+1) \cdot x + 1 \\ x &= x \cdot 1 + 0 \end{aligned}$$

sukzessive $\mathrm{ggT}(f, g) = \mathrm{ggT}(g, x) = \mathrm{ggT}(x, 1) = \mathrm{ggT}(1, 0) = 1$.

Der euklidische Algorithmus liefert weiter die Koeffizienten aus dem Satz von Bézout. Es gilt

$$\begin{aligned} \mathrm{ggT}(f, g) &= 1 \\ &= g - (x+1)x \\ &= g - (x+1)(f - xg) \\ &= -(x+1)f + (x^2 + x + 1)g. \end{aligned}$$

11.18 Satz

Sei $f \in K[x] \setminus \{0\}$. Sind dann f und f' teilerfremd, so hat f keine mehrfachen Nullstellen.

Beweis

Sei $d = \text{ggT}(f, f')$ und habe f die mehrfache Nullstelle a. Mit Satz 11.9 folgt dann $f'(a) = 0$. Nach Satz 11.5 gilt somit $(x - a)/f$ und $(x - a)/f'$, d.h. $(x - a)/d \Longrightarrow d \neq 1$.

Bemerkung

Der Ring $\mathbb{Z}$ der ganzen Zahlen und der Polynomring $K[x]$ weisen gewisse Analogien auf:

1. Die multiplikativ invertierbaren Elemente in $\mathbb{Z}$ bzw. in $K[x]$ sind $\{+1, -1\}$ bzw. $\{f \in K \mid f \neq 0\}$.
2. Für ein Element $z \in \mathbb{Z}$ gilt $1 \cdot z \in \mathbb{N}$ oder $-1 \cdot z \in \mathbb{N}$. Ein Polynom $0 \neq f \in K[x]$ kann durch ein Element $0 \neq \alpha \in K$ normiert werden.
3. In beiden Ringen ist Division mit Rest definiert. Ebenso ist der Satz von Bézout bzw. der euklidische Algorithmus in $\mathbb{Z}$ und $K[x]$ analog.
4. Einer Primzahl $p \in \mathbb{Z}$ entspricht ein irreduzibles, normiertes Polynom $f \in K[x]$. Für $a, b \in \mathbb{Z}$ sowie $g, h \in K[x]$ gilt dann jeweils $p/ab \Longrightarrow p/a$ oder p/b sowie $f/gh \Longrightarrow f/g$ oder f/h. Weiter hat eine positive ganze Zahl $z \in \mathbb{N} \setminus \{0\}$ bzw. eine normiertes Polynom $f \in K[x] \setminus K$ eine eindeutige Primfaktorzerlegung bzw. eine eindeutige Darstellung als Produkt irreduzibler, normierter Faktoren.

12 Direkte Summen

12.1 Definition

Seien $U_1, \ldots, U_m$ Unterräume eines Vektorraums, so dass jeder Vektor $v \in V$ eine eindeutige Darstellung $v = u_1 + \ldots + u_m$ mit $u_i \in U_i$ für $1 \leq i \leq m$ hat. Dann ist $V = U_1 \oplus \ldots \oplus U_m$ eine direkte Summe.

Beispiel

Sei $e_1, \ldots, e_m$ eine Basis von V. Dann ist $U_i = \langle e_i \rangle = Ke_i = \{\lambda e_i \mid \lambda \in K\} \subseteq V$ für $1 \leq i \leq m$ ein Unterraum. Wegen der Eindeutigkeit der Basisdarstellung folgt $V = U_1 \oplus \ldots \oplus U_m$.

12.2 Satz

Sei $V = U_1 + \ldots + U_m$ die Summe der Unterräume $U_i \subseteq V$ für $1 \leq i \leq m$. Sei weiter $W_i = U_1 + \ldots + U_{i-1} + U_{i+1} + \ldots + U_m$ für $1 \leq i \leq m$. Dann ist äquivalent

1. $V = U_1 \oplus \ldots \oplus U_m$
2. $U_i \cap W_i = \{0\}$ für $1 \leq i \leq m$

Beweis

1 $\Longrightarrow$ **2** Sei $u_i \in U_i \cap W_i$ für $1 \leq i \leq m$. Dann ist einerseits $u_i \in W_i$, d.h. es gilt $u_i = u_1 + \ldots + u_{i-1} + u_{i+1} + \ldots + u_m$ mit $u_j \in U_j$ für $j \neq i$. Daher folgt $0 = u_1 + \ldots + u_{i-1} - u_i + u_{i+1} + \ldots + u_m$. Andererseits gilt $u_i \in U_i$, also

$$0 + \ldots + 0 = 0 = u_1 + \ldots + u_m$$

mit $0 \in U_k$ und $u_k \in U_k$ für $1 \leq k \leq m$. Da nun V aber eine direkte Summe ist, ist die Darstellung von $0 \in V$ eindeutig. Man erhält also $u_k = 0$ für $1 \leq k \leq m$, d.h. insbesondere $u_i = 0$.

2 $\Longrightarrow$ **1** Sei $v \in V$. Dann gilt $v = u_1 + \ldots + u_m$ mit $u_i \in U_i$ für $1 \leq i \leq m$. Sei nun $v = \widetilde{u}_1 + \ldots + \widetilde{u}_m$ mit $\widetilde{u}_i \in U_i$ für $1 \leq i \leq m$ eine weitere Darstellung. Wegen $u_i, \widetilde{u}_i \in U_i$ folgt $u_i - \widetilde{u}_i \in U_i$ für $1 \leq i \leq m$. Mit $0 = v - v = (u_1 - \widetilde{u}_1) + \ldots + (u_m - \widetilde{u}_m)$ gilt für $1 \leq i \leq m$ daher

$$\widetilde{u}_i - u_i = (u_1 - \widetilde{u}_1) + \ldots + (u_{i-1} - \widetilde{u}_{i-1}) + (u_{i+1} - \widetilde{u}_{i+1}) + \ldots + (u_m - \widetilde{u}_m) \in W_i,$$

also $\widetilde{u}_i - u_i \in W_i \cap U_i = \{0\} \Longrightarrow \widetilde{u}_i = u_i$.

12.3 Definition / Bemerkung

Sei $\alpha \in \mathrm{End}(V)$. Ein Unterraum $U \subseteq V$ heißt dann α-invariant, wenn $\alpha(U) \subseteq U$ gilt. In diesem Fall ist $\alpha : U \longrightarrow U$ wieder wohldefiniert, d.h. $\alpha \in \mathrm{End}(U)$. Die Einschränkung von α auf U bezeichnet man mit $\alpha|_U$.

Bemerkung

Gewöhnlich beschreibt man Endomorphismen $\alpha \in \mathrm{End}(V)$ mit $n = \dim V < \infty$ durch Matrizen A bzgl. einer festen Basis B. Dann gilt mit der Koordinatenabbildung $\varphi : V \longrightarrow K^n$ dieser Basis für alle Vektoren $v \in V$ stets $\varphi\big(\alpha(v)\big) = A\varphi(v)$. Damit wird $\alpha \circ \cdots \circ \alpha$ durch $A \cdots A$ beschrieben.

Bemerkung

Sei $V = U_1 \oplus \ldots \oplus U_m$ mit $\dim V < \infty$ eine direkte Summe und $b_{i1}, \ldots, b_{in_i}$ für $1 \leq i \leq m$ eine Basis von U_i. Dann ist $b_{11}, \ldots, b_{1n_1}, \ldots, b_{m1}, \ldots, b_{mn_m}$ eine Basis von V. Insbesondere gilt also $\dim V = \dim U_1 + \ldots + \dim U_m$.

Beweis

Da $V = U_1 \oplus \ldots \oplus U_m$ eine direkte Summe ist, hat jeder Vektor $v \in V$ eine eindeutige Darstellung $v = u_1 + \ldots + u_m$ mit $u_i \in U_i$ für $1 \leq i \leq m$. Weiter ist $u_i = \sum_{k=1}^{n_i} \lambda_{ik} b_{ik}$ für $1 \leq i \leq m$ eine eindeutige Basisdarstellung. Insgesamt ist also $v \in V$ eine eindeutige Linearkombination

$$v = \sum_{i=1}^{m} \sum_{k=1}^{n_i} \lambda_{ik} b_{ik}.$$

Daher ist $b_{11}, \ldots, b_{1n_1}, \ldots, b_{m1}, \ldots, b_{mn_m}$ nach Satz 4.14 eine Basis von V.

12.4 Satz

Sei $\dim V < \infty$ und $\alpha \in \mathrm{End}(V)$. Für $1 \leq i \leq m$ sei V eine direkte Summe α-invarianter Unterräume U_i. Sei $\alpha|_{U_i}$ bzgl. einer Basis $b_{i1}, \ldots, b_{in_i}$ durch $A_i \in M_{n_i \times n_i}(K)$ dargestellt. Dann wird α bzgl. der Basis $b_{11}, \ldots, b_{1n_1}, \ldots, b_{m1}, \ldots, b_{mn_m}$ von V durch die Blockdiagonalmatrix

$$M = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_m \end{pmatrix}$$

dargestellt.

Beweis

Man betrachte $\alpha(b_{ij})$ für $1 \leq i \leq m$ und $1 \leq j \leq n_i$. Da U_i ein α-invarianter Unterraum ist, folgt für $b_{ij} \in U_i$ stets $\alpha(b_{ij}) \in U_i$. Daher gilt

$$\begin{aligned} \alpha(b_{ij}) &= \sum_{k=1}^{n_i} \lambda_{ijk} b_{ik} \\ &= \sum_{k=1}^{n_1} 0 \cdot b_{1k} + \ldots + \sum_{k=1}^{n_{i-1}} 0 \cdot b_{(i-1)k} + \sum_{k=1}^{n_i} \lambda_{ijk} b_{ik} + \sum_{k=1}^{n_{i+1}} 0 \cdot b_{(i+1)k} + \ldots + \sum_{k=1}^{n_m} 0 \cdot b_{mk} \end{aligned}$$

und es folgt die Behauptung.

12.5 Definition

Ein Endomorphismus $\alpha \in \mathrm{End}(V)$ eines endlich-dimensionalen Vektorraums V heißt diagonalisierbar, wenn α bzgl. einer geeigneten Basis durch eine Diagonalmatrix dargestellt wird.

Eine quadratische Matrix $A \in M_{n\times n}(K)$ heißt diagonalisierbar, wenn die lineare Abbildung $v \in K^n \longmapsto Av$ diagonalsierbar ist.

Bemerkung

1. Ein Endomorphismus $\alpha \in \mathrm{End}(V)$ eines endlich-dimensionalen Vektorraums V ist genau dann diagonalisierbar, wenn V eine Basis aus Eigenvektoren von α besitzt.

2. Eine quadratische Matrix $A \in M_{n\times n}(K)$ ist genau dann diagonalisierbar, wenn A ähnlich zu einer Diagonalmatrix ist.

Beweis

1 Sei $b_1, \ldots, b_n$ eine Basis aus Eigenvektoren. Dann gilt $\alpha(b_i) = 0 \cdot b_1 + \ldots + 0 \cdot b_{i-1} + \lambda_i \cdot b_i + 0 \cdot b_{i+1} + \ldots + 0 \cdot b_n$, d.h. α wird bzgl. $b_1, \ldots, b_n$ durch eine Diagonalmatrix beschrieben.

Sei nun α diagonalisierbar. Dann existiert eine Basis B, so dass $M_\alpha(B, B)$ eine Diagonalmatrix ist. Jeder Vektor $b \in B$ ist dann ein Eigenvektor.

2 Sei A diagonalisierbar. Dann wird $v \in K^n \longmapsto Av$ bzgl. einer Basis B durch eine Diagonalmatrix D dargestellt. Sei dann die Basistranfomration der Standardbasis von K^n zu B durch T beschrieben. Nach Korollar 7.12 gilt nun $A = T^{-1}DT$, d.h. A und D sind ähnlich.

Seien nun A ähnlich zu einer Diagonalmatrix D. Dann existiert eine invertierbare Matrix T mit $D = T^{-1}AT$. Man interpretiere nun T als Basistransformation der Standardbasis von K^n zu einer Basis B. Die Abbildung $v \in K^n \longmapsto Av$ wird dann bzgl. B durch D beschrieben.

12.6 Satz

Sei $n = \dim V < \infty$ und $\alpha \in \mathrm{End}(V)$. Dann gilt:

1. Seien $\lambda_1, \ldots, \lambda_s$ die verschiedenen Eigenwerte von α mit den geometrischen Vielfachheiten $g_1, \ldots, g_s$. Dabei gelte $g_1 + \ldots + g_s = n$. Dann ist α diagonalisierbar.

2. Wenn α genau n verschiedene Eigenwerte hat, ist α diagonalisierbar.

Beweis

1 Sei E_{λ_i} für $1 \leq i \leq s$ der Eigenraum zu λ_i und $W_i = E_{\lambda_1} + \ldots + E_{\lambda_{i-1}} + E_{\lambda_{i+1}} + \ldots + E_{\lambda_s}$. Sei dann $0 \neq v \in W$ definiert durch $v = e_1 + \ldots + e_{i-1} + e_{i+1} + \ldots + e_s$. Wegen $v \neq 0$ seien dabei o.B.d.A. $e_1, \ldots, e_j$ Eigenvektoren – d.h. $e_1, \ldots, e_j \notin \{0\}$ – und $e_{j+1}, \ldots, e_s \in \{0\}$. Es folgt

$$\begin{aligned} \alpha(v) &= \alpha(e_1) + \ldots + \alpha(e_{i-1}) + \alpha(e_{i+1}) + \ldots + \alpha(e_j) \\ &= \lambda_1 e_1 + \ldots + \lambda_{i-1} e_{i-1} + \lambda_{i+1} e_{i+1} + \ldots + \lambda_j e_j. \end{aligned}$$

Sei nun $v \in E_{\lambda_i}$, d.h. $\alpha(v) = \lambda_i v$. Dann gilt

$$0 = \alpha(v) - \alpha(v) = \lambda_1 e_1 + \ldots + \lambda_{i-1} e_{i-1} - \lambda_i v + \lambda_{i+1} e_{i+1} + \ldots + \lambda_j e_j.$$

Da $\lambda_1, \ldots, \lambda_j$ paarweise verschieden sind, ist dies nach Satz 10.2 eine Linearkombination linear unabhängiger Vektoren und es ist höchstens ein Eigenwert $\lambda_i = 0$. Damit erhält man einen Widerspruch und es folgt $E_{\lambda_i} \cap W_i = \{0\}$.

Nach Satz 12.2 gilt daher $E_{\lambda_1} + \ldots + E_{\lambda_s} = E_{\lambda_1} \oplus \ldots \oplus E_{\lambda_s}$, d.h. $\dim E_{\lambda_1} \oplus \ldots \oplus E_{\lambda_s} = \dim E_{\lambda_1} + \ldots + \dim E_{\lambda_s} = g_1 + \ldots + g_s = n = \dim V \Longrightarrow E_{\lambda_1} \oplus \ldots \oplus E_{\lambda_s} = V$. Daher besitzt V eine Basis aus Eigenvektoren und es folgt die Behauptung.

2 Seien $v_1, \ldots, v_n$ Eigenvektoren zu den Eigenwerten $\lambda_1, \ldots, \lambda_n$. Mit Satz 10.2 und $\dim V = n$ bildet $v_1, \ldots, v_n$ eine Basis von V, d.h. α ist diagonalisierbar.

Bemerkung

1. Satz 12.6 gilt analog für Matrizen.

2. Ein diagonalisierbarer Endomorphismus $\alpha \in \mathrm{End}(V)$ für $n = \dim V < \infty$ hat nicht notwendig n verschiedene Eigenwerte. Etwa wird $\alpha = \mathrm{id}$ stets durch eine Diagonalmatrix dargestellt, jedoch ist auch für $n \geq 2$ der einzige Eigenwert $\lambda = 1$.

3. Sei $\alpha \in \mathrm{End}(V)$ diagonalisierbar und durch
$$A = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$
dargestellt. Dann zerfällt $\chi_\alpha(x) = \chi_A(x) = (x - \lambda_1) \cdots (x - \lambda_n)$ in Linearfaktoren.

4. Dem charakteristischen Polynom $\chi_\alpha(x)$ lässt sich i.A. nicht ablesen, ob α diagonalisierbar ist. Offenbar haben etwa
$$A = \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} \text{ und } B = \begin{pmatrix} 1 & 1 \\ & 1 \end{pmatrix}$$

das gleiche charakteristische Polynom $(x-1)^2$. Dabei ist aber B – wegen

$$E_1 = \text{Kern} \begin{pmatrix} & 1 \\ & \end{pmatrix} = \langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \rangle$$

und da $\lambda = 1$ der einzige Eigenwert ist – nicht diagonalisierbar.

13 Minimalpolynome

Definition

Sei V ein K-Vektorraum und $\alpha \in \text{End}(V)$. Für $f = a_0 + a_1 x + \ldots + a_n x^n \in K[x]$ definiert man $f(\alpha)$ durch $f(\alpha)(v) = \sum_{i=1}^{n} a_i \alpha^i(v)$ mit $\alpha^i(v) = \alpha\big(\alpha^{i-1}(v)\big)$ und $\alpha^0(v) = v$. Für eine quadratische Matrix $A \in M_{n\times n}(K)$ gilt $f(A) = \sum_{i=1}^{n} a_i A^i$.

Bemerkung

Sei $\alpha \in \text{End}(V)$ fest gewählt. Dann ist $\varphi : f \in K[x] \longmapsto f(\alpha) \in \text{End}(V)$ ein K-linearer Ringhomomorphismus.

Beweis

Zunächst sind $K[x]$ und $\text{End}(V)$ jeweils K-Vektorräume und Ringe. Aus den Ringeigenschaften von $\text{End}(V)$ und $\alpha \in \text{End}(V)$ folgt weiter $f(\alpha) \in \text{End}(V)$, d.h. φ ist wohldefiniert.

Seien nun $f, g \in K[x]$ o.B.d.A. für $m \leq n$ durch $f = \sum_{i=1}^{n} a_i x^i$ und $g = \sum_{i=1}^{m} b_i x^i$ definiert. Dann gilt $f + g = \sum_{i=1}^{m} (a_i + b_i) x^i + \sum_{i=m+1}^{n} a_i x^i$. Aus den Eigenschaften von V folgt nun für alle $v \in V$ wegen $\alpha^i(v) \in V$ stets

$$\begin{aligned} \varphi(f)(v) + \varphi(g)(v) &= \sum_{i=1}^{n} a_i \alpha^i(v) + \sum_{i=1}^{m} b_i \alpha^i(v) \\ &= \sum_{i=1}^{m} (a_i + b_i) \alpha^i(v) + \sum_{i=m+1}^{n} a_i \alpha^i(v) \\ &= \varphi(f+g)(v), \end{aligned}$$

d.h. $\varphi(f) + \varphi(g) = \varphi(f+g)$.

Für $ax^n, bx^m \in K[x]$ und $\lambda \in K$ folgt für alle $v \in V$ weiter

$$\begin{aligned} \varphi(ax^n)\big(\varphi(bx^m)(v)\big) &= a\alpha^n\big(b\alpha^m(v)\big) &= ab\alpha^{n+m}(v) \\ &= \varphi(abx^{n+m})(v) &= \varphi(ax^n \cdot bx^m)(v), \end{aligned}$$

d.h. $\varphi(ax^n) \circ \varphi(bx^m) = \varphi(ax^n \cdot bx^m)$, sowie

$$\begin{aligned} \varphi(\lambda a x^n)(v) &= \lambda a \alpha^n(v) \\ &= \lambda \varphi(ax^n)(v), \end{aligned}$$

d.h. $\varphi(\lambda a x^n) = \lambda \varphi(ax^n)$. Damit folgt $\varphi(f \cdot g) = \varphi(f) \circ \varphi(g)$ und $\varphi(\lambda f) = \lambda \varphi(f)$ aus der Additivität von φ.

Zuletzt gilt für $v \in V$ stets $\varphi(1)(v) = v = (\,id)(v)$, also $\varphi(1) = 1$.

Bemerkung

Endomorphismen $f(\alpha), g(\alpha)$ mit $f, g \in K[x]$ und $\alpha \in \mathrm{End}(V)$ kommutieren.

Beweis

Man betrachte den K-linearen Ringhomomorphismus $\varphi : f \in K[x] \longmapsto f(\alpha) \in \mathrm{End}(V)$. Es folgt dann aus der Kommutativität in $K[x]$ stets $f(\alpha) \circ g(\alpha) = \varphi(f) \circ \varphi(g) = \varphi(f \cdot g) = \varphi(g \cdot f) = \varphi(g) \circ \varphi(f) = g(\alpha) \circ f(\alpha)$.

Beispiel

Es gilt etwa $\alpha \circ f(\alpha) = f(\alpha) \circ \alpha$. Dies ist einsichtig mit $\alpha = g(\alpha)$ für $g = x$.

13.1 Lemma

Sei $n = \dim V < \infty$ und $\alpha \in \mathrm{End}(V)$. Dann existiert ein $0 \neq f \in K[x]$ mit $f(\alpha) = 0$.

Beweis

Dem Endomorphismenraum $\mathrm{End}(V)$ entspricht der Vektorraum $M_{n\times n}(K)$ aller quadratischer $(n \times n)$-Matrizen. Nach Lemma 6.2 gilt daher $\dim \mathrm{End}(V) = \dim M_{n\times n}(K) = n^2$.

Damit sind die n^2+1 Endomorphismen $1, \alpha, \ldots, \alpha^{n^2}$ linear abhängig, d.h. es existieren $a_i \in K$ mit $a_0 + a_1\alpha + \ldots + a_{n^2}\alpha^{n^2} = 0$ und $a_i \neq 0$ für ein $1 \leq i \leq n^2$. Setze $f = a_0 + a_1 x + \ldots + a_{n^2} x^{n^2}$.

13.2 Bemerkung / Definition

Sei $\dim V < \infty$ und $\alpha \in \mathrm{End}(V)$. Da $\varphi : f \in K[x] \longmapsto f(\alpha) \in \mathrm{End}(V)$ ein Ringhomomorphismus ist, ist $I = \mathrm{Kern}\,\varphi = \{f \in K[x] \mid f(\alpha) = 0\}$ ein Ideal von $K[x]$.

Da nach Lemma 13.1 $I \neq \{0\}$ gilt, existiert nach Satz 11.4 ein normiertes Polynom m_α mit $I = K[x]m_\alpha$. Damit gilt für jedes Polynom $f \in K[x]$ mit $f(\alpha) = 0$ stets $f \in K[x]m_\alpha \Longrightarrow f = hm_\alpha$ für ein $h \in K[x] \Longrightarrow m_\alpha / f$. Also ist m_α ist das kleinste normierte Polynom, das α annulliert. Man nennt m_α das Minimalpolynom von α. Das Minimalpolynom quadratischer Matrizen definiert man analog.

Bemerkung

Lemma 13.1 gibt eine Möglichkeit an, das Minimalpolynom von $\alpha \in \mathrm{End}(V)$ für $\dim V < \infty$ zu bestimmen. Man finde dazu $k = \max\{k \in \mathbb{N} \mid 1, \alpha, \ldots, \alpha^{k-1} \text{ linear unabhängig}\}$. Dann gilt $\mathrm{rang}\, m_\alpha = k$. Die Koeffizienten a_i von m_α erhält man aus der nicht-trivialen Linearkombination $a_0 + a_1\alpha + \ldots + a_k\alpha^k = 0$.

13.3 Lemma

Sei $\alpha \in \mathrm{End}(V)$ und $f, g \in K[x]$. Dann gilt:

1. α-invariante Unterräume von V sind $f(\alpha)$-invariant.
2. $\mathrm{Bild}\, f(\alpha)$ und $\mathrm{Kern}\, f(\alpha)$ sind α-invariant.
3. Sei $V = U_1 \oplus \ldots \oplus U_m$ direkte Summe α-invarianter Unterräume. Dann ist $\mathrm{Kern}\, f(\alpha) = \bigoplus_{i=1}^{m} U_i \cap \mathrm{Kern}\, f(\alpha)$.
4. Für f/g gilt $\mathrm{Kern}\, f(\alpha) \subseteq \mathrm{Kern}\, g(\alpha)$ und $\mathrm{Bild}\, g(\alpha) \subseteq \mathrm{Bild}\, f(\alpha)$.
5. $\mathrm{Kern}\, f(\alpha) \cap \mathrm{Kern}\, g(\alpha) = \mathrm{Kern}\, \mathrm{ggT}(f, g)(\alpha)$.

Beweis

1 Sei U α-invariant. Dann gilt für $u \in U$ auch $\alpha^k(u) \in U$ für alle $k \in \mathbb{N}$, denn $\alpha^0(u) = u \in U$ und $\alpha^{k+1}(u) = \alpha\big(\alpha^k(u)\big) \in U$ für $\alpha^k(u) \in U$.

Da U ein Unterraum ist, liegt dann auch jede Linearkombination $f(\alpha)(u) = \sum_{i=0}^{k} \lambda_i \alpha^i(u)$ der $\alpha^i(u)$ in U.

2 Sei $v \in \operatorname{Bild} f(\alpha)$, also $v = f(\alpha)(w)$ für ein $w \in V$. Dann ist auch $\alpha(w) \in V$ und daher $\alpha(v) = \alpha\big(f(\alpha)(w)\big) = f(\alpha)\big(\alpha(w)\big) \in \operatorname{Bild} f(\alpha)$.

Sei nun $v \in \operatorname{Kern} f(\alpha)$, also $f(\alpha)(v) = 0$. Dann ist auch $f(\alpha)\big(\alpha(v)\big) = \alpha\big(f(\alpha)(v)\big) = \alpha(0) = 0$, also $\alpha(v) \in \operatorname{Kern} f(\alpha)$.

3 Wegen $U_i \cap \operatorname{Kern} f(\alpha) \subseteq \operatorname{Kern} f(\alpha)$ für alle $i = 1, \ldots, m$, gilt

$$\bigoplus_{i=1}^{m} U_i \cap \operatorname{Kern} f(\alpha) \subseteq \operatorname{Kern} f(\alpha).$$

Es bleibt also $\operatorname{Kern} f(\alpha) \subseteq \bigoplus_{i=1}^{m} U_i \cap \operatorname{Kern} f(\alpha)$ zu zeigen.

Sei dazu $v \in \operatorname{Kern} f(\alpha)$. Wegen $V = U_1 \oplus \ldots \oplus U_m$ hat dann v eine eindeutige Darstellung $v = u_1 + \ldots + u_m$ mit $u_i \in U_i$. Es folgt $0 = f(\alpha)(v) = f(\alpha)(u_1 + \ldots + u_m) = f(\alpha)(u_1) + \ldots + f(\alpha)(u_m)$.

Nach 1 ist U_i auch $f(\alpha)$-invariant, d.h. es ist $f(\alpha)(u_i) \in U_i$. Wegen der Eindeutigkeit der Darstellung der $0 \in V$ gilt dann $f(\alpha)(u_i) = 0$, also $u_i \in \operatorname{Kern} f(\alpha) \Rightarrow u_i \in U_i \cap \operatorname{Kern} f(\alpha) \Rightarrow v = \sum_{i=1}^{m} u_i \in \bigoplus_{i=1}^{m} U_i \cap \operatorname{Kern} f(\alpha)$.

4 Sei $f, g \in K[x]$ und f ein Teiler von g, d.h. $g = f \cdot h$ für ein $h \in K[x]$.

Sei weiter $v \in \operatorname{Kern} f(\alpha)$, d.h. $f(\alpha)(v) = 0$. Dann gilt

$$g(\alpha)(v) = (h \cdot f)(\alpha)(v) = h(\alpha)\big(f(\alpha)(v)\big) = h(\alpha)(0) = 0,$$

also $v \in \operatorname{Kern} g(\alpha)$.

Sei nun $v \in \operatorname{Bild} g(\alpha)$, d.h. $v = g(\alpha)(w)$ für ein $w \in V$. Dann ist auch $h(\alpha)(w) \in V$ und es folgt

$$v = g(\alpha)(w) = (f \cdot h)(\alpha)(w) = f(\alpha)\big(h(\alpha)(w)\big) \in \operatorname{Bild} f(\alpha).$$

5 Sei $d = \operatorname{ggT}(f, g)$. Wegen d/f und d/g folgt nach 4 zunächst $\operatorname{Kern} d(\alpha) \subseteq \operatorname{Kern} f(\alpha) \cap \operatorname{Kern} g(\alpha)$.

Nun gibt es nach dem Satz von Bézout $r, s \in K[x]$ mit $r \cdot f + s \cdot g = d$. Für $v \in \operatorname{Kern} f(\alpha) \cap \operatorname{Kern} g(\alpha)$, d.h. $f(\alpha)(v) = 0$ und $g(\alpha)(v) = 0$, gilt dann $d(\alpha)(v) = (r \cdot f + s \cdot g)(\alpha)(v) = r(\alpha)\big(f(\alpha)(v)\big) + s(\alpha)\big(g(\alpha)(v)\big) = r(\alpha)(0) + s(\alpha)(0) = 0$, also $v \in \operatorname{Kern} d(\alpha)$.

Korollar

Sind $f, g \in K[x]$ teilerfremd, so ist $\operatorname{Kern} f(\alpha) \cap \operatorname{Kern} g(\alpha) = \{0\}$.

Beweis

Es gilt f, g teilerfremd $\Rightarrow \operatorname{ggT}(f,g) = 1 \Rightarrow \operatorname{ggT}(f,g)(\alpha) = \mathrm{id} \Rightarrow \operatorname{Kern} f(\alpha) \cap \operatorname{Kern} g(\alpha) = \operatorname{Kern} \mathrm{id} = \{0\}$.

Bemerkung

Im Folgenden sei stets $\dim V < \infty$.

13.4 Lemma

Sei ein Endomorphismus $\alpha \in \operatorname{End}(V)$ gegeben, für dessen Minimalpolynom $\operatorname{grad} m > 1$ gilt. Sei weiter $p \in K[x]$ ein normierter Teiler von m mit $1 \leq \operatorname{grad} p < \operatorname{grad} m$. Dann ist $U = \operatorname{Bild} \frac{m}{p}(\alpha)$ ein α-invarianter Unterraum mit $\{0\} \neq U \subsetneq V$. Ferner ist p das Minimalpolynom von $\alpha|_U$.

Beweis

Da p ein Teiler von m ist, gilt $m = pq$ für ein $q \in K[x]$. Damit ist $U = \operatorname{Bild} \frac{m}{p}(\alpha) = \operatorname{Bild} q(\alpha)$ nach Lemma 13.3 α-invariant.

Weiter gilt $p(\alpha)(U) = p(\alpha)\big(\operatorname{Bild} q(\alpha)\big) = p(\alpha)\big(q(\alpha)(V)\big) = (pq)(\alpha)(V) = m(\alpha)(V) = 0(V) = \{0\}$. Mit $1 \leq \operatorname{grad} p$ ist $p \neq 0$. Wegen $\operatorname{grad} p < \operatorname{grad} m$ folgt damit $p(\alpha) \neq 0 \Longrightarrow p(\alpha)(V) \neq \{0\}$. Insgesamt ergibt sich also $U \subsetneq V$.

Nun ergibt sich aus $0 \neq m = pq$ zunächst $q \neq 0$. Wegen $\operatorname{grad} m = \operatorname{grad} p + \operatorname{grad} q \geq 1 + \operatorname{grad} q > \operatorname{grad} q$ folgt dann weiter $q(\alpha) \neq 0$. Damit gilt $U = q(\alpha)(V) \neq \{0\}$.

Sei schließlich $\widetilde{p} \in K[x]$ das Minimalpolynom von $\alpha|_U : U \longrightarrow U$. Mit $p(\alpha|_U)(U) = p(\alpha)(U) = \{0\}$ folgt dann einerseits $\widetilde{p}/p$. Andererseits gilt $(\widetilde{p}q)(\alpha)(V) = \widetilde{p}(\alpha)\big(q(\alpha)(V)\big) = \widetilde{p}(\alpha)(U) = \{0\}$, d.h. $pq = m/\widetilde{p}q \Longrightarrow \widetilde{p}q = pqh$ für ein $h \in K[x] \Longrightarrow \widetilde{p} = ph \Longrightarrow p/\widetilde{p}$. Da nun p und $\widetilde{p}$ normiert sind, folgt $p = \widetilde{p}$.

13.5 Satz

Sei $m \in K[x]$ das Minimalpolynom von $\alpha \in \operatorname{End}(V)$. Dabei sei $m = pq$ für $\operatorname{ggT}(p,q) = 1$. Dann gilt

1. $\operatorname{Kern} p(\alpha) = \operatorname{Bild} q(\alpha)$ und $\operatorname{Kern} q(\alpha) = \operatorname{Bild} p(\alpha)$
2. $V = \operatorname{Kern} p(\alpha) \oplus \operatorname{Kern} q(\alpha)$ und $V = \operatorname{Bild} p(\alpha) \oplus \operatorname{Bild} q(\alpha)$

Sei nun $m = \prod_{i=1}^{n} p_i^{e_i}$ die eindeutige Zerlegung von m in verschiedene normierte und irreduzible Polynome. Für $U_i = \operatorname{Kern} p_i^{e_i}(\alpha)$ gilt dann

3. $V = \bigoplus_{i=1}^{n} U_i$
4. U_i ist α-invariant
5. $p_i^{e_i}$ ist das Minimalpolynom von $\alpha|_{U_i}$

Beweis

Sei zunächst $v \in \operatorname{Bild} p(\alpha)$, d.h. $v = p(\alpha)(w)$ für ein $w \in V$. Es folgt

$$q(\alpha)(v) = q(\alpha)\big(p(\alpha)(w)\big) = (qp)(\alpha)(v) = m(\alpha)(v) = 0,$$

d.h. $v \in \operatorname{Kern} q(\alpha)$. Analog gilt für $v \in \operatorname{Bild} q(\alpha)$ stets $v \in \operatorname{Kern} p(\alpha)$. Damit erhalten wir

$$(*_1) \quad \operatorname{Bild} p(\alpha) \subseteq \operatorname{Kern} q(\alpha) \text{ und } \operatorname{Bild} q(\alpha) \subseteq \operatorname{Kern} p(\alpha).$$

Da p und q teilerfremd sind, gilt mit $(*_1)$ weiter

$$(*_2) \quad \operatorname{Kern} p(\alpha) \cap \operatorname{Kern} q(\alpha) = \{0\} \text{ und } \operatorname{Bild} p(\alpha) \cap \operatorname{Bild} q(\alpha) = \{0\}.$$

Nun gilt offenbar $\operatorname{Bild} p(\alpha) + \operatorname{Bild} q(\alpha) \subseteq V$. Andererseits existieren nach Bézout $r, s \in K[x]$ mit $pr+qs = \mathrm{ggT}(p,q) = 1$. Damit folgt für $v \in V$ stets $v = (pr+qs)(\alpha)(v) = p(\alpha)\big(r(\alpha)(v)\big) + q(\alpha)\big(s(\alpha)(v)\big) \in \operatorname{Bild} p(\alpha) + \operatorname{Bild} q(\alpha)$. Insgesamt gilt mit $(*_1)$ also

$$(*_3) \quad V = \operatorname{Bild} p(\alpha) + \operatorname{Bild} q(\alpha) \text{ und } V = \operatorname{Kern} p(\alpha) + \operatorname{Kern} q(\alpha).$$

Mit $(*_2)$ und $(*_3)$ sowie Satz 12.2 erhalten wir dann $V = \operatorname{Kern} p(\alpha) \oplus \operatorname{Kern} q(\alpha)$ und $V = \operatorname{Bild} p(\alpha) \oplus \operatorname{Bild} q(\alpha)$. Insbesondere gilt also

$$(*_4) \quad \dim \operatorname{Kern} p(\alpha) + \dim \operatorname{Kern} q(\alpha) = \dim \operatorname{Bild} p(\alpha) + \dim \operatorname{Bild} q(\alpha).$$

Wegen $(*_1)$ gilt nun $\dim \operatorname{Bild} p(\alpha) \leq \dim \operatorname{Kern} q(\alpha)$ und $\dim \operatorname{Bild} q(\alpha) \leq \dim \operatorname{Kern} p(\alpha)$. Aus $(*_4)$ folgt daher $\dim \operatorname{Bild} p(\alpha) = \dim \operatorname{Kern} q(\alpha)$ und $\dim \operatorname{Bild} q(\alpha) = \dim \operatorname{Kern} p(\alpha)$. Nach $(*_1)$ erhält man damit $\operatorname{Kern} p(\alpha) = \operatorname{Bild} q(\alpha)$ und $\operatorname{Kern} q(\alpha) = \operatorname{Bild} p(\alpha)$.

Bisher haben wir die Aussagen 1 und 2 bewiesen. Weiter folgt 4 unmittelbar aus Lemma 13.3. Es bleibt also 3 und 5 zu zeigen. Diese Aussagen weisen wir durch vollständige Induktion über n nach. Für $n = 1$ ist wegen $m = p_1^{e_1}$ und $U_1 = \operatorname{Kern} p_1^{e_1}(\alpha) = \operatorname{Kern} m(\alpha) = V$ – also $\alpha|_{U_1} = \alpha$ – der Fall klar.

Sei nun $m = p_1^{e_1} \cdot (p_2^{e_2} \cdots p_n^{e_n})$. Da die p_i für $1 \leq i \leq n$ irreduzibel und verschieden sind, sind nach Satz 11.12 die Polynome $p_1^{e_1}$ und $p_2^{e_2} \cdots p_n^{e_n}$ teilerfremd. Mit 2 folgt daher $V =$

$\operatorname{Kern} p_1^{e_1}(\alpha) \oplus \operatorname{Kern}(p_2^{e_2} \cdots p_n^{e_n})(\alpha)$. Aus 1 ergibt sich nun $U_1 = \operatorname{Kern} p_1^{e_1}(\alpha) = \operatorname{Bild} \frac{m}{p_1^{e_1}}(\alpha)$. Daher ist $p_1^{e_1}$ nach Lemma 13.4 das Minimalpolynom von $\alpha|_{U_1}$.

Nach 1 und Lemma 13.4 ist weiter $U = \operatorname{Kern}(p_2^{e_2} \cdots p_n^{e_n})(\alpha) = \operatorname{Bild} \frac{m}{p_1^{e_1}}(\alpha)$ ein α-invarianter Unterraum und $p_2^{e_2} \cdots p_n^{e_n}$ das Minimalpolymon von $\alpha|_U$. Die Anwendung der Induktionsannahme auf U liefert dann die Behautung.

Bemerkung

Nach Satz 12.4 lässt sich nun ein Endomorphismus $\alpha \in \operatorname{End}(V)$ mit einer Primzerlegung $m = \prod_{i=1}^{n} p_i^{e_i}$ des Minimalpolynoms durch eine Blockdiagonalmatrix

$$\begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_n \end{pmatrix}$$

darstellen, so dass für $1 \leq i \leq n$ und $U_i = \operatorname{Kern} p_i^{e_i}(\alpha)$ der Endomorphismus $\alpha|_{U_i}$ durch die Matrix A_i berschrieben wird. Weiter ist dann $p_i^{e_i}$ das Minimalpolynom von $\alpha|_{U_i}$. Daher ist die Untersuchung von α auf die Untersuchung solcher Endomorphismen zurückgeführt, deren Minimalpolynome Primpotenzen sind.

13.6 Definition

Sei $\alpha \in \operatorname{End}(V)$ und $U = \langle v, \alpha(v), \alpha^2(v), \ldots \rangle = \langle \{\alpha^i(v) \mid i \in \mathbb{N}\} \rangle$ für $v \in V$. Dann heißt U der von v erzeugte α-zyklische Unterraum von V. Man schreibt $U = \langle v \rangle_\alpha$.

13.7 Lemma

Sei $\alpha \in \operatorname{End}(V)$ und $v \in V$. Dann ist

1. $U = \langle v \rangle_\alpha$ der kleinste α-invariante Unterraum von V, der v enthält.

Sei weiter $m \in K[x]$ das Minimalpolynom von $\alpha|_U$ mit $\operatorname{grad} m = d$. Dann gilt

2. $\langle v \rangle_\alpha = \langle v, \alpha(v), \ldots, \alpha^{d-1}(v) \rangle$.

Beweis

1 Sei $w \in \langle v \rangle_\alpha$, d.h. $w = \sum_{i=0}^{n} a_i \alpha^i(v)$. Dann gilt auch $\alpha(w) = \alpha\big(\sum_{i=0}^{n} a_i \alpha^i(v)\big) = \sum_{i=0}^{n} a_i \alpha^{i+1}(v) \in \langle v \rangle_\alpha$, d.h. $\langle v \rangle_\alpha$ ist α-invariant. Sei nun W ein α-invarianter Unterraum von V mit $v \in W$.

Dann gilt offenbar $\{v, \alpha(v), \alpha^2(v), \ldots\} \subseteq W$, also $\langle v\rangle_\alpha = \langle v, \alpha(v), \alpha^2(v), \ldots\rangle \subseteq W$.

2 Da zunächst U ein α-invarianter Unterraum ist, gilt $\alpha|_U \in \mathrm{End}(U)$. Für $v \in \langle v\rangle_\alpha = U$ ist dann $\alpha(v) = \alpha|_U(v)$ und daher

$$m(\alpha)(v) = m(\alpha|_U)(v) = 0.$$

Sei nun $x^i = q_i m + r_i$ mit $q_i, r_i \in K[x]$ und $\operatorname{grad} r_i < \operatorname{grad} m = d$ für $i \in \mathbb{N}$ eine Division mit Rest. Dann ist

$$\begin{aligned} \alpha^i(v) &= (q_i m + r_i)(\alpha)(v) &&= q_i(\alpha)\big(m(\alpha)(v)\big) + r_i(\alpha)(v) \\ &= q_i(\alpha)(0) + r_i(\alpha)(v) &&= r_i(\alpha)(v). \end{aligned}$$

Wegen $\operatorname{grad} r_i < d$ gilt weiter $r_i = a_0 + a_1 x + \ldots + a_{d-1}x^{d-1}$. Daher folgt

$$\alpha^i(v) = r_i(\alpha)(v) = a_0 v + a_1\alpha(v) + \ldots + a_{d-1}\alpha^{d-1}(v),$$

also $\alpha^i(v) \in \langle v, \alpha(v), \ldots, \alpha^{d-1}(v)\rangle$ für alle $i \in \mathbb{N}$. Dies liefert $\langle v\rangle_\alpha = \langle v, \alpha(v), \alpha^2(v), \ldots\rangle = \langle v, \alpha(v), \ldots, \alpha^{d-1}(v)\rangle$.

13.8 Lemma

Sei $\alpha \in \mathrm{End}(V)$ und $U = \langle v\rangle_\alpha$ mit $v \in V$. Für das Minimalpolynom m von $\alpha|_U$ gilt dann $\dim U = \operatorname{grad} m$.

Beweis

Sei $\dim U = n$ und $\operatorname{grad} m = d$. Nach Lemma 13.7 gilt dann $V = \langle v, \alpha(v), \ldots, \alpha^{d-1}(v)\rangle$, d.h. $n \leq d$. Wegen $\dim U = n$ sind nun $v, \alpha(v), \ldots, \alpha^n(v) \in \langle v\rangle_\alpha = U$ linear abhängig. Daher existieren $a_1, \ldots, a_n \in K$ mit

$$\sum_{i=0}^{n} a_i\alpha^i(v) = 0$$

und $a_i \neq 0$ für ein $1 \leq i \leq n$. Man setze $f = \sum\limits_{i=0}^{n} a_i x^i \in K[x]$, d.h. $f(\alpha)(v) = 0$ und $\operatorname{grad} f \leq n$. Sei nun $w = \sum\limits_{i=0}^{n} b_i\alpha^i(v) \in U$. Dann gilt

$$f(\alpha)(w) = f(\alpha)(\sum_{i=0}^{n} b_i\alpha^i(v)) = \sum_{i=0}^{n} b_i f(\alpha)\big(\alpha^i(v)\big) = \sum_{i=0}^{n} b_i\alpha^i\big(f(\alpha)(v)\big) = \sum_{i=0}^{n} b_i\alpha^i(0) = 0,$$

also $f(\alpha|_U)(U) = f(\alpha)(U) = \{0\}$. Wegen $f \neq 0$ folgt daher $d = \operatorname{grad} m \leq \operatorname{grad} f \leq n$. Insgesamt erhalten wir also $d = n$.

13.9 Lemma

Sei das Minimalpolynom von $\alpha \in \mathrm{End}(V)$ eine Primpotenz p^k mit $k \geq 1$. Dann existiert ein α-zyklischer Unterraum U von V mit $\dim U = \operatorname{grad} p^k$. Insbesondere ist V dann α-zyklisch, falls $\operatorname{grad} p^k = \dim V$ gilt.

Beweis

Wegen $\operatorname{grad} p^{k-1} < \operatorname{grad} p^k$ und $p^{k-1} \neq 0$ gilt $p^{k-1}(\alpha) \neq 0$. Daher existiert ein $v \in V$ mit $p^{k-1}(\alpha)(v) \neq 0$. Sei dann $U = \langle v \rangle_\alpha$ und m das Minimalpolynom von $\alpha|_U$. Nach Lemma 13.8 gilt $\dim U = \operatorname{grad} m$.

Nun gilt aber $p^k(\alpha|_U)(U) = p^k(\alpha)(U) = \{0\}$, d.h. m/p^k. Da p irreduzibel ist, folgt $m = p^i$ für ein $0 \leq i \leq k$. Wegen $p^{k-1}(\alpha)(v) \neq 0 \Longrightarrow p^{k-1}(\alpha) \neq 0$ ist dabei $i = k$. Insgesamt erhalten wir also $\dim U = \operatorname{grad} p^k$. Für $\operatorname{grad} p^k = \dim V$ folgt schließlich $U = V$.

13.10 Satz

Sei m das Minimalpolynom von $\alpha \in \mathrm{End}(V)$. Dann gilt $\operatorname{grad} m \leq \dim V$.

Beweis

Wir zeigen die Aussage durch vollständige Induktion über $\dim V$. Für $\dim V = 1$ gilt dabei $V = \langle v \rangle = Kv$. Dann ist $\alpha(v) = \lambda v$ für $\lambda \in V$. Wegen $w \in V \Longrightarrow w = \mu v \Longrightarrow \alpha(w) - \lambda w = \alpha(\mu v) - \lambda \mu v = \mu \alpha(v) - \lambda \mu v = \mu \lambda v - \lambda \mu v = 0$ gilt $(x - \lambda)(\alpha)(V) = \{0\}$, d.h. $\operatorname{grad} m \leq \operatorname{grad}(x - \lambda) = 1 = \dim V$.

Nun betrachten wir zwei Fälle:

1 Sei $V = U \oplus W$ die direkte Summe α-invarianter Unterräume mit $\dim U \geq 1$ und $\dim W \geq 1$. Wegen $\dim V = \dim(U \oplus W) = \dim U + \dim W$ gilt dann $\dim U < \dim V$ und $\dim W < \dim V$. Seien nun m_U bzw. m_W die Minimalpolynome von $\alpha|_U$ bzw. $\alpha|_W$. Nach Induktionsannahme gelte dabei $\operatorname{grad} m_U \leq \dim U$ und $\operatorname{grad} m_W \leq \dim W$.

Sei nun $v \in V = U \oplus W$, d.h. $v = u + w$ für $u \in U$ und $w \in W$. Es folgt

$$\begin{aligned}
(m_U m_W)(\alpha)(v) &= (m_U m_W)(\alpha)(u+w) \\
&= (m_W m_U)(\alpha)(u) + (m_U m_W)(\alpha)(w) \\
&= m_W(\alpha)\big(m_U(\alpha)(u)\big) + m_U(\alpha)\big(m_W(\alpha)(w)\big) \\
&= m_W(\alpha)\big(m_U(\alpha|_U)(u)\big) + m_U(\alpha)\big(m_W(\alpha|_W)(w)\big) \\
&= m_W(\alpha)(0) + m_U(\alpha)(0) \\
&= 0,
\end{aligned}$$

also $(m_U m_W)(\alpha)(V) = \{0\}$. Wegen $m_U \neq 0$ und $m_W \neq 0$ folgt damit $\operatorname{grad} m \leq \operatorname{grad} m_U m_W = \operatorname{grad} m_U + \operatorname{grad} m_W \leq \dim U + \dim W = \dim V$.

2 Sei nun V keine direkte Summe α-invarianter Unterräume U mit $\dim U \geq 1$. Nach Lemma 13.5 gilt für $m = pq$ mit $\operatorname{ggT}(p,q) = 1$ aber stets $V = \operatorname{Bild} p(\alpha) \oplus \operatorname{Bild} q(\alpha)$. Es folgt daher $\dim \operatorname{Bild} p(\alpha) = 0$ oder $\dim \operatorname{Bild} q(\alpha) = 0$. Sei o.B.d.A. $\dim \operatorname{Bild} p(\alpha) = 0 \iff \operatorname{Bild} p(\alpha) = \{0\} \iff p(\alpha) = 0$.

Wegen $p \neq 0$ folgt damit $p = m \iff q = 1$. Daher ist m eine Primpotenz p^k. Für $k = 0$ ist die Aussage wegen $m = 1 \Longrightarrow \operatorname{grad} m = 0$ trivial. Sei also $k \geq 1$. Dann existiert nach Lemma 13.9 ein Unterraum U mit $\operatorname{grad} m = \dim U \leq \dim V$.

13.11 Lemma

Das Minimalpolynom von $\alpha \in \operatorname{End}(V)$ habe die Form p^k mit $p \in K[x]$ irreduzibel und normiert. Sei $n = \dim V = \operatorname{grad} p^k$. Dann ist V keine direkte Summe α-invarianter Unterräume U mit $\dim U < \dim V$. Ferner gilt $\operatorname{Kern} p^i(\alpha) = \operatorname{Bild} p^{k-i}(\alpha)$.

Beweis

Sei $V = U \oplus W$ direkte Summe α-invarianter Unterräume mit $\dim U < \dim V$ und $\dim W < \dim V$. Seien dann m_U und m_W die Minimalpolynome von $\alpha|_U$ und $\alpha|_W$. Dann gilt

$$p^k(\alpha|_U)(U) = p^k(\alpha)(U) = \{0\} \Longrightarrow m_U / p^k$$

und analog m_W / p^k.

Mit Satz 13.10 gilt weiter $\operatorname{grad} m_U \leq \dim U < \dim V = \operatorname{grad} p^k$ und analog $\operatorname{grad} m_W < \operatorname{grad} p^k$. Da p irreduzibel ist, sind m_U und m_W Teiler von p^{k-1} – d.h. $p^{k-1}(\alpha)(U) = \{0\}$ und $p^{k-1}(\alpha)(W) = \{0\}$. Daher gilt für $v \in V$ mit $v = u + w$ für $u \in U$ und $w \in W$ stets

$p^{k-1}(\alpha)(v) = p^{k-1}(\alpha)(u+w) = p^{k-1}(\alpha)(u) + p^{k-1}(\alpha)(w) = 0$, also $p^{k-1}(\alpha)(V) = \{0\}$ – im Widerspruch zu p^k minimal.

Nun zeigen wir $\operatorname{Kern} p^i(\alpha) = \operatorname{Bild} p^{k-i}(\alpha)$. Zunächst gilt $w = p^{k-i}(\alpha)(v) \Longrightarrow p^i(\alpha)(w) = p^i(\alpha)\big(p^{k-i}(\alpha)(v)\big) = p^k(\alpha)(v) = 0$, d.h. $\operatorname{Bild} p^{k-i}(\alpha) \subseteq \operatorname{Kern} p^i(\alpha)$.

Wegen $\dim V = \operatorname{grad} p^k$ existiert nach Lemma 13.9 ein $v \in V$ mit $V = \langle v \rangle_\alpha$. Mit $\dim V = n$ bildet dann $v, \alpha(v), \ldots, \alpha^{n-1}(v)$ eine Basis von V. Sei nun $w \in \operatorname{Kern} p^i(\alpha)$ mit $w = \sum_{j=0}^{n-1} a_j \alpha^j(v)$ für $a_j \in K$.

Setze $g = \sum_{j=0}^{n-1} a_j x^j \in K[x]$, also $w = g(\alpha)(v)$. Dann gilt $0 = p^i(\alpha)(w) = p^i(\alpha)\big(g(\alpha)(v)\big) = (p^i g)(\alpha)(v)$. Da aber V von v erzeugt wird, folgt $u \in V \Longrightarrow u = \sum_{j=0}^{n-1} \lambda_j \alpha^j(v)$ und damit

$$\begin{aligned}(p^i g)(\alpha)(u) &= (p^i g)(\alpha)\big(\sum_{j=0}^{n-1} \lambda_j \alpha^j(v)\big) &&= \sum_{j=0}^{n-1} \lambda_j (p^i g)(\alpha)\big(\alpha^j(v)\big) \\ &= \sum_{j=0}^{n-1} \lambda_j (\alpha^j)\big((p^i g)(\alpha)(v)\big) &&= \sum_{j=0}^{n-1} \lambda_j (\alpha^j)(0) \qquad = \ 0.\end{aligned}$$

Daher gilt $(p^i g)(\alpha) = 0 \Longrightarrow p^k / p^i g \Longrightarrow p^{k-i} / g$ und mit Lemma 13.3 ist $\operatorname{Bild} g(\alpha) \subseteq \operatorname{Bild} p^{k-i}(\alpha)$. Insgesamt existiert also für jedes $w \in \operatorname{Kern} p^i(\alpha)$ ein Polynom $g \in K[x]$ mit $w \in \operatorname{Bild} g(\alpha) \subseteq \operatorname{Bild} p^{k-i}(\alpha)$, d.h. $\operatorname{Kern} p^i(\alpha) \subseteq \operatorname{Bild} p^{k-i}(\alpha)$.

Definition

Sei $\alpha \in \operatorname{End}(V)$ und U ein α-invarianter Unterraum von V. Dann ist ein α-invariantes Komplement W von U ein α-invarianter Unterraum mit $V = U \oplus W$.

13.12 Satz

Sei das Minimalpolynom von $\alpha \in \operatorname{End}(V)$ eine Primpotenz p^k. Sei weiter U ein α-zyklischer Unterraum von V mit $\dim U = \operatorname{grad} p^k$. Dann existiert ein α-invariantes Komplement W von U.

Beweis

Offenar gilt $U \cap \{0\} = \{0\}$. Daher existiert ein bzgl. Inklusion maximaler α-invarianter Unterraum W mit $U \cap W = \{0\}$. Nach Satz 12.2 gilt dann $U + W = U \oplus W$. Für $U \oplus W = V$ folgt unmittelbar die Behauptung.

Sei also $U \oplus W < V$ ein echter Unterraum und $v \in V$ mit $v \notin U \oplus W$. Wegen $p^0(\alpha)(v) = v \notin U \oplus W$ und $p^k(\alpha)(v) = 0 \in U \oplus W$ existiert eine minimale Potenz $1 \leq s \leq k$ mit $p^s(\alpha)(v) \in U \oplus W$.

Dann existiert eine eindeutige Darstellung $p^s(\alpha)(v) = u + w$ mit $u \in U$ und $w \in W$. Es folgt $0 = p^k(\alpha)(v) = p^{k-s}(\alpha)\big((p^s)(\alpha)(v)\big) = p^{k-s}(\alpha)(u+w) = p^{k-s}(\alpha)(u) + p^{k-s}(\alpha)(w)$. Da U und W jeweils α-invariant sind, folgt $p^{k-s}(\alpha)(u) \in U$ und $p^{k-s}(\alpha)(w) \in W$. Mit der Eindeutigkeit der Darstellung von $0 \in U \oplus W$ erhält man dann $p^{k-s}(\alpha)(u) = p^{k-s}(\alpha)(w) = 0$.

Sei nun m_U das Minimalpolynom von $\alpha|_U$. Es folgt $p^k(\alpha|_U) = p^k(\alpha)(U) = \{0\} \Longrightarrow m_U / p^k$. Da p irreduzibel ist, gilt somit $m_U = p^i$ für $i \leq k$. Da U weiter α-zyklisch ist, erhält man $\dim U = \operatorname{grad} m_U$ mit Lemma 13.8. Aus der Voraussetzung $\dim U = \operatorname{grad} p^k$ folgt dann $m_U = p^k$.

Lemma 13.11 angewandt auf U und $\alpha|_U \in \operatorname{End}(U)$ liefert damit

$$\operatorname{Kern} p^{k-s}(\alpha|_U) = \operatorname{Bild} p^{k-(k-s)}(\alpha|_U) = \operatorname{Bild} p^s(\alpha|_U).$$

Es folgt $p^{k-s}(\alpha|_U)(u) = p^{k-s}(\alpha)(u) = 0 \Longrightarrow u \in \operatorname{Kern} p^{k-s}(\alpha|_U) \Longrightarrow u \in \operatorname{Bild} p^s(\alpha|_U) \Longrightarrow u = p^s(\alpha|_U)(y) = p^s(\alpha)(y)$ für ein $y \in U$. Man setze nun $z = v - y$, d.h. $p^s(\alpha)(z) = p^s(\alpha)(v) - p^s(\alpha)(y) = u + w - u = w \in W$.

Sei $p^{s-1}(\alpha)(z) \in U \oplus W$. Da U α-invariant ist, folgt mit $y \in U$ auch $p^{s-1}(\alpha)(y) \in U$. Insgesamt ergibt sich also $p^{s-1}(\alpha)(v) = p^{s-1}(\alpha)(v) - p^{s-1}(\alpha)(y) + p^{s-1}(\alpha)(y) = p^{s-1}(\alpha)(z) + p^{s-1}(\alpha)(y) \in U \oplus W$ – ein Widerspruch zur Wahl von s. Daher gilt $p^{s-1}(\alpha)(z) \notin U \oplus W$.

Man setze nun $W_1 = W + \langle z \rangle_\alpha$. Man betrachte dann für $d = \operatorname{grad} p^s - 1$ die Division mit Rest $x^i = q_i p^s + r_i$ mit $\operatorname{grad} r_i < d + 1$. Dabei gilt $\alpha^i(z) = q_i(\alpha)\big(p^s(\alpha)(z)\big) + r_i(\alpha)(z) = q_i(\alpha)(w) + r_i(\alpha)(z)$ und $r_i = a_0 + a_1 x + \ldots + a_d x^d$. Da W α-invariant ist, ist W nach Lemma 13.3 auch $q_i(\alpha)$-invariant – d.h. $q_i(\alpha)(w) = \widetilde{w} \in W$. Es folgt $\alpha^i(z) = \widetilde{w} + a_0 z + a_1 \alpha(z) + \ldots + a_d \alpha^d(z) \in W + \langle z, \alpha(z), \ldots, \alpha^d(z) \rangle$. Da dies für alle $i \in \mathbb{N}$ gilt, erhält man $W_1 = W + \langle z \rangle_\alpha = W + \langle z, \alpha(z), \ldots, \alpha^d(z) \rangle$.

Da W und $\langle z \rangle_\alpha$ jeweils α-invariant sind, folgt $w_1 \in W_1 \Longrightarrow w_1 = \widetilde{w} + \widetilde{z}$ mit $\widetilde{w} \in W$ und $\widetilde{z} \in \langle z \rangle_\alpha \Longrightarrow \alpha(w_1) = \alpha(\widetilde{w}) + \alpha(\widetilde{z}) \in W + \langle z \rangle_\alpha = W_1$. Daher ist auch W_1 ein α-invarianter Unterraum. Sei nun $z \in W$. Dann gilt mit $y \in U$ auch $v = z + y \in U + W$ – im Widerspruch zur Wahl von v, d.h. $z \notin W$. Wegen $z \in W + \langle z \rangle_\alpha = W_1$ und $W \subseteq W + \langle z \rangle_\alpha = W_1$ folgt $\dim W < \dim W_1$.

Nun ist aber W ein bzgl. Inklusion maximaler α-invarianter Unterraum mit $U \cap W = \{0\}$, d.h. es existiert ein $0 \neq \overline{u} \in U \cap W_1$. Insbesondere ist $\overline{u} \in W_1 \Longrightarrow \overline{u} = \overline{w} + b_0 z + b_1 \alpha(z) + \ldots + b_d \alpha^d(z)$ mit $\overline{w} \in W$ und $b_i \in K$. Damit ist $\overline{u} = \overline{w} + g(\alpha)(z)$ für ein Polynom $g = b_0 + b_1 x + \ldots + b_d x^d \in K[x]$ mit $\operatorname{grad} g \leq d = \operatorname{grad} p^s - 1 \Longrightarrow \operatorname{grad} g < \operatorname{grad} p^s$. Es gilt weiter $g(\alpha)(z) = \overline{u} - \overline{w}$.

Sei nun nach Bézout $t = \operatorname{ggT}(g, p^k) = Ag + Bp^k$ für $A, B \in K[x]$. Da U bzw. W α-invariant ist,

ist U bzw. W nach Lemma 13.3 auch $A(\alpha)$-invariant. Damit gilt $t(\alpha)(z) = A(\alpha)\big(g(\alpha)(z)\big) + B(\alpha)\big(p^k(\alpha)(z)\big) = A(\alpha)(\overline{u} - \overline{w}) + B(\alpha)(0) = A(\alpha)(\overline{u}) - A(\alpha)(\overline{w}) \in U + W$.

Man nehme $g = 0$ an. Dann ergibt sich durch $\overline{u} = \overline{w} + g(\alpha)(z) = \overline{w} \in W \Longrightarrow \overline{u} \in U \cap W \Longrightarrow u = 0$ ein Widerspruch, d.h. $g \neq 0$. Aus $t = \mathrm{ggT}(g, p^k)/g$ folgt daher $\operatorname{grad} t \leq \operatorname{grad} g < \operatorname{grad} p^s = s \cdot \operatorname{grad} p$. Da p irreduzibel ist, erhält man weiter $t = \mathrm{ggT}(g, p^k)/p^k \Longrightarrow t = p^i$ mit $i \cdot \operatorname{grad} p = \operatorname{grad} p^i = \operatorname{grad} t < s \cdot \operatorname{grad} p \Longrightarrow i < s$.

Da nun U ein α-invarianter Unterraum ist, ist U auch $p^i(\alpha)$-invariant. Daher folgt $y \in U \Longrightarrow p^i(\alpha)(y) \in U$. Mit $p^i(\alpha)(z) = t(\alpha)(z) \in U + W$ gilt schließlich $p^i(\alpha)(v) = p^i(\alpha)(v - y + y) = p^i(\alpha)(z + y) = p^i(\alpha)(z) + p^i(\alpha)(y) \in U + W$. Dies liefert wegen $i < s$ einen Widerspruch zur Wahl von s – d.h. $V = U \oplus W$.

13.13 Definition

Ein Vektorraum V mit $\alpha \in \mathrm{End}(V)$ heißt α-unzerlegbar, falls V keine direkte Summe α-invarianter Unterräume $U \neq \{0\}$ ist.

13.14 Satz

Sei das Minimalpolynom von $\alpha \in \mathrm{End}(V)$ eine Primpotenz p^k.

1. Dann ist V genau dann α-unzerlegbar, wenn $\operatorname{grad} p^k = \dim V$ gilt.
2. Sei V α-unzerlegbar und für $0 \leq i \leq k$ sei $U_i = \operatorname{Kern} p^i(\alpha)$. Dann gilt $\dim U_i = i \cdot \operatorname{grad} p$. Weiter ist U_i α-unzerlegbar und jeder α-invariante Unterraum ist einer der U_i.

Beweis

1 Sei zunächst $\dim V = \operatorname{grad} p^k$. Dann ist nach Lemma 13.11 V keine direkte Summe α-invarianter Unterräume U mit $\dim U < \dim V$ und somit α-unzerlegbar.

Sei nun V α-unzerlegbar. Da p irreduzibel ist, gilt $\operatorname{grad} p^k > 0$. Nach Satz 13.10 gilt allgemein $\dim p^k \leq \dim V$. Man nehme dabei nun $\dim p^k < \dim V$ an.

Nach Lemma 13.9 existiert ein α-zyklischer Unterraum U mit $\dim U = \operatorname{grad} p^k$, d.h. $0 < \dim U < \dim V$. Nach Satz 13.12 existiert ein α-invariantes Komplement W von U mit $\dim V = \dim U + \dim W \Longrightarrow 0 < \dim W < \dim V$. Insgesamt ist also V im Widerspruch zur Annahme α-zerlegbar.

2 Sei $v \in U_i = \operatorname{Kern} p^i(\alpha)$. Dann folgt $p^{i+1}(\alpha)(v) = p(\alpha)\big(p^i(\alpha)(v)\big) = p(\alpha)(0) = 0 \Longrightarrow v \in \operatorname{Kern} p^{i+1}(\alpha) = U_{i+1}$, d.h.

$$(*) \quad U_i \subseteq U_{i+1}.$$

Da V α-unzerlegbar ist, gilt nach 1 nun $\dim V = \operatorname{grad} p^k$ und mit Lemma 13.11 gilt folgt $U_i = \operatorname{Bild} p^{k-i}(\alpha)$ und $U_{i+1} = \operatorname{Bild} p^{k-i-1}(\alpha)$. Man erhält daher

$$U_i = p^{k-i}(\alpha)(V) = p(\alpha)\big(p^{k-i-1}(\alpha)(V)\big) = p(\alpha)(U_{i+1}).$$

Somit ist die lineare Abbildung $p(\alpha)|_{U_{i+1}} : u \in U_{i+1} \longmapsto p(\alpha)(u) \in U_i$ surjektiv und mit der Dimensionsformel folgt $\dim U_{i+1} = \dim \operatorname{Bild} p(\alpha)|_{U_{i+1}} + \dim \operatorname{Kern} p(\alpha)|_{U_{i+1}} = \dim U_i + \dim \operatorname{Kern} p(\alpha)|_{U_{i+1}}$.

Nun gilt $v \in \operatorname{Kern} p(\alpha)|_{U_{i+1}} \Longleftrightarrow v \in U_{i+1} \ \wedge \ p(\alpha)(v) = 0 \Longleftrightarrow v \in U_{i+1} \cap \operatorname{Kern} p(\alpha) = U_1$, d.h. $\operatorname{Kern} p(\alpha)|_{U_{i+1}} = U_{i+1} \cap U_1$. Mit $(*)$ folgt sukzessive $U_1 \subseteq \ldots \subseteq U_{i+1}$ und man erhält $\operatorname{Kern} p(\alpha)|_{U_{i+1}} = U_1$, also

$$(**) \quad \dim U_{i+1} = \dim U_i + \dim U_1 \Longleftrightarrow \dim U_{i+1} - \dim U_i = \dim U_1.$$

Die Summation von $(**)$ für alle $0 \leq i \leq k-1$ liefert dann $\dim U_k - \dim U_0 = \dim U_k - \dim U_{k-1} + \dim U_{k-1} - \ldots - \dim U_1 + \dim U_1 - \dim U_0 = k \cdot \dim U_1$. Dabei gilt $\dim U_k = \dim \operatorname{Kern} p^k(\alpha) = \dim \operatorname{Kern} 0 = \dim V = \operatorname{grad} p^k$ sowie $\dim U_0 = \dim \operatorname{Kern} p^0(\alpha) = \dim \operatorname{Kern} \mathrm{id} = \dim\{0\} = 0$, d.h.

$$\operatorname{grad} p^k = k \cdot \dim U_1 \Longrightarrow \operatorname{grad} p = \dim U_1.$$

Die Summation von $(**)$ für alle $0 \leq i \leq j-1$ liefert analog $\dim U_j = j \cdot \dim U_1 = j \cdot \operatorname{grad} p$.

Wegen $U_i = \operatorname{Bild} p^{k-i}(\alpha) = \operatorname{Bild} \frac{p^k}{p^i}(\alpha)$ ist p^i nach Lemma 13.4 das Minimalpolynom von $\alpha|_{U_i}$. Daher ist U_i wegen $\dim U_i = i \cdot \operatorname{grad} p = \operatorname{grad} p^i$ nach 1 α-unzerlegbar.

Sei nun $U \subseteq V$ ein beliebiger α-invarianter Unterraum und $\overline{m}$ das Minimalpolynom von $\alpha|_U$. Dabei gilt $U_k = V$, d.h. wir können o.B.d.A. $U \subsetneq V$ annehmen. Mit Satz 13.10 folgt dann $\operatorname{grad} \overline{m} \leq \dim U < \dim V = \operatorname{grad} p^k$. Wegen $p^k(\alpha)(U) = \{0\} \Longrightarrow \overline{m}/p^k$ erhält man dann $\overline{m} = p^i$ für $i < k$. Damit ergibt sich

$$(***) \quad U = \operatorname{Kern} p^i(\alpha|_U) \subseteq \operatorname{Kern} p^i(\alpha) = U_i.$$

Die Aussage folgt nun durch vollständige Induktion über $\dim V$. Für $\dim V = 1$ ist der Fall wegen $\dim U < \dim V \Longrightarrow \dim U = 0 \Longrightarrow U = \{0\} = U_0$ klar. Mit $i < k$ gilt zuerst $\dim U_i = i \cdot \operatorname{grad} p < k \cdot \operatorname{grad} p = \operatorname{grad} p^k = \dim V$. Weiter ist p^i das Minimalpolynom von $\alpha|_{U_i}$ und U_i ist α-unzerlegbar, d.h. insbesondere $\alpha|_{U_i}$-unzerlegbar.

Da U wegen $(***)$ ein α-invarianter Unterraum von U_i ist, erhält man für $j \leq i$ nach Induktionsannahme $U = \operatorname{Kern} p^j(\alpha|_{U_i}) = \{v \in U_i \mid p^j(\alpha|_{U_i})(v) = 0\} = \{v \in U_i \mid p^j(\alpha)(v) =$

$0\} = \{v \in V \mid p^j(\alpha)(v) = 0\} \cap U_i = \operatorname{Kern} p^j(\alpha) \cap \operatorname{Kern} p^i(\alpha) = \operatorname{Kern} \operatorname{ggT}(p^j, p^i)(\alpha) = \operatorname{Kern} p^j(\alpha) = U_j$.

13.15 Lemma

Sei V für $\alpha \in \operatorname{End}(V)$ ein α-unzerlegbarer Vektorraum. Dann ist V α-zyklisch und das Minimalpolynom von α ist eine Primpotenz p^k mit $\operatorname{grad} p^k = \dim V$.

Beweis

Sei m das Minimalpolynom von α. Für $m = pq$ mit $\operatorname{ggT}(p, q) = 1$ gilt nach Satz 13.5 nun $V = \operatorname{Bild} p(\alpha) \oplus \operatorname{Bild} q(\alpha)$. Da V aber α-unzerlegbar ist, folgt o.B.d.A. $\operatorname{Bild} p(\alpha) = \{0\} \Longrightarrow p = m \Longrightarrow q = 1$. Daher ist m eine Primpotenz. Nach Satz 13.4 gilt dann $\operatorname{grad} m = \dim V$ und V ist nach Lemma 13.9 α-zyklisch.

Bemerkung

Ein α-zyklischer Vektorraum V mit $\alpha \in \operatorname{End}(V)$ ist jedoch i.A. nicht α-unzerlegbar.

13.16 Lemma

Sei $\alpha \in \operatorname{End}(V)$. Dann ist V eine direkte Summe α-unzerlegbarer Unterräume U_j. Sei $m = \prod p_i^{k_i}$ die Primzerlegung des Minimalpolynoms von α. Setze $V_{p_i} = \operatorname{Kern} p_i^{k_i}(\alpha)$. Dann gilt $V = \bigoplus V_{p_i}$ und jeder Unterraum V_{p_i} ist eine direkte Summe gewisser U_j. Weiter ist $p_i^{k_i}$ das Minimalpolynom von $\alpha|_{V_{p_i}}$.

Beweis

Zunächst ist V entweder selbst α-unzerlegbar oder eine direkte Summe $V = U \oplus W$ mit $\dim U < \dim V$ und $\dim W < \dim V$. Nun beginne man für U und W erneut. Wegen $\dim V < \infty$ und $\dim \bigoplus U_j = \sum \dim U_j$ erhält man schließlich eine endliche Zerlegung. Weiter folgt aus Lemma 13.5 unmittelbar $V = \bigoplus \operatorname{Kern} p_i^{k_i}(\alpha) = \bigoplus V_{p_i}$.

Sei nun U_j ein α-unzerlegbarer Summand von $V = \bigoplus U_j$ und sei m_{U_j} das Minimalpolynom von $\alpha|_{U_j}$. Nach Lemma 13.5 ist m_{U_j} eine Primpotenz. Wegen $m(U_j) = \{0\} \Longrightarrow m_{U_j}/m$ ist daher $m_{U_j} = p_r^l$ für $l \leq k_r$. Es folgt $u \in U_j \Longrightarrow p_r^l(\alpha)(u) = 0 \Longrightarrow p_r^{k_r}(\alpha)(u) = p_r^{k_r - l}(\alpha)\big(p_r^l(\alpha)(u)\big) = p_r^{k_r - l}(\alpha)(0) = 0 \Longrightarrow u \in \operatorname{Kern} p_r^{k_r}(\alpha) = V_{p_r}$, also $U_j \subseteq V_{p_r}$. Wegen $V = \bigoplus V_{p_i}$ gilt dann $V_{p_r} \cap V_{p_s} = \{0\} \Longrightarrow U_j \cap V_{p_s} = \{0\}$ für alle $s \neq r$.

Damit folgt für ein festes i und alle j stets $V_{p_i} \cap U_j = U_j$ oder $V_{p_i} \cap U_j = \{0\}$. Damit ist $\bigoplus(V_{p_i} \cap U_j)$ eine direkte Summe gewisser U_j. Nach Lemma 13.15 und Lemma 13.7 sind die U_j α-invariant. Mit Lemma 13.3 folgt damit schließlich $V_{p_i} = \operatorname{Kern} p_i^{k_i}(\alpha) = \bigoplus \big(U_j \cap \operatorname{Kern} p_i^{k_i}(\alpha)\big) = \bigoplus(U_j \cap V_{p_i})$. Zuletzt ist $p_i^{k_i}$ wegen $V_{p_i} = \operatorname{Kern} p_i^{k_i}(\alpha)$ nach Lemma 13.5 das Minimalpolynom von $\alpha|_{V_{p_i}}$.

Bemerkung

Die Zerlegung $V = \bigoplus U_j$ in Lemma 13.16 ist in der Regel nicht eindeutig. Wir sehen jedoch in Lemma 13.17, dass die Dimensionen der U_j und die Minimalpolynome der $\alpha|_{U_i}$ eindeutig sind.

13.17 Lemma

Sei das Minimalpolynom von $\alpha \in \operatorname{End}(V)$ eine Primpotenz p^k. Sei $V = U_1 \oplus \ldots \oplus U_n$ eine direkte Summe α-unzerlegbarer Unterräume. Dann ist p^{k_i} mit $k_i \leq k$ das Minimalpolynom von $\alpha|_{U_i}$ und es gilt $\dim U_i = k_i \cdot \operatorname{grad} p$. Dabei ist das Tupel $(k_1, \ldots, k_n)$ bis auf Permutation eindeutig. Ferner gilt $k_i = k$ für mindestens ein i.

Beweis

Da stets $p^k(\alpha)(U_i) = \{0\}$ gilt, ist das Minimalpolynom von $\alpha|_{U_i}$ ein Teiler von p^k und damit eine Primpotenz p^{k_i} mit $k_i \leq k$. Mit der α-Unzerlegbarkeit von U_i und Lemma 13.15 erhält man $\dim U_i = \operatorname{grad} p^{k_i} = k_i \cdot \operatorname{grad} p$.

Man nehme nun $k_i < k \iff k_i \leq k-1$ für alle $1 \leq i \leq n$ an. Dann gilt $p^{k-1}(\alpha)(U_i) = p^{k-1-k_i}(\alpha)\big(p^{k_i}(\alpha)(U_i)\big) = \{0\}$ und daher $v \in V = U_1 \oplus \ldots \oplus U_n \Longrightarrow v = u_1 + \ldots + u_n \Longrightarrow p^{k-1}(\alpha)(v) = p^{k-1}(\alpha)(u_1) + \ldots + p^{k-1}(\alpha)(u_n) = 0 \Longrightarrow p^{k-1}(\alpha)(V) = \{0\} \Longrightarrow p^{k-1}(\alpha) = 0$. Mit $\operatorname{grad} p^{k-1} < \operatorname{grad} p^k$ liefert dies einen Widerspruch, d.h. es gilt $k_i = k$ für ein $1 \leq i \leq n$.

Sei nun o.B.d.A. $k_1 \leq \ldots \leq k_n$ und $0 \leq t \leq k$. Nach Lemma 13.15 ist U_i α-zyklisch, d.h. es gilt $U_i = \langle u \rangle_\alpha$ für ein $u \in U_i$. Nach Lemma 7.4 gilt $p^t(\alpha)(U_i) = p^t(\alpha)(\langle u \rangle_\alpha) = \langle p^t(\alpha)(u) \rangle_\alpha$, d.h. $p^t(\alpha)(U_i)$ ist ebenfalls α-zyklisch.

Mit $p^t(\alpha)(U_i) = \operatorname{Bild} p^t(\alpha|_{U_i}) = \operatorname{Bild} \frac{p^{k_i}}{p^{k_i-t}}(\alpha|_{U_i})$ und Lemma 13.4 ist p^{k_i-t} für $t \leq k_i$ das Minimalpolynom von $\alpha|_{p^t(\alpha)(U_i)}$. Nach Lemma 13.8 gilt dann für $t \leq k_i$ auch $\dim p^t(\alpha)(U_i) = \operatorname{grad} p^{k_i-t} = (k_i - t) \cdot \operatorname{grad} p$. Für $t > k_i$ folgt $p^t(\alpha)(U_i) = p^{t-k_i}(\alpha)\big(p^{k_i}(\alpha)(U_i)\big) = \{0\}$, also $\dim p^t(\alpha)(U_i) = 0$.

Sei nun $v \in V = U_1 \oplus \ldots \oplus U_n$ definiert durch $v = u_1 + \ldots + u_n$. Dann gilt $p^t(\alpha)(v) = p^t(\alpha)(u_1) +$

$\ldots + p^t(\alpha)(u_n) \in p^t(\alpha)(U_i) + \ldots + p^t(\alpha)(U_n) \Longrightarrow p^t(\alpha)(V) = p^t(\alpha)(U_i) + \ldots + p^t(\alpha)(U_n)$. Nach Lemma 13.7 ist U_i α-invariant und damit nach Lemma 13.3 auch $p^t(\alpha)$-invariant, d.h. $p^t(\alpha)(u_i) \in U_i$. Daher ist die Darstellung $p^t(\alpha)(v) = p^t(\alpha)(u_1) + \ldots + p^t(\alpha)(u_n)$ eindeutig und es folgt $p^t(\alpha)(V) = p^t(\alpha)(U_i) \oplus \ldots \oplus p^t(\alpha)(U_n)$. Insgesamt erhält man also

$$\dim p^t(\alpha)(V) = \sum_{i=1}^{n} \dim p^t(\alpha)(U_i) = \sum_{k_i \geq t} (k_i - t) \cdot \operatorname{grad} p.$$

Für $0 \leq r \leq k$ sei $\pi(r) = |\{i \mid k_i = r\}|$. Dann gilt $\pi(r)(r-t) = \sum\limits_{k_i = r} (k_i - t)$ und daher

$$\sum_{k_i \geq t} (k_i - t) = \sum_{r=t}^{k} \pi(r)(r-t) \Longrightarrow \frac{\dim p^t(\alpha)(V)}{\operatorname{grad} p} = \sum_{r=t}^{k} \pi(r)(r-t).$$

Für $t = k-1$ erhält man $\frac{\dim p^{k-1}(\alpha)(V)}{\operatorname{grad} p} = \pi(k-1) \cdot (k-1-k+1) + \pi(k) \cdot (k-k+1) = \pi(k)$. Damit ist $\pi(k)$ eindeutig bestimmt. Sukzessive ergibt sich nun für $0 \leq t < k-1$ eindeutig $\pi(k-1), \ldots, \pi(1)$. Mit $\dim U_i = k_i \cdot \operatorname{grad} p$ und $V = U_1 \oplus \ldots \oplus U_n \Longrightarrow \dim U_i > 0$ gilt schließlich $k_i \cdot \operatorname{grad} p > 0 \Longrightarrow k_i \neq 0 \Longrightarrow \pi(0) = 0$. Somit ist die Zuordnung π von der gewählten Zerlegung $V = U_1 \oplus \ldots \oplus U_n$ unabhängig und es folgt die Behauptung.

13.18 Lemma / Definition

Sei $V = \langle v \rangle_\alpha$ ein α-zyklischer Vektorraum der Dimension n und sei $m = a_0 + a_1 x + \ldots + a_{n-1}x^{n-1} + x^n$ das Minimalpolynom von α. Dann besitzt α bzgl. der Basis $v, \alpha(v), \ldots, \alpha^{n-1}(v)$ die Matrixdarstellung

$$\begin{pmatrix} & & & & -a_0 \\ 1 & & & & -a_1 \\ & 1 & & & -a_2 \\ & & 1 & & -a_3 \\ & & & \ddots & \vdots \\ & & & 1 & -a_{n-1} \end{pmatrix}.$$

Eine solche Matrix heißt Frobenius-Matrix zum Polynom m.

Beweis

Sei $\alpha^{i-1}(v) = e_i$. Für $1 \leq i \leq n-1$ gilt dann $\alpha(e_i) = \alpha\big(\alpha^{i-1}(v)\big) = \alpha^i(v) = e_{i+1}$. Damit ergeben sich die ersten $n-1$ Spalten der Darstellungsmatrix. Die letzte Spalte erhält man mit $a_0 v + a_1\alpha(v) + \ldots + a_{n-1}\alpha^{n-1}(v) + \alpha^n(v) = m(\alpha)(v) = 0 \Longrightarrow \alpha(e_n) = \alpha\big(\alpha^{n-1}(v)\big) = \alpha^n(v) = -a_0 v - a_1\alpha(v) - \ldots - a_{n-1}\alpha^{n-1}(v) = -a_0 e_1 - a_1 e_2 - \ldots - a_{n-1}e_n$.

13.19 Lemma

Das Minimalpolynom m der Frobenius-Matrix

$$A = \begin{pmatrix} & & & & -a_0 \\ 1 & & & & -a_1 \\ & 1 & & & -a_2 \\ & & 1 & & -a_3 \\ & & & \ddots & \vdots \\ & & & 1 & -a_{n-1} \end{pmatrix}$$

ist $a_0 + a_1 x + \ldots + a_{n-1}x^{n-1} + x^n$.

Beweis

Man betrachte die lineare Abbildung $v \in K^n \longmapsto Av$ und die Standardbasis $e_1, \ldots, e_n$ von K^n. Für $1 \leq i < n$ gilt dann $Ae_i = e_{i+1}$, d.h. $K^n = \langle e_1, \ldots, e_n \rangle = \langle e_1 \rangle_A$. Nach Lemma 13.8 gilt dann $\operatorname{grad} m = \dim K^n = n$. Mit $(a_0 + a_1 x + \ldots + a_{n-1}x^{n-1} + x^n)(A)e_1 = a_0 e_1 + a_1 A e_1 + \ldots + a_{n-1}A^{n-1}e_1 + A^n e_1 = a_0 + a_1 e_2 + \ldots + a_{n-1}e_n + Ae_n = 0$ und $K^n = \langle e_1 \rangle_A$ folgt dann die Behauptung.

13.20 Lemma

Sei V für $\alpha \in \operatorname{End}(V)$ ein α-zyklischer Vektorraum. Dann stimmen Minimalpolynom und charakteristisches Polynom von α überein.

Beweis

Sei $a_0 + a_1 x + \ldots + a_{n-1}x^{n-1} + x^n$ das Minimalpolynom von α. Nach Lemma 13.18 lässt sich dann α durch

$$A = \begin{pmatrix} & & & & -a_0 \\ 1 & & & & -a_1 \\ & 1 & & & -a_2 \\ & & 1 & & -a_3 \\ & & & \ddots & \vdots \\ & & & 1 & -a_{n-1} \end{pmatrix}$$

beschreiben. Mit Bemerkung 10.7 folgt

$$\chi_\alpha(x) = \chi_A(x) = \det(x1_n - A) = \det \begin{pmatrix} x & & & & & a_0 \\ -1 & x & & & & a_1 \\ & -1 & x & & & a_2 \\ & & -1 & x & & a_3 \\ & & & \ddots & & \vdots \\ & & & & 1 & x + a_{n-1} \end{pmatrix}.$$

Durch Enwicklung der Determinante nach der erste Zeile erhält man zunächst

$$\chi_\alpha(x) = x \cdot \det \begin{pmatrix} x & & & & a_1 \\ -1 & x & & & a_2 \\ & -1 & x & & a_3 \\ & & \ddots & & \vdots \\ & & & 1 & x + a_{n-1} \end{pmatrix} + (-1)^{1+n} a_0 \cdot \det \begin{pmatrix} -1 & x & & & \\ & -1 & x & & \\ & & -1 & x & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}$$

und durch vollständige Induktion über n folgt

$$\begin{aligned} \chi_\alpha(x) &= x \cdot (a_1 + a_2 x + \ldots + a_{n-1}x^{n-2} + x^{n-1}) + (-1)^{1+n} a_0 (-1)^{n+1} \\ &= a_0 + a_1 x + \ldots + a_{n-1}x^{n-1} + x^n. \end{aligned}$$

13.21 Satz (Frobenius-Normalform)

Sei $\dim V < \infty$ und $\alpha \in \mathrm{End}(V)$. Sei $m = p_1^{e_1} \cdots p_r^{e_r}$ die Primfaktorzerlegung des Minimalpolynoms von α. Dann lässt sich α durch eine Blockdiagonalmatrix

$$\begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_s \end{pmatrix}$$

beschreiben. Dabei sind die A_k Frobenius-Matrizen zu Polynomen $p_i^{f_i}$ mit $1 \leq f_i \leq e_i$ und eindeutig bis auf ihre Reihenfolge. Weiter gibt es für alle $1 \leq i \leq r$ eine Matrix A_k zu $p_i^{e_i}$.

Beweis

Nach Satz 13.5 ist $V = \bigoplus_{i=1}^{r} U_i$ die direkte Summe α-invarianter Unterräume $U_i = \mathrm{Kern}\, p_i^{e_i}(\alpha)$, wobei $p_i^{e_i}$ das Minimalpolynom von $\beta_i = \alpha|_{U_i}$ ist. Daher lässt sich α nach Satz 12.4 durch eine Blockdiagonalmatrix aus den Darstellungsmatrizen B_i von β_i beschreiben.

Sei nun $U_i = \bigoplus_{j=1}^{t} U_{ij}$ eine direkte Summe β_i-unzerlegbarer Unterräume. Da U_{ij} nach Lemma 13.15 β_i-zyklisch ist, wird β_i nach Satz 12.4 und Lemma 13.18 durch eine Blockdiagonalmatrix aus Frobenius-Matrizen A_k dargestellt. Mit Lemma 13.17 sind die A_k bis auf Permutation eindeutig und es existiert eine Matrix A_k zu $p_i^{e_i}$.

Bemerkung

Für die quadratische Matrix $A \in M_{n\times n}(K)$ betrachte man den Endomorphismus $\alpha : v \in K^n \longmapsto Av$. Dann lässt sich α durch eine Matrix F in Frobenius-Normalform beschreiben und A ist ähnlich zu F.

13.22 Satz (Cayley-Hamilton)

Sei χ_A das charakteristische Polynom von $A \in M_{n\times n}(K)$. Dann gilt $\chi_A(A) = 0$, d.h. das Minimalpolynom von A ist eine Teiler von χ_A. Ferner ist jeder irreduzible Teiler von χ_A ein Teiler des Minimalpolynoms von A.

Beweis

Zunächst ist A ähnlich zu einer Matrix

$$A' = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_s \end{pmatrix}$$

in Frobenius-Normalform. Mit Lemma 13.6 gilt dann $\chi_A(x) = \chi_{A'}(x)$ und mit Satz 9.17 folgt

$$\det(1_n x - A') = \det \begin{pmatrix} 1_{n_1}x - A_1 & & \\ & \ddots & \\ & & 1_{n_x}x - A_s \end{pmatrix} = \det(1_{n_1}x - A_1) \cdots \det(1_{n_s}x - A_s),$$

d.h. $\chi_A(x) = \chi_{A_1}(x) \cdots \chi_{A_s}(x)$. Da nach Satz 13.21 zu jedem Teiler $p_i^{e_i}$ des Minimalpolynoms m von A eine Frobenius-Matrix A_k existiert, gilt nach Lemma 13.19 und Lemma 13.20 dann $m/\chi_A \Longrightarrow \chi_A(A) = 0$.

Andererseits sind alle A_k Frobenius-Matrizen zu Teilern $p_i^{f_i}$ von m. Die irreduziblen Teiler sind daher gerade die p_i, also Teiler von m.

Bemerkung

Der „Beweis“ $\chi_A(x) = \det(1_n x - A)$ „$\Longrightarrow$“ $\chi_A(A) = \det(A1_n - A) = \det(A - A) = \det 0 = 0$ von Satz 13.22 ist falsch.

13.23 Lemma

Das Minimalpolynom eines Endomorphismus $\alpha \in \mathrm{End}(V)$ zerfällt genau dann in Linearfaktoren $m_\alpha = (x - \lambda_1)^{f_1} \cdots (x - \lambda_r)^{f_r}$, wenn das charakteristische Polynom in Linearfaktoren $\chi_\alpha = (x - \lambda_1)^{e_1} \cdots (x - \lambda_r)^{e_r}$ für $e_i \geq f_i$ zerfällt.

Beweis

Sei $\lambda \in K$ eine Nullstelle des Minimalpolynoms m_α. Mit $m_\alpha / \chi_\alpha \Longrightarrow \chi_\alpha = m_\alpha \cdot h$ für ein $h \in K[x]$ folgt dann $\chi_\alpha(\lambda) = m_\alpha(\lambda) \cdot h(\lambda) = 0 \cdot h(\lambda) = 0$. Damit ist nach Satz 11.5 jeder Linearfaktor $x - \lambda$ von m_α ein Linearfaktor von χ_α.

Andererseits sind die Linearfaktoren von χ_α irreduzible Teiler und damit auch Teiler von m_α. Wegen m_α / χ_α folgt schließlich $f_i \leq e_i$ für $1 \leq i \leq r$.

13.24 Lemma

Seien paarweise ähnliche Matrizen $A_1, \ldots, A_r, B_1, \ldots, B_r$ und invertierbare Matrizen $T_1, \ldots, T_r$ mit $A_i = T_i^{-1} B_i T_i$ gegeben. Dann sind auch die Blockdiagonalmatrizen

$$A = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_r \end{pmatrix} \text{ und } B = \begin{pmatrix} B_1 & & \\ & \ddots & \\ & & B_r \end{pmatrix}$$

ähnlich.

Beweis

Mit

$$A = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_r \end{pmatrix} = \begin{pmatrix} T_1^{-1} B_1 T_1 & & \\ & \ddots & \\ & & T_r^{-1} B_r T_r \end{pmatrix}$$

folgt

$$A = \begin{pmatrix} T_1^{-1} & & \\ & \ddots & \\ & & T_r^{-1} \end{pmatrix} B \begin{pmatrix} T_1 & & \\ & \ddots & \\ & & T_r \end{pmatrix}.$$

Dabei gilt

$$\begin{pmatrix} T_1^{-1} & & \\ & \ddots & \\ & & T_r^{-1} \end{pmatrix} = \begin{pmatrix} T_1^{-1} & & \\ & \ddots & \\ & & T_r^{-1} \end{pmatrix} = \begin{pmatrix} T_1 & & \\ & \ddots & \\ & & T_r \end{pmatrix}^{-1}$$

und die Behauptung folgt.

13.25 Bemerkung / Definition

Offenbar ist λ der einzige Eigenwert der Matrix

$$J_{n,\lambda} = \begin{pmatrix} \lambda & & & \\ 1 & \lambda & & \\ & \ddots & \ddots & \\ & & 1 & \lambda \end{pmatrix}.$$

Daher heißt $J_{n,\lambda}$ Jordan-Matrix zum Eigenwert λ.

13.26 Lemma

Sei A eine Frobenius-Matrix zu $(x-\lambda)^n$. Dann ist A ähnlich zu der Jordan-Matrix $J_{n,\lambda}$.

Beweis

Sei $e_1, \ldots, e_n$ die Standardbasis K^n. Für $1 \leq i \leq n-1$ gilt dann $Ae_i = e_{i+1}$, d.h. $K^n = \langle e_1 \rangle_A$ ist A-zyklisch. Man setze nun $v_i = (A - \lambda 1_n)^{i-1} e_1$ für $1 \leq i \leq n-1$.

Für $b_i \in K$ und $0 = b_1 v_1 + \ldots + b_n v_n = b_1 e_1 + b_2(A - \lambda 1_n)e_1 + \ldots + b_n(A - \lambda 1_n)^{n-1} e_1 = \big(b_1 \cdot 1_n + b_2(A - \lambda 1_n) + \ldots + b_n(A - \lambda 1_n)^{n-1}\big)e_1$ sei dann

$$B = b_1 \cdot 1_n + b_2(A - \lambda 1_n) + \ldots + b_n(A - \lambda 1_n)^{n-1}.$$

Offenbar gilt $B \in \langle A^0, A^1, \ldots, An-1 \rangle$, d.h. $B = d_0 A^0 + \ldots + d_{n-1} A^{n-1}$ für $d_i \in K$. Es folgt

$$0 = Be_1 = d_0 A^0 e_1 + \ldots + d_{n-1} A^{n-1} e_1.$$

Wegen $K^n = \langle e_1 \rangle_A$ ist dies eine Basisdarstellung, d.h. $d_0 = \ldots = d_{n-1} = 0$. Man erhält also $B = 0$ und damit

$$b_1 \cdot 1_n + b_2(A - \lambda 1_n) + \ldots + b_n(A - \lambda 1_n)^{n-1} = 0.$$

Da K^n aber A-zyklisch ist, hat das Minimalpolynom von A Grad n und es folgt $b_1 + b_2(x - \lambda) + \ldots + b_n(x - \lambda)^{n-1} = 0 \Longrightarrow b_1 = \ldots = b_n = 0$. Insgesamt sind also $v_1, \ldots, v_n$ linear unabhängig und bilden damit eine Basis von K^n.

Nun gilt $Av_i = A(A - \lambda 1_n)^{i-1} e_1 = (A - \lambda 1_n + \lambda 1n)(A - \lambda 1_n)^{i-1} e_1 = (A - \lambda 1_n)^i e_1 + \lambda(A - \lambda 1_n)^{i-1} e_1 = (A - \lambda 1_n)^i e_1 + \lambda v_i$ und daher $Av_i = v_{i+1} + \lambda v_i$ für $1 \leq i \leq n-1$. Da nach Lemma 13.19 $(x - \lambda)^n$ das Minimalpolynom von A ist, folgt $Av_n = (A - \lambda 1_n)^n e_1 + \lambda v_n = \lambda v_n$.

Der Endomorphismus $v \in K^n \longmapsto Av$ wird somit bzgl. der Basis $v_1, \ldots, v_n$ durch die Jordanmatrix $J_{n,\lambda}$ dargestellt. Da A die Standardbasisdarstellung dieses Endomorphismus ist, folgt die Behauptung.

13.27 Satz (Jordan-Normalform)

Sei $\dim V < \infty$ und $\alpha \in \mathrm{End}(V)$. Sei $\chi_\alpha = (x - \lambda_1)^{e_1} \cdots (x - \lambda_r)^{e_r}$ eine Zerlegung des charakteristischen Polynomes von α in paarweise verschiedene Linearfaktoren. Dann wird α durch eine Blockdiagonalmatrix

$$\begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_s \end{pmatrix}$$

beschrieben. Dabei sind die J_k Jordan-Matrizen J_{k_i,λ_i} für $1 \leq k_i \leq e_i$.

Beweis

Jeder Linearfaktor ist ein irreduzibler und normierter Teiler. Damit folgt die Aussage unmittelbar aus Lemma 13.23, Satz 13.21, Lemma 13.26 und Lemma 13.24.

13.28 Satz

Sei $\chi_A = (x - \lambda_1)^{l_1} \cdots (x - \lambda_r)^{l_r}$ die Zerlegung des charakteristischen Polynoms einer quadratischen Matrix $A \in M_{n \times n}(K)$ in paarweise verschiedene Linearfaktoren. Dann gilt $\det A = \lambda_1^{l_1} \cdots \lambda_r^{l_r}$.

Beweis

Nach Satz 13.27 ist A ähnlich zu einer Jordan-Normalform J, d.h. es gilt $\chi_J = \chi_A$ und $\det J = \det A$. Da weiter J eine untere Dreiecksmatrix ist, gilt $\chi_J = (x-\lambda_1)^{e_1} \cdots (x-\lambda_r)^{e_r}$ und $\det J = \lambda_1^{e_1} \cdots \lambda_r^{e_r}$. Dabei gibt jede Potenz e_i die Häufigkeit des Auftretens des Eigenwerts λ_i auf der Hauptdiagonalen von J an. Insgesamt folgt die Behauptung über $l_i = e_i$ für $1 \leq i \leq r$.

14 Bilinearformen

Bemerkung

Bilinearformen sind besondere Multilinearformen, die bereits in Kapitel 9 behandelt wurden.

14.1 Definition

Sei V ein K-Vektorraum. Eine Abbildung $\varphi : V \times V \longrightarrow K$ heißt Bilinearform, wenn für alle $w \in V$ die Abbildungen $v \in V \longmapsto \varphi(v, w) \in K$ und $v \in V \longmapsto \varphi(w, v) \in K$ linear sind.

Bemerkung

Sei $v_1, \ldots, v_n$ eine Basis von V. Sei weiter φ eine Bilinearform auf V. Dann beschreibt $C = (c_{ij}) \in M_{n\times n}(K)$ mit $c_{ij} = \phi(v_i, v_j)$ die Abbildung φ. Umgekehrt beschreibt jede Matrix $C = (c_{ij}) \in M_{n\times n}(K)$ eine Bilinearform zu einer festgewählten Basis. Man legt $\varphi(v_i, v_j) = c_{ij}$ fest und setzt φ bilinear auf V fort.

Bemerkung

Sei $C \in M_{n\times n}(K)$ bzgl. einer Basis $v_1, \ldots, v_n$ die Darstellungsmatrix einer Bilinearform φ auf V. Seien weiter $v, w \in V$ durch $v = \sum_{i=1}^{n} \alpha_i v_i$ und $w = \sum_{j=1}^{n} \beta_j w_j$ definiert. Für die Koordinatenvektoren

$$\alpha = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \text{ und } \beta = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}$$

gilt dann $\varphi(v, w) = \alpha^t C \beta$.

Beweis

Sei $(C\beta)_i$ die i-te Koordinate des Spaltenvektors $C\beta$. Es folgt

$$\begin{aligned}
\varphi(v,w) &= \varphi(\sum_{i=1}^{n}\alpha_i v_i, \sum_{j=1}^{n}\beta_j w_j) &&= \sum_{i=1}^{n}\alpha_i\varphi(v_i, \sum_{j=1}^{n}\beta_j w_j)\\
&= \sum_{i=1}^{n}\alpha_i\big(\sum_{j=1}^{n}\beta_j\varphi(v_i,v_j)\big) &&= \sum_{i=1}^{n}\alpha_i(\sum_{j=1}^{n}\beta_j c_{ij})\\
&= \sum_{i=1}^{n}\alpha_i(\sum_{j=1}^{n}c_{ij}\beta_j) &&= \sum_{i=1}^{n}\alpha_i(C\beta)_i\\
&= \alpha^t C\beta.
\end{aligned}$$

Korollar

Auf dem Standardvektorraum K^n wird jede Bilinearform φ durch $\varphi(a,b) = a^tCb$ für eine Matrix C beschrieben.

Bemerkung

Sei die Bilinearform φ durch C und C' dargestellt. Dann existiert eine invertierbare Matrix S mit $C' = S^tCS$.

Beweis

Sei $C = (c_{ij})$ bzgl. der Basis $v_1,\ldots,v_n$ und $C' = (c'_{ij})$ bzgl. der Basis $v'_1,\ldots,v'_n$ dargestellt mit $v'_i = \sum_{k=1}^{n} s_{ki}v_k$. Dann gilt

$$\begin{aligned}
c'_{ij} = \varphi(v'_i,v'_j) &= \varphi(\sum_{k=1}^{n}s_{ki}v_k, \sum_{h=1}^{n}s_{hj}v_h) &&= \sum_{k=1}^{n}s_{ki}\big(\sum_{h=1}^{n}s_{hj}\varphi(v_k,v_h)\big)\\
&= \sum_{k=1}^{n}s_{ki}(\sum_{h=1}^{n}s_{hj}c_{kh}) &&= \sum_{k=1}^{n}s_{ki}(\sum_{h=1}^{n}c_{kh}s_{hj})
\end{aligned}$$

Sei nun S die Matrix (s_{hj}). Dann ist

$$\sum_{h=1}^{n}c_{kh}s_{hj}$$

der Eintrag in der k-ten Zeile und j-ten Spalte von CS und

$$c'_{ij} = \sum_{k=1}^{n}s_{ki}(\sum_{h=1}^{n}c_{kh}s_{hj})$$

der Eintrag in der i-ten Zeile und j-ten Spalte von S^tCS, d.h. $C' = S^tCS$. Die Spalten von S bilden dabei die Koordinatenvektoren von $v'_1, \ldots, v'_n$ bzgl. der Basis $v_1, \ldots, v_n$ und sind daher linear unabhängig. Daher ist S invertierbar.

14.2 Definition

Zwei Matrizen $C, C' \in M_{n\times n}(K)$ heißen kongruent, falls es eine invertierbare Matrix $S \in M_{n\times n}(K)$ mit $C' = S^tCS$ gibt.

14.3 Lemma

Kongruenz von Matrizen ist eine Äquivalenzrelation.

Beweis

[Reflexivität] $C = 1_n^t C 1_n$.

[Symmetrie] Sei $C' = S^tCS$. Wegen $1_n = 1_n^t = (SS^{-1})^t = (S^{-1})^tS^t$ ist $(S^t)^{-1} = (S^{-1})^t$. Damit folgt $C = 1_nC1_n = ((S^{-1})^tS^t)C(SS^{-1}) = (S^{-1})^t(S^tCS)S^{-1} = (S^{-1})^tC'S^{-1}$.

[Transitivität] Sei $C' = S^tCS$ und $C'' = T^tC'T$. Dann ist $C'' = T^tC'T = T^t(S^tCS)T = (T^tS^t)C(ST) = (ST)^tC(ST)$.

14.4 Definition

Eine Matrix $C \in M_{n\times n}(K)$ heißt symmetrisch bzw. antisymmetrisch, falls $C^t = C$ bzw. $C^t = -C$ gilt.

14.5 Lemma

Eine Matrix ist genau dann symmetrisch bzw. antisymmetrisch, falls eine (und damit alle) zu ihr kongruente Matrix symmetrisch bzw. antisymmetrisch ist. Daher ist Symmetrie bzw. Antisymmetrie invariant unter Kongruenz.

Beweis

Für $C^t = C$ und $S \in M_{n\times n}(K)$ gilt stets $(S^tCS)^t = S^tC^tS^{tt} = S^tCS$. Analog folgt $(S^tDS)^t = -(S^tDS)$ für $D^t = -D$.

Bemerkung

Ist $2 = 1 + 1 \neq 0$ in einem Körper K, so ist jede Matrix $A \in M_{n\times n}(K)$ Summe einer symmetrischen und einer antisymmetrischen Matrix.

Beweis

Da $2 \in K$ invertierbar ist, folgt $A = 1 \cdot A = (2 \cdot 2^{-1}) \cdot A = (1 + 1) \cdot 2^{-1} \cdot A = 2^{-1} \cdot A + 2^{-1} \cdot A = 2^{-1} \cdot A + 2^{-1} \cdot A^t + 2^{-1} \cdot A - 2^{-1} \cdot A^t = 2^{-1} \cdot (A + A^t) + 2^{-1} \cdot (A - A^t)$ mit $(A + A^t)^t = A^t + A^{tt} = A^t + A = A + A^t$ und $(A - A^t)^t = A^t - A^{tt} = A^t - A = -(A - A^t)$.

14.6 Definition

Eine Bilinearform φ auf V heißt symmetrisch bzw. antisymmetrisch, falls $\varphi(v, w) = \varphi(w, v)$ bzw. $\varphi(v, w) = -\varphi(w, v)$ für alle $v, w \in V$ gilt.

Lemma

Bilinearformen $\varphi : V \times V \longrightarrow K$ sind schon dann alternierend, wenn $\varphi(v, v) = 0$ für alle $v \in V$.

Beweis

Seien $v, w \in V$ linear abhängig, d.h. $v = \lambda w$. Dann gilt $\varphi(v, w) = \varphi(v, \lambda v) = \lambda\varphi(v, v) = \lambda 0 = 0$.

14.7 Lemma

Eine Bilinearform $\varphi : V \times V \longrightarrow K$ mit $0 \neq 2 \in K$ ist genau dann antisymmetrisch, wenn sie alternierend ist.

Beweis

Sei φ antisymmetrisch und seien $v, w \in V$ linear abhängig, d.h. $v = \lambda w$. Dann gilt $\lambda\varphi(v, v) = \varphi(v, \lambda v) = \varphi(v, w) = -\varphi(w, v) = -\varphi(\lambda v, v) = -\lambda\varphi(v, v) \Longrightarrow 2\varphi(v, w) = 2\lambda\varphi(v, v) = 0 \Longrightarrow \varphi(v, w) = 0$.

Sei nun φ alternierend. Nach Lemma 9.4 gilt dann $\varphi(u,w) = \operatorname{sign}(12)\,\varphi(w,v) = -\varphi(w,v)$ für alle $v, w \in V$.

14.8 Lemma

Sei für $\dim V < \infty$ eine Bilinearform φ auf V mit darstellender Matrix C gegeben. Dann ist φ genau dann symmetrisch bzw. antisymmetrisch, falls C symmetrisch bzw. antisymmetrisch ist.

Beweis

Sei $\varphi : V^2 \longrightarrow K$ bzgl. der Basis $v_1, \ldots, v_n$ von V durch $C = (c_{ij})$ dargestellt. Falls φ symmetrisch ist, folgt $c_{ij} = \varphi(v_i, v_j) = \varphi(v_j, v_j) = c_{ji}$ und damit $C = C^t$.

Sei nun C symmetrisch und $v, w \in V$ definiert durch $v = \sum_{i=1}^{n} \lambda_i v_i$ und $w = \sum_{j=1}^{n} \mu_j v_j$. Dann gilt auch

$$\begin{aligned} \varphi(v,w) &= \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i\mu_j\varphi(v_i,v_j) &&= \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i\mu_j c_{ij} \\ &= \sum_{j=1}^{n}\sum_{i=1}^{n} \mu_j\lambda_i c_{ji} &&= \sum_{j=1}^{n}\sum_{i=1}^{n} \mu_j\lambda_i\varphi(v_j,v_j) = \varphi(w,v). \end{aligned}$$

Die Aussage über Antisymmetrie folgt analog.

14.9 Bemerkung / Definition

Sei für $\dim V < \infty$ die Bilinearform $\varphi : V^2 \longrightarrow K$ durch die Matrizen C und C' dargestellt. Dann existiert eine invertierbare Matrix S mit $C' = S^t C S$ und mit Satz 7.16 folgt $\operatorname{rang} C = \operatorname{rang} C'$. Daher definiert man $\operatorname{rang}\varphi = \operatorname{rang} C$.

Dabei heißt φ regulär, falls $\operatorname{rang}\varphi = \dim V$ gilt. Für $\dim V = n$ gilt $C \in M_{n\times n}(K)$ und mit Satz 7.17 folgt φ regulär $\Longleftrightarrow \operatorname{rang}\varphi = n \Longleftrightarrow \operatorname{rang} C = n \Longleftrightarrow C$ invertierbar.

14.10 Lemma

Sei φ eine Bilinearform auf V mit $\dim V < \infty$. Dann ist äquivalent:

1. φ ist regulär.
2. Aus $\varphi(u,w) = 0$ für alle $u \in V$ folgt $w = 0$.
3. Aus $\varphi(u,w) = 0$ für alle $w \in V$ folgt $u = 0$.

Beweis

Sei o.B.d.A. $V = K^n$ und somit $\varphi(v, w) = u^t C w$ für $C \in M_{n\times n}(K)$.

1 $\Longrightarrow$ **2** Sei φ regulär und weiter $\varphi(u, w) = u^t C w = 0$ für alle $u \in V$. Dann gilt insbesondere $e_i^t C w = 0$ für die Standardbasis $e_1, \ldots, e_n$, d.h. $Cw = 0$. Nun ist C nach Bemerkung 14.9 invertierbar, d.h. die lineare Abbildung $v \longmapsto Cv$ ist nach Satz 7.9 injektiv. Damit folgt $Cw = 0 \Longrightarrow w = 0$.

2 $\Longrightarrow$ **1** Sei φ nicht regulär, also C nicht invertierbar. Dann ist $v \longmapsto Cv$ nach Korollar 7.6 und Satz 7.9 nicht injektiv, d.h. es existiert $0 \neq w \in V$ mit $Cw = 0$. Dann gilt $\varphi(u, w) = u^t C w = 0$ für alle $u \in V$.

1 $\Longleftrightarrow$ **3** Wegen C invertierbar $\Longleftrightarrow C^t$ invertierbar und $u^t C w \in K \Longrightarrow w^t C^t u = (u^t C w)^t = u^t C w$ folgt die Ausssage analog.

14.11 Lemma

Sei $\varphi : V \times V \longrightarrow K$ für $0 \neq 2 \in K$ eine symmetrische Bilinearform mit $\varphi(v, v) = 0$ für alle $v \in V$. Dann gilt $\varphi = 0$.

Beweis

Für alle $v, w \in V$ gilt $0 = \varphi(v + w, v + w) = \varphi(v, v) + \varphi(v, w) + \varphi(w, v) + \varphi(w, w) = \varphi(v, w) + \varphi(w, v) = 2\,\varphi(v, w) \Longrightarrow \varphi(v, w) = 0$.

14.12 Satz (Jacobi)

Sei für $0 \neq 2 \in K$ und $n = \dim V < \infty$ eine symmetrische Bilinearform $\varphi : V \times V \longrightarrow K$ gegeben. Dann wird φ bzgl. einer geeigneten Basis durch eine Diagonalmatrix beschrieben.

Beweis

Für $\varphi = 0$ ist jede darstellende Matrix $C = 0$ und daher diagonal. Man betrachte nun $\varphi \neq 0$. Nach Lemma 14.11 existiert dann ein $v \in V$ mit $\delta = \varphi(v, v) \neq 0$. Man ergänze v zu einer Basis $v_1, \ldots, v_{n-1}, v$ von V und setze $\lambda_i = \varphi(v, v_i)$ sowie $w_i = v_i - \frac{\lambda_i}{\delta} v$ für $1 \leq i \leq n - 1$. Es folgt

$$\varphi(v, w_i) = \varphi(w_i, v) = \varphi(v_i - \frac{\lambda_i}{\delta} v, v) = \varphi(v_i, v) - \frac{\lambda_i}{\delta}\varphi(v, v) = \lambda_i - \frac{\lambda_i}{\delta}\delta = 0$$

für $1 \leq i \leq n - 1$. Weiter gilt offenbar $\langle w_1, \ldots, w_{n-1}, v\rangle = \langle v_1, \ldots, v_{n-1}, v\rangle = V$.

Nun setzen wir $U = \langle w_1, \ldots, w_{n-1} \rangle$ und beweisen die Aussage durch vollständige Induktion über $\dim V$. Insbesondere ist dabei $\varphi : U \times U \longrightarrow K$ eine symmetrische Bilinearform, d.h. nach Annahme existiert eine Basis $u_1, \ldots, u_{n-1}$ von U mit $\varphi(u_i, u_j) = 0$ für $i \neq j$.

Wegen $u_i \in \langle w_1, \ldots, w_{n-1} \rangle$ sowie $\varphi(v, w_i) = \varphi(w_i, v) = 0$ gilt auch $\varphi(v, u_i) = \varphi(u_i, v) = 0$ für $1 \leq i \leq n-1$. Schließlich folgt die Behauptung, da $v, u_1, \ldots, u_{n-1}$ wegen $v \notin \langle u_1, \ldots, u_{n-1} \rangle$ nach Lemma 4.12 eine Basis von V ist.

Definition / Bemerkung

Man schreibt $\mathrm{GL}_n(K) = \{W \in M_{n\times n}(K) \mid W \text{ invertierbar}\}$ und $\mathrm{SL}_n(K) = \{W \in M_{n\times n}(K) \mid \det W = 1\}$. Dabei ist $\mathrm{GL}_n(K)$ eine multiplikative Gruppe und $\mathrm{SL}_n(K)$ eine echte Untergruppe von $\mathrm{GL}_n(K)$.

14.13 Satz (Trägheitssatz von Sylvester)

Sei $S \in M_{n\times n}(\mathbb{R})$ symmetrisch, dann existiert $W \in \mathrm{GL}_n(K)$ mit

$$W^t S W = \begin{pmatrix} \underbrace{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}}_{t} & & & \\ & \underbrace{\begin{matrix} -1 & & \\ & \ddots & \\ & & -1 \end{matrix}}_{r-t} & & \\ & & \begin{matrix} 0 & & \\ & \ddots & \\ & & 0 \end{matrix} & \end{pmatrix}$$

für $r = \operatorname{rang} S$. Dabei ist $t \in \mathbb{N}$ eindeutig bestimmt.

Beweis

Sei $\varphi : V \times V \longrightarrow K$ die durch S dargestellte Bilinearform. Nach Satz 14.12 findet man eine Basis $v_1, \ldots, v_n$ mit $\varphi(v_i, v_j) = 0$ für $i \neq j$. Dann gilt $|\{i \in \mathbb{N} \mid \varphi(v_i, v_i) \neq 0\}| = \operatorname{rang} \varphi = \operatorname{rang} S = r$. Nun existiere o.B.d.A. ein $t \in \mathbb{N}$ mit $\varphi(v_i, v_i) > 0$ für $1 \leq i \leq t$ sowie $\varphi(v_i, v_i) < 0$ für $t+1 \leq i \leq r$ und $\varphi(v_i, v_i) = 0$ für $i \geq r+1$.

Man setze $\delta_i = \varphi(v_i, v_i)$ sowie $w_i = \frac{1}{\sqrt{\delta_i}} v_i$ für $1 \leq i \leq t$ und $w_i = \frac{1}{\sqrt{-\delta_i}} v_i$ für $t+1 \leq i \leq r$ und $w_i = v_i$ für $i \geq r+1$. Dann bildet $w_1, \ldots, w_n$ offenbar eine Basis von V mit $\varphi(w_i, w_j) = 0$ für $i \neq j$. Weiter gilt $\varphi(w_i, w_i) = 1$ für $1 \leq i \leq t$ und $\varphi(w_i, w_i) = -1$ für $t+1 \leq i \leq r$ sowie $\varphi(w_i, w_i) = 0$ für $i \geq r+1$.

Es bleibt zu zeigen, dass $t \in \mathbb{N}$ von der Konstruktion unabhängig ist. Sei dazu $u_1, \ldots, u_n$ eine weitere Basis von V mit $\varphi(u_i, u_j) = 0$ für $i \neq j$ sowie $\varphi(u_i, u_i) = 1$ für $1 \leq i \leq t'$ und $\varphi(u_i, u_i) = -1$ für $t'+1 \leq i \leq r$ und $\varphi(u_i, u_i) = 0$ für $i \geq r+1$.

Nun setze man $U = \langle u_1, \ldots, u_{t'} \rangle$ und $W = \langle w_{t+1}, \ldots, w_n \rangle$. Für $v \in U$ mit $v = \sum_{i=1}^{t'} \alpha_i u_i$ folgt

$$\varphi(v, v) = \sum_{i=1}^{t'} \sum_{j=1}^{t'} \alpha_i \alpha_j \varphi(u_i, u_j) = \sum_{i=1}^{t'} \alpha_i^2 \varphi(u_i, u_i) = \sum_{i=1}^{t'} \alpha_i^2.$$

Daher gilt $\varphi(v, v) \geq 0$ mit $\varphi(v, v) = 0 \iff \forall_{i=1,\ldots,t'} \ \alpha_i = 0 \iff v = 0$. Analog folgt $\varphi(w, w) \leq 0$ für $w \in W$, wobei wegen $\varphi(w_i, w_i) = 0$ für $i \geq r+1$ auch $\varphi(w, w) = 0$ mit $w \neq 0$ gelten kann.

Daher ergibt sich $v \in U \cap W \Longrightarrow 0 \leq \varphi(v, v) \leq 0 \Longrightarrow \varphi(v, v) = 0 \Longrightarrow v = 0 \Longrightarrow U \cap W = \{0\}$ und somit $n = \dim V \geq \dim(U + W) = \dim U + \dim W - \dim(U \cap W) = \dim U + \dim W = t' + n - t \Longrightarrow t \geq t'$. Für $\overline{W} = \langle w_1, \ldots, w_t \rangle$ und $\overline{U} = \langle u_{t'+1}, \ldots, u_n \rangle$ erhält man analog $t' \geq t$, also insgesamt $t = t'$.

14.14 Definition / Bemerkung

Eine Bilinearform $\varphi : V \times V \longrightarrow \mathbb{R}$ auf dem reellen Vektorraum V heißt positiv definit, falls für alle $0 \neq v \in V$ stets $\varphi(v, v) > 0$ gilt. Dabei ist φ genau dann positiv definit, wenn für alle $v \in V$ stets $\varphi(v, v) \geq 0$ und $\varphi(v, v) = 0 \Longrightarrow \varphi(v, v) = 0$ gilt.

14.15 Definition

Eine positiv definite, symmetrische Bilinearform auf einem reellen Vektorraum Skalarprodukt oder inneres Produkt. Ein reeller Vektorraum mit definiertem Skalarprodukt φ heißt euklidischer Raum. Man schreibt dabei statt $\varphi(u, w)$ oft uw. Der Betrag, die Norm oder die Länge von v ist $\sqrt{\varphi(v, v)} = \sqrt{vv} = |v|$.

14.16 Bemerkung

Auf $V = \mathbb{R}^n$ wird durch $\varphi(v, w) = v^t w$ das Standardskalarprodukt definiert. In diesem Fall stimmt $|v|$ mit der gewöhnlichen Länge von v überein.

14.17 Definition / Bemerkung

Sei V ein komplexer Vektorraum. Dann heißt eine Abbildung $\varphi : V \times V \longrightarrow \mathbb{C}$ hermitesche Form, falls für alle $v, v', w \in V$ und $\alpha \in \mathbb{C}$ stets

1. $\varphi(v + v', w) = \varphi(v, w) + \varphi(v', w)$
2. $\varphi(\alpha v, w) = \alpha\, \varphi(v, w)$
3. $\varphi(v, w) = \overline{\varphi(w, v)}$

gilt. Dabei ist $\overline{z} = a - bi$ für $z = a + bi \in \mathbb{C}$, d.h. für alle $v \in V$ gilt $\varphi(v, v) = \overline{\varphi(v, v)} \Longrightarrow \varphi(v, v) \in \mathbb{R}$. Daher heißt φ positiv definit, falls für alle $0 \neq v \in V$ stets $\varphi(v, v) > 0$ gilt. Ein komplexer Vektorraum mit positiv definiter hermitescher Form heißt unitärer Raum. Man schreibt wieder uw statt $\varphi(u, w)$ und der Betrag, die Norm oder die Länge von v ist ebenfalls $\sqrt{vv} = |v|$. Man beachte, dass hermitesche Formen weder bilinear noch symmetrisch sind.

14.18 Bemerkung

Für eine hermitesche Form $\varphi : V \times V \longrightarrow \mathbb{C}$ mit $v, w, w' \in V$ und $\alpha \in \mathbb{C}$ folgt

1. $\varphi(v, \alpha w) = \overline{\varphi(\alpha w, v)} = \overline{\alpha\, \varphi(w, v)} = \overline{\alpha}\, \overline{\varphi(w, v)} = \overline{\alpha}\, \varphi(v, w)$
2. $\varphi(v, w + w') = \overline{\varphi(w + w', v)} = \overline{\varphi(w, v)} + \overline{\varphi(w', v)} = \varphi(v, w) + \varphi(v, w')$

Bemerkung

Sei $v_1, \ldots, v_n$ eine Basis des komplexen Vektorraums V. Dann wird jede hermitesche Form $\varphi : V \times V \longrightarrow \mathbb{C}$ durch eine Matrix $C = (c_{ij}) \in M_{n\times n}(\mathbb{C})$ mit $c_{ij} = \varphi(v_i, v_j)$ dargestellt. Insbesondere gilt dabei $c_{ij} = \varphi(v_i, v_j) = \overline{\varphi(v_j, v_i)} = \overline{c_{ji}}$, d.h. $C = \overline{C}^t$.

14.19 Definition

Man schreibt für $C \in M_{n\times n}(\mathbb{C})$ statt $\overline{C}^t$ oft C^*, wobei C^* die zu C adjungierte Matrix heißt. Fall $C = C^*$ gilt, heißt C hermitesch.

Bemerkung

Für $C \in M_{n\times n}(\mathbb{C})$ gilt

1. $(C^*)^* = \overline{\overline{C}^t}^t = C$
2. $(C^* + D^*)^* = \overline{\overline{C^t} + \overline{C}^t}^t = \overline{\overline{C}^t}^t + \overline{\overline{D}^t}^t = C + D$
3. $(C \cdot D)^* = \overline{C \cdot D}^t = (\overline{C} \cdot \overline{D})^t = \overline{D}^t \cdot \overline{C}^t = D^* \cdot C^*$

14.20 Definition

Zwei Matrizen $A, B \in M_{n\times n}(\mathbb{C})$ heißen komplex kongruent, falls eine Matrix $M \in \mathrm{GL}_n(\mathbb{C})$ mit $B = M^*AM$ existiert.

14.21 Lemma

Komplexe Kongruenz ist eine Äquivalenzrelation, die verträglich mit der hermiteschen Eigenschaft ist.

Beweis

Sei $R = \{(A, B) \in M_{n\times n}(\mathbb{C}) \times M_{n\times n}(\mathbb{C}) \mid B = M^*AM \text{ für } M \in \mathrm{GL}_n(\mathbb{C})\}$. Zunächst gilt $1_n = 1_n^* = (MM^{-1})^* = (M^{-1})^*M^*$. Daher ist R wegen $A = 1_n^*A1_n$ reflexiv. Für $B = M^*AM$ folgt $A = (M^{-1})^*M^*AMM^{-1} = (M^{-1})^*BM^{-1}$, d.h. R ist symmetrisch. Da weiter für $D = N^*BN$ stets $D = N^*(M^*AM)N = (MN)^*A(MN)$ gilt, ist R transitiv. Für $A = A^*$ folgt schließlich $B^* = (M^*AM)^* = M^*A^*M^{**} = M^*AM = B$.

14.22 Bemerkung

Sei $v_1, \ldots, v_n$ eine Basis des komplexen Vektorraums V und C die darstellende Matrix der hermiteschen Form φ bzgl. dieser Basis. Für $v = \sum_{i=1}^{n} \alpha_i v_i$ und $w = \sum_{j=1}^{n} \beta_j v_j$ mit

$$\alpha = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \text{ und } \beta = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}$$

gilt dann

$$\varphi(v, w) = \sum_{i=1}^{n} \alpha_i \sum_{j=1}^{n} \overline{\beta_j}\, \varphi(v_i, v_j) = \sum_{i=1}^{n} \alpha_i (\sum_{j=1}^{n} c_{ij}\, \overline{\beta_j}) = \alpha^t\, C\, \overline{\beta}.$$

14.23 Lemma

Sei $\dim V < \infty$ ein komplexer Vektorraum und $C, C' \in M_{n\times n}(\mathbb{C})$ hermitesche Matrizen. Dann beschreiben C und C' genau dann die gleiche hermitesche Form auf V bzgl. geeigneter Basen, wenn sie komplex kongruent sind.

Beweis

Man betrachte o.B.d.A. den Standardvektorraum $V = \mathbb{C}^n$. Dann gilt für $\varphi(v,w) = v^t C\overline{w}$ offenbar $\varphi(v+v',w) = \varphi(v,w)+\varphi(v',w)$ und $\varphi(\alpha v,w) = \alpha\,\varphi(v,w)$. Daher und wegen $\varphi(v,w) = v^t\,C\,\overline{w} = (v^t\,C\,\overline{w})^t = \overline{w}^t\,C^t\,v = \overline{w^t\,C^*\,\overline{v}} = \overline{w^t\,C\,\overline{v}} = \overline{\varphi(w,v)}$ definiert φ eine hermitesche Form, die bzgl. der Standardbasis $e_1,\ldots,e_n$ durch $C = (c_{ij})$ beschrieben wird.

Sei nun φ bzgl. einer Basis $m_1,\ldots,m_n$ mit

$$m_i = \begin{pmatrix} m_{1i} \\ \vdots \\ m_{ni} \end{pmatrix}$$

durch $C' = (c'_{ij})$ dargestellt, d.h. $m_i = \sum\limits_{k=1}^{n} m_{ki}e_k$. Es folgt

$$c'_{ij} = \varphi(m_i,m_j) = \sum_{k=1}^{n} m_{ki}\big(\sum_{h=1}^{n}\overline{m_{hj}}\,\varphi(e_k,e_h)\big) = \sum_{k=1}^{n} m_{ki}(\sum_{h=1}^{n} c_{kh}\,\overline{m_{hj}}).$$

Für $M = (\overline{m_{ki}}) \in M_{n\times n}(\mathbb{C})$ erhält man damit $C' = M^*CM$. Da weiter die Spalten $m_1,\ldots,m_n$ von M linear unabhängig sind und somit M invertierbar ist, folgt die Behauptung.

Seien nun C und C' komplex kongruent mit $C' = M^*CM$ für eine invertierbare Matrix $M = (\overline{m_{ki}})$. Dann ergibt sich analog $c'_{ij} = \varphi(m_i,m_j)$ für

$$m_i = \begin{pmatrix} m_{1i} \\ \vdots \\ m_{ni} \end{pmatrix}.$$

Da nun $m_1,\ldots,m_n$ als linear unabhängige Spalten von M eine Basis von $\mathbb{C}^n$ bilden, folgt die Behauptung.

14.24 Definition / Bemerkung

Sei V ein reeller Vektorraum. Dann bildet $V_{\mathbb{C}} = V\times V$ mit komponentenweiser Addition sowie $(a+ib)(v,w) = (av-bw, aw+bv)$ für $v,w\in V$ und $a,b\in\mathbb{R}$ einen komplexen Vektorraum. Wegen $i(v,0) = (0+i)(v,0) = (0,v)$ interpretiert man dabei $\{(v,0) \mid v\in V\}$ als V und

identifiziert $V \times V$ mit $V + iV$. Dann hat jedes Element $(v, w) \in V \times V$ eine eindeutige Darstellung $v + iw \in V + iV$ und es gilt $(a + ib)(v + iw) = av + biv + aiw - bw = (av - bw) + i(aw + bv)$. Man beachte, dass $V \times \{0\}$ ein Unterraum des reellen Vektorraums $V \times V$ – aber wegen $(a + ib)(v, 0) = (av, bv) \notin V \times \{0\}$ für $b \neq 0$ nur eine Teilmenge von $V_{\mathbb{C}}$ – ist.

14.25 Lemma

Sei V ein reeller Vektorraum. Dann ist jede Basis von V eine Basis von $V_{\mathbb{C}}$. Insbesondere gilt also $\dim V = \dim V_{\mathbb{C}}$.

Beweis

Sei I eine Indexmenge und $\{v_i \mid i \in I\}$ eine Basis von V. Seien weiter $v + iw \in V_{\mathbb{C}}$, also $v, w \in V$. Dann gilt $v = \sum a_j v_j$ und $w = \sum b_j v_j$ für $a_j, b_j \in \mathbb{R}$, d.h. $v + iw = \sum (a_j + ib_j)v_j$. Daher ist $\{v_i \mid i \in I\}$ ein Erzeugendensystem von $V_{\mathbb{C}}$.

Es bleibt die lineare Unabhängigkeit der v_i in $V_{\mathbb{C}}$ zu zeigen. Sei dazu $J \subseteq I$ eine endliche Teilmenge mit $\sum_{j \in J} (a_j + ib_j)v_j = 0$ für $a_j, b_j \in \mathbb{R}$. Es folgt $0 = v + iw \in V_{\mathbb{C}}$ für $v = \sum_{j \in J} a_j v_j \in V$ und $w = \sum_{j \in J} b_j v_j \in V$. Nach Bemerkung 14.24 ist die Darstellung der $0 \in V_{\mathbb{C}}$ eindeutig, d.h. $v = w = 0$. Da die v_i in V jedoch linear unabhängig sind, ergibt sich $a_j = b_j = 0$ für alle $j \in J$ und die Behauptung folgt.

14.26 Lemma

Seien V, W reelle Vektorräume und $\alpha \in \mathrm{Hom}(V, W)$. Dann gibt es genau ein $\beta \in \mathrm{Hom}(V_{\mathbb{C}}, W_{\mathbb{C}})$ mit $\beta|_V = \alpha$ – d.h. α hat eine eindeutige Fortsetzung auf $V_{\mathbb{C}}$.

Beweis

[Eindeutigkeit] Für $\beta \in \mathrm{Hom}(V_{\mathbb{C}}, W_{\mathbb{C}})$ mit $\beta|_V = \alpha$ folgt für alle $v, w \in V$ stets $\beta(v + iw) = \beta(v) + i\beta(w) = \alpha(v) + i\alpha(w)$.

[Existenz] Mit $\beta : v + iw \in V_{\mathbb{C}} \longmapsto \alpha(v) + i\alpha(w) \in W_{\mathbb{C}}$ und $v \in V$ gilt $\beta(v) = \beta(v + i0) + \alpha(v) + i\alpha(0) = \alpha(v)$. Es folgt weiter

$$\begin{aligned} \beta\big((a + ib)(v + iw)\big) &= \beta\big((av - bw) + i(aw + bv)\big) &&= \alpha(av - bw) + i\alpha(aw + bv) \\ &= (a + ib)\alpha(v) + (ia - b)\alpha(w) &&= (a + ib)\big(\alpha(v) + i\alpha(w)\big) \\ &= (a + ib)\beta(v + iw) \end{aligned}$$

und

$$\begin{aligned}\beta\big(v+iw)+(v'+iw')\big) &= \beta\big((v+v')+i(w+w')\big) &&= \alpha(v+v')+i\alpha(w+w')\\ &= \alpha(v)+i\alpha(w)+\alpha(v')+i\alpha(w') &&= \beta(v+iw)+\beta(v'+iw'),\end{aligned}$$

d.h. $\beta \in \mathrm{Hom}(V_{\mathbb{C}}, W_{\mathbb{C}})$.

14.27 Lemma

Sei auf den reellen Vektorraum V eine symmetrische Bilinearform $\varphi : V \times V \longrightarrow \mathbb{R}$ definiert. Dann existiert genau eine hermitesche Form $\psi : V_{\mathbb{C}} \times V_{\mathbb{C}} \longrightarrow \mathbb{C}$ mit $\psi|_{V\times V} = \varphi$.

Beweis

[Eindeutigkeit] Für $\psi|_{V\times V} = \varphi$ und $v+iw, v'+iw' \in V_{\mathbb{C}}$ folgt $\psi(v+iw, v'+iw') = \psi(v,v') + i\,\psi(w,v') + \bar{i}\,\psi(v,w') + i\,\bar{i}\,\psi(w,w') = \psi(v,v') + i\,\psi(w,v') - i\,\psi(v,w') + \psi(w,w') = \varphi(v,v') + \varphi(w,w') + i\,\varphi(w,v') - i\,\varphi(v,w')$.

[Existenz] Mit $\psi : (v+iw, v'+iw') \in V_{\mathbb{C}}\times V_{\mathbb{C}} \longmapsto \varphi(v,v')+\varphi(w,w')+i\,\varphi(w,v')-i\,\varphi(v,w')$ gilt für $v, v' \in V$ zunächst $\psi(v,v') = \psi(v+i0, v'+i0) = \varphi(v,v') + \varphi(0,0) + i\,\varphi(0,v') - i\,\varphi(v,0) = \varphi(v,v')$. Nun zeigen wir, dass ψ eine hermitesche Form ist. Dazu gilt

$$\begin{aligned}&\psi\big((v+iw)+(v'+iw'), v''+iw''\big)\\ =\;& \psi\big((v+v')+i(w+w'), v''+iw''\big)\\ =\;& \varphi(v+v', v'') + \varphi(w+w', w'') + i\,\varphi(w+w', v'') - i\,\varphi(v+v', w'')\\ =\;& \varphi(v,v'') + \varphi(w,w'') + i\,\varphi(w,v'') - i\,\varphi(v,w'')\\ +\;& \varphi(v',v'') + \varphi(w',w'') + i\,\varphi(w',v'') - i\,\varphi(v',w'')\\ =\;& \psi(v+iw, v''+iw'') + \psi(v'+iw', v''+iw'')\end{aligned}$$

und

$$\begin{aligned}
& \psi\big((a+ib)(v+iw), v'+iw'\big) \\
= \;& \psi\big((av-bw)+i(aw+bv), v'+iw'\big) \\
= \;& \varphi(av-bw, v') + \varphi(aw+bv, w') + i\,\varphi(aw+bv, v') - i\,\varphi(av-bw, w') \\
= \;& a\big(\varphi(v, v') + \varphi(w, w') + i\,\varphi(w, v') - i\,\varphi(v, w')\big) \\
+ \;& ib\big(i\,\varphi(w, v') - i\,\varphi(v, w') + \varphi(v, v') + \varphi(w, w')\big) \\
= \;& (a+ib)\psi(v+iw, v'+iw').
\end{aligned}$$

Wegen $\varphi(v, w) \in \mathbb{R}$ gilt $\varphi(v, w) = \overline{\varphi(v, w)}$ und mit $\varphi(v, w) = \varphi(w, v)$ folgt weiter

$$\begin{aligned}
& \psi(v+iw, v'+iw') \\
= \;& \varphi(v, v') + \varphi(w, w') + i\,\varphi(w, v') - i\,\varphi(v, w') \\
= \;& \overline{\varphi(v', v) + \varphi(w', w) - i\,\varphi(v', w) + i\,\varphi(w', v)} \\
= \;& \overline{\psi(v'+iw', v+iw,)}.
\end{aligned}$$

14.28 Lemma

Sei φ eine symmetrische Bilinearform auf dem reellen Vektorraum V und ψ die hermitesche Fortsetzung auf $V_{\mathbb{C}}$. Dann ist φ genau dann positiv definit, wenn ψ es ist.

Beweis

Sei ψ positiv definit. Dann ist wegen $\varphi(v, v) = \psi(v+i0, v+i0) > 0$ für alle $0 \neq v \in V$ auch φ positiv definit. Sei nun φ positiv definit und $0 \neq v+iw \in V_{\mathbb{C}}$, d.h. $v \neq 0 \Longrightarrow \varphi(v, v) > 0$ oder $w \neq 0 \Longrightarrow \varphi(w, w) > 0$. Dann gilt $\psi(v+iw, v+iw) = \varphi(v, v)+\varphi(w, w)+i\,\varphi(w, v)-i\,\varphi(v, w) = \varphi(v, v) + \varphi(w, w) + i\,\varphi(v, w) - i\,\varphi(v, w) = \varphi(v, v) + \varphi(w, w) > 0$.

14.29 Satz (Ungleichung von Cauchy-Schwarz)

Sei V ein euklidischer oder unitärer Vektorraum mit Skalarprodukt $V \times V \longrightarrow K$. Dann gilt $|vw| \leq |v| \cdot |w|$ für alle $v, w \in V$.

Beweis

Das Skalarprodukt φ eines euklidischen Vektorraums V sei auf die positiv definite hermitesche Form $\psi : V_{\mathbb{C}} \times V_{\mathbb{C}} \longrightarrow \mathbb{C}$ fortgesetzt. Dann gilt für $v, w \in V$ stets $|vw| = |\varphi(v,w)| = |\psi(v,w)|$ und $\sqrt{\psi(v,v)} = \sqrt{\varphi(v,v)} = |v|$ sowie $\sqrt{\psi(w,w)} = \sqrt{\varphi(w,w)} = |w|$. Damit folgt die Ungleichung für euklidische Räume aus der Ungleichung für unitäre Räume.

Sei also o.B.d.A. V unitär mit Skalarprodukt ψ. Für $w = 0$ ist die Aussage trivial, d.h. man betrachtet $0 \neq w \in V \Longleftrightarrow \psi(w,w) > 0$. Da ψ positiv definit ist, folgt für $\lambda \in \mathbb{C}$ stets

$$0 \leq \psi(v - \lambda w, v - \lambda w) = \psi(v,v) - \lambda\,\psi(w,v) - \overline{\lambda}\,\psi(v,w) + \lambda\,\overline{\lambda}\,\psi(w,w).$$

Für $\lambda = \frac{\psi(v,w)}{\psi(w,w)}$ folgt wegen $\psi(w,w) \in \mathbb{R} \Longleftrightarrow \overline{\psi(w,w)} = \psi(w,w)$ dann

$$\overline{\lambda} = \overline{\left(\frac{\psi(v,w)}{\psi(w,w)}\right)} = \frac{\overline{\psi(v,w)}}{\psi(w,w)} = \frac{\psi(w,v)}{\psi(w,w)}.$$

Insgesamt erhält man also

$$\begin{aligned} 0 &\leq \psi(v,v) - \tfrac{\psi(v,w)}{\psi(w,w)}\,\psi(w,v) - \tfrac{\psi(w,v)}{\psi(w,w)}\,\psi(v,w) + \tfrac{\psi(v,w)}{\psi(w,w)}\,\psi(w,v) \\ &= \psi(v,v) - \tfrac{\psi(w,v)}{\psi(w,w)}\,\psi(v,w). \end{aligned}$$

Die Multiplikation mit $\psi(w,w) > 0$ liefert schließlich

$$\begin{aligned} 0 &\leq \psi(v,v)\,\psi(w,w) - \psi(w,v)\,\psi(v,w) \\ &= \psi(v,v)\,\psi(w,w) - \overline{\psi(v,w)}\,\psi(v,w) \\ &= |v|^2\,|w|^2 - |vw|^2 \end{aligned}$$

und es folgt die Behauptung.

Bemerkung

Die Vektoren $v, w \in V$ eines euklidischen oder unitären Raumes sind genau dann linear abhängig, wenn $|vw| = |v|\,|w|$ gilt.

Beweis

Sei $v = \lambda w$. Dann gilt

$$\begin{aligned} |vw|^2 &= (vw)(\overline{vw}) &&= (vw)(wv) \\ &= \big(v(\lambda v)\big)(wv) &&= \overline{\lambda}\,(vv)(wv) \\ &= (vv)\,\overline{\lambda}\,(wv) &&= (vv)\big(w(\lambda v)\big) \\ &= (vv)(ww) &&= |v|\,|w|. \end{aligned}$$

Nun sei $|vw| = |v|\,|w|$. Da für $w = 0$ die Vektoren v und w stets linear abhängig sind, untersuchen wir $w \neq 0 \Longleftrightarrow ww \neq 0$. Dann folgt $(v - \lambda w)(v - \lambda w) = 0$ für $\lambda = \frac{vw}{ww}$ analog zu Satz 14.29, also $v - \lambda w = 0 \Longleftrightarrow v = \lambda w$.

14.30 Korollar

Für $x_1, \ldots, x_n, y_1, \ldots, y_n \in \mathbb{R}$ gilt stets $(\sum_{i=1}^{n} x_i y_i)^2 \leq (\sum_{i=1}^{n} x_i^2)(\sum_{i=1}^{n} y_i^2)$.

Beweis

Sei durch $\varphi : (v, w) \in \mathbb{R}^n \times \mathbb{R}^n \longmapsto v^t w \in \mathbb{R}$ das Standardskalarprodukt auf $\mathbb{R}^n$ definiert. Sei weiter

$$v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \text{ und } w = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Wegen $v^t w \in \mathbb{R}$ ist nun $|v^t w|^2 = (v^t w)^2$. Mit der Ungleichung von Cauchy-Schwarz folgt dann insgesamt $(\sum_{i=1}^{n} x_i y_i)^2 = (v^t w)^2 = |v^t w|^2 = |\varphi(v, w)|^2 \leq \varphi(v, v)\,\varphi(w, w) = (\sum_{i=1}^{n} x_i^2)(\sum_{i=1}^{n} y_i^2)$.

14.31 Satz

Sei V ein euklidischer oder unitärer Raum. Für das Skalarprodukt und $v, w \in V$ sowie $\lambda \in \mathbb{K}$ gilt stets

1. [positive Definitheit] $|v| \geq 0$ mit $|v| = 0 \Leftrightarrow v = 0$
2. [Homogenität] $|\lambda v| = |\lambda|\,|v|$ mit $|\lambda|^2 = \lambda\,\overline{\lambda}$, also insbesondere $|v| = |-v| = |iv|$
3. [Dreiecksungleichung] $|v + w| \leq |v| + |w|$ mit $|v + w| = |v| + |w| \Longleftrightarrow v = 0$ oder $w = \mu v$ für $0 \leq \mu \in \mathbb{R}$

Beweis

Es sei o.B.d.A. V unitär. Dann folgen die Aussagen für euklidische Räume unmittelbar aus der Betrachtung ihrer komplexen Fortsetzung.

1 Die Behauptung ist äquivalent zu der Definition der positiven Definitheit des Skalaprodukts.

2 Es gilt $|\lambda v|^2 = (\lambda v)(\lambda v) = \lambda\,\overline{\lambda}\,(vv) = |\lambda|^2\,|v|^2$ und die Aussage folgt.

3 Für eine komplexe Zahl $z = a + ib \in \mathbb{Z}$ gilt $z + \overline{z} = a + ib + a - ib = 2a = 2\,\mathrm{Re}(z)$ und $\mathrm{Re}(z) = a \leq \sqrt{a^2 + b^2} = |z|$. Daher folgt die Behauptung mit der Ungleichung von Cauchy-Schwarz aus

$$
\begin{array}{rlllll}
(*) & |v+w|^2 & = & (v+w)(v+w) & = & vv + wv + vw + ww \\
 & & = & vv + \overline{vw} + vw + ww & = & |v|^2 + 2\,\mathrm{Re}(vw) + |w|^2 \\
 & & \leq & |v|^2 + 2|vw| + |w|^2 & \leq & |v|^2 + 2|v|\,|w| + |w|^2 \\
 & & = & (|v| + |w|)^2. & &
\end{array}
$$

Für $v = 0$ erhält man $|v + w| = |0 + w| = |w| = |0| + |w| = |v| + |w|$ und für $w = \mu v$ mit $0 \leq \mu \in \mathbb{R}$ folgt $|v + w| = |v + \mu v| = |(1 + \mu)v| = |1 + \mu|\,|v| = (1 + \mu)\,|v| = |v| + \mu\,|v| = |v| + |\mu|\,|v| = |v| + |\mu v| = |v| + |w|$.

Es gelte nun $|v + w| = |v| + |w|$. Mit $(*)$ erhält man dann $\mathrm{Re}(vw) = |vw|$ und $|vw| = |v|\,|w|$. Für $z = a + ib \in \mathbb{C}$ gilt dabei $a = \mathrm{Re}(z) = |z| = \sqrt{a^2 + b^2} \Longrightarrow a \geq 0 \;\wedge\; b = 0 \Longrightarrow z = a \in \mathbb{R}_0^+$, d.h. $vw \in \mathbb{R}_0^+$. Weiter sind v und w wegen $|vw| = |v|\,|w|$ nach Cauchy-Schwarz linear abhängig. Dann folgt insgesamt entweder $v = 0$ oder $w = \mu v \Longrightarrow \overline{\mu}\,|v|^2 = \overline{\mu}\,(vv) = v(\mu v) = vw \geq 0 \Longrightarrow \mu \geq 0$.

Bemerkung

Aus der Dreiecksungleichung folgt sofort $\big||v| - |w|\big| \leq |v - w|$.

Beweis

Es gilt $|v| = |v - w + w| \leq |v - w| + |w| \Longrightarrow |v| - |w| \leq |v - w|$ und $|w| = |w - v + v| \leq |w - v| + |v| = |-(w - v)| + |v| = |v - w| + |v| \Longrightarrow -(|v| - |w|) = |w| - |v| \leq |v - w|$. Insgesamt folgt die Behauptung.

14.32 Definition

Ein Element v eines euklidischen oder unitären Raums heißt normiert, falls $|v| = 1$ gilt.

Bemerkung

Für einen beliebigen Vektor $v \neq 0$ ist dabei $\frac{1}{|v|}v$ wegen $|\frac{1}{|v|}v| = |\frac{1}{|v|}|\,|v| = \frac{1}{|v|}\,|v| = 1$ stets normiert.

14.33 Definition

Seien $v, w \neq 0$ Elemente eines euklidischen Raums. Dann definiert $\cos\varphi = \frac{vw}{|v|\,|w|}$ den Zwischenwinkel φ von v und w. In einem euklidischen oder unitären Raum heißen v und w senkrecht oder orthogonal, falls $vw = 0$ gilt. In diesem Fall schreibt man auch $v \perp w$.

Bemerkung

Für Elemente eines euklidischen Raums gilt $vw \in \mathbb{R}$, also $|vw| = vw$ oder $|vw| = -vw$. Mit $|vw| \leq |v|\,|w|$ folgt $-1 \leq \frac{vw}{|v|\,|w|} \leq 1$, d.h. $\cos\varphi = \frac{vw}{|v|\,|w|}$ ist wohldefiniert.

14.34 Definition

Sei V ein euklidischer oder unitärer Raum und $0 \notin M \subseteq V$ eine Teilmenge. Dann heißt M ein Orthogonalsystem, falls je zwei verschiedene Elemente aus M orthogonal sind. Gilt zusätzlich $|v| = 1$ für alle $v \in M$, so heißt M ein Orthonormalsystem. Ist eine Basis ein Orthogonal- bzw. ein Orthonormalsystem, so nennt man diese Basis eine Orthogonal- bzw. eine Orthonormalbasis.

14.35 Satz

Orthogonalsysteme sind linear unabhängig.

Beweis

Wir zeigen, dass jeweils endlich viele Elemente eines Orthogonalsystems linear unabhängig sind. Seien also $v_i \neq 0$ für $1 \leq i \leq n$ paarweise orthogonal und $\sum_{i=1}^{n} a_i v_i = 0$ für $a_i \in K$. Für

$1 \leq j \leq n$ folgt

$$0 = 0v_j = (\sum_{i=1}^{n} a_i v_i)v_j = \sum_{i=1}^{n} a_i(v_i v_j) = a_j(v_j v_j) = a_j|v_j|^2$$

mit $|v_j| \neq 0$, also $a_j = 0$.

14.36 Satz

Jeder euklidische oder unitäre Raum V endlicher Dimension besitzt eine Orthonormalbasis $v_1, \ldots, v_n$. Für jeden Vektor $v \in V$ gilt dann $v = \sum_{i=1}^{n}(vv_i)v_i$.

Beweis

Wir beweisen die Aussage für euklidische Räume mit Hilfe des Trägheitssatz von Sylvester. Da dieser Satz entsprechend für hermitesche Matrizen gilt, folgt damit die Behauptung für unitäre Räume analog. Sei also V ein euklidischer Raum mit einem Skalarprodukt φ und $\dim V = n$.

Dann ist φ insbesondere eine symmetrische Bilinearform und wird daher durch eine symmetrische Matrix $C \in M_{n\times n}(\mathbb{R})$ dargestellt. Diese ist nach Satz 14.13 kongruent zu einer Matrix

$$D = \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{pmatrix} \in M_{n\times n}(\mathbb{R}),$$

d.h. φ wird bzgl. einer Basis $v_1, \ldots, v_n$ durch D beschrieben. Da φ positiv definit ist, gilt $D = 1_n$ und $v_1, \ldots, v_n$ ist eine Orthonormalbasis.

Für einen Vektor $v \in V$ mit $v = \sum_{i=1}^{n} a_i v_i$ gilt dann

$$vv_j = (\sum_{i=1}^{n} a_i v_i)v_j = \sum_{i=1}^{n} a_i(v_i v_j) = a_j(v_j v_j) = a_j|v_j|^2 = a_j$$

für $1 \leq j \leq n$. Daher folgt $v = \sum_{i=1}^{n}(vv_i)v_i$.

14.37 Satz (Orthonormalisierungsverfahren von Gram-Schmidt)

Sei V ein euklidischer oder unitärer Raum endlicher Dimension. Dann lässt sich jedes Orthonormalsystem zu einer Orthonormalbasis fortsetzen.

Beweis

Sei $v_1, \ldots, v_m$ ein Orthonormalsystem. Dann sind $v_1, \ldots, v_m$ nach Satz 14.35 linear unabhängig und lassen sich daher zu einer Basis $v_1, \ldots, v_m, v_{m+1}, \ldots, v_n$ von V fortsetzen. Für $0 \leq k \leq n-1$ setzt man

$$e_{k+1} = \frac{v_{k+1} - \sum_{i=1}^{k} (v_{k+1} e_i) e_i}{|v_{k+1} - \sum_{i=1}^{k} (v_{k+1} e_i) e_i|}.$$

Offenbar ist dabei e_1 wegen $|v_1| = 1 \Longrightarrow e_1 = v_1$ wohldefiniert mit $\langle e_1 \rangle = \langle v_1 \rangle$. Seien nun $e_1, \ldots, e_{\widetilde{n}}$ mit $\langle e_1, \ldots, e_{\widetilde{n}} \rangle = \langle v_1, \ldots v_{\widetilde{n}} \rangle$ für ein $\widetilde{n} < n$ wohldefiniert. Mit $\sum_{i=1}^{\widetilde{n}} (v_{\widetilde{n}+1} e_i) e_i \in \langle e_1, \ldots, e_{\widetilde{n}} \rangle = \langle v_1, \ldots v_{\widetilde{n}} \rangle$ und $v_{\widetilde{n}+1} \notin \langle v_1, \ldots v_{\widetilde{n}} \rangle$ folgt dann

$$0 \neq v_{\widetilde{n}+1} - \sum_{i=1}^{\widetilde{n}} (v_{\widetilde{n}+1} e_i) e_i \in \langle v_1, \ldots v_{\widetilde{n}+1} \rangle.$$

Daher ist $e_{\widetilde{n}+1}$ wohldefiniert mit $e_{\widetilde{n}+1} \in \langle v_1, \ldots v_{\widetilde{n}+1} \rangle$. Andererseits gilt

$$v_{\widetilde{n}+1} = |v_{\widetilde{n}+1} - \sum_{i=1}^{\widetilde{n}} (v_{\widetilde{n}+1} e_i) e_i| \cdot e_{\widetilde{n}+1} + \sum_{i=1}^{\widetilde{n}} (v_{\widetilde{n}+1} e_i) e_i \in \langle e_1, \ldots, e_{\widetilde{n}+1} \rangle.$$

Insgesamt erhält man also $\langle e_1, \ldots, e_{\widetilde{n}+1} \rangle = \langle v_1, \ldots, v_{\widetilde{n}+1} \rangle$. Durch vollständige Induktion zeigt sich somit, dass e_{k+1} für $0 \leq k \leq n-1$ wohldefiniert ist.

Für $1 \leq i \leq \widetilde{m} < m$ gilt nun $v_{\widetilde{m}+1} v_i = 0$ und $|v_{\widetilde{m}+1}| = 1$. Mit $e_j = v_j$ für $1 \leq j \leq \widetilde{m}$ folgt dann

$$e_{\widetilde{m}+1} = \frac{v_{\widetilde{m}+1} - \sum_{i=1}^{\widetilde{m}} (v_{\widetilde{m}+1} e_i) e_i}{|v_{\widetilde{m}+1} - \sum_{i=1}^{\widetilde{m}} (v_{\widetilde{m}+1} e_i) e_i|} = \frac{v_{\widetilde{m}+1} - \sum_{i=1}^{\widetilde{m}} (v_{\widetilde{m}+1} v_i) v_i}{|v_{\widetilde{m}+1} - \sum_{i=1}^{\widetilde{m}} (v_{\widetilde{m}+1} v_i) v_i|} = \frac{v_{\widetilde{m}+1}}{|v_{\widetilde{m}+1}|} = v_{\widetilde{m}+1}$$

und durch vollständige Induktion erhält man $e_j = v_j$ für $1 \leq j \leq m$.

Zuletzt gilt für $1 \leq i \leq n$ stets $|e_i| = 1$ – also auch $e_i \neq 0$. Daher bildet $e_1, \ldots, e_n$ nach Definition 14.34 und Satz 14.35 genau dann eine Orthonormalbasis, wenn $e_i e_j = 0$ für $1 \leq i, j \leq n$ und $i \neq j$ gilt. Dazu weisen wir durch vollständige Induktion nach, dass die Vektoren $e_1, \ldots, e_k$ für $1 \leq k \leq n$ orthogonal sind. Für $k = 1$ ist die Aussage trivial. Sei nun $e_i e_j = 0$ für $1 \leq i, j \leq k$ und $i \neq j$. Dann folgt $\sum_{i=1}^{k} (v_{k+1} e_i)(e_i e_j) = (v_{k+1} e_j)(e_j e_j) = (v_{k+1} e_j) |e_j|^2 = (v_{k+1} e_j)$

und damit

$$e_{k+1}e_j = \frac{v_{k+1} - \sum_{i=1}^{k}(v_{k+1}e_i)e_i}{|v_{k+1} - \sum_{i=1}^{k}(v_{k+1}e_i)e_i|}\, e_j = \frac{v_{k+1}e_j - \sum_{i=1}^{k}(v_{k+1}e_i)(e_ie_j)}{|v_{k+1} - \sum_{i=1}^{k}(v_{k+1}e_i)e_i|} = 0.$$

Beispiel

1. [Winkel zwischen Vektoren] Man betrachte den Vektorraum $V = \{f : [0,1] \longrightarrow \mathbb{R} \mid f$ stetig$\}$ über $\mathbb{R}$ und die Abbildung $\varphi(f,g) = \int_0^1 f(x)g(x)\mathrm{dx}$. Aus den elementaren Eigenschaften der Integralrechnung folgen für φ sofort Bilinearität und Symmetrie. Weiter gilt für $0 \neq f \in V$ stets $f^2(x) \geq 0$ und es existiert ein $x \in [0,1]$ mit $f^2(x) > 0$. Da nun f stetig ist, folgt damit $\varphi(f,f) = \int_0^1 f^2(x)\mathrm{dx} > 0$ und φ ist positiv definit.

   Der Winkel zwischen $f : x \longmapsto 1 \in V$ und $g : x \longmapsto x \in V$ ist dann durch

$$\cos\alpha = \frac{\varphi(f,g)}{\sqrt{\varphi(f,f)}\,\sqrt{\varphi(g,g)}} = \frac{\int_0^1 x\,\mathrm{dx}}{\sqrt{\int_0^1 1\,\mathrm{dx}}\,\sqrt{\int_0^1 x^2\,\mathrm{dx}}} = \frac{\sqrt{3}}{2}$$

   gegeben.

2. [Orthonormalisierung] Die Menge $V = \{f : x \longmapsto ax + b \mid a,b \in \mathbb{R}\}$ ist bzgl. $\varphi(f,g) = \int_0^1 f(x)g(x)\mathrm{dx}$ ein euklidischer Raum. Dabei ist $\{1,x\}$ eine Basis von V und $\{1\}$ wegen $|1|^2 = \varphi(1,1) = \int_0^1 1\,\mathrm{dx} = 1$ ein Orthonormalsystem.

   Wir setzen nun $\{1\}$ zu einer Orthonormalbasis $\{e_1,e_2\}$ von V fort. Dazu sei $e_1 = 1$ und

$$e_2 = \frac{x - \varphi(x,1)\cdot 1}{|x - \varphi(x,1)\cdot 1|} = \frac{x - \frac{1}{2}}{|x - \frac{1}{2}|}$$

   mit $|x - \frac{1}{2}|^2 = \int_0^1 (x - \frac{1}{2})^2\,\mathrm{dx} = \left[\frac{1}{3}(x - \frac{1}{2})^3\right]_0^1 = \frac{1}{12}$, also $e_2 = 2\sqrt{3}x - \sqrt{3}$.

3. [Orthogonalsystem in unitärem Raum] Sei der unitäre Raum $V = \{f : \mathbb{R} \longrightarrow \mathbb{C} \mid f$ periodisch über 2π und f stetig in $x \neq 2\pi k$ für $k \in \mathbb{Z}\}$ mit dem Skalarprodukt $\psi(f,g) = \int_0^{2\pi} f(x)\,\overline{g(x)}\mathrm{dx}$ gegeben. Dabei hat V keine endliche Dimension.

   Für $k \in \mathbb{Z}$ setzt man $f_k(x) = \mathrm{e}^{ikx}$ mit $i^2 = -1$. Mit $k \neq l$ gilt dann $\psi(f_k,f_l) = \int_0^{2\pi} \mathrm{e}^{ikx}\,\overline{\mathrm{e}^{ilx}}\,\mathrm{dx} = \int_0^{2\pi} \mathrm{e}^{ikx}\,\mathrm{e}^{-ilx}\,\mathrm{dx} = \int_0^{2\pi} \mathrm{e}^{i(k-l)x}\,\mathrm{dx} = \left[\frac{\mathrm{e}^{i(k-l)x}}{i(k-l)}\right]_0^{2\pi} = 0$. Daher bildet $\{f_k \mid k \in \mathbb{Z}\}$ ein Orthogonalsystem. Jedoch ist $\{f_k \mid k \in \mathbb{Z}\}$ keine Basis von V.

14.38 Definition

Sei V ein euklidischer oder unitärer Raum. Die Teilmengen $M \subseteq V$ und $N \subseteq V$ heißen orthogonal, falls für alle $m \in M$ und $n \in N$ stets $m \perp n$ gilt. Man schreibt dann $M \perp N$. Dabei verwendet man statt $\{v\} \perp M$ die Bezeichnung $v \perp M$. Für die Teilmenge $M \subseteq V$ wird $M^\perp = \{v \in V \mid v \perp M\}$ das orthogonale Komplement genannt.

Bemerkung

Sei $M \subseteq V$ Teilmenge eines euklidischen oder unitären Raums V. Dann gilt $0 \in M^\perp$, also $M^\perp \neq \emptyset$. Seien nun $v, w \in M^\perp$, d.h. es gilt $vm = 0$ und $wm = 0$ für alle $m \in M$. Es folgt $(\lambda v + \mu w)m = \lambda(vm) + \mu(wm) = 0$, also $\lambda v + \mu w \in M^\perp$. Insgesamt ist also $M^\perp$ ein Unterraum.

14.39 Lemma

Für Teilmengen $M \subseteq V$ und $N \subseteq V$ eines euklidischen oder unitären Raums gilt

1. $M \subseteq N \Longrightarrow N^\perp \subseteq M^\perp$
2. $M^\perp = \langle M \rangle^\perp$
3. $M \perp N \Longleftrightarrow \langle M \rangle \perp \langle N \rangle$

Beweis

1 Sei $M \subseteq N$ und $v \in N^\perp$. Dann gilt mit $v \perp N$ insbesondere $v \perp M$, also $v \in M^\perp$.

2 Nach 1 gilt wegen $M \subseteq \langle M \rangle$ zunächst $\langle M \rangle^\perp \subseteq M^\perp$. Andererseits ist jeder Vektor $w \in \langle M \rangle$ eine Linearkombination $w = \sum \lambda_i m_i$ für $m_i \in M$. Mit $v \in M^\perp$ folgt dann $vw = v(\sum \lambda_i m_i) = \sum \overline{\lambda_i}\,(vm_i) = 0$, also $v \in \langle M \rangle^\perp$. Daher folgt insgesamt $M^\perp = \langle M \rangle^\perp$.

3 Sei $M \perp N$, d.h. $M \subseteq N^\perp$. Nach 2 gilt dann $M \subseteq \langle N \rangle^\perp$, also $M \perp \langle N \rangle$. Damit folgt analog $\langle N \rangle \subseteq M^\perp = \langle M \rangle^\perp \Longrightarrow \langle N \rangle \perp \langle M \rangle$. Es gelte nun andererseits $\langle M \rangle \perp \langle N \rangle$. Dann folgt $M \subseteq \langle M \rangle \subseteq \langle N \rangle^\perp \Longrightarrow M \perp \langle N \rangle$ und damit $N \subseteq \langle N \rangle \subseteq M^\perp \Longrightarrow N \perp M$.

14.40 Satz

Sei V ein euklidischer oder unitärer Raum endlicher Dimension und U ein Unterraum von V. Dann gilt

1. $V = U \oplus U^\perp$
2. $(U^\perp)^\perp = U$

Beweis

1 Sei $v_1, \ldots, v_m$ eine Orthonormalbasis von U. Mit Satz 14.37 ergänze man $v_1, \ldots, v_m$ zu einer Orthonormalbasis $v_1, \ldots, v_m, v_{m+1}, \ldots, v_n$ von V. Für $W = \langle v_{m+1}, \ldots, v_n \rangle$ gilt dann

$V = U + W$. Da nun $v_1, \ldots, v_n$ insbesondere ein Orthogonalsystem ist, gilt $\{v_1, \ldots, v_m\} \perp \{v_{m+1}, \ldots, v_n\}$ und nach Lemma 14.39 folgt $U \perp W$. Damit erhält man $W \subseteq U^\perp$, also $V = U + U^\perp$. Aus der positiven Definitheit des Skalarprodukts folgt nun die Behauptung mit $u \in U \cap U^\perp \Longrightarrow uu = 0 \Longrightarrow u = 0 \Longrightarrow U \cap U^\perp = \{0\}$.

2 Nach 1 gilt $V = U \oplus U^\perp$ sowie $V = U^\perp \oplus (U^\perp)^\perp$ und es folgt $\dim U + \dim U^\perp = \dim V = \dim U^\perp + \dim(U^\perp)^\perp \Longrightarrow \dim U = \dim(U^\perp)^\perp$. Damit folgt die Behauptung aus $U \perp U^\perp \Longrightarrow U \subseteq (U^\perp)^\perp$.

15 Adjungierte und normale Endomorphismen

Bemerkung

In diesem Kapitel sei V stets ein euklidischer oder unitärer Raum.

Beispiel

Sei $V = \mathbb{R}^n$ versehen mit dem Standardskalarprodukt $vw = v^t w$. Wir suchen nun Endomorphismen, die Kongruenzabbildungen sind und Länge sowie Winkel erhalten. Insbesondere muss für eine solche Abbildung $\alpha \in \mathrm{End}(V)$ gelten

$$(*) \quad \forall_{v,w \in V} \ \alpha(v)\alpha(w) = vw,$$

denn dann ist etwa $|v| = |\alpha(v)|$. Sei nun α durch $A \in M_{n\times n}(\mathbb{R})$ dargestellt, d.h. es ist $\alpha(v) = Av$. Die Bedingung $(*)$ besagt dann

$$v^t(A^tA)w = (Av)^t(Aw) = \alpha(v)^t\alpha(w) = \alpha(v)\alpha(w) = vw = v^t w$$

und daher $v^t(1_n - A^tA)w = 0$. Setzt man in dieses Matrixprodukt für v und w die Standardbasisvektoren ein, so erhält man $1_n - A^tA = 0$ und damit $1_n = A^tA$. Insbesondere gilt also $A^tA = AA^t$.

15.1 Definition

Sei $\alpha \in \mathrm{End}(V)$. Dann heißt $\alpha^* \in \mathrm{End}(V)$ mit $\alpha(v)w = v\alpha^*(w)$ für alle $v, w \in V$ zu α adjungiert.

Beispiel / Bemerkung

Sei $V = \mathbb{C}^n$ mit dem Standardskalarprodukt $vw = v^t\,\overline{w}$ und $\alpha(v) = Av$ für $A \in M_{n\times n}(\mathbb{C})$. Dann wird α^* wegen

$$(*)\quad v(\overline{A}^t\,w) = v^t\,(\overline{\overline{A}^t\,w}) = v^t(A^t\,\overline{w}) = (v^tA^t)\,\overline{w} = (Av)^t\,\overline{w} = \alpha(v)^t\,\overline{w} = \alpha(v)w$$

durch $A^* = \overline{A}^t$ beschrieben. Dabei werden in $(*)$ die Eigenschaften von Matrixprodukt und Skalarprodukt verwendet, wobei natürlich nur das Matrixprodukt assoziativ ist.

15.2 Satz

Zu jedem $\alpha \in \mathrm{End}(V)$ existiert höchstens ein adjungierter Endomorphismus α^*. Für $\dim V < \infty$ gibt es stets eine Adjungierte von α.

Beweis

[Eindeutigkeit] Seien α^* und β zu α adjungiert und sei $\gamma = \alpha^* - \beta$. Es folgt $0 = \alpha(v)w - \alpha(v)w = v\alpha^*(w) - v\beta(w) = v\big(\alpha^*(w) - \beta(w)\big) = v\gamma(w)$ und für $v = \gamma(w)$ erhält man daher $0 = \gamma(w)\gamma(w) = |\gamma(w)|^2 \Longrightarrow \alpha^*(w) - \beta(w) = \gamma(w) = 0 \Longrightarrow \alpha^*(w) = \beta(w)$ für alle $w \in V$.

[Existenz] Sei $e_1, \ldots, e_n$ eine Orthonormalbasis von V und sei $\alpha^*(w) = \sum\limits_{i=1}^{n} \big(w\alpha(e_i)\big)e_i$. Mit den Eigenschaften des Skalarprodukts folgt nun

$$\begin{aligned}\alpha^*(v+w) &= \sum_{i=1}^{n}\big((v+w)\alpha(e_i)\big)e_i &&= \sum_{i=1}^{n}\big(v\alpha(e_i) + w\alpha(e_i)\big)e_i\\ &= \sum_{i=1}^{n}\big(v\alpha(e_i)\big)e_i + \sum_{i=1}^{n}\big(w\alpha(e_i)\big)e_i &&= \alpha^*(v) + \alpha^*(w)\end{aligned}$$

und

$$\alpha^*(\lambda w) = \sum_{i=1}^{n}\big((\lambda w)\alpha(e_i)\big)e_i = \sum_{i=1}^{n}\lambda\big(w\alpha(e_i)\big)e_i = \lambda\sum_{i=1}^{n}\big(w\alpha(e_i)\big)e_i = \lambda\alpha^*(w),$$

also $\alpha^* \in \mathrm{End}(V)$. Es bleibt zu zeigen, dass α^* zu α adjungiert ist. Da nach Satz 14.36 nun $v = \sum\limits_{i=1}^{n}(ve_i)e_i$ für alle $v \in V$ gilt, folgt

$$\begin{aligned}\alpha(v)w &= \alpha\big(\sum_{i=1}^{n}(ve_i)e_i\big)w &&= \big(\sum_{i=1}^{n}(ve_i)\alpha(e_i)\big)w\\ &= \sum_{i=1}^{n}(ve_i)\big(\alpha(e_i)w\big) &&= \sum_{i=1}^{n}(ve_i)\big(\overline{w\alpha(e_i)}\big)\\ &= \sum_{i=1}^{n}v\Big(\big(\alpha(e_i)w\big)e_i\Big) &&= v\Big(\sum_{i=1}^{n}\big(\alpha(e_i)w\big)e_i\Big) \;=\; v\alpha^*(w).\end{aligned}$$

15.3 Satz

Sei α^* die Adjungierte von $\alpha \in \mathrm{End}(V)$. Dann existiert $\alpha^{**} = (\alpha^*)^*$ mit $\alpha^{**} = \alpha$.

Beweis

Für alle $v, w \in V$ gilt $\alpha^*(v)w = \overline{w\alpha^*(v)} = \overline{\alpha(w)v} = v\alpha(w)$, d.h. $\alpha^{**}(w) = \alpha(w)$.

15.4 Satz

Seien α^* und β^* die Adjungierten von α und β. Dann existiert $(\alpha\beta)^*$ mit $(\alpha\beta)^* = \beta^*\alpha^*$.

Beweis

Es gilt $\alpha\beta(v)w = \alpha\big(\beta(v)\big)w = \beta(v)\alpha^*(w) = v\beta^*\big(\alpha^*(w)\big) = v\beta^*\alpha^*(w)$.

15.5 Satz

Sei α^* die Adjungierte von $\alpha \in \mathrm{End}(V)$. Dann gilt $\mathrm{Kern}\,\alpha^* = (\mathrm{Bild}\,\alpha)^\perp$ und für $\dim V < \infty$ ist $V = \mathrm{Bild}\,\alpha \oplus \mathrm{Kern}\,\alpha^* = \mathrm{Bild}\,\alpha^* \oplus \mathrm{Kern}\,\alpha$.

Beweis

Es gilt $w \in \mathrm{Kern}\,\alpha^* \iff \alpha^*(w) = 0$ und damit $v\alpha^*(w) = v0 = \overline{0}(vv) = 0(vv) = 0$ für alle $v \in V$. Umgekehrt gilt $v\alpha^*(w) = 0 \;\forall\; v \in V \Longrightarrow |\alpha^*(w)|^2 = \alpha^*(w)\alpha^*(w) = 0 \Longrightarrow \alpha^*(w) = 0 \Longrightarrow w \in \mathrm{Kern}\,\alpha^*$. Man erhält also $w \in \mathrm{Kern}\,\alpha^* \iff \alpha(v)w = v\alpha^*(w) = 0 \;\forall\; v \in V \iff uw = 0 \;\forall\; u \in \mathrm{Bild}\,\alpha \iff v \in (\mathrm{Bild}\,\alpha)^\perp$ und daher $\mathrm{Kern}\,\alpha^* = (\mathrm{Bild}\,\alpha)^\perp$.

Mit $\dim V < \infty$ folgt nach Satz 14.40 nun $V = \mathrm{Bild}\,\alpha \oplus (\mathrm{Bild}\,\alpha)^\perp = \mathrm{Bild}\,\alpha \oplus \mathrm{Kern}\,\alpha^*$ und wegen $\alpha^{**} = \alpha$ ergibt sich analog $V = \mathrm{Bild}\,\alpha^* \oplus (\mathrm{Bild}\,\alpha^*)^\perp = \mathrm{Bild}\,\alpha^* \oplus \mathrm{Kern}\,\alpha^{**} = \mathrm{Bild}\,\alpha^* \oplus \mathrm{Kern}\,\alpha$.

15.6 Definition

Sei α^* die Adjungierte von $\alpha \in \mathrm{End}(V)$. Dann heißt α normal, falls $\alpha\alpha^* = \alpha^*\alpha$ gilt.

15.7 Lemma

Ein Endomorphismus $\alpha \in \mathrm{End}(V)$ mit Adjungierter α^* ist genau dann normal, wenn $\alpha^*(v)\alpha^*(w) = \alpha(v)\alpha(w)$ für alle $v, w \in V$ gilt.

Beweis

Sei zunächst α normal. Dann gilt

$$\alpha(v)\alpha(w) = v\alpha * \big(\alpha(w)\big) = v\alpha\big(\alpha^*(w)\big) = \overline{\alpha\big(\alpha^*(w)\big)v} = \overline{\alpha^*(w)\alpha^*(v)} = \alpha^*(v)\alpha^*(w)$$

für alle $v, w \in V$. Es gelte nun $\alpha^*(v)\alpha^*(w) = \alpha(v)\alpha(w)$ für $v, w \in V$. Dann folgt

$$v\alpha^*\big(\alpha(w)\big) = \alpha(v)\alpha(w) = \alpha^*(v)\alpha^*(w) = v\alpha^{**}\big(\alpha^*(w)\big) = v\alpha\big(\alpha^*(w)\big),$$

d.h. $v\big((\alpha^*\alpha - \alpha\alpha^*)(w)\big) = v\alpha^*\big(\alpha(w)\big) - v\alpha\big(\alpha^*(w)\big) = 0$. Für $v = (\alpha^*\alpha - \alpha\alpha^*)(w)$ erhält man $|(\alpha^*\alpha - \alpha\alpha^*)(w)|^2 = \big((\alpha^*\alpha - \alpha\alpha^*)(w)\big)\big((\alpha^*\alpha - \alpha\alpha^*)(w)\big) = 0 \Longrightarrow (\alpha^*\alpha - \alpha\alpha^*)(w) = 0$. Daher gilt $\alpha^*\alpha(w) = \alpha\alpha^*(w)$ für alle $w \in V$, d.h. $\alpha\alpha^* = \alpha^*\alpha$.

15.8 Satz

Sei $\alpha \in \mathrm{End}(V)$ normal. Dann gilt

1. Ist v ein Eigenvektor von α mit Eigenwert λ, so ist v ein Eigenvektor von α^* mit Eigenwert $\overline{\lambda}$.
2. Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal.

Beweis

1 Sei v ein Eigenvektor von α mit Eigenwert λ. Dann gilt $0 = \alpha(v) - \lambda v$ und daher $0 = |\alpha(v) - \lambda v|^2 = (\alpha(v) - \lambda v)(\alpha(v) - \lambda v) = \alpha(v)\alpha(v) - \lambda\big(v\alpha(v)\big) - \overline{\lambda}\,\big(\alpha(v)v\big) + \lambda\,\overline{\lambda}\,(vv)$. Nach Satz 15.3 und Lemma 15.7 folgt damit

$$\begin{aligned} 0 &= \alpha^*(v)\alpha^*(v) - \lambda\big(\alpha^*(v)v\big) - \overline{\lambda}\,\big(v\alpha^*(v)\big) + \lambda\,\overline{\lambda}\,(vv) \\ &= (\alpha^*(v) - \overline{\lambda}\,v)(\alpha^*(v) - \overline{\lambda}\,v) \\ &= |\alpha^*(v) - \overline{\lambda}\,v|^2, \end{aligned}$$

d.h. $\alpha^*(v) - \overline{\lambda}\,v = 0 \Longrightarrow \alpha^*(v) = \overline{\lambda}\,v$.

2 Seien $v \in V$ und $w \in V$ Eigenvektoren von α mit den Eigenwerten λ und μ. Dann gilt $\lambda(vw) = (\lambda v)w = \alpha(v)w = v\alpha^*(w) = v(\overline{\mu}\, w) = \mu(vw)$ und man erhält

$$(\lambda - \mu)(vw) = \lambda(vw) - \mu(vw) = 0.$$

Für $\lambda \neq \mu$ gilt daher $vw = 0$ und es folgt die Behauptung.

15.9 Satz

1. Für $\alpha \in \mathrm{End}(V)$ sei $\alpha = \alpha^*$. Dann sind die Eigenwerte von α reell.
2. Die Eigenwerte einer symmetrischen Matrix $A \in M_{n\times n}(\mathbb{R})$ bzw. einer hermiteschen Matrix $A \in M_{n\times n}(\mathbb{C})$ sind reell.

Beweis

1 Ist V euklidsch, so betrachte man $V_{\mathbb{C}}$. Sei also o.B.d.A. V ein unitärer Raum und v ein Eigenvektor von α mit Eigenwert λ. Nach Satz 15.8 ist dann v ein Eigenvektor von α^* mit Eigenwert $\overline{\lambda}$, d.h. $\lambda v = \alpha(v) = \alpha^*(v) = \overline{\lambda}\, v$. Damit gilt $\lambda = \overline{\lambda} \Longrightarrow \lambda \in \mathbb{R}$.

2 Wir betrachten die symmetrische Matrix $A \in M_{n\times n}(\mathbb{R})$, da die Aussage für hermitesche Matrizen analog gilt. Dann folgt mit dem Stanardskalarprodukt

$$(Av)w = (Av)^t w = (v^t A^t)w = (v^t A)w = v^t(Aw) = v(Aw)$$

und daher $\alpha = \alpha^*$ für $\alpha : v \longmapsto Av$. Da nach 1 die Eigenwerte von α reell sind, folgt die Behauptung.

15.10 Lemma

Sei $\alpha \in \mathrm{End}(V)$ normal mit $\dim V < \infty$ und sei U ein α-invarianter Unterraum von V.

1. Dann ist $U^\perp$ ein α^*-invarianter Unterraum.
2. Ist auch $U^\perp$ α-invariant, so ist U α^*-invariant mit $(\alpha|_U)^* = \alpha^*|_U$. Insbesondere ist dann $\alpha|_U$ normal.

Beweis

1 Sei $w \in U^\perp$, d.h. $uw = 0$ für alle $u \in U$. Da U α-invariant ist, gilt insbesondere $0 = \alpha(u)w = u\alpha^*(w)$ $\forall\ u \in U$ und daher $\alpha^*(w) \in U^\perp$. Also ist $U^\perp$ α^*-invariant.

2 Sei nun $U^\perp$ auch α-invariant. Nach 1 und wegen $U^{\perp\perp} = U$ ist dann U α^*-invariant. Für $u, v \in U$ folgt damit $\alpha|_U(u)v = \alpha(u)v = u\alpha^*(v) = u\alpha^*|_U(v)$, d.h. $(\alpha|_U)^* = \alpha^*|_U$. Da zuletzt α normal ist, folgt für alle $u \in U$ stets $\alpha|_U\big(\alpha^*|_U(u)\big) = \alpha\big(\alpha^*(u)\big) = \alpha^*\big(\alpha(u)\big) = \alpha^*|_U\big(\alpha|_U(u)\big)$ und daher $\alpha|_U \circ \alpha^*|_U = \alpha^*|_U \circ \alpha|_U$.

15.11 Lemma

Sei $\alpha \in \mathrm{End}(V)$ normal mit $\dim V < \infty$ und sei v ein Eigenvektor von α. Dann ist $W = \{v\}^\perp$ ein α-invarianter Unterraum. Weiter ist $\alpha|_W$ normal.

Beweis

Nach Lemma 14.39 gilt zunächst $W = \{v\}^\perp = \langle v\rangle^\perp$. Sei nun λ der Eigenwert von v, d.h. nach Satz 15.8 gilt $\alpha^*(v) = \overline{\lambda}\, v$. Daher ist $\langle v\rangle$ ein α^*-invarianter Unterraum und mit Lemma 15.10 ist $\langle v\rangle^\perp$ α^{**}-invariant. Wegen $\alpha^{**} = \alpha$ und $\langle v\rangle^\perp = W$ folgt damit die Behauptung. Weiter ist $W^\perp$ wegen $W^\perp = \langle v\rangle^{\perp\perp} = \langle v\rangle$ und $\alpha(v) = \lambda v$ auch α-invariant. Nach Lemma 15.10 ist damit $\alpha|_W$ normal.

Bemerkung

Sei $v_1, \ldots, v_n$ eine Orthonormalbasis von V, d.h. $v_i v_i = 1$ und $v_i v_j = 0$ für $i \neq j$. In diesem Fall schreibt man $v_i v_j = \delta_{ij}$.

15.12 Satz (Spektralsatz für unitäre Räume)

Sei V ein unitärer Raum endlicher Dimension und sei $\alpha \in \mathrm{End}(V)$. Dann ist α genau dann normal, wenn eine Orthonormalbasis aus Eigenvektoren von α existiert.

Beweis

Es existiere zunächst eine Orthonormalbasis $v_1, \ldots, v_n$ von V aus Eigenvektoren von α, d.h. $\alpha(v_i) = \lambda_i v_i$ für $1 \leq i \leq n$ und $\lambda_i \in \mathbb{C}$. Nach 15.2 existiert nun wegen $\dim V < \infty$ ein adjungierter Endomorphismus $\alpha^* \in \mathrm{End}(V)$. Sei dann $\alpha^*(v_j) = \sum_{k=1}^{n} \mu_{jk} v_k$ für $1 \leq j \leq n$ und

$\mu_{jk} \in \mathbb{C}$. Es folgt

$$\begin{aligned}
\lambda_i\delta_{ij} &= \lambda_i(v_iv_j) &&= (\lambda_iv_i)v_j \\
&= \alpha(v_i)v_j &&= v_i\alpha^*(v_j) \\
&= v_i(\sum_{k=1}^{n}\mu_{jk}v_k) &&= \sum_{k=1}^{n}\overline{\mu_{jk}}\,(v_iv_k) \\
&= \sum_{k=1}^{n}\overline{\mu_{jk}}\,\delta_{ik} &&= \overline{\mu_{ji}}
\end{aligned}$$

für $1 \leq i \leq n$ und daher $\mu_{ii} = \overline{\lambda_i}$ sowie $\mu_{ij} = 0$ mit $i \neq j$. Man erhält also $\alpha^*(v_i) = \sum_{k=1}^{n}\mu_{ik}v_k = \overline{\lambda_i}\,v_i$ und somit

$$\begin{aligned}
\alpha\big(\alpha^*(v_i)\big) &= \alpha(\overline{\lambda_i}\,v_i) &&= \overline{\lambda_i}\,\alpha(v_i) &&= \overline{\lambda_i}\,(\lambda_iv_i) \\
&= (\overline{\lambda_i}\,\lambda_i)v_i &&= (\lambda_i\,\overline{\lambda_i})v_i &&= \lambda_i(\overline{\lambda_i}\,v_i) \\
&= \lambda_i\alpha^*(v_i) &&= \alpha^*(\lambda_iv_i) &&= \alpha^*\big(\alpha(v_i)\big).
\end{aligned}$$

Für jeden Vektoren $v \in V$ mit $v = \sum_{i=1}^{n}a_iv_i$ folgt nun

$$\begin{aligned}
\alpha\big(\alpha^*(v)\big) &= \alpha\big(\alpha^*(\sum_{i=1}^{n}a_iv_i)\big) &&= \sum_{i=1}^{n}a_i\alpha\big(\alpha^*(v_i)\big) \\
&= \sum_{i=1}^{n}a_i\alpha^*\big(\alpha(v_i)\big) &&= \alpha^*\big(\alpha(\sum_{i=1}^{n}a_iv_i)\big) &&= \alpha^*\big(\alpha(v)\big),
\end{aligned}$$

d.h. α ist normal.

Nun sei α sei normal. Dabei ist $\mathbb{C}$ algebraisch abgeschlossen, d.h. das charakteristische Polynom von α hat stets komplexe Nullstellen. Somit hat α hat stets Eigenvektoren mit komplexen Eigenwerten. Die Aussage beweisen wir damit durch vollständige Induktion über $\dim V$. Dabei ist für $\dim V = 1$ jeder normierte Eigenvektor ein Basis der gesuchten Art.

Man betrachte nun den normierten Eigenvektor v von α. Nach Lemma 15.11 ist dann $U = \{v\}^\perp$ ein α-invarianter Unterraum von V und $\alpha|_U$ ist ein normaler Endomorphismus. Mit Lemma 14.39 und Satz 14.40 gilt $V = \langle v\rangle \oplus U$ und daher $\dim U = \dim V - 1$. Nach Induktionsannahme existiert nun eine Orthonormalbasis $v_2, \ldots, v_n$ von U aus Eigenvektoren von $\alpha|_U$. Dabei sind $v_2, \ldots, v_n$ insbesondere Eigenvektoren von α.

Die Vektoren $v, v_2, \ldots, v_n$ sind also insgesamt normierte Eigenvektoren. Nach Satz 14.35 bleibt daher zu zeigen, dass $v, v_2, \ldots, v_n$ ein Orthogonalsystem ist. Da dies eine Menge von Eigenvektoren ist, gilt zuerst $0 \notin \{v, v_2, \ldots, v_n\}$. Weiter ist $v_2, \ldots, v_n$ ein Orthogonalsystem und es gilt $v \in \langle v\rangle = \langle v\rangle^{\perp\perp} = (\{v\}^\perp)^\perp = U^\perp$.

15.13 Lemma

Sei V ein euklidischer Vektorraum sowie $\alpha \in \mathrm{End}(V)$ normal und sei $\widehat{\alpha}$ die komplexe Fortsetzung von α. Dann gilt:

1. $\widehat{\alpha^*} = \widehat{\alpha}^*$
2. $\widehat{\alpha}$ ist normal

Sei weiter $v = v_1 + iv_2$ mit $v_1, v_2 \in V$ ein Eigenvektor von $\widehat{\alpha}$ mit Eigenwert $\lambda \notin \mathbb{R}$. Dann gilt:

3. $\overline{v} = v_1 - iv_2$ ist ein Eigenvektor von $\widehat{\alpha}$ mit Eigenwert $\overline{\lambda}$
4. v und $\overline{v}$ sind orthogonal
5. $\overline{v}$ ist genau dann normiert, wenn v normiert ist

Beweis

1 Nach Lemma 14.26 gilt $\widehat{\alpha}(v_1 + iv_2) = \alpha(v_1) + i\alpha(v_2)$ und $\widehat{\alpha^*}(w_1 + iw_2) = \alpha^*(w_1) + i\alpha^*(w_2)$. Es folgt

$$\begin{aligned}
\widehat{\alpha}(v_1 + iv_2)(w_1 + iw_2) &= \big(\alpha(v_1) + i\alpha(v_2)\big)(w_1 + iw_2) \\
&= \alpha(v_1)w_1 + i\big(\alpha(v_2)w_1\big) - i\big(\alpha(v_1)w_2\big) + \alpha(v_2)w_2 \\
&= v_1\alpha^*(w_1) + i\big(v_2\alpha^*(w_1)\big) - i\big(v_1\alpha^*(w_2)\big) + v_2\alpha^*(w_2) \\
&= (v_1 + iv_2)\big(\alpha^*(w_1) + i\alpha^*(w_2)\big) \\
&= (v_1 + iv_2)\widehat{\alpha^*}(w_1 + iw_2)
\end{aligned}$$

und damit $\widehat{\alpha}^* = \widehat{\alpha^*}$.

2 Es gilt

$$\begin{aligned}
\widehat{\alpha}\big(\widehat{\alpha}^*(v_1 + iv_2)\big) &= \widehat{\alpha}\big(\alpha^*(v_1) + i\alpha^*(v_2)\big) \\
&= \alpha\big(\alpha^*(v_1)\big) + i\alpha\big(\alpha^*(v_2)\big) \\
&= \alpha^*\big(\alpha(v_1)\big) + i\alpha^*\big(\alpha(v_2)\big) \\
&= \widehat{\alpha^*}\big(\alpha(v_1) + i\alpha(v_2)\big) \qquad = \widehat{\alpha^*}\big(\widehat{\alpha}(v_1 + iv_2)\big)
\end{aligned}$$

für alle $v_1, v_2 \in V$, d.h. $\widehat{\alpha} \circ \widehat{\alpha}^* = \widehat{\alpha^*} \circ \widehat{\alpha}$. Nach 1 folgt nun die Behauptung.

3 Sei $\widehat{\alpha}(v_1 + iv_2) = \lambda(v_1 + iv_2)$ für $\lambda = \lambda_1 + i\lambda_2$ mit $\lambda_1, \lambda_2 \in \mathbb{R}$. Es folgt

$$\alpha(v_1) + i\alpha(v_2) = \widehat{\alpha}(v_1 + iv_2) = (\lambda_1 + i\lambda_2)(v_1 + iv_2) = (\lambda_1 v_1 - \lambda_2 v_2) + i(\lambda_1 v_2 + \lambda_2 v_1).$$

Da die Darstellung in $V_{\mathbb{C}}$ eindeutig ist, erhält man $\alpha(v_1) = \lambda_1 v_1 - \lambda_2 v_2$ und $\alpha(v_2) = \lambda_1 v_2 + \lambda_2 v_1$. Daher gilt

$$\widehat{\alpha}(v_1 - iv_2) = \alpha(v_1) - i\alpha(v_2) = (\lambda_1 v_1 - \lambda_2 v_2) - i(\lambda_1 v_2 + \lambda_2 v_1) = (\lambda_1 - i\lambda_2)(v_1 - iv_2),$$

d.h. $\widehat{\alpha}(\overline{v}) = \overline{\lambda}\,\overline{v}$.

4 Da $\lambda \notin \mathbb{R}$ ein echt komplexer Eigenwert ist, folgt $\lambda \neq \overline{\lambda}$. Nach Satz 15.8 sind dann v und $\overline{v}$ orthogonal.

5 Es gilt zunächst

$$|\overline{v}|^2 = (v_1 - iv_2)(v_1 - iv_2) = v_1 v_1 - i(v_2 v_1) + i(\overline{v_2 v_1}) + v_2 v_2.$$

Wegen $v_1, v_2 \in V$ gilt nun $v_2 v_1 \in \mathbb{R}$ und daher $\overline{v_2 v_1} = v_2 v_1$, d.h. $|\overline{v}|^2 = v_1 v_1 + v_2 v_2$. Analog folgt

$$|v|^2 = (v_1 + iv_2)(v_1 + iv_2) = v_1 v_1 + i(v_2 v_1) - i(\overline{v_2 v_1}) + v_2 v_2 = v_1 v_1 + v_2 v_2$$

und man erhält insgesamt $|\overline{v}| = |v|$. Damit folgt die Behauptung.

15.14 Satz (Spektralsatz für euklidische Räume)

Sei V ein euklidischer Vektorraum endlicher Dimension und $\alpha \in \mathrm{End}(V)$ normal. Dann besitzt V eine Orthonormalbasis bzgl. der α für $\lambda_i, a_i, b_i \in \mathbb{R}$ durch die Matrix

$$\begin{pmatrix}
\lambda_1 & & & & & & & \\
& \ddots & & & & & & \\
& & \lambda_m & & & & & \\
& & & a_1 & b_1 & & & \\
& & & -b_1 & a_1 & & & \\
& & & & & \ddots & & \\
& & & & & & a_r & b_r \\
& & & & & & -b_r & a_r
\end{pmatrix}$$

beschrieben wird.

Beweis

Wir beweisen die Aussage durch vollständige Induktion über $\dim V$. Für $\dim V = 1$ gilt dabei $V = \langle v \rangle$ mit $\alpha(v) = \lambda v$ für ein $\lambda \in \mathbb{R}$. Dann bildet $\frac{v}{|v|}$ eine Orthonormalbasis bzgl. der α durch (λ) beschrieben wird.

Nun unterscheiden wir zwei Fälle. Sei zuerst λ ein reeller Eigenwert von α mit normiertem Eigenvektor v. Nach Lemma 14.39 sowie Satz 14.40 gilt nun $V = \langle v\rangle \oplus \{v\}^\perp$ und nach Lemma 15.11 ist $\{v\}^\perp$ α-invariant sowie $\alpha|_{\{v\}^\perp}$ normal. Da weiter $\langle v\rangle$ offenbar α-invariant ist, wird α durch eine Blockdiagonalmatrix der Darstellungsmatrizen von $\alpha|_{\langle v\rangle}$ und $\alpha|_{\{v\}^\perp}$ beschrieben. Dabei wird $\alpha|_{\langle v\rangle}$ durch (λ) dargestellt. Mit $V = \langle v\rangle \oplus \{v\}^\perp$ gilt zudem $\dim\{v\}^\perp < \dim V$ und die Behauptung folgt durch Anwendung der Induktionsannahme auf $\{v\}^\perp$.

Jetzt betrachten wir einen Endomorphismus $\alpha \in \mathrm{End}(V)$ ohne reelle Eigenwerte. Sei dann λ ein Eigenwert der komplexen Fortsetzung $\widehat{\alpha}$ von α mit Eigenvektor $v = v_1 + iv_2$, d.h. $\alpha(v_1) + i\alpha(v_2) = \widehat{\alpha}(v_1 + iv_2) = \lambda(v_1 + iv_2) = (\lambda v_1) + i(\lambda v_2)$. Aus der Eindeutigkeit der Darstellung in $V_{\mathbb{C}}$ folgt damit $\alpha(v_1) = \lambda v_1$ und $\alpha(v_2) = \lambda v_2$. Da λ also auch ein Eigenwert von α ist, gilt $\lambda \notin \mathbb{R}$ und nach Lemma 15.13 ist somit $\overline{v} = v_1 - iv_2$ ein Eigenvektor von $\widehat{\alpha}$ mit Eigenwert $\overline{\lambda}$.

Man setze $w_1 = \frac{1}{\sqrt{2}}(v + \overline{v}) = \sqrt{2}v_1 \in V$ und $w_2 = \frac{1}{i\sqrt{2}}(v - \overline{v}) = \sqrt{2}v_2 \in V$ sowie $U = \langle w_1, w_2\rangle$. Seien dabei v und $\overline{v}$ o.B.d.A. normiert, d.h. $vv = \overline{v}\overline{v} = 1$. Da nach Lemma 15.13 nun nun $v\overline{v} = \overline{v}v = 0$ gilt, erhält man

$$|w_1|^2 = \big(\frac{1}{\sqrt{2}}(v + \overline{v})\big)\big(\frac{1}{\sqrt{2}}(v + \overline{v})\big) = \frac{1}{2}(vv + v\overline{v} + \overline{v}v + \overline{v}\,\overline{v}) = 1.$$

Analog folgt

$$|w_2|^2 = \big(\frac{1}{i\sqrt{2}}(v - \overline{v})\big)\big(\frac{1}{i\sqrt{2}}(v - \overline{v})\big) = \frac{1}{2}(vv - v\overline{v} - \overline{v}v + \overline{v}\,\overline{v}) = 1$$

und

$$w_1 w_2 = \big(\frac{1}{\sqrt{2}}(v + \overline{v})\big)\big(\frac{1}{i\sqrt{2}}(v - \overline{v})\big) = \frac{1}{-2i}(vv - v\overline{v} + \overline{v}v - \overline{v}\,\overline{v}) = 0.$$

Für $\lambda = a + ib$ mit $a, b \in \mathbb{R}$ gilt weiter

$$\begin{aligned}
\alpha(w_1) &= \widehat{\alpha}(w_1) &&= \widehat{\alpha}\big(\tfrac{1}{\sqrt{2}}(v + \overline{v})\big)\\
&= \tfrac{1}{\sqrt{2}}\big(\widehat{\alpha}(v) + \widehat{\alpha}(\overline{v})\big) &&= \tfrac{1}{\sqrt{2}}(\lambda v + \overline{\lambda}\,\overline{v})\\
&= \tfrac{1}{\sqrt{2}}\big((a + ib)v + (a - ib)\overline{v}\big) &&= \tfrac{1}{\sqrt{2}}\big(a(v + \overline{v}) + ib(v - \overline{v})\big)\\
&= a\tfrac{1}{\sqrt{2}}(v + \overline{v}) - b\tfrac{1}{i\sqrt{2}}(v - \overline{v}) &&= aw_1 - bw_2
\end{aligned}$$

sowie

$$\begin{aligned}
\alpha(w_2) &= \widehat{\alpha}(w_2) &&= \widehat{\alpha}\big(\tfrac{1}{i\sqrt{2}}(v - \overline{v})\big)\\
&= \tfrac{1}{i\sqrt{2}}\big(\widehat{\alpha}(v) - \widehat{\alpha}(\overline{v})\big) &&= \tfrac{1}{i\sqrt{2}}(\lambda v - \overline{\lambda}\,\overline{v})\\
&= \tfrac{1}{i\sqrt{2}}\big((a + ib)v - (a - ib)\overline{v}\big) &&= \tfrac{1}{i\sqrt{2}}\big(a(v - \overline{v}) + ib(v + \overline{v})\big)\\
&= a\tfrac{1}{i\sqrt{2}}(v - \overline{v}) + b\tfrac{1}{i\sqrt{2}}(v - \overline{v}) &&= aw_2 + bw_1.
\end{aligned}$$

Insgesamt ist also U ein α-invarianter Unterraum und $\alpha|_U$ wird bzgl. der Orthonormalbasis w_1, w_2 durch

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

beschrieben.

Nach Satz 15.2 existiert nun α^* und nach Satz 15.8 ist dabei v bzw. $\overline{v}$ ein Eigenvektor von $\widehat{\alpha}^*$ mit Eigenwert $\overline{\lambda}$ bzw λ. Es folgt

$$\begin{array}{rcccccl}
\alpha^*(w_1) & = & \widehat{\alpha^*}(w_1) & = & \widehat{\alpha}^*(w_1) & = & \widehat{\alpha}^*\big(\frac{1}{\sqrt{2}}(v+\overline{v})\big) \\
& & & = & \frac{1}{\sqrt{2}}\big(\widehat{\alpha}^*(v)+\widehat{\alpha}^*(\overline{v})\big) & = & \frac{1}{\sqrt{2}}(\overline{\lambda}\,v+\lambda\overline{v}) \\
& & & = & \frac{1}{\sqrt{2}}\big((a-ib)v+(a+ib)\overline{v}\big) & = & \frac{1}{\sqrt{2}}\big(a(v+\overline{v})-ib(v-\overline{v})\big) \\
& & & = & a\frac{1}{\sqrt{2}}(v+\overline{v})+b\frac{1}{i\sqrt{2}}(v-\overline{v}) & = & aw_1+bw_2
\end{array}$$

und

$$\begin{array}{rcccccl}
\alpha^*(w_2) & = & \widehat{\alpha^*}(w_2) & = & \widehat{\alpha}^*(w_2) & = & \widehat{\alpha}^*\big(\frac{1}{i\sqrt{2}}(v-\overline{v})\big) \\
& & & = & \frac{1}{i\sqrt{2}}\big(\widehat{\alpha}^*(v)-\widehat{\alpha}^*(\overline{v})\big) & = & \frac{1}{i\sqrt{2}}(\overline{\lambda}\,v-\lambda\overline{v}) \\
& & & = & \frac{1}{i\sqrt{2}}\big((a-ib)v-(a+ib)\overline{v}\big) & = & \frac{1}{i\sqrt{2}}\big(a(v-\overline{v})-ib(v+\overline{v})\big) \\
& & & = & a\frac{1}{i\sqrt{2}}(v-\overline{v})-b\frac{1}{i\sqrt{2}}(v-\overline{v}) & = & aw_2-bw_1.
\end{array}$$

Daher ist U auch α^*-invariant und nach Lemma 15.10 ist $U^\perp$ wegen $\alpha^{**}=\alpha$ wieder α-invariant. Da nach Satz 14.40 nun $V = U \oplus U^\perp$ gilt, wird α durch eine Blockdiagonalmatrix der Darstellungsmatrizen von $\alpha|_U$ und $\alpha|_{U^\perp}$ beschrieben.

Da weiter $U^\perp$ und $U^{\perp\perp} = U$ jeweils α-invariant sind, ist $\alpha|_{U^\perp}$ nach Lemma 15.10 normal. Die Induktionsannahme liefert daher wegen $V = U \oplus U^\perp \Longrightarrow \dim U^\perp < \dim V$ eine Orthonormalbasis $u_1, \ldots, u_n$ von $U^\perp$, so dass α bzgl. $w_1, w_2, u_1, \ldots, u_n$ durch eine Matrix der verlangten Art dargestellt wird. Mit $\{w_1, w_2\} \subseteq U = U^{\perp\perp} = \langle u_1, \ldots, u_n\rangle^\perp$ ist $w_1, w_2, u_1, \ldots, u_n$ schließlich eine Orthonormalbasis.

15.15 Korollar

Sei V ein euklidischer Vektorraum endlicher Dimension und $\alpha \in \mathrm{End}(V)$ selbstadjungiert, d.h. $\alpha^* = \alpha$. Dann besitzt V eine Orthonormalbasis aus Eigenvektoren von α.

Beweis

Da α wegen $\alpha^* = \alpha$ insbesondere normal ist, existiert nach Satz 15.14 eine Orthonormalbasis $v_1, \ldots, v_m, w_1, \ldots, w_r$ mit $\alpha(v_i) = \lambda_i v_i$ und $\alpha(w_j) = a_j w_j - b_j w_{j+1}$ sowie $\alpha(w_{j+1}) = b_j w_j + a_j w_{j+1}$ für $1 \leq i \leq m$ und $j = 1, 3, \ldots, r-1$ sowie $\lambda_i, a_j, b_j \in \mathbb{R}$.

Man betrachte nun o.B.d.A. w_1 sowie w_2 und setze $v = w_1 + iw_2 \in V_{\mathbb{C}}$. Für die komplexe Fortsetzung $\widehat{\alpha}$ von α folgt

$$\widehat{\alpha}(v) = \alpha(w_1) + i\alpha(w_2) = (aw_1 - bw_2) + i(bw_1 + aw_2) = (a + bi)(w_1 + iw_2).$$

Daher ist v ein Eigenvektor von $\widehat{\alpha}$ mit Eigenwert $\lambda = a + bi$ und nach Satz 15.8 ist v ein Eigenvektor von $\widehat{\alpha}^*$ mit Eigenwert $\overline{\lambda} = a - bi$. Man erhält also

$$(aw_1 + bw_2) + i(aw_2 - bw_1) = (a - bi)(w_1 + iw_2) = \overline{\lambda}\, v = \widehat{\alpha}^*(v) = \widehat{\alpha^*}(v) = \alpha^*(w_1) + i\alpha^*(w_2)$$

und wegen der Eindeutigkeit der darstellung in $V_{\mathbb{C}}$ folgt $\alpha^*(w_1) = aw_1 + bw_2$ sowie $\alpha^*(w_2) = aw_2 - bw_1$. Mit $\alpha = \alpha^*$ ergibts sich schließlich $b = 0$ und w_1 sowie w_2 sind Eigenvektoren von α.

16 Orthogonale und unitäre Endomorphismen

16.1 Definition

Sei V ein euklidischer bzw. unitärer Raum. Dann heißt $\alpha \in \mathrm{End}(V)$ orthogonal bzw. unitär, falls $\alpha(v)\alpha(w) = vw$ für alle $v, w \in V$ gilt. Dabei ist $\mathrm{O}(V) = \{\alpha \in \mathrm{End}(V) \mid \alpha \text{ orthogonal}\}$ bzw. $\mathrm{U}(V) = \{\alpha \in \mathrm{End}(V) \mid \alpha \text{ unitär}\}$.

Bemerkung

Die Bezeichnung „orthogonal“ ist dabei etwas unglücklich. Man betrachte etwa $V = \mathbb{R}^n$ mit dem Standardskalarprodukt und $\alpha \in \mathrm{End}(V)$ definiert durch $\alpha(v) = Av$ für

$$A = \begin{pmatrix} 2 & & \\ & \ddots & \\ & & 2 \end{pmatrix}.$$

Mit $v^t w = 0$ folgt dann auch $\alpha(v)^t \alpha(w) = (Av)^t (Aw) = (2v)^t (2w) = 4(v^t w) = 0$, aber für $v^t w \neq 0$ ist $\alpha(v)^t \alpha(w) = 4(v^t w) \neq v^t w$. Daher erhält α die Orthogonalität der Vektoren und ist dennoch nicht orthogonal.

16.2 Lemma

1. Orthogonale und unitäre Endomorphismen sind stets injektiv.
2. Für $\dim V < \infty$ sind $\mathrm{O}(V)$ und $\mathrm{U}(V)$ Untergruppen von $\mathrm{GL}(V) = \{\alpha \in \mathrm{End}(V) \mid \alpha$ invertierbar$\}$.

Beweis

1 Aus $v \in \operatorname{Kern}\alpha \iff \alpha(v) = 0$ und $\alpha \in \mathrm{O}(V)$ bzw. $\alpha \in \mathrm{U}(V)$ folgt $0 = \alpha(v)\alpha(v) = vv = |v|^2 \Longrightarrow v = 0$. Daher ist $\operatorname{Kern}\alpha$ trivial und α injektiv.

2 Sei nun $\dim V < \infty$ und $\alpha \in \mathrm{O}(V)$ bzw. $\alpha \in \mathrm{U}(V)$. Nach 1 ist dann α injektiv und im endlich-dimensionalen Fall sogar surjektiv, d.h. $\alpha \in \mathrm{GL}(V)$. Sei dabei $\alpha^{-1} = \beta$. Da α surjektiv ist, gibt es für $v, w \in V$ auch $\widetilde{v}, \widetilde{w} \in V$ mit $\alpha(\widetilde{v}) = v$ und $\alpha(\widetilde{w}) = w$. Es folgt $\beta(v)\beta(w) = \beta\big(\alpha(\widetilde{v})\big)\beta\big(\alpha(\widetilde{w})\big) = \widetilde{v}\widetilde{w} = \alpha(v)\alpha(w) = vw$, d.h. $\beta \in \mathrm{O}(V)$ bzw. $\beta \in \mathrm{U}(V)$.

Weiter seien $\alpha, \beta \in \mathrm{O}(V)$ bzw. $\alpha, \beta \in \mathrm{U}(V)$. Dann gilt $(\alpha\beta)(v)(\alpha\beta)(w) = \alpha\big(\beta(v)\big)\alpha\big(\beta(w)\big) = \beta(v)\beta(w) = vw$, d.h. $\alpha\beta \in \mathrm{O}(V)$ bzw. $\alpha\beta \in \mathrm{U}(V)$. Wegen $\mathrm{id}(v)\,\mathrm{id}(w) = vw$ – also $\mathrm{O}(V) \neq \emptyset$ bzw. $\mathrm{U}(V) \neq \emptyset$ – sind insgesamt $\mathrm{O}(V)$ und $\mathrm{U}(V)$ Untergruppen von $\mathrm{GL}(V)$.

16.3 Satz

Sei V euklidisch oder unitär mit $\alpha \in \mathrm{End}(V)$. Dann sind äquivalent:

1. α ist orthogonal bzw. unitär
2. aus $|v| = 1$ folgt stets $|\alpha(v)| = 1$
3. für alle $v \in V$ gilt $|v| = |\alpha(v)|$
4. Bilder unter α von Orthonormalsystemen sind wieder Orthonormalsysteme

Beweis

1 $\Longrightarrow$ **2** Sei α orthogonal bzw. unitär und $|v| = 1$. Dann gilt auch $1 = |v|^2 = vv = \alpha(v)\alpha(v) = |\alpha(v)|^2$, also $|\alpha(v)| = 1$.

2 $\Longrightarrow$ **3** Für $v = 0$ ist die Aussage trivial. Sei also $v \neq 0$ und $\mu = |v|$. Dann gilt $|\frac{1}{\mu}v| = 1$ und damit nach Voraussetzung auch $|\frac{1}{\mu}v| = |\alpha(\frac{1}{\mu}v)|$. Damit folgt

$$|\alpha(v)| = |\alpha(\mu\,\frac{1}{\mu}v)| = |\mu\,\alpha(\frac{1}{\mu}v)| = |\mu|\,|\alpha(\frac{1}{\mu}v)| = |\mu|\,|\frac{1}{\mu}v| = |\mu\,\frac{1}{\mu}v| = |v|.$$

3 $\Longrightarrow$ **1** Vorab gilt für eine komplexe Zahl $a+ib = z \in \mathbb{C}$ allgemein $z+\overline{z} = (a+ib)+(a-ib) = 2a = 2\,\mathrm{Re}(z)$ und weiter $\mathrm{Re}(-iz) = \mathrm{Re}(-ia+b) = b = \mathrm{Im}(z)$.

Seien nun $u, v \in V$ und $\lambda \in \mathbb{K}$. Dann gilt

$$\begin{aligned}
|\alpha(u+\lambda v)|^2 &= \alpha(u+\lambda v)\alpha(u+\lambda v) \\
&= \big(\alpha(u)+\lambda\alpha(v)\big)\big(\alpha(u)+\lambda\alpha(v)\big) \\
&= \alpha(u)\alpha(u) + \lambda\alpha(v)\alpha(u) + \overline{\lambda}\,\alpha(u)\alpha(v) + \lambda\,\overline{\lambda}\,\alpha(v)\alpha(v) \\
&= |\alpha(u)|^2 + \lambda\alpha(v)\alpha(u) + \overline{\lambda\alpha(v)\alpha(u)} + \lambda\,\overline{\lambda}\,|\alpha(v)|^2 \\
&= |\alpha(u)|^2 + 2\,\mathrm{Re}\big(\lambda\alpha(v)\alpha(u)\big) + \lambda\,\overline{\lambda}\,|\alpha(v)|^2
\end{aligned}$$

und analog $|u+\lambda v|^2 = |u|^2 + 2\,\mathrm{Re}(\lambda v u) + \lambda\,\overline{\lambda}\,|v|^2$. Mit $|v| = |\alpha(v)|$ für alle $v \in V$ folgt nun $|\alpha(u+\lambda v)|^2 = |u|^2 + 2\,\mathrm{Re}\big(\lambda\alpha(v)\alpha(u)\big) + \lambda\,\overline{\lambda}\,|v|^2$ und $|u+\lambda v|^2 = |\alpha(u+\lambda v)|^2$. Damit erhält man $\mathrm{Re}(\lambda v u) = \mathrm{Re}\big(\lambda\alpha(v)\alpha(u)\big)$. Für $\lambda = 1$ folgt $\mathrm{Re}(vu) = \mathrm{Re}\big(\alpha(v)\alpha(u)\big)$ und für $\lambda = -i$ folgt $\mathrm{Im}(vu) = \mathrm{Im}\big(\alpha(v)\alpha(u)\big)$. Daher ist $vu = \alpha(v)\alpha(u)$.

1 $\Longrightarrow$ **4** Sei nun U ein Orthonormalsystem und $v, \widetilde{v} \in \alpha(U)$ mit $\alpha(u) = v \neq \widetilde{v} = \alpha(\widetilde{u})$. Dann gilt auch $u \neq \tilde{u}$. Es folgt $v\widetilde{v} = \alpha(u)\alpha(\widetilde{u}) = u\widetilde{u} = 0$ und $|u| = 1 \Longrightarrow |v| = |\alpha(u)| = 1$.

4 $\Longrightarrow$ **2** Für alle $v \in V$ mit $|v| = 1$ ist $\{v\}$ ein Orthonormalsystem. Also ist auch $\{\alpha(v)\}$ ein Orthonormalsystem, d.h. $|\alpha(v)| = 1$.

16.4 Lemma

Sei V euklidisch und $\alpha \in \mathrm{O}(V)$. Dann ist die komplexe Fortsetzung $\widehat{\alpha} \in \mathrm{End}(V_{\mathbb{C}})$ von α unitär.

Beweis

Sei $v + iw \in V_{\mathbb{C}}$ mit $v, w \in V$. Dann folgt

$$\begin{aligned}
\widehat{\alpha}(v+iw)\widehat{\alpha}(v+iw) &= \big(\alpha(v)+i\alpha(w)\big)\big(\alpha(v)+i\alpha(w)\big) \\
&= \alpha(v)\alpha(v) + i\alpha(w)\alpha(v) - i\alpha(v)\alpha(w) + \alpha(w)\alpha(w) \\
&= vv + i(wv) - i(vw) + ww \\
&= (v+iw)(v+iw).
\end{aligned}$$

Daher gilt für alle $v + iw \in V_{\mathbb{C}}$ stets $|\widehat{\alpha}(v+iw)| = |v+iw|$ und nach Satz 16.3 ist $\widehat{\alpha}$ unitär.

16.5 Lemma

Sei V euklidisch oder unitär mit $\dim V < \infty$ und $\alpha \in \mathrm{End}(V)$. Dann ist äquivalent:

1. $\alpha^*\alpha = 1$
2. α ist orthogonal bzw. unitär

Beweis

1 $\Longrightarrow$ **2** Mit $\alpha^*\alpha = 1$ gilt $vw = v\alpha^*\big(\alpha(w)\big) = \alpha(v)\alpha(w)$, d.h. α ist orthogonal bzw. unitär.

2 $\Longrightarrow$ **1** Für alle $v, w \in V$ gilt $vw = \alpha(v)\alpha(w) = v\alpha^*\big(\alpha(w)\big)$, d.h. $0 = vw - v\alpha^*\big(\alpha(w)\big) = v\Big(w - \alpha^*\big(\alpha(w)\big)\Big)$. Also folgt für alle $w \in V$ mit $v = w - \alpha^*\big(\alpha(w)\big)$ insbesondere $|w - \alpha^*(\alpha(w))|^2 = \Big(w - \alpha^*\big(\alpha(w)\big)\Big)\Big(w - \alpha^*\big(\alpha(w)\big)\Big) = 0 \Longrightarrow w = 0 \Longrightarrow w = \alpha^*(\alpha(w))$. Daher ist $\alpha^*\alpha = 1$.

16.6 Definition

Eine Matrix $A \in M_{n\times n}(\mathbb{R})$ bzw. $A \in M_{n\times n}(\mathbb{C})$ heißt orthogonal bzw. unitär, falls $A^*A = 1_n$ gilt. Dabei gilt für $A \in M_{n\times n}(\mathbb{R})$ natürlich $A^* = \overline{A^t} = A^t$.

Bemerkung

Eine Matrix $A \in M_{n\times n}(\mathbb{R})$ bzw. $A \in M_{n\times n}(\mathbb{C})$ ist genau dann orthogonal bzw. unitär, wenn ihre Spalten bzgl. des Standardskalarprodukts ein Orthonormalsystem bilden.

Beweis

Sei $A = (a_{ij})$ orthogonal bzw. unitär und $e_1, \ldots, e_n$ die Standardbasis. Dann gilt $A^*A = 1_n$ und daher $(Ae_i)^t(\overline{Ae_j}) = e_i^t(A^t\overline{A})\overline{e_j} = e_i^t(\overline{A^*A})\overline{e_j} = e_i^t\overline{e_j} = \delta_{ij}$. Somit bilden die Spalten Ae_i von A ein Orthonormalsystem.

Seien nun $a_1, \ldots, a_n$ die Spalten der Matrix $A = (a_{ij})$ mit

$$\sum_{k=1}^{n} a_{ki}\,\overline{a_{kj}} = a_i^t\,\overline{a_j} = \delta_{ij}.$$

Es folgt $A^t\,\overline{A} = (\delta_{ij}) = 1_n$ und daher

$$A^*A = \overline{A^t\,\overline{A}} = \overline{1_n} = 1_n.$$

16.7 Lemma

Sei V ein euklidischer oder unitärer Raum endlicher Dimension. Sei weiter A die darstellende Matrix von $\alpha \in \mathrm{End}(V)$ bzgl. einer Orthonormalbasis. Dann ist α genau dann orthogonal bzw. unitär, wenn A orthogonal bzw. unitär ist.

Beweis

Sei $\dim V = n$ und $\varphi : V \longrightarrow K^n$ der Isomorphismus, der $v \in V$ auf den Koordinatenvektor bzgl. einer fixierten Orthonormalbasis $v_1, \ldots, v_n$ abbildet. Ist dann $\alpha \in \mathrm{End}(V)$ bzgl. dieser Basis durch $A \in M_{n\times n}(K)$ dargestellt, so gilt

$$\varphi\big(\alpha(v)\big) = A\varphi(v)$$

für alle $v \in V$. Da das Skalarprodukt bzgl. einer Orthonormalbasis durch 1_n dargestellt wird, gilt weiter

$$vw = \varphi(v)^t\,\overline{\varphi(w)}$$

für alle $v, w \in V$. Damit folgt nun

$$\alpha(v)\alpha(w) = \varphi\big(\alpha(v)\big)^t\,\overline{\varphi\big(\alpha(w)\big)} = \big(A\varphi(v)\big)^t\,\overline{A\varphi(w)\big)} = \varphi(v)^t(A^t\,\overline{A})\overline{\varphi(w)}$$

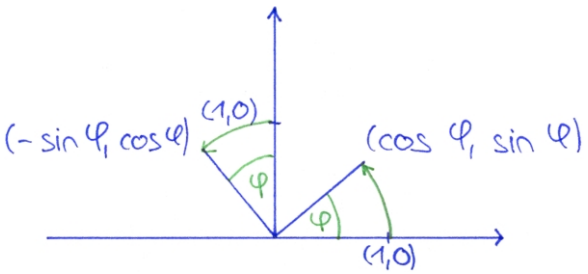
und man erhält α orthogonal bzw. unitär $\iff \varphi(v)^t\,\overline{\varphi(w)} = vw = \varphi(v)^t(A^t\,\overline{A})\overline{\varphi(w)}$. Mit der Standardbasis $e_1, \ldots, e_n$ von K^n gilt nun insbesondere

$$\delta_{ij} = e_i^t\,\overline{e_j} = \varphi(v_i)^t\,\overline{\varphi(v_j)} = \varphi(v_i)^t(A^t\,\overline{A})\overline{\varphi(v_j)} = e_i^t(A^t\,\overline{A})\overline{e_j} = a_{ij}$$

für $A^t\,\overline{A} = (a_{ij})$, d.h. $A^t\,\overline{A} = 1_n$. Andererseits gilt für $A^t\,\overline{A} = 1_n$ offenbar $\varphi(v)^t\,\overline{\varphi(w)} = \varphi(v)^t(A^t\,\overline{A})\overline{\varphi(w)}$. Insgesamt erhält man also α orthogonal bzw. unitär $\iff A^*A = \overline{A^t\,\overline{A}} = \overline{1_n} = 1_n$.

Beispiel

Wir betrachten eine Drehung in der Ebene um den Winkel φ.



Die durch die Bilder $(\cos\varphi, \sin\varphi)$ und $(-\sin\varphi, \cos\varphi)$ der Orthonormalbasisvektoren $(1,0)$ und $(0,1)$ beschriebene lineare Abbildung $\mathbb{R}^2 \longrightarrow \mathbb{R}^2$ besitzt die darstellende Matrix

$$A = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}.$$

Da die Spalten offenbar ein Orthonormalsystem bilden, ist die Matrix – und damit die Drehung – orthogonal.

16.8 Satz

Sei V ein euklidischer Raum endlicher Dimension und $\alpha \in \mathrm{End}(V)$ orthogonal. Dann wird α bzgl. einer geeigneten Orthonormalbasis durch die Matrix

$$\begin{pmatrix}
1 & & & & & & & & & \\
& \ddots & & & & & & & & \\
& & 1 & & & & & & & \\
& & & -1 & & & & & & \\
& & & & \ddots & & & & & \\
& & & & & -1 & & & & \\
& & & & & & \cos\varphi_1 & -\sin\varphi_1 & & \\
& & & & & & \sin\varphi_1 & \cos\varphi_1 & & \\
& & & & & & & & \ddots & \\
& & & & & & & & \cos\varphi_r & -\sin\varphi_r \\
& & & & & & & & \sin\varphi_r & \cos\varphi_r
\end{pmatrix}$$

beschrieben.

Beweis

Nach Lemma 16.5 gilt mit $\alpha \in \mathrm{O}(V)$ zuerst $\alpha^*\alpha = 1$ und damit insbesondere $\alpha^*\alpha = \alpha\alpha^*$. Da also α normal ist, existiert nach Satz 15.14 eine Orthonormalbasis bzgl. der α durch

$$\begin{pmatrix}
\lambda_1 & & & & & & \\
& \ddots & & & & & \\
& & \lambda_m & & & & \\
& & & a_1 & b_1 & & \\
& & & -b_1 & a_1 & & \\
& & & & & \ddots & \\
& & & & & a_r & b_r \\
& & & & & -b_r & a_r
\end{pmatrix}$$

beschrieben wird. Da aber nach Lemma 16.7 zudem $A^tA = 1_n$ gilt, folgt

$$\lambda_i^2 = 1 \Longleftrightarrow \lambda_i = 1 \ \vee \ \lambda_i = -1$$

für $1 \leq i \leq m$ und

$$\begin{pmatrix} a_j & -b_j \\ b_j & a_j \end{pmatrix} \begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix} = \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} \Longleftrightarrow a_j^2 + b_j^2 = 1$$

für $1 \leq j \leq r$. Die verlangte Reihenfolge der λ_i erhält man dabei durch Vertauschung der Basisvektoren. Weiter folgt $|a_j| \leq 1$ und daher $a_j = \cos\varphi_j$ für ein $\varphi_j \in [-\pi, \pi[$. Somit gilt $b_j^2 = 1 - a_j^2 = 1 - \cos^2\varphi_j = \sin^2\varphi_j$, d.h. $b_j = \sin\varphi_j$ oder $b_j = -\sin\varphi_j$. Indem man eventuell φ_j durch $-\varphi_j$ ersetzt, erhält man stets $a_j = \cos\varphi_j$ und $b_j = -\sin\varphi_j$.

16.9 Satz

Sei V ein euklidischer oder unitärer Raum und $\alpha \in \mathrm{End}(V)$ orthogonal bzw. unitär. Ist dann λ ein Eigenwert von α, so gilt $|\lambda| = 1$.

Beweis

Sei $v \in V$ ein Eigenvektor von $\alpha \in \mathrm{End}(V)$ zum Eigenwert λ. Dann gilt $|\lambda|^2\,|v|^2 = \lambda\,\overline{\lambda}(vv) = (\lambda v)(\lambda v) = \alpha(v)\alpha(v) = vv = |v|^2$. Da v ein Eigenvektor ist, gilt $v \neq 0 \Longleftrightarrow |v|^2 \neq 0$ und es folgt $|\lambda|^2 = 1 \Longleftrightarrow |\lambda| = 1$.

16.10 Definition

Zwei reelle bzw. komplexe Matrizen A und B heißen orthogonal bzw. unitär ähnlich, falls eine orthogonale bzw. unitäre Matrix T mit $B = T^*AT$ existiert.

Bemerkung

Seien A und B orthogonal bzw. unitär ähnlich mit $B = T^*AT$. Dann sind A und B wegen $T^*T = 1_n \Longleftrightarrow T^* = T^{-1}$ insbesondere ähnlich.

Bemerkung

Orthogonale bzw. unitäre Ähnlichkeit ist eine Äquivalenzrelation.

Beweis

[Reflexivität] Für jede reelle bzw. komplexe Matrix A gilt $A = 1_n A 1_n = 1_n^* A 1_n$ mit $1_n^* 1_n = 1_n$.

[Symmetrie] Sei $B = T^*AT$ mit $T^*T = 1_n$. Dann gilt auch $TT^* = 1_n$ und weiter $A = T(T^*AT)T^* = TBT^* = (T^*)^*B(T^*)$ mit $(T^*)^*(T^*) = TT^* = 1_n$.

[Transitivität] Sei $B = T^*AT$ und $C = S^*BS$. Es folgt $C = S^*T^*ATS = (TS)^*A(TS)$ mit $(TS)^*(TS) = S^*T^*TS = 1_n$.

16.11 Satz (Hauptachsentransformation)

Jede reelle symmetrische bzw. komplexe hermitesche Matrix ist orthogonal bzw. unitär ähnlich zu einer reellen Diagonalmatrix.

Beweis

Sei $A \in M_{n\times n}(\mathbb{K})$ symmetrisch bzw. hermitesch und $\alpha : \mathbb{K}^n \longrightarrow \mathbb{K}^n$ definiert durch $\alpha(v) = Av$. Dann ist α bzgl. des Standardskalarprodukts wegen $\alpha(v)w = (Av)^t\,\overline{w} = v^t(A^t\,\overline{w}) = v^t(\overline{A^*w}) = v^t(\overline{Aw}) = v\alpha(w)$ selbstadjungiert und daher insbersondere normal.

Somit besitzt $\mathbb{K}^n$ nach Korollar 15.15 bzw. Satz 15.12 eine Orthonormalbasis $v_1, \ldots, v_n$ aus Eigenvektoren von α mit den Eigenwerten $\lambda_1, \ldots, \lambda_n$, d.h. es gilt $v_i^t\,\overline{v_j} = \delta_{ij}$ und $\alpha(v_i) = \lambda_i v_i$. Dabei gilt nach Satz 15.9 für $1 \le i \le n$ stets $\lambda_i \in \mathbb{R}$.

Setze nun $T = (v_1 \cdots v_n) \in M_{n\times n}(\mathbb{R})$. Dann gilt

$$AT = (Av_1 \cdots Av_n) = (\lambda_1 v_1 \cdots \lambda_n v_n)$$

und mit $\overline{v_i}^t\, v_j = \overline{v_i^t\,\overline{v_j}} = \overline{\delta_{ij}} = \delta_{ij}$ folgt

$$T^*AT = \begin{pmatrix} \overline{v_1}^t \\ \vdots \\ \overline{v_1}^t \end{pmatrix} \begin{pmatrix} \lambda_1 v_1 & \cdots & \lambda_n v_n \end{pmatrix} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

Analog erhält man mit $\lambda_i = 1$ für alle $1 \le i \le n$ schließlich $T^*T = 1_n$.

Beispiel

Manche Kegelschnitte in der Ebene sind Gleichungen der Form

$$ax^2 + bxy + cy^2 = d.$$

Dann gilt $d = ax^2 + \frac{1}{2}bxy + \frac{1}{2}bxy + cy^2$ und damit

$$d = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & \frac{1}{2}b \\ \frac{1}{2}b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = v^t S v.$$

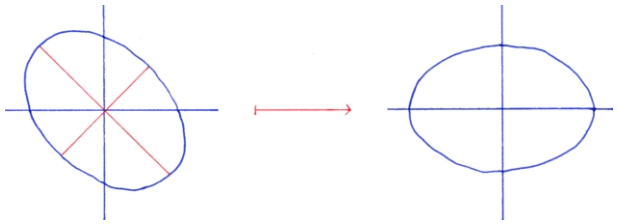
Da S symmetrisch ist, existiert nach Satz 16.11 eine orthogonale Matrix T mit $S = T^t D T$ für

$$D = \begin{pmatrix} d_1 & \\ & d_2 \end{pmatrix}.$$

Es folgt $d = v^t S v = v^t (T^t D T) v = (Tv)^t D (Tv)$. Dabei ist T invertierbar und damit $v \in \mathbb{R}^2 \longmapsto w = Tv$ ein Isomorphismus. In den neuen Koordinaten erhalten wir dann $d = w^t D w$ und mit

$$w = \begin{pmatrix} m \\ n \end{pmatrix}$$

ist $d = d_1 m^2 + d_2 n^2$. Für $d, d_1, d_2 > 0$ oder $d, d_1, d_2 < 0$ ist dies die Gleichung einer Ellipse. Dann beschreibt die durchgeführte Koordinatentransformation eine Kongruenzabbildung,



welche die Symmetrieachsen der Ellipse auf die Hauptachsen des Koordinatensystems abbildet.

16.12 Definition

Eine symmetrische bzw. hermitesche Matrix $S \in M_{n\times n}(\mathbb{K})$ heißt positiv definit, falls für alle $0 \neq v \in \mathbb{K}^n$ stets $v^t S \overline{v} > 0$ gilt.

16.13 Lemma

Eine symmetrische Bilinearform bzw. eine hermitesche Form ist genau dann positiv definit, wenn eine – und damit alle – darstellende Matrix positiv definit ist.

Beweis

Sei V ein reeller bzw. komplexer Vektorraum und die symmetrische Bilinearform bzw. hermitesche Form $\varphi : V \times V \longmapsto \mathbb{K}$ durch $S \in M_{n\times n}(\mathbb{K})$ dargestellt. Da die Koordinatenabbildung $v \in V \longmapsto \alpha \in \mathbb{K}^n$ bijektiv ist, folgt die Behauptung aus $\varphi(v, v) = \alpha^t S \overline{\alpha}$.

16.14 Satz

Eine symmetrische bzw. hermitesche Matrix ist genau dann positiv definit, wenn ihre Eigenwerte positiv sind.

Beweis

Sei $S \in M_{n\times n}(\mathbb{K})$ symmetrisch bzw. hermitesch. Nach Satz 16.11 existiert eine orthogonale Matrix T und eine reelle Diagonalmatrix D mit $S = T^*DT$. Einerseits sind S und D kongruent bzw. komplex kongruent und beschreiben daher die selbe symmetrische Bilinearform bzw. hermitesche Form. Nach Lemma 16.13 ist daher S genau dann positiv definit, wenn D positiv definit ist. Andererseits sind S und D ähnlich und besitzen daher die selben Eigenwerte.

Insgesamt ist die Aussage also nur für

$$D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

zu zeigen. Sei zuerst D positiv definit. Dann gilt mit der Standardbasis $e_1, \ldots, e_n$ insbesondere $\lambda_i = e_i^t D\,\overline{e_i} > 0$. Sei nun $\lambda_i > 0$ für $1 \leq i \leq n$ und

$$0 \neq v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n.$$

Dann gilt $x_j \neq 0 \Longleftrightarrow |x_j|^2 > 0$ für ein $1 \leq j \leq n$ und es folgt

$$v^t D\,\overline{v} = \sum_{i=1}^{n} \lambda_i x_i\,\overline{x_i} = \sum_{i=1}^{n} \lambda_i |x_i|^2 > 0.$$

16.15 Satz (Polarzerlegung von Matrizen)

Jede reelle bzw. komplexe invertierbare Matrix A hat eine eindeutige Darstellung $A = BC$ als Produkt einer symmetrischen bzw. hermiteschen positiv definiten Matrix B und einer orthogonalen bzw. unitären Matrix C.

Beweis

Wir betrachten o.B.d.A. die komplexe Matrix $A \in M_{n\times n}(\mathbb{C})$; der reelle Fall folgt analog. Zunächst untersuchen wir nun die Existenz der Darstellung $A = BC$. Da $AA^* \in M_{n\times n}(\mathbb{C})$

wegen $(AA^*)^* = A^{**}A^* = AA^*$ hermitesch ist, existiert nach Satz 16.11 eine unitäre Matrix T und eine reelle Diagonalmatrix $\widetilde{D}$ mit $AA^* = T\widetilde{D}T^*$.

Mit dem Standardskalarprodukt folgt $v^t AA^* \overline{v} = (A^t v)^t(\overline{A^t v}) = |A^t v|^2 \geq 0$. Da weiter A – und damit auch A^t – invertierbar ist, folgt für $v \neq 0$ auch $A^t v \neq 0 \Longrightarrow |A^t v| > 0 \Longrightarrow v^t AA^* \overline{v} > 0$. Daher ist AA^* positiv definit und nach Satz 16.14 sind die Eigenwerte der hermiteschen Matrix AA^* positiv. Da AA^* und $\widetilde{D}$ insbesondere ähnlich sind, erhält man $\lambda_i > 0$ für jeden Diagonaleintrag λ_i von $\widetilde{D}$.

Sei dann D die Diagonalmatrix mit den positiven Diagonaleinträgen $\sqrt{\lambda_i}$, d.h. $D^2 = \widetilde{D}$ und $\det D > 0$. Mit $TT^* = 1_n$ gilt auch T sowie T^* invertierbar $\Longleftrightarrow \det T \neq 0$ sowie $\det T* \neq 0$ und es folgt $\det TDT^* = \det T \det D \det T^* \neq 0 \Longleftrightarrow TDT^*$ invertierbar. Für $B = TDT^*$ und $C = (TDT^*)^{-1}A$ gilt nun offenbar $A = BC$.

Dabei ist D eine reelle Diagonalmatrix, d.h. man erhält $D^* = D$. Also ist B wegen $B^* = (TDT^*)^* = T^{**}D^*T^* = TDT^* = B$ einerseits hermitesch. Da andererseits B unitär ähnlich zu D ist, sind die Eigenwerte von B die positiven Diagonaleinträge von D. Nach Satz 16.14 ist B somit positiv definit.

Schließlich ist C wegen

$$
\begin{aligned}
CC^* &= \big((TDT^*)^{-1}A\big)\big((TDT^*)^{-1}A\big)^* &&= (TDT^*)^{-1}AA^*\big((TDT^*)^{-1}\big)^* \\
&= (TDT^*)^{-1}(T\widetilde{D}T^*)\big((TDT^*)^{-1}\big)^* &&= (TDT^*)^{-1}(TDDT^*)\big((TDT^*)^{-1}\big)^* \\
&= (TDT^*)^{-1}(TDT^*TDT^*)\big((TDT^*)^{-1}\big)^* &&= (TDT^*)\big((TDT^*)^{-1}\big)^* \\
&= B(B^{-1})^* &&= B^*(B^{-1})^* \\
&= (B^{-1}B)^* &&= 1_n^* \\
&= 1_n
\end{aligned}
$$

unitär. Nun untersuchen wir die Eindeutigkeit der Darstellung $A = BC$. Sei dazu $A = B'C'$ eine weitere Darstellung der verlangten Art. Dann gilt $BBBB^* = BCC^*B^* = BC(BC)^* = AA^* = (B'C')(B'C')^* = B'C'C'^*B'^* = B'B'^* = B'B'$ und daher $B^2 = B'^2$.

Da weiter B hermitesch und positiv definit ist, existiert nach Satz 16.11 und Satz 16.14 eine Basis $v_1, \ldots, v_n$ von $\mathbb{C}^n$ aus Eigenvektoren mit positiven Eigenwerten $\lambda_1, \ldots, \lambda_n$. Dabei gilt $B'^2 v_i = B^2 v_i = B(Bv_i) = B(\lambda_i v_i) = \lambda_i(Bv_i) = \lambda_i^2 v_i$.

Man setze nun $u = B'v_i - \lambda_i v_i$. Es folgt

$$\begin{aligned} B'u &= B'(B'v_i - \lambda_i v_i) &&= B'^2 v_i - B'\lambda_i v_i \\ &= \lambda_i^2 v_i - \lambda_i B'v_i &&= \lambda_i(\lambda_i v_i - B'v_i) \\ &= \lambda_i(-u) &&= (-\lambda_i)u \end{aligned}$$

mit $-\lambda_i < 0$. Da jedoch B' ebenfalls hermitesch und positiv definit ist, sind nach Satz 16.14 alle Eigenwerte positiv. Damit gilt $u = 0 \iff B'v_i = \lambda_i v_i = Bv_i$ und die Bilder der Basisvektoren $v_1, \ldots, v_n$ unter B und B' stimmen überein, d.h. es gilt $Bv = B'v$ für jeden Vektor $v \in \mathbb{C}$ und daher $B' = B = TDT^* \iff B'^{-1} = B^{-1} = (TDT^*)^{-1}$. Zuletzt folgt $C' = B'^{-1}A = B^{-1}A = C$.

17 Faktorräume

17.1 Definition

Sei U ein Unterraum eines Vektorraums V. Auf V erhält man durch $v \sim w \iff v - w \in U$ eine Äquivalenzrelation. Für $v \in V$ bezeichnet $[v]$ die Äquivalenzklasse von v, d.h. $[v] = \{w \in V \mid v - w \in U\} = \{v + u \mid u \in U\} = v + U$. Der Faktorraum (Quotientenraum) $V|_U$ von V nach U ist dann die Menge aller Äquivalenzklassen. Dabei schreibt man statt $[v]$ häufig auch $\overline{v}$.

Bemerkung

Da die Gruppe $(V, +)$ abelsch ist, ist $(U, +)$ ein Normalteiler. Daher ist $V|_U$ nach Satz 2.11 eine Faktorgruppe mit der wohldefinierten Addition $[v]+[w] = [v+w]$ und neutralem Element U.

17.2 Lemma

Durch $\lambda[v] = [\lambda v]$ für alle $\lambda \in K$ und $v \in V$ wird eine wohldefinierte Abbildung $K \times V|_U \longrightarrow V|_U$ erklärt, die $V|_U$ zu einem K-Vektorraum macht.

Beweis

$[K \times V|_U \longrightarrow V|_U$ ist wohldefiniert] Sei $[v] = [\widetilde{v}]$, d.h. $v \sim \widetilde{v} \Longrightarrow v - \widetilde{v} \in U$. Mit $\lambda \in K$ ist dann auch $\lambda v - \lambda\widetilde{v} = \lambda(v - \widetilde{v}) \in U \Longrightarrow [\lambda v] = [\lambda\widetilde{v}]$.

$[V|_U$ ist Vektorraum] $V|_U$ ist nach obiger Bemerkung eine abelsche Gruppe mit neutralem Element $[0] = U$. Die restlichen Vektorraumeigenschaften folgen sofort aus den entsprechenden Eigenschaften von V. Für alle $\lambda, \mu \in K$ und $v, w \in V$ gilt $\lambda(\mu[v]) = \lambda[\mu v] = [\lambda(\mu v)] = [(\lambda\mu)v] = (\lambda\mu)[v]$ und $(\lambda + \mu)[v] = [(\lambda + \mu)v] = [\lambda v + \mu v] = [\lambda v] + [\mu v] = \lambda[v] + \mu[v]$ sowie $\lambda[v + w] = [\lambda(v + w)] = [\lambda v + \lambda w] = [\lambda v] + [\lambda w] = \lambda[v] + \lambda[w]$ und $1 \cdot [v] = [1 \cdot v] = [v]$.

17.3 Satz

Sei U ein Unterraum eines Vektorraums V. Dann ist $\varphi : v \in V \longmapsto [v] \in V|_U$ ein Epimorphismus der Vektorräume V und $V|_U$ mit $\operatorname{Kern}\varphi = U$.

Beweis

Wegen $\varphi(v + w) = [v + w] = [v] + [w] = \varphi(v) + \varphi(w)$ und $\varphi(\lambda v) = [\lambda v] = \lambda[v] = \lambda\varphi(v)$ ist φ ein offenbar surjektiver Homomorphismus von Vektorräumen. Dabei gilt $\operatorname{Kern}\varphi = \{v \in V \mid \varphi(v) = [0]\} = \{v \in V \mid [v] = U\} = \{v \in V \mid v \in U\} = U$.

Korollar

Für $\dim V < \infty$ ist $\dim V|_U = \dim V - \dim U$.

Beweis

Nach dem Dimensionssatz für lineare Abbildungen gilt für den Epimorphismus $\varphi : v \in V \longmapsto [v] \in V|_U$ mit $\dim V < \infty$ die Beziehung $\dim V = \dim \operatorname{Bild}\varphi + \dim \operatorname{Kern}\varphi = \dim V|_U + \dim U$.

17.4 Lemma

Sei $\alpha \in \operatorname{End}(V)$ und U ein α-invarianter Unterraum. Dann wird durch $\overline{\alpha}([v]) = [\alpha(v)]$ ein Endomorphismus $\overline{\alpha} \in \operatorname{End}(V|_U)$ definiert.

Beweis

[$\overline{\alpha}$ ist wohldefiniert] Sei $[v] = [\widetilde{v}]$, d.h. $v - \widetilde{v} \in U$. Da aber U α-invariant ist, gilt dann auch $\alpha(v) - \alpha(\widetilde{v}) = \alpha(v - \widetilde{v}) \in U$ und es folgt $\overline{\alpha}([v]) = [\alpha(v)] = [\alpha(\widetilde{v})] = \overline{\alpha}([\widetilde{v}])$.

[$\bar{\alpha} \in \mathrm{End}(V|_U)$] Sei $\lambda \in K$ und $v, w \in V$. Dann gilt $\overline{\alpha}(\lambda[v]) = \overline{\alpha}([\lambda v]) = [\alpha(\lambda v)] = [\lambda\alpha(v)] = \lambda[\alpha(v)] = \lambda\,\overline{\alpha}([v])$ und $\overline{\alpha}([v]+[w]) = \overline{\alpha}([v+w]) = [\alpha(v+w)] = [\alpha(v)+\alpha(w)] = [\alpha(v)]+[\alpha(w)] = \overline{\alpha}([v]) + \overline{\alpha}([w])$.

17.5 Satz

Sei $\dim V < \infty$ und K algebraisch abgeschlossen. Dann lässt sich jeder Endomorphismus $\alpha \in \mathrm{End}(V)$ durch eine obere Dreiecksmatrix beschreiben.

Beweis

Sei $\dim V = n$ und $\alpha \in \mathrm{End}(V)$. Dabei ist α genau dann durch eine obere Dreiecksmatrix beschreibbar, wenn eine Basis $v_1, \ldots, v_n$ mit $\alpha(v_i) \in \langle v_1, \ldots, v_i\rangle$ existiert. Die Existenz einer solchen Basis zeigen wir durch vollständige Induktion über n. Dabei ist für $n = 1$ die Aussage klar. Nun schließen wir von $n-1$ auf n.

Das charakteristische Polynom von α über dem abgeschlossenen Körper K hat mindestens eine Nullstelle, d.h. α hat einen Eigenwert in K und somit einen Eigenvektor v_1. Man setze $U = \langle v_1\rangle$ und betrachte den Faktorraum $V|_U$ sowie den Endomorphismus $\overline{\alpha} \in \mathrm{End}(V|_U)$ mit $\overline{\alpha}(\overline{v}) = \overline{\alpha(v)}$ und den Epimorphismus $\varphi : v \in V \longmapsto \overline{v} \in V|_U$.

Wegen $\dim V|_U = \dim V - \dim U = n-1$ gibt es nach Induktionsannahme eine Basis $\overline{v}_2, \ldots, \overline{v}_n$ mit $\overline{\alpha}(\overline{v}_i) \in \langle \overline{v}_2, \ldots, \overline{v}_i\rangle$ für $2 \le i \le n$, d.h.

$$\overline{\alpha}(\overline{v}_i) = \sum_{k=2}^{i} \lambda_k \overline{v}_k.$$

Nun wählen wir für $2 \le i \le n$ gewisse Vektoren $v_i \in \overline{v}_i$ und setzen

$$w = \alpha(v_i) - \sum_{k=2}^{i} \lambda_k v_k.$$

Es folgt

$$\varphi(w) = \varphi(\alpha(v_i)) - \sum_{k=2}^{i} \lambda_k \varphi(v_k) = \overline{\alpha(v_i)} - \sum_{k=2}^{i} \lambda_k \overline{v}_k = \overline{\alpha}(\overline{v}_i) - \overline{\alpha}(\overline{v}_i) = 0$$

und daher $w \in \mathrm{Kern}\,\varphi = U = \langle v_1\rangle \Longrightarrow w = \lambda_1 v_1$. Also gilt $\alpha(v_i) = w + \sum\limits_{k=2}^{i} \lambda_k v_k = \sum\limits_{k=1}^{i} \lambda_k v_k \in \langle v_1, \ldots, v_i\rangle$ für $2 \le i \le n$. Da aber v_1 ein Eigenvektor ist, gilt $\alpha(v_i) \in \langle v_1, \ldots, v_i\rangle$ sogar für $1 \le i \le n$.

Es bleibt zu zeigen, dass $v_1, \ldots, v_n$ eine Basis ist. Wegen $\dim V = n$ ist dabei nur die lineare Unabhängigkeit zu prüfen. Sei dazu

$$\sum_{k=1}^{n} \mu_k v_k = 0 \Longrightarrow \sum_{k=1}^{n} \mu_k \overline{v}_k = \sum_{k=1}^{n} \mu_k \varphi(v_k) = \varphi(\sum_{k=1}^{n} \mu_k v_k) = \varphi(0) = 0.$$

Mit $\overline{v}_1 = U = 0$ und da $\overline{v}_2, \ldots, \overline{v}_n$ nach Wahl linear unabhängig sind, folgt

$$0 = \sum_{k=1}^{n} \mu_k \overline{v}_k = \sum_{k=2}^{n} \mu_k \overline{v}_k \Longrightarrow \forall_{2 \leq k \leq n}\ \mu_k = 0.$$

Man erhält $\mu_1 v_1 = 0$ und da v_1 ein Eigenvektor ist, folgt $\mu_1 = 0$. Damit sind insgesamt $v_1, \ldots, v_n$ linear unabhängig.

18 Theorie der linearen Blockcodes

Bemerkung

Wir behandeln die Übertragung von Wörtern einer festen Länge n aus einem Alphabet F, wobei manche Buchstaben durch Übertragungsfehler zufällig geändert werden. Durch den Einbau von Redundanz lässt sich eine begrenzte Anzahl von Fehlern entdecken und evtl. korrigieren.

Beispiel

Bei der einfachen Übertragung der Wörter 00, 01, 10, 11 erkennt man keine Übertragungsfehler.

1. An die 2-buchstabigen Wörter lässt sich ihre Quersumme als Prüfbit anhängen, d.h.

$$\begin{aligned} 00 &\longrightarrow 000 \\ 01 &\longrightarrow 011 \\ 10 &\longrightarrow 101 \\ 11 &\longrightarrow 110 \end{aligned}$$

   Da sich die übertragenen Wörter nun an mindesten 2 Stellen unterscheiden, kann man einzelne Übetragungsfehler erkennen. Diese sind aber nicht korrigierbar, etwa kann man 111 nicht eindeutig zuordnen.

2. Durch die Codierung

$$\begin{aligned} 00 &\longrightarrow 000\ 000 \\ 01 &\longrightarrow 000\ 111 \\ 10 &\longrightarrow 111\ 000 \\ 11 &\longrightarrow 111\ 111 \end{aligned}$$

unterscheiden sich die Wörter nun an mindestens 3 Stellen, d.h. einzelne Übertragungsfehler sind erkennbar und korrigierbar. Dazu ordnet man eindeutig einem an einer Stelle fehlerhaft übertragenen Wort dasjenige zu, von dem es sich an nur einer Stelle unterscheidet. Durch die längere Übertragung steigt allerdings auch die Fehlerwahrscheinlichkeit.

3. Die kürzere Codierung

$$\begin{aligned} 00 &\longrightarrow 00000 \\ 01 &\longrightarrow 00111 \\ 10 &\longrightarrow 11100 \\ 11 &\longrightarrow 11011 \end{aligned}$$

liefert diesselben Korrekturmöglichkeiten wie unter 2.

Bemerkung

Bei einer Codierung der Wörter 00, 01, 10 und 11 durch Code-Wörter der Länge 4 ist eine Fehlerkorrektur nicht mehr möglich.

Beweis

Wir nehmen an, es gebe 4 Codewörter der Länge 4, die sich an mindestens 3 Stellen unterscheiden. Zu jedem dieser Codewörter betrachten wir die Menge der 4 Wörter, die sich von dem Codewort an nur einer Stelle unterscheiden. Unterscheidet sich ein Wort von zwei Codewörtern an nur einer Stelle, so unterscheiden sich diese zwei Codewörter an höchstens 2 Stellen – ein Widerspruch. Daher sind die Mengen disjunkt und man erhält also 20 verschiedene Wörter. Mit dem Alphabet $\{0, 1\}$ lassen sich aber höchstens 16 verschiedene Wörter der Länge 4 bilden.

Beispiel

Die ISBN-Nummern der Bücher verwenden das Alphabet $F = \{0, \ldots, 9, X\} = \mathbb{Z}|_{11\mathbb{Z}}$ mit $X = 10$. Eine typische ISBN-Nummer ist

$$\underbrace{3}_{\text{Land}} - \underbrace{11}_{\text{Verlag}} - \underbrace{015\ \ 873}_{\text{Buchnummer}} - \underbrace{6}_{\text{Prüfziffer}}$$

Dabei ist

$$C = \{a_1 \cdots a_{10} \mid \sum_{i=1}^{10} i a_i = 0 \text{ und } a_i \neq X \text{ für } i \leq 9\}$$

die Menge aller zulässigen Codewörter. Wegen $1 + 10 = 0$ erhält man die Umformulierung

$$\sum_{i=1}^{10} i\, a_i = 0 \Longleftrightarrow a_{10} = \sum_{i=1}^{9} i\, a_i.$$

1. Ein einzelner Fehler wird erkannt: Es trete an der i-ten Stelle ein Fehler auf, d.h. statt a_i erhalte man $\widetilde{a}_i$. Es folgt

$$\begin{aligned}
& a_1 \cdots a_i \cdots a_{10},\ a_1 \cdots \widetilde{a}_i \cdots a_{10} \in C \\
\Longrightarrow\ & 1 \cdot a_1 + \ldots + i \cdot a_i + \ldots + 10 \cdot a_{10} = 0 = 1 \cdot a_1 + \ldots + i \cdot \widetilde{a}_i + \ldots + 10 \cdot a_{10} \\
\Longrightarrow\ & i \cdot a_i = i \cdot \widetilde{a}_i \\
\Longrightarrow\ & i \cdot (a_i - \widetilde{a}_i) = 0 \\
\Longrightarrow\ & a_i = \widetilde{a}_i.
\end{aligned}$$

2. Man kann auch erkennen, ob zwei Ziffern ihre Position vertauscht haben: Sei nach einem Übertragungsfehler a_i an der Stelle j und a_j an der Stelle a_i. Analog zu 1 gilt

$$\begin{aligned}
& a_1 \cdots a_i \cdots a_j \cdots a_{10},\ a_j \cdots a_i \cdots a_i \cdots a_{10} \in C \\
\Longrightarrow\ & i \cdot a_i + j \cdot a_j = i \cdot a_j + j \cdot a_i \\
\Longrightarrow\ & (i - j) \cdot (a_i - a_j) = 0 \\
\Longrightarrow\ & a_i = a_j.
\end{aligned}$$

Bemerkung

Für jede Primpotenz q gibt es bis auf Isomorphie genau einen Körper $\mathbb{F}_q$ mit genau q Elementen. Für $q = p \in \mathbb{P} = \{p \in \mathbb{N} \mid p \text{ prim}\}$ ist dies der Körper $\mathbb{F}_p = \mathbb{Z}|_{p\mathbb{Z}}$. In der Theorie der linearen Blockcodes betrachtet man dann Unterräume des Vektorraums $\mathbb{F}_q^n$ und interpretiert die Vektoren als n-stellige Wörter mit Alphabet $\mathbb{F}_q$. Im Folgenden wird meistens $q = 2$ sein.

18.1 Definition

Sei K ein Körper und $0 \neq n \in \mathbb{N}$. Auf dem Vektorraum K^n definiert man für

$$u = \begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix} \text{ und } v = \begin{pmatrix} b_1 & \cdots & b_n \end{pmatrix}$$

eine Abstandsfunktion $d(\cdot, \cdot)$ durch $d(u, v) = |\{i \mid a_i \neq b_i\}|$, d.h. durch die Anzahl der Positionen, in denen sich u und v unterscheiden. Dabei wird d auch Henning-Abstand genannt.

18.2 Satz

$d(\cdot,\cdot)$ ist eine translations-invariante Metrik, d.h. es gilt

1. $d(u,v) \geq 0$ und $d(u,v) = 0 \Longleftrightarrow u = v$
2. $d(u,v) = d(v,u)$
3. $d(u,v) \leq d(u,w) + d(w,v)$
4. $d(u+w, v+w) = d(u,v)$

Beweis

Offenbar gelten 1 und 2. Unterscheidet sich nun u von w in i und v von w in j Positionen, so können sich u und v in höchstens $i+j$ Positionen unterscheiden und es folgt 3.

Sei schließlich

$$u = \begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix},\ v = \begin{pmatrix} b_1 & \cdots & b_n \end{pmatrix} \text{ und } v = \begin{pmatrix} c_1 & \cdots & c_n \end{pmatrix}.$$

Da sich $a_i + c_i$ und $b_i + c_i$ genau dann unterscheiden, wenn sich a_i und b_i unterscheiden, gilt 4.

18.3 Definition

Für $u \in K^n$ heißt $d(u,0)$ das Gewicht von u. Man schreibt auch $d(u,0) = w(u)$.

18.4 Definition

Ein Code C ist ein Unterraum des Vektorraums $\mathbb{F}_q^n$. Das Minimalgewicht von C ist $d = d(C) = \min\{w(u) \mid 0 \neq u \in C\}$. Für $\dim C = k$ nennt man C einen (n,k,d)-Code.

Bemerkung

Aus Satz 18.2 folgt sofort $d(u,v) = d(u-v, v-v) = d(u-v, 0) = w(u-v)$. Daher ist $d(C)$ ist der kleinste Abstand, den zwei verschiedene Codewörter haben können.

Bemerkung

Sei $d = 2e+1$ für $e \in \mathbb{N}$ und C ein (n,k,d)-Code. Dann kann C bis zu e Fehlern erkennen und korrigieren.

Beweis

Sei $w \in \mathbb{F}_q^n$ ein empfangenes Wort, das sich an höchstens e Stellen von einem Codewort $u \in C$ unterscheidet, d.h. $d(u,w) \leq e$. Dann gilt für alle anderen Codewörter $v \in C$ stets $2e+1 = d \leq d(u,v) \leq d(u,w) + d(w,v) \leq e + d(w,v) \Longrightarrow d(w,v) \geq e+1$, d.h. u ist eindeutig bestimmt.

Bemerkung

Man sucht (n,k,d)-Codes mit fester Wortlänge n und einem maximalen Wert für k – d.h. einem großen Wortschatz – und d – d.h einer guten Fehlerkorrektur.

18.5 Satz (Singleton-Schranke)

Sei C ein (n,k,d)-Code. Dann gilt $k + d \leq n + 1$.

Beweis

Man betrachte die Abbildung $\varphi : (a_1, \ldots, a_n) \in C \longmapsto (a_1, \ldots, a_{n-d+1}) \in \mathbb{F}_q^{n-d+1}$, die die letzten $d-1$ Stellen der Codewörter abschneidet. Dann ist φ offenbar linear. Wegen $w(u) \geq d$ für alle $0 \neq u \in C$ unterscheidet sich u an mindestens d Stellen von 0. Damit hat aber auch $\varphi(u)$ an mindestens einer Stelle einen von 0 verschiedenen Eintrag, d.h. $\varphi(u) \neq 0$ bzw. $\operatorname{Kern} \varphi = \{0\}$. Da also φ injektiv ist, gilt $k = \dim C = \dim \varphi(C) \leq \dim \mathbb{F}_q^{n-d+1} = n - d + 1$ und die Behauptung folgt.

Bemerkung

Sei U ein Unterraum von $\mathbb{F}_q^n$ mit $\dim U = k$. Dann gilt $|U| = q^k$.

Beweis

Die Anzahl der Elemente des Unterraums U ist die Anzahl der verschiedenen Linearkombinationen der $k = \dim U$ Basisvektoren, d.h. die Anzahl der verschiedenen k-Tupel aus den q Körperelementen – also $|U| = q^k$.

18.6 Satz (Henning-Schranke)

Sei C ein (n, k, d)-Code und $e \in \mathbb{N}$ mit $2e + 1 \leq d$. Dann gilt

$$\sum_{j=0}^{e} \binom{n}{j} \cdot (q-1)^j \leq q^{n-k}.$$

Beweis

Sei $B(c)$ die Kugel in $\mathbb{F}_q^n$ um $c \in C$ mit Henningradius e, d.h. $B(c) = \{u \in \mathbb{F}_q^n \mid d(u, c) \leq e\}$. Dann sind je zwei Kugeln $B(c)$ und $B(c')$ disjunkt, denn sei $v \in B(c) \cap B(c') \Longrightarrow d(v, c) \leq e \wedge$ $\wedge\, d(v', c) \leq e \Longrightarrow 2e + 1 \leq d \leq d(v, v') \leq d(v, c) + d(c, v') \leq 2e$ – ein Widerspruch. Daher gilt

$$\sum_{c \in C} |B(c)| \leq \left|\mathbb{F}_q^n\right| = q^n.$$

Weiter ist $u \in B(c) \longmapsto u - c \in B(0)$ bijektiv, d.h. es gilt $|B(c)| = |B(0)|$. Mit $|C| = q^k$ folgt dann

$$q^k \cdot |B(0)| = \sum_{c \in C} |B(c)| \leq q^n.$$

Setze $S(j) = \{u \in \mathbb{F}_q^n \mid w(u) = j\}$ für $0 \leq j \leq e$. Dann ist $B(0) = S(0) \,\dot\cup\, \ldots \,\dot\cup\, S(e)$ eine disjunkte Vereiningung der $S(j)$, d.h. $|B(0)| = \sum\limits_{j=1}^{e} |S(j)|$. Für $v \in S(j)$ werden j beliebige Positionen mit $q-1$ von 0 verschiedenen Körperelementen besetzt, d.h. $|S(j)| = \binom{n}{j} \cdot (q-1)^j$. Somit gilt

$$q^k \cdot \sum_{j=1}^{e} \binom{n}{j} \cdot (q-1)^j = q^k \cdot |B(0)| \leq q^n$$

und es folgt die Behauptung.

Bemerkung

Die Codes, für welche die Henning-Schranke scharf ist (d.h. für die $\sum\limits_{j=0}^{e} \binom{n}{j} \cdot (q-1)^j = q^{n-k}$ gilt), heißen perfekte Codes. Diese perfekten Codes sind bekannt und existieren nur für kleine Werte von d.

18.7 Lemma

$\mathbb{F}_q^m$ besitzt genau $\frac{q^m - 1}{q-1}$ eindimensionale Unterräume.

Beweis

Sei x die Anzahl der eindimensionalen Unterräume $\langle v_i \rangle$ von $\mathbb{F}_q^m$. Wegen $\dim\langle v_i \rangle = 1$ ist $v_i \neq 0$, d.h. $\langle v_i \rangle = \{\lambda v_i \mid \lambda \in \mathbb{F}_q\}$ besitzt $q-1$ von 0 verschiedene Elemente. Andererseits gilt $0 \neq v \in \langle v_i \rangle \cap \langle v_k \rangle \Longrightarrow \lambda v_i = v = \mu v_k$ mit $\mu, \lambda \neq 0 \Longrightarrow v_i = (\lambda^{-1}\mu) v_k \Longrightarrow \langle v_i \rangle = \langle v_k \rangle$, d.h. jeder Vektor $v \neq 0$ liegt in genau einem Unterraum $\langle v_i \rangle$. In $\mathbb{F}_q^m$ liegen aber $q^m - 1$ von 0 verschiedene Vektoren, d.h. $q^m - 1 = x \cdot (q-1)$.

18.8 Satz (Henning-Code)

Sei $0 < m \in \mathbb{N}$ und $n = \frac{q^m-1}{q-1}$. Wähle dann aus jedem der n eindimensionalen Unterräume von $\mathbb{F}_q^m$ ein $v_i \neq 0$ und setze

$$C = \{\begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix} \in \mathbb{F}_q^n \mid \sum_{i=1}^{n} a_i v_i = 0\}.$$

Dann ist C ein $(n, n-m, 3)$-Code.

Beweis

Wir betrachten die lineare Abbildung

$$\varphi : \begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix} \in \mathbb{F}_q^n \longmapsto \sum_{i=1}^{n} a_i \, v_i \in \mathbb{F}_q^m.$$

Für $v \in \mathbb{F}_q^m$ existiert ein $\lambda \in \mathbb{F}_q$ und ein $1 \leq i \leq n$ mit $\lambda v_i = v$, d.h. $\varphi \begin{pmatrix} 0 & \cdots & \lambda & \cdots & 0 \end{pmatrix} = v$. Daher ist φ surjektiv. Weiter ist $C = \operatorname{Kern} \varphi$, d.h. C ist ein Unterraum von $\mathbb{F}_q^n$ und es gilt $n = \dim \mathbb{F}_q^n = \dim \operatorname{Bild} \varphi + \dim \operatorname{Kern} \varphi = \dim \mathbb{F}_q^m + \dim C = m + k \Longrightarrow k = n - m$.

Da je zwei Vektoren der v_i verschiedene Unterräume von $\mathbb{F}_q^m$ aufspannen, sind sie linear unabhängig und damit nur trivial zu 0 kombinierbar. Also ist $w(c) \geq 3$ für alle $0 \neq c \in C$. Es gilt aber $v_1 + v_2 \notin \langle v_1 \rangle$ und $v_1 + v_2 \notin \langle v_2 \rangle$, d.h. $v_1 + v_2 \in \langle v_i \rangle \Longrightarrow v_1 + v_2 = \lambda \, v_i$ für $3 \leq i \leq n$. Also folgt $v_1 + v_2 - \lambda \, v_i = 0$ und damit

$$c = \begin{pmatrix} 1 & 1 & 0 & \cdots & -\lambda & \cdots & 0 \end{pmatrix} \in C$$

mit $w(c) = 3$. Daher ist $d = 3$.

Bemerkung

1. Da Henning-Codes $(n, n-m, 3)$-Codes sind, können sie einen Fehler korrigieren.

2. Henning-Codes sind perfekte Codes, denn wegen $d = 3$ und $e \in \mathbb{N}$ gilt für die Henning-Schranke stets $e = 1$ und es folgt $q^m = 1 + \frac{q^m-1}{q-1} \cdot (q-1) = 1 + n \cdot (q-1) = \binom{n}{0} \cdot (q-1)^0 + \binom{n}{1} \cdot (q-1)^1 = \sum_{j=0}^{1} \binom{n}{j} \cdot (q-1)^j \leq q^{n-k} = q^m$ – d.h. $\sum_{j=0}^{1} \binom{n}{j} \cdot (q-1)^j = q^{n-k}$.

Beispiel

1. Sei $q = 2$ und $m = 2$. Dann ist $n = 3$ und $k = 1$. Die Vektoren

$$v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix},\ v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ und } v_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

erzeugen die drei eindimensionalen Unterräume, d.h. $C = \{\begin{pmatrix} 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}\}$.

2. Sei $q = 3$ und $m = 2$, d.h. $n = 4$. Dann ist

$$v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix},\ v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix},\ v_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ sowie } v_4 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

und es folgt $C = \{\begin{pmatrix} 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 & 2 & 0 \end{pmatrix}, \begin{pmatrix} 2 & 2 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 2 & 1 & 0 & 2 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 2 & 2 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 2 & 0 & 2 & 1 \end{pmatrix}\}$.

18.9 Satz (Simplex-Code)

Sei $0 \neq m \in \mathbb{N}$ und $n = \frac{q^m-1}{q-1}$. Wähle dann aus jedem der n eindimensionalen Unterräume des Spaltenraums $\mathbb{F}_q^m$ ein $v_i \neq 0$. Dann ist

$$C = \{\begin{pmatrix} z^t v_1 & \cdots & z^t v_n \end{pmatrix} \mid z \in \mathbb{F}_q^m\}$$

ein (n, m, q^{m-1})-Code, in dem je zwei verschiedene Wörter genau den Abstand q^{m-1} haben.

Beweis

Wir betrachten die lineare Abbildung

$$\varphi : z \in \mathbb{F}_q^m \longmapsto \begin{pmatrix} z^t v_1 & \cdots & z^t v_n \end{pmatrix} \in \mathbb{F}_q^n$$

mit dem Unterraum $\operatorname{Bild} \varphi = C$ von $\mathbb{F}_q^n$ und untersuchen wir $\operatorname{Kern} \varphi$. Sei also

$$\begin{pmatrix} z^t v_1 & \cdots & z^t v_n \end{pmatrix} = 0,$$

d.h. $z^t v_i = 0$ für $1 \leq i \leq n$. Da jeder Vektor $v \in \mathbb{F}_q^m$ in einem Unterraum $\langle v_i \rangle$ liegt, folgt $z^t v = z^t(\lambda_i v_i) = \lambda_i (z^t v_i) = 0$. Wählt man für v die Standardbasisvektoren, so erhält man schließlich $z = 0 \Longrightarrow \operatorname{Kern} \varphi = \{0\}$. Damit ist $\varphi : \mathbb{F}_q^m \longrightarrow C$ bijektiv und es gilt $\dim C = \dim \mathbb{F}_q^m = m$.

Haben je zwei Wörter $c, \widetilde{c} \in C$ den Abstand q^{m-1}, so folgt unmittelbar $d(C) = q^{m-1}$. Weiter ist $d(c, \widetilde{c}) = q^{m-1}$ für alle $c, \widetilde{c} \in C$ mit $c \neq \widetilde{c}$ äquivalent zu $w(c) = q^{m-1}$ für alle $0 \neq c \in C$. Im Folgenden zeigen wir dann $w(c) = q^{m-1}$ mit $c \neq 0$ für $q = 2$. Sei also $0 \neq c = \begin{pmatrix} z^t v_1 & \cdots & z^t v_n \end{pmatrix} \in C$, d.h. $z \neq 0$. Dann ist $w(c) = |\{i \mid z^t v_i \neq 0\}| = n - |\{i \mid z^t v_i = 0\}|$.

Nun gilt in $\mathbb{F}_2^m$ stets $\langle v_i \rangle = \{0, v_i\}$, d.h. $\mathbb{F}_2^m = \{0, v_1, \ldots, v_n\}$. Daher ist $|\{i \mid z^t v_i = 0\}| = |U \setminus \{0\}|$ für $U = \{v \in \mathbb{F}_2^m \mid z^t v = 0\}$. Dabei ist aber U der Kern der linearen und surjektiven Abbildung $v \in \mathbb{F}_2^m \longmapsto z^t v \in \mathbb{F}_2$, d.h. es gilt $\dim U = \dim \mathbb{F}_2^m - \dim \mathbb{F}_2 = m - 1$. Daher folgt $|\{i \mid z^t v_i = 0| = |U \setminus \{0\}| = |U| - 1 = q^{\dim U} - 1 = 2^{m-1} - 1$ und mit $n = \frac{q^m - 1}{q-1}$ erhält man für $q = 2$ schließlich $w(c) = n - (2^{m-1} - 1) = (2^m - 1) - 2^{m-1} + 1 = 2^{m-1}$.

18.10 Satz (allgemeiner Reed-Solomon-Code)

Sei $d, n \in \mathbb{N}$ und q eine Primpotenz mit $2 \leq d \leq n \leq q$. Seien weiter $a_1, \ldots, a_n \in \mathbb{F}_q$ paarweise verschieden. Man setze

$$v_i = \begin{pmatrix} 1 & a_i & a_i^2 & \cdots & a_i^{d-2} \end{pmatrix} \in \mathbb{F}_q^{d-1}.$$

Dann ist

$$C = \{\begin{pmatrix} \lambda_1 & \ldots & \lambda_n \end{pmatrix} \in \mathbb{F}_q^n \mid \sum_{i=1}^{n} \lambda_i v_i = 0\}$$

ein $(n, n-d+1, d)$-Code.

Beweis

Zunächst betrachten wir die Matrix

$$A = \begin{pmatrix} v_1 \\ \vdots \\ v_{d-1} \end{pmatrix} = \begin{pmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{d-2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_{d-1} & a_{d-1}^2 & \cdots & a_{d-1}^{d-2} \end{pmatrix} \in M_{d-1 \times d-1}(\mathbb{F}_q).$$

Dann ist $\det A$ eine Vandermondesche Determinante, d.h. $\det A = \prod_{i<j} (a_j - a_i)$. Da die a_i paarweise verschieden sind, folgt $\det A \neq 0 \Longrightarrow A$ ist invertierbar $\Longrightarrow v_1, \ldots, v_{d-1}$ sind linear unabhängig. Die Reihenfolge der a_i ist jedoch beliebig gewählt, d.h. jeweils $d-1$ verschiedene Vektoren v_i sind linear unabhängig.

Die lineare Abbildung

$$\varphi : \begin{pmatrix} \lambda_1 & \ldots & \lambda_n \end{pmatrix} \in \mathbb{F}_q^n \longmapsto \sum_{i=1}^{n} \lambda_i v_i \in \mathbb{F}_q^{d-1}$$

definiert nun den Unterraum $C = \operatorname{Kern} \varphi$ von $\mathbb{F}_q^n$. Da etwa $v_1, \ldots, v_{d-1}$ linear unabhängig sind und deshalb eine Basis von $\mathbb{F}_q^{d-1}$ bilden, gilt $\langle v_1, \ldots, v_n \rangle = \langle v_1, \ldots, v_{d-1} \rangle = \mathbb{F}_q^{d-1}$. Da $\operatorname{Bild} \varphi$

die Menge aller möglichen Linearkombinationen der v_i ist, erhält man somit $\operatorname{Bild}\varphi = \mathbb{F}_q^{d-1}$ und weiter $\dim C = \dim \mathbb{F}_q^n - \dim \mathbb{F}_q^{d-1} = n - d + 1$.

Andererseits folgt aus der linearen Unabhängkeit von jeweils $d-1$ Vektoren v_i, dass bei einer nicht-trivialen Linearkombination $\sum \lambda_i v_i = 0$ mindestens d Skalare λ_i von 0 verschieden sind. Weiter sind stets d Vektoren aus $\mathbb{F}_q^{d-1}$ linear abhängig, d.h. es existieren nicht-triviale Linearkombinationen $\sum \lambda_i v_i = 0$ mit genau d von 0 verschiedenen Skalaren λ_i. Insgesamt erhält man also $d(C) = d$.

Bemerkung

1. Für Reed-Solomon-Codes ist die Singleton-Schranke scharf.

2. Reed-Solomon-Codes werden für die Codierung von CDs verwendet.