Lineare Algebra II

Prof. Dr. Dirk Ferus

Wintersemester 2001/2

22. Oktober 2004

Inhaltsverzeichnis

Literatur

- Lorenz, F.: Lineare Algebra I,II, BI Hochschultaschenbücher
- Hoffman, K.; Kunze,R.: Linear Algebra, Prentice Hall
- Jänich, K.: Lineare Algebra, Springer
- Köcher, M.: Lineare Algebra und analytische Geometrie, Springer
  (Enthält viele interessante historische Fakten)
- Klingenberg, W.: Lineare Algebra und Geometrie, Springer Hochschultext
  (Insbesondere für die Geometrie)
- Walter, R.: Lineare Algebra und analytische Geometrie, Vieweg 1985
  (Insbesondere für die Quadriken)

0 Vorbemerkungen

Wir wollen im ersten Teil dieses Semesters die Struktur von Endomorphismen genauer untersuchen: Wie „funktioniert“ ein Endomorphismus? Wir werden sehen, daß das auf folgende Fragen hinausläuft.

- Lassen sich Endomorphismen „klassifizieren“, d.h. in Klassen einteilen, deren Elemente jeweils wesentliche Eigenschaften gemeinsam haben?
- Lassen sich die Klassen durch einfache „Invarianten“ charakterisieren?
- Gibt es in den Klassen jeweils einen ausgezeichneten Vertreter, eine *Normalform* der Klassenkameraden.
- Wie sehen entsprechende Resultate für Matrizen aus?

Um diese Fragen besser zu verstehen, betrachten wir zwei einfachere Beispiele. Das zweite hat mit Endomorphismen nichts, das erste gar nichts zu tun.

Beispiel 1. Sei K ein Körper. Wir betrachten nicht Endmorphismen, sondern die endlich-dimensionalen K-Vektorräume. Wir nennen zwei solche *äquivalent*, wenn sie isomorph sind. Damit haben wir eine Klasseneinteilung in *Isomorphieklassen* vorgenommen und gewissermaßen definiert, worauf es uns *nicht* ankommt. Die Elemente einer Klasse – isomorphe Vektorräume – sind von unserem damit eingenommenen Standpunkt aus „gleich“. Uns interssiert nicht, ob so ein Vektorraum aus Pfeilklassen besteht oder aus Funktionen oder aus n-tupeln. Die Elemente in einer Klasse sind charakterisiert durch eine einzige Invariante, nämlich durch eine natürliche Zahl: die Dimension. Zwei endlich-dimensionale K-Vektorräume sind genau dann äquivalent, wenn sie dieselbe Dimension haben. Und in der Klasse der n-dimensionalen K-Vektorräume gibt es einen Standard-Vertreter, den K^n.

□

Beispiel 2. Seien V, W zwei endlich-dimensionale K-Vektorräume. Wir betrachten die Menge $\mathrm{Hom}(V, W)$ der linearen Abbildungen von V nach W. Wir nennen zwei solche Abbildungen äquivalent, wenn sie sich nur durch Automorphismen von V und W unterscheiden:

$$f \sim \tilde{f} \iff \exists_{\psi \in \mathrm{Aut}(V),\, \phi \in \mathrm{Aut}(W)}\, f = \phi \circ \tilde{f} \circ \psi.$$

Offenbar haben dann f und $\tilde{f}$ denselben Rang.

Wählt man zu f eine Basis $\mathbf{v} = (v_1, \ldots, v_n)$ von V, so daß $(v_{r+1}, \ldots, v_n)$ eine Basis von $\mathrm{Kern}\, f$ ist, so ist das r-tupel $(f(v_1) =: w_1, \ldots, f(v_r) =: w_r)$ linear unabhängig. Ergänzen wir es zu einer Basis $\mathbf{w} = (w_1, \ldots, w_m)$ von W so „funktioniert“ f also so: Es bildet die ersten r Vektoren der V-Basis auf die ersten r Vektoren der W-Basis ab, alle weiteren auf 0. Die Darstellungsmatrix ist

$$A^{\mathbf{v}}_{\mathbf{w}}(f) = \begin{pmatrix} 1 & & 0 & 0 & \ldots & \ldots & 0 \\ & \ddots & & \vdots & & & \vdots \\ 0 & & 1 & 0 & \ldots & \ldots & 0 \\ 0 & \ldots & 0 & 0 & \ldots & \ldots & 0 \\ \vdots & & \vdots & \vdots & & & \vdots \\ 0 & \ldots & 0 & 0 & \ldots & \ldots & 0 \end{pmatrix}.$$

Die Zahl der Einsen dabei ist gerade der Rang von f.

Hat nun $\tilde{f}$ denselben Rang wie f, so wählen wir entsprechende Basen $(\tilde{v}_1, \ldots, \tilde{v}_n)$ und $(\tilde{w}_1, \ldots, \tilde{w}_m)$ und definieren Automorphismen $\psi \in \mathrm{Aut}(V)$ durch

$$\psi(\tilde{v}_j) = v_j$$

und $\phi \in \mathrm{Aut}(W)$ durch

$$\phi(w_j) = \tilde{w}_j.$$

Dann ist

$$\phi \circ f \circ \psi(\tilde{v}_j) = \phi(f(v_j)) = \begin{cases} \phi(w_j) \\ \phi(0) \end{cases} = \begin{cases} \tilde{w}_j \\ 0 \end{cases} = \tilde{f}(\tilde{v}_j).$$

Also sind f und $\tilde{f}$ äquivalent. Die linearen Abbildungen von V nach W lassen sich also bis auf Automorphismen von Urbild- und Zielraum klassifizieren. Die Klassen sind durch die Rang-Invariante charakterisiert. Es gibt zwar keinen Standard-Vertreter in einer solchen Klasse, aber eine Standard-Matrixdarstellung.

Den hier geschilderten Sachverhalt kann man für $V = K^n, W = K^m$ und $M(m, n; K)$ statt $\mathrm{Hom}(V, W)$ so aussprechen: Zwei Matrizen $A, \tilde{A} \in M(m, n; K)$ heißen *äquivalent*, wenn es invertierbare Matrizen $S \in M(m, m; K)$ und $T \in M(n, n; K)$ gibt, so daß

$$\tilde{A} = SAT$$

gilt. Zwei Matrizen in $M(m, n; K)$ sind genau dann äquivalent, wenn sie denselben Rang haben. In jeder Äquivalenzklasse gibt es einen ausgezeichneten Vertreter, nämlich die Matrix mit r Einsen auf der „Diagonalen“ und lauter Nullen sonst. Jede Matrix A vom Rang r kann man auf diese Normalform bringen, indem man sie von links und rechts mit invertierbaren Matrizen multipliziert.

Eine dritte Variante der Geschichte sieht so aus: Setzt man

$$V_0 := \mathrm{Kern}\, f = \mathrm{Spann}(v_{r+1}, \ldots, v_n), \quad V_1 := \mathrm{Spann}(v_1, \ldots, v_r),$$
$$W_1 := \mathrm{Spann}(w_1, \ldots, w_r), \quad W_0 = \mathrm{Spann}(w_{r+1}, \ldots, w_m),$$

so hat man

$$V = V_0 \oplus V_1, \quad W = W_0 \oplus W_1,$$
$$f|V_0 : V_0 \to W_0 \text{ ist die Null-Abbildung},$$
$$f|V_1 : V_1 \to W_1 \text{ ist ein Isomorphismus}.$$

Jede lineare Abbildung $fV \to W$ zerfällt in zwei Bausteine: Die Nullabbildung und einen Isomorphismus. Genauer zerfällt nicht nur f, sondern auch Definitionsbereich und Zielbereich zerfallen.

□

Wenn wir Endomorphismen *eines* Raumes V betrachten, scheint eine andere Äquivalenzrelation angebrachter:

Definition 1 (Ähnlichkeit). Wir nennen $f, \tilde{f} \in \mathrm{End}(V)$ *ähnlich*, wenn es *einen* Automorphismus $\phi \in \mathrm{Aut}(V)$ gibt, so daß

$$\tilde{f} = \phi^{-1} \circ f \circ \phi.$$

Entsprechend nennen wir zwei (n, n)-Matrizen A und $\tilde{A}$ *ähnlich*, wenn es eine invertierbare (n, n)-Matrix S gibt, so daß

$$\tilde{A} = S^{-1}AS.$$

Jetzt hat man also nur noch *einen* Automorphismus oder eine invertierbare Matrix zur Verfügung und damit weniger Flexibilität, die Klassen werden kleiner und mehr: man schaut genauer hin.

Bemerkung. Es ist natürlich egal, ob man $\tilde{f} = \phi^{-1} \circ f \circ \phi$ oder $\tilde{f} = \phi \circ f \circ \phi^{-1}$ schreibt. Die zweite Version ist aus mathematischen Gründen schöner, weil

$$\phi \mapsto \phi \dots \phi^{-1}$$

ein Homomorphismus von $\mathrm{GL}(V)$ in $\mathrm{GL}(\mathrm{End}(V))$ ist. Die erste Version hat insbesondere bei Matrizen vielfach den Vorteil, daß die Spalten der Matrix S eine einfache geometrische Bedeutung haben.

Mit dieser Situation wollen wir uns nun beschäftigen. Grob gesprochen wollen wir den Endomorphismus f in „ einfache Bausteine“ zerlegen, um dadurch eine Übersicht zu erhalten, was in einer Klasse möglich ist.

Die Zerlegung werden wir, ähnlich wie oben, beschreiben mit Hilfe der Zerlegung des Vektorraumes V in die direkte Summe von Unterräumen, die nun aber f-invariant sein sollen. Wir präzisieren diese Begriffe:

Definition 2. Seien $U, V_1, \dots, V_k$ Untervektorräume von V.

(i) V heißt die *direkte Summe* von $V_1, \dots, V_k$, geschrieben

$$V = V_1 \oplus \dots \oplus V_k,$$

wenn sich jedes $v \in V$ schreiben läßt als

$$v = v_1 + \dots + v_k$$

mit *eindeutig bestimmten* $v_i \in V_i$.

(ii) Ist V ein Euklidischer oder unitärer Vektorraum, so nennen wir

$$V = V_1 \oplus \dots \oplus V_k,$$

eine *orthogonale direkte Summe*, wenn für alle $i \neq j$

$$V_i \perp V_j,$$

d.h. für alle $v_i \in V_i, v_j \in V_j$ ist $\langle v_i, v_j \rangle = 0$.

(iii) U heißt *f-invariant*, wenn $f(U) \subset U$.

Hat man eine Zerlegung von V in f-invariante Unterräume gefunden, so sind die

$$f_i := f|_{V_i} : V_i \to V_i$$

die oben angesprochenen „Bausteine“ von f.

Was es bedeutet, daß diese Bausteine „einfach“ sind, hängt von der Situation ab, wir müssen das im einzelnen noch vereinbaren.

„Einfach“ ist sicher, wenn $f_i = 0$ oder $f_i = \mathrm{id}$, aber das läßt sich nicht immer erreichen. „Einfach“ ist vielleicht auch noch, wenn $f_i = \lambda_i \,\mathrm{id}$, d.h. wenn

$$f(v_i) = \lambda_i v_i \text{ für alle } v_i \in V_i.$$

Das führt auf die Begrifffe *Eigenwert* und *Eigenvektor*, die wir zunächst betrachten wollen.

1 Eigenwerte und -vektoren

Sei V ein endlich-dimensionaler Vektorraum über dem Körper K.

Definition 3 (Eigenvektor und Eigenwert). Sei $f \in \mathrm{End}(V)$. Ein Vektor $v \in V \backslash \{0\}$ heißt ein *Eigenvektor* von f, wenn es ein $\lambda \in K$ gibt, für das

$$f(v) = \lambda v.$$

Der Skalar λ heißt in diesem Fall ein *Eigenwert* von f.

Eigenwerte und -vektoren eine Matrix $A \in M(n, n; K)$ sind definiert als die Eigenwerte und -vektoren der zughörigen linearen Abbildung $\in \mathrm{End}(K^n)$.

Beachten Sie, daß Eigenvektoren nach Definition $\neq 0$ sind. Dagegen kann 0 sehr wohl Eigenwert sein, nämlich dann, wenn $\mathrm{Kern}\, f \neq \{0\}$ ist.

Die Gleichung $f(v) = \lambda v$ läßt sich folgendermaßen umschreiben:

$$f(v) = \lambda v \iff f(v) - \lambda v = 0 \iff (f - \lambda\,\mathrm{id})(v) = 0.$$

Die Eigenvektoren zum Eigenwert λ sind also gerade die nicht-trivialen Lösungen des homogenen linearen Gleichungssystems

$$\boxed{(f - \lambda\,\mathrm{id})(v) = 0.}$$

Insbesondere sind nicht-triviale Vielfache eines Eigenvektors und allgemeiner nicht-triviale Linearkombinationen von Eigenvektoren zum selben Eigenwert λ wieder Eigenvektoren.

Definition 4 (Eigenraum). Seien $f \in \mathrm{End}(V)$ und $\lambda \in K$. Dann definieren wir

$$\mathrm{Eig}(f, \lambda) := \mathrm{Kern}(f - \lambda\,\mathrm{id}).$$

Ist dieser Raum $\neq 0$, so nennen wir ihn *den Eigenraum von f zum Eigenwert λ.*

Die Eigenwerte von f sind also gerade jene λ, für die

$$\mathrm{Eig}(f, \lambda) \neq \{0\}.$$

Bemerkung. Die Eigenwerte von f sind die λ, für welche

$$\mathrm{Kern}(f - \lambda\,\mathrm{id}) \neq \{0\},$$

d.h. für welche $f - \lambda\,\mathrm{id}$ nicht injektiv ist. Man nennt λ einen *Spektralwert*, wenn $f - \lambda\,\mathrm{id}$ nicht *bijektiv* ist. Jeder Eigenwert ist also ein Spektralwert. Im endlich-dimensionalen Fall ist umgekehrt jeder Spektralwert ein Eigenwert, nicht aber im Fall unendlich-dimensionaler Räume. In der Funktionalanalysis spricht man deshalb auch von *Spektraltheorie* statt von Eigenwerttheorie.

Satz 1 (Unabhängigkeit von Eigenvektoren). *Sei $f \in \mathrm{End}(V)$ und seien $v_1, \ldots, v_k$ Eigenvektoren von f zu paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_k$. Dann sind $v_1, \ldots, v_k$ linear unabhängig.*

Beweis. Seien $x_1, \ldots, x_k \in K$ mit

$$x_1 v_1 + \ldots + x_k v_k = 0.$$

Wir müssen zeigen, daß $x_1 = \ldots = x_k = 0$. Durch wiederholte Anwendung von f auf diese Gleichung erhalten wir wegen $f^i(v_j) = \lambda_j^i$

$$\begin{array}{rcrl} x_1 v_1 & + \ldots + & x_k v_k & = 0 \\ \lambda_1 x_1 v_1 & + \ldots + & \lambda_k x_k v_k & = 0 \\ & & \ldots & \\ \lambda_1^{k-1} x_1 v_1 & + \ldots + & \lambda_k^{k-1} x_k v_k & = 0 \end{array}$$

Das ist ein homogenes lineares Gleichungssystem, allerdings mit „vektoriellen“ Zeilen. Wir machen daraus ein gewöhnliches mit skalaren Zeilen, indem wir ein beliebiges $\omega \in V^*$ auf alle Zeilen anwenden:

$$\begin{array}{rcrl} x_1 \omega(v_1) & + \ldots + & x_k \omega(v_k) & = 0 \\ \lambda_1 x_1 \omega(v_1) & + \ldots + & \lambda_k x_k \omega(v_k) & = 0 \\ & & \ldots & \\ \lambda_1^{k-1} x_1 \omega(v_1) & + \ldots + & \lambda_k^{k-1} x_k \omega(v_k) & = 0 \end{array}$$

Die Systemmatrix ist die Vandermondesche Matrix, und die hat für paarweise verschiedene λ_i eine Determinante $\neq 0$. Also hat das System nur die triviale Lösung

$$x_1 \omega(v_1) = \ldots = x_k \omega(v_k) = 0. \tag{1}$$

Das gilt für jedes ω. Zu jedem i können wir aber ein ω_i so wählen, daß $\omega_i(v_i) \neq 0$. (Warum gibt es das?) Dann folgt $x_i = 0$. □

Lemma 1 (Diagonalisierung von Endomorphismen). *Seien $f \in \mathrm{End}(V)$ und*

$$\mathbf{v} = (v_1, \ldots, v_n)$$

eine Basis von V. Dann sind folgende Aussagen äquivalent:

(i) Die Darstellungsmatrix von f bezüglich $\mathbf{v}$ ist diagonal:

$$A_{\mathbf{v}}^{\mathbf{v}}(f) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

(ii) Die Basis $\mathbf{v}$ besteht aus Eigenvektoren von f und es ist

$$f(v_i) = \lambda_i v_i.$$

Beweis. Trivial. □

Definition 5 (Diagonalisierbarkeit von Endomorphismen). $f \in \mathrm{End}(V)$ heißt *diagonalisierbar*, wenn V eine Basis aus Eigenvektoren von f besitzt.

Sei f diagonalisierbar. Natürlich kann es sein, daß $\lambda_i = \lambda_j$ für $i \neq j$: (Ein extremer Fall ist $f = \mathrm{id}$, dann sind alle Eigenwerte = 1.) Numeriert man in diesem Fall die Eigenwerte anders, nämlich so daß $\lambda_1, \ldots, \lambda_k$ die sämtlichen *paarweise verschiedenen* Eigenwerte von f sind, so gilt bei passender Numerierung der Basisvektoren einer „Eigenbasis“:

$$\begin{array}{lll} f(v_1) = \lambda_1 v_1, & f(v_2) = \lambda_1 v_2, \ldots & f(v_{i_1}) = \lambda_1 v_{i_1} \\ f(v_{i_1+1}) = \lambda_2 v_{i_1+1}, & \ldots\ldots & f(v_{i_2}) = \lambda_2 v_{i_2} \\ & \ldots & \\ f(v_{i_{k-1}+1}) = \lambda_k v_{i_{k-1}+1}, & \ldots\ldots & f(v_n) = \lambda_k v_n \end{array}$$

Die Vektoren

$$v_{i_{j-1}+1}, \ldots, v_{i_j}$$

spannen dann gerade den Eigenraum $\mathrm{Eig}(f, \lambda_{i_j})$ auf und V ist die direkte Summe der Eigenräume. Die Umkehrung ist auch klar. Also hat man:

Lemma 2. *Seien $\lambda_1, \ldots, \lambda_k$ die sämtlichen paarweise verschiedenen Eigenwerte von $f \in \mathrm{End}(V)$. Dann ist f genau dann diagonalisierbar, wenn*

$$V = \mathrm{Eig}(f, \lambda_1) \oplus \ldots \oplus \mathrm{Eig}(f, \lambda_k).$$

In diesem Fall besteht f also wirklich aus einfachen Bausteinen. Deshalb ist die Frage interessant, ob jedes $f \in \mathrm{End}(V)$ diagonalisierbar ist. Dazu muß man besser verstehen, wie man die Eigenwerte und Eigenvektoren von f finden kann. Bevor wir dieses Problem in Angriff nehmen, betrachten wir aber noch den Matrixfall.

Lemma 3 (Diagonalisierung von Matrizen). *Sei $A \in M(n, n; K)$ eine beliebige Matrix und sei $S = (v_1, \ldots, v_n) \in M(n, n; K)$ eine* invertierbare *Matrix mit den Spaltenvektoren $v_1, \ldots, v_n$. Dann sind folgende Aussagen äqivalent:*

(i) Es gilt

$$S^{-1}AS = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}. \tag{2}$$

(ii) Die Vektoren $(v_1, \ldots, v_n)$ bilden eine Basis des K^n aus Eigenvektoren von A mit

$$Av_i = \lambda_i v_i.$$

Beweis. Die Gleichung (2) ist äquivalent zu

$$A \underbrace{Se_i}_{=v_i} = \lambda_i \underbrace{Se_i}_{=v_i}.$$

Daraus folgt die Behauptung. □

Definition 6 (Diagonalisierbarkeit von Matrizen). Die Matrix $A \in M(n, n; K)$ heißt *diagonalisierbar*, wenn K^n eine Basis aus Eigenvektoren von A besitzt.

Nun zu den schon angesprochenen Fragen: Gibt es zu jedem f eine Basis aus Eigenvektoren? Wieviele Eigenwerte gibt es zu einem gegebenen f und wie kann man sie finden? Wieviel (linear unabhängige) Eigenvektoren kann man zu einem Eigenwert finden?

Satz 2 (Charakteristische Gleichung). *Seien $f \in \mathrm{End}(V)$ und $\lambda \in K$. Dann gilt:*

$$\lambda \textit{ ist ein Eigenwert von } f \iff \det(f - \lambda \,\mathrm{id}) = 0.$$

Entsprechendes gilt für Matrizen.
Man nennt

$$\det(f - \lambda \,\mathrm{id}) = 0 \tag{3}$$

die charakteristische Gleichung von f.

Beweis. Klar. □

Die Determinante des Endomorphismus $f - \lambda\,\mathrm{id}$ ist gleich der Determinante einer Darstellungsmatrix von diesem Endomorphismus. Wählt man also in V eine Basis und bezeichnet die Darstellungsmatrix von f mit A, so sind die Eigenwerte von f gerade die Lösungen von

$$\det(A - \lambda E) = 0. \tag{4}$$

Ist $A = (a_{ij})$, so ist

$$A - \lambda E = \begin{pmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & & \ddots & \\ a_{n1} & & & a_{nn} - \lambda \end{pmatrix}$$

Durch Entwicklung nach der ersten Spalte und vollständige Induktion sieht man, daß (3) bzw. (4) eine Gleichung der Form

$$(-\lambda)^n + c_{n-1}(-\lambda)^{n-1} + \ldots + c_1(-\lambda) + c_0 = 0$$

ist. Also ist das sogenannte *charakteristische Polynom von f*

$$\boxed{\chi_f(\lambda) := \det(f - \lambda\,\mathrm{id})}$$

ein Polynom n-ten Grades.

Beispiel 3. Wir bestimmen über $K = \mathbb{R}$ die Eigenwerte und Eigenvektoren der Matrix

$$A = \begin{pmatrix} 1 & 2 \\ 5 & 4 \end{pmatrix}.$$

Die charakteristische Gleichung ist

$$0 = \det \begin{pmatrix} 1 - \lambda & 2 \\ 5 & 4 - \lambda \end{pmatrix} = (1 - \lambda)(4 - \lambda) - 10 = \lambda^2 - 5\lambda - 6$$

Daraus ergibt sich

$$\lambda_{1,2} = \frac{5}{2} \pm \sqrt{\frac{25}{4} + 6} = \frac{5}{2} \pm \frac{7}{2} = \begin{cases} 6 \\ -1 \end{cases}.$$

Eigenvektoren zum Eigenwert 6. Wir lösen die Gleichung $(A - \lambda E)v = 0$:

$$\begin{pmatrix} -5 & 2 \\ 5 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Die Zeilen sind linear abhängig. Natürlich, denn so haben wir λ gerade bestimmt. Wir brauchen also nur z.B. die erste Zeile zu betrachten und finden

$$\mathrm{Eig}(A; 6) = \{t \begin{pmatrix} 2 \\ 5 \end{pmatrix} \mid t \in \mathbb{R}\}.$$

Eigenvektoren zum Eigenwert -1. Aus

$$\begin{pmatrix} 2 & 2 \\ 5 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

finden wir ebenso

$$\mathrm{Eig}(A; -1) = \{t \begin{pmatrix} 1 \\ -1 \end{pmatrix} \mid t \in \mathbb{R}\}.$$

Eine Basis des $\mathbb{R}^2$ aus Eigenvektoren ist

$$b_1 = \begin{pmatrix} 2 \\ 5 \end{pmatrix}, \quad b_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Setzt man

$$S = \begin{pmatrix} 2 & 1 \\ 5 & -1 \end{pmatrix},$$

so wird

$$S^{-1} = \begin{pmatrix} 1/7 & 1/7 \\ 5/7 & -2/7 \end{pmatrix} \quad \text{und} \quad S^{-1}AS = \begin{pmatrix} 6 & 0 \\ 0 & -1 \end{pmatrix}.$$

□

Bemerkung. Wir haben früher schon einmal am Rande bemerkt, daß der Polynombegriff nicht so ganz einfach ist. Für $K = \mathbb{Z}_2$ sind zum Beispiel die beiden Abbildungen

$$\lambda \mapsto 1 + \lambda + \lambda^2$$

und

$$\lambda \mapsto 1$$

gleich, weil $1 + 1 = 0$. Ist $1 + x + x^2$ nun ein Polynom vom Grad 2 oder vom Grad 0?

In der Algebra unterscheidet man zwischen

1. dem „formalen“ *Polynom* $p(X) = a_0 + a_1X + \ldots + a_nX^n$, das präzise gesagt ein Element von $K^{[\mathbb{N}]}$ ist, also eine Abbildung $\mathbb{N} \to K, i \mapsto a_i$, die nur endlich oft einen Wert $\neq 0$ annimmt. Das X ist hierbei ein Symbol, das einem die Notation erleichtert und keine inhaltliche Bedeutung hat. Man nennt es eine *Unbestimmte*.
2. der dadurch vermittelten *Polynomabbildung*

$$p_K : K \to K, \lambda \mapsto a_0 + a_1\lambda + \ldots + a_n\lambda^n.$$

Bei unendlichen Körpern, also zum Beispiel bei $\mathbb{Q}$, $\mathbb{R}$ oder $\mathbb{C}$, ist die Abbildung $p(X) \mapsto p_K$ injektiv (Identitätssatz für Polynomfunktionen) und man kann auf diese Unterscheidung verzichten.
Im allgemeinen aber muß man, um das charakteristische Polynom eines Endomorphismus zu definieren, zunächst den Begriff der Determinantenformen von Vektorräumen verallgemeinern auf Moduln über kommutativen Ringen.
Wir ignorieren diese Problematik für den Augenblick, indem wir uns auf unendliche Körper beschränken. Die Ergebnisse bleiben allgemein richtig, wenn man die richtige Definition des charakteristischen Polynoms und der Ordnung von Nullstellen von Polynomen zugrunde legt. Wir kommen später darauf zurück.

Wir fragen nun, ob es zu $f \in \mathrm{End}(V)$ eine Basis von V aus Eigenvektoren von f gibt.

Zunächst halten wir fest, daß zum Beispiel für $K = \mathbb{R}$ überhaupt kein Eigenwert existieren muß:

Beispiel 4. Die Matrix

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

hat das charakteristischen Polynom $\chi_f(\lambda) = \lambda^2 + 1$, und das hat keine Nullstelle in $\mathbb{R}$. Also gibt es keine Eigenvektoren und erst recht keine Basis aus solchen. A ist nicht diagonalisierbar.

□

Gibt es umgekehrt $n = \dim V$ *verschiedene* Eigenwerte, so gibt es dazu Eigenvektoren, und die sind nach Satz 1 linear unabhängig, also eine Basis.

Was passiert bei mehrfachen Nullstellen? Dazu beweisen wir:

Satz 3 (Algebraische und geometrische Vielfachheit). *Seien* $f \in \mathrm{End}(V)$, $\lambda_0 \in K$ *und*

$$g := \dim \mathrm{Eig}(f, \lambda_0) > 0.$$

Man nennt g *die* geometrische Vielfachheit *des Eigenwertes* λ_0. *Dann ist*

$$\chi_f(\lambda) = (\lambda_0 - \lambda)^g p(\lambda)$$

mit einem Polynom $p(\lambda)$. *Die* algebraische Vielfachheit *von* λ_0, *also die Vielfachheit der Nullstelle* λ_0 *von* $\chi_f(\lambda)$, *ist mindestens so groß wie die geometrische.*

Beweis. Sei $\mathbf{v} := (v_1, \ldots, v_n)$ eine Basis von V, so daß $(v_1, \ldots, v_g)$ eine Basis von $\mathrm{Eig}(f, \lambda_0)$ ist. Dann gilt

$$A^{\mathbf{v}}_{\mathbf{v}}(f) = \begin{pmatrix} \begin{bmatrix} \lambda_0 & & 0 \\ & \ddots & \\ 0 & & \lambda_0 \end{bmatrix} & B \\ 0 & C \end{pmatrix}$$

mit einer $g \times g$-Matrix in der linken oberen Ecke. Daher ist

$$\chi_f(\lambda) = (\lambda_0 - \lambda)^g \det(C - \lambda E).$$

□

Also liefert eine k-fache Nullstelle des charakteristischen Polynoms höchstens k linear unabhängige Eigenvektoren.

Beispiel 5. Die Matrix

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

hat das charakteristische Polynom $\chi_A(\lambda) = (3 - \lambda)^2$ mit 3 als doppelter Nullstelle. Aber

$$\mathrm{Eig}(A, 3) = \mathrm{Kern} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

hat die Dimension $1 < 2$. Die geometrische Vielfachheit ist kleiner als die algebraische.

□

Satz 4. *Notwendig und hinreichend für die Diagonalisierbarkeit eines Endomorphismus* $f \in \mathrm{End}(V)$ *ist, daß das charakteristische Polynom in Linearfaktoren zerfällt und die algebraischen Vielfachheiten gleich den geometrischen sind, d.h daß folgendes gilt:*

$$\chi_f(\lambda) = (\lambda_1 - \lambda)^{a_1} \ldots (\lambda_k - \lambda)^{a_k} \tag{5}$$

mit paarweise verschiedenen $\lambda_1, \ldots, \lambda_k$ *und*

$$\dim \mathrm{Eig}(f, \lambda_i) = a_i \quad \textit{für alle } i \in \{1, \ldots, k\}.$$

Beweis. Ist f diagonalisierbar, so gibt es eine Basis, bezüglich der die Darstellungsmatrix von der Form

$$A(f) = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_k \end{pmatrix}$$

mit

$$A_i = \begin{pmatrix} \lambda_i & & 0 \\ & \ddots & \\ 0 & & \lambda_i \end{pmatrix}$$

ist. Dabei ist A_i quadratisch mit $g_i := \dim \mathrm{Eig}(f, \lambda_i)$ Zeilen. Es folgt, daß χ_f von der angegebenen Gestalt mit $a_i = g_i$ ist.

Zerfällt umgekehrt das charakteristische Polynom wie in (5) und sind die geometrischen Vielfachheiten gleich den algebraischen, so wählt man in jedem Eigenraum eine Basis. Weil

$$g_1 + \ldots + g_k = a_1 + \ldots + a_k = \dim V$$

und Eigenvektoren zu verschiedenen Eigenwerten linear unabhängig sind, setzen sich diese Basen zu einer „Eigenbasis“ des ganzen Raumes zusammen, d.h. f ist diagonalisierbar. □

Später werden wir Normalformen herleiten für Endomorphismen, deren charakteristisches Polynom nicht in Linearfaktoren zerfällt oder deren geometrische Vielfachheiten „zu klein“ sind.

Zuvor betrachten wir aber die Situation für spezielle Endomorphismen in Vektorräumen mit einem Skalarprodukt.

2 Strukturätze in unitären und Euklidischen Räumen

2.1 Normale Endomorphismen in unitären Räumen

In diesem Abschnitt sei $(V, \langle ., . \rangle)$ zunächst ein Euklidischer oder ein endlich-dimensionaler unitärer Vektorraum über $\mathbb{R}$ bzw. $\mathbb{C}$. Später beschränken wir uns auf den unitären Fall.

Das folgende Lemma macht allerdings vom Skalarprodukt keinen Gebrauch, es beruht auf dem Fundamentalsatz der Algebra, daß jedes komplexe Polynom in Linearfaktoren zerfällt.

Lemma 4. *Ist f ein Endomorphismus eines n-dimensionalen $\mathbb{C}$-Vektorraumes, so besitzt $\chi_f(\lambda)$ genau n Eigenwerte, wenn man sie mit der algebraischen Vielfachheit zählt:*

$$\chi_f(\lambda) = (\lambda_1 - \lambda) \cdot \ldots \cdot (\lambda_n - \lambda).$$

Insbesondere ist $\det f = \det(f - 0\,\mathrm{id}) = \lambda_1 \cdot \ldots \cdot \lambda_n$ *das Produkt der Eigenwerte.*

Für die Komposition von Endomorphismen schreiben wir zur Vereinfachung auch fg statt $f \circ g$.

Definition 7 (Normaler Endomorphismus). Seien $f \in \mathrm{End}(V)$ und f^* die Adjungierte zu f, also

$$\langle f(v), w \rangle = \langle v, f^*(w) \rangle.$$

Dann heißt f *normal*, wenn

$$ff^* = f^*f.$$

Diese Definition ist zunächst ziemlich unmotiviert. Wir sehen später, daß im unitären Fall die normalen Endomorphismen dadurch charakterisiert sind, daß sie eine Orthonormalbasis aus Eigenvektoren besitzen, also „orthonormal diagonalisierbar“ sind.

Immerhin gibt es wichtige Klassen von Endomorphismen, die normal sind:

Beispiel 6. Selbstadjungierte ($f = f^*$), schiefadjungierte ($f = -f^*$) und orthogonale (bzw. unitäre) ($f^{-1} = f^*$) Endomorphismen sind (offenbar) normal. Wegen

$$\langle \lambda v, w\rangle = \lambda\langle v, w\rangle = \langle v, \bar\lambda w\rangle$$

ist $(\lambda\,\mathrm{id})^* = \bar\lambda\,\mathrm{id}$. Ist f normal, so ist daher

$$(f - \lambda\,\mathrm{id})(f - \lambda\,\mathrm{id})^* = (f - \lambda\,\mathrm{id})(f^* - \bar\lambda\,\mathrm{id}) = ff^* - \lambda f^* - \bar\lambda f - |\lambda|^2\,\mathrm{id}\,.$$

Ist also f normal, so auch $f - \lambda\,\mathrm{id}$ für alle $\lambda \in \mathbb{C}(\mathbb{R})$.

□

Lemma 5. *Sei $f \in \mathrm{End}(V)$ normal. Dann gilt*

(i) $\langle f(v), f(w)\rangle = \langle f^*(v), f^*(w)\rangle$ *für alle* $v, w \in V$,

(ii) $\mathrm{Kern}\, f = \mathrm{Kern}\, f^*$,

(iii) $\mathrm{Kern}\, f = (\mathrm{Bild}\, f)^\perp$,

(iv) $f(v) = \lambda v \iff f^*(v) = \bar\lambda v$. *(Im Euklidischen Fall ist $\bar\lambda = \lambda$.)*

Beweis. *Zu (i).*

$$\langle f(v), f(w)\rangle = \langle v, f^*f(w)\rangle = \langle v, ff^*(w)\rangle = \langle f^*(v), f^*(w)\rangle.$$

Zu (ii).

$$f(v) = 0 \iff \langle f(v), f(v)\rangle = 0 \underset{(i)}{\iff} \langle f^*(v), f^*(v)\rangle = 0 \iff f^*(v) = 0.$$

Zu (iii). Es gilt

$$\begin{aligned} f(v) = 0 &\underset{(ii)}{\iff} f^*(v) = 0 \\ &\iff \forall_w \langle w, f^*(v)\rangle = 0 \\ &\iff \forall_w \langle f(w), v\rangle = 0 \\ &\iff v \in (\mathrm{Bild}\, f)^\perp. \end{aligned}$$

Zu (iv). Weil $f - \lambda\,\mathrm{id}$ normal ist, ist

$$\begin{aligned} f(v) = \lambda v &\iff \langle (f - \lambda\,\mathrm{id})(v), (f - \lambda\,\mathrm{id})(v)\rangle \\ &\underset{(i)}{\iff} \langle (f - \lambda\,\mathrm{id})^*(v), (f - \lambda\,\mathrm{id})^*(v)\rangle \\ &\iff f^*(v) = \bar\lambda v. \end{aligned}$$

□

Lemma 6. *Seien $f \in \mathrm{End}(V)$ normal und $U \subset V$ ein f-invarianter Unterraum. Wir bezeichnen mit*

$$f_U : U \to U, u \mapsto f(u)$$

den von $f|_U$ induzierten Endomorphismus von U. Dann gilt

(i) $U^\perp$ ist f^*-invariant.

(ii) Ist U auch f^* invariant, so ist also auch $U^\perp$ invariant unter f und f^*. In diesem Fall sind $f_U \in \mathrm{End}(U)$ und $f_{U^\perp} \in \mathrm{End}(U^\perp)$ normal.

Beweis. Zu (i). Für $u \in U$ und $v \in U^\perp$ gilt

$$\langle u, f^*(v)\rangle = \langle f(u), v\rangle = 0.$$

Zu (ii). Es genügt zu zeigen, daß f_U normal ist. Zunächst gilt für $u_1, u_2 \in U$, daß

$$\langle u_1, (f_U)^*(u_2)\rangle = \langle f_U(u_1), u_2\rangle = \langle f(u_1), u_2\rangle = \langle u_1, f^*(u_2)\rangle = \langle u_1, (f^*)_U(u_2)\rangle.$$

Daher ist

$$(f^*)_U = (f_U)^*.$$

Daraus folgt für $u \in U$

$$f_U(f_U)^*(u) = f_U(f^*)_U(u) = f_U f^*(u) = f f^*(u) = f^* f(u) = (f^*)_U f_U(u) = (f_U)^* f_U(u).$$

Also ist f_U normal. □

Bemerkung. Mit dem nachstehenden Satz 5 zeigt man unter Verwendung der Vandermondeschen Matrix aus dem letzten Semester, daß es zu jedem normalen Endomorphismus f auf einem endlich-dimensionalen unitären Vektorraum ein Polynom $G(X)$ gibt, so daß $f^* = G(f)$ ist. Unter Benutzung dieses Resultats zeigt man dann dasselbe auch in Euklidischen Vektorräumen. Damit ist aber jeder f-invariante Unterraum U auch f^*-invariant und mit U auch $U^\perp$ f-invariant. Das ist eine wesentliche Verschärfung des obigen Lemmas, die sich aber erst im Nachhinein ergibt.

Satz 5 (Normale Endomorphismen in unitären Räumen). *Seien $(V, \langle .,.\rangle)$ ein n-dimensionaler unitärer Vektorraum, $1 \le n < \infty$, und $f \in \mathrm{End}(V)$. Dann besitzt V eine Orthonormalbasis aus Eigenvektoren von f genau dann, wenn f normal ist.*

Beweis. **A.** Zunächst nehmen wir an, daß f normal ist und zeigen die Existenz einer orthonormalen Eigenbasis durch vollständige Induktion über n.

$n = 1$. Dann gibt es in V einen Einheitsvektor v. Der ist eine ONBasis und erfüllt $f(v) = \lambda v$.

$(n-1) \to n$. Sei der Satz bewiesen für normale Endomorphismen von $(n-1)$-dimensionalen unitären Räumen. Nach dem Fundamentalsatz der Algebra besitzt das charakteristische Polynom von f eine Nullstelle. Also gibt es einen Eigenvektor v_1 von f, und o.E. $\|v_1\| = 1$. Sei $f(v_1) = \lambda v_1$ und sei

$$U := (\mathbb{C}v_1).$$

Dann ist U invariant unter f und nach Aussage (iv) von Lemma 5 auch unter f^*. Nach Lemma 6 ist daher

$$V = U \oplus U^\perp$$

als direkte Summe f invarianter Unterräume, auf denen f normal ist. Nach Induktionsvoraussetzung besitzt $U^\perp$ eine Orthonormalbasis $(v_2, \ldots, v_n)$ aus Eigenvektoren von $f_{U^\perp}$, also von f und $(v_1, \ldots, v_n)$ ist dann eine Orthonormalbasis von V aus Eigenvektoren von f.

B. Besitzt umgekehrt V eine ONBasis aus Eigenvektoren von f, so ist die zugehörige Darstellungsmatrix A diagonal:

$$A = \begin{pmatrix} \lambda_1 & \ldots & 0 \\ & \ddots & \\ 0 & \ldots & \lambda_n \end{pmatrix}$$

Die Darstellungsmatrix von f^* bezüglich dieser Basis ist

$$A^* = \bar{A}^T = \begin{pmatrix} \overline{\lambda_1} & \ldots & 0 \\ & \ddots & \\ 0 & \ldots & \overline{\lambda_n} \end{pmatrix}.$$

Weil der Komposition der Endomorphismen das Produkt der Darstellungsmatrizen entspricht und die Matrizen A und A^* kommutieren, ist f normal. □

Aus dem Satz folgt, daß die Eigenräume von f paarweise orthogonal sind, d.h. daß Eigenvektoren zu verschiedenen Eigenwerten orthogonal sind. Wir beweisen das noch einmal direkt (auch für den Euklidischen Fall):

Lemma 7. *Sei f ein normaler Endomorphismus eines Euklidischen oder unitären Raumes. Seien $v, w \in V \backslash \{0\}$ und*

$$f(v) = \lambda v, \quad f(w) = \mu w$$

mit $\lambda \neq \mu$. Dann gilt

$$\langle v, w \rangle = 0.$$

Eigenvektoren eines normalen Endomorphismus zu verschiedenen Eigenwerten sind orthogonal.

Beweis.

$$\lambda \langle v, w \rangle = \langle \lambda v, w \rangle = \langle f(v), w \rangle = \langle v, f^*(w) \rangle = \langle v, \bar{\mu} w \rangle = \mu \langle v, w \rangle.$$

Die Normalität haben wir für das vorletzte Gleichheitszeichen benötigt. Es folgt die Behauptung. □

Eine andere Formulierung des Satzes 5 ist der

Satz 6 (Spektralzerlegungssatz für normale Endomorphismen). *Sei f ein normaler Endomorphismus des unitären Raumes $(V, \langle ., . \rangle)$. Seien $\lambda_1, \ldots, \lambda_k \in \mathbb{C}$ die* verschiedenen *Eigenwerte von f. Dann gilt*

$$V = \mathrm{Eig}(f, \lambda_1) \oplus \ldots \oplus \mathrm{Eig}(f, \lambda_k)$$

also orthogonale direkte Summe.

Aus Satz 5 folgt unmittelbar die Diagonalisierbarkeit von selbstadjungierten, schiefadjungierten oder unitären Endomorphismen in unitären Räumen.

Satz 7 (Selbstadjungierte Endomorphismen in unitären Räumen). *Seien $(V, \langle ., . \rangle)$ ein n-dimensionaler unitärer Vektorraum und $f \in \mathrm{End}(V)$ selbstadjungiert. Dann besitzt f eine Orthonormalbasis $\mathbf{b}$ aus Eigenvektoren von f. Die Eigenwerte sind alle reell.*

$$A_{\mathbf{b}}^{\mathbf{b}}(f) = \begin{pmatrix} \lambda_1 & \dots & 0 \\ & \ddots & \\ 0 & \dots & \lambda_n \end{pmatrix}, \qquad \lambda_k \in \mathbb{R}.$$

Beweis. Die Existenz einer ONBasis aus Eigenvektoren folgt aus Satz 5.

Ist $v \neq 0$ mit $f(v) = \lambda v$, so folgt

$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle f(v), v\rangle = \langle v, f(v)\rangle = \langle v, \lambda v\rangle = \bar{\lambda}\langle v, v\rangle.$$

Daher $\lambda = \bar{\lambda}$, also $\lambda \in \mathbb{R}$. □

Satz 8 (Schiefadjungierte Endomorphismen in unitären Räumen). *Seien $(V, \langle ., . \rangle)$ ein n-dimensionaler unitärer Vektorraum und $f \in \mathrm{End}(V)$ schiefadjungiert. Dann besitzt f eine Orthonormalbasis $\mathbf{b}$ aus Eigenvektoren von f. Die Eigenwerte sind alle rein imaginär (eventuell 0).*

$$A_{\mathbf{b}}^{\mathbf{b}}(f) = \begin{pmatrix} i\lambda_1 & \dots & 0 \\ & \ddots & \\ 0 & \dots & i\lambda_n \end{pmatrix}, \qquad \lambda_k \in \mathbb{R}.$$

Beweis. Die Existenz einer ONBasis aus Eigenvektoren folgt aus Satz 5.

Ist $v \neq 0$ mit $f(v) = \lambda v$, so folgt

$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle f(v), v\rangle = -\langle v, f(v)\rangle = -\langle v, \lambda v\rangle = -\bar{\lambda}\langle v, v\rangle.$$

Daher $\lambda = -\bar{\lambda}$, also $\lambda \in i\mathbb{R}$. □

Satz 9 (Unitäre Automorphismen in unitären Räumen). *Seien $(V, \langle ., . \rangle)$ ein n-dimensionaler unitärer Vektorraum und $f \in \mathrm{End}(V)$ unitär. Dann besitzt f eine Orthonormalbasis $\mathbf{b}$ aus Eigenvektoren von f. Die Eigenwerte sind alle vom Betrage 1, lassen sich also schreiben in der Form $\lambda_k = e^{i\phi_k} = \cos\phi_k + i\sin\phi_k$ mit $\phi_k \in \mathbb{R}$.*

$$A_{\mathbf{b}}^{\mathbf{b}}(f) = \begin{pmatrix} e^{i\phi_1} & \dots & 0 \\ & \ddots & \\ 0 & \dots & e^{i\phi_n} \end{pmatrix}, \qquad \phi_k \in \mathbb{R}.$$

Beweis. Die Existenz einer ONBasis aus Eigenvektoren folgt aus Satz 5.

Ist $v \neq 0$ mit $f(v) = \lambda v$, so folgt

$$\lambda\bar{\lambda}\langle v, v\rangle = \langle \lambda v, \lambda v\rangle = \langle f(v), f(v)\rangle = \langle v, v\rangle.$$

Daher $|\lambda|^2 = \lambda\bar{\lambda} = 1$. □

Wir geben nun Matrixversionen der Normalform-Sätze für normale Endomorphismen in unitären Räumen.

Wir erinnern: Eine Matrix $A \in M(n,n;\mathbb{C})$ heißt

- normal, wenn $A^*A = AA^*$,
- selbstadjungiert, wenn $A^* := \overline{A^T} = A$,
- schiefadjungiert, wenn $A^* = -A$,
- unitär, wenn $A^*A = AA^* = E$, wo E die Einheitsmatrix bezeichnet.

Satz 10 (Normale komplexe Matrizen). *Sei $A \in M(n,n;\mathbb{C})$ normal. Dann gibt es eine unitäre Matrix $S \in M(n,n;\mathbb{C})$, so daß*

$$S^*AS = \begin{pmatrix} \lambda_1 & \dots & 0 \\ & \ddots & \\ 0 & \dots & \lambda_n \end{pmatrix}. \tag{6}$$

Das heißt, jede normale Matrix ist unitär ähnlich *zu einer Diagonalmatrix.*
Ist A selbstadjungiert, schiefadjungiert bzw. unitär, so sind die λ_j reell, rein imaginär bzw. komplexe Zahlen vom Betrage 1.

Beweis. Wir betrachten den $\mathbb{C}^n$ mit der kanonischen hermiteschen Metrik

$$\langle v, w\rangle = \sum v_k \overline{w_k}$$

als unitären Vektorraum. Dann ist die Multiplikation mit A ein normaler Endomorphismus f_A. Also gibt es eine ONBasis $(b_1, \dots, b_n)$ aus Eigenvektoren von f_A:

$$Ab_j = \lambda_j b_j.$$

Bezeichnet S die Matrix mit den b_j als Spalten, so ist S unitär und es gilt für die Standardbasis $(e_1, \dots, e_n)$ des $\mathbb{C}^n$

$$Se_j = b_j.$$

Also hat man

$$ASe_j = \lambda_j Se_j$$

und durch Multiplikation mit $S^* = S^{-1}$ folgt (6). Die Diagonalelemente λ_j sind die Eigenwerte von f_A, und damit folgen die Aussagen des Satzes darüber aus den entsprechenden Sätzen für Endomorphismen. □

Bemerkung. Der Beweis gibt explizit an, wie man eine Matrix S finden kann: Man bestimmt die Eigenwerte und dazu eine ONBasis aus Eigenvektoren. Dabei sind Eigenvektoren zu *verschiedenen* Eigenwerten automatisch orthogonal, ist ein Eigenraum höherdimensional, so muß man eine Basis darin orthonormieren. Die Eigenvektoren bilden dann gerade die Spalten er gesuchten unitären Matrix S.

2.2 Normale Endomorphismen in Euklidischen Räumen

Die Diagonalisierung normaler Endomorphismen im unitären Fall beruhte wesentlich darauf, daß das charakteristische Polynom über den komplexen Zahlen in Linearfaktoren zerfällt.

Im Euklidischen Fall ist das nicht mehr wahr. Die Drehung der Euklidischen Ebene um einen Winkel $\neq 0, \pi$ hat keine reellen Eigenwerte, also keine „Eigenbasis“. Aber ein reelles Polynom zerfällt in lineare und quadratische Faktoren, und dem entspricht eine Zerlegung des Vektorraumes als direkte Summe von höchstens 2-dimensionalen Unterräumen, auf denen die Endomorphismen sich einfach verhalten.

Wir untersuchen den reellen Fall unter Zuhilfenahme der Ergebnisse im Komplexen. Die Idee ist folgende:

Wir identifizieren Endomorphismen mit Matrizen, betrachten also $V = \mathbb{R}^n$. Reelle Matrizen betrachten wir als spezielle komplexe Matrizen. Aus einem nicht-reellen Eigenwert $\mu \in \mathbb{C}\backslash\mathbb{R}$ und einem zugehörigen Eigenvektor $v \in \mathbb{C}^n$ konstruieren wir dann einen zweidimensionalen Unterraum von $\mathbb{R}^n$, der invariant unter A und $A^* = A^T$ ist, und auf dem sich A sehr einfach darstellt.n

Zunächst betrachten wir die Beziehungen zwischen den kanonischen Skalarprodukten auf $\mathbb{R}^n$ und $\mathbb{C}^n$.

Lemma 8. *Wir betrachten $\mathbb{R}^n$ als Teilmenge von $\mathbb{C}^n$ und schreiben Vektoren $z \in \mathbb{C}^n$ als*

$$z = x + iy, \quad x, y \in \mathbb{R}^n.$$

Die kanonischen Skalarprodukte auf $\mathbb{R}^n$ bzw. $\mathbb{C}^n$ seien

$$\langle x, y\rangle_{\mathbb{R}} = \sum_{j=1}^{n} x_j y_j, \quad x, y \in \mathbb{R}^n$$
$$\langle z, w\rangle_{\mathbb{C}} = \sum_{j=1}^{n} z_j \overline{w_j}, \quad z, w \in \mathbb{C}^n.$$

Dann gilt für $x, y, u, v \in \mathbb{R}^n$

$$\langle x + iy, u + iv\rangle_{\mathbb{C}} = \langle x, u\rangle_{\mathbb{R}} + \langle y, v\rangle_{\mathbb{R}} + i(\langle y, u\rangle_{\mathbb{R}} - \langle x, v\rangle_{\mathbb{R}}). \tag{7}$$

Insbesondere ist

$$\langle x + iy, x - iy\rangle_{\mathbb{C}} = \|x\|^2 - \|y\|^2 + 2i\langle x, y\rangle_{\mathbb{R}}. \tag{8}$$

Beweis. Für reelle x, u ist $\langle x, u\rangle_{\mathbb{C}} = \langle x, u\rangle_{\mathbb{R}}$. Die Behauptung folgt dann aus der Sesquilinearität.

□

Lemma 9. *Ist*

$$p(\lambda) = a_n\lambda^n + \ldots + a_1\lambda + a_0, \quad a_j \in \mathbb{R},$$

ein reelles *Polynom vom Grad n, so gibt es $\lambda_1, \ldots \lambda_r, \mu_1, \ldots, \mu_s \in \mathbb{C}$, so daß gilt*

$$\lambda_1, \ldots \lambda_r \in \mathbb{R},$$
$$\mu_1, \ldots, \mu_s \in \mathbb{C},$$
$$\mu_j \neq \overline{\mu_k} \quad \textit{für alle } j, k \in \{1, \ldots, s\},$$
$$p(\lambda) = a_n(\lambda - \lambda_1) \cdot \ldots \cdot (\lambda - \lambda_r)(\lambda - \mu_1)(\lambda - \overline{\mu_1}) \cdot \ldots \cdot (\lambda - \mu_s)(\lambda - \overline{\mu_s}).$$

Insbesondere ist $\mu_k \in \mathbb{C}\backslash\mathbb{R}$ (wähle $j = k$). Die Zerlegung ist eindeutig bis auf die Numerierung der λ und μ.

Beweis. Weil die Koeffizienten a_i reell sind ist

$$\overline{p(\lambda)} = \overline{\sum a_k \lambda^k} = \sum a_k \bar{\lambda}^k = p(\bar{\lambda}).$$

Ist μ also eine nicht-reelle Nullstelle von $p(\lambda)$, so ist auch $\bar{\mu}$ eine solche, und das Polynom läßt sich durch $(\lambda - \mu)(\lambda - \overline{\mu})$ teilen. Durch Induktion über die Anzahl der nicht-reellen Nullstellen folgt die Behauptung aus dem Fundamentalsatz der Algebra. □

Vereinbarung. Für Endomorphismen reeller Vektorräume oder (quadratische) reelle Matrizen bezeichnen wir auch die *komplexen* Nullstellen des charakteristischen Polynoms als *Eigenwerte*. Dann gilt auch für reelle Endomorphismen, daß die Determinante das Produkt der Eigenwerte ist.

Satz 11 (Normale reelle Matrizen). *Sei $A \in M(n,n;\mathbb{R})$ eine normale Matrix. Sei*

$$\chi_A(\lambda) = (-1)^n(\lambda - \lambda_1)\cdot\ldots\cdot(\lambda - \lambda_r)(\lambda - \mu_1)(\lambda - \overline{\mu_1})\cdot\ldots\cdot(\lambda - \mu_r)(\lambda - \overline{\mu_r})$$

die Faktorzerlegung des charakteristischen Polynoms von A gemäß Lemma 9. Dann gibt es eine orthogonale Matrix $S \in O(n)$, so daß

$$S^*AS = \begin{pmatrix} \lambda_1 & \ldots & 0 & 0 & \ldots & 0 \\ & \ddots & & \vdots & & \vdots \\ 0 & \ldots & \lambda_r & 0 & \ldots & 0 \\ 0 & \ldots & 0 & M_1 & \ldots & 0 \\ \vdots & & \vdots & & \ddots & \\ 0 & \ldots & 0 & 0 & \ldots & M_s \end{pmatrix} \tag{9}$$

mit folgenden reellen 2×2-Matrizen M_j: Ist $\mu_j = \alpha_j + i\beta_j$, so ist

$$M_j = \begin{pmatrix} \alpha_j & -\beta_j \\ \beta_j & \alpha_j \end{pmatrix}.$$

Jede normale reelle Matrix ist also orthogonal ähnlich *zu einer Matrix der Form* (9).

Bemerkung. Schreibt man $\alpha + i\beta = Re^{i\phi}$, so ist

$$\begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} = R\begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}.$$

Das ist eine *Drehstreckung*.

Beweis des Satzes. Durch vollständige Induktion über n

$\underline{n = 1.}$ Nichts zu beweisen.

$\underline{(n-1) \to n.}$ *Fall A.* Hat A einen reellen Eigenwert λ_1 mit zugehörigem Eigenvektor v_1, so setzen wir $U = \mathbb{R}v_1$ und erhalten eine Zerlegung

$$\mathbb{R}^n = U \oplus U^\perp$$

in A- und A^*-invariante Unterräume. Nach Lemma 6 ist die Induktion komplett.

Fall B. Hat A keine reellen Eigenwerte, so gibt es einen komplexen Eigenwert $\mu = \alpha + i\beta \in \mathbb{C}$ mit $\beta \neq 0$ und einen zugehörigen Eigenvektor $v = x + iy \in \mathbb{C}^n$ mit $x, y \in \mathbb{R}^n$. Aus $Av = \mu v$ folgt

$$A\bar{v} = \overline{Av} = \overline{\mu v} = \bar{\mu}\bar{v}.$$

Weil Eigenvektoren zu verschiedenen Eigenwerten orthogonal sind, folgt

$$\langle x + iy, x - iy \rangle_{\mathbb{C}} = 0.$$

Aus (8) folgt daher, daß x und y orthogonal und von gleicher Länge sind. O.E. können wir also annehmen, daß sie orthonormal sind. Weiter ist

$$Ax + iAy = A(x + iy) = (\alpha + i\beta)(x + iy) = \alpha x - \beta y + i(\beta x + \alpha y).$$

Also folgt

$$\begin{aligned} Ax &= \alpha x - \beta y \\ Ay &= \beta x + \alpha y. \end{aligned}$$

Ebenso folgt, vgl. Lemma 5 Teil (iv), daß

$$A^*(x + iy) = \overline{(\alpha + i\beta)}(x + iy),$$

also

$$\begin{aligned} A^*x &= \alpha x + \beta y \\ A^*y &= -\beta x + \alpha y. \end{aligned}$$

Daher ist $U = \text{Spann}(x, y) \subset \mathbb{R}^n$ invariant unter A und A^* und nach Lemma 6 sind wir fertig. □

Nun betrachten wir die Endomorphismus- Version dieses Satzes:

Satz 12 (Normale Endomorphismen in Euklidischen Vektorräumen). *Sei f ein normaler Endomorphismus des Euklidischen Vektorraumes $(V, \langle ., . \rangle)$. Dann gibt es Unterräume $V_1, \ldots, V_r, V_{r+1}, \ldots, V_{r+s}$ von V mit folgenden Eigenschaften:*

$$\begin{gathered} \dim V_j = 1 \text{ für } j \in \{1, \ldots, r\}, \\ \dim V_j = 2 \text{ für } j \in \{r+1, \ldots, r+s\}, \\ f(V_j) \subset V_j \text{ und } f^*(V_j) \subset V_j \text{ für alle } j, \\ V = V_1 \oplus \ldots \oplus V_{r+s}, \\ V_j \perp V_l \text{ für } j \neq l. \end{gathered}$$

Insbesondere ist f auf den $V_1, \ldots, V_r$ jeweils eine Streckung und auf den $V_{r+1}, \ldots, V_{r+s}$ eine Drehstreckung.

Beweis. Wir wählen eine Orthonomalbasis $\mathbf{v}$ von V und bezeichnen mit $\Phi : \mathbb{R}^n \to V$ die zugehörige Koordinatenabbildung. Dann ist die Darstellungsmatrix A von f eine reelle normale $n \times n$-Matrix mit

$$f\Phi(x) = \Phi(Ax)$$

und es gibt eine orthogonale Matrix $S = (b_1 \ldots b_n)$, so daß S^*AS die Normalform aus Satz 11 hat. Dann gilt

$$Ab_j = \lambda_j b_j \text{ für } j \in \{1, \ldots, r\},$$

also

$$f\Phi(b_j) = \Phi(Ab_j) = \Phi(\lambda_j b_j) = \lambda_j \Phi(b_j).$$

Ebenso sieht man, daß

$$f(\Phi(b_{r+2j-1})) = \alpha_j \Phi(b_{r+2j-1}) - \beta_j \Phi(b_{r+2j}),$$
$$f(\Phi(b_{r+2j})) = +\beta_j \Phi(b_{r+2j-1}) + \alpha_j \Phi(b_{r+2j}).$$

Daher leisten

$$V_j := \mathbb{R}\Phi(b_j) \qquad j \in \{1, \ldots, r\},$$
$$V_{r+j} := \mathrm{Spann}(b_{r+2j-1}, b_{r+2j}) \qquad j \in \{1, \ldots, s\}.$$

das Gewünschte. □

Nun spezialisieren wir diesen Satz. Für selbstadjungierte Endomorphismen eines Euklidischen Vektorraumes sind alle Nullstellen des charakterstischen Polynoms reell. Daher braucht man keine 2-dimensionalen invarianten Unterräume:

Satz 13 (Hauptachsentransformation). *Sei $f \in \mathrm{End}(V)$ ein selbstadjungierter Endomorphismus des Euklidischen Vektorraumes $(V, \langle ., . \rangle)$. Dann besitzt V eine Orthonormalbasis aus Eigenvektoren von f, d.h. f ist „orthogonal diagonalisierbar“.*

Beispiel 7. Der Name rührt von folgendem Sachverhalt her: Ist f ein selbstadjungierter Endomorphismus des $\mathbb{R}^2$, so ist

$$\langle f(v), v \rangle = c$$

die Gleichung eines Kegelschnittes: Für $f = \mathrm{id}$ und $c = r^2$ erhält man z.B. die Gleichung eines Kreises. Benutzt man nun eine ON-Basis des $\mathbb{R}^2$ aus Eigenvektoren von f und die dadurch gegebenen Koordinaten, so schreibt sich die Gleichung als

$$\lambda_1 x^2 + \lambda_2 y^2 = c.$$

Ist etwa $c = 1$, $\lambda_1 = \frac{1}{a^2}$ und $\lambda_2 = \frac{1}{b^2}$, so hat man eine Ellipsengleichung

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

Die Eigenrichtungen von f sind gerade die Hauptachsen dieser Ellipse.

□

Schiefadjungierte Endomorphismen führen zu schiefadjungierten komplexen Matrizen, und die haben rein-imaginäre Eigenwerte. Damit folgt:

Satz 14 (Schiefadjungierte Endomorphismen in Euklidischen Vektorräumen). *Sei $f \in \mathrm{End}(V)$ ein schiefadjungierter Endomorphismus des Euklidischen Vektorraumes $(V, \langle ., . \rangle)$. Dann besitzt V eine Orthonormalbasis, bezüglich derer f eine Darstellungsmatrix folgender Form bekommt:*

$$\begin{pmatrix}
0 & \ldots & 0 & 0 & \ldots & 0 \\
 & \ddots & & \vdots & & \vdots \\
0 & \ldots & 0 & 0 & \ldots & 0 \\
0 & \ldots & 0 & \begin{bmatrix} 0 & -\beta_1 \\ \beta_1 & 0 \end{bmatrix} & \ldots & 0 \\
\vdots & 0 & \vdots & & \ddots & \\
0 & \ldots & 0 & 0 & \ldots & \begin{bmatrix} 0 & -\beta_s \\ \beta_s & 0 \end{bmatrix}
\end{pmatrix}$$

Bei orthogonalen Automorphismen ist die Darstellungsmatrix (komplex) unitär, die Eigenwerte sind vom Betrag 1, also von der Form $\cos\phi + i\sin\phi$. Das liefert den

Satz 15 (Orthogonale Automorphismen in Euklidischen Vektorräumen). *Sei $f \in \mathrm{End}(V)$ ein orthogonaler Automorphismus des Euklidischen Vektorraumes $(V, \langle .,. \rangle)$. Dann besitzt V eine Orthonormalbasis, bezüglich derer f eine Darstellungsmatrix folgender Form bekommt:*

$$\begin{pmatrix} \pm 1 & \dots & 0 & & 0 & \dots & & 0 & \\ & \ddots & & & \vdots & & & \vdots & \\ 0 & \dots & \pm 1 & & 0 & \dots & & 0 & \\ 0 & \dots & 0 & \begin{bmatrix} \cos\phi_1 & -\sin\phi_1 \\ \sin\phi_1 & \cos\phi_1 \end{bmatrix} & & \dots & & 0 & \\ \vdots & 0 & \vdots & & & \ddots & & & \\ 0 & \dots & 0 & & 0 & \dots & \begin{bmatrix} \cos\phi_r & -\sin\phi_r \\ \sin\phi_r & \cos\phi_r \end{bmatrix} & & \end{pmatrix}.$$

Der Vektorraum zerfällt also in die direkte Summe paarweise orthogonaler Unterräume, die entweder Fixgeraden von f sind, oder Ebenen, die von f in sich gedreht werden. Ist n ungerade, so muß es wenigstens einen Eigenwert ± 1 geben.

Beispiel 8. Die orthogonale Matrix

$$\begin{pmatrix} \cos\phi & \sin\phi \\ \sin\phi & -\cos\phi \end{pmatrix}$$

hat die Determinante -1. Weil $\lambda\bar{\lambda} \geq 0$, kann sie keine imaginären Eigenwerte haben, die Eigenwerte sind also $+1$ und -1 (kann man natürlich auch direkt ausrechnen). Daher ist die Matrix ist diagonalisierbar!

Um einen Eigenvektor von $\lambda = 1$ zu bestimmen, muß man den Kern der Matrix

$$\begin{pmatrix} \cos\phi - 1 & \sin\phi \\ \sin\phi & -\cos\phi - 1 \end{pmatrix}$$

bestimmen.Der Fall $\sin\phi = 0$ ist trivial. Sei also $\sin\phi \neq 0$. Wir wissen, daß der Rang < 2 ist. Also brauchen wir nur die erste Zeile zu betrachten, die nach unserer Annahme jedenfalls nicht trivial ist. Wir sehen , daß

$$\begin{pmatrix} \sin\phi \\ 1 - \cos\phi \end{pmatrix} = \begin{pmatrix} \sin 2\frac{\phi}{2} \\ 1 - \cos 2\frac{\phi}{2} \end{pmatrix} = \begin{pmatrix} 2\sin\frac{\phi}{2}\cos\frac{\phi}{2} \\ 2\sin^2\frac{\phi}{2} \end{pmatrix} = 2\sin\frac{\phi}{2}\begin{pmatrix} \cos\frac{\phi}{2} \\ \sin\frac{\phi}{2} \end{pmatrix}.$$

ein Eigenvektor ist. Also ist auch

$$\begin{pmatrix} \cos\frac{\phi}{2} \\ \sin\frac{\phi}{2} \end{pmatrix}$$

ein Eigenvektor zu $\lambda = 1$. Die Matrix liefert eine Spiegeluung an der Geraden in dieser Richtung.

Im Gegensatz zu dieser Situation sind die Drehungen

$$\begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}$$

nur bei $\phi \in \pi\mathbb{Z}$ diagonalisierbar.

□

Beispiel 9. Beachte, daß für $A \in \mathrm{O}(n)$ stets $1 = \det(AA^*) = (\det A)^2$, also $\det A = \pm 1$. Die orientierungserhaltenden orthogonalen Matrizen, also die mit Determinante=+1, bilden eine Untergruppe von $\mathrm{O}(n)$, die sogenannte *Spezielle orthogonale Gruppe* $\mathrm{SO}(n)$. Sei $A \in \mathrm{SO}(3)$. Hat A einen nicht-reellen Eigenwert $\lambda_1 = \cos\phi + i\sin\phi$, so ist das konjugiert-komplexe $\lambda_2 = \cos\phi - i\sin\phi$ ein weiterer Eigenwert und wegen

$$1 = \det A = \lambda_1\lambda_2\lambda_3 = (\cos^2\phi + \sin^2\phi)\lambda_3$$

ist der dritte Eigenwert =+1. Der zugehörige Eigenraum ist eine (punktweise) feste Gerade (Achse), um die A um den Vektor ϕ dreht. Sind alle Eigenwerte von A reell, so sind sie $(1,1,1)$ und $A = E$ oder sie sind $(-1,-1,+1)$, d.h. A ist die Drehung um die Fixgerade zum Eigenwert $+1$ um 180^0. Die Matrizen aus SO(3) liefern also genau die Drehungen des $\mathbb{R}^3$ um eine feste Achse durch 0.

□

2.3 Polar- und Iwasawa-Zerlegung

Definition 8. Sei f ein selbstadjungierter Endomorphismus eines Euklidischen oder unitären Vektorraums $(V, \langle .,. \rangle)$. Gilt

$$\langle f(v), v\rangle > 0 \text{ für alle } v \in V\backslash\{0\},$$

so heißt *f positiv definit*. Gilt

$$\langle f(v), v\rangle \geq 0 \text{ für alle } v \in V,$$

so heißt *f positiv semidefinit*. Entsprechend definiert man *negativ (semi)definit*.

Satz 16 (Rayleigh-Quotient). *Sei f selbstadjungiert. Sei $(v_1, \ldots, v_n)$ eine ON-Basis aus Eigenvektoren von f und $f(v_i) = \lambda_i v_i$. Die λ_i sind reell, und wir nehmen an, daß*

$$\lambda_1 \leq \lambda_i \leq \lambda_n \text{ für alle } i.$$

Dann gilt für alle $v \neq 0$

$$\lambda_1 \leq \frac{\langle f(v), v\rangle}{\langle v, v\rangle} \leq \lambda_n. \tag{10}$$

Der Quotient in der Mitte heißt der Rayleigh-Quotient.
Für $v = v_1$ bzw. $v = v_n$ nimmt er offenbar sein Minimum λ_1 bzw. sein Maximum λ_n an. Insbesondere ist f positiv (semi-)definit genau dann, wenn alle Eigenwerte von f positiv (nicht negativ) sind.

Beweis. Sei $v = \sum x_i v_i$ dann ist

$$\langle f(v), v\rangle = \langle \sum x_i\lambda_i v_i, \sum x_j v_j\rangle = \sum \lambda_i x_i x_j \langle v_i, v_j\rangle = \sum \lambda_i x_i^2.$$

Daraus folgt

$$\langle f(v), v\rangle \leq \lambda_n \sum x_i^2 = \lambda_n \langle v, v\rangle,$$
$$\langle f(v), v\rangle \geq \lambda_1 \sum x_i^2 = \lambda_1 \langle v, v\rangle,$$

und damit (10). Schließlich ist $\lambda_1 > 0$ (oder ≥ 0) genau dann, wenn *alle* Eigenwerte positiv (≥ 0) sind. □

Satz 17 (Quadratwurzelsatz). *Sei f ein positiv semidefiniter Endomorphismus eines Euklidischen oder unitären Vektorraumes $(V, \langle .,. \rangle)$. Dann gibt es genau ein positiv semidefinites $g \in \mathrm{End}(V)$ mit*

$$f = g^2.$$

Wir nennen g die Wurzel aus f. Ist f positiv definit, so auch g. Weiter kommutieren f und g

$$fg = gf.$$

Beweis. Existenz. Sei $(v_1, \ldots, v_n)$ eine orthonormale Eigenbasis zu f und

$$f(v_i) = \lambda_i v_i, \qquad \lambda_{i \geq 0}.$$

Sei $g \in \mathrm{End}(V)$ definiert durch

$$g(v_i) = \sqrt{\lambda_i} v_i, \qquad i \in \{1, \ldots, n\}.$$

Dann ist offenbar $(v_1, \ldots, v_n)$ eine orthonormale Eigenbasis zu g und alle Eigenwerte von g sind reell ≥ 0. Also ist g selbstadjungiert, positiv semidefinit und $g^2(v_i) = \lambda_i v_i$ für alle i, also $g^2 = f$.

Einzigkeit. Sei $g \in \mathrm{End}(V)$ positiv semidefinit mit $g^2 = f$ und $v \in V$ mit $g(v) = \mu v$. Dann ist

$$f(v) = g^2(v) = \mu^2 v.$$

Also sind die verschiedenen Eigenwerte μ_i von g gerade die nicht-negativen Wurzeln aus den verschiedenen Eigenwerten λ_i von f und

$$\mathrm{Eig}(g, \mu_i) \subset \mathrm{Eig}(f, \lambda_i).$$

Wegen

$$V = \mathrm{Eig}(g, \mu_1) \oplus \ldots \oplus \mathrm{Eig}(g, \mu_k)$$

muß hier Gleichheit stehen, und damit ist g eindeutig festgelegt.

f ist genau dann positiv definit, wenn alle $\lambda_i > 0$. In diesem Fall ist aber auch g positiv definit.

Schließlich kommutieren f und g:

$$f \circ g = g^2 \circ g = g \circ g^2 = g \circ f.$$

□

Satz 18 (Polarzerlegung). *Sei V Euklidisch oder unitär und endlich dimensional. Jedes $f \in \mathrm{Aut}(V)$ ist das Produkt aus einem positiv definiten h und einem unitären (orthogonalen) Endomorphismus k*

$$f = hk.$$

Diese Zerlegung ist eindeutig.

Beweis. Vorbemerkung. Im Beweis brauchen wir die folgenden Identitäten:
Für $f, g \in \mathrm{End}(V)$ ist

$$(f^*)^* = f,$$
$$(fg)^* = g^* f^*,$$

und für $f, g \in \mathrm{Aut}(V)$ gilt

$$\begin{aligned}(fg)^{-1} &= g^{-1}f^{-1},\\ f^* &\in \mathrm{Aut}(V),\\ (f^*)^{-1} &= (f^{-1})^*.\end{aligned}$$

Wir beweisen nur die letzte:

$$\begin{aligned}(f^*)^{-1} = (f^{-1})^* &\iff \forall_{v,w}\langle v, (f^*)^{-1}(w)\rangle = \langle v, (f^{-1})^*(w)\rangle = \langle f^{-1}(v), w\rangle\\ &\iff \forall_{v,w}\langle f(v), (f^*)^{-1}(w)\rangle = \langle f^{-1}(f(v)), w\rangle = \langle v, w\rangle.\end{aligned}$$

Aber

$$\langle f(v), (f^*)^{-1}(w)\rangle = \langle v, f^*(f^*)^{-1}(w)\rangle = \langle v, w\rangle.$$

Existenz von h und k. Wir betrachten $H := ff^*$. Dafür gilt

$$H^* = (ff^*)^* = f^{**}f^* = ff^* = H.$$

und

$$\langle H(v), v\rangle = \langle ff^*(v), v\rangle = \langle f^*(v), f^*(v)\rangle > 0$$

für $v \neq 0$. Also ist H selbstadjungiert und positiv definit. Wir bezeichnen mit h die Wurzel aus H. Dann gilt für $k := h^{-1}f$

$$\begin{aligned}\langle k(v), k(v)\rangle &= \langle h^{-1}f(v), h^{-1}f(v)\rangle\\ &= \langle f(v), (h^{-1})^*h^{-1}f(v)\rangle\\ &= \langle f(v), (h^*)^{-1}h^{-1}f(v)\rangle\\ &= \langle f(v), (h^2)^{-1}f(v)\rangle\\ &= \langle f(v), (H)^{-1}f(v)\rangle\\ &= \langle f(v), (ff^*)^{-1}f(v)\rangle\\ &= \langle f(v), (f^*)^{-1}f^{-1}f(v)\rangle\\ &= \langle f(v), (f^*)^{-1}(v)\rangle\\ &= \langle f(v), (f^{-1})^*(v)\rangle\\ &= \langle f^{-1}f(v), v\rangle\\ &= \langle v, v\rangle.\end{aligned}$$

Also ist k unitär (orthogonal), h positiv definit und

$$f = hk.$$

Eindeutigkeit. Aus $f = hk$ folgt

$$ff^* = (hk)(hk)^* = h\underbrace{kk^*}_{=\mathrm{id}}h^* = h^2.$$

Aber diese Gleichung bestimmt nach dem Wurzelsatz h eindeutig, und damit ist auch k eindeutig. □

Wir formulieren diesen Satz für Matrizen. Dazu bezeichnen wir für ein festes n und $\mathbb{R}$ oder $\mathbb{C}$ als Körper

G die Gruppe $\mathrm{GL}(n, \mathbb{R})$ bzw. $\mathrm{GL}(n, \mathbb{C})$ der invertierbaren Matrizen,

P die Menge der positiv definiten selbstadjungierten (n, n)-Matrizen über $\mathbb{R}$ bzw. $\mathbb{C}$,

K die Gruppe $O(n)$ bzw. $U(n)$,

A die Gruppe der diagonalen (n, n)-Matrizen mit positiven reellen Diagonalelementen,

Ndie Gruppe der oberen (n, n)-Dreiecksmatrizen mit Einsen auf der Diagonalen.

K erinnert an „kompakt“ und A an „abelsch“; später werden Sie vielleicht lernen, warum.

Satz 19 (Polarzerlegung von Matrizen). *Es gilt mit den vorstehenden Bezeichnungen*

$$G = PK.$$

Genau soll das bedeuten: Jedes $f \in G$ läßt sich eindeutig schreiben als $f = hk$ mit $h \in P$ und k in K.

Der Beweis ist klar, wenn man (n, n)-Matrizen als Endomorphismen des $\mathbb{R}^n$ oder $\mathbb{C}^n$ interpretiert.

Beispiel 10. Für

$$G = \mathrm{GL}(1, \mathbb{C}) = \{(z) \,|\, 0 \neq z \in \mathbb{C}\}$$

ist

$$P = \{(r) \,|\, r \in \mathbb{R}, r > 0\},$$
$$K = \{(b) \,|\, b \in \mathbb{C}, |b| = 1\} = \{(e^{i\phi}) \,|\, \phi \in \mathbb{R}\}.$$

Der Satz sagt dann nichts anderes, als daß man die komplexe Zahl $z \neq 0$ eindeutig schreiben kann in der Form

$$z = re^{i\phi}$$

mit $r > 0$ und $\phi \in [0, 2\pi[$. Das ist die Darstellung in Polarkoordinaten, und daher kommt der Name *Polarzerlegung.*

□

Einen weiteren Zerlegungssatz wollen wir bei der Gelegenheit auch formulieren, obwohl er nichts mit der Eigenwerttheorie zu tun hat.

Satz 20 (QR-Zerlegung von Matrizen, Iwasawa-Zerlegung). *Im Sinne des letzen Satzes gilt*

$$G = KAN.$$

Beweis. Zur Existenz. Die Existenz folgt aus dem leicht modifizierten Orthonormalisierungsverfahren von Gram-Schmidt. Ein System von n linear unabhängige Vektoren $v_1, \ldots, v_n$ wird zunächst orthogonalisiert, indem man von v_k eine geeignete Linearkombination von $v_1, \ldots, v_{k-1}$ abzieht. (Beim Originalverfahren wird jeder „orthogonalisierte Vektor“ gleich normiert und eine Linearkombination dieser neuen Vektoren abgezogen. Aber weil der Spann von $(v_1, \ldots, v_{k-1})$ gleich dem der neuen Vektoren $(w_1, \ldots, w_{k-1})$ ist, kann man wie oben modifizieren. Der Nachteil ist, daß man die nötigen Koeffizienten nicht einfach als Skalarprodukte hinschreiben kann, aber das stört uns im Augenblick nicht.) Man erhält orthogonale Vektoren $\tilde{w_1}, \ldots, \tilde{w_n}$. Und durch Division mit der jeweiligen Norm erhält man ein Orthonormalsystem $w_1, \ldots, w_n$. Schreibt man die Vektoren als Spalten in eine Matrix so bedeutet das

$$(\tilde{w}_1, \ldots, \tilde{w}_n) = (v_1, \ldots, v_n) \begin{pmatrix} 1 & \ldots & * \\ & \ddots & \vdots \\ 0 & & 1 \end{pmatrix}$$
$$(w_1, \ldots, w_n) = (v_1, \ldots, v_n) \begin{pmatrix} 1 & \ldots & * \\ & \ddots & \vdots \\ 0 & & 1 \end{pmatrix} \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix}.$$

Durch Rechtsmultiplikation mit den Inversen der beiden letzten Matrizen ergibt sich die Behauptung. Beachte, daß P und A Gruppen sind: die Inversen sind wieder von derselben Gestalt, die Inverse einer oberen Dreiecksmatrix (mit Einsen auf der Diagonalen) ist wieder eine solche, die Inverse einer Diagonalmatrix ist wieder diagonal.

Zur Eindeutigkeit. Seien $k_1, k_2 nK$, $a_1, a_2 \in A$ und $n_1, n_2 \in N$. Aus $k_1 a_1 n_1 = k_2 a_2 n_2$ folgt

$$k_2^{-1} k_1 = k_2 a_2 n_2 n_1^{-1} a_1^{-1}.$$

Die linke Seite ist orthogonal bzw. unitär, die rechte eine obere Dreiecksmatrix mit reellen positiven Diagonalelementen:

$$\begin{pmatrix} d_1 & \dots & * \\ & \ddots & \vdots \\ 0 & & d_n \end{pmatrix}.$$

Weil diese Matrix orthogonal bzw. unitär ist, ist die erste Spalte ein Einheitsvektor, also $d_1 = 1$, und damit $d_{1j} = 0$ für $j > 1$. Dann ist aber die zweite Spalte der zweite Einheitvektor usw. Die rechte Seite ist also die Einheitsmatrix E und daher $k_1 = k_2$. Mit einem ähnlichen Argument folgt aus $a_1 n_1 = a_2 n_2$, also aus $a_2^{-1} a_1 = n_2 n_1^{-1}$, daß $a_1 = a_2$ und daher $n_1 = n_2$ ist. □

Bemerkung. Dieser Satz ist von großer theoretischer und praktischer Bedeutung.

Die theoretische Bedeutung gehört in den Bereich der Liegruppen und dort nennt man die Zerlegung Iwasawa-Zerlegung. Sie liefert insbesondere Einsichten in die topologische Struktur der Gruppe G.

Die praktische Bedeutung lernen Sie in der Praktischen Mathematik kennen. Sie beruht auf der Existenz von numerisch stabilen Algorithmen zur Herstellung dieser Zerlegung (Householder 1964). Hat man ein lineares Gleichungssystem

$$Ax = b$$

mit invertierbarem A, und hat man A geschrieben als $A = QR$ mit unitärem Q und oberer Dreiecksmatrix R, so ist das System wegen $Q^{-1} = Q^*$ (leicht zu bilden!) äquivalent zu

$$Rx = Q^* b.$$

Das ist aber in Zeilenstufenform, und die (eindeutige) Lösung läßt sich einfach rekursiv bestimmen. Die Bezeichnungswahl $A = QR$ mit „R“ für „rechte Dreiecksmatrix“ ist in der numerischen Mathematik üblich und führt zu dem Namen „QR-Zerlegung“.

Beispiel 11. Wendet man Satz 18 auf f^* an, so folgt $f^* = h_1 k_1$ und $f = k_1^* h_1^* = k_1^{-1} h_1$. Also kann man Satz 19 auf schreiben als

$$G = KP.$$

Vergleicht man das mit

$$G = KAN,$$

so haben die entsprechenden Zerlegungen kaum etwas miteinander zu tun. Es gilt zum Beispiel

$$\begin{aligned} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} &= \begin{pmatrix} \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} \frac{3}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \sqrt{2} & 0 \\ 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

□

3 Intermezzo: Ein Kapitel Algebra

Für das Studium von Endomorphismen ist das charakteristische Polynom von zentraler Bedeutung. Insbesondere ist spielt seine Zerlegung in „einfache Bausteine“ eine Rolle. Wir schalten deshalb hier einen Abschnitt über Polynomalgebren ein.

Definition 9 (Algebra). Eine *K-Algebra* ist ein K-Vektorraum $\mathcal{A}$ zusammen mit einer *Multiplikation*

$$\mathcal{A} \times \mathcal{A} \to \mathcal{A},$$

die folgende drei Axiome erfüllt:

(A1) Für alle $a, b, c \in \mathcal{A}$ ist

$$a(b + c) = ab + ac \text{ und } (a + b)c = ac + bc.$$

(A2) Für alle $a, b \in \mathcal{A}$ und $\lambda \in K$ ist

$$\lambda(ab) = (\lambda a)b = a(\lambda b)$$

(A3) Für alle $a, b, c \in \mathcal{A}$ ist

$$a(bc) = (ab)c$$

Gibt es in $\mathcal{A}$ ein Element 1 mit

(A4) Für alle $a \in \mathcal{A}$ ist

$$1a = a1 = a,$$

so heißt $\mathcal{A}$ eine *unitäre* Algebra. 1 ist dann eindeutig bestimmt und heißt das *Einselement* von $\mathcal{A}$.

$\mathcal{A}$ heißt *kommutativ*, wenn gilt:

(A5) Für alle $a, b \in \mathcal{A}$ ist

$$ab = ba,$$

Beispiel 12. $M(n, n; K)$ bzw $\mathrm{End}(V)$ sind auf offensichtliche Weise unitäre aber im allgemeinen nicht kommutative K-Algebren. Der Körper K selbst ist eine unitäre kommutative K-Algebra.

□

Beispiel 13.

$$\mathcal{A} := \{f : \mathbb{R} \to \mathbb{R} \text{ Abbildung} \mid \exists_{r>0}\, f|\mathbb{R}\backslash[-r, r] = 0\}$$

ist mit den offensichtlichen Rechenoperationen eine kommutative Algebra, aber sie ist nicht unitär.

□

Das Beispiel, welches uns am meisten interessiert, gibt die folgende Definition. Wir erinnern, daß $\mathbb{N} = \{0, 1, 2, \ldots\}$.

Definition 10 (Polynomalgebra). Die Menge

$$K[X] := \{F : \mathbb{N} \to K, i \mapsto F_i \mid F_i = 0 \text{ für alle bis auf endlich viele } i \in \mathbb{N}\}$$

ist mit der wertweisen Addition und Skalarmultiplikation ein K-Vektorraum. Man hat auch eine wertweise Multiplikation der Elemente von $K[X]$ miteinander, aber wir definieren ein anderes Produkt, das sogenannte *Cauchyprodukt*:
Für $F, G \in K[X]$ sei

$$(F \odot G)_i := \sum_{j=0}^{i} F_j G_{i-j}.$$

damit wird $K[X]$ eine unitäre, kommutative K-Algebra, die *Polynomalgebra (in einer Unbestimmten) über K*.

Das Einselement ist gegeben durch

$$1_i := \delta_{0i} = \begin{cases} 1 & \text{für } i = 0, \\ 0 & \text{sonst.} \end{cases}$$

Die Elemente von $K[X]$ heißen *Polynome* über K.

Der Nachweis der Algebra-Axiome ist eine leichte Übung.

Definition 11. Für $F \in K[X]$ heißt

$$\operatorname{Grad} F = \sup\{i \in \mathbb{N} \mid F_i \neq 0\}$$

der *Grad* von F. Nach der üblichen Konvention ist der Grad von 0 dann $-\infty$.
Polynome vom Grad < 1 nennt man auch *skalare Polynome*.

Lemma 10 (Rechenregeln für den Grad). *Für $F, G \in K[X]$ gilt*

(i) $\operatorname{Grad}(F \odot G) = \operatorname{Grad} F + \operatorname{Grad} G$.

(ii) $\operatorname{Grad}(F + G) \leq \max\{\operatorname{Grad} F, \operatorname{Grad} G\}$.

Beweis. *Zu (i).* Ist $F = 0$ oder $G = 0$, so ist auch $F \odot G = 0$ und auf beiden Seiten steht $-\infty$.

Ist $\operatorname{Grad} F = m > 0$ und $\operatorname{Grad} G = n > 0$, so ist für $N \geq m + n$

$$(F \odot G)_N = \sum_{k=0}^{N} F_k G_{N-k}.$$

aber $F_k = 0$ für $k > m$ und $G_{N-k} = 0$ für $N - k > n$. Für $N = m + n$ findet man

$$(F \odot G)_{m+n} = F_m G_n \neq 0.$$

Für $N > m + n$ folgt aus $N - k \leq n$, daß $k > m$, also ist $(F \odot G)_N = 0$.

Zu (ii). Klar □

Bemerkung. Ein „klassisches“ Polynom

$$a_0 + a_1 X + \ldots + a_n X^n \tag{11}$$

ist natürlich eindeutig bestimmt durch die Koeffizienten $(a_0, \ldots, a_n)$, die man durch Nullen zu einer unendlichen Folge $(a_0, \ldots, a_n, 0, \ldots)$ fortsetzen kann. Das ist gerade die Wertetabelle einer Funktion F aus der Definition. Die Addition von Polynomen und die Multiplikation

mit Skalaren entspricht einfach der entsprechenden Operation für die Koeffizientenfolge. Und die Multiplikation

$$(a_0 + a_1X + \ldots + a_nX^n)(b_0 + b_1X + \ldots + b_mX^m) = \sum_{k=0}^{m+n} \left(\sum_{j=0}^{k} a_j b_{k-j} \right) X^k,$$

entspricht gerade dem Cauchyprodukt. Unsere Definition liefert also Polynome als „formale“ Ausdrücke der Form (11). Sie erzwingt, daß zum Beispiel für $K = \mathbb{Z}_2$ die Polynome $1+X+X^2$ und 1 voneinander verschieden sind. Das folgende Lemma erlaubt es, Polynome wieder in der vertrauten Form zu schreiben. Das ist aber eben nur „formal“. Auf den Zusammenhang mit Polynomfunktionen kommen wir später noch einmal zurück.

Lemma 11. *Wir definieren $X \in K[X]$ (selbes X) durch*

$$X_i := \delta_{1i} = \begin{cases} 1 & \text{für } i = 1, \\ 0 & \text{sonst.} \end{cases}$$

Wir setzen $X^0 = 1 \in K[X]$ und

$$X^n := \underbrace{X \odot \ldots \odot X}_{n \text{ Faktoren}}$$

für $n > 0$. Dann gilt

$$(X^n)_i = \delta_{ni} = \begin{cases} 1 & \text{für } i = n, \\ 0 & \text{sonst.} \end{cases} \tag{12}$$

Daher ist für $F \in K[X]$ vom Grad n

$$F = \sum_{k=0}^{n} F_k X^k. \tag{13}$$

Wenn wir die „Unbestimmte“ X also so interpretieren, können wir die abstrakt definierten Polynome wieder in der gewohnten Form schreiben.

Beweis des Lemmas. Wir zeigen (12) durch vollständige Induktion über n. Für $n = 0$ und $n = 1$ ist die Behauptung klar.

$\underline{n \to n+1.}$ Es ist

$$(X^{n+1})_i = (X^n \odot X)_i = \sum_{j=0}^{i} \delta_{nj} \delta_{1\ i-j}.$$

Die Summanden sind immer 0, ausgenommen $n = j$ und $1 = i - j = i - n$. Dieser Fall tritt nur auf für $i = n + 1$ und dann ist die Summe $= 1$, andernfalls aber $= 0$.

$\underline{\text{Zu (13).}}$ Nach (12) ist für alle i

$$\left(\sum_{k=0}^{n} F_k X^k \right)_i = \sum_{k=0}^{n} F_k X_i^k = F_i.$$

Dabei haben wir benutzt, daß Summe und *Skalar*multiplikation in $K[X]$ wertweise definiert sind. □

Konvention. Wir schreiben in Zukunft für das Cauchyprodukt von Polynomen einfach FG oder $F(X)G(X)$ statt $F \odot G$.

Unser Ziel ist der nachstehende Zerlegungssatz für Polynome, aber dafür brauchen wir noch einen Begriff:

Definition 12. Ein Polynom $P \in K[X]$ heißt ein *Primpolynom*, wenn

$$\operatorname{Grad} P > 0$$

und P nicht das Produkt zweier Polynome $G, H \in K[X]$ von positivem Grad ist, wenn also gilt

$$\forall_{G,H \in K[X]} \, (P = GH \implies \operatorname{Grad} G = 0 \text{ oder } \operatorname{Grad} H = 0).$$

Weiter heißt ein Polynom *normiert*, wenn der höchste nicht verschwindende Koeffizient $= 1$ ist.

Satz 21 (Primfaktorzerlegung). *Sei $F \in K[X]$ ein normiertes Polynom positiven Grades. Dann gibt es bis auf die Numerierung eindeutig bestimmte normierte Primpolynome $P_1, \ldots, P_k \in K[X]$ mit*

$$F = P_1 \ldots P_k.$$

Ist F nicht normiert und F_n der höchste nicht verschwindende Koeffizient, so hat man eine Zerlegung

$$F = F_n \, P_1 \ldots P_k.$$

Bemerkung. Skalare Faktoren kann man natürlich in einem Produkt von Polynomen beliebig hin und her schieben. Das wird im Satz durch die Forderung der Normiertheit verhindert. Abgesehen davon hat man also das Polynom F in kleinste Einzelteile zerlegt: Die Faktoren lassen sich nicht weiter aufspalten.

Beispiel 14. Für $a \in K$ ist $F(X) = X - a$ ein Primpolynom, wie sofort aus der Gradformel folgt.

□

Beispiel 15. Für $K = \mathbb{R}$ ist $F(X) = X^2 + 1$ ein Primpolynom. Gäbe es nämlich eine nicht-triviale Zerlegung, so müßten die Faktoren von Grad 1 sein:

$$X^2 + 1 = (aX + b)(uX + v) = auX^2 + (av + bu)X + bv.$$

Also

$$\begin{aligned} au &= 1 \\ av + bu &= 0 \\ bv &= 1 \end{aligned}$$

Multipliziert man die mittlere Gleichung mit uv und setzt die beiden anderen ein, so folgt $u^2 + v^2 = 0$, also (K=ℝ!) $u = v = 0$. Widerspruch!

□

Existenzbeweis zum Satz 21. Das ist einfach. Wir benutzen vollständige Induktion über $n := \operatorname{Grad} F$.

$\underline{n = 1.}$ Dann ist $F(X) = F_0 + X$. Nach Lemma 10 kann das nicht das Produkt zweier Polynome von positivem Grad sein. Also ist $P_1 := F$ ein normiertes Primpolynom.

$\underline{n \to (n+1).}$ Der Satz sei schon bewiesen für alle Polynome vom Grad $\leq n$ und F sei normiert vom Grad $n + 1$.

1. Fall: F ist ein Primpolynom. Dann ist nichts zu zeigen.

2. Fall: F ist kein Primpolynom. Dann gibt es also Polynome G vom Grad $r > 0$ und H vom Grad $s > 0$ mit $F = GH$. Insbesondere ist $r + s = n + 1$ und

$$G_r H_s = F_{n+1} = 1.$$

Wir setzen

$$\tilde{G}(X) := \frac{1}{G_r} G(X), \quad \tilde{H}(X) := \frac{1}{H_s} H(X).$$

Das sind dann normierte Polynome vom Grad $r \leq n$ bzw. $s < N$ mit $\tilde{G}\tilde{H} = F$. Nach Induktionsvoraussetzung sind sie Produkte normierter Primpolynome. Also gilt dasselbe für F. □

Die Eindeutigkeit der Zerlegung ist wichtig und viel schwieriger zu sehen. Ein wesentliches Hilfsmittel zum Beweis ist der Divisionsalgorithmus für Polynome, den Sie im reellen Fall schon aus der Schule kennen, und den wir jetzt beschreiben. Zur Vorbereitung dient das folgende

Lemma 12. *Seien $F, D \in K[X]$ zwei Polynome mit*

$$0 \leq \operatorname{Grad} D \leq \operatorname{Grad} F.$$

Dann gibt es ein Polynom $G \in K[X]$ mit

$$\operatorname{Grad}(F - DG) < \operatorname{Grad} F.$$

Beweis. Seien

$$F(X) = \sum_{k=0}^{m-1} F_k X^k + F_m X^m, \quad F_m \neq 0,$$
$$D(X) = \sum_{k=0}^{n-1} D_k X^k + D_n X^n, \quad D_n \neq 0.$$

Nach Voraussetzung ist $m \geq n$. Wir setzen

$$G := \frac{F_m}{D_n} X^{m-n}$$

und erhalten

$$F - DG = \sum_{k=0}^{m-1} F_k X^k - \sum_{k=0}^{n-1} \frac{F_m}{D_n} D_k X^{m-n+k}.$$

Die rechte Seite hat offenbar Grad $< m$. □

Satz 22 (Euklidischer Algorithmus). *Seien $F, D \in K[X]$ und $D \neq 0$. Dann gibt es eindeutig bestimmte $Q, R \in K[X]$ mit*

$$F = DQ + R,$$
$$\operatorname{Grad} R < \operatorname{Grad} D.$$

Beweis. Einzigkeit. Sei $F = DQ + R = D\tilde{Q} + \tilde{R}$ mit $\operatorname{Grad} R < \operatorname{Grad} D, \operatorname{Grad} \tilde{R} < \operatorname{Grad} D$. Dann hat man

$$R - \tilde{R} = D(\tilde{Q} - Q),$$

also

$$\underbrace{\operatorname{Grad}(R-\tilde{R})}_{<\operatorname{Grad} D} = \operatorname{Grad} D + \operatorname{Grad}(\tilde{Q}-Q).$$

Das geht nur, wenn $\operatorname{Grad}(\tilde{Q}-Q) = -\infty$, also $\tilde{Q} = Q$ und dann auch $\tilde{R} = R$.

Existenz.

1. Fall. $\operatorname{Grad} F < \operatorname{Grad} D$. In diesem Fall ist

$$f = D\,0 + F.$$

2. Fall. $\operatorname{Grad} F \geq \operatorname{Grad} D$. Dann benutzen wir vollständige Induktion über $n := \operatorname{Grad} F$.

$n = 0$. Dann ist $\operatorname{Grad} D = 0$, also $D \in K\backslash\{0\}$, und

$$F = D\left(\frac{1}{D}F\right) + 0.$$

$n \to (n+1)$. Sei die Behauptung für Polynome vom Grad $\leq n$ bereits bewiesen und sei $\operatorname{Grad} F = n+1$. Nach dem Lemma 12 gibt es dann $G \in K[X]$ mit

$$\operatorname{Grad}(F - DG) < n+1.$$

Nach Induktionsvoraussetzung gibt es also $\tilde{Q}, R$ mit

$$F - DG = D\tilde{Q} + R, \quad \operatorname{Grad}(r) < \operatorname{Grad} D.$$

Also ist

$$F = D\underbrace{(\tilde{Q}+G)}_{=:Q} + R, \quad \operatorname{Grad}(R) < \operatorname{Grad} D.$$

□

Definition 13. Seien $F, G \in K[X]$ und $G \neq 0$. Wir nennen G einen *Teiler von* F und sagen G *teilt* F, wenn es ein $H \in K[X]$ gibt, für das $F = GH$.

Notation:

$$G|F.$$

Satz 23 (ggT-Algorithmus). *Seien $F, G \in K[X]$ und*

$$0 \leq \operatorname{Grad} G \leq \operatorname{Grad} F.$$

Definiere (endliche) rekursive Folgen Q_k, R_k mit Hilfe des Euklidischen Algorithmus durch

$$R_{-1} := F, \quad R_0 := G,$$
$$R_k =: R_{k+1}Q_k + R_{k+2}.$$

Weil $\operatorname{Grad} R_{k+2} < \operatorname{Grad} R_{k+1}$, wird der Rest R_k schließlich $= 0$ und die Konstruktion bricht ab. Sei $R_n \in K[X]$ der letzte von 0 verschiedene Rest. Dann gilt

(i)

$$R_n|F \text{ und } R_n|G,$$

und für alle $T \in K[X]$ gilt

$$T|F \text{ und } T|G \implies T|R_n.$$

(ii) Es gibt $A, B \in K[X]$ mit

$$R_n = AF + BG.$$

Beweis. Zu (i). Wir haben

$$\begin{aligned} F = R_{-1} &= R_0Q_0 + R_1 \\ G = R_0 &= R_1Q_1 + R_2 \\ &\dots \\ R_{n-3} &= R_{n-2}Q_{n-2} + R_{n-1} \\ R_{n-2} &= R_{n-1}Q_{n-1} + R_n \\ R_{n-1} &= R_nQ_n \end{aligned}$$

Betrachtet man dies System „von unten“, so sieht man,

$$R_n|R_{n-1},\ R_n|R_{n-2},\ \dots,\ R_n|R_0 = G,\ R_n|R_{-1} = F.$$

Aus $T|F = R_{-1}$ und $T|G = R_0$ folgt mit der ersten Gleichung $T|R_1$ usw. und schließlich $T|R_n$.

Zu (ii). Jeder Rest ist offenbar „Linearkombination“ der beiden vorangehenden mit Koeffizienten in $K[X]$. Daraus folgt (ii). □

Lemma 13. *Sind $F, G \in K[X]\backslash\{0\}$, so gibt es genau ein normiertes $D \in K[X]\backslash\{0\}$ mit folgenden Eigenschaften:*

(i) $D|F$ und $D|G$

(ii) Für alle $T \in K[X]$ gilt

$$T|F \wedge T|G \implies T|D.$$

Beweis. Existenz. Sei $R_n(X) = \sum_{k=0}^{m} R_{nk}X^k$ zu F und G bestimmt wir im vorstehenden Satz mit höchstem Koeffizienten $R_{nm} \neq 0$. Sei

$$D := \frac{1}{R_{nm}} \sum_{k=0}^{m} R_{nk}X^k$$

die „Normierung“ von R_n. Dann leistet D das Gewünschte.

Eindeutigkeit. Haben D und $\tilde{D}$ beide diese Eigenschaften, so gilt $D|\tilde{D}$ und $\tilde{D}|D$. Also ist $\operatorname{Grad} D = \operatorname{Grad} \tilde{D}$ und weil beide normiert sind, ist $\operatorname{Grad}(D - \tilde{D}) < \operatorname{Grad} D$. Aus

$$D|D - \tilde{D}$$

folgt dann aber $D - \tilde{D} = 0$. □

Definition 14. Seien $F, G \in K[X]\backslash\{0\}$ und D dazu wie im Lemma. D heißt *er größte gemeinsame Teiler von F und G*. Notation:

$$D = ggT(F, G) \text{ oder } D = (F, G).$$

F und G heißen *teilerfremd*, wenn $ggT(F, G) = 1$.

Satz 24. *Seien $F, G, P \in K[X]$ und P ein Primpolynom. Dann gilt*

$$P|FG \implies P|F \text{ oder } P|G.$$

Beweis. Sei $D = ggT(P, F)$. Weil P ein Primpolynom ist, folgt aus $D|P$, daß

$$P = \lambda D \text{ mit einem } \lambda \in K\backslash\{0\}$$
$$\text{oder}$$
$$D = 1.$$

Im ersten Fall folgt aus $D|F$ auch $P|F$.

Im zweiten Fall gibt es $A, B \in K[X]$ mit

$$1 = AP + BF$$

und daher

$$G = APG + BFG.$$

Weil aber $P|APG$ und $P|BFG$ folgt $P|G$. □

Durch Induktion folgt daraus sofort:

Korollar 1. *Seien $F_1, \ldots, F_n, P \in K[X]$ und P ein Primpolynom. Dann gilt*

$$P|F_1 \cdot \ldots \cdot F_n \implies \exists_j\, P|F_j.$$

Lemma 14. *Sind P, Q* normierte *Primpolynome, so gilt*

$$P|Q \implies P = Q.$$

Beweis. Wegen $P|Q$ gibt es $H \in K[X]$ mit $Q = PH$. Weil Q ein Primpolynom ist, ist P oder H skalar, und weil P als Primpolynom nicht skalar ist, ist $H \in K$. Weil P und Q normiert sind, folgt aus $Q = HP$ aber $H = 1$, also $Q = P$. □

Damit kommen wir zum noch ausstehenden Beweis der Eindeutigkeit in der Primfaktorzerlegung für Polynome.

Eindeutigkeitsbeweis zum Satz 21. Wir müssen zeigen: Sind

$$P_1, \ldots, P_m, Q_1, \ldots, Q_n \in K[X]$$

normierte Primpolynome mit

$$P_1 \ldots P_m = Q_1 \ldots Q_n,$$

so ist $m = n$ und nach eventueller Umnumerierung gilt

$$(P_1, \ldots, P_m) = (Q_1, \ldots, Q_m).$$

Wir zeigen das durch Induktion über $m + n$. Der Fall $m + n = 2$, also $m = n = 1$ ist klar. Aus

$$P_1 \ldots P_m = Q_1 \ldots Q_n,$$

folgt aber $P_m | Q_1 \ldots Q_n$. Nach dem Korollar folgt

$$P_m | Q_j \text{ für ein } j \in \{1, \ldots, n\}.$$

Sei o.E. $P_m | Q_n$. Nach dem Lemma ist dann

$$P_m = Q_n,$$

also

$$P_1 \ldots P_{m-1} = Q_1 \ldots Q_{n-1}.$$

Nach Induktionsvoraussetzung folgt $m - 1 = n - 1$ und die $P's$ und $Q's$ sind gleich bis auf Numerierung. □

Eine wichtige Konsequenz aus der Eindeutigkeit ist das

Korollar 2. *Sind $F, G \in K[X]$ mit $G|F$ und ist*

$$F = a\, P_1 \ldots P_k, \quad a \in K$$

mit normierten P_i die Primfaktorzerlegung von F, so gibt es $b \in K$ und eine Teilmenge $\{i_1, \ldots, i_m\} \subset \{1, \ldots, k\}$, so daß

$$G = b\, P_{i_1} \ldots P_{i_m}.$$

Beweis. Sei $F = GH$ und $G = b\, Q_1 \ldots Q_m$ bzw. $H = c\, R_1 \ldots R_j$ die Primpolynomzerlegungen von G und H, so folgt

$$a\, P_1 \ldots P_k = bc\, Q_1 \ldots Q_m R_1 \ldots R_j$$

Aus der Eindeutigkeit folgt, daß die $Q's$ eine Teilmenge der $P's$ sind. □

Wir betrachten noch einen weiteren Begriff aus der Algebra:

Definition 15 (Ideal). Sei $\mathcal{A}$ eine kommutative unitäre K-Algebra. $J \subset \mathcal{A}$ heißt ein *Ideal*, wenn es abgeschlossen gegenüber Addition und Skalarmultiplikation ist, und wenn gilt

$$\forall_{a \in J} \forall_{b \in \mathcal{A}}\ ab \in J.$$

Die Teilmenge J ist also multiplikativ „absorbierend“.

Beispiel 16. Ist $\mathcal{A}$ eine kommutative unitäre K-Algebra und $a \in \mathcal{A}$, so ist

$$(a) := \{ab \,|\, b \in \mathcal{A}\}$$

offenbar ein Ideal, das von a erzeugte *Hauptideal*.

□

Satz 25. *Sei $J \neq \{0\}$ ein Ideal in $K[X]$ und D das normierte Polynom kleinsten Grades in J. Dann gilt*

$$J = (D).$$

Jedes Ideal in $K[X]$ ist also ein Hauptideal. Man nennt $K[X]$ auch eine Hauptidealalgebra.

Beweis. Sei $F \in J$. Dann gibt es $Q, R \in K[X]$ mit

$$F = DQ + R, \quad \operatorname{Grad} R < \operatorname{Grad} D$$

Aber $R = F - DG \in J$ und nach Wahl von D folgt $R = 0$. Also ist jedes $F \in J$ ein Vielfaches von D. Umgekehrt liegen die DG natürlich in J, weil J ein Ideal ist. □

Wir kommen nun zum Begriff der Polynomfunktionen zurück, den wir aber gleich etwas erweitern.

Satz 26 (Einsetzung). *Sei $\mathcal{A}$ eine unitäre K-Algebra. Dann definieren wir für $a \in \mathcal{A}$*

$$\eta_a : K[X] \to \mathcal{A}$$

wie folgt: Für $F(X) = \sum_{k=0}^{n} F_k X^k$ sei

$$\eta_a(F) := F(a) := \sum_{k=0}^{n} F_k a^k.$$

Beachten Sie:

- *F ist von Hause aus eine Abbildung $F : \mathbb{N} \to K$, so daß $F(a)$ bisher nicht definiert war. $F(X)$ war nur eine andere Schreibweise für F.*
- *Die Rechenoperationen auf der rechten Seite (Skalarmultiplikation, Potenz und Summe) spielen sich in $\mathcal{A}$ ab.*

Dann ist η_a ein Homomorphismus von K-Algebren:

$$\eta_a(F + G) = \eta_a(F) + \eta_a(G), \quad \eta_a(F \odot G) = \eta_a(F)\eta_a(G),$$
$$\eta_a(cF) = c\eta_a(F) \text{ für } c \in K.$$

Daher liefert die Abbildung

$$K[X] \to \mathcal{A}^{\mathcal{A}}, F(X) \mapsto \eta_{\cdot}(F)$$

ebenfalls einen Homomorphismus von Algebren, der jedem Polynom eine Selbstabbildung von $\mathcal{A}$ zuordnet. Für $\mathcal{A} = K$ ist $\eta_{\cdot}(F)$ die zu dem Polynom gehörende Polynomfunktion.

Beweis. Leicht. □

Definition 16 (Nullstellen, Minimalpolynom). Sei $F \in K[X]$ ein Polynom und $\mathcal{A}$ eine unitäre K-Algebra (z.B. $\mathcal{A} = K$).

(i) Wir nennen $a \in \mathcal{A}$ eine *Nullstelle* oder *Wurzel* von F, falls

$$F(a) = \eta_a(F) = 0.$$

(ii) Für $a \in \mathcal{A}$ ist die Menge

$$N_a := \{F \in K[X] \mid F(a) = 0\}$$

ein Ideal in $K[X]$, das *Nullstellenideal* an der Stelle a.

(iii) Ist $N_a \neq \{0\}$, so gibt es nach Satz 25 ein eindeutig bestimmtes normiertes Polynom $\mu_a(X) \in K[X]$ mit

$$N_a = (\mu_a).$$

$\mu_a(X)$ heißt das *Minimalpolynom von* a. Jedes Polynom mit Nullstelle a hat also $\mu_a(X)$ als Teiler.

Wir betrachten dazu zwei Beispiele.

Beispiel 17. Seien $\mathcal{A} = K$ und $a \in K$. Dann ist $X - a$ offenbar ein normiertes Polynom kleinsten positiven Grades mit a als Nullstelle. Also ist

$$\mu_a(X) = X - a.$$

Für jedes $F \in K[X]$ gilt dann

$$F(a) = 0 \iff (X - a) | F(X).$$

Dabei ist $\Rightarrow$ die Tatsache, daß das Minimalpolynom das Nullstellenideal erzeugt, während $\Leftarrow$ eine triviale Folge der Definition des Teilers ist.

□

Beispiel 18. Im Gegensatz dazu betrachten wir den Fall $\mathcal{A} = M(2, 2; \mathbb{R})$ und

$$a = A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

- Das Minimalpolynom von A ist nicht $X - A$, denn das liegt gar nicht in $K[X]$.
- Weil $A \neq \lambda E$, ist das Minimalpolynom von A auch nicht von der Form $X - \lambda$, sondern von höherem Grad.
- Andrerseits ist $A^2 = -E$. Also ist

$$\mu_A = X^2 + 1$$

das Minimalpolynom.

Weil $A^4 = E$, ist $X^4 - 1 \in N_A$ und daher $X^2 + 1$ ein Teiler von $X^4 - 1$. Aber das wußten Sie schon.

□

Die folgende Definition macht nur Sinn für den Fall $\mathcal{A} = K$. (Warum?)

Definition 17 (Nullstellenordnung). Ist $a \in K$ eine Nullstelle von $F \in K[X]\backslash\{0\}$, so heißt

$$\operatorname{ord}(F, a) := \#\{j \mid (X-a)^j | F\}$$

die *Ordnung* oder *Multiplizität* der Nullstelle a von F.

Offenbar tritt $X - a$ dann genau j-mal in der Primfaktorzerlegung von F auf. Aus der Gradformel folgt:

Lemma 15. *Seien $F \in K[X]$, $F \neq 0$, und $a_1, \ldots, a_k$ Nullstellen von F. Dann gilt*

$$\sum_{i=1}^{k} \operatorname{ord}(F, a_i) \leq \operatorname{Grad} F.$$

Die Anzahl der mit Multiplizitäten gezählten Nullstellen eines Polynoms ist höchstens gleich seinem Grad.

4 Strukturätze im allgemeinen Fall

Im weiteren sei V stets ein endlich-dimensionaler K-Vektorraum über einem Körper K.

4.1 Die verallgemeinerte Determinante und das charakteristische Polynom

Die Determinante von Matrizen kann man verstehen als Abbildung

$$\det : M(n, n; K) = K^n \times \ldots \times K^n \to K.$$

Wenn man aber das charakteristische Polynom definieren möchte, stehen in den Spalten plötzlich „Vektoren“, deren Komponenten nicht in K, sondern in $K[X]$ liegen. Und das Ergebnis soll ebenfalls ein Polynom in $K[X]$ sein. Dazu ist eine Erweiterung des Determinantenbegriffes erforderlich.

Abweichend von unserem Zugang zur Determinanten hätten wir auch zunächst die Determinante einer Matrix $A = (a_{ij})$ durch die Leibnizformel

$$\det(a_1, \ldots, a_n) = \sum_{\sigma \in \mathcal{S}_n} \operatorname{sign} \sigma \, a_{\sigma(1)1} \ldots a_{\sigma(n)n}$$

definieren und dann die folgenden Determinanteneigenschaften beweisen können:

1. $\det(a_1 + \tilde{a}_1, a_2, \ldots, a_n) = \det(a_1, a_2, \ldots, a_n) + \det(\tilde{a}_1, a_2, \ldots, a_n)$ und entsprechend in allen anderen aSpalten.
2. Für $\lambda \in K$ gilt $\det(\lambda a_1, a_2, \ldots, a_n) = \lambda \det(a_1, a_2, \ldots, a_n)$ und entsprechend in allen anderen Spalten.
3. Falls zwei Spalten gleich sind, ist

$$\det(a_1, \ldots, a_n) = 0.$$

4. $\det(E) = 1$.

5. $\det(AB) = \det A \det B$.

Ersetzt man den Körper K durch einen kommutativen Ring mit Einselement, insbesondere durch $K[X]$, so kann man die Determinante wie oben definieren, und es bleiben alle diese Eigenschaften erhalten. Ausführlicher finden Sie das behandelt z.B. in *F. Lorenz, Lineare Algebra I, Anhang zu Kapitel IV.*

Definition 18 (Charakteristisches Polynom). (i) Für $A = (a_{ij}) \in M(n, n; K)$ ist *das charakteristische Polynom* definiert durch

$$\chi_A(X) := \det \begin{pmatrix} a_{11} - X & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} - X \end{pmatrix}$$

als ein Element von $K[X]$.

(ii) Weil die Darstellungsmatrizen eines Vektorraum-Endomorphismus ähnlich zueinander sind ($A \to SAS^{-1}$) folgt aus dem Multiplikationssatz, daß man das charakteristische Polynom für Endomorphismen einfach mittels einer beliebigen Darstellungsmatrix definieren kann.

Damit haben wir also das charakteristische Polynom als Element von $K[X]$ definiert. Insbesondere macht es Sinn von Grad und Nullstellenordnung zu sprechen. Und wir können nicht nur Werte aus K, sondern aus jeder unitären K-Algebra einsetzen, zum Beispiel Endomorphismen: $\chi_f(f) \in \mathrm{End}(V)$ macht Sinn.

Wir wiederholen noch einmal:

Definition 19 (Minimalpolynom). Für $f \in \mathrm{End}(V)$ haben wir einen Einsetzungshomomorphismus

$$\eta_f : K[X] \to \mathrm{End}(V), F(X) \mapsto F(f).$$

Das Nullstellenideal

$$N_f = \{F \in K[X] \mid F(f) = 0\}$$

ist nicht trivial, wie wir gleich sehen werden. Es wird also erzeugt von einem eindeutig bestimmten normierten Polynom $\mu_f(X)$, dem *Minimalploynom von* f. Analog definiert man $\mu_A(X)$ für quadratische Matrizen A.

Wir werden sehen, daß Minimalpolynom und charakteristisches Polynom enge Beziehungen haben, und daß ihre Primfaktorzerlegung wesentlich mit Struktur des Endomorphismus f zu tun hat.

Beispiel 19. Der Endomorphismus A des $\mathbb{R}^8$ mit der Matrix

$$A = \begin{pmatrix} [\lambda] & & & & \\ & \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} & & & \\ & & \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix} & & \\ & & & \begin{bmatrix} \mu & 1 \\ 0 & \mu \end{bmatrix} \end{pmatrix}$$

mit $\lambda \neq \mu$ hat das charakteristische Polynom

$$\chi_A(X) = (X - \lambda)^6 (X - \mu)^2.$$

Es ist leicht nachzurechnen, daß das folgende Polynom A als Nullstelle hat.

$$\mu_A(X) = (X-\lambda)^3(X-\mu)^2.$$

Es ist sogar das Minimalpolynom. Das Minimalpolynom ist nämlich ein Teiler hiervon, also von der Form $(X-\lambda)^a(X-\mu)^b$ mit $a \leq 3$ und $b \leq 2$. Aber $(A-\lambda E)^2(A-\mu E)^1 \neq 0$.

Es ist ratsam, diese Matrix bei der folgenden Normalformtheorie immer als Beispiel parat zu halten.

□

Beispiel 20. Sei $f \in \mathrm{End}(V)$ ein Endomorphismus des n-dimensionalen K-Vektorraums V. Weil $\dim \mathrm{End}(V) = n^2$, sind die n^2+1 Potenzen $\mathrm{id} = f^0, f = f^1, \ldots, f^{n^2}$ linear abhängig. Es gibt also $a_0, \ldots, a_{n^2}$, die nicht alle 0 sind, so daß

$$\sum_{k=0}^{n^2} a_k f^k = 0.$$

Also ist f eine Nullstelle des Polynoms

$$\sum_{k=0}^{n^2} a_k X^k \in K[X].$$

Dieses Polynom liegt also im Nullstellenideal N_f. Also ist

$$\mathrm{Grad}\, \mu_f \leq n^2.$$

□

Nun wollen wir zeigen, daß auch das charakteristische Polynom

$$\chi_f = \det(f - X\,\mathrm{id}) \in K[X]$$

den Endomorphismus f annulliert. Dazu betrachten wir zunächst einen Spezialfall.

Lemma 16 (f-zyklische Basis). *Seien $f \in \mathrm{End}(V)$ und $v \in V$, so daß*

$$\mathbf{b} := (v, f(v), \ldots, f^{n-1}(v)) \tag{14}$$

eine Basis von V ist[1] *Dann ist $f^n(v)$ eine Linearkombination dieser Basisvektoren*

$$f^n(v) = -\sum_{k=0}^{n-1} c_k f^k(v) \tag{15}$$

und die Darstellungsmatrix $A = A_{\mathbf{b}}^{\mathbf{b}}(f)$ gegeben durch

$$A = \begin{pmatrix} 0 & 0 & \ldots & 0 & -c_0 \\ 1 & 0 & \ldots & 0 & -c_1 \\ 0 & \ddots & & \vdots & \vdots \\ \vdots & & \ddots & 0 & \vdots \\ 0 & \ldots & 0 & 1 & -c_{n-1} \end{pmatrix}. \tag{16}$$

Das charakteristische Polynom von f ist

$$\chi_f(X) = (-1)^n \left(X^n + \sum_{k=1}^{n-1} c_k X^k \right). \tag{17}$$

[1] Wir nennen (14) in diesem Fall eine *f-zyklische Basis von V*. Im allgemeinen besitzt V allerdings keine f-zyklische Basis, denken Sie an $f = \mathrm{id}_V$.

Beweis. Wir müssen nur die letzte Formel beweisen. Die folgt aber einfach aus der Entwicklung nach der letzten Spalte. □

Beispiel 21. In der Theorie der gewöhnlichen linearen Differentialgleichungen schreibt man die Differentialgleichung

$$y^{(n)} + c_{n-1}y^{(n-1)} + \ldots + c_1 y' + c_0 y = 0 \tag{18}$$

durch Einführung von dummy-Variablen um in ein lineares Differentialgleichungs*system*: Man setzt $y = y_0$ und

$$\begin{aligned} y_0' &= y_1 \\ y_1' &= y_2 \\ &\vdots \\ y_{n-2}' &= y_{n-1} \\ y_{n-1}' &= -c_0 y_0 - c_1 y_1 + \ldots + c_{n-1} y_{n-1} \end{aligned}$$

Im Matrixschreibweise ist dieses System gegeben durch

$$\begin{pmatrix} y_0 \\ \vdots \\ \vdots \\ y_{n-1} \end{pmatrix}' = \begin{pmatrix} 0 & 1 & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \\ -c_0 & -c_1 & \ldots & -c_{n-1} \end{pmatrix} \begin{pmatrix} y_0 \\ \vdots \\ \vdots \\ y_{n-1} \end{pmatrix}.$$

Aus dem Lemma folgt mit $\det A^T = \det A$, daß das charakteristische Polynom der Systemmatrix bis aufs Vorzeichen gerade das sogenannte charakteristische Polynom von (18) ist, nämlich

$$\lambda^n + c_{n-1}\lambda^{n-1} + \ldots + c_1\lambda + c_0.$$

□

Lemma 17. *Ist $f \in \mathrm{End}(V)$ und besitzt V eine f-zyklische Basis, so ist*

$$\chi_f(f) = 0. \tag{19}$$

Beweis. Sei

$$\mathbf{b} := (v, f(v), \ldots, f^{n-1}(v)) \tag{20}$$

eine f-zyklische Basis von V und $f^n(v)$ gegeben durch (15). Aus (16) folgt dann

$$\chi_f(f)(v) = (-1)^n \left(f^n(v) + \sum_{k=0}^{n-1} c_k f^k(v) \right) = 0.$$

Durch Anwenden von f auf diese Gleichung folgt

$$0 = f(f^n(v) + \sum_{k=0}^{n-1} c_k f^k(v)) = f^n(f(v)) + \sum_{k=0}^{n-1} c_k f^k(f(v))) = \chi_f(f)(f(v))$$

und weiter

$$\chi_f(f)(f^k(v)) = 0$$

für alle $k \in \{0, \ldots, n-1\}$. Also annulliert $\chi_f(f)$ alle Basisvektoren und ist deshalb $= 0$. □

Die Gleichung (19) bleibt auch richtig, wenn V keine f-zyklische Basis besitzt. Wir zeigen zuvor noch ein Lemma, von dem wir den schwierigeren Teil (ii) allerdings erst später benötigen.

Lemma 18 (Charakteristisches Polynom bei partiell invarianter Zerlegung). *Sei* $f \in \mathrm{End}(V)$ *und sei*

$$V = U \oplus W$$

mit einem f*-invarianten Unterraum* U

$$f(U) \subset U.$$

Sei $\pi : V \to V$ *die entsprechende Projektion auf* W*, also für alle* $v \in V$

$$v = \underbrace{v - \pi(v)}_{\in U} + \underbrace{\pi(v)}_{\in W}.$$

Wir bezeichnen mit χ_U *bzw.* χ_W *das charakteristische Polynom von* $f_U \in \mathrm{End}(U)$ *bzw.* $\pi \circ f|_W \in \mathrm{End}(W)$*, vgl. Lemma 6 für die Bezeichnung* f_U*. Dann gilt:*

(i) $\chi_f(X) = \chi_U(X)\chi_W(X)$.

(ii) Für alle $G(X) \in K[X]$ *und* $w \in W$ *gilt*

$$\pi G(f)(w) = G(\pi f)(w).$$

Beweis. *Zu (i).* Weil $f(U) \subset U$ ist die Darstellungsmatrix von f bezüglich einer aus Basen von U und W zusammengesetzen Basis von V von der Form

$$\begin{pmatrix} A & B \\ 0 & C \end{pmatrix},$$

wobei A die Darstellungsmatrix von $f|_U \in \mathrm{End}(U)$ und C die von $\pi \circ f|_W \in \mathrm{End}(W)$ ist. Aus der Determinantenformel für solche Matrizen folgt (i).

Zu (ii). Es genügt zu zeigen, daß für alle k und $w \in W$

$$\pi f^k(w) = (\pi f)^k(w).$$

Das zeigen wir mit Induktion über k.

$k = 0$. Trivial.

$k \to (k+1)$. Es gilt

$$\begin{aligned} \pi f^{k+1}(w) = \pi f(f^k(w)) &= \pi f(\underbrace{f^k(w) - \pi f^k(w)}_{=:u \in U} + \pi f^k(w)) \\ &= \pi(\underbrace{f(u)}_{\in U = \mathrm{Kern}\,\pi}) + \pi f(\pi f)^k(w) \quad \text{nach Ind.vor.} \\ &= (\pi f)^{k+1}(w) \end{aligned}$$

□

Satz 27 (Cayley-Hamilton). *Für* $f \in \mathrm{End}(V)$ *gilt*

$$\chi_f(f) = 0. \tag{21}$$

Beweis. Wir zeigen

$$\chi_f(f)(v) = 0$$

für alle $v \in V$. Sei also $v \in V$ und o.E. $v \neq 0$. Wir bilden die Vektoren $f(v), f^2(v), \ldots$. Es seien

$$v, f(v), \ldots, f^{r-1}(v) \tag{22}$$

linear unabhängig und

$$f^r(v) = -\sum_{k=0}^{n-1} c_k f^k(v).$$

Wir setzen

$$U := \text{Spann}(v, f(v), \ldots, f^{r-1}(v))$$

und wählen einen komplementären Unterraum W:

$$V = U \oplus W.$$

Dann ist $f(U) \subset U$ und wir können das Lemma anwenden. Mit den dort vereinbarten Bezeichnungen gilt

$$\chi_f(f)(v) = \chi_W(f)\chi_U(f)(v) = \chi_W(f)\underbrace{\chi_U(f|_U)(v)}_{=0} = 0$$

nach Lemma 17, weil $v \in U$ liegt. □

Bemerkung 1. Auf der Suche nach dem Minimalpolynom von f ist der Satz von Cayley-Hamilton hilfreich: Er besagt, daß $\chi_f(X)$ im Nullstellenideal von f liegt. Dieses Ideal ist ein Hauptideal, erzeugt vom Minimalpolynom. Das Minimalpolynom ist also ein Teiler des charakteristischen Polynoms. Hat man die Primfaktorzerlegung des charakteristischen Polynoms, so erhält man daraus das Minimalpolynom durch Weglassen von Faktoren, vgl. Korollar 2. Welche, kann man ausprobieren. Siehe dazu die Argumentation in Beispiel 19.

Bemerkung 2. Hier ist der kürzeste Beweis für den Satz von Cayley-Hamilton für Matrizen:

$$\chi_A(A) = \det(A - AE) = \det 0 = 0.$$

Warum ist er falsch? Das erste Gleichheitszeichen impliziert, daß es egal ist, ob man erst die Determinante von $A - XE$ bildet und dann den Einsetzunghomomorphismus anwendet, oder ob man erst einsetzt und dann die Determinate bildet:

$$\eta_A(\chi_A(X)) = \det(\eta_A(A - XE)).$$

Warum ist das egal? Was soll das η_A rechts bedeuten? Nach unserer Definition ist das Argument der erweiterten Determinante eine Matrix, deren Komponenten Polynome in X sind. Einsetzen von A liefert eine Matrix, deren Komponenten Matrizen sind. Also ist

$$\eta_A(A - XE) \in M(n, n; M(n, n; K)) = M(n, n; K)^n \times \ldots \times M(n, n; K)^n$$

und jedenfalls nicht

$$\eta_A(A - XE) = 0 \in M(n, n; K).$$

Man kann das tatsächlich zu einem Beweis ausbauen, aber der ist auch nicht so einfach. Vgl. *F. Lorenz, Lineare Algebra I.*

4.2 Trigonalisierung

Definition 20 (Trigonalisierbarkeit). Ein Endomorphismus $f \in \text{End}(V)$ heißt *trigonalisierbar*, wenn es eine Basis $\mathbf{b}$ gibt, bezüglich der f eine Darstellungsmatrix in oberer Dreiecksgestalt hat:

$$A_{\mathbf{b}}^{\mathbf{b}}(f) = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \tag{23}$$

Satz 28 (Trigonalisierbarkeit). *Für $f \in \mathrm{End}(V)$ sind folgende Aussagen äquivalent:*

(i) *f ist trigonalisierbar.*

(ii) *Es gibt eine Basis $(v_1, \ldots, v_n)$ von V, so daß gilt*

$$f(v_k) \in \mathrm{Spann}(v_1, \ldots, v_k) \text{ für alle } k.$$

(iii) *Es gibt eine Flagge*

$$\{0\} = V_0 \subset V_1 \subset \ldots \subset V_n = V$$

von f-invarianten Unterräumen von V mit $\dim V_k = k$.

(iv) *Das charakteristische Polynom von f zerfällt in Linearfaktoren:*

$$\chi_f(X) = (\lambda_1 - X) \ldots (\lambda_n - X).$$

Beweis. Wir zeigen

$$(i) \implies (iv) \implies (iii) \implies (ii) \implies (i).$$

Zu (i) $\implies$ (iv). Trivial.

Zu (iv) $\implies$ (iii). Durch vollständige Induktion über $n := \dim V$.

$n = 1$. Nichts zu zeigen.

$n \implies (n+1)$. Sei v_1 ein Eigenvektor zum Eigenwert λ_1. Ergänze zu einer Basis $(v_1, \ldots, v_{n+1})$ von V und setze $W := \mathrm{Spann}(v_2, \ldots, v_{n+1})$. Sei $V_1 := Kv_1$ und sei $\pi : V \to W$ die Projektion zur Zerlegung $V = V_1 \oplus W$. Setze $g : \pi \circ f|_W : W \to W$. Dann ist nach Lemma 18

$$\chi_f(X) = \chi_{f_{V_1}}(X)\chi_g(X),$$

also

$$\chi_g(X) = (\lambda_2 - X) \ldots (\lambda_{n+1} - X).$$

Nach Induktionsvoraussetzung besitzt W eine g-invariante Flagge

$$0 = W_0 \subset W_1 \subset \ldots \subset W_n = W.$$

Dann ist also $f(W_k) \subset V_1 + W_k$ und

$$\{0\} =: V_0 \subset V_1 + W_0 \subset V_2 := V_1 + W_1 \subset \ldots \subset V_{n+1} := V_1 + W_n = V$$

ist eine f-invariante Flagge wachsender Dimension.

Zu (iii) $\implies$ (ii). Wähle für $1 \leq i \leq n$

$$v_i \in V_i \backslash V_{i-1}.$$

Dann sind die v_i linear unabhängig und es gilt

$$V_k = \mathrm{Spann}(v_1, \ldots, v_k),$$
$$f(v_i) \in f(V_i) = V_i.$$

Zu (ii) $\implies$ (i). Trivial.

□

4.3 Primzerlegung, Diagonalisierbarkeit

In diesem Abschnitt wollen wir zeigen, daß ein Endomorphismus sich in „einfache Bausteine“ zerlegen läßt, die genau der Primfaktorzerlegung seines Minimalpolynoms entsprechen.

Lemma 19 (Zerlegung des Minimalpolynoms induziert f-invariante direkte Summe). *Seien $f \in \mathrm{End}(V)$ und $H_1, H_2 \in K[X]$ zwei normierte teilerfremde Polynome. Für $G(X) = H_1(X)H_2(X)$ gelte*

$$G(f) = 0.$$

Setze

$$V_1 := \mathrm{Kern}\, H_1(f), \quad V_2 := \mathrm{Kern}\, H_2(f).$$

Dann gilt:

(i) $V_1 = H_2(f)(V), \quad V_2 = H_1(f)(V).$

(ii) V_1 *und* V_2 *sind* f*-invariant:* $f(V_i) \subset V_i$.

(iii) $V = V_1 \oplus V_2$,

(iv) Ist speziell $G = \mu_f$ *das Minimalpolynom von* f*, so gilt für* $f_i := f|_{V_i} \in \mathrm{End}(V_i)$

$$\mu_{f_i} = H_i, \quad i \in \{1, \ldots, 2\}.$$

Beweis. Im Beweis benutzen wir wiederholt, daß für beliebige Polynome $F_1, F_2 \in K[X]$ gilt

$$F_1(f)F_2(f) = F_2(f)F_1(f).$$

Zu (i). Weil $ggT(H_1, H_2) = 1$ ist, gibt es nach Satz 23 $A_1, A_2 \in K[X]$ mit

$$1 = A_1(X)H_1(X) + A_2(X)H_2(X),$$

also

$$\mathrm{id} = A_1(f)H_1(f) + A_2(f)H_2(f),$$

Daher gilt für $v \in V_1$

$$v = A_1(f)\underbrace{H_1(f)(v)}_{=0} + A_2(f)H_2(f)(v) = H_2(f)A_2(f)(v) \in H_2(f)(V).$$

Also $V_1 \subset H_2(f)(V)$. Andrerseits ist

$$H_1(f)H_2(f)(V) = G(f)(V) = 0,$$

also

$$H_2(f)(V) \subset \mathrm{Kern}\, H_1(f) = V_1.$$

Damit folgt

$$V_1 = H_2(f)(V)$$

und analog $V_2 = H_1(f)(V)$.

Zu (ii). Es gilt

$$H_i(f)f(V_i) = fH_i(f)(V_i) = \{0\}.$$

Daher ist

$$f(V_i) \subset \mathrm{Kern}\, H_i = V_i.$$

Zu (iii). Wie im Beweis von (i) findet man für $v \in V$

$$v = A_1(f)H_1(f)(v) + A_2(f)H_2(f)(v) = \underbrace{H_1(f)A_1(f)(v)}_{\in V_2} + \underbrace{H_2(f)A_2(f)(v)}_{\in V_1}.$$

Also ist

$$V = V_1 + V_2.$$

Andrerseits folgt aus

$$v = A_1(f)H_1(f)(v) + A_2(f)H_2(f)(v)$$

$v = 0$, falls $v \in V_1 \cap V_2 = \operatorname{Kern} H_1(f) \cap \operatorname{Kern} H_2(f)$. Daher ist $V_1 \cap V_2 = 0$ und

$$V = V_1 \oplus V_2.$$

Zu (iv). Nach Definition ist $H_1(f)(V_1) = \{0\}$, also

$$H_1(f_1) = 0. \tag{24}$$

Andrerseits ist

$$H_2(f)(V) = V_1.$$

Daraus folgt

$$(\mu_{f_1} H_2)(f) = \mu_{f_1}(f)H_2(f) = 0,$$

also

$$\mu_f | \mu_{f_1} H_2.$$

Nach Voraussetzung war aber $\mu_f = H_1 H_2$ und wir erhalten

$$H_1 H_2 | \mu_{f_1} H_2.$$

Daher gilt

$$H_1 | \mu_{f_1}. \tag{25}$$

Aus (24) und (25) folgt die Behauptung. □

Satz 29 (Primzerlegungssatz, 1. Normalform). *Seien $f \in \mathrm{End}(V)$ und seien $P_1, \dots P_k \in K[X]$ paarweise verschiedene normierte Primpolynome. Seien weiter $m_1, \dots, m_k \in \mathbb{N}\backslash\{0\}$. Für*

$$G(X) := P_1(X)^{m_1} \dots P_k(X)^{m_k} \tag{26}$$

gelte

$$G(f) = 0.$$

Insbesondere gilt dies, wenn (26) die Primfaktorzerlegung des Minimalpolynoms $\mu_f = G$ ist. Wir setzen

$$V_i := \mathrm{Kern}\, P_i(f)^{m_i}.$$

Dann gilt:

(i) Die V_i sind f-invariante Unterräume.

(ii) $V = V_1 \oplus \dots \oplus V_k$.

(iii) Ist $G = \mu_f$ und $f_i := f|_{V_i} \in \mathrm{End}(V_i)$, so ist

$$\mu_{f_i}(X) = P_i(X)^{m_i}.$$

Wählt man in jedem V_i eine Basis, so setzen sich diese zu einer Basis $\mathbf{b}$ von V zusammen, bezüglich der die Darstellungsmatrix von f Blockdiagonalgestalt hat:

$$A_{\mathbf{b}}^{\mathbf{b}}(f) = \begin{pmatrix} A_1 & & & 0 \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_k \end{pmatrix}.$$

Dabei ist A_i eine $(\dim V_i)$-reihige quadratische Matrix.

Beweis. Durch vollständige Induktion über k.

$\underline{k = 1.}$ Nichts zu beweisen.

$\underline{k \to (k+1).}$ Dann schreiben wir

$$G(X) = \underbrace{P_1(X)^{m_1}}_{=:H_1(X)} \underbrace{(P_2(X)^{m_2} \dots P_{k+1}(X)^{m_{k+1}})}_{=:H_2(X)}.$$

Aus dem Lemma 19 folgt mit

$$V_1 := \mathrm{Kern}\, P_1(f)^{m_1}, \quad W := \mathrm{Kern}\, H_2(f),$$

daß

$$V = V_1 \oplus W$$

mit f-invarianten Unterräumen V_1 und W. Ist $G = \mu_f$ das Minimalpolynom von f, so ist $P_1(X)^{m_1}$ das von f_{k+1} und $H_2(X)$ das von $f|_W \in \mathrm{End}(V)$.

Nach Induktionsannahme angewendet auf $f|_W \in \mathrm{End}(W)$ und H_2 spaltet auch W als direkte Summe f-invarianter Unterräume und es gilt die entsprechende Aussage über Minimalpolynome. □

Bevor wir ein Beispiel betrachten, konstatieren wir noch eine Folgerung aus dem Lemma 19.

Lemma 20. *Ist $f \in \mathrm{End}(V)$ und ist $Q \in K[X]$ ein normiertes Primpolynom, so gilt*

$$\operatorname{Kern} Q(f) \neq \{0\} \implies Q | \mu_f.$$

Ist insbesondere $\lambda \in K$ ein Eigenwert von f, so ist $X - \lambda$ ein Primfaktor von μ_f.

Beweis. Wäre Q kein Teiler von μ_f, so wären μ_f und Q teilerfremd, weil Q ein Primpolynom ist. Weil $G(X) = Q(X)\mu_f(X)$ f als Nullstelle hat, folgte aus Lemma 19, daß

$$V = \underbrace{\operatorname{Kern} Q(f)}_{\neq\{0\}} \oplus \underbrace{\operatorname{Kern} \mu_f}_{=V}.$$

Aber das ist ein Widerspruch. □

Beispiel 22 (Hauptvektoren). Das charakteristische Polynom von $f \in \mathrm{End}(V)$ zerfalle in Linearfaktoren:

$$\chi_f(X) = \pm(X - \lambda_1)^{n_1} \dots (X - \lambda_k)^{n_k}$$

mit paarweise verschiedenen Eigenwerten λ_i der algebraischen Vielfachheiten $n_i > 0$. (Das ist nach dem Fundamentalsatz der Algebra z.B. für $K = \mathbb{C}$ immer der Fall.) Nach dem Lemma ist das Minimalplynom dann

$$\mu_f(X) = (X - \lambda_1)^{m_1} \dots (X - \lambda_k)^{m_k}, \quad 0 < m_i \leq n_i.$$

> (Bemerkung. Auch im allgemeinen Fall ist jeder Primfaktor des charakteristischen Polynoms mit positivem Exponenten im Minimalpolynom vertreten; aber das werden wir erst später einsehen.)

Es gilt

$$V = V_1 \oplus \dots \oplus V_k$$

mit f-invarianten Unterräumen

$$V_i = \operatorname{Kern}(f - \lambda_i \operatorname{id})^{m_i} \supset \operatorname{Eig}(f, \lambda_i).$$

Ist f diagonalisierbar, so gilt

$$V = \operatorname{Eig}(f, \lambda_1) \oplus \dots \oplus \operatorname{Eig}(f, \lambda_k).$$

Dann ist also $V_i = \operatorname{Eig}(f, \lambda_i)$ und V_i hat eine Basis aus Eigenvektoren zu λ_i.

Wir haben aber schon gesehen, daß möglicherweise die geometrische Vielfachheit mancher Eigenwerte kleiner ist als die algebraische: $\dim \operatorname{Eig}(f, \lambda_i) < n_i$. In diesem Fall ist $\operatorname{Eig}(f, \lambda_i) \subsetneq V_i$ und es gibt keine Basis von V_i aus Eigenvektoren zu λ_i. Aber es gibt immerhin eine Basis aus Vektoren, die die Gleichung

$$(f - \lambda_i \operatorname{id})^{m_i} v = 0$$

erfüllen. Bei $m_i = 1$ ist das gerade die Eigenvektorgleichung, bei $m_i > 1$ hat diese Gleichung aber außer den Eigenvektoren möglicherweise weitere Lösungen, die man *Hauptvektoren* nennt. Ist $m \geq 1$ der kleinste Exponent mit

$$(f - \lambda_i \operatorname{id})^{m} v = 0,$$

so heißt v ein Hauptvektor der *Stufe* m. Die Hauptvektoren der Stufe 1 sind also gerade die Eigenvektoren.

Bezüglich einer Basis aus Hauptvektoren hat die Darstellungsmatrix dann die im Satz angegebene Blockdiagonalgestalt. Wir werden im später untersuchen, wie man die Hauptvektoren wählen muß und kann, um die einzelnen Blöcke noch möglichst einfach zu gestalten.

□

Beispiel 23. Wir konkretisieren das anhand der Matrix aus Beispiel 19:

$$A = \begin{pmatrix} \begin{bmatrix}\lambda\end{bmatrix} & & & & \\ & \begin{bmatrix}\lambda & 1\\ 0 & \lambda\end{bmatrix} & & & \\ & & \begin{bmatrix}\lambda & 1 & 0\\ 0 & \lambda & 1\\ 0 & 0 & \lambda\end{bmatrix} & & \\ & & & \begin{bmatrix}\mu & 1\\ 0 & \mu\end{bmatrix} \end{pmatrix}$$

mit $\lambda \neq \mu$. Dafür gilt:

- Die algebraische Vielfachheit von λ ist 6, seine geometrische 3. Die kanonischen Basisvektoren e_1, e_2, e_4 sind Eigenvektoren und e_3, e_5, e_6 Hauptvektoren zum Eigenwert λ. Es gilt

  $$(f - \lambda \operatorname{id})^2(e_3) = 0 = (f - \lambda \operatorname{id})^2(e_5),$$
  $$(f - \lambda \operatorname{id})^3(e_6) = 0.$$

  Das sieht man leicht:

  $$\begin{aligned} f(e_6) = \lambda e_6 + e_5 &\implies (f - \lambda \operatorname{id})e_6 = e_5,\\ f(e_5) = \lambda e_5 + e_4 &\implies (f - \lambda \operatorname{id})e_5 = e_4,\\ f(e_4) = \lambda e_4 &\implies (f - \lambda \operatorname{id})e_4 = 0. \end{aligned}$$

  Also sind e_3, e_5 Hauptvektoren der Stufe 2 und e_6 ist eine Hauptvektor der Stufe 3.

- Die algebraische Vielfachheit von μ ist 2, die geometrische 1. Ein Eigenvektor ist e_7, ein Hauptvektor der Stufe 2 ist e_8.

□

Korollar 3 (Diagonalisierbarkeit). *Der Endomorphismus $f \in \operatorname{End}(V)$ ist genau dann diagonalisierbar, wenn sein Minimalpolynom das Produkt paarweise verschiedener Linearfaktoren ist:*

$$\mu_f(X) = (X - \lambda_1) \dots (X - \lambda_k), \quad \lambda_i \text{ paarweise verschieden.}$$

Beweis. Sei f diagonalisierbar und

$$G(X) := (X - \lambda_1) \dots (X - \lambda_k),$$

wobei die λ_i die paarweise verschiedenen Eigenwerte von f sind. Für $v \in \operatorname{Eig}(f, \lambda_i)$ ist folgt dann

$$\begin{aligned} G(f)(v) &= (f - \lambda_1 \operatorname{id}) \dots (f - \lambda_k \operatorname{id})(v)\\ &= (f - \lambda_1 \operatorname{id}) \dots \widehat{(f - \lambda_i \operatorname{id})} \dots (f - \lambda_k \operatorname{id}) \underbrace{(f - \lambda_i \operatorname{id})(v)}_{=0}. \end{aligned}$$

Weil f diagonalisierbar ist, erzeugen die Eigenvektoren ganz V. Also ist dann $G(f) = 0$. Nach dem Lemma 20 enthält μ_f alle Primfaktoren $X - \lambda_i$ und daher ist $G = \mu_f$.

Sei nun umgekehrt $\mu_f(X) = (X - \lambda_1) \dots (X - \lambda_k)$. Nach dem Primzerlegungssatz ist dann

$$V = V_1 \oplus \dots \oplus V_k$$

und $X - \lambda_i$ das Minimalpolynom von $f|_{V_i} \in \mathrm{End}(V_i)$. Das bedeutet aber $V_i = \mathrm{Eig}(f, \lambda_i)$. Deshalb ist f diagonalisierbar. □

4.4 Rationale Normalform

In diesem Abschnitt wollen wir die Primzerlegung eines Endomorphismus weiter „verfeinern“. Wir wollen die invarianten Unterräume, die den verschiedenen Primfaktoren entsprechen, in weitere invariante Unterräume zerlegen. Das gibt die sogenannte *rationale Normalform*. Später betrachten wir dann den Spezialfall, daß das charkateristische Polynom in Linearfaktoren zerfällt, und erhalten aus der rationalen die berühmte *Jordansche Normalform*.

Die weitere Zerlegung der f-invarianten Teilräume aus der Primzerlegung benutzt f-zyklischen Unterräume, die wir schon im Beweis des Satzes von Cayley-Hamilton kennengelernt haben.

Definition 21. Seien $f \in \mathrm{End}(V)$ und $v \in V$. Dann heißt

$$Z(f, v) := \mathrm{Spann}\{f^k(v) \,|\, k \in \mathbb{N}\} = \{G(f)v \,|\, G(X) \in K[X]\}$$

der von v erzeugte *f-zyklische Unterraum*.

Wir erinnern an Lemma 16. Sind

$$v, f(v), \dots, f^{k-1}(v) \tag{27}$$

linear unabhängig, aber ist $f^k(v)$ von den vorangehenden linear abhängig, so ist $Z(f, v)$ offenbar f-invariant, und offenbar der kleinste f-invariante Unterraum, der v enthält. Die Vektoren (27) bilden eine (f-zyklische) Basis. Bezüglich dieser ist die Darstellungsmatrix von $f|_{Z(f,v)} \in \mathrm{End}(V)$ von der Form

$$A = \begin{pmatrix} 0 & 0 & \dots & 0 & -c_0 \\ 1 & 0 & \dots & 0 & -c_1 \\ 0 & \ddots & & \vdots & \vdots \\ \vdots & & \ddots & 0 & \vdots \\ 0 & \dots & 0 & 1 & -c_{n-1} \end{pmatrix}. \tag{28}$$

Es war

$$\chi_A(X) = (-1)^n \{X^n + \sum_{k=0}^{n-1} c_k X^k\}.$$

Wir halten noch einige Eigenschaften des Minimalpolynoms auf zyklischen Unterräumen fest.

Lemma 21. *Seien $f \in \mathrm{End}(V)$ und $v \in V$. Für das Minimalpolynom von*

$$f|_{Z(f,v)} \in \mathrm{End}(Z(f, v))$$

schreiben wir kurz $\mu_{(f,v)}$ und für das charakteristische Polynom $\chi_{(f,v)}$.
Dann gilt

(i) $\mu_{(f,v)}(X) = \pm\chi_{(f,v)}(X)$.

(ii) $\dim Z(f,v) = \text{Grad } \mu_{(f,v)}(X)$.

(iii) $\mu_{(f,v)} | \mu_f$.

(iv) $\mu_{(f,v)}(X)$ *ist das normierte Polynom* $G(X)$ *kleinsten positiven Grades, für welches* $G(f)(v) = 0$*, erzeugt also das Ideal*

$$\{G(X) \in K[X] \,|\, G(f)(v) = 0\}.$$

(v) Für $v, w \in V$ *und* $H(X) \in K[X]$ *gilt*

$$\mu_{(f,v)}(X) = \mu_{(f,w)}(X) \implies \mu_{(f,H(f)v)}(X) = \mu_{(f,H(f)w)}(X).$$

Beweis. Zu *(i).* Sei k maximal, so daß $v, f(v), \ldots, f^{k-1}(v)$ linear unabhängig sind.

$$\mu_{(f,v)}(X) = \sum_{j=0}^{m} a_j X^j, \quad a_m = 1.$$

Dann ist

$$0 = \mu_{(f,v)}(f)(v) = \sum_{j=0}^{m} a_j f^j(v).$$

Also sind $v, \ldots, f^m(v)$ linear abhängig. Nach Definition von $Z(f,v)$ ist dann aber $m \geq k$. Andrerseits folgt aus $\mu_{(f,v)} | \chi_{(f,v)}$, daß $k \leq m$, also $k = m$. Weil $\mu_{(f,v)}$ und $(-1)^k \chi_{(f,v)}$ beide normiert sind, folgt *(i)*.

Zu *(ii).* Triviale Folge aus *(i)*.

Zu *(iii).* Weil $\mu_f(f) = 0$, ist erst recht $\mu_f(f|_{Z(f,v)}) = \mu_f(f)|_{Z(f,v)} = 0$. Daraus folgt *(iii)*.

Zu *(iv).* Aus $H(f)(v) = 0$ folgt $0 = f^k H(f)(v) = H(f)(f^k(v))$ und damit

$$H(f)|Z(f,v) = 0.$$

Also ist $\mu_{(f,v)}(X) | H(X)$. Da andererseits aber $\mu_{(f,v)}(v) = 0$ folgt die Behauptung.

Zu *(v).* Ist $G(X) \in K[X]$, so gilt

$$G(f)H(f)(w) = 0 \implies \mu_{(f,w)} | G(X)H(X) \implies \mu_{(f,v)} | G(X)H(X) \implies G(f)H(f)(v) = 0.$$

Natürlich gilt auch die Umkehrung, d.h.

$$\{G \in K[X] \,|\, G(f)H(f)(v) = 0\} = \{G \in K[X] \,|\, G(f)H(f)(w) = 0\}.$$

Aus *(iv)* folgt daraus die Behauptung. □

Wegen des Primzerlegungssatzes können wir uns im folgenden auf den Fall beschränken, daß das Minimalpolynom eine Primpolynompotenz ist. Das macht die Formulierungen einfacher und die Struktur durchsichtiger.

Satz 30 (Rationale Normalform: Spezielle Version). *Sei $f \in \mathrm{End}(V)$ ein Endomorphismus, dessen Minimalpolynom die Potenz eines normierten Primpolynoms ist:*

$$\mu_f(X) = P(X)^m, \quad \mathrm{Grad}\, P(X) = l.$$

(Im komplexen Fall ist also $\mu_f(X) = (X-\lambda)^m$ für einen Eigenwert λ.)
Dann gibt es eindeutig bestimmte positive natürliche Zahlen r und

$$m = m_1 \geq m_2 \geq \ldots \geq m_r \geq 1,$$

und dazu $v_1, \ldots, v_r \in V$, so daß

$$V = Z(f, v_1) \oplus \ldots \oplus Z(f, v_r)$$

und (mit der Bezeichnung aus Lemma 21)

$$\mu_{(f,v_i)} = P(X)^{m_i} \text{ für alle } i \in \{1, \ldots, r\}.$$

Insbesondere ist nach dem Lemma 21

$$\dim Z(f, v_i) = l\, m_i \quad \text{für alle } i \in \{1, \ldots, r\}.$$

Beweis. **Teil A: Existenz.** Durch vollständige Induktion über $\dim V$.

<u>$\dim V = 1$.</u> Dann ist

$$f = \lambda\, \mathrm{id}, \quad \mu_f = X - \lambda$$

und $V = Z(f, v)$ für jedes $v \neq 0$. Insbesondere ist $r = m_1 = 1$.

<u>Induktionsschluß.</u> Wir nehmen an, daß $\dim V > 1$ und der Satz für Vektorräume kleinerer Dimension bereits bewiesen sei.

<u>1. Schritt: Abspalten eines möglichst großen f-zyklischen Unterraums.</u> Das Minimalpolynom von f ist $P(X)^m$. Also gibt es einen Vektor $v_1 \in V$ mit

$$P(f)^{m-1}(v_1) \neq 0,$$

und wir setzen $m_1 = m$ und

$$U := Z(f, v_1).$$

Das Minimalpolynom $\mu_{(f,v_1)}(X)$ ist dann ein Teiler von $P(X)^m$, und nach Wahl von v_1 ist

$$\mu_{(f,v_1)}(X) = P(X)^m.$$

Nach Lemma 21 ist $\dim Z(f, v_1) = l\, m$, und einen größeren f-zyklischen Unterraum gibt es nicht.

<u>2. Schritt: Anwenden der Induktionsvoraussetzung auf einen komplementären Unterrraum.</u>
Ist $U = V$, so ist nichts mehr zu zeigen. Sei also $U \neq V$ und

$$V = U \oplus W.$$

Wenn W ein f-invarianter Unterraum ist, ist man mit der Induktion so gut wie fertig. Aber leider ist es völlig unklar, ob man einen zu U komplementären *f-invarianten* Unterraum finden kann!

Wir bezeichnen mit $\pi : V \to W$ die Projektion und definieren $\tilde{f} \in \mathrm{End}(W)$ durch

$$\tilde{f}(w) := \pi f(w).$$

Was ist das Minimalpolynom von $\tilde{f}$? Wir haben für alle $w \in W$

$$\mu_f(\tilde{f})(w) = \mu_f(\pi f)(w) \underset{\text{Lemma 18}}{=} \pi\mu_f(f)(w) = 0.$$

Deshalb ist $\mu_{\tilde{f}}(X)$ ein Teiler von $\mu_f(X) = P(X)^m$, und es gibt ein $\tilde{m} \in \{1, \ldots, m\}$ mit

$$\mu_{\tilde{f}}(X) = P(X)^{\tilde{m}}.$$

Nach Induktionsvoraussetzung gibt es daher positive natürliche Zahlen r und

$$m \geq \tilde{m} = m_2 \geq \ldots \geq m_r$$

und Vektoren $w_2, \ldots, w_r \in W$, so daß

$$W = Z(\tilde{f}, w_2) \oplus \ldots \oplus Z(\tilde{f}, w_r)$$

und

$$\mu_{(\tilde{f}, w_i)} = P(X)^{m_i} \text{ für alle } i \in \{2, \ldots, r\}.$$

Insbesondere merken wir an, daß

$$P(\tilde{f})^{m_i - 1}(w_i) \neq 0, \tag{29}$$

denn sonst wäre $P(\tilde{f})^{m_i-1}(\tilde{f}^k(w_i)) = \tilde{f}^k P(\tilde{f})^{m_i-1}(w_i) = 0$ für alle k, also $P(\tilde{f})^{m_i-1} = 0$.

3. Schritt: Korrektur der w_i. Die Räume $Z(f, w_i) \subset V$ (ohne Tilde über dem f) leisten noch nicht das Gewünschte. Zum Beispiel ist im allgemeinen

$$P(f)^{m_i}(w_i) \neq 0.$$

Wir wissen nur, daß

$$\pi P(f)^{m_i}(w_i) = P(\tilde{f})^{m_i}(w_i) = \mu_{(\tilde{f}, w_i)}(\tilde{f})(w_i) = 0,$$

also

$$P(f)^{m_i}(w_i) \in U.$$

Wir werden die w_i deshalb noch um einen Anteil in dem (f-invarianten) Unterraum U korrigieren. Diese Korrektur ist das Hauptproblem des Beweises.

Wir möchten $u_i \in U$ finden, so daß für $v_i := w_i - u_i$

$$P(f)^{m_i}(v_i) = P(f)^{m_i}(w_i - u_i) = 0, \tag{30}$$

d.h. so daß

$$P(f)^{m_i}(u_i) = P(f)^{m_i}(w_i) \tag{31}$$

ist.

Beachten Sie, daß die Elemente von $U = Z(f, v_1)$ Linearkombinationen der $f^i(v)$, also von der Form

$$G(f)(v_1) \text{ mit } G(X) \in K[X]$$

sind. Deshalb gibt es $G_i(X) \in K[X]$ mit

$$G_i(f)(v_1) = P(f)^{m_i}(w_i). \tag{32}$$

Es genügt nun zu zeigen, daß

$$P(X)^{m_i} | G_i(X), \tag{33}$$

denn dann ist $G_i(X) = P(X)^{m_i} H_i(X)$, und aus (32) und

$$G_i(f)(v_1) = P(f)^{m_i} \underbrace{H_i(f)(v_1)}_{=:u_i}$$

folgt (31).

Zum Beweis von (33): Weil $m_i \leq \tilde{m} \leq m$ genügt es zu zeigen, daß

$$P(X)^m | P(X)^{m-m_i} G_i(X).$$

Weil $P(X)^m = \mu_{(f,v_1)}(X)$ genügt es, zu zeigen, daß

$$P(f)^{m-m_i} G_i(f) | U = 0.$$

Weil $U = Z(f, v_1)$ ist, genügt dafür der Nachweis, daß

$$P(f)^{m-m_i} G_i(f)(v_1) = 0.$$

Aber es gilt

$$P(f)^{m-m_i} G_i(f)(v_1) = P(f)^{m-m_i} P(f)^{m_i}(w_i) = P(f)^m(w_i) = \mu_f(f)(w_i) = 0.$$

4. Schritt: Die $v_i = w_i - u_i$ erfüllen alle Wünsche. Wir wählen also Vektoren $u_i \in U$ zu den w_i wie im vorigen Schritt und setzen

$$v_i = w_i - u_i, \quad i \in \{2, \ldots, r\}. \tag{34}$$

Dann gilt nach Konstruktion jedenfalls

$$P(f)^{m_i}|_{Z(f,v_i)} = 0.$$

Und weil

$$\pi P(f)^{m_i-1}(v_i) = \pi P(f)^{m_i-1}(w_i - u_i) = \pi P(f)^{m_i-1}(w_i) - 0 = P(\tilde{f})^{m_i-1}(w_i) \underset{(29)}{\neq} 0,$$

ist $P(f)^{m_i-1}|_{Z(f,v_i)} \neq 0$ und

$$P(X)^{m_i} = \mu_{(f,v_i)}.$$

Weiter zeigen wir

$$V = Z(f, v_1) + \ldots + Z(f, v_r). \tag{35}$$

Sei $v \in V$. Dann gibt es $A_i(X) \in K[X]$, so daß

$$\begin{aligned}
v &= A_1(f)v_1 + A_2(\tilde{f})w_2 + \ldots + A_r(\tilde{f})w_r \\
&\underset{\text{Lemma 18}}{=} A_1(f)v_1 + \pi A_2(f)w_2 + \ldots + \pi A_r(f)w_r \\
&= A_1(f)v_1 + A_2(f)w_2 + \ldots + A_r(f)w_r - u \quad u \in U \\
&= \underbrace{A_1(f)v_1 - u + A_2(f)u_2 + \ldots + A_r(f)u_r}_{\in U = Z(f,v_1)} + A_2(f)v_2 + \ldots + A_r(f)v_r
\end{aligned}$$

Das beweist (35).

Aus Lemma 21 folgt $\dim Z(f, v_1) = \operatorname{Grad} P(X)^{m_1}$ und für $i \geq 2$

$$\dim Z(f, v_i) = \operatorname{Grad} P(X)^{m_i} = \dim Z(\tilde{f}, w_i).$$

Weil $V = U \oplus W$ und $W = Z(\tilde{f}, w_2) \oplus \ldots \oplus Z(\tilde{f}, w_r)$ folgt aus Dimensionsgründen, daß

$$V = Z(f, v_1) \oplus \ldots \oplus Z(f, v_r). \tag{36}$$

Damit ist der Existenzbeweis abgeschlossen.

Teil B: Eindeutigkeit. Es sei $\mu_f(X) = P(X)^m$ mit einem normierten Primpolynom $P(X)$ und es seien

$$V = Z(f, v_1) \oplus \ldots \oplus Z(f, v_r)$$

und

$$V = Z(f, w_1) \oplus \ldots \oplus Z(f, w_s)$$

zwei Zerlegungen mit

$$\mu_{(f,v_i)}(X) = P(X)^{m_i}, \quad \mu_{(f,w_j)}(X) = P(X)^{n_j}$$

für alle $i \in \{1, \ldots, r\}, j \in \{1, \ldots, s\}$. Weiter seien

$$m = m_1 \geq m_2 \geq \ldots \geq m_r \geq 1, \quad m = n_1 \geq n_2 \geq \ldots \geq n_s \geq 1.$$

Wir wollen durch Induktion zeigen, daß $m_i = n_i$ für alle $i \in \{1, \ldots, \min(r, s)\}$. Nach Lemma 21 folgt für diese i dann

$$\dim Z(f, v_i) = \operatorname{Grad} P(X)^{m_i} = \operatorname{Grad} P(X)^{n_i} = \dim Z(f, w_i),$$

und damit $r = s$.

$\underline{m_1 = m = n_1}$ ist bereits vorausgesetzt.

$\underline{i \to i+1.}$ Es sei $i < \min(r, s)$ und bereits gezeigt, daß $m_j = n_j$ für alle $j \in \{1, \ldots, i\}$. Dann gilt

$$\begin{aligned} P(f)^{m_{i+1}} V &= P(f)^{m_{i+1}} Z(f, v_1) \oplus \ldots \oplus P(f)^{m_{i+1}} Z(f, v_r) \\ &= Z(f, P(f)^{m_{i+1}}(v_1)) \oplus \ldots \oplus Z(f, P(f)^{m_{i+1}}(v_i)) \oplus \ldots \oplus Z(f, P(f)^{m_{i+1}}(v_r)) \\ &= Z(f, P(f)^{m_{i+1}}(v_1)) \oplus \ldots \oplus Z(f, P(f)^{m_{i+1}}(v_i)), \end{aligned} \tag{37}$$

weil $P(f)^{m_{i+1}}(v_j) = 0$ für $j \geq i+1$. Weiter ist

$$P(f)^{m_{i+1}} V = Z(f, P(f)^{m_{i+1}}(w_1)) \oplus \ldots \oplus Z(f, P(f)^{m_{i+1}}(w_s)). \tag{38}$$

Nach Induktionsvoraussetzung ist $m_j = n_j$ für alle $j \in \{1, \ldots, i\}$, also

$$\mu_{(f,v_j)}(X) = P(X)^{m_j} = \mu_{(f,w_j)}(X).$$

Aus Lemma 21 ergibt sich daher für alle $j \in \{1, \ldots, i\}$:

$$\mu_{(f,P(f)^{m_{i+1}}(v_j))}(X) = \mu_{(f,P(f)^{m_{i+1}}(w_j))}(X)$$

und

$$\dim Z(f, P(f)^{m_{i+1}}(v_j)) = \dim Z(f, P(f)^{m_{i+1}}(w_j)).$$

Vergleich von (37) und (38) liefert

$$\begin{aligned} &\dim Z(f, P(f)^{m_{i+1}}(v_1)) + \ldots + \dim Z(f, P(f)^{m_{i+1}}(v_i)) \\ &\quad = \dim P(f)^{m_{i+1}} V \\ &\quad = \dim Z(f, P(f)^{m_{i+1}}(w_1)) + \ldots + \dim Z(f, P(f)^{m_{i+1}}(w_i)). \end{aligned}$$

Es folgt

$$\dim Z(f, P(f)^{m_{i+1}}(w_{i+1})) = \ldots = \dim Z(f, P(f)^{m_{i+1}}(w_s)) = 0.$$

Daher ist $P(f)^{m_{i+1}}(w_{i+1}) = 0$ und deshalb $P(X)^{n_{i+1}} | P(X)^{m_{i+1}}$, also $n_{i+1} \leq m_{i+1}$. Ebenso findet man $m_{i+1} \leq n_{i+1}$, also

$$m_{i+1} = n_{i+1}.$$

□

Wir beseitigen jetzt die Bedingung, daß μ_f nur *einen* Primteiler hat:

Satz 31 (Rationale Normalform: Allgemeine Version). *Seien V ein n-dimensionaler K-Vektorraum und $f \in \mathrm{End}(V)$. Dann gilt:*

(i) Es gibt $r \in \mathbb{N}$, Vektoren $v_1, \ldots, v_r \in V$ und normierte Primpolynompotenzen positiven Grades $P_1(X)^{m_1}, \ldots, P_r(X)^{m_r}$ so daß gilt:

(a) $V = Z(f, v_1) \oplus \ldots \oplus Z(f, v_r)$.

(b) $\mu_{(f,v_i)} = P_i(X)^{m_i}$ für alle $i \in \{1, \ldots, r\}$.

Die Familie

$$(P_1(X)^{m_1}, \ldots, P_r(X)^{m_r}) \tag{39}$$

ist durch f (bis auf die Numerierung) eindeutig bestimmt. Sie heißt die Familie der Elementarteiler *von f.*

(ii) Sind die v_i und $P_i(X)^{m_i}$ gegeben wie in (i), so ist

$$\chi_f(X) = (-1)^n \prod_{i=1}^{r} P_i(X)^{m_i}$$

das charakteristische Polynom von f. Ist

$$P_i(X)^{M_i}$$

die höchste vorkommende Potenz von $P_i(X)$ und sind $P_1(X), \ldots, P_k(X)$ die in der Familie vorkommenden <u>verschiedenen</u> Primpolynome, so ist

$$\mu_f(X) = \prod_{i=1}^{k} P_i(X)^{M_i}$$

das Minimalpolynom von f.

Beweis.

<u>*Zu (i).*</u> Die Existenz folgt unmittelbar aus den Sätzen 29 und 31.

Zur Eindeutigkeit: Sei eine Zerlegung von V in f-zyklische Unterräume wie in (i) gegeben und seien $\mu_{(f,v_i)} = P_i(X)^{m_i}$ die zugehörigen Minimalpolynome. Sei M_i der höchste vorkommende Exponent von $P_i(X)$, und sei die Numerierung so gewählt, daß $P_1, \ldots, P_k$ die *paarweise verschiedenen* Primpolynome sind. Dann ist

$$\mu_f(X) = \prod_{i=1}^{k} P_i(X)^{M_i}.$$

Wir setzen $V_i := \mathrm{Kern}\, P_i(f)^{M_i}$. Dann ist

$$V = \bigoplus_{1}^{k} V_i$$

die Primzerlegung zu f. Die $P_i(X)$ sind also die normierten Primteiler von μ_f. Weiter gilt

$$P_j|P_i \implies P_j = P_i \implies Z(f, v_j) \subset V_i.$$

Da sowohl die $Z(f, v_j)$ wie die V_i eine direkte Summenzerlegung von V bilden, folgt

$$V_i = \bigoplus_{P_j|P_i} Z(f, v_j).$$

Aus dem Eindeutigkeitsteil von Satz 30 angewendet auf $f|_{V_i}$ folgt also die Eindeutigkeit der auftretenden Potenzen von $P_i(X)$.

Zu (ii). Klar. □

Korollar 4. *Jeder Primteiler des charakteristischen Polynoms $\chi_f(X)$ ist auch Teiler des Minimalpolynoms $\mu_f(X)$.*

Wir untersuchen noch die Matrixversion der rationalen Normalform. Betrachtet man in V die Basis, die sich aus den offensichtlichen f-zyklischen Basen der $Z(f, v_i)$ zusammensetzt, so hat die Darstellungsmatrix von f also Blockdiagonalgestalt

$$\begin{pmatrix} B_1 & & & 0 \\ & B_2 & & \\ & & \ddots & \\ 0 & & & B_r \end{pmatrix}.$$

Die Blöcke sind von der Form

$$B_i = \begin{pmatrix} 0 & 0 & \dots & 0 & -c_{i0} \\ 1 & 0 & \dots & 0 & -c_{i1} \\ 0 & \ddots & & \vdots & \vdots \\ \vdots & & \ddots & 0 & \vdots \\ 0 & \dots & 0 & 1 & -c_{i\,d_i-1} \end{pmatrix} \tag{40}$$

mit

$$d_i := \dim Z(f, v_i) = \operatorname{Grad} P(X)^{m_i}.$$

Das Minimalpolynom von B_i ist

$$\mu_{B_i}(X) = X^{d_i} + \sum_{k=1}^{d_i-1} c_k X^k = P(X)^{m_i}.$$

Daher ergibt sich aus Satz 29 und Satz 30

Satz 32 (Rationale Normalform: Matrixversion). *Sei $A \in M(n,n;K)$. Dann gibt es eine invertierbare Matrix $S \in M(n,n;K)$, so daß*

$$S^{-1}AS = \begin{pmatrix} B_1 & & & 0 \\ & B_2 & & \\ & & \ddots & \\ 0 & & & B_r \end{pmatrix}$$

mit Blöcken der Form

$$B_i = \begin{pmatrix} 0 & 0 & \dots & 0 & -c_{i0} \\ 1 & 0 & \dots & 0 & -c_{i1} \\ 0 & \ddots & & \vdots & \vdots \\ \vdots & & \ddots & 0 & \vdots \\ 0 & \dots & 0 & 1 & -c_{i\,d_i-1} \end{pmatrix},$$

deren Minimalpolynome Potenzen der Primfaktoren von μ_A sind. Die B_i sind bis auf die Numerierung eindeutig bestimmt.
Jeder Endomorphismus eines n-dimensionalen K-Vektorraumes besitzt bezüglich einer geeigneten Basis eine Darstellungsmatrix dieser Form.

Beispiel 24. Sei $f = \mathrm{id}$ die Identität auf einem n-dimensionalen Vektorraum V. Sei $(v_1, \dots, v_n)$ eine beliebige Basis. Das Minimalpolynom ist

$$\mu_{\mathrm{id}}(X) = X - 1.$$

Die rationale Normalform ist gegeben durch

$$V = Z(f, v_1) \oplus \dots \oplus Z(f, v_n) = Kv_1 \oplus \dots \oplus Kv_n$$

mit den zugehörigen Minimalpolynomen $\mu_{(\mathrm{id},v_i)}(X) = X - 1$. Es ist also

$$r = n \text{ und } m_1 = \dots = m_r = 1.$$

□

Beispiel 25. Der Endomorphismus des $\mathbb{R}^3$ mit der Matrix

$$A = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$$

hat das Minimalpolynom

$$\mu_A(X) = (X - \lambda)^3 = X^3 - 3\lambda X^2 + 3\lambda^2 X - \lambda^3.$$

Er besitzt daher eine Darstellungsmatrix der Form

$$S^{-1}AS = \begin{pmatrix} 0 & 0 & \lambda^3 \\ 1 & 0 & -3\lambda^2 \\ 0 & 1 & 3\lambda \end{pmatrix}.$$

Eine Basis dafür ist gegeben durch die A-zyklische Basis

$$\begin{aligned} v_1 &= e_3 \\ v_2 &= Av_1 = e_2 + \lambda e_3 \\ v_3 &= Av_2 = e_1 + \lambda e_2 + \lambda(e_2 + \lambda e_3) = e_1 + 2\lambda e_2 + \lambda^2 e_3. \end{aligned}$$

Die Matrix S enthält diese Vektoren als Spalten:

$$S = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 2\lambda \\ 1 & \lambda & \lambda^2 \end{pmatrix}$$

□

Beispiel 26. Die Matrix A aus Beispiel 19 hat entsprechend die rationale Normalform

$$A = \begin{pmatrix} \begin{bmatrix} \lambda \end{bmatrix} & & & & \\ & \begin{bmatrix} 0 & -\lambda^2 \\ 1 & 2\lambda \end{bmatrix} & & & \\ & & \begin{bmatrix} 0 & 0 & \lambda^3 \\ 1 & 0 & -3\lambda^2 \\ 0 & 1 & 3\lambda \end{bmatrix} & & \\ & & & \begin{bmatrix} 0 & -\mu^2 \\ 1 & 2\mu \end{bmatrix} \end{pmatrix}.$$

□

4.5 Ähnlichkeitsklassifikation

Wir erinnern an folgende Definition

Definition 22 (Ähnlichkeit). Wir nennen $f, g \in \mathrm{End}(V)$ *ähnlich*, wenn es *einen* Automorphismus $\phi \in \mathrm{Aut}(V)$ gibt, so daß

$$g = \phi^{-1} f \phi.$$

Ensprechend nennen wir zwei (n,n)-Matrizen A und B *ähnlich*, wenn es eine invertierbare (n,n)-Matrix S gibt, so daß

$$A = S^{-1} B S.$$

Zwei Endomorphismen von V sind mit anderen Worten genau dann ähnlich, wenn sie bezüglich geeigneter Basen dieselbe Darstellungsmatrix besitzen.

Die Resultate der beiden letzten Abschnitte erlauben es, die Endomorphismen bis auf Ähnlichkeit vollständig zu klassifizieren.

Satz 33 (Ähnlichkeitsklassifikation). *Sei V ein n-dimensionaler K-Vektorraum.*

(i) Zwei Endomorphismen von V sind genau dann ähnlich, wenn sie bei geeigneter Numerierung dieselbe Familie von Elementarteilern haben.

(ii) Jede Familie $(P_1(X)^{m_1}, \ldots, P_r(X)^{m_r})$ normierter Primpolynompotenzen mit

$$\sum_{i=1}^{r} \mathrm{Grad}\, P_i(X)^{m_i} = \dim V \tag{41}$$

ist die Elementarteilerfamilie eines Endomorphismus von V.

Man sagt, die Elementarteiler sind ein vollständiges Invariantensystem für die Ähnlichkeitsklassifizierung von Endomorphismen.

Beweis. *Zu (i).* Die Elementarteiler bestimmen die Darstellungsmatrix auf „dem“ jeweils zugehörigen f-zyklischen Unterraum: Wählt man in $Z(f, v_i)$ die kanonische zyklische Bassis, so hat $f|_{Z(f,v_i)}$ die Darstellungsmatrix

$$\begin{pmatrix} 0 & 0 & \ldots & 0 & -c_0 \\ 1 & 0 & \ldots & 0 & -c_1 \\ 0 & \ddots & & \vdots & \vdots \\ \vdots & & \ddots & 0 & \vdots \\ 0 & \ldots & 0 & 1 & -c_{n-1} \end{pmatrix}. \tag{42}$$

Dabei sind die c_k die Koeffizienten von $P_i(X)^{m_i}$. Haben f und g dieselben Elementarteiler, so haben sie bezüglich geeigneter Basen also dieselbe Darstellungsmatrix und sind damit einander ähnlich.

Umgekehrt haben ähnliche Endomorphismen dieselben Elementarteiler. Um das zu sehen, beachten Sie, daß $(\phi^{-1} f \phi)^k = \phi^{-1} f^k \phi$, also

$$G(\phi^{-1} f \phi) = \phi^{-1} G(f) \phi$$

für alle $G(X) \in K[X]$. Daher ist

$$Z(\phi^{-1} f \phi, \phi^{-1}(v)) = \phi^{-1} Z(f, v)$$

und für $G(X) \in K[X]$ gilt

$$G(f)(v) = 0 \iff G(\phi^{-1} f \phi)(\phi^{-1}(v)) = 0.$$

Nach Lemma 21 ist daher $\mu_{(f,v)}(X) = \mu_{(\phi^{-1} f \phi, \phi^{-1}(v))}(X)$. Der „$f$-Zerlegung“

$$V = Z(f, v_1) \oplus \ldots \oplus Z(f, v_r)$$

mit den Minimalpolynomen $P_i(X)^{m_i}$ entspricht daher die „$\phi^{-1} f \phi$-Zerlegung“

$$V = Z(\phi^{-1} f \phi, \phi^{-1}(v_1)) \oplus \ldots \oplus Z(\phi^{-1} f \phi, \phi^{-1}(v_r))$$

mit *denselben* Minimalpolynomen. Daher sind die Elementarteiler von f und $\phi^{-1} f \phi$ gleich.

Zu (ii). Ist eine Familie von normierten Primpolynompotenzen gegeben, so kann man eine Blockdiagonalmatrix mit Blöcken (42) hinschreiben, die gerade diese Familie als Elementarteiler hat. Nach Wahl einer beliebigen Basis von V definert diese Matrix einen Endomorphismus mit den gewünschten Elementarteilern. □

Beispiel 27. Die beiden reellen Matrizen

$$A = \begin{pmatrix} \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} & & \\ & \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} & \\ & & \end{pmatrix}, \qquad B = \begin{pmatrix} \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} & & \\ & [2] & \\ & & [2] \end{pmatrix}$$

haben dasselbe charakteristische Polynom und dasselbe Minimalpolynom

$$\chi_A(X) = (X-2)^4 = \chi_B(X), \qquad \mu_A(X) = (X-2)^2 = \mu_B(X).$$

Aber ihre Elementarteiler sind verschieden, nämlich

$$((X-2)^2, (X-2)^2)$$

bzw.

$$((X-2)^2, X-2, X-2).$$

Sie sind also nicht ähnlich.

□

Beispiel 28. Gegeben seien die Elementarteiler $X^2 - 2X + 2, (X^2 - 2X + 2)^2, X + 3$ über $K = \mathbb{R}$. Sie gehören zu Endomorphismen f des $\mathbb{R}^7$ mit dem charakteristischen Polynom

$$\chi_f(X) = -(X^2 - 2X + 2)^3 (X + 3)$$

und dem Minimalpolynom

$$\mu_f(X) = -(X^2 - 2X + 2)^2 (X + 3).$$

Eine Darstellungsmatrix ist wegen

$$(X^2 - 2X + 2)^2 = X^4 - 4X^3 + 8X^2 - 8X + 4$$

gegeben durch

$$\begin{pmatrix} \begin{bmatrix} 0 & 0 & 0 & -4 \\ 1 & 0 & 0 & 8 \\ 0 & 1 & 0 & -8 \\ 0 & 0 & 1 & 4 \end{bmatrix} & & \\ & \begin{bmatrix} 0 & -2 \\ 1 & 2 \end{bmatrix} & \\ & & -3 \end{pmatrix}.$$

□

4.6 Jordansche Normalform

Die Darstellungsmatrix von $f \in \mathrm{End}(V)$ bezüglich einer f-zyklische Basis $(v, f(v), \ldots, f^{n-1}(v))$ von V war von der Form

$$\begin{pmatrix} 0 & 0 & \ldots & 0 & -c_0 \\ 1 & 0 & \ldots & 0 & -c_1 \\ 0 & \ddots & & \vdots & \vdots \\ \vdots & & \ddots & 0 & \vdots \\ 0 & \ldots & 0 & 1 & -c_{n-1} \end{pmatrix}$$

mit dem Minimalpolynom

$$\mu_f(X) = X^n + \sum_{k=0}^{n-1} c_k X^k.$$

Falls das Minimalpolynom von der Form $\mu_f(X) = (X - \lambda)^n$ ist, kann man eine andere Basis wählen, die noch eine einfachere Darstellungsmatrix produziert, vgl. das folgende Lemma.

Ist allgemeiner f ein Endomorphismus mit in Linearfaktoren zerfallendem Minimalpolynom, so kann man die rationale Normalform aus Satz 32 entsprechend vereinfachen und erhält die sogenannte *Jordansche Normalform*.

Lemma 22. *Sei $f \in \mathrm{End}(V)$ mit einer f-zyklische Basis $(v, f(v), \ldots, f^{n-1}(v))$ von V. Sei*

$$\mu_f = (X - \lambda)^n.$$

Dann ist auch

$$((f - \lambda \,\mathrm{id})^{n-1}(v), \ldots, (f - \lambda \,\mathrm{id})(v), v) \tag{43}$$

eine Basis von V, und bezüglich dieser hat f die Darstellungsmatrix

$$\begin{pmatrix} \lambda & 1 & 0 & \ldots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \ldots & \ldots & 0 & \lambda \end{pmatrix}.$$

Matrizen dieser Form nennt man elementare Jordanmatrizen.

Beweis. Wären die Vektoren (43) linear abhängig, so gäbe es ein nicht-triviales Polynom $G(X) \in K[X]$ vom Grad $\leq n - 1$, für das

$$G(f)(v) = 0.$$

Nach Lemma 21 wäre dann $\mathrm{Grad}\, \mu_f \leq n - 1$. Widerspruch!

Die Form der Darstellungsmatrix folgt aus

$$f(f - \lambda \,\mathrm{id})^j(v) = (f - \lambda \,\mathrm{id})^{j+1}(v) + \lambda(f - \lambda \,\mathrm{id})^j(v)$$

und $(f - \lambda \,\mathrm{id})^n = 0$. □

Nach Korollar 4 zerfällt das charakteristische Polynom von $f \in \mathrm{End}(V)$ genau dann in Linearfaktoren, wenn das Minimalpolynom in Linearfaktoren zerfällt. Nach Satz 28 ist das äquivalent zur Trigonalisierbarkeit von f. In diesem Fall sind die Elementarteiler von f von der Form $(X - \lambda_i)^{m_i}$, und man kann in Satz 32 die Blöcke einfacher gestalten.

Satz 34 (Jordansche Normalform). *Das charakteristische Polynom von $f \in \mathrm{End}(V)$ zerfalle in Linearfaktoren. Dann gibt es eine Basis $\mathbf{b}$ von V, so daß*

$$A_{\mathbf{b}}^{\mathbf{b}}(f) = \begin{pmatrix} J_1 & & & 0 \\ & J_2 & & \\ & & \ddots & \\ 0 & & & J_r \end{pmatrix} \tag{44}$$

eine Blockdiagonalmatrix mit elementaren Jordanmatrizen

$$J_i = \begin{pmatrix} \lambda_i & 1 & & 0 \\ 0 & \lambda_i & \ddots & \\ & \ddots & \ddots & 1 \\ 0 & & 0 & \lambda_i \end{pmatrix}.$$

als Blöcken ist. Diese Darstellung ist eindeutig bis auf die Reihenfolge der Blöcke. Sie heißt die Jordansche Normalform *von f.*
Jede quadratische Matrix mit in Linearfaktoren zerfallendem charakteristischen Polynom ist ähnlich zu einer solchen Matrix der Form (44).

Beweis. Erledigt. □

Bemerkungen.

1. Die λ_i sind natürlich die Eigenwerte von f, aber verschiedene der elementaren Jordanblöcke J_i können durchaus denselben Eigenwert enthalten. Der Eigenwert λ_i kommt in so vielen Blöcken vor, wie es seine geometrische Vielfachheit angibt.
2. Jeder Block J_i mit Diagonale λ_i und Zeilenzahl m_i entspricht einem Elementarteiler $(X - \lambda_i)^{m_i}$ und umgekehrt.
3. Die Zeilenzahl M_i des größten Jordanblocks mit Diagonale λ_i enthält, ist der Exponent von $(X - \lambda_i)$ im Minimalpolynom von f.
4. Insgesamt steht ein Eigenwert so oft auf der Diagonalen, wie seine algebraische Vielfachheit angibt.

Beispiel 29. Die Matrix

$$A = \begin{pmatrix} \begin{bmatrix}\lambda\end{bmatrix} & & & & \\ & \begin{bmatrix}\lambda & 1 \\ 0 & \lambda\end{bmatrix} & & & \\ & & \begin{bmatrix}\lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda\end{bmatrix} & & \\ & & & \begin{bmatrix}\mu & 1 \\ 0 & \mu\end{bmatrix} \end{pmatrix}$$

aus Beispiel 19 mit $\lambda \neq \mu$ ist bereits in Jordanscher Normalform. Das charakteristische Polynom

$$\chi_A(X) = (X - \lambda)^6 (X - \mu)^2$$

und das Minimalpolynom

$$\mu_A(X) = (X - \lambda)^3 (X - \mu)^2.$$

sind sofort ablesbar, ebenso die Elementarteiler

$$X-\lambda, (X-\lambda)^2, (X-\lambda)^3, (X-\mu)^2.$$

□

Beispiel 30. Bestimme die Jordansche Normalform der Matrix

$$A = \begin{pmatrix} 3 & 2 & 10 \\ -10 & 12 & 11 \\ -8 & 4 & 20 \end{pmatrix}.$$

Berechnung von $\det(A - XE)$ liefert

$$\chi_A(X) = -(X-7)(X-14)^2.$$

Das Minimalpolynom ist also

$$(X-7)(X-14) \text{ oder } (X-7)(X-14)^2.$$

Den Rang von

$$A - 14E = \begin{pmatrix} -11 & 2 & 10 \\ -10 & -2 & 11 \\ -8 & 4 & 6 \end{pmatrix}$$

kann man mit bloßem Auge bestimmen: Die erste Spalte plus die Hälfte der zweiten ist das Negative der dritten. Also ist der Rang =2 und der Eigenwert 14 hat die geometrische Vielfachheit $3-2=1$. Die Matrix A ist deshalb nicht diagonalisierbar. Darum ist $\mu_A(X) = (X-7)(X-14)^2$. Die zugehörige Familie von Elementarteiler ist

$$(X-7, (X-14)^2)$$

und die Jordansche Normalform

$$J = \begin{pmatrix} [7] & 0 \\ 0 & \begin{bmatrix} 14 & 1 \\ 0 & 14 \end{bmatrix} \end{pmatrix}.$$

□

Beispiel 31. Die konkrete Berechnung einer Basis, bezüglich der f in Jordanscher Normalform erscheint, ist schwierig. Zunächst bestimmt man die Eigenwerte λ_i.

Ist die algebraische Vielfachheit von λ_i gleich der geometrischen, so bestimmt man eine Basis für den Eigenraum $\mathrm{Eig}(f, \lambda_i)$.

Im Fall eines 2-fachen Eigenwertes λ_i mit geometrischer Vielfachheit 1 bestimmt man einen Eigenvektor v_i. Dann hat die Gleichung

$$(f - \lambda_i \,\mathrm{id}) w_i = v_i$$

eine Lösung w_i. Die Vektoren $(v_i = (f - \lambda_i \,\mathrm{id}) w_i, w_i)$ bilden dann eine Basis von

$$\mathrm{Kern}(f-\lambda_i)^2.$$

Ein Verfahren zur Berechnung der „Jordanbasis“ im allgemeinen Fall findet man z.B. in den neueren Auflagen des Fischer.

□

Beispiel 32 (Matrixexponential). Ein beliebtes Beispiel für die Anwendung der Jordanschen Normalform stammt aus der Theorie der gewöhnlichen Differentialgleichungen. Tatsächlich braucht man dafür allerdings nur die Primzerlegung und die Bemerkung aus Beispiel 22. Das Problem ist folgendes: Gegeben sei eine komplexe (n, n)-Matrix A. Gesucht sind Lösungen $x : \mathbb{R} \to \mathbb{C}^n$ für die Differentialgleichung

$$\frac{dx}{dt} = Ax.$$

Im Fall $n = 1$ und $A = (a)$ sind die Lösungen gegeben durch

$$x(t) = e^{ta}v,$$

wobei $v \in \mathbb{C}$ beliebig ist und den Anfangswert für $t = 0$ repräsentiert.

Für beliebiges n ist

$$x(t) = e^{tA}v$$

mit beliebigem Anfangswert $v \in \mathbb{C}^n$. Das Problem ist nur, daß für eine Matrix A die Berechnung von

$$e^{tA} = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k$$

ein erhebliches Problem bedeutet.

Ist allerdings v ein Eigenvektor von A zum Eigenwert λ, so ist

$$e^{tA}v = e^{t\lambda E + t(A - \lambda \,\mathrm{id})}v = e^{t\lambda}e^{t(A-\lambda E)}v.$$

(Dabei benutzt man wesentlich, daß die Matrizen λE und $(A - \lambda \,\mathrm{id})$ kommutieren; im allgemeinen ist $e^{A+B} \neq e^A e^B$!) Nun kommt der Vorteil dieser Zerlegung:

$$e^{t(A-\lambda E)}v = \sum_{k=0}^{\infty} \frac{t^k}{k!} (A - \lambda E)^k v = v + \underbrace{(A - \lambda E)v}_{=0} + \frac{t^2}{2!} \underbrace{(A - \lambda E)^2 v}_{=0} + \ldots = v,$$

so daß

$$x(t) = e^{t\lambda}v.$$

Ist also A diagonalisierbar, so findet man mit einer Basis von Eigenvektoren auf diese Weise n linear unabhängige Lösungen, aus denen sich jede belliebige Lösung dann linear kombinieren läßt.

Was tun, wenn A nicht diagonalisierbar ist? Aus dem Primzerlegungssatz folgt die Existenz einer Basis von Hauptvektoren von A. Ist aber v ein Hauptvektor zum Eigenwert λ, d.h. gilt

$$(A - \lambda E)^m v = 0$$

für ein $m > 0$, so folgt

$$e^{t(A-\lambda E)}v = \sum_{k=0}^{\infty} \frac{t^k}{k!} (A - \lambda E)^k v = \sum_{k=0}^{m-1} \frac{t^k}{k!} (A - \lambda E)^k v. \tag{45}$$

Man kommt also mit endlichen Summen aus. Aus dem charakteristischen Polynom von A ergeben sich die Eigenwerte, deren algebraische Vielfachheit man als m in der Hauptvektorgleichung nehmen kann, und die Bestimmung der Hauptvektorbasis erfordert das Lösen linearer Gleichungssysteme.

Kennt man eine Matrix S, für die $S^{-1}AS$ Jordansche Normalform hat, so sind die weiteren Rechnungen sehr einfach: die Spalten v_i dieser Matrix sind dann nämlich Hauptvektoren mit

$$(A - \lambda_i E)v_i = \begin{cases} v_{i+1} \\ oder \\ 0. \end{cases}$$

□

5 Vektorgeometrie

Die Geometrie, insbesondere die ebene Euklidische Geometrie, ist seit der Antike Gegenstand mathematischen Interesses gewesen. Aus der unendlichen Fülle von Resultaten geben wir hier eine ziemlich „gemischte“ Auswahl.

5.1 Affine Räume

Vektorräume sind für viele geometrische oder physikalische Sachverhalte eigentlich nicht die richtigen Modelle, weil sie einen ausgezeichneten Punkt, den Nullpunkt haben. Das „richtige“ mathematische Modell ist deshalb oft ein *affiner Raum*.

Definition 23. Seien V ein K-Vektorraum und $\mathcal{A}$ eine nicht-leere Menge. Mit $\mathcal{S}(\mathcal{A})$ bezeichnen wir die Permutationsgruppe von $\mathcal{A}$, d.h. die Gruppe der bijektiven Abbildungen von $\mathcal{A}$ auf sich. Sei

$$\tau : V \to \mathcal{S}(\mathcal{A}), x \mapsto \tau_x$$

ein Gruppenhomomorphismus der additiven Gruppe von V in $\mathcal{S}(\mathcal{A})$, also

$$\tau_{x+y}(Q) = \tau_x(\tau_y(Q)). \quad \text{und } \tau_{-x} = \tau_x^{-1},$$

Für jedes $Q \in \mathcal{A}$ sei weiter

$$\tau(Q) : V \to \mathcal{A}, x \mapsto \tau_x(Q) \quad \text{bijektiv.}$$

Die Umkehrabbildung bezeichnen wir mit $P \mapsto Q - P$, so daß also

$$\tau_x(Q) = P \iff P - Q = x.$$

Der Vektor $P-Q$ heißt der *Verbindungsvektor* von Q nach P und die Abbildung $\tau_x : \mathcal{A} \to \mathcal{A}$ *die Translation um* x.

Ein Tripel $(\mathcal{A}, V, \tau)$ aus einer Menge, einem Vektorraum und einer Translationsabbildung mit den obigen Eigenschaften heißt ein *affiner Raum* über V. Man sagt dann auch einfach, $\mathcal{A}$ sei ein affiner Raum. Den Vektorraum V nennt man den *Richtungsvektorraum* oder den *Tangentialraum* von $\mathcal{A}$. Die Dimension von $\mathcal{A}$ wird definiert als die Dimension von V.

Bemerkung. Nach Wahl eines Punktes $O \in \mathcal{A}$, eines *Ursprungs*, liefert

$$\tau(O) : V \to \mathcal{A}$$

eine Bijektion. Also „ist“ ein affiner Raum *fast* ein Vektorraum. Aber jeder gewählte Ursprung liefert einen anderen Vektorraum.

Beispiel 33. Affine Unterräume eines Vektorraums sind affine Räume: Ist $U \subset V$ ein Untervektorraum, so ist $\mathcal{A} := v + U$ für jedes $v \in V$ ein affiner Raum über U. Die Translationsabbildung ist für $x \in U$ und $Q \in v + U$ gegeben durch die Addition in V:

$$\tau_x(Q) := Q + x.$$

Insbesondere (Fall $U = V$) ist jeder Vektorraum ein affiner Raum.

□

Definition 24. Eine nicht-leere Teilmenge $\mathcal{A}' \subset \mathcal{A}$ heißt ein *affiner Unterraum* des affinen Raumes $(\mathcal{A}, V, \tau)$, wenn es einen Vektorunterraum $V' \subset V$ gibt, so daß für ein (und dann für alle) $Q' \in \mathcal{A}'$ gilt

$$\mathcal{A}' = \{\tau_{x'}(Q') \,|\, x' \in V'\}.$$

Dann ist $(\mathcal{A}', V', \tau|_{V'} : V' \to \mathcal{S}(\mathcal{A}'))$ ein affiner Raum über V'.

Ein affiner Unterraum der Dimension 1 heißt eine *(affine) Gerade*, ein affiner Unterraum der Dimension 2 eine *(affine) Ebene*.

Zwei Geraden heißen *parallel*, wenn sie denselben Richtungsvektorraum haben.

Bemerkung. In einem affinen Raum kann man zwar den Verbindungsvektor $x = P - Q$ zweier Punkte bilden, aber man kann Punkte nicht addieren. „Addieren" kann man Vektoren $x \in V$ zu Punkten $Q \in \mathcal{A}$, indem man $\tau_x(Q)$ bildet. Wir schreiben auch

$$\tau_x(Q) =: Q + x.$$

Natürlich kann man erst recht keine Linearkombinationen von Punkten aus $\mathcal{A}$ bilden, aber es gibt Ausnahmen:

Satz 35 (und Definition: Baryzentrischer Kalkül). *Seien $P_1, \ldots, P_n \in \mathcal{A}$ Punkte des affinen Raums $\mathcal{A}$ und seien $\lambda_1, \ldots, \lambda_n \in K$. Es gelte*

$$\sum_{i=1}^{n} \lambda_i = 1.$$

Sei weiter $Q \in \mathcal{A}$ beliebig. Dann ist

$$\sum_{i=1}^{n} \lambda_i P_i := Q + \sum_{i=1}^{n} \lambda_i (P_i - Q)$$

unabhängig von der Wahl von Q. Es heißt der Schwerpunkt *(oder das* Baryzentrum*)* der Punkte $P_1, \ldots, P_n \in \mathcal{A}$ mit Gewichten $\lambda_1, \ldots, \lambda_n \in K$ *oder eine* baryzentrische Linearkombination *der Punkte P_i.*
Ist $\frac{1}{n} \in K$ (ist z.B. $K = \mathbb{R}$) und $\lambda_i = \frac{1}{n}$ für alle i, so bekommt man den vertrauten geometrischen Schwerpunkt.

$$\sigma(P_1, \ldots, P_n) := \sum_{k=1}^{n} \frac{1}{n} P_k.$$

(Beachten Sie, daß der Ausdruck $\sum_{k=1}^{n} P_k$ und damit auch $\frac{1}{n} \sum_{k=1}^{n} P_k$ nicht definiert ist.)

Beweis der Unabhängigkeit von Q. Sei $P \in \mathcal{A}$ ein weiterer Punkt. Dann ist

$$\begin{aligned}
Q + \sum_{i=1}^{n} \lambda_i (P_i - Q) &\underset{(*)}{=} Q + \sum_{i=1}^{n} \lambda_i ((P_i - P) + (P - Q)) \\
&= Q + \sum_{i=1}^{n} \lambda_i (P_i - P) + \left(\sum_{i=1}^{n} \lambda_i \right) (P - Q) \\
&= Q + (P - Q) + \sum_{i=1}^{n} \lambda_i (P_i - P) \\
&\underset{(**)}{=} P + \sum_{i=1}^{n} \lambda_i (P_i - P).
\end{aligned}$$

Dabei haben wir in der Gleichung (**) benutzt, daß

$$Q + (P - Q) = \tau_{P-Q}(Q) = \tau_{\tau(Q)^{-1}(P)}(Q) = \tau(Q)\tau(Q)^{-1}(P) = P.$$

Die Gleichung (*) folgt aus

$$\tau_{(A-P)+(P-Q)}(Q) = \tau_{(A-P)}\tau_{(P-Q)}(Q) = \tau_{(A-P)}(P) = A = \tau_{A-Q}(Q),$$

denn daraus ergibt sich wegen der Bijektivität von $\tau(Q)$, daß

$$(A-P)+(P-Q) = A-Q.$$

□

Ebenso kann man zeigen, daß

$$\sum \lambda_i P_i := \sum \lambda_i (P_i - Q)$$

unabhängig ist von der Wahl von Q, falls $\sum \lambda_i = 0$. Das Resultat ist dann allerdings nicht ein Punkt von $\mathcal{A}$, sondern ein Vektor in V.

Definition 25 (Allgemeine Lage, Dreieck). Die Punkte $P_0, \ldots, P_k \in \mathcal{A}$ eines affinen Raumes heißen *in allgemeiner Lage*, wenn die Vektoren $P_1 - P_0, \ldots, P_k - P_0 \in V$ linear unabhängig sind. Drei Punkte sind in allgemeiner Lage, wenn sie nicht auf einer Geraden liegen. Andernfalls heißen sie *kollinear*. Ein *Dreieck* in einem affinen Raum ist ein Tripel (A, B, C) von Punkten in allgemeiner Lage.

Satz 36. *Ist $n = \dim V$ und sind $P_0, \ldots, P_n \in \mathcal{A}$ in allgemeiner Lage, so läßt sich jedes $P \in \mathcal{A}$ eindeutig darstellen als*

$$P = \sum_{i=0}^{n} \lambda_i P_i$$

mit $\sum_{i=1}^{n} \lambda_i = 1$.

Beweis. Selbst. □

Satz 37. *Für $P_0, \ldots, P_k \in \mathcal{A}$ ist*

$$\{\sum_{i=0}^{k} \lambda_i P_i \,|\, \sum_{i=0}^{k} \lambda_i = 1\}$$

ein affiner Unterraum über dem Vektorraum

$$\text{Spann}(P_1 - P_0, \ldots, P_k - P_0)$$

Beweis. Selbst □

Definition 26 (Verbindungsgerade, Strecke, Mittelpunkt). Seien $P, Q \in \mathcal{A}$ zwei Punkte eines affinen Raums.

(i) Falls $P \neq Q$, heißt

$$g(PQ) = \{(1-\lambda)P + \lambda Q \,|\, \lambda \in K\}$$

die *Gerade durch P und Q* oder die *Verbindungsgerade von P und Q*.

(ii) Falls $K = \mathbb{R}$ heißt

$$PQ = \{(1-\lambda)P + \lambda Q \,|\, 0 \leq \lambda \leq 1\}$$

die *Strecke zwischen P und Q*.

(iii) Falls $K = \mathbb{R}$ heißt

$$\sigma(P,Q) = \frac{1}{2}P + \frac{1}{2}Q = P + \frac{1}{2}(P-P) + \frac{1}{2}(Q-P) = P + \frac{1}{2}(Q-P)$$

der *Mittelpunkt der Strecke* PQ.

Lemma 23. *Für* $P_1, \ldots, P_n \in \mathcal{A}$ *und* $1 \leq k < n$ *ist*

$$\frac{k}{n}\sigma(P_1,\ldots,P_k) + \frac{n-k}{n}\sigma(P_{k+1},\ldots,P_n) = \sigma(P_1,\ldots,P_n).$$

Beweis. Für $Q \in \mathcal{A}$ ist

$$\begin{aligned}
&\frac{k}{n}\sigma(P_1,\ldots,P_k) + \frac{n-k}{n}\sigma(P_{k+1},\ldots,P_n) \\
&\qquad = Q + \frac{k}{n}\left(\sigma(P_1,\ldots,P_k) - Q\right) + \frac{n-k}{n}\left(\sigma(P_{k+1},\ldots,P_n) - Q\right) \\
&\qquad = Q + \left(\frac{k}{n}\sum_1^k \frac{1}{k}(P_i - Q) + \frac{n-k}{n}\sum_{k+1}^k \frac{1}{n-k}(P_i - Q)\right) \\
&\qquad = Q + \sum_1^n \frac{1}{n}(P_i - Q) \\
&\qquad = \sigma(P_1,\ldots,P_n).
\end{aligned}$$

□

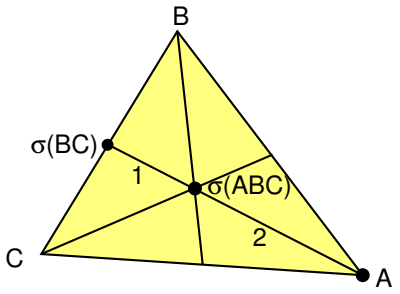
Satz 38 (Schwerpunktsatz). *Für* $A, B, C \in \mathcal{A}$ *ist*

$$\frac{1}{3}A + \frac{2}{3}\sigma(B,C) = \sigma(A,B,C).$$

Den Schwerpunkt der drei Punkte A, B, C *kann man also ermitteln, indem man zunächst den Mittelpunkt* $\sigma(B,C)$ *der Strecke* BC *ermittelt und dann den Schwerpunkt von* A *und* $\sigma(B,C)$ *mit den Gewichten* $\frac{1}{3}$ *und* $\frac{2}{3}$ *bildet. Insbesondere liegt* $\sigma(ABC)$ *auf der Geraden*

$$g(A\sigma(BC)),$$

der Seitenhalbierenden *von* BC *und teilt die Strecke* $A\sigma(BC)$ *im Verhältnis* $\frac{1}{3} : \frac{2}{3} = 1 : 2$ *teilt. Da dies für jede Seitenhalbierende gleichermaßen gilt, schneiden sich insbesondere die drei Seitenhalbierenden in einem Punkt.*



Definition 27 (Affine Abbildungen). Seien $(\mathcal{A}, V, \tau)$ und $(\mathcal{A}', V', \tau)$ zwei affine Räume zu Vektorräumen V und V' über demselben Körper K. Eine Abbildung $f : \mathcal{A} \to \mathcal{A}'$ heißt *affin*, wenn

$$f(\sum_{i=0}^n \lambda_i P_i) = \sum_{i=0}^n \lambda_i f(P_i)$$

für alle baryzentrischen Linearkombinationen.

Satz 39. *Seien $(\mathcal{A}, V, \tau)$ und $(\mathcal{A}', V', \tau)$ zwei affine Räume wie in der Definition.*

(i) Seien $\phi : V \to V'$ linear und $Q \in \mathcal{A}, Q' \in \mathcal{A}'$. Dann ist

$$f : \mathcal{A} \to \mathcal{A}', P \mapsto \tau_{\phi(P-Q)}(Q')$$

eine affine Abbildung mit $f(Q) = Q'$.

(ii) Jede affine Abbildung ist von dieser Form: Ist $f : \mathcal{A} \to \mathcal{A}'$ affin und $Q \in \mathcal{A}$, so ist

$$\phi : V \to V', x \mapsto f(\tau_x(Q)) - f(Q)$$

linear und unabhängig von der Wahl von Q. Es gilt

$$f(P) = \tau_{\phi(P-Q)}(f(Q)).$$

Beweis. Selbst □

5.2 Euklidische affine Räume

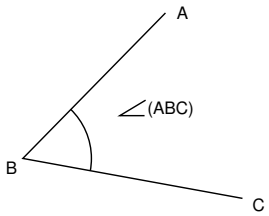
Definition 28. Ein *Euklidischer affiner Raum* $(\mathcal{A}, V, \tau)$ ist ein affiner Raum über einem Euklidischen Vektorraum $(V, \langle ., . \rangle)$. In diesem Fall definiert man den *Abstand* von Punkten $A, B \in \mathcal{A}$ oder die *Länge der Strecke AB* durch

$$|AB| := |A - B| := \sqrt{\langle A - B, A - B \rangle}.$$

(Wir schreiben $|.|$ und nicht $\|.\|$ für die Norm.)

Für drei Punkte $A \neq B \neq C$ definiert man den Winkel $\angle(ABC) \in [0, \pi]$ zwischen den Strecken BA und BC durch

$$\cos \angle(ABC) = \cos \angle(A-B, C-B) = \frac{\langle A - B, C - B \rangle}{|A - B||C - B|}.$$



Zwei Geraden $g = P + \mathbb{R}v$ und $h = P + \mathbb{R}w$ mit Schnittpunkt P schneiden sich unter dem Winkel

$$\angle(v, w).$$

Sie stehen senkrecht aufeinander, wenn $\angle(v, w) = \pi/2$.

Lemma 24 (Translationsinvarianz). *Sei $(\mathcal{A}, V, \tau)$ ein Euklidischer affiner Raum. Dann ist für jedes $x \in V$ die Translation $\tau_x : \mathcal{A} \to \mathcal{A}$ längen- und winkeltreu. Das heißt, es gilt für alle $A, B, C \in V$*

$$|\tau_x(A) - \tau_x(B)| = |A - B|$$

und, falls $A \neq B \neq C$,

$$\angle(\tau_x(A)\tau_x(B)\tau_x(C)) = \angle(ABC).$$

Beweis. Selbst □

In der *ebenen* Euklidischen Geometrie braucht man häufig eine Senkrechte zu einer gegebenen Geraden. Dafür ist eine Drehung um 90^0 hilfreich, die wir im folgenden Lemma beschreiben.

Lemma 25 (Komplexe Struktur). *Sei $(V, \langle ., . \rangle)$ ein zweidimensionaler Euklidischer Vektorraum. Dann gibt es genau zwei orthogonale Abbildungen J von $(V, \langle ., . \rangle)$ mit*

$$J^2 = -\operatorname{id}.$$

Diese unterscheiden sich nur durch das Vorzeichen. Jedes solche J nennt man eine komplexe Struktur *für $(V, \langle ., . \rangle)$. Ist V orientiert, so gibt es genau eine komplexe Struktur für $(V, \langle ., . \rangle)$, so daß (a, Ja) für alle $a \neq 0$ eine positiv orientierte Basis ist.*

Für $V = \mathbb{R}^2 = \mathbb{C}$ ist die Multiplikation mit i eine komplexe Struktur, daher kommt der Name.

Beweis des Lemmas. Wir wählen eine Orthonormalbasis von V. Bezüglich dieser hat jede orthogonale Abbildung die Darstellungsmatrix der Form

$$A = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \quad \text{oder} \quad B = \begin{pmatrix} \cos\phi & \sin\phi \\ \sin\phi & -\cos\phi \end{pmatrix}.$$

Wir haben früher gezeigt, daß B diagonalisierbar ist, also ist $B^2 \neq -E$. Die Matrix A ist eine Drehung um den Winkel ϕ. Also ist $A^2 = -E$ genau dann, wenn $\phi = \pm\frac{\pi}{2} \mod 2\pi$. Die Matrix einer komplexen Struktur J ist also bezüglich einer ONB gegeben durch

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \cos\pi/2 & -\sin\pi/2 \\ \sin\pi/2 & \cos\pi/2 \end{pmatrix}$$

oder deren Negatives. □

Mittels einer komplexen Struktur kann findet man zu einem beliebigen Vektor a einen dazu orthogonalen Ja. Das erlaubt es, gewisse geometrische Sachverhalte bequem rechnerisch beschreiben.

Beispiel 34. Gegeben seien zwei Punkte A, B und eine Gerade $g = P + \mathbb{R}v$ in einer Euklidischen Ebene. Dann gilt

$$A \text{ und } B \text{ liegen auf derselben Seite der Geraden } g \iff \langle Jv, A-P\rangle\langle Jv, B-P\rangle \geq 0 \quad (46)$$

für eine (und dann auch für die andere) komplexe Struktur J.

□

Vereinbarung. Wenn ich im zweidimensionalen Euklidischen Vektorraum zu einem Vektor explizit einen dazu orthogonalen brauche, werde ich gewöhnlich eine komplexe Struktur J benutzen, ohne vorab zu erklären, daß ich jetzt eine (von den beiden) auswähle. Ich gestatte mir das immer dann kommentarlos, wenn es egal ist, *welche* ich wähle, d.h. wenn die Orientierung keine Rolle spielt.

Satz 40 (und Definition: Lot). *Sei $\mathcal{A}$ eine Euklidische Ebene. Seien $g \subset \mathcal{A}$ eine Gerade und $P \in \mathcal{A}$. Dann gibt es genau eine Gerade, die senkrecht auf g steht und durch P geht. Sie heißt* das Lot $\perp_g (P)$ von P auf g. *Das Lot schneidet g in genau einem Punkt Q, dem* Fußpunkt des Lotes von P auf g. *Wir nennen $|Q - P|$ den* Abstand der Geraden g vom Punkt P. *Für alle $X \in g$ gilt*

$$|X - P| \geq |Q - P|.$$

Beweis. Sei

$$g = R + \mathbb{R}v \text{ mit } v \in V, \|v\| = 1.$$

Dann ist

$$l = P + \mathbb{R}\, Jv$$

die einzige Gerade durch P senkrecht zu g. Berechnung des Schnittpunkts:

$$R + sv = P + tJv \iff R - P = (-s)v + tJv.$$

Aber (v, Jv) ist eine ONBasis von V, und deshalb gibt es ein eindeutig bestimmtes Paar (s, t) mit dieser Eigenschaft. Es ist

$$|Q - P| = t = |\langle Q - P, Jv\rangle|.$$

Für $X \in g$

$$|X-P|^2 = \langle X-P, v\rangle^2 + \langle X-P, Jv\rangle^2 \geq \langle X-P, Jv\rangle^2 \geq (\langle X-Q, Jv\rangle + \underbrace{\langle Q-P, Jv\rangle}_{=\pm|Q-P|})^2 \geq |Q-P|^2.$$

Daraus folgt die Ungleichung. □

5.3 Winkelsätze

Sei $(\mathcal{A}, V, \tau)$ eine Euklidische Ebene.

Lemma 26. *Seien $a, b \in V\backslash\{0\}$ und $\gamma = \angle(a, b)$. Dann gilt*

$$\sin\gamma = \frac{|\langle Ja, b\rangle|}{|a|\,|b|}.$$

Beweis. Es gilt

$$\begin{aligned}\sin^2\gamma &= 1 - \frac{\langle a, b\rangle^2}{|a|^2|b|^2} \\ &= \frac{1}{|a|^2|b|^2}(|a|^2|b|^2 - \langle a, b\rangle^2).\end{aligned}$$

Nun ist (a, Ja) eine orthogonale Basis von V und

$$|a|^2 b = \langle a, b\rangle a + \langle Ja, b\rangle Ja.$$

Also ist

$$|a|^2|b|^2 = \langle a, b\rangle^2 + \langle Ja, b\rangle^2.$$

Einsetzen liefert

$$\sin^2\gamma = \frac{1}{|a|^2|b|^2}\,\langle Ja, b\rangle^2.$$

Da für $0 \leq \gamma \leq \pi$ der Sinus nicht-negativ ist, folgt die Behauptung. □

Ein wichtiges Axiom der synthetischen ebenen Geometrie ist der folgende

Satz 41 (Winkelzerlegung). *Seien V ein 2-dimensionaler Euklidischer Vektorraum und $a, b, c \in V$ mit folgenden Eigenschaften*

(i) a und c liegen auf verschiedenen Seiten von $\mathbb{R}b$, d.h.

$$\langle a, Jb\rangle\langle c, Jb\rangle < 0,$$

(ii) a, b, c liegen in einer Halbebene, d.h. es gibt $v \in V\backslash\{0\}$ mit

$$a, b, c \in \{x \in V \mid \langle x, v\rangle \geq 0\}.$$

Dann gilt

$$\angle(a, b) + \angle(b, c) = \angle(a, c).$$

Beweis. Ist ist

$$\angle(a, b) + \angle(b, c) \leq \pi, \tag{47}$$

so braucht man nur zu zeigen, daß

$$\cos(\angle(a, b) + \angle(b, c)) = \cos\angle(a, c). \tag{48}$$

Die Ungleichung (47) ist zwar unter den gemachten Voraussetzungen anschaulich „klar“, aber der Beweis erfordert eine Fallunterscheidung:

1. Fall: $\langle a, b\rangle \geq 0$ und $\langle c, b\rangle \geq 0$. Dann ist $\angle(a, b) \leq \pi/2$ und $\angle(b, c) \leq \pi/2$. Also gilt (47) und wir zeigen (48). Es ist

$$\begin{aligned}
\cos(\angle(a, b) + \angle(b, c)) &= \cos\angle(a, b)\cos\angle(b, c) - \sin\angle(a, b)\sin\angle(b, c)\\
&= \frac{1}{|a||b|^2|c|}\left(\langle a, b\rangle\langle b, c\rangle - |\langle a, Jb\rangle\langle c, Jb\rangle|\right)\\
&= \frac{1}{|a||b|^2|c|}\left(\langle a, b\rangle\langle b, c\rangle + \langle a, Jb\rangle\langle c, Jb\rangle\right) = \frac{1}{|a||c|}\left(\langle a, c\rangle\right)\\
&= \cos\angle(a, c).
\end{aligned}$$

Dabei haben wir benutzt, daß $(\frac{b}{|b|}, \frac{Jb}{|b|})$ eine Orthonormalbasis ist.

2. Fall: $\langle a, b\rangle < 0$ und $\langle c, b\rangle \geq 0$. Wähle v wie im Satz, o.E. $|v| = 1$. Dann gilt

$$\underbrace{\langle a, b\rangle}_{<0} = \underbrace{\langle a, v\rangle\langle b, v\rangle}_{\geq 0} + \langle a, Jv\rangle\langle b, Jv\rangle$$

Also ist

$$\langle a, Jv\rangle\langle b, Jv\rangle < 0.$$

Wir können daher den ersten Fall auf a, v, b anwenden:

$$\angle(a, v) + \angle(v, b) = \angle(a, b).$$

Ebenso folgt

$$\angle(a, v) + \angle(v, c) = \angle(a, c).$$

und

$$\angle(v, b) + \angle(b, c) = \angle(v, c).$$

Insgesamt folgt damit

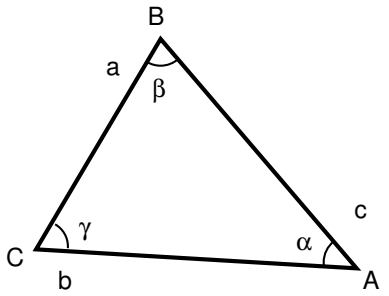
$$\angle(a, b) + \angle(b, c) = \angle(a, v) + \angle(v, b) + \angle(b, c) = \angle(a, v) + \angle(v, c) = \angle(a, c).$$

3. Fall: $\langle a, b\rangle \geq 0 > \langle c, b\rangle$. Analog.

4. Fall: $\langle a, b\rangle < 0$ und $\langle c, b\rangle < 0$. Dieser Fall kann nicht auftreten.

□

Konvention. Gegeben sei ein Dreieck (A, B, C) in $\mathcal{A}$. Wir nennen A, B, C die *Eckpunkte* des Dreiecks.
Unter den *Seiten* des Dreiecks verstehen wir je nach Zusammenhang die Strecken BC, AC und AB oder die entsprechenden Geraden $g(BC), g(AC)$ und $g(AB)$.



Wir vereinbaren folgende Bezeichnungen:

$$a := B - C, \quad b := C - A, \quad c := A - B$$
$$\alpha := \angle(CAB), \quad \beta := \angle(ABC), \quad \gamma = \angle(BCA).$$

Satz 42 (Winkelsumme im Dreieck). *Die Winkelsumme im Dreieck ist π: Mit den vereinbarten Bezeichnungen am Dreieck ist*

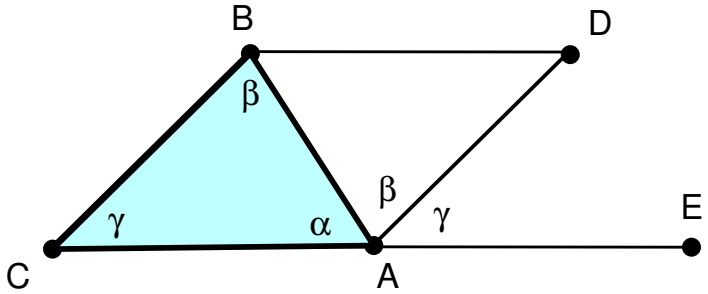
$$\alpha + \beta + \gamma = \pi.$$

Beweis. Wir definieren

$$D := A + B - C$$
$$E := 2A - C$$

und erhalten

$$B = A + (D - A) + (C - A)$$
$$D = A + (B - A) + (E - A)$$



Mit dem Winkelzerlegungssatz folgt daher

$$\angle(CAE) = \angle(CAB) + \angle(BAD) + \angle(DAE).$$

Aber

$$\angle(CAE) = \angle(C - A, E - A) = \angle(C - A, A - C) = \pi,$$
$$\angle(CAB) = \alpha,$$
$$\angle(BAD) = \angle(B - A, D - A) = \angle(B - A, B - C) = \angle(ABC) = \beta,$$
$$\angle(DAE) = \angle(D - A, E - A) = \angle(B - C, A - C) = \angle(BCA) = \gamma.$$

Daraus folgt die Behauptung. □

Satz 43 (Cosinussatz). *Im Dreieck (A, B, C) gilt*

$$|c|^2 = |a|^2 + |b|^2 - 2|a|\,|b| \cos\gamma.$$

Beweis.

$$|c|^2 = \langle A-B, A-B\rangle = \langle (B-C)+(C-A), (B-C)+(C-A)\rangle$$
$$= |a|^2 + |b|^2 - 2\langle a, b\rangle = |a|^2 + |b|^2 - 2|a|\,|b|\cos\angle(a,b).$$

□

Als Korollar erhält man den ehrwürdigsten Satz der Geometrie:

Satz 44 (Pythagoras). *Im rechtwinkligen Dreieck ist die Summe der Kathetenquadrate gleich dem Hypothenusenquadrat: Für $\gamma = \pi/2$ gilt*

$$|c|^2 = |a|^2 + |b|^2$$

Das ist klar, weil $\cos \pi/2 = 0$. Über andere Beweise des Pythagoras haben wir in den Übungen im letzten Semester gesprochen.

Satz 45 (Sinussatz). *Im Dreieck (A, B, C) gilt*

$$\frac{\sin\alpha}{|a|} = \frac{\sin\beta}{|b|} = \frac{\sin\gamma}{|c|}.$$

Beweis. Es gilt $a + b + c = 0$ und daher mit Lemma 26

$$\frac{\sin\alpha}{|a|} == \frac{|\langle Jb, c\rangle|}{|a|\,|b|\,|c|} = \frac{|-\langle Ja + Jc, c\rangle|}{|b|\,|c|} = \frac{|\langle Ja, c\rangle|}{|b|\,|c|} = \frac{\sin\beta}{|b|}.$$

Ebenso erhält man die zweite Gleichung. □

5.4 Kreise

In diesem Abschnitt sei $\mathcal{A}$ eine Euklidische Ebene.

Definition 29 (Mittelsenkrechte). Für zwei verschiedene Punkte $A, B \in \mathcal{A}$ heißt

$$m_{AB} := \{\sigma(A; B) + tJ(B-A) \mid t \in \mathbb{R}\}$$

die *Mittelsenkrechte* auf der Strecke AB.

Lemma 27. *Seien $A, B \in V$, $A \neq B$ und $P \in \mathcal{A}$. Dann gilt*

$$P \in m_{AB} \iff |P-A| = |P-B|.$$

Beweis. Es ist

$$P = \sigma(A, B) + s(B-A) + tJ(B-A).$$

Daher ist

$$|P-A| = |\frac{1}{2}A + \frac{1}{2}B - A + s(B-A) + tJ(B-A)|$$
$$= |(\frac{1}{2}+s)(B-A) + tJ(B-A)|$$
$$= |B-A|\sqrt{(\frac{1}{2}+s)^2 + t^2}.$$

Ebenso folgt

$$|P-B| = |B-A|\sqrt{(\frac{1}{2}-s)^2+t^2}.$$

Also

$$|P-A| = |P-B| \iff s = 0 \iff P \in m_{AB}.$$

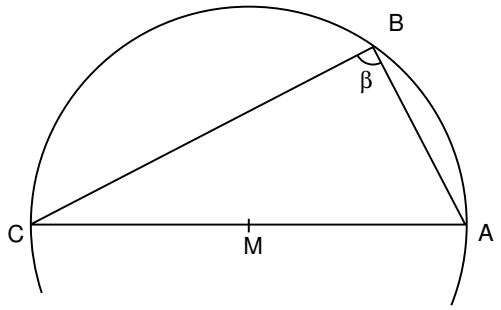
□

Definition 30 (Kreis). Der Kreis um $M \in \mathcal{A}$ mit Radius $r > 0$ ist die Menge

$$k = k_r(M) = \{P \in \mathcal{A} \,|\, |P-M| = r\}.$$

Satz 46 (Thales). *Sei (A,B,C) ein Dreieck. Setze $r := \frac{1}{2}|C-A|$ und $M = \sigma(C,A)$. Der Kreis $k = k_r(M)$ enthält dann A und C. Auch B liegt auf diesem Kreis genau dann, wenn der Winkel β bei B ein rechter ist:*

$$\beta = \frac{\pi}{2}.$$



Beweis. B liegt genau dann auch auf dem Kreis, wenn M auf der Mittelsenkrechten m_{BC} liegt, dh. wenn $\sigma(C,A) - \sigma(B,C)$ senkrecht auf CB steht, also wenn

$$\begin{aligned} 0 &= \langle \sigma(C,A) - \sigma(B,C), B-C\rangle \\ &= \langle \frac{1}{2}C + \frac{1}{2}A - \frac{1}{2}B - \frac{1}{2}C, B-C\rangle \\ &= \frac{1}{2}\langle A-B, B-C\rangle. \end{aligned}$$

Aber das bedeutet gerade, daß $\beta = \frac{\pi}{2}$. □

Lemma 28 (Kreistangente). *Seien k ein Kreis mit Mittelpunkt M und Radius r und $P \in k$. Dann gibt es genau eine Gerade g mit*

$$g \cap k = \{P\}.$$

Diese heißt die Tangente in P an k. *Es gilt*

$$g = \{P + tJ(P-M) \,|\, t \in \mathbb{R}\}.$$

Die Tangente steht also senkrecht auf dem „Radius“ MP.

Beweis. Sei

$$g = \{P + tv \,|\, t \in \mathbb{R}\}$$

eine Gerade durch P. Dann ist $P + tv \in k$ genau dann, wenn

$$|P + tv - M|^2 = t^2|v|^2 + 2t\langle P - M, v\rangle + \underbrace{|P - M|}_{=r}^2 = r^2,$$

also

$$t^2|v|^2 + 2t\langle P - M, v\rangle = 0.$$

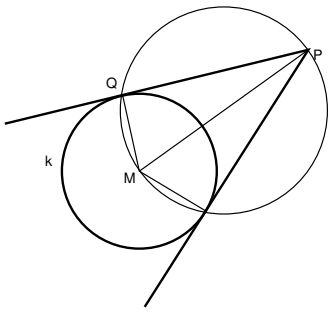
Die Gerade schneidet k nur in P genau dann, wenn diese Gleichung nur die Lösung $t = 0$ hat, also genau dann, wenn $\langle P - M, v\rangle = 0$. □

Satz 47 (Kreistangente). *Sei k ein Kreis mit Mittelpunkt M und Radius r, und sei P ein Punkt außerhalb des Kreises:*

$$|P - M| > r.$$

Dann gibt es genau zwei Tangenten an k durch P. Diese berühren k in den Schnittpunkten von k mit dem Kreis um $\sigma(M, P)$ durch M.

Beweis. Eine Gerade durch P ist Tangente an k im Punkt $Q \in k$ genau dann, wenn das Dreieck (P, Q, M) bei Q einen rechten Winkel hat. Nach dem Satz von Thales bedeutet das, daß Q ein Schnittpunkt von k mit dem „Thaleskreis“ um $\sigma(M, P)$ durch M ist. Der Satz folgt deshalb aus dem nachstehenden Lemma.



Lemma 29. *Zwei Kreise $k_0 = k_{r_0}(M_0)$ und $k_1 = k_{r_1}(M_1)$ mit*

$$|r_0 - r_1| < |M_0 - M_1| < r_0 + r_1$$

schneiden sich in genau zwei Punkten.

Beweis. Wir setzen $a := \frac{M_1 - M_0}{M_1 - M_0}$ und $M_1 =: M_0 + sa$. Aus der Voraussetzung folgt

$$\begin{aligned} s < r_0 + r_1, &\quad \text{also } s - r_0 < r_1, \\ r_0 - r_1 < s, &\quad \text{also } r_0 - s < r_1, \\ r_1 - r_0 < s, &\quad \text{also } r_1 < s + r_0. \end{aligned}$$

Insgesamt liefert das

$$(r_0 - s)^2 < r_1^2 < (r_0 + s)^2.$$

Nun ist

$$k_0 = \{P(t) := M_0 + r_0(\cos t\, a + \sin t\, Ja) \,|\, 0 \le t \le 2\pi\}$$

und

$$|P(t) - M_1|^2 = (r_0 \cos t - s)^2 + r_0^2 \sin^2 t = r_0^2 + s^2 - 2r_0 s \cos t.$$

Diese Funktion nimmt im Intervall $[0, 2\pi]$ jeden Wert r_1^2 zwischen $(r_0 - s)^2$ und $(r_0 + s)^2$ genau zweimal an. □

Satz 48 (Sehnensatz). *Seien k ein Kreis mit Mittelpunkt M und Radius r und P ein Punkt, $P \notin k$. Sei g eine Gerade durch P, die k in den Punkte Q, R schneidet. Dann gilt*

(i) Das Produkt der Sehnenabschnitte $|P-Q|$ *und* $|P-R|$ *ist*

$$\pm|Q-P| \cdot |R-P| = |P-M|^2 - r^2.$$

Es ist also unabhängig von der Geraden durch P. Das Vorzeichen ist durch die rechte Seite bestimmt, hängt also davon ab, ob P innerhalb oder außerhalb des Kreises liegt. (Sehnensatz)

(ii) Ist $|P-M| > r$ und $\tilde{Q} \in k$, so daß $g(P, \tilde{Q})$ eine Tangente an k ist, so gilt

$$|Q-P| \cdot |R-P| = |P-\tilde{Q}|^2.$$

Das Produkt der Sehnenabschnitte ist gleich dem Quadrat des Tangentenabschnittes. *(Sehnen-Tangenten-Satz)*

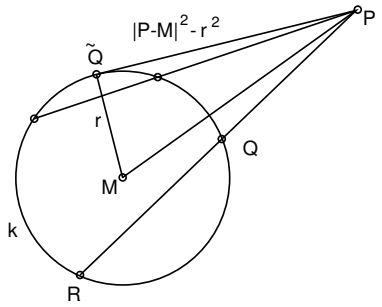
Beweis. Wir schreiben die Gerade g in der Form

$$P + tv$$

mit $|v| = 1$. Dann sind die Schnittpunkte mit dem Kreis gegeben durch $P + t_i v$, wobei die t_i die beiden Lösungen von

$$\langle P - M + tv, P - M + tv \rangle = r^2$$

sind.



Die Sehnenabschnitt sind wegen $|v| = 1$ dann gerade die $|t_i|$. Wir müssen also zeigen, daß $t_1 t_2 = |P-M|^2 - r^2$. Aber es gilt

$$\begin{aligned} r^2 &= \langle P - M + tv, P - M + tv \rangle \\ &= \langle P - M, P - M \rangle + 2t\langle P - M, v \rangle + t^2. \end{aligned}$$

Das Produkt der Nullstellen der quadratischen Gleichung $t^2 + at + b = 0$ ist aber nach Vieta gerade der konstante Term b, in unserem Falle also

$$t_1 t_2 = |P-M|^2 - r^2.$$

□

5.5 Schnittpunktsätze im Dreieck

Wir erinnern an den

Satz 49 (Schwerpunktsatz). *Die Seitenhalbierenden $g(A, \sigma(B,C))$ usw. eines Dreiecks (A,B,C) schneiden sich im Schwerpunkt $\sigma(A,B,C)$. Dieser teilt die Seitenhalbierenden im Verhältnis $1:2$. Es gilt*

$$\sigma(A,B,C) - A = 2(\sigma(B,C) - \sigma(A,B,C)$$

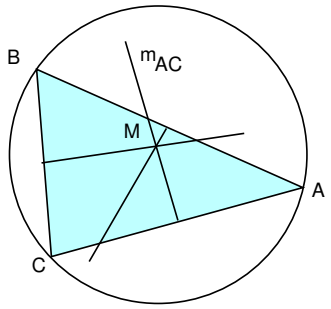
und ensprechend für die anderen Ecken.

Vgl. Satz 38.

Satz 50 (Umkreis). *Sei (A,B,C) ein Dreieck. Dann gibt es genau einen Kreis durch A,B,C. Dieser heißt der* Umkreis *des Dreiecks. Sein Mittelpunkt ist der gemeinsame Schnittpunkt der Mittelsenkrechten auf den drei Seiten AB, BC und AC.*

Beweis. Weil keine Seiten des Dreiecks parallel sind, sind keine Mittelsenkrechten parallel, und deshalb schneiden sich die Mittelsenkrechten m_{AB} und m_{BC} in einem Punkt M.

Dieser hat nach dem Lemma 27 gleichen Abstand von A und B ebenso wie von B und C, also von allen drei Eckpunkten. Daher geht der Kreis um M vom Radius $|A-M|$ durch alle drei Eckpunkte, und die Mittelsenkrechte auf AC geht ebenfalls durch M. Umgekehrt folgt aus dem Lemma 27, daß der Mittelpunkt eines Kreises durch die drei Eckpunkte auf allen drei Mittelsenkrechten liegen muß. Daraus folgt die Eindeutigkeit. □



Korollar 5. *Durch drei nicht kollineare Punkte einer affinen Euklidischen Ebene geht genau ein Kreis.*

Definition 31. Die Lote von den Eckpunkten eines Dreiecks (A,B,C) auf die gegenüberliegende Seite heißen die *Höhen* des Dreiecks.

$$h_A := \perp_{g(B,C)} (A)$$

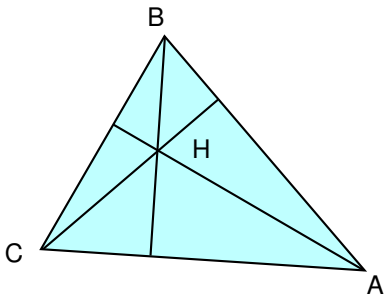
und h_B, h_C analog.

Satz 51 (Höhenschnittpunkt). *Die drei Höhen eines Dreiecks (A,B,C) schneiden sich in einem Punkt H.*

1. Beweis.

Es ist

$h_A = \{A+tJ(B-C) \,|\, t \in \mathbb{R}\}, h_B = \{B+tJ(C-A) \,|\, t \in \mathbb{R}\}.$



Weil die Eckpunkte in allgemeiner Lage sind, sind die Höhen nicht parallel, schneiden sich also in einem Punkt. Der Schnittpunkt dieser beiden Geraden ist dann

$$H = A + tJ(B - C) = B + sJ(C - A)$$

mit geeigneten $s, t \in \mathbb{R}$.

Es folgt

$$\langle H - C, B - C\rangle = \langle A - C, B - C\rangle$$
$$\langle H - C, A - C\rangle = \langle B - C, A - C\rangle$$

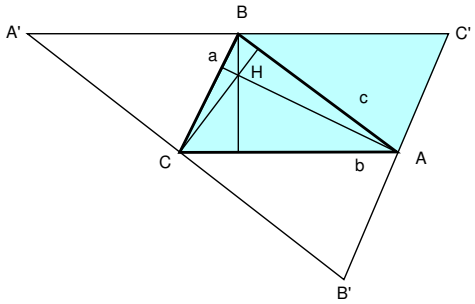
Also folgt

$$\langle H - C, B - A\rangle = \langle H - C, (B - C) - (A - C)\rangle = \langle H - C, B - C\rangle - \langle H - C, A - C\rangle = 0.$$

Das bedeutet aber, daß die Gerade $g(CH)$ die Höhe h_C ist. Diese geht also auch durch den Punkt H. □

2. Beweis (Gauß).

Durch die Ecken A, B, C des Dreiecks legt man Parallelen zu den gegenüberliegenden Seiten. Es entsteht ein neues Dreieck (A', B', C'). Weiter entstehen sehr viele Parallelogramme, von denen in der Abbildung eines gefärbt ist. An diesen Parallelogrammen erkennt man, daß A, B, C die Mittelpunkte der neuen Dreiecksseiten sind, daß und daß die Höhen des alten Dreiecks gerade die Mittelsenkrechten des neuen sind.



Weil die Mittelsenkrechten sich in einem Punkt schneiden, tun es auch die Höhen des ursprünglichen Dreiecks.

Wir halten noch fest, daß der Höhenabschnitt BH des kleinen Dreiecks im großen Dreieck der Abschnitt der Mittelsenkrechten von der Seitenmitte bis zum Umkreismittelpunkt M' ist. □

Der Höhenschnittpunkt hat anders als der Schwerpunkt und der Umkreismittelpunkt keine unmittelbare geometrische Bedeutung im Dreieck (A, B, C). Der Beweis von Gauß verschafft ihm aber eine!

Ohne Beweis führen wir hier an:

Satz 52 (Inkreis). *Die drei Winkelhalbierenden eines nicht-ausgearteten Dreiecks schneiden sich in einem Punkt W. Dieser hat von allen Seiten gleichen Abstand r, so daß der Kreis um W mit Radius r alle Seiten tangiert. Der Kreis heißt der* Inkreis *des Dreiecks.*

Satz 53 (Euler 1763). *Seien M, S und H die Schnittpunkte der Mittelsenkrechten, Seitenhalbierenden und Höhen des Dreiecks (A, B, C). Dann gilt*

$$H = M + 3(S - M).$$

Insbesondere liegen die drei Punkte also auf einer Geraden. Ist $S \neq M$, so ist diese Gerade eindeutig bestimmt und heißt die Eulergerade. *Der Schwerpunkt teilt dann die Strecke MH im Verhältnis $1 : 2$.*
Bemerkung. Man kann zeigen, daß genau dann zwei der drei Punkte M, S, H zusammenfallen, wenn alle drei zusammenfallen. Dies ist genau für gleichseitige Dreiecke der Fall.

Beweis. Es gilt

$$3(S - M) = (A - M) + (B - M) + (C - M).$$

Deshalb gilt für den Punkt $P := M + 3(S - M)$

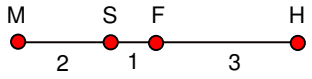
$$\begin{aligned}\langle P - A, B - C\rangle &= \langle M - A + 3(S - M), B - C\rangle = \langle (B - M) + (C - M), B - C\rangle \\ &= \langle (B - M) + (C - M), (B - M) - (C - M)\rangle = |B - M|^2 - |C - M|^2 = 0,\end{aligned}$$

weil M der Umkreismittelpunkt ist. Die Gerade durch P und A steht also senkrecht auf BC und P liegt auf der Höhe h_A. Aus Symmetriegründen liegt P dann auf allen drei Höhen, d.h. $P = H$. □

Satz 54 (Feuerbach 1822). *Seien M, S und H die Schnittpunkte der Mittelsenkrechten, Seitenhalbierenden und Höhen des Dreiecks (A, B, C). Dann ist*

$$F := \sigma(M, H) = M + \frac{3}{2}(S - M)$$

ein weiterer Punkt auf der Eulergeraden:



Der Kreis um F von halben Umkreisradius enthält dann folgende neun Punkte:

- *die drei Seitenmitten,*
- *die drei Höhenfußpunkte,*
- *die Mittelpunkte der Höhenabschnitte AH, BH, CH.*

Dieser Kreis heißt der Feuerbachkreis *des Dreiecks.*
Ohne Beweis führen wir an, daß der Feuerbachkreis außerdem den Inkreis des Dreiecks berührt.

1. Beweis. Aus

$$F = M + \frac{1}{2}((A - M) + (B - M) + (C - M)) = -\frac{1}{2}M + \frac{1}{2}A + \frac{1}{2}B + \frac{1}{2}B$$

folgt

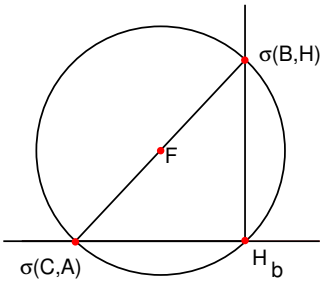
$$F - \sigma(C, A) = \frac{1}{2}(B - M).$$

Daher ist $2|F - \sigma(A, B)|$ der Umkreisradius r. Gleiches gitl für die anderen Seitenmitten. Also liegen sie alle auf dem Kreis um F vom halben Umkreisradius.

Weiter ist

$$F - \sigma(B, H) = \sigma(M, H) - \sigma(B, H) = \frac{1}{2}(M + H - (B + H)) = -\frac{1}{2}(B - M).$$

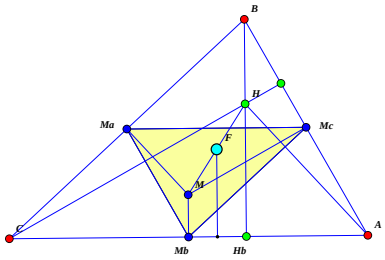
Damit sind $\sigma(C, A)$ und $\sigma(B, H)$ antipodale Punkte auf dem Feuerbachkreis. Zusammen mit dem Fußpunkt H_b der Höhe h_B bilden diese Punke ein rechtwinkliges Dreieck, und nach dem Satz von Thales liegt der Fußpunkt auf dem Feuerbachkreis.



□

2. Beweis.

Wir betrachten die Mittelsenkrechte m_{CA} und die Höhe h_B. Das sind parallele Geraden, die $g(C,A)$ in $M_b = \sigma(C,A)$ und H_b schneiden. Weil $F = \sigma(M,H)$, liegt F auf der Mittelsenkrechten $m_{M_b\ H_b}$.



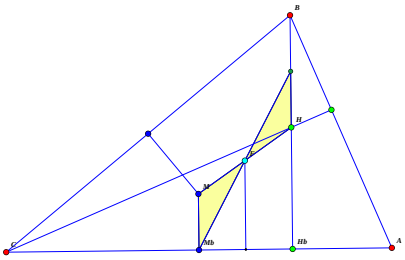
Aus dem Beweis von Gauß über den Höhenschnittpunkt folgt, daß

$$2|M_b - M| = |B - H|.$$

Also ist

$$|M_b - M| = |\sigma(B,H) - H|,$$

und F halbiert die Strecke $M_b\sigma(B,H)$.

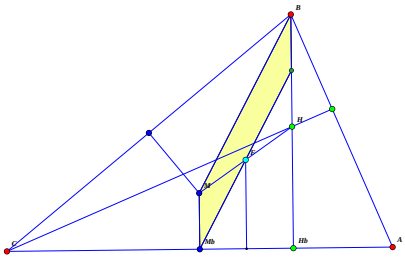


Aber es ist auch

$$|M_b - M| = |B - \sigma(B,H)|,$$

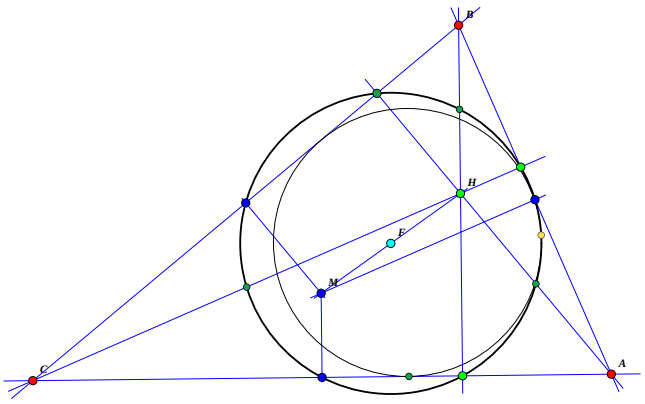
und wir erhalten ein Parallelogramm $(M_b, \sigma(B,H), B, M)$. Daraus folgt, daß

$$\frac{1}{2}|B - M| = |M_b - F| = |F - \sigma(B,H)|.$$



Nun ist $r = |B - M|$ ist der Umkreisradius. Deshalb geht der Kreis um F mit Radius $r/2$ durch M_b, H_b und $\sigma(B,H)$. Aus Symmetriegründen gilt das für alle anderen Seiten dann auch.

□



5.6 Dreiecksfläche

Wenn wir die Fläche des Dreiecks definieren als halbe Grundseite mal Höhe, so finden wir:

Satz 55 (Dreiecksfläche). *Die Fläche F des Dreiecks (A, B, C) ist gegeben duch*

$$F = \frac{1}{2}\,|\langle Ja, b\rangle| = \frac{1}{2}|a|\,|b|\sin\gamma.$$

Beweis. Die Höhe ist

$$h = |\langle \frac{Jb}{|b|}, a\rangle|.$$

Daher folgt

$$F = \frac{1}{2}\,h|b| = \frac{1}{2}|\langle Jb, a\rangle| \underset{\text{Lemma 26}}{=} \frac{1}{2}|a||b|\sin\gamma$$

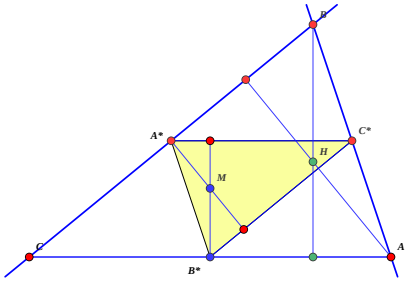
□

5.7 Feuerbachkreis und Eulergerade: Weihnachtsversion

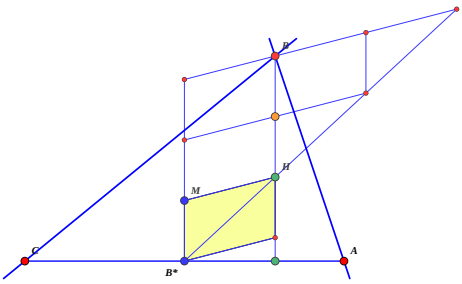
Voraussetzungen: Geometrie des Parallelogramms, Strahlensatz, Umkreis, Höhenschnittpunkt mit Beweis von Gauß (vgl. erste Figur).

Das Dreieck (A^*, B^*, C^*) zwischen den Seitenmitten hat die halben Seitenlängen des Dreiecks (A, B, C). Daher ist

$$|B - H| = 2|M - B^*|. \quad (*)$$



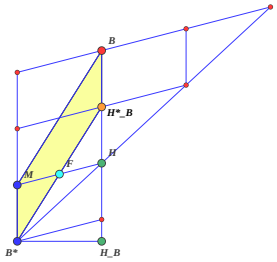
Wir betrachten das Parallelogramm $(B^*, M, H, B^* + (H - M))$ und ergänzen zu dem abgebildeten Netz aus Parallelogrammen. Wegen $(*)$ ist B ein Knotenpunkt



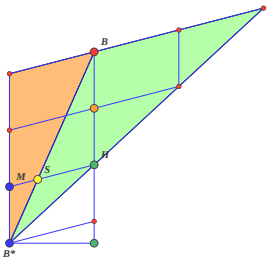
Der Schnittpunkt F von MH und $B^*H_B^*$ ist der Mittelpunkt der Strecken MH und $B^*H_B^*$. Sei $r := |B - M|$ der Umkreisradius. Dann gilt

$$\frac{r}{2} = |F - B^*| = |H_B^* - F| = |F - H_B|.$$

Der Kreis um F mit Radius $r/2$ enthält also die Seitenmitten, die Fußpunkte der Höhen und die Mitten der Höhenabschnitte zwischen H und den Ecken. (Feuerbachkreis)



Der Schnittpunkt S von MH und der Seitenhalbierenden BB^* teilt die Strecke MH im Verhältnis $1 : 2$.
Also schneiden sich die Seitenhalbierenden in *einem* Punkt, und dieser liegt auf der Eulergeraden durch M und H. (Satz von Euler).



5.8 Ellipsen, Hyperbeln, Parabeln

Sei $(\mathcal{A}, V, \tau)$ eine Euklidische affine Ebene.

Definition 32. Seien (e_1, e_2) eine Orthonormalbasis von V und $Q \in \mathcal{A}$. Dann heißt die Abbildung

$$\mathbb{R}^2 \to \mathcal{A}, (x, y) \mapsto \tau_{(xe_1+ye_2)}(Q)$$

ein (in Q zentriertes) Euklidisches Koordinatensystem. Wir bezeichnen die Umkehrabbildung mit

$$(x, y) : \mathcal{A} \to \mathbb{R}^2, P \mapsto (x(P), y(P)).$$

und nennen (x, y) *Euklidische Koordinaten.*

Definition 33 (Kegelschnitte). (i) Eine *Ellipse* ist eine Teilmenge $E \subset \mathcal{A}$, die in geeigneten Euklidischen Koordinaten gegeben ist durch eine Gleichung

$$E = \{P \in \mathcal{A} \,|\, \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1\}$$

mit $a, b, \in \mathbb{R}$, $0 < b \leq a$. Dabei schreiben wir zur Vereinfachung x und y statt $x(P)$ und $y(P)$.

Für $a = b$ ist E ein Kreis vom Radius a. Wir nennen a und b die *(Längen der) Halbachsen* und die Punkte $(\pm a, 0)$ und $(0, \pm b)$ die *Scheitel* von E.

(ii) Eine *Hyperbel* ist eine Teilmenge $E \subset \mathcal{A}$, die in geeigneten Euklidischen Koordinaten gegeben ist durch eine Gleichung

$$H = \{P \in \mathcal{A} \,|\, \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1\}$$

mit positiven $a, b, \in \mathbb{R}$. Die Punkte $(\pm a, 0)$ heißen die *Scheitel* der Hyperbel.

Die Teilmengen der Hyperbel mit $x \geq a$ bzw. $x \leq -a$ heißen die beiden *Zweige* oder *Äste* der Hyperbel. Die Geraden

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0,$$

also $bx \pm ay = 0$, heißen die *Asymptoten* der Hyperbel.

Beachte, daß auf der Hyperbel $y^2 = \frac{b^2}{a^2}(x^2 - a^2)$ also

$$y = \pm\frac{b}{a}\sqrt{x^2 - a^2} \approx \pm\frac{b}{a}x \text{für } |x| \to \infty.$$

(iii) Eine *Parabel* ist eine Teilmenge $E \subset \mathcal{A}$, die in geeigneten Euklidischen Koordinaten gegeben ist durch eine Gleichung

$$P = \{(x, y) \in \mathbb{R}^2 \,|\, y^2 = 2px\}$$

mit $p > 0$. Der Punkt $(0, 0)$ heißt der *Scheitel*, $2p$ der *Parameter* und die Gerade $y = 0$ die *Achse* der Parabel.

(iv) Wir benutzen einstweilen den Begriff *Kegelschnitt* als Sammelbezeichnung für Ellipse, Hyperbel und Parabel. Später werden wir diese Definition noch etwas erweitern und den Namen erklären.

Satz 56 (und Definition: Tangente). *Sei κ ein Kegelschnitt und $P \in \kappa$. Dann gibt es genau eine Gerade g durch P, mit*

$$g \cap \kappa = \{P\}$$

und folgender Zusatzbedingung:

- *g nicht parallel zu einer Asymptote, falls κ Hyperbel,*
- *g nicht parallel zur Achse, falls κ Parabel.*

Diese Gerade heißt die Tangente in P an κ.

Beweis. Wir zeigen das für den Falle der Hyperbel, die anderen Fälle gehen ähnlich. Wir nehmen o.E. an, daß $\mathcal{A} = \mathbb{R}^2$ und

$$\kappa = H = \{(x, y) \in \mathbb{R}^2 \mid \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1\}.$$

Sei $P = (x_0, y_0) \in H$ und sei

$$g = \{(x_0 + tu, y_0 + tv) \mid t \in \mathbb{R}\}, \quad (u, v) \neq (0, 0)$$

eine Gerade durch P. Wir bestimmen die Schnittpunkte mit H:

$$\begin{aligned} 1 &= \frac{(x_0 + tu)^2}{a^2} - \frac{(y_0 + tv)^2}{b^2} \\ &= \frac{x_0^2}{a^2} + 2t\frac{x_0 u}{a^2} + t^2\frac{u^2}{a^2} - \frac{y_0^2}{b^2} - 2t\frac{y_0 v}{b^2} - t^2\frac{v^2}{b^2} \\ &= t^2\left(\frac{u^2}{a^2} - \frac{v^2}{b^2}\right) + 2t\left(\frac{x_0 u}{a^2} - \frac{y_0 v}{b^2}\right) + \underbrace{\frac{x_0}{a^2} - \frac{y_0}{b^2}}_{=1}. \end{aligned}$$

Wenn die Gerade nicht parallel zu einer Asymptotenrichtung ist, ist $\frac{u^2}{a^2} - \frac{v^2}{b^2} \neq 0$, und die Gleichung hat genau dann die einzige Lösung $t = 0$, wenn

$$\frac{x_0 u}{a^2} - \frac{y_0 v}{b^2} = 0$$

d.h. wenn

$$\left(\frac{x_0}{a^2}, -\frac{y_0}{b^2}\right)$$

ein Normalenvektor von g ist. Also gibt es genau eine solche Gerade g. □

Der vorstehende Beweis gibt auch die Gleichung für die Tangente an. Wir schreiben sie zusammen mit den Formeln für die beiden anderen Fälle auf:

Satz 57 (Tangentengleichung). *Sei κ ein Kegelschnitt in Normalform in der Bezeichnung aus Satz 56. Dann ist die Tangente an κ in $(x_0, y_0) \in \kappa$ gegeben durch die Gleichung*

$$\begin{aligned} \frac{(x - x_0)x_0}{a^2} \pm \frac{(y - y_0)y_0}{b^2} = 0 \qquad &\textit{(Ellipse/Hyperbel),} \\ (y - y_0)y_0 - p(x - x_0) = 0. \qquad &\textit{(Parabel).} \end{aligned}$$

Satz 58 (Konstruktion von Kegelschnitten). *Sei k ein Kreis vom Radius R mit Mittelpunkt M und $F \notin k$.*
Für $P \in k$ bezeichne $m_P := m_{FP}$ die Mittelsenkrechte von FP und

$$X(P) := m_P \cap g(M, P).$$

Wir setzen

$$\kappa := \bigcup_{P \in k} X(P).$$

Dann gilt

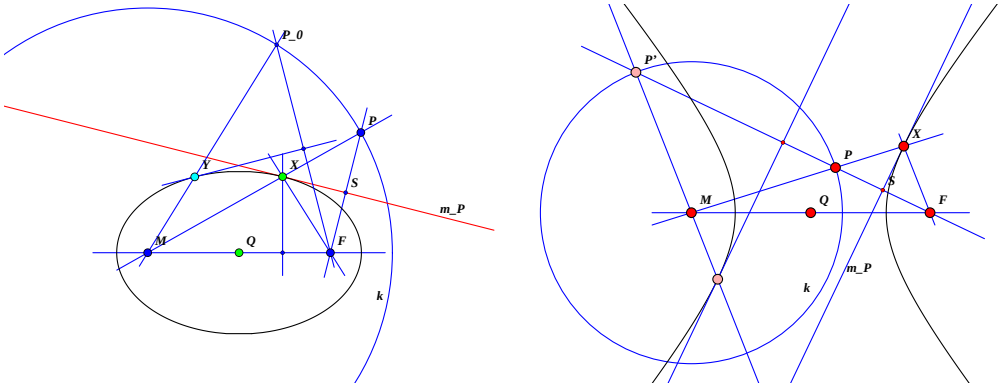
(i) Ist $|F - M| < R$, so ist κ eine Ellipse und

$$\kappa = \{X \,|\, |X - M| + |X - F| = R\}.$$

(ii) Ist $|F - M| > R$, so ist κ eine Hyperbel und

$$\kappa = \{X \,|\, ||X - M| - |X - F|| = \pm R\}.$$

In beiden Fällen ist m_P für alle $P \in k$ die Tangente an κ in $X(P)$.



Beweis. Weil $F \notin k$, ist $M \notin X(P)$ für alle P. Sei also $X \neq M$. Dann ist $X \in \kappa$ genau dann, wenn $X \in X(P)$ für ein $P \in k$. Die einzig möglichen Kandidaten sind aber

$$P_{pm} = M \pm \frac{R}{|X - M|}(X - M).$$

Deshalb ist $X \in \kappa$ genau dann, wenn

$$|X - F| = \left| X - (M + \frac{R}{|X - M|}(X - M)) \right| = |X - M| \left| 1 + \frac{R}{|X - M|} \right| = ||X - M| + R|$$

oder

$$|X - F| = ||X - M| - R|.$$

Nun führen wir Koordinaten ein: Wir setzen

$$Q := \sigma(M, F), f := |F - Q|, e_1 := \frac{1}{f}(F - Q), e_2 := Je_1,$$

und für $X \in \mathcal{A}$

$$x = x(X) := \langle X - Q, e_1 \rangle,$$
$$y = y(X) := \langle X - Q, e_2 \rangle.$$

Dann ist

$$|X-F| = \sqrt{(x-f)^2+y^2}, \quad |X-M| = \sqrt{(x+f)^2+y^2},$$

und nach kurzer Rechnung findet man

$$\begin{aligned} |X-F| = ||X-M|+R| &\iff 4xf + R^2 = 2R\sqrt{(x+f)^2+y^2}, \\ |X-F| = ||X-M|-R| &\iff 4xf + R^2 = -2R\sqrt{(x+f)^2+y^2}. \end{aligned}$$

Also

$$X \in \kappa \iff (4xf+R^2)^2 = 4R^2((x+f)^2+y^2).$$

Wiederum nach kurzer Rechnung folgt, daß $X \in \kappa$ genau dann, wenn

$$\frac{x^2}{R^2/4} + \frac{y^2}{(R^2-4f^2)/4} = 1. \tag{49}$$

F liegt außerhalb des Kreises, wenn $|F-M| = 2f > R$. In diesem Fall erhält man eine Hyperbel, andernfalls eine Ellipse.

Warum ist m_P die Tangente? Beachte zunächst, daß $X(P)$ auf der Geraden $g(M,P)$ liegt. Nach dem unten stehenden Lemma schneidet der Kreis um $X(P)$ durch $F \notin k$ und P den Kreis k nur im Punkt P.

Annahme: Ein weiterer Punkt $X(P_0)$ von κ liegt auf m_P. Dann schneidet der Kreis um $X(P_0)$ durch F den Kreis k genau in P_0. Weil sein Mittelpunkt aber auf $m_P = m_{FP}$ liegt, schneidet er auch in P. Widerspruch!

Annahme: m_P ist parallel zu einer Asymptote. Wir zeigen, daß dann

$$\langle P-F, P-M\rangle = 0. \tag{50}$$

Dann ist $g(P,M)$ parallel zu m_P und die beiden Geraden schneiden sich gar nicht oder fallen zusammen. Letzteres ist unmöglich, weil $P \notin m_P$. Ersteres ist ebenfalls unmöglich, weil nach Voraussetzung $X(P) \in g(M,P) \cap m_P$. Widerspruch!

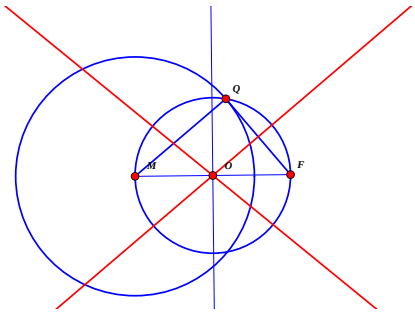
Weil wir die Asymptoten mittels der Koordinatendarstellung definiert haben, müssen wir darauf zurückgreifen. Wir betrachten mit Blick auf (49) die Hyperbel

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1,$$

wo $a^2 = R^2/4, b^2 = (4f^2-R^2)/4$, also $R = 2a$ und $f = \sqrt{a^2+b^2}$.

Wir betrachten das Dreieck mit den Ecken $M = (-f,0), F = (+f,0)$ und Q auf dem Thaleskreis um 0 durch M und F. Dann ist die Strecke QF genau dann senkrecht zur Asymptotenrichtung $(a, \pm b)$, wenn mit $z = |Q-M|$ gilt

$$\frac{b^2}{a^2} = \frac{4f^2-z^2}{z^2}$$



Auflösen nach z^2 liefert unter Benutzung von $f^2 = a^2+b^2$

$$z^2 = 4a^2 = R^2.$$

Also liegt $Q = P$ dann auf dem Kreis k und $g(M,P)$ ist parallel zu m_P. Aber dann ist $X(P) = \emptyset$. □

Lemma 30. *Seien $k_1 = k_{r_1}(M_1)$ und $k_2 = k_{r_2}(M_2)$ zwei Kreise, und sei $P \in k_1 \cap k_2$. Die Punkte p, M_1, M_2 seien kollinear. Dann folgt*

$$k_1 \cap k_2 = \{P\} \text{ oder } k_1 = k_2.$$

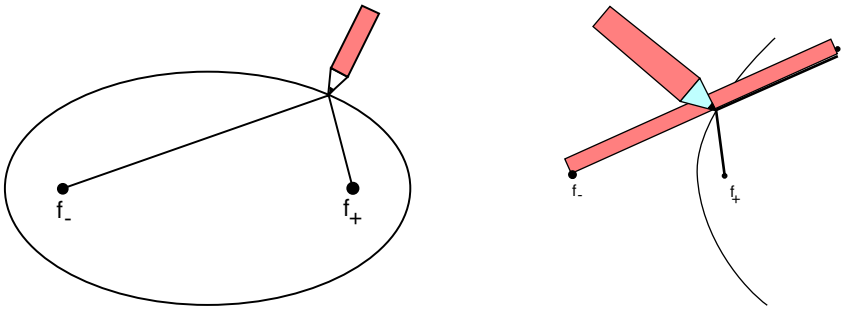
Beweis. Wir nehmen an, daß $Q \in k_1 \cap k_2$ und $Q \neq P$. Wir müssen zeigen, daß dann $k_1 = k_2$. Nach Voraussetzung gilt

$$\{M_1, M_2\} \subset m_{PQ} \cap g(M_1, P).$$

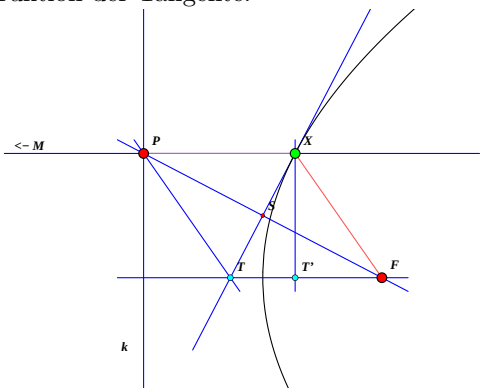
Weil $P \in g(M_1, P)$ und $P \notin m_{PQ}$, ist $m_{PQ} \neq g(M_1, P)$. Also haben die beiden Geraden höchstens einen Punkt gemeinsam, und es folgt $M_1 = M_2$. Weil P auf beiden Kreisen liegt, folgt $k = k_0$. □

Bemerkungen.

1. Insbesondere für die Ellipse nennt man das durch den Satz 58 gegebene Konstruktionsverfahren auch die „Gärtner-Konstruktion“:



2. In der obigen Konstruktion ist $g(P, M)$ die Gerade durch P senkrecht zum Kreis, so daß man die Konstruktion auch ohne explizite Benutzung des Punktes M beschreiben kann. Dann kann man den Kreis k auch durch eine Gerade ersetzen, und die Konstruktion liefert eine *Parabel* $y^2 = 2px$. Sie ist der Ort aller Punkte, die gleichen Abstand von F und der *Leitlinie* k haben. p ist der Abstand zwischen F und der Leitlinie. Der *Scheitel* halbiert die Strecke TT' zwischen dem Lotpunkt T' des Berührungspunktes X auf die *Achse* und dem Schnittpunkt T der Tangente und der Achse. Das ist eine bequeme Methode zur Konstruktion der Tangente.



3. Der Beweis des Satzes liefert die Halbachsen als Funktionen von R und $f := \frac{1}{2}|F - M|$.

$$a = \frac{R}{2}, \qquad b = \sqrt{|\frac{R^2}{4} - f^2|}.$$

Für die Parabel ist p der Abstand $|k, F|$ von F zur Leitgeraden k. Insbesondere folgt daraus, daß man jeden Kegelschnitt auf diese Weise konstruieren kann.

4. Die Punkte F und $F' := M$ nennt man die *Brennpunkte* des Kegelschnitts. In den Standard-Koordinaten sind sie gegeben wie folgt:

$$\begin{aligned} \text{Ellipse:} \quad & (\pm\sqrt{a^2-b^2}, 0), \\ \text{Hyperbel:} \quad & (\pm\sqrt{a^2+b^2}, 0), \\ \text{Parabel:} \quad & (\frac{p}{2}, 0). \end{aligned}$$

5. Ist X ein Punkt auf der Mittelsenkrechten m zweier Punkte F und F', so halbiert m den Winkel zwischen den Verbindungsgeraden zu F und F'. Daher schließt in der Konstruktion des Satzes 58 die Tangente in $X = X(P)$ mit den beiden Geraden $g(X, F)$ und $g(X, F')$, den sogenannten Brennstrahlen, denselben Winkel ein. Man erhält mit dem Reflexionsgesetz (Einfallswinkel=Ausfallswinkel):

Die von einem Brennpunkt F ausgehenden Strahlen treffen sich nach Reflexion im anderen Brennpunkt F'.

   Bei Hyperbel und Parabel bedarf das allerdings einer Interpretation: Bei der Hyperbel muß man die rückwärtige Verlängerung der reflektierten Strahlen betrachten. Bei der Parabel werden die achsenparallelen Strahlen in den Brennpunkt reflektiert (Parabolantenne).

Warum „Kegelschnitt“? Wir betrachten im $\mathbb{R}^3$ einen Kegel mit der x-Achse als Rotationsachse:

$$y^2 + z^2 - x^2 = 0.$$

Jetzt drehen wir die Achse des Kegels um den Winkel ϕ in der xz-Ebene und erhalten einen neuen Kegel mit der Gleichung

$$y^2 + (z\,\cos\phi - x\,\sin\phi)^2 - (z\,\sin\phi + x\,\cos\phi)^2 = 0$$

Der Schnitt dieses Kegels mit der affinen Ebene $z = p \neq 0$ hat in den Euklidischen Koordinaten x, y die Gleichung

$$\begin{aligned} 0 =& y^2 + (p\,\cos\phi - x\,\sin\phi)^2 - (p\,\sin\phi + x\,\cos\phi)^2 \\ =& y^2 + x^2(\sin^2\phi - \cos^2\phi) - 4px\cos\phi\sin\phi + p^2(\cos^2\phi - \sin^2\phi) \\ =& y^2 - x^2\cos 2\phi - 2px\sin 2\phi + p^2\cos 2\phi. \end{aligned}$$

Für $\cos 2\phi = 0$ ist $\sin 2\phi = \pm 1$ und man bekommt

$$0 = y^2 \mp 2px,$$

also eine Parabel. Andernfalls können wir weiter schließen

$$\begin{aligned} 0 &= y^2 - \cos 2\phi \underbrace{(x + p\frac{\sin 2\phi}{\cos 2\phi})^2}_{\tilde{x}} + p^2(\frac{\sin^2 2\phi}{\cos 2\phi} + \cos 2\phi) \\ &= y^2 - \cos 2\phi\,\tilde{x}^2 + \frac{p^2}{\cos 2\phi}. \end{aligned}$$

Das ist je nach Vorzeichen von $\cos 2\phi$ eine Hyperbel oder Ellipse.

Schneidet man den Kegel mit der Ebene $z = 0$, so erhält man „degenerierte“ Kegelschnitte, nämlich

$$y^2 = \cos 2\phi\, x^2.$$

Für $-\phi/4 < \phi < \pi/4$ ist das ein Paar sich schneidender Geraden, für $\phi = \pm\pi/4$ fallen diese zusammen („Doppelgerade“), für $\cos 2\phi < 0$ besteht der Schnitt nur aus dem Nullpunkt.

5.9 Transformationsgruppen

Die Begriffe der Euklidischen Geometrie sind invariant unter bijektiven affinen Abbildungen mit orthogonalem linearen Anteil (=Bewegungen). Dagegen sind affine Begriffe, wie etwa der Schwerpunkt, invariant unter allgemeinen bijektiven Transformationen.

Es ist eine Entdeckung des 19. Jahrhunderts, daß verschiedene „Geometrien“ durch ihre Invarianzgruppen charakterisiert sind und mit diesen Gruppen studiert werden sollten. Dieser Sachverhalt stand zum Beispiel bei der Klassifikation der normalen Endomorphismen eines Euklidischen Vektorraumes im Hintergrund: Das war eben eine Klassifikation bis auf die Operation der Gruppe $\mathrm{O}(V)$.

Ich will zwei Beispiele aus der Geometrie dafür geben, daß die Benutzung der „richtigen“ Invarianzgruppe eines Problems sehr hilfreich sein kann.

5.9.1 Hyperbeltangenten, $\mathrm{SL}(2)$ und $\mathrm{O}(1,1)$

In einem Euklidischen Vektorraum V sind die orthogonalen Transformationen insbesondere Volumen- (oder Flächen)treu. Aber es gibt eine viel größere Gruppe von Transformationen mit dieser Eigenschaft, gebildet von allen Transformationen mit Determinante $= 1$. Das folgt aus dem Determinanten-Multiplikationssatz und war gewissermaßen die Leitidee bei der Definition der Determinante. Deshalb ist für Fragen des Raum- oder Flächeninhalts eigentlich keine Euklidische Struktur erforderlich, man braucht nur eine Determinantenform. Die Automorphismen mit Determinante $= 1$ bilden dann eine Untergruppe von $\mathrm{Aut}(V) = \mathrm{GL}(V)$, die sogenannte *spezielle lineare Gruppe* $\mathrm{SL}(V)$.

Allgemein gibt der Absolutbetrag der Determinanten die Volumenverzerrung eines Endomorphismus an.

Wir benutzen das für einen einfachen Beweis zu folgendem Satz über Hyperbeln. Der Einfachheit halber sei unser affiner Raum der $\mathbb{R}^2$.

Satz 59. *Der Flächeninhalt des Dreiecks zwischen der Tangente an die Hyperbel*

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \tag{51}$$

und den Asymptoten ist unabhängig vom Berührpunkt der Tangente.

Beweis.

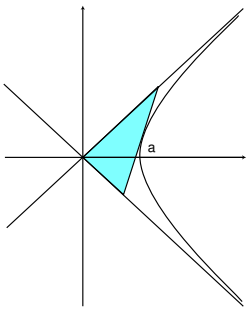
Zunächst bildet die lineare Abbildung des $\mathbb{R}^2$ mit Matrix

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$$

die Hyperbel

$$x^2 - y^2 = 1 \tag{52}$$

bijektiv auf die Hyperbel (51) ab. Dabei gehen die Asymptoten und Tangenten in Asymptoten und Tangenten über, und der Flächeninhalt wird mit dem konstanten Faktor ab verzerrt. Deshalb genügt der Nachweis für die Hyperbel (52).



Offenbar können wir uns auch auf Tangenten an den rechten Ast ($x > 0$) beschränken, weil nämlich die Drehung mit Zentrum 0 um den Winkel π einerseits die Tangenten und Asymptoten und andrerseits auch den Flächeninhalt erhält.

Auf dem $\mathbb{R}^2$ hat man eine symmetrische aber indefinite Bilinearform

$$\langle x, y\rangle = x_1 y_1 - x_2 y_2,$$

die sogenante Lorentzmetrik. Die Einheitshyperbel (52) ist dann gegeben durch die Gleichung

$$\langle x, x\rangle = 1,$$

wie der Kreis durch die entsprechende Euklidische Gleichung. Die Gruppe der Matrizen, die diese Bilinearform erhalten, die sogenannte *Lorentzgruppe* $\mathrm{O}(1,1)$, enthält insbesondere die Untergruppe der Matrizen

$$A_t := \begin{pmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{pmatrix}, \quad t \in \mathbb{R}, \tag{53}$$

wie man mittels $\cosh^2 - \sinh^2 = 1$ sofort nachrechnet. Sie ist das Pendant zur Euklidischen Drehgruppe der Matrizen

$$\begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$$

und heißt auch die Gruppe der *Lorentz-* oder *hyperbolischen Rotationen*. Diese Matrizen in $\mathrm{O}(1,1)$ bilden natürlich die Hyperbel (52) auf sich ab. Sie erhalten die Asymptoten, denn die sind durch

$$\langle x, x\rangle = 0$$

charakterisiert. Die Abbildungen A_t erhalten darüber hinaus auch die Äste. Schließlich ist jeder Punkt auf dem rechten Ast von (52) von der Form

$$(\cosh t, \sinh t)$$

für ein geeignetes t. Die lineare Abbildung A_t bildet also den Punkt $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ in einen beliebig vorgegebenen Punkt der Hyperbel ab: so wie eine geeignete Rotation um den Mittelpunkt eines Kreises einen festen Punkt auf der Peripherie in einen beliebig vorgegebenen anderen transportiert. Die Abbildung A_t bildet dann das Tangenten-Asymptoten-Dreieck im Punkt $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ in das entsprechende Dreieck an der Stelle $(\cosh t, \sinh t)$ ab. Und wegen

$$\det A_t = \cosh^2 t - \sinh^2 t = 1$$

bleibt der Flächeninhalt erhalten (und zwar $= 1$). Damit ist der Satz bewiesen. □

Bemerkungen.

1. So richtig kann man diesen Beweis erst würdigen, wenn man versucht hat, in einem Punkt (x_0, y_0) der Hyperbel (51) die Tangente auszurechnen, dann ihre Schnittpunkte mit den Asymptoten und dann den Flächeninhalt des Dreiecks.

2. Tatsächlich haben wir aber viel mehr gewonnen als einen netten Satz und einen hübschen Beweis, nämlich die Analogie zwischen Euklidischer Geometrie und Lorentzscher Geometrie. Beide sind Modelle für eine nicht-degenerierte Bilinearform und deren Invarianzgruppe. Daß man jeden Punkt des Kreises in jeden anderen durch eine Rotation überführen kann, ist völlig selbstverständlich. Daß im richtigen Rahmen für die Hyperbel dasselbe gilt, haben wir jetzt gelernt.

Die Dreiteilung des Winkels. Es ist bekannt, daß das klassische Problem der Dreiteilung eines beliebig vorgegebenen Winkels (insbesondere des Winkels von 60^0) mit Zirkel und Lineal nicht lösbar ist. Aber schon um 400 n. Chr. wußte Pappus, daß dieses Problem mit Hilfe einer Hyperbel sehr einfach zu lösen ist.

Lemma 31. *Seien $P \neq Q$ zwei Punkte auf einem Ast einer Hyperbel und seien P^* und Q^* die Schnittpunkte der Sekante $g(P,Q)$ mit den beiden Asymptoten. Dann ist der Mittelpunkt $M = \sigma(P,Q)$ der Strecke PQ gleichzeitig der Mittelpunkte der Strecke P^*Q^*:*

$$\sigma(P,Q) = \sigma(P^*,Q^*).$$

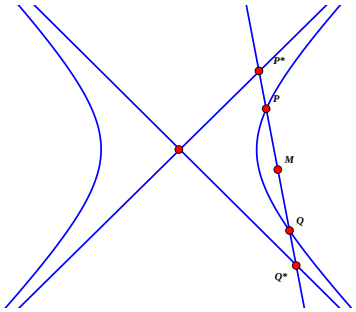
Beweis.

Sei $\langle .,. \rangle$ wieder das Lorentzprodukt. Wie oben zeigt man, daß es genügt, den Satz für die Hyperbel

$$\langle x, x\rangle = 1$$

im $\mathbb{R}^2$ zu zeigen. Wir setzen

$$v := P - M.$$



Dann ist $P = M + v$ und $Q = M - v$. Weil diese beiden Punkte auf der Hyperbel liegen, hat man

$$\langle M+v, M+v\rangle = 1 = \langle M-v, M-v\rangle.$$

Es folgt $\langle M,M\rangle + 2\langle M,v\rangle + \langle v,v\rangle = \langle M,M\rangle - 2\langle M,v\rangle + \langle v,v\rangle$, also

$$\langle M, v\rangle = 0.$$

Ist $P^* = M + tv$, so ist, weil P^* auf einer Asymptote liegt

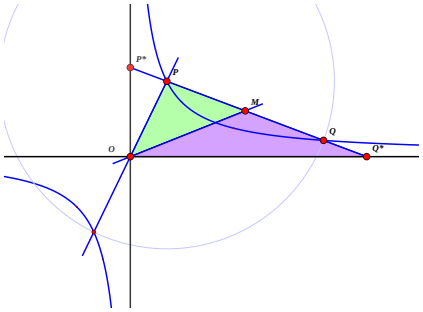
$$0 = \langle M+tv, M+tv\rangle = \langle M,M\rangle + 2t\underbrace{\langle M,v\rangle}_{=0} + t^2\langle v,v\rangle = \langle M,M\rangle - 2t\underbrace{\langle M,v\rangle}_{=0} + t^2\langle v,v\rangle = \langle M-tv, M-tv\rangle$$

Damit ist liegt $M - tv$ auf der anderen Asymptote, und wir haben $Q^* = M - tv$. Aus $|P^* - M| = |tv| = |Q^* - M|$ folgt die Behauptung. □

Satz 60 (Dreiteilung des Winkels). *Sei der Winkel $\phi \in]0, \pi/2[$ gegeben. Wir wählen P auf dem rechten Ast der Hyperbel*

$$x^2 - y^2 = 1$$

so daß $g(O,P)$ mit der positiven x-Achse den Winkel ϕ einschließt.
(Beachte: Diesmal ist es wichtig, daß die Asymptoten der Hyperbel orthogonal zueinander stehen.) Sei $r := |P - O|$. Wir schlagen um P einen Kreis vom Radius $2r$, der die Hyperbel im Punkt Q trifft. Sei $M = \sigma(P,Q)$. Dann drittelt die Gerade $g(O,M)$ den Winkel ϕ.



Beweis. Nach dem Lemma ist M der Mittelpunkt der Strecke P^*Q^*, so daß der Kreis um M durch P^* gerade der Thaleskreis über P^*Q^* ist. Insbesondere liegt O auf diesem Kreis, so daß $|Q^* - M| = |O - M|$ ist: Das Dreieck OMQ^* ist gleichschenklig. NAch Konstruktion ist auch das Dreieck OPM gleichschenklig, und daraus folgt die Behauptung sehr leicht elementargeometrisch:

$$\pi - 2\angle(MOQ^*) + \angle(OMP) = \pi,$$

und daher

$$\angle(OMP) = 2\angle(MOQ^*).$$

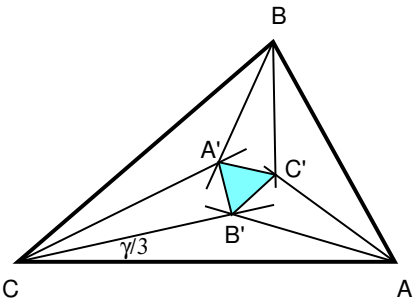
□

5.9.2 Der Satz von Morley und die affine Gruppe des $\mathbb{C}^1$

Hier geht es um folgenden ebenso hübschen wie „exotischen“ Satz:

Satz 61 (Morley, um 1899). *Die paarweisen Schnittpunkte der benachbarten Winkeldrittelnden eines beliebigen Dreiecks bilden die Ecken eines* gleichseitigen *Dreiecks.*

Natürlich kann man das mittels Winkeldrittel-Formeln trigonometrisch beweisen, vgl. Klingenberg. Aber im Gegensatz etwa zu den Schnittpunkten der Winkelhalbierenden (Inkreis) bleibt die Rolle der Winkeldrittelnden offen, und die Bedeutung des Satzes wird durch den Beweis nicht klarer. Der folgende Beweis ist nicht kürzer, vor allem, weil wir erst gewisse elementare Sachverhalte erklären müssen. Er ist „vom höheren Standpunkt aus“ allerdings viel durchsichtiger und erklärt die Rolle der Schnittpunkte der Drittelnden und damit den Satz viel besser.



Eine aparte ON-Basis im $\mathbb{C}^n$. Im Vektorraum $\mathbb{C}^n$ mit dem kanonischen hermiteschen Produkt

$$\langle z, w\rangle = \sum_{k=1}^{n} z_k \bar{w}_k$$

gibt es neben der kanonischen noch eine besonders interessante ONBasis, die aus einer primitiven n-ten Einheitwurzel konstruiert wird, also aus $\xi \in \mathbb{C}$ mit $\xi^n = 1$. *Primitiv* bedeutet, daß $\xi^k \neq 1$ für $0 < k < n$. Das hat zur Folge, daß

$$\xi^k = 1 \iff k \equiv 0 \mod n.$$

Insbesondere kann man

$$\xi = e^{2\pi i/n} = \cos\frac{2\pi}{n} + i\,\sin\frac{2\pi}{n}$$

wählen.

Lemma 32. *Sei ξ eine primitive n-te Einheitswurzel. Dann ist*

$$\frac{1}{\sqrt{n}}(\xi^{(k-1)(l-1)})_{k,l=1,\dots,n} = \frac{1}{\sqrt{n}}\begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \xi & \xi^2 & \dots & \xi^{n-1} \\ 1 & \xi^2 & \xi^4 & \dots & \xi^{2(n-1)} \\ \vdots & & & & \vdots \\ 1 & \xi^{n-1} & \xi^{2(n-1)} & \dots & \xi^{(n-1)(n-1)} \end{pmatrix}.$$

unitär, d.h. die Spalten dieser Matrix bilden eine ONBasis von $\mathbb{C}^n$.

Beweis des Lemmas. Zunächst ist

$$\bar{\xi}\xi = 1 = \xi^{n-1}\xi,$$

also

$$\bar{\xi} = \xi^{n-1}.$$

Weiter ist jedes $\zeta = \xi^k$ ebenfalls eine n-te Einheitswurzel, weil $(\xi^k)^n = (\xi^n)^k = 1$. Und es gilt

$$(1 + \zeta + \zeta^2 + \ldots + \zeta^{n-1})(1 - \zeta) = 1 - \zeta^n = 0,$$

also, wenn $\zeta \neq 1$,

$$1 + \zeta + \zeta^2 + \ldots + \zeta^{n-1} = 0.$$

Deshalb erhalten wir

$$\begin{aligned}\sum_{\mu=1}^{n} \xi^{(k-1)(\mu-1)} \bar{\xi}^{(l-1)(\mu-1)} &= \sum_{\mu=1}^{n} \xi^{(k-1)(\mu-1)} \big(\xi^{n-1}\big)^{(l-1)(\mu-1)} \\ &= \sum_{\mu=1}^{n} \xi^{(\mu-1)(k-1+(n-1)(l-1))} = \sum_{\mu=1}^{n} \xi^{(\mu-1)(k-1-(l-1))} \\ &= \sum_{\mu=1}^{n} (\xi^{k-l})^{\mu-1} = \begin{cases} 0 & \text{für } k \neq l \\ n & \text{für } k = l \end{cases}\end{aligned}$$

□

Bemerkung. Um den lästigen Faktor $\frac{1}{\sqrt{n}}$ loszuwerden, betrachten wir im weiteren auf $\mathbb{C}^n$ das hermitesche Skalarprodukt

$$\langle z, w \rangle := \frac{1}{n} \sum_{i=1}^{n} z_i \bar{w}_i.$$

Dann ist die oben betrachtete Matrix ohne den Vorfaktor unitär.

Fourierentwicklung von Dreiecken. Wie erkennt man ein gleichseitiges Dreieck? In einem Euklidischen Vektorraum V mit gegebener ONBasis $(e_1, \ldots, e_n)$ kann man jeden Vektor $v \in V$ „fourierentwickeln“, also schreiben als

$$v = \sum_{k=1}^{n} c_k e_k, \qquad c_k = \langle v, e_k \rangle.$$

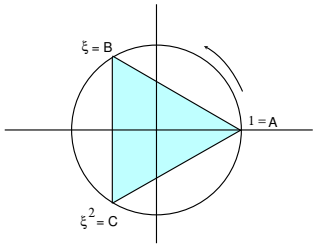
Dasselbe gilt auch in unitären Vektorräumen über $\mathbb{C}$. Im letzten Semester habe ich auch erklärt, was das mit den Fourierreihen aus der Analysis zu tun hat.

Ein Dreieck in der Ebene $\mathbb{R}^2 = \mathbb{C}$ ist gegeben durch sein Ecken z_1, z_2, z_3, also durch einen Punkt $z \in \mathbb{C}^3$. Es läßt sich daher nach der eben eingeführten ON-Basis „fourierentwickeln“

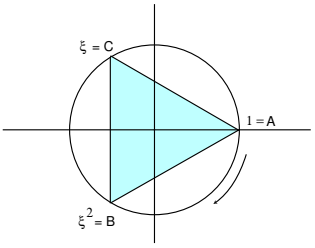
$$z = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = c_1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ \xi \\ \xi^2 \end{pmatrix} + c_3 \begin{pmatrix} 1 \\ \xi^2 \\ \xi \end{pmatrix}, \qquad c_k = \langle z, \begin{pmatrix} 1 \\ \xi^{k-1} \\ \xi^{2(k-1)} \end{pmatrix} \rangle.$$

Dabei ist $\xi = e^{2\pi i/3}$, $c_k \in \mathbb{C}$, und wir haben $\xi^4 = \xi$ benutzt.

Der Vektor $\begin{pmatrix} 1 \\ \xi \\ \xi^2 \end{pmatrix}$ repräsentiert das gleichseitige Standarddreieck mit den dritten Einheitwurzeln als Ecken.



Der Vektor $\begin{pmatrix} 1 \\ \xi^2 \\ \xi \end{pmatrix}$ repräsentiert dasselbe Dreieck, aber mit einer anderen Numerierung der Ecken.



Der erste Vektor $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ schließlich ist ein zum Punkt degeneriertes Dreieck und $c_1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ ist der Schwerpunkt des gegebenen Dreiecks.

Nach Translation (= Positionierung des Schwerpunktes in 0) ist also jedes Dreieck eine Linearkombination aus einem „linksdrehenden“ und einem „rechtsdrehenden“ gleichseitigen Standarddreieck. Entsprechend dem Lemma läßt sich allgemeiner jedes n-Eck aus Standard-n-Ecken linear kombinieren.

Offenbar gilt

$$c_2 = 0 \text{ oder } c_3 = 0 \implies z \text{ ist gleichseitig.}$$

Dieses Kriterium werden wir in unserem Beweis des Satzes von Morley verwenden.

(Weil durch Translation (um c_1) und Drehstreckung (mit einem Faktor c_2) sich jedes linksdrehende gleichseitige Dreieck aus $(1, \xi, \xi^2)$ erhalten läßt und die Fourierkoeffienten eindeutig bestimmt sind, ist ein linksdrehendes Dreieck *genau dann* gleichseitig, wenn $c_3 = 0$. Analog für rechtsdrehende.)

Die affine Gruppe von $\mathbb{C}^1$. Wie charakterisiert man die Schnittpunkte der Winkeldrittelnden? Dreht man in A in positiver Richtung um den Winkel $2\alpha/3$, so landet der Punkt B' gerade auf seinem Spiegelbild bezüglich CA. Dreht man deshalb im Punkt C ebenfalls positiv um den Winkel $2\gamma/3$, so kehrt der Spiegelpunkt wieder nach B' zurück. Bezeichnet man die Drehungen mit f_1 bzw. f_3, so ist B' also ein Fixpunkt, und zwar der einzige, von $f_3 f_1$. Die Charakterisierung der Eckpunkte unseres hoffentlich gleichseitigen Dreieckes sind also Fixpunkte gewisser Abbildungen, die wir zunächst genauer beschreiben:

Die Abbildungen

$$f : \mathbb{C}^1 \to \mathbb{C}^1, z \mapsto az + b$$

mit $a \neq 0$ bilden bezüglich der Komposititon gerade die *affine Gruppe* $\mathrm{Aff}(\mathbb{C}^1)$ des affinen Raumes $\mathbb{C}^1$. Wir nennen

$$\delta(f) := a$$

den *Drehstreckungsanteil* von f und

$$\tau(f) := b$$

den *Translationsanteil* von f.

Die Abbildung $\delta : \mathrm{Aff}(\mathbb{C}^1) \to \mathbb{C}\backslash\{0\}$ erfüllt

$$\delta(f_1 \circ f_2) = \delta(f_1)\delta(f_2), \tag{54}$$

sie ist ein Gruppenhomomorphismus in die multiplikative Gruppe der komplexen Zahlen $\neq 0$.

Ist $\delta(f) = 1$, so nennt man f eine *Translation*.

Wir beachten, daß jedes $f(z) = az + b$, welches keine Translation ist, genau einen *Fixpunkt* $f(z) = z$ hat, nämlich

$$\mathrm{Fix}(f) := \frac{b}{1-a}.$$

Lemma 33 (A. Connes 1998). *Seien $f_1, f_2, f_3 \in \mathrm{Aff}(\mathbb{C}^1)$. Die Produkte*

$$f_1f_2,\ f_2f_3,\ f_3f_1 \text{ und } f_1f_2f_3$$

seien keine *Translationen. Wir setzen $\xi := \delta(f_1f_2f_3)$. Dann sind die folgenden Bedingungen äquivalent:*

(i) $f_1^3f_2^3f_3^3 = \mathrm{id}$

(ii) *ξ ist eine promitive 3. Einheitswurzel und mit $\alpha = \mathrm{Fix}(f_1f_2)$, $\beta = \mathrm{Fix}(f_2f_3)$, $\gamma = \mathrm{Fix}(f_3f_1)$ gilt*

$$\alpha + \beta\xi + \gamma\xi^2 = 0. \tag{55}$$

Beweis des Lemmas von Connes. Nach Voraussetzung ist $f_1f_2f_3$ keine Translation, also $\xi \neq 1$. Darum ist ξ genau dann eine primitive 3. Einheitwurzel, wenn $\xi^3 = 1$.

Es gilt

$$f_1^3f_2^3f_3^3 = \mathrm{id} \iff \delta(f_1^3f_2^3f_3^3) = 1 \wedge \tau(f_1^3f_2^3f_3^3) = 0.$$

Aber nach (54) ist

$$\delta(f_1^3f_2^3f_3^3) = (\delta(f_1)\delta(f_2)\delta(f_3))^3 = \xi^3.$$

Also bleibt zu zeigen, daß

$$b := \tau(f_1^3f_2^3f_2^3) = 0 \iff \alpha + \beta\xi + \gamma\xi^2 = 0. \tag{56}$$

Sei $f_i(z) = a_iz + b_i$. Die Fixpunkte der zweifachen Produkte sind dann

$$\alpha = \frac{a_1b_2 + b_1}{1 - a_1a_2},\ \beta = \frac{a_2b_3 + b_2}{1 - a_2a_3},\ \gamma = \frac{a_3b_1 + b_3}{1 - a_3a_1},$$

und der Translationsanteil ist

$$b = (a_1^2 + a_1 + 1)b_1 + a_1^3(a_2^2 + a_2 + 1)b_2 + (a_1a_2)^3(a_3^2 + a_3 + 1)b_3.$$

Ein längliche Rechnung (vgl. unten) liefert dann

$$b = -\xi a_1^2a_2(a_1 - \xi)(a_2 - \xi)(a_3 - \xi)(\alpha + \beta\xi + \gamma\xi^2).$$

Weil nach Voraussetzung die zweifachen Produkte keine Translationen sind, ist $a_ka_{k+1} \neq 1$, also gilt wegen $\xi = a_1a_2a_3$, daß $a_k - \xi \neq 0$ für alle k. Es folgt (56). □

Beweis des Satzes von Morley. Eine Drehung um einen Punkt $w \in \mathbb{C}$ kann man so realisieren: Erst Translation von w in 0, dann Multiplikation mit einer Zahl a vom Betrag 1 und schließlich Rücktranslation. Sie ist also von der Form

$$f(z) = a(z - w) + w = az + w(1 - a),$$

d.h. sie ist in $\mathrm{Aff}(\mathbb{C}^1)$. Definieren wir f_1, f_3 wie eingangs und f_2 als Drehung in B im positiven Sinne um $2\beta/3$, so sind die zweifachen Produkte keine Translationen: sie sind

nämlich $\neq$ id und haben, wie oben festgestellt, einen Fixpunkt, nämlich jeweils einen der Punkte $C' =: \alpha, A' =: \beta, B' =: \gamma$. Der Rotationsanteil von $f_1 f_2 f_3$ ist

$$\xi = e^{\frac{2}{3}\alpha i} e^{\frac{2}{3}\beta i} e^{\frac{2}{3}\gamma i} = e^{\frac{2}{3}(\alpha+\beta+\gamma)i} = e^{\frac{2}{3}\pi i} \neq 1.$$

Also ist auch $f_1 f_2 f_3$ keine Translation.

Schließlich ist

$$f_1^3 f_2^3 f_3^3 = \mathrm{id}\,.$$

Nach der nebenstehenden Skizze ist nämlich zunächst

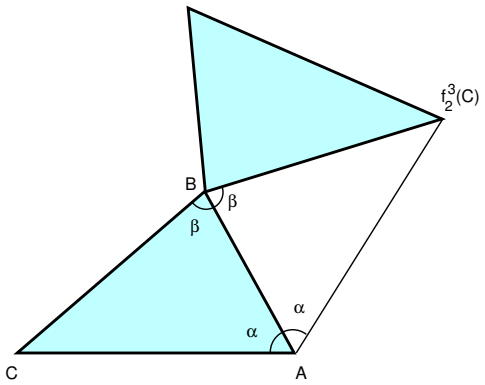
$$f_1^3 f_2^3 f_3^3(C) = f_1^3 f_2^3(C) = C.$$

Ebenso sieht man (zyklische Vertauschung), daß

$$f_2^3 f_3^3(A) = A$$

und daher

$$f_1^3 f_2^3 f_3^3(A) = A.$$



Aber eine affine Abbildung mit zwei Fixpunkten ist die Identität:

$$f_1^3 f_2^3 f_3^3 = \mathrm{id}\,.$$

Nach dem Lemma von Connes folgt daher

$$0 = \alpha + \beta\xi + \gamma\xi^2 = \alpha + \beta\bar{\xi}^2 + \gamma\bar{\xi} = \langle \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}, \begin{pmatrix} 1 \\ \xi^2 \\ \xi \end{pmatrix} \rangle.$$

In der Fourierzerlegung von $\begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}$ kommt also nur eines der gleichseitigen Standarddreiecke vor, es ist also selber gleichseitig. □

Bemerkungen.

1. Der dargestellte Beweis (nicht nur des Lemmas) stammt vom Fields-Medal-Preisträger Alain Connes unter Mitwirkung von Napoleon Bonaparte, vgl. den Artikel von Connes in der Festschrift zum 40jährigen Bestehen des Institut des Hautes Études Scientifiques (IHES) in Bures [Les Relations entre les Mathématiques et la Physique Théorique, Bures-sur Yvette].

2. Die im Beweis des Lemmas von Connes unterschlagene Rechnung dient nur dem Nachweis, daß die beiden Ausdrücke für b gleich sind. Das kann man zum Beispiel von einem Mathematica Programm prüfen lassen. Ein solches (von Ekki Tjaden) fügen wir hier an:

```
(*morley2.m*)(*zum Satz von Morley/Beweis Lemma von Connes*)
(*ekki : Mon Sep 24 15 : 21 : 13 MDT 2001*)(*general nonsense*)
si = Simplify;
    (*************************************************)
    (*Notation der affinen Abbildungen f[z] = a z + b
      ** als Liste {a, b} mit 2 Eintraegen.
      ** Dann haben wir Funktionen fuer
      ** A - Teil, B - Teil und Fixpunkt*)
    A[f : {_, _}] := f[[1]]
    B[f : {_, _}] := f[[2]]
    Fix[f : {_, _}] := B[f]/(1 - A[f])
    (*Die Komposition/Multiplikation ist dann f*g == Mult[f, g]*)
    Mult[f : {_, _}, g : {_, _}] := {A[f] A[g], A[f] B[g] + B[f]}
    MultList[ff : {{_, _} ..}] := Fold[Mult, {1, 0}, ff]
    (*****************************************)
    (*Nun zum gegebenen Dreieck.** Wir drehen jeweils mit 2/3 des Dreieckwinkels um diese Ecke*)
    (*Hier sind die Drehungen-- einfach so modulo 3--*)
    f[0] = {a[0], b[0]};
    f[1] = {a[1], b[1]};
    f[2] = {a[2], b[2]};
    f[n_] := f[n - 3] /; n > 2

    Unprotect[Power, Times];
    (*Ueber die a[j] wissen wir ..*)
    a[0] a[1] a[2] := \[Xi]
    (*..das ist eine dritte Einheitswurzel.
  ** Falls irgendwelche hoeheren Potenzen vorkommen,
  ** dann gibt es noch die folgenden Moeglichkeiten.*)
    a[0]^n_ a[1] a[2] := \[Xi] a[0]^(n - 1)
    a[0] a[1]^n_ a[2] := \[Xi]  a[1]^(n - 1)
    a[0] a[1] a[2]^n_ := \[Xi]  a[2]^(n - 1)
    a[0] a[1]^m_ a[2]^k_ := \[Xi]  a[1]^(m - 1) a[2]^(k - 1)
    a[0]^n_ a[1] a[2]^k_ := \[Xi]  a[0]^(n - 1) a[2]^(k - 1)
    a[0]^n_ a[1]^m_ a[2] := \[Xi]  a[0]^(n - 1) a[1]^(m - 1)
    a[0]^n_ a[1]^m_ a[2]^k_ := \[Xi] ^Min[n, m, k]*
a[0]^(n - Min[n, m, k]) a[1]^(m - Min[n, m, k]) a[2]^(k - Min[n, m, k])

(*Und ueber \[Xi]  wissen wir ..*)
    \[Xi] ^2 := -1 - \[Xi]
    \[Xi] ^n_ := \[Xi] ^Mod[n, 3] /; n > 2
    Protect[Power, Times];

    (*Die folgende Abbildung ist die Identitaet,
   ** weil zuviele Fixpunkte da sind*)

    myid = si[MultList[{f[0], f[0], f[0], f[1], f[1], f[1], f[2], f[2], f[2]}]];
    (*Nur so-- als Test--*)
    1 == A[myid];
    (*Der B - Teil davon ist also eine evtl.komplizierte Darstellung von 0*)
    myzero = si[B[myid]];
    (*****************************************)
    (*Nun zum inneren Dreieck mit der Ecke q[j + 2]  ** als Fixpunkt von f[j]*f[j + 1]*)

    Do[q[j] = si[Fix[Mult[f[j + 1], f[j + 2]]]], {j, 0, 2}];

    (*Dieser Testwert sollte fuer ein gleichseitiges Dreieck 0 sein*)
    testvalue = q[1] + \[Xi]  q[2] + \[Xi] ^2 q[0];

    (*Etwas angefuettert behaupten wir, dass dieser Ausdruck 0 ist.*)
    claimedzero = a[0]^2 a[1] Product[\[Xi] - a[j], {j, 0, 2}] testvalue;
    (*Wir bringen alles auf einen Bruchstrich
  ** und schauen uns den Zaehler an ...*)
    Behauptung =0 == si[Numerator[Together[myzero - claimedzero]]]
    If[Behauptung == True, Print["Alles ok!"], Print["Hmm, so ein Mist."]]

    Out[362]=
      True
      Alles ok!
```

6 Quadriken

Quadriken sind verallgemeinerte Kegelschnitte: Teilmengen eines affinen Raumes, die man durch eine quadratische Gleichung beschreiben kann. Wir wollen in diesem Abschnitt den Zusammenhang zwischen den geometrischen und den algebraischen Daten untersuchen und wir wollen Quadriken unter gewissen Voraussetzungen klassifizieren. „Klassifikation“ setzt eine Äquivalenzrelation voraus, und die wiederum hängt zusammen mit den gegebenen Strukturen. Betrachtet man Quadriken als Teilmengen eines beliebigen affinen Raumes, so ist die Äquivalenz unter affinen Transformationen die natürliche Wahl. Betrachtet man hingegen Quadriken (wie zum Beispiel oben die Kegelschnitte) in einem *Euklidischen* affinen Raum, so bietet sich die Äquivalenz unter *längentreuen* affinen Transformationen an. Im „rein affinen“ Fall erweist sich überdies der Grundkörper als wichtig.

Zur Vereinfachung der Notation nehmen wir an, daß in unserem affinen Raum ein Nullpunkt ausgezeichnet ist, d.h. wir betrachten den Fall $\mathcal{A} = V$.

Generalvoraussetzung. Sei V ein n-dimensionaler K-Vektorraum der Charakteristik $\neq 2$, d.h. $1 + 1 \neq 0$, und $n < \infty$.

6.1 Definition der Quadrik und eindeutige Darstellung

Definition 34 (Quadrik). Eine Teilmenge $Q \subset V$ heißt eine *Quadrik* oder *Hyperfläche 2. Ordnung*,

- wenn sie nicht in einer affinen Hyperebene enthalten ist

und wenn gilt: Es gibt

- eine nicht-triviale symmetrische Bilinearform $\beta : V \times V \to K$,
- eine Linearform $\omega : V \to K$ und
- $\alpha \in K$,

so daß

$$Q = \{x \in V \mid \beta(x, x) + 2\omega(x) + \alpha = 0\}.$$

Wir schreiben zur Abkürzung $Q = Q(\beta, \omega, \alpha)$ oder

$$Q : \ \beta(x, x) + 2\omega(x) + \alpha = 0.$$

Beachten Sie, daß (β, ω, α) durch Q offenbar *nicht* eindeutig bestimmt ist, man kann alle drei mit einem Faktor $\neq 0$ multiplizieren. Wir zeigen später, daß das die einzige Mehrdeutigkeit ist.

Beispiel 35. Bezüglich einer Basisdarstellung $x = \sum x_i b_i$ ist

$$\beta(x, x) + 2\omega(x) + \alpha = \sum_{i,j} x_i x_j \beta(b_i, b_j) + 2 \sum_i x_i \omega(b_i) + \alpha.$$

Eine Quadrik ist dann also gegeben durch eine quadratische Gleichung für die Komponenten x_i. Daher der Name.

□

Beispiel 36. Kegelschnitte im Sinne der Definition 33 sind Quadriken.

□

Beispiel 37 (Zylinder, Kegel). Der Zylinder

$$Z := \{(x, y, z) \,|\, x^2 + y^2 = 1\} \subset \mathbb{R}^3$$

und der Doppelkegel

$$C := \{(x, y, z) \,|\, x^2 + y^2 = z^2\} \subset \mathbb{R}^3$$

sind Quadriken.

□

Schnitt von Quadriken mit Geraden. Als erstes wollen wir die Frage klären, wieweit (β, ω, α) durch die Menge Q festgelegt ist. Dazu untersuchen wir den Schnitt von Geraden $x + Kv$ mit Q. Sei also $Q = Q(\beta, \omega, \alpha) \subset V$ eine Quadrik und seien $x, v \in V, v \neq 0$. Dann gilt

$$\begin{aligned} &\beta(x + tv, x + tv) + 2\omega(x + tv) + \alpha \\ &\quad = \beta(v, v)t^2 + 2(\beta(x, v) + \omega(v))t + (\beta(x, x) + 2\omega(x) + \alpha). \end{aligned} \tag{57}$$

Als Polynom vom Grad ≤ 2 hat das keine, eine oder zwei Nullstellen, oder es ist 0. Also gilt für eine Gerade $g = x + Kv$.

$$\#(g \cap Q) \in \{0, 1, 2\} \quad \text{oder} \quad g \subset Q.$$

Etwas genauer: Sei $x \in Q$. Dann verschwindet der konstante Term in (57) und $t = 0$ ist eine Nullstelle. Wir erhalten:

$$\#(g \cap Q) = \begin{cases} \infty, & \text{falls } \beta(v, v) = 0 \wedge \beta(x, v) + \omega(v) = 0, \\ 1, & \text{falls } \beta(v, v) \neq 0 \wedge \beta(x, v) + \omega(v) = 0, \\ 1, & \text{falls } \beta(v, v) = 0 \wedge \beta(x, v) + \omega(v) \neq 0, \\ 2, & \text{falls } \beta(v, v) \neq 0 \wedge \beta(x, v) + \omega(v) \neq 0. \end{cases}$$

Ist $\beta(v, v) \neq 0$, so ist $\beta(v, .) \neq 0$ und $\beta(x, v) + \omega(v) = 0$ kann nicht für alle $x \in Q$ gelten, sonst läge Q in einer affinen Hyperebene. Bei $\beta(v, v) \neq 0$ gibt es also stets eine Gerade in Richtung v mit $\#(g \cap Q) = 2$, bei $\beta(v, v) = 0$ niemals. Vgl. auch Satz/Definition 56 über die Tangenten an die Kegelschnitte.

Definition 35 (Ausnahmerichtungen). Seien $Q = Q(\beta, \omega, \alpha) \subset V$ eine Quadrik und $v \in V \backslash \{0\}$. Gilt eine der beiden folgenden äquivalenten Bedingungen:

(i) Es gibt eine Gerade in Richtung v, die die Quadrik genau zweimal schneidet:

$$\exists_{x \in V} \#((x + Kv) \cap Q) = 2,$$

(ii) $\beta(v, v) \neq 0$,

so heißt Kv eine *Nicht-Ausnahmerichtung*, andernfalls eine *Ausnahmerichtung*.

Die Ausnahmerichtungen sind also durch die Menge Q allein bestimmt. Damit bestimmt Q schon eine Eigenschaft von β.

Beispiel 38. Die Hyperbel hat 2 Ausnahmerichtungen, nämlich die Asymptotenrichtungen. Die Ellipse hat keine Ausnahmerichtung. Die Parabel hat eine Ausnahmerichtung, nämlich die Achsrichtung. Die Quadrik

$$Q := \{(x, y) \in \mathbb{R}^2 \,|\, xy = 0\}$$

besteht aus den beiden Achsen des $\mathbb{R}^2$, und diese sind die beiden Ausnahmerichtungen.

□

Definition 36 (Sehnenmittelpunkt, Diametralhyperebene). Sei $v \in V \backslash \{0\}$, so daß Kv keine Ausnahmerichtung von $Q = Q(\beta, \omega, \alpha)$ ist.

(i) Zu jedem $x \in Q$ gibt es genau ein $\tilde{x} \in Q$ mit

$$(x + Kv) \cap Q = \{x, \tilde{x}\}.$$

(Der Fall $x = \tilde{x}$ ist möglich.) Sei

$$m(x, v) := \sigma(x, \tilde{x}) = \frac{1}{2}(x + \tilde{x})$$

der Mittelpunkt der Sehne $x\tilde{x}$.

(ii) Wir setzen

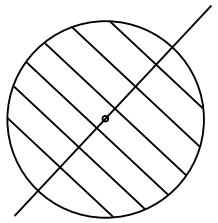
$$D_v := \{y \in V \,|\, \beta(y, v) + \omega(v) = 0\}.$$

Weil $\beta(v, .) \neq 0$, ist D_v eine affine Hyperebene, die sogenannte *Diametralhyperebene* zu v.

Beispiel 39. Sei $Q = \{(x_1, x_2) \,|\, x_1^2 + x_2^2 = r^2\} \subset \mathbb{R}^2$ ein Kreis und $v = (v_1, v_2) \neq 0$. Dann ist

$$D_v = \{x | x_1 v_1 + x_2 v_2 = 0\}.$$

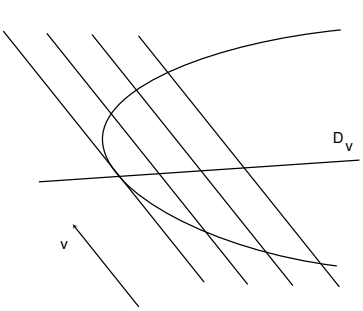
Das ist die Gerade durch 0 senkrecht zu v, eine „Durchmessergerade“. Sie enthält alle Mittelpunkte der Kreissehnen parallel zu v. Die Diametralhyperebenen sind eine Verallgemeinerung der Kreisdurchmesser:

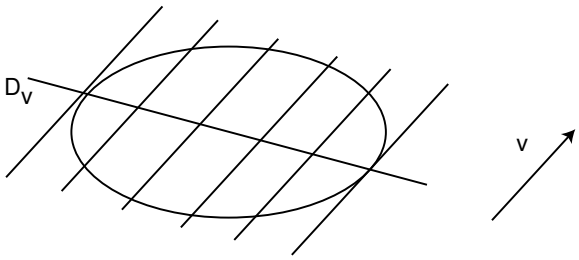


□

Satz 62 (Diametralhyperebene). *D_v ist die einzige affine Hyperebene von V mit*

$$\{m(x, v) \,|\, x \in Q\} \subset D_v.$$





Beweis. *Zu („⊂“).* Sei $x \in Q$ und $\tilde{x}$ der zweite Schnittpunkt von $x + Kv$ mit Q. Nach Voraussetzung ist $\beta(v,v) \neq 0$ und nach (57) ist

$$\tilde{x} = x - 2\,\frac{\beta(x,v) + \omega(v)}{\beta(v,v)}\,v.$$

Also ist

$$m(x,v) = x - \frac{\beta(x,v) + \omega(v)}{\beta(v,v)}\,v$$

und

$$\beta(m(x,v),v) + \omega(v) = \beta(x,v) - (\beta(x,v) + \omega(v)) + \omega(v) = 0.$$

Zu („Einzigkeit“). Annahme: Es gibt eine weitere Hyperebene D, die alle Mittelpunkt $m(x,v)$ enthält. Dann wäre $D \cap D_v$ ein affiner Unterraum der Kodimension 2, der alle $m(x,v)$ enthält. Dann läge aber Q in der affinen Hyperebene

$$(D \cap D_v) + Kv.$$

Widerspruch. □

Korollar 6. *Ist*

$$Q = Q(\beta,\omega,\alpha) = Q(\tilde{\beta},\tilde{\omega},\tilde{\alpha})$$

eine Quadrik in V, so ist

$$\beta(u,v) = 0 \iff \tilde{\beta}(u,v) = 0.$$

Beweis. Sei o.E. $u \neq 0 \neq v$. Ist Kv keine Ausnahmerichtung von Q, so ist D_v durch Q allein bestimmt. Der Richtungsvektorraum von D_v ist aber gerade der Kern von $\beta(.,v)$ und von $\tilde{\beta}(.,v)$. Also folgt die Behauptung. Gleiches gilt, falls Ku keine Ausnahmerichtung von Q ist.

Seien also Ku und Kv Ausnahmerichtungen und o.E. $Ku \neq Kv$. Dann gilt also

$$\beta(u,u) = \beta(v,v) = 0 = \tilde{\beta}(u,u) = \tilde{\beta}(v,v).$$

Aus $\beta(u,v) = 0$ folgt dann

$$\beta(u+v,u+v) = \beta(u,u) + 2\beta(u,v) + \beta(v,v) = 0.$$

Weil $Ku \neq Kv$ ist $u + v \neq 0$ und deshalb $K(u+v)$ eine Ausnahmerichtung. Also ist auch

$$0 = \tilde{\beta}(u+v,u+v) = 2\tilde{\beta}(u,v).$$

Damit haben wir auch für Ausnahmerichtungen gezeigt:

$$\beta(u,v) = 0 \implies \tilde{\beta}(u,v) = 0.$$

□

Damit haben wir eine enge Beziehung zwischen β und $\tilde{\beta}$ gefunden. Die folgenden Lemmata runden das ab.

Lemma 34. *Seien $\omega, \tilde{\omega} \in V^*$ zwei Linearformen auf V mit*

$$\operatorname{Kern}\omega = \operatorname{Kern}\tilde{\omega}.$$

Dann gibt es ein $\lambda \in K\backslash\{0\}$ mit

$$\tilde{\omega} = \lambda\omega.$$

Beweis. Ist $\omega = 0$, so ist $\operatorname{Kern}\tilde{\omega} = \operatorname{Kern}\omega = V$, also $\tilde{\omega} = 0 = \omega$. Andernfalls gibt es eine Basis $(b_1, \ldots, b_n)$ von V, so daß $(b_2, \ldots, b_n)$ eine Basis von $\operatorname{Kern}\omega = \operatorname{Kern}\tilde{\omega}$ ist. Dann ist

$$\omega(b_1) \neq 0 \neq \tilde{\omega}(b_1)$$

und

$$\tilde{\omega} = \frac{\tilde{\omega}(b_1)}{\omega(b_1)}\omega,$$

wie man sofort auf der Basis nachprüft. □

Lemma 35. *Sind $\beta, \tilde{\beta} : V \times V \to K$ nicht-triviale symmetrische Bilinearformen mit*

$$\operatorname{Kern}\beta(u,.) = \operatorname{Kern}\tilde{\beta}(u,.) \quad \text{für alle } u \in V,$$

so gibt es ein $\lambda \in K\backslash\{0\}$ mit

$$\tilde{\beta} = \lambda\beta.$$

Beweis. Es genügt zu zeigen: Es gibt ein $\lambda \in K$, so daß

$$\tilde{\beta}(u,.) = \lambda\beta(u,.)$$

für alle $u \in V$ mit $\beta(u,.) \neq 0$. Wenn $\beta(u,.) = 0$, ist nämlich $\tilde{\beta}(u,.) = 0$, und die Gleichheit gilt auch dann.

Nach Lemma 34 gibt es zunächst zu jedem $u \in V$ mit $\beta(u,.) \neq 0$ ein eindeutig bestimmtes $\lambda_u \in K\backslash\{0\}$ mit

$$\tilde{\beta}(u,.) = \lambda_u\beta(u,.).$$

Wir müssen zeigen, daß λ_u unabhängig von u ist. Seien dazu $u_1, u_2 \in V$ mit $\beta(u_i,.) \neq 0$. Dann gibt es ein $v \in V$ mit

$$\beta(u_1, v) \neq 0 \neq \beta(u_2, v). \tag{58}$$

Zunächst gibt es nämlich v_i mit $\beta(u_i, v_i) \neq 0$. Ist $\beta(u_1, v_2) \neq 0$ oder $\beta(u_2, v_1) \neq 0$, so wählen wir $v := v_2$ bzw. $v := v_1$. Andernfalls tut es $v = v_1 + v_2$.

Aus (58) folgt nun

$$\lambda_{u_i}\beta(u_i, v) = \tilde{\beta}(u_i, v) = \tilde{\beta}(v, u_i) = \lambda_v\beta(u_i, v),$$

also

$$\lambda_{u_1} = \lambda_v = \lambda_{u_2}.$$

□

Damit ergibt sich nun der

Satz 63 (Eindeutige Darstellung von Quadriken). *Ist*

$$Q = Q(\beta, \omega, \alpha) = Q(\tilde{\beta}, \tilde{\omega}, \tilde{\alpha})$$

eine Quadrik in V, so gibt es $\lambda \neq 0$ mit

$$\tilde{\beta} = \lambda\beta,\ \tilde{\omega} = \lambda\omega,\ \tilde{\alpha} = \lambda\alpha.$$

Beweis. Nach Korollar 6 ist

$$\operatorname{Kern} \beta(u, .) = \operatorname{Kern} \tilde{\beta}(u, .)$$

für alle $u \in V$. Nach Lemma 35 gibt es $\lambda \neq 0$ mit

$$\tilde{\beta} = \lambda\beta.$$

Daher gilt

$$\lambda\beta(x, x) + 2\tilde{\omega}(x) + \tilde{\alpha} = 0 \iff x \in Q \iff \beta(x, x) + 2\omega(x) + \alpha = 0.$$

Für $x \in Q$ folgt

$$(\tilde{\omega} - \lambda\omega)(x) + \tilde{\alpha} - \lambda\alpha = 0.$$

Weil Q nicht in einer Hyperebene liegt, ist also

$$\tilde{\omega} = \lambda\omega, \qquad \tilde{\alpha} = \lambda\alpha.$$

□

6.2 Affine Klassifikation von Quadriken I

Affine Bilder von Quadriken sind wieder Quadriken. Weil affine Transformationen f umkehrbar sind, genügt es zu zeigen, daß mit Q auch $f^{-1}(Q)$ eine Quadrik ist. Das ist etwas einfacher zu formulieren.

Sei also $f : V \to V$ eine affine Transformation, also eine (bijektive!) Abbildung der Form

$$f(x) = \phi(x) + a$$

mit $\phi \in \mathrm{GL}(V)$ und $a \in V$. Sei weiter $Q = Q(\beta, \omega, \alpha) \subset V$ eine Quadrik.

Wir betrachten die Menge $f^{-1}(Q)$. Wir finden

$$\begin{aligned} f(x) \in Q \iff 0 =& \beta(f(x), f(x)) + 2\omega(f(x)) + \alpha \\ =& \beta(\phi(x), \phi(x)) + 2(\beta(\phi(x), a) + \omega(\phi(x))) + \beta(a, a) + 2\omega(a) + \alpha. \end{aligned}$$

Definition 37. Sind $\omega \in V^*$ eine Linearform und β eine symmetrische Bilinearform auf V, und ist $\phi \in \mathrm{Aut}(V)$, so definieren

$$\phi^*\omega(x) := \omega(\phi(x)) \tag{59}$$

eine neue Linearform und

$$\phi^*\beta(x, y) := \beta(\phi(x), \phi(y)) \tag{60}$$

eine neue symmetrische Bilinearform.

Damit könne wir das Ergebnis der obigen Rechnung so formulieren:

Lemma 36. *Ist $Q(\beta, \omega, \alpha)$ eine Quadrik und*

$$f(x) = \phi(x) + a$$

eine affine Transformation, so ist

$$f^{-1}(Q) = Q(\phi^*\beta,\ \phi^*(\beta(a, .) + \omega),\ \beta(a, a) + 2\omega(a) + \alpha).$$

Definition 38 (Affine Äquivalenz). Zwei Quadriken in V heißen *(affin) äquivalent*, wenn sie durch eine affine Transformation

$$f(x) = \phi(x) + a$$

mit $\phi \in \mathrm{GL}(V)$ und $a \in V$ in einander übergehen.

Fragen. Wieviele Äquivalenzklassen gibt es? Gibt es in jeder Klasse eine „Normalform“? Wie sieht die aus?

Wir teilen die Quadriken zunächst grob in drei „Typen“ ein.

6.2.1 Affine Typen von Quadriken

Sei wieder V ein endlich-dimensionaler Vektorrraum über einem Körper der Charakteristik $\neq 2$.

Definition 39 (Mittelpunkt). Sei $Q \subset V$ eine Quadrik.

(i) $z \in V$ heißt *ein Mittelpunkt von* Q, wenn

$$\forall_{v \in V} \left(z + v \in Q \implies z - v \in Q\right).$$

(ii) Q heißt *Mittelpunktsquadrik*, wenn Q einen Mittelpunkt hat.

Bezeichne mit $\Sigma_z : V \to V, x \mapsto 2z - x$ die *Punktspiegelung* an z. Das ist eine affine Abbildung. z ist ein Mittelpunkt von Q, wenn Q invariant unter Σ_z ist.

Beispiel 40. Eine Ellipse oder Hyperbel hat offensichtlich einen Mittelpunkt.

Bei einem elliptischen Zylinder

$$\{(x, y, z) \in \mathbb{R}^3 \mid \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1\}$$

ist jeder Punkt auf der Achse $x = y = 0$ ein Mittelpunkt.

Die Parabel hat keinen Mittelpunkt.

□

Lemma 37. *Sei $Q \subset V$ eine Quadrik und $f : V \to V$ eine affine Transformation. Dann ist $f(z)$ genau dann eine Mittelpunkt von $f(Q)$, wenn z ein Mittelpunkt von Q ist.*

Beweis. Bezeichnet Σ_z die Punktspiegelung an z, so ist

$$f(\Sigma_z(x)) = f(2z - x) = 2f(z) - f(x) = \Sigma_{f(z)}(f(x)).$$

Daraus folgt die Behauptung. □

Satz 64. *Seien $Q = Q(\beta, \omega, \alpha)$ eine Quadrik und $z \in V$. Dann sind folgende Aussagen äquivalent:*

(i) z ist Mittelpunkt von Q.

(ii) $\beta(z, .) + \omega = 0$, d.h. alle Diametralhyperebenen gehen durch z.

(iii) Es gibt $\gamma \in K$ mit

$$Q = \{x \in V \mid \beta(x - z, x - z) = \gamma\}.$$

Beweis. $\underline{(i) \implies (ii)}$. Sei $z + v \in Q$. Dann gilt

$$\beta(z + v, z + v) + 2\omega(z + v) + \alpha = 0,$$
$$\beta(z - v, z - v) + 2\omega(z - v) + \alpha = 0,$$

und daraus folgt durch Subtraktion

$$4\beta(z,v) + 4\omega(v) = 0.$$

Ist also $x \in Q$, so folgt

$$\beta(z, x-z) + \omega(x-z) = 0$$

oder

$$\beta(z,x) + \omega(x) = \beta(z,z) + \omega(z).$$

Weil Q nach Definition der Quadrik nicht in einer affinen Hyperebene liegt, folgt daraus

$$\beta(z,.) + \omega = 0.$$

(ii) $\Longrightarrow$ (iii). Aus $\beta(z,.) + \omega = 0$ folgt für $x \in V$

$$\beta(x-z, x-z) = \beta(x,x) - 2\beta(z,x) + \beta(z,z) = \beta(x,x) + 2\omega(x) + \alpha + (\beta(z,z) - \alpha).$$

Also ist

$$x \in Q \iff \beta(x-z, x-z) = \beta(z,z) - \alpha.$$

(iii) $\Longrightarrow$ (i). Wir haben

$$\beta((z-v)-z, (z-v)-z) = \beta(-v,-v) = \beta(v,v) = \beta((z+v)-z, (z+v)-z).$$

Gilt also *(iii)*, so folgt (i). □

Satz 65 (und Definition). *Seien Q eine Quadrik und $z \in Q$ ein Punkt auf Q. Dann gilt:*

$$z \text{ ist Mittelpunkt von } Q \iff \forall_{x \in Q}\, z + K(x-z) \subset Q.$$

Ein solches z heißt eine Spitze von Q.

Beweis. *Zu ($\Leftarrow$.).* Trivial nach Definition eines Mittelpunktes.

Zu ($\Longrightarrow$.). Ist $z \in Q$ ein Mittelpunkt, so folgt nach dem letzten Satz:

$$x \in Q \iff \beta(x-z, x-z) = \gamma = \beta(z-z, z-z) = 0.$$

Aus der rechten Seite folgt aber für alle $t \in K$

$$\beta(z + t(x-z) - z, z + t(x-z) - z) = t^2\beta(x-z, x-z) = 0.$$

□

Bemerkung. Sind z_0, z_1 Mittelpunkte von Q und ist z_0 eine Spitze, so ist

$$\Sigma_{z_1}(z_0) = -z_0 + 2z_1 \in Q.$$

Also ist nach dem letzten Satz auch

$$z_0 + \frac{1}{2}(-z_0 + 2z_1 - z_0) = z_1 \in Q.$$

Ist also *ein* Mittelpunkt von Q eine Spitze, so sind auch alle anderen Spitzen.

Definition 40 (Affine Typen von Quadriken). Wir teilen die Quadriken in folgende drei Typen ein:

1. Eine Quadrik heißt *echte Mittelpunktsquadrik*, wenn sie (wenigstens) einen Mittelpunkt, aber keine Spitze besitzt.
2. Eine Quadrik heißt *Kegel*, wenn sie eine Spitze besitzt.
3. Eine Quadrik, die keinen Mittelpunkt besitzt, heißt ein *Paraboloid*.

Nach Lemma 37 ist der Typ einer Quadrik invariant unter affinen Transformationen.

6.2.2 Mittelpunktsquadriken

Wir untersuchen zunächst die Äquivalenz von Mittelpunktsquadriken in zwei Schritten: Zuerst zeigen wir, daß jede solche äquivalent zu einer Quadrik mit $\omega = 0$ und $\alpha \in \{0, 1\}$ ist. Dieser Schritt entspricht im wesentlichen der quadratischen Ergänzung zur Beseitigung linearer Glieder in der Darstellung durch eine quadratische Gleichung. Im zweiten Schritt untersuchen wir, wann zwei Quadriken von dieser speziellen Form äquivalent sind.

1. Ist $Q(\beta, \omega, \alpha)$ eine Mittelpunktsquadrik, so läßt sich $a = z$ nach Satz 64 so wählen, daß $\beta(a, .) + \omega = 0$. Dann wird (mit $\phi = \mathrm{id}$)

$$f^{-1}(Q) = Q(\beta, 0, \omega(a) + \alpha). \tag{61}$$

   Jede Mittelpunktsquadrik ist also *translationsäquivalent* zu einer Quadrik

$$Q = Q(\beta, 0, \alpha).$$

   Umgekehrt ist für jede Quadrik mit solchen Daten offenbar 0 ein Mittelpunkt. Vgl. auch Satz 64.

2. Ist $Q = Q(\beta, 0, \alpha)$ ein Kegel, so gibt es $z \in Q$ mit $0 = \beta(z, .) + \omega = \beta(z, .)$. Dann ist $0 = \beta(z, z) + \alpha$, also $\alpha = 0$. Umgekehrt folgt aus $\alpha = 0$, daß 0 eine Spitze, also Q ein Kegel ist.

3. Jede echte Mittelpunktsquadrik ist also translationsäquivalent zu einer Quadrik $Q(\beta, 0, 1)$ und jeder Kegel translationsäquivalent zu einer Quadrik $Q(\beta, 0, 0)$.

4. Sind zwei echte Mittelpunktsquadriken

$$Q_1 = Q(\beta_1, 0, \alpha_1) \quad \text{und} \quad Q_2 = Q(\beta_2, 0, \alpha_2)$$

   affin äquivalent, so gibt es $f = \phi(.) + a$ und $\lambda \in K \backslash \{0\}$ mit

$$\begin{aligned} \lambda\beta_2 &= \phi^*\beta_1, \\ \lambda\omega_2 = 0 &= \phi^*\beta_1(a, .), \\ \lambda\alpha_2 &= \beta_1(a, .) + \alpha_1. \end{aligned}$$

   Es folgt $\beta_1(a, .) = 0, \lambda = \frac{\alpha_1}{\alpha_2}$ und

$$\frac{\alpha_1}{\alpha_2}\beta_2 = \phi^*\beta_1.$$

   Umgekehrt impliziert die letzte Gleichung offenbar (mit $a = 0$) die affine Äquivalenz.

5. Ebenso sieht man, daß zwei Kegel

$$Q_1 = Q(\beta_1, 0, 0) \quad \text{und} \quad Q_2 = Q(\beta_2, 0, 0)$$

   genau dann äquivalent sind, wenn es ein $\lambda \neq 0$ gibt, so daß

$$\lambda\beta_2 = \phi^*\beta_1.$$

Definition 41. Zwei symmetrische Bilinearformen $\beta, \tilde{\beta}$ auf V heißen *äquivalent*, wenn es ein $\phi \in \mathrm{Aut}(V)$ mit

$$\tilde{\beta} = \phi^*\beta$$

gibt.

Die Frage nach der Äquivalenz von (Mittelpunkts)Quadriken hängt also wesentlich an der Frage, wann zwei symmetrische Bilinearformen auf V äquivalent sind.

6.2.3 Paraboloide

Satz 66. *Eine Quadrik $Q = Q(\beta, \omega, \alpha)$ ist genau dann ein Paraboloid, wenn es ein $a \in V$ gibt, für das*

$$\beta(a, .) = 0 \text{ und } \omega(a) \neq 0.$$

Beweis. Zu ($\Leftarrow$). Nach Voraussetzung gibt es $a \in V$ mit $\beta(a, .) = 0$ und $\omega(a) \neq 0$. Gäbe es einen Mittelpunkt b, so wäre $\beta(b, .) + \omega = 0$, also insbesondere

$$\beta(b, a) + \omega(a) = 0.$$

Widerspruch!

Zu ($\Rightarrow$). Wir setzen

$$V_0 := \{x \in V \mid \beta(x, .) = 0\}.$$

V_0 heißt auch der *Annihilator* von β. Wir wählen einen zu V_0 komplementären Unterraum:

$$V = V_0 \oplus V_1.$$

Dann ist

$$\beta' : V_1 \to V^*, x_1 \mapsto \beta(x_1, .) \text{ injektiv.}$$

Andrerseits folgt aus $\beta(x, .)|_{V_1} = 0$, daß für alle $y_i \in V_i$ auch

$$\beta(x, y_0 + y_1) = \underbrace{\beta(x, y_0)}_{=0} + \beta(x, y_1) = 0.$$

Also ist sogar

$$\beta' : V_1 \to V_1^*, x_1 \mapsto \beta(x_1, .)|_{V_1} \text{ injektiv}$$

und damit surjektiv. Es gibt also ein $x_1 \in V_1$ mit

$$\beta(x_1, .)|_{V_1} = -\omega|_{V_1}.$$

Wäre $\omega|_{V_0} = 0$, so wäre für alle $y_i \in V_i$

$$\beta(x_1, y_0 + y_1) = \beta(x_1, y_1) = -\omega(y_1) == -\omega(y_0) - \omega(y_1) = -\omega(y_0 + y_1),$$

d.h.

$$\beta(x_1, .) + \omega = 0,$$

und x_1 ein Mittelpunkt. Widerspruch! Also ist $\omega|_{V_0} \neq 0$, und es gibt ein a mit $\beta(a, .) = 0$, aber $\omega(a) \neq 0$. □

Sei nun $Q = Q(\beta, \omega, \alpha)$ ein Paraboloid und $\beta(a, .) = 0, \omega(A) \neq 0$. Setze $\lambda := -\frac{\alpha}{2\omega(a)}$ und $f := \mathrm{id} + \lambda a$. Dann wird $f^{-1}(Q) = Q(\beta, \omega, 0)$.

Korollar 7. *Jedes Paraboloid ist also affin äquivalent zu einem Paraboloid*

$$Q = Q(\beta, \omega, 0).$$

Dabei ist $\omega \neq 0$, sonst wäre 0 eine Spitze.

Notwendig für die affine Äquivalenz zweier Paraboloide $Q_i = Q(\beta_i, \omega_i, 0)$ ist daher, daß für ein $\lambda \in K\backslash\{0\}$ die symmetrischen Bilinearformen β_1 und $\lambda\beta_2$ äquivalent sind. Das ist aber auch hinreichend:

Satz 67. *Zwei Paraboloide $Q_1 = Q(\beta_1, \omega_1, 0)$ und $Q_2 = Q(\beta_2, \omega_2, 0)$ sind genau dann affin äqivalent, wenn für ein $\lambda \in K\backslash\{0\}$ die Formen $\lambda\beta_2$ und β_1 äquivalent sind.*

Beweis. Die Notwendigkeit war bereits gezeigt. Sei also

$$\lambda\beta_2 = \phi^*\beta_1.$$

Dann hat man folgende affinen Äquivalenzen:

$$Q(\beta_1, \omega_1, 0) \sim Q(\phi^*\beta_1, \phi^*\omega_1, 0) \sim Q(\lambda\beta_2, \phi^*\omega_1, 0).$$

Die Behauptung folgt deshalb aus dem nachstehenden Lemma: □

Lemma 38. *Je zwei Paraboloide*

$$Q_1 = (\beta, \omega_1, 0) \text{ und } Q_2 = (\beta, \omega_2, 0)$$

sind affin äquivalent.

Beweis. Zunächst gibt es ein e mit

$$\beta(e, .) = 0, \quad \omega_1(e) \neq 0 \neq \omega_2(e).$$

Nach Satz 66 wählen wir ein e_1 mit $\beta(e_1, .) = 0, \omega_1(e_1) \neq 0$. Ist $\omega_2(e_1) \neq 0$, so wählen wir $e = e_1$. Andernfalls wählen wir ein e_2 mit $\beta(e_2, .) = 0, \omega_2(e_2) \neq 0$. Ist $\omega_1(e_2) \neq 0$, so wählen wir $e = e_2$. Andernfalls leistet $e = e_1 + e_2$ das Gewünschte.

Wir schreiben nun

$$V = V_0 \oplus V_1$$

mit $V_0 := \operatorname{Kern}\omega_1$ und $V_1 := Ke$ und entsprechend

$$x = x_0 + x_1.$$

Damit definieren wir $\phi \in \operatorname{End}(V)$ durch

$$\phi(x) := x_0 + \frac{\omega_2(x)}{\omega_1(e)}e.$$

Dann ist $\phi \in \operatorname{Aut}(V)$, denn aus $\phi(x) = 0$ folgt x_0 und $\omega_2(x_0 + x_1) = \omega_2(x_1) = 0$. Also ist $x_1 \in Ke \cap \operatorname{Kern}\omega_2$ und daher auch $x_1 = 0$.

Schließlich folgt aus $\beta(a, .) = 0$ und $\omega_1(x_0) = 0$, daß

$$\begin{aligned} \beta(\phi(x), \phi(x)) &= \beta(x_0, x_0) = \beta(x, x) \\ \omega_1(\phi(x)), &= \omega_1(x_0) + \omega_2(x) = \omega_2(x). \end{aligned}$$

Also ist $\phi^*\beta = \beta$ und $\phi^*\omega_1 = \omega_2$. □

6.3 Affine Klassifikation von Quadriken II: Quadratische Formen

Wie wir in den vorangehenden Abschnitten gesehen haben, ist die Klassifikation der Quadriken im wesentlichen reduziert auf die Klassifikation der symmetrischen Bilinearformen. Damit beschäftigen wir uns jetzt.

Vorbemerkungen.

1. Ist β eine symmetrische Bilinearform und definiert man

$$q(x) := \beta(x, x),$$

so erfüllt q offenbar

$$q(\lambda x) = \lambda^2 q(x) \text{ für alle } x \in V,\ \lambda \in K. \tag{62}$$

q bestimmt β eindeutig, weil

$$\beta(x, y) = \frac{1}{2}(q(x + y) - q(x) - q(y)). \tag{63}$$

Zum Beweis berechnet man $\beta(x + y, x + y)$. Eine Abbildung $q : V \to K$ für die (62) gilt und für die (63) eine (symmetrische) Bilinearform definiert, nennt man eine *quadratische Form*. Symmetrische Bilinearformen und quadratische Formen entsprechen sich also eineindeutig. Wir verwenden diese Namen synonym.

2. Seien β eine quadratische Form auf V. Dann wird β bezüglich einer Basis $(e_1, \ldots, e_n)$ dargestellt durch die eindeutig bestimmte (symmetrische) Matrix

$$(\beta(e_i, e_j)).$$

Ist $\phi \in \mathrm{Aut}(V)$, so ist $(\phi(e_1), \ldots, \phi(e_n))$ ebenfalls eine Basis, bezüglich der β dargestellt wird durch die Matrix

$$(\beta(\phi(e_i), \phi(e_j))) = (\phi^* \beta(e_i, e_j)).$$

Zwei quadratische Formen sind also genau dann äquivalent, wenn sie bezüglich geeigneter Basen dieselbe Darstellungsmatrix haben.

3. Sei

$$B = (\beta(e_i, e_j))_{i,j=1,\ldots,n}.$$

die Darstellungsmatrix von β bezüglich der Basis $e_1, \ldots, e_n$ eine Basis von V. Sei $\tilde{e}_1, \ldots, \tilde{e}_n$ eine weitere Basis und $e_j = \sum s_{ij} \tilde{e}_i$. Dann gilt mit $S := (s_{ij})$ und $\tilde{B} = (\beta(\tilde{e}_i, \tilde{e}_j))$ die Basistransformation

$$B = S^* \tilde{B} S.$$

Dabei ist S^* die transponierte Matrix.

Beachten Sie im Gegensatz dazu: Ist $f \in \mathrm{End}(V)$ mit Darstellungsmatrizen A bzw. $\tilde{A}$ bezüglich der obigen Basen, so gilt

$$A = S^{-1} \tilde{A} S.$$

Die Normalformenprobleme für Endomorphismen und quadratische Formen sind also voneinander verschieden.

Satz 68 (Diagonalisierung symmetrischer Bilinearformen). *Sei $\beta : V \times V \to K$ eine quadratische Form. Dann gibt es ein eindeutig bestimmtes $s \in \{0, \dots, n\}$ und eine Basis $(e_1, \dots, e_n)$ mit*

$$\beta(e_i, e_j) \neq 0 \iff i = j \leq s.$$

Ist $x = \sum_{i=1}^{n} x_i e_i$, so folgt also

$$\beta(x, x) = \sum_{i=1}^{s} a_i x_i^2, \quad a_i \neq 0.$$

Man nennt $n - s$ den Nullitätsindex $i_0(\beta)$ *von β.*

Beweis. 1. Schritt. Sei $(e_1, \dots, e_n)$ eine Basis wie im Satz, und sei $\beta^* : V \to V^*, x \mapsto \beta(x, .)$. Dann liegt $x = \sum x_i e_i$ genau dann im Kern von β^*, wenn für alle $i \in \{1, \dots, n\}$

$$0 = \beta(x, e_i) = x_i \beta(e_i, e_i).$$

Nach Wahl der Basis bedeutet das aber gerade $x_1 = \dots = x_s = 0$. Mit anderen Worten ist $\operatorname{Kern} \beta^*$ gerade der Spann von $(e_{s+1}, \dots, e_n)$ und $s = n - \dim \operatorname{Kern} \beta^*$ ist eindeutig bestimmt.

2. Schritt. Wir wählen eine Basis $e_1, \dots, e_n$ von V so, daß $(e_{s+1}, \dots, e_n)$ eine Basis des Kerns von β^* ist. Wir nehmen an, daß $\beta \neq 0$, also $s \geq 1$.

3. Schritt. Wir können ohne Einschränkung annehmen, daß

$$\beta(e_1, e_1) \neq 0.$$

Andernfalls gibt es ein $j > 1$ mit $\beta(e_1, e_j) \neq 0$, und wir ersetzen e_1 durch $\tilde{e}_1 = e_1 + \lambda e_j$, wo

$$\lambda = \begin{cases} 1, & \text{falls } \beta(e_j, e_j) = 0, \\ -\frac{\beta(e_1, e_j)}{\beta(e_j, e_j)} & \text{falls } \beta(e_j, e_j) \neq 0. \end{cases}$$

Dann bleibt die lineare Unabhängigkeit erhalten und

$$\beta(\tilde{e}_1, \tilde{e}_1) = \lambda(2\beta(e_1, e_j) + \lambda \beta(e_j, e_j)) \neq 0.$$

4. Schritt. Wir können ohne Einschränkung annehmen, daß

$$\beta(e_1, e_j) = 0 \text{ für alle } j \geq 2.$$

Wir ersetzen sonst e_j für $2 \leq j \leq s$ durch

$$\tilde{e}_j := e_j - \frac{\beta(e_i, e_j)}{\beta(e_1, e_1)} e_1.$$

Die lineare Unabhängigkeit bleibt erhalten, und es ist

$$\beta(e_1, \tilde{e}_j) = \beta(e_1, e_j) - \beta(e_1, e_j) = 0.$$

5. Schritt. Sei nun $(e_1, \dots, e_n)$ eine Basis und $k \in \{1, \dots, s\}$ mit

$$\beta(e_i, e_j) = 0 \text{ für alle } 1 \leq i \leq k, j \neq i,$$

und

$$\beta(e_i, e_i) \neq 0 \text{ für alle } 1 \leq i \leq k.$$

Falls $k = s$ sind wir fertig. Andernfalls könne wir o.E. annehmen, daß $\beta(e_{k+1}, e_{k+1}) \neq 0$, vgl. das Argument aus dem 3. Schritt. Weiter können wir o.E. annehmen, daß

$$\beta(e_{k+1}, e_j) = 0 \text{ für alle } j > k+1,$$

vgl. das Argument aus dem 4. Schritt. Fortsetzung dieses Verfahrens liefert die gewünschte Basis. □

Bemerkung. Hat man eine Basis $(e_1, \ldots, e_n)$, die β diagonalisiert

$$\beta(x, x) = \sum_{i=1}^{s} a_i x_i^2 \quad \text{für } x = \sum x_i e_i,$$

und ersetzt man e_i durch $\tilde{e}_i := \lambda_i e_i$ mit $\lambda_i \neq 0$, so ist

$$x = \sum (x_i/\lambda_i)\tilde{e}_i = \sum \tilde{x}_i \tilde{e}_i$$

und

$$\beta(x, x) = \sum_{i=1}^{s} a_i \lambda_i^2 \tilde{x}_i^2. = \sum_{i=1}^{s} \tilde{a}_i \tilde{x}_i^2.$$

Zwei diagonal dargestellte quadratische Formen

$$\beta(x, x) = \sum_{i=1}^{s} a_i x_i^2, \quad \tilde{\beta}(x, x) = \sum_{i=1}^{s} \tilde{a}_i x_i^2$$

mit $a_i \neq 0 \neq \tilde{a}_i$ sind offenbar äquivalent, wenn die $\tilde{a}_i/a_i$ im zugrundliegenden Körper K Quadrate sind.

Wir behandeln die Fälle $K = \mathbb{R}$ und $K = \mathbb{C}$.

6.3.1 Reelle quadratische Formen und Quadriken

Falls $K = \mathbb{R}$, kann man auf V ein Euklidisches Skalarprodukt $\langle ., . \rangle$ wählen. Dann gibt es ein selbstadjungiertes $f \in \mathrm{End}(V)$ mit

$$\beta(x, y) = \langle f(x), y \rangle.$$

Wir wissen, daß V dann eine ON-Basis aus Eigenvektoren von f besitzt, und diese diagonalisiert offenbar die Bilinearfom β. Der Beweis des letzten Satzes 68 ist im reellen Fall also ganz einfach.

Die Eigenwerte von f sind durch β keineswegs bestimmt, weil das Skalarprodukt beliebig gewählt werden kann. Allerdings ist die Multiplizität des Eigenwertes 0 gerade der Nullitätsindex $i_0(\beta)$. Und die Anzahl der positiven bzw. der negativen Eigenwerte ist ebenfalls durch β bestimmt. Im folgenden beweisen wir das ohne Bezug auf ein Skalarprodukt direkt für die quadratische Form.

Ist $(e_1, \ldots, e_n)$ eine β diagonalisierende Basis von V mit $\beta(e_i, e_i) = a_i$, so ist

$$V = V_+ \oplus V_- \oplus V_0,$$

wobei

$$\begin{aligned} V_+ &:= \mathrm{Spann}\{e_i \,|\, a_i > 0\}, \\ V_- &:= \mathrm{Spann}\{e_i \,|\, a_i < 0\}, \\ V_0 &:= \mathrm{Spann}\{e_i \,|\, a_i = 0\}. \end{aligned}$$

Dann ist aber $\beta|_{V_+ \times V_+}$ positiv definit und V_+ ist ein maximal-dimensionaler Unterraum von V mit dieser Eigenschaft. Gäbe es nämlich einen größer-dimensionalen V'_+, so wäre nach dem Dimensionssatz $V'_+ \cap (V_- \oplus V_0) \neq \{0\}$. Widerspruch! Daher ist zwar nicht V_+, wohl aber die Dimension von V_+ unabhängig von der diagonalisierenden Basis. Entsprechendes gilt für V_-. Dagegen ist der Raum $V_0 = \mathrm{Kern}(x \mapsto \beta(x, .))$ und nicht nur seine Dimension eindeutig bestimmt.

Die Dimension von V_+ nennt man auch den *Positivitätsindex* $i_+(\beta)$ von β. Entsprechend heißt $i_-(\beta) := \dim V_-$ der *Negativitätsindex*. Die Differenz $i_+(\beta) - i_-(\beta)$ bezeichnet man auch als den *Trägheitsindex* von β.

Es gilt

$$i_+(\beta) + i_-(\beta) + i_0(\beta) = n.$$

Lemma 39. *Die Indizes einer reellen quadratischen Form sind invariant unter der Operation von* $\mathrm{Aut}(V)$*: Es gilt für alle* $\phi \in \mathrm{Aut}(V)$

$$i_j(\phi^*(\beta)) = i_j(\beta), \quad j \in \{+, -, 0\}.$$

Beweis. Leicht. □

Satz 69 (Klassifikation reeller quadratischer Formen, Trägheitssatz von Sylvester). *Sei β eine quadratische Form auf einem n-dimensionalen reellen Vektorraum V. Dann gibt es eindeutig bestimmte Zahlen $p = i_+(\beta)$ und $q = i_-(\beta)$ und eine Basis $(e_1, \ldots, e_n)$ von V, so daß für allex $= \sum_{i=1}^n x_i e_i$ gilt*

$$\beta(x,x) = x_1^2 + \ldots + x_p^2 - x_{p+1}^2 - \ldots - x_{p+q}^2.$$

Zwei reelle quadratische Formen β_1, β_2 sind genau dann äquivalent, wenn sie dieselben Indizes haben, d.h. wenn

$$i_+(\beta_1) = i_+(\beta_2) \text{ und } i_-(\beta_1) = i_-(\beta_2).$$

Beweis. Betrachte eine diagonalisierende Basis $(e_1, \ldots, e_n)$ mit

$$\beta(e_i, e_i) > 0 \text{ für } i \le p, \quad \beta(e_i, e_i) < 0 \text{ für } p+1 \le i \le p+q,$$

wobei $p = i_+(\beta), q := i_-(\beta)$. Ersetze e_i für $i \le s = p+q$ durch $\frac{1}{\sqrt{|\beta(e_i,e_i)|}} e_i$. Das liefert die Normalform. Die Gleichheit der Indizes ist nach dem vorangehenden Lemma notwendig für die Äquivalenz der Formen. Sie ist auch hinreichend, weil zwei Formen mit gleicher Darstellungsmatrix äquivalent sind. □

Satz 70 (Affine Klassifikation reeller Quadriken). *Zwei Quadriken*

$$Q_1 = Q(\beta_1, \omega_1, \alpha_1) \text{ und } Q_2 = Q(\beta_2, \omega_2, \alpha_2)$$

in einem n-dimensionalen reellen Vektorraum V sind genau dann affin äquivalent, wenn sie zum selben Typ gemäß Definition 40 gehören und überdies gilt:

(i) Q_1 und Q_2 sind Kegel bzw. Paraboloide und

$$\{i_+(\beta_1), i_-(\beta_1)\} = \{i_+(\beta_2), i_-(\beta_2)\},$$

oder

(ii) Q_1 und Q_2 sind echte Mittelpunktsquadriken und bei Darstellung als

$$Q_1 = Q(\beta_1, 0, \alpha_1), \quad Q_2 = Q(\beta_2, 0, \alpha_2)$$

mit $\alpha_1\alpha_2 > 0$ haben β_1 und β_2 denselben Positivitäts- und Negativitätsindex:

$$i_+(\beta_1) = i_+(\beta_2), \quad i_-(\beta_1) = i_-(\beta_2).$$

Die entsprechenden Normalformen sind (mit $i_+ =: p$, $i_+ + i_- = s$):

Echte Mittelpunktsquadriken: $\sum_{i=1}^p x_i^2 - \sum_{i=p+1}^s x_i^2 = 1, \quad 0 \le p \le s \le n,$

Kegel: $\sum_{i=1}^p x_i^2 - \sum_{i=p+1}^s x_i^2 = 0, \quad 0 \le p \le s \le n,$

Paraboloide: $\sum_{i=1}^p x_i^2 - \sum_{i=p+1}^s x_i^2 = 2x_n, \quad 0 \le p \le s < n.$

Nach geeigneter Translation und bezüglich einer geeignet gewählten Basis wird jede Quadrik Q durch eine solche Gleichung gegeben.
(Aber nicht jede solche Gleichung definiert eine Quadrik, weil z.B. $\sum x_i^2 = -1$ die leere Menge oder $\sum_1^n x_i^2 = 0$ eine einpunktige Menge definieren.)

Beweis. Zu (i). Jedes Paraboloid besitzt nach Translation (bei der sich das β nicht ändert) eine Darstellung der Form $Q = Q(\beta, \omega, 0)$. Je zwei Paraboloide in dieser, also in beliebiger Form sind affin äquivalent, wenn die quadratischen Formen β_1 und $\lambda\beta_2$ für ein geeignetes $\lambda \in \mathbb{R}\backslash\{0\}$ äquivalent sind. Multipliziert man β mit einem positiven Faktor, so ändern sich $i_+(\beta)$ und $i_-(\beta)$ offenbar nicht, multipliziert man mit einem negativen Faktor, so vertauschen sich diese beiden Zahlen. Daraus folgt die Behauptung über die affine Äquivalenz.

Für die Normalform sei $Q = Q(\beta, \omega, 0)$ und $a \in V$ mit $\beta(a, .) = 0$ und $\omega(a) \neq 0$, vgl. Satz 66. Betrachte β eingeschränkt auf den Kern von ω. Nach dem Satz von Sylvester gibt es dann $0 \leq p, q \leq p+q \leq n-1$ und eine Basis $(e_1, \ldots, e_{n-1})$ von $\operatorname{Kern}\omega$, so daß für $x = \sum_{i=1}^{n-1} x_i e_i$ gilt

$$\beta(x, x) = x_1^2 + \ldots + x_p^2 - x_{p+1}^2 - \ldots - x_{p+q}^2$$

Ergänze die obige Basis durch $e_n := -\frac{1}{\omega(a)}\, a$ zu einer Basis von V. Dann folgt

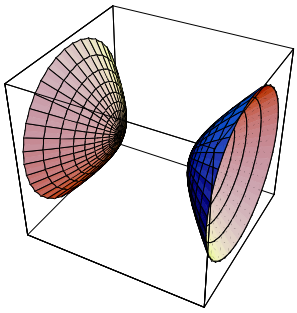
$$\begin{aligned}\beta(x, x) + 2\omega(x) &= x_1^2 + \ldots + x_p^2 - x_{p+1}^2 - \ldots - x_{p+q}^2 + 2\omega(x_n e_n) \\ &= x_1^2 + \ldots + x_p^2 - x_{p+1}^2 - \ldots - x_{p+q}^2 - 2x_n.\end{aligned}$$

Damit ist (i) für Paraboloide bewiesen. Für Kegel sind die Argumente ähnlich, aber einfacher.

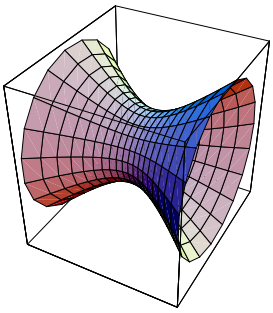
Zu (ii). Hier argumentiert man wie für (i). Nur ist der Faktor λ nicht frei, sondern in der Form $Q_i = Q_i(\beta_i, 0, \alpha_i)$ durch α_1/α_2 bestimmt. □

Daß es bei echten Mittelpunktsquadriken nicht genügt, daß die beiden Index*mengen* übereinstimmen, zeigt das folgende Beispiel.

Beispiel 41. Die Quadriken



und



$x^2 - y^2 - z^2 = 1$ $\qquad$ $x^2 - y^2 - z^2 = -1$

haben dasselbe β und damit dieselben Indizes, sind aber offenbar nicht affin äquivalent. Man erhält ein sogenanntes *zweischaliges* bzw. *einschaliges* Hyperbolid. Bei richtiger Normierung der Gleichungen, nämlich

$$-x^2 + y^2 + z^2 + 1 = 0, \quad x^2 - y^2 - z^2 + 1 = 0,$$

sind dann auch die Indizes nicht mehr gleich.

□

Beispiel 42. Im dreidimensionalen reellen affinen Raum findet man die folgenden affinen Äquivalenzklassen von Quadriken:

Echte Mittelpunktsquadriken	
$x^2 + y^2 + z^2 = 1$	Ellipsoid
$x^2 + y^2 - z^2 = 1$	einschaliges Hyperboloid
$x^2 - y^2 - z^2 = 1$	zweischaliges Hyperboloid
$x^2 + y^2 = 1$	elliptischer Zylinder
$x^2 - y^2 = 1$	hyperbolischer Zylinder
$x^2 = 1$	Paar paralleler Ebenen
Kegel	
$x^2 + y^2 - z^2 = 0$	Doppelkegel
$x^2 - y^2 = 0$	Paar sich schneidender Ebenen
Paraboloide	
$x^2 + y^2 = 2z$	elliptisches Paraboloid
$x^2 - y^2 = 2z$	hyperbolisches Paraboloid
$x^2 = 2z$	parabolischer Zylinder

□

6.3.2 Komplexe quadratische Formen und Quadriken

Satz 71 (Klassifikation der komplexen quadratischen Formen). *Zwei quadratische Formen auf dem n-dimensionalen komplexen Vektorraum V sind genau dann ähnlich, wenn sie denselben Nullitätsindex haben.*
Bezüglich einer geeigneten Basis hat eine solche Form die Darstellung

$$\beta(x,x) = \sum_{i=1}^{s} x_i^2.$$

Beweis. Sei β eine Form mit Nullitätsindex s. Nach Satz 68 gibt es eine Basis $(b_1, \ldots, b_n)$ mit

$$\beta(b_j, b_k) = a_j \delta_{jk}$$

und

$$a_j = 0 \iff j > s.$$

Für $1 \leq j \leq s$ sei $c_j \in \mathbb{C}\backslash\{0\}$ mit

$$c_j^2 = a_j$$

und

$$e_j := b_j / c_j.$$

Setze weiter $e_j = b_j$ für $j > s$. Dann folgt

$$\beta(e_j, e_k) = \delta_{jk}.$$

Dann ist aber

$$\beta(x,x) = \sum_{j=1}^{s} x_j^2$$

für $x \sum_{j=1}^{n} x_j e_j$. □

Wie im letzen Abschnitt folgt daraus

Satz 72 (Klassifikation der komplexen Quadriken). *Zwei Quadriken*

$$Q_1 = Q(\beta_1, \omega_1, \alpha_1) \text{ und } Q_2 = Q(\beta_2, \omega_2, \alpha_2)$$

in einem n-dimensionalen komplexen Vektorraum V sind genau dann affin äquivalent, wenn sie zum selben Typ gemäß Definition 40 gehören und β_1 und β_2 denselben Nullitätsindex haben.
Die Normalformen sind

Echte Mittelpunktsquadriken:	$\sum_{i=1}^{s} x_i^2 = 1,$	$1 \leq s \leq n,$
Kegel:	$\sum_{i=1}^{s} x_i = 0,$	$1 \leq s \leq n,$
Paraboloide:	$\sum_{i=1}^{s} x_i^2 = 2x_n,$	$1 \leq s < n.$

Nach geeigneter Translation und bezüglich einer geeignet gewählten Basis wird jede Quadrik Q durch eine solche Gleichung gegeben.

Beispiel 43. Im 3-dimensionalen komplexen Raum hat man folgende affine Äquivalenzklassen von Quadriken:

Echte Mittelpunktsquadriken
$x^2 + y^2 + z^2 = 1$
$x^2 + y^2 = 1$
$x^2 = 1$
Kegel
$x^2 + y^2 + z^2 = 0$
$x^2 + y^2 = 0$
Paraboloide
$x^2 + y^2 = 2z$
$x^2 = 2z$

□

6.4 Klassifikation im Euklidischen

Das Problem der Euklidischen Klassifikation von Quadriken ist differenzierter als der affine Fall. Zwei Quadriken in einem Euklidischen affinen Raum heißen *kongruent*, wenn es eine längentreue affine Transformation (orthogonale Abbildung und Translation) gibt, die die eine in die andere überführt. Kreis und Ellipse sind affin äquivalent, nicht aber kongruent.

Wie oben sieht man: Zwei Quadriken sind genau dann kongruent, wenn sie bezüglich geeigneter Euklidischer Koordinaten (bezogen auf Orthonormalbasen) dieselbe Darstellung haben.

Echte Mittelpunktsquadriken. Wir betrachten eine echte Mittelpunktsquadrik. Nach geeigneter Translation (längentreu!) hat sie die Darstellung

$$\beta(x,x) + \alpha = 0$$

mit einer symmetrischen Bilinearform β. Nach Normierung dieser Gleichung können wir

$$\alpha = -1$$

erhalten. Nun gibt es einen selbstadjungierten Endomorphismus $f \in \mathrm{End}(V)$ mit

$$\beta(x,y) = \langle f(x), y\rangle \quad \text{für alle } x, y \in V.$$

Für die Euklidische Klassifikation müssen wir also folgende Frage beantworten:

Wann gibt es zu selbstadjungierten $f, \tilde{f} \in \mathrm{End}(V)$ ein $g \in \mathrm{O}(V)$ mit

$$\{x \,|\, \langle f(x), x\rangle = 1\} = g\left(\{x \,|\, \langle \tilde{f}(x), x\rangle = 1\}\right)?$$

Die letzte Gleichung bedeutet

$$\langle \tilde{f}(x), x\rangle = 1 \iff \langle g^* f g(x), x\rangle = 1.$$

Lemma 40. *Sei V ein Euklidischer Vektorraum. Die selbstadjungierten $f_1, f_2 \in \mathrm{End}(V)$ seien nicht negativ semidefinit. Dann gilt*

$$\forall_{x\in V}\, (\langle f_1(x), x\rangle = 1 \iff \langle f_2(x), x\rangle = 1)$$

genau dann, wenn $f_1 = f_2$.

Der folgende Beweis benutzt ein wenig Analysis.

Beweis. Ist $x \in V$ mit $\langle f_1(x), x\rangle = \lambda > 0$, so folgt

$$\langle f_1(x/\sqrt{\lambda})\,, x/\sqrt{\lambda}\,\rangle = 1 = \langle f_2(x/\sqrt{\lambda})\,, x/\sqrt{\lambda}\,\rangle.$$

Also gilt

$$\langle f_1(x), x\rangle = \langle f_2(x), x\rangle \quad \text{für alle } x \text{ mit } \langle f_1(x), x\rangle > 0.$$

oder

$$\langle (f_1 - f_2)(x), x\rangle = 0 \quad \text{für alle } x \text{ mit } \langle f_1(x), x\rangle > 0.$$

Nach Voraussetzung können wir ein $x \in V$ wählen, für das $\langle f_1(x), x\rangle > 0$. Sei $v \in V$ beliebig. Weil

$$\mathbb{R} \to \mathbb{R}, t \mapsto \langle f_1(x+tv), x+tv\rangle = \langle f_1(x), x\rangle + 2t\langle f_1(x), v\rangle + t^2\langle f_1(v), v\rangle$$

stetig ist, gibt es ein $\delta > 0$, so daß

$$\langle f_1(x+tv), x+tv\rangle > 0 \quad \text{für alle } t \text{ mit } |t| < \delta.$$

Also gilt

$$\langle (f_1-f_2)(x), x\rangle + 2t\langle (f_1-f_2)(x), v\rangle + t^2\langle (f_1-f_2)(v), v\rangle = 0$$

für $|t| < \delta$. Das quadratisches Polynom ist deshalb Null und man erhält für alle $v \in V$

$$\langle (f_1-f_2)(v), v\rangle = 0.$$

Daraus folgt aber $f_1 = f_2$. □

Korollar 8. *Jede echte Mittelpunktsquadrik in einem Euklidischen Vektorraum hat nach Translation eine Darstellung*

$$Q = Q(\beta, 0, -1).$$

Die Form β ist eindeutig bestimmt und definiert ein eindeutig bestimmtes selbstadjungiertes $f \in \mathrm{End}(V)$ mit $\beta(x,x) = \langle f(x), x\rangle$ für alle x. Zwei echte Mittelpunktsquadriken sind genau dann kongruent, wenn die zugehörigen Endomorphismen orthogonal ähnlich sind, d.h. wenn sie dasselbe charakteristische Polynom (oder dieselben Eigenwerte mit denselben Multiplizitäten) haben.
Die Normalform bezüglich einer geeigneten ON-Basis ist

$$\sum_{i=1}^{n} \lambda_i x_i^2 = 1, \quad \lambda_i \in \mathbb{R}$$

mit bis auf die Reihenfolge eindeutig bestimmten λ_i.

Beispiel 44. Im Euklidischen 3-dimensionalen Raum treten zum Beispiel an die Stelle der affinen Klasse

$$x^2 + y^2 + z^2 = 1$$

unendlich viele Klassen inäquivalenter Ellipsoide

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1, \quad 0 < a \le b \le c.$$

□

Ohne Beweis geben wir noch kurz die Ergebnisse der Kongruenz-Klassifikation für die beiden anderen Typen:

Kegel. Diese haben in einem geeigneten Euklidischen Koordinatensystem die Gleichung

$$\sum_{i=1}^{n} \lambda_i x_i^2 = 0.$$

Die Kongruenzklasse legt $(\lambda_1, \ldots, \lambda_n)$ eindeutig bis auf Reihenfolge und Skalierung mit einem Faktor $\lambda \neq 0$ fest.

Paraboloide. Diese haben in einem Euklidischen Koordinatensystem die Gleichung

$$\sum_{i=1}^{n-1} \lambda_i x_i^2 = 2x_n.$$

Die Kongruenzklasse bestimmt $(\lambda_1, \ldots, \lambda_{n-1})$ eindeutig bis auf Numerierung.

7 Der Fundamentalsatz über Kollineationen

Ziel dieses Abschnittes ist der Beweis eines Satzes über geradentreue Abbildungen eines Vektorraumes in sich. Verallgemeinerungen dieses Satzes firmieren unter den Namen *Fundamentalsatz der affinen bzw. projektiven Geometrie*. Diese Verallgemeinerungen sind allerdings nicht sehr erheblich, der wesentliche Sachverhalt wird bereits in unserer Version deutlich.

In diesem Abschnitt sei V ein Vektorraum über einem Körper K der Charakteristik $\neq 2$.

Definition 42. Eine Abbildung $f : V \to V$ heißt eine *Kollineation*, wenn sie bijektiv ist und Geraden auf Geraden abbildet. Anders gesagt: Kollineare Punkte gehen in kollineare Punkte, daher der Name.

Beispiel 45. Affine Transformationen $f(x) = \phi(x) + a$ mit $\phi \in \mathrm{Aut}(V)$ und $a \in V$ sind Kollineationen.

□

Beispiel 46. Für $V = \mathbb{C}^2$ als Vektorraum über $\mathbb{C}$, ist die Abbildung $f : V \to V$ mit

$$f(z_1, z_2) := (\bar{z_1}, \bar{z_2})$$

eine Kollineation, denn

$$f((a_1, a_2) + \lambda(z_1, z_2)) = (\bar{a_1} + \bar{\lambda}\bar{z_1}, \bar{a_2} + \bar{\lambda}\bar{z_2}) = (\bar{a_1}, \bar{a_2}) + \bar{\lambda}(\bar{z_1}, \bar{z_2}).$$

Also geht die Gerade $\{(a_1, a_2) + \lambda(z_1, z_2) \mid \lambda \in \mathbb{C}\}$ in die Gerade $\{(\bar{a_1}, \bar{a_2}) + \mu(\bar{z_1}, \bar{z_2}) \mid \mu \in \mathbb{C}\}$ über.

Beachte, daß f nicht $\mathbb{C}$-linear ist. Es gilt zwar $f(z + w) = f(z) + f(w)$, aber

$$f(\lambda z) = \bar{\lambda} f(z).$$

□

Definition 43. Ein Automorphismus des Körpers K ist eine Bijektion $\sigma : K \to K$ mit

$$\sigma(\lambda + \mu) = \sigma(\lambda) + \sigma(\mu),$$
$$\sigma(\lambda\mu) = \sigma(\lambda)\sigma(\mu),$$

für alle $\lambda, \mu \in K$.

Es ist leicht zu sehen, daß für einen solchen Automorphismus stets

$$\sigma(0) = 0 \text{ und } \sigma(1) = 1.$$

Beispiel 47. Natürlich ist $\mathrm{id} : K \to K$ ein Automorphismus, der sogenannte triviale Automorphismus von K.

Die Konjugation $\mathbb{C} \to \mathbb{C}, z \mapsto \bar{z}$ ist ein Automorphismus.

□

Satz 73 (Fundamentalsatz über Kollineationen).
Sei V ein K-Vektorraum der Dimension ≥ 2 und $f : V \to V$ eine Kollineation mit $f(0) = 0$. Dann gibt es einen Automorphismus $\sigma : K \to K$, so daß für alle $x, y \in V$ und $\lambda \in K$

$$f(x + y) = f(x) + f(y),$$
$$f(\lambda x) = \sigma(\lambda) f(x).$$

Bemerkungen.

1. Ist $\sigma = \mathrm{id}$, so ist also f eine *lineare* Abbildung. Wir werden noch zeigen, daß zum Beispiel $K = \mathbb{R}$ nur den trivialen Automorphismus besitzt.
2. Die Bedingung $f(0) = 0$ ist nicht sehr einschneidend: Ist $g : V \to V$ eine beliebige Kollineation, so definiert
$$f(x) := g(x) - g(0)$$
eine Kollineation mit $f(0) = 0$. Für diese gilt dann der Satz.
3. Aus der Definition der Kollineation folgt unmittelbar für alle x
$$f(Kx) = Kf(x). \tag{64}$$
4. Parallel Geraden sind dadurch charakterisiert, daß sie sich nicht schneiden. Weil f injektiv ist, schneiden sich dann auch die Bilder nicht. Kollineationen bilden also parallele Gerade auf parallele Geraden ab.
5. Weil wir es im folgenden oft mit linearer Unabhängigkeit zu tun haben, vereinbaren wir die Notation
$$x \wedge y \neq 0 \iff x \text{ und } y \text{ linear unabhängig.}$$

Dem Beweis des Fundamentalsatzes stellen wir noch ein Lemma voran:

Lemma 41. *Unter den Voraussetzungen des Fundamentalsatzes gilt:*
$$x \wedge y \neq 0 \implies f(x) \wedge f(y) \neq 0.$$

Beweis des Lemmas. Sind $f(x)$ und $f(y)$ linear abhängig, so ist einer der Vektoren ein Vielfaches des anderen, d.h sie liegen auf einer Geraden durch $0 = f(0)$. Weil $x \neq 0$ ist $f(Kx) = Kf(x)$ *die* Gerade durch $f(0)$ und $f(x)$, auf der dann also auch $f(y)$ liegt. Das bedeutet wegen der Injektivität von f aber, daß $y \in Kx$ im Widerspruch zur Voraussetzung. □

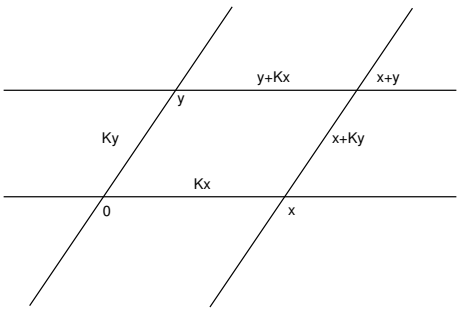
Beweis des Fundamentalsatzes. Wir zeigen nun zunächst
$$f(x + y) = f(x) + f(y). \tag{65}$$

0. Fall: $x = 0$ oder $y = 0$. Dann ist die Behauptung trivial.

1. Fall: $x \wedge y \neq 0$. Dann sind Kx und $y + Kx$ voneinander verschiedene parallele Geraden, die durch f auf ebensolche abgebildet werden:

$$f(y + Kx) = f(y) + f(Kx) = f(y) + Kf(x).$$

Analoges gilt, wenn man x und y vertauscht. Nun ist $x+y$ gerade der Schnittpunkt der Geraden $y+Kx$ und $x+Ky$ und $f(x)+f(y)$ der Schnittpunkt der Geraden $f(y) + Kf(x)$ und $f(x) + Kf(y)$.



Daher haben wir

$$\begin{aligned}\{f(x + y)\} &= f((y + Kx) \cap (x + Ky)) = f(y + Kx) \cap f(x + Ky) \\ &= (f(y) + Kf(x)) \cap (f(x) + Kf(y)) = \{f(x) + f(y)\}.\end{aligned}$$

2. Fall: $x = \mu y \neq 0$ und $x + y \neq 0$. Nach der Dimensionsvoraussetzung gibt es ein $z \in V$ mit $x \wedge z \neq 0$. Dafür folgt

$$y \wedge z \neq 0, \quad (x + y) \wedge z \neq 0, \quad x \wedge (y + z) \neq 0.$$

Aus dem 1. Fall ergibt sich deshalb

$$\begin{array}{c} f(x + y + z) = f((x + y) + z) = f(x + y) + f(z) \\ \| \\ f(x + (y + z)) = f(x) + f(y + z) = f(x) + f(y) + f(z). \end{array}$$

Auch in diesem Fall gilt also (65).

3. Fall: $x + y = 0, x \neq 0$. Wie oben wählen wir ein z mit $x \wedge z \neq 0$. Dann ist auch $y \wedge z \neq 0$ und

$$f(z) = f((x + y) + z) = f(x + (y + z)) = f(x) + f(y + z) = f(x) + f(y) + f(z).$$

Es folgt

$$f(x + y) = f(0) = 0 = f(x) + f(y).$$

Damit ist (65) in allen Fällen bewiesen.

Wir beweisen nun

$$f(\lambda x) = \sigma(\lambda) f(x) \tag{66}$$

für einen Körperautomorphismus σ.

Für $x \neq 0$ ist auch $f(x) \neq 0$, und daher gibt es nach (64) eine eindeutig bestimmte Funktion $\sigma(., x) : K \to K$ mit

$$f(\lambda x) = \sigma(\lambda, x) f(x).$$

Wir zeigen zunächst, daß diese unabhängig von x ist, daß also für alle $x \neq 0 \neq y$ und alle λ

$$\sigma(\lambda, x) = \sigma(\lambda, y). \tag{67}$$

1. Fall: $x \wedge y \neq 0$. Dann haben wir

$$\begin{array}{c} f(\lambda(x + y)) = f(\lambda x) + f(\lambda y) = \sigma(\lambda, x) f(x) + \sigma(\lambda, y) f(y) \\ \| \\ \sigma(\lambda, x + y) f(x + y) = \sigma(\lambda, x + y) f(x) + \sigma(\lambda, x + y) f(y) \end{array}$$

Weil $f(x)$ und $f(y)$ linear unabhängig sind, folgt

$$\sigma(\lambda, x) = \sigma(\lambda, x + y) = \sigma(\lambda, y).$$

2. Fall: $x = \mu y \neq 0$. Wähle ein z mit $x \wedge z \neq 0$. Dann folgt $y \wedge z \neq 0$ und aus dem 1. Fall

$$\sigma(\lambda, x) = \sigma(\lambda, z) = \sigma(\lambda, y).$$

Damit ist (67) bewiesen, und wir definieren

$$\sigma(\lambda) := \sigma(\lambda, x)$$

für ein $x \in V$. Dann ist für alle x (offenbar auch für $x = 0$) und $\lambda \in K$

$$f(\lambda x) = \sigma(\lambda) f(x).$$

Schließlich zeigen wir, daß σ ein Körperautomorphismus ist.

Es gilt für $x \neq 0$

$$\begin{aligned} f((\lambda+\mu)x) &= \sigma(\lambda+\mu)f(x) \\ \| & \\ f(\lambda x + \mu x) &= \sigma(\lambda)f(x) + \sigma(\mu)f(x), \end{aligned}$$

also

$$\sigma(\lambda+\mu) = \sigma(\lambda) + \sigma(\mu).$$

Schließlich finden wir

$$\sigma(\lambda\mu)f(x) = f((\lambda\mu)x)) = f(\lambda(\mu x)) = \sigma(\lambda)f(\mu x) = \sigma(\lambda)\sigma(\mu)f(x).$$

□

Zur Abrundung dieses Resultates zeigen wir noch:

Satz 74. *Die Körper $\mathbb{Q}$ und $\mathbb{R}$ haben nur den trivialen Körperautomorphismus $\sigma = \mathrm{id}$.*

Beweis. Sei σ ein Automorphismus von K. Aus $\sigma(1) = 1$ und der Additivität folgt für alle $n \in \mathbb{Z}$

$$\sigma(n) = n\sigma(1) = n.$$

Weiter folgt aus der Multiplikativität, daß

$$\sigma(\frac{\lambda}{\mu}) = \frac{\sigma(\lambda)}{\sigma(\mu)}$$

für alle $\lambda, \mu \neq 0$ in K. Insbesondere folgt für alle $p, q \neq 0$ in $\mathbb{Z}$, daß

$$\sigma(\frac{p}{q}) = \frac{p}{q}.$$

Damit ist die Behauptung für $\mathbb{Q}$ bewiesen.

Im Falle $K = \mathbb{R}$ wissen wir immerhin auch, daß $\sigma(\lambda) = \lambda$, falls $\lambda \in \mathbb{Q}$. Da jedes $\lambda \in \mathbb{R}$ Grenzwert einer Folge rationaler Zahlen ist, genügt es also zu zeigen, daß σ stetig ist.

Für $\lambda \leq \mu$ in $\mathbb{R}$ gibt es ein $\xi \in \mathbb{R}$ mit $\mu - \lambda = \xi^2$. Dann ist aber

$$\sigma(\mu) - \sigma(\lambda) = \sigma(\mu - \lambda) = \sigma(\xi)^2 \geq 0$$

Also ist σ monoton:

$$\lambda \leq \mu \implies \sigma(\lambda) \leq \sigma(\mu).$$

Seien nun $\lambda_1, \lambda_2 \in \mathbb{R}$ mit

$$\lambda_1 \leq \lambda_2.$$

Dann gilt für $\tilde{\lambda}_1, \tilde{\lambda}_2 \in \mathbb{Q}$ mit

$$\tilde{\lambda}_1 \leq \lambda_1 \leq \lambda_2 \leq \tilde{\lambda}_2,$$

daß

$$\tilde{\lambda}_1 = \sigma(\tilde{\lambda}_1) \leq \sigma(\lambda_1) \leq \sigma(\lambda_2) \leq \sigma(\tilde{\lambda}_2) = \tilde{\lambda}_2.$$

Also ist

$$0 \leq \sigma(\lambda_2) - \sigma(\lambda_1) \leq \sigma(\tilde{\lambda}_2) - \sigma(\tilde{\lambda}_1) = \tilde{\lambda_2} - \tilde{\lambda_1}.$$

Weil es Folgen rationaler Zahlen gibt, die von unten gegen λ_1 bzw. von oben gegen λ_2 konvergieren, folgt durch Grenzübergang

$$0 \leq \sigma(\lambda_2) - \sigma(\lambda_1) \leq \lambda_2 - \lambda_1.$$

Daraus ergibt sich

$$|\sigma(\lambda_2) - \sigma(\lambda_1)| \leq |\lambda_2 - \lambda_1|$$

für alle $\lambda_1, \lambda_2 \in \mathbb{R}$ und damit die Stetigkeit von σ. □

Der folgende Satz ist nicht schwer zu zeigen:

Satz 75. *Die Konjugation $\sigma : \mathbb{C} \to \mathbb{C}, z \mapsto \bar{z}$ ist der einzige nicht-triviale Automorphismus von $\mathbb{C}$ auf sich, für den $\sigma(\mathbb{R}) \subset \mathbb{R}$. (*Nicht-trivial *heißt* $\neq$ id.*)*

Aber die naheliegende Vermutung, daß die Konjugation überhaupt der einzige Körperautomorphismus von $\mathbb{C}$ auf sich sei, ist falsch. Gegenbeispiele kann man mit Hilfe des Auswahlaxioms konstruieren, aber explizit hingeschrieben hat noch niemand eines.

Schlußbemerkung. Der „Fundamentalsatz der affinen Geometrie“ folgt aus dem vorstehenden Satz über Kollineationen ganz unmittelbar nach Wahl eines Ursprungs. Den entsprechenden Satz der projektiven Geometrie beweist man, indem man zunächst zeigt, daß eine Kollineation k-dimensionale Unterräume auf ebensolche abbildet. Dann bildet sie insbesondere projektive Hyperebenen auf projektive Hyperebenen ab. Nach Entfernung einer projektiven Hyperebene ist man aber wieder in der affinen Situation und kann den obigen Satz anwenden.

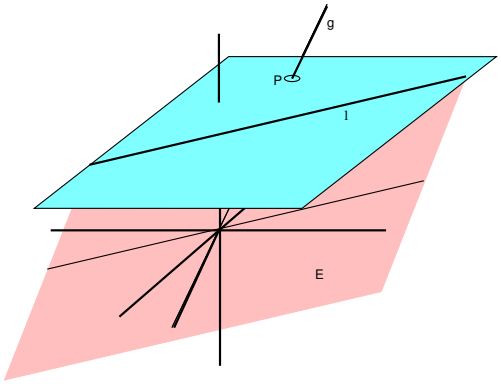
8 Projektive Geometrie: Ein Ausblick

In der ebenen affinen Geometrie spielt der Schnittpunkt von Geraden eine wichtige Rolle, und der Ausnahmefall paralleler Geraden führt immer wieder zu umständlichen Fallunterscheidungen. Dem kann man abhelfen, indem man zum affinen Raum „unendlich ferne Punkte“ hinzufügt. Das klingt mysteriös, läßt sich aber auf die folgende Weise sehr anschaulich realisieren:

Wir betrachten die affine Ebene

$$\mathcal{A}: \quad z = 1$$

im $\mathbb{R}^3$. Dann gibt es eine offensichtliche Bijektion zwischen den Punkten $P \in \mathcal{A}$ und den Geraden $g \subset \mathbb{R}^3$ durch den Nullpunkt 0, *die nicht in der (x,y)-Ebene liegen.*
Ebenso entsprechen sich affine Geraden l in $\mathcal{A}$ und Ebenen $E \subset \mathbb{R}^3$ durch 0, wobei man wieder die (x,y)-Ebene ausnehmen muß.
Zwei voneinander verschiedene Geraden in $\mathcal{A}$ schneiden sich in genau einem Punkt, oder sie sind parallel. Im $\mathbb{R}^3$ interpretiert trifft die Schnittgerade der entsprechenden Ebenen die Eben $\mathcal{A}$, oder sie liegt in der (x,y)-Ebene.



Wir definieren nun $\mathcal{P}$ als Menge der Geraden durch 0 (=*Nullpunktsgeraden*) des $\mathbb{R}^3$ und identifizieren die Punkte von $\mathcal{A}$ mit der Teilmenge der Geraden nicht in der (x,y)-Ebene. Die Geraden der (x,y)-Ebene nennen wir *die unendlich fernen Punkte* von $\mathcal{A}$. Geraden in $\mathcal{P}$ entsprechen Ebenen im $\mathbb{R}^3$, und die (x,y)-Ebene heißt *die unendlich ferne Gerade.*

Dann schneiden sich *je* zwei verschiedene Geraden von $\mathcal{P}$ in genau einem Punkt. Dieser ist genau dann ein unendlich ferner Punkt, wenn die Geraden affine Parallelen sind.

Definition 44. Sei V ein $(n+1)$-dimensionaler K-Vektorraum.

(i) Der *projektive Raum $\mathcal{P}(V)$ über V* ist

$$\mathcal{P}(V) := \{g \mid g \text{ Nullpunktsgerade in } V\}.$$

Die Elemente von $P(V)$ bezeichnen wir als die *Punkte* von $P(V)$.

(ii) Ein *k-dimensionaler (projektiver) Unterraum $\mathcal{L}$ von $\mathcal{P}(V)$* ist eine Teilmenge der Form

$$\mathcal{L} = \{g \in \mathcal{P}(V) \mid g \subset V'\},$$

wobei $V' \subset V$ ein $(k+1)$-dimensionaler Vektorunterraum ist.

Ein 1-dimensionaler Unterraum heißt *(projektive) Gerade*, ein $(n-1)$-dimensionaler eine *(projektive) Hyperebene.*

Satz 76 (Dimensionssatz). *Sind $\mathcal{L}_1$ und $\mathcal{L}_2$ zwei Unterräume von $\mathcal{P}(V)$ und bezeichnet $\mathcal{L}_1 + \mathcal{L}_2$ den kleinsten Unterraum von $\mathcal{P}(V)$ der beide enthält, so gilt*

$$\dim \mathcal{L}_1 + \dim \mathcal{L}_2 = \dim(\mathcal{L}_1 + \mathcal{L}_2) + \dim(\mathcal{L}_1 \cap \mathcal{L}_2).$$

Übergang zur affinen Geometrie. Ist

$$\mathcal{H} = \{g \in \mathcal{P}(V) \,|\, g \subset V'\} \subset \mathcal{P}(V)$$

eine projektive Hyperebene und $x_0 \in V \backslash V'$, so ist

$$\mathcal{A} := x_0 + V'$$

eine affine Hyperebene. Die Abbildung

$$\mathcal{A} \to \mathcal{P}(V),\ x \mapsto g(0, x) = Kx$$

ist dann eine Bijektion von $\mathcal{A}$ auf $\mathcal{P}(V) \backslash \mathcal{H}$. Das Komplement einer projektiven Hyperebene „ist“ also ein affiner Raum.

Projektiver Abschluß. Umgekehrt kann man „durch Hinzunahme einer unendlich fernen Hyperebene“ jeden affinen Raum in einen projektiven Raum gleicher Dimension einbetten.

Genauer: Ist $\mathcal{A}$ ein affiner Raum über dem K-Vektorraum V und wählt man einen Basispunkt $Q \in \mathcal{A}$, so hat man eine Bijektion

$$\mathcal{A} \to V, P \mapsto (P - Q).$$

Wir definieren $\tilde{V} := V \times K$. Dann ist

$$\mathcal{A} \to V \times \{1\} \subset \tilde{V}, P \mapsto (P - Q) \times \{1\}$$

ebenfalls eine Bijektion und damit auch

$$\mathcal{A} \to \mathcal{P}(\tilde{V}), P \mapsto K((P - Q) \times \{1\})$$

eine Bijektion von $\mathcal{A}$ auf $\mathcal{P}(\tilde{V}) \backslash \mathcal{P}(V \times \{0\})$.

Homogene Koordinaten. Zur Beschreibung von Punkten und Punktmengen in einem projektiven Raum $\mathcal{P}(V)$ benutzt man statt der Geraden in V gern einfach nur Punkt $\neq 0$. Man nennt $x \in V \backslash \{0\}$ eine *homogene Koordinate* für

$$[x] := Kx \in \mathcal{P}(V).$$

Beispiel 48. Im n-dimensionalen reellen projektiven Raum

$$\mathbb{R}P^n := \mathcal{P}(\mathbb{R}^{n+1})$$

ist $(x_1, \ldots, x_{n+1}) \neq 0$ eine homogene Koordinate für $[x_1, \ldots, x_{n+1}]$. Zwei Punkte $(x_1, \ldots, x_{n+1})$ und $(y_1, \ldots, y_{n+1})$ beschreiben denselben Punkt von $\mathbb{R}P^n$ genau dann, wenn

$$(x_1, \ldots, x_{n+1}) = \lambda(y_1, \ldots, y_{n+1})$$

für ein $\lambda \in \mathbb{R} \backslash \{0\}$. Insbesondere ist zum Beispiel

$$[x_1, \ldots, x_{n+1}] = [\frac{x_1}{x_{n+1}}, \ldots, \frac{x_n}{x_{n+1}}, 1],$$

sofern $x_{n+1} \neq 0$. Die Punkte, die man so beschreiben kann, sind also gerade das Komplement der Hyperebene

$$\mathcal{H} = \{[x_1, \ldots, x_{n+1}] \,|\, x_{n+1} = 0\}.$$

□

Beispiel 49 (Das Sphärenmodell für $\mathbb{R}P^n$). Jeder Punkt auf $\mathbb{R}P^n$, d.h. jede Nullpunktsgerade im $\mathbb{R}^{n+1}$ bestimmt eindeutig ein Paar antipodischer Punkte auf der Einheitssphäre $x_1^2+\ldots+x_{n+1}^2=1$. Also kann man sich $\mathbb{R}P^n$ auch vorstellen als die Sphäre mit identifizierten Antipoden. Oder als die obere Halbsphäre mit Antipodenidentifikation auf dem Äquator.

Insbesondere ist die reelle projektive Gerade $\mathbb{R}P^1$ zu interpretieren als der Halbkreis mit Identifizierung der beiden einzigen „Äquatorpunkte“, und das ergibt wieder eine Kreislinie.

Die reelle projektive Ebene ist die identifizierte Kugelfläche $S^2/\sim$. Eine projektive Gerade entspricht einem Großkreis auf S^2. Je zwei verschiedene solche schneiden sich in einem Paar von Antipodenpunkten, d.h. in *einem* Punkt von $\mathbb{R}P^2$.

□

Projektivitäten = Projektive Transformationen. Ein Automorphismus $\phi \in \mathrm{Aut}(V)$ induziert eine bijektive Abbildung der Nullpunktsgeraden von V auf die Nullpunktsgeraden von V, also eine Bijektion von $\mathcal{P}(V)$ auf sich. Die so induzierten Abbildungen $f : \mathcal{P}(V) \to \mathcal{P}(V)$ heißen *Projektivitäten* oder *projektive Transformationen* von $\mathcal{P}(V)$. Beachten Sie daß ϕ und $\lambda\phi$ dieselbe Transformation auf $\mathcal{P}(V)$ induzieren. Die projektiven Transformationen bilden eine Gruppe, die projektive Gruppe des $\mathcal{P}(V)$.

Beispiel 50. Wir betrachten den $\mathbb{R}P^n$ mit der unendliche fernen Hyperebene $x_{n+1}=0$, also mit dem eingebetteten affinen Raum

$$\mathbb{R}^n = \mathbb{R}^n \times \{1\}.$$

Jeder Automorphismus von $\mathbb{R}^{n+1}$ wird bezüglich der Standardbasis beschrieben durch eine Matrix

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

wobei A eine $(n \times n)$-Matrix ist und $a = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$, $b = (b_1 \ldots b_n)$ und $c \in \mathbb{R}$.

Falls $b = 0$ und $c = 1$ entspricht f auf dem $\mathbb{R}^n$ einer affinen Abbildung $Ax + a$, aber die projektive Gruppe ist größer. Insbesondere können Punkte von $\mathbb{R}^n$ auch in unendlich ferne Punkte abgebildet werden.

□

Wir wollen nicht systematisch auf Quadriken in $\mathcal{P}(V)$ eingehen, aber ein Beispiel in der reellen projektiven Ebene betrachten.

Beispiel 51. Wir betrachten wie im vorhergehenden Beispiel den $\mathbb{R}^2 = \mathbb{R}^2 \times \{1\}$ als Teilmenge von $\mathbb{R}P^2$. Den Kreis

$$K = \{(x, y) \,|\, x^2 + y^2 = 1\} \subset \mathbb{R}^2$$

identifizieren wir also mit der Menge

$$\{[x, y, 1] \,|\, x^2 + y^2 = 1\} = \{[x, y, z] \,|\, (\frac{x}{z})^2 + (\frac{y}{z})^2 = 1\} = \{[x, y, z] \,|\, x^2 + y^2 = z^2\}.$$

Nun betrachten wir die projektive Transformation

$$f : \mathbb{R}P^2 \to \mathbb{R}P^2, [x, y, z] \mapsto [z, y, x]$$

und fragen nach $f^{-1}(K)$. Das sollte bei vernünftigen Definitionen dann eine zum Kreis K projektiv äquivalente Quadrik sein.

Es gilt

$$f([x,y,z]) \in K \iff [z,y,x] \in K \iff z^2+y^2=x^2$$

Wenn wir $f^{-1}(K) \cap \mathbb{R}^2$ betrachten, also die Punkte in $f^{-1}(K)$ mit $z=1$, so erhalten wir $1+y^2=x^2$ oder

$$y^2-x^2=1.$$

Das ist also eine Hyperbel.

Weil es mehr projektive als affine Transformationen gibt sind projektiv gesehen mehr Quadriken äquivalent, offenbar Ellipsen und Hyperbeln. Und auch Parabeln, wie man mit der Transformation

$$f : \mathbb{R}P^2 \to \mathbb{R}P^2, [x,y,z] \mapsto [x-z, \sqrt{2}y, x+z]$$

zeigt. Diese bildet die Parabel $y^2=2x$ in den obigen Kreis ab. Beachten Sie: In der projektiven Ebene ist die Parabel gegeben durch

$$\{[x,y,z] \,|\, (\frac{y}{z})^2 = 2\frac{x}{z}\} = \{[x,y,z] \,|\, y^2 = 2xz\}.$$

Zur Parabel gehört also auch der unendlich ferne Punkt

$$[x,0,0]$$

mit $x \neq 0$.

□

Dualität. Sei $n+1 := \dim V$ und V^* der Dualraum von V. Das ist auch ein $n+1$ dimensionaler Vektorraum, und man kann $\mathcal{P}(V^*)$ betrachten.

Ein Punkt in $\mathcal{P}(V^*)$ ist eine Gerade $K\omega = [\omega]$, wobei $\omega \in V^*$ eine nicht-verschwindende Linearform auf V ist. Alle $\lambda\omega$ mit $\lambda \in K\backslash\{0\}$ haben denselben Kern, und dieser ist eine Hyperebene in V, definiert also eine projektive Hyperebene in $\mathcal{P}(V)$. Sie besteht aus allen Geraden, auf denen ω verschwindet.

Umgekehrt: Zu einer Hyperebene $V' \subset V$ gibt es zwar viele $\omega \in V^*$ mit $\operatorname{Kern}\omega = V'$, aber diese unterscheiden sich alle nur um einen skalaren Faktor. Man hat also eine Bijektion

$$\text{Punkte in } \mathcal{P}(V^*) \;\leftrightarrow\; \text{Hyperebenen in } \mathcal{P}(V)$$

Eine Gerade $\gamma \subset \mathcal{P}(V^*)$ besteht aus allen $[\omega] = [a_1\omega_1 + a_2\omega_2]$ mit $a_1, a_2 \in K$, nicht beide $= 0$. Dabei sind ω_1 und ω_2 zwei linear unabhängige Linearformen auf V. Es ist leicht zu sehen, daß

$$\bigcap_{\omega\in\gamma} \operatorname{Kern}\omega = \operatorname{Kern}\omega_1 \cap \operatorname{Kern}\omega_2 =: U',$$

wobei $U' \subset V$ ein Unterraum der Dimension $n+1-2$ ist (Lösungsraum zweier unabhängiger homogener linearer Gleichungen), also einem $(n-2)$-dimensionalen projektiven Unterraum $\mathcal{L} \subset \mathcal{P}(V)$ entspricht. Den Punkten der Geraden $\gamma \subset \mathcal{P}(V^*)$ entspricht ein *Bündel* von Hyperebenen in $\mathcal{P}(V)$, nämlich allen Hyperebenen, die $\mathcal{L}$ enthalten.

Beispiel 52. Im Fall der projektiven Ebene ($n=2$) entsprechen die Punkte von $\mathcal{P}(V^*)$ den Geraden von $\mathcal{P}(V)$ und die Geraden von $\mathcal{P}(V^*)$ den Punkten von $\mathcal{P}(V)$. Daß ein Punkt auf einer Geraden liegt, bedeutet, daß die duale Gerade durch den dualen Punkt geht.

□

Es ist eine beliebte Methode der projektiven Geometrie, Sätze über $\mathcal{P}(V^*)$ dual in $\mathcal{P}(V)$ zu interpretieren (und umgekehrt).

Wir können dafür hier nur ein ganz triviales Beispiel geben. Es ist leicht, den folgenden Satz zu beweisen:

Satz A: Durch zwei verschiedene Punkte einer projektiven Ebene geht genau eine Gerade.

Seien nun zwei verschieden Geraden in der projektiven Ebene $\mathcal{P}(V)$ gegeben. Diesen entsprechen zwei verschiedene Punkte in $\mathcal{P}(V^*)$. Nach dem Satz geht durch diese beiden Punkte genau eine Gerade in $\mathcal{P}(V^*)$. Dieser Geraden entspricht in $\mathcal{P}(V)$ ein Punkt, der auf den beiden gegebenen Geraden liegt. Einem weiteren Schnittpunkt der Geraden würde eine weitere Gerade durch die beiden Punkte in $\mathcal{P}(V^*)$ entsprechen, die es aber nach dem Satz nicht gibt. Also folgt durch Dualisierung von Satz A der

Satz B: Zwei Geraden einer projektiven Ebene schneiden sich in genau einem Punkt.

Zum Schluß ... noch ein Beispiel, das zwar nicht in die orthodoxe projektive Geometrie gehört, sehr wohl aber eine wichtige Rolle in der Differentialgeometrie und Topologie spielt.

Beispiel 53 ($\mathbb{R}P^3 = \mathrm{SO}(3)$). Unter der Orthogonalprojektion vom $\mathbb{R}^4$ auf den $\mathbb{R}^3$ entspricht die obere Hälfte der S^3 gerade der Vollkugel

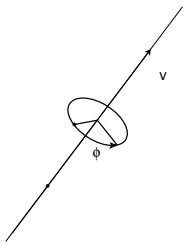
$$B := \{(x, y, z) \,|\, x^2 + y^2 + z^2 \leq 1\}$$

im $\mathbb{R}^3$. Die äquatoriale S^2 von S^3 entspricht dabei der Randsphäre $x^2+y^2+z^2 = 1$. Deshalb können wir $\mathbb{R}P^3$ interpretieren als die dreidimensionale Einheitsvollkugel mit Identifikation $(x, y, z) \sim -(x, y, z)$ auf dem Rand.

Wir erinnern daran, daß SO(3) die Gruppe der orthogonalen Transformationen des $\mathbb{R}^3$ mit Determinante= $+1$ war. Die drei (komplexen) Eigenwerte einer solchen Transformation sind vom Betrag 1 und multiplizieren sich zur Determiante 1. Weiter ist mit λ auch $\bar{\lambda}$ ein Eigenwert. Daraus ergibt sich sofort, daß $+1$ ein einfacher oder dreifacher Eigenwert ist. Im letzteren Fall ist die Transformation natürlich die Identität. Also ist jedes $A \in \mathrm{SO}(3)$ eine Drehung um eine Achse.

Mit dieser Information zeigt man, daß folgende Abbildung surjektiv und im wesentlichen bijektiv ist:

Jedem $v \in B$ ordnen wir die „positive“ Drehung um den Winkel $\phi = \pi\|v\|$ um die „orientierte“ Achse $\mathbb{R}v$ zu, vgl. die Abbildung.
Dem Nullvektor ordnen wir die Identität zu.
Dann ist die Abbildung auf dem Inneren von B injektiv. Nur antipodische Randpunkte liefern dieselbe orthogonale Transformation: Beim Drehwinkel π ist es egal, mit welcher Orientierung man dreht.



Wir können die Elemente von SO(3) also eineindeutig kodieren durch die Vektoren in der Einheitskugel mit Antipoden identifizierung auf dem Rand. Damit haben wir die die Punkte von $\mathbb{R}P^3$ bijektiv auf die Drehungen aus SO(3) abgebildet.

□

ENDE