VII. EUKLIDISCHE RÄUME

Im Kapitel II haben wir gesehen, wie $\mathbf{R}^2$ und $\mathbf{R}^3$ mit einem natürlichen Skalarprodukt ausgestattet werden können. Damit haben wir Begriffe wie Orthogonalität und Winkel betrachtet. Jetzt arbeiten wir mit diesen Begriffen in höheren Dimensionen, wobei wir wieder die axiomatische Methode verwenden.

Definition: Ein **Skalarprodukt** (oder **inneres Produkt**) auf einem reelen Vektorraum V ist eine Abbildung

$$(\ |\ ): V \times V \to \mathbf{R}$$

sodaß
1) die Abbildung **bilinear** ist, d.h.

$$(\lambda_1 x_1 + \lambda_2 x_2 | \mu_1 y_1 + \mu_2 y_2) = \lambda_1 \mu_1 (x_1|y_1) + \lambda_1 \mu_2 (x_1|y_2) + \lambda_2 \mu_1 (x_2|y_1) + \lambda_2 \mu_2 (x_2|y_2)$$

$$(\lambda_1, \lambda_2, \mu_1, \mu_2 \in \mathbf{R}, x_1, x_2, y_1, y_2 \in V);$$

2) $(\ |\ )$ ist **symmetrisch** d.h. $(x|y) = (y|x) \quad (x, y \in V)$;
3) $(\ |\ )$ ist **positiv definit** d.h. $(x|x) > 0$ für $x \neq 0$.

(**Bemerkung:** Aus Eigenschaft 3) folgt: Sei $x \in V$ mit der Eigenschaft, daß $(x|y) = 0$ für jedes $y \in V$. Dann gilt: $x = 0$.)

Ein euklidischer Vektorraum ist ein Vektorraum V über $\mathbf{R}$, zusammen mit einem Skalarprodukt $(\ |\ )$ darauf.

Beispiel: I. $\mathbf{R}^n$ mit dem Skalarprodukt

$$(x|y) = \xi_1 \eta_1 + \cdots + \xi_n \eta_n$$

ist ein euklidischer Raum;
II. Die folgenden Abbildungen auf $\mathrm{Pol}\,(n)$ sind Skalarprodukte:

$$(p|q) := \int_0^1 p(t) q(t) dt.$$

$$(p|q) := a_0 b_0 + a_1 b_1 + \cdots + a_n b_n$$

(wobei

$$p(t) = a_0 + a_1 t + \cdots + a_n t^n$$
$$q(t) = b_0 + b_1 t + \cdots + b_n t^n).$$

In euklidischen Räumen kann man die Grundbegriffe (Länge, Winkel, Orthogonalität) der euklidischen Geometrie einführen.

Länge: Die **Länge** $\|x\|$ eines Vektors $x \in V$ ist folgendermaßen definiert:

$$\|x\| := \sqrt{(x|x)}$$

(Z.B. in $\mathbf{R}^n$ gilt:

$$\|x\| = (\xi_1^2 + \cdots + \xi_n^2)^{\frac{1}{2}}).$$

Einige Eigenschaften:
I.

$$|(x|y)| \leq \|x\|\|y\| \qquad (x, y \in V)$$

(genannt die CAUCHY-SCHWARZ-Ungleichung – nach dem klassischen Fall

$$\xi_1\eta_1 + \cdots + \xi_n\eta_n \leq (\xi_1^2 + \cdots + \xi_n^2)^{\frac{1}{2}}(\eta_1^2 + \cdots + \eta_n^2)^{\frac{1}{2}} \quad \text{in} \quad \mathbf{R}^n).$$

Beweis: Die quadratische Funktion

$$\begin{aligned} \lambda \mapsto\ & (x + \lambda y | x + \lambda y) \\ &= \lambda^2\|y\|^2 + 2\lambda(x|y) + \|x\|^2 \end{aligned}$$

ist nirgends negativ. Daher ist die Diskrimante nicht positiv, d.h. $4(x|y)^2 - 4\big(\|y\|^2\|x\|^2\big) \leq 0$. q.e.d.

II. $\|x + y\| \leq \|x\| + \|y\|$ (Dreiecksungleichung).

Beweis:

$$\begin{aligned} \|x + y\|^2 &= (x + y | x + y) \\ &= \|x\|^2 + 2(x|y) + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \quad \text{(nach I.)} \\ &= \big(\|x\| + \|y\|\big)^2 \text{q.e.d.} \end{aligned}$$

Winkel: Seien x, y Vektoren aus einem euklidischen Raum V, wobei $x \neq 0$, $y \neq 0$. Dann gilt (nach I. oben)

$$-1 \leq \frac{(x|y)}{\|x\|\|y\|} \leq 1.$$

Daher gibt es ein eindeutig bestimmtes $\theta \in [0, \pi]$ mit $\cos\theta = (x|y)/\|x\|\|y\|$.

θ heißt der **Öffnungswinkel** zwischen x und y (geschrieben $\angle(x, y)$). Falls $x = 0$ oder $y = 0$, dann setzen wir $\angle(x, y) = 0$.

Orthogonalität: Zwei Vektoren $x, y \in V$ sind **orthogonal** (geschrieben $x \perp y$) falls $(x|y) = 0$. (D.h. falls $\angle(x, y) = \pi/2$ oder $x = 0$ oder $y = 0$). **Ein Orthonormalsystem** in V (kurz ON-System) ist eine Menge von Einheitsvektoren, die paarweise orthogonal sind.

Beispiel: Die kanonische Basis von $\mathbf{R}^n$ ist ein ON-System.

Bemerkung: Die Vektoren eines ON-Systems sind linear unabhängig.

Beweis: Sei $\{x_1, \dots, x_n\}$ ein ON-System. Dann gilt:

$$\begin{aligned}
\sum_{k=1}^{n} \lambda_k x_k = 0 &\Rightarrow \Big(\sum_{k=1}^{n} \lambda_k x_k | x_i\Big) = 0 \quad (i = 1, \dots, n) \\
&\Rightarrow \sum \lambda_k (x_k | x_i) = 0 \\
&\Rightarrow \lambda_i (x_i | x_i) = 0 \\
&\Rightarrow \lambda_i = 0 \qquad\qquad\qquad \text{q.e.d.}
\end{aligned}$$

Insbesondere ist in einem euklidischen Raum V mit $\dim V = n$ jedes ON-System $\{x_1, \dots, x_n\}$ mit n-Elementen automatisch eine Basis. Solche Basen (die sogenannten **orthonormalen Basen**) haben die angenehme Eigenschaft, daß die Koordinaten eines Vektors x mit Hilfe des Skalarproduktes sofort ausgerechnet werden können. Es gilt nämlich: Sei $x \in V$. Dann hat x die Darstellung

$$x = \sum_{k=1}^{n} (x | x_k) x_k$$

bezüglich der orthonormalen Basis $(x_1, \dots, x_n)$ von V. Denn aus $x = \sum_{i=1}^{n} \lambda_i x_i$ folgt:

$$\lambda_k = \left(\sum_{i=1}^{n} \lambda_i x_i | x_k\right) = (x | x_k).$$

Die Existenz von orthonormalen Basen folgt aus dem folgenden Satz, dessen Beweis eine natürliche geometrische Konstruktion, die wir fürher kennengelernt haben, verwendet.

Satz: Sei $(x_1, \dots, x_k)$ ein ON-System in einem euklidischen Raum V. Dann existiert $x_{k+1}, \dots, x_n$ in V, sodaß $(x_1, \dots, x_n)$ eine ON-Basis für V ist.

Beweis: Wähle $z \notin [x_1, \dots, x_k]$. Definiere

$$\tilde{x}_{k+1} = z - \sum_{i=1}^{n} (z | x_i) x_i$$

Dann gilt: $\tilde{x}_{k+1} \neq 0$ (warum?) und $\tilde{x}_{k+1} \perp x_i$ für $i = 1, \dots, k$. Wir definieren dann $x_{k+1} = \tilde{x}_{k+1} / \|\tilde{x}_{k+1}\|$.
$(x_1, \dots, x_{k+1})$ ist ein ON-System. Wir fahren weiter fort, bis wir ein ON-System mit n-Elementen (und daher eine ON-Basis) bekommen.

Korollar: Jeder endlich dimensionale euklidische Raum hat eine ON-Basis $(x_1, \dots, x_n)$.

Beispiel: Wir wenden diese Methode an, um eine ON-Basis für Pol (2) bezüglich des Skalarproduktes

$$(p|q) = \int\limits_0^1 p(t)q(t)dt$$

zu erhalten.

Wir wählen:

$$\begin{aligned}
x_0 &= 1 \\
\tilde{x}_1(t) &= t - (t|1)1 = t - \frac{1}{2}, \\
x_1(t) &= \sqrt{12}\left(t - \frac{1}{2}\right), \\
\tilde{x}_2(t) &= t^2 - (t^2|x_0)x_0(t) - (t^2|x_1)x_1(t), \\
&= t^2 - t + \frac{1}{6}, \\
x_2(t) &= 6\sqrt{5}\left(t^2 - t + \frac{1}{6}\right).
\end{aligned}$$

Wiederum sind lineare Abbildungen zwischen euklidischen Räumen, die ihre Struktur respektieren, besonders interessant: Daher die

Definition: Eine lineare Abbildung $f\colon V \to V_1$ (wobei V und V_1 euklidisch sind) heißt eine **Isometrie,** falls

$$(fx|fy) = (x|y) \quad (x, y \in V).$$

Wir notieren einige einfache Eigeschaften solcher Abbildungen. Einige von diesen Ergebnissen haben wir schon im Kapitel II kennengelernt. Normalerweise kann man die Beweise für $V = \mathbf{R}^2$ auf den allgemeinen Fall sofort übertragen.

I. Eine Isometrie f respektiert die Längen von Vektoren.
Denn $\|fx\|^2 = (fx|fy) = (x|y) = \|x\|^2$.
Anderseits ist eine lineare Abbildung f mit der Eigenschaft, daß $\|f(x)\| = \|x\| \quad (x \in V)$ automatisch eine Isometrie, (denn

$$(x|y) = \frac{\left(\|x+y\|^2 - \|x-y\|^2\right)}{4}).$$

II. Eine Isometrie $f\colon V \to V_1$ ist immer injektiv, (denn aus $fx = 0$ folgt $\|x\| = \|fx\| = 0$ d.h. $x = 0$). f ist genau dann surjektiv, (und daher bijektiv, d.h. ein linearer Isomorphismus), wenn $\dim V_1 = \dim V$.

III. Eine Abbildung $f: V \to V_1$ mit der Eigenschaft, daß $\|f(x) - f(y)\| = \|x - y\| \quad (x, y \in V)$ hat immer die Gestalt $x \mapsto fx + u$ (wobei f eine lineare Isometrie ist und $u \in V_1$).

IV. Eine lineare Isometrie bildet orthonormale Systeme in orthonormale Systeme ab. Insbesondere bildet sie ON-Basen in ON-Basen ab, falls $\dim V = \dim V_1$.

Andererseits gilt: wenn f eine lineare Abbildung ist, sodaß f **eine** Basis von V in eine ON-Basis von V_1 abbildet, dann ist f eine Isometrie.

Daraus folgt: Jeder n-dimensionale euklidische Raum V ist isometrisch isomorph zu $\mathbf{R}^n$. (Wähle eine lineare Abildung f von $\mathbf{R}^n$ in V, sodaß f die kanonische Basis für $\mathbf{R}^n$ in eine ON-Basis $(x_1, \dots, x_n)$ von V abbildet.)

Sei jetzt A eine $n \times n$ Matrix. Da die Bilder der kanonischen Basis von $\mathbf{R}^n$ bezüglich f_A gerade die Spalten von A sind, sehen wir, daß f_A genau dann eine Isometrie ist, wenn die Spalten von A eine orthonormale Basis für $\mathbf{R}^n$ bilden. Anders ausgedrückt: A erfüllt die Bedingung:

$$A^t A = I.$$

Daraus folgt, daß $A^t = A^{-1}$ bzw. $AA^t = I$. (Solche Matrizen nennt man **orthonormal**.)

Typische Beispiele von Isometrien sind:

Spiegelung: Das sind Abbildungen mit Matrizen der Gestalt

$$\begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & -1 & & & \\ & & & & . & & \\ & & & & & 1 & \\ & & & & & & \ddots & \\ & & & & & & & 1 \end{bmatrix}$$

Drehungen: Sie haben Matrizen der Gestalt

$$\begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \cos\theta & -\sin\theta & & \\ & & & \sin\theta & \cos\theta & & \\ & & & & & 1 & \\ & & & & & & \ddots & \\ & & & & & & & 1 \end{bmatrix}$$

bzgl. einer geeigneten orthonormalen Basis. Wir werden später sehen, daß alle linearen Isometrien aus Abbildungen dieser Gestalten zusammengesetzt sind.

Orthogonale Projektionen:

Im Kapitel II haben wir gesehen, daß jeder Teilraum eines Vektorraumes einen Komplementärraum besitzt und dementsprechend eine Projektion, die den ganzen Raum auf den Teilraum projiziert. Allerdings gibt es viele solche Projektionen. Wir werden jetzt sehen, daß die Geometrie eines euklidischen Raumes eine wichtige Projektion – die sogenannte orthogonale Projektion – aus diesen Projektionen auszeichnet.

Definition: Sei M ein Teilraum eines euklidischen Raumes. Die Menge

$$M^{\perp} := \big\{x \in V : (x|y) = 0 \quad \text{für} \quad y \in M\big\}$$

heißt **orthogonales Komplement** von M.

Satz:
1) $M^{\perp}$ ist eine Teilraum von V.
2) M und $M^{\perp}$ sind Komplementärräume d.h. $V = M \oplus M_1$. Genauer, sei $(x_1, \dots, x_n)$ eine ON-Basis für V, sodaß $M = [x_1, \dots, x_r]$. Dann gilt:

$$M^{\perp} = [x_{r+1}, \dots, x_n].$$

Beweis: Da $x_i \in M^{\perp}$ für $i = r+1, \dots, n$, ist es klar, daß $[x_{r+1}, \dots, x_n] \subseteq M^{\perp}$. Sei jetzt $x \in M^{\perp}$. x hat die Darstellung

$$x = \sum_{k=1}^{n} (x|x_k)x_k = \sum_{k=r+1}^{n} (x|x_k)x_k \in [x_{r+1}, \dots, x_n]$$

da $(x|x_k) = 0 \quad (k = 1, \dots, r)$. q.e.d.

Sei jetzt M ein Teilraum des euklidischen Raumes V. Nach diesem Satz hat jedes $x \in V$ eine eindeutige Darstellung $y + z (y \in M, z \in M^{\perp})$. y heißt die **orthogonale Projektion** von x auf M (geschrieben $P_M x$). (D.h. P_M ist die Projektion auf M entlang $M^{\perp}$.)

Bemerkungen: I. Bezüglich der orthonormalen Basis, die wir in dem obigen Beweis konstruierten, hat P_M die Matrix

$$\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$$

Andererseits gilt: Eine lineare Abbildung, die eine solche Darstllung bezgl. einer it orthonormalen Basis besitzt, ist eine Projektion dieser Art.

II. Der Punkt $P_M(x)$ ist die beste Approximation zu x aus M d.h.

$$\|x - P_M(x)\| < \|x - y\| \quad \big(y \in M, y \neq P_M(x)\big).$$

III. $x - 2P_M(x)$ ist das Spiegelbild von x an $M^{\perp}$. Daher heißt die lineare Abbildung Id $-2P_M$ eine **Spiegelung**.

Orthogonale Zerlegungen: Sei V ein euklidischer Raum. Eine Folge $V_1, \dots, V_r$ von Teilräumen von V ist eine **orthogonale Zerlegung**, falls
1) $V_i \perp V_j (i \neq j)$ (d.h. $x_i \perp x_j$ falls $x_i \in V_i, x_j \in V_j$);
2) jedes $x \in V$ hat eine Darstellung

$$x_1 + \cdots + x_r$$

(wobei $x_i \in V_i$ für jedes i).

Nach 1) ist diese Zerlegung eindeutig bestimmt. Wir schreiben dann $V = V_1 \oplus, \dots, \oplus V_r$. Eine wichtige Bemerkung ist dann: Sei für jedes i $(x_1^i, \dots, x_{n_i}^i)$ eine orthonormale Basis für V_i. Dann ist

$$(x_1^1, x_2^1, \dots, x_{n_1}^1, x_1^2, \dots, x_1^r, \dots, x_{n_r}^r)$$

eine orthonomale Basis für V.

Jetzt betrachten wir den Begriff der adjungierten Abbildung. Falls f eine lineare Abbildung von V in V_1 ist, dann heißt eine lineare Abbildung $g\colon V_1 \to V$ **zu f adjungiert,** falls

$$\big(f(x)|y\big) = \big(x|g(y)\big) \quad (x \in V, y \in V_1).$$

Satz: Es gibt genau eine lineare Abbildung $g\colon V_1 \to V$, sodaß f und g adjungiert sind.

Beweis: Zunächst wählen wir orthonormale Basen $(x_1, \dots, x_n)$ bzw. $(y_1, \dots, y_m)$ für V bzw. V_1 und bezeichnen mit $A = [a_{ij}]$ die Matrix von f bezüglich dieser Basen d.h. $a_{ij} = \big(f(x_i)|y_j\big)$. Falls ein solches g existiert, dann gilt

$$a_{ij} = \big(f(x_j)|y_i\big) = \big(x_j|g(y_i)\big)$$

und das ist das (j,i)-te Element der Matrix, die g darstellt. Daraus folgt, daß die Matrix von g gleich A^t sein muß. Wir müssen jetzt nur zeigen, daß die Abbildung g, mit Matrix A^t die Bedingung

$$\big(f(x)|y\big) = \big(x|g(y)\big)$$

erfüllt.

Dazu rechnen wir einfach nach:

$$\begin{aligned}
\big(f(x)|y\big) &= \left(f\Big(\sum_{j=1}^{n} (x|x_j)x_j\Big) | \sum_{k=1}^{m} (y|y_k)y_k \right) \\
&= \left(\sum_{i=1}^{m} \Big(\sum_{j=1}^{n} a_{ij}(x|x_j)\Big) y_i | \sum_{k=1}^{m} (y|y_k)y_k \right) \\
&= \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{m} a_{ij}(x|x_j)(y|y_k)(y_i|y_k) \\
&= \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}(x|x_j)(y|y_i).
\end{aligned}$$

Analog gilt:

$$\begin{aligned}\big(x|g(y)\big) &= \left(\sum_{j=1}^{n}(x|x_j)x_j|g\Big(\sum_{i=1}^{m}(y|y_i)y_i\Big)\right)\\ &= \sum_{i=1}^{m}\sum_{j=1}^{n}a_{ij}(x|x_j)(y|y_i) \qquad \text{q.e.d.}\end{aligned}$$

Die Abbildung g, die wir in diesem Beweis konstruiert haben, bezeichnet man mit f^t. Eine Abbildung $f\colon V \to V$ heißt **selbst-adjungiert**, falls $f = f^t$. Das bedeutet, daß die Matrix von f bezüglich einer (und daher jeder) orthonormalen Basis symmetrisch ist.
Mit dieser Schreibweise sehen wir, daß eine Abbildung $f \in L(V)$ genau dann eine Isometrie ist, wenn $f^t f = \mathrm{Id}$ (d.h. $f^{-1} = f^t$). Denn f ist eine Isometrie, wenn

$$(x|y) = \big(f(x)|f(y)\big) = \big(f^t f(x)|y\big)$$

gilt für jedes y. Daher gilt $f^t f(x) = x$ für jedes x d.h. $f^t f = \mathrm{Id}$.

Besondere Beispiele von selbst-adjungierten Abbildungen sind Abbildungen, die eine diagonale Matrix bezüglich einer *orthonormalen* Basis besitzen. Der wichtigste Satz dieses Kapitels besagt, daß *alle* selbst-adjungierten Abbildungen diese Eigenschaft haben. Wir bringen einen Beweis, der einfache Hilfsmittel aus der Analysis verwendet.

Satz: Sei f eine selbst-adjungierte lineare Abbildung von V in V. Dann besitzt V eine orthonormale Basis $(x_1, \ldots, x_n)$, wobei jedes x_i ein Eigenvektor von f ist. Bezüglich dieser Basis hat f die Matrix $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, wobei λ_i die Eigenwerte von f sind.

Beweis: Wir betrachten die Funktion:

$$\varphi\colon x \to \frac{\big(f(x)|x\big)}{(x|x)}$$

auf $V \setminus \{0\}$, Es gilt: $\varphi(\lambda x) = \varphi(x) \quad (\lambda \neq 0)$.

Nach einem Satz aus der Analysis, existiert $x_1 \in V$, sodaß $\|x_1\| = 1$ und $\varphi(x_1) \geq \varphi(x)$ $\big(\|x\| = 1\big)$. Die Funktion

$$\psi\colon t \to \varphi(x_1 + ty) \quad (y \neq 0)$$

hat daher ein Minimum für $t = 0$. Daraus folgt:

$$\psi'(0) = 0.$$

Aber $\psi'(0) = 2\big(y|f(x_1)\big) - 2\lambda_1(y|x_1)$ wobei $\lambda_1 := \big(x_1|f(x_1)\big)$ (siehe unten). Daraus folgt:

$$\big(y|f(x_1)\big) = \lambda_1(y|x_1).$$

Da diese Gleichung für jedes $y \in V$ gilt, kann man schließen, daß $f(x_1) = \lambda_1 x_1$ d.h. x_1 ist ein Eigenvektor von f (mit Eigenwert $\lambda_1 = \big(f(x_1)|x_1\big)$).

Seien jetzt $V_1 := [x_1]^\perp$. Dann gilt $f(V_1) \subseteq V_1$. Denn aus $x \in V_1$ folgt $(x|x_1) = 0$ und daher

$$\big(f(x)|x_1\big) = \big(x|f(x_1)\big) = (x|\lambda_1 x_1) = 0$$

d.h. $f(x) \in V_1$.

Da $f|_{V_1}$ jetzt eine selbst-adjungierte Abbildung von V_1 in sich selbst ist, kann man wieder die gleiche Methode anwenden, um einen Eigenwert λ_2 und Eigenvektor x_2 in V_1 zu finden. Ein natürlicher Induktionsbeweis liefert jetzt das Ergebnis.

Berechnung der Ableitung von ψ:

$$\begin{aligned}
\psi'(0) &= \lim_{t\to 0} \frac{1}{t}\big(\psi(t) - \psi(0)\big) \\
&= \lim_{t\to 0} \frac{1}{t}\left[\varphi(x_1 - ty) - \varphi(x_1)\right] \\
&= \lim_{t\to 0} \frac{1}{t}\left[\frac{(x_1|x_1)(f(x_1)|x_1)+t(f(y)|x_1)+t(f(x_1)|y)+t^2(f(y)|y)-((x_1|x_1)+t(y|x_1)+t(x_1|y)+t^2(y|y))(f(x_1)|x_1)}{(x_1+ty|x_1+ty)(x_1|x_1)}\right] \\
&= 2\big(y|f(x_1)\big) - 2(y|x_1)\lambda_1.
\end{aligned}$$

Die Funktion φ heißt RAYLEIGH Funktion von f. Aus der Beweis bekommt man folgende Charakterisierung der Eigenwerte:

Satz: Seien $\lambda_1 < \cdots < \lambda_n$ die Eigenwerte der selbst-adjungierten Abbildung f. Dann gilt

$$\begin{aligned}
\lambda_1 &= \min\{\varphi(x) : x \in V \setminus \{0\}\} \\
\lambda_n &= \max\{\varphi(x) : x \in V \setminus \{0\}\}.
\end{aligned}$$

Als eine Anwendung dieser Theorie bringen wir eine kurze Behandlung der Kegelschnitte in höheren Dimensionen. Das sind Teilmengen von $\mathbf{R}^n$ der Gestalt:

$$\left\{(\xi_1, \dots, \xi_n) : \sum_{i,j=1}^{n} a_{ij}\xi_i\xi_j + 2b_1\xi_1 + \cdots + 2b_n\xi_n + c = 0\right\}$$

wobei die a_{ij}, b_i und c aus $\mathbf{R}$ sind.

Zunächst schreiben wir diese Gleichung in der Gestalt

$$\big(f_A(x)|x\big) + 2(b|x) + c = 0$$

wobei $x = (\xi_1, \dots, \xi_n)$, und $b = (b_1, \dots, b_n)$.

Ohne Verlust der Allgemeinheit können wir annehmen, daß A symmetrisch ist (und daher f_A selbst-adjungiert). Denn ersetzen wir a_{ij} durch $\dfrac{(a_{ij} + a_{ji})}{2}$, so bleibt die Gleichung unverändert. Der Einfachheit wegen werden wir annehemen, daß der Kegelschnitt **zentral** ist d.h. $b = 0$ (geometrisch bedeutet das, daß der Kegelschnitt einen Punkt der Zentralsymmetrie, den wir als Urspung wählen können, besitzt). Wir können jetzt unser Ergebnis über die Diagonalisierbarkeit von selbst-adjungierten Abbildungen verwenden, um eine Klassizifierung für solche Kegelschnitte in $\mathbf{R}^n$ zu gewinnen:

Satz: Sei Q ein zentraler Kegelschnitt in $\mathbf{R}^n$ wie oben. Dann existiert eine orthonomrale Basis $(x_1, \dots, x_n)$ und reele Zahlen $\lambda_1, \dots, \lambda_n$ in $\mathbf{R}$, sodaß

$$Q = \{x = \eta_1 x_1 + \cdots + \eta_n x_n \colon \lambda_1 \eta_1^2 + \dots \lambda_n \eta_n^2 + c = 0\}$$

(Anders ausgedrückt: die Isometrie, die $(x_1, \dots, x_n)$ in die kanonische Basis $(e_1, \dots, e_n)$ abbildet, bildet Q in den Kegelschnitt

$$\left\{(\xi_1, \dots, \xi_n) \in \mathbf{R}^n \colon \lambda_1 \xi_1^2 + \cdots + \lambda_n \xi_n^2 + c = 0\right\}$$

ab.)

Beweis: Wähle für $(x_1, \dots, x_n)$ eine orthonormale Basis, bestehend aus Eigenvektoren von f_A. $\lambda_1, \dots \lambda_n$ sind dann die Eigenwerte von f_A.

Beispiel: In $\mathbf{R}^3$ hat jeder zentrale Kegelschnitt (bis auf eine Isometrie) eine der folgenden Gestalten:

$\dfrac{\xi_1^2}{a^2} + \dfrac{\xi_2^2}{b^2} + \dfrac{\xi_3^2}{c^2} + d = 0$ ($d < 0$, Ellipsoid; $d = 0$ Punkt; $d > 0, \emptyset$)

$\dfrac{\xi_1^2}{a^2} + \dfrac{\xi_2^2}{b^2} - \dfrac{\xi_3^2}{c^2} + d = 0$ ($d < 0$, einschaliges Hyporboloid; $d = 0$, Kegel; $d > 0$, zweischaliges Hyperboloid)

$\dfrac{\xi_1^2}{a^2} + \dfrac{\xi_2^2}{b^2} + d = 0$ ($d < 0$, elliptischer Zylinder; $d = 0$ Gerade; $d > 0, \emptyset$))

$\dfrac{\xi_1^2}{a^2} - \dfrac{\xi_2^2}{b^2} + d = 0$ ($d \neq 0$, hyperbolischer Zylinder; $d = 0$ zwei Ebenen)

$\dfrac{\xi_1^2}{a^2} + d = 0$ ($d < 0$, zwei parallele Ebenen; $d = 0$ eine Ebene; $d > 0, \emptyset$)

Pseudoinverse: Sei jetzt f eine lineare Abbildung zwischen den euklidischen Räumen V, V_1. Wir zerlegen V und V_1 wie folgt:

$$\begin{aligned} V &= (\operatorname{Ker} f) \oplus (\operatorname{Ker} f)^\perp \\ V_1 &= (fV) \oplus (fV)^\perp. \end{aligned}$$

Behauptung: Die Einschränkung $\tilde{f}$ von f auf $(\operatorname{Ker} f)^\perp$ ist ein Isomorphismus von $(\operatorname{Ker} f)^\perp$ auf fV.

Beweis: $\tilde{f}$ ist injektiv: Seien $x, y \in (\operatorname{Ker} f)^\perp$ mit $\tilde{f}x = \tilde{f}y$. Dann gilt:

$$f(x-y) = \tilde{f}(x) - \tilde{f}(y) = 0 \quad \text{d.h.} \quad x - y \in \operatorname{Ker} f.$$

Aus $(\operatorname{Ker} f) \cap (\operatorname{Ker} f)^\perp = \{0\}$ folgt $x = y$.

$\tilde{f}$ ist surjektiv: Sei $x \in f(V)$. Es existiert $x \in V$ mit $fx = z$. x hat eine Darstellung $x_1 + x_2$ mit $x_1 \in \operatorname{Ker} f$, $x_2 \in (\operatorname{Ker} f)^\perp$. Dann gilt

$$z = fx = f(x_1 + x_2) = f(x_1) + f(x_2) = \tilde{f}(x_2) \qquad \text{q.e.d.}$$

Wir definieren eine lineare Abbildung $g: V_1 \to V$ wie folgt:

$$g = \tilde{f}^{-1} P_{f(V)}$$

g hat die folgenden Eigenschaften:
1) $f \circ g \circ f = f$;
2) $g \circ f \circ g = g$;
3) für $y \in V_1$ ist $x = g(y)$ die "beste" Lösung der Gleichung $f(x) = y$ d.h.

$$\|f(x) - y\| \leq \|f(z) - y\| \quad (z \in V)$$

und x ist das Element mit kleinster Norm, das diese Eigenschaft hat.

g heißt die **Pseudo-Inverse** von f – geschrieben $f^\dagger$.

Um eine bequeme Darstellung der Pseudoinversen zu gewinnen, brauchen wir folgendes Ergebnis.

Satz: Sei $f: V \to V_1$ eine lineare Abbildung zwischen euklidischen Räumen. Dann gibt es orthonormale Basen $(x_1, \dots, x_n)$ bzw. $(y_1, \dots, y_n)$, sodaß f eine Matrix der Gestalt

$$\begin{bmatrix} A & 0 \\ 0 & \partial \end{bmatrix} \quad \text{wobei} \quad A = \operatorname{diag}(\mu_1, \dots, \mu_r).$$

besitzt. (Die μ_i sind von f eindeutig (bis auf Permutationen) bestimmt und heißen die **Singulärwerte** von f).

Die entsprechende Aussage für Matrizen ist: Sei A eine $m \times n$ Matrix. Dann gibt es orthonomale Matrizen Q_1 und Q_2 und $\mu_1, \dots, \mu_r$ mit

$$Q_1 A Q_2 = \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} \quad \text{wobei} \quad A = \operatorname{diag}(\mu_1, \dots, \mu_r).$$

Beweis: Die Abbildung $f^t f$ von V in V ist selbsadjungiert (denn $\big(f^t f(x)|y\big) = \big(f(x)|f(y)\big) = \big(x|f^t f(y)\big)$.

Wir wählen daher eine orthonormale Basis $(x_1, \ldots, x_n)$ wobei jedes x_i ein Eigenvektor ist, etwa mit Eigenwert λ_i. Wir numerieren die x_i so, daß eventuelle Nulleigenwerte am Schluß sind, d.h.

$$\lambda_i \neq 0 \quad (i = 1, \ldots, r), \quad \lambda_{r+1} = \cdots = \lambda_n = 0.$$

Dann gilt:

$$f(x_i) \perp f(x_j) \quad (i \neq j, 1 \leq i \leq r, \quad 1 \leq j \leq r).$$

Denn

$$\big(f(x_i)|f(x_j)\big) = \big(f^t f(x_i)|x_i\big) = \lambda_i(x_i|x_j) = 0.$$

Außerdem gilt: $\|f(x_i)\| = \sqrt{\lambda_i}$. Sei $y_i = \dfrac{f(x_i)}{\sqrt{\lambda_i}}$ $(i = 1, \ldots, r)$. Daraus folgt, daß $(y_1, \ldots, y_r)$ ein orthonormales System ist. Wir erweitern zu einer orthonormalen Basis $(y_1, \ldots, y_m)$. $(x_1, \ldots, x_n)$ bzw. $(y_1, \ldots, y_m)$ sind die gesuchten Basen.

Aus diesem Beweis sieht man, daß f genau dann injektiv ist, wenn $r = n$. Die Matrix von f ist dann

$$\begin{bmatrix} A_1 \\ 0 \end{bmatrix} \quad \text{wobei} \quad A = \operatorname{diag}(\mu_1, \ldots, \mu_r).$$

wobei die μ_i alle ungleich null sind. Die Matrix von$f^\dagger$ ist dann

$$[\, A^{-1} \; 0 \,]$$

Aber dies ist die Matrix der Abbildung $(f^t f)^{-1} f^t$. Damit haben wir folgenden Satz bewiesen:

Satz: Falls f injektiv ist, dann auch $f^t f$ und die Pseudo-Inverse von f ist die Abbildung $(f^t f)^{-1} f^t$.

Beispiel (least square fitting of data):
Gegeben sind die Punkte $(t_1, y_1), (t_1, y_2), \ldots, (t_n, y_n)$ in $\mathbf{R}^2$. Wir bestimmen c, d, sodaß die Punkte "so gut wie möglich" auf der Geraden $y = ct + d$ liegen, d.h. c, d sind "Lösungen" des Systems

$$\begin{bmatrix} t_1 & 1 \\ \vdots & \vdots \\ t_n & 1 \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

Genauer: Wir suchen c, d, sodaß der Fehler

$$(y_1 - ct_1 - d_1)^2 + \cdots + (y_n - ct_n - d_n)^2$$

minimal ist.

Dazu berechnen wir die Pseudoinverse der Matrix

$$A = \begin{bmatrix} t_1 & 1 \\ \vdots & \vdots \\ t_n & 1 \end{bmatrix}$$

Es gilt:

$$A^tA = \begin{bmatrix} n & & \bar{t} & \\ \bar{t} & t_1^2 + & \dots & +t_n^2 \end{bmatrix}$$

wobei $\bar{t} = t_1 + \cdots + t_n$.

Diese Matrix ist invertierbar, falls $n(t_1^2 + \cdots + t_n^2) \neq \bar{t}^2$. Es gilt dann

$$A^\dagger = (A^tA)^{-1}A^t$$

und die Lösung unseres Problems ist

$$\begin{bmatrix} c \\ d \end{bmatrix} = A^\dagger \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

Statt das allgemeine Ergebnis durchzurechnen, betrachten wir ein konkretes Beispiel: Wir suchen eine Approximaxion für die Daten

$$\begin{array}{cccccc} t & 0 & 1 & 2 & 3 & 5 \\ y & 0 & 2 & 3 & 4 & 6. \end{array}$$

Die Gleichung ist dann

$$\begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 5 & 1 \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 3 \\ 4 \\ 6 \end{bmatrix}$$

$$A^tA = \begin{bmatrix} 0 & 1 & 2 & 3 & 5 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 5 & 1 \end{bmatrix} = \begin{bmatrix} 39 & 11 \\ 11 & 5 \end{bmatrix}$$

$$(A^tA)^{-1} = \frac{1}{74} \begin{bmatrix} 5 & -11 \\ -11 & 39 \end{bmatrix}$$

Daher gilt:

$$\begin{bmatrix} c \\ d \end{bmatrix} = \frac{1}{74} \begin{bmatrix} 5 & -11 \\ -11 & 39 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 & 3 & 5 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \\ 3 \\ 4 \\ 6 \end{bmatrix} = \begin{bmatrix} \frac{85}{74} \\ \\ \frac{210}{37} \end{bmatrix}$$

Hermitesche Vektorräume: Für gewissen Zwecke ist es sinnvoll, diese Konzepte im Rahmen von komplexen Vektorräumen zu studieren. Insbesondere kann man dann einen rein algebraischen Beweis der Diagonalisierbarkeit von symmetrischen Matizen bringen. Wir definieren daher **einen hermiteschen Vektorraum** als einen komplexen Vektorraum V, zusammen mit einer Abbildung $(\ |\ )$ von $V \times V$ in $\mathbf{C}$, sodaß

1) $(x|x) \geq 0 \quad (x \in V)$.
2) $(x|y) = \overline{(y|x)} \quad (x, y \in V)$.
3) $(\lambda x + \mu y|z) = \lambda(x|z) + \mu(y|z) \quad (\lambda, \mu \in \mathbf{C},\ x, y, z \in V)$.
4) $(x|x) = 0$ impliziert $x = 0$.

Beispiele von solchen komplexen "Skalarprodukten" sind:

$$\big((\lambda_1, \ldots, \lambda_n), (\mu_1, \ldots, \mu_n)\big) \mapsto \lambda_1\overline{\mu}_1 + \cdots + \lambda_n\overline{\mu}_n$$

in $\mathbf{C}^n$ bzw.

$$(p, q) \mapsto \int p(t)\overline{q(t)}dt$$

in $\mathrm{Pol}_{\mathbf{C}}(n)$.

Wie in einem euklidischen Raum, so definieren wir folgende Begriffe für einen hermiteschen Raum:

a) die **Länge** $\|x\|$ eines Vektors x durch

$$\sqrt{(x|x)}$$

Es gilt dann:

$$\big|(x|y)\big| \leq \|x\|\|y\|$$
$$\|x + y\| \leq \|x\| + \|y\|.$$

(Beweise wie im Fall von euklidischen Räumen.)

b) **orthogonale Systeme bzw. orthonormale Basen:**
Wiederum können wir die Gram-Schmidtsche Methode anwenden, um die Existenz von orthonormalen Basen nachzuweisen. Daher ist jeder hermitesche Raum zu einem Raum der Gestalt $\mathbf{C}^n$ isomorph.

c) **adjungierte Abbildungen:**
Falls $f \in L(V, V_1)$, dann ist eine lineare Abbildung $g\colon V_1 \to V$ zu f adjungiert, falls

$$\big(f(x)|y\big) = \big(x|g(y)\big) \quad (x \in V,\ y \in V_1).$$

Falls f die Matrix $a = [a_{ij}]$ bezüglich **orthonormalen** Basen $(x_1, \ldots, x_n)$ für V bzw. $(y_1, \ldots, y_m)$ für V_1 hat, dann hat g die Matrix

$$A^* = [\overline{a}_{ji}]$$

bzgl. $(y_1, \ldots, y_m)$ und $(x_1, \ldots, y_n)$.

d) eine lineare Abbildung $f\colon V \to V$ heißt **hermitesch**, falls $f = f^*$ d.h. $\big(f(x)|y\big) = \big(x|f(y)\big) \quad (x, y \in V)$.
$T\colon V \to V$ heißt **unitär**, falls $(Tx|Ty) = (x|y) \quad (x, y \in V)$.

Jetzt untersuchen wir die Eigenwerte von hermiteschen und unitären Abbildungen (bzw. Matrizen).

Satz: Sei $f \in L(V)$, λ ein Eigenwert von f. Dann gilt:
1) λ ist reell, falls f hermitesch;
2) $|\lambda| = 1$ falls f unitär.

Beweis: Sei $x\,(\neq 0)$ ein entsprechender Eigenvektor. Dann gilt:
1) $\big(f(x)|x\big) = (\lambda x|x) = \lambda(x|x)$.
Aber $\big(f(x)|x\big) = \big(x|f(x)\big) = \big(\overline{f(x)|x}\big)$ und ist daher reell.

2) $\big(f(x)|f(x)\big) = (\lambda x|\lambda x) = |\lambda|^2(x|x)$.
Aber $\big(f(x)|f(x)\big) = (x|x)$ – daher gilt $|\lambda|^2 = 1$.

Satz: Seien λ_1, λ_2 verschiedene Eigenwerte einer hermiteschen Abbildung f, x_1 bzw. x_2 entsprechende Eigenvektoren. Dann gilt: $x_1 \perp x_2$.

Beweis: Aus $\big(f(x_1)|x_2\big) = \big(x_1|f(x_2)\big)$ folgt:

$$\lambda_1(x_1|x_2) = \lambda_2(x_1|x_2)$$

d.h. $(\lambda_1 - \lambda_2)(x_1|x_2) = 0$ oder $(x_1|x_2) = 0$.

Unser Hauptziel ist es jetzt, einen algebraischen Beweis der Diagonalisierbarkeit einer hermiteschen Matrix zu bringen. Dazu brauchen wir einige Lemata:

Hilfsatz: Sei $f \in L(V)$. Dann gilt

$$\operatorname{Ker} f = \operatorname{Ker} f^* f$$

.

Beweis: $f^*f(x) = 0 \Rightarrow \big(f^*f(x)|x\big) = 0$
Daher ist $\|f(x)\|^2 = \big(f(x)|f(x)\big) = \big(f^*f(x)|x\big) = 0$.

Korollar: Sei A eine $m \times n$ Matrix. Dann gilt

$$\operatorname{Rg}(A) = \operatorname{Rg}(A^*A) = \operatorname{Rg}(A^*)$$
$$\operatorname{Korang}(A) = \operatorname{Korang}(A^*A) = \operatorname{Korang}(A^*).$$

Korollar: Sei $f \in L(V)$ hermitesch. Dann gilt:

$$\operatorname{Ker} f = \operatorname{Ker}(f^r)$$

für jede natür;liche Zahl r.

Satz: Sei $f \in L(V)$ eine hermitesche Abbildung. Dann existiert eine orthonormale Basis (x_i) für V, sodaß jedes x_i ein Eigenvektor von f ist. Bezüglich dieser Basis hat f die Matrizendarstellung

$$\operatorname{diag}\{\lambda_1, \ldots, \lambda_n\}$$

wobei die λ_i die Eigenwerte von f sind (und daher alle aus $\mathbf{R}$ sind).

Beweis: Sei χ_f die charakteristische Funktion von f. Wir zerlegen sie in die Gestalt:

$$\chi_f(\lambda) = (\lambda - \lambda_1)^{n_1} \dots (\lambda - \lambda_r)^{n_r}$$

wobei die λ_i die Eigenwerte von f sind. Nach dem Satz von Cayley-Hamilton gilt:

$$(f - \lambda_1 Id)^{n_1} \dots (f - \lambda_r Id)^{n_r} = 0.$$

Wir behaupten daß

$$(f - \lambda_1 Id) \dots (f - \lambda_r Id) = 0$$

Denn aus der ersten Gleichung folgt, daß

$$\begin{aligned} [(f - \lambda_2 Id)^{n_2} \dots (f - \lambda_r Id)^{n_r}]\,(x) &\in \operatorname{Ker}(f - \lambda_1 Id)^{n_1} \\ &= \operatorname{Ker}(f - \lambda_1 Id). \qquad (x \in V) \end{aligned}$$

Daher gilt:

$$(f - \lambda_1 Id)(f - \lambda_2 Id)^{n_2} \dots (f - \lambda_r Id)^{n_r} = 0$$

Wegen Kommutativität der Faktoren können wir systematisch die Indizes der Faktoren $(f - \lambda_i Id)^{n_i}$ auf eins reduzieren.
Das bedeutet, daß die Minimalfunktion von f ein Produkt von verschiedenen linearen Faktoren ist, d.h. die Jordansche Normalform von f hat die Gestalt:

$$\operatorname{diag}(\lambda_1, \dots, \lambda_1, \lambda_2, \dots, \lambda_r, \dots, \lambda_r)$$

Sei Jetzt $V_i = \operatorname{Ker}(f - \lambda_i \text{ Id})$. Dann gilt: $V_i \perp V_j (i \neq j)$ da Eigenvektoren von f, die verschiedenen Eigenwerten entsprechen, orthogonal sind.

Daher können wir eine orthonormale Basis für V konstruieren, indem wir orthonormale Basen für jedes V_i zusammenstellen. Bezüglich dieser Basis hat f die Matrix $\operatorname{diag}(\lambda_1, \dots, \lambda_1, \lambda_2, \dots, \lambda_r)$.

In der Sprache der Matrizen kann man dieses Ergebnis forlgendermaßen ausdrücken: Sei A eine hermitesche $n \times n$ Matrix. Dann sind die Eigenwerte von A reell und es existieren n orthonormale Vektoren $X_1, \dots, X_n$, die alle Eigenvektoren von A sind. Falls U die Matrix $[X_1, \dots, X_n]$ bezeichnet, dann ist U unitär und es gilt:

$$U^{-1} A U = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$$

wobei die λ_i die (reellen) Eigenwerte von A sind.

Wir können folgende Verschärfung dieses Satzes, die später nützlich sein wird, angeben.

Satz: Seien f, g kommutierende hermitesche Abbildungen auf V. Dann besitzt V eine orthonormale Basis $(x_1, \dots, x_n)$, wobei jedes x_i ein Eigenvektor sowohl von f als auch von g ist.

Beweis: Seien $\lambda_1, \dots, \lambda_r$ die Eigenwerte von f und setze $V_i = \operatorname{Ker}(f - \lambda_i \operatorname{Id})$. Wir wissen, daß V die direkte Summe $V_1 \oplus \cdots \oplus V_r$ ist. Wir behaupten, daß $g(V_i) \subseteq V_i$ gilt für jedes i. Denn

$$\begin{aligned} x \in V_i \Rightarrow f(x) &= \lambda_i x \\ \Rightarrow (g \circ f)(x) &= \lambda_i g(x) \\ \Rightarrow (f \circ g)(x) &= \lambda_i g(x) \quad \text{wegen d. Kommutativität} \\ \text{d.h.} \quad g(x) &\in V_i. \end{aligned}$$

$g|_{V_i}$ ist jetzt eine hermitesche Abbildung auf V_i – daher hat V_i eine orthonormale Basis, die aus Eigenvektoren von g besteht. Durch Zusammensetzung der Basen für die V_i bekommen wir die gesuchte orthonormale Basis für V.

Wir betrachten jetzt eine Verallgemeinerung von hermiteschen Abbildungen – die normalen Abbildungen. $f \in L(V)$ ist **normal**, falls f und f^* kommutieren. Ein typisches Beispiel ist eine Abbildung f_A wobei A die Diagonalmatrix

$$\operatorname{diag}(\lambda_1 \dots, \lambda_n)$$

ist ($\lambda_1, \dots, \lambda_n \in \mathbf{C}$). Denn $A^* = \operatorname{diag}(\overline{\lambda}_1, \dots, \overline{\lambda}_n)$ und $A^*A = AA^*$. Wir werden jetzt sehen, daß die normalen Abbildungen genau diejenigen Abbildungen sind, die solche Darstellungen bzgl. einer geeigneten Wahl einer orthonormalen Basis besitzen.

Zunächst einige Hilfsüberlegungen:

Hilfsatz: Sei f eine lineare Abbildung auf V. f ist genau dann normal, wenn $f = g + ih$, wobei g, h hermitesch sind und g, h kommutieren.

Beweis: Zunächst die Bemerkung, daß jedes $f \in L(V)$ eine eindeutig bestimmte Darstellung $g + ih$ besitzt, wobei g und h hermitesch sind (wähle $g = \frac{1}{2}(f + f^*)$, $h = \dfrac{1}{2i}(f - f^*)$). Da $f^* = g - ih$, sieht man leicht, daß f genau dann mit f^* kommutiert, wenn g mit h kommutiert.

Satz: Sei f eine normale Abbildung. Dann besitzt V eine orthonormale Basis (x_i) derart, daß jedes x_i ein Eigenvektor von f ist. Bezüglich dieser Basis hat f die Darstellung

$$\operatorname{diag}(\lambda_1, \dots, \lambda_n)$$

wobei die λ_i die Eigenwerte von f sind.

Beweis: Sei $f = g + ih$, wobei g, h hermitesch sind. Wir wählen eine orthonormale Basis $x_1, \dots, x_n$, sodaß die x_i Eigenvektoren sowohl von g als auch von h sind, etwa

$$g(x_i) = \alpha_i x_i, \quad h(x_i) = \beta_i x_i.$$

Dann gilt: $f(x_i) = (g + ih)(x_i) = (\alpha_i + i\beta_i)x_i$. q.e.d.

Korollar: Sei $f \in L(V)$ eine unitäre Abbildung. Dann existiert eine orthonormale Basis (x_i) für V, sodaß die Matrix von f die Gestalt

$$\mathrm{diag}\{\cos\theta_1 + i\sin\theta_1, \ldots, \cos\theta_n + i\sin\theta_n\}$$

hat.

Der Beweis besteht aus den zwei einfachen Bemerkungen:
a) Eine unitäre Abbildung ist normal.
b) Die Eigenwerte einer unitären Abbildung haben Absolutbetrag 1 und daher die Gestalt

$$(\cos\theta + i\sin\theta) \quad (\theta \in [0, 2\pi[).$$

Mit Hilfe dieses Satzes können wir eine Verallgemeinerung der Klassifizierung von Isometrien auf $\mathbf{R}^2$ und $\mathbf{R}^3$ gewinnen.

Satz: Sei $f\colon \mathbf{R}^n \to \mathbf{R}^n$ eine lineare Isometrie. Dann existiert eine orthonormale Basis (x_i) von $\mathbf{R}^n$, sodaß die Matrix von f die Blockgestalt

$$\mathrm{diag}(I_r - I_s\, D_{\theta_1} \ldots D_{\theta_t})$$

besitzt. (D.h. f ist die Zusammensetzung von s Spiegelungen und t Drehungen.

N.B. Wenn wir dieses Ergebnis auf $\mathbf{R}^2$, bzw. $\mathbf{R}^3$ anwenden, so bekommen wir wieder die Klassifizierung der zwei- bzw. dreidimensionalen linearen Isometrien (vgl. Kap. II).