Teil 3. Euklid'sche Vektorräume. Unitäre Vektorräume

3A. Bilinearformen.

Sei K ein Körper, sei V ein K-Vektorraum. Eine *Bilinearform* φ auf V ist eine Abbildung $\varphi\colon V \times V \to K$ mit den folgenden beiden Eigeschaften:

(1) Seien $v, w_1, w_2 \in V$ und $\alpha_1, \alpha_2 \in K$, so ist

$$\varphi(v, \alpha_1 w_1 + \alpha_2 w_2) = \alpha_1 \varphi(v, w_1) + \alpha_2 \varphi(v, w_2).$$

(2) Seien $v_1, v_2, w \in V$ und $\alpha_1, \alpha_2 \in K$, so ist

$$\varphi(\alpha_1 v_1 + \alpha_2 v_2, w) = \alpha_1 \varphi(v_1, w) + \alpha_2 \varphi(v_2, w).$$

(Bedingung (1) besagt, daß für jedes $v \in V$ die Abbildung $\varphi(v, -)\colon V \to K$ linear ist; Bedingung (2) besagt, daß für jedes $w \in V$ die Abbildung $\varphi(-, w)\colon V \to K$ linear ist; daher der Name "bilinear"). Man sagt, daß die Bilinearform $\varphi\colon V \times V \to K$ *symmetrisch* ist, falls $\varphi(v, w) = \varphi(w, v)$ für alle $v, w \in V$ gilt.

Ist $\varphi\colon V \times V \to K$ bilinear, ist $v_1, \dots, v_n$ eine Basis von V und bilden wir $a_{ij} = \varphi(v_i, v_j)$, so erhalten wir eine $(n \times n)$-Matrix $A = (a_{ij})_{ij}$ und es gilt: Sind v, w beliebige Vektoren von V, etwa $v = \sum_i \alpha_i v_i$ und $w = \sum_i \beta_i v_i$, so ist

$$\varphi(v, w) = \begin{bmatrix} \alpha_1 & \dots & \alpha_n \end{bmatrix} \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix}.$$

Ist $A = (a_{ij})_{ij}$ eine $(n \times n)$-Matrix mit Koeffizienten in K, so definieren wir eine Abbildung $\varphi_A\colon K^n \times K^n \to K$ durch

$$\varphi_A(v, w) := {}^t v \cdot A \cdot w;$$

man rechnet sofort nach, daß φ_A eine Bilinearform auf K^n ist. Und, wie wir gerade gesehen haben, erhält man jede Bilinearform auf dem Vektorraum K^n auf diese Weise. Beachte: Genau dann ist φ_A eine symmetrische Bilinearform, wenn A eine symmetrische Matrix ist.

3B. Euklid'sche Vektorräume.

Sei jetzt $K = \mathbb{R}$. Sei V ein $\mathbb{R}$-Vektorraum. Eine Bilinearform φ auf V heißt *positiv definit,* falls $\varphi(v, v) > 0$ für alle $v \neq 0$ gilt. Ist V ein reeller Vektorraum mit einer positiv definiten symmetrischen Bilinearform φ, so nennt man (V, φ) einen *euklid'schen Vektorraum,* statt $\varphi(v, w)$ schreibt man häufig einfach $\langle v, w \rangle$, und man nennt $\langle -, - \rangle$ oder eben φ das *innere Produkt.*

Achtung. Bisher haben wir spitze Klammern $\langle v_1, \dots, v_n \rangle$ verwendet, um den von den Vektoren $v_1, \dots, v_n$ erzeugten Unterraum zu bezeichnen. Im Fall eines euklid'schen Vektorraums würde dies (zumindest bei

Paaren von Vektoren) zu Unklarheiten führen. Im Folgenden werden wir den von $v_1, \ldots, v_n$ erzeugten Unterraum mit $\mathrm{span}(v_1, \ldots, v_n)$ bezeichnen!

Beispiel 1. Sei $A = (a_{ij})_{ij}$ eine Diagonalmatrix (also $a_{ij} = 0$ für $i \neq j$). Genau dann ist φ_A positiv definit, wenn $a_{ii} > 0$ für alle i gilt. Denn:

$$\varphi_A(\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}, \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}) = \sum_{i=1}^{n} a_{ii} c_i^2.$$

Ist also A eine derartige Matrix, so ist $(\mathbb{R}^n, \varphi_A)$ ein euklid'scher Vektorraum. Insbesondere ist der Fall $A = I_n$ wichtig, man nennt $(\mathbb{R}^n, \varphi_{I_n})$ den *kanonischen n-dimensionalen euklid'schen Vektorraum,* und φ_{I_n} das (kanonische) *innere Produkt auf* $\mathbb{R}^n$.

Beispiel 2. Seien $a < b$ reelle Zahlen. Sei $C([a, b])$ die Menge der stetigen reellwertigen Funktionen $[a, b] \to \mathbb{R}$. Definiere auf $C([a, b])$ eine Bilinearform

$$\langle f, g \rangle = \int_a^b f(t) g(t) \,\mathrm{d}t \quad \text{für} \quad f, g \in C([a, b]).$$

Dies ist eine positiv definite symmetrische Bilinearform.

Sei $V = (V, \langle -, - \rangle)$ ein euklid'scher Vektorraum.

Norm. Ist $v \in V$, so setze $||v|| = \sqrt{\langle v, v \rangle}$. Man nennt dies die *Norm* des Vektors v (oder auch seine "Länge"). Man nennt v *normiert,* falls $||v|| = 1$ gilt. Wir werden uns später noch genauer mit Rechenregeln, die die Norm betreffen, beschäftigen. Im Augenblick ist vor allem die folgende Regel wichtig:

$$||\gamma v|| = |\gamma| \cdot ||v|| \quad \text{für alle} \quad v \in V \quad \text{und} \quad \gamma \in \mathbb{R}.$$

Die Norm ist mit Hilfe der Bilinearform $\langle -, - \rangle$ definiert. Mit Hilfe der Norm dreier Vektoren kann man umgekehrt das innere Produkt berechnen, denn es gilt:

$$\langle v, w \rangle = \frac{1}{2}\big(||v + w||^2 - ||v||^2 - ||w||^2\big)$$

Das Normieren von Vektoren: *Ist $v \neq 0$, so gibt es genau* **zwei** *normierte Vektoren v' mit* $\mathrm{span}(v) = \mathrm{span}(v')$, *nämlich die beiden Vektoren* $\frac{1}{||v||}v$ *und* $\frac{-1}{||v||}v$. Beweis: Ganz allgemein gilt: $||\lambda v|| = |\lambda| \cdot ||v||$, für $\lambda \in \mathbb{R}$ und $v \in V$. Also ist $||\lambda v|| = 1$ genau dann, wenn $|\lambda| = \frac{1}{||v||}$ gilt, also wenn $\lambda = \frac{1}{||v||}$ oder $\lambda = -\frac{1}{||v||}$. Wenn man sagt, daß man *einen Vektor v normiert,* so meint man, daß man v durch $\frac{1}{||v||}v$ ersetzt.

Orthogonalität. Zwei Vektoren v, w in V mit $\langle v, w \rangle = 0$ heißen *orthonal. Sind die Vektoren $v_1, \ldots, v_m$ paarweise orthogonal und alle von Null verschieden, und ist* *eine Linearkombination* $w = \sum_{i=1}^m c_i v_i$ *gegeben, so erhält man den Koeffizienten* c_j *durch*

$$c_j = \frac{\langle w, v_j\rangle}{\langle v_j, v_j\rangle};$$

insbesondere sieht man: *diese Folge* $v_1, \ldots, v_m$ *ist linear unabhängig.* Ist hier v_j ein normierter Vektor, so ist der Nenner 1, also ist dann $c_j = \langle w, v_j\rangle$. (Ist u ein normierter Vektor, so nennt man $\langle w, u\rangle u$ die *Projektion des Vektors* w *in Richtung* u.) Man nennt $v_1, \ldots, v_n$ eine *Orthonormalfolge* in V, falls diese Vektoren normiert und paarweise orthogonal sind. Ist diese Folge zusätzlich ein Erzeugendensystem von V (also eine Basis), so spricht man von einer *Orthonormalbasis.*

Satz. *Jeder endlich-dimensionale euklid'sche Vektorraum* V *besitzt eine Orthonormalbasis.* Stärker: *Jede Orthonormalbasis eines Unterraums von* V *läßt sich zu einer Orthonalbasis von* V *verlängern.*

Beweis: Sei U ein Unterraum von V. Sei $u_1, \ldots, u_m$ eine Orthonormalbasis von U. Wir können annehmen $U \subset V$, wähle $v \in V \setminus U$. Bilde

$$v' = v - \sum_{i=1}^n \langle v, u_i\rangle u_i.$$

Es ist $\langle v', u_j\rangle = 0$ für $1 \le j \le n$, denn

$$\begin{aligned}\langle v', u_j\rangle &= \langle v - \sum_{i=1}^n \langle v, u_i\rangle u_i, u_j\rangle \\ &= \langle v, u_j\rangle - \sum_{i=1}^n \langle v, u_i\rangle\langle u_i, u_j\rangle \\ &= \langle v, u_j\rangle - \langle v, u_j\rangle\langle u_j, u_j\rangle = 0\end{aligned}$$

Nun müssen wir v' nur noch normieren: Setze $u_{m+1} = \frac{1}{||v'||}v'$. Natürlich gilt:

$$\operatorname{span}(u_1, \ldots, u_m, v) = \operatorname{span}(u_1, \ldots, u_m, v') = \operatorname{span}(u_1, \ldots, u_m, u_{m+1})$$

Mit Induktion erhalten wir eine Orthonormalbasis von ganz V, die die gegebene Orthonormalbasis von U fortsetzt.

Dieses Verfahren wird das **Gram-Schmidt'sche Orthonormalisierungsverfahren** genannt. Ist also eine Basis $v_1, \ldots, v_n$ von V gegeben, so erhält man eine Orthonormalbasis $u_1, \ldots, u_n$, so daß für alle $1 \le m \le n$ gilt:

$$(*) \qquad \operatorname{span}(v_1, \ldots, v_m) = \operatorname{span}(u_1, \ldots, u_m)$$

Das Verfahren erfodert jeweils zwei Schritte:

Sei schon eine Orthonormalfolge $u_1, \ldots, u_m$ mit $(*)$ gegeben.

- **Orthogonalisieren.** Setze

$$v' = v - \sum_{i=1}^n \langle v_{m+1}, u_i\rangle u_i.$$

- **Normieren.** Setze
$$u_{m+1} = \frac{1}{||v'||}v'.$$
Dann ist $u_1, \ldots, u_{m+1}$ eine Orthonormalfolge mit
$$\mathrm{span}(v_1, \ldots, v_{m+1}) = \mathrm{span}(u_1, \ldots, u_{m+1}).$$

Die Existenz einer Orthonormalbasis in einem endlich-dimensionalen euklid'schen Vektorraum bedeutet, dass jeder n-dimensionale euklid'sche Vektorraum zum $\mathbb{R}^n$ mit dem kanonischen inneren Produkt "isomorph" ist. Sind zwei euklid'sche Vektorräume $(V, \langle -,-\rangle_V)$ und $(W, \langle -,-\rangle_W)$ gegeben, so nennt man eine Abbildung $f\colon V \to W$ einen *Isomorphismus* (von euklid'schen Vektorräumen) wenn f bijektiv und linear (also ein Vekotrraum-Isomorphismus) ist und zusätzlich gilt
$$\langle f(v_1), f(v_2)\rangle_W = \langle v_1, v_2\rangle_V$$
für alle $v_1, v_2 \in V$ (man sagt, dass f mit dem inneren Produkt [genauer: den beiden inneren Produkten $\langle -,-\rangle_V$ und $\langle -,-\rangle_W$] verträglich ist).

Sei $V = (V, \langle -,-\rangle)$ ein euklid'scher Raum mit einer Orthonormalbasis $\mathcal{V} = (v_1, \ldots, v_n)$. Dann liefert die Zuordnung $e_i \mapsto v_i$ einen Isomorphismus f von $\mathbb{R}^n$ mit dem kanonischen inneren Produkt auf $V, \langle -,-\rangle)$ (mit $f(\sum \lambda e_i) = \sum \lambda_i v_i$).

Beweis: Da $\mathcal{V}$ eine Basis von V ist, ist natürlich f ein Vektorraum-Isomorphism. Dass f mit dem inneren Produkt verträglich ist, folgt aus der Orthonormalität der Basis $\mathcal{V}$.

Ist U ein Unterraum, so bezeichnet man mit $U^\perp$ die Menge der Vektoren $v \in V$ mit $\langle u, v\rangle = 0$ für alle $u \in U$. *Dies ist ein Unterraum, und es ist $U \cap U^\perp = 0$.* Wir werden gleich sehen, daß für endlich-dimensionales U auch $U + U^\perp = V$ gilt: Wähle eine Orthonormalbasis $u_1, \ldots, u_m$ von U. Dann ist
$$v' = v - \sum\nolimits_{i=1}^{n} \langle v, u_i\rangle u_i$$
in $U^\perp$ und es gilt $v \in U + v' \subseteq U + U^\perp$. Also ist auch $V = U + U^\perp$. (Wir können also $V = U \oplus U^\perp$ schreiben).

3C. Orthogonale Endomorphismen.

Ist $(V, \langle -,-\rangle)$ ein euklid'scher Vektorraum, so nennt man einen Endomorphismus $f\colon V \to V$ *orthogonal*, falls
$$\langle f(v), f(w)\rangle = \langle v, w\rangle \quad \text{für alle} \quad v, w \in V$$
gilt. *Genau dann ist f orthogonal, wenn $||f(v)|| = ||v||$ für alle $v \in V$ gilt,* wenn also f "normerhaltend" ist. (Beweis: Es sollte klar sein, daß ein orthogonaler Endomorphismus normerhaltend ist. Umgekehrt haben wir gesehen, daß man das innere Produkt mit Hilfe der Norm von Vektoren berechnen kann).

Typische Beispiele orthogonaler Endomorphismen des kanonischen euklid'schen Vektorraums $\mathbb{R}^2$ sind die Drehungen $f_A\colon \mathbb{R}^2 \to \mathbb{R}^2$ mit

$$A = \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix}$$

und $t \in \mathbb{R}$; wegen der Periodizität von Sinus und Cosinus reicht es, $0 \le t < 2\pi$ zu betrachten. Man nennt A die *Drehmatrix zum Winkel* t. Ist $t = 0$, so ist $A = I_2$ die Einheitsmatrix, ist $t = \pi$, so ist $A = -I_2$ ebenfalls eine Skalarmatrix (dies ist gerade die Punktspiegelung am Urspung).

Satz (Hauptsatz über orthogonale Endomorphismen). *Sei V ein endlich-dimensionaler euklid'scher Vektorraum. Sei $f\colon V \to V$ ein orthogonaler Endomorphismus von V. Dann besitzt V eine Orthonormalbasis, so daß f bezüglich dieser Basis eine Matrizendarstellung der Form $\bigoplus_{i=1}^m A_i$ hat, wobei die Matrizen A_i (1×1)-Matrizen der Form $[1]$ oder $[-1]$ oder (2×2)-Drehmatrizen zu Winkeln $0 < t_i < \pi$ oder $\pi < t_i < 2\pi$ sind.*

Bevor wir mit dem Beweis beginnen, ist an den Fundamentalsatz der Algebra zu erinnern.

Nachtrag zum Fundamentalsatz der Algebra. Der Fundamentalsatz der Algebra impliziert, daß der Grad eines jeden irreduziblen Polynoms $h \in \mathbb{R}[T]$ 1 oder 2 ist. Eine Folgerung daraus ist:

Lemma. *Sei $V \neq 0$ ein endlich-dimensionaler reeller Vektorraum. Sei $f\colon V \to V$ ein Endomorphismus. Besitzt f keine Eigenvektoren, so gibt es in V einen zweidimensionalen f-invarianten Unterraum.*

Beweis: Sei $\chi_f(T)$ das charakteristische Polynom von f, zerlege $\chi_f = h_t \cdots h_1$ mit normierten irreduziblen Polynomen h_i. Nach dem Fundamentalsatz der Algebra haben diese Polynome h_i den Grad 1 oder 2 (und Grad 1 kann hier gar nicht auftreten, da wir voraussetzen, daß f keine Eigenvektoren besitzt). Wähle $v \neq 0$ in V. Bilde $v_0 = v$, $v_i = h_i(f)(v_{i-1})$ für $1 \le i \le t$. Es ist $v_0 \neq 0$, aber $v_n = h_t(f) \cdots h_1(f)(v) = \chi_f(f)(v) = 0$ (nach dem Satz von Cayley-Hamilton). Also gibt es ein $1 \le i \le n$ mit $v_{i-1} \neq 0$, und $v_i = 0$. Sei $w = v_{i-1}$. Betrachte den Unterraum U, der von w und $f(w)$ erzeugt wird. Wir zeigen, daß U f-invariant ist. Sei $h_i(T) = T^2 + cT + d$ mit reellen Zahlen c, d. Nach Konstruktion ist

$$0 = h_i(f)(w) = (f^2 + cf + d \cdot 1)(w) = f^2(w) + cf(w) + dw,$$

und demnach ist $f^2(w) = -cf(w) - dw$ eine Linearkombination von $f(w)$ und w, also wieder ein Element von U. Andererseits ist U auf jeden Fall von Null verschieden, denn $w \neq 0$. Wäre U eindimensional, so hätten wir einen eindimensionalen f-invarianten Unterraum, also

Eigenvektoren. Da wir dies ausgeschlossen haben, muß U zweidimensional sein.

Nun zum Beweis des Hauptsatzes über orthogonale Endomorphismen. Wir beginnen mit drei Vorbemerkungen:

(1) *Ist γ Eigenwert von f, so ist $\gamma = 1$ oder -1.* Beweis: Sei v zugehöriger Eigenvektor, also $v \neq 0$ und $f(v) = \gamma \cdot v$. Es ist

$$||v|| = ||f(v)|| = ||\gamma \cdot v|| = |\gamma| \cdot ||v||.$$

Wegen $||v|| \neq 0$ folgt $1 = |\gamma|$, also $\gamma = 1$ oder $\gamma = -1$.

(2) Insbesondere sehen wir, daß ein orthogonaler Endomorphismus injektiv ist (denn 0 ist kein Eigenwert). Ist also f ein orthogonaler Endomorphismus eines endlichdimensionalen Vektorraums (und das wird hier vorausgesetzt), so ist f auch surjektiv.

(3) *Ist U ein f-invarianter Unterraum von V, so ist auch $U^{\perp}$ f-invariant.* Beweis: Sei $v \in U^{\perp}$. Zu zeigen ist: $f(v) \in U^{\perp}$. Sei also $u \in U$. Da die Einschränkung von f auf U ein orthogonaler Endomorphismus von U ist, ist diese Einschränkung surjektiv: es gibt also $u' \in U$ mit $f(u') = u$. Also

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

Nun der eigentliche Beweis, mit Induktion nach der Dimension $\dim V = n$.

Wir betrachten zuerst den Fall, daß V keine f-invarianten Unterräume außer 0 und V besitzt (dies schließt insbesondere den Fall $n = 1$ ein). Nach dem Fundamentalsatz der Algebra ist demnach $n = 1$ oder $n = 2$.

Ist $n = 1$, so ist jeder von Null verschiedene Vektor $v \in V$ ein Eigenvektor, die zugehörige Matrizendarstellung von f ist $[1]$ oder $[-1]$, weil 1 und –1 die einzigen möglichen Eigenwerte sind.

Sei nun $n = 2$. Da wir voraussetzen, daß V keine f-invarianten Unterräume hat, kann f keine Eigenvektoren besitzen. Sei v_1, v_2 eine beliebige Orthonormalbasis von V. Sei $f(v_1) = av_1 + bv_2$ mit $a, b \in \mathbb{R}$. Wegen

$$1 = ||v_1|| = ||f(v_1)|| = ||av_1 + bv_2|| = \sqrt{a^2 + b^2}$$

folgt also $a^2 + b^2 = 1$. Da v_1, v_2 orthogonal sind, müssen auch die Vektoren $f(v_1)$, $f(v_2)$ orthogonal sein. Und natürlich ist mit v_2 auch $f(v_2)$ normiert. In V gibt es nur zwei normierte Vektoren, die zu $f(v_1) = av_1 + bv_2$ orthogonal sind, nämlich $-bv_1 + av_2$ und $bv_1 - av_2$. Ist $f(v_2) = bv_1 - av_2$, so rechnet man sofort nach, daß das charakteristische Polynom von f das Polynom $(T + a)(T - a) - b^2 = T^2 - 1 = (T-1)(T+1)$ ist, aber dies impliziert, daß f Eigenvektoren besitzt, unmöglich. Also ist $f(v_2) = -bv_1 + av_2$, und demnach ist die darstellende Matrix eine Drehmatrix mit Winkel wie angegeben.

Wir betrachten nun den Fall, daß V einen f-invarianten Unterraum U besitzt mit $0 \subset U \subset V$. Dann ist nach (3) auch $U^\perp$ f-invariant. Betrachten wir die Einschränkung von f auf U, so gibt es (nach Induktion) eine Orthonormalbasis $u_1, \dots, u_m$ bezüglich derer $f|U$ die angegebene Form hat; die darstellende Matrix sei mit B bezeichnet. Betrachten wir entsprechend die Einschränkung von f auf $U^\perp$, so gibt es (nach Induktion) eine Orthonormalbasis $u_{m+1}, \dots, u_n$ bezüglich derer $f|U^\perp$ die angegebene Form hat; hier sei die Matrix mit C bezeichnet. Bezüglich der Basis $u_1, \dots, u_n$ hat f die Matrizen-Darstellung

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

und mit B und C hat auch A die gewünschte Form.

3D. Unitäre Vektorräume, unitäre Endomorphismen.

Wir betrachten nun einen komplexen Vektorraum V. Zur Erinnerung: Ist $\gamma \in \mathbb{C}$, so schreibt man $\overline{\gamma}$ für die zu γ *konjugiert-komplexe* Zahl (ist $\gamma = a + bi$ mit $a, b \in \mathbb{R}$, so ist $\overline{\gamma} = a - bi$). Die Zuordnung $\gamma \mapsto \overline{\gamma}$ ist ein Körperautomorphismus von $\mathbb{C}$ (es gilt also $\overline{\gamma_1 + \gamma_2} = \overline{\gamma_1} + \overline{\gamma_2}$ und $\overline{\gamma_1 \cdot \gamma_2} = \overline{\gamma_1} \cdot \overline{\gamma_2}$).

Eine *hermite'sche Form* φ auf V ist eine Abbildung $\varphi\colon V \times V \to \mathbb{C}$ mit den folgenden beiden Eigeschaften:

(1) Seien $v_1, v_2, w \in V$ und $\alpha_1, \alpha_2 \in \mathbb{C}$. Dann ist

$$\varphi(\alpha_1 v_1 + \alpha_2 v_2, w) = \alpha_1 \varphi(v_1, w) + \alpha_2 \varphi(v_2, w).$$

(2) Für $v, w \in V$ gilt

$$\varphi(w, v) = \overline{\varphi(v, w)}.$$

Eine hermite'sche Form ist also in der ersten Variablen linear, in der zweiten ist sie wegen (1) und (2) immer noch additiv $\varphi(v, w_1 + w_2) = \varphi(v, w_1) + \varphi(v, w_2)$ und es gilt $\varphi(v, \gamma w) = \overline{\gamma}\varphi(v, w)$, für $v, w, w_1, w_2 \in V, \gamma \in \mathbb{C}$.

Typisches Beispiel: Sei $H \in M(n \times n, \mathbb{C})$. Man nennt H *hermite'sche Matrix,* falls gilt ${}^t\overline{H} = H$ (d.h.: Es ist $H = (h_{ij})$ mit $h_{ji} = \overline{h_{ij}}$ für alle i, j). Setze

$$\varphi(v, w) = {}^t v \cdot A \cdot \overline{w}.$$

Dann ist φ eine hermite'sche Form auf $\mathbb{C}^n$.

Eine hermite'sche Form φ auf V heißt *positiv definit*, falls für alle von Null verschiedenen $v \in V$ gilt $\varphi(v, v) > 0$ (dies macht Sinn als Ungleichung in $\mathbb{R}$, denn wegen der Eigenschaft (2) ist $\varphi(v, v)$ eine reelle Zahl!). Ist φ eine positiv definite hermite'sche Form auf V, so nennt man $V = (V, \varphi)$ einen *unitären Raum* (und man schreibt meist $\langle v, w \rangle$ statt $\varphi(v, w)$ und nennt auch dies wieder ein inneres Produkt).

Sei nun V ein unitärer Raum. Man nennt $v_1, \ldots, v_n$ eine *Orthonormalbasis* von V, falls gilt $\langle v_i, v_i \rangle = 1$, und $\langle v_i, v_j \rangle = 0$ für alle $i \neq j$. Ist U ein Unterraum, so setzt man $U^\perp = \{v \in V \mid \langle u, v \rangle = 0 \quad \text{für alle} \quad u \in U\}$ und nennt dies das *orthogonale Komplement* von U. Es ist $U \cap U^\perp = 0$. Ist U endlich-dimensional, so ist $U + U^\perp = V$.

Ganz analog wie im Fall eines euklid'schen Vektorraums zeigt man für einen unitären Vektorraum V:

Orthonormalisierung. Sei U ein Unterraum von V. Sei $u_1, \ldots, u_m$ eine Orthonormalbasis von U. Ist $v \in V$, so setze

$$v' = v - \sum_{i=1}^{m} \langle v, u_i \rangle u_i.$$

Es ist $v' \in U^\perp$. Ist $v \in U$, so ist $v' = 0$. Ist $v \notin U$, so sind die Vektoren $u_1, \ldots, u_m, v'$ linear unabhängig. Daraus folgt: Jeder endlich-dimensionale unitäre Vektorraum besitzt eine Orthonormalbasis. Ist V ein endlich-dimensionaler unitärer Vektorraum und U ein Unterraum von V, so läßt sich jede Orthonormalbasis von U zu einer Orthonormalbasis von V fortsetzen.

Ist $f\colon V \to V$ ein Endomorphismus des unitären Vektorraums V, so nennt man f *unitär*, falls gilt

$$\langle f(v), f(w) \rangle = \langle v, w \rangle \quad \text{für alle} \quad v, w \in V.$$

Eigenwerte. Sei f unitärer Endomorphismus des unitären Vektorraums V. Ist γ ein Eigenwert von f, so ist $|\gamma| = 1$. Insbesondere ist die Abbildung f injektiv. Ist V endlich-dimensional, so ist f bijektiv und das charakteristische Polynom von f hat die Form $(T - \gamma_1) \cdots (T - \gamma_n)$ mit $|\gamma_i| = 1$ für $1 \leq i \leq n$.

Orthogonales Komplement. Sei f unitärer Endomorphismus des unitären Vektorraums V. Ist U ein endlich-dimensionaler Unterraum von V, der f-invariant ist, so ist auch $U^\perp$ f-invariant.

Diagonalisierbarkeit. (Hauptsatz über unitäre Endomorphismen.) *Sei V ein endlich-dimensionaler unitärer Vektorraum und $f\colon V \to V$ ein unitärer Endomorphismus. Dann gibt es eine Orthonormalbasis von V, so daß die zugehörige Matrizendarstellung von f eine Diagonalmatrix ist, deren Diagonalkoeffizienten den Betrag 1 haben.*

3E. Orthogonale Matrizen, unitäre Matrizen.

Orthogonale Matrizen.

Eine reelle $(n \times n)$-Matrix A heißt *orthogonal,* falls A invertierbar ist mit

$$A^{-1} = {}^tA.$$

Die Menge der orthogonalen $(n \times n)$-Matrizen bildet eine Untergruppe $\mathcal{O}(n)$ der Gruppe $\mathrm{GL}(n, \mathbb{R})$.

Ist A eine orthogonale Matrix, so ist $|\det A| = 1$. (Denn es ist $1 = \det A^{-1}A = \det({}^tA \cdot A) = (\det A)^2$.). Ist $\det A = 1$, so nennt man A *orientierungserhaltend* (oder eigentlich), ist $\det A = -1$, so nennt man A *orientierungsumkehrend* (oder uneigentlich).

(1) Genau dann ist A orthogonal, wenn die Spalten von A eine Orthonormalbasis des kanonischen euklid'schen Raums $\mathbb{R}^n$ bilden, und auch genau dann, wenn die Transponierten der Zeilen von A eine Orthonormalbasis des kanonischen euklid'schen Raums $\mathbb{R}^n$ bilden.

(1′) Folgerung: Ist V ein n-dimensionaler euklid'scher Raum und sind zwei Orthonormalbasen $\mathcal{A}$ und $\mathcal{B}$ gegeben, so ist die Basiswechselmatrix $T_{\mathcal{B}}^{\mathcal{A}}$ eine orthogonale Matrix.

(2) Sei A eine reelle $(n \times n)$-Matrix. Genau dann ist A orthogonal, wenn die Abbildung $f_A \colon \mathbb{R}^n \to \mathbb{R}^n$ ein orthogonaler Endomorphismus (bezüglich des kanonischen inneren Produkts auf $\mathbb{R}^n$) ist.

(3) Ist V ein euklid'scher Vektorraum mit Orthonormalbasis $v_1, \dots, v_n$, und ist $f \colon V \to V$ orthogonaler Endomorphismus, so ist die Matrizendarstellung von f bezüglich dieser Basis eine orthogonale Matrix.

(4) **Umformulierung des Haupsatzes über orthogonale Endomorphismen.** *Ist A eine orthogonale Matrix, so gibt es eine orthogonale Matrix P mit $P^{-1}AP = \bigoplus_{i=1}^m C_i$, wobei die Matrizen C_i (1×1)-Matrizen $[1]$ oder $[-1]$ oder (2×2)-Drehmatrizen zu Winkeln $0 < t_i < \pi$ oder $\pi < t_i < 2\pi$ sind.*

Unitäre Matrizen.

Sei A eine komplexe $(n \times n)$-Matrix. Die Matrix A heißt *unitär,* falls A invertierbar ist mit

$$A^{-1} = {}^t\overline{A}.$$

Die Menge der unitären $(n \times n)$-Matrizen bildet eine Untergruppe $\mathcal{U}(n)$ der Gruppe $\mathrm{GL}(n, \mathbb{C})$.

Ist A eine unitäre Matrix, so ist $|\det A| = 1$. (Denn es ist $\det \overline{A} = \overline{\det A}$, also $1 = \det A^{-1}A = \det({}^t\overline{A} \cdot A) = \overline{d} \cdot d$ mit $d = \det A$).

(1) Genau dann ist A unitär, wenn die Spalten von A eine Orthonormalbasis des kanonischen hermite'schen Raums $\mathbb{C}^n$ bilden, und auch genau dann, wenn die Transponierten der Zeilen von A eine Orthonormalbasis des kanonischen hermite'schen Raums $\mathbb{C}^n$ bilden.

(1′) Folgerung: Ist V ein n-dimensionaler hermite'scher Raum und sind zwei Orthonormalbasen $\mathcal{A}$ und $\mathcal{B}$ gegeben, so ist die Basiswechselmatrix $T_{\mathcal{B}}^{\mathcal{A}}$ eine unitäre Matrix.

(2) Sei A eine komplexe $(n \times n)$-Matrix. Genau dann ist A unitär, wenn die Abbildung $f_A \colon \mathbb{C}^n \to \mathbb{C}^n$ ein unitärer Endomorphismus (bezüglich des kanonischen inneren Produkts auf $\mathbb{C}^n$) ist.

(3) Ist V ein hermite'scher Vektorraum mit Orthonormalbasis $v_1, \ldots, v_n$, und ist $f\colon V \to V$ unitärer Endomorphismus, so ist die Matrizendarstellung von f bezüglich dieser Basis eine unitäre Matrix.

(4) **Umformulierung des Haupsatzes über unitäre Endomorphismen.** *Ist A eine unitäre Matrix, so gibt es eine unitäre Matrix P, so daß $P^{-1}AP$ eine Diagonalmatrix ist, deren Diagonalkoeffizienten den Betrag 1 haben.*

Orthogonale Matrizen sind unitäre Matrizen:

Jede orthogonale Matrix A ist auch eine unitäre Matrix (jede reelle Matrix kann natürlich als komplexe Matrix aufgefaßt werden). Wie wir wissen, gibt es eine orthogonale Matrix P mit $P^{-1}AP = \bigoplus_{i=1}^m C_i$, wobei die Matrizen C_i (1×1)-Matrizen $[1]$ oder $[-1]$ oder (2×2)-Drehmatrizen $D(\alpha_i)$ zu Winkeln $0 < \alpha_i < \pi$ oder $\pi < \alpha_i < 2\pi$ sind. *Die Drehmatrix* $D(\alpha) = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$ *ist (als komplexe Matrix) ähnlich zur Diagonalmatrix mit den Diagonalkoeffizienten* $\cos(\alpha) + i\sin(\alpha)$ *und* $\cos(\alpha) - i\sin(\alpha)$; es ist nämlich

$$\underbrace{\begin{bmatrix} \frac{-i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-i}{\sqrt{2}} \end{bmatrix}}_{{}^t\overline{P}} \cdot \underbrace{\begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}}_{D(\alpha)} \cdot \underbrace{\begin{bmatrix} \frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \end{bmatrix}}_{P} = \begin{bmatrix} \cos\alpha + i\sin\alpha & 0 \\ 0 & \cos\alpha - i\sin\alpha \end{bmatrix}$$

dabei ist $P = \begin{bmatrix} \frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \end{bmatrix}$ eine unitäre Matrix.

3F. Cauchy-Schwarz.

Wir betrachten nun euklid'sche und unitäre Vektorräume. Wir scheiben K für den jeweiligen Grundkörper: Im euklid'schen Fall ist $K = \mathbb{R}$, im unitären Fall ist $K = \mathbb{C}$.

Satz (Cauchy-Schwarz). *Sei V ein euklid'scher oder unitärer Vektorraum. Dann gilt für alle $v, w \in V$*

$$|\langle v, w\rangle| \leq ||v|| \cdot ||w||.$$

Das Gleichheitszeichen gilt genau dann, wenn die Vektoren v, w linear abhängig sind.

Beweis: Ist $w = 0$, so wird $0 \leq 0$ behauptet. Sei also $w \neq 0$. Sei $c = \frac{\langle v,w\rangle}{\langle w,w\rangle}$, also $\overline{c} = \frac{\langle w,v\rangle}{\langle w,w\rangle}$. Es ist

$$\begin{aligned} 0 \leq \langle v - cw, v - cw\rangle &= \langle v, v\rangle - c\langle w, v\rangle - \overline{c}\langle v, w\rangle + c\overline{c}\langle w, w\rangle \\ &= \langle v, v\rangle - \frac{\langle v, w\rangle}{\langle w, w\rangle}\langle w, v\rangle - \frac{\langle w, v\rangle}{\langle w, w\rangle}\langle v, w\rangle + \frac{\langle v, w\rangle}{\langle w, w\rangle}\frac{\langle w, v\rangle}{\langle w, w\rangle}\langle w, w\rangle \\ &= \langle v, v\rangle - \frac{\langle v, w\rangle}{\langle w, w\rangle}\langle w, v\rangle. \end{aligned}$$

Multiplizieren wir mit $\langle w, w\rangle$, so erhalten wir

$$\begin{aligned} 0 &\le \langle v, v\rangle\langle w, w\rangle - \langle v, w\rangle\langle w, v\rangle \\ &= \langle v, v\rangle\langle w, w\rangle - \langle v, w\rangle\overline{\langle v, w\rangle} \\ &= \langle v, v\rangle\langle w, w\rangle - |\langle v, w\rangle|^2, \end{aligned}$$

also

$$|\langle v, w\rangle|^2 \le \langle v, v\rangle\langle w, w\rangle.$$

Wurzelziehen liefert die erste Behauptung.

Wann gilt das Gleichheitszeichen? Einerseits, wenn $w = 0$ gilt. Ist $w \neq 0$, so gilt das Gleichheitszeichen genau dann, wenn $0 = \langle v - cw, v - cw\rangle$, wenn also $v = cw$. Die Bedingungen $w = 0$ oder ($w \neq 0$ und $w = cv$ für ein $c \in K$) besagen gerade, daß die Vektoren v, w linear abhängig sind.

Folgerung 1 (Dreiecksungleichung): $||v + w|| \le ||v|| + ||w||$, für $v, w \in V$.

Beweis: Es ist

$$\begin{aligned} ||v + w||^2 &= \langle v + w, v + w\rangle \\ &= \langle v, v\rangle + 2\langle v, w\rangle + \langle w, w\rangle \\ &\le ||v||^2 + 2||v|| \cdot ||w|| + ||w||^2 = \big(||v|| + ||w||\big)^2, \end{aligned}$$

zieht man links und rechts die Wurzel, so folgt die Behauptung.

Insgesamt sieht man, daß $||-||$ die Rechenregeln für eine "Norm" erfüllt:

(N0) $||-||$ ist eine Abbildung $V \to \mathbb{R}_{\ge 0}$.

(N1) $||v|| = 0$ genau dann, wenn $v = 0$.

(N2) $||\lambda v|| = |\lambda| \cdot ||v||$, für $\lambda \in K$ und $v \in V$.

(N3) **(Dreiecksungleichung)** $||v + w|| \le ||v|| + ||w||$, für alle $v, w \in V$.

Folgerung 2. Sei V ein euklid'scher Vektorraum. Für von Null verschiedene Vektoren $v, w \in V$ gilt

$$-1 \le \frac{\langle v, w\rangle}{||v|| \cdot ||w||} \le 1,$$

demnach gibt es eine eindeutig bestimmte reelle Zahl $0 \le \alpha \le \pi$ mit

$$(*) \qquad \cos\alpha = \frac{\langle v, w\rangle}{||v|| \cdot ||w||}.$$

Man nennt α den (unorientierten) Winkel zwischen v und w und schreibt auch $\alpha = \measuredangle\{v, w\}$. *Dies ist der übliche Winkelbegriff* (in $\mathbb{R}^2$ und $\mathbb{R}^3$).

Beweis: Seien also $v, w \in V$ von Null verschiedene Vektoren. Multiplizieren wir v oder w mit einer positiven reellen Zahl, so ändert sich weder der Wert $\frac{\langle v, w\rangle}{||v|| \cdot ||w||}$, noch das übliche Winkelmaß, also können wir voraussetzen $||v|| = ||w|| = 1$. Sind v, w linear abhängig, so ist nun entweder $w = v$ und $\measuredangle\{v,w\} = 0$, oder aber $w = -v$ und $\measuredangle\{v,w\} = \pi$.

Seien nun v, w linear unabhängig. Setze $v_1 = v$ und wähle einen zu v_1 orthogonalen Einheitsvektor $v_2 \in \operatorname{span}\{v_1, v_2\}$ (es gibt genau zwei solche Vektoren). Ist V n-dimensional, so können wir v_1, v_2 zu einer Orthonormalbasis $v_1, \ldots, v_n$ von V fortsetzen und wir können dann V mit dem kanonischen euklid'schen Raum $\mathbb{R}^n$ vermöge $v_i \mapsto e_i$ identifizieren. Dabei entspricht v_1 dem kanonischen Basisvektor e_1, und v_2 entspricht einem Vektor der Länge 1 in dem von e_1, e_2 aufgespannten Unterraum. Auch wenn V unendlich-dimensional ist, können wir den von v_1, v_2 aufgespannten Unterraum mit dem kanonischen 2-dimensionalen euklid'schen Raum $\mathbb{R}^2$ identifizieren, und zwar v_1 mit dem kanonischen Basisvektor e_1, und v_2 mit einem Vektor v_2' der Länge 1 in $\mathbb{R}^2$.

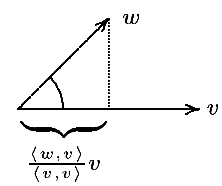
Wegen $||v_2'|| = 1$ ist $v_2' = \begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix}$ für ein $0 \le \alpha < 2\pi$ und natürlich ist

$$\left\langle \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix} \right\rangle = \cos\alpha$$

(hier ist $\langle -, - \rangle$ das kanonische innere Produkt). Der neu-eingeführte Winkelbegriff ordnet dem Paar $\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix}$ allerdings nicht die Zahl α, sondern die Zahl $\min(\alpha, 2\pi - \alpha)$ zu, daher wurde oben betont, daß es sich hier um **nicht-orientierte** Winkel handelt.

Die **Projektion von w in Richtung** $v \neq 0$ ist der Vektor

$$||w|| \cdot \cos\measuredangle\{v,w\} \cdot \frac{1}{||v||} \cdot v = \frac{\langle w,v\rangle}{\langle v,v\rangle} v,$$



seine Länge ist $||w|| \cdot |\cos\measuredangle\{v,w\}| = \frac{\langle v,w\rangle}{||v||}$. (Dies liefert eine Interpretation für das häufige Auftreten des Terms $\frac{\langle v,w\rangle}{\langle v,v\rangle} v$ in Formeln. Zum Beispiel in der Spiegelungsformel $\sigma_a(v) = v - \frac{2\langle v,a\rangle}{\langle a,a\rangle} a$, hier wird von v das Doppelte von $\frac{\langle v,a\rangle}{\langle a,a\rangle} a$ subtrahiert: subtrahiert man nur $\frac{\langle v,a\rangle}{\langle a,a\rangle} a$, so subtrahiert man gerade die Projektion von v in Richtung a und man landet in $(\mathbb{R}a)^\perp$; zweimaliges Subtrahieren liefert gerade das Spiegelbild.)

Den **Abstand eines Punkts v von einer Hyperebene H** kann man auf diese Weise berechnen: Eine Hyperebene im $\mathbb{R}^n$ ist von der Form $u + L$, mit $u \in \mathbb{R}^n$ und L ein $(n-1)$-dimensionaler Unterraum, oder auch die Menge der Lösungen einer Gleichung $\sum a_i X_i = b$ mit $[a_1, \ldots, a_n] \neq 0$. Hier ist L die Menge der Lösungen der zugehörigen homogenen Gleichung $\sum a_i X_i = 0$, insbesondere ist $L = (\mathbb{R}a)^\perp$. Der Abstand von v zu $u + L$ ist gleich dem Abstand von $v - u$ zu L. Zu berechnen ist also die Länge der Projektion des Vektors $v - u$ in Richtung von a.

Dies liefert auch ein Verfahren, um den **Abstand zweier paralleler Hyperebenen** H, H' zu berechnen. Sei also $H = u + L$, und $H' = u' + L$ mit $u, u' \in \mathbb{R}^3$ und L ein $(n-1)$-dimensionaler Unterraum. Der Abstand von H und H' ist gleich dem Abstand von $u - u'$ von L.

Die Gleichung $(*)$ liefert eine **geometrische Interpretation des inneren Produkts**, denn man kann sie folgendermaßen umschreiben:

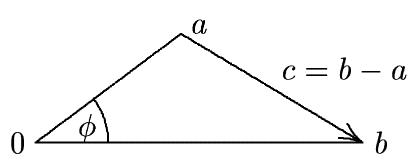
$$\langle v, w\rangle = ||v|| \cdot ||w|| \cdot \cos \measuredangle\{v, w\},$$

man erhält also den Wert $\langle v, w\rangle$, indem man die Längen von v und von w miteinander und mit dem Cosinus des eingeschlossenen Winkels multipliziert. Oder, wenn wir Klammern setzen: $\langle v, w\rangle = ||v|| \cdot \big(||w|| \cdot \cos \measuredangle\{v, w\}\big)$, es wird also die Länge von v mit der Länge der Projektion von w in Richtung v multipliziert. Oder auch: $\langle v, w\rangle = ||w|| \cdot \big(||v|| \cdot \cos \measuredangle\{v, w\}\big)$, hier wird die Länge von w mit der Länge der Projektion von v in Richtung w multipliziert. Alle diese Interpretation verdeutlichen das Orthogonalitätskriterium: *Für zwei von Null verschiedene Vektoren v, w ist $\measuredangle\{v, w\} = \frac{\pi}{2}$ genau dann, wenn $\langle v, w\rangle = 0$ gilt.*

Die Gleichung $(*)$ wird in der Elementargeometrie der "Cosinus-Satz" genannt: *Sei ein Dreieck mit den Seiten $||a||, ||b||, ||c||$ gegeben, der Winkel ϕ liege gegenüber der Seite $||c||$. Dann ist*

$$||c||^2 = ||a||^2 + ||b||^2 - 2||a|| \cdot ||b|| \cos \phi .$$

Betrachtet man nämlich das folgende Dreieck:



so ist

$$\begin{aligned}
||c||^2 &= \langle c, c\rangle \\
&= \langle b - a, b - a\rangle \\
&= \langle b, b\rangle + \langle a, a\rangle - 2\langle a, b\rangle \\
&= ||a||^2 + ||b||^2 - 2||a|| \cdot ||b|| \cos \phi .
\end{aligned}$$

Das Vektorprodukt im $\mathbb{R}^3$. Seien $v = {}^t[v_1, v_2, v_3], w = {}^t[w_1, w_2, w_3] \in \mathbb{R}^3$. Man definiert

$$v \times w = {}^t[v_2w_3 - v_3w_2, v_3w_1 - v_1w_3, v_1w_2 - v_2w_1],$$

man erhält auf diese Weise eine Abbildung $\mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$. Es gelten folgende Rechenregeln: (a) Genau dann sind die Vektoren v, w linear abhängig, wenn $v \times w = 0$. (Denn wir bilden die drei Unterdeterminanten der (2×3)-Matrix $\begin{bmatrix} v_1 & w_1 \\ v_2 & w_2 \\ v_3 & w_3 \end{bmatrix}$.)

(b) $\langle v, v \times w \rangle = 0$, $\langle w, v \times w \rangle = 0$. (Nachrechnen!) Der Vektor $v \times w$ steht demnach senkrecht auf v und auf w. Sind also v, w linear unabhängig, so erhält man auf diese Weise einen von Null verschiedenen Vektor, der auf der von $v, w, 0$ aufgespannten Ebene senkrecht steht. (Eine wichtige Aufgabenstellung.)

(c) Es ist $||v \times w||^2 = ||v||^2 ||w||^2 - \langle v, w \rangle^2$. (Nachrechnen). Man sieht also, daß hier die Cauchy-Schwarz'sche Ungleichung präzisiert wird: die Differenz der Quadrate der beiden Seiten ist durch $||v \times w||^2$ gegeben.

(d) Schreiben wir die Gleichung (c) um, so sehen wir:

$$\begin{aligned} ||v \times w||^2 &= ||v||^2 ||w||^2 - \langle v, w \rangle^2 \\ &= ||v||^2 ||w||^2 (1 - \cos^2 \measuredangle\{v, w\}) \\ &= ||v||^2 ||w||^2 \sin^2 \measuredangle\{v, w\} \end{aligned}$$

Also ist

$$||v \times w|| = ||v|| \cdot ||w|| \cdot |\sin \measuredangle\{v, w\}|.$$

In der Statistik wird bei der Beurteilung der sogenannten "linearen Regression" auf geometrische Begriffsbildungen zurückgegriffen:

Geometrische Interpretation des Korrelations-Koeffizienten. Seien reelle Zahlen $x_1, \ldots, x_n$ mit Mittelwert $\overline{x}$ und entsprechend seien $y_1, \ldots, y_n$ mit Mittelwert $\overline{y}$ gegeben. Der lineare Korrelations-Koeffizient r_{xy} ist nichts anderes als $\cos \phi$, wobei ϕ der Winkel zwischen den Vektoren

$$(x_1, \ldots, x_n) - (\overline{x}, \ldots, \overline{x}) \quad \text{und} \quad (y_1, \ldots, y_n) - (\overline{y}, \ldots, \overline{y})$$

ist.

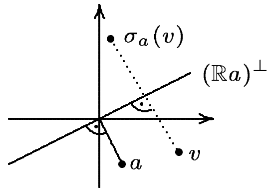
3G. Die Gruppe $\mathcal{O}(V)$ wird von Spiegelungen erzeugt.

Sei V ein euklid'scher Vektorraum. Sei $\mathcal{O}(V)$ die Menge der orthogonalen Automorphismen von V in sich. Dies ist eine Untergruppe der Gruppe $\mathrm{GL}(V)$ aller Automorphismen von V.

Spiegelungen. Sei $0 \neq a \in V$. Definiere eine Abbildung $\sigma_a \colon V \to V$ durch

$$\sigma_a(v) = v - \frac{2\langle v, a \rangle}{\langle a, a \rangle} a \quad \text{für} \quad v \in V.$$

Dies ist ein orthogonaler Endomorphismus, mit $\sigma_a^2 = 1$, *es ist* $\sigma_a(a) = -a$ *und* $\sigma_a(v) = v$ *für alle* $v \in (\mathbb{R}a)^\perp$. Beweis: Nachrechnen. Man nennt dies die *Spiegelung* an der Hyperebene $(\mathbb{R}a)^\perp$.



Lemma. *Sind* $u \neq v \in V$, *mit* $||u|| = ||v||$, *so gilt* $\sigma_{v-u}(v) = u$. (Es gibt also eine Spiegelung, die die beiden Vektoren u und v vertauscht.)

Beweis: Ganz allgemein gilt: $\langle u+v, v-u\rangle = ||v||^2 - ||u||^2$. Wegen $||v|| = ||u||$ folgt hieraus $\langle v, v-u\rangle = -\langle u, v-u\rangle$, also ist $\frac{2\langle v,v-u\rangle}{\langle v-u,v-u\rangle} = 1$. Berechnet man nun $\sigma_{v-u}(v)$, so erhält man

$$\sigma_{v-u}(v) = v - \frac{2\langle v, v-u\rangle}{\langle v-u, v-u\rangle}(v-u) = v - (v-u) = u.$$

Man sieht also: Betrachte die *Sphäre* $\Sigma_r = \{x \in \mathbb{R}^n \mid ||x|| = r\}$ *mit Radius* r. Zu je zwei Elementen $u, v \in \Sigma_r$ gibt es $\phi \in \mathcal{O}(n)$ mit $\phi(u) = v$; man sagt: Die Gruppe $\mathcal{O}(V)$ operiert *transitiv* auf der Menge Σ_r.

Satz. *Jeder orthogonale Endomorphismus des* $\mathbb{R}^n$ *kann als Produkt von höchstens* n *Spiegelungen geschrieben werden.*

Beweis: Sei $\phi \in \mathcal{O}(n)$. Ist $\phi = 1$, so ist ϕ "Produkt von 0 Spiegelungen". Sei nun $u \in \mathbb{R}$ mit $\phi(u) \neq u$. Die Vektoren $u, \phi(u)$ haben gleiche Länge, also gibt es σ_a mit $\sigma_a(u) = \phi(u)$. Es ist also $\sigma_a \circ \phi(u) = u$. Sei $U = \langle u\rangle^\perp$. Da u unter $\sigma_a \circ \phi$ auf sich abgebildet wird, gilt das gleiche für den Unterraum U der Dimension $n-1$. $\phi' = \sigma_a \circ \phi \mid U$. Nach Induktion ist ϕ' Produkt von Spiegelungen σ'_{a_i} mit $a_i \in U$ und $2 \leq i \leq t$ mit $t \leq n$. Dies sind Einschränkungen der entsprechenden Spiegelungen $\sigma_{a_i} \in \mathcal{O}(V)$, also ist

$$\sigma_a \circ \phi(x) = \sigma_{a_2} \circ \cdots \circ \sigma_{a_t}(x) \quad \text{für alle} \quad x \in U.$$

Aber natürlich auch für $x = u$, demnach für alle $x \in V$. Wir sehen:

$$\sigma_a \circ \phi = \sigma_{a_2} \circ \cdots \circ \sigma_{a_t}.$$

Multiplizieren wir von links mit $\sigma_a^{-1} = \sigma_a$, so erhalten wir

$$\phi = \sigma_{a_1} \circ \sigma_{a_2} \circ \cdots \circ \sigma_{a_t},$$

dabei haben wir $a_1 = a$ gesetzt.

3H. Iwasawa-Zerlegung oder QR-Zerlegung.

Satz. *Jede Matrix $A \in M(n \times n, \mathbb{R})$ kann als Produkt $A = PB$ einer orthogonalen Matrix P und einer oberen Dreiecksmatrix B geschrieben werden.*

Beweis (mit Induktion): Sei v_1 die erste Spalte von A, sei $d_1 = ||v_1||$. Sei $\phi\colon \mathbb{R}^n \to \mathbb{R}^n$ eine orthogonale Abbildung mit $\phi(v_1) = d_1 e_1$ und sei Q die $(n \times n)$-Matrix mit $\phi = f_Q$. Da ϕ orthogonale Abbildung ist, ist A eine orthogonale Matrix. Es ist $f_{QA}(e_1) = \phi \circ f_A(e_1) = d_1 e_1$. Die Matrix QA hat demnach die Form

$$QA = \begin{bmatrix} d_1 & c \\ 0 & A' \end{bmatrix}$$

mit einer Matrix $A' \in M((n-1) \times (n-1), \mathbb{R})$. Nach Induktion läßt sich $A' = P'B'$ schreiben, mit P' orthogonal und B' obere Dreiecksmatrix. Also

$$A = Q^{-1} \begin{bmatrix} d_1 & c \\ 0 & A' \end{bmatrix} = Q^{-1} \begin{bmatrix} 1 & 0 \\ 0 & P' \end{bmatrix} \begin{bmatrix} d_1 & c \\ 0 & B' \end{bmatrix}.$$

Die ersten beiden Matrizen Q^{-1} und $\begin{bmatrix} 1 & 0 \\ 0 & P' \end{bmatrix}$ sind orthogonale Matrizen, also ist auch das Produkt orthogonal.

Folgerung (Iwasawa-Zerlegung oder QR-Zerlegung). *Jede invertierbare Matrix $A \in M(n \times n, \mathbb{R})$ kann als Produkt $A = PDU$ einer orthogonalen Matrix P, einer Diagonalmatrix U und einer unipotenten oberen Dreiecksmatrix U geschrieben werden.*

Eindeutigkeit: Wenn wir zusätzlich voraussetzen, daß die Diagonalkoeffizienten von D positiv sind (was wir können), so sind alle drei Matrizen P, D, U eindeutig bestimmt. Beweis: Übungsaufgabe 1 von Blatt 9. Der Satz kann auch als eine unmittelbare Folgerung aus dem Gram-Schmidt'schen Orthonormalisierungsverfahren angesehen werden, besser: als eine Umformulierung. Siehe ebenfalls Übungsaufgabe 1 von Blatt 9.

Statt der Reihenfolge "orthogonal-diagonal-unipotent" kann man natürlich ebenso gut die Reihenfolge "orthogonal-unipotent-diagonal" erzielen. Hier die Vorstellung, die man im Fall $n = 2$ für die Iwasawa-Zerlegung "orthogonal-unipotent-diagonal" haben sollte: Wir beginnen mit zwei linear unabhängigen Vektoren v_1, v_2, drehen oder spiegeln so, daß v_1 ein positiv-Vielfaches des ersten Einheitsvektors ist, ersetzen dann den zweiten Vektor durch seine Projektion in Richtung e_2 und als letztes wird skaliert: