Kapitel VI: So einfach wie möglich, aber nicht einfacher (Normalformen von Endomorphismen)

0 Wiederholung einiger Grundbegriffe der linearen Algebra I

0.1 Definition:

Ein Körper K ist eine Menge zusammen mit zwei Verknüpfungen $+, \cdot$, so dass gilt:

(a) $(K, +)$ ist eine kommutative Gruppe, d.h.

1.) Es gilt das Assoziativgesetz $(a + b) + c = a + (b + c)$
2.) Es gibt ein neutrales Element $0 \in K$ mit $a + 0 = 0 + a = a$
3.) Es gibt inverse Elemente $-a \in K$ mit $a + (-a) = (-a) + a = 0$
4.) Es gilt das Kommutativgesetz $a + b = b + a$

(b) $(K \setminus \{0\}, \cdot)$ ist eine kommutative Gruppe (neutrales Element $1 \in K$, inverses Element $a^{-1} \in K$)

(c) Es gelten die Distributivgesetze:

1.) $a \cdot (b + c) = ab + ac$
2.) $(a + b) \cdot c = ac + bc$

0.2 Beispiele: (für Körper)

(a) $\mathbb{Q}$, der Körper der rationalen Zahlen

(b) $\mathbb{R}$, der Körper der reellen Zahlen

(c) $\mathbb{C}$, der Körper der komplexen Zahlen

(d) $\mathbb{F}_p$, endlicher Körper mit p Elementen (p Primzahl). $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$

Im Folgenden sei K ein Körper.

0.3 Definition:

Ein K-Vektorraum V ist eine Menge mit einer Verknüpfung $+ : V \times V \to V$ und einer Abbildung $\cdot : K \times V \to V$ („skalare Multiplikation“), so dass gilt:

(a) $(V, +)$ ist eine kommutative Gruppe (neutrales Element 0)

(b) Es gilt das Assoziativgesetz $a \cdot (b \cdot v) = (a \cdot b) \cdot v$

(c) Es gelten die Distributivgesetze:

1.) $(a + b) \cdot v = av + bv$
2.) $a \cdot (v + w) = av + aw$

(d) Es gilt die Treue: $1 \cdot v = v$

0.4 Beispiele: (für Vektorräume)

(a) K^n = Menge der n-Tupel mit Einträgen aus K (mit komponentenweiser Addition und skalarer Multiplikation)

(b) $K[x]$ = Menge der Polynome in einer Unbestimmten x mit Koeffizienten aus K

$K[x]_{\leq n}$ = Menge der Polynome vom Grad (kleiner gleich) n

(c) $Mat_{m,n}(K)$ = Menge der $m \times n$-Matrizen über K

Im Folgenden sei V ein K-Vektorraum.

0.5 Definition:

(a) Eine Menge $E \subseteq V$ heißt ein Erzeugendensystem von V, wenn es für jeden Vektor $v \in V$ endlich viele Vektoren $w_1, \ldots, w_n \in E$ und Zahlen $a_1, \ldots, a_n \in K$ gibt mit $v = a_1 w_1 + \ldots + a_n w_n$.

(b) Eine Menge $E \subseteq V$ heißt linear unabhängig, wenn für $w_1, \ldots, w_n \in E$ und $a_1, \ldots, a_n \in K$ aus $a_1 w_1 + \ldots + a_n w_n = 0$ folgt, dass $a_1 = \ldots = a_n = 0$ sein muss.

(c) Eine Menge $E \subseteq V$ heißt eine Basis von V, wenn sie ein Erzeugendensystem von V und linear unabhängig ist.

0.6 Satz:

Ist V endlich erzeugt, so besitzt V eine Basis. Je zwei Basen haben gleich viele Elemente. Diese Elementezahl heißt die Dimension von V und wird mit $dim_K(V)$ bezeichnet.

0.7 Beispiele: (für Basen und Dimensionen)

(a) Der K-Vektorraum K^n erfüllt $dim_K(K^n) = n$. Eine Basis von K^n ist z.B. die Standardbasis $E = \{e_1, \ldots, e_n\}$ mit $e_i = (0, \ldots, 0, \underbrace{1}_{\text{i-te Stelle}}, 0, \ldots, 0)$.

(b) Der K-Vektorraum $K[x]$ besitzt die Basis $\{1, x, x^2, \ldots\}$. Der K-Vektorraum $K[x]_{\leq n}$ erfüllt $dim_K(K[x]_{\leq n}) = n + 1$ und besitzt z.B. die Basis $\{1, x, x^2, \ldots, x^n\}$.

(c) Der K-Vektorraum $Mat_{m,n}(K)$ erfüllt $dim_K(Mat_{m,n}(K)) = m \cdot n$ und besitzt z.B. die Basis $E = \{A_{ij} \mid 1 \leq i \leq m, 1 \leq j \leq n\}$ mit

$$A_{ij} = \begin{array}{c} \text{j-te Spalte} \\ \begin{pmatrix} & \vdots & \\ \cdots & 1 & \cdots \\ & \vdots & \end{pmatrix} \end{array} \text{ i-te Zeile}$$

0.8 Definition:

Seien V, W zwei K-Vektorräume.

(a) Eine Abbildung $f : V \to W$ heißt K-lineare Abbildung oder ein Homomorphismus von Vektorräumen, wenn gilt: $f(a_1 v_1 + a_2 v_2) = a_1 f(v_1) + a_2 f(v_2)$ für $a_1, a_2 \in K$ und $v_1, v_2 \in V$.

(b) Eine K-lineare Abbildung $f : V \to V$ heißt ein Endomorphismus von V.

0.9 Satz:

Sei V ein endlich-dimensionaler K-Vektorraum mit Basis $B = \{v_1, \ldots, v_n\}$ und W ein endlich-dimensionaler K-Vektorraum mit Basis $C = \{w_1, \ldots, w_m\}$.
Ferner sei $f : V \to W$ eine K-lineare Abbildung.

(a) Man kann der Abbildung f wie folgt eine Darstellungsmatrix $M_C^B(f) \in Mat_{m,n}(K)$ zuordnen. Für $j = 1, \ldots, n$ schreibe $f(v_j) = a_{1j}w_1 + \ldots + a_{mj}w_m$. Dann setze die Tupel $(a_{1j}, \ldots, a_{mj})$ in die Spalten einer Matrix, d.h. setze $M_C^B(f) = (a_{ij})$.

(b) Umgekehrt kann man jeder Matrix $A = (a_{ij}) \in Mat_{m,n}(K)$ eine lineare Abbildung $f_A : V \to W$ zuordnen, indem man f_A durch

$$f_A(c_1v_1 + \ldots + c_nv_n) = \sum_{j=1}^{n} c_j f_A(v_j) = \sum_{j=1}^{n} c_j \sum_{i=1}^{m} a_{ij}w_i = (w_1, \ldots, w_m) \cdot A \cdot \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$ definiert.

0.10 Beispiele: (für lineare Abbildungen)

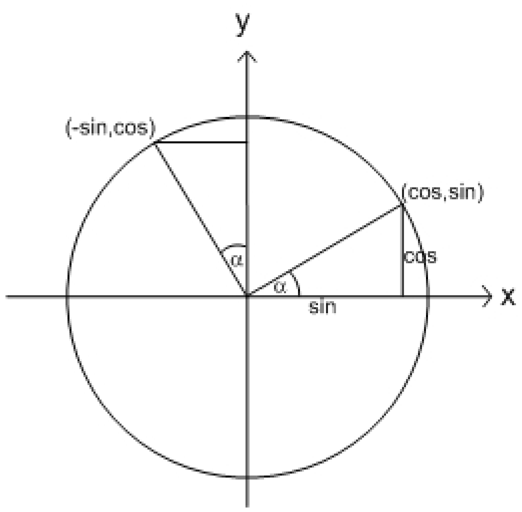
(a) Die Nullabbildung $f : V \to W$ mit $f(v) = 0$ für alle $v \in V$.

(b) Die Identität / identische Abbildung $\begin{aligned} id_V : V &\to V \\ v &\mapsto v \end{aligned}$

(c) Drehung um $0 \in \mathbb{R}^2$ um den Winkel $\alpha \in [0, 2\pi)$:

$\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ mit der Matrix $M_E(\varphi) = \begin{pmatrix} cos(\alpha) & -sin(\alpha) \\ sin(\alpha) & cos(\alpha) \end{pmatrix}$ bezüglich $E = \{e_1, e_2\}$.

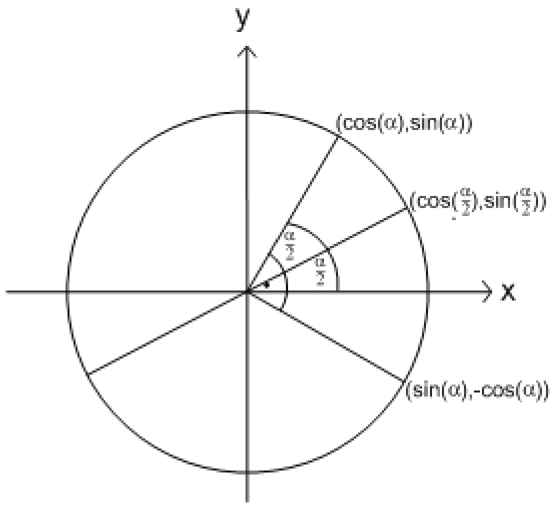
Skizze:



(d) Spiegelung an einer Geraden $G = \mathbb{R} \cdot (cos(\frac{\alpha}{2}), sin(\frac{\alpha}{2}))$ in $\mathbb{R}^2$:

$\psi : \mathbb{R}^2 \to \mathbb{R}^2$ mit der Matrix $M_E(\psi) = \begin{pmatrix} cos(\alpha) & sin(\alpha) \\ sin(\alpha) & -cos(\alpha) \end{pmatrix}$

Skizze:



(e) Ist $V = U_1 \oplus U_2$ die direkte Summe zweier Untervektorräume U_1, U_2, so heißt

$$\begin{aligned} pr_U : V &\to U_1 \\ v &\mapsto u_1 \text{ für } v = u_1 + u_2 \end{aligned}$$

die Projektion von V auf U_1 längs U_2.

Bezüglich geeigneter Basen gilt $M_C^B(pr_{U_1}) = \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & 1 & \cdots & 0 \end{pmatrix}$

Man kann pr_{U_1} auch als Endomorphismus $\begin{aligned} \Phi : V &\to V \\ v &\mapsto pr_{U_1}(v) \end{aligned}$ auffassen. In diesem Fall gilt $\Phi^2 = \Phi$.

0.11 Definition:

(a) Zwei Matrizen $A, B \in Mat_{m,n}(K)$ heißen äquivalent, wenn es Matrizen $S \in GL_m(K)$ und $T \in GL_n(K)$ gibt mit $A = SBT$.

(b) Zwei Matrizen $A, B \in Mat_n(K)$ heißen ähnlich, wenn es eine Matrix $T \in GL_n(K)$ gibt mit $A = TBT^{-1}$.

0.12 Satz:

(a) Zwei Matrizen $A, B \in Mat_{m,n}(K)$ sind äquivalent, wenn sie dieselbe lineare Abbildung $f : K^m \to K^n$ bezüglich verschiedener Basen B_1, B_2 von K^m und C_1, C_2 von K^n repräsentieren, d.h. wenn gilt $A = M_{B_2}^{B_1}(f)$ und $B = M_{C_2}^{C_1}(f)$.

(b) Zwei Matrizen $A, B \in Mat_n(K)$ sind ähnlich, wenn sie denselben Endomorphismus $f : V \to V$ bezüglich zwei verschiedener Basen C_1, C_2 von V repräsentieren, d.h. wenn gilt $A = M_{C_1}(f)$ und $B = M_{C_2}(f)$.

Dies folgt aus der Transformationsformel.

0.13 <u>Definition:</u>

Sei $A = (a_{ij}) \in Mat_n(K)$. Dann heißt die Zahl

$$det(A) = \sum_{\sigma \in S_n} sign(\sigma) \cdot a_{1,\sigma(1)} \cdot a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}$$

die <u>Determinante</u> von A.

0.14 <u>Satz:</u> (Rechenregeln für die Determinante)

(a) det ist linear in den Zeilen und Spalten von A.

(b) det ist alternierend in den Zeilen und Spalten von A.

(c) Man kann $det(A)$ durch Entwicklung nach einer Zeile oder Spalte berechnen.

(d) Es gilt der Determinantenmultiplikationssatz:
$det(AB) = det(A) \cdot det(B)$

(e) Es gilt $det(A) \neq 0 \Leftrightarrow f_A$ ist bijektiv $\Leftrightarrow A$ ist invertierbar

19 Aufgeräumte Matrizen oder so ähnlich (Diagonalisierbarkeit)

A. Wiederholung: Eine Basis aus Eigenvektoren

Im Folgenden sei K ein Körper, V ein endlich-dimensionaler K-Vektorraum mit Basis $B = \{v_1, \dots, v_n\}$ und $f : V \to V$ ein Endomorphismus.

19.1 Definition:

(a) Ein Element $\lambda \in K$ heißt Eigenwert von f, wenn es einen Vektor $v \in V \setminus \{0\}$ gibt mit $f(v) = \lambda v$.

(b) Ist $\lambda \in K$ ein Eigenwert von f und $v \in V$ mit $f(v) = \lambda v$, so heißt v ein Eigenvektor von f zum Eigenwert λ.

(c) Sei $\lambda \in K$ ein Eigenwert von f. Die Menge aller Eigenvektoren von f zum Eigenwert λ bildet einen K-Untervektorraum von V. Er heißt der Eigenraum von f zum Eigenwert λ und wird mit $Eig(f, \lambda)$ bezeichnet.

19.2 Beispiele: (für Eigenwerte, Eigenvektoren und Eigenräume)

(a) $Eig(f, 1) = \{v \in V \mid f(v) = v\}$ ist die Menge der Fixpunkte von f.

(b) $Eig(f, 0) = \{v \in V \mid f(v) = 0\} = Kern(f)$ ist der Kern von f.

(c) Sei $K = \mathbb{R}$, $V = \mathbb{R}^2$ und $f : \mathbb{R}^2 \to \mathbb{R}^2$ die Drehung um 0 um den Winkel $\alpha \in [0, 2\pi)$. Es gelte also $M_E(f) = \begin{pmatrix} cos(\alpha) & -sin(\alpha) \\ sin(\alpha) & cos(\alpha) \end{pmatrix}$

Für $\alpha \notin \{0, \pi\}$ besitzt f dann keine Eigenwerte.
Für $\alpha = 0$ gilt $f = id_V$ und $Eig(f, 1) = V$.
Für $\alpha = \pi$ gilt $f = -id_V$ und $Eig(f, -1) = V$.

(d) Sei $K = \mathbb{R}$, $V = \mathbb{R}^2$ und $f : \mathbb{R}^2 \to \mathbb{R}^2$ die Spiegelung an der Geraden $G = \mathbb{R} \cdot (cos(\frac{\alpha}{2}), sin(\frac{\alpha}{2}))$. Es gelte also $M_E^E(f) = \begin{pmatrix} cos(\alpha) & sin(\alpha) \\ sin(\alpha) & -cos(\alpha) \end{pmatrix}$

Dann besitzt f die Eigenwerte $\lambda = 1$ und $\lambda = -1$.
Es gilt $Eig(f, 1) = \mathbb{R} \cdot (cos(\frac{\alpha}{2}), sin(\frac{\alpha}{2})) =: \mathbb{R} \cdot v_1$ und
$Eig(f, -1) = \mathbb{R} \cdot (-sin(\frac{\alpha}{2}), cos(\frac{\alpha}{2})) =: \mathbb{R} \cdot v_2$
Die Vektoren v_1, v_2 bilden eine Basis von $\mathbb{R}^2$ aus Eigenvektoren.

19.3 Definition:

(a) Der Endomorpismus $f : V \to V$ heißt diagonalisierbar, wenn es eine Basis B von V gibt, so dass $M_B^B(f)$ eine Diagonalmatrix ist.

(b) Eine Matrix $A \in Mat_n(K)$ heißt diagonalisierbar, wenn $f_A : K^n \to K^n$ diagonalisierbar ist, d.h. wenn es eine Matrix $T \in GL_n(K)$ gibt, so dass TAT^{-1} eine Diagonalmatrix ist.

19.4 Satz:

Ein Endomorphismus $f : V \to V$ ist genau dann diagonalisierbar, wenn es eine Basis B von V gibt, die aus Eigenvektoren von f besteht.
(In diesem Fall ist $M_B^B(f)$ eine Diagonalmatrix.)

19.5 Definition:

(a) Sei $A \in Mat_n(K)$. Das Polynom $\chi_A(x) = det(A - x \cdot I_n) \in K[x]$ heißt das charakteristische Polynom von A.

(b) Das charakteristische Polynom von $f : V \to V$ ist wie folgt definiert:
Wähle eine Basis B von V und setze $\chi_f(x) = \chi_{M_B^B(f)}(x) \in K[x]$.

(Nach der Transformationsformel und dem Determinantenmultiplikationssatz hängt $\chi_f(x)$ nicht von der Wahl der Basis B ab.)

19.6 Satz: (Berechnung von Eigenwerten und Eigenräumen)

(a) Eine Zahl $\lambda \in K$ ist genau dann ein Eigenwert von f, wenn sie eine Nullstelle von $\chi_f(x)$ ist.

(b) Ist $\lambda \in K$ ein Eigenwert von f, so gilt $Eig(f, \lambda) = Kern(f - \lambda \cdot id_V)$

(Insbesondere kann man dann $Eig(f, \lambda)$ als Lösungsmenge eines LGS berechnen.)

19.7 Beispiele:

(a) Sei $K = \mathbb{R}$, $V = \mathbb{R}^2$ und $f : \mathbb{R}^2 \to \mathbb{R}^2$ die Drehung um 0 um den Winkel $\alpha \in [0, 2\pi)$.

Dann gilt $M_E^E(f) = \begin{pmatrix} cos(\alpha) & -sin(\alpha) \\ sin(\alpha) & cos(\alpha) \end{pmatrix}$

und $\chi_f(x) = det \begin{pmatrix} cos(\alpha) - x & -sin(\alpha) \\ sin(\alpha) & cos(\alpha) - x \end{pmatrix}$

$= cos^2(\alpha) - 2cos(\alpha)x + x^2 + sin^2(\alpha) = x^2 - 2cos(\alpha) + 1$

Die Diskriminante der quadratischen Gleichung $x^2 - 2cos(\alpha) + 1 = 0$ ist $4cos^2(\alpha) - 4 \leq 0$.
Also besitzt f nur Eigenwerte für $\alpha \in \{0, \pi\}$ (siehe Beispiel 19.2).

$(ax^2 + bx + c = 0 \Rightarrow x_{1/2} = \dfrac{-b \pm \sqrt{b^2 - 4ac}}{2a} \Rightarrow$ Diskriminante: $b^2 - 4ac)$

(b) Sei nun $g : \mathbb{R}^2 \to \mathbb{R}^2$ die Spiegelung an der Geraden $G = \mathbb{R} \cdot (cos(\frac{\alpha}{2}), sin(\frac{\alpha}{2}))$. Dann gilt

$\chi_g(x) = det \begin{pmatrix} cos(\alpha) - x & sin(\alpha) \\ sin(\alpha) & -cos(\alpha) - x \end{pmatrix} = -cos^2(\alpha) + x^2 - sin^2(\alpha)$

$= x^2 - 1 = (x - 1)(x + 1)$

Also besitzt g die Eigenwerte $\lambda = 1$ und $\lambda = -1$. Für die zugehörigen Eigenräume erhält man aus den LGS $M_E^E(g) \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$ bzw. $M_E^E(g) \cdot \begin{pmatrix} x \\ y \end{pmatrix} = -\begin{pmatrix} x \\ y \end{pmatrix}$

die in Beispiel 19.2.d angegebenen Lösungen.

B. Wie kann man die Diagonalisierbarkeit eines Endomorphismus testen?

19.8 Bemerkung:

Aus Satz 19.4 und Satz 19.6 ergibt sich ein Verfahren, wie man die Diagonalisierbarkeit von f testen kann. Wir verwenden dazu auch Satz 15.14:

Sind $\lambda_1, \ldots, \lambda_m \in K$ die paarweise verschiedenen Eigenwerte von f, so ist $Eig(f, \lambda_1) \oplus \ldots \oplus Eig(f, \lambda_m)$ eine direkte Summe von Untervektorräumen.

Um f auf Diagonalisierbarkeit zu testen, verfahre also wie folgt:

(a) Berechne $\chi_f(x)$ und finde die Nullstellen $\lambda_1, \ldots, \lambda_m \in K$ dieses Polynoms.

(b) Für $i = 1, \ldots, m$ berechne $Eig(f, \lambda_i) = Kern(f - \lambda_i \cdot id_V)$.

(c) Füge die Basen der Eigenräume $Eig(f, \lambda_i)$ zusammen und prüfe, ob die Vereinigung eine Basis von V darstellt.

Das Problem ist allerdings, dass nicht klar ist wie man Schritt (a) durchführen soll.

19.9 Definition:

Schreibe $\chi_f(x) = (-1)^n (x - \lambda_1)^{\mu_1} \cdots (x - \lambda_m)^{\mu_m} \cdot p(x)$, wobei $\lambda_1, \ldots, \lambda_m \in K$ die paarweise verschiedenen Eigenwerte von f sind und $p(x) \in K[x]$ keine Nullstellen hat. Dann heißt $\mu_i \geq 1$ die Vielfachheit oder Multiplizität des Eigenwerts λ_i für $i = 1, \ldots, m$.

19.10 Satz:

Zerlege das charakteristische Polynom von f wie in Definition 19.9, d.h. schreibe $\chi_f(x) = (-1)^n (x - \lambda_1)^{\mu_1} \cdots (x - \lambda_m)^{\mu_m} \cdot p(x)$ mit $\mu_i \geq 1$ und $p(x) \in K[x]$ ohne Nullstellen und $\lambda_1, \ldots, \lambda_m \in K$ paarweise verschieden. Für $i = 1, \ldots, m$ gilt dann $1 \leq dim_K(f, \lambda_i) \leq \mu_i$.

Beweis: Da $\lambda_1, \ldots, \lambda_m \in K$ die paarweise verschiedenen Eigenwerte von f sind, gilt $Eig(f, \lambda_i) \neq 0$ und somit $dim_K(Eig(f, \lambda_i)) \geq 1$. Sei nun $\{v_1, \ldots, v_s\}$ eine Basis von $Eig(f, \lambda_i)$. Ergänze sie zu einer Basis $B = \{v_1, \ldots, v_s, w_1, \ldots, w_t\}$ von V. Dann gilt:

$$M_B^B(f) = \left(\begin{array}{ccc|c} \lambda_i & & & \\ & \ddots & & * \\ & & \lambda_i & \\ \hline & 0 & & A \end{array} \right) \text{ und somit } \chi_f(x) = det(M_B^B(f) - x \cdot I_n) = (\lambda_i - x)^s \cdot \chi_A(x)$$

Also ist die Vielfachheit μ_i der Nullstelle λ_i von $\chi_f(x)$ mindestens $s = dim_K(Eig(f, \lambda_i))$. qed.

19.11 <u>Satz:</u> (1. Charakterisierung der Diagonalisierbarkeit)

Für einen Endomorphismus $f : V \to V$ sind die folgenden Bedingung äquivalent:

(a) f ist diagonalisierbar.

(b) Es gibt n linear unabhängige Eigenwerte von f.

(c) Sind $\lambda_1, \ldots, \lambda_m \in K$ die paarweise verschiedenen Eigenwerte von f, so gilt $dim_K(Eig(f, \lambda_1)) + \ldots + dim_K(Eig(f, \lambda_m)) = dim_K(V) = n.$

(d) $\chi_f(x)$ zerfällt in Linearfaktoren, d.h. es gilt $\chi_f(x) = (-1)^n (x - \lambda_1)^{\mu_1} \cdots (x - \lambda_m)^{\mu_m}$ mit $\lambda_1, \ldots, \lambda_m \in K$ paarweise verschieden und für $i = 1, \ldots, m$ gilt $dim_K(Eig(f, \lambda_i)) = \mu_i$.

<u>Beweis:</u> „(a)⇔(b)“ Vgl. Satz 15.11.b

„(b)⇔(c)“ Nach Satz 15.14.b ist $U = Eig(f, \lambda_1) \oplus \ldots \oplus Eig(f, \lambda_m)$ eine direkte Summe.
Also gilt $dim_K(U) = dim_K(Eig(f, \lambda_1)) + \ldots + dim_K(f, \lambda_m))$.
Die Bedingung (b) ist äquivalent mit $U = V$, also mit $dim_K(U) = dim_K(V)$
und dies ist gerade die Bedingung (c).

„(c)⇔(d)“ Folgt aus Satz 19.10:
In der Ungleichungskette $dim_K(Eig(f, \lambda_1)) + \ldots + dim_K(f, \lambda_m)) \leq \mu_1, \ldots, \mu_m \leq n$
gilt genau dann Gleichheit, wenn für $i = 1, \ldots, m$ gilt:
$dim_K(Eig(f, \lambda_i)) = \mu_i$ und wenn in der Darstellung
$\chi_f(x) = (-1)^n (c - \lambda_1)^{\mu_1} \cdots (x - \lambda_m)^{\mu_m} p(x)$ gilt $p(x) = 1$ qed.

19.12 <u>Korollar:</u>

Besitzt f n paarweise verschiedene Eigenwerte, so ist f diagonalisierbar.

<u>Beweis:</u> Wegen $dim_K(Eig(f, \lambda_i)) \geq 1$ gilt Gleichheit in
$n \leq dim_K(Eig(f, \lambda_1)) + \ldots + dim_K(Eig(f, \lambda_m)) \leq \mu_1 + \ldots + \mu_m$. qed.

19.13 <u>Beispiel:</u>

Betrachte die $\mathbb{Q}$-lineare Abbildung $f : \mathbb{Q}^3 \to \mathbb{Q}^3$ mit Matrix

$$M_E(f) = \begin{pmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{pmatrix}$$

1.) Es gilt $\chi_f(x) = det \begin{pmatrix} -x & 0 & -2 \\ 1 & 2-x & 1 \\ 1 & 0 & 3-x \end{pmatrix} = (2-x) \cdot [(-x)(3-x) + 2]$

$$= (2-x)(x^2 - 3x + 2) = (2-x)(x-1)(x-2) = (-1)^3 (x-1)(x-2)^2$$

2.) Es gilt: $Eig(f, 1) = \mathbb{Q} \cdot (-2, 1, 1)$

3.) $Eig(f, 2) = \mathbb{Q} \cdot (-1, 0, 1) \oplus \mathbb{Q} \cdot (0, 1, 0)$
Also ist $\mu_1 = 1, \mu_2 = 2$ und f ist diagonalisierbar und $V = Eig(f, 1) \oplus Eig(f, 2)$.
In der Basis $B = \{(-2, 1, 1), (-1, 0, 1), (0, 1, 0)\}$ von $\mathbb{Q}^3$ gilt:

$$M_B(f) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

20 Ein kleines Polynom mit großen Kenntnissen (Das Minimalpolynom)

Im Folgenden sei K ein Körper, V ein n-dimensionaler K-Vektorraum mit Basis $B = \{v_1, \ldots, v_n\}$ und $f : V \to V$ ein Endomorphismus mit Darstellungsmatrix $M_B(f) \in Mat_n(K)$.

A. Wie man einen Endomorphismus in ein Polynom einsetzt

20.1 Bemerkung:

Sei $p \in K[x]$ ein Polynom, $A \in Mat_m(K)$ und $f \in End_K(V)$.
Schreibe $p = c_o + c_1 x + \ldots + c_k x^k$ mit $c_i \in K$. Was bedeutet $p(f)$ bzw. $p(A)$?

(a) Man kann $A^2, A^3, \ldots \in Mat_n(K)$ berechnen. Damit berechnet man $c_i A^i$ für $i \geq 1$.
(Man multipliziert A mit $c_i \in K$, indem man alle Einträge von A mit c_1 multipliziert.)
Nun bildet man $p(A) = c_k A^K + \ldots + c_1 A + c_0 I_m$.

Die Abbildung $\begin{aligned} s : K[x] &\to Mat_m(K) \\ p &\mapsto p(A) \end{aligned}$ ist K-linear und ein Homomorphismus von Ringen, d.h.

1.) $s(p_1 + p_2) = (p_1 + p_2)(A) = p_1(A) + p_2(A) = s(p_1) + s(p_2)$
2.) $s(cp) = (cp)(A) = c \cdot p(A) = c \cdot s(p)$ für $c \in K$
3.) $s(p_1 \cdot p_2) = (p_1 \cdot p_2)(A) = p_1(A) \cdot p_2(A) = s(p_1) \cdot s(p_2)$

Sie heißt der Substitutionshomomorphismus oder der Einsetzungshomomorphismus von A in Polynome aus $K[x]$.

(b) Auch $f^2, f^3, \ldots$ sind Endomorphismen von V. Dann kann man $c_i f^i \in End_K(V)$ bilden.
Also folgt $p(f) = c_k f^k + \ldots + c_1 f + c_o \cdot id_V \in End_K(V)$
Die Darstellungsmatrix von $p(f)$ ist $p(M_B(f))$.

Die Abbildung $\begin{aligned} \sigma : K[x] &\to End_K(V) \\ p &\mapsto p(f) \end{aligned}$ ist K-linear und ein Ringhomomorphismus.
Sie heißt der Substitutionshomomorphismus von f in Polynome aus $K[x]$.

20.2 Beispiel:

Sei $f : \mathbb{R}^3 \to \mathbb{R}^3$ die Drehung um den Winkel $\alpha = \frac{\pi}{3} = 60°$ um die z-Achse. Es gilt:

$$M_E(f) = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} =: A$$

(a) Für $p = x^3 + 1$ gilt $p(A) = A^3 + I_3 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}$

(b) Für $q = x^6 - 1$ gilt $q(f) = f^6 - id_V = id_V - id_V = 0$ und $q(A) = A^6 - I_3 = 0$.

20.3 Satz:

Sei $A \in Mat_n(K)$ und $f \in End_K(V)$.

(a) Die Menge $Kern(s) = \{p \in K[x] \mid p(A) = 0\}$ ist ein Ideal von $K[x]$, d.h.

1.) $Kern(s)$ ist ein K-Untervektorraum von $K[x]$.

2.) $K[x] \cdot Kern(s) \subseteq Kern(s)$.

(b) Die Menge $Kern(\sigma) = \{p \in K[x] \mid p(f) = 0\}$ ist ein Ideal von $K[x]$.

Beweis:

(a) Sei $p_1, p_2 \in Kern(s)$ und $c_1, c_2 \in K$. Dann gilt:
$(c_1p_1 + c_2p_2)(A) = c_1p_1(A) + c_2p_2(A) = 0$. Für $q \in K[x]$ gilt:
$(qp_1)(A) = q(A)p_1(A) = q(A) \cdot 0 = 0$, also $q \cdot p \in Kern(s)$.

(b) Analog zu (a). qed.

20.4 Definition:

(a) Sei $I = Kern(\sigma) = \{p \in K[x] \mid p(f) = 0\}$. Dann ist I ein Hauptideal von $K[x]$, d.h. es gibt ein $\mu_f \in K[x]$ mit $I =< \mu_f >= K[x] \cdot \mu_f$.
Das Polynom μ_f ist dabei eindeutig bestimmt, wenn man festlegt, dass es normiert sein soll, das heißt, dass für den Gradkoeffizienten $LC(\mu_f) = 1$ gelten soll.
Das Polynom μ_f heißt das Minimalpolynom von f.

(b) Entsprechend heißt das normierte Polynom $\mu_A \in K[x]$ mit $Kern(s) =< \mu_A >$ das Minimalpolynom von A.

20.5 Beispiele:

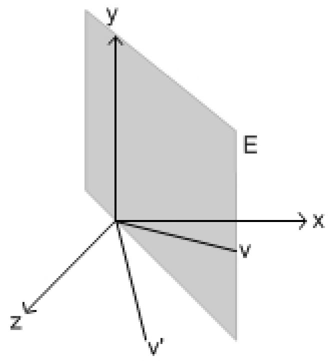
(a) Sei $A = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \in Mat_n(K)$ die Nullmatrix. Dann gilt $\mu_A = x$.

(b) Sei $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \in Mat_3(K)$. Berechne $A^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ und $A^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

Also gilt $x^3 \in< \mu_A >$.
Wegen $A \neq 0$ und $A^2 \neq 0$ folgt $\mu_A = x^3$.

(c) Für $A = I_m$ gilt $\mu_{I_m} = x - 1$.

(d) Sei $f : \mathbb{R}^3 \to \mathbb{R}^3$ die Spiegelung an der Ebene $E : x = z$.
Skizze:



Es gilt $f^2 = id_{\mathbb{R}^3}$, also $x^2 - 1 \in < \mu_f >$

$\Rightarrow \mu_f \in \{1, x-1, x+1, x^2-1\}$

Wegen $f \neq \pm id_{\mathbb{R}^3}$ folgt $\mu_f = x^2 - 1$.

20.6 Satz:

Für die Matrix $A = M_B(f) \in Mat_n(K)$ gilt $\mu_A = \mu_f$.

Beweis: Für jedes Polynom $p \in K[x]$ ist $p(A)$ die Darstellungsmatrix von $p(f)$ bezüglich der Basis B. Genau dann gilt $p(f) = 0$, wenn $p(A) = 0$ gilt.
Dies zeigt $Kern(s) = Kern(\sigma)$ und somit $\mu_A = \mu_f$. qed.

20.7 Satz: (Berechnung des Minimalpolynoms)

Sei $A \in Mat_n(K)$. Betrachte die folgenden Schritte:

(a) Ist $A = 0$, so gibt $\mu_A = x$ aus und stoppe.
Andernfalls setze $i = 0$.

(b) Erhöhe i um eins und berechne $I_n, A^0, A^1, \ldots, A^i$.

(c) Prüfe, ob es $c_o, \ldots, c_{i-1} \in K$ gibt mit $c_0 I_n + c_1 A + \ldots + c_{i-1}A^{i-1} + A^i = 0$.
(Betrachte hierzu A^i als sehr lange Zeile und reduziere mit dem Gauß-Verfahren.)

(d) Gibt es in (c) solche Zahlen $c_0, \ldots, c_{i-1} \in K$, so gib
$\mu_A = c_0 + c_1 x + \ldots + c_{i-1}x^{i-1} + x^i$ aus und stoppe.
Andernfalls fahre mit (b) fort.

Dies ist ein Algorithmus, der das Minimalpolynom μ_A von A berechnet.

Beweis: *Endlichkeit:* Folgt daraus, dass $A^0, A^1, \ldots, A^{n^2}$ auf jeden Fall linear abhängig sind.

Korrektheit: Folgt aus der Definition von μ_A:
Als Erzeugendes des Hauptideals $\{p \in K[x] \mid p(A) = 0\}$ ist μ_A das normierte Polynom kleinsten Grades mit $\mu_A(A) = 0$. qed.

20.8 Beispiel:

Sei $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & -1 \end{pmatrix} \in Mat_3(\mathbb{Q})$

(a) $1 \neq 0$

(b) I_n, A sind schon berechnet.

(c) I_3 und A sind linear unabhängig.

(b) $i = 2, A^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

(c) Betrachte: $\begin{array}{c} I_3 \\ A \\ A^2 \end{array} \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 2 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}$

und erhalte $A^2 - I_3 = 0, c_0 = -1, c_1 = 0$.

(d) Gib $\mu_A = x^2 - 1$ aus und stoppe.

B. Was haben χ_f und μ_f miteinander zu tun?

20.9 Definition:

Ein K-Untervektorraum $U \subseteq V$ heißt f-invariant, wenn $f(U) \subseteq U$ gilt.

20.10 Satz:

Sei $U \subseteq V$ ein f-invarianter K-Untervektorraum und sei $f|_U : U \to U$.

(a) Die Abbildung f induziert einen Endomorphismus $\begin{array}{rcl} \overline{f} : V/U & \to & V/U \\ v+U & \mapsto & f(v)+U \end{array}$.

(b) Es gilt $\chi_f = \chi_{f|_U} \cdot \chi_{\overline{f}}$.

Beweis:

(a) Sei $v_1, v_2 \in V$ mit $v_1 + U = v_2 + U$. Dann gilt $v_1 - v_2 \in U$, also
$f(v_1) - f(v_2) = f(v_1 - v_2) \in U$.
Hieraus folgt $f(v_1) + U = f(v_2) + U$. D.h. $\overline{f}$ ist wohldefiniert.
Die K-linearität von $\overline{f}$ ist klar.

(b) Wähle eine Basis $\{u_1, \dots, u_l\}$ von U.
Ergänze sie zu einer Basis $\{u_1, \dots, u_l, w_1, \dots, w_m\}$ von V.
Für $i = 1, \dots, l$ schreibe $f(u_i) = a_{1i}u_1 + \dots + a_{li}u_l$ mit $a_{ji} \in K$.
Für $i = 1, \dots, m$ schreibe $f(w_i) = b_{1i}u_1 + \dots + b_{li}u_k + \dots + c_{1i}w_1 + \dots + c_{mi}w_m$
mit $b_{ji}, c_{ji} \in K$. Die Menge $\{\overline{w_1}, \dots, \overline{w_m}\}$ mit $\overline{w_i} = w_i + U$ ist eine Basis von V/U.
Es gilt $\overline{f}(\overline{w_i}) = c_{1i}\overline{w_1} + \dots + c_{mi}\overline{w_m}$.

Insgesamt ist $A = (a_{ij}) \in Mat_l(K)$ die Matrix von $f\mid_U$ bezüglich der Basis $\{u_1, \ldots, u_l\}$ und $C = (c_{ij}) \in Mat_m(K)$ die Matrix von $\overline{f}$ bezüglich der Basis $\{\overline{w_1}, \ldots, \overline{w_m}\}$.
Die Matrix von f bezüglich der Basis $\{u_1, \ldots, u_l, w_1, \ldots, w_m\}$ ist

$M = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$ mit $B = (b_{ij})$. Also folgt:

$$\chi_f(x) = det \begin{pmatrix} A - xI_l & B \\ 0 & C - xI_m \end{pmatrix} = det(A - xI_l) \cdot det(C - xI_m) = \chi_{f|_U}(x) \cdot \chi_{\overline{f}}(x)$$ qed.

20.11 <u>Theorem:</u> (Der Satz von Cayley-Hamilton)

Es gilt $\chi_f(f) = 0$.

<u>Beweis:</u> Sei $\chi_f(x) = c_0 + c_1x + \ldots + c_nx^n$ mit $c_i \in K$.
Für $g = \chi_f(f) = c_0 \cdot id_V + c_1f + \ldots + c_nf^n$ müssen wir also $g = 0$ beweisen. Sei $v \in V \setminus \{0\}$.

<u>1. Schritt:</u> Konstruiere einen f-invarianten K-Untervektorraum $U \subseteq V$ wie folgt:
Sei $u_1 = v, u_2 = f(u_1) = f(v), u_3 = f(u_2) = f^2(v), \ldots$. Dann sei $1 \leq l \leq n$ mit $u_{l+1} \in < u_1, \ldots, u_l >$. Wähle die kleinste solche Zahl l und schreibe $u_{l+1} = a_1u_1 + \ldots + a_lu_l$ mit $u_i \in K$.
Setze $U =< u_1, \ldots, u_l >$. Dies ist ein f-invarianter K-Untervektorraum, denn $f(u_i) = u_{i+1} \in U$ für $i = 1, \ldots, l-1$ und $f(u_l) = u_{l+1} = a_1u_1 + \ldots + a_lu_l \in U$.

<u>2. Schritt:</u> Zeige $(\chi_{f|_U})(f\mid_u)(v) = 0$.
Die Matrix von $f\mid_U : U \to U$ bezüglich der Basis $\{u_1, \ldots, u_l\}$ ist

$$A = \begin{pmatrix} 0 & 0 & \cdots & \cdots & 0 & a_1 \\ 1 & 0 & & & \vdots & \vdots \\ 0 & 1 & \ddots & & \vdots & \vdots \\ \vdots & 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & & \ddots & 0 & \vdots \\ 0 & 0 & \cdots & 0 & 1 & a_l \end{pmatrix}$$

Dann folgt $\chi_{f|_U}(x) = \chi_A(x) = (-1)^l[x^l - a_lx^{l-1} - \ldots - a_2x - a_1]$
$\Rightarrow (-1)^l(\chi_{f|_U})(f\mid_U)(v) = (f^l - a_lf^{l-1} - \ldots - a_1)(v)$
$= -a_1u_1 - a_2u_2 - \ldots - a_lu_l + u_{l+1} = 0$

<u>3. Schritt:</u> Zeige $g(v) = (\chi_f)(f)(v) = 0$.
Nach Satz 20.10 gilt $\chi_f(f) = \chi_{\overline{f}}(f) \cdot \chi_{f|_U}(f)$ und somit
$g(v) = (\chi_f)(f)(v) = (\chi_{\overline{f}}(f))(\chi_{f|_U}(f\mid_U)(v)) = (\chi_{\overline{f}}(f))(0) = 0$ qed.

20.12 <u>Korollar:</u>

Das Polynom χ_f ist ein Vielfaches von μ_f.

<u>Beweis:</u> Nach dem Theorem gilt $\chi_f \in \{p \in K[x] \mid p(f) = 0\}$ und μ_f erzeugt dieses Ideal. qed.

20.13 <u>Beispiel:</u>

(a) Für $f = id_V$ gilt $\chi_f(x) = (1 - x)^n$ und $\mu_f(x) = x - 1$.

(b) Für $f = 0$ gilt $\chi_f(x) = (-1)^nx^n$ und $\mu_f(x) = x$.

(c) Für die Spiegelung $f : \mathbb{R}^2 \to \mathbb{R}^2$ an der x-Achse gilt:

$M_E(f) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, denn $f(e_1) = e_1$ und $f(e_2) = -e_2$,

also $\chi_f(x) = (1-x)(-1-x) = x^2 - 1$ und $\mu_f = x^2 - 1$, denn $\mu_f \mid \chi_f$.

In der Tat gilt $M_E(f)^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ und $M_E(f)^2 - I_2 = 0$.

Insgesamt sehen wir also, dass μ_f ein echter oder unechter Teiler von χ_f sein kann.

20.14 Satz:

Die Nullstellen von μ_f sind genau die Nullstellen von χ_f, also genau die Eigenwerte von f. (Aber in χ_f können die Linearfaktoren $(x - \lambda_i)$ eine höhere Vielfachheit haben.)

Beweis: Nach Korollar 20.12 ist jede Nullstelle von μ_f auch Nullstelle von χ_f. Wir müssen also noch zeigen, dass jede Nullstelle λ von χ_f, also jeder Eigenwert λ von f, auch Nullstelle von μ_f ist.

Sei $v \in V \setminus \{0\}$ ein Eigenvektor zum Eigenwert λ. Schreibe
$\mu_f = c_0 + c_1 x + \ldots + c_l x^l$ mit $c_i \in K$. Dann gilt:
$0 = \mu_f(f)(v) = (c_0 + c_1 f + \ldots + c_l f^l)(v) = c_0 + c_1\lambda v + c_2\lambda^2 v + \ldots + c_l\lambda^l v$
$= \underbrace{\mu_f(\lambda)}_{\in K} \cdot \underbrace{v}_{\in V \setminus \{0\}}$ Wegen $v \neq 0$ folgt $\mu_f(\lambda) = 0$. qed.

C. Was weiß das Minimalpolynom denn nun wirklich?

20.15 Theorem: (2. Charakterisierung der Diagonalisierbarkeit)

Ein Endomorphismus $f : V \to V$ ist genau dann diagonalisierbar, wenn μ_f ein Produkt paarweise verschiedener Linearfaktoren ist.
In diesem Fall gilt $\mu_f = (x - \lambda_1) \cdots (x - \lambda_m)$, wobei $\{\lambda_1, \ldots, \lambda_m\}$ genau die (paarweise verschiedenen) Eigenwerte von f sind.

Beweis: „$\Rightarrow$“ Sei f diagonalisierbar.
Nach der 1. Charakterisierung der Diagonalisierbarkeit gilt $\chi_f = (-1)^n (x - \lambda_1)^{\mu_1} \cdots (x - \lambda_m)^{\mu_m}$, wobei $\{\lambda_1, \ldots, \lambda_m\}$ die verschiedenen Eigenwerte sind. Es genügt zu zeigen, dass $(f - \lambda_1 \cdot id_V) \cdots (f - \lambda_m \cdot id_V) = 0$ gilt. (Dann folgt, dass μ_f das Produkt $(x - \lambda_1) \cdots (x - \lambda_m)$ teilt. Nach Satz 20.14 ist μ_f gleich diesem Produkt.)
Wähle eine Basis B von V bestehend aus Eigenvektoren von f und erhalte:

$$M_B(f) = \left(\begin{array}{ccc|c|ccc} \lambda_1 & & & & & & \\ & \ddots & & 0 & & 0 & \\ & & \lambda_1 & & & & \\ \hline & 0 & & \ddots & & 0 & \\ \hline & & & & \lambda_m & & \\ & 0 & & 0 & & \ddots & \\ & & & & & & \lambda_m \end{array}\right)$$

$$\underbrace{\qquad\qquad}_{\mu_1 \text{ Spalten}} \qquad \underbrace{\qquad\qquad}_{\mu_m \text{ Spalten}}$$

Die Elemente auf der Hauptdiagonalen von $(A - \lambda_1 I_n) \cdots (A - \lambda_m I_n)$ sind $(\lambda_i - \lambda_1) \cdots (\lambda_i - \lambda_m) = 0$. Also folgt $(A - \lambda_1 I_n) \cdots (A - \lambda_m I_n) = 0$ und somit $(f - \lambda_1 \cdot id_V) \cdots (f - \lambda_m \cdot id_V) = 0$.

„$\Leftarrow$" Sei nun $\mu_f = (x - \lambda_1) \cdots (x - \lambda_m)$, wobei $\{\lambda_1, \ldots, \lambda_m\}$ nach Satz 20.14 genau die verschiedenen Eigenwerte von f sind.

<u>Zwischenbehauptung:</u> $V = Kern(f - \lambda_1 \cdot id_V) \oplus Bild(f - \lambda_1 \cdot id_V)$

<u>Beweis:</u> Teile $(x - \lambda_2) \cdots (x - \lambda_m)$ mit Rest durch $(x - \lambda_1)$ und erhalte $(x - \lambda_2) \cdots (x - \lambda_m) = q \cdot (x - \lambda_1) + r$ mit $r \in K$. Wegen $\lambda_i \neq \lambda_1$ für $i \neq 1$ ist $r \neq 0$.
Für $v \in V$ folgt $rV = (f - \lambda_2 \cdot id_V) \cdots (f - \lambda_m \cdot id_V)(v) - q(f)(f - \lambda_1 \cdot id_V)(v)$
Hierbei gilt $(f - \lambda_1 \cdot id_V)[(f - \lambda_2 \cdot id_V) - (f - \lambda_m \cdot id_V)(v)] = (\mu_f(f))(v) = 0$
Dies zeigt $(f - \lambda_2 \cdot id_V) \cdots (f - \lambda_m \cdot id_V)(v) \in Kern(f - \lambda_1 \cdot id_V)$. Ferner gilt: $q(f)(f - \lambda_1 \cdot id_V)(v) = (f - \lambda_1 \cdot id_V)(q(f)(v)) \in Bild(f - \lambda_1 \cdot id_V)$.
Insgesamt folgt $r \cdot v \in Kern(f - \lambda_1 \cdot id_V) + Bild(f - \lambda_1 \cdot id_V)$, also $V = Kern(f - \lambda_1 \cdot id_V) + Bild(f - \lambda_1 \cdot id_V)$.
Aus der Dimensionsformel für lineare Abbildungen folgt:
$dim_K(V) = dim_K(Kern(f - \lambda_1 \cdot id_V)) + dim_K(Bild(f - \lambda_1 \cdot id_V))$ und somit $dim_K(Kern(f - \lambda_1 \cdot id_V)) \cap Bild(f - \lambda_1 \cdot id_V)) = 0$. Dies zeigt die Zwischenbehauptung.

Nun zeigen wir die Behauptung, dass f diagonalisierbar ist, mit vollständiger Induktion nach $dim_K(V) = n$.

<u>$n = 1$:</u> $\mu_f(x - \lambda_1) \Rightarrow f = \lambda_1 \cdot id_V$ ist offensichtlich diagonalisierbar.

<u>$n \to n + 1$:</u> Da λ_1 ein Eigenwert von f ist, gilt $dim_K(Kern(f - \lambda_1 \cdot id_V)) \geq 1$, also $dim_K(Bild(f - \lambda_1 \cdot id_V)) < n$ nach der Zwischenbehauptung.
Betrachte den Untervektorraum $U = Bild(f - \lambda_1 \cdot id_V)$. Jedes $u \in U$ ist vo der Form $u = f(w) - \lambda_1 \cdot w$ für ein $w \in V$. Dann gilt $f(a) = f(f(w)) - \lambda_1 f(w) = (f - \lambda_1 \cdot id_V)(f(w)) \in U$.
Daher ist U ein f-invarianter Untervektorraum.
Aus $\chi_{f|_U} \mid \chi_f$ folgt, dass $\chi_{f|_U}$ in Linearfaktoren zerfällt und somit zerfällt auch $\mu_{f|_U}$ in Linearfaktoren. Es gilt auch $\mu_{f|_U} \mid \mu_f$, d.h. $\mu_{f|_U}$ zerfällt in paarweise verschiedene Linearfaktoren.
Nun können wir die Induktionsvoraussetzung auf die Abbildung $f\mid_U : U \to U$ anwenden und erhalten, dass $f\mid_U$ diagonalisierbar ist.
Wähle eine Basis $\{u_1, \ldots, u_l\}$ von U bestehend aus Eigenvektoren von f.
Nach der Zwischenbehauptung gibt es eine Basis $\{w_1, \ldots, w_k\}$ von $Kern(f - \lambda_1 \cdot id_V) = Eig(f, \lambda_1)$, so dass $\{u_1, \ldots, u_l, w_1, \ldots, w_k\}$ eine Basis von V ist. Diese Basis besteht aus Eigenvektoren von f. qed.

20.16 <u>Korollar:</u> (Diagonalisierbarkeitstest)

Sei $A \in Mat_n(K)$. Betrachte die folgenden Schritte:

(a) Berechne $\chi_A = det(A - xI_n)$. Teste, ob χ_A in Linearfaktoren zerfällt.
Ist dies nicht der Fall, so gib „A ist nicht diagonalisierbar" aus und stoppe.

(b) Schreibe $\chi_A(x) = (x - \lambda_1)^{\mu_1} \cdots (x - \lambda_m)^{\mu_m}$.
Prüfe nun, ob $(A - \lambda_1 I_n) \cdots (A - \lambda_m I_n)$ die Nullmatrix ist.
Ist dies nicht der Fall, so gib „A ist nicht diagonalisierbar" aus und stoppe.

(c) Gib „A ist diagonalisierbar" aus und stoppe.

<u>Beweis:</u> Wenn A diagonalisierbar ist und $\chi_A(x) = (x - \lambda_1)^{\mu_1} \cdots (x - \lambda_m)^{\mu_m}$, dann muss nach dem Theorem $\mu_A(x) = (x - \lambda_1) \cdots (x - \lambda_m)$ sein. Dies kann man prüfen, indem man $\mu_A(A) = 0$ nachprüft. qed.

20.17 Beispiel:

Wir testen die Diagonalisierbarkeit der Abbildung $f : \mathbb{Q}^3 \to \mathbb{Q}^3$ mit

$M_E(f) = \begin{pmatrix} 3 & 4 & 3 \\ -1 & 0 & -1 \\ 1 & 2 & 3 \end{pmatrix}$ 1.) $\chi_f = det \begin{pmatrix} 3-x & 4 & 3 \\ -1 & -x & -1 \\ 1 & 2 & 3-x \end{pmatrix}$

$= (3-x)[x^2 - 3x + 2] - 4[x-2] + 3[x-2] = (x-2)[(3-x)(x-1) - 1]$
$= (x-2)(-x^2 + 4x + 4) = -(x-2)^3$

2.) Teste, ob $\mu_f = x - 2$ gilt: $M_E(f) = 2 \cdot I_3 \neq 0$, also ist f nicht diagonalisierbar.

21 Wie man wenigstens teilweise aufräumt - Eine kleine Putzanleitung (Trigonalisierbarkeit)

Sei K ein Körper, V ein n-dimensionaler K-Vektorraum mit Basis $B = \{v_1, \ldots, v_n\}$.
Sei A eine obere Dreiecksmatrix, d.h. wenn $A = (a_{ij})$ ist, so gilt $a_{ij} = 0$, falls $i > j$, also

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}$$

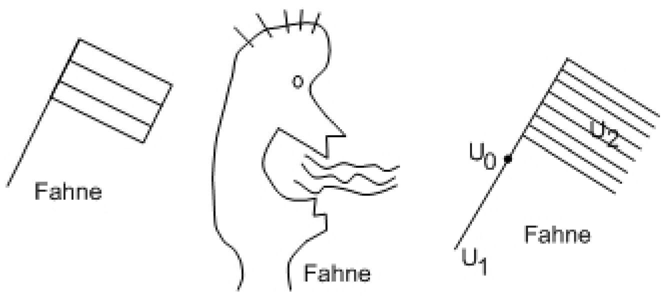
21.1 Bemerkung:

Sei $A = M_B(f)$ eine obere Dreiecksmatrix. Für $i = 1, \ldots, n$ betrachte $U_i =< v_1, \ldots, v_i >$.
Es gilt $f(v_i) = a_{1i}v_1 + \ldots + a_{ii}v_i \in U$.

21.2 Definition:

(a) Eine Kette von Untervektorräumen $\{0\} = U_0 \subseteq U_1 \subseteq \ldots \subseteq U_n = V$ mit $dim_K(U_i) = i$ heißt eine Fahne von V.

(b) Eine Fahne $U_0 \subseteq \ldots \subseteq U_n = V$ mit $f(U_i) \subseteq U_i$ für $i = 0, \ldots, n$ heißt eine f-invariante Fahne von V.
Skizze:



21.3 Satz:

Sei $f : V \to V$ ein Endomorphismus.
Genau dann gibt es eine f-invariante Fahne $\{0\} = U_0 \subseteq \ldots \subseteq U_n = V$, wenn es eine Basis B von V gibt, so dass $M_B(f)$ eine obere Dreiecksmatrix ist.

Beweis: „$\Rightarrow$“ Durch Basisergänzung erhalten wir $U_i =< v_1, \ldots, v_i >= U_{i-1}+ < v_i >$ für $i = 1, \ldots, n$. Bezüglich der Basis $B = \{v_1, \ldots, v_n\}$ besitzt $M_B(f)$ dann eine obere Dreiecksgestalt.

„$\Leftarrow$“ Sei $B = \{v_1, \ldots, v_n\}$. Nach Voraussetzung ist $U_i =< v_1, \ldots, v_i >$ ein f-invarianter Untervektorraum mit $dim_K(U_i) = i$. Also ist $U_0 \subseteq \ldots \subseteq U_n = V$ eine f-invariante Fahne. qed.

21.4 Definition:

(a) Ein Endomorphismus $f : V \to V$ heißt trigonalisierbar, wenn es eine Basis B von V gibt, so dass $M_B(f)$ eine obere Dreiecksmatrix ist.

(b) Eine Matrix $A \in Mat_n(K)$ heißt trigonalisierbar, wenn es eine invertierbare Matrix $S \in GL_n(K)$ gibt, so dass SAS^{-1} eine obere Dreiecksmatrix ist.

21.5 Satz: (Trigonalisierungssatz)

Ein Endomorphismus ist genau dann trigonalisierbar, wenn χ_f in Linearfaktoren zerfällt.

Beweis: „$\Rightarrow$“ Sei B eine Basis von V, so dass $M_B(f)$ eine obere Dreiecksmatrix ist. Schreibe $M_B(f) = (a_{ij})$. Dann gilt:

$$\chi_f = det \begin{pmatrix} a_{11}-x & a_{12} & \cdots & a_{1n} \\ & \ddots & \ddots & \vdots \\ & & \ddots & a_{n-1,n} \\ 0 & & & a_{nn}-x \end{pmatrix} = (a_{11}-x)\cdots(a_{nn}-x) = (-1)^n(x-a_{11})\cdots(x-a_{nn}).$$

„$\Leftarrow$“ Wir schließen mit vollständiger Induktion nach $dim_K(V) = n$.

$n = 1$: klar

$n \to n+1$: Sei $v_1 \in V \setminus \{0\}$ ein Eigenvektor zum Eigenwert λ_1. Ergänze v_1 zu einer Basis $B = \{v_1, v_2, \ldots, v_n\}$ von V. Dann gilt:

$$M_B(f) = \left(\begin{array}{c|ccc} \lambda_1 & a_{12} & \cdots & a_{1n} \\ \hline 0 & & & \\ \vdots & & \tilde{A} & \\ 0 & & & \end{array}\right) \text{ mit } \tilde{A} = \begin{pmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

Sei $W =< v_2, \ldots, v_n >$ und $g : W \to W$, $v_j \mapsto a_{2j}v_2 + \ldots + a_{nj}v_n$.

Aus $\chi_f = (\lambda_1 - x) \cdot \chi_g$ folgt, dass χ_g auch in Linearfaktoren zerfällt, d.h. $\chi_g = (-1)^{n-1}(x-\lambda_1)^{\mu_1-1}(x-\lambda_2)^{\mu_2}\cdots(x-\lambda_m)^{\mu_m}$. Nach der Induktionsvoraussetzung ist g trigonalisierbar, d.h. es gibt eine g-invariante Fahne $\{0\} = \tilde{U_0} \subseteq \tilde{U_1} \subseteq \ldots \subseteq \tilde{U}_{n-1} = W$.

Behauptung: Durch $U_i = \tilde{U}_{n-1} + < v_1 >$ für $i = 2, \ldots, n$ und $U_1 =< v_1 >$ ist eine f-invariante Fahne von V gegeben. (Nach Satz 21.3 ist dann die Behauptung des Satzes bewiesen.)

Beweis: Es gilt $f(a_1v_1 + w) = \lambda_1 a v_1 + h(w) + g(w)$ mit $a \in K, w \in W$

und $h : W \to < v_1 >$, $v_i \mapsto a_{1j}v_1$

Dies zeigt $f(a_1v_1 + w) \in \widetilde{U_{i-1}} + < v_1 >$, falls $w \in \tilde{U}_{n-1}$ gilt.
Also gilt $f(U_i) \subseteq U_i$, d.h. die Fahne $\{0\} = U_0 \subseteq \ldots \subseteq U_n = V$ ist f-invariant. qed.

21.6 <u>Korollar:</u> (Trigonalisierungsalgorithmus)

Sei $B = \{v_1, \ldots, v_n\}$ eine Basis von V und es gelte $\chi_f = (-1)^n(x-\lambda_1)^{\mu_1} \cdots (x-\lambda_m)^{\mu_m}$. Betrachte die folgenden Instruktionen:

(a) Ist $n = 1$, so gib $S = (1) \in Mat_1(K)$ aus und stoppe.

(b) Bestimme einen Eigenvektor $\tilde{v_1} \in V \setminus \{0\}$ zum Eigenwert λ_1 und ergänze ihn durch Elemente von B zu einer Basis von $\tilde{B} = \{\tilde{v_1}, \ldots, \tilde{v_n}\}$.

(c) Schreibe $M_{\tilde{B}}(f) = \begin{pmatrix} \lambda_1 & a_{12} & \cdots & a_{1n} \\ 0 & & & \\ \vdots & & \tilde{A} & \\ 0 & & & \end{pmatrix}$ und berechne rekursiv eine invertierbare Matrix $\tilde{S} \in GL_{n-1}(K)$, so dass $\tilde{S}\tilde{A}\tilde{S}^{-1}$ eine obere Dreiecksmatrix ist.

(d) Gib die Matrix $S = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \tilde{S} & \\ 0 & & & \end{pmatrix} \cdot T^B_{\tilde{B}}$ aus und stoppe.

Dies ist ein Algorithmus, der eine Matrix $S \in GL_n(K)$ berechnet, so dass $SM_B(f)S^{-1}$ eine obere Dreiecksmatrix ist.

<u>Beweis:</u> $n = 1$: klar

$n \to n+1$: Sei $a = (a_{12}, \ldots, a_{1n})$ und $\tilde{S}^{-1} = (\underbrace{s_1, \ldots, s_{n-1}}_{\text{Spalten}})$. Dann gilt:

$$SM_B(f)S^{-1} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \tilde{S} & \\ 0 & & & \end{pmatrix} \cdot \underbrace{T^B_{\tilde{B}} \cdot M_B(f) \cdot T^{\tilde{B}}_B}_{M_{\tilde{B}}(f)} \cdot \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \tilde{S}^{-1} & \\ 0 & & & \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \tilde{S} & \\ 0 & & & \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 & a_{12} & \cdots & a_{1n} \\ 0 & & & \\ \vdots & & \tilde{A} & \\ 0 & & & \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \tilde{S}^{-1} & \\ 0 & & & \end{pmatrix}$$

$$= \begin{pmatrix} \lambda_1 & a_{12} & \cdots & a_{1n} \\ 0 & & & \\ \vdots & & \tilde{S}\tilde{A} & \\ 0 & & & \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \tilde{S}^{-1} & \\ 0 & & & \end{pmatrix} = \begin{pmatrix} \lambda_1 & as_1 & \cdots & as_{n-2} \\ 0 & & & \\ \vdots & & \tilde{S}\tilde{A}\tilde{S}^{-1} & \\ 0 & & & \end{pmatrix}$$

und dies ist eine obere Dreiecksmatrix. qed.

21.7 Beispiel:

Betrachte die Matrix $A = \begin{pmatrix} 3 & 4 & 3 \\ -1 & 0 & -1 \\ 1 & 2 & 3 \end{pmatrix} \in Mat_3(\mathbb{Q})$.

Es gilt $\chi_A(x) = -(x-2)^3$ und $dim_Q(Eig(A,2)) = 2$. Die ursprüngliche Basis ist $B = \{e_1, e_2, e_3\}$
(b) Berechne $v_1 = (1,-1,1)$ und $\tilde{B} = \{v_1, e_2, e_3\}$ Es folgt

$$T_B^{\tilde{B}} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \text{ und } T_{\tilde{B}}^B = (T_B^{\tilde{B}})^{-1} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

(c) Neue Matrix $T_{\tilde{B}}^B A T_B^{\tilde{B}} = \left(\begin{array}{c|cc} 2 & 4 & 3 \\ \hline 0 & 4 & 2 \\ 0 & -2 & 0 \end{array}\right)$ und somit $\tilde{A} = \begin{pmatrix} 4 & 2 \\ -2 & 0 \end{pmatrix}$

(b) Berechne $\widehat{v_1} = (1,-1)$ und $\widehat{B} = \{\widehat{v_1}, \widehat{e_2}\}$.

$$\Rightarrow T_B^{\{\widehat{e_1},\widehat{e_2}\}} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}, T_{\{\widehat{e_1},\widehat{e_2}\}}^B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

(d) $\widehat{S} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$

(d) $S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}, S^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix}$

Also ist $SAS^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 3 & 4 & 3 \\ -1 & 0 & -1 \\ 1 & 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 3 \\ 0 & 2 & 2 \\ 0 & 0 & 2 \end{pmatrix}$.

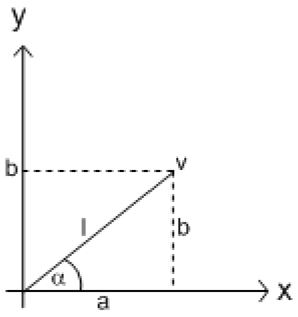
Kapitel VII: Die Vermessung der Welt (Bilinearformen und Skalarprodukte)

22 Pro und Contra Formen mit zwei Argumenten (Bilinearformen)

A. Wieso so und nicht anders?

22.1 Bemerkung: (Zur Längenmessung)

Gegeben sei ein Vektor $v = (a, b) \in \mathbb{R}^2$
Skizze:



Der Satz von Pythagoras: $l = \sqrt{a^2 + b^2}$

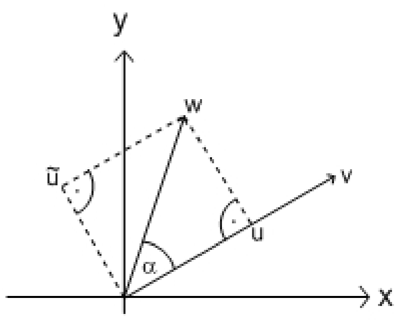
Also ist $\|v\| = \sqrt{a^2 + b^2}$ die Länge von v.

22.2 Definition:

Zu einem Vektor $v = (a, b) \in \mathbb{R}^2 \setminus \{0\}$ heißt $e_v = \dfrac{v}{\|v\|}$ der Einheitsvektor in Richtung v.

22.3 Bemerkung: (Zur Definition der Winkelmessung))

Sei $v = (a, b)$ und $w = (c, d)$ in $\mathbb{R}^2 \setminus \{0\}$
Skizze:



Gesucht ist eine „Formel“ für $\alpha = \measuredangle(v, w)$,

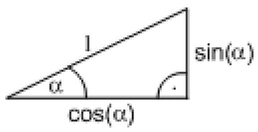
die nur von den Koordinaten a, b, c, d abhängt.

Wir suchen als „Hilfsfunktion“ eine Abbildung $\Phi : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$

mit deren Hilfe wir Winkel messen können.

(a) Die Frage, ob zwei Vektoren aufeinander „senkrecht“ stehen, sollte nicht von der Länge abhängen. Also: Sind v und w senkrecht zueinander, so gilt $\Phi(v, w) = 0$

(b) Definition von $cos(\alpha)$:



Dies liefert $u = \|w\| \cdot cos(\alpha) \cdot e_v$. Nach der Definition der Vektoraddition gilt $w = u + \tilde{u}$. Hätte man nun die Eigenschaft $\Phi(v, u + \tilde{u}) = \Phi(v, u) + \Phi(v, \tilde{u})$, so könnte man (a) verwenden (also $\Phi(v, \tilde{u}) = 0$) und hätte $\Phi(v, w) = \Phi(v, u)$.
Dies liefert $\Phi(v, w) = \Phi(\|v\| \cdot e_v, \|w\| \cdot cos(\alpha) \cdot e_v)$

(c) Nun hätten wir gerne die Eigenschaften $\Phi(\lambda v, w) = \Phi(v, \lambda w) = \lambda\Phi(v, w)$ für $\lambda \in \mathbb{R}$.
Dann wäre $\Phi(v, w) = \|v\| \cdot \|w\| \cdot cos(\alpha) \cdot \Phi(e_v, e_v)$

(d) Hätte man nun $\Phi(v, v) = \|v\|^2$, so wäre $\Phi(e_v, e_v) = \Phi(\frac{v}{\|v\|}, \frac{v}{\|v\|}) = \frac{1}{\|v\|^2} \cdot \Phi(v, v) = 1.$
und somit $\Phi(v, w) = \|v\| \cdot \|w\| \cdot cos(\alpha)$.

Also könnte man den Winkel α mit $\cos(\alpha) = \dfrac{\Phi(v, w)}{\|v\| \cdot \|w\|}$ messen.
Nun hätten wir gerne eine solche Funktion Φ und eine schöne Formel für $\Phi(v, w)$ in Abhängigkeit der Koordinaten a, b, c, d.

(a) Sind v und w senkrecht zueinander, so gilt $\Phi(v, w) = 0$

(b) $\Phi(v, w_1 + w_2) = \Phi(v, w_2) + \Phi(v, w_2)$
Analog sollte auch $\Phi(v_1 + v_2, w) = \Phi(v_1, w) + \Phi(v_2, w)$ gelten.

(c) $\Phi(\lambda v, w) = \Phi(v, \lambda w) = \lambda\Phi(v, w)$ für $\lambda \in R$.
((b) & (c) $\Rightarrow \Phi$ ist eine lineare Funktion von v und w.)

(d) $\Phi(v, v) = \|v\|^2$

Also kann man aus diesen Bedingungen eine Formel für Φ ableiten?

Es gilt $\Phi(v + w, v + w) = \|v + w\|^2 = \|(a + c, b + d)\|^2 = a^2 + c^2 + b^2 + d^2 + 2(ac + bd)$
und $\Phi(v + w, v + w) = \Phi(v, v + w) + \Phi(w, v + w) = \Phi(v, v) + \Phi(v, w) + \underbrace{\Phi(w, v)}_{= \Phi(v, w)} + \Phi(w, w)$ (d.h. Φ <u>symmetrisch</u>)
$= \|(a, b)\|^2 + \|(c, d)\|^2 + a^2 + b^2 + c^2 + d^2 + 2\Phi(v, w)$
Also müsste $\Phi(v, w) = ac + bd$ sein.

Nun prüft man leicht nach, dass (a)-(d) in der Tat gelten.
In $\mathbb{R}^3$ erhält man entsprechend für $v = (a_1, a_2, a_3)$ und $w = (b_1, b_2, b_3)$ die Forderungen $\|v\| = \sqrt{a_1 + a_2 + a_4}$ und $\Phi(v, w) = a_1b_1 + a_2b_2 + a_3b_3$.

B. Allgemeiner, allgemeiner!

Im Folgenden sei K ein Körper und V ein (n-dimensionaler) K-Vektorraum.

22.4 Definition:

(a) Eine Abbildung $\Phi : V \times V \to K$ heißt <u>bilinear</u> oder <u>Bilinearform</u>, wenn folgendes gilt:

1.) $\Phi(v, a_1w_1 + a_2w_2) = a_1\Phi(v, w_1) + a_2\Phi(v, w_2)$ für $a_1, a_2 \in K$ und $v, w_1, w_2 \in V$.
2.) $\Phi(a_1v_1 + a_2v_2, w) = a_1\Phi(v_1, w) + a_2\Phi(v_2, w)$ für $a_1, a_2 \in K$ und $v_1, v_2, w \in V$.

(b) Eine Bilinearform Φ heißt <u>symmetrisch</u>, wenn für alle $v, w \in V$ gilt $\Phi(v, w) = \Phi(w, v)$.

22.5 Beispiel:

Sei $B = \{v_1, \dots, v_n\}$ eine Basis von V. Für $v = a_1v_1 + \dots + a_nv_n$ und $w = b_1v_1 + \dots + b_nv_n$ mit $a_i, b_j \in K$ sei $\Phi(v,w) = a_1b_1 + a_2b_2 + \dots + a_nb_n$.

Dann ist $\begin{array}{rcl} \Phi : V \times V & \to & K \\ (v,w) & \mapsto & a_1b_1 + \dots + a_nb_n \end{array}$ eine symmetrische Bilinearform auf V.

Sie heißt die Standardbilinearform bezüglich der Basis B.

22.6 Definition:

Sei $B = \{v_1, \dots, v_n\}$ eine Basis von V und $\Phi : V \times V \to K$ eine Bilinearform auf V.
Dann heißt $G_B(\Phi) = (\Phi(v_i, v_j))_{i,j} \in Mat_n(K)$ die Gramsche Matrix (oder die Strukturmatrix) von Φ bezüglich der Basis B.

22.7 Satz:

Sei B eine Basis von V.

(a) Zu einer Matrix $A = (a_{ij}) \in Mat_n(K)$ kann man eine Bilinearform $\Phi_A : V \times V \to K$ definieren durch $\Phi_A(b_1v_1 + \dots + b_nv_n, c_1v_1 + \dots + c_nv_n) = (b_1, \dots, b_n) \cdot A \cdot \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in K$

(b) Die beiden Zuordnungen $A \to \Phi_A$ und $\Phi \to G_B(\Phi)$ sind invers zueinander.

(c) Die Menge $Bil(V)$ aller Bilinearformen auf V ist ein K-Vektorraum bezüglich $(\Phi + \Psi)(v,w) = \Phi(v,w) + \Psi(v,w)$ und $(\lambda \cdot \Phi)(v,w) = \lambda(\Phi(v,w))$

(d) Die Abbildung $\begin{array}{rcl} \varphi : Bil(V) & \overset{\cong}{\to} & Mat_n(K) \\ \Phi & \mapsto & G_B(\Phi) \end{array}$ ist ein Isomorphismus von K-Vektorräumen.

(e) Die Standardbilinearform Φ erfüllt $G_B(\Phi) = I_n$.

Beweis:

(a) Zu zeigen: Φ_A ist bilinear. Seien $b_i, b'_j, c_k \in K$. Dann gilt:
$\Phi_A((b_1v_1 + \dots + b_nv_n) + (b'_1v_1 + \dots + b'_nv_n), c_1v_1 + c_nv_n)$

$$= \Phi_A((b_1 + b'_1)v_1 + \dots + (b_n + b'_n)v_n, c_1v_1 + \dots + c_nv_n) = (b_1 + b'_1, \dots, b_n + b'_n) \cdot A \cdot \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$

$$= (b_1, \dots, b_n) \cdot A \cdot \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} + (b'_1, \dots, b'_n) \cdot A \cdot \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$

$$= \Phi_A(b_1v_1 + \dots + b_nv_n, c_1v_1 + \dots + c_nv_n) + \Phi_A(b'_1v_1 + \dots + b'_nv_n, c_1v_1 + \dots + c_nv_n)$$

Für $\lambda \in K$ gilt $\Phi_A(\lambda(b_1v_1 + \dots + b_nv_n), c_1v_1 + \dots + c_nv_n)$

$$= (\lambda b_1, \dots, \lambda b_n) \cdot A \cdot \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = \lambda \Phi_A(b_1v_1 + \dots + b_nv_n, c_1v_1 + \dots + c_nv_n)$$

Die Linearität im zweiten Argument folgt analog.

(b) <u>Zu zeigen:</u> $G_B(\Phi_A) = A$. Es gilt: $G_B(\Phi_A) = (\Phi_A(v_i, v_j))_{i,j} = (e_i A e_j^{tr})_{i,j}$

$$= \left((0, \ldots, 0, 1, 0, \ldots, 0) \cdot \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \right)_{i,j} = (a_{ij})_{i,j} = A$$

<u>Zu zeigen:</u> $\Phi_{G_B(\Phi)} = \Phi$ Sei $v = b_1 v_1 + \ldots + b_n v_n$ und $w = c_1 v_1 + \ldots + c_n v_n \in V$

Dann gilt: $\Phi_{G_B(\Phi)}(v, w) = \underbrace{(b_1, \ldots, b_n)}_{\sum_i b_i e_i} G_B(\Phi) \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = \sum_{i=1}^n \sum_{j=1}^n b_i c_j (e_i \cdot G_B(\Phi) \cdot e_j^{tr})$

$= \sum_{i=1}^n \sum_{j=1}^n b_i c_j \Phi(v, w) = \Phi(b_1 v_1 + \ldots + b_n v_n, c_1 v_1 + \ldots + c_n v_n) = \Phi(v, w)$.

(c) Man prüft leicht nach, dass die so definierten Abbildungen $\Phi + \Psi$ und $\lambda \Phi$ wieder Bilinearformen auf V sind.

(d) Φ ist nach (b) bijektiv. zz: φ ist K-linear:
Für $\Phi, \Psi \in Bil(V)$ ist zu zeigen $\varphi(\Phi + \Psi) = \varphi(\Phi) + \varphi(\Psi)$ Es gilt:
$\varphi(\Phi + \Psi) = G_B(\Phi + \Psi) = ((\Phi + \Psi)(v_i, v_j))_{i,j} = (\Phi(v_i, v_j) + \Psi(v_i, v_j))_{i,j}$
$= (\Phi(v_i, v_j))_{i,j} + (\Psi(v_i, v_j))_{i,j} = G_B(\Phi) + G_B(\Psi) = \varphi(\Phi) + \varphi(\Psi)$ und
$\varphi(\lambda \Phi) = G_B(\lambda \Phi) = ((\lambda \Phi)(v_i, v_j))_{i,j} = (\lambda \Phi(v_i, v_j))_{i,j} = \lambda (\Phi(v_i, v_j))_{i,j} = \lambda G_B(\Phi) = \lambda \varphi(\Phi)$.

(e) Es gilt: $G_B(\Phi) = (\Phi(v_i, v_j))_{i,j} = (\delta_{ij})_{i,j} = I_n$ qed.

<u>Ziel:</u> Stelle eine Basistransformationsformel für Bilinearformen auf

22.8 <u>Satz:</u> (Basistransformationsformel für Bilinearformen)

Seien B, C zwei Basen von V und sei $\Phi : V \times V \to K$ eine Bilinearform. Dann gilt:
$G_B(\Phi) = (T_C^B)^{tr} G_C(\Phi) T_C^B$

<u>Beweis:</u> Sei $B = \{v_1, \ldots, v_n\}$ und $C = \{w_1, \ldots, w_n\}$.
Seien $v, w \in V$ mit $v = a_1 v_1 + \ldots + a_n v_n = a'_1 w_1 + \ldots + a'_n w_n$ und
$w = b_1 v_1 + \ldots + b_n v_n = b'_1 w_1 + \ldots + b'_n w_n$ mit $a_i, a'_j, b_k, b'_l \in K$. Es gilt:

$$\begin{pmatrix} a'_1 \\ \vdots \\ a'_n \end{pmatrix} = T_C^B \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \text{ und } \begin{pmatrix} b'_1 \\ \vdots \\ b'_n \end{pmatrix} = T_C^B \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

Also folgt: $(a_1, \ldots, a_n) \cdot G_B(\Phi) \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \Phi(v, w) = (a'_1, \ldots, a'_n) \cdot G_C(\Phi) \cdot \begin{pmatrix} b'_1 \\ \vdots \\ b'_n \end{pmatrix}$

$= (a_1, \ldots, a_n) \cdot (T_C^B)^{tr} \cdot G_C(\Phi) \cdot T_C^B \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$

Setzt man nun für v und w die Basisvektoren v_i, w_j ein so folgt:
$(G_B(\Phi))_{i,j} = ((T_C^B)^{tr} G_C(\Phi) T_C^B)_{i,j}$ für alle i, j.
Also folgt die gewünschte Gleichheit der Matrizen. qed.

22.9 Bemerkung: (Vergleich der beiden Transformationsformeln)

(a) Die Basistransformationsformel für lineare Abbildungen lautet:
$M_B(f) = (T_C^B)^{-1} M_C(f) T_C^B$

(b) Die Basistransformationsformel für Bilinearformen lautet:
$G_B(\Phi) = (T_C^B)^{tr} G_C(\Phi) T_C^B$

C. Kann man mit Bilinearformen bereits Winkel messen?

Die Definition $cos(\alpha) = \frac{\Phi(v,w)}{\|v\|\|w\|}$ macht keinen Sinn. Auch $\|v\| = \sqrt{\Phi(v,w)}$ funktioniert noch nicht richtig. Aber die Bedingung (a) kann man bereits studieren:

22.10 Definition:

Sei $\Phi : V \times V \to K$ eine Bilinearform auf V.

(a) Zwei Vektoren $v, w \in V$ heißen orthogonal (bzw. senkrecht zueinander), wenn $\Phi(v, w) = 0$ gilt. Schreibweise: $v \perp_\Phi w$

(b) Ist $v \in V$ und $U \subseteq V$, so heißt v orthogonal zu U, wenn für alle $u \in U$ gilt: $\Phi(v, w) = 0$. Schreibweise: $v \perp_\Phi U$

(c) Zwei Teilmengen $U, W \subseteq V$ heißten orthogonal, wenn für alle $u \in U$ und $w \in W$ gilt: $\Phi(u, w) = 0$. Schreibweise: $U \perp_\Phi W$.

22.11 Beispiel:

(a) Sei $\Phi : \mathbb{Q}^2 \times \mathbb{Q}^2 \to \mathbb{Q}$ die Standardbilinearform.
Sei $v = (a, b) \in \mathbb{Q}^2 \setminus \{0\}$. Für $w \in \mathbb{Q}^2$ gilt genau dann $\Phi(v, w) = 0$, wenn w von der Form $w = (\lambda b, -\lambda a)$ mit $\lambda \in \mathbb{Q}^2$ ist.

(b) Sei $\Phi : \mathbb{Q}^2 \times \mathbb{Q}^2 \to \mathbb{Q}$ die Bilinearform mit $G_E(\Phi) = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix}$ Für $v = (2, -1)$ und jeden Vektor $w = (a, b) \in \mathbb{Q}^2$ gilt $\Phi(v, w) = (2, -1) \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = 0$
Also gilt $v \perp_\Phi \mathbb{Q}^2$.

(c) Sei p eine Primzahl und $\Psi : \mathbb{F}_p^n \times \mathbb{F}_p^n \to \mathbb{F}_p$ die Standardbilinearform.
Ferner sei $n \geq p$. Der Vektor $v = (\underbrace{1, \ldots, 1}_{p \text{ Stück}}, 0, \ldots, 0)$ steht auf sich selbst senkrecht,

denn $\Phi(v,v) = (1, \ldots, 1, 0, \ldots, 0) \cdot \begin{pmatrix} 1 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \underbrace{1 + \ldots + 1}_{p \text{ mal}} = 0$

Frage: Wie kann man diese pathologischen Fälle vermeiden?

22.12 Definition:

Eine Bilinearform $\Phi : V \times V \to K$ heißt nicht ausgeartet, wenn aus $v \in V$ und $v \perp_\Phi V$ folgt, dass $v = 0$ gelten muss.
Mit anderen Worten, der Nullvektor ist der einzige, der auf allen Vektoren senkrecht steht.

22.13 Satz:

Die Standardbilinearform $\Phi : V \times V \to K$ ist nicht ausgeartet.

Beweis: Sei $v = a_1 v_1 + \ldots + a_n v_n$ mit $a_1, \ldots, a_n \in K$ und $a_i \neq 0$.

Für den Vektor $w = \frac{1}{a_i} v_i$ gilt dann $\Phi(v,w) = (a_1, \ldots, a_n) I_n \cdot \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \frac{1}{a_i} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = a_i \frac{1}{a_i} = 1 \neq 0.$

Also gibt es zu jedem $v \in V \setminus \{0\}$ ein $w \in W$ mit $\Phi(v,w) \neq 0$ qed.

22.14 Satz:

Sei $B = \{v_1, \ldots, v_n\}$ eine Basis von V. Eine Bilinearform $\Phi : V \times V \to K$ ist genau dann nicht ausgeartet, wenn ihre Gramsche Matrix $G_B(\Phi) \in Mat_n(K)$ invertierbar ist.

Beweis: „$\Rightarrow$“ Sei $G_B(\Phi) = (a_{ij})$ und $b_1, \ldots, b_n \in K$ mit $b_1(a_{11}, \ldots, a_{1n}) + \ldots + b_n(a_{n1}, \ldots, a_{nn} = 0$.
Zeige: $b_1 = \ldots = b_n = 0$. Wegen $\Phi(v_i, v_j) = a_{ij}$ erhalten wir

$$\left.\begin{array}{lcl} b_1\Phi(v_1,v_1) + \ldots + b_n\Phi(v_n,v_1) & = & 0 \\ \vdots & & \vdots \\ b_1\Phi(v_1,v_n) + \ldots + b_n\Phi(v_n,v_n) & = & 0 \end{array}\right\} (*)$$

Dies bedeutet, dass der Vektor $v = b_1 v_1 + \ldots + b_n v_n$ die Gleichungen
$\Phi(v, v_1) = 0, \ldots, \Phi(v, v_n) = 0$ erfüllt. Somit ergibt sich $\Phi(v,w) = 0$ für jeden Vektor $w \in V$.
Da Φ nicht ausgeartet ist, folgt $v = 0$, also $b_1 = \ldots = b_n = 0$.
Somit sind die Zeilen von $G_B(\Phi)$ linear unabhängig und $G_B(\Phi)$ invertierbar.

„$\Leftarrow$“ Sei $v \in V$ mit $v \perp_\Phi V$. Schreibe $v = b_1v_1 + \ldots + b_nv_n$ mit $b_i \in K$.
Dann gilt: $\Phi(v, v_1) = \ldots = \Phi(v, v_n) = 0$. Setzt man hier die Koordinatendarstellung von V ein, so sieht man, dass $(*)$ gilt. Da $G_B(\Phi)$ nach Voraussetzung invertierbar ist, sind die Zeilen von $G_B(\Phi)$ linear unabhängig und es folgt $b_1 = \ldots = b_n = 0$.
Also ergibt sich $v = 0$. qed.

22.15 Beispiel:

(a) Die Standardbilinearform Φ ist nicht ausgeartet, da I_n invertierbar ist.

(b) Die Bilinearform $\Phi : \mathbb{Q}^2 \times \mathbb{Q}^2 \to \mathbb{Q}$ mit $G_E(\Phi) = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix}$

ist ausgeartet, denn $det \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix} = 0$.

Ziel: Definiere „orthogonales Komplement“, „orthogonale direkte Summe“

22.16 Definition:

Sei $\Phi : V \times V \to K$ eine Bilinearform und $W \subseteq V$ eine Teilmenge.
Dann heißt $W^\perp = \{v \in V \mid v \perp_\Phi W\}$ das orthogonale Komplement von W bezüglich Φ.

22.17 Satz: (Eigenschaften des orthogonalen Komplements)

Sei $\Phi : V \times V \to K$ eine Bilinearform und seien $Q, W \subseteq V$ zwei Teilmengen.

(a) $W^\perp$ ist stets ein K-Untervektorraum von V.

(b) Für $v \in V \setminus \{0\}$ und $a \in K \setminus \{0\}$ gilt: $(av)^\perp = v^\perp$

(c) Gilt $U \subseteq W$, so gilt $W^\perp \subseteq U^\perp$.

(d) Es gilt $W^\perp =< W >^\perp$.

(e) Ist Φ symmetrisch, so gilt $W \subseteq (W^\perp)^\perp$ und $((W^\perp)^\perp)^\perp = W^\perp$.

Beweis:

(a) Seien $a_1, a_2 \in K$ und $v_1, v_2 \in W^\perp$. Für $w \in W$ gilt
$\Phi(a_1v_1 + a_2v_2, w) = a_1 \underbrace{\Phi(v_1, w)}_{=0} + a_2 \underbrace{\Phi(v_2, w)}_{=0} = 0$
Dies zeigt $a_1v_1 + a_2v_2 \in W^\perp$.

(b) Für $w \in W$ gilt $w \in v^\perp \Leftrightarrow \Phi(w, v) = 0 \underset{a\neq 0}{\Leftrightarrow} a\Phi(w, v) = 0 \Leftrightarrow \Phi(w, av) = 0 \Leftrightarrow w \in (av)^\perp$

(c) Ist $v \in W^\perp$, so gilt $\Phi(v, w) = 0$ für alle $w \in W$. Insbesondere folgt $\Phi(v, w) = 0$ für alle $w \in U$, also $v \in U^\perp$.

(d) „$\supseteq$“ Aus $W \subseteq< W >$ und (c) folgt $< W >^\perp \subseteq W^\perp$

„$\subseteq$“ Sei $v \in W^\perp$. Für jedes $w \in< W >$ ist zu zeigen, dass $\Phi(v, w) = 0$ gilt. Sei $w \in< W >$.
Dann gibt es $w_1, \ldots, w_n \in W$ und $a_1, \ldots, a_n \in K$ mit $w = a_1w_1 + \ldots + a_nw_n$.
Dann folgt $\Phi(v, w) = a_1 \underbrace{\Phi(v, w_1)}_{=0} + \ldots + a_1 \underbrace{\Phi(v, w_n)}_{=0} = 0$

(e) 1.) Zu zeigen: $W \subseteq (W^\perp)^\perp$

Sei $w \in W$. Für alle $v \in W^\perp$ gilt $\Phi(v, w) = 0$. Da Φ symmetrisch ist, folgt $\Phi(w, v) = 0$ für alle $v \in W^\perp$. Also gilt $w \in (W^\perp)^\perp$.

2.) Zu zeigen: $((W^\perp)^\perp)^\perp = W^\perp$

„$\subseteq$“ Nach (1) gilt $W \subseteq (W^\perp)^\perp$. Wendet man jetzt (c) an, so folgt $((W^\perp)^\perp)^\perp \subseteq W^\perp$.

„$\supseteq$“ Wir wenden die erste Behauptung auf $U = W^\perp$ an und erhalten
$W^\perp = U \subseteq (U^\perp)^\perp = ((W^\perp)^\perp)^\perp$.

qed.

Problem: Ist Φ ausgeartet, so kann es für einen Untervektorraum $W \neq 0$ passieren, dass $W^\perp = V$ gilt.

22.18 Satz: (Dimensionssatz für orthogonale Komplemente)

Sei $\Phi : V \times V \to K$ eine nicht ausgeartete Bilinearform und sei $U \subseteq V$ ein K-Untervektorraum. Dann gilt: $dim_K(U^\perp) + dim_K(U) = dim_K(V)$.

Beweis: Sei $B = \{v_1, \ldots, v_m\}$ eine Basis von U. Ergänze sie zu einer Basis $C = \{v_1, \ldots, v_n\}$ von V. Für eine Vektor $v = a_1v_1 + \ldots + a_nv_n$ mit $a_i \in K$ gilt $v \in U^\perp$ genau dann, wenn für $i = 1, \ldots, m$ gilt $\Phi(v, v_i) = 0$. Dies bedeutet:

$$(*) \begin{cases} a_1\Phi(v_1, v_1) + \ldots + a_n\Phi(v_n, v_1) & = \quad 0 \\ \vdots & \quad \vdots \\ a_1\Phi(v_1, v_m) + \ldots + a_n\Phi(v_n, v_m) & = \quad 0 \end{cases}$$

Betrachte dies als LGS für die Koeffizienten $a_1, \ldots, a_n$ von v. Die Koeffizientenmatrix dieses LGS besteht gerade aus den ersten m Spalten von $G_C(\Phi)$. Da Φ nicht ausgeartet ist, sind diese m Spalten linear unabhängig. Die Koeffizientenmatrix des LGS $(*)$ hat also den Rang m und es gilt:

$$dim_K(U^\perp) = dim_K\{(a_1, \ldots, a_n) \in K^n \mid (a_1, \ldots, a_n) \text{ erfüllt } (*)\} = n - m = dim_K(V) - dim_K(U).$$

qed.

22.19 Bemerkung:

Obwohl $U^\perp$ die „richtige“ Dimension besitzt, muss nicht $V = U \oplus U^\perp$ gelten:
In Beispiel 22.11.c sahen wir eine nicht ausgeartete Bilinearform $\Phi : V \times V \to K$ und eine Vektor $v \neq 0$ mit $\Phi(v, v) = 0$.
Also in diesem Fall gilt $< v > \cap < v >^\perp \neq < 0 >$ und somit $< v > + < v >^\perp \subsetneq V$. Für solche Bilinearformen macht auch unsere geplante Definition $\|v\|^2 = \Phi(v, v)$ keinen Sinn.

22.20 Korollar:

Ist Φ eine nicht ausgeartete und symmetrische Bilinearform, so gilt für jeden Untervektorraum $U \subseteq V$ die Gleichung $U = (U^\perp)^\perp$.

Beweis: Die Inklusion $U \subseteq (U^\perp)^\perp$ gilt nach Satz 22.17.e.
Nach Satz 22.19 gilt $dim_K((U^\perp)^\perp) = n - dim_K(^\perp) = n - (n - dim_K(U)) = dim_K(U)$.
Also folgt $U = (U^\perp)^\perp$. qed.

Frage: Wie kann man prüfen, ob eine Bilinearform $\Phi : V \times V \to K$ symmetrisch ist?

22.21 Satz:

Sei $B = \{v_1, \ldots, v_n\}$ eine Basis von V und $\Phi : V \times V \to K$ eine Bilinearform. Genau dann ist Φ symmetrisch, wenn $G_B(\Phi)$ eine symmetrische Matrix ist, d.h. wenn $G_B(\Phi) = G_B(\Phi)^{tr}$ gilt.

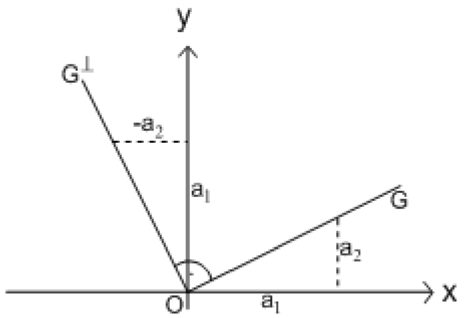
Beweis: „$\Rightarrow$“ Sei $G_B(\Phi) = (\Phi(v_i, v_j)) = (a_{ij})_{i,j}$.
Für alle i, j gilt $(G_B(\Phi)^{tr})_{i,j} = a_{ji} = \Phi(v_j, v_i) = \Phi(v_i, v_j) = a_{ij} = (G_B(\Phi))_{i,j}$.

„$\Leftarrow$“ Da $G_B(\Phi)$ symmetrisch ist, gilt $\Phi(v_i, v_j) = \Phi(v_j, v_i)$ für alle i, j. Seien nun $v = a_1v_1 + \ldots + a_nv_n$ und $w = b_1v_1 + \ldots + b_nw_n$ zwei Vektoren aus V mit $a_i, b_j \in K$. Dann gilt:
$\Phi(v, w) = \sum_{i=1}^{n} a_i\Phi(v_i, w) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_ib_j\Phi(v_i, v_j) = \sum_{j=1}^{n}\sum_{i=1}^{n} b_ja_i\Phi(v_j, v_i) = \sum_{j=1}^{n} b_j\Phi(v_j, v) = \Phi(w, v)$.

qed.

22.22 Beispiel:

(a) Sei $K = \mathbb{R}, V = \mathbb{R}^2$ und $\begin{array}{rcl}\Phi : \mathbb{R}^2 \times \mathbb{R}^2 & \to & \mathbb{R} \\ ((a,b),(c,d)) & \mapsto & ac + bd\end{array}$ die Standardbilinearform.
Sei $G \subseteq \mathbb{R}^2$ eine Gerade durch O. Dann gibt es $(u_1, u_2) \in \mathbb{R}^2 \setminus \{0\}$ mit $G = \mathbb{R} \cdot (a_1, a_2)$.
Es folgt: $G^\perp = \{(b_1, b_2) \in \mathbb{R}^2 \mid a_1b_1 + a_2b_2 = 0\} = \mathbb{R} \cdot (-a_2, a_1)$
Es gilt: $dim_{\mathbb{R}}(G^\perp) = 1 = dim_{\mathbb{R}}(\mathbb{R}^2) - dim_{\mathbb{R}}(G) = 2 - 1$.
Skizze:



(b) Sei $K = \mathbb{R}, V = \mathbb{R}^3$ und $\Phi : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ die Standardbilinearform.
Sei $G = \mathbb{R} \cdot (a_1, a_2, a_3)$ eine Gerade durch O, wobei z.B. $a_1 \neq 0$ gelte.
Dann folgt
$G^\perp = \{(b_1, b_2, b_3) \in \mathbb{R}^3 \mid a_1b_1 + a_2b_2 + a_3b_3 = 0\} = \mathbb{R} \cdot (-a_2, a_1, 0) + \mathbb{R} \cdot (0, a_3, -a_2)$
Somit gilt $dim_{\mathbb{R}}(G^\perp) = 2 = dim_{\mathbb{R}}(\mathbb{R}^3) - dim_{\mathbb{R}}(G) = 3 - 1$
Das orthogonale Komplement von G ist eine Ebene durch O.

Fazit:

- Wir brauchen eine Bilinearform $\Phi : V \times V \to K$.
- Ist Φ nicht ausgeartet, so kann man „orthogonal“ und „orthogonales Komplement“ definieren.
- Φ sollte symmetrisch sein.
- Problem:
  - Es kann $v \in V \setminus \{0\}$ geben mit $\Phi(v, v) = 0$.
  - $\|v\| = \sqrt{\Phi(v, w)}$ funktioniert nicht richtig!

23 Winkelzüge oder normgerechtes Verhalten? (Skalarprodukte)

A. Kennen wir das Skalarprodukt nicht schon?

In diesem Abschnitt sei $K = \mathbb{R}$ und V ein reeller Vektorraum, d.h. ein $\mathbb{R}$-Vektorraum.
Sei $n = dim_{\mathbb{R}}(V)$.

23.1 Definition:

(a) Eine Bilinearform $\Phi : V \times V \to \mathbb{R}$ heißt positiv definit, wenn für alle $v \in V \setminus \{0\}$ gilt: $\Phi(v, v) > 0$.

(b) Ein Skalarprodukt auf V ist eine positiv definite, symmetrische Bilinearform $\Phi : V \times V \to \mathbb{R}$.

(c) Ein reeller Vektorraum V zusammen mit einem Skalarprodukt Φ auf V heißt ein euklidischer Vektorraum.

23.2 Bemerkung:

Ist $\Phi : V \times V \to \mathbb{R}$ ein Skalarprodukt, so ist Φ nicht ausgeartet:
Für $v \in V$ mit $v \perp_\Phi V$ gilt insbesondere $\Phi(v, v) = 0$. Hieraus folgt $v = 0$, da Φ positiv definit ist.

23.3 Beispiel:

Sei $B = \{v_1, \ldots, v_n\}$ eine Basis von V. Sei $\Phi : V \times V \to \mathbb{R}$ die Standardbilinearform bezüglich B.
Dann ist Φ ein Skalarprodukt auf V, denn für $v = a_1v_1 + \ldots + a_nv_n$ gilt
$\Phi(v, v) = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} a_i a_j \delta_{ij} = a_1^2 + \ldots + a_n^2 \geq 0.$
Aus $\Phi(v, v) = 0$ folgt $a_1 = \ldots = a_n = 0$, also $v = 0$.
Auf einem reellen Vektorraum heißt die Standardbilinearform daher auch das Standardskalarprodukt bezüglich B.
Das Skalarprodukt auf $V = \mathbb{R}^n$ bezüglich E ist das in Kapitel I betrachtete.
Schreibweise: $< v, w >$ statt $\Phi(v, w)$

Fragen:

- Gibt es außer dem Standardskalarprodukt noch weitere Skalarprodukte auf einem reellen Vektorraum?
- Wie kann man an Hand der Gramschen Matrix erkennen, ob Φ ein Skalarprodukt ist?

B. Können wir jetzt endlich Längen und Winkel messen?

23.4 Definition:

Sei V ein reeller Vektorraum.

(a) Eine Norm auf V ist eine Abbildung $\|\cdot\| : V \to \mathbb{R}$ mit folgenden Eigenschaften:

1.) Für $v \in V$ gilt: $\|v\| \geq 0$ und $\|v\| = 0 \Leftrightarrow v = 0$. (Positivität)

2.) Für $v \in V$ und $a \in \mathbb{R}$ gilt: $\|a \cdot v\| = |a| \cdot \|v\|$. (Homogenität)

3.) Für $v, w \in V$ gilt: $\|v + w\| \leq \|v\| + \|w\|$. (Dreiecksungleichung)

(b) Sei $\Phi : V \times V \to \mathbb{R}$ ein Skalarprodukt. Dann heißt $\begin{aligned} \|\cdot\|_\Phi : V &\to \mathbb{R} \\ v &\mapsto \sqrt{\Phi(v,v)} \end{aligned}$

die zu Φ assoziierte Norm.

(c) Die zum Standardskalarprodukt $< \cdot, \cdot >$ gehörige Norm $\|\cdot\|$ heißt auch die euklidische Norm auf $V = \mathbb{R}^n$.

(d) Ist V ein reeller Vektorraum und $\|\cdot\| : V \to \mathbb{R}$ eine Norm auf V, so heißt $(V, \|\cdot\|)$ ein normierter Vektorraum.

23.5 Satz:

Sei V ein reeller Vektorraum und $\Phi : V \times V \to \mathbb{R}$ ein Skalarprodukt auf V.

(a) Es gilt die Cauchy-Schwarzsche Ungleichung: $|\Phi(v,w)| \leq \|v\|_\Phi \cdot \|w\|_\Phi$ für $v, w \in V$.

(b) Die zu Φ assoziierte Norm ist eine Norm auf V. Insbesondere gilt die Dreiecksungleichung: $\|v + w\|_\Phi \leq \|v\|_\Phi + \|w\|_\Phi$ für alle $v, w \in V$.

Beweis:

(a) Für $w = 0$ sind beide Seiten gleich Null. Sei also $w \neq 0$.
Für alle $a \in \mathbb{R}$ gilt $0 \leq \Phi(v - aw, v - aw) = \Phi(v,v) - 2a\Phi(v,w) + a^2\Phi(w,w)$.
Wegen $\Phi(w,w) > 0$ können wir $a = \dfrac{\Phi(v,w)}{\Phi(w,w)}$ verwenden und erhalten:

$$0 \leq \Phi(v,v) - 2\frac{\Phi(v,w)^2}{\Phi(w,w)} + \frac{\Phi(v,w)^2}{\Phi(w,w)} = \Phi(v,v) - \frac{\Phi(v,w)^2}{\Phi(w,w)}$$

$$\Rightarrow 0 \leq \Phi(v,v)\Phi(w,w) - \Phi(v,w)^2 \Rightarrow \Phi(v,w)^2 \leq \|v\|_\Phi^2 \cdot \|w\|_\Phi^2$$

$$\Rightarrow |\Phi(v,w)| \leq \|v\|_\Phi \cdot \|w\|_\Phi$$

(b) *Positivität:* klar

Homogenität: $\|av\|_\Phi = \sqrt{\Phi(av,av)} = \sqrt{a^2\Phi(v,v)} = |a| \cdot \sqrt{\Phi(v,v)} = |a| \cdot \|v\|_\Phi$

Es gilt $\|v + w\|_\Phi^2 = \Phi(v + w, v + w) = \Phi(v,v) + 2\Phi(v,w) + \Phi(w,w)$
$\underset{(a)}{\leq} \|v\|_\Phi^2 + 2\|v\|_\Phi \cdot \|w\|_\Phi + \|w\|_\Phi^2 = (\|v\|_\Phi + \|w\|_\Phi)^2$

Hieraus folgt die Behauptung. qed.

Also können wir in einem euklidischen Vektorraum Längen messen. Die Winkelmessung funktioniert wie folgt:

23.6 Korollar:

Für alle $v, w \in V \setminus \{0\}$ gilt: $-1 \leq \dfrac{\Phi(v,w)}{\|v\|_\Phi \cdot \|w\|_\Phi} \leq 1$.

Also gibt es einen eindeutig bestimmten Winkel $\alpha \in [0, 2\pi)$ mit $cos(\alpha) = \dfrac{\Phi(v,w)}{\|v\|_\Phi \cdot \|w\|_\Phi}$.

Dieser Winkel α heißt der Winkel zwischen v und w (bezüglich Φ).
Er wir auch mit $\alpha\measuredangle(v,w)$ notiert.

Beweis: Dies folgt sofort aus der Cauchy-Schwarzschen Ungleichung. qed.

23.7 Definition:

Sei V ein euklidischer Vektorraum und $\Phi : V \times V \to \mathbb{R}$ ein Skalarprodukt auf V. Sei $W \subseteq V$ eine Menge von Vektoren.

(a) Die Vektoren aus W heißen (paarweise) orthogonal, wenn für $w_1, w_2 \in W$ gilt: $\Phi(w_1, w_2) = 0$.

(b) Die Vektoren aus W heißen orthonormal, wenn sie orthogonal sind und für alle $w \in W$ gilt: $\|w\|_\Phi = 1$.

(c) Ist B eine Basis von V und ist B orthogonal, so heißt B eine Orthogonalbasis (OGB) von V.

(d) Ist B eine Basis von V und ist B orthonormal, so heißt B eine Orthonormalbasis (ONB) von V.

Frage: Besitzt jeder euklidische Vektorraum eine ONB?

23.8 Bemerkung:

Ist $B = \{v_1, \ldots, v_n\}$ eine OGB, so ist $\{\dfrac{v_1}{\|v_1\|_\Phi}, \ldots, \dfrac{v_n}{\|v_n\|_\Phi}\}$ eine ONB von V.

Wegen $v_i \neq 0$ ist $\|v_i\|_\Phi \neq 0$. Ferner gilt $\|\dfrac{v_i}{\|v_i\|_\Phi}\|_\Phi = \dfrac{1}{\|v_i\|_\Phi} \cdot \|v_i\|_\Phi = 1$

Diesen Prozess nennt man Normieren.

23.9 Theorem: (Das (Gram-)Schmidtsche Orthonormalisierungsverfahren)

Sei $B = \{v_1, \ldots, v_n\}$ eine Basis eines euklidischen Vektorraumes V mit Skalarprodukt $\Phi : V \times V \to \mathbb{R}$. Betrachte die folgenden Instruktionen:

(a) Setze $w_1 = \dfrac{v_1}{\|v_1\|}$.

(b) Für $i = 2, \ldots, n$ berechne der Reihe nach $\tilde{v}_i = \Phi(v_i, w_1)w_1 + \ldots + \Phi(v_i, w_{i-1})w_{i-1}$ und setze $w_i = \dfrac{v_i - \tilde{v}_i}{\|v_i - \tilde{v}_i\|}$

(c) Gib $C = \{w_1, \ldots, w_n\}$ aus und stoppe.

Dies ist ein Algorithmus, der eine ONB C von V berechnet.

Beweis: *Endlichkeit:* Klar.

Korrektheit: Wir schließen mit vollständiger Induktion nach n.

$n = 1$: w_1 ist normiert und $\{w_1\}$ ist orthogonal. Die Menge $\{w_1\}$ ist also eine ONB von $< v_1 >$.

$n > 1$: Nach Induktionsvoraussetzung ist $\{w_1, \ldots, w_{n-1}\}$ eine ONB von $< v_1, \ldots, v_{n-1} >$.
Offenbar gilt $v_n \notin < v_1, \ldots, v_{n-1} > = < w_1, \ldots, w_{n-1} >$. Somit ist $v_n \neq \tilde{v}_n$ und w_n ist wohldefiniert.
Für $i = 1, \ldots, n-1$ berechnen wir $\Phi(w_1, w_i) = \frac{1}{\|v_n - \tilde{v}_n\|} \cdot \Phi(v_n - \tilde{v}_n, w_i)$

$$= \frac{1}{\|v_n - \tilde{v}_n\|} \cdot (\Phi(v_n, w_i) - \Phi(v_n, w_1) \underbrace{\Phi(w_1, w_i)}_{= 0 \text{ nach IV}} - \ldots - \Phi(v_n, w_{n-1}) \underbrace{\Phi(w_n, w_i)}_{= 0 \text{ nach IV}})$$

$$= \frac{1}{\|v_n - \tilde{v}_n\|} \cdot (\Phi(v_n, w_i) - \Phi(v_n, w_i) \underbrace{\Phi(w_i, w_i)}_{= 1 \text{ nach IV}}) = 0$$

Somit gilt $w_n \perp_\Phi < w_1, \ldots, w_{n-1} >$. Außerdem gilt $\|w_n\| = 1$. Nun folgt die Behauptung aus
$< w_1, \ldots, w_n > = < v_1, \ldots, v_{n-1}, w_n > = < v_1, \ldots, v_{n-1}, v_n > = V$ qed.

23.10 Beispiel:

Sei $K = \mathbb{R}, V = \mathbb{R}^3$ und $B = \{v_1, v_2, v_3\}$ mit $v_1 = (1, 2, 0), v_2 = (0, 1, 2), v_3 = (0, 1, 0)$.
Wir berechnen aus B eine ONB von $\mathbb{R}^3$ berzüglich $< \cdot, \cdot >$:

(a) $w_1 = \frac{v_1}{\|v_1\|} = \frac{v_1}{\sqrt{5}} = (\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}, 0)$

(b) $\tilde{v}_2 = < v_2, w_1 > w_1 = \frac{2}{\sqrt{5}} w_1 = (\frac{2}{5}, \frac{4}{5}, 0)$

$w_2 = \frac{v_2 - \tilde{v}_2}{\|v_2 - \tilde{v}_2\|} = \frac{1}{\|v_2 - \tilde{v}_2\|} \cdot (-\frac{2}{5}, \frac{1}{5}, 2) = (-\frac{2}{\sqrt{105}}, \frac{1}{\sqrt{105}}, \frac{10}{\sqrt{105}})$

denn $\|v_2 - \tilde{v}_2\| = \sqrt{\frac{4}{25} + \frac{1}{25} + \frac{100}{25}} = \frac{1}{5}\sqrt{105}$

(b) $\tilde{v}_3 = < v_3, w_1 > w_1 + < v_3, w_2 > w_2 = \frac{2}{\sqrt{5}} w_1 + \frac{1}{\sqrt{105}} w_2 = (\frac{2}{5}, \frac{4}{5}, 0) + (-\frac{2}{105}, \frac{1}{105}, \frac{10}{105})$

$= (\frac{40}{105}, \frac{85}{105}, \frac{10}{105}) = (\frac{8}{21}, \frac{17}{21}, \frac{2}{21})$

$w_3 = \frac{1}{\|v_3 - \tilde{v}_3\|}(v_3 - \tilde{v}_3) = \frac{1}{\|v_3 - \tilde{v}_3\|}(-\frac{8}{21}, \frac{4}{21}, -\frac{2}{21}) = (-\frac{4}{\sqrt{21}}, \frac{2}{\sqrt{21}}, -\frac{1}{\sqrt{21}})$

denn $\|v_3 - \tilde{v}_3\| = \sqrt{\frac{64}{441} + \frac{16}{441} + \frac{4}{441}} = \frac{1}{21}\sqrt{84} = \frac{2}{\sqrt{21}}$

(c) Es gilt $C = \{w_1, w_2, w_3\} = \{(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}, 0), (-\frac{2}{\sqrt{105}}, \frac{1}{\sqrt{105}}, \frac{10}{\sqrt{105}}), (-\frac{4}{\sqrt{21}}, \frac{2}{\sqrt{21}}, -\frac{1}{\sqrt{21}})\}$

Wir prüfen z.B. $< w_2, w_3 > = \frac{8}{\sqrt{105}\sqrt{21}} + \frac{2}{\sqrt{105}\sqrt{21}} - \frac{10}{\sqrt{105}\sqrt{21}} = 0$

23.11 Korollar:

Ist $W \subseteq V$ eine orthonormale Menge von Vektoren, so kann man W zu einer ONB von V ergänzen.

Beweis: Sei $W = \{w_1, \ldots, w_m\}$.

(a) Die Vektoren in W sind $\mathbb{R}$-linear unabhängig.
Sind $a_1, \ldots, a_m \in \mathbb{R}$ mit $a_1 w_1 + \ldots + a_m w_m = 0$, so folgt
$0 = \Phi(a_1 w_1 + \ldots + a_m w_m, a_1 w_1 + \ldots + a_m w_m) = a_1^2 + \ldots + a_m^2$, also $a_1 = \ldots = a_m = 0$.
Insbesondere erhalten wir $m \leq n = dim_{\mathbb{R}}(V)$.

(b) Ergänze W zu einer Basis $\{w_1, \ldots, w_m, v_{m+1}, \ldots, v_n\}$ von V. Wende das Gram-Schmidt-Verfahren an. Erhalte eine ONB $\{w_1, \ldots, w_m, w_{m+1}, \ldots, w_n\}$ von V, wie man durch Anwendung des Algorithmus sieht. qed.

Frage: Funktioniert die Bildung orthogonaler Komplemente und orthogonaler direkter Summen richtig, wenn Φ ein Skalarprodukt ist?

23.12 Satz:

Sei V ein euklidischer Vektorraum mit Skalarprodukt $\Phi : V \times V \to \mathbb{R}$ und sei $U \subseteq V$ ein $\mathbb{R}$-Untervektorraum. Dann gilt $V = U \oplus U^\perp$
In diesem Fall schreiben wir $V = U \bigoplus\!\!\!\!\perp U^\perp$ und nennen dies eine orthogonale direkte Summe.

Beweis: Wir wissen bereits, dass $dim_{\mathbb{R}}(U) + dim_{\mathbb{R}}(U^\perp) = n = dim_{\mathbb{R}}(V)$ gilt, denn Φ ist nicht ausgeartet. Zu zeigen ist also noch $U \cap U^\perp = \{0\}$.
Sei $v \in U \cap U^\perp$. Wegen $v \in U^\perp$ gilt insbesondere $v \perp v$, also $\Phi(v, v) = 0$. Da Φ positiv definit ist, folgt $v = 0$. qed.

C. Und sonst? Was kann man jetzt sonst noch alles messen?

Idee: Im $\mathbb{R}^n$ kann man Volumina mit Determinanten messen!

23.13 Bemerkung: (Wiederholung)

(a) Seien $v = (a_1, a_2)$ und $w = (b_1, b_2)$ zwei linear unabhängige Vektoren in $\mathbb{R}^2 \setminus \{0\}$. Dann ist die Fläche des von v und w aufgespannten Parallelogramms gegeben durch:

$$Vol(v, w) = |a_1 b_2 - a_2 b_1|$$

(b) Seien $v = (a_1, a_2, a_3), w = (b_1, b_2, b_3)$ und $u = (c_1, c_2, c_3)$ lineare unabhängige Vektoren in $\mathbb{R}^3 \setminus \{0\}$. Dann ist das Volumen des duch v, w und u aufgespannten Spats (oder Parallelepipeds) gegeben durch

$$Vol(v, w, u) = \left| det \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} \right| = |det(A)| \text{ mit } A = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$$

Betrachte $v = a_1e_1 + a_2e_2 + a_3e_3$ mit (a_1, a_2, a_3) als Koordinatentupel bezüglich der Basis $\{e_1, e_2, e_3\}$. Dann kann man die Gramsche Matrix des Standardskalarprodukts bezüglich der Basis E bilden und erhält bezüglich der Basis $B = \{v, w, u\}$ die Gramsche Matrix

$$G_B(\Phi) = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}^{tr} \cdot G_E(\Phi) \cdot \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} = AA^{tr}$$

Also folgt $Vol(v, w, u) = |det(A)| = \sqrt{det(AA^{tr})} = \sqrt{det(G_B(\Phi))}$

23.14 Definition:

Sei V ein euklidischer Vektorraum mit Skalarprodukt $\Phi : V \times V \to \mathbb{R}$ und sei $W = \{v_1, \dots, v_m\} \subseteq V$ mit $m \leq n$. Dann heißt

$$G(v_1, \dots, v_m) = det \begin{pmatrix} \Phi(v_1, v_1) & \cdots & \Phi(v_1, v_m) \\ \vdots & & \vdots \\ \Phi(v_m, v_1) & \cdots & \Phi(v_m, v_m) \end{pmatrix} \in \mathbb{R}$$

die Gramsche Determinante von W.

23.15 Satz:

Sei V ein eulidischer Vektorraum mit Skalarprodukt $\Phi : V \times V \to \mathbb{R}$ und sei $W = \{v_1, \dots, v_m\} \subseteq V$.

(a) Es gilt $G(v_1, \dots, v_m) \geq 0$.

(b) Genau dann gilt $G(v_1, \dots, v_m) = 0$, wenn $\{v_1, \dots, v_m\}$ linear abhängig ist.

Beweis:

(a) Sei $B = \{w_1, \dots, w_n\}$ eine ONB von V (existiert nach 23.9). Für $i = 1, \dots, m$ schreibe $v_i = a_{i1}w_1 + \ldots + a_{in}w_n$ mit $a_{ij} \in \mathbb{R}$ und setze $A = (a_{ij}) \in Mat_{m,n}(\mathbb{R})$. Dann gilt:

$$G(v_1, \dots, v_m) = det(\Phi(v_i, v_j))_{i,j}$$

Hierbei gilt: $\Phi(v_i, v_j) = \Phi(a_{i1}w_1 + \ldots + a_{in}w_n, a_{j1}w_1 + \ldots + a_{jn}w_n)$
$= \sum_{k=1}^{n} \sum_{l=1}^{n} a_{ik}a_{jl}\Phi(w_k, w_l) = a_{i1}a_{j1} + \ldots + a_{in}a_{jn}$.

Also folgt: $G(v_1, \dots, v_m) = det(a_{i1}a_{j1} + \ldots + a_{in}a_{jn})_{i,j} = det(AA^{tr})$.

Ergänze A zur Matrix $C = \begin{pmatrix} A \\ e_{m+1} \\ \vdots \\ e_n \end{pmatrix} \in Mat_n(\mathbb{R})$ und erhalte $det(CC^{tr})$

$$= det\left(\begin{pmatrix} A \\ e_{m+1} \\ \vdots \\ e_n \end{pmatrix} \cdot \begin{pmatrix} A^{tr} & e_{m+1}^{tr} & \cdots & e_m^{tr} \end{pmatrix} \right) = det \left(\begin{array}{c|ccc} AA^{tr} & & 0 & \\ \hline & 1 & & \\ 0 & & \ddots & \\ & & & 1 \end{array}\right) = det(AA^{tr})$$

Dies liefert $G(v_1, \dots, v_m) = det(AA^{tr}) = det(CC^{tr}) = det(C) \cdot det(C^{tr}) = det(C^2) \geq 0$

(b) „⇒“ Es gelte $G(v_1, \dots, v_m) = 0$. Angenommen, W ist linear unabhängig. Dann wähle $\{w_{i_1}, \dots, w - i_{n-m}\} \subseteq B$ so, dass $\{v_1, \dots, v_m, w_{i_1}, \dots, w_{i_{n-m}}\}$ linear unabhängig ist.

Bilde nun die Matrix $C = \begin{pmatrix} A \\ e_{m+1} \\ \vdots \\ e_n \end{pmatrix}$ und erhalte (wie eben)

$0 = G(v_1, \dots, v_m) = det(AA^{tr}) = det(CC^{tr}) = det(C)^2$. Dies widerspricht der Annahme, dass C invertierbar ist.

„⇐“ Klar. qed.

23.16 Definition:

Sei V ein euklidischer Vektorraum mit Skalarprodukt $\Phi : V \times V \to \mathbb{R}$ und sei $V = \{v_1, \dots, v_m\}$ eine Menge von linear unabhängigen Vektoren in V. Dann heißt $Vol(v_1, \dots, v_m) = \sqrt{G(v_1, \dots, v_m)}$ das Volumen des von W aufgespannten m-dimensionalen Spats (oder Parallelepipeds).

23.17 Satz: (Die Ungleichung von Hadamard)

Sei $W = \{v_1, \dots, v_m\}$ eine Menge von Vektoren in einem euklidischen Vektorraum V mit Skalarprodukt $\Phi : V \times V \to \mathbb{R}$. Dann gilt: $Vol(v_1, \dots, v_m) \le \|v_1\|_\Phi \cdots \|v_m\|_\Phi$
Hierbei gilt Gleichheit genau dann, wenn W bezüglich Φ orthogonal ist.

Beweis: Indem man die Ungleichung quadriert, sieht man, dass es genügt
$G(v_1, \dots, v_m) \le \Phi(v_1, v_1) \cdots \Phi(v_m, v_m)$ zu zeigen:
Wir zeigen, dass für jedes $r \le m$ und für die eindeutig bestimmte Zerlegung $v_r = v_r' + v_r''$ mit $v_r' \in < v_1, \dots, v_{r-1} >$ und $v_r'' \in < v_1, \dots, v_m >^\perp$ gilt:
$G(v_1, \dots, v_m) \underset{(1)}{=} G(v_1, \dots, v_{r-1}) \cdot \Phi(v_r'', v_r'') \underset{(2)}{\le} G(v_1, \dots, v_{r-1}) \cdots \Phi(v_r, v_r)$.
Dann folgt die Behauptung des Satzes mit Induktion nach r.

Es gilt $\Phi(v_r, v_r) = \Phi(v_r' + v_r'', v_r' + v_r'') = \Phi(v_r', v_r') + 2\underbrace{\Phi(v_r', v_r'')}_{=0} + \Phi(v_r'', v_r'')$

$$= \Phi(v_r', v_r') + \Phi(v_r'', v_r'') \ge \Phi(v_r', v_r').$$

Dies zeigt (2) und die Tatsache, dass in (2) genau dann Gleichheit gilt, wenn $v_r \perp < v_1, \dots, v_{r-1} >$ erfüllt ist.

Nun beweisen wir noch (1). Dazu verwenden wir $\Phi(v_i, v_r') = \Phi(v_i, v_r)$ für $i = 1, \dots, r-1$ und $\Phi(v_r, v_r) = \Phi(v_r', v_r') + \Phi(v_r'', v_r'')$. Dies liefert:

$$G(v_1, \dots, v_r) = det \left(\begin{array}{ccc|c} \Phi(v_1, v_1) & \cdots & \Phi(v_1, v_{r-1}) & \Phi(v_1, v_r') \\ \vdots & & \vdots & \vdots \\ \Phi(v_{r-1}, v_1) & \cdots & \Phi(v_{r-1}, v_{r-1}) & \Phi(v_{r-1}, v_r') \\ \hline \Phi(v_r', v_1) & \cdots & \Phi(v_r', v_{r-1}) & \Phi(v_r, v_r') + \Phi(v_r'', v_r'') \end{array} \right)$$

Schreibe $v'_r = c_1 v_1 + \ldots + c_{r-1} v_{r-1}$ mit $c_1, \ldots, c_{r-1} \in \mathbb{R}$. Subtrahiere von der letzten Spalte das c_i-fache der i-ten Spalte für $i = 1, \ldots, r-1$. Die letzte Spalte ist dann:

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \\ \Phi(v''_r, v''_r) \end{pmatrix}$$

Entwicklung nach der neuen letzten Spalte ergibt $G(v_1, \ldots, v_r) = G(v_1, \ldots, v_{r-1}) \cdot \Phi(v''_r, v''_r)$.

qed.

24 Dame, König, Ass, Vektorraum (Hierarchie der Vektorräume)

Sei V ein (n-dimensionaler) reeller Vektorraum.

24.1 Bemerkung:

Ist V ein euklidischer Vektorraum mit Skalarprodukt $\Phi : V \times V \to \mathbb{R}$, so wird V durch $\begin{array}{rcl} \|\cdot\|_\Phi : V & \to & \mathbb{R} \\ v & \mapsto & \sqrt{\Phi(v,v)} \end{array}$ zu einem normierten Vektorraum. Jedoch nicht jeder normierte reelle Vektorraum entsteht auf diese Weise, d.h. nicht jede Norm kommt von einem Skalarprodukt her (vgl. Übungen).

24.2 Satz: (Die Parallelogramm-Gleichung)

(a) Ist V euklidisch mit Skalarprodukt $\Phi : V \times V \to \mathbb{R}$ und $\|\cdot\|_\Phi : V \to \mathbb{R}$ die zu Φ assoziierte Norm, so gilt:

$$\|v+w\|_\Phi^2 + \|v-w\|_\Phi^2 = 2\|v\|_\Phi^2 + 2\|w\|_\Phi^2$$

(b) Ist $\|\cdot\| : V \to \mathbb{R}$ eine Norm auf V, für die die Parallelogramm-Gleichung stets gilt, so existiert ein Skalarprodukt $\Phi : V \times V \to \mathbb{R}$ mit $\|\cdot\| = \|\cdot\|_\Phi$.

Beweis:

(a) Es gilt $\|v+w\|_\Phi^2 + \|v-w\|_\Phi^2 = \Phi(v+w, v+w) + \Phi(v-w, v-w)$
$= [\Phi(v,v) + 2\Phi(v,w) + \Phi(w,w)] + [\Phi(v,v) - 2\Phi(v,w) + \Phi(w,w)] = 2\Phi(v,v) + 2\Phi(w,w)$
$= 2\|v\|_\Phi^2 + 2\|w\|_\Phi^2$

(b) Setze $\Phi(v,w) = \frac{1}{2}[\|v+w\|^2 - \|v\|^2 - \|w\|^2]$ für $v, w \in V$ und erhalte eine Abbildung $\Phi : V \times V \to \mathbb{R}$.

Φ *symmetrisch:* Klar.

Φ *positiv definit:* $\Phi(v,v) = \frac{1}{2}[\|2v\|^2 - 2\|v\|^2] = \|v\|^2 \geq 0$

Φ *bilinear:* Seien $v_1, v_2, w \in V$. Dann gilt: $2\Phi(v_1+v_2, w) - 2\Phi(v_1, w) - 2\Phi(v_2, w)$
$= \|v_1+v_2+w\|^2 - \|v_1+v_2\|^2 - \|w\|^2 - \|v_1+w\|^2 + \|v_1\|^2 + \|w\|^2 - \|v_2+w\|^2 + \|v_2\|^2 + \|w\|^2$
$= -\frac{1}{2}\|v_1+v_2+2w\|^2 - \frac{1}{2}\|v_1-v_2\|^2 + \frac{1}{2}\|v_1+v_2+2w\|^2 + \frac{1}{2}\|v_1+v_2\|^2 - \|v_1+v_2\|^2 + \|v_1\|^2 + \|v_2\|^2$
$= -\frac{1}{2}\|v_1-v_2\|^2 - \frac{1}{2}\|v_1+v_2\|^2 + \|v_1\|^2 + \|v_2\|^2 = 0.$

Wegen der Symmetrie von Φ gilt damit auch $\Phi(v, w_1+w_2) = \Phi(v, w_1) + \Phi(v, w_2)$ für $v, w_1, w_2 \in V$.

Seien nun $a \in \mathbb{R}$ und $v, w \in V$. Zu zeigen: $\Phi(av, w) = a\Phi(v,w)$.

1.) Für $a \in \mathbb{N}_+$ schließe mit vollständiger Induktion nach a:

$\underline{a = 1}$: Klar.

$\underline{a > 1}$: $\Phi(av, w) = \Phi((a-1)v + v, w) = \Phi((a-1)v, w) + \Phi(v, w)$
$\underset{\text{I.V.}}{=} (a-1)\Phi(v,w) + \Phi(v,w) = a\Phi(v,w)$

2.) Für $a \in \mathbb{Z}, a < 0$ verwende $\Phi(av, w) = -\Phi(-av, w)$ nach der Parallelogrammgleichung: $2\Phi(av, w) + 2\Phi(-av, w)$
$= \|av + w\|^2 - \|av\|^2\|w\|^2 + \| - av + w\|^2 - \| - av\|^2 - \|w\|^2$
$= \|av + w\|^2 + \|av - w\|^2 - 2\|av\|^2 - 2\|w\|^2 = 0.$

3.) Für $a = \frac{p}{q} \in \mathbb{Q}$ mit $q > 0$ verwende $\Phi(av, w) = \frac{1}{q^2}\Phi(pv, qw) \underset{(2)}{=} \frac{pq}{q^2}\Phi(v, w) = a\Phi(v, w)$.

4.) Für $a \in \mathbb{R}$ beliebig verwende eine Folge $q_1, q_2, \ldots$ in $\mathbb{Q}$ mit $\lim\limits_{i \to \infty} q_i = a$ und die Stetigkeit von Φ. qed.

24.3 Definition:

(a) Eine Metrik auf V ist eine Abbildung $d : V \times V \to \mathbb{R}$ mit folgenden Eigenschaften:

1.) $d(v, w) \geq 0$ und $d(v, w) = 0 \Leftrightarrow v = w$
2.) $d(v, w) = d(w, v)$
3.) $d(v, w) \leq d(v, u) + d(u, w)$ für $u, v, w \in V$

(b) Ein Vektorraum V zusammen mit einer Metrik auf V heißt ein metrischer (Vektor)raum.

24.4 Satz:

Ist $(V, \|\cdot\|)$ ein normierter Vektorraum, so wird durch $\begin{array}{rcl} d : V \times V & \to & \mathbb{R} \\ (v, w) & \mapsto & \|v - w\| \end{array}$ eine Metrik auf V definiert. Also ist jeder normierte Vektorraum ein metrischer Vektorraum.

Beweis:

1.) $d(v, w) = \|v - w\| \geq 0$ und $d(v, w) = \|v - w\| = 0 \Leftrightarrow v - w = 0 \Leftrightarrow v = w$
2.) $d(v, w) = \|v - w\| = \|w - v\| = d(w, v)$
3.) $d(v, w) = \|v - w\| \leq \|v - u\| + \|u - w\| = d(v, u) + d(u, w)$ qed.

24.5 Bemerkung:

(a) Nicht jede Metrik auf einem reellen Vektorraum kommt von einer Norm.
Beispiel: Sei $\mathscr{C}(\mathbb{R}, \mathbb{R})$ der Vektorraum der stetigen Abbildungen $f : \mathbb{R} \to \mathbb{R}$ und $[f]_k = max\{|f(x)| : -k \leq x \leq k\}$ für $k \geq 0$. Dann setze $d(f, g) = \sum\limits_{k \geq 0} 2^{-k} \dfrac{[f - g]_k}{1 + [f - g]_k}$.
Dies ist eine Metrik auf V, für die es eine Norm $\|\cdot\|$ gibt mit $d(f, g) = \|f - g\|$.

(b) Es ergibt sich eine Hierarchie der reellen Vektorräume:

$$\begin{array}{c} \text{euklidischer Vektorraum} \\ \Downarrow \quad \not\Uparrow \\ \text{normierter Vektorraum} \\ \Downarrow \quad \not\Uparrow \\ \text{metrischer Raum} \\ \Downarrow \quad \not\Uparrow \\ \text{(topologischer Raum)} \end{array}$$

Kapitel VIII: Die Achse des Guten (Die Hauptachsentransformation)

25 Wie man pervertiert invertiert indem man ungeniert transponiert (Orthogonale Abbildungen und Matrizen)

A. Über die Treue

In diesem Abschnitt sei V ein (n-dimensionaler) euklidischer Vektorraum mit Skalarprodukt $\Phi : V \times V \to \mathbb{R}$.

Frage: Welche Endomorphismen $f : V \to V$ erhalten Längen und Winkel?

25.1 Definition:

(a) Eine $\mathbb{R}$-lineare Abbildung $f : V \to V$ heißt orthogonal, wenn für alle $v, w \in V$ gilt: $\Phi(f(v), f(w)) = \Phi(v, w)$.

(b) Eine Abbildung $f : V \to V$ heißt längentreu, wenn für alle $v, w \in V$ gilt $\|v - w\|_\Phi = \|f(v) - f(w)\|_\Phi$.

(c) Eine Abbildung $f : V \to V$ heißt winkeltreu, wenn für alle $v, w \in V$ gilt: $\measuredangle(f(v), f(w)) = \measuredangle(v, w)$

25.2 Bemerkung:

Eine orthogonale $\mathbb{R}$-lineare Abbildung $f : V \to V$ ist längen- und winkeltreu, denn für alle $v, w \in V$ gilt:

$$\begin{aligned}\|f(v) - f(w)\|_\Phi &= \sqrt{\Phi(f(v) - f(w), f(v) - f(w))} = \sqrt{\Phi(f(v - w), f(v - w))}\\ &= \sqrt{\Phi(v - w, v - w)} = \|v - w\|_\Phi\end{aligned}$$

$$\measuredangle(f(v), f(w)) = arccos\frac{\Phi(f(v), f(w))}{\|f(v)\|_\Phi \cdot \|f(w)\|_\Phi} = arccos\frac{\Phi(v, w)}{\|v\|_\Phi \cdot \|w\|_\Phi} = \measuredangle(v, w)$$

Frage: Wie erkennt man an einer Darstellungsmatrix von f, ob f orthogonal ist?

25.3 Lemma:

Sei $B = \{v_1, \ldots, v_n\}$ eine ONB von V und $f : V \to V$ ein Endomorphismus. Genau dann ist f orthogonal, wenn für $i, j \in \{1, \ldots, n\}$ gilt: $\Phi(f(v_i), f(v_j)) = \Phi(v_i, v_j)$.

Beweis: „$\Rightarrow$“ Klar.

„$\Leftarrow$“ Seien $v, w \in V$. Schreibe $v = a_1v_1 + \ldots + a_nv_n$ und $w = b_1v_1 + \ldots + b_nv_n$ mit $a_i, b_j \in \mathbb{R}$. Dann gilt:

$$\begin{aligned}\Phi(f(v), f(w)) &= \Phi(a_1f(v_1) + \ldots + a_nf(v_n), b_1f(v_1) + \ldots + b_nf(v_n))\\ &= \sum_{i=1}^n\sum_{j=1}^n a_ib_j\Phi(f(v_i), f(v_j)) = \sum_{i=1}^n\sum_{j=1}^n a_ib_j\Phi(v_i, v_j)\\ &= \Phi(a_1v_1 + \ldots + a_nv_n, b_1v_1 + \ldots + b_nv_n) = \Phi(v, w)\end{aligned}$$

qed.

25.4 Satz:

Sei $B = \{v_1, \ldots, v_n\}$ eine ONB von V. Genau dann ist ein Endomorphismus $f : V \to V$ orthogonal, wenn die Matrix $M_B(f)$ invertierbar ist und $M_B(f)^{-1} = M_B(f)^{tr}$.
Insbesondere ist jeder orthogonale Endomorphismus ein Automorphismus.

Beweis: Schreibe $M_B(f) = (a_{ij})$, d.h. es gelte $f(v_j) = a_{1j}v_1 + \ldots + a_{nj}v_n$ für $i = 1, \ldots, n$.
Dann gilt:

f orthogonal $\underset{25.3}{\Leftrightarrow}$ Für alle $i, j \in \{1, \ldots, n\}$ gilt $\Phi(f(v_i), f(v_j)) = \Phi(v_i, v_j) = \delta_{ij}$

$\Leftrightarrow$ Für alle $i, j \in \{1, \ldots, n\}$ gilt $\sum\limits_{k=1}^{n}\sum\limits_{l=1}^{n} a_{ki}a_{lj}\underbrace{\Phi(v_k, v_l)}_{\delta_{kl}}$

$$= \sum_{k=1}^{n} a_{ki}a_{kj} = (a_{1i}, \ldots, a_{ni}) \cdot \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix} = [\underbrace{e_i M_B(f)^{tr}}_{\substack{\text{i-te Zeile} \\ \text{von } M_B(f)^{tr}}}] \cdot [\underbrace{M_B(f)e_j}_{\substack{\text{j-te Spalte} \\ \text{von } M_B(f)}}] = \delta_{ij}$$

$\Leftrightarrow$ Für $i, j \in \{1, \ldots, n\}$ ist der (i, j)-Eintrag von $M_B(f)^{tr} \cdot M_B(f)$ gleich δ_{ij}

$\Leftrightarrow M_B(f) \cdot M_B(f)^{tr} = I_n$

$\Leftrightarrow M_B(f)$ ist invertierbar und $M_B(f)^{-1} = M_B(f)^{tr}$ qed.

25.5 Definition:

Eine Matrix $A \in Mat_n(\mathbb{R})$ heißt orthogonal, wenn sie invertierbar ist und $A^{-1} = A^{tr}$ gilt.
Sie ist also genau dann orthogonal, wenn die assoziierte lineare Abbildung $f_A : V \to V$ orthogonal ist.

25.6 Beispiel: (Die orthogonalen Endomorphismen des $\mathbb{R}^2$)

Sei $V = \mathbb{R}^2, E$ die Standardbasis, $\begin{array}{rcl} \Phi : \mathbb{R}^2 \times \mathbb{R}^2 & \to & \mathbb{R} \\ (v, w) & \mapsto & < v, w > \end{array}$ das Standardskalarprodukt bezüglich E und sei $f : \mathbb{R}^2 \to \mathbb{R}^2$ eine $\mathbb{R}$-lineare Abbildung.

Schreibe $M_E(f) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ mit $a, b, c, d \in \mathbb{R}$. Die Abbildung f ist genau dann orthogonal,

wenn $M_E(f)^{tr} \cdot M_E(f) = \begin{pmatrix} a & c \\ b & d \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ gilt. Dies liefert die Gleichungen:

$$\begin{array}{crcl} (I) & a^2 + c^2 & = & 1 \\ (II) & b^2 + d^2 & = & 1 \\ (III) & ab + cd & = & 0 \end{array}$$

Somit gibt es ein $\varphi \in [0, 2\pi)$ mit $a = cos(\varphi)$ und $v = sin(\varphi)$, sowie ein $\psi \in [0, 2\pi)$ mit $b = sin(\psi)$ und $d = cos(\psi)$.

Gleichung (III) liefert dann $sin(\varphi + \psi) = cos(\varphi)sin(\psi) + sin(\varphi)cos(\psi) = 0$.

Dies zeigt $\varphi + \psi \in \{0, \pi, 2\pi, 3\pi\}$ und daher $\psi \in \{-\varphi, \pi - \varphi, 2\pi - \varphi, 3\pi - \varphi\}$. Damit erhalten wir:

$$M_E(f) = \underbrace{\begin{pmatrix} cos(\varphi) & sin(\varphi) \\ sin(\varphi) & -cos(\varphi) \end{pmatrix}}_{\text{Spiegelung}} \text{ oder } M_E(f) = \underbrace{\begin{pmatrix} cos(\varphi) & -sin(\varphi) \\ sin(\varphi) & cos(\varphi) \end{pmatrix}}_{\text{Drehung um den Winkel } \varphi}$$

Im ersten Fall gilt $\chi_f(x) = x^2 - 1 = (x-1)(x+1)$ und in einer geeigneten ONB B von $\mathbb{R}^2$ gilt:

$$M_B(f) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Im zweiten Fall gilt $\chi_f(x) = x^2 - 2cos(\varphi)x + 1$. Für $\varphi \neq 0$ ist f nicht trigonalisierbar.

25.7 Satz: (Eigenschaften orthogonaler Abbildungen)

In obiger Situation gilt:

(a) Die Eigenwerte von f sind in $\{1, -1\}$ enthalten.

(b) Es gilt: $Eig(f, 1) \bigoplus_\Phi Eig(f, -1)$.

(c) Auch die Umkehrabbildung $f^{-1} : V \to V$ ist orthogonal.

(d) Es gilt: $det(f) \in \{1, -1\}$.

(e) Die Spalten von $M_B(f)$ bilden eine ONB von $\mathbb{R}^n$ bezüglich $< \cdot, \cdot >$.
Ebenso bilden die Zeilen von $M_B(f)$ eine ONB von $\mathbb{R}^n$ bezüglich $< \cdot, \cdot >$.

Beweis:

(a) Sei $\lambda \in \mathbb{R}$ ein Eigenwert von f und $v \in V \setminus \{0\}$ mit $f(v) = \lambda v$.
Dann gilt: $\|f(v)\| = \sqrt{\Phi(f(v), f(v))} = \sqrt{\Phi(v, v)} = \|v\|$ und $\|f(v)\| = \|\lambda v\| = |\lambda| \cdot \|v\|$.
Wegen $v \neq 0$ folgt $|\lambda| = 1$, also $\lambda \in \{1, -1\}$.

(b) Sei $v \in Eig(f, 1)$ und $w \in Eig(f, -1)$. Dann gilt:
$\Phi(v, w) = \Phi(f(v), f(w)) = \Phi(v, -w) = -\Phi(v, w)$.
Hieraus ergibt sich $\Phi(v, w) = 0$, also $v \perp_\Phi w$.

(c) Nach Satz 25.5 ist f bijektiv, also ist $f^{-1} : V \to V$ eine wohldefinierte, $\mathbb{R}$-lineare Abbildung. Für $v, w \in V$ gilt $\Phi(f^{-1}(v), f^{-1}(w)) = \Phi(f(f^{-1}(v)), f(f^{-1}(w))) = \Phi(v, w)$.
Also ist f^{-1} orthogonal.

(d) Sei B eine ONB von V. Nach Satz 25.5 gilt $M_B(f)^{tr} \cdot M_B(f) = I_n$. Also folgt:
$1 = det(I_n) = det(M_B(f)^{tr} \cdot M_B(f)) = det(M_B(f)^{tr}) \cdot det(M_B(f)) = det(M_B(f))^2 = det(f)^2$
und somit $det(f) \in \{1, -1\}$.

(e) Sei $B = \{v_1, \ldots, v_n\}$ eine ONB von V. Für $i, j = 1, \ldots, n$ gilt:
$\Phi(f(v_i), f(v_j)) = \Phi(v_i, v_j) = \delta_{ij}$, d.h. auch $C = \{f(v_1), \ldots, f(v_n)\}$ ist eine ONB von V.
Schreibe nun $M_B(f) = (a_{ij})$ mit $a_{ij} \in \mathbb{R}$, d.h. es gelte $f(v_i) = a_{1j}v_1 + \ldots + a_{nj}v_n$ für $j = 1, \ldots, n$.

$$\Rightarrow \delta_{ij} = \Phi(f(v_i), f(v_j)) = \Phi(a_{1i}v_1 + \ldots + a_{ni}v_n, a_{1j}v_1 + \ldots + a_{nj}v_n)$$

$$= \sum_{k=1}^{n}\sum_{l=1}^{n} a_{ki}a_{lj}\Phi(\underbrace{v_k, v_l}_{\delta_{kl}}) = \sum_{k=1}^{n} a_{ki}a_{kj} =< \begin{pmatrix} a_{1i} \\ \vdots \\ a_{ni} \end{pmatrix}, \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix} >$$

Also sind die Spalten von $M_B(f)$ eine ONB von $\mathbb{R}^n$ bezüglich $< \cdot, \cdot >$.
Nach (c) ist auch $f^{-1} : V \to V$ eine orthogonale Abbildung und ihre Matrix bezüglich B ist $M_B(f^{-1}) = M_B(f)^{-1} = M_B(f)^{tr}$. Nach dem bereits bewiesenen sind die Spalten von $M_B(f)^{tr}$ also eine ONB von $\mathbb{R}^n$ bezüglich $< \cdot, \cdot >$. Die Spalten von $M_B(f)^{tr}$ sind genau die Zeilen von $M_B(f)$. qed.

25.8 Beispiel: (Orthogonale Endomorphismen des $\mathbb{R}^3$)

Sei $V = \mathbb{R}^3$ mit Standardskalarprodukt $< \cdot, \cdot >$ bezüglich E und sei $f : \mathbb{R}^3 \to \mathbb{R}^3$ eine orthogonale Abbildung. Das Polynom $\chi_f \in \mathbb{R}[x]$ hat die Gestalt $\chi_f = -x^3 + ax^2 + bx \pm 1$. Es folgt $\lim\limits_{x\to\infty} \chi_f(x) = -\infty$ und $\lim\limits_{x\to-\infty} \chi_f(x) = \infty$. Somits besitzt $\chi_f(x)$ nach dem Zwischenwertsatz eine Nullstelle λ.

Nach Satz 25.7.a folgt $\lambda \in \{1, -1\}$. Sei $v_1 \in \mathbb{R}^3 \setminus \{0\}$ ein Eigenvektor zum Eigenwert λ. Indem wir gegebenenfalls v_1 durch $\frac{v_1}{\|v_1\|}$ ersetzen, können wir $\|v_1\| = 1$ annehmen.

N ach Korollar 23.11 können wir v_1 zu einer ONB $B = \{v_1, v_2, v_3\}$ von $\mathbb{R}^3$ ergänzen. Für den $\mathbb{R}$-Untervektorraum $W = \mathbb{R}v_2 + \mathbb{R}v_3$ gilt dann $f(v_2) \perp v_1$ wegen
$< f(v_2), v_1 >=< f(v_2), \pm f(v_1) >=< v_2, \pm v_1 >= 0$ und $f(v_3) \perp v_1$ wegen
$< f(v_3), v_1 >=< f(v_3), \pm f(v_1) >=< v_3, \pm v_1 >= 0$. Dies zeigt $f(v_2), f(v_3) \in< v_1 >^\perp= W$, also $f(W) \subseteq W$. Daher besitzt $M_B(f)$ die Gestalt

$$M_B(f) = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & a & b \\ 0 & c & d \end{pmatrix} \text{ mit } a, b, c, d \in \mathbb{R}$$

Nach Satz 27.7.d gilt $det(M_B(f)) = \lambda \cdot det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \pm 1$, also $det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \pm 1$. Es ergeben sich vier Fälle:

(a) $\lambda = -1, det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = -1$: Die orthogonale Abbildung $g : \mathbb{R}^2 \to \mathbb{R}^2$ mit Matrix $M_E(g) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ ist nach Beispiel 25.6 eine Spiegelung und hat zwei Eigenwerte $\lambda' = 1, \lambda'' = -1$. Man kann daher $\{v_2, v_3\}$ durch eine ONB $\{v_2', v_3'\}$ von W so ersetzen, dass die Matrix von f bezüglich $B' = \{v_1, v_2', v_3'\}$ die Gestalt

$$M_B'(f) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

hat. Dies ist eine Drehung um 180° um die v_3'-Achse.

(b) $\lambda = 1, det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = -1$. Betrachte wieder die Abbildung $g : \mathbb{R}^2 \to \mathbb{R}^2$ mit Matrix $M_E(g) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Sie ist eine Spiegelung. Daher gibt es eine ONB $B'\{v_1, v_2', v_3'\}$ von V mit

$$M_B'(f) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Dies ist eine Spiegelung an der Ebene $\mathbb{R}v_1 + \mathbb{R}v_3'$.

(c) $\lambda = -1, det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = 1$. Die Abbildung $g : \mathbb{R}^2 \to \mathbb{R}^2$ mit Matrix $M_E(g) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ ist nach Beispiel 25.6 eine Drehung. Also gibt es einem Winkel $\varphi \in [0, 2\pi)$ und eine ONB $\{\tilde{v}_2, \tilde{v}_3\}$ von W, so dass die Matrix von f bezüglich $\tilde{B} = \{v_1, \tilde{v}_2, \tilde{v}_3\}$ die Gestalt

$$M_{\tilde{B}}(f) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & cos(\varphi) & -sin(\varphi) \\ 0 & sin(\varphi) & cos(\varphi) \end{pmatrix}$$

hat. Hier ist f eine <u>Drehspiegelung</u> mit Drehachse $\mathbb{R}v_1$, Drehwinkel φ und Spiegelung an der Ebene $\mathbb{R}\tilde{v}_2 + \mathbb{R}\tilde{v}_3$.

(Spezialfall: $\varphi = \pi \Rightarrow M_B(f) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \Rightarrow f$ ist die Punkstpiegelung an O.)

(d) $\lambda = 1, det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = 1$. Die Abbildung $g : \mathbb{R}^2 \to \mathbb{R}^2$ mit Matrix $M_E(g) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ ist eine Drehung um einen Winkel $\varphi \in [0, 2\pi)$. Also gibt es eine ONB $\tilde{B} = \{v_1, \tilde{v}_2, \tilde{v}_3\}$ mit

$$M_{\tilde{B}}(f) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & cos(\varphi) & -sin(\varphi) \\ 0 & sin(\varphi) & cos(\varphi) \end{pmatrix}$$

Hier ist f eine Drehung um den Winkel φ um die $\mathbb{R}v_1$-Achse.
(Der Fall $\varphi = \pi$ entspricht auch dem Fall (a)).

<u>Fazit:</u> Im $\mathbb{R}^3$ gibt es folgende orthogonale Abbildungen:

- Drehungen (Um Achsen durch O).
- Spiegelungen (An Ebenen durch O).
- Drehspiegelungen (inkl. Punktspiegelung an O) (Drehachse durch O, Spiegelachse senkrecht dazu und auch durch O)

25.9 <u>Theorem:</u> (Normalform orthogonaler Abbildungen)

Sei V ein n-dimensionaler euklidischer Vektorraum mit Skalarprodukt $\Phi : V \times V \to \mathbb{R}$ und sei $f : V \to V$ eine orthogonale, $\mathbb{R}$-lineare Abbildung. Dann gibt es eine ONB $B = \{v_1, \ldots, v_n\}$ von V, so dass die Matrix von f bezüglich B die Gestalt

$$M_B(f) = \left(\begin{array}{ccc|ccc|ccc} 1 & & 0 & & & & & & \\ & \ddots & & & 0 & & & 0 & \\ 0 & & 1 & & & & & & \\ \hline & & & -1 & & 0 & & & \\ & 0 & & & \ddots & & & 0 & \\ & & & 0 & & -1 & & & \\ \hline & & & & & & Q_1 & & 0 \\ & 0 & & & 0 & & & \ddots & \\ & & & & & & 0 & & Q_t \end{array}\right)$$

besitzt, wobei $Q_i \in Mat_2(\mathbb{R})$ von der Form $Q_i = \begin{pmatrix} cos(\varphi_i) & -sin(\varphi_i) \\ sin(\varphi_i) & cos(\varphi_i) \end{pmatrix}$ ist mit $\varphi \in (0, 2\pi) \setminus \{\pi\}$. Mit anderen Worten, die Matrix Q_i ist ein <u>Drehkästchen</u>.

<u>Beweis:</u> Wir schließen mit vollständiger Induktion nach n:

<u>$n = 1$:</u> Klar.

<u>$n > 1$:</u> Es gibt einen f-invarianten $\mathbb{R}$-Untervektorraum $W \subseteq V$ mit $dim_{\mathbb{R}}(W) \in \{1, 2\}$ (vgl. Übungen). Weil f injektiv ist, folgt aus $f(W) \subseteq W$ bereits $f(W) = W$. Für $w \in W$ und $v \in W^{\perp}$ gilt:

$$\Phi(f(v), w) = \Phi(f^{-1}(f(v)), \underbrace{f^{-1}(w)}_{\in W}) = \Phi(v, \underbrace{f^{-1}(w)}_{\in W}) = 0$$

Hieraus folgt $f(W^{\perp}) \subseteq W^{\perp}$ und daher $f(W^{\perp}) = W^{\perp}$. Also sind $g = f_{|W} : W \to W$ und $h = f_{|W^{\perp}} : W^{\perp} \to W^{\perp}$ zwei orthogonale Abbildungen.
Für eine Vektor $v = w + \tilde{w} \in V$ mit $w \in W$ und $\tilde{w} \in W^{\perp}$ gilt: $f(v) = g(w) + h(\tilde{w})$. Wir wenden nun die Induktionsvoraussetzung auf $W^{\perp}$ an und erhalten eine ONB $\tilde{B}$ von $W^{\perp}$, so dass $M_{\tilde{B}}(h)$ die behauptete Gestalt besitzt. Jetzt gibt es zwei Fälle:

<u>1. Fall:</u> $dim_{\mathbb{R}}(W) = 1$. Dann besitzt W eine Basis $\{v\}$ mit einem Eigenvektor v von f, also mit $f(v) \in \{v, -v\}$ und mit $\|v\|_{\Phi} = 1$.
Bezüglich der ONB $B = \{v\} \cup \tilde{B}$ von V besitzt $M_B(f)$ dann die Gestalt

$$M_B(f) = \begin{pmatrix} \pm 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & M_{\tilde{B}}(h) & \\ 0 & & & \end{pmatrix}$$

Bringt man die Vektoren in B als in die richtige Reihenfolge, so hat die Matrix von f die behauptete Gestalt.

2. Fall: Ist $dim_{\mathbb{R}}(W) = 2$, so gibt es nach Beispiel 25.6 eine ONB $\hat{B} = \{v_1, v_2\}$ von W, so dass $M_{\hat{B}}(g)$ die Gestalt

$$M_{\hat{B}}(g) = \begin{pmatrix} cos(\varphi) & -sin(\varphi) \\ sin(\varphi) & cos(\varphi) \end{pmatrix} \text{ mit } \varphi \in (0, 2\pi) \setminus \{\pi\}$$

$$\text{oder } M_{\hat{B}}(g) = \begin{pmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{pmatrix} \text{ hat.}$$

Bezüglich der Basis $B = \hat{B} \cup \tilde{B}$ hat $M_B(f)$ dann die Gestalt

$$M_B(f) = \begin{pmatrix} Q & 0 \\ 0 & M_{\tilde{B}}(h) \end{pmatrix} \text{ mit einem Drehkästchen Q oder}$$

$$M_B(f) = \left(\begin{array}{cc|c} \pm 1 & 0 & 0 \\ 0 & \pm 1 & \\ \hline 0 & & M_{\tilde{B}}(h) \end{array}\right)$$

Ordnet man die Basis B also geeignet an, so hat $M_B(f)$ die behauptete Gestalt. qed.

B. Und was hat das alles mit Fußball zu tun?

Sie V ein endlich-dimensionaler euklidischer Vektorraum mit Skalarprodukt $\Phi : V \times V \to \mathbb{R}$.

25.10 Bemerkung:

Sei $f : \mathbb{R}^n \to \mathbb{R}^n$ eine $\mathbb{R}$-lineare Abbildung. Die Determinante

$$|det(f)| = |det(M_E(f))| = det\,|(f(e_1), \ldots, f(e_n))|$$

misst das Volumen des Bildes des Einheitsquaders, d.h. das Volumen von

$$Q = \{c_1 f(e_1) + \ldots + c_n f(e_n) \mid 0 \leq c_i \leq 1\}$$

Somit ist $|det(f)|$ ein Maß für die Volumenänderung bei der Anwendung der Abbildung f.

Was aber ist die Bedeutung von $Vorz(det(f)) \in \{1, -1\}$?

25.11 Definition:

Sei V ein n-dimensionaler $\mathbb{R}$-Vektorraum und $f : V \to V$ ein Automorphismus.
Dann heißt die Abbildung f

(a) orientierungstreu (oder orientierungserhaltend), wenn $det(f) > 0$ gilt und

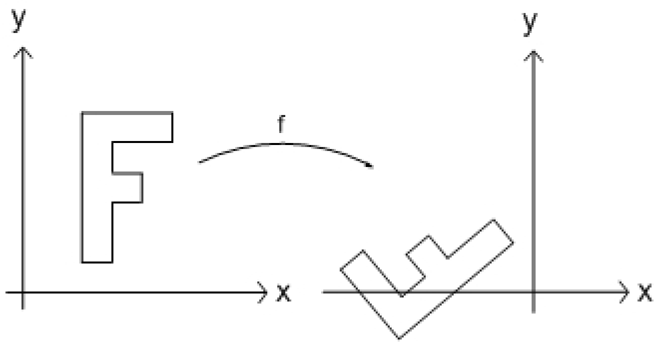
(b) orientierungsuntreu, wenn $det(f) < 0$ gilt.

25.12 Beispiel:

(a) Betrachte die $\mathbb{R}$-lineare Abbildung $f : \mathbb{R}^2 \to \mathbb{R}^2$ mit

$$M_E(f) = \begin{pmatrix} -1 & -\frac{1}{4} \\ 1 & -\frac{1}{2} \end{pmatrix}$$

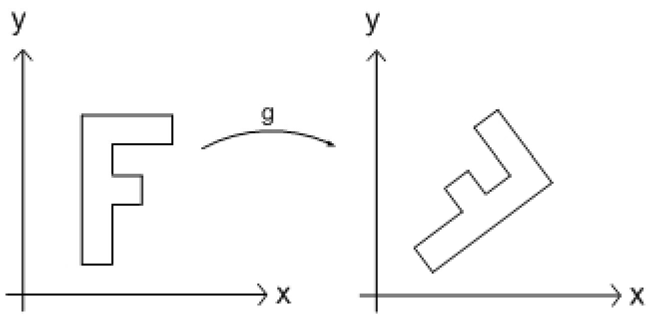
Es gilt $det(f) = \frac{3}{4} > 0$



(b) Betrachte nun $g : \mathbb{R}^2 \to \mathbb{R}^2$ mit

$$M_E(g) = \begin{pmatrix} -\frac{1}{4} & 1 \\ 1 & \frac{1}{2} \end{pmatrix}$$

Es gilt $det(g) = -\frac{9}{8} < 0$



25.13 Definition:

Seien $B = \{v_1, \ldots, v_n\}$ und $C = \{w_1, \ldots, w_n\}$ zwei Basen eines $\mathbb{R}$-Vektorraums V. Sei $f : V \to V$ ein eindeutig bestimmter Automorphismus mit $f(v_i) = w_i$ für $i = 1, \ldots, n$.

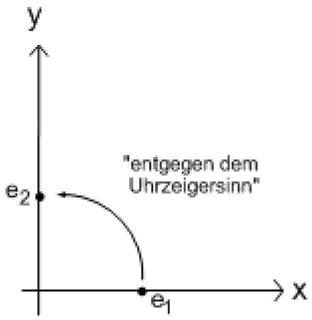
(a) Die Basen B und C heißen gleichorientiert, wenn $det(f) > 0$ gilt.
Schreibweise: $B \underset{or}{\sim} C$
Es ist leicht zu sehen, dass $\underset{or}{\sim}$ eine Äquivalenzrelation ist.

(b) Sei M die Menge aller Basen von V. Dann ist $\underset{or}{\sim}$ eine Äquivalenzrelation auf M.
Eine Orientierung auf V ist eine Äquivalenzklasse gleich orientierter Basen.
Schreibweise: $[B]$ sei die Äquivalenzklasse von B.
Es gibt genau zwei Orientierungen: $M = [B] \dot{\cup} [C]$

25.14 Beispiel: (Orientierungen im $\mathbb{R}^2$ und $\mathbb{R}^3$)

(a) Die Standardbasis von $\mathbb{R}^2$, also $E = \{e_1, e_2\}$, hat eine Äquivalenzklasse $[E]$, die wir die Orientierung entgegen dem Uhrzeigersinn (oder die mathematisch-positive Orientierung) nennen.

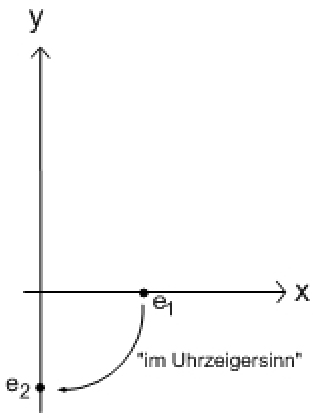
Skizze:



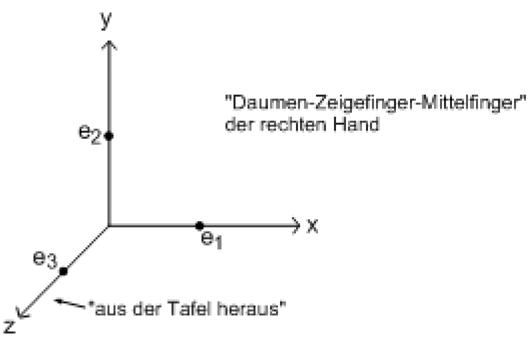
Die Basis $B = \{e_1, -e_2\}$ hat eine andere Orientierung,

denn $\begin{array}{rcl} f : \mathbb{R}^2 & \to & \mathbb{R}^2 \\ e_1 & \mapsto & e_1 \\ e_2 & \mapsto & -e_2 \end{array}$ erfüllt $det(f) = det \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = -1 < 0$

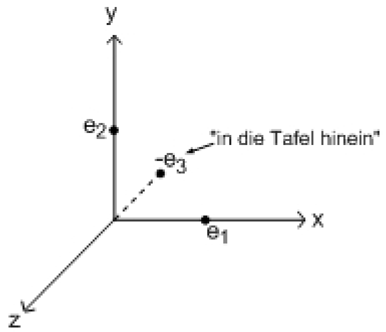
Die Orientierung $[B]$ heißt die Orientierung im Uhrzeigersinn.
Skizze:



(b) Die Standardbasis $E = \{e_1, e_2, e_3\}$ von $\mathbb{R}^3$ besitzt die rechte-Hand-Orientierung (oder mathematisch-positive Orientierung) $[E]$ von $\mathbb{R}^3$.
Skizze:



Die Basis $B = \{e_1, e_2, -e_3\}$ von $\mathbb{R}^3$ erfüllt $[B] \neq [E]$. Ihre Orientierung heißt die linke-Hand-Orientierung.
Skizze:



25.15 Satz: (Der Satz vom Fußball)

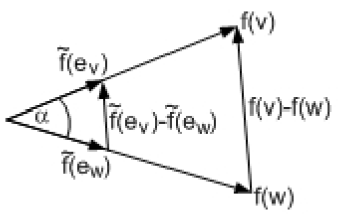
Bei einem Fußballspiel liegt der Ball zu Beginn der ersten und der zweiten Halbzeit jeweils am Anstoßpunkt. Dann gibt es genau 2 Punkte auf der Balloberfläche, die sich zu beiden Zeitpunkten genau an der gleichen Stelle des umgebenden Raumes befinden.
(Oder alle Punkte der Balloberfläche befinden sich zweimal exakt an derselben Stelle im umgebenden Raum.)

Beweis: Das Koordinatensystem von $\mathbb{R}^3$ sei so gewählt, dass der Mittelpunkt des Balles im Ursprung O liegt. O.E. sei das Koordinatensystem so gewählt, dass der Radius des Balles gerade 1 ist. Die Balloberfläche ist dann $S = \{(a, b, c) \in \mathbb{R}^3 \mid a^2 + b^2 + c^2 = 1\}$.
Sei $\tilde{f} : S \to S$ die Abbildung, die jedem Punkt der Balloberfläche zu Beginn des Spieles seine Position zu Beginn der zweiten Halbzeit zuordnet. Die Abbildung $\tilde{f}$ ist offenbar längentreu.
Nun definieren wir eine Abbildung $f : \mathbb{R}^3 \to \mathbb{R}^3$ durch $f(v) = \|v\| \cdot \tilde{f}(\frac{v}{\|v\|})$. Dann gilt:
$\|f(v)\| = \|v\| \cdot \|\tilde{f}(\frac{v}{\|v\|})\| = \|v\|$ für alle $v \in \mathbb{R}^3$.

Behauptung 1: Die Abbildung f erfüllt: $< f(v), f(w) >=< v, w >$ für alle $v, w \in \mathbb{R}^3$.

Beweis: Es gilt: $< f(v), f(w) >= -\frac{1}{2}[\|f(v) - f(w)\|^2 - \|f(v)\|^2 - \|f(w)\|^2]$
$$= -\frac{1}{2}[\underbrace{d(f(v), f(w))^2}_{\text{Distanz von } f(v) \text{ und } f(w)} - \|v\|^2 - \|w\|^2]$$

Skizze:



Es gilt: $\|f(v) - f(w)\|^2 = \|f(v)\|^2 + \|f(w)\|^2 - 2 \cdot \|f(v)\| \cdot \|f(w)\| \cdot cos(\alpha)$
$= \|v\|^2 + \|w\|^2 - 2 \cdot \|v\| \cdot \|w\| \cdot cos(\alpha) = \|v\|^2 + \|w\|^2 - 2 \cdot \|v\| \cdot \|w\| \cdot < e_v, e_w >$
denn $cos(\alpha) = \dfrac{< e_v, e_w >}{1 \cdot 1} =< e_v, e_w >$
Dies liefert $< f(v), f(w) >= \|v\| \cdot \|w\| \cdot cos(\alpha) =< v, w > \Rightarrow$ Behauptung 1

Behauptung 2: Die Abbildung f ist $\mathbb{R}$-linear.

Beweis: Schreibe $v \in \mathbb{R}^3$ als $v =< v, e_1 > e_1+ < v, e_2 > e_2+ < v, e_3 > e_3$
und $f(v) =< f(v), e_1 > e_1+ < f(v), e_2 > e_2+ < f(v), e_3 > e_3$.
Wegen Behauptung 1 ist $\{f(e_1), f(e_2), f(e_3)\}$ wieder eine ONB von $\mathbb{R}^3$. Somit gilt auch:

$$f(v) =< f(v), f(e_1) > f(e_1)+ < f(v), f(e_2) > f(e_2)+ < f(v), f(e_3) > f(e_3)$$

$$=< v, e_1 > f(e_1)+ < v, e_2 > f(e_2)+ < v, e_3 > f(e_3)$$

Seien nun $v, w \in \mathbb{R}^3$ und $\lambda, \mu \in \mathbb{R}$. Dann gilt:

$$f(\lambda v + \mu w) =< \lambda v + \mu w, e_1 > f(e_1)+ < \lambda v + \mu w, e_2 > f(e_2)+ < \lambda v + \mu w, e_3 > f(e_3)$$

$$= \lambda < v, e_1 > f(e_1) + \mu < w, e_1 > f(e_1) + \lambda < v, e_2 > f(e_2) + \mu < w, e_2 > f(e_2)$$
$$+\lambda < v, e_3 > f(e_3) + \mu < w, e_3 > f(e_3) = \lambda f(v) + \mu f(w) \Rightarrow \underline{\text{Behauptung 2}}$$

Also ist $f : \mathbb{R}^3 \to \mathbb{R}^3$ eine orthogonale, $\mathbb{R}$-lineare Abbildung. Ferner ist f orientierungserhaltend, wie man sieht, indem man sich in den Ball ein Koordinatensystem eingebaut denkt. Dies liefert $det(f) = 1$.

Nach Beispiel 25.8 ergibt sich, dass f einen Eigenwert $\lambda = 1$ besitzt. Die beiden Vektoren der Länge 1 in $Eig(f, 1)$ sind also Fixpunkte von f. Mögliche Matrizen für f:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & cos(\alpha) & -sin(\alpha) \\ 0 & sin(\alpha) & cos(\alpha) \end{pmatrix} \text{ oder } \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ oder } \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Insgesamt ist f eine Drehung um eine geeignete Achse und hat genau 2 Fixpunkte. qed.

C. Gibt es auch nicht-lineare, längentreue Abbildungen?

25.16 Beispiel:

Sei $a \in V$ ein fester Vektor. Dann heißt die Abbildung $\tau_a : V \to V,\ v \mapsto v + a$ die Translation (oder die Verschiebung) um a. Die Abbildung τ_a ist für $a \neq 0$ nicht $\mathbb{R}$-linear, denn
$\tau_a(v + w) = v + w + a \neq \tau_a(v) + \tau_a(w) = v + a + w + a = v + w + 2a$.

Die Abbildung τ_a ist bijektiv und es gilt $(\tau_a)^{-1} = \tau_{-a}$.

Die Abbildung τ_a ist längentreu, denn $\|\tau_a(v) - \tau_a(w)\| = \|(v + a) - (w + a)\| = \|v - w\|$ für alle $v, w \in V$.

25.17 Satz:

Sei V ein euklidischer Vektorraum mit Skalarprodukt $\Phi : V \times V \to \mathbb{R}$ und sei $f : V \to V$ eine längentreue Abbildung. Dann gibt es einen festen Vektor $a \in V$ und eine orthogonale, $\mathbb{R}$-lineare Abbildung $g : V \to V$ mit $f = \tau_a \circ g$. Insbesondere ist f bijektiv.

Beweis: Sei $a = f(0)$ und $g = \tau_{-a} \circ f$.

Dann ist $g : V \to V$ eine längentreue Abbildung mit $g(0) = \tau_{-a}(f(0)) = \tau_{-a}(a) = 0$.
Insbesondere gilt: $\|g(v)\| = d(g(v), g(0)) = d(v, 0) = \|v\|$. Für alle $v \in V$ folgt somit:

$$\begin{aligned}\Phi(g(v), g(w)) &= -\tfrac{1}{2}[\|g(v) - g(w)\|^2 - \|g(v)\|^2 - \|g(w)\|^2] \\ &= -\tfrac{1}{2}[\underbrace{d(g(v), g(w))^2}_{= d(-v,w)^2} - \|v\|^2 - \|w\|^2] \\ &= -\tfrac{1}{2}[\|v - w\|^2 - \|v\|^2 - \|w\|^2] = \Phi(v, w)\end{aligned}$$

Nun zeigen wir noch, dass g $\mathbb{R}$-linear ist ($\Rightarrow$ Satz):

Sei $B = \{v_1, \dots, v_n\}$ eine ONB von V bezüglich Φ. Dann folgt, dass auch $\{g(v_1), \dots, g(v_n)\}$ eine ONB von V ist. Ein Vektor $v \in V$ besitzt Darstellungen:

$$g(v) = \sum_{i=1}^{n} \Phi(g(v), g(v_i))g(v_i) = \sum_{i=1}^{n} \Phi(v, v_i)g(v_i)$$

Für $v, w \in V$ und $\lambda, \mu \in \mathbb{R}$ erhalten wir:

$$\begin{aligned}g(\lambda v + \mu w) = \sum_{i=1}^{n} \Phi(\lambda v + \mu w, v_i)g(v_i) &= \lambda \sum_{i=1}^{n} \Phi(v, v_i)g(v_i) + \mu \sum_{i=1}^{n} \Phi(w, v_i)g(v_i) \\ &= \lambda g(v) + \mu g(w)\end{aligned}$$

qed.

25.18 Definition:

Sei V ein reeller Vektorraum und sei $d : V \times V \to \mathbb{R}$ eine Metrik.
Eine Abbildung $f : V \to V$ heißt eine Isometrie von V, wenn für alle $v, w \in V$ gilt:
$d(f(v), f(w)) = d(v, w)$.

25.19 Korollar:

Ist V ein euklidischer Vektorraum, so ist jede Isometrie von V die Komposition einer $\mathbb{R}$-linearen, orthogonalen Abbildung $g : V \to V$ mit einer Translation.

26 Gespiegelt und doch gleich! (Symmetrische Matrizen)

A. Sind symmetrische Matrizen etwas Besseres, oder halten sie sich nur dafür?

26.1 Theorem:

Ist $A \in Mat_n(\mathbb{R})$ eine symmetrische Matrix, so besitzt A mindestens einen reellen Eigenwert.

Beweis: (Unter Verwendung von Analysis II)

Betrachte die Abbildung
$$\begin{aligned} q : \mathbb{R}^n &\to \mathbb{R} \\ (x_1, \ldots, x_n) &\mapsto (x_1, \ldots, x_n) \cdot A \cdot (x_1, \ldots, x_n)^{tr} = q(x_1, \ldots, x_n) \end{aligned}.$$
Dies ist eine Polynomfunktion vom Grad 2 in den Unbestimmten $x_1, \ldots, x_n$.
Sei $S = \{(a_1, \ldots, a_n) \in \mathbb{R}^n \mid a_1^2 + \ldots + a_n^2 = 1\}$ die Einheitssphäre. Da S eine kompakte Teilmenge von $\mathbb{R}^n$ und q eine stetige Funktion ist, nimmt q auf S ein Maximum an.
Sei also $v = (x_1, \ldots, x_n) \in S$ mit $vAv^{tr} \geq wAw^{tr}$ für alle $w \in S$.

Behauptung: Ist $w \in S$ und $w \perp v$, so gilt $w \perp Av^{tr}$.

Beweis: Sei $\lambda \in [0, 1]$ und $\mu = \sqrt{1 - \lambda^2}$ und sei $u = \lambda v + \mu w$.
Wegen $w \perp v$ gilt dann $\|u\|^2 =< \lambda v + \mu w, \lambda v + \mu w >= \lambda^2 \|v\|^2 + \mu^2 \|w\|^2 = 1$, also $u \in S$. Nun gilt:

$$vAv^{tr} \geq uAu^{tr} = (\lambda v + \mu w)A(\lambda v + \mu w)^{tr} = \lambda^2 \cdot vAv^{tr} + \lambda\mu \underbrace{vAw^{tr} + \lambda\mu\, wAv^{tr}}_{\text{gleich, da } A \text{ symmmetrisch}} + \mu^2 wAw^{tr}$$
$$= \lambda^2 vAv^{tr} + \mu^2 wAw^{tr} + 2\lambda\mu wAv^{tr}$$

Hieraus folgt: $2\lambda\mu wAw^{tr} \leq (\underbrace{1 - \lambda^2}_{= \mu^2})vAv^{tr} - \mu^2 wAw^{tr}$

Für $\mu \neq 0, \lambda \neq 1$ folgt: $2\lambda wAv^{tr} = \mu(\underbrace{vAv^{tr} - wAw^{tr}}_{\geq 0, \text{ da } vAv^{tr} \text{ maximal war}})$

Angenommen, $w \perp Av^{tr}$ gilt nicht, d.h. angenommen $wAv^{tr} \neq 0$.
O.E. sei dann $wAv^{tr} > 0$ (sonst ersetze w durch $-w$).
Dann folgt $0 < 2\sqrt{\underbrace{1 - \mu^2}_{\to 1}}(wAv^{tr}) \leq \underbrace{\mu}_{\to 0}(vAv^{tr} - wAw^{tr})$
und für $\mu \to 0$ folgt ein Widerspruch. $\Rightarrow$ Behauptung

Betrachte nun den $\mathbb{R}$-Untervektorraum $W = (\mathbb{R} \cdot v)^\perp$. Es gilt $W^\perp = \mathbb{R} \cdot v$ und $Av^{tr} \in W^\perp = \mathbb{R} \cdot v$.
Somit gibt es ein $\lambda \in \mathbb{R}$ mit $Av^{tr} = \lambda \cdot v^{tr}$, d.h. v ist ein Eigenvektor von A. Wegen $v \neq 0$ besitzt A den Eigenwert $\lambda \in \mathbb{R}$. qed.

26.2 Korollar:

Sei V ein n-dimensionaler $\mathbb{R}$-Vektorraum und $f : V \to V$ ein Endomorphismus. Gibt es eine Basis B von V, so dass $M_B(f)$ symmetrisch ist, so besitzt f einen Eigenwert $\lambda \in \mathbb{R}$.

Frage: Welche Endomorphismen besitzen symmetrische Darstellungsmatrizen?

26.3 Satz:

Sei V ein n-dimensionaler, euklidischer Vektorraum mit Skalarprodukt $\Phi : V \times V \to \mathbb{R}$ und sei $f : V \to V$ eine $\mathbb{R}$-lineare Abbildung. Dann sind die folgenden Bedingungen äquivalent:

(a) Es gibt eine ONB B von V, so dass $M_B(f)$ symmetrisch ist.

(b) Für jede ONB B von V ist $M_B(f)$ symmetrisch.

(c) Für alle $v, w \in V$ gilt: $\Phi(f(v), w) = \Phi(v, f(w))$.

(d) Ist $B = \{v_1, \ldots, v_n\}$ eine ONB von V, so gilt $\Phi(f(v_i), v_j) = \Phi(v_i, f(v_j))$ für alle i, j.

Sind diese Bedingungen erfüllt, so heißt f ein selbstadjungierter Endomorphismus.

Beweis:„b⇒a“ Klar.

„d⇒b“ Sei $M_B(f) = (a_{ij})$, d.h. es gelte $f(v_j) = a_{1j}v_1 + \ldots + a_{nj}v_n$. Dann gilt:

$$\Phi(f(v_i), v_j) = \Phi(a_{1i}v_1 + \ldots + a_{ni}v_n, v_j) = a_{ji} \text{ und}$$

$$\Phi(v_i, f(v_j)) = \Phi(v_i, a_{1j}v_1 + \ldots + a_{nj}v_n) = a_{ij}$$

Also folgt $a_{ji} = a_{ij}$ für alle i, j, d.h. $M_B(f)$ ist symmetrisch.

„c⇒d“ Klar.

„a⇒c“ Sei $M_B(f) = (a_{ij})$ mit $a_{ij} = a_{ji}$ für alle i, j. Schreibe $B = \{v_1, \ldots, v_n\}$ und $v = b_1v_1 + \ldots + b_nv_n, w = c_1v_1 + \ldots + c_nv_n$ mit $b_k, c_l \in \mathbb{R}$. Dann gilt:

$$\Phi(f(v), w) = \Phi(f(b_1v_1 + \ldots + b_nv_n), c_1v_1 + \ldots + c_nv_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} b_ic_j\Phi(f(v_i), v_j)$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} b_ic_j\Phi(a_{1i}v_1 + \ldots + a_{ni}v_n, v_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} b_ic_ja_{ji} = \sum_{i=1}^{n}\sum_{j=1}^{n} b_ic_ja_{ij}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} b_ic_j\Phi(v_i, a_{1j}v_1 + \ldots + a_{nj}v_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} b_ic_j\Phi(v_i, f(v_j))$$

$$= \Phi(b_1v_1 + \ldots + b_nv_n, f(c_1v_1 + \ldots + c_nv_n)) = \Phi(v, f(w))$$

qed.

26.4 Theorem:

Sei V ein n-dimensionaler, euklidischer Vektorraum und $f : V \to V$ ein selbstadjungierter Endomorphismus. Dann gibt es eine ONB B von V, so dass $M_B(f)$ eine Diagonalmatrix ist.

Beweis: Wir schließen mit vollständiger Induktion nach n:

$n = 1$: $M_B(f)$ ist stets eine Diagonalmatrix.

$n > 1$: Sei C eine ONB von V. Nach Satz 26.3 ist $M_C(f)$ eine symmetrische Matrix. Nach Korollar 26.2 besitzt f dann einen Eigenwert $\lambda \in \mathbb{R}$.
Sei $v \in V \setminus \{0\}$ ein Eigenvektor zum Eigenwert λ. Setze $W = (\mathbb{R} \cdot v)^\perp$. Dann gilt: $V = W \bigoplus \mathbb{R} \cdot v$ und $dim_{\mathbb{R}}(W) = n - 1$.

Ferner gilt: $f(W) \subseteq W$, denn für alle $w \in W$ gilt:
$\Phi(f(w), v) = \Phi(w, f(v)) = \Phi(w, \lambda v) = \lambda\Phi(w, v) = 0$, also $f(w) \in (\mathbb{R} \cdot v)^\perp = W$.

Nun wenden wir die Induktionsvoraussetzung auf die Abbildung $g = f_{|W} : W \to W$ an. Somit gibt es eine ONB $\tilde{B} = \{v_2, \ldots, v_n\}$ von W mit $M_{\tilde{B}}(g)$ als Diagonalmatrix.
Die Menge $\hat{B} = \{v, v_2, \ldots, v_n\}$ ist eine OGB von V. Ersetzen wir $v_1 = \frac{v}{\|v\|}$, so ist $B = \{v_1, \ldots, v_n\}$ eine ONB von V.
Bezüglich der Basis B hat

$$M_B(f) = \left(\begin{array}{c|ccc} \lambda & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & M_{\tilde{B}}(g) & \\ 0 & & & \end{array}\right)$$

Diagonalgestalt. qed.

26.5 Korollar: (Algorithmus zur Diagonalisierung selbstadjungierter Abbildungen)

Sei $f : V \to V$ ein selbstadjungierter Endomorphismus.

(a) Berechne $\chi_f(x)$ und eine Faktorisierung $\chi_f(x) = (-1)^n(x - \lambda_1)^{\mu_1} \cdots (x - \lambda_m)^{\mu_m}$ mit paarweise verschiedenen $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$ und mit $\mu_i > 0$.

(b) Für $i = 1, \ldots, m$ berechne eine Basis B_i von $Eig(f, \lambda_i)$.

(c) Für $i = 1, \ldots, m$ wende das Schmidtsche Orthonormalisierungsverfahren auf B_i an und erhalte eine ONB C_i von $Eig(f, \lambda_i)$.

(d) Gib $C = C_1 \cup \ldots \cup C_m$ aus und stoppe.

Dies ist ein Algorithmus, der eine ONB C von V berechnet, so dass $M_C(f) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_m \end{pmatrix}$ eine Diagonalmatrix ist.

Beweis: Es ist nur zu zeigen, dass $V = Eig(f, \lambda_1) \bigoplus \ldots \bigoplus Eig(f, \lambda_m)$ gilt.

Schreibe $\chi_f(x) = (-1)^n(x - \lambda_1)^{\mu_1} \cdots (x - \lambda_m)^{\mu_m}$ wie im Algorithmus. Nach dem Theorem gibt es eine ONB $B = \{v_1, \ldots, v_n\}$ von V bestehend aus Eigenvektoren. Also müssen jeweils μ_i dieser Vektoren eine ONB von $Eig(f, \lambda_i)$ bilden. Wegen $v_k \perp v_l$ für $k = l$ folgt $Eig(f, \lambda_i) \perp Eig(f, \lambda_j)$ für $i \neq j$. qed.

26.6 Beispiel:

Gegeben sei die $\mathbb{R}$-lineare Abbildung $f : \mathbb{R}^3 \to \mathbb{R}^3$ mit

$$A = M_E(f) = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{1}{3} & -\frac{14}{15} & \frac{2}{15} \\ \frac{2}{3} & \frac{2}{15} & -\frac{11}{15} \end{pmatrix}$$

1.) A ist symmetrisch.
2.) A ist orthonormal, denn die Spalten stehen aufeinander senkrecht und haben die Länge 1.

Wir wollen f bzw. A diagonalisieren:

(a) $\chi_f(x) = -x^3 - x^2 + x + 1 = -(x+1)^2(x-1)$

(b) Berechne $Eig(f,1) = \mathbb{R} \cdot (5,1,2)$ und $Eig(f,-1) = \mathbb{R} \cdot (0,2,-1) \oplus \mathbb{R} \cdot (1,-1,-2)$

(c) Orthonormalisiere $Eig(f,1)$ und erhalte $< (\frac{5}{\sqrt{30}}, \frac{1}{\sqrt{30}}, \frac{2}{\sqrt{30}}) >$
Orthonormalisiere $Eig(f,-1)$ und erhalte $< (0, \frac{2}{\sqrt{5}}, -\frac{1}{\sqrt{5}}), (\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}) >$

(d) Also ist $C = \{(\frac{5}{\sqrt{30}}, \frac{1}{\sqrt{30}}, \frac{2}{\sqrt{30}}), (0, \frac{2}{\sqrt{5}}, -\frac{1}{\sqrt{5}}), (\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}})\}$ eine ONB von $\mathbb{R}^3$ mit

$$M_C(f) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

B. Und was bedeutet dies für symmetrische Matrizen?

26.7 Satz: (Diagonalisierung symmetrischer Matrizen)

Sei $A \in Mat_n(\mathbb{R})$ eine symmetrische Matrix. Dann gibt es eine orthogonale Matrix $T \in O_n(\mathbb{R})$, so dass $T^{tr}AT = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_m \end{pmatrix}$ eine Diagonalmatrix mit $\lambda_i \in \mathbb{R}$ ist.

Beweis: Sei $f : \mathbb{R}^n \to \mathbb{R}^n$ der Endomorphismus mit Matrix $M_E(f) = A$.
Da E eine ONB bezüglich $< \cdot, \cdot >$ darstellt, ist f eine selbstadjungierte, $\mathbb{R}$-lineare Abbildung.
Also gibt es eine ONB B von $\mathbb{R}^n$ bezüglich der $M_B(f)$ eine Diagonalmatrix ist.
Sei T die Transformationsmatrix von E nach B. Dann ist T orthogonal und es folgt
$M_B(f) = T^{-1}AT = T^{tr}AT$. qed.

Frage: Wie kann man die Diagonalisierung explizit durchführen?

- erste Möglichkeit: Verwende die Methode aus Korollar 26.5
- zweite Möglichkeit: Simultane Zeilen- und Spaltenumformungen

26.8 Definition:

$$\text{Sei } A = \begin{array}{cc} \begin{array}{c} \\ i \\ \\ j \\ \\ \end{array} & \left(\begin{array}{ccccc} \ddots & \vdots & & \vdots & \\ \cdots & a & \cdots & b & \cdots \\ & \vdots & \ddots & \vdots & \\ \cdots & b & \cdots & c & \cdots \\ & \vdots & & \vdots & \ddots \end{array}\right) \\ & \begin{array}{ccccc} \quad & i & \quad & j & \quad \end{array} \end{array} \in Mat_n(\mathbb{R}) \text{ symmetrisch mit } a, b, c \in \mathbb{R} \text{ und } a \neq 0.$$

Führe die folgenden beiden Operationen durch:

1.) Addiere das $-\frac{b}{a}$-fache der i-ten Zeile zur j-ten Zeile.

2.) Addiere das $-\frac{b}{a}$-fache der i-ten Spalte zur j-ten Spalte

Das Ergebnis ist die symmetrische Matrix:

$$\tilde{A} = \begin{pmatrix} \ddots & \vdots & & \vdots & \\ \cdots & a & \cdots & 0 & \cdots \\ & \vdots & \ddots & \vdots & \\ \cdots & 0 & \cdots & d & \cdots \\ & \vdots & & \vdots & \ddots \end{pmatrix} \text{ mit } d = -\frac{b^2}{a}$$

Die kombinierte Operation nennt man eine simultane Zeilen- und Spaltenoperation.

26.9 Bemerkung: (Diagonalisierung mit simultanen Zeilen- und Spaltenumformungen)

Gegeben sei eine symmetrische Matrix $A \in Mat_n(\mathbb{R})$. Dann kann man A mit simultanen Zeilen- und Spaltenoperationen in Diagonalgestalt überführen.
Wenden wir dabei eine elementare Zeilenoperation C_i^{tr} an, so müssen wir anschließend sofort die entsprechende Spaltenoperation C_i durchführen.
Gehe dabei wie folgt vor:

(a) Ist $a_{ii} \neq 0$, so kann man die i-te Zeile/Spalte mit Hilfe von a_{ii} „ausräumen“.

(b) Ist $a_{ii} = 0$ und ein weiteres $a_{jj} = 0$, und gibt es in der i-ten Zeile/Spalte ein Element $b_{ij} = b_{ji} \neq 0$, so addiere die j-te Zeile/Spalte zur i-ten und erhalte auf der Diagonalen $\tilde{a}_{ij} = 2b_{ij}$. Wende dann (a) an.

Wir führen wie folgt Buch:

$$\begin{array}{ccccc|ccccccc} & & & I_n & & & & & A & & \\ & & C_1^{tr} & I_n & & & & C_1^{tr} & A & C_1 & \\ & C_2^{tr} & C_1^{tr} & I_n & & & C_2^{tr} & C_1^{tr} & A & C_1 & C_2 \\ & & & \vdots & & & & & \vdots & & \\ C_r^{tr} & \cdots & C_1^{tr} & I_n & = T^{tr} & D = & C_r^{tr} & \cdots C_1^{tr} & A & C_1 & \cdots C_r \end{array}$$

Das Verfahren berechnet auch die Transformationsmatrix $T \in GL_n(\mathbb{R})$.

Warnung: T ist im Allgemeinen nicht orthogonal!

26.10 Beispiel:

Gegeben sei $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \in Mat_2(\mathbb{R})$ mit $a \neq 0$.

$$\begin{array}{c|c} I_2 & A \\ \begin{pmatrix} 1 & 0 \\ -\frac{b}{a} & 1 \end{pmatrix} & \begin{pmatrix} a & 0 \\ 0 & c - \frac{b^2}{a} \end{pmatrix} \\ \shortparallel & \\ T^{tr} & \end{array}$$

Also gilt: $D = T^{tr}AT$ mit $T = \begin{pmatrix} 1 & -\frac{b}{a} \\ 0 & 1 \end{pmatrix}$

26.11 Beispiel:

Wir wollen die Matrix $A = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 4 \\ 2 & 4 & 4 \end{pmatrix} \in Mat_3(\mathbb{Q})$ diagonalisieren:

$$\begin{array}{ccc|ccc}
1 & 0 & 0 & 1 & 2 & 2 \\
0 & 1 & 0 & 2 & 1 & 4 \\
0 & 0 & 1 & 2 & 4 & 4 \\
\hline
1 & 0 & 0 & 1 & 0 & 2 \\
-2 & 1 & 0 & 0 & -3 & 0 \\
0 & 0 & 1 & 2 & 0 & 4 \\
\hline
1 & 0 & 0 & 1 & 0 & 0 \\
-2 & 1 & 0 & 0 & -3 & 0 \\
-2 & 0 & 1 & 0 & 0 & 0 \\
 & \shortparallel & & & \shortparallel & \\
 & T^{tr} & & & D &
\end{array}$$

Also gilt $D = T^{tr}AT$ mit $T = \begin{pmatrix} 1 & -2 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

26.12 Satz:

Sei $A \in Mat_n(\mathbb{R})$ eine symmetrische Matrix und sei $\Phi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ die symmetrische Bilinearform mit $G_E(\Phi) = A$.

(a) Es gibt eine Basis $B = \{v_1, \ldots, v_n\}$ von $\mathbb{R}^n$, so dass $G_B(\Phi)$ die folgende Gestalt besitzt:

$$G_B(\Phi) = \left(\begin{array}{c|c|c}
\begin{matrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{matrix} & 0 & 0 \\
\hline
0 & \begin{matrix} -1 & & 0 \\ & \ddots & \\ 0 & & -1 \end{matrix} & 0 \\
\hline
0 & 0 & \begin{matrix} 0 & & \\ & \ddots & \\ & & 0 \end{matrix}
\end{array}\right) \quad \text{mit } k, l \geq 0$$

$$\underbrace{\hphantom{xxxxxx}}_{k} \quad \underbrace{\hphantom{xxxxxx}}_{l}$$

(b) Es gibt eine Matrix $T \in GL_N(\mathbb{R})$, so dass $T^{tr}AT$ die in (a) angegebene Gestalt besitzt.

Beweis:

(a) Wähle gemäß Theorem 26.4 eine ONB $\tilde{B} = \{\tilde{v}_1, \ldots, \tilde{v}_n\}$ von $\mathbb{R}^n$ (bezüglich $< \cdot, \cdot >$) bestehend aus Eigenvektoren von A, so dass $M_{\tilde{B}}(f_A)$ eine Diagonalmatrix ist. Mit anderen Worten, es gibt eine Matrix $S \in O_n(\mathbb{R})$, so dass $S^{tr}AS = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ eine Diagonalmatrix ist.

O.E. sei $\{\tilde{v}_1, \ldots, \tilde{v}_n\}$ so numeriert, dass $\lambda_1, \ldots, \lambda_k > 0$ und $\lambda_{k+1}, \ldots, \lambda_{k+l}$ und $\lambda_{k+l+1} = \ldots = \lambda_n = 0$ gilt.

Definiere nun $B = \{v_1, \ldots, v_n\}$ durch

$$v_i = \begin{cases} \frac{1}{\sqrt{|\lambda_i|}} \cdot \tilde{v}_i & , \text{ für } i = 1, \ldots, k+l \\ \tilde{v}_i & , \text{ für } i = k+l+1, \ldots, n \end{cases}$$

Dann folgt: $\Phi(v_i, v_j) = v_i \cdot G_E(\Phi) \cdot v_j^{tr} = v_i A v_j^{tr}$

$$= \begin{cases} \frac{1}{\sqrt{|\lambda_i|}} \cdot v_i A \tilde{v}_j^{tr} & , \text{ für } j = 1, \ldots, k+l \\ v_i A \tilde{v}_j^{tr} & , \text{ für } j > k+l \end{cases} = \begin{cases} \frac{\lambda_j}{\sqrt{|\lambda_i|}} \cdot < v_i, \tilde{v}_j^{tr} > & , \text{ für } j = 1, \ldots, k+l \\ \lambda_j \cdot < v_i, \tilde{v}_j^{tr} > & , \text{ für } j > k+l \end{cases}$$

$$= \begin{cases} \frac{\lambda_i}{|\lambda_i|} \cdot \delta_{i,j} & , \text{ für } j = 1, \ldots, k+l \\ 0 & , \text{ für } j > k+l \end{cases} = \begin{cases} 1 \cdot \delta_{i,j} & , \text{ für } j = 1, \ldots, k \\ -1 \cdot \delta_{i,j} & , \text{ für } j = k+1, \ldots, k+l \\ 0 & , \text{ für } j = k+l+1, \ldots, n \end{cases}$$

Somit besitzt $G_B(\Phi)$ die angegebene Gestalt.

(b) Folgt aus (a) und der Transformationsformel für Gramsche Matrizen. qed.

27 Spray more - get more! Der Hauptachseneffekt (Die Hauptachsentransformation)

Im Folgenden sei $\Phi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ eine symmetrische Bilinearform. Das Standardskalarprodukt sei $< \cdot, \cdot >: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$.

Fragen:

- Wie kann man die Gramsche Matrix von Φ auf möglichst einfache Gestalt transformieren?
- Wie kann man dies geometrisch interpretieren?
- Wie kann man algorithmisch entscheiden, ob Φ positiv definit ist?

A. Wo verlaufen die Hauptachsen? Wo laufen sie denn hin?

27.1 Theorem: (Die Hauptachsentransformation für symmetrische Bilinearformen)

Sei $\Phi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ eine symmetrische Bilinearform.

(a) Dann gibt es eine ONB $B = \{v_1, \ldots, v_n\}$ von $\mathbb{R}^n$ bezüglich $< \cdot, \cdot >$, so dass gilt: $\Phi(v_i, v_j) = 0$ falls $i \neq j$. Mit anderen Worten, $G_B(\Phi)$ ist eine Diagonalmatrix.

(b) Es gibt eine OGB $C = \{w_1, \ldots, w_n\}$ von $\mathbb{R}^n$, so dass $G_C(\Phi)$ die folgende Gestalt besitzt:

$$G_C(\Phi) = \begin{pmatrix} I_k & 0 & 0 \\ 0 & -I_l & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ mit } k, l \geq 0$$

Beweis:

(a) Vgl. Satz 26.7

(b) Vgl. Satz 26.12 qed.

27.2 Beispiel:

Sei $\Phi : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ eine symmetrische Bilinearform und $A = G_E(\Phi) = \begin{pmatrix} a & b \\ b & a \end{pmatrix}$ mit $a, b \in \mathbb{R}$

1.) Wir berechnen $\chi_A(x) = x^2 - 2ax + (a^2 - b^2)$ und erhalten $\lambda_1 = a + b, \lambda_2 = a - b$ als Eigenwerte von A.

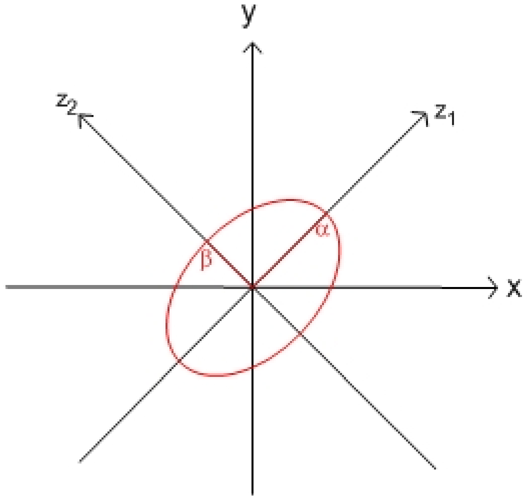
2.) Wir berechnen $Eig(A, a+b) = \mathbb{R} \cdot (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}) =: \mathbb{R} \cdot v_1$ und $Eig(A, a-b) = \mathbb{R} \cdot (-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}) =: \mathbb{R} \cdot v_2$. Dabei ist $B = \{v_1, v_2\}$ eine ONB von $\mathbb{R}^2$.

3.) In der Basis B gilt $G_B(\Phi) = \begin{pmatrix} a+b & 0 \\ 0 & a-b \end{pmatrix}$

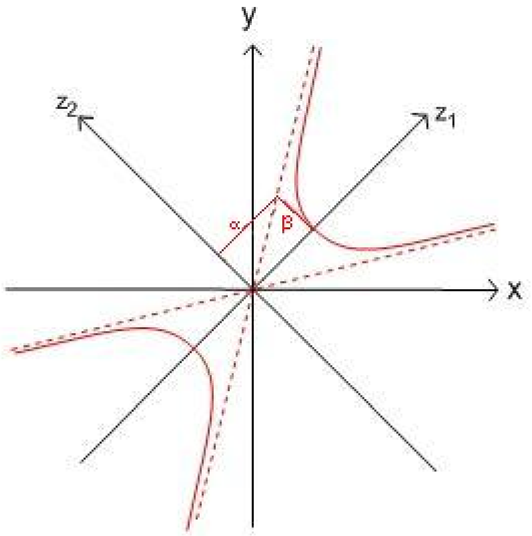
Jetzt betrachten wir die quadratische Form $q : \mathbb{R}^2 \to \mathbb{R}$, $x, y \mapsto (x,y) \cdot A \cdot (x,y)^{tr}$, also $q(x,y) = ax^2 + 2bxy + ay^2$.

Sei $C \subseteq \mathbb{R}^2$ die Kurve $C = \{(x,y) \in \mathbb{R}^2 \mid q(x,y) = 1\}$, die sogenannte Nullstellenmenge von $q-1$.

Skizze: 1. Fall: $\lambda_1, \lambda_2 > 0$



2. Fall: $\lambda_1 > 0, \lambda_2 < 0$



Was haben unsere Ausgangsdaten a, b (bzw. $\lambda_1 = a+b, \lambda_2 = a-b$) mit den Halbachsenlängen (=halbe Länge der Hauptachsen) α, β zu tun?

Seien z_1, z_2 die Koordinaten eines Punktes (x,y) bezüglich der ONB $B = \{v_1, v_2\}$. Dann gilt:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \quad \text{„Drehung um 45"}$$

Die Gleichung der quadratischen Form q in den neuen Koordinaten z_1, z_2 lautet:
$q(z_1, z_2) = (a+b)z_1^2 + (a-b)z_2^2$. O.E. gelte $\lambda_1 = a+b > 0$.

Setze $\alpha = \frac{1}{\sqrt{\lambda_1}} = \frac{1}{\sqrt{a+b}}$ und $\beta = \frac{1}{\sqrt{|\lambda_2|}} = \frac{1}{\sqrt{|a-b|}}$

Dann gilt: $q(z_1, z_2) = \frac{z_1^2}{\alpha^2} \pm \frac{z_2^2}{\beta^2}$ mit $\begin{cases} + & \text{, falls } \lambda_2 = a-b > 0 \\ - & \text{, falls } \lambda_2 = a-b < 0 \end{cases}$

Die Gleichung $\frac{z_1^2}{\alpha^2} \pm \frac{z_2^2}{\beta^2} = 1$ ist die Gleichung einer Ellipse bzw. Hyperbel mit den Halbachsenlängen α, β (siehe §28).

B. Und wann definiert Φ ein Skalarprodukt?

27.3 Satz:

Eine symmetrische Bilinearform $\Phi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ ist genau dann ein Skalarprodukt, wenn die Eigenwerte von $G_E(\Phi)$ alle positiv sind.
($\lightning$: Bei der Transformation $T^{tr} \cdot G_E(\Phi) \cdot T$ können sich die Eigenwerte ändern!)

Beweis: Sei $B = \{v_1, \ldots, v_n\}$ eine ONB von $\mathbb{R}^n$ bezüglich $< \cdot, \cdot >$, so dass
$G_B(\Phi) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ eine Diagonalmatrix ist. Die Zahlen $\lambda_1, \ldots, \lambda_n$ sind die (nicht notwendig verschiedenen) Eigenwerte von $G_B(\Phi)$. Wegen $G_E(\Phi) = T^{-1} G_B(\Phi) T$ mit $T \in O_n(\mathbb{R})$ sind $\lambda_1, \ldots, \lambda_n$ auch die Eigenwerte von $G_E(\Phi)$.

Für einen Vektor $v = a_1 v_1 + \ldots + a_n v_n$ mit $a_i \in \mathbb{R}$ gilt:

$$\Phi(v, v) = \Phi(a_1 v_1 + \ldots + a_n v_n, a_1 v_1 + \ldots + a_n v_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j \Phi(v_i, v_j)$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j \lambda_i \cdot \delta_{ij} = \sum_{i=1}^{n} a_i^2 \lambda_i$$

Diese Summe ist genau dann für alle $(a_1, \ldots, a_n) \in \mathbb{R}^n \setminus \{0\}$ positiv, wenn $\lambda_1 > 0, \ldots, \lambda_n > 0$ gilt. qed.

27.4 Definition:

Sei V ein $\mathbb{R}$-Vektorraum und $\Phi : V \times V \to \mathbb{R}$ eine Bilinearform. Dann heißt

$$V_0 = \{v \in V \mid \Phi(v, w) = 0 \text{ für alle } w \in V\}$$

der Ausartungsraum von Φ.
Offenbar ist V_0 ein $\mathbb{R}$-Untervektorraum von V.

27.5 Satz:

Sei $\Phi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ eine symmetrische Bilinearform. Sei W_0 der Ausartungsraum von Φ. Dann gibt es $\mathbb{R}$-Untervektorräume W_+, W_- von $\mathbb{R}^n$ mit folgenden Eigenschaften:

(a) $\mathbb{R}^n = W_0 \oplus W_+ \oplus W_-$ ist eine orthogonale direkte Summe bezüglich $< \cdot, \cdot >$ und bezüglich Φ.

(b) Für $w \in W_+ \setminus \{0\}$ gilt $\Phi(w, w) > 0$ und für $w \in W_- \setminus \{0\}$ gilt $\Phi(w, w) < 0$.

Beweis: Sei $B = \{w_1, \ldots, w_n\}$ eine ONB bezüglich $< \cdot, \cdot >$, die aus Eigenvektoren von $G_E(\Phi)$ besteht. Schreibe $\Phi(w_i, w_j) = \lambda_i \delta_{ij}$ für $i = 1, \ldots, n$ mit $\lambda_i \in \mathbb{R}$.
O.E. gelte dabei $\lambda_1, \ldots, \lambda_k > 0$ und $\lambda_{k+1}, \ldots, \lambda_{k+l} < 0$ und $\lambda_{k+l+1} = \ldots = \lambda_n = 0$.
Setze $W_+ =< w_1, \ldots, w_k >$ und $W_- =< w_{k+1}, \ldots, w_{k+l} >$. Mit diesen beiden Vektorräumen weisen wir nun (a) und (b) nach:

(a) Bezüglich $< \cdot, \cdot >$ gilt: $\mathbb{R}^n = W_0 \bigoplus W_+ \bigoplus W_-$ nach Konstruktion.
Bezüglich Φ folgt die Orthogonalität der direkten Summe aus $Eig(f_A, \lambda_i) \perp_\Phi Eig(f_A, \lambda_j)$ falls $i \neq j$.

(b) Zuerst zeigen wir $W_0 =< w_{k+l+1}, \ldots, w_n >$. Dabei ist „$\supseteq$" klar.
„$\subseteq$" Sei $v = a_1 w_1 + \ldots + a_n w_n \in V^\perp$ mit $a_i \in \mathbb{R}$. Für $i = 1, \ldots, k + l$ gilt
$\Phi(v, w_i) = a_1 \Phi(w_1, w_i) + \ldots + a_n \Phi(w_n, w_i) = a_i \Phi(w_i, w_i) = a_i \lambda_i = 0 \Rightarrow a_i = 0$.
$\Rightarrow v \in < w_{k+l+1}, \ldots, w_n > = W_0$
Nun zeigen wir: $\Phi(w, w) > 0$ für $w \in W_+ \setminus \{0\}$:
Sei $w = a_1 w_1 + \ldots + a_k w_k$ mit $a_i \in \mathbb{R}$.
$\Rightarrow \Phi(w, w) = \sum_{i=1}^{k} \sum_{j=1}^{k} a_i a_j \Phi(w_i, w_j) = a_1^2 \lambda_1 + \ldots + a_k^2 \lambda_k > 0$
Analog folgt $\Phi(w, w) < 0$ für $w \in W_- \setminus \{0\}$. qed.

27.6 Bemerkung:

Sei V ein endlich-dimensionaler, reeller Vektorraum und $\Phi : V \times V \to \mathbb{R}$ eine symmetrische Bilinearform.

(a) Die Mengen $C_0 = \{v \in V \mid \Phi(v, v) = 0\}$ und $C_+ = \{v \in V \mid \Phi(v, v) > 0\} \cup \{0\}$ und $C_- = \{v \in V \mid \Phi(v, v) < 0\} \cup \{0\}$ sind im Allgemeinen keine Untervektorräume von V.

(b) Es gibt im Allgemeinen sehr viele Untervektorräume $V_+ \subseteq C_+$ und $V_- \subseteq C_-$ mit $V = V_0 \oplus V_+ \oplus V_-$

27.7 Theorem: (Der Trägheitssatz von Sylvester)

Sei V eine endlich-dimensionaler, reeller Vektorraum und $\Phi : V \times V \to \mathbb{R}$ eine symmetrische Bilinearform.
Sei V_0 der Ausartungsraum von Φ und seien $V_+, V_- \subseteq V$ $\mathbb{R}$-Untervektorräume mit $V_+ \subseteq \{v \in V \mid \Phi(v, v) > 0\} \cup \{0\}$ und $V_- \subseteq \{v \in V \mid \Phi(v, v) < 0\} \cup \{0\}$ und $V = V_0 \bigoplus V_+ \bigoplus V_-$.
Dann sind die Zahlen $r_+ = dim_\mathbb{R}(V_+)$ und $r_- = dim_\mathbb{R}(V_-)$ unabhängig von der Wahl der Zerlegung, d.h. ist $V = V_0 \bigoplus \tilde{V}_+ \bigoplus \tilde{V}_-$, so folgt $r_+ = dim_\mathbb{R}(\tilde{V}_+)$ und $r_- = dim_\mathbb{R}(\tilde{V}_-)$.

- Die Zahl r_+ heißt der Trägheitsindex (oder Positivitätsindex) von Φ.
- Die Zahl r_- heißt der Negativitätsindex von Φ.
- Die Zahl $r_+ - r_-$ heißt die Signatur von Φ.
- Die Zahl $r_0 = dim_\mathbb{R}(V_0)$ heißt der Nullindex von Φ.
- Die Zahl $r_+ + r_-$ heißt der Rang von Φ.

Beweis:
Zwischenbehauptung: Ist $W \subseteq V$ ein $\mathbb{R}$-Untervektorraum mit $W \subseteq \{v \in V \mid \Phi(v,v) > 0\} \cup \{0\}$, so gilt $W \cap (V_- \oplus V_0) = \{0\}$.

Beweis: Sei $v \in W \setminus \{0\}$. Gilt $v \in V_- \oplus V_0$, so schreibe $v = v_- + v_0$ mit $v_- \in V_-$ und $v_0 \in V_0$. Dann gilt: $0 < \Phi(v,v) = \Phi(v_- + v_0, v_- + v_0) = \Phi(v_-, v_-) + \Phi(v_0, v_0) < 0 \lightning \Rightarrow$ Behauptung

Nun wenden wir die Dimensionsformel an und erhalten:

$$dim_{\mathbb{R}}(W) + dim_{\mathbb{R}}(V_-) + dim_{\mathbb{R}}(V_0) \leq n$$

Sei nun $V = V_0 \oplus \tilde{V}_+ \oplus \tilde{V}_-$ eine weitere Zerlegung der gewünschten Art. Setzt man $W = \tilde{V}_+$, so folgt $dim_{\mathbb{R}}(\tilde{V}_+) \leq n - dim_{\mathbb{R}}(V_-) - dim_{\mathbb{R}}(V_0) = dim_{\mathbb{R}}(V_+) = r_+$
Vertauscht man die Rollen von V_+, V_- mit $\tilde{V}_+, \tilde{V}_-$, so folgt analog $dim_{\mathbb{R}}(V_+) \leq dim_{\mathbb{R}}(\tilde{V}_+)$, also insgesamt $dim_{\mathbb{R}}(\tilde{V}_+) = r_+$.
Wegen $r_+ + r_- = n - dim_{\mathbb{R}}(V_0) = dim_{\mathbb{R}}(\tilde{V}_+) + dim_{\mathbb{R}}(\tilde{V}_-) = r_+ + dim_{\mathbb{R}}(\tilde{V}_-)$ folgt:
$dim_{\mathbb{R}}(\tilde{V}_-) = r_-$. qed.

27.8 Satz: (Das Hauptminorenkriterium für positive Definitheit)

Sei $\Phi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ eine Bilinearform und $A = G_E(\Phi) = (a_{ij})$ ihre Gramsche Matrix. Dann sind die folgenden Bedingungen äquivalent:

(a) Die Abbildung Φ ist ein Skalarprodukt.

(b) Die Matrix A ist symmetrisch und für $k = 1, \ldots, n$ gilt: $\underbrace{det(a_{ij})_{i,j=1,\ldots,k}}_{\text{Hauptminoren}} > 0$

Beweis: „(a)⇒(b)" Wähle $S \in O_n(\mathbb{R})$ so, dass $S^{tr}AS = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ gilt mit $\lambda_i \in \mathbb{R}$. Nach Satz 27.3 gilt $\lambda_i > 0$ für $i = 1, \ldots, n$.
Insbesondere folgt $det(A) = \lambda_1 \cdots \lambda_n > 0$. Die Matrix $A_k = (a_{ij})_{i,j=1,\ldots,k}$ ist die Gramsche Matrix der Einschränkung von Φ auf
den $\mathbb{R}$-Untervektorraum $\{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid x_{k+1} = \ldots = x_n = 0\} \cong \mathbb{R}^k$. Also folgt analog $det(A_k) > 0$.

„(b)⇒(a)" Wir schließen mit vollständiger Induktion nach n:

$n = 1$: Klar.

$n > 1$: Wähle $\tilde{T} \in GL_{n-1}(\mathbb{R})$ mit $\tilde{T}^{tr} A_{n-1} \tilde{T} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_{n-1} \end{pmatrix}$ mit $\lambda_i > 0$ für $i = 1, \ldots, n-1$.

Setze $T = \left(\begin{array}{ccc|c} & & & 0 \\ & \tilde{T} & & \vdots \\ & & & 0 \\ \hline 0 & \cdots & 0 & 1 \end{array}\right) \in GL_n(\mathbb{R})$. Dann folgt: $T^{tr}AT = \left(\begin{array}{ccc|c} \lambda_1 & & 0 & \beta_1 \\ & \ddots & & \vdots \\ 0 & & \lambda_{n-1} & \beta_{n-1} \\ \hline \beta_1 & \cdots & \beta_{n-1} & \beta_n \end{array}\right) =: B$

Nach Voraussetzung ist $det(B) = det(A) \cdot det(T)^2 > 0$.

Sei nun $S = \left(\begin{array}{ccc|c} & & & \gamma_1 \\ & I_{n-1} & & \vdots \\ & & & \gamma_{n-1} \\ \hline 0 & \cdots & 0 & 1 \end{array}\right)$ mit $\gamma_i = -\dfrac{\beta_i}{\lambda_i}$.

Dann folgt $S^{tr}BS = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ mit $det(B) \cdot det(S)^2 = \lambda_1 \cdots \lambda_n > 0$, also $\lambda_n > 0$

Nach Satz 27.3 folgt, dass Φ positiv definit ist. qed.

28 Störe meine Kreise nicht! (Kegelschnitte)

A. Ellipsen - Die Theorie vom Einheitsei

Im Folgenden betrachten wir die Zeichenebene und identifizieren sie mit $\mathbb{R}^2$

28.1 Definition:

(a) Seien $x_0, y_0 \in \mathbb{R}$ und $r \in \mathbb{R}$. Die Menge der Punkte

$$K = \{(x, y) \in \mathbb{R}^2 \mid (x - x_0)^2 + (y - y_0)^2 = r^2\}$$

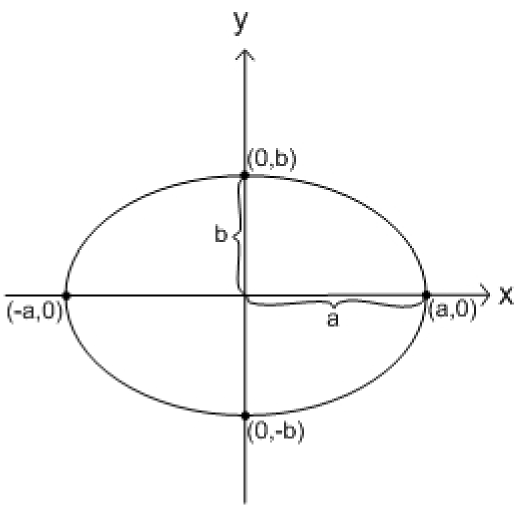
heißt der Kreis mit Mittelpunkt (x_0, y_0) und Radius r.

(b) Seien $a, b \in \mathbb{R}$ mit $0 < b \leq a$. Die Menge der Punkte

$$E = \{(x, y) \in \mathbb{R}^2 \mid \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1\}$$

heißt die Ellipse mit den Halbachsen a, b und den Scheitelpunkten $(a, 0), (-a, 0), (0, b), (0, -b)$.

Skizze:



28.2 Bemerkung:

(a) Der Kreis K ist im Fall $(x_0, y_0) = (0, 0)$ eine Ellipse mit den Halbachsen r und r.

(b) Man kann allgemeinere Ellipsen definieren durch

$$\hat{E} = \{(x, y) \in \mathbb{R}^2 \mid \frac{(c_1 x + c_2 y + d_1)^2}{a^2} + \frac{(c_3 x + c_4 y + d_2)^2}{b^2} = 1\} \text{ mit } \begin{pmatrix} c_1 & c_2 \\ c_3 & c_4 \end{pmatrix} \in SO_2(\mathbb{R})$$

und $d_1, d_2 \in \mathbb{R}$

Ellipsen mit solchen allgemeinen Gleichungen werden später untersucht.

28.3 Definition:

Seien $a, b \in \mathbb{R}$ mit $0 < b \leq a$.

(a) Sei $c = \sqrt{a^2 - b^2} \geq 0$. Dann heißen die Punkte $F = (c, 0)$ und $F' = (-c, 0)$ die Brennpunkte von E.

(b) Die Zahl $e = \frac{c}{a}$ erfüllt $0 \leq e < 1$. Sie heißt die Exzentrizität der Ellipse E. Es gilt $e = 0$ genau dann, wenn E ein Kreis ist.

28.4 Satz: (Die Abstandseigenschaft der Ellipse)

Für jeden Punkt $P \in E$ gilt: $\|P - F\| + \|P - F'\| = 2a$.

Beweis: Sei $P = (x, y) \in E$. Aus $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$ folgt $x^2b^2 + y^2a^2 - a^2b^2 = 0$.

$\Rightarrow [(x-c)^2 + y^2] \cdot [(x+c)^2 + y^2] = [x^2 + y^2 + c^2 + 2xc] = (x^2 + y^2 + a^2 - b^2)^2 - 4x^2(a^2 - b^2)$

$= x^4 + y^4 + a^4 + b^4 + 2x^2y^2 + 2x^2a^2 - 2x^2b^2 + 2y^2a^2 - 2y^2b^2 - 2a^2b^2 - 4x^2a^2 + 4x^2b^2$

$= x^4 + y^4 + a^4 + b^4 + 2x^2y^2 - 2x^2a^2 + 2x^2b^2 + 2y^2a^2 - 2y^2b^2 - 2a^2b^2$

$= x^4 + y^4 + a^4 + b^4 + 2x^2y^2 - 2x^2a^2 - 2x^2b^2 - 2y^2a^2 - 2y^2b^2 - 2a^2b^2$

$= (x^2 + y^2 - a^2 - b^2)^2$

$\Rightarrow \sqrt{(x-c)^2 + y^2} \cdot \sqrt{(x+c)^2 + y^2} = a^2 + b^2 - x^2 - y^2$, da $x^2 + y^2 \leq a^2 + b^2$

$\Rightarrow (\sqrt{(x-c)^2 + y^2} \cdot \sqrt{(x+c)^2 + y^2})^2 = (x-c)^2 + y^2 + (x+c)^2 + y^2 + 2\sqrt{(x-c)^2 + y^2} \cdot \sqrt{(x+c)^2 + y^2}$

$= 2x^2 + 2y^2 + 2a^2 - 2b^2 + 2a^2 + 2b^2 - 2x^2 - 2y^2 = 4a^2$

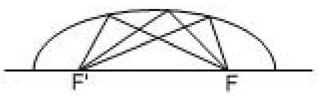
$\Rightarrow \underbrace{\sqrt{(x-c)^2 + y^2}}_{\|P - F\|} + \underbrace{\sqrt{(x+c)^2 + y^2}}_{\|P - F'\|} = 2a$

qed.

28.5 Korollar: (Die Fadenkonstruktion der Ellipse)

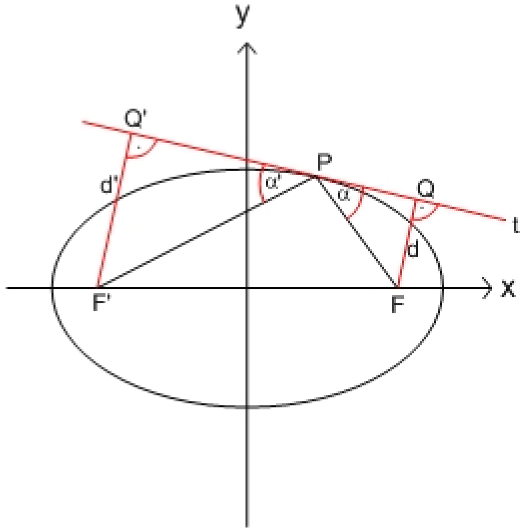
Spannt man einen Faden der Länge $2a$ zwischen den Brennpunkten $F = (c, 0)$ und $F'(-c, 0)$ mit $c = \sqrt{a^2 - b^2}$, so ergeben die Spannpunkte gerade die Ellipse E.

Skizze:



28.6 Satz: (Die Reflexionseigenschaft der Ellipse)

Verlässt ein Lichtstrahl einen Brennpunkt der Ellipse und wird er an der Ellipse reflektiert, so geht der reflektierte Strahl durch den anderen Brennpunkt der Ellipse.
Skizze:



Beweisskizze: Gegeben sei ein Punkt $P = (x_0, y_0) \in E$.
Die Tangente t durch P an E hat die Gleichung $\frac{x_0}{a^2}x + \frac{y_0}{b^2}y - 1 = 0$.
Jetzt fällen wir die Lote von F und F' auf t und erhalten $d = \frac{b}{b'}(a - ex_0)$ und $d' = \frac{b}{b'}(a + ex_0)$ mit $b' = \sqrt{\frac{x_0^2 b^2}{a^2} + \frac{y_0^2 a^2}{b^2}}$ und $e = \frac{\sqrt{a^2 - b^2}}{a}$
Ferner gilt: $\|F - P\| = a - ex_0$ und $\|F' - P\| = a + ex_0$. Also sind die Dreiecke $\triangle FPQ$ und $\triangle F'PQ'$ ähnlich. Es folgt $\alpha = \alpha'$. qed.

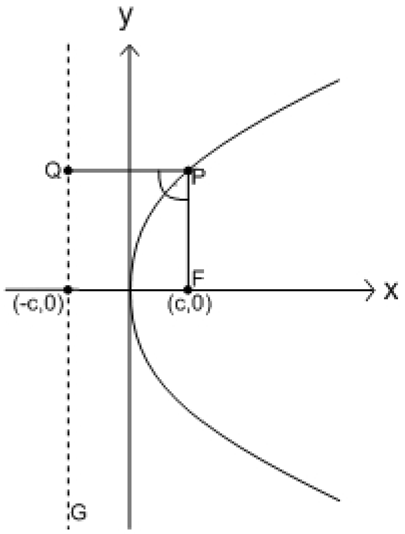
B. Parabeln - Ganz ohne Lessing

28.7 Definition:

Sei $c > 0$. Die Menge der Punkte

$$C = \{(x, y) \in \mathbb{R}^2 \mid y^2 = 4cx\}$$

heißt eine Parabel mit Brennpunkt $F = (c, 0)$ und Scheitelpunkt $O = (0, 0)$.
Skizze:



28.8 Bemerkung:

(a) Ändert man das Koordinatensystem und verschiebt man das Bild von C, so erhält man allgemeinere Parabeln (z.B. $y = x^2$, vgl. später).

(b) Die Parabel ist der Grenzfall einer Menge von Ellipsen mit $e \to 1$. Dabei wird ein Brennpunkt der Ellipse festgehalten und der andere nach ∞ bewegt.

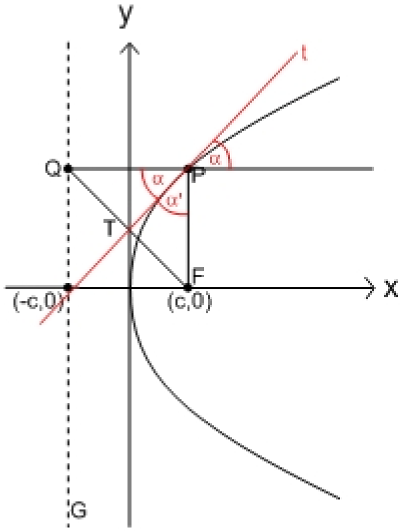
28.9 Satz: (Die Abstandseigenschaft der Parabel)

Die Parabel C besteht aus allen Punkten P, die von der Geraden $G = (-c, 0) + \mathbb{R}(0, 1)$ und von $F = (c, 0)$ gleich weit entfernt sind.

Beweis: Sei $P = (x_0, y_0) \in C$. Dann gilt:
$\|P - F\| = \sqrt{(x_0 - c)^2 + y_0^2} = \sqrt{x_0^2 - 2x_0c + c^2 + 4x_0c} = x_0 + c = \|P - Q\|$,
wobei Q der Fußpunkt des Lots von P auf G ist. qed.

28.10 Satz: (Die Reflexionseigenschaft der Parabel)

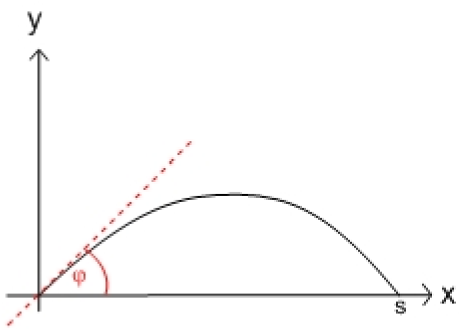
Jeder Lichtstrahl, der auf einer Parallelen zur x-Achse aus dem Unendlichen kommt und an der Parabel C reflektiert wird, verläuft anschließend durch den Brennpunkt $F = (c, 0)$.
Skizze:



Beweis: Der Lichtstrahl sei $L = (0, y_0) + \mathbb{R}(1, 0)$. Sein Schnittpunkt mit C sei $P = (x_0, y_0)$. O.E. gelte $y_0 > 0$. Dann ist C in der Nähe von P gegeben durch
$y = \sqrt{4cx} = 2\sqrt{cx} \Rightarrow y' = 2\sqrt{x} \cdot \dfrac{1}{\sqrt{c}} = \sqrt{\dfrac{c}{x}}$
Die Gleichung der Tangente an C durch P lautet damit $t(x) = y_0 + \sqrt{\dfrac{c}{x_0}}(x - x_0)$
Der Schnittpunkt mit der y-Achse ist:
$T = (0, y_0 - \sqrt{\dfrac{c}{x_0}}x_0) = (0, y_0 - \sqrt{cx_0}) = (0, \frac{1}{2}y_0)$ (=Mittelpunkt von F und Q)
Nach Satz 28.9 gilt $\|F - P\| = \|F - Q\|$. Somit ist das Dreieck $\triangle FPQ$ gleichschenklig.
Wegen $T = (0, \frac{1}{2}y_0) = \frac{1}{2}(F + Q) = \frac{1}{2}((c, 0) + (-c, y_0))$ ist T der Mittelpunkt der Seite FQ dieses Dreiecks. Somit folgt $\alpha = \alpha'$. qed.

28.11 Bemerkung:

Vom Ursprung $O = (0, 0)$ aus werde im Winkel φ zum Boden (x-Achse) eine Kanonenkugel abgeschossen. Unter Vernachlässigung des Luftwiderstandes ist der Winkel φ so zu wählen, dass die erzielte Schussweite s maximal wird.



Sei $\overrightarrow{v} = (v_x, v_y) = (v \cdot cos(\varphi), v \cdot sin(\varphi))$ mit $v = \|\overrightarrow{v}\|$ der Geschwindigkeitsvektor der Kugel beim Abschießen. Sei t die Zeit.

Es gilt: $y''(t) = -g$ mit g =Gravitationskonstante$\approx 9,81\frac{m}{s^2} \Rightarrow y'(t) = -gt + v_y$
$\Rightarrow y(t) = -\frac{1}{2}gt^2 + v_yt$ und $x(t) = v_xt$.

Dies zeigt: $y(t) = -\frac{1}{2}g\frac{1}{v_x^2}x(t) + \frac{v_y}{v_y}x(t)$ für alle $t \geq 0$.

Die Flugbahn liegt also auf der Parabel $C = \{(x,y) \in \mathbb{R}^2 \mid y^2 + \frac{g}{2v_x^2}x^2 - \frac{v_y}{v_y}x = 0\}$
Für den Auftreffpunkt $(s,0) = (x(t_0),0)$ gilt $y(t_0) = 0 \Rightarrow \frac{1}{2}gt_0^2 = v_y t_0 \underset{t_0 \neq 0}{\Rightarrow} t_0 = \frac{2v_y}{y}$

und $s = x(t_0) = \frac{2}{g}v_x v_y = \frac{2}{g}v^2 \cdot sin(\varphi)cos(\varphi) = \frac{v^2}{g}sin(2\varphi)$.
Dies ist maximal für $2\varphi = \frac{\pi}{2}$, also für $\varphi = \frac{\pi}{4} = 45$

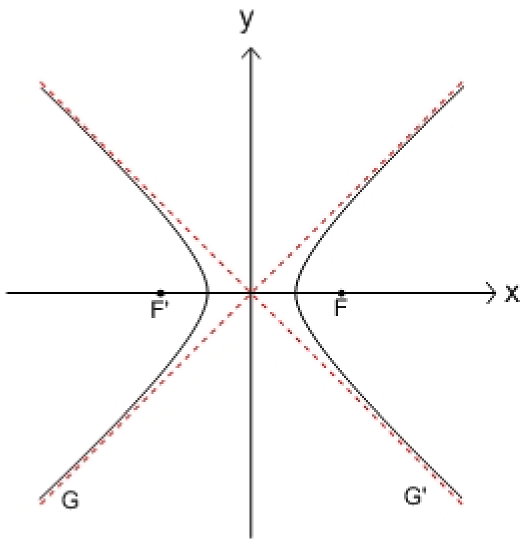
C. Hyperbeln - Besser als jeder Boomerang

28.12 Definition:

Seien $a, b \in \mathbb{R}_+$. Ferner sei $c = \sqrt{a^2+b^2}$. Dann heißt die Menge der Punkte

$$H = \{(x,y) \in \mathbb{R}^2 \mid \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1\}$$

eine Hyperbel mit den Brennpunkten $F = (c,0)$ und $F' = (-c,0)$.
Skizze:



Durch Koordinatentransformation und Verschiebungen erhält man weitere Hyperbeln (vgl. später).

28.13 Satz: (Die Abstandseigenschaft der Hyperbel)

Für jeden Punkt $P \in H$ gilt: $|\|P - F\| - \|P - F'\|| = 2a$.

Beweis: Sei $P = (x_0, y_0)$.

Aus $\dfrac{x_0^2}{a^2} - \dfrac{y_0}{b^2}$ folgt: $b^2x_0^2 - a^2y_0^2 = a^2b^2$, also:

$[(x_0-c)^2 + y_0^2)][(x_0+c)^2 + y_0^2] = ((x_0-c)^2(x_0+c))^2 + y_0^2(x_0-c)^2 + y_0^2(x_0+c)^2 + y_0^4$
$= (x_0^2 - a^2 - b^2) + y_0^2x_0^2 - y_0^2 2x_0c + y_0^2c^2 + y_0^2x_0^2 + 2y_0^2x_0c + y_0^2c^2 + y_0^4$
$= (x_0^2 - a^2 - b^2) + 2y_0^2x_0^2 + 2y_0^2c^2 + y_0^4$
$= x_0^4 + a^4 + b^4 - 2a^2x_0^2 - 2b^2x_0^2 + 2a^2b^2 + 2y_0^2x_0^2 + 2y_0^2a^2 + 2y_0^2b^2 + y_0^4$
$= x_0^4 + y_0^4 + a^4 + b^4 + 2x_0^2y_0^2 - 2a^2x_0^2 + 2b^2x_0^2 - 2a^2y_0^2 + 2b^2y_0^2 - 2a^2b^2$
$= (x_0^2 + y_0^2 - a^2 + b^2)^2$
Denn: $2b^2x_0^2 - 2a^2y_0^2 - 2a^2b^2 = 2(b^2x_0^2 - a^2y_0^2 - a^2b^2) = 0$

Hieraus folgt: $\sqrt{[(x_0-c)^2+y_0^2)][(x_0+c)^2+y_0^2]} = x_0^2+y_0^2-a^2+b^2$

$\Rightarrow 2x_0^2+2y_0^2+2c^2-2\sqrt{[(x_0-c)^2+y_0^2)][(x_0+c)^2+y_0^2]} = 4a^2$

$\Rightarrow (\sqrt{[(x_0-c)^2+y_0^2)]}-\sqrt{[(x_0+c)^2+y_0^2]})^2$

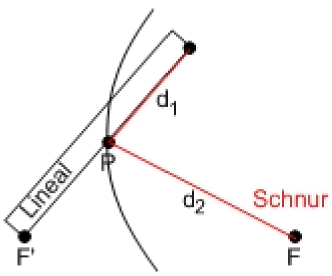
$= x_0^2-2x_0c+c^2+y_0^2-2\sqrt{[(x_0-c)^2+y_0^2)]}\sqrt{[(x_0+c)^2+y_0^2]}+x_0^2+2x_0c+c^2+y_0^2$

$= \left|\sqrt{[(x_0-c)^2+y_0^2)]}-\sqrt{[(x_0+c)^2+y_0^2]}\right| = 2a$

$\Rightarrow |\|P-F\|-\|P-F'\|| = 2a$ qed.

28.14 Korollar: (Die Fadenkonstruktion der Hyperbel)

Ein Linear der Länge $l >> 2a$ wird mit einem Kantenende am Punkt $F' = (-c, 0)$ befestigt. Am anderen Ende des Lineals wird eine Schnur der Länge $l - 2a$ festgemacht. Das zweite Ende der Schnur wird am Punkt $F = (c, 0)$ fixiert.
Die Spannpunkte der Schnur, die auf der Kante des Lineals liegen, bilden einen Ast der Hyperbel H. Vertauscht man die Rollen von F und F' und wiederholt die Prozedur, so erhält man den anderen Ast von H.
Skizze:



Beweis: Es gilt: $d_1 + d_2 = l - 2a$ und

$\|P-F'\| = l - d_1 = d_1 + d_2 + 2a - d_1 = 2a + d_2$ und $\|P-F\| = d_2$

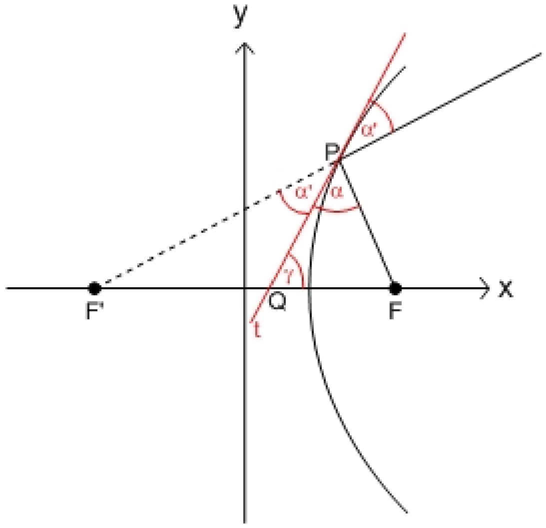
Also folgt: $\|P-F\| - \|P-F'\| = 2a + d_2 - d_2 = 2a$ qed.

28.15 Bemerkung:

Für $x \to \infty$ und $x \to -\infty$ nähert sich H an die Geraden $G : y = \frac{b}{a}x$ und $G' : y = -\frac{b}{a}x$ an, denn $y = \pm\frac{b}{a}\sqrt{x^2-a^2} \approx \pm\frac{b}{a}x$.
Diese beiden Geraden heißen die Asymptoten von H.

28.16 Satz: (Die Reflexionseigenschaft der Hyperbel)

Wird ein von einem Brennpunkt F der Hyperbel H ausgehender Lichtstrahl an der Hyperbel reflektiert, so scheint der reflektierte Strahl aus dem anderen Brennpunkt F' zu kommen.
Skizze:



Beweis: Sei $P = (x_0, y_0)$ der Punkt auf H, an dem der Lichtstrahl reflektiert wird.
O.E. gelte $y_0 > 0$.
In der Nähe von P besitzt H die Gleichung:

$$y = \frac{b}{a}\sqrt{x^2 - a^2} \Rightarrow y'(x_0) = \frac{b}{a}\frac{1}{2\sqrt{x_0^2 - a^2}}2x_0 = \frac{b}{a}\frac{1}{\sqrt{x_0^2 - a^2}}x_0 \cdot \frac{a}{b} \cdot \frac{b}{a} = \frac{b^2 x_0}{a^2 y_0}$$

Daher lautet die Gleichung der Tangente t an H in P:

$$y = \frac{b^2 x_0}{a^2 y_0}(x - x_0) + y_0$$

Der Durchstroßpunkt $Q = (q, 0)$ von t auf der x-Achse efüllt:

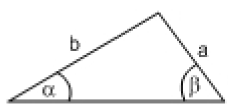
$$\frac{b^2 x_0}{a^2 y_0}q = \frac{b^2 x_0^2}{a^2 y_0} - \frac{b^2 y_0^2}{b^2 y_0} = \frac{b^2}{y_0}\left[\frac{x_0^2}{a^2} - \frac{y_0^2}{b^2}\right] = \frac{b^2}{y_0} \text{ und somit } q = \frac{a^2}{x_0}$$

Nun wenden wir den Sinussatz in den Dreiecken $\triangle F'QP$ und $\triangle FPQ$ an und erhalten:

$$\frac{q + c}{sin(\alpha')} = \frac{\sqrt{(x_0 + c) + y_0}}{sin(\pi - \gamma)} = \frac{\sqrt{(x_0 + c) + y_0}}{sin(\gamma)} \text{ sowie } \frac{c - q}{sin(\alpha)} = \frac{\sqrt{(x_0 - c) + y_0}}{sin(\gamma)}$$

Hieraus kann man nachrechnen, dass $sin(\alpha) = sin(\alpha')$ gilt, also $\alpha = \alpha'$. qed.

[Sinussatz:



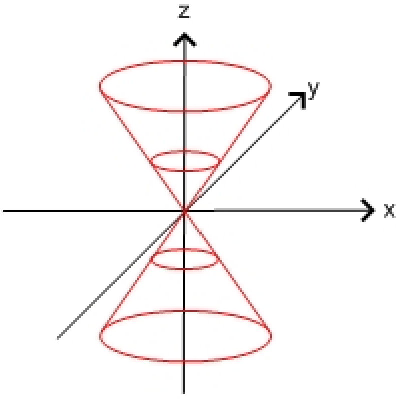
$$\frac{a}{sin(\alpha)} = \frac{b}{sin(\beta)}$$

]

D. Autsch! Ein Kreiskegel wird geschnitten!

28.17 Definition:

Die Menge der Punkte
$$K = \{(x,y,z) \in \mathbb{R}^3 \mid x^2 + y^2 = z^2\}$$
heißt der Kreiskegel mit Spitze $(0,0,0)$.
Skizze:



28.18 Bemerkung:

(a) Eine invertierbare Matrix $A \in GL_n(\mathbb{R})$ entspricht einer bijektiven, $\mathbb{R}$-linearen Abbildung $\varphi : \mathbb{R}^n \to \mathbb{R}^n$. Das Bild der Standardbasis $E = \{e_1, \ldots, e_n\}$ unter φ ist eine Basis $B = \{v_1, \ldots, v_n\}$ von $\mathbb{R}^n$.
Sei nun eine Punktmenge $M \subseteq \mathbb{R}^n$ definiert durch eine Gleichung:
$$M = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid f(x_1, \ldots, x_n) = 0\}$$
Dabei sind $x_1, \ldots, x_n$ die Koordinaten eines Punktes $P = (x_1, \ldots, x_n)$ bezüglich der Standardbasis E.
Welche Gleichungen erfüllen die Koordinaten der Punkte von M bezüglich der neuen Basis B? Es gilt:
$$P = (x_1, \ldots, x_n) = x_1 e_1 + \ldots + x_n e_n \text{ und } v_i = A \cdot e_i \Rightarrow e_i = A^{-1} \cdot v_i$$
$$\Rightarrow P = x_1 A^{-1} v_1 + \ldots + x_n A^{-1} v_n = A^{-1}(x_1 v_1 + \ldots + x_n v_n)$$
Also ist $A^{-1}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \tilde{x}_1 \\ \vdots \\ \tilde{x}_n \end{pmatrix}$ das Koordinatentupel von P bezüglich der neuen Basis B.

(b) Wir drehen jetzt die Basis $E = (e_1, e_2, e_3)$ vom $\mathbb{R}^3$ um einen Winkel $\varphi \in [0, 2\pi)$ um die y-Achse.
$$\Rightarrow A = \begin{pmatrix} cos(\varphi) & 0 & sin(\varphi) \\ 0 & 1 & 0 \\ -sin(\varphi) & 0 & cos(\varphi) \end{pmatrix} \text{ und } A^{-1} = \begin{pmatrix} cos(\varphi) & 0 & -sin(\varphi) \\ 0 & 1 & 0 \\ sin(\varphi) & 0 & cos(\varphi) \end{pmatrix} \Rightarrow \begin{pmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{z} \end{pmatrix} = A^{-1}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$
$$\text{also } \begin{cases} \tilde{x} = cos(\varphi)x - sin(\varphi)z \\ \tilde{y} = y \\ \tilde{z} = sin(\varphi)x + cos(\varphi)z \end{cases} \text{ und } \begin{cases} x = cos(\varphi)\tilde{x} + sin(\varphi)\tilde{z} \\ y = \tilde{y} \\ z = -sin(\varphi)\tilde{x} + cos(\varphi)\tilde{z} \end{cases}$$

In diesem neuen Koordinatensystem hat der Kreiskegel K also die Gleichung:

$$[cos(\varphi)\tilde{x} + sin(\varphi)\tilde{z}]^2 + \tilde{y}^2 = [-sin(\varphi)\tilde{x} + cos(\varphi)\tilde{z}]^2$$

$$\Leftrightarrow (cos^2(\varphi) - sin^2(\varphi))\tilde{x}^2 + 4sin(\varphi)cos(\varphi)\tilde{x}\tilde{z} + \tilde{y}^2 = (cos^2(\varphi) - sin^2(\varphi))\tilde{z}$$

28.19 Satz: (Die Ebenenschnitte des Kreiskegels)

Schneidet man den Kreiskegel mit der Ebene $E = \{(\tilde{x}, \tilde{y}, \tilde{z}) \in \mathbb{R}^3 \mid \tilde{z} = 1\}$, so ergeben sich in Abhängigkeit von φ die folgenden Schnittkurven:

(a) $0 \leq \varphi < \frac{\pi}{4}$: Ellipse

(b) $\varphi = \frac{\pi}{4}$: Parabel

(c) $\frac{\pi}{4} < \varphi \leq \frac{\pi}{2}$: Hyperbel

(Aus Symmetriegründen liefern andere Winkel φ nichts neues.)

Beweis:

(a) Für $0 \leq \varphi < \frac{\pi}{4}$ gilt: $\alpha^2 := cos^2(\varphi) - sin^2(\varphi) > 0$
Nach einer quadratischen Ergänzung lautet die Gleichung von $K \cap E$ dann:
$\alpha^2\tilde{x}^2 + 4sin(\varphi)cos(\varphi)\tilde{x} + \tilde{y}^2 = \alpha^2$
$\Leftrightarrow \alpha^2(\tilde{x}^2 + \frac{4}{\alpha^2}sin(\varphi)cos(\varphi)\tilde{x} + (\frac{2}{\alpha^2}sin(\varphi)cos(\varphi))^2) + \tilde{y}^2 = \alpha^2 + \frac{4}{\alpha^2}sin^2(\varphi)cos^2(\varphi)$
$\Leftrightarrow \alpha^2(\tilde{x}^2 + \frac{2}{\alpha^2}sin(\varphi)cos(\varphi))^2 + \tilde{y}^2 = \alpha^2 + \frac{4}{\alpha^2}sin^2(\varphi)cos^2(\varphi)$
$= \frac{1}{\alpha^2}[\alpha^4 + 4sin^2(\varphi)cos^2(\varphi)] = \frac{1}{\alpha^2}[cos^4(\varphi) - 2cos^2(\varphi)sin^2(\varphi) + sin^4(\varphi) + 4sin^2(\varphi)cos^2(\varphi)]$
$= \frac{1}{\alpha^2}[cos^2(\varphi) + sin^2(\varphi)]^2 = \frac{1}{\alpha^2}$
Setze $\overline{x} = \tilde{x} + \frac{2}{\alpha^2}sin(\varphi)cos(\varphi)$ und $\overline{y} = \tilde{y}$ und $\overline{z} = \tilde{z}$ und $a = \frac{1}{\alpha^2}$ und $b = \frac{1}{\alpha}$
Dann lautet die Gleichung von $K \cap E$ in der $(\overline{x}, \overline{y})$-Ebene:

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

(b) Für $\varphi = \frac{\pi}{4}$ und $\tilde{z} = 1$ ergibt sich für $K \cap E$ die Gleichung:

$$4sin(\frac{\pi}{4})cos(\frac{\pi}{4})\tilde{x} + \tilde{y}^2 = 0 \Leftrightarrow 4 \cdot \frac{\sqrt{2}}{2} \cdot \frac{\sqrt{2}}{2}\tilde{x} + \tilde{y}^2 = 0 \Leftrightarrow \tilde{y}^2 = -2\tilde{x}$$

(c) Die analoge Rechnung wie im Beweis von (a) liefert:

$$\alpha^2 = sin^2(\varphi) - cos^2(\varphi) > 0 \text{ und } \frac{\overline{x}^2}{a^2} - \frac{\overline{y}^2}{b^2} = 1$$

qed.

29 Die hohe Kunst der quadratischen Ergänzung (Quadriken)

A. Einige Definitionen - quadratisch, praktisch, gut

In diesem Abschnitt sei K ein Körper mit $char(K) \neq 2$.

29.1 Definition:

(a) Eine Summe der Form

$$f = \sum_{1 \leq i \leq j \leq n} \alpha_{ij} x_i x_j + \sum_{1 \leq k \leq n} \alpha_{0k} x_k + \alpha_{00} \text{ mit } \alpha_{lm} \in K$$

heißt ein quadratisches Polynom über K.
Dabei soll mindestens ein α_{ij} mit $1 \leq i \leq j \leq n$ ungleich 0 sein.

(b) Sei f ein quadratisches Polynom über K. Dann heißt die Punktmenge

$$Q = \{(p_1, \dots, p_n) \in K^n \mid f(p_1, \dots, p_n) = 0\}$$

eine Quadrik in K^n (genauer, die durch f definierte Quadrik).

29.2 Bemerkung:

(a) Sei f ein quadratisches Polynom wie in Definition 29.1. Wir definieren:

$$\overline{x} = (1, x_1, \dots, x_n) \text{ und } \overline{A} = \begin{pmatrix} a_{00} & \cdots & a_{nn} \\ \vdots & & \vdots \\ a_{n0} & \cdots & a_{nn} \end{pmatrix} \in Mat_{n+1}(K)$$

mit $a_{ii} = \alpha_{ii}$ für $i = 0, \dots, n$ und $a_{ij} = a_{ji} = \frac{1}{2}\alpha_{ij}$ für $1 \leq i \leq j \leq n$

Dann ist $\overline{A}$ eine symmetrische Matrix und es gilt:

$$f(x_1, \dots, x_n) = \overline{x}^{tr} \cdot A \cdot \overline{x} = \begin{pmatrix} 1 & x_1 & \cdots & x_n \end{pmatrix} \cdot A \cdot \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Somit is die durch f definierte Quadrik gegeben durch

$$Q = \{(p_1, \dots, p_n) \in K^n \mid \overline{p}^{tr} \overline{A} \overline{p} = 0\} \text{ mit } \overline{p} = (1, p_1, \dots, p_n) \in K^{n+1}$$

Wir sagen auch, dass $\overline{A}$ die Quadrik Q beschreibt.

(b) Zu einem Vektor

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n \text{ heißt } \overline{x} = \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^{n+1}$$

der erweiterte Spaltenvektor zu x.

(c) Zu einer Matrix $A \in Mat_n(K)$ heißt

$$\overline{A} = \begin{pmatrix} a_{00} & a_{01} & \cdots & a_{0n} \\ a_{10} & & & \\ \vdots & & A & \\ a_{n0} & & & \end{pmatrix} \in Mat_{n+1}(K)$$

eine erweiterte Matrix.

29.3 Beispiel:

(a) Sei $K = \mathbb{R}$ und $0 < b \leq a$. Die durch das quadratische Polynom $f = \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1$ definierte Quadrik in $\mathbb{R}^2$ ist die Ellipse (hier schreiben wir x statt x_1 und y statt x_2). Sie ist die durch die erweiterte Matrix

$$\overline{A} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \frac{1}{a^2} & 0 \\ 0 & 0 & \frac{1}{b^2} \end{pmatrix} \in Mat_3(\mathbb{R}) \text{ beschriebene Quadrik, denn:}$$

$$\overline{x}^{tr}\overline{A}\overline{x} = \begin{pmatrix} 1 & x & y \end{pmatrix} \overline{A} \begin{pmatrix} 1 \\ x \\ y \end{pmatrix} = f$$

(b) Sei $K = \mathbb{R}$ und $c > 0$. Die durch das quadratische Polynom $f = y^2 - 4cx$ definierte Quadrik in $\mathbb{R}^2$ ist die Parabel. Sie wird beschrieben durch die erweiterte Matrix

$$\overline{A} = \begin{pmatrix} 0 & -2c & 0 \\ -2c & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \in Mat_3(\mathbb{R}), \text{ denn } \begin{pmatrix} 1 & x & y \end{pmatrix} \overline{A} \begin{pmatrix} 1 \\ x \\ y \end{pmatrix} = -4cx + y^2 = f$$

(c) Sei $K = \mathbb{R}$ und seien $a, b \in \mathbb{R}_+$. Die durch das quadratische Polynom $f = \frac{x^2}{a^2} - \frac{y^2}{b^2} - 1$ definierte Quadrik in $\mathbb{R}^2$ ist die Hyperbel. Sie wird beschrieben durc die erweiterte Matrix

$$\overline{A} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \frac{1}{a^2} & 0 \\ 0 & 0 & -\frac{1}{b^2} \end{pmatrix} \in Mat_3(\mathbb{R}), \text{ denn } \begin{pmatrix} 1 & x & y \end{pmatrix} \overline{A} \begin{pmatrix} 1 \\ x \\ y \end{pmatrix} = f$$

29.4 Definition: (affine Abbildungen, Affinität)

(a) Sei $v \in K^n$. Die Abbildung $\begin{array}{rccc} \tau_v : & K^n & \to & K^n \\ & x & \mapsto & x + v \end{array}$ heißt die Translation von v.

(b) Eine Abbildung $\varphi : K^n \to K^n$ heißt eine affine Abbildung, wenn es einen Vektor $v \in K^n$ und eine K-lineare Abbildung $\psi : K^n \to K^n$ gibt, mit $\varphi(x) = (\tau_v \circ \psi)(x)$.
Anders ausgedrückt, für alle $x \in K^n$ muss $\varphi(x) = v + \psi(x)$ gelten.
(Hierbei ist $\varphi(0) = v$.)

(c) Zu einer affinen Abbildung $\varphi : K^n \to K^n$ kann man wie folgt eine beschreibende Matrix definieren: Sei $A = (a_{ij}) = M_E(\psi)$ und $v = \varphi(0) = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$. Dann gilt also:

$$\varphi(x_1, \ldots, x_n) = A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \text{ für alle } (x_1, \ldots, x_n) \in K^n. \text{ Setzt man}$$

$$\overline{A} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ b_1 & & & \\ \vdots & & A & \\ b_n & & & \end{pmatrix}, \text{ so folgt: } \overline{A} \cdot \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 1 \\ y_1 \\ \vdots \\ y_n \end{pmatrix} \text{ mit } (y_1, \ldots, y_n) = \varphi(x_1, \ldots, x_n)$$

(d) Eine Affine Abbildung $\varphi : K^n \to K^n$ heißt eine Affinität, wenn φ bijektiv ist. Da eine Translation stets bijektiv ist, gilt:
φ Affinität $\Leftrightarrow A$ invertierbar $\Leftrightarrow \overline{A}$ invertierbar

29.5 Satz:

Sei $Q \subseteq K^n$ eine Quadrik und $\varphi : K^n \to K^n$ eine Affinität.
Dann ist $\varphi(Q) \subseteq K^n$ ebenfalls eine Quadrik.

Beweis: Schreibe $\varphi(v) = Bv + b$ mit $b \in K^n$ und $B \in GL_n(K)$. Setze

$$\overline{B} = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline & & & \\ b & & B & \\ & & & \end{array}\right) \in GL_{n+1}(K)$$

Dann ist φ auch gegeben durch $\varphi(\overline{v}) = \overline{B}\overline{v}$ für $\overline{v} = \begin{pmatrix} 1 \\ v \end{pmatrix} \in K^{n+1}$. Somit gilt $\varphi^{-1}(\overline{w}) = \overline{C}\overline{w}$ für $w \in K^{n+1}$ mit

$$\overline{C} = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline & & & \\ -B^{-1}b & & B^{-1} & \\ & & & \end{array}\right) = \overline{B}^{-1}$$

Sei Q beschrieben durch die Matrix $\overline{A} \in Mat_{n+1}(K)$. Dann gilt:

$w = \varphi(v) \in \varphi(Q) \Leftrightarrow v \in Q \Leftrightarrow \overline{v}^{tr}\overline{A}\overline{v} = 0 \Leftrightarrow (\overline{w}^{tr}\overline{C}^{tr})\overline{A}(\overline{C}\overline{w}) = 0$

$\Leftrightarrow \overline{w}^{tr}\overline{D}\overline{w} = 0$ mit $\overline{D} = \overline{C}^{tr}\overline{AC}$

$\Leftrightarrow w$ ist in der durch g definierten Quadrik enthalten, wobei $g(x_1, \ldots, x_n) = \overline{x}^{tr}\overline{D}\overline{x}$ gilt. qed.

29.6 Beispiel:

Sei $Q \subseteq \mathbb{R}^2$ die durch $f(x_1, x_2) = x_1^2 + 9x_2^2 - 6x_1x_2 + 128x_1 - 46x_2 - 129$ definierte Quadrik. Um welche Art von Kegelschnitt handelt es sich?

Die Quadrik Q wird beschrieben durch die Matrix

$$\overline{A} = \begin{pmatrix} -129 & 64 & -23 \\ 64 & 1 & -3 \\ -23 & -3 & 9 \end{pmatrix}$$

Nun führen wir die Koordinatentransformation $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ mit $\tilde{x}_1 = x_1 - 3x_2$ und $\tilde{x}_2 = x_2$ durch. Die Quadrik $\varphi(Q)$ ist dann definiert durch $g(\tilde{x}_1, \tilde{x}_2) = \tilde{x}_1^2 + 128\tilde{x}_1 + 338\tilde{x}_2 - 129$ (Ersetze in f: $x_1 = \tilde{x}_1 + 3\tilde{x}_2, x_2 = \tilde{x}_2$)

Als nächstes führen wir die Koordinatentransformation $\psi : \mathbb{R}^2 \to \mathbb{R}^2$ durch mit $\overline{x}_1 = \tilde{x}_1 + 64$ und $\overline{x}_2 = 169\tilde{x}_2 - \frac{4225}{2}$. Dann ist $(\psi \circ \varphi)(Q)$ definiert durch $h(\overline{x}_1, \overline{x}_2) = \overline{x}_1 + 2\overline{x}_2$
(Ersetze in g: $\tilde{x}_1 = \overline{x}_1 - 64, \tilde{x}_2 = \frac{1}{169}\overline{x}_2 + \frac{25}{2}$)

Also ist $(\psi \circ \varphi)(Q)$ und daher auch Q eine Parabel.

Satz:

Die Menge der Punkte $(x_1, x_2) \in \mathbb{R}^2$, die einer quadratischen Gleichung

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + 2b_1x_1 + 2b_2x_2 + c = 0$$

genügt, ist stets ein Kegelschnitt.

Frage: Wie findet man (systematisch) ein Koordinatensystem (bzw. eine Affinität), in dem eine gegebene Quadrik Q eine möglichst einfache definierende quadratische Gleichung besitzt?

B. Klasse! Eine Kassifikation der reellen Quadriken im n-dimensionalen Raum!

Im Folgenden sei $K = \mathbb{R}$. Ferner sei eine Quadrik $Q \subseteq \mathbb{R}^n$ gegeben, die beschrieben werde durch eine symmetrische Matrix

$$\overline{A} = \left(\begin{array}{c|ccc} a_{00} & a_{01} & \cdots & a_{0n} \\ \hline a_{10} & & & \\ \vdots & & A & \\ a_{n0} & & & \end{array}\right) \in Mat_{n+1}(\mathbb{R}) \text{ mit } A = (a_{ij}) \in Mat_n(\mathbb{R})$$

29.7 Bemerkung:

Nach der Hauptachsentransformation symmetrischer Matrizen gibt es eine orthogonale Matrix $T \in O_n(\mathbb{R})$, so dass

$$D = T^{tr}AT = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

eine Diagonalmatrix ist.

In dem man T um die Multiplikation mit

$$\begin{pmatrix} \varepsilon_1 & & 0 \\ & \ddots & \\ 0 & & \varepsilon_n \end{pmatrix} \text{ mit } \varepsilon_1 = \begin{cases} 1 & \text{, falls } \lambda_i = 0 \\ \dfrac{1}{\sqrt{|\lambda_i|}} & \text{, falls } \lambda_i \neq 0 \end{cases}$$

abändert und gegebenenfalls eine Zeilen-/Spaltenpermutation durchführt, erhält man nun eine Matrix $S \in GL_n(\mathbb{R})$, so dass $\hat{D} = S^{tr}AS$ die Gestalt

$$\hat{D} = \begin{pmatrix} I_k & 0 & 0 \\ 0 & -I_l & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

besitzt (vgl. 26.4 und 26.12). Nun setze

$$\overline{S} = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & S & \\ 0 & & & \end{array}\right) \in GL_{n+1}(\mathbb{R})$$

Führt man nun die durch S gegebene Koordinatentransformation durch, so wird $\varphi(Q)$ definiert durch

$$g(\tilde{x}_1, \ldots, \tilde{x}_n) = \tilde{x}_1^2 + \ldots + \tilde{x}_k^2 - \tilde{x}_{k+1}^2 - \ldots - \tilde{x}_{k+l}^2 + 2(c_{01}\tilde{x}_1 + \ldots + c_{0n}\tilde{x}_n) + c_{00} \text{ mit } c_{00}, \ldots, c_{0n} \in \mathbb{R}$$

Also wird $\varphi(Q)$ beschrieben durch die Matrix

$$\tilde{D} = \overline{S}^{tr}\overline{AS} = \left(\begin{array}{c|ccc} c_{00} & c_{01} & \cdots & c_{0n} \\ \hline c_{01} & I_k & 0 & 0 \\ \vdots & 0 & -I_l & 0 \\ c_{0n} & 0 & 0 & 0 \end{array}\right)$$

Frage: Kann man auch $c_{01}, \ldots, c_{0n}$ „wegtransformieren“?

29.8 Theorem: (Affine Hauptachsentransformation reeller Quadriken)

Gegeben sei eine Quadrik $Q \subseteq \mathbb{R}$, die beschrieben werde durch eine symmetrische Matrix $A \in Mat_{n+1}(\mathbb{R})$. Dann gibt es eine Affinität $\varphi : \mathbb{R}^n \to \mathbb{R}^n$, so dass $\varphi(Q) \subseteq \mathbb{R}^n$ eine Quadrik ist, die durch eine Gleichung in Hauptachsenform definiert wird.
Dies ist eine Gleichung von einer der folgenden Bauarten:

(a) $\tilde{x}_1^2 + \ldots + \tilde{x}_k^2 - \tilde{x}_{k+1}^2 - \ldots - \tilde{x}_{k+l}^2 = 0$

(b) $\tilde{x}_1^2 + \ldots + \tilde{x}_k^2 - \tilde{x}_{k+1}^2 - \ldots - \tilde{x}_{k+l}^2 - 1 = 0$

(c) $\tilde{x}_1^2 + \ldots + \tilde{x}_k^2 - \tilde{x}_{k+1}^2 - \ldots - \tilde{x}_{k+l}^2 + 2\tilde{x}_{k+l+1} = 0$

Beweis: Wir starten mit einem Koordinatensystem, in dem Q durch eine Matrix $\tilde{D}$ wie am Ende von Bemerkung 29.7 beschrieben wird.
Nun wenden wir eine Translation an, d.h. die Matrix

$$\overline{U} = \left(\begin{array}{c|cccccccc}
1 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\
\hline
-c_{10} & 1 & & & & & & & \\
\vdots & & \ddots & & & & & 0 & \\
-c_{k0} & & & \ddots & & & & & \\
c_{(k+1)0} & & & & \ddots & & & & \\
\vdots & & & & & \ddots & & & \\
c_{(k+l)0} & & & & & & \ddots & & \\
0 & & & & & & & \ddots & \\
\vdots & & 0 & & & & & & \ddots \\
0 & & & & & & & & 1
\end{array}\right)$$

Wir erhalten $\overline{C} = \overline{U}^{tr}\tilde{D}\overline{U}$ mit einer symmetrischen Matrix $\overline{C} \in Mat_{n+1}(\mathbb{R})$ der Gestalt

$$\overline{C} = \left(\begin{array}{c|cccccc}
d_{00} & 0 & \cdots & 0 & c_{0(k+l+1)} & \cdots & c_{0n} \\
\hline
0 & I_k & 0 & 0 & 0 & \cdots & 0 \\
\vdots & 0 & -I_l & 0 & & & \vdots \\
0 & 0 & 0 & 0 & & & \vdots \\
c_{0(k+l+1)} & 0 & & & \ddots & & \vdots \\
\vdots & \vdots & & & & \ddots & \vdots \\
c_{0n} & 0 & \cdots & \cdots & \cdots & \cdots & 0
\end{array}\right)$$

Nun unterscheiden wir drei Fälle:

(a) $d_{00} = c_{0(k+l+1)} = \ldots = c_{0n} = 0$

In diesem Fall wird Q definiert durch:

$$\begin{pmatrix} 1 & \tilde{x}_1 & \cdots & \tilde{x}_n \end{pmatrix} \overline{C} \begin{pmatrix} 1 \\ \tilde{x}_1 \\ \vdots \\ \tilde{x}_n \end{pmatrix} = \tilde{x}_1^2 + \ldots + \tilde{x}_k^2 - \tilde{x}_{k+1}^2 - \ldots - \tilde{x}_{k+l}^2$$

(b) $d_{00} \neq 0, c_{0(k+l+1)} = \ldots = c_{0n} = 0$. O.E. gelte $d_{00} < 0$
(sonst multipliziere die Gleichung von Q mit -1 und vertausche k und l)

Setze $\delta = \sqrt{-d_{00}}$ und multipliziere die Gleichung von Q mit $\frac{1}{\delta^2}$. Führe dann die Koordinatentransformation $\tilde{x}_i \mapsto \delta\tilde{x}_i$ durch. Dann lautet die neue Gleichung von Q:

$$\tilde{x}_1^2 + \ldots + \tilde{x}_k^2 - \tilde{x}_{k+1}^2 - \ldots - \tilde{x}_{k+l}^2 - 1 = 0, \text{ denn:}$$

$$\begin{pmatrix} 1 & \tilde{x}_1 & \cdots & \tilde{x}_n \end{pmatrix} \overline{C} \begin{pmatrix} 1 \\ \tilde{x}_1 \\ \vdots \\ \tilde{x}_n \end{pmatrix} = \begin{pmatrix} 1 & \tilde{x}_1 & \cdots & \tilde{x}_n \end{pmatrix} \begin{pmatrix} d_{00} \\ \tilde{x}_1 \\ \vdots \\ \tilde{x}_k \\ -\tilde{x}_{k+1} \\ \vdots \\ -\tilde{x}_{k+l} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$= d_{00} + \tilde{x}_1^2 + \ldots + \tilde{x}_k^2 - \tilde{x}_{k+1}^2 - \ldots - \tilde{x}_{k+l}^2 \mid \cdot \frac{1}{\delta^2}$$

$$\rightsquigarrow -1 + \frac{1}{\delta^2}\tilde{x}_1^2 + \ldots + \frac{1}{\delta^2}\tilde{x}_k^2 - \frac{1}{\delta^2}\tilde{x}_{k+1}^2 - \ldots - \frac{1}{\delta^2}\tilde{x}_{k+l}^2$$

In der Sprache der Matrizen ist diese Transformation gegeben durch:

$$\overline{D} = \overline{V}^{tr}\overline{C}\overline{V} \text{ mit } \overline{V} = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & \delta I_k & 0 & 0 \\ \vdots & 0 & -\delta I_l & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

(c) Es gebe ein $r \in \{k+l+1, \ldots, n\}$ mit $c_{0r} \neq 0$. O.E. gelte $c_{0(k+l+1)} \neq 0$ (ansonsten permutiere $c_{0(k+l+1)}, \ldots, c_{0n}$ entsprechend)

Führe nun die folgende Koordinatentransformation durch:

$\tilde{x}_i \mapsto \tilde{x}_i$ für $c \neq k+l+1$ und
$\tilde{x}_{k+l+1} \mapsto c_{0(k+l+1)}\tilde{x}_{k+l+1} + \ldots + c_{0n}\tilde{x}_n + \frac{1}{2}d_{00}$

In den neuen Koordinaten lautet die Gleichung von Q dann:

$$\tilde{x}_1^2 + \ldots + \tilde{x}_k^2 - \tilde{x}_{k+1}^2 - \ldots - \tilde{x}_{k+l}^2 + 2\tilde{x}_{k+l+1} = 0, \text{ denn:}$$

$$\begin{pmatrix} 1 & \tilde{x}_1 & \cdots & \tilde{x}_n \end{pmatrix} \overline{C} \begin{pmatrix} 1 \\ \tilde{x}_1 \\ \vdots \\ \tilde{x}_n \end{pmatrix}$$

$$= d_{00} + c_{0(k+l+1)}\tilde{x}_{k+l+1} + \ldots + c_{0n}\tilde{x}_n + \tilde{x}_1^2 + \ldots + \tilde{x}_k^2 - \tilde{x}_{k+1}^2 - \ldots - \tilde{x}_{k+l}^2 + c_{0(k+l+1)}\tilde{x}_{k+l+1} + \ldots + c_{0n}\tilde{x}_n$$

$$= \underbrace{d_{00} + 2c_{0(k+l+1)}\tilde{x}_{k+l+1} + \ldots + c_{0n}\tilde{x}_n}_{2\tilde{x}_{k+l+1}} + \tilde{x}_1^2 + \ldots + \tilde{x}_k^2 - \tilde{x}_{k+1}^2 - \ldots - \tilde{x}_{k+l}^2$$

In der Sprache der beschreibenden Matrizen ist diese Koordinatentransformation gegeben durch $\overline{E} = \overline{W}^{tr}\overline{C}\overline{W}$, wobei $\overline{W}$ den Zeilen-/Spaltenoperationen entspricht, bei denen man mit Hilfe von $c_{0(k+l+1)}$ nacheinander $d_{00}, c_{0(k+l+2)}, \ldots, c_{0n}$ beseitigt.

Es ergibt sich:

$$\overline{E} = \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & I_k & & & & & & \vdots \\ \vdots & & -I_l & & & & & \vdots \\ 0 & & & 0 & & & & \vdots \\ 1 & & & & & & & \vdots \\ 0 & & & & & & & \vdots \\ \vdots & & & & & & & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \end{pmatrix}$$

qed.

Kapitel IX: Zurück in die Realität (Anwendungen)

30 EZNE TOTAC KAPUTIE UERERTRAGUXG (Kodierungstheorie)

A. Die allgemeine Situation

30.1 Bemerkung:

Eine Nachricht soll über einen „verrauschten“ Übertragungskanal übermittelt werden. Der Empfänger erhält also eine eventuell veränderte Nachricht. Der Empfänger möchte:

(a) feststellen, ob die Nachricht verändert wurde

(b) gegebenenfalls die urspüngliche Nachricht wieder herstellen

Damit dies gelingt, soll der Sender geeignete „Zusatzinformationen“ übertragen.

Skizze: (K endliches Alphabet $\widehat{=}$ endlicher Körper)

Nachricht		**Kodierer**		**Übertragungskanal**
$(a_1, \ldots, a_n) \in K^n$	$\rightarrow$	aus $(a_1, \ldots, a_n) \in K^n$ wird $(b_1, \ldots, b_m) \in K^m$	$\rightarrow$	zu $(b_1, \ldots, b_m)$ wird ein Fehler $(c_1, \ldots, c_m)$ „addiert“

	Dekodierer		**Ausgabe**
$\rightarrow$	1.) aus dem empfangenen $(b_1 + c_1, \ldots, b_m + c_m)$ wird das „wahrscheinlichste“ $(b'_1, \ldots, b'_m)$ ermittelt 2.) zu $(b'_1, \ldots, b'_m)$ wird ein passendes $(a'_1, \ldots, a'_n)$ ermittelt	$\rightarrow$	$(a'_1, \ldots, a'_n)$

30.2 Beispiel: (Der Wiederholungscode)

Sei $r \geq 2$ und C die Menge der möglichen Nachrichteneinheiten (Buchstaben, Elemente von $\mathbb{F}_p$, ...). Um $a \in C$ zu ermitteln, senden wir $\underbrace{(a, a, \ldots, a)}_{r\text{-mal}} \in C^r$ über den Kanal.

Dekodierer: Wähle als wahrscheinlichste Nachricht $a' \in C$ diejenige, die im empfangenen Tupel am häufigsten vorkommt.

Vorteil: Indem man $r >> 0$ wählt, kann man die Wahrscheinlichkeit einer Fehldekodierung sehr klein machen.

Nachteil: Um eine relativ kleine Nachricht zu übertragen, müssen sehr viele Daten über den Kanal geschickt werden.

30.3 Beispiel: (Der ISBN-Code)

Es soll eine 9-stellige Buchnummer übertragen werden: $(a_1, \ldots, a_9) \in \{0, \ldots, 9\}^9$

Kodierer: Ein Prüfsymbol $a_{10} \in \{0, 1 \ldots, 9, X\}$ wird wie folgt berechnet:

$$a_{10} \equiv -10a_1 - 9a_2 - 8a_3 - \ldots - 2a_9 (mod 11), \text{ wobei } X = \overline{10} \text{ gelte}$$

(a) Der ISBN-Code erkennt alle Einzelfehler.

(b) Der ISBN-Code erkennt alle Vertauschungsfehler (Transpositionen)

Wir nehmen z.B. an, dass die ersten beiden Zahlen vertauscht wurden, d.h. dass das Tupel $(a_2, a_1, a_3, \ldots, a_{10})$ epfangen wurde. O.E. gelte $a_1 \neq a_2$. Wäre das empfangene Tupel ein Codewort, so müsste gelten:

$$10a_2 + 9a_1 + 8a_3 + \ldots 2a_9 + a_{10} \equiv 0 (mod 11)$$

Es gilt sicher:

$$10a_1 + 9a_2 + 8a_3 + \ldots 2a_9 + a_{10} \equiv 0 (mod 11)$$

Die Differenz ergibt: $a_2 - a_1 \equiv 0 (mod 11)$. Wegen $a_1, a_2 \in \{0, \ldots, 9\}$ gilt:
$a_2 - a_1 \in \{-9, -8, \ldots, 8, 9\}$. Dies liefert $a_2 - a_1 = 0$, also $a_1 = a_2 \lightning$

Also kann das emfangene Tupel $(a_2, a_1, a_3, \ldots, a_{10})$ den Prüfsummentest nicht bestehen.

Fragen: 1.) Wie kann man die „Zusatzinformationen“ mathematisch exakt messen?

2.) Wie kann man die Fehlererkennugs- und Korrekturfähigkeit eines Codes definieren und messen?

B. Das mathematische Modell

Die Buchstabenmenge sei ein endlicher Körper K.

30.4 Satz:

Sei K ein endlicher Körper. Die Charakteristik von K ist die kleinste Zahl $p > 0$ mit $\underbrace{1_K + 1_K + \ldots + 1_K}_{p\text{-mal}} = 0_K$. Sie wird mit $char(K)$ bezeichnet.

(a) Die Zahl $p = char(K)$ ist eine Primzahl.

(b) Sei $p = char(K)$. Dann gibt es eine Zahl $e > 0$ mit $\#K = p^e$.

(c) Zu jeder Primzahlpotenz $q = p^e$ gibt es bis auf die Isomorphie genau einen Körper K mit $\#K = q$. Er wird mit $\mathbb{F}_q$ bezeichnet.

Beweis:

(a) Sei $char(K) = ab$ mit $a, b > 1$.
$\Rightarrow (\underbrace{1_K + \ldots + 1_K}_{a\text{-mal}}) \cdot (\underbrace{1_K + \ldots + 1_K}_{b\text{-mal}}) = 0_K \Rightarrow a \cdot 1_K = 0_K$ oder $b \cdot 1_K = 0_K \lightning$
zur Minimalität von $char(K)$.

(b) Die Menge $\{0_K, 1_K, 2_K, \ldots, (p-1)_K\} \subseteq K$ ist ein Teilkörper von K, der zu $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ isomorph ist.
Identifiziere $\mathbb{F}_p$ nun mit diesem Teilkörper, d.h. es gilt $\mathbb{F}_p \subseteq K$. Betrachte nun K als $\mathbb{F}_p$-Vektorraum.

Dann folgt: $e = dim_{\mathbb{F}_p}(K) < \infty$. Sei $\{v_1, \ldots, v_e\}$ eine Basis dieses Vektorraumes. Es folgt, dass $\begin{array}{rcl}\varphi : K & \to & (\mathbb{F}_p)^e \\ v_i & \mapsto & e_i\end{array}$ ein Isomorphismus von Vektorräumen ist.
Somit folgt: $\#K = \#(\mathbb{F}_p)^e = p^e$

(c) Für $e = 1$ gibt es den Körper $K = \mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$. Für $e > 1$ wird die Existenz und Eindeutigkeit von $\mathbb{F}_p$ in der Vorlesung „Algebra und Zahlentheorie“ bewiesen. qed.

30.5 Definition:

Die Menge der Codewörter (d.h. der kodierten Nachrichteneinheiten) eines Codes sei $C \subsetneq (\mathbb{F}_p)^n$. Dann heißt

$$R = \frac{log_p(\#C)}{n}$$

die Informationsrate des Codes C.

30.6 Bemerkung:

Offenbar gilt $0 \leq R \leq 1$. Der Satz von Shannon besagt, dass es zu jeder Informationsrate $R < 1$ für $n >> 0$ Codes gibt, bei denen die Wahrscheinlichkeit einer Fehldekodierung beliebig klein gemacht werden kann.

Problem: Wie findet man solche „guten“ Codes?

Im Folgenden sei $e > 0, p$ eine Primzahl, $q = p^e$ und $K = \mathbb{F}_q$ (meist: $q = p, e = 1$).

30.7 Definition:

Sei $K = \mathbb{F}_q$ und $V = (\mathbb{F}_q)^n$. Für zwei Tupel $v = (a_1, \ldots, a_n) \in \mathbb{F}_q$ und $w = (b_1, \ldots, b_n) \in \mathbb{F}_q$ heißt

$$d(v, w) = \{i \mid a_i \neq b_i\}$$

der Hamming-Abstand von v und w.

30.8 Satz:

Sei $K = \mathbb{F}_q$. Dann ist der Hamming-Abstand

$$\begin{array}{rcl} d : \mathbb{F}_q^n \times \mathbb{F}_q^n & \to & \mathbb{R} \\ (v, w) & \mapsto & d(v, w) \end{array}$$

eine Metrik auf $\mathbb{F}_q^n$.

Beweis: Offenbar gilt: $d(v, w) \geq 0$ und $d(v, w) = 0 \Leftrightarrow v = w$. Weiter ist $d(v, w) = d(w, v)$ klarerweise erfüllt. Zu zeigen ist also die Dreiecksungleichung:

Seien $v = (a_1, \ldots, a_n), w = (b_1, \ldots, b_n), u = (c_1, \ldots, c_n) \in \mathbb{F}_q^n$. O.E. gebe es ein $i_0 \in \{1, \ldots, n\}$ mit $a_i \neq b_i$ für $1 \leq i \leq i_0$ und $a_i = b_i$ für $i_0 + 1 \leq i \leq n$. Nach der Definition von $d(v, w)$ gilt somit $d(v, w) = i_0$. Jetzt definieren wir:

1.) $i_1 =$Anzahl der Stellen in $\{1, \ldots, i_0\}$, an denen sich w und u unterscheiden

2.) $i_2 =$Anzahl der Stellen in $\{i_0 + 1, \ldots, n\}$, an denen sich w und u unterscheiden (und somit auch v und u)

An den $i_0 - i_1$ Stellen in $\{1, \ldots, i_0\}$, an denen w und u gleich sind, unterscheiden sich v und u sicher.
Insgesamt folgt: $d(v, w) = i_0, d(w, u) = i_1 + i_2$ und $d(v, u) \geq (i_0 - i_1) + i_2$

$\Rightarrow d(v, u) + d(u, w) \geq (i_0 - i_1) + i_2 + i_1 + i_2 = i_0 + 2i_2 \geq i_0 = d(v, w)$ qed.

30.9 Definition:

Sei $C \subseteq \mathbb{F}_q^n$.

(a) Die Zahl $d(C) = min\{d(x, y) \mid x, y \in C\}$ heißt der Minimalabstand von C.

(b) Für $v \in \mathbb{F}_q^n$ und $r \geq 0$ heißt $B_r(v) = \{w \in \mathbb{F}_q^n \mid d(v, w) \leq r\}$ die Kugel vom Radius r um v.

(c) Sei $t \geq 0$. Eine Teilmenge $C \subseteq \mathbb{F}_q^n$ heißt ein t-fehlerkorrigierender Code, wenn $d(C) \geq 2t + 1$ gilt.

30.10 Satz:

Sei $C \subseteq \mathbb{F}_q^n$ ein t-fehlerkorrigierender Code.

(a) Zu jedem $v \in \mathbb{F}_q^n$ gibt es höchstens ein $c \in C$ mit $d(v, c) \leq t$.

(b) Die Kugeln $B_t(c)$ mit $c \in C$ sind paarweise disjunkt.

Beweis:

(a) Seien $c, c' \in C$ mit $d(v, c) \leq t, d(v, c') \leq t$
$\Rightarrow d(c, c') \leq d(c, v) + d(c', v) \leq 2t < 2t + 1 \leq d(C) \Rightarrow c = c'$

(b) Folgt sofort aus (a). qed.

30.11 Beispiel:

Sei $K = \mathbb{F}_2$ und $V = \mathbb{F}_2^5$. Der Code $C = \{c_1, c_2, c_3, c_4\}$ mit
$c_1 = (0, 0, 0, 0, 0), c_2 = (1, 1, 1, 0, 0), c_3 = (0, 0, 1, 1, 1), c_4 = (1, 1, 0, 1, 1)$ erfüllt $d(C) = 3$.

Insbesondere ist C 1-fehlerkorrigierend. Es gilt:

$$R = \frac{1}{5} log_2(\#C) = \frac{1}{5} log_2(4) = 0,4$$

C. Lineare Codes

30.12 Definition:

Ein $\mathbb{F}_q$-Untervektorraum $C \subseteq \mathbb{F}_q^n$ heißt ein linearer Code der Länge n und der Dimension $k = dim_{\mathbb{F}_q}(C)$.
Ist $d = d(C)$ der Minimalabstand von C, so heißt C auch ein $[n, k, d]$-Code.

30.13 Beispiel:

Sei $C \subseteq \mathbb{F}_q^n$ der von $\{(1,0,0,1),(0,1,1,1)\}$ erzeugte $\mathbb{F}_q$-Untervektorraum. Dann gilt:

$$C = \{(0,0,0,0),(1,0,0,1),(0,1,1,1),(1,1,1,0)\}$$

Also ist C ein $[4,2,2]$-Code.

$$\begin{array}{rrcl} Kodierer: & (0,0) & \mapsto & (0,0,0,0) \\ & (0,1) & \mapsto & (0,1,1,1) \\ & (1,0) & \mapsto & (1,0,0,1) \\ & (1,1) & \mapsto & (1,1,1,0) \end{array}$$

Dekodierer: Zu jedem empfangenen $v \in \mathbb{F}_q^n$ gibt es genau ein $c \in C$, so dass sich v und c an höchstens einer der ersten drei Stellen unterscheiden und die gleiche vierte Stelle haben.
Bei ≤ 1 Übertragungsfehlern in den ersten drei Stellen wird v korrekt dekodiert.

30.14 Satz:

Sei $C \subseteq \mathbb{F}_q^n$ ein $[n,k,d]$-Code.

(a) Es gilt: $R = \frac{k}{n}$

(b) Es gilt: $d(C) = min\{\#\{i \mid c_i \neq 0\} \mid (c_1,\ldots,c_n) \in C \setminus \{0\}\}$
$=$ kleinste Anzahl der Einträge $\neq 0$ eines Codewortes $\neq (0,\ldots,0)$

Beweis:

(a) $R = \dfrac{log_q(\#C)}{n} = \dfrac{log_q(q^k)}{n} = \dfrac{k}{n}$

(b) Für $w = (a_1,\ldots,a_n), w = (b_1,\ldots,b_n) \in \mathbb{F}_q^n$ gilt:

$$d(v,w) = \#\{i \mid a_i \neq b_i\} = \#\{i \mid a_i - b_i \neq 0\} = d(\underbrace{v-w}_{\in C},0)$$

Also gilt: $d(C) = min\{d(v,w) \mid v,w \in C, v \neq w\} = min\{d(u,0) \mid u \in C \setminus \{0\}\}$ qed.

30.15 Beispiel:

Der Code C aus Beispiel 30.13 erfüllt $d(C) = 2$.
Ein $[u,k,d]$-Code ist $\lfloor \dfrac{d-1}{2} \rfloor$-Fehlerkorrigierend.

Fragen: 1.) Wie kann man „gute“ Codes konstruieren?
2.) Gibt es „beliebig gute“ lineare Codes?

30.16 Bemerkung: (erweiterte Codes)

Sei C ein $[n,k,d]$-Code. Definiere

$$\overline{C} = \{(c_1,\ldots,c_{n+1}) \in \mathbb{F}_q^{n+1} \mid (c_1,\ldots,c_n) \in C, c_1 + \ldots + c_{n+1} = 0\}$$

Dann ist $\overline{C}$ ein linearer Code der Länge $n+1$ und der Dimension k. Es gilt:

$$d(\overline{C}) \in \{d(C), d(C)+1\}$$

Der Code $\overline{C}$ heißt der erweiterte Code zu C.

30.17 Definition:

Sei $C \subseteq \mathbb{F}_q^n$ ein $[n,k,d]$-Code. Dann ist

$$\widehat{C} = \{(c_1,\ldots,c_{n-1}) \in \mathbb{F}_q^{n-1} \mid \text{ es gibt ein } c_n \in \mathbb{F}_q \text{ mit } (c_1,\ldots,c_n) \in C\}$$

ein linearer Code der Länge $n-1$. Er heißt der punktierte Code zu C.

30.18 Satz:

Sei $C \subseteq \mathbb{F}_q^n$ ein $[n,k,d]$-Code. Dann ist $\widehat{C} \in \mathbb{F}_q^{n-1}$ ein $[n-1,k',d']$-Code mit $k' \in \{k,k-1\}$ und $d' \in \{d,d-1\}$.

Beweis: „$d' \in \{d,d-1\}$" Die Codewörter $(c_1,\ldots,c_n) \in C \setminus \{0\}$ mit minimalem $d(c,0)$ erfüllen

$$d((c_1,\ldots,c_{n-1}),(0,\ldots,0)) \in \{d,d-1\}$$

„$k' \in \{k,k-1\}$" Zuerst zeigen wir $k' \leq k$: Sei $\{v_1,\ldots,v_k\}$ eine Basis von C. Der Vektor v_i' entstehe aus v_i durch Streichen des letzten Eintrags. Dann ist $\{v_1',\ldots,v_k'\}$ ein Erzeugendensystem von $\widehat{C}$. Gilt dabei $k' < k$, so kann man $\lambda_1,\ldots,\lambda_k \in \mathbb{F}_q$ finden mit $\lambda_1 v_1' + \ldots + \lambda_k v_k' = 0$ und $(\lambda_1,\ldots,\lambda_k) \neq (0,\ldots,0)$. Also gibt es ein Codewort $w = \lambda_1 v_1 + \ldots + \lambda_k v_k = (0,\ldots,0,\alpha)$ mit $\alpha \in \mathbb{F}_q$. Wegen $(\lambda_1,\ldots,\lambda_k) \neq (0,\ldots,0)$ gilt $\alpha \neq 0$. Wählt man nun eine Basis $\{v_{i_1}',\ldots,v_{i_l}'\}$ von $\widehat{C}$, so ist $\{v_{i_1}',\ldots,v_{i_l}',w\}$ eine Basis von C.
Es folgt: $l = dim_{\mathbb{F}_q}(\widehat{C}) = k-1 = k'$ qed.

30.19 Satz: (Die Singleton-Schranke)

Sei $C \subseteq \mathbb{F}_q^n$ ein $[n,k,d]$-Code. Dann gilt:

$$k \leq n-d+1$$

Gilt hier Gleichheit, so heißt C ein MDS-Code („maximum distance separable"). Die Zahl $n-k$ heißt auch die Redundanz von C. Es gilt also: $n-k \geq d-1$.

Beweis: Sei $\tilde{C} \subseteq \mathbb{F}_q^{n-d+1}$ der Code, der entsteht, indem man C $(d-1)$-mal punktiert. Dann folgt: $d(\tilde{C}) \geq d-(d-1) \geq 1$. Also werden durch das Punktieren keine zwei Codewörter von C gleich.
Es folgt: $\#\tilde{C} = \#C = q^k$ und somit ist $\tilde{C}$ ein $[n-d+1,k,\tilde{d} \geq 1]$-Code.
Somit erhalten wir: $q^k = \#\tilde{C} = \#\mathbb{F}_q^{n-d+1} \Rightarrow k \leq n-d+1$ qed.

30.20 Definition:

Sei $C \subseteq \mathbb{F}_q^n$ ein $[n,k,d]$-Code.

(a) Ist $\{v_1,\ldots,v_k\}$ eine Basis von C, so heißt die Matrix

$$G = \begin{pmatrix} v_1 \\ \vdots \\ v_k \end{pmatrix} \in Mat_{k,n}(\mathbb{F}_q)$$

eine Generatormatrix von C.

(b) Der lineare Code

$$C^{\perp} = \{v \in \mathbb{F}_q^n \mid \underbrace{< v, c >}_{\text{Standardbilinearform}} = 0 \text{ für alle } c \in C\} \subseteq \mathbb{F}_q^n$$

heißt der zu C duale Code.

(c) Eine Generatormatrix von $C^{\perp}$ heißt auch eine Parity-Check-Matrix (oder Kontrollmatrix) von C.

30.21 Bemerkung:

(a) Da $< \cdot, \cdot >: \mathbb{F}_q^n \times \mathbb{F}_q^n \to \mathbb{F}_q$ nicht ausgeartet ist, gilt:

$$dim_{\mathbb{F}_q}(C^{\perp}) = n - k$$

(b) Es gibt viele Vektoren $v \in \mathbb{F}_q^n$ mit $< v, v >= 0$. Es kann sogar $C^{\perp} = C$ passieren („selbst duale Codes“).

(c) $(C^{\perp})^{\perp} = C$

(d) Sei $H \in Mat_{n-k,n}(\mathbb{F}_q)$ eine Parity-Check-Matrix von C. Für $v \in \mathbb{F}_q^n$ gilt dann:

$v \in C \Leftrightarrow < v, w >= 0$ für alle $w \in C^{\perp}$

$\Leftrightarrow < v, w_i >= 0$ für eine Basis $\{w_1, \dots, w_{n-k}\}$ von $C^{\perp} \Leftrightarrow H \cdot v = 0$

Man kann mit einer Parity-Check-Matrix also prüfen, ob $v \in C$ gilt.

30.22 Beispiel:

Sei $C =< (1, 0, 0, 1), (0, 1, 1, 1) >\subseteq \mathbb{F}_2^4$.

(a) Die Matrix $G = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix} \in Mat_{2,4}(\mathbb{F}_2)$ ist eine Generatormatrix von C.

(b) Es gilt $C^{\perp} =< (1, 1, 0, 1), (1, 0, 1, 1) >$ und $H = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix}$ ist eine Parity-Check-Matrix von C.

30.23 Beispiel: (Hamming-Codes)

Sei p eine Primzahl, $e > 0$, $q = p^e$, $n = 1 + q + q^2 + \ldots + q^{k-1} = \dfrac{q^k - 1}{q - 1}$

Auf $\mathbb{F}_q \setminus \{0\}$ betrachten wir eine Äquivalenzrelation:

$$(a_1, \dots, a_k) \sim (b_1, \dots, b_k) \Leftrightarrow \text{ es gibt ein } \lambda \in \mathbb{F}_q \text{ mit } (a_1, \dots, a_k) = \lambda \cdot (b_1, \dots, b_k)$$

Die Äquivalenzklassen $[(a_1, \dots, a_k)]$ entsprechen eindeutig den Geraden durch 0 in $\mathbb{F}_q^n$.

Behauptung 1: Es gibt genau n Äquivalenzklassen.

Beweis: $\#(\mathbb{F}_q^k \setminus \{0\}) = q^k - 1$. Jede Äquivalenzklasse besitzt $q-1$ Elemente. Also gibt es $n = \dfrac{q^k-1}{q-1}$ Äquivalenzklassen. $\Rightarrow$ Behauptung 1

Wähle aus jeder Äquivalenzklasse einen Repräsentanten $p_i \in \mathbb{F}_q^k$. Dann bilden wir den Matrix $G = \left(\begin{array}{c|c|c|c} p_1 & p_2 & \cdots & p_n \end{array} \right) \in Mat_{k,n}(\mathbb{F}_q)$.
Sei $\tilde{C}$ der Code mit Generatormatrix G. Dann heißt $C = (\tilde{C})^\perp$ der $[n, n-k]$-Hamming-Code. Er besitzt die Länge n und die Dimension $n-k$.

Behauptung 2: Die Kugeln $B_1(c)$ mit $c \in C$ bilden eine disjunkte Überdeckung von $\mathbb{F}_q^n$, d.h. es gilt: $\mathbb{F}_q^n = \dot{\bigcup\limits_{c \in C}} B_1(c)$.
Insbesondere kann jedes empfangene Wort eindeutig dekodiert werden (C ist ein „vollkommener Code“). Ferner folgt: $d(C) = 3$.
Somit ist C ein 1-fehlerkorrigierender $[n, n-k, 3]$-Code.

Beweis: Für jedes $c \in C$ enthält $B_1(c)$ genau $1 + n(q-1) = q^k$ Vektoren.
Es sind $\#C = q^{n-k}$ Kugeln. Diese Kugeln enthalten also insgesamt maximal $q^{n-k} \cdot q^k = q^n$ Vektoren. Somit genügt es zu zeigen, dass $B_1(c), B_1(c')$ disjunkt sind für $c \neq c'$. Andernfalls wäre $d(C) \leq 2$, also $d(c - c', 0) \leq 2$.
Da $G = \left(\begin{array}{c|c|c} p_1 & \cdots & p_n \end{array} \right)$ eine Parity-Check-Matrix von C ist, gibt es dann $\lambda_i, \lambda_j \in \mathbb{F}_q$ mit $G \cdot (c - c') = G \cdot (\underbrace{\lambda_i e_i + \lambda_j e_j}_{\text{höchstens 2 Einträge} \neq 0}) = \lambda_i p_i + \lambda_j p_j = 0$.
Da die Geraden $< p_1 >, \ldots, < p_n >$ paarweise verschieden sind, folgt $\lambda_i = \lambda_j = 0$ im Widerspruch zu $c \neq c'$ $\Rightarrow$ Behauptung 2

Anwendung der Codierungstheorie:

(a) 1969 - Mariner-Raumsonde (Mars): [32,6,16]-Hadamard-Code

(b) 1980/81 - Voyager-Raumsonden (äußere Planeten): Golay-Codes

(c) seit 1980 - CD-Player, CD-Rom, Harddisc: Reed-Solomon-Codes

(d) seit 1990 - Internet: CRC-Codes („cyclic redundancy check“)

(e) seit 2000 - Handy, DVD, Digital-TV, ADSL-Modem: CIRC-Codes („cross interlaced Reed-Solomon codes“)

31 Wir basteln Züge (Rekursionsgleichungen)

A. Versuch und Irrtum

31.1 Beispiel:

Wir bauen Züge. Es gibt drei Arten von Waggons:

- Typ I: Länge 1 Einheit
- Typ A: Länge 2 Einheiten
- Typ B: Länge 2 Einheiten

I.) Aufgabe: Wie viele Züge der Länge 12 gibt es?
Ein Zug ist eine geordnete Folge von Waggons mit einem Anfang und einem Ende.

(a) Bestimme die Züge der Länge 3:

III, IA, AI, IB, BI

(b) Sei t_n die Anzahl der Züge der Länge n. Finde t_n für kleine n:

n	1	2	3	4
t_n	1	3	5	11

(Länge 4: IIII, IIA, IIB, IAI, IBI, AII, BII, AA, BB, AB, BA)

(c) Finde eine Rekursionsformel für t_n!
Betrachte dazu 3 Fälle, abhängig vom ersten Waggon:

$t_n = t_{n-1} + 2t_{n-2}$ mit Anfangsbedinungen $t_1 = 1, t_2 = 3$

(d) Berechne die Antwort durch Auswerten der Rekursionsformel:

n	1	2	3	4	5	6	7	8	9	10	11	12
t_n	1	3	5	11	21	43	85	171	341	683	1365	2731

Es gibt also 2731 Züge der Länge 12.

II.) Was ist die tiefere Bedeutung der Zahl 2731? Wie hängt t_n von n ab? Wie hängt 2731 von 12 ab? Gibt es eine Formel für t_n in Abhängigkeit von n?

- Tipp 1: Betrachte $t_2, t_4, t_6, \ldots$ und $t_1, t_3, t_5, \ldots$
- Tipp 2: Die Folge wächst ungefähr wie 2^n

$$t_n = \frac{1}{3}(2^{n+1} + (-1)^n) = \begin{cases} \frac{1}{3}(2^{n+1} + 1) & \text{, falls } n \text{ gerade} \\ \frac{1}{3}(2^{n+1} - 1) & \text{, falls } n \text{ ungerade} \end{cases}$$

Frage: Wie löst man die Rekursion $t_n = t_{n-1} + 2t_{n-2}$ mit Anfangsbedingungen $t_1 = 1, t_2 = 2$ systematisch?

Idee: Versuche, einfachere verwandte Aufgaben zu lösen!
Relaxiere die Anfangsbedingunen, z.B. betrachte den allgemeinen Fall $t_1 = 1, t_2$ beliebig:

(a) $t_2 = 1$: $1, 1, 3, 5, 11, 21, \ldots$

(b) $t_2 = 2$: $1, 2, 4, 8, 16, 31, \ldots \longrightarrow t_n = 2^{n-1}$

(c) $t_2 = 3$: $1, 3, 5, 11, 21, 43, \ldots$

(d) $t_2 = 4$: $1, 4, 6, 14, 26, 54, \ldots$

(e) $t_2 = -1$: $1, -1, 1, -1, 1, -1, \ldots \longrightarrow t_n = (-1)^{n-1}$

(f) $t_2 = -2$: $1, -2, 0, -4, -4, -12, \ldots$

(g) $t_2 = \frac{1}{2}$: $1, \frac{1}{2}, \frac{5}{2}, \frac{17}{2}, \frac{31}{2}, \ldots$

- Für $t_2 = 2$ und $t_2 = -1$ ist die Folge (t_n) einfach zu beschreiben.
- Hängt die „Zugfolge“ (c) irgendwie von (b) und (e) ab?

Ja: $t_n = \frac{4}{3} \cdot 2^{n-1} - \frac{1}{3} \cdot (-1)^{n-1}$
Ansatz: $t_n = a \cdot 2^{n-1} + b \cdot (-1)^{n-1}$ und damit $1 = t_1 = a + b, 3 = t_2 = 2a - b$

$$\Rightarrow \left\{ \begin{array}{ccccc} a & + & b & = & 1 \\ 2a & - & b & = & 3 \end{array} \right. \Rightarrow 3a = 4, a = \frac{4}{3} \Rightarrow b = 1 - a = -\frac{1}{3}$$

III.) Was hat dies alles mit Linearer Algebra zu tun?
1. Methode: x^n-Ansatz

- Lasse die Anfangsbedingungen ganz weg und setze $t_n = x^n$ in die Rekursionsgleichung ein.
  $\Rightarrow x^n = x^{n-1} + 2x^{n-2} \underset{x \neq 0}{\Rightarrow} x^2 - x - 2 = 0 \Rightarrow (x-2)(x+1) = 0$
  $\Rightarrow x = 2$ oder $x = 1 \Rightarrow t_n = 2^n$ oder $t_n = (-1)^n$ sind Lösungen der Rekursion
- Nun versuche, die Anfangsbedingungen zu erfüllen:
  $t_n = a \cdot 2^n + b \cdot (-1)^n, t_1 = 1, t_2 = 2 \Rightarrow a = \frac{2}{3}, b = \frac{1}{3}$
  $\Rightarrow t_n = \frac{1}{3} \cdot 2^{n+1} + \frac{1}{3} \cdot (-1)^n$

2. Methode: Ansatz mit Matrizen:

$$\begin{pmatrix} t_{n-1} \\ t_n \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} t_{n-2} \\ t_{n-1} \end{pmatrix} \text{ liefert } \begin{pmatrix} t_{n-1} \\ t_n \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 \\ 2 & 1 \end{pmatrix}}_{:= A} \begin{pmatrix} t_{n-2} \\ t_{n-1} \end{pmatrix}$$

$$\begin{pmatrix} t_1 \\ t_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, A \cdot \begin{pmatrix} t_1 \\ t_2 \end{pmatrix}, A^2 \cdot \begin{pmatrix} t_1 \\ t_2 \end{pmatrix}, \ldots$$

Trick: $A = T \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} T^{-1}, A^{100} = T \begin{pmatrix} \lambda_1^{100} & 0 \\ 0 & \lambda_2^{100} \end{pmatrix} T^{-1}$

(a) $\chi_A(x) = det(A - x \cdot I_2) = det \begin{pmatrix} -x & 1 \\ 2 & 1-x \end{pmatrix} = x^2 - x - 2 = (x-2)(x+1)$

Also besitzt A die Eigenwerte $\lambda_1 = 2$ und $\lambda_2 = -1$ Daher gilt: $dim_{\mathbb{Q}}(Eig(A,2)) = dim_{\mathbb{Q}}(Eig(A,-1)) = 1$ und A ist diagonalisierbar.

(b) Berechnung der Eigenräume:

$Eig(A,2) = Kern(f_A - 2 \cdot id_{\mathbb{Q}^2})$

Betrachte: $\begin{pmatrix} -2 & 1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow Eig(A,2) =< (1,2) >$

$Eig(A,-1) =< (1,-1) >$

(c) Schreibe $\begin{pmatrix} t_1 \\ t_2 \end{pmatrix}$ als Linearkombination der Eigenvektoren:

$$\begin{pmatrix} t_1 \\ t_2 \end{pmatrix} = \frac{4}{3} \begin{pmatrix} 1 \\ 2 \end{pmatrix} - \frac{1}{3} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Nun folgt:

$$\begin{aligned} \begin{pmatrix} t_{n-1} \\ t_n \end{pmatrix} &= A^{n-2} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \frac{4}{3} \cdot A^{n-2} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} - \frac{1}{3} \cdot A^{n-2} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ &= \frac{4}{3} \cdot 2^{n-2} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} - \frac{1}{3} \cdot (-1)^{n-2} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} \end{aligned}$$

Und wir erhalten:

$$t_n = \frac{4}{3} \cdot 2^{n-2} \cdot 2 - \frac{1}{3} \cdot (-1)^{n-2} \cdot (-1) = \frac{1}{3}(2^{n+1} + (-1)^n)$$

31.2 Beispiel:

Wie versuchen, mit der gleichen Methode die Rekursion $t_n = 2t_{n-1} - 2t_{n-2}$ mit $t_1 = 1, t_2 = 3$ zu lösen.

(a) Wertetabelle:

n	1	2	3	4	5	6	7	8	9	10	11	12	13
t_n	1	3	4	2	-4	-12	-16	-8	16	48	64	32	-64

(b) Finde ein Muster! Es gibt einen 4er-Rhytmus:

$t_{n+1} = (-4) \cdot t_n$

Woher kommt dieses Muster?

<u>x^n-Methode:</u> Setze $t_n = x^n$ in die Rekursion ein:

$x^n = 2x^{n-1} - 2x^{n-2} \underset{x\neq 0}{\Rightarrow} x^2 - 2x + 2 = 0$

$\Rightarrow (x-1)^2 = -1 \Rightarrow x - 1 = \pm i \Rightarrow x = 1 \pm i$

Versuche die Folge (t_n) als Linearkombination von $(1+i)^n$ und $(1-i)^n$ darzustellen:

$t_n = a(1+i)^n + b(1-i)^n$ mit $a, b \in \mathbb{C}$

$$\left\{\begin{array}{rcll} 1 &=& a(1+i) &+\ b(1-i) \\ 3 &=& a2i &+\ b(-2i) \end{array}\right. \Rightarrow \left\{\begin{array}{rcll} 1 &=& (b-\frac{3}{2}i)(i+1) &+\ b(1-i) \\ -\frac{3}{2}i &=& a &-\ b \end{array}\right.$$

$$\Rightarrow \left\{\begin{array}{rcl} 1 &=& 2b - \frac{3}{2}i + \frac{3}{2} \\ a &=& b - \frac{3}{2}i \end{array}\right. \Rightarrow \left\{\begin{array}{rcl} b &=& -\frac{1}{4} + \frac{3}{4}i \\ a &=& -\frac{1}{4} - \frac{3}{4}i \end{array}\right.$$

Es folgt: $t_n = (-\frac{1}{4} - \frac{3}{4}i)(1+i)^n + (-\frac{1}{4} + \frac{3}{4}i)(1-i)^n = 2Re[(-\frac{1}{4} + \frac{3}{4}i)(1-i)^n]$

31.3 Beispiel:

Wir wollen die folgende lineare 3-Term-Rekursion lösen:

$$t_{n+2} = 2t_{n+1} + t_n - 2t_{n-1} \text{ mit } t_0 = 1, t_1 = 4, t_2 = 4$$

(a) Raten:

n	0	1	2	3	4	5	6	7	$\cdots$
t_n	1	4	4	10	16	34	64	130	$\cdots$

$$t_n = \begin{cases} 2^n & , n \text{ gerade} \\ 2^n + 2 & , n \text{ ungerade} \end{cases} \Rightarrow t_n = 2^n + 1 + (-1)^{n+1}$$

(b) <u>x^n-Methode:</u> $x^{n+1} = 2x^{n+1} + x^n - 2x^{n-1}$

$\underset{x\neq 0}{\Rightarrow} x^3 - 2x^2 - x + 2 = 0 \Rightarrow (x^2-1)(x-2) = 0 \Rightarrow (x-2)(x+1)(x-1) = 0$

$\Rightarrow x \in \{-1, 1, 2\} \Rightarrow u_n = (-1)^n, v_n = 1^n,_n = 2^n$

Setze $t_n = au_n + bv_n + cw_n$ mit $a, b, c \in \mathbb{R}$

$t_0 = 1, t_1 = 4, t_2 = 4 \Rightarrow a = -1, b = 1, c = 1 \Rightarrow t_n = -(-1)^n + 1 + 2^n$

(c) Matrix-Eigenwert-Methode:

$$\begin{pmatrix} t_n \\ t_{n+1} \\ t_{n+2} \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & 1 & 2 \end{pmatrix}}_{:= A} \begin{pmatrix} t_{n-1} \\ t_n \\ t_{n+1} \end{pmatrix} \text{ mit } \begin{pmatrix} t_0 \\ t_1 \\ t_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \\ 4 \end{pmatrix}$$

$$\begin{pmatrix} t_n \\ t_{n+1} \\ t_{n+2} \end{pmatrix} = A^n \cdot \begin{pmatrix} t_0 \\ t_1 \\ t_2 \end{pmatrix}$$

1.) Finde die Eigenwerte und Eigenvektoren von A:

$\chi_A(x) = det(A - x \cdot I_3) = -x^3 + 2x^2 + x - 2 = -(x-1)(x+1)(x-2)$

Die Eigenwerte sind also $\lambda_1 = 1, \lambda_2 = -1, \lambda_3 = 2$

2.) Berechnung der Eigenräume:

$Eig(A,-1) =< (1,-1,1) >, Eig(A,1) =< (1,1,1) >, Eig(A,2) =< (1,2,4) >$

3.) Schreibe den Anfangsvektor als Linearkombination der Eigenvektoren:

$$\begin{pmatrix} 1 \\ 4 \\ 4 \end{pmatrix} = a \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} + b \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + c \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} \Rightarrow a = -1, b = 1, c = 1$$

4.) Wende A^n an:

$$\begin{pmatrix} t_n \\ t_{n+1} \\ t_{n+1} \end{pmatrix} = A^n \begin{pmatrix} 1 \\ 4 \\ 4 \end{pmatrix} = -A^n \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} + A^n \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + A^n \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}$$

$$= -(-1)^n \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + 2^n \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}$$

$\Rightarrow t_n = -(-1)^n + 1 + 2^n$

B. Die schlaue Theorie

Gegeben sei ein System linearer Differenzengleichungen:

$$(*) \qquad \begin{pmatrix} x_1(k+1) \\ \vdots \\ x_n(k+1) \end{pmatrix} = A \cdot \begin{pmatrix} x_1(k) \\ \vdots \\ x_n(k) \end{pmatrix} \text{ mit } A \in Mat_n(K) \text{ und } \begin{pmatrix} x_1(0) \\ \vdots \\ x_n(0) \end{pmatrix} \in K^n$$

$$\text{Schreibe } x(k) = \begin{pmatrix} x_1(k) \\ \vdots \\ x_n(k) \end{pmatrix} \in K^n$$

31.4 Bemerkung:

(a) Der Vektor $x(0)$ ist vorgegeben. Es gilt: $x(k) = A^k \cdot x(0)$ für jedes $k \geq 0$. Jedoch ist A^k unter Umständen schwierig zu berechnen.

(b) Angenommen die Matrix A ist diagonalisierbar:
Schreibe $A = TDT^{-1}$ mit $T \in GL_n(K)$ und D als Diagonalmatrix.

$\Rightarrow A^k = \underbrace{(TDT^{-1}) \cdots (TDT^{-1})}_{k\text{-mal}} = TD^kT^{-1}$ und D^k ist leicht zu berechnen.

31.5 Satz: (Lösung linearer Differenzengleichungssysteme)

Gegeben sei ein System linearer Differenzengleichungen $(*)$. Die Matrix A sei diagonalisierbar. Betrachte die folgenden Schritte:

(a) Berechne $\chi_A(x) \in K[x]$ und eine Faktorisierung $\chi_A(x) = (\lambda_1 - x) \cdots (\lambda_n - x)$ mit (nicht notwendigerweise verschiedenen) $\lambda_1, \ldots, \lambda_n \in K$.

(b) Berechne die Eigenräume $Eig(A, \lambda_i)$ und finde so eine Basis $B = \{v_1, \ldots, v_n\}$ von K^n bestehend aus Eigenvektoren von A, d.h. es gelte $Av_i = \lambda_i v_i$

(c) Stelle den Vektor $x(0)$ als Linearkombination der Basis B dar:
$x(0) = c_1v_1 + \ldots + c_nv_n$ mit $c_i \in K$.

(d) Gib $x(k) = c_1\lambda_1^k v_1 + \ldots + c_n\lambda_n^k v_n$ aus.

Dies ist ein Algorithmus, der eine Formel für die Lösung $x(k)$ von $(*)$ berechnet.

Beweis: Es gilt: $x(k) = A^k \cdot x(0) = A^k(c_1v_1 + \ldots + c_nv_n) = c_1(A^kv_1) + \ldots + c_n(A^kv_n)$
$= c_1\lambda_1^k v_1 + \ldots + c_n\lambda_n^k v_n$ qed.

31.6 Beispiel:

Gesucht sei die Lösung von $\left\{ \begin{array}{lclcl} x(k) & = & x(k-1) & + & 3y(k-1) \\ y(k) & = & 2x(k-1) & + & 2y(k-1) \end{array} \right.$ mit $x(0) = a_1, y(0) = a_2$

(a)
$$\begin{pmatrix} x(k) \\ y(k) \end{pmatrix} = A \cdot \begin{pmatrix} x(k-1) \\ y(k-1) \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 2 & 2 \end{pmatrix} \cdot \begin{pmatrix} x(k-1) \\ y(k-1) \end{pmatrix}$$

$\Rightarrow \chi_A(x) = x^2 - 3x - 4 = (x-4)(x+1)$

$\Rightarrow \lambda_1 = 4, \lambda_2 = -1$

(b) $Eig(A,4) =< (1,1) >, Eig(A,-1) =< (3,-2) >$

(c)

$$\text{Schreibe} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = c_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + c_2 \begin{pmatrix} 3 \\ -2 \end{pmatrix} = \begin{pmatrix} c_1 + 3c_2 \\ c_1 - 2c_2 \end{pmatrix}$$

$\Rightarrow c_1 = \frac{2}{5}a_1 + \frac{3}{5}a_1, c_2 = \frac{1}{5}a_1 - \frac{1}{5}a_2$

(d)

$$\begin{pmatrix} x(k) \\ y(k) \end{pmatrix} = (\frac{2}{5}a_1 + \frac{3}{5}a_1) \cdot 4^k \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} + (\frac{1}{5}a_1 - \frac{1}{5}a_2) \cdot (-1)^k \cdot \begin{pmatrix} 3 \\ -2 \end{pmatrix}$$

Insgesamt gilt also:

$$\left\{ \begin{array}{rclcl} x(k) & = & \frac{1}{5}(2a_1 + 3a_2)4^k & + & \frac{3}{5}(a_1 - a_2)(-1)^k \\ y(k) & = & \frac{1}{5}(2a_1 + 3a_2)4^k & - & \frac{2}{5}(a_1 - a_2)(-1)^k \end{array} \right.$$

32 Das Letzte (Ausblick)

A. Die Singulärwertzerlegung

32.1 Satz:

Sei $A \in Mat_{m,n}(\mathbb{R})$. Dann gibt es orthogonale Matrizen $U \in O_m(\mathbb{R}), V \in O_n(\mathbb{R})$ und eine Matrix $S \in Mat_{m,n}(\mathbb{R})$ der Form

$$S = \left(\begin{array}{ccc} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \\ \hline 0 & \cdots & 0 \end{array}\right) \text{ oder } S = \left(\begin{array}{ccc|c} \lambda_1 & & 0 & 0 \\ & \ddots & & \vdots \\ 0 & & \lambda_m & 0 \end{array}\right)$$

mit folgenden Eigenschaften:

(a) $A = USV^{tr}$

(b) $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_s \geq 0$ mit $s = min\{m, n\}$ (<u>Singulärwerte</u> von A)

(c) $r = \#\{i \mid \lambda_i \neq 0\} = Rang(A)$

(d) Die ersten r Spalten von V bilden eine ONB des Zeilenraumes von A.

(e) Die restlichen $n - r$ Spalten von V bilden eine ONB des Kerns von f_A.

(f) Die ersten r Spalten von U bilden eine ONB des Spaltenraumes von A.

(g) Die restlichen $n - r$ Spalten von U bilden eine ONB des Kerns von $f_{A^{tr}} = (f_A)^*$.

<u>Idee:</u> AA^{tr} ist tausendmal schöner wie A!

- $(AA^{tr})^{tr} = (A^{tr})^{tr} A^{tr} = AA^{tr} \Rightarrow AA^{tr}$ ist symmetrisch.
- AA^{tr} hat nur Eigenwerte ≥ 0.

B. Pseudoinverse

Sei $A \in Mat_{m,n}(\mathbb{R}$ und $A = USV^{tr}$ ihre Singulärwertzerlegung (SVD)

32.2 Definition:

Sei $S^+ = \left(\begin{array}{ccc|c} \frac{1}{\lambda_1} & & 0 & 0 \\ & \ddots & & \vdots \\ 0 & & \frac{1}{\lambda_r} & 0 \\ \hline 0 & \cdots & 0 & 0 \end{array}\right) \in Mat_{m,n}(\mathbb{R})$ mit $r = Rang(A)$ und $A^+ = US^+V^{tr}$.

Dann heißt A^+ die Pseudoinverse (oder Moore-Penrose-Inverse) von A.

32.3 Satz:

Betrachte ein LGS $A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$ mit $A \in Mat_{m,n}(\mathbb{R})$ und $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{R}^m$

(a) Ist $Ax = b$ unlösbar, so ist $x^+ = A^+b$ die lineare Ausgleichslösung kleinster Menge, d.h. $\|Ax^+ - b\|$ ist minimal.

(b) Ist $Ax = b$ lösbar, so ist $x^+ = A^+b$ die Lösung kleinster Länge.

C. Multilineare Algebra

- Bilinearformen: $\Phi : V \times V \to K$ linear im 1. und 2. Argument.
- Determinantenfunktionen: $det : \underbrace{K^n \times \ldots \times K^n}_{\in Mat_n(K)} \to K$ linear in jedem Argument (und alternierend)
- multilineare Abbildungen: $\varphi : \underbrace{V \times \ldots \times V}_{n\text{-mal}} \to W$ linear in jedem Argument

Hallo liebe Mitstudenten!

Fehler beim Abtippen der Vorlesungsmitschrift lassen sich natürlich nicht grundsätzlich vermeiden. Solltet ihr Fehler finden so teilt mir diese doch bitte mit (sowohl Rechtschreibfehler als auch inhaltliche Fehler, wie z.B. falsche Indizes o.ä.). Am besten immer mit Seitenzahl und Nummer des entsprechenden Satzes/Korollars/... unter der Emailadresse:

patrick.vorderstemann@uni-dortmund.de

Mit freundlichen Grüßen,
Patrick