U Unitäre Räume

Im Vorkapitel hatten wir in der Euklidischen Ebene bzw. dem Euklidischen Raum Winkel und Längen betrachtet und diese Begriffe, insbesondere das Senkrechtstehen, über das Skalarprodukt und die daraus gewonnene Norm auf den $\mathbb{R}^2$ bzw. den $\mathbb{R}^3$ übertragen. Dies können wir nun auf eine große Klasse von Vektorräumen ausdehnen. Wir erhalten so die „Unitären Räume" bzw. die „Hilberträume", die in der Funktionalanalysis und deren Anwendungen eine wichtige Rolle spielen. Ein großer Teil der Quantenmechanik beispielsweise ist die Theorie von Homomorphismen in gewissen Hilberträumen.

Vereinbarung *Im ganzen Kapitel betrachten wir nur reelle oder komplexe Vektorräume, d.h. der Grundkörper ist $\mathbb{R}$ oder $\mathbb{C}$. Für eine komplexe Zahl $a = a_1 + ia_2$ ist $\overline{a} = a_1 - ia_2$ die konjugiert komplexe.*

Alle Sätze und Definitionen sind — sofern nicht ausdrücklich etwas anderes gesagt ist — für den komplexen Fall formuliert, gelten aber auch für den reellen Fall, wobei dann ein „konjugiert Strich" gegenstandslos ist.

Definition U.1 *Es sei U ein Vektorraum über $K := \mathbb{R}$ oder $K := \mathbb{C}$. Eine Abbildung $\langle \cdot, \cdot \rangle : U \times U \to K$ heißt ein positiv definites, hermitesches Skalarprodukt, wenn S1 bis S4 gelten:*

S1: $\forall_{x,y \in U}\ \langle x, y\rangle = \overline{\langle y, x\rangle}$ *(hermitesch),*

S2: $\forall_{x,y_1,y_2 \in U}\ \langle x, y_1 + y_2\rangle = \langle x, y_1\rangle + \langle x, y_2\rangle$,

S3: $\forall_{x,y \in U}\forall_{\alpha \in K}\ \langle x, \alpha y\rangle = \alpha \langle x, y\rangle$,

S4: $\forall_{x \in U}\ \langle x, x\rangle \geq 0, = 0 \Leftrightarrow x = 0$ *(positiv definit).*

Bemerkung *S1 liefert insbesondere, daß für jedes $x \in U$ der Wert $\langle x, x\rangle$ reell ist, sodaß S4 eine sinnvolle Forderung darstellt.*

Im reellen Fall sind dies genau die schon im Vorkapitel notierten Regeln für das dort im $\mathbb{R}^2$ bzw. $\mathbb{R}^3$ betrachtete Skalarprodukt. Im komplexen Fall ergibt sich eine vernünftige Verallgemeinerung erst, wenn man S_1 und S_3 die oben notierte Form gibt.

Ein – aber keineswegs das einzige – Skalarprodukt in $\mathbb{C}^2$ ist etwa:

Für $x = (\xi_1, \xi_2),\ y = (\eta_1, \eta_2)$ setze $\langle x, y\rangle = \overline{\xi}_1\eta_1 + \overline{\xi}_2\eta_2$.

So wird $\langle x, x\rangle \geq 0$ auch für komplexe Vektoren, während dies ohne das Konjugieren nicht erreicht würde.

Dies führt uns auf das *Standard-Skalarprodukt* im $\mathbb{R}^n$ oder $\mathbb{C}^n$: Für den Spaltenvektor $x \in \mathbb{C}^n$ bezeichne $x^H := \overline{x}^T$ den dazu transponierten Zeilenvektor, dessen Komponenten zudem noch konjugiert komplex genommen sind. x^H heißt der zu x *hermitesch adjungierte* Vektor.

Dann liefert im Sinne der Matrixmultiplikation

$$\langle x, y\rangle := x^H \cdot y$$

ein Skalarprodukt im $\mathbb{C}^n$, das „Standard-Skalarprodukt".

Definition U.2 *Ein reeller oder komplexer Raum mit Skalarprodukt heißt „euklidischer" bzw. „unitärer" Raum.*

Zunächst notieren wir einige Rechenregeln für das Skalarprodukt:

Satz U.3 *(i) Es ist $\langle x, 0\rangle = \langle 0, y\rangle = 0$.*

(ii) $\langle x, \alpha y + \beta z\rangle = \alpha \langle x, y\rangle + \beta \langle x, z\rangle$.

(iii) $\langle \alpha y + \beta z, x\rangle = \overline{\alpha} \langle y, x\rangle + \overline{\beta} \langle z, x\rangle$.
D.h. $\langle \cdot, \cdot\rangle$ *ist „konjugiert linear" im ersten Argument.*

Beweis:

(i) Für alle $x, y, z \in U$ ist $\langle x, 0\rangle = \langle x, 0 \cdot y\rangle = 0 \langle x, y\rangle = 0$ und damit $\langle 0, z\rangle = \overline{\langle z, 0\rangle} = \overline{0} = 0$.

(ii) Dies folgt unmittelbar aus S2, S3.

(iii) Mit S1 und (ii) ist $\langle \alpha y + \beta z, x\rangle = \overline{\langle x, \alpha y + \beta z\rangle} = \overline{\alpha \langle x, y\rangle + \beta \langle x, z\rangle} = \overline{\alpha}\ \overline{\langle x, y\rangle} + \overline{\beta}\ \overline{\langle x, z\rangle} = \overline{\alpha} \langle y, x\rangle + \overline{\beta} \langle z, x\rangle$. □

Da $\langle x, x\rangle \geq 0$ für jedes x, können wir hieraus stets die eindeutig bestimmte nicht negative Quadratwurzel ziehen:

Definition U.4 *Sei* $(U, \langle \cdot, \cdot\rangle)$ *ein Raum mit Skalarprodukt. Dann heißt die Abbildung* $|\cdot| : U \to \mathbb{R} : x \to |x| := \sqrt{\langle x, x\rangle}$ *die „Norm" von* x.

Satz U.5 *Für die Norm gelten*

N1: $\forall_{x \in U}\ |x| \geq 0, = 0 \Leftrightarrow x = 0,$

N2: $\forall_{x \in U} \forall_{\alpha \in K}\ |\alpha x| = |\alpha||x|$, *wobei* $|\alpha|$ *der Betrag der reellen oder komplexen Zahl* α *ist,*

N3: $\forall_{x,y \in U}\ |x + y| \leq |x| + |y|$ *(Dreiecksgleichung).*

Beweis:

N1: folgt unmittelbar aus S4.

N2: $|\alpha x| = \sqrt{\langle \alpha x, \alpha x\rangle} = \sqrt{\overline{\alpha} \cdot \alpha \langle x, x\rangle} = \sqrt{|\alpha|^2 \cdot \langle x, x\rangle} = |\alpha| \cdot |x|$.

Zum Beweis der Dreiecksungleichung N3 benötigen wir

Satz U.6 *In einem unitären Raum gilt die* CAUCHY-SCHWARZ*sche-Ungleichung:* $\forall_{x,y \in U}\ |\langle x, y\rangle| \leq |x| \cdot |y|$, *wobei* $=$ *nur für linear abhängige* x, y *gilt.*

Beweis: Ist $\langle x, y\rangle = 0$, so ist nichts zu beweisen. Sei also $\langle x, y\rangle \neq 0$ und damit auch $x \neq 0$, $y \neq 0$. Wir setzen $\alpha := \frac{\langle x,y\rangle}{|\langle x,y\rangle|}|y|, \beta := \frac{|\langle x,y\rangle|}{|y|}$. Dann ist (Mausefallenbeweis!)

$$\begin{aligned}
0 &\leq |\alpha x - \beta y|^2 \\
&= \langle \alpha x - \beta y, \alpha x - \beta y\rangle \\
&= \overline{\alpha}\alpha|x|^2 - \overline{\alpha}\beta \langle x, y\rangle - \overline{\beta}\alpha \langle y, x\rangle + \overline{\beta}\beta|y|^2 \\
&= |y|^2 \cdot |x|^2 - \frac{\overline{\langle x, y\rangle}}{|\langle x, y\rangle|}|y| \cdot \frac{|\langle x, y\rangle|}{|y|} \cdot \langle x, y\rangle \\
&\quad - \frac{|\langle x, y\rangle|}{|y|} \cdot \frac{\langle x, y\rangle}{|\langle x, y\rangle|} \cdot |y| \cdot \overline{\langle x, y\rangle} + \frac{|\langle x, y\rangle|^2}{|y|^2} \cdot |y|^2 \\
&= |y|^2 \cdot |x|^2 - |\langle x, y\rangle|^2.
\end{aligned}$$

Daraus folgt, die behauptete Ungleichung und Gleichheit kann nur eintreten, wenn $\alpha x - \beta y = 0$ ist.

Damit erhalten wir auch die Dreiecksungleichung N3:

$$\begin{aligned}|x+y|^2 &= \langle x+y, x+y\rangle \\ &= |x|^2 + \langle x,y\rangle + \langle y,x\rangle + |y|^2 \\ &\overset{*}{\leq} |x|^2 + |x||y| + |x||y| + |x|^2 \\ &= (|x|+|y|)^2.\end{aligned}$$

($*$ nach CAUCHY-SCHWARZ.)

$\boxed{1}$

Definition U.7 *(i) In einem unitären Raum $(U, \langle\cdot,\cdot\rangle)$ heißen $x, y \in U$ orthogonal $(x \perp y)$, wenn $\langle x, y\rangle = 0$. Hierbei sind $x = 0$ und/oder $y = 0$ zugelassen.*

(ii) Zwei Familien $(x_i | i \in I), (y_i | i \in I)$ heißen ein „Biorthonormalsystem (BONS)", wenn für alle $i, j \in I : \langle x_i, y_j\rangle = \delta_{ij}$ (Kroneckersymbol).

(iii) Eine Familie $(x_i | i \in I)$ heißt „Orthonormalsystem (ONS)", wenn sie mit sich selbst ein Biorthonormalsystem bildet, d.h. für alle $i \in I$ $\langle x_i, x_j\rangle = \delta_{ij}$. Ist diese Familie eine Basis, so sprechen wir von „Orthonormal-Basis (ONB)".

Satz U.8 *(i) Bilden $(x_i | i \in I), (y_i | i \in I)$ ein Biorthonormalsystem, so sind je endlich viele der x_i bzw. der y_i linear unabhängig.*

(ii) In einem Orthonormalsystem $(x_i | i \in I)$ sind je endlich viele linear unabhängig.

Beweis:

(i) Sei $(x_1, \ldots, x_n), (y_1, \ldots, y_n)$ ein solcher endlicher Teil des Biorthonormalsystems, sodaß also $\langle x_i, x_j\rangle = \delta_{ij}$ $(i, j = 1, \ldots, n)$. Wir betrachten eine Linearkombination $0 = \sum_{i=1}^n \alpha_i x_i$. Bilden wir das Skalarprodukt mit y_j zu einem festen j, so folgt $0 = \langle 0, y_j\rangle = \langle \sum_{i=1}^n \alpha_i x_i, y_j\rangle = \sum_{i=1}^n \overline{\alpha}_i \langle x_i, y_j\rangle = \sum_{i=1}^n \overline{\alpha}_i \delta_{ij} = \overline{\alpha}_j$. Da j beliebig war, sind somit alle Koeffizienten $\alpha_j = 0$, d.h. $(x_1, \ldots, x_n)$ unabhängig. Völlig analog schließt man für die y.

(ii) folgt aus (i), da Orthonormalsysteme ja spezielle Biorthonormalsysteme sind. □

Über die Existenz von Orthonormalsystemen gibt der folgende Satz Auskunft, der ein wichtiges Konstruktionsprinzip enthält:

Satz U.9 (Orthogonalisieren nach E. SCHMIDT) *Es sei $(u_1, u_2, \ldots)$ eine endliche oder unendliche Folge von Vektoren in einem unitären Raum $(U, \langle\cdot,\cdot\rangle)$, sodaß je endlich viele linear unabhängig sind. Dann existiert eine orthonormierte Folge $(x_1, x_2, \ldots)$, sodaß*

(i) $\forall_n$ $\mathrm{span}(u_1, \ldots, u_n) = \mathrm{span}(x_1, \ldots, x_n)$,

(ii) $\forall_{ij}$ $\langle x_i, x_y\rangle = \delta_{ij}$, d.h. die $(x_1, x_2, \ldots)$ ein ONS bilden.

Beweis: Setze $y_1 := u_1, x_1 := \frac{1}{|y_1|} y_1$.

Setze $y_{n+1} := u_{n+1} - \sum_{i=1}^n \langle x_i, u_{n+1}\rangle x_i$, $x_{n+1} := \frac{1}{|y_{n+1}|} y_{n+1}$, $(n = 1, 2, \ldots)$. Wir haben zu zeigen:

1. $\forall_n$ $y_n \neq 0$
2. $\forall_n$ $\langle x_n, x_n\rangle = 1$
3. $\forall_{m,n,m<n}$ $\langle x_n, x_m\rangle = 0$

4. $\forall_n \; \mathrm{span}(u_1, \ldots, u_n) = \mathrm{span}(x_1, \ldots, x_n)$.

Alle diese Behauptungen gelten für $n = 1$.

Wir nehmen an, sie gelten bis zu einer Nummer n und zeigen, daß sie dann auch für $n+1$ gelten:

1. Wegen 4. für n ist $\sum_{i=1}^n \langle x_i, u_{n+1}\rangle x_i \in \mathrm{span}(x_1, \ldots, x_n) = \mathrm{span}(u_1, \ldots, u_n)$. Da $(u_1, \ldots, u_{n+1})$ linear unabhängig, ist somit $y_{n+1} \neq 0$.

2. $\langle x_{n+1}, x_{n+1}\rangle = \left\langle \frac{1}{|y_{n+1}|} y_{n+1}, \frac{1}{|y_{n+1}|} y_{n+1} \right\rangle = \frac{1}{|y_{n+1}|^2} \cdot |y_{n+1}|^2 = 1$.

3. Sei $m \leq n$. Dann ist

$$\begin{aligned} |y_{n+1}| \langle x_m, x_{n+1}\rangle = \langle x_m, y_{n+1}\rangle &= \langle x_m, u_{n+1}\rangle - \sum_{i=1}^n \langle x_i, u_{n+1}\rangle \langle x_m, x_i\rangle \\ &= \langle x_m, u_{n+1}\rangle - \sum_{i=1}^n \langle x_i, u_{n+1}\rangle \delta_{mi} \\ &= \langle x_m, u_{n+1}\rangle - \langle x_m, u_{n+1}\rangle = 0. \end{aligned}$$

4. Wegen 4. für n ist $x_{n+1} \in \mathrm{span}(x_1, \ldots, x_n, u_{n+1}) = \mathrm{span}(u_1, \ldots, u_n, u_{n+1})$, also $\mathrm{span}(x_1, \ldots, x_{n+1}) \subset \mathrm{span}(u_1, \ldots, u_{n+1})$. Da die $(x_1, \ldots x_{n+1})$ als ONS unabhängig sind, folgt die Gleichheit. □

Wenden wir diesen Satz gleich an:

Satz U.10 *Sei $(U, \langle\cdot,\cdot\rangle)$ ein endlich-dimensionaler unitärer Raum. Dann gelten*

(i) U besitzt eine Orthonormalbasis.

(ii) Jedes ONS $(x_1, \ldots, x_m)$ in U läßt sich zu einer ONB ergänzen.

(iii) Zu jedem Unterraum $V \subset U$ gibt es ein orthogonales Komplement $V^\perp$, d.h. einen Unterraum $V^\perp \subset U$, der komplementär zu V ist, und für den zusätzlich gilt $\forall_{x \in V} \forall_{y \in V^\perp} \langle x, y\rangle = 0$.

Beweis:

(i) Wähle irgendeine Basis von U und orthogonalisiere nach E. SCHMIDT.

(ii) Ergänze $(x_1, \ldots, x_m)$ irgendwie zu einer Basis und orthogonalisiere.

(iii) Wähle $(x_1, \ldots, x_m)$ als ONB von V. Ergänze nach (ii) zu ONB $(x_1, \ldots, x_m, x_{m+1}, \ldots, x_n)$ von U und setze $V^\perp := \mathrm{span}(x_{m+1}, \; \ldots, x_n)$. Dann sind nach Konstruktion diese beiden Unterräume komplementär. Ferner hat $x \in V$ die Form $x = \sum_{i=1}^m \alpha_i x_i$ und $y \in V^\perp$ die Form $y = \sum_{j=m+1}^n \beta_j x_j$. Somit ist $\langle x, y\rangle = \left\langle \sum_{i=1}^m \alpha_i x_i, \sum_{j=m+1}^n \beta_j x_j \right\rangle = \sum_{i=1}^m \sum_{j=m+1}^n \overline{\alpha}_i \beta_j \langle x_i, x_j\rangle = 0$, da $\langle x_i, x_j\rangle = 0$ für $i \leq m, j \geq m+1$. □

Mit Orthonormalbasen läßt sich besonders bequem rechnen. Als Beispiele zeigen wir

Satz U.11 *Sei $(U, \langle\cdot,\cdot\rangle)$ ein endlich-dimensionaler unitärer Raum, $(u_1, \ldots, u_n)$ ONB von U. Dann gelten:*

(i) Ist $x = \sum_{i=1}^n \xi_i u_i, \quad y = \sum_{j=1}^n \eta_j u_j$, so ist $\langle x, y\rangle = \sum_{i=1}^n \overline{\xi}_i \eta_j$.

(ii) Ist $x = \sum_{i=1}^n \xi_i u_i$, so ist $|x| = \left(\sum_{i=1}^n \overline{\xi}_i \xi_i\right)^{1/2} = \left(\sum_{i=1}^n |\xi_i|^2\right)^{1/2}$.

(iii) Für jedes x ist $x = \sum_{i=1}^n \langle u_i, x\rangle u_i$ und $|x| = \left(\sum_{i=1}^n |\langle u_i, x\rangle|^2\right)^{1/2}$.

Beweis:

(i)

$$\begin{aligned}
\langle x, y\rangle &= \left\langle \sum_{i=1}^{n} \xi_i u_i, \sum_{j=1}^{n} \eta_j u_j \right\rangle = \sum_{i,j=1}^{n} \overline{\xi}_i \eta_j \langle u_i, u_j\rangle \\
&= \sum_{i,j=1}^{n} \overline{\xi}_i \eta_j \delta_{ij} = \sum_{i=1}^{n} \overline{\xi}_i \eta_i .
\end{aligned}$$

(ii) Siehe (i) für $x = y$.

(iii) Betrachten wir $y := x - \sum_{i=1}^{n} \langle u_i, x\rangle\, u_i$. Dann ist für jedes j:

$$\begin{aligned}
\langle u_j, y\rangle &= \langle u_j, x\rangle - \sum_{i=1}^{n} \langle u_i, x\rangle \langle u_j, u_i\rangle = \langle u_j, x\rangle - \sum_{i=1}^{n} \langle u_i, x\rangle\, \delta_{ij} \\
&= \langle u_j, x\rangle - \langle u_j, x\rangle = 0.
\end{aligned}$$

Die u_j bilden eine Basis von U. Damit hat jedes $z \in U$ die Form $z = \sum_{j=1}^{n} \alpha_j u_j$. Damit folgt dann $\langle z, y\rangle = \left\langle \sum_{j=1}^{n} \alpha_j u_j, y \right\rangle = \sum_{j=1}^{n} \overline{\alpha}_j \langle u_j, y\rangle = 0$. Da dies für jeden Vektor $z \in U$ gilt, gilt dies auch für $z = y$, womit dann $\langle y, y\rangle = 0$, d.h. $y = 0$ folgt. Das weitere folgt mit (i) und (ii) □

Einige wichtige Anwendungen seien noch explizit behandelt.

Satz und Definition U.12 (Proximum) *Es sei V ein m-dimensionaler Unterraum eines unitären Raumes $(U, \langle\cdot,\cdot\rangle), y \in U$. Dann existiert genau ein $y_0 \in V$, sodaß $y - y_0 \in V^{\perp}$. Für y_0 gilt*

$$|y - y_0| < |y - x| \text{ für jedes } x \in V, x \neq y_0.$$

Man nennt daher y_0 „Proximum zu y“ in V. Ist $(x_1, \ldots, x_m)$ eine ONB von V, so ist das Proximum $y_0 = \sum_{j=1}^{m} \langle x_j, y\rangle\, x_j$.

Beweis: Wähle eine ONB $(x_1, \ldots, x_m)$ von V und setze damit $y_0 := \sum_{j=1}^{m} \langle x_j, y\rangle \in V$. Ist $y \in V$, so ist offenbar $y = y_0$ und alle weiteren Aussagen gelten trivialerweise. Sei also $y \notin V$. Wir orthonormieren das System $(x_1, \ldots, x_m, y)$ und erhalten ein ONS $(x_1, \ldots, x_m, x_{m+1})$, mit dem dann

$$y = \sum_{j=1}^{m+1} \langle x_j, y\rangle\, x_j = y_0 + \langle x_{m+1}, y\rangle\, x_{m+1},$$

sodaß $y - y_0 \in \operatorname{span}(x_{m+1}) \subset (\operatorname{span}(x_1, \ldots, x_m))^{\perp} = V^{\perp}$. Ist y_0' irgendein Element von V mit $y - y_0' \in V^{\perp}$, so ist $y_0' - y_0 \in V$, da beide in V; ferner ist $y_0' - y_0 = (y - y_0) - (y - y_0') \in V^{\perp}$, da beide in $V^{\perp}$. Somit ist $y_0' - y_0 \in V \cap V^{\perp} = 0$. Damit ist $y_0' = y_0$. Schließlich gilt für y_0 die Proximum-Eigenschaft; denn für jedes $x \in V$ ist

$$\begin{aligned}
|y - x|^2 &= |(y - y_0) - (x - y_0)|^2 \\
&= \langle (y - y_0) - (x - y_0), (y - y_0) - (x - y_0)\rangle \\
&= |y - y_0|^2 + |x - y_0|^2 - (\langle (y - y_0), (x - y_0)\rangle + \langle (x - y_0), (y - y_0)\rangle).
\end{aligned}$$

Bei den letzten beiden Skalarprodukten ist je ein Faktor in V, der andere in $V^{\perp}$, sodaß die Skalarprodukte $= 0$ sind. Damit ist

$$|y - x|^2 = |y - y_0|^2 + |x - y_0|^2 \geq |y - y_0|^2,$$

wobei $=$ genau für $x = y_0$ eintritt. Unser zu irgendeiner ONB von V konstruiertes y_0 ist also das Proximum, dadurch eindeutig bestimmt und somit unabhängig von der gewählten ONB. □

Satz und Definition U.13 *Es sei V ein m-dimensionaler Unterraum eines unitären Raumes U. Für $y \in V$ sei $Py := y_0$ das Proximum zu y in V. Dann ist die Abbildung $P : U \to U$ ein Homomorphismus. Es ist $\operatorname{im} P = V$, $\ker P = V^\perp$, P ist idempotent, d.h. $P^2 = P$, ferner ist für alle $y, z \in U : \langle Py, z\rangle = \langle y, Pz\rangle$. Man nennt P den „orthogonalen Projektor" auf V.*

Beweis: Mit der eben gewonnenen Darstellung für y_0 ist $Py = y_0 = \sum_{j=1}^m \langle x_j, y\rangle\, x_j$ und damit trivialerweise homomorph. Die Aussagen über Kern, Bild und Idempotenz folgen unmittelbar aus dem vorigen Satz. Für die letzte Aussage zerlegen wir $y = Py + y^\perp, z = Pz + z^\perp$ mit $y^\perp, z^\perp \in V^\perp$, womit dann

$$\begin{aligned} \langle Py, z\rangle &= \left\langle Py, Pz + z^\perp\right\rangle = \langle Py, Pz\rangle\,, \\ \langle y, Pz\rangle &= \left\langle Py + y^\perp, Pz\right\rangle = \langle Py, Pz\rangle\,. \end{aligned}$$

□

Die eben behandelte Aufgabe, beliebige Elemente eines unitären Raumes U durch Elemente eines Unterraumes V zu approximieren, ist insbesondere von Interesse, wenn U ein nicht mehr endlich-dimensionaler Funktionen-Raum ist und für V Unterräume „schöner" Funktionen gewählt werden, etwa trigonometrische oder gewöhnliche Polynome beschränkten Grades (Fourier-Approximation, Approximation im quadratischen Mittel.) Typischerweise interessiert hier nicht nur das Proximum bezüglich eines festen Unterraumes, sondern die ganze Folge der Proxima bzgl. einer aufsteigenden Folge von Unterräumen.

Betrachten wir daher die folgende Situation: $(U, \langle\cdot,\cdot\rangle)$ sei ein unitärer Raum, darin $(u_1, u_2, \dots)$ eine beliebige Folge von Elementen, sowie $V_m := \operatorname{span}(u_1, \dots u_m)$, $(m = 1, 2, \dots)$. Zu $y \in U$ sei y_m das Proximum an y in V_m.

Nach Satz und Definition U.12 ist das Proximum y_m charakterisiert als das Element von V_m, für das $y - y_m \in V^\perp$, oder äquivalent

$$\langle u_i, y - y_m\rangle = 0 \qquad (i = 1, \dots m).$$

Stellt man y_m durch die u_j dar, so folgt

Lemma U.14 *$y_m = \sum_{j=1}^m \alpha_j u_j$ ist Proximum an y in V_m genau, wenn die Koeffizienten das System*

$$\sum_{j=1}^m \langle u_i, u_j\rangle\, \alpha_j = \langle u_i, y\rangle \qquad (i = 1, .., m)$$

erfüllen.

Beweis: Es gilt $0 = \langle u_i, y - y_m\rangle$ genau dann, wenn
$\langle u_i, y\rangle = \langle u_i, y_m\rangle = \left\langle u_i, \sum_{j=1}^m \alpha_j u_j\right\rangle = \sum_{j=1}^m \langle u_i, u_j\rangle\, \alpha_j$. □

Die Matrix dieses Systems hat in Position (i, j) das Skalarprodukt $\langle u_i, u_j\rangle$ stehen.

Definition U.15 *Die Matrix $(\langle u_i, u_j\rangle)$ heißt „Gramsche Matrix", deren Determinante $G(u_1, \dots, u_m) := \det(\langle u_i, u_j\rangle)$ heißt „Gramsche Determinante".*

Wir stellen einige Eigenschaften der Gramschen Determinante zusammen.

Satz U.16 *(i) Für jede Permutation π ist $G(u_{\pi(1)}, \dots, u_{\pi(m)}) = G(u_1, \dots, u_m)$.*

(ii) $G(u_1, \dots, u_{j-1}, u_j + \lambda u_k, u_{j+1}, \dots, u_m) = G(u_1, \dots, u_j, \dots, u_m)$ für $\lambda \in \mathbb{C}, k \neq j$.

(iii) $G(u_1, \dots, u_{j-1}, \lambda u_j, u_{j+1}, \dots, u_m) = |\lambda|^2 G(u_1, \dots, u_j, \dots, u_m)$.

(iv) Ist $u_{m+1} = v_m + d_m$ mit $v_m \in V_m, d_m \perp V_m$, so ist $G(u_1, \dots, u_m, u_{m+1}) = |d_m|^2 \cdot G(u_1, \dots, u_m)$

(v) $G(u_1, \dots, u_m) \geq 0, = 0 \Leftrightarrow (u_1, \dots, u_m)$ linear abhängig.

(vi) $G(u_1, \dots, u_m) \leq |u_1|^2 |u_2|^2 \cdots |u_m|^2$. Dabei gilt $=$ genau wenn ein $u_i = 0$ oder alle u_i paarweise orthogonal sind.

Beweis:

(i) Jede Permutation von m Elementen kann man aus sogenannten Transpositionen, die genau zwei Elemente vertauschen, aufbauen. Vertauscht man etwa u_r und u_s, so erhält man die neue Gramsche Matrix, wenn man in der alten die Zeilen r, s und die Spalten r, s vertauscht, was sich realisieren läßt, indem man von links und von rechts mit der selben Permutationsmatrix multipliziert. Da deren Determinante ± 1 ist, bleibt dann die Gramsche Determinante selbst ungeändert.

(ii) Man erhält die modifizierte Matrix, indem man zunächst zu Zeile j das $\overline{\lambda}$-fache von Zeile k addiert und danach zu Spalte j noch das λ-fache von Spalte k.

(iii) Aus Zeile j kann man den Faktor $\overline{\lambda}$, aus Spalte j den Faktor λ herausziehen.

(iv) Es ist nach (ii), da $v_m \in V_m$:

$$\begin{aligned} G(u_1, \dots, u_m, u_{m+1}) &= G(u_1, \dots u_m, v_m + d_m) = G(u_1, \dots, u_m, d_m) \\ &= \det \left(\begin{array}{ccc|c} \langle u_1, u_1 \rangle & \cdots & \langle u_1, u_m \rangle & \vdots \\ \vdots & \ddots & \vdots & \langle u_i, d_m \rangle \\ \langle u_m, u_1 \rangle & \cdots & \langle u_m, u_m \rangle & \vdots \\ \hline \cdots & \langle d_m, u_j \rangle & \cdots & \langle d_m, d_m \rangle \end{array} \right) \\ &= |d_m|^2 G(u_1, \dots, u_m), \end{aligned}$$

da $\langle d_m, u_j \rangle = \langle u_j, d_m \rangle = 0, \quad i, j \leq m$.

(v) und (vi) zeigt man zusammen simultan durch Induktion.
$m = 1$: $G(u_1) = \det(\langle u_1, u_1 \rangle) = |u_1|^2$, woraus die Behauptungen für $m = 1$ folgen.
$m \Rightarrow m + 1$: Nach (iv) ist $G(u_1, \dots, u_m, u_{m+1}) = G(u_1, \dots, u_m, d_m) = |d_m|^2 \cdot G(u_1, \dots, u_m)$. Damit ist jedenfalls $G(u_1, \dots, u_{m+1})$ reell, ≥ 0. Ferner haben wir $G(u_1, \dots, u_{m+1}) = 0$ genau dann, wenn entweder $d_m = 0$, d.h. $u_{m+1} \in V_m = \operatorname{span}(u_1, \dots, u_m)$ oder wenn $G(u_1, \dots, u_m) = 0$, d.h. $(u_1, \dots, u_m)$ linear abhängig, also insgesamt genau, wenn $(u_1, \dots, u_{m+1})$ linear abhängig. Dies beweist (v) für $m+1$. Es ist $u_{m+1} = v_m + d_m$, wobei $\langle v_m, d_m \rangle = 0$, sodaß $|u_{m+1}|^2 = |v_m|^2 + |d_m|^2 \geq |d_m|^2$. Also ist

$$\begin{aligned} G(u_1, \dots, u_{m+1}) &= |d_m|^2 \cdot G(u_1, \dots, u_m) \\ &\leq |u_{m+1}|^2 \cdot G(u_1, \dots, u_m) \\ &\leq |u_{m+1}|^2 \cdot |u_1|^2 \cdots |u_m|^2 \end{aligned}$$

nach Induktionsannahme. Da $|d_m| = |u_{m+1}|$ genau wenn $u_m = 0$ ist, d.h. $u_{m+1} = d_m \in V_m^{\perp}$, folgt auch der Zusatz. □

Bemerkung U.17 *Die Aussage von Satz U.16(v) ist eine Verallgemeinerung der Ungleichung von* CAUCHY-SCHWARZ*:*

$$\begin{aligned} 0 \le G(u_1, u_2) &= \det \begin{pmatrix} \langle u_1, u_1\rangle & \langle u_1, u_2\rangle \\ \langle u_2, u_1\rangle & \langle u_2, u_2\rangle \end{pmatrix} \\ &= |u_1|^2 \cdot |u_2|^2 - \langle u_2, u_1\rangle \langle u_1, u_2\rangle = |u_1|^2 |u_2|^2 - |\langle u_1, u_2\rangle|^2. \end{aligned}$$

Wenden wir diese Aussagen an!

Satz U.18 *Es sei $(u_1, \ldots, u_m)$ eine Basis von $V \subset U$, $y \in U$, $y_0 \in V$ des Proximum und $\triangle_\mu$ bezeichne das algebraische Komplement (mit Vorzeichen!) zur Position $(m+1, \mu)$ in der Gramschen Matrix zu $(u_1, \ldots, u_m, y)$. Dann ist $\triangle_{m+1} = G(u_1, \ldots, u_m) > 0$ und damit*

$$|y - y_0| = \left(\frac{G(u_1, \ldots, u_m, y)}{G(u_1, \ldots, u_m)} \right)^{\frac{1}{2}}$$

und

$$\hat{y} := \frac{-1}{\Delta_{m+1}} \cdot \sum_{\mu=i}^{m} \Delta_\mu u_\mu = y_0$$

des Proximum.

Beweis: Die Darstellung für $|y - y_0|$ folgt unmittelbar aus Satz U.16(iv). Es ist $\hat{y}$ in V, ferner für $1 \le i \le m$,

$$\begin{aligned} \langle u_i, y - \hat{y}\rangle &= \langle u_i, y\rangle + \frac{1}{\Delta_{m+1}} \cdot \sum_{\mu=i}^{m} \Delta_\mu \langle u_i, u_\mu\rangle \\ &= \frac{1}{\Delta_{m+1}} \Big(\sum_{\mu=i}^{m} \Delta_\mu \langle u_i, u_\mu\rangle + \Delta_{m+1} \langle u_i, y\rangle\Big) = 0, \end{aligned}$$

da hier die algebraischen Komplemente zur letzten Zeile der Gramschen Matrix mit den Elementen einer anderen Zeile kombiniert werden. Folglich ist $y - \hat{y} \in V_m^\perp$, also $\hat{y}$ das Proximum. □

Wir hatten zu Spaltenvektoren $x \in \mathbb{C}^n$ den hermitesch adjungierten Vektor $x^H := (\overline{x^T})$ eingeführt, der also ein Zeilenvektor ist mit den gegenüber x komplex-konjugierten Komponenten. Damit war

$$\langle x, y\rangle := x^H \cdot y$$

das „Standard-Skalarprodukt" auf $\mathbb{C}^n$. Wir verallgemeinern dies für beliebige komplexe Matrizen:

Definition U.19 (Adjungierte Matrix) *Sei $A \in \mathbb{C}^{n\times m}$. Dann heißt $A^H := (\overline{A^T}) = (\overline{A}^T) \in \mathbb{C}^{m\times n}$ die hermitesch adjungierte Matrix. Man erhält sie, indem man A transponiert und alle Elemente konjugiert komplex nimmt.*

Dazu einige Rechenregeln:

Lemma U.20 *(i) Für die Einheitsmatrix I ist $I^H = I$.*
(ii) $(A \cdot B)^H = B^H A^H$
(iii) $(A^{-1})^H = (A^H)^{-1}$ (sofern A^{-1} existiert).

(iv) $\det(A^H) = \overline{(\det A)}$

(v) $A^H A$ *ist die Gramsche Matrix zu den Spalten* $(a_1, \ldots, a_m)$ *von* A*, somit* $\det(A^H A) = G(a_1, \ldots, a_m)$.

Beweis:

(i) ist trivial.

(ii) $(A \cdot B)^H = \overline{(AB)^T} = \overline{B^T A^T} = B^H A^H$

(iii) Dies folgt direkt mit (i) und (ii)

(iv) Es ist $\det A^T = \det A$ und etwa aus dem Entwicklungssatz von LAPLACE folgt $\det \overline{A} = \overline{\det A}$.

(v) Die Zeilen von A^H sind die hermitesch adjungierten der Spalten von A, somit steht in Position (i, j) von $A^H A$ das Element

$$a_i^H \cdot a_j = \langle a_i, a_j \rangle$$

woraus (v) direkt folgt. □

Das liefert uns etwa die Ungleichung von HADAMARD:

Satz U.21 *Für eine Matrix* $A \in \mathbb{C}^{n \times n}$ *mit Spalten* $(a_1, \ldots, a_n)$ *gilt mit der Norm aus dem Standard-Skalarprodukt*

$$|\det A| \leq |a_1| \cdot |a_2| \cdots |a_n|,$$

wobei $=$ *genau dann eintritt, wenn ein* $a_i = 0$ *oder alle paarweise orthogonal sind.*

Beweis: Es ist

$$\begin{aligned} |\det A|^2 &= \overline{\det A} \cdot \det A = \det(A^H A) \\ &= G(a_1, \ldots, a_n) \leq |a_1|^2 \cdots |a_n|^2 \end{aligned}$$

nach Satz U.16, aus dem auch die Aussage über $=$ folgt. □

Eine weitere Anwendung kommt aus der Ausgleichsrechnung:

Problem U.22 (Approximative Lösung eines linearen Gleichungssystems) *Gegeben sei das System* $Ax = b$ *mit* $A \in \mathbb{C}^{n \times m}$*. Wir nehmen an:*

(i) $b \notin \operatorname{im} A$*, d.h. das System ist nicht lösbar.*

(ii) $m < n$ *und* $\operatorname{rg} A = m$*, d.h. die Spalten* a_i *von* A *sind linear unabhängig, erzeugen aber nicht den* $\mathbb{C}^n$*.*

Gesucht ist ein $x_0 \in \mathbb{C}^m$*, so daß* $|Ax_0 - b|$ *minimal.*

Lösung: Setze $U := \mathbb{C}^n$, $V := \operatorname{span}(a_1, \ldots, a_m) = \operatorname{im} A$, $b_0 :=$ Proximum an b in V. Dann ist $b_0 \in \operatorname{im} A$, somit $Ax = b_0$ lösbar und wegen der Rangbedingung gibt es genau eine Lösung x_0. Dann ist $|Ax_0 - b| = |b_0 - b| \leq |v - b|$ für alle $v \in \operatorname{im} A$, d.h. $|Ax_0 - b|$ minimal.

Dieses Proximum berechnet man nach Lemma U.14 aus der Basis $(u_1, \ldots, u_m) := (a_1, \ldots, a_m)$ von $\operatorname{im} A$, als $b_0 = \sum_{j=1}^m \alpha_j a_j = Ax_0$ mit $x_0 = (\alpha_j)$. Dabei ist zu lösen

$$\sum_{j=1}^m \langle a_i, a_j \rangle \, \alpha_j = \langle a_i, b \rangle \quad (i = 1, \ldots, m),$$

d.h.

$$A^H A x_0 = A^H b$$

Wir haben also

Satz U.23 *Das Problem U.22 hat unter den gemachten Voraussetzungen genau eine Lösung x_0. Sie ist die eindeutig bestimmte Lösung der sog. „Normalen-Gleichung“:*

$$A^H Ax = A^H b.$$

Mittels Eigenschaften der adjungierten Matrix lassen sich einige Klassen von Matrizen beschreiben, die für die weiteren Überlegungen wichtig sind.

Definition U.24 *Es sei $A \in \mathbb{C}^{n\times n}$.*

(i) A heißt „normal“, wenn $A^H A = AA^H$

(ii) A heißt „hermitesch“ oder „selbstadjungiert“, wenn $A^H = A$. Im reellen Fall bedeutet dies $A^T = A$. Man nennt dann A „symmetrisch“.

(iii) A heißt „unitär“, wenn $A^H = A^{-1}$. Im reellen Fall bedeutet dies $A^T = A^{-1}$ und man nennt dann A auch „orthogonal“.

Die unter (ii) und (iii) genannten Matrizen sind natürlich auch normal.

Die folgenden Aussagen lassen sich mit den schon bereitgestellten Mitteln leicht beweisen.

Lemma U.25 *(i) Ist A selbstadjungiert, so ist $\det A$ reell.*

(ii) A ist genau dann unitär, wenn die Spalten eine ONB bilden.

(iii) Ist A unitär, so ist $|\det A| = 1$.

2

Jede Matrix $F \in K^{n\times m}$ liefert über $f : x \mapsto Fx$ einen Homomorphismus $f \in \mathrm{Hom}(K^m, K^n)$. Insbesondere liefert die zu F adjungierte Matrix $F^H \in K^{m\times n}$ über $f^H : y \mapsto F^H y$ einen Homomorphismus $f^H \in \mathrm{Hom}(K^n, K^m)$. Den Zusammenhang von f und f^H wollen wir genauer studieren. Dabei werden sich wichtige Folgerungen über Eigenwerte und über die durch quadratische Formen gegebenen geometrischen Gebilde ergeben.

Bezeichnen $\langle\cdot,\cdot\rangle_n$ bzw. $\langle\cdot,\cdot\rangle_m$ die Standard-Skalarprodukte im $\mathbb{C}^n$ bzw. $\mathbb{C}^m$, so gilt: Für

$$x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_m \end{pmatrix}, \qquad y = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_m \end{pmatrix}, \qquad F = (\alpha_{ij})_{n,m}$$

ist

$$\begin{aligned} \langle y, Fx\rangle_n &= \sum_{i=1}^{n} \overline{\eta}_i \Big(\sum_{j=1}^{m} \alpha_{ij}\xi_i\Big) \\ &= \sum_{j=1}^{m} \Big(\overline{\sum_{i=1}^{n} \overline{\alpha_{ij}}\eta_i}\Big)\xi_j \\ &= \left\langle F^H y, x\right\rangle_m . \end{aligned}$$

Schreiben wir dies mit den eingeführten Homomorphismen f bzw. f^H, so haben wir: Für alle $x \in \mathbb{C}^n, y \in \mathbb{C}^m$ ist $\langle y, f(x)\rangle_n = \left\langle f^H(y), x\right\rangle_m$. Dieser Sachverhalt läßt sich verallgemeinern.

Satz und Definition U.26 (Hermitesch adjungierter Homomorphismus) *Es seien U, V zwei endlich-dimensionale K-Vektorräume über $K = \mathbb{R}$ oder $K = \mathbb{C}$ mit Skalarprodukten $\langle\cdot,\cdot\rangle_U$ bzw. $\langle\cdot,\cdot\rangle_V$, ferner $f \in \mathrm{Hom}(U,V)$. Dann gibt es genau einen Homomorphismus $f^H \in \mathrm{Hom}(V,U)$, sodaß für alle $u \in U, v \in V$ gilt: $\langle v, f(u)\rangle_V = \langle f^H(v), u\rangle_U$. f^H heißt der zu f (hermitesch) adjungierte Homomorphismus. Es gelten die Rechenregeln:*

$$\begin{aligned} \mathrm{id}_U^H &= \mathrm{id}_U, \\ (\alpha f + \beta g)^H &= \overline{\alpha} f^H + \overline{\beta} g^H \\ (fg)^H &= g^H f^H. \end{aligned}$$

Beweis: Es seien $(u_1, \ldots, u_m)$ bzw. $(v_1, \ldots, v_n)$ ONB in U bzw. V und damit $f(u_j) = \sum_{i=1}^n \alpha_{ij} v_i$ $(j = 1, \ldots, m)$. Dann ist also $F := (\alpha_{ij})_{n,m}$ Matrixdarstellung von f bezüglich dieser Basen. Wir definieren einen Homomorphismus $f^H \in \mathrm{Hom}(V,U)$ durch $f^H(v_i) := \sum_{j=1}^m \overline{\alpha}_{ij} u_j$ $(i = 1, \ldots, n)$. Er wird also bezüglich der verwendeten Basen durch die adjungierte Matrix F^H dargestellt.

Nun rechnet man nach:

$$\begin{aligned} \langle v_\nu, f(u_\mu)\rangle_V &= \left\langle v_\nu, \sum_{i=1}^n \alpha_{i\mu} v_\nu \right\rangle_V \\ &= \sum_{i=1}^n \alpha_{i\mu} \langle v_\nu, v_\mu)\rangle_V = \alpha_{\nu\mu} \\ \langle f^H(v_\nu), u_\mu\rangle_U &= \left\langle \sum_{j=1}^m \overline{\alpha}_{\nu j} u_j, u_\mu \right\rangle_U \\ &= \sum_{j=1}^m \alpha_{\nu j} \langle u_j, u_\mu)\rangle_U = \alpha_{\nu\mu}. \end{aligned}$$

Somit stimmen diese beiden Skalarprodukte je auf den Paaren von Basisvektoren überein, woraus direkt die Übereinstimmung für beliebige Elemente folgt. f^H besitzt also die geforderte Eigenschaft.

Die Eindeutigkeit folgt so: Gilt für $f' \in \mathrm{Hom}(V,U)$ ebenfalls $\langle v, f(u)\rangle_V = \langle f'(v), u\rangle_U$ für alle u und v, so auch

$$0 = \langle f^H(v), u\rangle_U - \langle f'(v), u\rangle_U = \langle f^H(v) - f'(v), u\rangle_U .$$

Hier kann man speziell $u := f^H(v) - f'(v)$ wählen, sodaß dann $f^H(v) = f'(v)$ folgt.

Damit ist Existenz und Eindeutigkeit des hermitesch adjungierten Homomorphismus gezeigt. Die Rechenregeln seien als Übung gestellt. □

Satz U.27 *Sind U, V endlich-dimensionale unitäre Räume, $f \in \mathrm{Hom}(U,V)$, $f^H \in \mathrm{Hom}(V,U)$ der adjungierte und sind B_U und B_V Orthonormalbasen in U bzw. V, so ist die Matrixdarstellung F^H von f^H bezüglich B_V, B_U die adjungierte Matrix zur Darstellung F von f bzgl. B_U, B_V.*

Im weiteren betrachten wir nun Endomorphismen eines unitären Raumes und insbesondere deren Eigenwerte und Eigenvektoren. Dabei beschränken wir uns auf sog. „normale“ Endomorphismen, die aus zwei Gründen interessant sind:

- Ihre Eigenwert-Theorie ist relativ übersichtlich und
- viele in den Anwendungen — etwa der Quantenmechanik — zu untersuchenden Endomorphismen sind normal.

Definition und Satz U.28 *Es sei U ein unitärer Raum, $f \in \mathrm{Hom}(U,U)$, $f^H \in \mathrm{Hom}(U,U)$ sein adjungierter. Dann heißt f „normal“, wenn*

$$f^H f = f f^H.$$

Ist sogar $f^H f = f f^H = \mathrm{id}_U$, d.h. $f^H = f^{-1}$, so heißt f „unitär“ (im reellen Fall „orthogonal“), ist $f^H = f$, so heißt f „selbstadjungiert“ oder „hermitesch“ (im reellen „symmetrisch“).

Gehen wir mittels einer ONB in U zu einer Matrixdarstellung über, so gilt (nach Satz U.27):

- *f normal $\Longleftrightarrow F^H F = F F^H$,*
- *f unitär $\Longleftrightarrow F^H F = F F^H = I \Longleftrightarrow F^H = F^{-1}$,*
- *f selbstadjungiert $\Longleftrightarrow F^H = F$.*

Wir sprechen dann von normalen, unitären, selbstadjungierten Matrizen.

Zum Studium der Eigenwerte und Eigenräume normaler Endomorphismen brauchen wir noch ein paar Vorarbeiten.

Lemma U.29 *Für $g \in \mathrm{Hom}(U,U)$ gelten:*

(i) $(\mathrm{im}\, g)^\perp = \ker g^H$.

(ii) Ist g normal, so ist $\ker g = \ker g^H$.

(iii) Ist g normal, so ist $(\mathrm{im}\, g)^\perp = \ker g$, also $U = \ker g \bigoplus \mathrm{im}\, g$.

Beweis:

(i) $y \in (\mathrm{im}\, g)^\perp \iff \forall_{x\in U}\ 0 = \langle y, g(x)\rangle = \left\langle g^H(y), x\right\rangle \iff g^H(y) = 0 \iff y \in \ker g^H$.

(ii) Es ist $\langle g(x), g(x)\rangle$ stets reell, somit für normales g

$$\begin{aligned} \langle g(x), g(x)\rangle &= \left\langle g^H g(x), x\right\rangle \\ &= \left\langle x, g^H g(x)\right\rangle \\ &= \left\langle x, g g^H(x)\right\rangle \\ &= \left\langle g^H(x), g^H(x)\right\rangle \end{aligned}$$

sodaß $g(x) = 0 \Longleftrightarrow g^H(x) = 0$, also die Kerne übereinstimmen.

(iii) Dies faßt nur (i) und (ii) zusammen. □

In Satz und Definition U.13 hatten wir orthogonale Projektoren über die Proximum-Bildung eingeführt. Die können wir nun folgendermaßen charakterisieren.

Satz U.30 *Für $P \in \mathrm{Hom}(U,U)$ sind äquivalent:*

(i) P ist orthogonaler Projektor (auf $V = \mathrm{im}\, P$),

(ii) $P^2 = P$ und $P^H = P$, d.h. P ist idempotent und selbstadjungiert.

Beweis:

(i) $\Longrightarrow$ (ii): Die ist schon mit Satz und Definition U.13 gezeigt.

(ii) $\Longrightarrow$ (i): Es sei also $P^2 = P$ und $P^H = P$. Nach Lemma U.29 ist $\ker P = (\mathrm{im}\, P)^\perp$ und $U = \mathrm{im}\, P \bigoplus \ker P$.

Zerlege $x \in U$ als $x = x_1 + x_2$ mit $x_1 \in \mathrm{im}\, P,\ x_2 \in \ker P$. Dann ist

$$Px = P(x_1 + x_2) = Px_1 + 0 = Px_1.$$

Da $x_1 \in \mathrm{im}\, P$ existiert ein $y \in U$, sodaß $x_1 = Py$, somit ist $Px_1 = P(Py) = Py = x_1$. Folglich ist $x - Px = (x_1 + x_2) - x_1 = x_2 \in \ker P = (\mathrm{im}\, P)^\perp$ und Px ist das Proximum an x, d.h. P orthogonaler Projektor. □

Satz U.31 *Es sei $f \in \mathrm{Hom}(U,U)$, normal, $\lambda \in \mathbb{C}$. Dann gelten:*

(i) $(f - \lambda\,\mathrm{id}_U)^H = f^H - \overline{\lambda}\,\mathrm{id}_U$ *und* $f - \lambda\,\mathrm{id}_U$ *ist normal.*

(ii) Es ist $f(x) = \lambda x$ *genau, wenn* $f^H(x) = \overline{\lambda}x$.

(iii) Eigenvektoren zu verschiedenen Eigenwerten von f sind orthogonal.

Beweis:

(i) Wegen der Rechenregeln aus Satz und Definition U.26 brauchen wir nur die Normalität zu beweisen:

$$\begin{aligned}(f - \lambda\,\mathrm{id}_U)(f - \lambda\,\mathrm{id}_U)^H &= (f - \lambda\,\mathrm{id}_U)(f^H - \overline{\lambda}\,\mathrm{id}_U)\\ &= ff^H - \lambda f^H - \overline{\lambda}f + \lambda\overline{\lambda}\,\mathrm{id}_U\\ &= (f - \lambda\,\mathrm{id}_U)^H(f - \lambda\,\mathrm{id}_U),\end{aligned}$$

da f selbst normal ist.

(ii) Dies ist die Aussage von Lemma U.29(ii) angewendet auf $g := f - \lambda\,\mathrm{id}_U$.

(iii) Sind $\lambda \neq \lambda'$ Eigenwerte von f mit Eigenvektoren x, x'. Dann ist

$$\lambda\,\langle x', x\rangle = \langle x', \lambda x\rangle = \langle x', f(x)\rangle = \langle f^H(x'), x\rangle = \left\langle \overline{\lambda}' x', x\right\rangle = \lambda'\,\langle x', x\rangle\,.$$

Da $\lambda \neq \lambda'$ ist, muß notwendig $\langle x', x\rangle = 0$ sein. □

Damit erhalten wir nun als zentrale Aussage über Eigenwerte und Eigenvektoren normaler Endomorphismen den folgenden

Satz U.32 (Spektralsatz) *U sei endlich-dimensionaler komplexer unitärer Raum, $\neq (0), f \in \mathrm{Hom}(U,U)$. Dann sind äquivalent:*

(i) f ist normal

(ii) f hat eine ONB aus Eigenvektoren.

(iii) Es gibt orthogonale Projektoren P_i und Zahlen λ_i $(i = 1, \ldots, m)$, sodaß

$$\begin{array}{lrcl}(\alpha) & \sum_{i=1}^m P_i &=& \mathrm{id}_U,\\ (\beta) & P_iP_j &=& 0 \quad (i \neq j)\\ (\gamma) & \sum_{i=1}^m \lambda_i P_i &=& f\end{array}$$

So eine Darstellung heißt „Spektraldarstellung“.

Beweis: (i) $\Rightarrow$ (ii): Als Endomorphismus in einem endlich-dimensionalen komplexen Raum $U(\neq (0))$ besitzt f mindestens einen Eigenwert λ, da das charakteristische Polynom mindestens eine Nullstelle hat. Es ist dann $x \neq 0$ genau dann Eigenvektor zu λ, wenn $x \in E(f;\lambda) := \ker(f - \lambda\,\mathrm{id}_U)$, dem Eigenraum von f zu λ. Dies ist ein endlich-dimensionaler Unterraum von $U, \neq 0$, der somit eine ONB besitzt. Betrachten wir nun alle — höchstens $\dim U$ vielen — Eigenräume zu den verschiedenen Eigenwerten von f und wählen für jeden eine ONB, so geben diese alle zusammen ein ONS $(x_1, \ldots, x_m)$ in U. Denn nach Konstruktion sind alle normiert und paarweise orthogonal, soweit sie zum selben Eigenwert gehören, und nach Satz U.31 orthogonal, sofern sie in verschiedenen Eigenräumen liegen. Wir setzen $V := \mathrm{span}(x_1, \ldots, x_m)$ und zeigen, daß $V = U$:

Die x_i sind Eigenvektoren von f, somit nach Satz U.31 auch Eigenvektoren von f^H, sodaß dann auch $f^H(V) \subset V$. Falls $V \subsetneq U$ (sonst ist ja nichts mehr zu beweisen,) bilden wir das orthogonale Komplement $V^\perp$, womit dann $U = V \oplus V^\perp$ ist. Für $v \in V, w \in V^\perp$ ist dann

$$\langle v, f(w)\rangle = \langle f^H(v), w\rangle = 0,$$

da $f^H(v) \in V, w \in V^\perp$. Da hier v, w sonst beliebig gewählt werden können, ist also $f(w) \in V^\perp$ für $w \in V^\perp$. Also bildet f den Unterraum $V^\perp$ in sich selbst ab und (s. oben) die Einschränkung $f_{|V^\perp}$ hat somit einen Eigenwert λ_0 mit zugehörigem Eigenvektor $x_0 \in V^\perp$. x_0 gehört damit zu den Eigenvektoren von f und liegt damit - nach Konstruktion von V - in dem Vektorraum V, was aber $x_0 \in V \cap V^\perp$ also $x_0 = 0$ impliziert, obwohl x_0 als Eigenvektor $\neq 0$ ist. Also führt die Annahme, daß $V \subsetneq U$ zum Widerspruch, d.h. es ist $V = U$.

(ii) $\Rightarrow$ (iii) : Es sei $(x_1, \ldots, x_n)$ eine ONB aus Eigenvektoren. Wir haben damit als Eigenräume $V_i := E(f; \lambda_i) = \mathrm{span}(x_j | x_j$ Eigenvektor zum Eigenwert $\lambda_i)$.

Haben wir genau m verschiedene Eigenwerte, so folgt trivialerweise $V_i \perp V_j$ für $i, j \leq m, i \neq j$, d.h. die Elemente aus verschiedenen Eigenräumen sind paarweise orthogonal, und somit auch

$$U = V_1 \oplus \cdots \oplus V_m$$

Die orthogonalen Projektoren P_i auf dem jeweiligen Raum V erfüllen dann (iii):

Zu α : Sei $U \ni u = v_1 + \cdots + v_m$ entsprechend den V_i zerlegt, d.h. $v_i \in V_i, v_i \in V_j^\perp$ für $j \neq i$, so ist $P_i v_i = v_i, P_i v_j = 0$ $(i \neq j)$ und damit

$$\begin{aligned}\sum_{i=1}^m P_i u &= \textstyle\sum_{i=1}^m P_i(\sum_{j=1}^m v_j) = \sum_{i,j=1}^m P_i v_j \\ &= \textstyle\sum_{i=1}^m P_i v_i = \sum_{i=1}^m v_i = u.\end{aligned}$$

Zu β: $P_i P_j u = P_i v_j = 0$ $(i \neq j)$.

Zu γ: $\sum_{i=1}^m \lambda_i P_i u = \sum_{i=1}^m \lambda_i v_i = \sum_{i=1}^m f(v_i) = f(\sum_{i=1}^m v_i) = f(u)$.

(iii) $\Rightarrow$ (i) : Nach den Rechenregeln für adjungierte ist

$$f^H = (\sum_{i=1}^m \lambda_i P_i)^H = \sum_{i=1}^m \overline{\lambda}_j P_j^H = \sum_{i=1}^m \overline{\lambda}_i P_i,$$

da die Projektoren selbstadjungiert sind. Dann ist

$$f^H f = (\sum_{i=1}^m \overline{\lambda}_i P_i)(\sum_{j=1}^m \lambda_i P_i) = \sum_{i,j=1}^m \overline{\lambda}_i \lambda_j P_i P_j = \sum_{i=1}^m |\lambda|^2 P_i$$

und ebenso erhält man diese Darstellung für ff^H. □

Deuten wir dies für Matrizen: Eine normale Matrix F besitzt also eine ONB $(x_1, \ldots, x_n)$ aus Eigenvektoren, ist folglich diagonalisierbar. Stellen wir F bezüglich dieser Basis dar, so ist also, wenn die Matrix X gerade die x_j als Spalten hat.

$$\begin{aligned}D := X^{-1}FX &= X^{-1}(Fx_1, \ldots, Fx_n) = X^{-1}(\lambda_1 x_1, \ldots, \lambda_n x_n) \\ &= (\lambda_1 X^{-1}x_1, \ldots, \lambda_n X^{-1}x_n) = (\lambda_1 e_1, \ldots, \lambda_n e_n)\end{aligned}$$

gerade diese Diagonalmatrix.

Die Projektoren P_i werden bezüglich dieser Basis ebenfalls durch Diagonalmatrizen dargestellt, wobei alle Diagonalelemente 0 oder 1 sind und in dem Projektor zum Eigenraum $E(f; \lambda)$ genau die Diagonalelemente gleich 1 sind, für die in D gerade $\lambda_i = \lambda$ gilt.

Für weitere Anwendungen brauchen wir zunächst

Lemma U.33 *Sind* $(P_1, \ldots, P_m)$ *orthogonale Projektoren mit* $\sum_{i=1}^m P_i = \mathrm{id}_U$, $P_iP_j = 0$ $(i \neq j)$ *und sind mit* $\lambda_i, \gamma_i \in \mathbb{C}$ $\quad f = \sum_i \lambda_i P_i$, $g = \sum_j \gamma_j P_j$, *so ist* $fg = gf = \sum_{i=1}^m \lambda_i \gamma_i P_i$.

Beweis:

$$\begin{aligned} fg &= (\sum_i \lambda_i P_i)(\sum_j \gamma_j P_j) \\ &= \sum_{i,j} \lambda_i \gamma_j P_i P_j = \sum_i \lambda_i \gamma_i P_i \end{aligned}$$

und die selbe Darstellung ergibt sich natürlich für gf. □

Dies liefert uns den

Satz U.34 *Es seien* $(P_1, \ldots, P_m)$ *orthogonale Projektoren in* U *mit* $\sum_{i=1}^m P_i = \mathrm{id}_U$, $P_iP_j = 0$ $(i \neq j)$, *ferner* $f = \sum_{i=1}^m \lambda_i P_i$ *mit Koeffizienten* $\lambda_i \in \mathbb{C}$. *Dann gelten:*

(i) Sind alle $\lambda_i \neq 0$, *so existiert* f^{-1} *und hat die Darstellung*

$$f^{-1} = \sum_{i=1}^m \frac{1}{\lambda_i} P_i.$$

(ii) Ist $p(T) = \sum_{j=0}^k \alpha_j T^j \in \mathbb{C}[T]$ *ein Polynom mit komplexen Koeffizienten, so ist*

$$p(f) := \sum_{j=0}^k \alpha_j f^j = \sum_{i=1}^m p(\lambda_i) P_i.$$

Sind dabei die Zahlen $p(\lambda_i) \neq 0$ *für alle* i, *so ist* $p(f)$ *invertierbar und*

$$(p(f))^{-1} = \sum_{i=1}^m \frac{1}{p(\lambda_i)} P_i.$$

(iii) Sind p, q *zwei solche Polynome, dabei* $q(\lambda_i) \neq 0$ *für alle* i, *so existiert* $(q(f))^{-1}$, *ist mit* $p(f)$ *vertauschbar und mit der rationalen Funktion* $r(T) := \frac{p(T)}{q(T)}$ *ist*

$$r(f) := p(f)(q(f))^{-1} = (q(f))^{-1} p(f) = \sum_{i=1}^m r(\lambda_i) P_i.$$

(iv) Einen Ausblick in weitere Anwendungen gibt etwa

$$e^{if} := \sum_{j=1}^m \exp(i\lambda_j) P_j.$$

Dies ist ein unitärer Homomorphismus, sofern alle λ_j *reell sind.* ($\exp(i\lambda)$ *bezeichnet die komplexe Exponentialfunktion!)*

Beweis:

(i) $g := \sum_{i=1}^m \frac{1}{\lambda_i} P_i$ ist wohldefiniert und mit Lemma U.33 bekommen wir $fg = gf = \sum_{i=1}^m \lambda_i \frac{1}{\lambda_i} P_i = \mathrm{id}_U$.

(ii) Induktive Anwendung von Lemma U.33 liefert zunächst $f^j = \sum_{i=1}^m \lambda_i^j P_i$ und damit auch die behauptete Darstellung für $p(f)$. Für die weitere Aussage wende man(i) an

(iii) Stelle $p(f)$ und $(q(f))^{-1}$ nach (ii) dar und wende Lemma U.33 an.

(iv) $e^{if} = \sum_{j=1}^m \exp(i\lambda_j)P_j$ ist jedenfalls wohldefiniert (und normal). Ferner ist nach den Rechenregeln für adjungierte bei *reellen* λ_j

$$\begin{aligned}(e^{if})^H = \sum_{j=1}^m \overline{\exp(i\lambda_j)}P_j &= \sum_{j=1}^m \exp(-i\lambda_j)P_j \\ &= \sum_{j=1}^m \frac{1}{\exp(i\lambda_j)}P_j = (e^{if})^{-1},\end{aligned}$$

sodaß ein unitärer Homomorphismus vorliegt. □

Betrachten wir noch einige wichtige Spezialfälle normaler Homomorphismen.

Nach Definition und Satz U.28 heißt ein Endomorphismus selbstadjungiert ($K = \mathbb{C}$) bzw. symmetrisch ($K = \mathbb{R}$), wenn $f^H = f$, bzw. für Matrizen $F^H = F$. Selbstadjungierte Endomorphismen sind insbesondere normal, sodaß die bisherigen Sätze wieder gelten. Darüber hinaus erhalten wir

Satz U.35 *(i) Alle Eigenwerte eines selbstadjungierten Endomorphismus bzw. einer selbstadjungierten Matrix sind reell.*

(ii) Für den Fall $K = \mathbb{R}$ (er ist im Spektralsatz nicht erfaßt!) gilt: Jede symmetrische (reelle) Matrix F besitzt eine ONB aus Eigenvektoren im $\mathbb{R}^n$. Sie läßt sich mittels einer (reellen) orthogonalen Matrix X auf Diagonalform D transformieren:

$$X^{-1}FX = X^HFX = D.$$

Beweis:

(i) : Ist x ($\neq 0$) Eigenvektor zum Eigenwert λ des selbstadjungierten Endomorphismus f, so gilt

$$\lambda \langle x,x\rangle = \langle x,\lambda x\rangle = \langle x,f(x)\rangle = \left\langle f^H(x),x\right\rangle = \langle f(x),x\rangle = \langle \lambda x,x\rangle = \overline{\lambda}\langle x,x\rangle\,.$$

Wegen $\langle x,x\rangle \neq 0$ muß λ reell sein.
(Man hätte auch eleganter über die Spektraldarstellung argumentieren können.)

(ii) : Jede reelle symmetrische Matrix F definiert einen selbstadjungierten Endomorphismus f im $\mathbb{C}^n$. Dieser besitzt eine ONB $(x_1,\dots,x_n)$ des $\mathbb{C}^n$ aus Eigenvektoren zu den sämtlich *reellen* Eigenwerten λ_i ($i = 1,\dots,n$). Nun ist x Eigenvektor zu λ genau wenn

$$x \in \ker(F - \lambda I) \subset \mathbb{C}^n.$$

Somit gibt es zu λ genau

$$\dim\ker(F - \lambda I) = n - \mathrm{rg}(F - \lambda I)$$

viele unabhängige Eigenvektoren. Da F und λ reell sind, können wir auch den Kern von $(F - \lambda I)$ im $\mathbb{R}^n$ bestimmen und der hat (über $\mathbb{R}$) die selbe Dimension. Wählen wir also für jeden solchen reellen Kern, d.h. reellen Eigenraum eine ONB, so ergeben alle diese zusammen eine (reelle) ONB des $\mathbb{R}^n$ aus Eigenvektoren von F. □

Als weiteren Spezialfall betrachten wir die unitären ($\mathbb{C}$) bzw. orthogonalen ($\mathbb{R}$) Endomorphismen, für die $f^H = f^{-1}$ bzw. $f^Hf = ff^H = \mathrm{id}$ ist.

Satz U.36 *Es sei U ein n-dimensionaler (reeller oder komplexer) unitärer Raum und $f \in \mathrm{Hom}(U,U)$. Dann sind äquivalent:*

(i) f ist unitär bzw. orthogonal.

(ii) *Ist* $(x_1, \dots x_n)$ *eine ONB, so auch* $(f(x), \dots, f(x_n))$.

(iii) *Es gibt eine ONB* $(x_1, \dots, x_n)$, *für die auch* $(f(x_1), \dots, f(x_n))$ *eine ONB ist.*

(iv) f *ist isometrisch, d.h. für alle* $x \in U$ *ist* $|f(x)| = |x|$

(v) *Für alle* $x, y \in U$ *ist* $\langle f(x), f(y)\rangle = \langle x, y\rangle$.

(vi) f *ist normal und alle Eigenwerte sind vom Betrag* $= 1$.

Beweis: (i) $\Rightarrow$ (ii): Ist f unitär oder orthogonal, so gilt auf einer ONB:

$$\langle f(x_i), f(x_j)\rangle = \left\langle f^H f(x_i), x_j\right\rangle = \langle x_i, x_j\rangle = \delta_{ij}$$

sodaß auch die Bilder eine ONB bilden.

(ii) $\Rightarrow$ (iii): Was in (ii) für jede ONB gefordert wird, braucht in (iii) nur für eine zu gelten.

(iii) $\Rightarrow$ (iv): Sei $(x_1, \dots, x_n)$ eine ONB, die unter f wieder in eine ONB überführt wird. Dann ist für $x = \sum_{i=1}^n \alpha_i x_i$

$$\begin{aligned}
|f(x)|^2 &= \left\langle \sum_{i=1}^n \alpha_i f(x_i), \sum_{j=1}^n \alpha_j f(x_j)\right\rangle \\
&= \sum_{i,j} \overline{\alpha}_i \alpha_j \langle f(x_i), f(x_j)\rangle \\
&= \sum_{i,j} \overline{\alpha}_i \alpha_j \langle x_i, x_j\rangle \\
&= \left\langle \sum_i \alpha_i x_i, \sum_j \alpha_j x_j \right\rangle = |x|^2.
\end{aligned}$$

(iv) $\Rightarrow$ (v): Durch Nachrechnen verifiziere man die Identitäten

– reell: $|x+y|^2 - |x-y|^2 = 4\langle x, y\rangle$

– komplex: (i=imaginäre Einheit)

$$|x+y|^2 - |x-y|^2 + i(|x+iy|^2 - |x-iy|^2) = 4\langle x, y\rangle\,,$$

sodaß f mit dem Erhalten der Norm auch das Skalarprodukt unverändert läßt.

(v) $\Rightarrow$ (i): Für alle x, y folgt aus (v): $\langle x, y\rangle = \langle f(x), f(y)\rangle = \left\langle f^H f(x), y\right\rangle$, sodaß stets $f^H f(x) = x$, d.h. $f^H f = \mathrm{id}_U$ und damit hier $f^H = f^{-1}$ folgt. Also ist f unitär.

Bleibt noch die Äquivalenz von (vi) mit den übrigen Aussagen zu zeigen: Ist f unitär, so auch normal und nach (iv) folgt für Eigenvektoren: $|x| = |f(x)| = |\lambda x| = |\lambda| \cdot |x|$, sodaß $|\lambda| = 1$ sein muß.

Setzen wir (vi) voraus, so ist f normal, hat also eine ONB $(x_1, \dots, x_n)$ aus Eigenvektoren, wobei sogar alle Eigenwerte λ_i vom Betrag 1 sind. Dann ist

$$\begin{aligned}
\langle f(x_i), f(x_j)\rangle &= \langle \lambda_i x_i, \lambda_j x_j\rangle \\
&= \overline{\lambda}_i \lambda_j \langle x_i, x_j\rangle \\
&= \begin{cases} 0, & i \neq j, \\ |\lambda_i|^2, & i = j, \end{cases} \\
&= \langle x_i, x_j\rangle\,,
\end{aligned}$$

d.h. wir haben (iii) gezeigt und f ist also unitär. $\square$

Die Übersetzung dieses Sachverhaltes auf Matrizen liefert

Satz U.37 *Es sei $F \in \mathbb{C}^{n\times n}$ oder $\in \mathbb{R}^{n\times n}$. Dann sind äquivalent:*

(i) F ist unitär, d.h. $F^H = F^{-1}$ (in $\mathbb{C}$) bzw. $F^T = F^{-1}$ (in $\mathbb{R}$).

(ii) Ist $(x_1, \ldots, x_n)$ ONB, so auch $(Fx_1, \ldots, Fx_n)$.

(iii) Die Spalten von F bilden eine ONB.

(iv) F ist isometrisch, d.h. für alle x ist $|Fx| = |x|$.

(v) Es ist stets $\langle Fx, Fy\rangle = \langle x, y\rangle$.

(vi) F ist normal und alle Eigenwerte haben Betrag $= 1$.

Der Beweis folgt direkt aus dem vorigen Satz. Bei (iii) wurde benutzt, daß die kanonischen Einheitsvektoren eine ONB bilden, ferner gilt ja bekanntlich: „Die Spalten sind …“

Betrachten wir noch den reellen Fall, d.h. orthogonale Matrizen $F \in \mathbb{R}^{n\times n}$. Der Endomorphismus $\mathbb{C}^n \longrightarrow \mathbb{C}^n, x \longmapsto Fx$ ist unitär, besitzt also eine Basis aus Eigenvektoren zu Eigenwerten, die sämtlich vom Betrag 1 sind. Für reelle Eigenwerte, die sind dann $+1$ oder -1, können wir wie bei den symmetrischen Matrizen auf reelle Eigenvektoren schließen.

Echt komplexe Eigenwerte haben die Form $\lambda = e^{i\varphi} = \cos\varphi + i\sin\varphi$ mit $\varphi \neq k\pi$, d.h. $\sin\varphi \neq 0$.

Sei dazu $x := u - iv$ mit $u, v \in \mathbb{R}^n$ ein Eigenvektor. Dann ist also

$$Fx = Fu - iFv = e^{i\varphi}(u - iv) = e^{i\varphi}x.$$

Da F, u, v alle reell sind, ist dann mit $\overline{x} := u + iv$ auch

$$F\overline{x} = \overline{Fx} = \overline{e^{i\varphi}(u - iv)} = e^{-i\varphi}(u + iv) = e^{-i\varphi}\overline{x}$$

d.h. $\overline{x}$ ist Eigenwert von F zum konjugiert komplexen Eigenwert. Da $\sin\varphi \neq 0$ ist, sind $e^{i\varphi}$ und $e^{-i\varphi}$ verschiedene Eigenwerte, also die Eigenvektoren x und $\overline{x}$ orthogonal. Damit ist

$$0 = \langle x, \overline{x}\rangle = \langle u - iv, u + iv\rangle = \langle u, u\rangle - \langle v, v\rangle + i(\langle u, v\rangle + \langle v, u\rangle).$$

Da hier alle notierten Skalarprodukte reelle Werte haben, ist insbesondere

$$\langle u, u\rangle = \langle v, v\rangle \quad \text{und} \quad \langle u, v\rangle = \langle v, u\rangle = 0.$$

Es sind also u und v orthogonal und von der selben Norm, die wir Œ als $= 1$ wählen können. Rechnet man noch $e^{i\varphi}(u - iv)$ nach Real- und Imaginärteil aus, so folgt:

Zu jedem Eigenwert $e^{i\varphi} \neq \pm 1$ ist auch $e^{-i\varphi}$ Eigenwert, und zu solchem Paar gibt es Vektoren $u, v \in \mathbb{R}^n$ mit $|u| = |v| = 1, \langle u, v\rangle = 0$, sodaß

$$\begin{aligned} Fu &= \cos\varphi\, u + \sin\varphi\, v \\ Fv &= -\sin\varphi\, u + \cos\varphi\, v \end{aligned}$$

d.h.

$$F(u|v) = (u|v)\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$

F bewirkt also in der von u und v aufgespannten (reellen) Ebene eine Drehung um den Winkel φ.

Nun kann man folgendermaßen eine reelle ONB zu F gewinnen: Zunächst wähle eine ONB zu dem Eigenraum zu $+1$, dann eine zum Eigenraum zu -1. Wie bei symmetrischen Matrizen kann man hier reelle Eigenvektoren wählen. Dann wähle für jeden Eigenraum zu einem komplexen Eigenwert mit positivem Imaginärteil eine (komplexe) ONB. Nach dem eben Gezeigten sind dann die jeweiligen konjugiert komplexen Vektoren Eigenvektoren zu den konjugiert komplexen Eigenwerten und jedes solche Paar konjugierter $x, \overline{x}$ Vektoren liefert eine solche reelle ONB (u, v) für eine reelle Ebene, in der F als Drehung wirkt. Wegen $\mathrm{span}(x, \overline{x}) = \mathrm{span}(u, v)$ erhalten wir so insgesamt eine ONB des $\mathbb{R}^n$. Wir haben also

Satz U.38 *Zu jeder orthogonalen Matrix $F \in \mathbb{R}^{n\times n}$ gibt es eine ONB $(x_1, \ldots, x_n)$ im $\mathbb{R}^n$, bezüglich der F die folgende Gestalt hat:*

$$X^H F X = \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & 0 & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & 0 & & & & R_1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & R_s \end{pmatrix}$$

wobei die Matrix X die Spalten x_i hat und die R_σ 2×2-Matrizen der Form

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

sind. Dies bedeutet, daß der $\mathbb{R}^n$ zerlegt wird in eine direkte (orthogonale) Summe

$$\mathbb{R}^n = V^+ \oplus V^- \oplus V_1 \oplus \cdots \oplus V_s,$$

wobei V^+, V^- die Eigenräume von F zu $+1$ bzw. -1 sind – hierauf wirkt F als Identität bzw. als Punktspiegelung – und in 2-dimensionale Räume V_σ, auf denen F als Drehung wirkt.

Die hier behandelten Typen von Endomorphismen oder Matrizen bilden teilweise Gruppen bezüglich der Komposition bzw. der Matrix-Multiplikation; d.h. mit je zweien ist auch das Produkt vom selben Typ, ferner existiert die Inverse und ist ebenfalls vom selben Typ.

Wir definieren hier nur die wichtigsten dieser Gruppen und lassen das Nachrechnen der Gruppeneigenschaften als Übung.

Definition U.39 *Die allgemeine lineare Gruppe*

$$\mathrm{GL}(n) := \left\{ F \in K^{n\times n} \,\middle|\, \det F \neq 0 \right\}$$

Die orthogonale Gruppe

$$\mathrm{O}(n) := \left\{ F \in \mathbb{R}^{n\times n} \,\middle|\, F \text{ orthogonal} \right\}$$

Die unitäre Gruppe

$$\mathrm{U}(n) := \left\{ F \in \mathbb{C}^{n\times n} \,\middle|\, F \text{ unitär} \right\}.$$

Dazu betrachtet man jeweils noch die „speziellen Gruppen", $\mathrm{SL}(n)$, $\mathrm{SO}(n)$, $\mathrm{SU}(n)$ die aus den invertierbaren, orthogonalen, unitären Matrizen mit Determinante $= +1$ bestehen.

3

Unsere Ergebnisse über selbstadjungierte Endomorphismen wollen wir nun noch anwenden, um einiges über symmetrische Bilinearformen, bzw. quadratische Formen zu erfahren. Insbesondere erhalten wir daraus eine Klassifizierung der Kurven bzw. Flächen zweiter Ordnung.

Ein selbstadjungierter Endomorphismus f eines endlich-dimensionalen unitären Raumes U besitzt ja eine ONB aus Eigenvektoren, die sämtlich zu *reellen* Eigenwerten gehören.

Definition U.40 (Signatur) *Sei f selbstadjungiert, $\in \operatorname{Hom}(U,U)$, V_+, bzw. V_-, bzw. V_0 sei der von allen Eigenvektoren zu positiven bzw. negativen Eigenwerten bzw. zum Eigenwert 0 aufgespannte Unterraum. Dann heißt das Zahlentripel (k_+, k_0, k_-) die Signatur von f, wobei $k_+ := \dim V_+,\ \ k_0 := \dim V_0,\ \ k_- := \dim V_-$. Es ist $k_+ + k_0 + k_- = \dim U$.*

Für selbstadjungiertes f ist für alle x : $\langle x, f(x)\rangle = \left\langle f^H(x), x\right\rangle = \langle f(x), x\rangle = \overline{\langle x, f(x)\rangle}$, d.h. $\langle x, f(x)\rangle$ stets reell. Zudem ist, wie man sofort nachrechnet: $\langle x, f(x)\rangle > 0$ für $x \in V_+, x \neq 0$, ferner $= 0$ für $x \in V_0$ und < 0 für $x \in V_-, x \neq 0$.

Definition U.41 (Definit) *Sei f selbstadjungiert, $\in \operatorname{Hom}(U,U), V \subset U$ ein Unterraum. Dann heißt*

– *f auf V positiv definit: $\Leftrightarrow \forall x \in V,\ x \neq 0$ ist $\langle x, f(x)\rangle > 0$*

– *f auf V negativ definit: $\Leftrightarrow \forall x \in V,\ x \neq 0$ ist $\langle x, f(x)\rangle < 0$.*

Lassen wir ≥ 0 bzw. ≤ 0 zu, so reden wir von positiv bzw. negativ semidefinit. Ist kein Unterraum V spezifiziert, so ist $V := U$ gemeint.

Auf den Räumen V_+ bzw. V_- ist also f positiv bzw. negativ definit, auf $V_+ + V_0$ bzw. $V_- + V_0$ positiv bzw. negativ semidefinit. Diese Räume sind auch maximal unter dieser Bedingung:

Satz U.42 *Ist f positiv (negativ) definit auf V, so ist $\dim V \leq k_+$ $(\dim V \leq k_-)$. Ist f positiv (negativ) semidefinit auf V, so ist $\dim V \leq k_+ + k_0$ $(\dim V \leq k_- + k_0)$.*

Beweis: Sei f positiv definit auf V. Dann ist $\langle x, f(x)\rangle > 0$ für alle $x \in V, x \neq 0$. Ist $x \in V_0 + V_-$, so ist $\langle x, f(x)\rangle \leq 0$. Damit ist $V \cap (V_0 + V_-) = (0)$, somit $k_+ + k_0 + k_- = \dim U \geq \dim(V + V_0 + V_-) = \dim V + \dim V_0 + \dim V_- = \dim V + k_0 + k_-$, d.h. $\dim V \leq k_+$. □

Diese Definitionen und Aussagen übetragen sich sofort auf selbstadjungierte Matrizen $F = F^H \in \mathbb{C}^{n\times n}$, wenn wir den durch $x \longmapsto Fx$ definierten Endomorphismus von $\mathbb{C}^n$ – der mit dem Standard-Skalarprodukt versehen – betrachten. Wir sprechen dann von *positiv, negativ, (semi-)definiten Matrizen, bzw. von der Signatur der Matrix F.*

Satz U.43 (Trägheitssatz von Sylvester) *Ist $F \in \mathbb{C}^{n\times n}$ selbstadjungiert, $G \in \mathbb{C}^{n\times n}$ invertierbar, so ist auch $G^H F G$ selbstadjungiert und F und $G^H F G$ haben dieselbe Signatur.*

Beweis: Es ist $(G^H F G)^H = G^H F^H G^{HH} = G^H F G$, d.h. $G^H F G$ ist selbstadjungiert. Es seien (k_+, k_0, k_-) bzw. (k'_+, k'_0, k'_-) die Signaturen von F bzw. von $G^H F G$ und V_+, V_0, V_- bzw. V'_+, V'_0, V'_- die entsprechend Definition U.40 gebildeten Unterräume. Dann haben wir für $x \in V'_+, x \neq 0$:

$$0 < \left\langle x, G^H F G x\right\rangle = \langle Gx, FGx\rangle\,.$$

Somit ist F positiv definit auf $GV'_+ := \{y|$ Es existiert ein $x \in V'_+, y = Gx\}$. Nach Satz U.42 ist dann $\dim GV'_+ \leq k_+$ und da G invertierbar ist, ist $\dim GV'_+ = \dim V'_+$ also $k'_+ \leq k_+$. Nun ist $F = (G^{-1})^H (G^H F G) G^{-1}$ und mit dem selben Schluß wie eben erhalten wir $k_+ \leq k'_+$, sodaß beide gleich sein müssen. Analog verfährt man bei k_0, k_-. □

Eine weitere Anwendung von Satz U.42 liefert uns Abschätzungen der Eigenwerte selbstadjungierter Endormorphismen.

Satz U.44 (1. Separationssatz) *Es sei U ein n-dimensionaler unitärer Raum, $f \in \mathrm{Hom}(U, U)$ selbstadjungiert. Seine sämtlich reellen Eigenwerte seien entsprechend der Größe numeriert: $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$. P sei ein orthogonaler Projektor in U mit $m := \mathrm{rg}\, P < n$. Dann gilt: $f' := PfP \in \mathrm{Hom}(U, U)$ ist selbstadjungiert und mit $V := \mathrm{im}\, P$ ist $f'_{|V^\perp} = 0, f'(V) \subset V$ und für die Eigenwerte $\lambda'_1 \leq \cdots \leq \lambda'_m$ von $f'_{|V}$ gilt $\lambda_j \leq \lambda'_j \leq \lambda_{j+(n-m)}$ $(1 \leq j \leq m)$.*

Beweis: $(PfP)^H = P^H f^H P^H = PfP$, da f und P selbstadjungiert sind. Damit ist auch f' selbstadjungiert. Nach Satz und Definition U.13 ist $\ker P = V^\perp$, somit $f'_{|V^\perp} = 0$, und wegen $V = \mathrm{im}\, P$ trivialerweise $\mathrm{im}\, f' \subset V$. Damit ist $f'_{|V} \in \mathrm{Hom}(V, V)$ auch selbstadjungiert. Dazu sei $(x'_1, \ldots, x'_m)$ Basis aus Eigenvektoren zu Eigenwerten $\lambda'_1 \leq \cdots \leq \lambda'_m$. Ferner ist P als Projektor idempotent, somit $Px = x$ für $x \in V$. Dann gilt auf V auch

$$\langle x, f'(x) \rangle = \langle x, PfPx \rangle = \langle Px, fPx \rangle = \langle x, f(x) \rangle .$$

Zur Abschätzung der Eigenwerte wähle $j : \lambda \leq j \leq m$ und bilde die Unterräume von V:

$$V_j^+ := \mathrm{span}(x'_1, \ldots, x'_j), \quad V_j^- := \mathrm{span}(x'_{j+1}, \ldots, x'_m)$$

und die Homomorphismen

$$g := \lambda'_j \,\mathrm{id} - f, \quad g' := \lambda'_j \,\mathrm{id} - f'.$$

Für $1 \leq i \leq m$ ist

$$g'(x'_i) = \lambda'_j x'_i - f'(x'_i) = (\lambda'_j - \lambda'_i) x'_i.$$

Da die λ'_i der Größe nach geordnet sind, ist also

- g' positiv semidefinit auf V_j^+, und
- g' negativ semidefinit auf V_j^-.

Diese Räume liegen in V, sodaß darin

$$\langle x, g'(x) \rangle = \lambda_j \langle x, x \rangle - \langle x, f'(x) \rangle = \lambda_j \langle x, x \rangle - \langle x, f(x) \rangle = \langle x, g(x) \rangle .$$

Somit ist auch g positiv semidefinit auf V_j^+ und negativ semidefinit auf V_j^-, sodaß g nach Satz U.42 wenigstens $j = \dim V_j^+$ viele nichtnegative und wenigstens $m - j = \dim V_j^-$ viele nichtpositive Eigenwerte besitzt. Da die Eigenwerte von g gerade die $\lambda'_j - \lambda_i (i = 1, n)$ sind, folgen daraus direkt die behaupteten Abschätzungen. □

Für Matrizen und spezielle Projektoren erhalten wir daraus den

Satz U.45 (2. Separationssatz) *F sei eine selbstadjungierte $n \times n$-Matrix, H die daraus durch Streichen von Zeile k und Spalte k entstehende $(n-1) \times (n-1)$-Untermatrix. Dann gilt für die Eigenwerte $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ von F und $\lambda'_1 \leq \lambda'_2 \leq \cdots \leq \lambda'_{n-1}$ von H*

$$\lambda_1 \leq \lambda'_1 \leq \lambda_2 \leq \lambda'_2 \leq \cdots \leq \lambda_{n-1} \leq \lambda'_{n-1} \leq \lambda_n .$$

Beweis: Es sei $P : \mathbb{C}^n \longmapsto \mathbb{C}^n$ der Projektor, der genau die k-te Komponente annulliert

$$P : \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_{n-1} \\ \alpha_n \\ \alpha_{n+1} \\ \vdots \\ \alpha_n \end{pmatrix} \mapsto \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_{n-1} \\ 0 \\ \alpha_{n+1} \\ \vdots \\ \alpha_n \end{pmatrix},$$

$V := \operatorname{im} P$ sein Bildraum. Setze

$$f \in \operatorname{Hom}(\mathbb{C}^n, \mathbb{C}^n) : x \longmapsto Fx,$$
$$f' \in \operatorname{Hom}(\mathbb{C}^n, \mathbb{C}^n) : x \longmapsto PFPx.$$

Dann ist H gerade die Matrixdarstellung von $f'_{|V}$ bezüglich der V-Basis $(e_1, \ldots, e_{k-1}, e_{k+1}, \ldots, e_n)$. Auf diese Konstellation paßt der Satz U.44 der die Behauptung liefert. □

Bemerkung U.46 *Analoge Resultate erhält man natürlich, wenn man symmetrisch 2, 3 oder mehr Zeilen und Spalten streicht.*

Diese Separationssätze liefern auch ein schönes Kriterium dafür, wann eine selbstadjungierte Matrix positiv definit ist.

Eine *„Haupt-Unterdeterminante“* einer Matrix ist die Determinante eines *„Haupt-Minors“*, d.h. einer durch symmetrisches Streichen von Zeilen und Spalten entstehenden Untermatrix.

Satz U.47 *$F := (\alpha_{ij}) \in \mathbb{C}^{n\times n}$ sei selbstadjungiert. Dann gelten*

(i) Ist F positiv definit, so sind alle Haupt-Unterdeterminanten positiv.

(ii) Sind alle Haupt-Unterdeterminanten der speziellen Form $\det F_k = \det(\alpha_{ij})_{k,k}$ $(1 \le k \le n)$ positiv, so ist F positiv definit.

Bei (ii) werden gerade die symmetrisch liegenden Untermatrizen betrachtet, die die linke obere Ecke von F ausfüllen.

Beweis: Da eine selbstadjungierte Matrix ähnlich zu einer Diagonalmatrix ist und ähnliche Matrizen die selben Eigenwerte haben, ist die Determinante einer solchen Matrix das Produkt ihrer Eigenwerte.

Zu (i): *F sei positiv definit:* Nach Satz U.42 sind dann alle Eigenwerte von F positiv, also nach dem Satz U.45 auch alle Eigenwerte von $(n-1)\times(n-1)$-Haupt-Minoren und durch Iterieren dieses Schlusses auch alle Eigenwerte beliebiger Haupt-Minoren. Als Produkte positiver Eigenwerte sind dann auch deren Determinanten positiv.

Zu (ii): *Die Determinanten* $\det F_k = \det(\alpha_{ij})_{k,k}$ *sind positiv für* $1 \le k \le n$:

Wir zeigen, daß dann alle F_k nur positive Eigenwerte besitzen, also insbesondere $F = F_n$ selbst, wonach F positiv definit ist.

F_k habe die Eigenwerte $\lambda_1^{(k)} \le \cdots \le \lambda_k^{(k)}$.

$k = 1$: Es ist $F_1 = (\alpha_{11})$ mit $\alpha_{11} = \det F_1 = \lambda_1^1$ Wegen $\det F_1 > 0$ ist also $\lambda_1^1 > 0$. Somit sind alle Eigenwerte von F_1 positiv.

$k \Longrightarrow k+1$: Sei $0 < \lambda_1^{(k)}$, d.h. F_k habe nur positive Eigenwerte. F_k entsteht aus F_{k+1} durch Streichen der $(k+1)$-ten Zeile und Spalte. Also gilt mit dem zweiten Separationssatz

$$\lambda_1^{(k+1)} \le \lambda_1^{(k)} \le \lambda_2^{(k+1)} \le \ldots$$

Somit kann höchstens der kleinste Eigenwert $\lambda_1^{(k+1)}$ von F_{k+1} nicht positiv sein. Die Determinanten-Bedingung besagt aber

$$0 < \det F_{k+1} = \lambda_1^{(k+1)} \lambda_2^{(k+1)} \cdots \lambda_{k+1}^{(k+1)},$$

sodaß wegen $\lambda_i^{(k+1)} > 0$ $(i \geq 2)$ auch $\lambda_1^{(k+1)} > 0$ sein muß, was wir zeigen wollten. □

Der Begriff der positiv definiten Matrix gibt uns auch einen Überblick über alle möglichen Skalarprodukte auf dem $\mathbb{C}^n$.

Satz U.48 *$\langle \cdot, \cdot \rangle$ sei das Standard-Skalarprodukt auf $\mathbb{C}^n$, d.h. $\langle x, y \rangle = \sum_i \overline{\xi_i} \eta_i$. Dann gelten:*

(i) Ist $F = (\varphi_{ij}) \in \mathbb{C}^{n \times n}$ selbstadjungiert und positiv definit, so ist

$$(x, y) := \langle x, Fy \rangle$$

ein Skalarprodukt.

(ii) Ist $(\cdot, \cdot)$ ein Skalarprodukt auf $\mathbb{C}^n$, so bilde mit den kanonischen Einheitsvektoren e_i die Werte $\varphi_{ij} := (e_i, e_j)$ $(i, j = 1, \ldots, n)$. Dann ist $F = \varphi_{(ij)} \in \mathbb{C}^{n \times n}$ eine selbstadjungierte positiv definite Matrix und

$$(x, y) = \langle x, Fy \rangle .$$

Beweis:

(i): Man rechne die Eigenschaften S1, ..., S4 nach. Die Definitheit von F braucht man für S4.

(ii): Für $F := (\varphi_{ij})$ ist $F^H := (\varphi_{ij}^H)$ mit $\varphi_{ij}^H = \overline{\varphi_{ji}} = \overline{(e_j, e_i)} = (e_i, e_j) = \varphi_{ij}$.
Also ist $F = F^H$ d.h. selbstadjungiert.
Damit wird für $x = \sum_i \xi_i e_i, \ y = \sum_j \eta_j e_j$

$$\begin{aligned}
\langle x, Fy \rangle &= \left\langle \sum_i \xi_i e_i, \sum_j \eta_j F e_j \right\rangle \\
&= \left\langle \sum_i \xi_i e_i, \sum_{j,k} \eta_j \varphi_{kj} e_k \right\rangle \\
&= \sum_{i,j,k} \overline{\xi_i} \eta_j \varphi_{kj} \langle e_i, e_k \rangle \\
&= \sum_{i,j} \overline{\xi_i} \eta_j \varphi_{ij} \\
&= \sum_{i,j} \overline{\xi_i} \eta_j (e_i, e_j) \\
&= (\sum_i \xi_i e_i, \sum_j \eta_j e_j) = (x, y).
\end{aligned}$$

Somit wird das Skalarprodukt $(\cdot, \cdot)$ über F richtig dargestellt. Für $x = y$ beweist dies auch die Definitheit von F. □

4

Die Überlegungen dieses Kapitels haben noch eine schöne geometrische Anwendung: Die allgemeine Gleichung einer Fläche zweiten Grades im $\mathbb{R}^n$ lautet

$$\sum_{i=1}^{n} \alpha_{ii}\xi_i^2 + 2 \cdot \sum_{\substack{i,j=1\\ i<j}}^{n} \alpha_{ij}\xi_i\xi_j + 2 \cdot \sum_{i=1}^{n} \beta_i\xi_i + \gamma = 0, \tag{U.49}$$

mit $\alpha_{ij}, \beta_i, \gamma \in \mathbb{R}$.

Im Falle $n = 2$ ist dies die Gleichung für den allgemeinen Kegelschnitt, hier ist natürlich das Wort „Fläche“ etwas unpassend, für $n = 3$ bekommen wir Flächen, die sog. Quadriken. Setzen wir A: $= (\alpha_{ij})_{n,n}$, wobei $a_{ij} := a_{ji}$ für $i > j$ gesetzt sei,

$$b := \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}, \qquad x := \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix},$$

so kann man mit dem Standard-Skalarprodukt im $\mathbb{R}^n$ die Gleichung (U.49) kompakt schreiben als

$$\langle x, Ax \rangle + 2\langle x, b \rangle + \gamma = 0 \tag{U.50}$$

Da A symmetrisch ist (nach Konstruktion), gibt es nach Satz U.35 eine orthogonale Matrix G, d.h. $G^H = G^{-1}$, sodaß $G^H AG =: L$ Diagonalform hat. Wählen wir nun als neues Koordinatensystem gerade die durch die Spalten von G gegebene ONB in $\mathbb{R}^n$, so erhält jeder Vektor $x \in \mathbb{R}^n$ bezüglich der neuen Basis die Darstellung $x' = G^H x$, bzw. $x = Gx'$.

Damit wird (U.50) zu

$$\begin{aligned} 0 &= \langle Gx', AGx' \rangle + 2\langle Gx', b \rangle + \gamma \\ &= \left\langle x', G^H AGx' \right\rangle + 2\left\langle x', G^H b \right\rangle + \gamma \\ &= \langle x', Lx' \rangle + 2\left\langle x', G^H b \right\rangle + \gamma, \end{aligned}$$

d.h. mit $c := G^H b$:

$$0 = \langle x', Lx' \rangle + 2\langle x', c \rangle + \gamma. \tag{U.51}$$

Damit haben wir wieder eine Gleichung der Gestalt (U.50), wobei jetzt aber die Matrix L Diagonalform hat, d.h. in der ausgeschriebenen Form (U.49) der zweite Term mit den „gemischten“ Produkten $\xi_i\xi_j (i \neq j)$ verschwunden ist. Als nächstes versuchen wir den linearen Term durch eine Translation der Form $x' = x'' - x_0$ zu beseitigen. Dies liefert

$$\begin{aligned} 0 &= \langle x'' - x_0, L(x'' - x_0) \rangle + 2\langle x'' - x_0, c \rangle + \gamma \\ &= \langle x'', Lx'' \rangle - (\langle x_0, Lx'' \rangle + \langle x'', Lx_0 \rangle - 2\langle x'', c \rangle) + (\langle x_0, Lx_0 \rangle + \gamma). \end{aligned}$$

Da L symmetrisch ist, ist $\langle x_0, Lx'' \rangle = \langle x'', Lx_0 \rangle$ und wir erhalten mit

$$\gamma' := \langle x_0, Lx_0 \rangle + \gamma$$

$$0 = \langle x'', Lx'' \rangle - 2\langle x'', Lx_0 - c \rangle + \gamma'.$$

Damit der lineare Term verschwindet, muß $Lx_0 - c = 0$ sein, d.h. $c \in \operatorname{im} L$. Dies ist nun nicht notwendig der Fall. Wir können aber folgendes erreichen: Die Signatur von A bzw. L sei $(k_+, k_-, k_0) = (s, t, n - (s + t))$. Durch geeignetes Numerieren erhalten wir

$$L = \left(\begin{array}{ccc|ccc|c} \lambda_1 & & & & & & \\ & \ddots & & & 0 & & 0 \\ & & \lambda_s & & & & \\ \hline & & & \lambda_{s+1} & & & \\ & 0 & & & \ddots & & 0 \\ & & & & & \lambda_{s+t} & \\ \hline & 0 & & & 0 & & 0 \end{array}\right),$$

wobei $\lambda_i > 0$ $(i = 1, ..., s)$, $\lambda_i < 0$ $(i = s+1, ..., s+t)$ und schließlich $n-(s+t)$ Eigenwerte $= 0$.

Dann können wir zwar nicht unbedingt $Lx_0 = c = 0$ erreichen, aber doch sichern, daß die ersten $s+t$ Komponenten verschwinden. Damit haben wir: Gehen wir zunächst von der Standardbasis im $\mathbb{R}^n$ über zu einer ONB aus Eigenvektoren zu A und führen anschließend noch die oben beschriebene Translation aus, so bekommt (U.49) die Gestalt:

$$\langle x, Lx \rangle + 2 \langle x, c \rangle + \gamma = 0 \tag{U.52}$$

Dabei ist L Diagonalmatrix der Form

$$L = \operatorname{diag}(\lambda_1, \ldots, \lambda_s, \lambda_{s+1}, \ldots, \lambda_{s+t}, 0, \ldots, 0)$$

und $\lambda_1, \ldots, \lambda_s$ sind die positiven, $\lambda_{s+1}, \ldots, \lambda_{s+t}$ die negativen Eigenwerte von A. Ferner hat C die Gestalt $c = (0, \ldots, 0, \gamma_{s+t+1}, \ldots, \gamma_n)^T$.

Die Gleichung (U.52) können wir noch folgendermaßen kompakt schreiben: Setzen wir zu $x = (\xi_1, \ldots, \xi_n)^T$: $\hat{x} := (\xi_1, \ldots, \xi_n, 1)^T$ und

$$F := \begin{pmatrix} L & c \\ c^T & \gamma \end{pmatrix}_{(n+1)\times(n+1)}$$

$$= \left(\begin{array}{ccc|ccc|c} \lambda_1 & & & & & & \\ & \ddots & & & 0 & & 0 \\ & & \lambda_{s+t} & & & & \\ \hline & & & 0 & & & \gamma_{s+t+1} \\ & 0 & & & \ddots & & \vdots \\ & & & & & 0 & \gamma_n \\ \hline & 0 & & \gamma_{s+t+1} & \cdots & \gamma_n & \gamma \end{array}\right)_{(n+1)\times(n+1)},$$

so kann man das in der Form

$$\langle \hat{x}, F\hat{x} \rangle = 0 \tag{U.53}$$

zusammenfassen.

Bemerkung U.54 *Im Falle $c = 0$ spricht man von einer Kurve bzw. Fläche mit Zentrum, im Falle $c \neq 0$ von einer Kurve bzw. Fläche ohne Zentrum.*

Wir wollen nun die Fläche $n = 2, 3$ genauer studieren:

Im Falle $n = 2$ reden wir statt von Flächen von Kurven und erhalten die bekannten Kegelschnitte.

Die wichtigsten Spezialfälle für $n = 2$ sind:

$s = 2$:

Die Kurve hat ein Zentrum, also $c = 0, \lambda_1, \lambda_2 > 0$.

(U.52) lautet: $\lambda_1\xi_1^2 + \lambda_2\xi_2^2 + \gamma = 0$ oder nach Multiplikation mit $\alpha := -\frac{1}{\gamma}$

$$(\frac{\lambda_1}{-\gamma})\xi_1^2 + (\frac{\lambda_2}{-\gamma})\xi_2^2 = 1.$$

Für $\gamma > 0$ hat dies keine Lösung, für $\gamma < 0$ ist es die Gleichung einer ***Ellipse*** mit den Halbachsen $a_1 := \sqrt{|\frac{\gamma}{\lambda_1}|}$, $a_2 := \sqrt{|\frac{\gamma}{\lambda_2}|}$. Für $\lambda_1 = \lambda_2$ gehört hierzu auch der *Kreis.*

$s = 1$, $t = 1$:

Die Kurve hat ein Zentrum, also $c = 0, \lambda_1 > 0, \lambda_2 < 0$. Gleichung (U.52) lautet: $|\lambda_1|\xi_1^2 - |\lambda_2|\xi_2^2 + \gamma = 0$ und dies liefert für $\gamma \neq 0$ analog oben eine *Hyperbel.*

$s = 1$, $t = 0$:

$c = \binom{0}{\gamma_2}$ mit $\gamma_2 \neq 0$. Die Kurve hat kein Zentrum.

Gleichung (U.52) lautet: $\lambda_1\xi_1^2 + 2\gamma_2\xi_2 + \gamma = 0$ und dies ist die Gleichung einer *Parabel.*

Die weiteren Fälle liefern entweder die schon behandelten Kurven (etwa $s = 0$, $t = 2$) oder Ausartungen zu Geraden oder Punkten. Die Details seien dem Leser überlassen.

Spezialfälle für $n = 3$:

$s = 3$:

$\lambda_1, \lambda_2, \lambda_3 > 0$, die Fläche hat ein Zentrum, d.h. $c = 0$. Die Gleichung (U.52) lautet: $\lambda_1\xi_1^2 + \lambda_2\xi_2^2 + \lambda_3\xi_3^2 + \gamma = 0$.

Für $\gamma > 0$ hat dies keine Lösung, für $\gamma \leq 0$ ist es die Gleichung eines ***Ellipsoids.*** Für den Fall, daß zwei oder gar alle drei Eigenwerte gleich sind, ergibt dies das *Rotationsellipsoid* bzw. die *Kugel.* Das Ellipsoid ist ein beschränktes Gebilde und der Schnitt mit jeder Ebene ist eine Ellipse (oder leer).

Im folgenden seien die wichtigsten Fälle tabellarisch aufgeführt, wobei wir die Matrix F aus (U.53) benutzen und dabei noch einige einfache Normierungen vornehmen. Die μ_i bezeichnen positive Zahlen. Die weiteren Fälle liefern außer Entartungen wie Ebenen, Geraden, Punkte, keine wesentlichen neuen Flächen und seien deshalb übergangen.

5

$$F = \begin{pmatrix} + & & & \\ & + & & \\ & & + & \\ & & & + \end{pmatrix}$$
(Leere Menge)
$$\mu_1\xi_1^2 + \mu_2\xi_2^2 + \mu_3\xi_3^2 + 1 = 0$$

$$F = \begin{pmatrix} + & & & \\ & + & & \\ & & + & \\ & & & - \end{pmatrix}$$
(Ellipsoid)
$$\mu_1\xi_1^2 + \mu_2\xi_2^2 + \mu_3\xi_3^2 - 1 = 0 \tag{1}$$

$$F = \begin{pmatrix} + & & & \\ & - & & \\ & & - & \\ & & & + \end{pmatrix}, \begin{pmatrix} + & & & \\ & + & & \\ & & - & \\ & & & - \end{pmatrix}$$
(Einschaliges Hyperboloid)
$$\mu_1\xi_1^2 + \mu_2\xi_2^2 - \mu_3\xi_3^2 - 1 = 0 \tag{2}$$

$$F = \begin{pmatrix} + & & & \\ & + & & \\ & & - & \\ & & & + \end{pmatrix}, \begin{pmatrix} + & & & \\ & - & & \\ & & - & \\ & & & - \end{pmatrix}$$
(Zweischaliges Hyperboloid)
$$\mu_1\xi_1^2 - \mu_2\xi_2^2 - \mu_3\xi_3^2 - 1 = 0 \tag{3}$$

$$F = \begin{pmatrix} + & & & \\ & + & & \\ & & 0 & - \\ & & - & 0 \end{pmatrix}$$
(Elliptisches Paraboloid)
$$\mu_1\xi_1^2 + \mu_2\xi_2^2 = 2\xi_3 \tag{4}$$

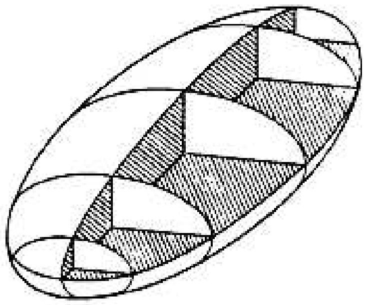
$$F = \begin{pmatrix} + & & & \\ & - & & \\ & & 0 & - \\ & & - & 0 \end{pmatrix}$$
(Hyperbolisches Paraboloid)
$$\mu_1\xi_1^2 - \mu_2\xi_2^2 = 2\xi_3 \tag{5}$$

$$F = \begin{pmatrix} + & & & \\ & + & & \\ & & - & \\ & & & 0 \end{pmatrix}, \begin{pmatrix} + & & & \\ & - & & \\ & & - & \\ & & & 0 \end{pmatrix}$$
(Elliptischer Doppelkegel)
$$\mu_1\xi_1^2 + \mu_2\xi_2^2 = \mu_3\xi_3^2 \tag{6}$$

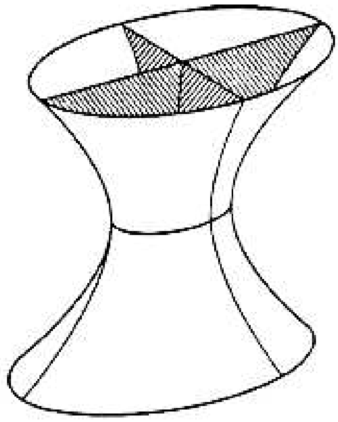
$$F = \begin{pmatrix} + & & & \\ & + & & \\ & & 0 & \\ & & & - \end{pmatrix}$$
(Elliptischer Zylinder)
$$\mu_1\xi_1^2 + \mu_2\xi_2^2 = 1 \tag{7}$$

$$F = \begin{pmatrix} + & & & \\ & - & & \\ & & 0 & \\ & & & - \end{pmatrix}, \begin{pmatrix} + & & & \\ & - & & \\ & & 0 & \\ & & & + \end{pmatrix}$$
(Hyperbolischer Zylinder)
$$\mu_1\xi_1^2 - \mu_2\xi_2^2 = 1 \tag{8}$$

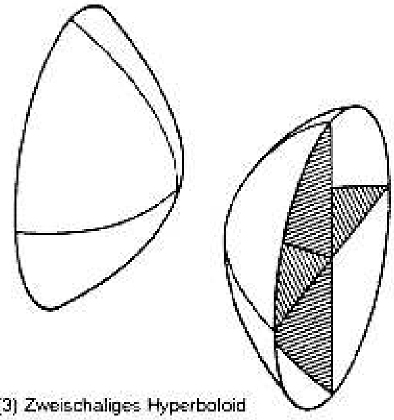
$$F = \begin{pmatrix} + & & & \\ & 0 & & - \\ & & 0 & \\ & - & & 0 \end{pmatrix}$$
(Parabolischer Zylinder)
$$\mu_1\xi_1^2 = 2\xi_2 \tag{9}$$



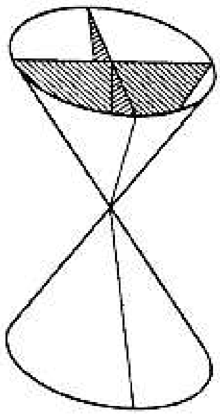
(1) Ellipsoid mit einigen Schnitten



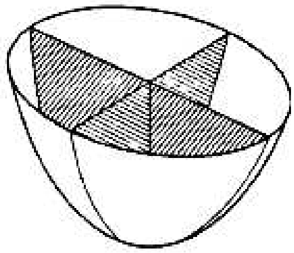
(2) Einschaliges Hyperboloid



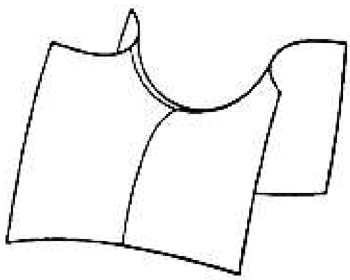
(3) Zweischaliges Hyperboloid



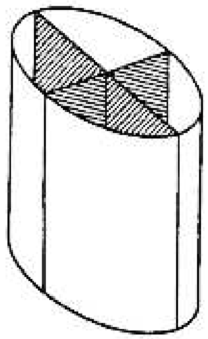
(6) Doppelkegel (elliptisch)



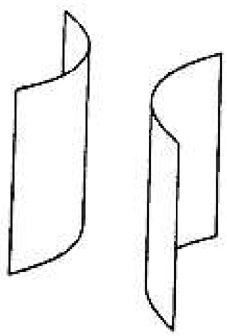
(4) Elliptisches Paraboloid



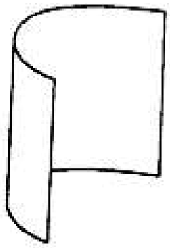
(5) Hyperbolisches Paraboloid



(7) Elliptischer Zylinder



(8) Hyperbolischer Zylinder



(9) Parabolischer Zylinder

BU Anleitung, Beispiele, Aufgaben

1

BU1: Hier zunächst einige Beispiele für unitäre Räume. $\mathbb{C}^n$ ist mit dem Standard-Skalarprodukt $\langle\cdot,\cdot\rangle$ ein euklidischer Raum, wobei

$$(x,y) \mapsto \langle x,y\rangle = \sum_{i=1}^{n} \overline{x_i} y_i$$

mit $x = (x_1,\ldots,x_n)^T$, $y = (y_1,\ldots,y_n)^T$. $\mathbb{R}^n$ ist mit dem Standard-Skalarprodukt:

$$(x,y) \mapsto \langle x,y\rangle = \sum_{i=1}^{n} x_i y_i,$$

wo $x = (x_1,\ldots,x_n)^T$, $y = (y_1,\ldots,y_n)^T$, ebenfalls ein euklidischer Raum. Rechnen Sie selbst die Axiome nach! Die aus diesem Skalarprodukt entstehende Norm ist gerade

$$|x| = \langle x,x\rangle^{1/2} = (\sum_{i=1}^{n} \overline{x_i} x_i)^{1/2} = (\sum_{i=1}^{n} |x_i|^2)^{1/2}.$$

Die Schwarzsche Ungleichung lautet:
$|\langle x,y\rangle| \leq |x||y|$, d.h. $|\sum_{i=1}^n \overline{x}_i y_i| \leq (\sum_{i=1}^n |x_i|^2)^{1/2}(\sum_{i=1}^n |y_i|^2)^{1/2}$.

BU2: Ein weiteres Beispiel ist folgendes: $C[a,b]$ ist der Raum, der auf einem reellen Intervall $[a,b]$ stetigen komplexwertigen Funktionen. $p \in C[a,b]$ nehme nur reelle positive Werte an. Dann ist für $f,g \in C[a,b]$

$$\langle f,g\rangle := \int_b^a p(t)\overline{f(t)}g(t)dt$$

ein Skalarprodukt auf $C[a,b]$.

Beweis:

1. $\langle f,f\rangle = \int_a^b p(t)|f(t)|^2 dt > 0$ und $= 0 \iff \forall_t f(t) = 0$, d.h. $f = 0$.
2. $$\begin{aligned} \langle f,\alpha g+\beta h\rangle &= \int_a^b p(t)\overline{f(t)}(\alpha g(t)+\beta h(t))dt \\ &= \alpha\int_a^b p(t)\overline{f(t)}g(t)dt + \beta\int_a^b p(t)\overline{f(t)}h(t)dt \\ &= \alpha\langle f,g\rangle + \beta\langle f,h\rangle\,. \end{aligned}$$
3. $$\begin{aligned} (g,f) &= \int_a^b p(t)\overline{g(t)}f(t)dt \\ &= \overline{(\int_a^b \overline{p(t)\overline{g(t)}f(t)}dt)} \\ &= \overline{(\int_a^b \overline{p(t)}g(t)\overline{f(t)}dt)} \\ &= \overline{(\int_a^b p(t)\overline{f(t)}g(t)dt)} \\ &= \overline{\langle f,g\rangle}, \end{aligned}$$

da $p(t)$ reell, d.h. $p(t) = \overline{p(t)}$. □

Die zugehörige Norm auf $C[a,b]$ lautet:

$$|f| := \left(\int_a^b p(t)\overline{f(t)}f(t)dt\right)^{1/2} = \left(\int_a^b p(t)|f(t)|^2 dt\right)^{1/2}$$

und die Schwarzsche Ungleichung

$$|\int_a^b p(t)\overline{f(t)}g(t)dt)| \leq (\int_a^b p(t)|f(t)|^2 dt)^{1/2}(\int_a^b p(t)|g(t)|^2 dt)^{1/2}.$$

Insbesondere ist $p(t) = 1$ eine zulässige „Gewichtsfunktion“. Dieses Skalarprodukt auf $C[a,b]$ bzw. auf dem größeren Raum $L^2[a,b]$, der auf $[a,b]$ quadratisch integrierbaren Funktionen, wird Ihnen vermutlich im Laufe Ihres Studiums noch öfter begegnen.

Das in BU1 angegebene Skalarprodukt ist keineswegs das einzige auf $\mathbb{C}^n$. Rechnen Sie selbst nach: Ist

$$P = \begin{pmatrix} p_1 & & \\ & \ddots & \\ & & p_n \end{pmatrix}$$

eine Diagonalmatrix mit reellen positiven Diagonalelementen, so ist mit dem Standardskalarprodukt $\langle\cdot,\cdot\rangle$ auch $\langle x,y\rangle_P := \langle x, Py\rangle$ ein Skalarprodukt. Was ist, wenn Sie — im Falle $n = 2$ — setzen $P = \left(\begin{smallmatrix} 2 & 1 \\ 1 & 2 \end{smallmatrix}\right)$?

Weiter bis 2.

2

BU3: Es seien $A, B \in \mathbb{R}^{n\times n}$. A habe die Spalten $a_1, \ldots, a_n$, ferner B die Zeilen $b_1^T, \ldots, b_n^T$. Dann ist das Element c_{ij} von $C := BA$ von der Form $c_{ij} = (b_i, a_j)$ $(i,j = 1,\ldots,n)$. Ist nun A invertierbar, $B := A^{-1}$, so ist $BA = I$, d.h. $c_{ij} = \delta_{ij}$ (Kronecker). Damit bilden aber für invertierbares A die Spalten von A und die Zeilen von A^{-1} ein Biorthonormalsystem. Was folgt für die Zeilen von A und die Spalten von A^{-1}?

Nun ein paar Beispiele zum Begriff der Orthogonalität:

BU4: Nehmen wir den $\mathbb{C}^n$ mit dem Standard-Skalarprodukt. Dann bilden die kanonischen Einheitsvektoren $(e_1,\ldots,e_n)$ eine Orthonormalbasis. Es ist ja

$$e_i = (0,\ldots,0,1,0,\ldots,0)^T = (\delta_{i1}, \delta_{i2}, \ldots, \delta_{ii}, \ldots, \delta_{in})^T$$

(δ_{ik} = Kronecker-Symbol). Damit ist

$$\langle e_i, e_j\rangle = \sum_{k=1}^n \delta_{ik}\delta_{jk} = \delta_{ij}.$$

(Rechnen Sie für $n = 3$ dies notfalls explizit!)

Sie bilden also ein Orthogonalsystem, und da sie überdies eine Basis bilden, also eine ONB. Ein etwas ungewohnteres Orthogonalsystem erhalten wir aus BU2:

BU5: Nach BU2 ist $C[0,2\pi]$ mit $\langle f,g\rangle = \int_0^{2\pi} \overline{f(t)}g(t)dt$ ein euklidischer Raum. Wir untersuchen die Funktionen $f_k(t) := \frac{1}{\sqrt{2\pi}}e^{ikt}$ ($i =$ imaginäre Einheit) für $k \in \mathbb{Z}$. Dann ist wegen
$\overline{e^{ikt}} = e^{-ikt}$:

$$\langle f_j, f_k\rangle = \frac{1}{2\pi}\int_0^{2\pi} e^{-ijt}e^{ikt}dt = \frac{1}{2\pi}\int_0^{2\pi} e^{i(k-j)t}dt = \frac{1}{2\pi}\begin{cases} 2\pi & j = k \\ 0 & j \neq k \end{cases} = \begin{cases} 1 & j = k \\ 0 & j \neq k \end{cases}$$

Diese Funktionen bilden also ein Orthonormalsystem in $C[0,2\pi]$.

Das Orthogonalisierungsverfahren von SCHMIDt finden Sie im folgenden Beispiel aufgeführt:

BU6: Wir nehmen den $\mathbb{R}^4$ mit dem Standard-Skalarprodukt: Wir setzen

$$u_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad u_2 = \begin{pmatrix} 3 \\ 3 \\ 1 \\ 1 \end{pmatrix}, \quad u_3 = \begin{pmatrix} 5 \\ 3 \\ 3 \\ 1 \end{pmatrix}, \quad u_4 = \begin{pmatrix} 8 \\ 4 \\ 2 \\ 2 \end{pmatrix}$$

Dann ist $y_1 := u_1$, $|y_1| = \sqrt{4} = 2$, $x_1 = \begin{pmatrix} 0{,}5 \\ 0{,}5 \\ 0{,}5 \\ 0{,}5 \end{pmatrix} = \frac{1}{|y_1|}y_1$.

$$y_2 := u_2 - \sum_{i=1}^{1} \langle x_i, u_2\rangle\, x_i = u_2 - \langle x_1, u_2\rangle\, x_1 = \begin{pmatrix} 3 \\ 3 \\ 1 \\ 1 \end{pmatrix} - 4\begin{pmatrix} 0,5 \\ 0,5 \\ 0,5 \\ 0,5 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}$$

Hier ist $|y_2| = 2$ und somit

$$x_2 := \frac{1}{y_2}y_2 = \begin{pmatrix} 0,5 \\ 0,5 \\ -0,5 \\ -0,5 \end{pmatrix}$$

Damit ist

$$y_3 := u_3 - \langle x_1, u_3\rangle\, x_1 - \langle x_2, u_3\rangle\, x_2 = \begin{pmatrix} 5 \\ 3 \\ 3 \\ 1 \end{pmatrix} - 6\begin{pmatrix} 0,5 \\ 0,5 \\ 0,5 \\ 0,5 \end{pmatrix} - 2\begin{pmatrix} 0,5 \\ 0,5 \\ -0,5 \\ -0,5 \end{pmatrix} = \begin{pmatrix} +1 \\ -1 \\ +1 \\ -1 \end{pmatrix}$$

Durch Normieren folgt wieder

$$x_3 := \begin{pmatrix} 0,5 \\ -0,5 \\ 0,5 \\ -0,5 \end{pmatrix}$$

Den letzten Vektor x_4 sollten Sie selbst bestimmen. Wir haben also — bzw. Sie werden — erhalten

$$x_1 = \begin{pmatrix} 0,5 \\ 0,5 \\ 0,5 \\ 0,5 \end{pmatrix}, \quad x_2 = \begin{pmatrix} 0,5 \\ 0,5 \\ -0,5 \\ -0,5 \end{pmatrix}, \quad x_3 = \begin{pmatrix} 0,5 \\ -0,5 \\ 0,5 \\ -0,5 \end{pmatrix}, \quad x_4 = \begin{pmatrix} 0,5 \\ -0,5 \\ -0,5 \\ 0,5 \end{pmatrix}.$$

Prüfen Sie nun selbst noch die restlichen Aussagen von Satz U.9.

Dieses Orthogonalisierungsverfahren ist natürlich nicht auf den $\mathbb{R}^n$ oder $\mathbb{C}^n$ beschränkt. Im vorigen Beispiel haben sie ein ONS in $C[0, 2\pi]$ gesehen. Denken Sie mal über folgende Aufgabe nach!

Wir betrachten $C[0,1]$ mit dem Skalarprodukt $\langle f, g\rangle = \int_0^1 \overline{f(t)}g(t)dt$. Darin sind durch $f_k(t) := t^k$ $(k = 0, 1, \dots)$ linear unabhängige Elemente gegeben, die wir also nach Satz U.9 orthonormieren können.

Was ist dafür zu tun?

Für die Bestimmung des Proximums gibt es einige wichtige Anwendungen:

BU7: Im $\mathbb{R}^2$ seien Vektoren x_0, x_1 gegeben, $x_1 \neq 0$. Dann beschreiben die Punkte von $G = \{x_0 + tx_1 \,|\, t \in \mathbb{R}\}$ eine Gerade durch den Punkt x_0. Von einem weiteren Punkt $y \in \mathbb{R}^2$ wollen wir den kleinsten euklidischen Abstand zu G wissen, d.h. gesucht ist

$$\delta := \inf\{|y - x| \,|\, x \in G\} = \inf\{|(y - x_0) - tx_1| \,|\, t \in \mathbb{R}\}$$

Zu bestimmen ist also das Proximum $t_0x_1 \in V := \mathrm{span}(x_1)$ zu $y - x_0$. Dazu orthogonalisieren wir $(x_1, y - x_0)$ nach E. SCHMIDT. Dies liefert das Orthonormalsystem

$$\begin{aligned} x_1' &:= \frac{1}{|x_1|} \cdot x_1, \\ x_2 &:= y - x_0 - \langle x_1', y - x_0\rangle\, x_1', \\ x_2' &:= \frac{1}{|x_2|} \cdot x_2. \end{aligned}$$

Damit ist dann

$$y - x_0 = \langle x_1', y - x_0\rangle \cdot x_1' + x_2 = \langle x_i', y - x_0\rangle\, x_1' + |x_2| \cdot x_2'.$$

Nach dem Beweis von Satz und Definition U.12 ist damit

$$\langle x_1', y - x_0\rangle\, x_1'$$

das Proximum und $|x_2| = \delta$. Man beachte, daß x_2' gar nicht explizit gebraucht wird.

Stellen Sie analoge Formeln für den Abstand eines Punktes von einer Ebene in $\mathbb{R}^3$ auf.

Weitere wichtige Beispiele findet man in der Approximationstheorie:

In einem homogenen elektrischen Feld $\mathcal{E} = (\epsilon_1, \epsilon_2, \epsilon_3)^T$ gilt für die zwischen zwei um einen Vektor $s = (\sigma_1, \sigma_2, \sigma_3)^T$ auseinanderliegenden Punkten gemessene Spannung U:

$$\langle \mathcal{E}, s\rangle = U \text{ d.h. } \epsilon_1\sigma_1 + \epsilon_2\sigma_2 + \epsilon_3\sigma_3 = U.$$

Messen wir nun für drei unabhängige Richtungen s_1, s_2, s_3 die jeweiligen Spannungen U_1, U_2, U_3, so läßt sich das Feld, d.h. $\epsilon_1, \epsilon_2, \epsilon_3$ aus dem Gleichungssystem

$$\epsilon_1\sigma_{i1} + \epsilon_2\sigma_{i2} + \epsilon_3\sigma_{i3} = U_i \qquad (i = 1, 2, 3)$$

berechnen. Da beim Messen Fehler auftreten, wird man weitere Messungen ausführen. Dies gibt weitere Gleichungen dieser Art, d.h. das Feld genügt einem System

$$\epsilon_1\sigma_{i1} + \epsilon_2\sigma_{i2} + \epsilon_3\sigma_{i3} = U_i \qquad (i = 1, 2, \dots, m).$$

Wir haben also $m > 3$ Gleichungen für drei Unbekannte. Theoretisch sollte dies System eindeutig lösbar sein, durch die eingetretenen Meßfehler wird in der Praxis dieses System keine Lösung besitzen. Man hilft sich indem man sagt: Bestimme

$\epsilon_1, \epsilon_2, \epsilon_3$ so, daß, wenn schon $\epsilon_1\sigma_{i1} + \epsilon_2\sigma_{i2} + \epsilon_3\sigma_{i3} - U_i = 0$ nicht möglich ist, so diese Ausdrücke doch möglichst klein werden.

Dies ergibt die Situation von Problem U.22 mit $A := (\sigma_{ij}),\ b := (U_i),\ x := (\epsilon_1, \epsilon_2, \epsilon_3)^T$. Betrachten wir dazu folgendes Beispiel:

$$A := \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & -1 \end{pmatrix}, \quad b := \begin{pmatrix} 3,1 \\ 0,9 \\ 0,9 \\ -0,9 \end{pmatrix}.$$

Man überzeugt sich mit Gauß-Elimination sofort, daß $ax = b$ nicht lösbar ist.

Stattdessen betrachten wir

$$A^H Ax = A^H b.$$

Es ist

$$A^H A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & -1 \end{pmatrix} = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix}, A^H b = \begin{pmatrix} 4 \\ 4 \\ 4 \end{pmatrix}.$$

Somit ist $x_0 = (1,1,1)^T$ die Lösung von $A^H Ax = A^H b$.

Unsere Meßfehler ergeben sich aus

$$|Ax_0 - b| = | \begin{pmatrix} 3 \\ 1 \\ 1 \\ -1 \end{pmatrix} - \begin{pmatrix} 3,1 \\ 0,9 \\ 0,9 \\ -0,9 \end{pmatrix} | = | \begin{pmatrix} -0,1 \\ 0,1 \\ 0,1 \\ -0,1 \end{pmatrix} | = 0,2.$$

Weiter bis 3.

3

Beispiele für normale Homomorphismen erhalten Sie aus Matrizen F, die mit ihrer adjungierten F^H vertauschbar sind: $FF^H = F^H F$. Ein interessantes Beispiel ist

BU9: In $C[0, 2\pi]$ betrachten wir den Teilraum U aller Funktionen f, die beliebig oft differenzierbar sind und für deren sämtliche Ableitungen einschließlich der Funktion selbst gilt $f^{(n)}(0) = f^{(n)}(2\pi)$ $(n = 0, 1, \ldots)$. Etwa die Funktionen $\sin kx$, $\cos kx$ gehören zu U. Wir versehen U mit dem Skalarprodukt

$$\langle f, g \rangle = \int_0^{2\pi} \overline{f(t)} g(t) dt$$

und untersuchen den Endomorphismus

$$\hat{D} : f(t) \mapsto \frac{df}{dt}(t).$$

Dies ist ein normaler Endomorphismus. Durch partielle Integration folgt nämlich

$$\left\langle \hat{D}(f), g \right\rangle = \int_0^{2\pi} \overline{f'(t)} g(t) dt = \overline{f(2\pi)} g(2\pi) - \overline{f(0)} g(0) - \int_0^{2\pi} \overline{f(t)} g'(t) dt = -\left\langle f, \hat{D}(g) \right\rangle,$$

da der ausintegrierte Bestandteil nach Voraussetzung über den Raum U wegfällt. Damit ist $\hat{D}^H = -\hat{D}$ und somit

$$\hat{D}\hat{D}^H = \hat{D}(-\hat{D}) = -\hat{D}\hat{D} = \hat{D}^H\hat{D}.$$

Nun ist mit $\hat{D}$ auch jedes Polynom in $\hat{D}$ mit komplexen Koeffizienten normal. Dies liefert uns hier beispielsweise

$$L := -\hat{D}^2 + \alpha\hat{D} + \beta\,\mathrm{id}_U$$

ist normal $(\alpha, \beta \in \mathbb{C})$ und L bewirkt für $f \in U$:

$$L(f) = -f'' + \alpha f' + \beta f.$$

Das Anwenden eines unitären oder orthogonalen Endomorphismus ändert Skalarprodukte nicht. Da Winkel und Längen hierüber definiert sind, sind diese Abbildungen also *längentreu* und *winkeltreu*. Sie sind also Verallgemeinerungen der aus Drehungen und Spiegelungen aufgebauten Abbildungen, die wir schon im Vorkapitel studiert hatten.

BU 10: Ein Beispiel für eine solche Abbildung im $\mathbb{C}^4$ wird etwa durch die Matrix

$$F = \frac{1}{2}\begin{pmatrix} 0 & 0 & 1+i & 1-i \\ 0 & 0 & 1-i & 1+i \\ 1+i & 1-i & 0 & 0 \\ 1-i & 1+i & 0 & 0 \end{pmatrix}$$

gegeben. Prüfen Sie für $f : x \mapsto Fx$ die einzelnen Bedingungen von Satz U.36 nach. Die Eigenwerte von F sind $+1, -1, +i, -i$ und die Vektoren

$$x_1 = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad x_2 = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}, \quad x_3 = \frac{1}{2}\begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}, \quad x_4 = \frac{1}{2}\begin{pmatrix} -1 \\ 1 \\ 1 \\ -1 \end{pmatrix}$$

sind die zugehörigen Eigenvektoren. Transformieren Sie F auf Diagonalform.

BU11: Sei U wieder der in BU9 beschriebene Raum der beliebig oft differenzierbaren Funktionen mit „periodischen Randbedingungen“ mit dem in BU9 notierten Skalarprodukt. Wir betrachten für eine beliebige Zahl α die Abbildung

$$T : U \to U, \quad f(t) \mapsto e^{i\alpha t} f(t).$$

— überzeugen Sie sich, daß tatsächlich für $f \in U$ auch $T(f) \in U$ ist! — Dieses T ist unitär; denn wir haben für $f, g \in U$:

$$\begin{aligned} \langle T(f), T(g)\rangle &= \int_0^{2\pi} \overline{(e^{i\alpha t} f(t))} \cdot (e^{i\alpha t} g(t))dt = \int_0^{2\pi} e^{-i\alpha t}\overline{f(t)}e^{i\alpha t} g(t)dt \\ &= \int_0^{2\pi} \overline{f(t)} g(t)dt = \langle f, g\rangle \end{aligned}$$

womit die Bedingung Satz U.36(v) gezeigt ist.

BU12: Vertauschen wir in der Matrix F aus BU10 die ersten beiden Spalten, so erhalten wir die Matrix

$$G := \frac{1}{2} \cdot \begin{pmatrix} 0 & 0 & 1+i & 1-i \\ 0 & 0 & 1-i & 1+i \\ 1-i & 1+i & 0 & 0 \\ 1+i & 1-i & 0 & 0 \end{pmatrix},$$

für die $G^H = G$ ist, d.h. G ist hermitesch oder selbstadjungiert. Die Eigenwerte sind $\lambda_1 = \lambda_2 = 1, \lambda_3 = \lambda_4 = -1$ mit Eigenvektoren

$$x_1 = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad x_2 = \frac{1}{2}\begin{pmatrix} i \\ -i \\ 1 \\ -1 \end{pmatrix}, \quad x_3 = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}, \quad x_4 = \frac{1}{2}\begin{pmatrix} 1 \\ -1 \\ i \\ -i \end{pmatrix},$$

die ein Orthonormalsystem bilden.

BU13: In $C[0,1]$ mit dem Skalarprodukt $\langle f, g \rangle = \int_0^1 f(t)g(t)dt$ ist der „Multiplikationsoperator"

$$\varphi : f(t) \mapsto tf(t)$$

selbstadjungiert, denn

$$\langle \varphi(f), g \rangle = \int_0^1 tf(t)g(t)dt = \int_0^1 f(t)tg(t)dt = \langle f, \varphi(g) \rangle .$$

Dieser Homomorphismus hat *keinen* Eigenwert - so etwas kommt nur im ∞-dimensionalen vor.

BU14: Verifizieren Sie, daß mit dem Differentialoperator $\hat{D}$ aus BU9 zwar $\hat{D}$ selbst nur normal und nicht selbstadjungiert ist, während der Operator

$$D := i\hat{D} : f \mapsto i\frac{df}{dt}$$

selbstadjungiert ist. Zeigen Sie, daß jedes $k \in Z$ ein Eigenwert ist mit Eigenvektor $f_k : f_k(t) = e^{ikt}$.

BU15: Schließlich sei noch ein Beispiel für die reelle Normalform durchgerechnet: Sei mit reellen α, β

$$F := \frac{1}{2} \begin{pmatrix} \alpha & 1-\beta & -\alpha & 1+\beta \\ 1+\beta & \alpha & 1-\beta & -\alpha \\ -\alpha & 1+\beta & \alpha & 1-\beta \\ 1-\beta & -\alpha & 1+\beta & \alpha \end{pmatrix}$$

Betrachten wir dies als Abbildung in $\mathbb{C}^4$, so erhalten wir als Eigenwerte $\lambda_1 = 1$, $\lambda_2 = -1$, $\lambda_3 = \alpha - i\beta$, $\lambda_4 = \alpha + i\beta$ mit den zugehörigen Eigenvektoren

$$u_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad u_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}, \quad u_3 = \begin{pmatrix} 1-i \\ 1+i \\ -1+i \\ -1-i \end{pmatrix}, \quad u_4 = \begin{pmatrix} 1+i \\ 1-i \\ -1-i \\ -1+i \end{pmatrix}$$

Nach dem Beweis von Definition und Satz U.28 haben wir nun, um zu der *reellen* Normalform zu kommen, die Vektoren

$$x_1 := u_1, \; x_2 := u_2, \; x_3 := \frac{1}{2}(u_3 + u_4), \; x_4 := \frac{1}{2i}(u_3 - u_4)$$

zu bilden, d.h.

$$x_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad x_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}, \quad x_3 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}, \quad x_4 = \begin{pmatrix} -1 \\ 1 \\ 1 \\ -1 \end{pmatrix},$$

die in der Tat ein Orthogonalsystem bilden — die Normierung ist weggelassen. Prüfen Sie nach, daß gilt:

$$Fx_1 = x_1, \quad Fx_2 = x_2, \quad Fx_3 = \alpha x_3 + \beta x_4, \quad Fx_4 = -\beta x_3 + \alpha x_4,$$

d.h. F hat bezüglich dieser Basis die Darstellung

$$\tilde{F} := \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & \alpha & -\beta \\ & & \beta & \alpha \end{pmatrix}.$$

Im Falle $\alpha^2 + \beta^2 = 1$ können Sie dies geometrisch über Drehungen und Spiegelungen interpretieren. Tun Sie das!

Weiter bis $\boxed{4}$.

$\boxed{4}$

BU16: Sei $f : \mathbb{C}^n \to \mathbb{C}^n$ gegeben durch $x \mapsto Fx$, wobei F eine Diagonalmatrix der Form $F = \operatorname{diag}(\lambda_1, \ldots, \lambda_m, \mu_1, \ldots, \mu_n, \nu_1, \ldots, \nu_k)_{m+n+k, m+n+k}$. Dabei seien diese Diagonalelemente reell und

$$\lambda_i > 0 \ (i = 1, \ldots, m), \ \mu_i = 0 \ (i = 1, \ldots, n), \ \nu_i < 0 \ (i = 1, \ldots, k).$$

Nun bestimmen Sie einmal die Unterräume, auf denen f positiv definit, positiv semidefinit, usw. ist.

Als nächstes nehmen Sie einen selbstadjungierten Endomorphismus, der bezüglich einer orthogonalen Basis aus Eigenvektoren die eben genannte Matrixdarstellung hat und behandeln dafür die selbe Fragestellung.

BU17: Ist A eine beliebige $n \times n$ Matrix, so ist AA^H selbstadjungiert (Beweis ?). Zeigen Sie

- AA^H ist positiv semidefinit,
- AA^H ist positiv definit, genau wenn A invertierbar ist.

Ein kleines Beispiel: Für

$$A = \begin{pmatrix} 1 & -2 \\ -1 & 2 \end{pmatrix}$$

ist

$$AA^H = \begin{pmatrix} 5 & -5 \\ -5 & 5 \end{pmatrix}$$

und

$$\langle AA^H x, x \rangle = 5(\xi_1 - \xi_2)^2 \geq 0.$$

BU18: Mit Satz U.47 hatten wir einen Zusammenhang zwischen der Definitheit und den Hauptunterdeterminaten hergestellt. Untersuchen Sie diesen Zusammenhang für negativ definite bzw. semidefinite Matrizen. Muß man, um aus den Hauptunterdeterminanten auf die Positiv-Definitheit schließen zu können, das Vorzeichen aller Hauptunterdeterminanten kennen? Wie steht das bei Semidefinit?

BU19: $\langle x, y\rangle_1$ sei das Standard-Skalarprodukt in $\mathbb{R}^n$ oder $\mathbb{C}^n$, A eine positiv definite hermitesche Matrix. Dann ist $\langle x, y\rangle_2 := \langle Ax, y\rangle_1$ wieder ein Skalarprodukt auf $\mathbb{R}^n$ oder $\mathbb{C}^n$ (Beweis?) und somit auch $|\cdot|_2 : |x|_2 := \sqrt{\langle Ax, x\rangle}$ eine Norm auf $\mathbb{R}^n$ oder $\mathbb{C}^n$. Vergleichen Sie die „Einheitskugeln“ bezüglich beider Normen und den Orthogonalitätsbegriff bezüglich beider Skalarprodukte.

Betrachten Sie speziell im $\mathbb{R}^2$ die Matrizen

$$A_1 := \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}, \quad A_2 := \frac{1}{2}\begin{pmatrix} 5 & -1 \\ -1 & 5 \end{pmatrix}$$

Was geschieht wenn Sie eine positiv-semidefinite, aber nicht definite Matrix A nehmen?

BU20: Sei $F_1 := \begin{pmatrix} 2 & -6 \\ -6 & -7 \end{pmatrix}$, $F_2 := G^H F_1 G$ mit $G = \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}$. Dann ist

$$F_2 = \begin{pmatrix} 25 & 0 \\ 0 & -50 \end{pmatrix}$$

Somit hat F_2 die Eigenwerte 25, -50 also einen positiven, einen negativen und keinen $= 0$. Nach dem Sylvesterschen Satz muß F_1 dieselbe Signatur haben.

Die Eigenwerte von F_1 sind die Nullstellen von

$$\begin{aligned}\det(F_1 - \lambda I) &= \det\begin{pmatrix} 2-\lambda & -6 \\ -6 & -7-\lambda \end{pmatrix} = (2-\lambda)(-7-\lambda) - 36 \\ &= \lambda^2 + 5\lambda - 50 = (\lambda - 5)(\lambda + 10).\end{aligned}$$

Damit hat F_1 die Eigenwerte 5 und -10, die zwar verschieden von denen von F_2 sind, aber die selbe Vorzeichenverteilung besitzen.

Weiter bis $\boxed{5}$.

$\boxed{5}$

BU21: Im $\mathbb{R}^3$ beschreibt die Gleichung $x^2+y^2 = \delta(z-z_0^2)$ einen Kreis-Doppelkegel um die z-Achse mit „Taille“ im Punkt $(0, 0, z_0)$. Für $\delta = \frac{R}{z_0^2}$ und $z_0 \to \infty$ erhalten wir als Grenzfall den Kreiszylinder mit Radius R um die z-Achse.

$ax + by + cz = d$ beschreibt eine allgemeine Ebene im $\mathbb{R}^3$. Für $c \neq 0$ können wir nach z auflösen und aus der Ebenen- und der Kegelgleichung z eliminieren.

Dies liefert dann gerade die Gleichung (U.49) für — wir sind jetzt im $\mathbb{R}^2$ — die Kurven zweiter Ordnung, die eben gerade die „Kegelschnitte“ sind.

Versuchen Sie einmal die obigen Parameter so zu wählen, daß Sie eine Ellipse, bzw. eine Parabel, bzw. eine Hyperbel erhalten. Wie kommen die „Sonderfälle“ zustande?

BU22: In der Normalformendarstellung für Flächen zweiter Ordnung hatten wir gesehen, daß man alle Komponenten des linearen Anteils c, die auf nichtverschwindende Eigenwerte treffen, eliminieren kann. Überlegen Sie einmal, daß man auch von den restlichen Komponenten noch alle mit evtl. einer Ausnahme zu Null machen kann. Was bedeutet dies geometrisch?

Zum Einüben sei noch folgender Weg empfohlen. Nehmen Sie einen Kegelschnitt in Normalform, etwa die Ellipse

$$\frac{x^2}{4} + \frac{y^2}{9} = 1,$$

die ihr Zentrum im Nullpunkt und Halbachsen der Längen 2 und 3 in Richtung der Koordinatenachsen hat. Jetzt verschieben und drehen Sie das Koordinatensystem und stellen die Ellipse bezüglich der neuen Koordinaten dar. Von der so erhaltenen quadratischen Form bilden Sie nun entsprechend dem Vorgehen im Skript die Normalform. Viel Spaß beim (Ver-)Rechnen.

BU23: Zum Schluß noch eine Aufgabe vor allem für die künftigen Lehrer unter Ihnen. a, b seien zwei feste Punkte im $\mathbb{R}^2$, α ein reeller Parameter, $|\cdot|$ die übliche Euklidnorm im $\mathbb{R}^2$.

Bestimmen Sie alle Punkte $x \in \mathbb{R}^2$, für die eine der folgenden Bedingungen gilt

1. $|x - a| + |x - b| = \alpha,$
2. $|x - a| - |x - b| = \alpha,$
3. $|x - a| \cdot |x - b| = \alpha,$
4. $\frac{|x-a|}{|x-b|} = \alpha.$

1., 2., 4. liefern Kurven zweiter Ordnung. Welche?