21. Euklidische und unitäre Räume

Wir wollen uns nun mit einem ganz anderen Thema beschäftigen, nämlich wie man Längen von Vektoren und Winkel zwischen zwei Vektoren berechnen (und überhaupt erst einmal definieren) kann. Zur Motivation betrachten wir dazu zunächst einmal den sehr einfachen Fall des Vektorraums $\mathbb{R}^2$, in dem sich diese beiden Fragen mit Hilfe von Elementargeometrie und Schulmathematik leicht beantworten lassen.

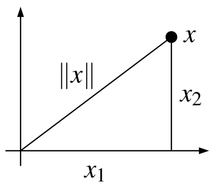
Beispiel 21.1 (Längen und Winkel in $\mathbb{R}^2$)**.** Wie ihr sicher aus der Schule wisst, ist das wesentliche Hilfsmittel für die Längen- und Winkelmessung in $\mathbb{R}^2$ das sogenannte *Skalarprodukt*

$$\langle x,y\rangle = x_1y_1 + x_2y_2 \in \mathbb{R} \qquad \text{für } x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},\ y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in \mathbb{R}^2.$$

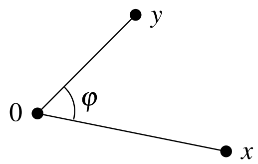
So ergibt sich z. B. wie im Bild unten links dargestellt aus dem Satz des Pythagoras, dass die Länge eines Vektors $x \in \mathbb{R}^2$ durch den Ausdruck

$$\sqrt{x_1^2 + x_2^2} = \sqrt{\langle x,x\rangle}$$

gegeben ist, den wir im Folgenden kurz als $||x||$ schreiben werden.



$$||x|| = \sqrt{\langle x,x\rangle}$$



$$\frac{y}{||y||} = e^{i\varphi} \cdot \frac{x}{||x||}$$

Wollen wir den Winkel φ zwischen zwei Vektoren $x,y \in \mathbb{R}^2\backslash\{0\}$ wie im Bild oben rechts berechnen, betrachten wir dazu am besten zunächst einmal die Vektoren $\frac{x}{||x||}$ und $\frac{y}{||y||}$, die in die gleiche Richtung wie x bzw. y zeigen, aber die Länge 1 haben. Fassen wir dann $x = x_1 + ix_2$ und $y = y_1 + iy_2$ als Elemente der komplexen Ebene $\mathbb{C} = \mathbb{R}^2$ auf, so folgt aus den Bemerkungen 6.5 und 9.11, dass

$$\frac{y}{||y||} = e^{i\varphi} \cdot \frac{x}{||x||}$$

ist, da sich die Winkel bei der komplexen Multiplikation addieren und $e^{i\varphi}$ eine Zahl mit Winkel φ und Betrag 1 ist. Einfache Umformungen in $\mathbb{C}$ ergeben nun wegen $||x||^2 = \overline{x}x$

$$e^{i\varphi} = \frac{y}{x} \cdot \frac{||x||}{||y||} = \frac{\overline{x}y}{\overline{x}x} \cdot \frac{||x||}{||y||} = \frac{\overline{x}y}{||x|| \cdot ||y||}.$$

Wegen $\overline{x}y = (x_1 - ix_2)(y_1 + iy_2) = x_1y_1 + x_2y_2 + i(x_1y_2 - x_2y_1)$ besagt der Realteil dieser Gleichung

$$\cos\varphi = \frac{x_1y_1 + x_2y_2}{||x|| \cdot ||y||},$$

woraus wir mit der obigen Definition des Skalarprodukts folgern, dass

$$\varphi = \arccos \frac{\langle x,y\rangle}{||x|| \cdot ||y||}$$

ist (beachte hierbei, dass der Arkuskosinus nur Werte zwischen 0 und π zurückliefert und aufgrund der Symmetrien der Kosinusfunktion damit den *unorientierten* Winkel zwischen x und y ergibt).

Sowohl Längen als auch Winkel lassen sich damit durch das Skalarprodukt ausdrücken. Wenn wir diese beiden Konzepte auch in anderen Vektorräumen definieren wollen, sollten wir den Begriff des Skalarprodukts also auf beliebige Vektorräume verallgemeinern.

Das Problem dabei ist jedoch, dass die Formel $\langle x,y\rangle = x_1y_1 + x_2y_2$ (oder eine entsprechend verallgemeinerte Version für höhere Dimensionen) explizit die Koordinaten der beiden Vektoren x und y benutzt. In einem allgemeinen Vektorraum gäbe es solche Koordinaten aber erst nach Wahl einer Basis – und die Formel würde natürlich auch unterschiedliche Ergebnisse liefern, wenn man die Koordinaten bezüglich verschiedener Basen nehmen würde. Wir schließen daraus, dass es in einem allgemeinen Vektorraum *kein natürlich definiertes* Skalarprodukt gibt, sondern dass ein Skalarprodukt eine *Zusatzstruktur* darstellt, die man zusätzlich zum Vektorraum erst einmal festlegen muss, bevor man mit konkreten Rechnungen anfangen kann.

Wir müssen diese Zusatzstruktur also zunächst erst einmal genauer definieren, d. h. konkret angeben, welche Eigenschaften ein Skalarprodukt haben soll. Klar ist, dass wir zwei Elementen eines K-Vektorraums V ein Element des zugrunde liegenden Körpers K zuordnen wollen, also formal eine Abbildung von $V \times V$ nach K betrachten müssen. Mit solchen Abbildungen wollen wir uns nun zunächst beschäftigen.

21.A Bilinearformen

Die erste wichtige Eigenschaft eines Skalarprodukts ist, dass es linear in beiden Vektoren ist. Derartige Abbildungen bezeichnet man als Bilinearformen.

Definition 21.2 (Bilinearformen)**.** Eine **Bilinearform** auf einem K-Vektorraum V ist eine Abbildung $b\colon V \times V \to K,\ (x,y) \mapsto b(x,y)$, die in beiden Komponenten eine lineare Abbildung ist, d. h. für die für alle $x_1,x_2,x,y_1,y_2,y \in V$ und $\lambda \in K$ die Eigenschaften

$$\begin{aligned} b(x_1+x_2,y) &= b(x_1,y)+b(x_2,y), \\ b(\lambda x,y) &= \lambda\, b(x,y), \\ b(x,y_1+y_2) &= b(x,y_1)+b(x,y_2), \\ b(x,\lambda y) &= \lambda\, b(x,y) \end{aligned}$$

gelten. Wie man leicht nachprüft, ist die Menge aller Bilinearformen auf V ein Unterraum von $\mathrm{Abb}(V \times V, K)$ (siehe Beispiel 13.3 (c)) und damit ein K-Vektorraum. Wir bezeichnen ihn mit $\mathrm{BLF}(V)$.

Beispiel 21.3.

(a) Die Abbildung

$$b\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R},\ \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) \mapsto x_1y_1 + x_1y_2 + x_2y_1 + 4x_2y_2$$

ist offensichtlich eine Bilinearform: Hält man y_1 und y_2 fest, so ist der gegebene Ausdruck eine lineare Abbildung in x_1 und x_2, und umgekehrt. Hingegen ist

$$b\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R},\ \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) \mapsto x_1y_1 + x_1 + y_1$$

keine Bilinearform: Da lineare Abbildungen nach Bemerkung 16.2 stets 0 auf 0 abbilden, kann b wegen

$$b\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = 1$$

bei festgehaltener zweiter Komponente y nicht linear im ersten Argument x sein.

(b) Ist $A \in K^{n\times n}$ eine quadratische Matrix, so ist

$$b_A\colon K^n \times K^n \to K,\ b(x,y) = x^{\mathsf{T}} A\, y \qquad (*)$$

nach den Rechenregeln für Matrizen aus Lemma 15.7 eine Bilinearform auf K^n – beachte, dass das Ergebnis hierbei als Produkt dreier Matrizen der Größen $1\times n$, $n \times n$ und $n \times 1$ eine

1×1-Matrix, also ein Element von K ist. Ist $A = (a_{i,j})_{i,j}$ und sind $x_1, \dots, x_n$ und $y_1, \dots, y_n$ die Koordinaten von x und y, so ist eine alternative Schreibweise für $(*)$ nach Definition 15.5

$$b(x,y) = (x_1 \ \cdots \ x_n) \cdot \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \sum_{i,j=1}^{n} x_i a_{i,j} y_j .$$

Wir wollen nun sehen, dass man in der Tat sogar jede mögliche Bilinearform auf K^n auf diese Art aus einer eindeutig bestimmten Matrix erhalten kann. Dies besagt der folgende Satz, der völlig analog zu Satz 16.23 und Folgerung 16.24 (a) über den Zusammenhang zwischen linearen Abbildungen und Matrizen ist, und der damit letztlich besagt, dass Bilinearformen auf K^n und $n \times n$-Matrizen über K „im Prinzip dasselbe" sind.

Satz und Definition 21.4 (Bilinearformen auf K^n und Matrizen)**.** *Es sei $n \in \mathbb{N}$.*

(a) *Zu jeder Bilinearform $b\colon K^n \times K^n \to K$ gibt es genau eine Matrix $A \in K^{n \times n}$ mit $b = b_A$ wie in Beispiel 21.3 (b), also mit*

$$b(x,y) = x^{\mathsf{T}} A\, y \quad \text{für alle } x,y \in K^n,$$

*nämlich $A = (b(e_i,e_j))_{i,j}$. Man nennt sie die **Gramsche Matrix** von b und bezeichnet sie mit A_b.*

(b) *Die Abbildung $K^{n \times n} \to \mathrm{BLF}(K^n)$, $A \mapsto b_A$ ist ein Isomorphismus mit Umkehrabbildung $\mathrm{BLF}(K^n) \to K^{n \times n}$, $b \mapsto A_b$.*

Beweis.

(a) Zunächst einmal legt die Bedingung $b = b_A$ die Matrix $A = (a_{i,j})_{i,j}$ eindeutig fest: Da wir den Eintrag in Zeile i und Spalte j von A als das Matrixprodukt $e_i^{\mathsf{T}} A\, e_j$ schreiben können, ist notwendigerweise

$$a_{i,j} = e_i^{\mathsf{T}} A\, e_j = b(e_i,e_j),$$

und damit $A = A_b$. Da b bilinear ist, folgt aus dieser Gleichung aber auch für alle Vektoren $x = x_1 e_1 + \cdots + x_n e_n$ und $y = y_1 e_1 + \cdots + y_n e_n$

$$b(x,y) = b\left(\sum_{i=1}^{n} x_i e_i, \sum_{j=1}^{n} y_i e_i\right) = \sum_{i,j=1}^{n} x_i b(e_i,e_j) y_j = x^{\mathsf{T}} A\, y,$$

und damit $b = b_A$.

(b) Die Abbildung $K^{n \times n} \to \mathrm{BLF}(K^n)$, $A \mapsto b_A$ ist linear, denn für alle $A,B \in K^{n \times n}$, $\lambda \in K$ und $x,y \in K^n$ gilt

$$\begin{aligned} & b_{A+B}(x,y) = x^{\mathsf{T}}(A+B)\,y = x^{\mathsf{T}} A\, y + x^{\mathsf{T}} B\, y = b_A(x,y) + b_B(x,y) \\ \text{und} \quad & b_{\lambda A}(x,y) = x^{\mathsf{T}}(\lambda A)\,y = \lambda\, x^{\mathsf{T}} A\, y = \lambda\, b_A(x,y). \end{aligned}$$

Da sie nach (a) auch bijektiv ist, ist sie damit ein Isomorphismus. Nach Definition der Gramschen Matrix ist ihre Umkehrabbildung genau $\mathrm{BLF}(K^n) \to K^{n \times n}$, $b \mapsto A_b$. □

Beispiel 21.5.

(a) Zur Bilinearform

$$b\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}, \ \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right) \mapsto x_1 y_1 + x_1 y_2 + x_2 y_1 + 4 x_2 y_2 \qquad (*)$$

aus Beispiel 21.3 (a) ist die zugehörige Gramsche Matrix

$$A_b = \begin{pmatrix} b(e_1,e_1) & b(e_1,e_2) \\ b(e_2,e_1) & b(e_2,e_2) \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix}.$$

Eine alternative Beschreibung von $A_b = (a_{i,j})_{i,j}$ ist offensichtlich, dass $a_{i,j}$ in einer Darstellung der Form $(*)$ von $b(x,y)$ genau der Koeffizient von $x_i y_j$ ist.

Umgekehrt können wir nun nach Satz 21.4 aus dieser Matrix auch die ursprüngliche Bilinearform durch die Formel

$$b(x,y) = (x_1 \ \ x_2) \cdot \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = x_1y_1 + x_1y_2 + x_2y_1 + 4x_2y_2$$

zurückgewinnen.

(b) Die Einheitsmatrix $E \in K^{n\times n}$ entspricht in der Korrespondenz aus Satz 21.4 genau der Bilinearform

$$b_E \colon K^n \times K^n \to K, \ (x,y) \mapsto x^{\mathsf T} y = x_1y_1 + \cdots + x_ny_n,$$

die wir in Beispiel 21.1 im Fall $K = \mathbb{R}$ und $n = 2$ schon beim gewöhnlichen Skalarprodukt auf $\mathbb{R}^2$ gesehen haben.

Genau wie bei linearen Abbildungen in Satz 16.26 und Folgerung 16.27 (a) können wir unsere Korrespondenz zwischen Bilinearformen und Matrizen nun unmittelbar von K^n auf beliebige endlich-dimensionale Vektorräume erweitern, indem wir dort eine Basis wählen und mit den Koordinatenvektoren bezüglich dieser Basis arbeiten.

Folgerung 21.6 (Bilinearformen auf V und Matrizen)**.** *Es seien V ein endlich-dimensionaler K-Vektorraum sowie $B = (x_1,\dots,x_n)$ eine Basis von V mit zugehöriger Koordinatenabbildung $\Phi_B \colon V \to K^n$ (siehe Konstruktion 16.17). Dann ist die Abbildung*

$$K^{n\times n} \to \mathrm{BLF}(V),\ A \mapsto b_A^B \qquad \text{mit } b_A^B(x,y) := \Phi_B(x)^{\mathsf T} A\, \Phi_B(y)$$

wieder ein K-Vektorraumisomorphismus mit Umkehrabbildung

$$\mathrm{BLF}(V) \to K^{n\times n},\ b \mapsto A_b^B \qquad \text{mit } A_b^B := (b(x_i,x_j))_{i,j}.$$

*Wie oben nennt man A_b^B die **Gramsche Matrix** von b bezüglich der Basis B.*

Beweis. Die Abbildung

$$\mathrm{BLF}(K^n) \to \mathrm{BLF}(V), \quad b \mapsto \Big((x,y) \mapsto b(\Phi_B(x), \Phi_B(y)) \Big),$$

die einer Bilinearform b auf K^n die Bilinearform auf V zuordnet, bei der man einfach in b die Koordinatenvektoren der Vektoren aus V einsetzt, ist offensichtlich ein Isomorphismus mit Umkehrabbildung

$$\mathrm{BLF}(V) \to \mathrm{BLF}(K^n), \quad b \mapsto \Big((x,y) \mapsto b(\Phi_B^{-1}(x), \Phi_B^{-1}(y)) \Big).$$

Verketten wir den Isomorphismus aus Satz 21.4 mit dieser Abbildung, erhalten wir also wie behauptet einen Isomorphismus $K^{n\times n} \to \mathrm{BLF}(K^n) \to \mathrm{BLF}(V)$, der eine Matrix A auf die Bilinearform $(x,y) \mapsto b_A(\Phi_B(x), \Phi_B(y)) = \Phi_B(x)^{\mathsf T} A\, \Phi_B(y)$ abbildet, und dessen Umkehrung einer Bilinearform b auf V die Matrix $(b(\Phi_B^{-1}(e_i), \Phi_B^{-1}(e_j)))_{i,j} = (b(x_i,x_j))_{i,j}$ zuordnet. □

Natürlich hängt die Gramsche Matrix einer Bilinearform wie in Folgerung 21.6 von der gewählten Basis ab. Wie das folgende Lemma zeigt, ist die Transformationsformel bei einem Basiswechsel jedoch eine andere als für Endomorphismen (siehe Bemerkung 19.3 (b)).

Lemma 21.7 (Verhalten von Gramschen Matrizen unter Basiswechsel)**.** *Es seien b eine Bilinearform auf einem endlich-dimensionalen K-Vektorraum V sowie B und B' zwei Basen von V. Dann gilt für die Gramschen Matrizen von b bezüglich B und B'*

$$A_b^{B'} = T^{\mathsf T} A_b^B\, T,$$

wobei $T = A^{B',B}$ die Basiswechselmatrix aus Definition 16.38 ist.

Beweis. Es seien $B = (x_1,\dots,x_n)$ und $B' = (x_1',\dots,x_n')$ die gewählten Basen. Nach Definition 16.38 enthält die k-te Spalte von $T = (a_{i,j})_{i,j}$ für $k = 1,\dots,n$ genau den Koordinatenvektor von x_k' bezüglich B, d. h. es gilt

$$x_k' = a_{1,k}x_1 + \cdots + a_{n,k}x_n.$$

Damit folgt für die Gramschen Matrizen mit der Formel aus Folgerung 21.6 sofort

$$A_b^{B'} = (b(x_k', x_l'))_{k,l} = \left(b\left(\sum_{i=1}^n a_{i,k} x_i, \sum_{j=1}^n a_{j,l} x_j \right) \right)_{k,l} = \left(\sum_{i,j=1}^n a_{i,k} b(x_i, x_j) a_{j,l} \right)_{k,l} = T^\mathsf{T} A_b^B T. \qquad \square$$

49

Bemerkung 21.8 (Verhalten von Matrizen unter Basiswechsel)**.** Bisher hatten wir Matrizen nahezu ausschließlich zur Beschreibung von linearen Abbildungen benutzt. Nach Definition ist eine Matrix aber zunächst einmal nichts weiter als ein rechteckiges Zahlenschema, ohne Vorgabe einer Bedeutung dieser Zahlen. In der Tat haben wir nun gesehen, dass man Matrizen auch noch für ganz andere Dinge verwenden kann, nämlich z. B. zur Darstellung von Bilinearformen.

Ohne weitere Informationen ergibt es daher keinen Sinn zu fragen, wie sich eine (quadratische) Matrix unter einem Basiswechsel transformiert. Die Antwort auf diese Frage hängt nach Bemerkung 19.3 (b) und Lemma 21.7 davon ab, welche Bedeutung die Einträge in der Matrix haben:

Bei einem Basiswechsel mit zugehöriger Basiswechselmatrix T transformiert sich ...
... eine zu einem *Endomorphismus* gehörige Matrix A in die Matrix $T^{-1}AT$,
... eine zu einer *Bilinearform* gehörige Matrix A in die Matrix $T^\mathsf{T}AT$.

Wie bei linearen Abbildungen oder Endomorphismen könnten wir uns nun schließlich auch bei Bilinearformen wieder nach einer Normalform fragen: Wie können wir zu einer gegebenen Bilinearform $b \in \mathrm{BLF}(V)$ eine Basis von V so wählen, dass die zugehörige Gramsche Matrix A_b^B möglichst einfach wird? Wir wollen diese Frage hier allerdings nicht in dieser vollen Allgemeinheit beantworten, da wir im Folgenden hauptsächlich an Bilinearformen mit noch weiteren speziellen Eigenschaften interessiert sind. Diese Eigenschaften wollen wir jetzt einführen.

21.B Skalarprodukte

Wir kommen nun zu den in der Einleitung zu diesem Kapitel bereits angekündigten Skalarprodukten. Sie sind als Bilinearformen mit den folgenden beiden Eigenschaften definiert, die wir auch gleich wieder analog für Matrizen einführen wollen.

Definition 21.9 (Symmetrie und positive Definitheit)**.** Es seien $b \in \mathrm{BLF}(V)$ eine Bilinearform auf einem K-Vektorraum V und $A \in K^{n \times n}$.

(a) Die Bilinearform b heißt **symmetrisch**, wenn $b(x,y) = b(y,x)$ für alle $x,y \in V$ gilt.

Die Matrix A heißt **symmetrisch**, wenn $A^\mathsf{T} = A$ gilt.

(b) Es sei nun zusätzlich $K = \mathbb{R}$.

Die Bilinearform b heißt dann **positiv definit**, wenn $b(x,x) > 0$ für alle $x \in V \backslash \{0\}$ gilt.

Die Matrix A heißt dann **positiv definit**, wenn $x^\mathsf{T} A x > 0$ für alle $x \in \mathbb{R}^n \backslash \{0\}$ gilt.

Bemerkung 21.10.

(a) Die Bedingung der positiven Definitheit lässt sich offensichtlich nur für einen geordneten Körper (siehe Kapitel 4.B) formulieren. Für uns ist hierbei eigentlich nur der Fall $K = \mathbb{R}$ interessant. Wir werden die Bedingung der positiven Definitheit in Konstruktion 21.18 aber noch etwas abändern, so dass sie dann auch im Fall $K = \mathbb{C}$ anwendbar ist.

(b) Da eine Bilinearform b in jedem Eintrag linear ist, gilt natürlich stets $b(0,0) = 0$. Eine positiv definite Bilinearform auf einem $\mathbb{R}$-Vektorraum V erfüllt damit also immer $b(x,x) \geq 0$ für alle $x \in V$. Diese Bedingung, die wir später für Skalarprodukte fordern werden, wird uns dann sicherstellen, dass wir aus $b(x,x)$ die Wurzel ziehen und so die Länge von x definieren können.

In manchen Fällen (siehe Satz 26.20) benötigt man allerdings auch noch die folgenden zur positiven Definitheit analogen Bedingungen: Eine Bilinearform $b \in \mathrm{BLF}(V)$ auf einem reellen Vektorraum V heißt ...

- ... **negativ definit**, wenn $b(x,x) < 0$ für alle $x \in V$ mit $x \neq 0$;
- ... **positiv semidefinit**, wenn $b(x,x) \geq 0$ für alle $x \in V$;
- ... **negativ semidefinit**, wenn $b(x,x) \leq 0$ für alle $x \in V$;
- ... **indefinit**, wenn sie weder positiv noch negativ semidefinit ist, also wenn es $x,y \in V$ gibt mit $b(x,x) < 0$ und $b(y,y) > 0$.

Entsprechende Eigenschaften definiert man natürlich auch für reelle quadratische Matrizen.

Als Erstes wollen wir nun die wohl erwartete Aussage zeigen, dass sich die in Definition 21.9 eingeführten Begriffe für Bilinearformen und Matrizen entsprechen.

Lemma 21.11 (Symmetrie und positive Definitheit bei Bilinearformen und Matrizen). *Es seien b eine Bilinearform auf einem endlich-dimensionalen K-Vektorraum V, B eine Basis von V, und A_b^B wie in Folgerung 21.6 die zugehörige Gramsche Matrix. Dann gilt:*

(a) *Die Bilinearform b ist genau dann symmetrisch, wenn die Matrix A_b^B symmetrisch ist.*

(b) *Im Fall $K = \mathbb{R}$ ist b genau dann positiv definit, wenn A_b^B positiv definit ist.*

Beweis. Es sei $B = (x_1, \dots, x_n)$.

(a) „$\Rightarrow$“: Ist b symmetrisch, so folgt natürlich sofort

$$(A_b^B)^\top = (b(x_j,x_i))_{i,j} = (b(x_i,x_j))_{i,j} = A_b^B.$$

„$\Leftarrow$“: Ist umgekehrt A_b^B symmetrisch, so gilt nach Folgerung 21.6 für alle $x,y \in V$

$$b(x,y) = \Phi_B(x)^\top A_b^B \Phi_B(y) \stackrel{(*)}{=} \Phi_B(y)^\top (A_b^B)^\top \Phi_B(x) = \Phi_B(y)^\top A_b^B \Phi_B(x) = b(y,x),$$

wobei wir in $(*)$ gemäß Lemma 15.7 (d) die transponierte 1×1-Matrix gebildet haben.

(b) Da die Koordinatenabbildung $\Phi_B \colon V \to \mathbb{R}^n$ ein Isomorphismus ist, ist die Matrix A_b^B genau dann positiv definit, wenn $\Phi_B(x)^\top A_b^B \Phi_B(x) > 0$ für alle $x \in V \backslash \{0\}$ gilt. Nach Folgerung 21.6 bedeutet dies genau $b(x,x) > 0$ für alle $x \in V \backslash \{0\}$, also dass b positiv definit ist. □

Bemerkung 21.12 (Invarianz von Symmetrie und positiver Definitheit). Sind $A, A' \in K^{n \times n}$ zwei quadratische Matrizen mit $A' = T^\top A\, T$ für ein $T \in \mathrm{GL}(n,K)$, so besagt Lemma 21.11 insbesondere, dass A' genau dann symmetrisch (bzw. im Fall $K = \mathbb{R}$ positiv definit) ist, wenn dies für A gilt: A' und A beschreiben nach Lemma 21.7 nämlich die gleiche Bilinearform bezüglich zweier evtl. verschiedener Basen, und nach Lemma 21.11 hängt es nur von dieser Bilinearform (aber eben nicht von der gewählten Basis) ab, ob die Matrix symmetrisch bzw. positiv definit ist.

Mit Hilfe der eingeführten Konzepte können wir nun Skalarprodukte auf reellen Vektorräumen definieren.

Definition 21.13 (Skalarprodukte). Es sei V ein $\mathbb{R}$-Vektorraum. Ein **Skalarprodukt** auf V ist eine positiv definite, symmetrische Bilinearform $b \colon V \times V \to \mathbb{R}$. Ein $\mathbb{R}$-Vektorraum V zusammen mit einem Skalarprodukt heißt ein **euklidischer Raum**.

Für $x,y \in V$ schreiben wir statt $b(x,y)$ dann auch $\langle x,y \rangle$. Die (wegen der positiven Definitheit existierende) Zahl

$$||x|| := \sqrt{\langle x,x \rangle} \quad \in \mathbb{R}_{\geq 0}$$

heißt in Verallgemeinerung von Beispiel 21.1 die **Norm** oder **Länge** von x (bezüglich b). Man nennt einen Vektor $x \in V$ **normiert** (bezüglich b), falls $||x|| = 1$ gilt.

Bemerkung 21.14.

(a) Ist V ein endlich erzeugter $\mathbb{R}$-Vektorraum und B eine Basis von V, so lässt sich ein Skalarprodukt auf V nach Lemma 21.11 also genau durch eine positiv definite, symmetrische Matrix $A \in \mathbb{R}^{n \times n}$ beschreiben bzw. definieren (nämlich durch die Gramsche Matrix des Skalarprodukts bezüglich der Basis B).

(b) Die Einschränkung eines Skalarprodukts auf einen Untervektorraum ist offensichtlich wieder ein Skalarprodukt.

Beispiel 21.15.

(a) Ist $A = \operatorname{diag}(\lambda_1, \dots, \lambda_n) \in \mathbb{R}^{n \times n}$ eine Diagonalmatrix, so ist A natürlich zunächst einmal symmetrisch. Gilt nun zusätzlich $\lambda_i > 0$ für alle $i = 1, \dots, n$, so ist A auch positiv definit, denn für alle $x \in \mathbb{R}^n \backslash \{0\}$ ist

$$x^\mathsf{T} A x = \sum_{i=1}^{n} \lambda_i x_i^2 > 0,$$

da in dieser Summe alle Terme nicht-negativ sind und mindestens einer positiv ist. Also ist die zugehörige Bilinearform

$$\langle x, y \rangle = x^\mathsf{T} A y = \sum_{i=1}^{n} \lambda_i x_i y_i$$

ein Skalarprodukt auf $\mathbb{R}^n$. Speziell für $A = E$ erhalten wir daraus die schon in Beispiel 21.5 (b) betrachtete Bilinearform

$$\langle x, y \rangle = x^\mathsf{T} y = \sum_{i=1}^{n} x_i y_i.$$

Sie ist die direkte Verallgemeinerung von Beispiel 21.1 und wird das **Standardskalarprodukt** auf $\mathbb{R}^n$ (und nach Bemerkung 21.14 (b) auch auf Unterräumen von $\mathbb{R}^n$) genannt.

Gilt analog $\lambda_i < 0$ / $\lambda_i \geq 0$ / $\lambda_i \leq 0$ für alle $i = 1, \dots, n$, so ist die oben konstruierte Bilinearform negativ definit / positiv semidefinit / negativ semidefinit. Sie ist indefinit, falls es unter den $\lambda_1, \dots, \lambda_n$ eine positive und eine negative Zahl gibt.

(b) Die in Beispiel 21.5 (a) schon betrachtete reelle Matrix $A = \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix}$ ist positiv definit: Für alle $x \in \mathbb{R}^2$ ist zunächst

$$x^\mathsf{T} A x = x_1^2 + x_1 x_2 + x_2 x_1 + 4x_2^2 = (x_1 + x_2)^2 + 3x_2^2 \geq 0,$$

und die Gleichheit kann hier nur gelten für $x_1 + x_2 = x_2 = 0$, also für $x = 0$. Da A auch symmetrisch ist, ist die durch diese Matrix (bezüglich der Standardbasis) definierte Bilinearform

$$\langle x, y \rangle = x^\mathsf{T} A y = x_1 y_1 + x_1 y_2 + x_2 y_1 + 4x_2 y_2$$

also ein Skalarprodukt auf $\mathbb{R}^2$.

(c) Betrachten wir reelle $m \times n$-Matrizen über ihre Einträge als Vektoren in $\mathbb{R}^{mn}$, so können wir das Standardskalarprodukt zweier Matrizen $A = (a_{i,j})_{i,j}$ und $B = (b_{i,j})_{i,j}$ im Vektorraum $\mathbb{R}^{m \times n}$ mit der Definition des Matrixprodukts und der Spur schreiben als

$$\langle A, B \rangle = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{i,j} b_{i,j} = \operatorname{Spur}\left(A^\mathsf{T} B\right).$$

Wir werden dies in Zukunft als das *Standardskalarprodukt auf Matrizenräumen* wählen.

Konstruktion 21.16 (Skalarprodukte auf Funktionenräumen)**.** Auch auf unendlich-dimensionalen Vektorräumen sind Skalarprodukte eine sehr nützliche Zusatzstruktur. Ein besonders wichtiges Beispiel dafür ist wie in Definition 11.7 der reelle Vektorraum $V = C^0([a,b])$ der stetigen Funktionen auf einem abgeschlossenen Intervall $[a,b]$ mit $a < b$. Wir behaupten, dass dann

$$\langle f, g \rangle := \int_a^b f(x) g(x)\, dx \qquad (*)$$

ein Skalarprodukt auf V definiert. Die Existenz dieses Integrals ergibt sich hierbei sofort aus Satz 12.12. Auch die Bilinearität folgt direkt aus den Eigenschaften des Integrals in Satz 12.13; z. B. ist für $f_1, f_2, g \in V$

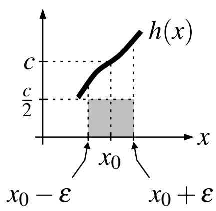
$$\begin{aligned}\langle f_1+f_2, g\rangle &= \int_a^b (f_1(x)+f_2(x))g(x)\,dx = \int_a^b f_1(x)g(x)\,dx + \int_a^b f_2(x)g(x)\,dx \\ &= \langle f_1, g\rangle + \langle f_2, g\rangle\,.\end{aligned}$$

Die Symmetrie ist natürlich offensichtlich. Für die positive Definitheit müssen wir zeigen, dass

$$\langle f, f\rangle = \int_a^b f(x)^2\,dx > 0$$

für alle $f \in V\backslash\{0\}$ gilt. Dies ergibt sich aus der folgenden etwas allgemeineren Aussage, die wir später noch mehrmals verwenden werden: Ist $h\colon [a,b] \to \mathbb{R}_{\geq 0}$ eine stetige Funktion, die an mindestens einer Stelle ungleich 0 ist (in unserem Fall $h(x) = f(x)^2$), so ist $\int_a^b h(x) > 0$.

Um diese Aussage zu zeigen, sei also $x_0 \in [a,b]$ mit $c := h(x_0) > 0$. Da die Funktion h nach Voraussetzung stetig ist, gibt es dann nach Bemerkung 8.8 wie im Bild rechts ein $\varepsilon \in \mathbb{R}_{>0}$, so dass $h(x) > \frac{c}{2}$ für alle $x \in [a,b]$ mit $|x - x_0| \leq \varepsilon$ gilt. Unter dem Graphen von h liegt also sicher ein positives Flächenstück (der Breite 2ε und Höhe $\frac{c}{2}$, falls x_0 nicht zu weit am Rand des Intervalls $[a,b]$ liegt). Da die Funktion h nirgends negativ ist, folgt damit wie behauptet $\int_a^b h(x)\,dx > 0$.



Damit definiert die Formel $(*)$ ein Skalarprodukt auf V – und somit nach Bemerkung 21.14 (b) auch auf allen Unterräumen davon, wie z. B. dem Unterraum $C^1([a,b])$ aller stetig differenzierbaren Funktionen oder dem Unterraum $\mathrm{Pol}([a,b],\mathbb{R})$ aller Polynomfunktionen auf $[a,b]$. Beachte, dass diese Formel ganz analog zum Standardskalarprodukt auf $\mathbb{R}^n$ in Beispiel 21.15 (a) ist, wenn wir die Summe über die Koordinaten von $\mathbb{R}^n$ durch ein Integral über alle Punkte im Definitionsintervall $[a,b]$ ersetzen. Auch die Interpretation z. B. der Norm $\|f\| = \sqrt{\langle f, f\rangle}$ ist in dem Sinne analog, dass diese Norm klein ist, wenn die Funktion „nur wenig von der Nullfunktion abweicht“. Man bezeichnet dieses Skalarprodukt daher auch als *Standardskalarprodukt auf* $C^0([a,b])$.

Beachte, dass die Stetigkeit der Funktionen für die Existenz dieses Skalarprodukts entscheidend ist: Wäre z. B. für einen gegebenen Punkt $x_0 \in [a,b]$ die unstetige Funktion

$$f\colon [a,b] \to \mathbb{R},\ x \mapsto \begin{cases} 0 & \text{für } x \neq x_0, \\ 1 & \text{für } x = x_0 \end{cases}$$

zugelassen, so wäre hier $\langle f, f\rangle = \int_a^b f(x)^2\,dx = 0$, obwohl f nicht die Nullfunktion ist. Wir hätten in diesem Fall also nur eine positiv semidefinite Bilinearform, aber kein Skalarprodukt (siehe auch Aufgabe 21.23 (c)).

Aufgabe 21.17. Es sei $n \in \mathbb{N}_{>0}$. Untersuche, ob die folgenden Abbildungen b Skalarprodukte auf dem reellen Vektorraum V sind:

(a) $V = \mathbb{R}^n$, $b(x,y) = x^{\mathsf{T}} A\, y$ mit $A = \begin{pmatrix} 2 & 1 & & & \\ 1 & 2 & 1 & & 0 \\ & \ddots & \ddots & \ddots & \\ & 0 & 1 & 2 & 1 \\ & & & 1 & 2 \end{pmatrix}$;

(b) $V = \mathbb{R}^{n\times n}$, $b(A,B) = \mathrm{Spur}(AB)$.

Im Rest dieses Kapitels wollen wir nun die grundlegenden Eigenschaften von euklidischen Räumen untersuchen. Da es in der Praxis öfters einmal vorkommt, werden wir den Begriff des Skalarprodukts aber zunächst noch auf den Fall von komplexen Vektorräumen erweitern.

Konstruktion 21.18 (Skalarprodukte im komplexen Fall)**.** Wollen wir die Definition 21.13 auf einen komplexen Vektorraum übertragen, so haben wir das Problem, dass die Bedingung der positiven Definitheit über $\mathbb{C}$ zunächst einmal keinen Sinn ergibt, da $\mathbb{C}$ kein geordneter Körper ist. Dies hat zur

Folge, dass wir die Norm eines Vektors nicht mehr wie gewohnt definieren können: Würden wir wie beim Standardskalarprodukt im Reellen auch in $\mathbb{C}^n$ die Formeln

$$\langle x,y\rangle = \sum_{i=1}^{n} x_i y_i \quad \text{und} \quad \|x\| = \sqrt{\langle x,x\rangle} = \sqrt{x_1^2+\cdots+x_n^2}$$

verwenden, so müssten wir hier die Wurzel aus einer im Allgemeinen komplexen Zahl $x_1^2+\cdots+x_n^2$ bilden – was nicht eindeutig möglich ist und auch nicht wie gewünscht zu einer nicht-negativen reellen Zahl als Länge eines Vektors führen würde. Die Lösung dieses Problems besteht darin, im Skalarprodukt grundsätzlich jede Koordinate *eines* Eintrags komplex zu konjugieren. Es spielt dabei keine Rolle, welchen Eintrag wir dafür wählen – wir werden in dieser Vorlesung immer den ersten nehmen, aber auch die umgekehrte Konvention des zweiten Eintrags ist in der Literatur zu finden. Durch diese komplexe Konjugation erhalten wir z. B. für das Standardskalarprodukt die Formeln

$$\langle x,y\rangle = \sum_{i=1}^{n} \overline{x_i}\, y_i \quad \text{und damit} \quad \|x\| = \sqrt{\langle x,x\rangle} = \sqrt{\overline{x_1}\,x_1+\cdots+\overline{x_n}\,x_n} = \sqrt{|x_1|^2+\cdots+|x_n|^2},$$

was wieder zu einer reellen, nicht-negativen Länge eines Vektors führt.

Mit dem Hintergrund dieser Idee sind die entsprechenden Abänderungen für beliebige Skalarprodukte auf komplexen Vektorräumen relativ offensichtlich. Wir werden die sich daraus ergebenden Definitionen und Resultate im Folgenden nur kurz auflisten; die Beweise dieser Aussagen sind völlig analog zu denen im reellen Fall.

Es sei also V ein $\mathbb{C}$-Vektorraum. Eine **Sesquilinearform** auf V ist eine Abbildung $s\colon V\times V\to\mathbb{C}$, so dass für alle $x_1,x_2,x,y_1,y_2,y\in V$ und $\lambda\in\mathbb{C}$ die Eigenschaften

$$\begin{aligned} s(x_1+x_2,y) &= s(x_1,y)+s(x_2,y),\\ s(\lambda x,y) &= \overline{\lambda}\, s(x,y),\\ s(x,y_1+y_2) &= s(x,y_1)+s(x,y_2),\\ s(x,\lambda y) &= \lambda\, s(x,y) \end{aligned}$$

gelten. Der einzige Unterschied zu Bilinearformen besteht also in der komplexen Konjugation des Skalars in der zweiten Zeile oben – und in der Tat kommt der Begriff „Sesquilinearform“ aus dem Lateinischen und bedeutet „eineinhalbfach lineare Form“.

Die Sesquilinearformen auf V bilden einen $\mathbb{C}$-Vektorraum, den wir mit $\mathrm{SLF}(V)$ bezeichnen wollen. Ist V endlich-dimensional und $B=(x_1,\dots,x_n)$ eine Basis von V, so ist $\mathrm{SLF}(V)$ wie in Folgerung 21.6 isomorph zu $\mathbb{C}^{n\times n}$ über die beiden zueinander inversen Isomorphismen

$$\begin{aligned} &\mathbb{C}^{n\times n}\to \mathrm{SLF}(V),\ A\mapsto s_A^B && \text{mit } s_A^B(x,y) := \overline{\Phi_B(x)}^{\mathsf{T}} A\, \Phi_B(y)\\ \text{und}\quad &\mathrm{SLF}(V)\to\mathbb{C}^{n\times n},\ s\mapsto A_s^B && \text{mit } A_s^B := (s(x_i,x_j))_{i,j}, \end{aligned}$$

wobei der Querstrich über $\Phi_B(x)\in\mathbb{C}^n$ bedeutet, dass jede Komponente des Vektors komplex konjugiert wird. Man bezeichnet A_s^B wieder als die *Gramsche Matrix* von s. Ist B' eine weitere Basis von V, so transformieren sich die Gramschen Matrizen analog zu Lemma 21.7 gemäß

$$A_s^{B'} = \overline{T}^{\mathsf{T}} A_s^B\, T \quad \text{mit } T = A^{B',B},$$

wobei $T=A^{B',B}$ die übliche Basiswechselmatrix ist.

Eine Sesquilinearform s auf einem $\mathbb{C}$-Vektorraum V heißt **hermitesch**, wenn $s(y,x)=\overline{s(x,y)}$ für alle $x,y\in V$. Ist dies der Fall, so erhalten wir daraus insbesondere mit $y=x$

$$s(x,x)=\overline{s(x,x)}, \quad \text{also} \quad s(x,x)\in\mathbb{R}$$

für alle $x\in V$. Wir können daher fragen, ob diese reelle Zahl immer nicht-negativ ist, und nennen eine hermitesche Sesquilinearform s analog zum reellen Fall **positiv definit**, wenn $s(x,x)>0$ für alle $x\in V$ mit $x\neq 0$.

Entsprechend heißt eine Matrix $A \in \mathbb{C}^{n\times n}$ **hermitesch**, wenn $\overline{A}^{\mathsf{T}} = A$. In diesem Fall ist $\overline{x}^{\mathsf{T}} A\, x \in \mathbb{R}$ für alle $x \in \mathbb{C}^n$, denn es ist

$$\overline{\overline{x}^{\mathsf{T}} A\, x} = x^{\mathsf{T}} \overline{A}\, \overline{x} \overset{(*)}{=} \overline{x}^{\mathsf{T}} \overline{A}^{\mathsf{T}} x = \overline{x}^{\mathsf{T}} A\, x,$$

wobei in $(*)$ nach Lemma 15.7 (d) die transponierte 1×1-Matrix gebildet wurde. Gilt sogar $\overline{x}^{\mathsf{T}} A\, x > 0$ für alle $x \in \mathbb{C}^n$ mit $x \neq 0$, so heißt A **positiv definit**. Wie in Lemma 21.11 ist eine Sesquilinearform auf einem endlich-dimensionalen Vektorraum genau dann hermitesch, bzw. eine hermitesche Sesquilinearform genau dann positiv definit, wenn ihre Gramsche Matrix zu einer beliebigen Basis diese Eigenschaft besitzt. Dabei ist eine hermitesche Diagonalmatrix immer reell, und wie in Beispiel 21.15 (a) genau dann positiv definit, wenn alle Diagonaleinträge positiv sind. Die anderen Definitheitsbegriffe aus Bemerkung 21.10 (b) definiert man natürlich sowohl für hermitesche Sesquilinearformen als auch für hermitesche Matrizen analog.

Ein **Skalarprodukt** auf einem $\mathbb{C}$-Vektorraum V ist nun eine positiv definite, hermitesche Sesquilinearform s auf V, die wir dann wieder als $\langle x,y\rangle := s(x,y)$ schreiben. Ein $\mathbb{C}$-Vektorraum zusammen mit einem Skalarprodukt heißt ein **unitärer Raum**. Für $x \in V$ nennen wir in diesem Fall wieder

$$\|x\| := \sqrt{\langle x,x\rangle} \quad \in \mathbb{R}_{\geq 0}$$

die *Norm* (oder *Länge*) von x.

Das **Standardskalarprodukt** auf $\mathbb{C}^n$ ist dasjenige, das bezüglich der Standardbasis der Einheitsmatrix entspricht, also

$$\langle x,y\rangle = \overline{x}^{\mathsf{T}} y = \overline{x_1}\, y_1 + \cdots + \overline{x_n}\, y_n,$$

wobei $x_1,\dots,x_n$ und $y_1,\dots,y_n$ die Koordinaten von x bzw. y sind. Beachte, dass die Norm eines Vektors $x \in \mathbb{C}^n$ in diesem Fall dann

$$\|x\| = \sqrt{\langle x,x\rangle} = \sqrt{|x_1|^2 + \cdots + |x_n|^2} = \sqrt{(\operatorname{Re} x_1)^2 + (\operatorname{Im} x_1)^2 + \cdots + (\operatorname{Re} x_n)^2 + (\operatorname{Im} x_n)^2},$$

und damit gleich seiner Norm in $\mathbb{R}^{2n}$ bezüglich des Standardskalarprodukts ist. Das Standardskalar-
50 produkt für Matrizen $A,B \in \mathbb{C}^{m\times n}$ ist $\langle A,B\rangle = \operatorname{Spur}\big(\overline{A}^{\mathsf{T}} B\big)$ analog zu Beispiel 21.15 (c).

Im Folgenden wollen wir den reellen und komplexen Fall in der Regel zusammen behandeln. Wie in der Analysis schreiben wir daher wieder $\mathbb{K}$ für einen der Körper $\mathbb{R}$ oder $\mathbb{C}$, und sprechen von einem *Vektorraum mit Skalarprodukt*, wenn wir einen euklidischen bzw. unitären Raum meinen. Die Formeln werden wir dabei wie in Konstruktion 21.18 mit der komplexen Konjugation schreiben, so dass sie dann für $\mathbb{K} = \mathbb{R}$ und $\mathbb{K} = \mathbb{C}$ gleichermaßen gelten – im reellen Fall ist die komplexe Konjugation dann zwar unnötig, sie schadet aber natürlich auch nicht.

Als erstes Resultat über Skalarprodukte wollen wir nun eine sehr wichtige Ungleichung beweisen.

Satz 21.19 (**Cauchy-Schwarz-Ungleichung**). *In jedem Vektorraum V mit Skalarprodukt gilt*

$$|\langle x,y\rangle| \leq \|x\| \cdot \|y\|$$

für alle $x,y \in V$, wobei die Gleichheit genau dann gilt, wenn x und y linear abhängig sind.

Beweis. Für $y = 0$ ist die Aussage des Satzes offensichtlich, denn dann sind beide Seiten gleich Null und die Vektoren linear abhängig. Andernfalls setzen wir $\lambda := \frac{\langle y,x\rangle}{\langle y,y\rangle}$ (und damit $\overline{\lambda} = \frac{\langle x,y\rangle}{\langle y,y\rangle}$ wegen der Symmetrie bzw. Hermitizität des Skalarprodukts) und folgern aus der positiven Definitheit, dass

$$\begin{aligned}
0 &\leq \langle x-\lambda y, x-\lambda y\rangle \qquad (*)\\
&= \langle x,x\rangle - \lambda\,\langle x,y\rangle - \overline{\lambda}\,\langle y,x\rangle + \lambda\,\overline{\lambda}\,\langle y,y\rangle\\
&= \langle x,x\rangle - \frac{\langle y,x\rangle\,\langle x,y\rangle}{\langle y,y\rangle} - \frac{\langle x,y\rangle\,\langle y,x\rangle}{\langle y,y\rangle} + \frac{\langle x,y\rangle\,\langle y,x\rangle}{\langle y,y\rangle}\\
&= \langle x,x\rangle - \frac{\langle x,y\rangle\,\langle y,x\rangle}{\langle y,y\rangle}\\
&= \|x\|^2 - \frac{|\langle x,y\rangle|^2}{\|y\|^2}
\end{aligned}$$

und damit $|\langle x,y\rangle|^2 \le \|x\|^2 \cdot \|y\|^2$ gilt. Wurzelziehen liefert nun die behauptete Ungleichung.

Gilt in dieser Ungleichung sogar die Gleichheit, gilt also die Gleichheit in $(*)$, so ergibt sich aus der positiven Definitheit des Skalarprodukts sofort $x - \lambda y = 0$, d. h. x und y sind linear abhängig. Sind umgekehrt x und y linear abhängig, gilt also $x = \mu\, y$ für ein $\mu \in \mathbb{K}$, so ist $\langle y,x\rangle = \langle y,\mu\, y\rangle = \mu\,\langle y,y\rangle$ und damit $\lambda = \frac{\langle y,x\rangle}{\langle y,y\rangle} = \mu$, so dass in $(*)$ oben $x - \lambda y = x - \mu y = 0$ ist und damit sowohl dort als auch in der Cauchy-Schwarz-Ungleichung sogar die Gleichheit gilt. □

Mit der Cauchy-Schwarz-Ungleichung können wir die folgenden Eigenschaften der Norm herleiten:

Satz 21.20 (Eigenschaften der Norm)**.** *In jedem Vektorraum V mit Skalarprodukt gilt für alle $x,y \in V$ und $\lambda \in \mathbb{K}$:*

(a) $\|\lambda x\| = |\lambda| \cdot \|x\|$*;*

(b) $\|x\| > 0$ *für alle* $x \neq 0$*;*

(c) $\|x+y\| \le \|x\| + \|y\|$ *(**Dreiecksungleichung**).*

Beweis.

(a) Es ist $\|\lambda x\| = \sqrt{\langle \lambda x, \lambda x\rangle} = \sqrt{\overline{\lambda}\lambda\,\langle x,x\rangle} = \sqrt{|\lambda|^2\,\langle x,x\rangle} = |\lambda| \cdot \|x\|$.

(b) folgt sofort aus der positiven Definitheit des Skalarprodukts.

(c) Es gilt

$$\begin{aligned}
\|x+y\|^2 &= \langle x+y, x+y\rangle \\
&= \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle \\
&= \|x\|^2 + \langle x,y\rangle + \overline{\langle x,y\rangle} + \|y\|^2 \\
&= \|x\|^2 + 2\,\mathrm{Re}\,\langle x,y\rangle + \|y\|^2 && \text{(Bemerkung 6.4)} \\
&\le \|x\|^2 + 2\,|\langle x,y\rangle| + \|y\|^2 && \text{(Lemma 6.9 (b))} \\
&\le \|x\|^2 + 2\,\|x\| \cdot \|y\| + \|y\|^2 && \text{(Satz 21.19)} \\
&= (\|x\| + \|y\|)^2,
\end{aligned}$$

woraus durch Wurzelziehen die Behauptung folgt. □

Bemerkung 21.21 (Normen in der linearen Algebra und Analysis)**.** In der Analysis werden wir später normierte Vektorräume definieren als reelle oder komplexe Vektorräume mit einer reellwertigen „Normabbildung" $\|\cdot\|$, die die drei Eigenschaften aus Satz 21.20 erfüllt (siehe Definition 23.1). In diesem Sinne sind Vektorräume mit Skalarprodukt also immer normierte Vektorräume aus der Sicht der Analysis. Wir werden allerdings sehen, dass es auch sehr viele Normen gibt, die nicht von einem Skalarprodukt kommen, z. B. in $\mathbb{R}^2$ die Summennorm $\|x\| = |x_1| + |x_2|$ oder die Maximumsnorm $\|x\| = \max\{|x_1|, |x_2|\}$.

Wir kommen nun zum Winkel zwischen zwei Vektoren. Man definiert ihn in der Regel nur im reellen Fall und orientiert sich dabei an der geometrischen Deutung aus Beispiel 21.1 im Fall des Standardskalarprodukts.

Konstruktion 21.22 (Winkel)**.** Es seien V ein euklidischer Raum und $x,y \in V\backslash\{0\}$. Nach der Cauchy-Schwarz-Ungleichung aus Satz 21.19 ist dann

$$\frac{|\langle x,y\rangle|}{\|x\| \cdot \|y\|} \le 1, \qquad \text{also} \qquad -1 \le \frac{\langle x,y\rangle}{\|x\| \cdot \|y\|} \le 1.$$

Die Zahl

$$\varphi = \arccos \frac{\langle x,y\rangle}{\|x\| \cdot \|y\|} \quad \in [0,\pi]$$

ist daher wohldefiniert; wir nennen sie in Analogie zu Beispiel 21.1 den (unorientierten) **Winkel** zwischen x und y.

Aufgabe 21.23 (Skalarprodukte aus positiv semidefiniten Formen)**.** Zu einer symmetrischen Bilinearform b auf einem reellen Vektorraum V sei $U_b = \{x \in V : b(x,x) = 0\}$. Man zeige:

(a) U_b ist im Allgemeinen kein Unterraum von V.

(b) Ist b jedoch positiv semidefinit, so ist U_b ein Unterraum, und $\overline{b}(\overline{x},\overline{y}) := b(x,y)$ ist ein wohldefiniertes Skalarprodukt auf V/U_b.

(c) Was ergibt sich aus dieser Konstruktion, wenn V der Vektorraum aller stückweise stetigen Funktionen auf einem Intervall $[a,b]$ und $b(f,g) = \int_a^b f(x)g(x)\,dx$ ist? Wie kann man sich in diesem Fall die Elemente von U_b und V/U_b anschaulich vorstellen?

Aufgabe 21.24. Es sei $f\colon V \to V$ ein Endomorphismus eines $\mathbb{K}$-Vektorraums mit Skalarprodukt, so dass $\langle f(x), x\rangle = 0$ für alle $x \in V$. Man zeige:

(a) Ist V ein unitärer Raum (also $\mathbb{K} = \mathbb{C}$), so ist f die Nullabbildung.

(b) Ist V ein euklidischer Raum (also $\mathbb{K} = \mathbb{R}$), so ist f nicht notwendig die Nullabbildung.

21.C Orthogonalität

Der mit Abstand wichtigste Fall der Winkeldefinition ist derjenige, in dem die beiden betrachteten Vektoren „senkrecht aufeinander stehen", also dieser Winkel gleich $\frac{\pi}{2}$ und nach Konstruktion 21.22 damit das Skalarprodukt gleich 0 ist. Im Gegensatz zu allgemeinen Winkeln können wir diesen Spezialfall auch wieder sowohl für reelle als auch für komplexe Vektorräume definieren.

Definition 21.25 (Orthogonale Vektoren)**.** Es sei V ein Vektorraum mit Skalarprodukt.

(a) Zwei Vektoren $x,y \in V$ heißen **orthogonal** bzw. **senkrecht** zueinander (in Zeichen: $x \perp y$), wenn $\langle x,y\rangle = 0$ gilt. Für $x \in V$ und einen Unterraum $U \leq V$ schreiben wir kurz $x \perp U$, falls $x \perp y$ für alle $y \in U$.

(b) Eine Familie $B = (x_1,\dots,x_n)$ von Vektoren in V heißt **orthogonal**, wenn $x_i \perp x_j$ für alle $i \neq j$ gilt, also wenn die gegebenen Vektoren paarweise zueinander senkrecht sind. Gilt zusätzlich $\|x_i\| = 1$ für alle i, so nennt man B **orthonormal**.

(c) Eine **Orthogonalbasis** bzw. **Orthonormalbasis** von V ist eine orthogonale bzw. orthonormale Familie, die gleichzeitig eine Basis von V ist.

Beispiel 21.26.

(a) Nach Definition ist der Nullvektor orthogonal zu jedem anderen Vektor.

(b) Die Standardbasis von $\mathbb{R}^n$ bzw. $\mathbb{C}^n$ ist natürlich eine Orthonormalbasis bezüglich des Standardskalarprodukts, da $\langle e_i, e_j\rangle = 0$ und $\|e_i\| = 1$ für alle $i,j = 1,\dots,n$ gilt.

Bemerkung 21.27.

(a) Die Orthogonalitätsrelation ist symmetrisch: Gilt $x \perp y$ für zwei Vektoren x und y in einem Vektorraum mit Skalarprodukt, ist also $\langle x,y\rangle = 0$, so folgt auch $\langle y,x\rangle = \overline{\langle x,y\rangle} = \overline{0} = 0$ und damit $y \perp x$.

(b) Es sei U ein Unterraum eines Vektorraums V mit Skalarprodukt. Sind dann $x \in V$ und $(x_1,\dots,x_n)$ eine Basis von U, so genügt es für die Bedingung $x \perp U$ zu überprüfen, dass $x \perp x_i$ für alle $i = 1,\dots,n$: In diesem Fall gilt nämlich $\langle x,x_i\rangle = 0$ für alle $i = 1,\dots,n$, also auch
$$\langle x, \lambda_1 x_1 + \cdots + \lambda_n x_n\rangle = \lambda_1 \langle x,x_1\rangle + \cdots + \lambda_n \langle x,x_n\rangle = 0$$
für alle $\lambda_1,\dots,\lambda_n \in \mathbb{K}$, und damit $\langle x,y\rangle = 0$ für alle $y \in U$.

(c) Natürlich ist jede orthonormale Familie auch orthogonal. Umgekehrt können wir aus einer orthogonalen Familie, die nicht den Nullvektor enthält, immer eine orthonormale Familie machen, indem wir jeden ihrer Vektoren normieren, also durch seine Länge dividieren.

Eine erste sehr nützliche Eigenschaft der Orthogonalität einer Familie ist, dass daraus bereits ihre lineare Unabhängigkeit folgt:

Lemma 21.28. *In einem Vektorraum mit Skalarprodukt ist jede orthogonale Familie, die nicht den Nullvektor enthält, linear unabhängig.*

Insbesondere ist sie damit also stets eine Orthogonalbasis des von ihr aufgespannten Unterraums.

Beweis. Es sei $B = (x_1, \dots, x_n)$ orthogonal mit $x_i \neq 0$ für alle i. Weiterhin sei $\lambda_1 x_1 + \cdots + \lambda_n x_n = 0$ mit $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ eine Linearkombination des Nullvektors. Bilden wir dann das Skalarprodukt von x_k für ein $k \in \{1, \dots, n\}$ mit dieser Gleichung, so folgt

$$0 = \langle x_k, \lambda_1 x_1 + \cdots + \lambda_n x_n \rangle = \lambda_1 \langle x_k, x_1 \rangle + \cdots + \lambda_n \langle x_k, x_n \rangle = \lambda_k \|x_k\|^2,$$

weil B orthogonal ist und damit $\langle x_k, x_i \rangle$ für alle $i \neq k$ gilt. Da weiterhin nach Voraussetzung $x_k \neq 0$ und damit aufgrund der positiven Definitheit $\|x_k\| \neq 0$ ist, folgt $\lambda_k = 0$. Dies gilt aber für alle k, d. h. die ursprüngliche Linearkombination ist trivial, und B damit linear unabhängig. □

Wir wollen nun sehen, dass Orthogonalbasen (und damit nach Bemerkung 21.27 (c) auch Orthonormalbasen) in jedem endlich-dimensionalen Vektorraum mit Skalarprodukt existieren und zu besonders schönen Eigenschaften führen. In der Tat gilt sogar noch mehr, nämlich eine „Basisergänzungseigenschaft" analog zu Folgerung 14.15. Um dies zu beweisen, brauchen wir die folgende Konstruktion der orthogonalen Projektion auf einen Unterraum.

Satz und Definition 21.29 (Orthogonale Projektionen)**.** *Es sei U ein endlich-dimensionaler Unterraum eines Vektorraums V mit Skalarprodukt. Ferner sei $(x_1, \dots, x_n)$ eine Orthogonalbasis von U (wir werden in Satz 21.31 noch sehen, dass eine solche Orthogonalbasis immer existiert). Dann gilt:*

(a) *Es gibt genau eine lineare Abbildung $f\colon V \to U$ mit $x - f(x) \perp U$ für alle $x \in V$, nämlich*

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

*Man nennt sie die **orthogonale Projektion** von V auf U.*

(b) *Die orthogonale Projektion $f(x)$ von x aus (a) ist der eindeutig bestimmte Punkt in U mit kleinstem Abstand zu x, d. h. es gilt $\|y - x\| > \|f(x) - x\|$ für alle $y \in U$ mit $y \neq f(x)$.*

Beweis.

(a) Für ein zunächst festes $x \in V$ seien $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ die Koordinaten des Vektors $f(x) \in U$ bezüglich der gegebenen Basis $(x_1, \dots, x_n)$, also $f(x) = \lambda_1 x_1 + \cdots + \lambda_n x_n$. Dann gilt

$$\begin{aligned} x - f(x) \perp U &\overset{21.27\,(b)}{\Leftrightarrow} \langle x_i, x - f(x) \rangle = 0 \text{ für alle } i = 1, \dots, n \\ &\Leftrightarrow \langle x_i, x - \lambda_1 x_1 - \cdots - \lambda_n x_n \rangle = 0 \text{ für alle } i = 1, \dots, n \\ &\overset{(*)}{\Leftrightarrow} \langle x_i, x \rangle - \lambda_i \langle x_i, x_i \rangle = 0 \text{ für alle } i = 1, \dots, n, \\ &\Leftrightarrow \lambda_i = \frac{\langle x_i, x \rangle}{\|x_i\|^2} \text{ für alle } i = 1, \dots, n, \end{aligned}$$

wobei wir in $(*)$ die Orthogonalität der Basis verwendet haben. Die Bedingung $x - f(x) \perp U$ für alle $x \in V$ ist also äquivalent zur angegebenen Formel für f. Da diese Vorschrift offensichtlich auch linear in x ist, ist Teil (a) des Satzes damit gezeigt.

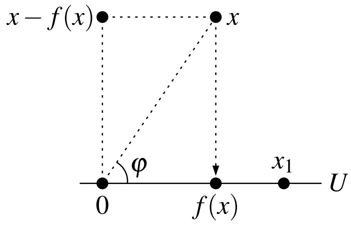
(b) Es sei $y \in U$ mit $y \neq f(x)$, also $y = f(x) + u$ für ein $u \in U \backslash \{0\}$. Dann gilt

$$\begin{aligned} \|y - x\|^2 &= \|f(x) + u - x\|^2 = \langle f(x) + u - x, f(x) + u - x \rangle \\ &= \underbrace{\langle f(x) - x, f(x) - x \rangle}_{= \|f(x) - x\|^2} + \underbrace{\langle u, u \rangle}_{= \|u\|^2} + \underbrace{\langle f(x) - x, u \rangle + \langle u, f(x) - x \rangle}_{= 0 \text{ wegen } f(x) - x \perp U}. \end{aligned}$$

Wegen der positiven Definitheit des Skalarprodukts ist aber $\|u\| > 0$, und damit wie behauptet $\|y - x\|^2 > \|f(x) - x\|^2$. □

Bemerkung 21.30 (Geometrische Deutung orthogonaler Projektionen)**.** Orthogonale Projektionen haben eine einfache geometrische Deutung, die im Bild unten rechts für einen eindimensionalen Unterraum $U = \mathrm{Lin}(x_1)$ von $V = \mathbb{R}^2$ dargestellt ist.

Die Abbildung konstruiert in diesem Fall zu einem $x \in V$ das Lot auf den Unterraum U; der so entstehende Punkt ist die orthogonale Projektion $f(x)$. Nach Konstruktion 21.22 ist dann nämlich



$$\|f(x)\| = \|x\| \cos\varphi = \|x\| \cdot \frac{\langle x_1, x\rangle}{\|x_1\| \cdot \|x\|} = \frac{\langle x_1, x\rangle}{\|x_1\|},$$

und damit wie in Satz 21.29

$$f(x) = \|f(x)\| \cdot \frac{x_1}{\|x_1\|} = \frac{\langle x_1, x\rangle}{\|x_1\|^2}\, x_1.$$

Beachte, dass $f(x)$ wie in Satz 21.29 auch anschaulich im Bild der Punkt in U ist, der am nächsten an x liegt, und dass die Differenz $x - f(x)$ senkrecht zu x_1 ist. Wir können also einen Vektor konstruieren, der auf (allen Vektoren von) U senkrecht steht, indem wir von einem beliebigen Vektor $x \notin U$ seine orthogonale Projektion auf U abziehen. Dies ist die Grundidee des folgenden Verfahrens zur Bestimmung von Orthogonalbasen.

Satz 21.31 (**Gram-Schmidtsches Orthogonalisierungsverfahren**)**.** *Es seien V ein endlich-dimensionaler Vektorraum mit Skalarprodukt und $U \le V$ ein Untervektorraum. Dann lässt sich jede Orthogonalbasis von U zu einer Orthogonalbasis von V ergänzen.*

Insbesondere besitzt also jeder endlich-dimensionale Vektorraum mit Skalarprodukt eine Orthogonalbasis (und damit nach Bemerkung 21.27 (c) auch eine Orthonormalbasis).

Beweis. Der Beweis dieses Satzes ist konstruktiv und erlaubt daher auch eine einfache (iterative) Konstruktion solcher Orthogonalbasen.

Es sei $(x_1, \dots, x_k)$ die gegebene Orthonormalbasis von U. Ist bereits $U = V$, so sind wir natürlich fertig. Ansonsten wählen wir einen beliebigen Vektor $x \in V \backslash U$ und setzen

$$x_{k+1} := x - \sum_{i=1}^{k} \frac{\langle x_i, x\rangle}{\|x_i\|^2}\, x_i.$$

Nach Satz 21.29 (a) ist dieser Vektor senkrecht zu U. Außerdem ist er wegen $x \notin U = \mathrm{Lin}(x_1, \dots, x_k)$ nicht der Nullvektor. Nach Lemma 21.28 ist $(x_1, \dots, x_{k+1})$ also eine Orthogonalbasis eines Unterraums $U' = \mathrm{Lin}(x_1, \dots, x_{k+1}) \supsetneq U$.

Ist jetzt $U' = V$, so sind wir fertig. Ansonsten setzen wir das obige Verfahren iterativ mit dem neuen
51 Unterraum U' fort, bis wir genügend Vektoren gefunden haben. □

Beispiel 21.32. Wir wollen eine Orthogonal- bzw. Orthonormalbasis von $\mathbb{R}^2$ für das Skalarprodukt

$$\langle x, y\rangle = x^{\mathsf{T}} \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix} y$$

aus Beispiel 21.15 (b) bestimmen. Den ersten Vektor können wir dabei (ungleich 0) beliebig wählen, z. B. $x_1 = e_1$. Für den zweiten Vektor starten wir mit einem beliebigen Element von $\mathbb{R}^2 \backslash \mathrm{Lin}(x_1)$, z. B. $x = e_2$, und subtrahieren davon wie im Beweis von Satz 21.31 seine orthogonale Projektion auf $\mathrm{Lin}(x_1)$. Mit

$$\langle x_1, e_2\rangle = (1\ \ 0) \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 1 \quad \text{und} \quad \|x_1\|^2 = \langle x_1, x_1\rangle = (1\ \ 0) \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1$$

erhalten wir so

$$x_2 = e_2 - \frac{\langle x_1, e_2\rangle}{\|x_1\|^2}\, x_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Die Familie (x_1,x_2) ist also eine Orthogonalbasis von $\mathbb{R}^2$ bezüglich des gegebenen Skalarprodukts. Um eine Orthonormalbasis zu bestimmen, müssen wir diese beiden Vektoren noch normieren: Wegen $\|x_1\|^2 = 1$ und

$$\|x_2\|^2 = \langle x_2, x_2 \rangle = (\,-1 \;\; 1\,) \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = 3$$

ist eine Orthonormalbasis

$$\left(\frac{x_1}{\|x_1\|}, \frac{x_2}{\|x_2\|} \right) = \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{3}} \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right).$$

Bemerkung 21.33 (Gram-Schmidt als Normalformensatz)**.** Es sei V ein endlich erzeugter Vektorraum mit Skalarprodukt $\langle x,y \rangle = b(x,y)$. Ist dann $B = (x_1,\dots,x_n)$ gemäß Satz 21.31 eine Orthonormalbasis von V, so ist die Gramsche Matrix von b bezüglich B nach Folgerung 21.6 gerade

$$A_b^B = \big(\langle x_i, x_j \rangle \big)_{i,j} = E_n,$$

denn es ist ja $\langle x_i, x_j \rangle$ gleich 1 für $i = j$ und 0 für $i \neq j$. In Analogie zu den Normalformaussagen aus Satz 17.29 und Folgerung 20.14 kann man die Existenz von Orthonormalbasen damit auch so auffassen, dass es zu jedem Skalarprodukt eine Basis gibt, bezüglich der die Gramsche Matrix die Einheitsmatrix ist, also diese sehr einfache Form hat.

Da sich Gramsche Matrizen unter Basiswechsel wie in Lemma 21.7 bzw. Konstruktion 21.18 verhalten, ist die entsprechende Aussage in Matrixform gemäß Lemma 21.11 also, dass es zu jeder positiv definiten, symmetrischen (im Fall $\mathbb{K} = \mathbb{R}$) bzw. hermiteschen (im Fall $\mathbb{K} = \mathbb{C}$) Matrix $A \in \mathbb{K}^{n \times n}$ eine invertierbare Matrix $T \in \mathrm{GL}(n,\mathbb{K})$ gibt mit $\overline{T}^{\mathsf{T}} A\, T = E$: Man muss in die Spalten von T einfach eine Orthonormalbasis von $\mathbb{K}^n$ bezüglich des Skalarprodukts $\langle x,y \rangle = \overline{x}^{\mathsf{T}} A\, y$ schreiben. Für das Skalarprodukt und die Orthonormalbasis aus Beispiel 21.32 ergibt sich also z. B.

$$T^{\mathsf{T}} A\, T = E \quad \text{mit} \quad A = \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix} \quad \text{und} \quad T = \begin{pmatrix} 1 & -\frac{1}{\sqrt{3}} \\ 0 & \frac{1}{\sqrt{3}} \end{pmatrix},$$

was man durch direkte Berechnung des Matrixprodukts auch sofort bestätigen kann.

Eine Verallgemeinerung dieser Aussage auf nicht notwendig positiv definite symmetrische Bilinearformen bzw. hermitesche Sesquilinearformen werden wir in Satz 22.35 bzw. Bemerkung 22.36 (b) kennenlernen.

Folgerung 21.34. *Für jede symmetrische bzw. hermitesche Matrix $A \in \mathbb{K}^{n \times n}$ gilt:*

(a) $\det A \in \mathbb{R}$.

(b) *Ist A zusätzlich positiv definit, so ist sogar* $\det A > 0$.

Beweis.

(a) Ist $\overline{A}^{\mathsf{T}} = A$, so folgt $\overline{\det A} = \det \overline{A} = \det \overline{A}^{\mathsf{T}} = \det A$ und damit $\det A \in \mathbb{R}$.

(b) Nach Bemerkung 21.33 gibt es eine invertierbare Matrix T mit $\overline{T}^{\mathsf{T}} A T = E$. Damit ist

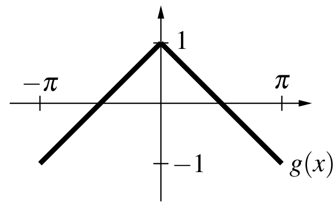
$$1 = \det E = \det \overline{T}^{\mathsf{T}} \cdot \det A \cdot \det T = \det A \cdot \det \overline{T} \cdot \det T = \det A \cdot |\det T|^2,$$

woraus $\det A > 0$ folgt. □

Aufgabe 21.35.

(a) Für welche $m,n \in \mathbb{N}$ definiert $\langle f,g \rangle := \sum_{i=0}^{m} f(i)\, g(i)$ ein Skalarprodukt auf $\mathrm{Pol}_n(\mathbb{R},\mathbb{R})$?

(b) Berechne für dieses Skalarprodukt eine Orthonormalbasis von $\mathrm{Pol}_2(\mathbb{R},\mathbb{R})$ im Fall $m = n = 2$.



Aufgabe 21.36. Es sei $V = C^0([-\pi,\pi])$ der reelle Vektorraum aller stetigen Funktionen auf dem Intervall $[-\pi,\pi]$ mit dem Standardskalarprodukt $\langle f,g\rangle := \int_{-\pi}^{\pi} f(x)g(x)\,dx$ wie in Konstruktion 21.16. Wir betrachten darin das Element $g \in V$ definiert durch $g(x) := 1 - \frac{2}{\pi}|x|$.

Für alle $n \in \mathbb{N}_{>0}$ seien weiterhin $f_n \in V$ mit $f_n(x) := \cos nx$, und $U_n := \mathrm{Lin}(f_1,\dots,f_n) \le V$.

(a) Zeige, dass $(f_1,\dots,f_n)$ für alle n eine Orthogonalbasis von U_n ist.

(b) Berechne für alle $n \in \mathbb{N}_{>0}$ die orthogonale Projektion g_n von g auf U_n (also die Funktion in U_n, die nach Satz 21.29 (b) von g den kleinsten Abstand hat, d. h. sie am besten approximiert).

(c) Zeichne die Funktionen g_n für kleine n mit einem Computer und vergleiche sie mit der ursprünglichen Funktion g.

Wir wollen nun noch drei interessante Anwendungen von Orthonormalbasen betrachten. Die erste betrifft Abbildungsmatrizen bzw. Basiswechselmatrizen: Bisher mussten wir zur Berechnung solcher Matrizen wie in Bemerkung 19.1 Koordinatendarstellungen von Vektoren berechnen und die dabei erhaltenen Koeffizienten in die Spalten einer Matrix schreiben. Dies erfordert die Lösung linearer Gleichungssysteme und ist daher recht rechenaufwendig. In einem Vektorraum mit Skalarprodukt ist die Berechnung solcher Matrizen bezüglich Orthonormalbasen dagegen deutlich einfacher:

Satz 21.37 (Abbildungsmatrizen zu Orthonormalbasen)**.** *Es seien V ein endlich-dimensionaler Vektorraum mit Skalarprodukt und $B = (x_1,\dots,x_n)$ eine Orthonormalbasis von V.*

(a) *Für alle $x \in V$ gilt*

$$x = \langle x_1,x\rangle\, x_1 + \cdots + \langle x_n,x\rangle\, x_n.$$

Mit anderen Worten ist in der Koordinatendarstellung $x = \lambda_1 x_1 + \cdots + \lambda_n x_n$ bezüglich B also $\lambda_i = \langle x_i,x\rangle$ für alle $i = 1,\dots,n$.

(b) *Ist $f\colon V \to V$ ein Endomorphismus, so ist die Abbildungsmatrix von f bezüglich B gleich $A_f^B = \big(\langle x_i, f(x_j)\rangle\big)_{i,j}$.*

(c) *Ist $B' = (y_1,\dots,y_n)$ eine weitere Basis von V, so ist die zugehörige Basiswechselmatrix gleich $A^{B',B} = \big(\langle x_i,y_j\rangle\big)_{i,j}$.*

Beweis.

(a) Bilden wir für $i = 1,\dots,n$ das Skalarprodukt von x_i mit der Gleichung $x = \lambda_1 x_1 + \cdots + \lambda_n x_n$, so erhalten wir wegen der Orthonormalität von B sofort $\langle x_i,x\rangle = \lambda_i \langle x_i,x_i\rangle = \lambda_i$.

(b) Nach Satz 16.26 ist der Eintrag in Zeile i und Spalte j von A_f^B genau die i-te Koordinate von $f(x_j)$ bezüglich B, nach (a) also $\langle x_i, f(x_j)\rangle$.

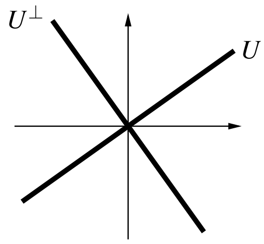
(c) Nach Definition 16.38 ist der Eintrag in Zeile i und Spalte j von $A^{B',B}$ genau die i-te Koordinate von y_j bezüglich B, nach (a) also $\langle x_i,y_j\rangle$. □

In der zweiten Anwendung von Orthonormalbasen geht es um Komplemente von Unterräumen in endlich erzeugten Vektorräumen: Wir wissen ja nach Beispiel 17.10 und Satz 17.11 bereits, dass solche Komplemente zwar immer existieren, aber in der Regel nicht eindeutig bestimmt sind. Falls im zugrundeliegenden Vektorraum aber ein Skalarprodukt gegeben ist, wollen wir jetzt sehen, dass es zu jedem Unterraum ein besonderes Komplement gibt – nämlich das sogenannte *orthogonale Komplement* – das immer eindeutig bestimmt ist.

Definition 21.38 (Orthogonales Komplement)**.** Es seien V ein Vektorraum mit Skalarprodukt und $U \leq V$ ein Unterraum von V. Dann nennen wir die Menge

$$\begin{aligned} U^\perp &:= \{x \in V : x \perp U\} \\ &= \{x \in V : \langle x,y\rangle = 0 \text{ für alle } y \in U\} \end{aligned}$$

das **orthogonale Komplement** von U. (Das Bild rechts illustriert dies für den Vektorraum $\mathbb{R}^2$ mit dem Standardskalarprodukt.)



Satz 21.39. *Es seien V ein endlich-dimensionaler Vektorraum mit Skalarprodukt und $U \leq V$ ein Unterraum. Dann ist das orthogonale Komplement $U^\perp$ ein Komplement von U im Sinne von Definition 17.8, d. h. es gilt $V = U \oplus U^\perp$.*

Beweis. Nach Satz 21.31 können wir eine Orthonormalbasis $(x_1,\dots,x_k)$ von U finden und zu einer Orthonormalbasis $(x_1,\dots,x_n)$ von V ergänzen. Nach Bemerkung 17.12 ist $U' := \mathrm{Lin}(x_{k+1},\dots,x_n)$ dann ein Komplement von U. Es genügt also zu zeigen, dass $U' = U^\perp$ gilt: Für einen Vektor $x \in V$ mit der Darstellung $x = \langle x_1,x\rangle x_1 + \cdots + \langle x_n,x\rangle x_n$ wie in Satz 21.37 (a) folgt unmittelbar

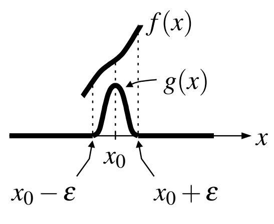
$$x \in U' = \mathrm{Lin}(x_{k+1},\dots,x_n) \Leftrightarrow \langle x_1,x\rangle = \cdots = \langle x_k,x\rangle = 0 \overset{21.27\,(b)}{\Leftrightarrow} x \in U^\perp. \qquad \square$$

Bemerkung 21.40. Der Beweis von Satz 21.39 zeigt auch, wie man das orthogonale Komplement eines Unterraums U in V berechnen kann: Man ergänzt eine Orthonormalbasis $(x_1,\dots,x_k)$ von U zu einer Orthonormalbasis $(x_1,\dots,x_n)$ von V; die dabei neu hinzu genommenen Vektoren $(x_{k+1},\dots,x_n)$ bilden dann eine Orthonormalbasis von $U^\perp$.

Da in diesem Fall natürlich auch $x_1,\dots,x_k$ die Vektoren $x_{k+1},\dots,x_n$ zu einer Orthonormalbasis von V ergänzen, folgt aus dieser Beschreibung auch sofort, dass $(U^\perp)^\perp = U$ gilt.

Beispiel 21.41 (Orthogonale Komplemente in unendlich-dimensionalen Vektorräumen)**.** Als „abschreckendes Beispiel" dafür, dass Satz 21.39 nicht so selbstverständlich ist, wie er vielleicht scheint, wollen wir kurz zeigen, dass diese Aussage für unendlich-dimensionale Vektorräume im Allgemeinen *falsch* ist, dass das orthogonale Komplement dann also nicht unbedingt immer ein Komplement im Sinne von Definition 17.8 ist. Dazu sei $V = C^0([a,b])$ der Vektorraum aller stetigen Funktionen auf dem Intervall $[a,b]$ mit dem Standardskalarprodukt $\langle f,g\rangle = \int_a^b f(x)g(x)\,dx$ aus Konstruktion 21.16. Wir betrachten darin den Unterraum $U = C^1([a,b]) \subsetneq V$ aller stetig differenzierbaren Funktionen und werden zeigen, dass $U^\perp = \{0\}$ gilt, so dass also insbesondere $U + U^\perp = U \neq V$, d. h. $U^\perp$ kein Komplement von U ist.

Dazu sei $f \in V\backslash\{0\}$ beliebig; wir werden zeigen, dass $f \notin U^\perp$ gilt. Wegen $f \neq 0$ gibt es ein $x_0 \in [a,b]$ mit $f(x_0) \neq 0$; ohne Einschränkung sei $f(x_0) > 0$. Da f stetig ist, gibt es dann nach Bemerkung 8.8 ein $\varepsilon \in \mathbb{R}_{>0}$, so dass $f(x) > 0$ für alle $x \in [a,b]$ mit $|x - x_0| < \varepsilon$ ist. Es sei nun $g\colon [a,b] \to \mathbb{R}_{\geq 0}$ eine stetig differenzierbare Funktion wie rechts im Bild, also mit $g(x_0) > 0$ und $g(x) = 0$ für $|x - x_0| \geq \varepsilon$. Dann ist $fg \geq 0$ auf $[a,b]$ und $f(x_0)g(x_0) > 0$, und damit $\langle f,g\rangle = \int_a^b f(x)g(x)\,dx > 0$ nach Konstruktion 21.16.



Wir haben also ein $g \in U$ gefunden mit $\langle f,g\rangle \neq 0$. Damit ist $f \notin U^\perp$, und da $f \in V\backslash\{0\}$ beliebig war, folgt wie behauptet $U^\perp = \{0\}$.

Unsere letzte Anwendung der Existenz von Orthonormalbasen in diesem Kapitel ist das folgende einfache Kriterium für die positive Definitheit einer Matrix.

Satz 21.42 (Hurwitz-Kriterium). *Es sei $A = (a_{i,j})_{i,j} \in \mathbb{K}^{n\times n}$ eine symmetrische (für $\mathbb{K} = \mathbb{R}$) bzw. hermitesche (für $\mathbb{K} = \mathbb{C}$) Matrix. Für $k = 1,\dots,n$ bezeichnen wir mit $A_k = (a_{i,j})_{i,j=1,\dots,k} \in \mathbb{K}^{k\times k}$ die Matrizen, die man erhält, wenn man von A nur die ersten k Zeilen und Spalten betrachtet. (Beachte, dass ihre Determinanten nach Folgerung 21.34 (a) in jedem Fall reell sind.)*

Dann ist A genau dann positiv definit, wenn $\det A_k > 0$ für alle $k = 1,\dots,n$.

Beweis. Es sei b die zu A gehörige symmetrische Bilinearform bzw. hermitesche Sesquilinearform auf $\mathbb{K}^n$, so dass also $A = (b(e_i, e_j))_{i,j}$ bzw. $b(x,y) = \overline{x}^\mathsf{T} A y$ für alle $x, y \in \mathbb{K}^n$ gilt.

„$\Rightarrow$“: Ist A positiv definit und b damit ein Skalarprodukt, so ist nach Bemerkung 21.14 (b) auch die Einschränkung von b auf den Unterraum $\mathrm{Lin}(e_1, \dots, e_k)$ ein Skalarprodukt. Die zugehörige Gramsche Matrix $(b(e_i, e_j))_{i,j=1,\dots,k} = A_k$ ist also ebenfalls positiv definit und hat damit nach Folgerung 21.34 (b) eine positive Determinante.

„$\Leftarrow$“: Wir zeigen die Aussage mit Induktion über n, der Induktionsanfang für $n = 1$ ist trivial.

Für den Induktionsschritt $n \to n+1$ bemerken wir zunächst, dass nach Annahme insbesondere $\det A_k > 0$ für $k = 1, \dots, n$ gilt. Damit ist A_n nach Induktionsvoraussetzung positiv definit, und wir können nach Satz 21.31 eine Orthonormalbasis $(x_1, \dots, x_n)$ für das zugehörige Skalarprodukt finden, das gerade die Einschränkung von b auf $\mathrm{Lin}(e_1, \dots, e_n)$ ist. Bilden wir nun wie im Gram-Schmidt-Verfahren in Satz 21.31 den Vektor

$$x_{n+1} := e_{n+1} - \sum_{i=1}^{n} b(x_i, e_{n+1})\, x_i \qquad \notin \mathrm{Lin}(x_1, \dots, x_n)$$

(beachte, dass hier wegen $b(x_i, x_i) = 1$ kein Nenner auftritt), so gilt wieder

$$b(x_j, x_{n+1}) = b(x_j, e_{n+1}) - \sum_{i=1}^{n} b(x_i, e_{n+1}) b(x_j, x_i) = b(x_j, e_{n+1}) - b(x_j, e_{n+1}) = 0$$

für alle $j = 1, \dots, n$, da $(x_1, \dots, x_n)$ orthonormal ist. Damit hat die Gramsche Matrix von b bezüglich der Basis $B = (x_1, \dots, x_{n+1})$ die Form

$$A_b^B = \overline{T}^\mathsf{T} A\, T = (b(x_i, x_j))_{i,j} = \left(\begin{array}{c|c} E_n & 0 \\ \hline 0 & b(x_{n+1}, x_{n+1}) \end{array}\right),$$

wobei wir wie üblich $T = (x_1 \,|\, \cdots \,|\, x_{n+1})$ gesetzt haben. Insbesondere folgt damit

$$b(x_{n+1}, x_{n+1}) = \det A_b^B = \det\left(\overline{T}^\mathsf{T} A\, T\right) = \overline{\det T} \cdot \det A \cdot \det T = |\det T|^2 \cdot \det A > 0.$$

Die Matrix A_b^B ist also eine Diagonalmatrix mit positiven Diagonaleinträgen, und damit nach Beispiel 21.15 (a) (bzw. Konstruktion 21.18) positiv definit. Daher ist auch b, und somit auch A positiv definit. □

Bemerkung 21.43. Die Determinanten von quadratischen Teilmatrizen einer gegebenen Matrix A werden oft auch *Minoren* von A genannt, die in Satz 21.42 auftretenden Determinanten $\det A_k$ bezeichnet man als *Hauptminoren* von A. Das Hurwitz-Kriterium ist daher auch unter dem Namen *Hauptminorenkriterium* bekannt.

Beispiel 21.44. Wir betrachten noch ein letztes Mal die symmetrische reelle Matrix

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

aus Beispiel 21.15 (b), die wir dort schon durch eine direkte Rechnung als positiv definit erkannt haben. Mit dem Hurwitz-Kriterium folgt dies (in der Notationen des Satzes) nun auch einfacher aus

$$\det A_1 = \det(1) = 1 > 0 \qquad \text{und} \qquad \det A_2 = \det A = 3 > 0.$$

Aufgabe 21.45. Für ein $\lambda \in \mathbb{R}$ sei $A = \begin{pmatrix} \lambda & 1 & 2 \\ 1 & 2 & 3 \\ 2 & 3 & 7 \end{pmatrix}$.

(a) Für welche λ ist A positiv definit? Für welche λ ist A negativ definit?

(b) Wir betrachten nun $V = \mathbb{R}^3$ als euklidischen Vektorraum mit dem Skalarprodukt b_A für den Wert $\lambda = 1$. Berechne zu $U = \mathrm{Lin}(e_1) \leq V$ eine Orthonormalbasis des orthogonalen Komplements $U^\perp$.

52

21.D Dualräume

Zu einem K-Vektorraum V haben wir bisher zwei Möglichkeiten betrachtet, aus Abbildungen neue Vektorräume zu bilden: den Vektorraum $\mathrm{End}(V)$ aller linearen Abbildungen $V \to V$ und den Vektorraum $\mathrm{BLF}(V)$ aller Bilinearformen $V \times V \to K$. Wir wollen nun zum Abschluss dieses Kapitels noch kurz eine dritte Konstruktion dieser Art untersuchen, nämlich den Raum $\mathrm{Hom}(V,K)$ aller linearen Abbildungen $V \to K$. Er ist natürlich ein Spezialfall des Vektorraums $\mathrm{Hom}(V,W)$ aller Morphismen in einen weiteren K-Vektorraum W, den wir in Abschnitt 16.C schon ausführlich betrachtet haben, hat aber trotzdem – gerade auch im Zusammenhang mit den in diesem Kapitel eingeführten Skalarprodukten – ein paar besondere Eigenschaften.

Definition 21.46 (Linearformen und Dualräume)**.** Eine lineare Abbildung von einem K-Vektorraum V in seinen Grundkörper K bezeichnet man als **Linearform** auf V. Der Vektorraum $\mathrm{Hom}(V,K)$ aller dieser Linearformen wird auch mit V^* bezeichnet und der **Dualraum** von V genannt.

Bemerkung 21.47 (Duale Basen)**.** Es sei V ein endlich-dimensionaler Vektorraum mit gegebener Basis $B = (x_1,\dots,x_n)$.

(a) Nach Folgerung 16.27 (a) gibt es einen Isomorphismus

$$V^* \to K^{1\times n},\ f \mapsto (f(x_1)\,|\cdots|\,f(x_n)),$$

der jeder Linearform $f\colon V \to K$ seine Abbildungsmatrix bezüglich der Basis B im Startraum zuordnet. Insbesondere ist also $\dim V^* = n = \dim V$, und V^* damit nach Abschnitt 16.B isomorph zu V.

(b) Unter dem Isomorphismus $V^* \cong K^{1\times n}$ aus (a) entspricht der i-te Einheitsvektor in $K^{1\times n} \cong K^n$ offensichtlich genau der Linearform, die auf x_i den Wert 1 und auf allen anderen Vektoren von B den Wert 0 annimmt; man bezeichnet sie in der Regel mit $x_i^* \in V^*$. Da Isomorphismen Basen auf Basen abbilden, ist $B^* := (x_1^*,\dots,x_n^*)$ damit eine Basis von V^*. Sie heißt die **duale Basis** zu B und ist also bestimmt durch

$$x_i^*(x_j) = \begin{cases} 1 & \text{für } i=j, \\ 0 & \text{für } i \neq j \end{cases}$$

für alle $i,j = 1,\dots,n$.

Beachte, dass wir in Bemerkung 21.47 (a) zwar gesehen haben, dass ein endlich-dimensionaler Vektorraum V immer isomorph zu seinem Dualraum V^* ist, dass wir einen konkreten Isomorphismus zwischen V und V^* aber erst dann hinschreiben können, wenn wir eine Basis von V (und damit einen Isomorphismus $V \cong K^n$) gewählt haben. Man sagt daher auch, dass es zwischen V und V^* zwar einen Isomorphismus, aber keinen *natürlichen Isomorphismus* gibt.

Diese Situation ändert sich, wenn V ein euklidischer Raum ist: In diesem Fall wollen wir jetzt zeigen, dass das gegebene Skalarprodukt eine natürliche Einbettung von V in V^* liefert, die im endlich-dimensionalen Fall sogar natürlichen Isomorphismus zwischen V und V^* liefert.

Satz 21.48 (Einbettung eines euklidischen Vektorraums in seinen Dualraum)**.** *Es sei V ein euklidischer Vektorraum.*

(a) *Es gibt einen natürlichen injektiven Morphismus $\Gamma\colon V \to V^*$, wobei $\Gamma(x) \in V^*$ für alle $x \in V$ definiert ist durch $\Gamma(x)(y) = \langle x,y\rangle$ für alle $y \in V$.*

(b) *Ist V endlich-dimensional, so ist die Abbildung Γ aus (a) sogar ein Isomorphismus.*

Beweis.

(a) Wir müssen die folgenden Dinge überprüfen:

- Für alle $x \in V$ ist $\Gamma(x)$ wirklich in V^*, also eine lineare Abbildung von V nach K, weil das Skalarprodukt im zweiten Eintrag linear ist: Für alle $y_1, y_2 \in V$ gilt
$$\Gamma(x)(y_1+y_2) = \langle x, y_1+y_2\rangle = \langle x, y_1\rangle + \langle x, y_2\rangle = \Gamma(x)(y_1) + \Gamma(x)(y_2).$$
Die Verträglichkeit mit der Skalarmultiplikation zeigt man analog.
- Die Abbildung $\Gamma\colon V \to V^*$ ist linear, weil das Skalarprodukt im ersten Eintrag linear ist: Für alle $x_1, x_2 \in V$ und $y \in V$ gilt
$$\Gamma(x_1+x_2)(y) = \langle x_1+x_2, y\rangle = \langle x_1, y\rangle + \langle x_2, y\rangle = \Gamma(x_1)(y) + \Gamma(x_2)(y),$$
und damit $\Gamma(x_1+x_2) = \Gamma(x_1) + \Gamma(x_2)$ in V^*. Auch hier folgt die Verträglichkeit mit der Skalarmultiplikation wieder analog.
- Der Morphismus Γ ist injektiv, d. h. $\operatorname{Ker}\Gamma = \{0\}$: Es sei $x \in V$ mit $\Gamma(x) = 0 \in V^*$, d. h. $\Gamma(x)$ ist die Nullabbildung von V nach K. Insbesondere ist dann also $\Gamma(x)(x) = 0 \in K$. Nach Definition von Γ bedeutet dies $\langle x, x\rangle = 0$, wegen der positiven Definitheit des Skalarprodukts also $x = 0$.

(b) Im endlich-dimensionalen Fall folgt aus der Dimensionsformel für Morphismen aus Folgerung 16.30 (c) sofort
$$\dim \operatorname{Im}\Gamma = \dim V - \dim \operatorname{Ker}\Gamma \overset{\text{(a)}}{=} \dim V \overset{21.47\,\text{(a)}}{=} \dim V^*,$$
und damit auch die Surjektivität von Γ. □

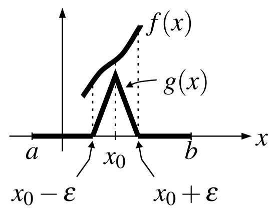
Bemerkung 21.49.

(a) Für einen endlich-dimensionalen euklidischen Raum V bildet der natürliche Isomorphismus Γ aus Satz 21.48 (b) jede Orthonormalbasis $B = (x_1, \dots, x_n)$ auf ihre duale Basis $B^* = (x_1^*, \dots, x_n^*)$ wie in Bemerkung 21.47 (b) ab, denn für alle $i, j = 1, \dots, n$ gilt
$$\begin{aligned}\Gamma(x_i)(x_j) = \langle x_i, x_j\rangle &= \begin{cases} 1 & \text{für } i = j, \\ 0 & \text{für } i \neq j \end{cases} \\ &= x_i^*(x_j),\end{aligned}$$
und damit $\Gamma(x_i) = x_i^*$.

(b) Für einen unitären Raum V ist die wie in Satz 21.48 konstruierte Abbildung $\Gamma\colon V \to V^*$ zwar noch definiert und bijektiv, aber nicht mehr linear (und damit kein Isomorphismus), da das Skalarprodukt im ersten Eintrag nicht linear ist und somit der obige Beweis der Linearität von Γ nicht funktioniert.

Beispiel 21.50. Analog zu Beispiel 21.41 verhalten sich unendlich-dimensionale Vektorräume auch in Satz 21.48 wieder anders als endlich-dimensionale, da die Einbettung Γ in den Dualraum dann kein Isomorphismus mehr ist. Als konkretes Beispiel dafür betrachten wir wieder einmal den Raum $V = C^0([a,b])$ mit dem Standardskalarprodukt. Dabei wählen wir a und b mit $a < 0 < b$ und betrachten die Auswerteabbildung
$$\delta\colon V \to \mathbb{R},\ g \mapsto g(0).$$
Sie ist offensichtlich linear, und damit ein Element des Dualraums V^*. Allerdings wollen wir jetzt sehen, dass sie nicht im Bild der natürlichen Abbildung $\Gamma\colon V \to V^*$ liegt, also dass es kein $f \in V$ gibt mit $\Gamma(f) = \delta$, d. h. mit $\Gamma(f)(g) = \delta(g)$, und damit mit
$$\int_a^b f(x)g(x)\,dx = g(0) \quad \text{für alle } g \in C^0([a,b]). \tag{$*$}$$
Wir zeigen dies mit einem Widerspruchsbeweis analog zu Beispiel 21.41 und nehmen also an, dass es doch eine solche stetige Funktion $f \in C^0([a,b])$ gibt.

Ist dann $x_0 \in [a,b] \backslash \{0\}$ beliebig, so muss notwendigerweise $f(x_0) = 0$ gelten: Wäre nämlich $f(x_0) \neq 0$, wobei wir ohne Einschränkung $f(x_0) > 0$ annehmen können, so gibt es nach Bemerkung 8.8 zunächst ein $\varepsilon \in \mathbb{R}_{>0}$ mit $f(x) > 0$ für alle $x \in [a,b]$ mit $|x - x_0| < \varepsilon$. Wir können nun wie im Bild rechts eine stetige Funktion $g\colon [a,b] \to \mathbb{R}_{\geq 0}$ wählen mit $g(x_0) > 0$, $g(0) = 0$ und $g(x) = 0$ für $|x - x_0| \geq \varepsilon$. Dann ist $fg \geq 0$ auf $[a,b]$ und $f(x_0)g(x_0) > 0$, und damit nach Konstruktion 21.16



$$\int_a^b f(x)g(x)\,dx > 0 = g(0)$$

im Widerspruch zu $(*)$. Also ist $f(x_0) = 0$ für alle $x_0 \neq 0$. Da f stetig ist, müsste f dann aber schon die Nullfunktion sein – was natürlich auch $(*)$ widerspricht. Also kann es keine solche Funktion f geben, d. h. die Abbildung $\Gamma\colon V \to V^*$ ist in diesem Fall nicht surjektiv.

Bemerkung 21.51 (Die „Deltafunktion"). Man kann das Ergebnis von Beispiel 21.50 auch so interpretieren: Da Γ eine natürliche Einbettung ist, können wir den Funktionenraum $C^0([a,b])$ über Γ als Teilmenge von $(C^0([a,b]))^*$ auffassen und so zu einem „erweiterten Funktionenraum" $(C^0([a,b]))^*$ übergehen (in dem „Funktionen" nicht mehr stetige Abbildungen von $[a,b]$ nach $\mathbb{R}$ sind, sondern lineare Abbildungen von $C^0([a,b])$ nach $\mathbb{R}$). Dies ist völlig analog z. B. zur Konstruktion der komplexen Zahlen aus den reellen in Kapitel 6: Dort haben wir zunächst einen neuen Körper $\mathbb{C} = \mathbb{R}^2$ eingeführt, dann eine natürliche Einbettung $\mathbb{R} \to \mathbb{C}$, $x \mapsto (x,0)$ gefunden, und dies benutzt, um $\mathbb{R}$ als Teilmenge von $\mathbb{C}$ bzw. $\mathbb{C}$ als eine Erweiterung von $\mathbb{R}$ aufzufassen.

Wie wir oben gesehen haben, existiert in diesem „erweiterten Funktionenraum" $(C^0([a,b]))^*$ nun z. B. das Element δ, das sich gemäß $(*)$ so verhält, als würde es $g(0)$ zurück liefern, wenn man δ mit einer Funktion g multipliziert und darüber integriert. Dieses Element δ, das insbesondere in der Physik eine große Rolle spielt, bezeichnet man oft als die „*Deltafunktion*", auch wenn es sich hierbei nicht um eine Funktion im eigentlichen, sondern nur im oben betrachteten erweiterten Sinne handelt.

Oft wird δ in der Physik ohne Erwähnung von Dualräumen als eine „Funktion" eingeführt, die die Eigenschaft von f in $(*)$ hat, also insbesondere (wie wir oben gesehen haben) überall außer in 0 gleich 0 ist und (mit $g = 1$) $\int_a^b \delta(x)\,dx = 1$ erfüllt – d. h. in 0 „gerade so unendlich" ist, dass die Fläche unter ihrem Graphen (der Breite 0 und Höhe ∞) gleich 1 ist. Dies erscheint natürlich zunächst unsinnig, lässt sich aber wie eben gesehen durch eine Erweiterung des Funktionenraums zu einer mathematisch exakten Theorie machen. Derartige Fragestellungen – also wie sich die lineare Algebra auf unendlich-dimensionalen Funktionenräumen verhält – sind die Inhalte der sogenannten Funktionalanalysis, in die ihr im zweiten Studienjahr eine Einführung hören könnt.

Aufgabe 21.52 (Duale Abbildungen). Zu einer linearen Abbildung $f\colon V \to W$ definiert man die *duale Abbildung* f^* zwischen den Dualräumen W^* und V^* durch

$$f^*\colon W^* \to V^*,\ \varphi \mapsto \varphi \circ f.$$

Man zeige:

(a) Für jede lineare Abbildung $f\colon V \to W$ ist $f^*\colon W^* \to V^*$ ebenfalls linear, und auch die Abbildung $\mathrm{Hom}(V,W) \to \mathrm{Hom}(W^*,V^*)$, $f \mapsto f^*$ ist linear.

(b) Sind V und W endlich-dimensional mit Basen B bzw. C, so gilt für die Abbildungsmatrix von f^* bezüglich der dualen Basen B^* und C^*

$$A_{f^*}^{C^*,B^*} = \left(A_f^{B,C}\right)^{\mathsf{T}}.$$

(c) Sind V und W endlich-dimensionale euklidische Vektorräume, so gibt es natürliche Isomorphismen $\Gamma_V\colon V \to V^*$ und $\Gamma_W\colon W \to W^*$ wie in Satz 21.48. Konstruieren wir damit aus f^* die Abbildung $g := \Gamma_V^{-1} \circ f^* \circ \Gamma_W\colon W \to V$, so gilt

$$\langle x, f(y)\rangle = \langle g(x), y\rangle \quad \text{für alle } x \in V \text{ und } y \in W.$$