a) Euklidische Vektorräume

Zu Beginn dieses Kapitels hatten wir für Vektoren in $\mathbb{R}^n$ ein Skalarprodukt definiert, das uns gestattete, Längen und Winkel zu berechnen. In diesem Abschnitt sollen diese Begriffe und ihr Umfeld für beliebige reelle Vektorräume erklärt werden. Da es hier um die Grundbegriffe der klassischen EUKLIDischen Geometrie geht, definieren wir:

Definition: Ein EUKLIDischer Vektorraum ist ein Paar $(V, \cdot)$ bestehend aus einem $\mathbb{R}$-Vektorraum V und einer Abbildung

$$\cdot : V \times V \to \mathbb{R}; \quad (\vec{v}, \vec{w}) \mapsto \vec{v} \cdot \vec{w}$$

mit folgenden Eigenschaften:
a) $\cdot$ ist linear in beiden Argumenten (kurz ausgedrückt *bilinear*), d.h.

$$(\lambda \vec{u} + \mu \vec{v}) \cdot \vec{w} = \lambda(\vec{u} \cdot \vec{w}) + \mu(\vec{u} \cdot \vec{w})$$

und

$$\vec{u} \cdot (\lambda \vec{v} + \mu \vec{w}) = \lambda(\vec{u} \cdot \vec{v}) + \mu(\vec{u} \cdot \vec{w})$$

für alle $\vec{u}, \vec{v}, \vec{w} \in V$ und $\lambda, \mu \in \mathbb{R}$.
b) $\cdot$ ist symmetrisch, d.h. $\vec{v} \cdot \vec{w} = \vec{w} \cdot \vec{v}$ für alle $\vec{v}, \vec{w} \in V$.
c) $\cdot$ ist positiv definit, d.h. $\vec{v} \cdot \vec{v} > 0$ für alle Vektoren $\vec{v} \neq \vec{0}$ aus V.

Die bilineare Abbildung $\cdot$ heißt *Skalarprodukt;* wie bei Produkten üblich, werden wir den Malpunkt oft weglassen. Gelegentlich, wenn Verwechslungen mit anderen Produkten zu befürchten sind, werden wir anstelle von $\vec{v} \cdot \vec{w}$ auch $(\vec{v}, \vec{w})$ schreiben; manchmal sieht man in der Literatur auch die Schreibweise $\langle \vec{v}, \vec{w} \rangle$.

Es ist nicht ganz sicher, ob EUKLID wirklich gelebt hat; das nebenstehende Bild aus dem 18. Jahrhundert ist mit Sicherheit reine Phantasie. EUKLIDs Lebensdaten werden mit etwa 325 bis etwa 265 angegeben; er soll eine Schule von Mathematikern in Alexandrien geleitet haben. EUKLID ist vor allem bekannt als Autor der *Elemente,* in denen er die Geometrie seiner Zeit systematisch darstellte und (nicht immer ganz korrekt) auf wenige Definitionen sowie die berühmten fünf Postulate zurückführte. Drei der 13 Bücher der Elemente befassen sich mit Zahlentheorie (u.a. dem EUKLIDischen Algorithmus) und eines mit irrationalen Zahlen.

Typisches Beispiel eines EUKLIDischen Vektorraums ist natürlich der $\mathbb{R}^n$ mit seinem Standardskalarprodukt

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = \sum_{i=1}^{n} v_i w_i \,.$$

Ein anderes Beispiel ist der Raum $C^0([0,\ 1],\ \mathbb{R})$ aller stetiger Funktionen vom Einheitsintervall $[0,\ 1]$ nach $\mathbb{R}$ mit dem Skalarprodukt

$$(f,\ g) \underset{\text{def}}{=} \int_0^1 f(t)g(t)\,dt\,.$$

(Dies ist ein typischer Fall, wo die Klammerschreibweise vorzuziehen ist, denn $f \cdot g$ ist bereits vergeben für die Funktion, die jedem $t \in [0,\ 1]$ den Wert $f(t) \cdot g(t)$ zuordnet.)

Die Eigenschaften *a)* und *b)* aus der Definition sind klar, und für *c)* ist zumindest klar, daß

$$(f,\ f) = \int_0^1 f(t)^2\,dt \geq 0$$

ist, da der Integrand nirgends negativ wird. Falls f nicht identisch verschwindet, gibt es einen Punkt $t_0 \in (0,\ 1)$ mit $f(t_0) = h \neq 0$. Wegen der Stetigkeit von f gibt es dazu eine Umgebung $(a,\ b)$, so daß $|f(t)| > h/2$ für $t \in (a,\ b)$. Damit ist

$$(f,\ f) = \int_0^1 f(t)^2\,dt \geq \int_a^b f(t)^2\,dt > \int_a^b \frac{h^2}{4}\,dt = \frac{h^2}{4}(b-a) > 0\,,$$

$(f,\ f) = 0$ ist also nur möglich, wenn f überall verschwindet.

Man beachte, daß hier die Stetigkeit von f eine wesentliche Rolle spielt: Für die Funktion f, die überall verschwindet außer im Punkt $1/2$, wo sie den Wert 1 annimmt, ist das Integral über $f(t)^2$ gleich null, obwohl die Funktion nicht überall verschwindet.

Da wir EUKLIDische Vektorräume eingeführt haben, um dort so etwas wie EUKLIDische Geometrie zu betreiben, sollten wir als nächstes deren Grundbegriffe definieren:

Definition: *a)* Die Länge eines Vektors $\vec{v}$ aus einem EUKLIDischen Vektorraum V ist $\|\vec{v}\| = \sqrt{\vec{v} \cdot \vec{v}}$.
b) Der Winkel zwischen zwei Vektoren $\vec{v}, \vec{w} \in V \setminus \{\vec{0}\}$ ist

$$\measuredangle(\vec{v}, \vec{w}) = \arccos\left(\frac{\vec{v} \cdot \vec{w}}{\sqrt{(\vec{v} \cdot \vec{v})(\vec{w} \cdot \vec{w})}}\right) .$$

Da der Arkuscosinus nur Werte zwischen null und π annimmt, im Gradmaß also zwischen null und 180°, ist der so definierte Winkel *unorientiert.* Im $\mathbb{R}^3$ konnten wir durch Auszeichnung von Rechtssystemen über die Dreifingerregel oder Korkenzieherregel orientierte Winkel zwischen null und 2π definieren; das ist in höherdimensionalen EUKLIDischen Vektorräumen nicht mehr möglich: Falls man dort einen orientierten Winkel definieren möchte, muß man zunächst eine Orientierung des Vektorraums auszeichnen. Eine entsprechende Theorie orientierter EUKLIDischer Vektorräume gibt es, sie wird aber im Rahmen dieser Vorlesung für nichts benötigt, so daß wir uns nicht weiter damit befassen müssen.

Wir müssen uns allerdings die Frage stellen, warum die obige Definition überhaupt einen Winkel erklärt: Da der Cosinus nur Werte zwischen -1 und 1 annimmt, ist sie offensichtlich nur sinnvoll, wenn das Argument des Arkuskosinus höchstens den Betrag eins hat.

Für das Standardskalarprodukt im $\mathbb{R}^2$ oder $\mathbb{R}^3$ ist das kein Problem: In §1 dieses Kapitels hatten wir das Skalarprodukt im $\mathbb{R}^3$ *definiert* als Produkt aus den Längen der beiden Vektoren und dem Cosinus des eingeschlossenen Winkels; wir hatten dann nachgerechnet, daß dies mit der Koordinatendefinition, die dem hier definierten Skalarprodukt entspricht, übereinstimmt.

Für beliebige EUKLIDische Vektorräume V müssen wir zeigen, daß für zwei Vektoren $\vec{v}, \vec{w} \in V$ stets

$$\left|\frac{\vec{v} \cdot \vec{w}}{\sqrt{(\vec{v} \cdot \vec{v})(\vec{w} \cdot \vec{w})}}\right| \le 1$$

ist. Mit der oben definierten Länge können wir dies auch umschreiben als

$$\left|\frac{\vec{v} \cdot \vec{w}}{\|\vec{v}\| \cdot \|\vec{w}\|}\right| \le 1 \quad \text{oder} \quad |\vec{v} \cdot \vec{w}| \le \|\vec{v}\| \cdot \|\vec{w}\| .$$

Diese Ungleichung werden wir im folgenden noch häufiger in den verschiedensten Situationen anwenden müssen; wir werden sie daher im nächsten Abschnitt gleich etwas allgemeiner beweisen, als es hier notwendig ist. Für den Augenblick muß die Frage der Wohldefiniertheit des Winkels noch offen bleiben.

Für EUKLIDische Vektorräume wie $\mathcal{C}^0\big([0, 1], \mathbb{R}\big)$ ist die „Länge" eines „Vektors" natürlich nichts mehr, was man sich geometrisch anschaulich vorstellen könnte; trotzdem ist sie oft eine nützliche Größe. Eine stetige Funktion $I\colon [0, 1] \to \mathbb{R}$ könnte beispielsweise einen zeitabhängigen Strom beschreiben, der durch einen Widerstand R fließt; nach dem OHMschen Gesetz fällt dort eine Spannung $U(t) = R \cdot I(t)$ ab, so daß die elektrische Leistung gleich $U(t) \cdot I(t) = R \cdot I(t)^2$ ist. Die elektrische Energie während des Zeitraums $[0, 1]$ ist daher gleich

$$\int_0^1 U(t) \cdot I(t)\, dt = \int_0^1 R \cdot I(t)^2\, dt = R \cdot \int_0^1 I(t)^2\, dt = R \cdot \|I\| ,$$

d.h. die „Länge" von I ist bis auf einen konstanten Faktor gerade gleich der Energie des Signals I.

b) Hermitesche Vektorräume

Für Vektorräume über den komplexen Zahlen sind die Definitionen aus dem vorigen Abschnitt nicht sonderlich geeignet; beispielsweise ist mit dem üblichen Skalarprodukt in $\mathbb{C}^2$

$$\begin{pmatrix} 1 \\ i \end{pmatrix} \cdot \begin{pmatrix} 1 \\ i \end{pmatrix} = 1^2 + i^2 = 0 .$$

Solche Vektoren mit Quadrat Null sind zwar in einigen Anwendungen (wie etwa der speziellen Relativitätstheorie) durchaus sinnvoll, meist möchte man aber das Skalarprodukt eines Vektors mit sich selbst als

Quadrat seiner Länge interpretieren, und die sollte für Vektoren ungleich dem Nullvektor positiv sein.

Wir haben bereits jeder komplexen Zahl ungleich Null eine positive reelle Zahl zugeordnet, nämlich ihren Betrag

$$|z| = \sqrt{z\overline{z}}\,.$$

Entsprechend können wir zwei komplexen Vektoren

$$\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \quad \text{und} \quad \vec{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$$

aus $\mathbb{C}^n$ das Produkt

$$\vec{v}\cdot\vec{w} = \sum_{\ell=1}^{n} v_\ell \overline{w_\ell}$$

zuordnen. Auch hier ist offensichtlich $\vec{v}\cdot\vec{v}$ für jeden Vektor außer dem Nullvektor eine positive reelle Zahl, denn wir addieren ja die Betragsquadrate der Komponenten des Vektors.

Dieses Produkt erfüllt nun allerdings nicht mehr die Bedingungen *a)* und *b)* aus dem vorigen Abschnitt: Es ist zwar noch linear im ersten Argument, nicht aber im zweiten, denn schon bei Multiplikation des zweiten Vektors mit einer komplexen Zahl $\lambda \neq \mathbb{R}$ ist

$$\vec{v}\cdot(\lambda\vec{w}) = \sum_{\ell=1}^{n} v_\ell \cdot \overline{\lambda w_\ell} = \overline{\lambda}\cdot\sum_{\ell=1}^{n} v_\ell\overline{w_\ell} = \overline{\lambda}\cdot(\vec{v}\cdot\vec{w}) \neq \lambda\cdot(\vec{v}\cdot\vec{w})\,.$$

Bezüglich der Addition gibt es keine Probleme, d.h.

$$\vec{u}\cdot(\lambda\vec{v}+\mu\vec{w}) = \overline{\lambda}\cdot(\vec{u}\cdot\vec{v}) + \mu\cdot(\vec{u}\cdot\vec{w})\,.$$

Die Symmetriebedingung ist allerdings verletzt, denn schon für zwei komplexe Zahlen z und w ist das Produkt $z\overline{w} = \overline{w}z = \overline{w\overline{z}}$ im allgemeinen verschieden von $w\overline{z}$, und entsprechend ist auch für zwei Vektoren

$$\vec{v}\cdot\vec{w} = \overline{\vec{w}\cdot\vec{v}}\,.$$

Der Begriff des HERMITEschen Vektorraums formalisiert die Eigenschaften des gerade eingeführten Produkts.

CHARLES HERMITE (1822–1901) war einer der bedeutendsten Mathematiker des neunzehnten Jahrhunderts. Zu seinen Resultaten zählen eine Vereinfachung des ABELschen Beweises, daß Gleichungen fünften Grades im allgemeinen nicht durch Wurzelausdrücke gelöst werden können, die explizite Lösung solcher Gleichungen durch elliptische Funktionen, der Nachweis, daß e eine transzendete Zahl ist, also keiner algebraischen Gleichung über $\mathbb{Q}$ genügt, eine Interpolationsformel und vieles mehr. HERMITE war auch ein sehr guter akademischer Lehrer; er unterrichtete an der École Polytechnique, dem Collège de France, der École Normale Supérieure und der Sorbonne.

Definition: Ein HERMITEscher Vektorraum ist ein Paar $(V,\cdot)$ bestehend aus einem $\mathbb{C}$-Vektorraum V und einer Abbildung

$$\cdot : V\times V \to \mathbb{C}; \qquad (\vec{v},\vec{w}) \mapsto \vec{v}\cdot\vec{w}$$

mit folgenden Eigenschaften:
a) $\cdot$ ist linear im ersten Argument, d.h.

$$(\lambda\vec{u}+\mu\vec{v})\cdot\vec{w} = \lambda(\vec{u}\cdot\vec{w}) + \mu(\vec{v}\cdot\vec{w})$$

für alle $\vec{u},\vec{v},\vec{w}\in V$ und $\lambda,\mu\in\mathbb{C}$.
b) $\cdot$ ist HERMITEsch symmetrisch, d.h. $\vec{v}\cdot\vec{w} = \overline{\vec{w}\cdot\vec{v}}$ für alle $\vec{v},\vec{w}\in V$.
c) $\cdot$ ist positiv definit, d.h. $\vec{v}\cdot\vec{v}$ ist eine positive reelle Zahl für alle Vektoren $\vec{v}\neq\vec{0}$ aus V.

Die bilineare Abbildung $\cdot$ heißt HERMITE*sches Skalarprodukt;* auch hier werden wir den Malpunkt nicht immer hinschreiben und bei Verwechslungsgefahr mit anderen Produkten auch gelegentlich $(\vec{v},\ \vec{w})$ schreiben.

Durch Kombination der Eigenschaften *a)* und *b)* kann man leicht ausrechnen, was anstelle der Linearität für das zweite Argument gilt:

$$\begin{aligned}\vec{u}\cdot(\lambda\vec{v}+\mu\vec{w}) &= \overline{(\lambda\vec{v}+\mu\vec{w})\cdot\vec{u}} = \overline{\lambda(\vec{v}\cdot\vec{u})+\mu(\vec{w}\cdot\vec{u})}\\ &= \overline{\lambda}\,\overline{(\vec{v}\cdot\vec{u})} + \overline{\mu}\,\overline{(\vec{w}\cdot\vec{u})} = \overline{\lambda}\,(\vec{u}\cdot\vec{v}) + \overline{\mu}\,(\vec{u}\cdot\vec{w})\,.\end{aligned}$$

Diese Eigenschaft, zusammen mit der Linearität im ersten Argument, bezeichnet man gelegentlich als *Sesquilinearität,* d.h. „anderthalbfache Linearität“, im Gegensatz zur echten Linearität in zwei Argumenten, der *Bilinearität.*

Typisches Beispiel eines HERMITEschen Vektorraums ist natürlich der $\mathbb{C}^n$ mit dem eingangs definierten Produkt. HERMITEsche Vektorräume von komplexwertigen Funktionen werden wir im nächsten Abschnitt betrachten.

Auch in HERMITEschen Vektorräumen werden wir $\|\vec{v}\| = \sqrt{\vec{v}\cdot\vec{v}}$ gelegentlich als *Länge* des Vektors $\vec{v}$ bezeichnen; auf die Definition von Winkeln hingegen wollen wir verzichten, da die Übernahme der entsprechenden Definition für EUKLIDische Vektorräume hier auf geometrisch nicht sinnvolle komplexe Winkel führen würde.

Ansonsten können und werden wir EUKLIDische und HERMITEsche Vektorräume im folgenden gleichzeitig behandeln: Ersetzt man in der Definition eines HERMITEschen Vektorraums überall $\mathbb{C}$ durch $\mathbb{R}$, so erhält man einen EUKLIDischen Vektorraum. Zwar stehen noch an vielen Stellen die Querstriche für komplexe Konjugation, aber da sich eine reelle Zahl unter komplexer Konjugation nicht ändert, sind die Konjugationsstriche nur überflüssig, nicht schädlich.

Als erstes Beispiel für die gemeinsame Behandlung von EUKLIDischen und HERMITEschen Vektorräumen wollen wir die bereits angekündigte Ungleichung $(\vec{v}\cdot\vec{w})^2 \le (\vec{v}\cdot\vec{v})(\vec{w}\cdot\vec{w})$ beweisen; im Hinblick auf den nächsten Abschnitt sind dabei die Voraussetzungen etwas schwächer formuliert als es für diese Vektorräume notwendig wäre:

Cauchy-Schwarzsche Ungleichung: Für $k = \mathbb{R}$ oder $\mathbb{C}$ sei auf dem k-Vektorraum V eine Abbildung $\cdot : V \times V \to k;\ (f,g) \mapsto f\cdot g$ gegeben mit den Eigenschaften
a) $(\lambda\vec{u} + \mu\vec{v})\cdot\vec{w} = \lambda(\vec{u}\cdot\vec{w}) + \mu(\vec{v}\cdot\vec{w})$
b) $\vec{w}\cdot\vec{v} = \overline{\vec{v}\cdot\vec{w}}$
c) $\|\vec{v}\|^2 = \vec{v}\cdot\vec{v} \ge 0$ für alle $\vec{v}\in V$.
Dann ist für alle Vektoren $\vec{v},\vec{w}\in V$

$$|\vec{v}\cdot\vec{w}| \le \|\vec{v}\|\cdot\|\vec{w}\| \ .$$

Der *Beweis* beruht auf folgendem Trick: Wegen *c)* ist für $\lambda,\mu\in\mathbb{C}$

$$\begin{aligned}(\lambda\vec{v}+\mu\vec{w})\cdot(\lambda\vec{v}+\mu\vec{w}) &= \lambda\overline{\lambda}\,\vec{v}\cdot\vec{v} + \lambda\overline{\mu}\,\vec{v}\cdot\vec{w} + \mu\overline{\lambda}\,\vec{w}\cdot\vec{v} + \mu\overline{\mu}\,\vec{w}\cdot\vec{w}\\ &= \lambda\overline{\lambda}\,\vec{v}\cdot\vec{v} + \lambda\overline{\mu}\,\vec{v}\cdot\vec{w} + \overline{\lambda}\mu\,\overline{\vec{v}\cdot\vec{w}} + \mu\overline{\mu}\,\vec{w}\cdot\vec{w}\end{aligned}$$

stets größer oder gleich null. Speziell für $\lambda = (\vec{w}\cdot\vec{w}) \in \mathbb{R}$ und $\mu = -(\vec{v}\cdot\vec{w})$ erhalten wir

$$(\vec{w}\cdot\vec{w})^2(\vec{v}\cdot\vec{v}) - 2(\vec{w}\cdot\vec{w})(\overline{\vec{v}\cdot\vec{w}})(\vec{v}\cdot\vec{w}) + (\vec{v}\cdot\vec{w})(\overline{\vec{v}\cdot\vec{w}})(\vec{w}\cdot\vec{w}) \ge 0\,,$$

also

$$(\vec{w}\cdot\vec{w})\big((\vec{w}\cdot\vec{w})(\vec{v}\cdot\vec{v}) - |\vec{v}\cdot\vec{w}|^2\big) \ge 0\,.$$

Ist hier $\vec{w}\cdot\vec{w} \ne 0$, können wir durch diese Zahl dividieren und die Behauptung ist bewiesen. Andernfalls können wir, falls wenigstens $\vec{v}\cdot\vec{v}$ nicht null ist, im obigen Argument die Rollen von $\vec{v}$ und $\vec{w}$ vertauschen und damit die Behauptung beweisen. Wenn schließlich $\vec{v}\cdot\vec{v}$ und $\vec{w}\cdot\vec{w}$ beide null sind, folgt aus

$$(\vec{v}\pm\vec{w})\cdot(\vec{v}\pm\vec{w}) = \pm(\vec{v}\cdot\vec{w} + \overline{\vec{v}\cdot\vec{w}}) = \pm 2\,\mathfrak{Re}(\vec{v}\vec{w}) \ge 0\,,$$

daß der Realteil von $\vec{v}\cdot\vec{w}$ verschwindet, und genauso verschwindet auch der Imaginärteil, da $(i\vec{v}\pm\vec{w})\cdot(i\vec{v}\pm\vec{w}) = \pm 2\,\mathfrak{Im}(\vec{v}\cdot\vec{w}) \ge 0$ ist. Somit ist $\vec{v}\cdot\vec{w} = 0$, und die Ungleichung gilt auch in diesem Fall. ■

Baron AUGUSTIN LOUIS CAUCHY (1789–1857) stellte als erster durch die exakte Definition von Begriffen wie *Konvergenz* und *Stetigkeit* die Analysis auf ein sicheres Fundament. In insgesamt 789 Arbeiten beschäftigte er sich u.a. auch mit komplexer Analysis, Variationsrechnung, Differentialgleichungen, FOURIER-Analysis, Permutationsgruppen, der Diagonalisierung von Matrizen und der theoretischen Mechanik. Als überzeugter Royalist hatte er häufig Schwierigkeiten mit den damaligen Regierungen; er lebte daher mehrere Jahre im Exil in Turin und später in Prag, wo er (mit sehr mäßigem Erfolg) den französischen Thronfolger unterrichtete.

Der deutsche Mathematiker KARL HERMAN AMANDUS SCHWARZ (1843–1921) beschäftigte sich hauptsächlich mit konformen Abbildungen und mit sogenannten Minimalflächen, d.h. Flächen mit vorgegebenen Eigenschaften, deren Flächeninhalt minimal ist. Im Rahmen einer entsprechenden Arbeit für die WEIERSTRASS-Festschrift von 1885 (im Falle eines durch Doppelintegrale definierten Skalarprodukts) bewies er die obige Ungleichung; CAUCHY hatte sie bereits in seinem Analysislehrbuch von 1821 für endlichdimensionale Vektoren bewiesen. SCHWARZ lehrte nacheinander in Halle, Zürich, Göttingen und Berlin.

In einem EUKLIDischen oder HERMITEschen Vektorraum V sind die drei Bedingungen *a)* bis *c)* erfüllt, dort gilt also die CAUCHY-SCHWARZsche Ungleichung. Tatsächlich gilt noch etwas mehr:

Zusatz: In einem EUKLIDischen oder HERMITEschen Vektorraum V gilt für alle $\vec{v}, \vec{w} \in V$
$$|\vec{v} \cdot \vec{w}| \le \|\vec{v}\| \cdot \|\vec{w}\|$$
mit Gleichheit genau dann, wenn $\vec{v}$ und $\vec{w}$ linear abhängig sind.

Beweis: Wir müssen nur noch zeigen, daß genau dann ein Gleichheitszeichen in der Ungleichung steht, wenn $\vec{v}$ und $\vec{w}$ linear abhängig sind. Falls $\vec{v}$ oder $\vec{w}$ gleich dem Nullvektor ist, sind die beiden Vektoren linear abhängig und auf beiden Seiten der Gleichung steht null; die Behauptung ist also richtig. Sind beide Vektoren ungleich $\vec{0}$, betrachten wir wie oben
$$\|(\vec{w} \cdot \vec{w}) \cdot \vec{v} - (\vec{v} \cdot \vec{w}) \cdot \vec{w})\|^2 = (\vec{w} \cdot \vec{w})\big((\vec{w} \cdot \vec{w})(\vec{v} \cdot \vec{v}) - (\vec{v} \cdot \vec{w})^2\big)\ .$$
In der CAUCHY-SCHWARZschen Ungleichung gilt genau dann das Gleichheitszeichen, wenn dies verschwindet. Wegen der positiven Definitheit des Produkts verschwindet die Länge aber genau dann, wenn der Vektor selbst der Nullvektor ist, d.h.
$$(\vec{w} \cdot \vec{w}) \cdot \vec{v} - (\vec{v} \cdot \vec{v}) \cdot \vec{w} = \vec{0}\ .$$
Diese Gleichung besagt, daß $\vec{v}$ und $\vec{w}$ linear abhängig sind. Sind umgekehrt $\vec{v}$ und $\vec{w}$ linear abhängig, so ist etwa $\vec{w} = \alpha\vec{v}$ mit einem $\alpha \in \mathbb{R}$ oder $\mathbb{C}$. Dann ist aber auch
$$\vec{w} \cdot \vec{w} = \alpha\overline{\alpha}\,(\vec{v} \cdot \vec{v})\,,$$
und zusammen mit $\vec{w} = \alpha\vec{v}$ folgt daraus sofort die obige Formel. ■

c) Die Räume $\mathrm{L}^2(D, \mathbb{R})$ und $\mathrm{L}^2(D, \mathbb{C})$

Der Grund für die etwas allgemeinere Formulierung der CAUCHY-SCHWARZschen Ungleichung liegt vor allem in der Anwendung auf Skalarprodukte, die durch Integrale definiert sind. Im vorigen Abschnitt hatten wir uns davon überzeugt, daß die stetigen Funktionen auf dem Einheitsintervall einen EUKLIDischen Vektorraum bilden, aber dieses

Resultat ist für viele Anwendungen in der technischen Informatik zu schwach:

1.) In der Signalverarbeitung spielen unter anderem Rechteckimpulse eine große Rolle; diese werden beschrieben durch Funktionen, die auf einem gewissen Intervall $[a,\ b]$ gleich einer Konstanten h sind und sonst überall null. Offensichtlich ist eine solche Funktion für $h \ne 0$ an den Stellen a und b nicht stetig.
2.) Bei der Zerlegung eines Signals in seine Frequenzen oder seiner Synthese daraus kann man nicht immer davon ausgehen, daß diese in einem beschränkten Bereich liegen; man muß gelegentlich über ganz $\mathbb{R}$ integrieren.
3.) Da $e^{i\varphi}$ soviel einfacher handhabbar ist als $\sin\varphi$ und $\cos\varphi$, möchte man selbst bei der Behandlung reeller Größen oft lieber komplexwertige Funktionen betrachten.

Wir hätten also gerne ein Skalarprodukt auf dem Vektorraum aller Funktionen von $\mathbb{R}$ nach $\mathbb{C}$. (Dies ist ein $\mathbb{C}$-Vektorraum, denn für zwei Funktionen $f, g\colon \mathbb{R} \to \mathbb{C}$ und zwei komplexe Zahlen λ, μ ist auch
$$\lambda f + \mu g\colon \begin{cases} \mathbb{R} \to \mathbb{C} \\ t \mapsto \lambda f(t) + \mu g(t) \end{cases}$$
eine solche Funktion.) Die auf den ersten Blick naheliegende Verallgemeinerung
$$(f,\ g) = \int\limits_{-\infty}^{\infty} f(t) \cdot \overline{g(t)}\, dt$$
des Skalarprodukts von $\mathcal{C}^0\big([0,\ 1], \mathbb{R}\big)$ zu einem HERMITEschen Skalarprodukt hat den Nachteil, daß dieses Integral im allgemeinen nicht einmal existiert.

(Da wir hier komplexwertige Integranden haben können, sei kurz an die Definition eines komplexwertigen Integrals erinnert: Für eine Funktion $w(t) = u(t) + iv(t)$ ist
$$\int\limits_a^b w(t)\, dt = \int\limits_a^b u(t)\, dt + i \int\limits_a^b v(t)\, dt\,,$$

Realteil und Imaginärteil werden also jeweils für sich integriert und dann wieder zusammengesetzt.)

Die Nichtexistenz eines solchen Integrals kann zwei Gründe haben: Entweder ist die Funktion von vornherein nicht integrierbar, d.h. sie hat keine Stammfunktion, oder aber das Integral divergiert. Der erste Fall braucht uns nicht sonderlich zu interessieren: Zu den Funktionen, die keine Stammfunktion haben, zählen in erster Linie seltsame Funktionen wie etwa die DIRICHLETsche Sprungfunktion, die auf $\mathbb{Q}$ eins ist und auf $\mathbb{R} \setminus \mathbb{Q}$ null; die Funktionen, mit denen man es in der Praxis zu tun hat, sind integrierbar.

Der zweite Fall ist ernster: Schon für das triviale Beispiel $f = g = 1$ existiert $(1, 1)$ offensichtlich nicht, da $\int_{-\infty}^{\infty} 1\,dt$ divergiert. Im Falle von $\mathcal{C}^0\big([0, 1], \mathbb{R}\big)$ haben wir aber gesehen, daß die „Länge" eines Vektors zumindest bei der dort bertrachteten Anwendung auf Ströme bis auf eine Konstante gleich der Energie des Signals war; da diese in jeder realen Anwendung endlich sein muß, spricht nur wenig dagegen, wenn wir uns auf Funktionen beschränken, für die das (f, f) entsprechende Integral existiert, d.h. wir beschränken uns auf den Raum

$$\widetilde{\mathrm{L}}^2(\mathbb{R}, \mathbb{C}) = \left\{ \begin{matrix} f\colon \mathbb{R} \to \mathbb{C} \\ \text{integrierbar} \end{matrix} \;\middle|\; \|f\|^2 \underset{\text{def}}{=} \int_{-\infty}^{\infty} f(t) \cdot \overline{f(t)}\, dt < \infty \right\}.$$

Allgemeiner sei Intervall $D \subseteq \mathbb{R}$ irgendein Intervall, das auch nach einer oder beiden Seiten hin unendlich sein darf; auch hierfür definieren wir

$$\widetilde{\mathrm{L}}^2(D, \mathbb{C}) = \left\{ \begin{matrix} f\colon D \to \mathbb{C} \\ \text{integrierbar} \end{matrix} \;\middle|\; \|f\|^2 \underset{\text{def}}{=} \int_D f(t) \cdot \overline{f(t)}\, dt < \infty \right\}$$

und entsprechend auch

$$\widetilde{\mathrm{L}}^2(D, \mathbb{R}) = \left\{ \begin{matrix} f\colon D \to \mathbb{R} \\ \text{integrierbar} \end{matrix} \;\middle|\; \|f\|^2 \underset{\text{def}}{=} \int_D f(t)^2\, dt < \infty \right\}.$$

Für ein endliches abgeschlossenes Intervall $D = [a, b]$ ist die Endlichkeitsbedingung natürlich automatisch erfüllt, und in diesem Fall gibt es

für zwei integrierbare Funktionen f und g auch keine Probleme mit der Existenz von

$$(f, g) \underset{\text{def}}{=} \int_D f(t) \cdot \overline{g(t)}\, dt\,.$$

Für allgemeinere D, insbesondere für unendliche Intervalle wie $\mathbb{R}$ selbst, muß die Existenz dieses Integrals erst noch gezeigt werden. Dazu gehen wir aus von der trivialen Ungleichung

$$(a - b)^2 = a^2 - 2ab + b^2 \geq 0 \quad \text{für alle } a, b \in \mathbb{R}\,,$$

die wir auch als

$$a \cdot b \leq \frac{a^2 + b^2}{2}$$

schreiben können. Insbesondere ist für zwei Funktionen f, g aus $\widetilde{\mathrm{L}}^2(D, \mathbb{C})$ oder $\widetilde{\mathrm{L}}^2(D, \mathbb{R})$ für jeden Punkt $t \in D$

$$\left| f(t) \cdot \overline{g(t)} \right| = |f(t)| \cdot |g(t)| \leq \frac{|f(t)|^2 + |g(t)|^2}{2} = \frac{f(t) \cdot \overline{f(t)} + g(t)\overline{g(t)}}{2}\,,$$

das nach Voraussetzung konvergente Integral

$$\int_D \frac{|f(t)|^2 + |g(t)|^2}{2}\, dt = \frac{1}{2} \int_D f(t) \cdot \overline{f(t)}\, dt + \frac{1}{2} \int_D g(t)\overline{g(t)}\, dt$$

ist also eine konvergente Majorante von

$$\int_D \left| f(t) \cdot \overline{g(t)} \right|\, dt\,.$$

Wenn das Integral über den Betrag einer Funktion konvergiert, konvergiert erst recht das über die Funktion selbst, d.h.

$$(f, g) \underset{\text{def}}{=} \int_D f(t) \cdot \overline{g(t)}\, dt$$

existiert.

Trotzdem ist $\widetilde{\mathrm{L}}^2(D, \mathbb{R})$ kein EUKLIDischer und $\widetilde{\mathrm{L}}^2(D, \mathbb{C})$ kein HERMITEscher Vektorraum, denn aus dem Verschwinden von (f, f) können wir nicht schließen, daß f die Nullfunktion sei: Zumindest jede Funktion f, die an nur endlich vielen Stellen von null verschieden ist, liegt auch in $\widetilde{\mathrm{L}}^2(\mathbb{R}, \mathbb{C})$ und erfüllt ebenfalls die Gleichung $(f, f) = 0$. Leider lassen sich solche Funktionen nur schwer ausschließen, denn beispielsweise sind

$$f_1\colon \begin{cases} \mathbb{R} \to \mathbb{C} \\ t \mapsto \begin{cases} 1 & \text{für } 0 \le t \le 1 \\ 0 & \text{sonst} \end{cases} \end{cases} \text{ und } f_2\colon \begin{cases} \mathbb{R} \to \mathbb{C} \\ t \mapsto \begin{cases} 1 & \text{für } 0 < t < 1 \\ 0 & \text{sonst} \end{cases} \end{cases}$$

(neben vielen anderen Funktionen) beides legitime Beschreibungen eines Rechteckimpulses, und $f_1 - f_2$ ist gerade die Funktion, die an den Stellen 0 und 1 den Wert 1 annimmt und sonst überall verschwindet.

Unter dem Gesichtspunkt von Anwendungen wird man meist sagen, daß beide Funktionen denselben Rechteckimpuls beschreiben; es ist sicherlich nicht möglich, zwei unterschiedliche Signale zu erzeugen, von denen eines besser durch f_1 und das andere besser durch f_2 beschrieben wird. (Es ist ohnehin in der Praxis unmöglich, einen echten Rechteckimpuls zu erzeugen, aber das ist ein anderes Problem.)

Offensichtlich sind die Funktionen $f\colon \mathbb{R} \to \mathbb{R}$ oder $\mathbb{C}$ sehr viel „feiner“ als real existierende Signale; es liegt daher nahe, Funktionen, die „dasselbe“ Signal beschreiben, miteinander zu identifizieren. Wenn wir $\|f\|^2 = \sqrt{(f, f)}$ als Energie interpretieren, bietet sich an, zwei Funktionen genau dann miteinander zu identifizieren, wenn $\|f - g\| = 0$ ist; wir definieren also

$$f \sim g \underset{\text{def}}{\Longleftrightarrow} \|f - g\| = 0\,.$$

Die Menge aller Funktionen, die zu f äquivalent sind, bezeichnen wir mit

$$[f] \underset{\text{def}}{=} \{g \mid f \sim g\}\,;$$

anschaulich interpretiert ist $[f]$ also das durch f und die anderen Elemente dieser Menge beschriebene Signal. Wir setzen

$$\mathrm{L}^2(D, \mathbb{C}) = \{[f] \mid f \in \widetilde{\mathrm{L}}^2(D, \mathbb{C})\}\,,\ \mathrm{L}^2(D, \mathbb{R}) = \{[f] \mid f \in \widetilde{\mathrm{L}}^2(D, \mathbb{R})\}$$

und wollen zeigen, daß dies ein HERMITEscher und ein EUKLIDischer Vektorraum sind. Nach Konstruktion von f ist klar, daß $\|f\| = 0$ genau dann, wenn $[f] = [0]$ ist, wir müssen also nur zeigen, daß die Vektorraumoperationen und das Produkt wohldefiniert ist. Dazu reicht, daß für $f_1 \sim f$ und $g_1 \sim g$ und zwei Körperelemente λ, μ auch $\lambda f_1 + \mu g_1 \sim \lambda f_1 + \mu g_1$ ist. Zunächst ist

$$\begin{aligned}
&\|(\lambda f_1 + \mu g_1) - (\lambda f_1 + \mu g_1)\|^2 = \|\lambda(f_1 - g) + \mu(g_1 - g)\|^2 \\
&= \big(\lambda(f_1 - f) + \mu(g_1 - g),\ \lambda(f_1 - g) + \mu(g_1 - g)\big) \\
&= \lambda\big(f_1 - f,\ \lambda(f_1 - g) + \mu(g_1 - g)\big) \\
&\qquad + \mu\big(g_1 - g,\ \lambda(f_1 - g) + \mu(g_1 - g)\big) \\
&= \lambda\overline{\lambda}\big(f_1 - f,\ f_1 - f\big) + \lambda\overline{\mu}\big(f_1 - f,\ f_1 - f\big) \\
&\qquad + \mu\overline{\lambda}\big(g_1 - g,\ f_1 - f\big) + \mu\overline{\mu}\big(g_1 - g,\ g_1 - g\big)\,.
\end{aligned}$$

In dieser Summe sind die Terme $(f_1 - f,\ f_1 - f)$ und $(g_1 - g,\ g_1 - g)$ nach Voraussetzung gleich null, und nach der CAUCHY-SCHWARZschen Ungleichung ist auch $|(f_1 - f,\ g_1 - g)| \le \|f_1 - f\| \cdot \|g_1 - g\| = 0$, genauso $|(g_1 - g,\ f_1 - f)|$. Also verschwindet die gesamte Summe.

d) Orthonormalbasen

Für den $\mathbb{R}^3$ wissen wir, daß man das (Standard-)Skalarprodukt auch ausrechnen kann als Produkt der Längen der beiden Vektoren mal dem Cosinus des eingeschlossenen Winkels. Insbesondere ist also das Skalarprodukt zweier von $\vec{0}$ verschiedener Vektoren genau dann gleich null, wenn der Cosinus des eingeschlossenen Winkels verschwindet, d.h. wenn der Winkel gleich $\pi/2$ oder $3\pi/2$ ist. Geometrisch gesehen heißt dies, daß die beiden Vektoren aufeinander senkrecht stehen oder, vornehmer ausgedrückt, *orthogonal* zueinander sind.

Diese Sprechweise wollen wir auch für beliebige EUKLIDische und HERMITEsche Vektorräume verwenden: Zwei Vektoren $\vec{v}, \vec{w} \in V$ heißen *orthogonal,* wenn $\vec{v} \cdot \vec{w} = 0$ ist.

Definition: *a)* Eine Basis $\mathcal{B}$ eines EUKLIDischen oder HERMITEschen Vektorraums heißt *Orthogonalbasis,* wenn jedes Element von $\mathcal{B}$ orthogonal zu allen übrigen ist.
b) Eine Orthogonalbasis $\mathcal{B}$ heißt *Orthonormalbasis,* wenn zusätzlich für jeden Vektor $\vec{b} \in \mathcal{B}$ gilt: $\vec{b} \cdot \vec{b} = 1$.

Standardbeispiel sind die Einheitsvektoren

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad \vec{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix},$$

die sowohl für den $\mathbb{R}^n$ als auch für $\mathbb{C}^n$ eine Orthonormalbasis bilden.

Tatsächlich hat *jeder* EUKLIDische oder HERMITEsche Vektorraum eine Orthonormalbasis; für den Beweis werden wir uns allerdings, wie immer, wenn von Basen die Rede ist, zur Vermeidung von logischen Schwierigkeiten auf den endlichdimensionalen Fall beschränken:

Satz: Jeder endlichdimensionale EUKLIDische oder HERMITEsche Vektorraum V hat eine Orthonormalbasis.

Beweis: Wir wissen bereits, daß V *irgendeine* Basis $\{\vec{b}_1, \dots, \vec{b}_n\}$ hat; daraus konstruieren wir zunächst eine *Orthogonalbasis* nach dem sogenannten GRAM-SCHMIDTschen Orthogonalisierungsverfahren. Dieses Verfahren besteht aus n Schritten, wobei im r-ten Schritt eine Orthogonalbasis $\{vc_1, \dots, \vec{c}_r\}$ von $[\vec{b}_1, \dots, \vec{b}_r]$ konstruiert wird.

Der erste Schritt ist der einfachste: Da es noch keine Orthogonalitätsbedingung für $\vec{c}_1$ gibt, können wir einfach $\vec{c}_1 = \vec{b}_1$ setzen.

Nachdem wir $r \geq 1$ Schritte durchgeführt haben, haben wir r linear unabhängige Vektoren $\vec{c}_1, \dots, \vec{c}_r$ mit $\vec{c}_i \cdot \vec{c}_j = 0$ für $i \neq j$ aus dem von $\vec{b}_1, \dots, \vec{b}_r$ aufgespannten Untervektorraum. Ist $r = n$, haben wir eine Orthogonalbasis; andernfalls muß ein auf den bisher konstruierten $\vec{c}_i$ senkrecht stehender Vektor $\vec{c}_{r+1}$ gefunden werden, der zusammen mit diesen den von $\vec{b}_1$ bis $\vec{b}_{r+1}$ erzeugten Untervektorraum erzeugt.

Da $\vec{c}_1, \dots, \vec{c}_r$ und $\vec{b}_1, \dots, \vec{b}_r$ denselben Untervektorraum erzeugen, gilt dasselbe für $\vec{c}_1, \dots, \vec{c}_r, \vec{b}_{r+1}$ und $\vec{b}_1, \dots, \vec{b}_{r+1}$; das Problem ist, daß $\vec{b}_{r+1}$ im allgemeinen nicht orthogonal zu den $\vec{c}_i$ sein wird. Wir dürfen $\vec{b}_{r+1}$ aber abändern durch einen beliebigen Vektor aus dem von $\vec{c}_1, \dots, \vec{c}_r$ aufgespannten Untervektorraum; also setzen wir

$$\vec{c}_{r+1} = \vec{b}_{r+1} + \lambda_1 \vec{c}_1 + \cdots + \lambda_r \vec{c}_r$$

und versuchen, die λ_i so zu bestimmen, daß dieser Vektor orthogonal zu $\vec{c}_1, \dots, \vec{c}_r$ wird.

Wegen der Orthogonalität der $\vec{c}_i$ ist

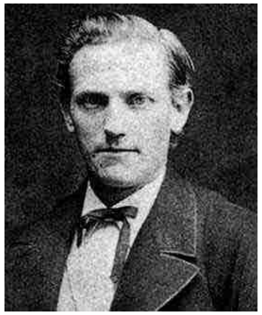
$$\vec{c}_{r+1} \cdot \vec{c}_i = \vec{b}_{r+1} \cdot \vec{c}_i + \sum_{j=1} \lambda_j (\vec{c}_j \cdot \vec{c}_i) = \vec{b}_{r+1} \cdot \vec{c}_i + \lambda_i (\vec{c}_i \cdot \vec{c}_i);$$

setzen wir daher $\lambda_i = -\dfrac{\vec{b}_{r+1} \cdot \vec{c}_i}{\vec{c}_i \cdot \vec{c}_i}$, so ist $\vec{v} \cdot \vec{c}_i = 0$ für alle $i = 1, \dots, r$.

Nach dem n-ten Schritt haben wir eine Orthogonalbasis $\{\vec{c}_1, \dots, \vec{c}_n\}$ von V konstruiert. Daraus wird die gewünschte *Orthonormalbasis* $\{\vec{e}_1, \dots, \vec{e}_n\}$, wenn wir jeden Vektor durch seine Länge dividieren, d.h.

$$\vec{e}_i = \frac{\vec{c}_i}{\|\vec{c}_i\|}.$$

■



Der dänische Mathematiker JØRGAN PEDERSEN GRAM (1850–1916) lehrte an der Universität Kopenhagen, war aber gleichzeitig auch noch geschäftsführender Direktor einer Versicherungsgesellschaft und Präsident des Verbands der dänischen Versicherungsunternehmen. Er publizierte anscheinend nur eine einzige mathematische Arbeit *Sur quelque théorèmes fondamentaux de l'algèbre moderne,* die 1874 erschien. Das GRAM-SCHMIDTsche Orthogonalisierungsverfahren, durch das er heute hauptsächlich bekannt ist, stammt wohl von LAPLACE (1749–1827) und wurde auch schon 1836 von CAUCHY verwendet.

ERHARD SCHMIDT (1876–1959) wurde in Estland geboren; er studierte in Berlin bei SCHWARZ und promovierte in Göttingen bei HILBERT.
ERHARD SCHMIDT ist einer der Begründer der modernen Funktionalanalysis; insbesondere geht die Verallgemeinerung EUKLIDischer und HERMITEscher Vektorräumen zu sogenannten HILBERT-Räumen, mit der wir uns im Rest dieses Semesters bei Anwendungen der linearen Algebra auf FOURIER-Analysis und Differentialgleichungen noch beschäftigen werden, auf ihn zurück. Er war Professor in Zürich, Erlangen, Breslau und Berlin.