Kapitel 2
Differential- und Integralrechnung mehrerer Veränderlicher

Im letzen Kapitel haben wir (homogene) lineare Funktionen mehrerer Veränderlicher betrachtet; damit lassen sich zwar viele wichtige Probleme beschreiben, aber in noch mehr Anwendungen hat man es mit nichtlinearen Funktionen zu tun. Bei diesen sind, wie im Eindimensionalen, Differentiation und Integration wesentliche Hilfsmittel zu einem besseren Verständnis.

Eines der ersten Ziele dieses Kapitels muß daher sein, Stetigkeit und Differenzierbarkeit für Funktionen mehrerer Veränderlicher zu erklären; sobald wir diese verstanden haben, müssen wir uns auch mit verschiedenen Verallgemeinerungen des Integrals auf die mehrdimensionale Analysis beschäftigen.

Im Eindimensionalen hat die Ableitung $f'(x)$ einer Funktion $f(x)$ die Eigenschaft, daß für kleine Werte von h

$$f(x+h) \approx f(x) + hf'(x)$$

ist, es geht also darum, die Funktion f lokal durch eine lineare Funktion zu approximieren. Ähnlich ist es auch im Mehrdimensionalen; wir wollen also Vektorräume und lineare Abbildungen nicht vergessen, sondern sie zu einem der Grundsteine der mehrdimensionalen Analysis machen.

Dazu brauchen wir noch mindestens ein weiteres Hilfsmittel: Bei der Definition von Grenzwerten, Stetigkeit *usw.* spielen Abstände eine große Rolle. Von diesen war bislang noch nicht die Rede, und für Vektorräume über beliebigen Körpern lassen sie sich in der Tat auch gar nicht sinnvoll definieren. In der Differential- und Integralrechnung interessieren

aber nur reelle (und später auch komplexe) Vektorräume; der nächste Paragraph zeigt, wie man solche Begriffe im Reellen definieren kann. Vom Inhalt her könnte dieser Paragraph genauso gut im vorigen Kapitel stehen; der Grund für seine Aufnahme ins zweite Kapitel liegt vor allem daran, daß wir EUKLIDische Vektorräume und ihre Verallgemeinerungen praktisch nur im Zusammenhang mit analytischen Fragen brauchen werden.

§1: Euklidische und normierte Vektorräume

a) Euklidische Vektorräume

Im letzten Kapitel hatten wir für die Definition von Determinanten im $\mathbb{R}^2$ und $\mathbb{R}^3$ das dortige Skalarprodukt betrachtet. Solche Skalarprodukte wollen wir nun für allgemeine reelle Vektorräume definieren. Wir beginnen mit einem etwas schwächeren Begriff, der sich für Vektorräume über beliebigen Körpern definieren läßt:

Definition: Eine *Bilinearform* auf dem k-Vektorraum V ist eine Abbildung $\cdot\colon V \times V \to k$ mit folgenden Eigenschaften:
a) $\cdot$ ist *bilinear,* d.h.

$$(\lambda\vec{v}_1 + \mu\vec{v}_2)\cdot\vec{w} = \lambda(\vec{v}_1\cdot\vec{w}) + \mu(\vec{v}_2\cdot\vec{w}) \qquad \text{und}$$
$$\vec{v}\cdot(\lambda\vec{w}_1 + \mu\vec{w}_2) = \lambda(\vec{v}\cdot\vec{w}_1) + \mu(\vec{v}\cdot\vec{w}_2)$$

für alle $\vec{v}_1, \vec{v}_2, \vec{w}, \vec{w}_1, \vec{w}_2 \in V$ und $\lambda, \mu \in k$.
b) $\cdot$ ist *symmetrisch,* d.h.

$$\vec{v}\cdot\vec{w} = \vec{w}\cdot\vec{v} \quad \text{für alle } \vec{v}, \vec{w} \in V\,.$$

Die Skalarprodukte in $\mathbb{R}^2$ und $\mathbb{R}^3$ sind natürlich Bilinearformen im Sinne dieser Definition; allgemeiner wird für jeden $\mathbb{R}^n$ durch

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = v_1w_1 + \cdots + v_nw_n$$

eine Bilinearform definiert.

Noch allgemeiner e rklärt diese Formel auch für die Vektorräume k^n über einem beliebigen Körper k eine Bilinearform, die auch für (nach unserem derzeitigen Kenntnsisstand) eher exotische Körper durchaus nützliche Anwendungen haben kann: Ist etwa $\mathbb{F}_2 = \{0, 1\}$ der Körper mit zwei Elementen (in dem Addition und Multiplikation modulo zwei definiert sind), so ist das Produkt eines Vektors $\vec{v} \in \mathbb{F}_2^n$ mit dem Vektor, dessen sämtliche Komponenten gleich eins sind, genau dann gleich null, wenn die Anzahl der Einsen im Vektor $\vec{v}$ gerade ist; ansonsten ist es null. Auf diese Weise läßt sich also eine Paritätsprüfung einfach und kompakt formulieren, und wenn man außer dem Vektor mit lauter Einsen als Komponenten noch weitere geeignete Vektoren wählt, lassen sich nicht nur Paritätsfehler, sondern auch noch andere Fehler erkennen und eventuell sogar korrigieren.

Verglichen mit dem Skalarprodukt, wie wir es aus $\mathbb{R}^2$ und $\mathbb{R}^3$ gewohnt sind, fehlt diesen Bilinearformen jedoch eine wesentliche Eigenschaft: In $\mathbb{R}^2$ und $\mathbb{R}^3$ ist das Skalarprodukt eines Vektors mit sich selbst gleich dem Quadrat der Länge und verschwindet somit genau dann, wenn der Vektor gleich dem Nullvektor ist. Dies gilt auch für die gerade definierte Bilinearform auf $\mathbb{R}^n$, denn

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = v_1^2 + \cdot + v_n^2$$

verschwindet als Summe von Quadraten genau dann, wenn jedes einzelne v_i verschwindet.

Über dem Körper $\mathbb{F}_2$ dagegen ist beispielsweise

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

und auch über den komplexen Zahlen ist etwa in $\mathbb{C}^2$

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

Auch über den reellen Zahlen lassen sich Bilinearformen finden, für die das Produkt eines Vektors mit sich selbst verschwinden kann, ohne daß der Vektor gleich dem Nullvektor sein müste: Für die Bilinearform

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \star \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \underset{\text{def}}{=} v_1 w_1 - v_2 w_2$$

etwa ist

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} \star \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 1^2 - 1^2 = 0\,.$$

Da wir im $\mathbb{R}^n$ (und später auch in allgemeineren Räumen) Längen und Abstände definieren wollen, um so auch dort analytische Grundbegriffe wie Konvergenz und Stetigkeit einführen zu können, müssen wir solche Bilinearformen ausschließen: Abstände dürfen nicht negativ sein, und der Abstand zwischen zwei Punkten soll nur dann verschwinden, wenn beide Punkte gleich sind.

Da man in beliebigen Körpern nicht von Positivität und Negativität reden kann, beschränken wir uns ab jetzt auf den Fall $k = \mathbb{R}$ und definieren:

Definition: *a)* Eine Bilinearform $V \times V \to \mathbb{R}$ auf einem reellen Vektorraum V heißt *positiv semidefinit,* wenn

$$\vec{v} \cdot \vec{v} \geq 0 \quad \text{für alle } \vec{v} \in V\,;$$

sie heißt *positiv definit* oder *Skalarprodukt,* wenn zusätzlich gilt

$$\vec{v} \cdot \vec{v} = 0 \Rightarrow \vec{v} = \vec{0}\,.$$

b) Ein EUKLIDischer Vektorraum ist ein Paar $(V, \cdot)$ bestehend aus einem $\mathbb{R}$-Vektorraum V und einem Skalarprodukt $\cdot : V \times V \to \mathbb{R}$.

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; in der Literatur findet man manchmal 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. EUKLID ist vor allem bekannt als Autor der *Elemente,* in denen er die Geometrie seiner Zeit systematisch darstellte und (in gewisser Weise) auf wenige Definitionen sowie die berühmten fünf Postulate zurückführte; diese Elemente entstanden um 300 v. Chr. und waren zwar nicht der erste, aber doch der erfolgreichste Versuch einer solchen Zusammenfassung. EUKLID arbeitete wohl am Museion in Alexandrien; außer den Elementen schrieb er auch ein Buch über Optik und weitere, teilweise verschollene Bücher.

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

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = v_1 w_1 + \cdots + v_n w_n \,.$$

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 Bilinearität des so definierten Produkts ist klar; außerdem ist es offensichtlich positiv semidefinit, denn

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

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

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

$(f,\,f) = 0$ 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 die Nullfunktion ist.

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 0° 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 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: Das Skalarprodukt im $\mathbb{R}^3$ hatten wir in Kapitel I, §*7, definiert* als Produkt aus den Längen der beiden Vektoren und dem Cosinus des eingeschlossenen Winkels; wir hatten dann nachgerechnet, daß dies mit der Koordinatendefinition des hier definierten Skalarprodukts 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

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

Cauchy-Schwarzsche Ungleichung: In einem EUKLIDischen Vektorraum $(V, \cdot)$ ist für alle Vektoren $\vec{v}, \vec{w} \in V$

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

mit Gleichheit genau dann, wenn $\vec{v}$ und $\vec{w}$ linear abhängig sind.

Der *Beweis* beruht auf folgendem Trick: Wegen der positiven Definitheit des Skalarprodukts ist für alle $\lambda, \mu \in \mathbb{R}$

$$(\lambda\vec{v} + \mu\vec{w}) \cdot (\lambda\vec{v} + \mu\vec{w}) = \lambda^2\, \vec{v} \cdot \vec{v} + 2\lambda\mu\, \vec{v} \cdot \vec{w} + \mu^2\, \vec{w} \cdot \vec{w} \ge 0 \,.$$

Speziell für $\lambda = (\vec{w} \cdot \vec{w})$ und $\mu = -(\vec{v} \cdot \vec{w})$ erhalten wir

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

also ist

$$\|\lambda\vec{v} + \mu\vec{w}\|^2 = \|\vec{w}\|^2 \left(\|\vec{w}\|^2 \|\vec{v}\|^2 - |\vec{v} \cdot \vec{w}|^2 \right) \ge 0 \,.$$

Für $\|\vec{w}\| = 0$ sagt uns diese Ungleichung nichts, aber da dann $\vec{w} = \vec{0}$ sein muß, ist die Behauptung in diesem Fall ohnehin klar: Beide Seiten der CAUCHY-SCHWARZschen Ungleichung sind null, und $\vec{w} = 0$ ist linear abhängig von $\vec{v}$.

Für $\|\vec{w}\| \neq 0$ können wir durch $\|\vec{w}\|^2$ dividieren und sehen, daß

$$\frac{\|\lambda\vec{v} + \mu\vec{w}\|^2}{\|w\|^2} = \|v\|^2 \|w\|^2 - (\vec{v} \cdot \vec{w})^2 \ge 0$$

ist. Im Falle der Gleichheit muß der Vektor $\lambda\vec{v} + \mu\vec{w}$ gleich dem Nullvektor sein, d.h. $\vec{v}$ und $\vec{w}$ sind linear abhängig. ■

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.

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

Arbeit, die während des Zeitraums [0, 1] verrichtet wird, 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.

Ähnliche physikalische Interpretationen gibt es bei fast allen Anwendungen von Vektorräumen, deren Elemente Funktionen sind.

b) Orthonormalbasen

Beim Standardskalarprodukt im $\mathbb{R}^n$ ist das Produkt zweier verschiedener Einheitsvektoren stets null und alle Einheitsvektoren haben die Länge eins. Diese nützliche Eigenschaft einer Basis wollen wir nun auch für beliebige EUKLIDische Vektorräume genauer betrachten.

Definition: *a)* Zwei Vektoren $\vec v, \vec w \in V$ eines EUKLIDischen Vektorraums heißen *orthogonal,* wenn $\vec v\cdot\vec w = 0$ ist.
b) Eine Basis $\mathcal{B}$ von V 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: $\vec b\cdot\vec b = 1$.

Der Leser sollte sich an dieser Stelle zumindest kurz überlegen, daß zwei zueinander orthogonale Vektoren den Winkel $\pi/2$ oder 90° einschließen, also auch im Sinne der Elementargeometrie aufeinander senkrecht stehen.

Die Nützlichkeit von Orthonormalbasen ergibt sich aus folgender Beobachtung: Ist $\mathcal{B}$ eine Basis des EUKLIDischen Vektorraums V und sind

$$\vec v = v_1\vec b_1 + \cdots + v_n\vec b_n \quad\text{und}\quad \vec w = w_1\vec b_1 + \cdots + w_n\vec b_n$$

die Basisdarstellungen zweier Vektoren aus V, so ist wegen der Bilinearitätseigenschaft

$$\vec v\cdot\vec w = \sum_{i=1}^n\sum_{j=1}^n v_iw_j(\vec b_i\cdot\vec b_j)\,,$$

eine Summe aus n^2 Termen, unter denen allerdings aufgrund der Symmetrie höchsten $\frac12 n(n+1)$ verschiedene sind.

Im Fall einer Orthogonalbasis, verschwindet $\vec b_i\cdot\vec b_j$ wann immer $i\neq j$ ist; dann reduziert sich die Summe auf

$$\vec v\cdot\vec w = \sum_{i=1}^n v_iw_i(\vec b_i\cdot\vec b_i)\,.$$

Für eine Orthonormalbasis schließlich ist $\vec b_i\cdot\vec b_i = 1$, die Summe vereinfacht sich also noch etwas weiter zu

$$\vec v\cdot\vec w = \sum_{i=1}^n v_iw_i\,.$$

Speziell im Endlichdimensionalen, wo wir $\mathcal{B} = \{\vec b_1,\dots,\vec b_n\}$ setzen können, läßt sich ein beliebiges Skalarprodukt also bezüglich einer Orthonormalbasis auf die gleiche Weise berechnen wie das Skalarprodukt im $\mathbb{R}^n$. Ein endlichdimensionaler EUKLIDischer Vektorraum, der eine Orthonormalbasis hat, unterscheidet sich also nicht wesentlich vom $\mathbb{R}^n$.

Tatsächlich hat *jeder* EUKLIDische Vektorraum eine Orthonormalbasis. Für den Beweis wollen wir uns allerdings, wie schon bei der Existenz von Basen auf den endlichdimensionalen Fall beschränken.

Satz: Jeder endlichdimensionale EUKLIDische 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 $\{\vec c_1,\dots,\vec c_r\}$ 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

$\vec{b}_1, \ldots, \vec{b}_r$ aufgespannten Untervektorraum. Ist $r = n$, haben wir eine Orthogonalbasis von V gefunden; 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 $\vec{c}_1, \ldots, \vec{c}_r$ und $\vec{b}_1, \ldots, \vec{b}_r$ denselben Untervektorraum erzeugen, gilt dasselbe für $\vec{c}_1, \ldots, \vec{c}_r, \vec{b}_{r+1}$ und $\vec{b}_1, \ldots, \vec{b}_{r+1}$; 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 $\vec{c}_1, \ldots, \vec{c}_r$ 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 $\vec{c}_1, \ldots, \vec{c}_r$ wird.

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

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

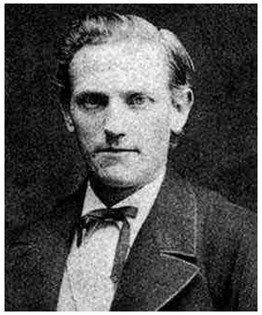
setzen wir daher $\lambda_i = -\dfrac{\vec{b}_{r+1} \cdot \vec{c}_i}{\vec{c}_i \cdot \vec{c}_i}$, so ist $\vec{v} \cdot \vec{c}_i = 0$ für alle $i = 1, \ldots, r$.

Damit haben wir eine Orthogonalbasis $\{\vec{c}_1, \ldots, \vec{c}_{r+1}\}$ von $[\vec{b}_1, \ldots, \vec{b}_{r+1}]$ konstruiert.

Nach dem n-ten Schritt haben wir demnach eine Orthogonalbasis $\{\vec{c}_1, \ldots, \vec{c}_n\}$ von V. Daraus erhalten wir die gewünschte *Orthonormalbasis* $\{\vec{e}_1, \ldots, \vec{e}_n\}$, indem wir jeden Basisvektor 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 Vektorräume zu sogenannten HILBERT-Räumen, mit der wir uns im nächsten Semester im Zusammenhang mit der FOURIER-Analysis und der Theorie der Differentialgleichungen beschäftigen werden, auf ihn zurück. Er war Professor in Zürich, Erlangen, Breslau und Berlin.

Als *Beispiel* zum GRAM-SCHMIDTschen Orthogonalisierungsverfahren betrachten wir den Untervektorraum V des $\mathbb{R}^5$, der von den drei Vektoren

$$\vec{b}_1 = \begin{pmatrix} 2\\2\\3\\4\\4 \end{pmatrix}, \quad \vec{b}_2 = \begin{pmatrix} 4\\4\\3\\2\\4 \end{pmatrix} \quad \text{und} \quad \vec{b}_3 = \begin{pmatrix} 3\\3\\3\\6\\1 \end{pmatrix}$$

aufgespannt wird.

Im ersten Schritt setzen wir einfach $\vec{c}_1 = \vec{b}_1$.

Im zweiten Schritt müssen wir einen Vektor $\vec{c}_2 = \vec{b}_2 + \lambda_1 \vec{c}_1$ finden, so daß $\vec{c}_1 \cdot \vec{c}_2 = 0$ ist. Da $\vec{b}_2 \cdot \vec{c}_1 = 49$ und $\vec{c}_1 \cdot \vec{c}_1 = 49$ beide gleich sind, müssen wir nach GRAM-SCHMIDT $\lambda_1 = -1$ setzen, d.h.

$$\vec{c}_2 = \vec{b}_2 - \vec{c}_1 = \vec{b}_2 - \vec{b}_1 = \begin{pmatrix} 2 \\ 2 \\ 0 \\ -2 \\ 0 \end{pmatrix} .$$

Im dritten Schritt suchen wir einen Vektor $\vec{c}_3 = \vec{b}_3 + \mu_1\vec{c}_1 + \mu_2\vec{c}_2$ mit $\vec{c}_3 \cdot \vec{c}_1 = \vec{c}_2 \cdot \vec{c}_1 = 0$. Da $\vec{b}_3 \cdot \vec{c}_1 = \vec{c}_1 \cdot \vec{c}_1 = 49$ ist, setzen wir $\mu_1 = -1$; der zweite Koeffizient μ_2 ist Null, da $\vec{b}_3$ und $\vec{c}_2$ bereits orthogonal sind. Damit wird

$$\vec{c}_3 = \vec{b}_3 - \vec{c}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 2 \\ -3 \end{pmatrix} .$$

Schließlich müssen wir diese Orthogonalbasis noch zur Orthonormalbasis machen; wegen

$$\vec{c}_1 \cdot \vec{c}_1 = 49, \quad \vec{c}_2 \cdot \vec{c}_2 = 12 \quad \text{und} \quad \vec{c}_3 \cdot \vec{c}_3 = 15$$

müssen wir die Vektoren $\vec{c}_i$ noch durch

$$\|\vec{c}_1\| = 7, \quad \|\vec{c}_2\| = \sqrt{12} = 2\sqrt{3} \quad \text{und} \quad \|\vec{c}_2\| = \sqrt{15}$$

dividieren und erhalten eine Orthonormalbasis $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ mit

$$\vec{e}_1 = \begin{pmatrix} 2/7 \\ 2/7 \\ 3/7 \\ 4/7 \\ 4/7 \end{pmatrix}, \quad \vec{e}_2 = \begin{pmatrix} \sqrt{3}/3 \\ \sqrt{3}/3 \\ 0 \\ -\sqrt{3}/3 \\ 0 \end{pmatrix}, \quad \vec{e}_3 = \begin{pmatrix} \sqrt{15}/15 \\ \sqrt{15}/15 \\ 0 \\ 2\sqrt{15}/15 \\ -\sqrt{15}/5 \end{pmatrix} .$$

c) Affine und Euklidische Geometrie

EUKLIDische Vektorräume haben ihren Namen natürlich davon, daß man dort so etwas wie EUKLIDische Geometrie betreiben kann. Die Objekte, mit denen es die EUKLIDische Geometrie zu tun hat, sind allerdings

keine Vektoren, sondern Punkte, Geraden, Ebenen *usw.;* genauso werden wir es im Hauptteil dieses Kapitels nicht mit der Differential- und Integralrechnung von Funktionen auf einer Menge von *Vektoren* zu tun haben, sondern mit Funktionen auf Punktmengen.

Wir müssen daher beispielsweise unterscheiden zwischen der *Ebenen* $\mathbb{R}^2$ und dem zweidimensionalen *Vektorraum* $\mathbb{R}^2$: Eine Ebene ist eine Menge von Punkten, während ein Vektorraum aus Vektoren besteht, die jeweils *zwei* Punkte verbinden. Erst wenn wir einen festen Punkt auszeichnen, können wir Vektoren und Punkte in eindeutige Beziehung zueinander setzen. Wir müssen daher unterscheiden zwischen dem Punkt (x, y) aus der Ebenen $\mathbb{R}^2$ und dem Vektor $\binom{x}{y}$ aus dem Vektorraum $\mathbb{R}^2$, der den Punkt $(0, 0)$ mit (x, y) verbindet, aber beispielsweise auch den Punkt $(2-x, 3)$ mit $(2, 3+y)$. Der Begriff des affinen Raums formalisiert diese Unterscheidung:

Definition: Ein affiner Raum über dem Körper k ist gegeben durch ein Paar (A, V) bestehend aus einer nichtleeren Menge A, deren Elemente wir als Punkte bezeichnen, und einem Vektorraum V, so daß gilt:

1.) Es gibt eine Abbildung $\varphi\colon \begin{cases} A \times A \to V \\ (\boldsymbol{x}, \boldsymbol{y}) \mapsto \overrightarrow{\boldsymbol{xy}} \end{cases}$, mit der Eigenschaft, daß

$$\overrightarrow{\boldsymbol{xy}} + \overrightarrow{\boldsymbol{yz}} = \overrightarrow{\boldsymbol{xz}} \quad \text{für alle } \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in A .$$

2.) Für jeden Punkt $O \in A$ ist die Abbildung

$$\varphi_O\colon A \to V; \quad \boldsymbol{x} \mapsto \varphi(O, \boldsymbol{x}) = \overrightarrow{O\boldsymbol{x}}$$

bijektiv; ihre Umkehrabbildung sei mit ψ_O bezeichnet.

Die Dimension des Vektorraums V bezeichnen wir auch als Dimension des affinen Raums.

Anschaulich gesehen ordnet die Abbildung φ zwei Punkten $\boldsymbol{x}, \boldsymbol{y} \in A$ den Verbindungsvektor $\overrightarrow{\boldsymbol{xy}}$ zu, und ψ beschreibt das *Abtragen* eines Vektors: In §1 hatten wir Vektoren eingeführt als Äquivalenzklassen von Pfeilen gleicher Länge und gleicher Richtung; nimmt man zum Vektor $\vec{v}$ den Pfeil, der im Punkt O beginnt, so hat dieser den Endpunkt $\psi_O(\vec{v})$.

Standardbeispiel eines affinen Raums ist der $\mathbb{R}^n$; hier ist A gleich der *Menge* $\mathbb{R}^n$, deren Elemente wir als Punkte $\boldsymbol{x} = (x_1, \ldots, x_n)$ schreiben, und V ist gleich dem *Vektorraum* $\mathbb{R}^n$, dessen Elemente wir als Spaltenvektoren schreiben. φ ist die Abbildung, die den Punkten

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

deren Verbindungsvektor

$$\varphi(\boldsymbol{x}, \boldsymbol{y}) = \overrightarrow{\boldsymbol{xy}} = \begin{pmatrix} y_1 - x_1 \\ \vdots \\ y_n - x_n \end{pmatrix}$$

zuordnet, und für $O = (a_1, \ldots, a_n)$ ist

$$\psi_O\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) = (a_1 + x_1, \ldots, a_n + x_n)\,.$$

Ganz entsprechend läßt sich für jeden beliebigen Körper k die Menge k^n zu einem affinen Raum machen; der zugehörige Vektorraum ist natürlich der *Vektorraum* k^n.

Weitere interessante Beispiele sind vor allem die Geraden, Ebenen *usw.* in diesem Raum; diese fassen wir zusammen unter der

Definition: Ein affiner Unterraum des affinen Raums (A, V) ist ein affiner Raum (B, U) mit $B \subseteq A$ und $U \le V$, wobei auch die Abbildungen φ und ψ_O Einschränkungen der entsprechenden Abbildungen von (A, V) sind.

Betrachten wir etwa die Gerade

$$\left\{(1 + \lambda, \lambda) \;\middle|\; \lambda \in \mathbb{R}\right\} \subseteq \mathbb{R}^2$$

durch den Punkt $(1, 0)$ mit Steigung 1. Hier ist B natürlich gleich der Menge aller Punkte auf der Geraden, also die gerade angegebene Menge selbst, und als Vektorraum nehmen wir den vom Richtungsvektor $\binom{1}{1}$ der Geraden aufgespannten Untervektorraum des $\mathbb{R}^2$.

Dieselbe Gerade können wir auch durch die Gleichung

$$y = x + 1 \quad \text{oder} \quad -x + y = 1$$

beschreiben. Allgemein können wir, wie wir uns als nächstes überlegen wollen, jeden affinen Unterraum eines k^n als Lösungsmenge eines linearen Gleichungssystems in n Unbekannten interpretieren, und umgekehrt ist auch jede solche Lösungsmenge, falls sie nicht leer ist, affiner Unterraum von k^n. (Mit k^n ist hier natürlich das Paar (k^n, k^n) gemeint, das wir kurz, wenn auch schlampig, als affinen Raum k^n bezeichnen.)

Beginnen wir mit der Lösungsmenge L eines linearen Gleichungssystems

$$A\vec{x} = \vec{b} \quad \text{mit} \quad A \in k^{m \times n}, \quad \vec{b} \in k^m$$

für einen Vektor $\vec{x} \in k^n$; mit anderen Worten mit der Lösungsmenge von m linearen Gleichungen in n Unbekannten. (Hier schreiben wir, aus alter Gewohnheit, die Punkte eines affinen Raums ausnahmsweise als Vektoren.) *Falls* es Lösungen gibt, ist die Differenz $\vec{y} = \vec{x}^{(1)} - \vec{x}^{(2)}$ zweier (nicht notwendigerweise verschiedener) Lösungen $\vec{x}^{(1)}$ und $\vec{x}^{(2)}$ eine Lösung des homogenen Gleichungssystems $A\vec{y} = \vec{0}$.

Die Lösungsmenge eines *homogenen* linearen Gleichungssystems in n Variablen ist aber Kern einer linearen Abbildung und daher (und auch aus vielen anderen Gründen) ein Untervektorraum U des k^n. Zusammen mit diesem Untervektorraum wird L zu einem affinen Raum, wobei die Abbildung $\varphi: L \times L \to U$ zwei Lösungen $\vec{x}^{(1)}$ und $\vec{x}^{(2)}$ deren Differenz $\vec{y} = \vec{x}^{(1)} - \vec{x}^{(2)}$ als „Verbindungsvektor“ zuordnet; entsprechend macht für jede Lösung $\vec{x}$ des inhomogenen Gleichungssystems die Abbildung $\psi_{\vec{x}}: U \to L$ eine Lösung $\vec{y}$ des homogenen Gleichungssystems durch Addition von $\vec{x}$ zu einer Lösung des inhomogenen Gleichungssystems.

Ist umgekehrt (B, U) ein affiner Unterraum von (k^n, k^n), so können wir eine Basis $\{\vec{b}_1, \ldots, \vec{b}_r\}$ von U wählen und diese zu einer Basis $\{\vec{b}_1, \ldots, \vec{b}_n\}$ des k^n ergänzen. Bezüglich dieser neuen Basis des k^n ist dann U die Lösungsmenge des homogenen linearen Gleichungssystems

$$x_{r+1} = 0, \quad x_{r+2} = 0, \quad \ldots, \quad x_n = 0\,.$$

Sind nun $(c_1, \ldots, c_n)$ und $(d_1, \ldots, d_n)$ zwei Punkte aus B, so liegt ihre Differenz (d.h. ihr „Verbindungsvektor“) in U, und somit muß für alle $i > r$ notwendigerweise $c_i = d_i$ sein. Damit ist B bezüglich der Basis