Matrix ist nach Definition die Determinante der Spaltenvektoren, und wenn man das λ-fache der j-ten Spalte zur i-ten Spalte addiert, ist

$$\begin{aligned}&\det\bigl(\vec v_1,\dots,\vec v_i+\lambda\vec v_j,\dots,\vec v_n\bigr)\\ &=\det\bigl(\vec v_1,\dots,\vec v_i,\dots,\vec v_n\bigr)+\lambda\det\bigl(\vec v_1,\dots,\vec v_j,\dots,\vec v_n\bigr)\\ &=\det\bigl(\vec v_1,\dots,\vec v_i,\dots,\vec v_n\bigr)\,,\end{aligned}$$

da der zweite Summand den Vektor $\vec v_j$ doppelt enthält und somit verschwindet.

Etwas entsprechendes gilt auch für Zeilen; das könnten wir zwar direkt nachrechnen, aber wenn wir uns daran erinnern, daß die Zeilen einer Matrix gleich den Spalten der transponierten Matrix sind, folgt das viel bequemer aus dem folgenden

Lemma: Für jede $n\times n$-Matrix A ist $\det A=\det {}^tA$.

Beweis: Nach Definition ist für $A=(a_{ij})$

$$\det A=\sum_{\pi\in\mathfrak S_n}\varepsilon(\pi)a_{1\pi(1)}\cdots a_{n\pi(n)}$$

und

$$\det {}^tA=\sum_{\pi\in\mathfrak S_n}\varepsilon(\pi)a_{\pi(1)1}\cdots a_{\pi(n)n}\,.$$

Ordnet man die Faktoren des Produkts $a_{\pi(1)1}\cdots a_{\pi(n)n}$ nach dem ersten Index, so erhält man

$$a_{\pi(1)1}\cdots a_{\pi(n)n}=a_{1\pi^{-1}(1)}\cdots a_{1\pi^{-1}(n)}$$

und damit

$$\det {}^tA=\sum_{\pi\in\mathfrak S_n}\varepsilon(\pi)a_{1\pi^{-1}(1)}\cdots a_{1\pi^{-1}(n)}\,.$$

Da mit π auch π^{-1} die gesamte Gruppe $\mathfrak S_n$ durchläuft und $\varepsilon(\pi)=\varepsilon(\pi^{-1})$ ist, stimmt die Summe dieser Terme mit der aus der Formel für $\det A$ überein, die beiden Determinanten sind also gleich. ■

Für Leser, denen nicht klar ist, warum $\varepsilon(\pi)$ und $\varepsilon(\pi^{-1})$ gleich sind, seien zur Übung des Umgangs mit Permutationen hier kurz zwei Beweise angegeben: Am einfachsten sieht man diese Formel aufgrund des obigen Lemmas, wonach für zwei beliebige Permutationen π und π' gilt: $\varepsilon(\pi\circ\pi')=\varepsilon(\pi)\cdot\varepsilon(\pi')$. Setzt man hier $\pi'=\pi^{-1}$, so folgt

$$1=\varepsilon(\text{Identität})=\varepsilon(\pi)\cdot\varepsilon(\pi^{-1})\,,$$

so daß die beiden Vorzeichen gleich sein müssen.

Alternativ können wir uns auch explizit überlegen, daß sich π^{-1} als Produkt von r Transpositionen schreiben läßt, falls dies für π der Fall ist; genauer gilt: Ist $\pi=\tau_1\circ\cdots\circ\tau_r$, so ist $\pi^{-1}=\tau_r\circ\cdots\circ\tau_1$, denn wie wir bereits im Abschnitt über Permutationen gesehen haben, ist

$$\pi\circ(\tau_r\circ\cdots\circ\tau_1)=(\tau_1\circ\cdots\circ\tau_r)\circ(\tau_r\circ\cdots\circ\tau_1)=\text{Identität}\,,$$

da jede Transposition zu sich selbst invers ist.

Aus obigem Lemma folgt insbesondere, daß man die Determinante einer Matrix auch als Determinante der Zeilenvektoren definieren kann und daß die Spaltenvektoren einer Matrix genau dann linear unabhängig sind, wenn es die Zeilenvektoren sind. Mit nur wenig mehr Aufwand könnte man so auch zeigen, daß der im Zusammenhang mit der Lösung linearer Gleichungssysteme definierte *Rang* einer Matrix genausogut über Zeilen- wie über Spaltenoperationen berechnet werden kann.

Zurück zu den Rechenregeln für Determinanten! Aus dem gerade bewiesenen Lemma folgt sofort:

Eine Determinante ändert ihren Wert nicht, wenn man ein Vielfaches einer Zeile zu einer anderen Zeile addiert, denn die Zeilen einer Matrix sind die Spalten der transponierten Matrix.

Ebenso zeigt man: *Eine Determinante wird mit -1 multipliziert, wenn man entweder zwei Zeilen oder zwei Spalten miteinander vertauscht,* denn für Spalten ist das gerade die definierende Eigenschaft (D1) der Determinanten.

Als letzte Regel sei noch aufgeführt: *Eine Determinante wird mit λ multipliziert, wenn* eine *ihrer Zeilen oder* eine *ihrer Spalten mit λ multipliziert werden.* Für Spalten folgt das sofort aus der Linearität in jedem Argument, für Zeilen folgt es daraus nach Transponieren.

g) Der Entwicklungssatz von Laplace

Hier geht es um ein Berechnungsverfahren, das die definierenden Eigenschaften der Determinante ausnutzt, die „Entwicklung nach einer Zeile oder Spalte".

Für die Entwicklung einer Matrix nach beispielsweise nach der j-ten Spalte schreiben wir den j-ten Spaltenvektor

$$\vec{v}_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix}$$

als Linearkombination der Einheitsvektoren

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

d.h. als

$$\vec{v}_j = a_{1j}\vec{e}_1 + \cdots + a_{nj}\vec{e}_n\,.$$

Dann ist wegen der Linearität der Determinante in ihrem j-ten Argument

$$\begin{aligned} \det A &= \det\big(\vec{v}_1, \dots, \vec{v}_j, \dots, \vec{v}_n\big) = \det\Big(\vec{v}_1, \dots, \sum_{i=1}^n a_{ij}\vec{e}_i, \dots, \vec{v}_n\Big) \\ &= \sum_{i=1}^n a_{ij} \det\big(\vec{v}_1, \dots, \vec{e}_i, \dots, \vec{v}_n\big)\,. \end{aligned}$$

Um daraus die Determinante von A zu berechnen, müssen wir die Determinanten

$$D_{ij} \underset{\text{def}}{=} \det\big(\vec{v}_1, \dots, \vec{e}_i, \dots, \vec{v}_n\big)$$

kennen, wobei $\vec{e}_i$ an der j-ten Stelle steht.

Dazu wenden wir die definierende Formel an: In den Produkten $a_{1\pi(1)} \cdots a_{n\pi(n)}$ ist für die gewählte Stelle j, an der wir den Vektor $\vec{e}_i$ eingefügt haben, $a_{j\ell} = 0$ für $\ell \neq i$ und $a_{ji} = 1$. Also verschwindet $a_{1\pi(1)} \cdots a_{n\pi(n)}$ für jede Permutation π mit $\pi(j) \neq i$.

Am einfachsten läßt sich dies im Fall $i = j = n$ weiter ausrechnen: Dann müssen wir nur Permutationen mit $\pi(n) = n$ betrachten, und die entsprechen eineindeutig den Permutationen auf der Menge aller Zahlen von 1 bis $n-1$; was übrigbleibt ist also (wegen $a_{nn} = 1$) gerade die Determinante jener $(n-1) \times (n-1)$-Matrix, die durch Streichung der letzten Spalte und der letzten Zeile aus A entsteht.

Der allgemeine Fall läßt sich durch Vertauschungen darauf zurückführen: Wir können zwar nicht einfach die j-te Spalte mit der n-ten vertauschen, denn dann steht ja die n-te Spalte an j-ter Stelle, aber wir können die j-te Spalte durch sukzessive Vertauschung mit ihrem rechten Nachbarn immer weiter nach außen bringen: Wir vertauschen zunächst die j-te Spalte mit der $(j+1)$-ten, dann die neue $(j+1)$-te (=alte j-te) Spalte mit der $(j+2)$-ten *usw.*, bis schließlich die alte j-te Spalte an n-ter Stelle steht, ohne daß sich vor oder nach der Stelle j irgendetwas an der relativen Reihenfolge der Spalten geändert hätte. Die Anzahl der dazu notwendigen Vertauschungen ist $(n-j)$, die Determinante wurde bei der gesamten Prozedur also mit $(-1)^{n-j}$ multipliziert.

Als nächstes müssen wir auch noch die i-te Zeile nach unten bringen; dazu verwenden wir die gerade gezeigte Formel $\det A = \det {}^tA$. Die i-te Zeile von A ist gleich der i-ten Spalte von ${}^tA$; wie wir oben gesehen haben, läßt sie sich durch $(n-i)$ Spaltenvertauschungen zur letzen Spalte machen, wobei die Determinante mit $(-1)^{n-i}$ multipliziert wird, und durch nochmaliges Transponieren erhalten wir schließlich jene Matrix, in der die i-te Zeile zur n-ten geworden ist, ohne daß sich vor oder nach der Stelle i irgendetwas an der relativen Reihenfolge der Zeilen geändert hätte.

Die Determinante der so entstehenden Matrix ist, wie wir oben beim Spezialfall $i = j = n$ gesehen haben, gleich der Determinante jener $(n-1) \times (n-1)$-Matrix A_{ij}, die durch Streichen der letzten Zeile und Spalte entsteht; bezogen auf die ursprüngliche Matrix A entsteht sie durch Streichen der i-ten Zeile und der j-ten Spalte. Also ist, wenn wir noch die Vertauschungen berücksichtigen,

$$D_{ij} = (-1)^{n-i} \cdot (-1)^{n-j} \cdot \det A_{ij} = (-1)^{i+j} \det A_{ij}\,,$$

denn $(n-i)+(n-j)=2n-(i+j)$ hat dieselbe Parität wie $(i+j)$. Fassen wir alles zusammen, erhalten wir die Formel

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij} \,.$$

Die Entwicklung nach einer Zeile geht ganz entsprechend: Da die i-te Zeile von A gleich der i-ten Spalte von ${}^tA$ ist, führt obige Rechnung für die transponierte Matrix auf die Formel

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij} \,.$$

Damit haben wir den folgenden Satz bewiesen, der trotz seines Namens bereits LEIBNIZ bekannt war:

Entwicklungssatz von Laplace: A sei eine $n \times n$-Matrix und A_{ij} sei jene $(n-1)\times(n-1)$-Matrix, die durch Streichung der i-ten Zeile und der j-ten Spalte von A entsteht. Dann ist für jedes feste j

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}$$

und für jedes feste i

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij} \,,$$

■

Die erste Formel aus diesem Satz bezeichnet man als die *Entwicklung der Determinante nach ihrer j-ten Spalte,* die zweite (die durch Anwendung der ersten auf die transponierte Matrix entsteht) entsprechend als die *Entwicklung nach der i-ten Zeile.*

PIERRE-SIMON DE LAPLACE (1749–1827) war einer der bedeutendsten französischen Wissenschaftler seiner Zeit; berühmt sind vor allem seine Anwendungen der Analysis auf die Wahrscheinlichkeitstheorie, die Himmelsmechanik und die Potentialtheorie. Bekannt wurde er auch durch die nach ihm und KANT benannte Theorie zur Entstehung des Universums. Als politischer Opportunist kam er gut durch die Wirren seiner Zeit; er saß unter anderem in dem Kommitee, das Maßeinheiten nach dem Dezimalsystem einführte, war kurze Zeit Innenminister und wurde schließlich sogar Graf und Marquis.

BARON GOTTFRIED WILHELM VON LEIBNIZ (1646–1716) gilt als der letzte Universalgelehrte, der das gesamte Wissen seiner Zeit überblickte. In der Mathematik ist er vor allem berühmt durch die Entwicklung der Infinitesimalrechnung (bezüglich derer es einen langen Prioritätsstreit mit NEWTON gab); Bezeichnungen wie $\frac{dy}{dx}$ und $\int f(x)\,dx$ gehen auf ihn zurück. Durch seine Begründung der symbolischen Logik legte er auch einen wesentlichen Grundstein der späteren Informatik. Weitere Arbeiten befassen sich mit den Naturwissenschaften und der Technik, der Philosophie, Theologie und der Geschichte.

Als erste Anwendung wollen wir noch einmal die SARRUSsche Regel für Determinanten von 3×3-Matrizen beweisen: Entwicklung nach der ersten Spalte zeigt, daß

$$\begin{aligned}\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \\ &= a_{11}(a_{22}a_{33}-a_{23}a_{32}) - a_{21}(a_{12}a_{33}-a_{13}a_{32}) \\ &\quad + a_{31}(a_{12}a_{23}-a_{13}a_{22})\end{aligned}$$

ist, was ausmultipliziert natürlich genau die alte Formel ergibt.

Rein theoretisch ist es gleichgültig, nach welcher Zeile oder Spalte man eine Determinante entwickelt, aber man wird natürlich versuchen, eine zu finden, in der möglichst viele Nullen stehen, so daß in obiger Summe möglichst wenige Summanden übrigbleiben. In der Tat überlegt man sich leicht, daß bei der Anwendung des LAPLACEschen Entwicklungssatzes auf eine vollbesetzte Determinante wie bei der definierenden Formel rund $n \cdot n!$ Recheenoperationen erfordert und somit gegenüber der direkten Anwendung dieser Formel nur den Vorteil bietet, diese Rechenoperationen klarer zu gliedern.

Interessant zur praktischen Berechnung von Determinanten, vorzugsweise für nicht allzu große n, ist die Formel daher nur, wenn man sie mit den Rechenregeln aus dem vorigen Abschnitt kombiniert, um sich Nullen und einfache Einträge zu verschaffen. Da dieselbe Art von Operationen auch zur LR-Zerlegung führen, wird dann der Übergang zwischen LAPLACEschem Entwicklungssatz und Berechnung via LR-Zerlegung fließend.

Betrachten wir als Beispiel die Determinante der Matrix

$$A = \begin{pmatrix} 1 & 2 & -2 & 1 \\ -1 & -3 & 2 & -2 \\ -4 & -2 & -1 & 0 \\ 3 & 2 & -1 & 1 \end{pmatrix}.$$

Sie hat zwar einen Eintrag Null, aber da in der dritten und vierten Spalte gleich in zwei Zeilen die Einträge entgegengesetzt gleich sind, ist es besser, zunächst die dritte Spalte durch die Summe von dritter und vierter Spalte zu ersetzen:

$$\begin{vmatrix} 1 & 2 & -2 & 1 \\ -1 & -3 & 2 & -2 \\ -4 & -2 & -1 & 0 \\ 3 & 2 & -1 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 2 & -1 & 1 \\ -1 & -3 & 0 & -2 \\ -4 & -2 & -1 & 0 \\ 3 & 2 & 0 & 1 \end{vmatrix}.$$

Wenn wir jetzt nach der dritten Spalte entwickeln, kommen nur zwei der Summanden wirklich vor: Der erste ($i = 1$ und $j = 3$) und der dritte ($i = j = 3$). In beiden Fällen ist $i + j$ gerade, der Vorfaktor $(-1)^{i+j}$ ist also jeweils +1. Die 3×3-Determinanten, die durch Streichung der dritten Spalte und der ersten *bzw.* dritten Zeile entstehen, können nach der Regel von SARRUS berechnet werden:

$$\begin{vmatrix} -1 & -3 & -2 \\ -4 & -2 & 0 \\ 3 & 2 & 1 \end{vmatrix} = (2 + 0 + 16) - (12 + 0 + 12) = -6$$

und

$$\begin{vmatrix} 1 & 2 & 1 \\ -1 & -3 & -2 \\ 3 & 2 & 1 \end{vmatrix} = (-3 - 12 - 2) - (-9 - 4 - 2) = -2\,.$$

Also ist $\det A = (-1)(-6) + (-1)(-2) = 8$.

Auch als Hilfmittel zur Berechnung von abstrakt gegebenen Determinanten ist der LAPLACEsche Entwicklungssatz oft nützlich, da er es bei geschickter Anwendung gelegentlich erlaubt, eine Rekursionsformel zu finden und damit eine allgemeine Formel für einen gewisse Klasse von Determinanten.

Als Beispiel betrachten wir die unter anderem in der Numerik wichtige VANDERMONDEsche Determinante

$$V(a_1, \ldots, a_n) = \begin{vmatrix} 1 & a_1 & a_1^2 & \ldots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \ldots & a_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & a_n^2 & \ldots & a_n^{n-1} \end{vmatrix}.$$

Der Franzose ALEXANDRE THÉOPHILE VANDERMONDE (1735–1796) war zunächst Musiker; erst im Alter von 35 Jahren begann er sich für Mathematik zu interessieren und publizierte in den Jahren 1771 und 1772 vier Arbeiten über Gleichungen, Determinanten und über das Problem, einen Springer so über ein Schachbrett zu bewegen, daß er jedes Feld genau einmal betritt. Die VANDERMONDEsche Determinante ist nirgends in seinem publizierten Werk zu finden; sie wurde erst um 1935 von HENRI LEBESGUE (1875–1941) nach ihm benannt.

Für die Anwendung des LAPLACEschen Entwicklungssatzes auf diese Determinante bietet sich an, nach der ersten Spalte zu entwickeln, denn diese besteht aus lauter Einsen, und die $(n-1)\times(n-1)$-Determinanten, die nach dem Entwicklungssatz auftreten, sind im wesentlichen wieder VANDERMONDEsche Determinanten. Allerdings entstehen dabei n Summanden, und für jeden von diesen muß die ganz Prozedur wiederholt werden *usw.*– die Berechnung der Determinante auf diese Weise ist also zumindest ziemlich aufwendig.

Ein besserer Ansatz ergibt sich, wenn wir die Einsen in der ersten Spalte dadurch ausnutzen, daß wir etwa die erste Zeile von jeder anderen Zeile subtrahieren. Der Wert der Determinante ändert sich dadurch nicht, d.h.

$$V(a_1,\dots,a_n) = \begin{vmatrix} 1 & a_1 & a_1^2 & \dots & a_1^{n-1} \\ 0 & a_2 - a_1 & a_2^2 - a_1^2 & \dots & a_2^{n-1} - a_1^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & a_n - a_1 & a_n^2 - a_1^2 & \dots & a_n^{n-1} - a_1^{n-1} \end{vmatrix} .$$

Wenn wir hier nach der ersten Spalte entwickeln, muß nur eine einzige $(n-1) \times (n-1)$-Determinante berücksichtigt werden, alle anderen haben den Vorfaktor null. Also ist

$$V(a_1,\dots,a_n) = \begin{vmatrix} a_2 - a_1 & a_2^2 - a_1^2 & \dots & a_2^{n-1} - a_1^{n-1} \\ a_3 - a_1 & a_3^2 - a_1^2 & \dots & a_3^{n-1} - a_1^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_n - a_1 & a_n^2 - a_1^2 & \dots & a_n^{n-1} - a_1^{n-1} \end{vmatrix}$$

gleich jeder Determinanten, die durch Streichung der ersten Spalte und der ersten Zeile entsteht.

Hier können wir in jeder Zeile die jeweils vorne stehende Differenz ausklammern, denn genau wie

$$x^k - 1 = (x-1)(x^{k-1} + x^{k-2} + \dots + x + 1)$$

durch $(x-1)$ teilbar ist, ist auch

$$a_i^k - a_1^k = (a_i - a_1)(a_i^{k-1} + a_i^{k-2}a_1 + a_i^{k-3}a_1^2 + \dots + a_i a_1^{k-2} + a_1^{k-1})$$

durch $(a_i - a_1)$ teilbar; den Quotienten schreiben wir kurz als $q_{i,k-1}$:

$$q_{i,k-1} \underset{\text{def}}{=} a_i^{k-1} + a_i^{k-2}a_1 + a_i^{k-3}a_1^2 + \dots + a_i a_1^{k-2} + a_1^{k-1} .$$

Wegen der Linearität der Determinante können wir jeden Faktor, den wir aus einer Zeile (oder Spalte) ausklammern, vor die Determinante ziehen und erhalten für $V(a_1,\dots,a_n)$ somit den Wert

$$(a_2 - a_1)(a_3 - a_1)\cdots(a_n - a_1) \begin{vmatrix} 1 & q_{21} & q_{22} & \dots & q_{2,n-2} \\ 1 & q_{31} & q_{32} & \dots & q_{3,n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & q_{n1} & q_{n2} & \dots & q_{n,n-2} \end{vmatrix} .$$

Die Nützlichkeit dieser Formel steht und fällt damit, daß wir die q_{ij} gut miteinander in Verbindung bringen können. Für verschiedene Indizes i

haben die entsprechenden Ausdrücke offensichtlich wenig miteinander zu tun; sie enthalten nicht einmal dieselben Variablen. Schreiben wir allerdings

$$\begin{aligned} q_{ij} &= a_i^j + a_i^{j-1}a_1 + a_i^{j-2}a_1^2 + \cdots a_i a_1^{j-1} + a_1^j \\ &= a_i^j + a_1(a_i^{j-1} + a_i^{j-2}a_1 + \cdots a_i a_1^{j-2} + a_1^{j-1}) , \end{aligned}$$

so sehen wir, daß

$$q_{ij} = a_i^j + a_1 q_{i,j-1} \quad \text{oder} \quad q_{ij} - a_1 q_{i,j-1} = a_i^j$$

ist. Subtrahieren wir also zuerst a_1 mal die vorletzte Spalte von der letzten, so werden die Einträge der letzten Spalte zu a_i^{n-2}. Entsprechend subtrahieren wir a_1 mal die $(n-2)$-te Zeile von der $(n-1)$-ten und erhalten lauter Einträge a_i^{n-3} und so weiter, bis schließlich die Subtraktion des a_1-fachen der ersten Spalte von der zweiten die Einträge der letzteren zu

$$q_{i1} - a_1 = (a_i + a_1) - a_1 = a_i$$

macht. Somit ist $V(a_1,\dots.a_n)$ gleich

$$(a_2 - a_1)(a_3 - a_1)\cdots(a_n - a_1) \begin{vmatrix} 1 & a_2 & a_2^2 & \dots & a_2^{n-2} \\ 1 & a_3 & a_3^2 & \dots & a_3^{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & a_n^2 & \dots & a_n^{n-2} \end{vmatrix} .$$

Die Determinante rechts ist offensichtlich wieder eine VANDERMONDEsche Determinante, allerdings mit um eins verminderter Zeilen- und Spaltenzahl und mit einer Variablen weniger.

Damit haben wir die Rekursionsformel

$$V(a_1,\dots,a_n) = (a_2 - a_1)(a_3 - a_1)\cdots(a_n - a_1)\, V(a_2,\dots,a_n) ,$$

die es erlaubt, die Berechnung von $V(a_1,\dots,a_n)$ auf eine einzige VANDERMONDEsche Determinante der Größe $(n-1) \times (n-1)$ zurückzuführen.

Zur vollständigen Berechnung von $V(a_1,\dots,a_n)$ fehlt uns jetzt nur noch ein Induktionsanfang; direktes Nachrechnen zeigt sofort, daß

$$V(a_n) = \det(1) \quad \text{und} \quad V(a_{n-1}, a_n) = \begin{vmatrix} 1 & a_{n-1} \\ 1 & a_n \end{vmatrix} = a_n - a_{n-1}$$

ist, also folgt induktiv

$$V(a_1,\dots,a_n) = \prod_{j<i}(a_i - a_j)\,.$$

h) Die Cramersche Regel

Determinanten können auch angewandt werden, um die Lösungen eines linearen Gleichungssystems vom Rang n aus n Gleichungen in n Unbekannten in geschlossener Form als Funktion der Koeffizienten darzustellen. Dazu schreiben wir das Gleichungssystem

$$A\vec{x} = \vec{b}$$

in der Form

$$x_1\vec{a}_1 + \cdots + x_n\vec{a}_n = \vec{b}\,,$$

wobei die $\vec{a}_i$ die Spaltenvektoren der Matrix A seien und $(x_1,\dots,x_n)$ eine Lösung des linearen Gleichungssystems.

Nun ersetzen wir in

$$\det A = \det\bigl(\vec{a}_1,\dots,\vec{a}_n\bigr)$$

rechts den Vektor $\vec{a}_i$ durch die rechte Seite $\vec{b}$ des Gleichungssystems. (Es gibt eigentlich keinen vernünftigen Grund, warum wir das tun sollten; auch dieser Trick wird, wie so viele, erst nachträglich durch das Ergebnis gerechtfertigt.) Die so entstehende Determinante ist

$$\begin{aligned}
&\det\bigl(\vec{a}_1,\dots,\vec{b},\dots,\vec{a}_n\bigr)\\
&= \det\Bigl(\vec{a}_1,\dots,\sum_{j=1}^{n} x_j\vec{a}_j,\dots,\vec{a}_n\Bigr)\\
&= \sum_{j=1}^{n} x_j \det\bigl(\vec{a}_1,\dots,\vec{a}_j,\dots,\vec{a}_n\bigr)\\
&= x_i \det\bigl(\vec{a}_1,\dots,\vec{a}_i,\dots,\vec{a}_n\bigr) = x_i \det A\,,
\end{aligned}$$

da für jeden Index $j \neq i$ der Vektor $\vec{a}_j$ zweimal als Argument der Determinante auftritt, so daß alle Summanden bis auf den i-ten verschwinden.

Falls $\det A = 0$ ist, nützt uns diese Formel überhaupt nichts; ist allerdings $\det A \neq 0$, wissen wir bereits, daß das Gleichungssystem eindeutig lösbar ist, und wir können die Komponenten dieser eindeutig bestimmten Lösung explizit ausdrücken durch die

CRAMERsche Regel:

$$x_i = \frac{\det\bigl(\vec{a}_1,\dots,\vec{b},\dots,\vec{a}_n\bigr)}{\det A}$$

Der Schweizer Mathematiker GABRIEL CRAMER (1704–1752) lehrte an der Universität Genf. Bekannt wurde er vor allem durch seine Arbeiten über Determinanten, er beschäftigte sich aber auch viel mit Analysis und Geometrie, insbesondere ist er Autor eines Buchs über algebraische Kurven. Weitere Arbeitsgebiete sind mathematische Methoden der Physik sowie die Geschichte der Mathematik. Die CRAMERsche Regel im Spezialfall $n = 2$ ist bereits 1545 in der *Ars magna* des italienischen Mathematikers GIROLAMO CARDANO (1501–1576) zu finden.

Die CRAMERsche Regel ist sicherlich kein Verfahren, das man oft anwendet zur Lösung eines einzelnen Gleichungssystems: Der Aufwand für die Berechnung von $n + 1$ Determinanten ist weit größer als eine Lösung nach dem GAUSS-Algorithmus. Falls man allerdings eine ganze Familie ähnlicher Gleichungssysteme hat, in der sich nur wenige Parameter ändern, kann es sich lohnen, mit Hilfe der CRAMERschen Regel eine (von den Parametern abhängige) Lösungsformel zu berechnen und dann diese anzuwenden.

i) Geschichte und Anwendungen von Determinanten

Determinanten erblickten das Licht der Welt im Jahr 1683, und zwar gleich zweimal: Der japanische Mathematiker SEKI benutzte sie, ohne Ihnen einen Namen zu geben, zur Lösung von Gleichungen höheren Grades und zeigte anhand von Beispielen, wie man sie für 2×2- bis 5×5-Matrizen berechnet. Im gleichen Jahr schrieb auch LEIBNIZ einen Brief an DE L'HÔPITAL, in dem er erwähnte, daß ein gewisses homogenes lineares Gleichungssystem in drei Variablen nichttrivial lösbar sei, da (in heutiger Terminologie) seine Determinante verschwinde.

TAKAKAZU SEKI KOWA (1642–1708) war Sohn eines Samurai, wurde aber schon sehr jung von einem Adeligen namens SEKI GOROZAYEMON adoptiert. Einer von dessen Dienern weckte das Interesse des neunjährigen SEKI an der Mathematik, woraufhin dieser eine große Bibliothek japanischer und chinesischer Mathematikbücher anschaffte, anhand derer er sich selbst in das Gebiet einarbeitete. Als Staatsbeamter und ab 1704 Zeremonienmeister des Shogun befaßte er sich weiterhin viel mit Mathematik und entdeckte außer Determinanten beispielsweise auch das NEWTON-Verfahren und (vor JAKOB BERNOULLI) die BERNOULLI-Zahlen.

LEIBNIZ sprach noch nicht von Determinanten, sondern von *Resultanten;* ein Begriff, den unabhängig davon 1772 auch LAPLACE benutzte, als er damit die Bahnen der inneren Planeten berechnete. Das *Wort* Determinante erschien erstmal 1801 in den *Disquisitiones arithmeticae* von GAUSS, der damit die Eigenschaften quadratischer Formen untersuchte. Auch CAUCHY, der 1812 den Multiplikationssatz bewies, sprach von Determinanten.

Heute bezeichnet man als *Resultanten* spezielle Determinaten, die zur Lösung nichtlinearer Gleichungssysteme benutzt werden und auf den englischen Mathematiker JAMES JOSEPH SYLVESTER (1814–1897) zurückgehen. Mit Hilfe solcher Resultanten gelang es beispielsweise 1985 zwei Mathematikern bei *General Motors* das inverse kinematische Problem für einen Roboterarm mit sechs Freiheitsgraden zu lösen, d.h. also ein Verfahren zu entwickeln, mit dem der Manipulator am Ende des Arm *automatisch* auf eine vorgegebene Position und Ausrichtung gebracht werden kann.

Weitere wichtige Anwendungen haben Determinanten heute auch in der Numerik, wo sie wohl *Konditionszahlen* definieren, die etwas über die Stabilität und Robustheit eines Verfahrens aussagen, als auch beispielsweise (wie die VANDERMONDEsche Determinante) bei der Approximation von Funktionen oder Datenpunkten durch Polynomfunktionen.

Ein weiterer Teil der vielfältigen Anwendungen von Determinanten

hängt damit zusammen, daß sie – wie wir schon am Beispiel der mit dem Spatprodukt identischen 3×3-Determinanten gesehen haben – als Volumina aufgefaßt werden können; dies wird uns auch im nächsten Kapitel wieder begegnen, wenn wir das Mehrfachintegrale von einem Koordinatensystem in ein anderes transformieren.

§6: Euklidische und Hermitesche Vektorräume

Zu Beginn des letzten Paragraphen hatten wir Vektor- und Skalarprodukte im $\mathbb{R}^3$ betrachtet. Zumindest das Vektorprodukt spielt in der Tat sonst nirgends eine Rolle, Skalarprodukte können aber selbst in unendlichdimensionalen reellen oder komplexen Vektorräumen sehr nützlich sein, wo sie beispielsweise die Energie oder Leistung eines Signals berechnen; daneben spielen sie eine wichtige Rolle in der Fehler- und Ausgleichsrechnung sowie in der Statistik.

Im letzten Paragraphen haben wir gesehen, daß das Skalarprodukt im $\mathbb{R}^3$ eng mit Längen und Winkeln zusammenhängt; da dies zwei der Grundbegriffe der EUKLIDischen Geometrie sind, werden wir reelle Vektorräume mit Skalarprodukt allgemein als EUKLIDische Vektorräume bezeichnen.

a) Euklidische Vektorräume

Wir beginnen mit einem etwas schwächeren Begriff als dem des Skalarprodukts, der sich im Gegensatz zu letzterem noch 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

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

eine Bilinearform definiert.

Noch allgemeiner erklärt diese Formel auch für die Vektorräume k^n über einem beliebigen Körper k eine Bilinearform, die auch für (nach unserem derzeitigen Kenntnisstand) 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

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

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 $\mathbb{R}^n$ mit seinem Standardskalarprodukt

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

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:

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

da der Integrand nirgends negativ wird. Falls f nicht identisch verschwindet, gibt es einen Punkt $t_0\in(0,\,1)$ mit $f(t_0)=h\neq0$. 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 $\mathbb{R}^3$ kann man durch Auszeichnung von Rechtssystemen über die Dreifingerregel oder Korkenzieherregel auch orientierte Winkel zwischen null und 2π definieren, aber 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, ob 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 im vorigen Paragraphen *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.

Da zwei Vektoren stets eine Ebene aufspannen, sollten wir erwarten, daß Entsprechendes auch für beliebige reelle Vektorräume gilt, was auch in der Tat der Fall ist. Mit dem Beweis wollen wir allerdings noch warten bis zum übernächsten Abschnitt, wo wir ohne nennenswerten zusätzlichen Aufwand gleich auch noch eine ähnliche Formel für komplexe Vektorräume beweisen können.

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) Hermitesche Vektorräume

Die unmittelbare Verallgemeinerung des Skalarprodukts des $\mathbb{R}^n$ auf $\mathbb{C}^n$ ist nicht sonderlich nützlich: Dann hätten wir nämlich zum Beispiel

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

Solche Vektoren mit Quadrat Null sind zwar in einigen Anwendungen (wie etwa der speziellen Relativitätstheorie) durchaus sinnvoll und nützlich, meist möchte man aber das Skalarprodukt eines Vektors mit sich selbst als Quadrat seiner Länge interpretieren, und die sollte für alle Vektoren außer 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 Forderungen aus der Definition eines Skalarprodukts: Es ist zwar noch linear im ersten Argument, nicht mehr aber im zweiten, denn schon bei Multiplikation des zweiten Vektors mit einer komplexen Zahl $\lambda \notin \mathbb{R}$ 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; somit wird die Linearitätsregel für das zweite Argument zu

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

Die Symmetriebedingung ist ebenfalls 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 diese Eigenschaften:

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

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

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.

$$\vec v\cdot\vec v>0$$

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 $(\vec v,\,\vec w)$ anstelle von $\vec v\cdot\vec w$ schreiben.

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 transzendente Zahl ist, also keiner algebraischen Gleichung über $\mathbb{Q}$ genügt, eine Interpolationsformel und vieles mehr. HERMITE galt als ein sehr guter akademischer Lehrer; er unterrichtete an der École Polytechnique, dem Collège de France, der École Normale Supérieure und der Sorbonne.

Durch Kombination der Eigenschaften *a)* und *b)* kann man leicht ausrechnen, was anstelle der Linearität für das zweite Argument gilt: Genau wie im obigen Beispiel ist

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

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 der $\mathbb{C}^n$ mit dem eingangs definierten Produkt, jedoch wird sich im nächsten Semester zeigen, daß gerade HERMITEsche Vektorräume von komplexwertigen Funktionen sehr interessante Anwendungen haben.

Hier sei nur als Beispiel für das Rechnen mit HERMITEschen Skalarprodukten in $\mathbb{C}^n$ gezeigt, mit dem wir auch nochmals das Rechnen mit Matrizen wiederholen können, und zwar geht es um ein Verfahren von G. PHILIPPE, veröffentlicht in *Quadrature,* Oct.-Déc. 2000, S. 23–34, wie man aus komplexen Vektoren reelle quadratische Matrizen A,B definieren kann, für die $AB=-BA$ ist, die also *antikommutieren.*

Dazu betrachten wir einen Vektor $\vec v\in\mathbb{C}^n$, wobei $\mathbb{C}^n$ sein Standard-HERMITEsches Produkt habe. Zu den Komponenten v_i von $\vec v$ betrachten wir die $n\times$-Matrix W mit Einträgen $w_{ij}=v_i\overline{v_j}$ und die Matrix

$$M=a\cdot E-2W\quad\text{mit}\quad a=\vec v\cdot\vec v\,.$$

Da aE als Diagonalmatrix mit W kommutiert, ist
$$M^2 = a^2E - 4aW + 4W^2 .$$
Der Eintrag an der Stelle ik von W^2 ist
$$\sum_{j=1}^{n} w_{ij}w_{jk} = \sum_{j=1}^{n} v_i\overline{v_j} \cdot v_j\overline{v_k} = v_i\overline{v}_k \sum_{j=1}^{n} v_j\overline{v_j} = w_{ik}(\vec{v}\cdot\vec{v}) = aw_{ik} ,$$
d.h. $M^2 = a^2E$.

Bezeichen A und B die reellen $n \times n$-Matrizen aus den Real- und Imaginärteilen von M, ist also $M = A + iB$, so ist demnach
$$(A + iB)^2 = A^2 - B^2 + i(AB + BA) = a^2E$$
eine reelle Matrix, der Imaginärteil $AB + BA$ muß also verschwinden. Das ist aber gleichbedeutend damit, daß $AB = -BA$ ist.

Als Beispiel betrachten wir den Vektor $\binom{1}{i}$ aus $\mathbb{C}^2$. Hier ist $a = 2$ und
$$M = \begin{pmatrix} 1 & -i \\ i & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + i\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} ,$$
und in der Tat ist
$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = -E \quad \text{und} \quad \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = E .$$
Kompliziertere Vektoren $\vec{v}$ führen zu interessanteren Beispielen.

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.

c) Die Cauchy-Schwarzsche Ungleichung

Als erstes Beispiel wollen wir das noch offene Problem aus Abschnitt *a)* lösen und zeigen, daß die Zahl, die wir dort als Cosinus eines Winkels definiert hatten, tatsächlich zwischen null und eins liegt. Natürlich zeigen wir dies gleich etwas allgemeiner so, daß wir auch eine Aussage für HERMITEsche Vektorräume bekommen, und wir verallgemeinern auch gleich noch etwas weiter in Hinblick auf die Tatsache, daß wir in Vektorräumen, die auch unstetige Funktionen enthalten, im allgemeinen kein wirkliches Skalar- *bzw.* HERMITEsches Produkt haben. Trotzdem werden Produkte in solchen Vektorräumen im nächsten Semester für die FOURIER-Theorie sehr nützlich sein.

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) $\|\vec{v}\|^2 = \vec{v}\cdot\vec{v} \geq 0$ für alle $\vec{v} \in V$.
Dann ist für alle Vektoren $\vec{v}, \vec{w} \in V$
$$|\vec{v}\cdot\vec{w}| \leq \|\vec{v}\|\cdot\|\vec{w}\| .$$

(Man beachte, daß hier nicht gefordert wird, daß $\vec{v}\cdot\vec{v} > 0$ für $\vec{v} \neq \vec{0}$.)

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 nichtnegativ. Speziell für $\lambda = (\vec{w}\cdot\vec{w}) \in \mathbb{R}$ und $\mu = -(\vec{v}\cdot\vec{w})$ erhalten wir
$$(\vec{w}\cdot\vec{w})^2(\vec{v}\cdot\vec{v}) - 2(\vec{w}\cdot\vec{w})(\overline{\vec{v}\cdot\vec{w}})(\vec{v}\cdot\vec{w}) + (\vec{v}\cdot\vec{w})(\overline{\vec{v}\cdot\vec{w}})(\vec{w}\cdot\vec{w}) \geq 0 ,$$
also
$$(\vec{w}\cdot\vec{w})\big((\vec{w}\cdot\vec{w})(\vec{v}\cdot\vec{v}) - |\vec{v}\cdot\vec{w}|^2\big) \geq 0 .$$
Ist hier $\vec{w}\cdot\vec{w} \neq 0$, 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.

d) Orthonormalbasen

Genau wie im Falle des $\mathbb{R}^3$ wollen wir auch für beliebige EUKLIDische oder HERMITEsche Vektorräume zwei Vektoren als *orthogonal* bezeichnen, wenn ihr (HERMITEsches) Skalarprodukt verschwindet. Viele Rechnungen vereinfachen sich, wenn man von einer Basis ausgeht, deren Vektoren paarweise orthogonal sind und möglicherweise zusätzlich noch die Länge eins habe; um diese Art von Basen soll es hier gehen:

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 $\mathbb{R}^n$ als auch für $\mathbb{C}^n$ bezüglich des jeweiligen Standardskalar- oder HERMITEschen Produkts eine Orthonormalbasis bilden.

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

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

Beweis: Zunächst wissen wir, daß V überhaupt eine Basis $\{\vec{b}_1,\dots,\vec{b}_n\}$ hat; daraus konstruieren wir schrittweise eine *Orthogonalbasis* nach dem sogenannten GRAM-SCHMIDTschen Orthogonalisierungsverfahren. Dieses liefert in seinem r-ten Schritt eine Orthogonalbasis $\{\vec{c}_1,\dots,\vec{c}_r\}$ des von dem ersten r Basisvektoren aufgespannten Untervektorraums $[\vec{b}_1,\dots,\vec{b}_r]$.

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 $r\ge 1$ 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 $i\ne j$ aus dem von

$\vec{b}_1, \ldots, \vec{b}_r$ 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 $\vec{c}_1, \ldots, \vec{c}_r$ 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 um 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 $i = 1, \ldots, r$.

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

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

■

Um auch den Umgang mit HERMITEschen Skalarprodukten und das Rechnen mit komplexen Zahlen zu üben, betrachten wir als Beispiel den von

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

aufgespannten Untervektorraum von $\mathbb{C}^4$.

Wie oben im Beweis können wir bei der Anwendung des GRAM-SCHMIDTschen Orthogonalisierungsverfahrens $\vec{c}_1 = \vec{b}_1$ setzen; im zweiten Schritt suchen wir einen Vektor $\vec{c}_2 = \vec{b}_2 + \lambda_1 \vec{c}_1$ mit der Eigenschaft, daß

$$\vec{c}_2 \cdot \vec{c}_1 = (\vec{b}_2 + \lambda_1 \vec{c}_1) \cdot \vec{c}_1 = \vec{b}_2 \cdot \vec{c}_1 + \lambda_1 (\vec{c}_1 \cdot \vec{c}_1) = 0$$

ist.

(Wir können natürlich auch fordern, daß

$$\vec{c}_1 \cdot \vec{c}_2 = \vec{c}_1 \cdot (\vec{b}_2 + \lambda_1 \vec{c}_1) = \vec{c}_1 \cdot \vec{b}_2 + \overline{\lambda}_1 (\vec{c}_1 \cdot \vec{c}_1)$$

verschwindet, aber dann erhalten wir zunächst eine Formel für $\overline{\lambda}_1$ und müssen noch komplex konjugieren, Daher ist zumindest im komlexen Fall die Forderung $\vec{c}_2 \cdot \vec{c}_1 = 0$ rechnerisch etwas bequemer.)

Wir brauchen also die HERMITEschen Skalarprodukte

$$\vec{b}_2 \cdot \vec{c}_1 = \begin{pmatrix} 1+2i \\ -2+i \\ -i \\ -1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ i \\ -i \\ -1 \end{pmatrix} \quad \text{und} \quad \vec{c}_1 \cdot \vec{c}_1 = \begin{pmatrix} 1 \\ i \\ -i \\ -1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ i \\ -i \\ -1 \end{pmatrix} .$$

Bei deren Berechnung darf man auf keinen Fall vergessen, die Einträge des jeweils zweiten Vektors komplex zu konjugieren, d.h.

$$\begin{aligned} \vec{b}_2 \cdot \vec{c}_1 &= (1+2i) \cdot \overline{1} + (-2+i) \cdot \overline{i} + (-i) \cdot (\overline{-i}) + (-1) \cdot (\overline{-1}) \\ &= (1+2i) \cdot 1 + (-2+i) \cdot (-i) + (-i) \cdot i + (-1) \cdot (-1) . \\ &= (1+2i) + (1+2i) + 1 + 1 = 4 + 4i \end{aligned}$$

Entsprechend ist

$$\begin{aligned} \vec{c}_1 \cdot \vec{c}_1 &= 1 \cdot \overline{1} + i \cdot \overline{i} + (-i) \cdot (\overline{-i}) + (-1) \cdot (\overline{-1}) \\ &= 1 \cdot 1 + i \cdot (-i) + (-i) \cdot i + (-1) \cdot (-1) = 1 + 1 + 1 + 1 = 4 , \end{aligned}$$

wobei das allerdings auch mit weniger Mühe einzusehen ist: Da für einen Vektor $\vec{v} \in \mathbb{C}^n$ mit Komponenten $v_1, \ldots, v_n$

$$\vec{v} \cdot \vec{v} = \sum_{i=1}^{n} v_i \cdot \overline{v_i} = \sum_{i=1}^{n} |v_i|^2$$

ist, hätte es genügt, einfach die Betragsquadrate der vier Einträge aufzusummieren, wobei im vorliegenden Fall trivial ist, daß diese allesamt eins sind.

Einsetzen der beiden Skalarprodukte in die Gleichung für λ_1 zeigt, daß

$$(4+4i)+4\lambda_1=0 \quad \text{und damit} \quad \lambda_1=\frac{-4-4i}{4}=-1-i$$

sein muß. Somit ist

$$\vec{c}_2=\vec{b}_2+\lambda_1\vec{c}_1=\begin{pmatrix}1+2i\\-2+i\\-i\\-1\end{pmatrix}-\begin{pmatrix}1+i\\-1+i\\1-i\\-1-i\end{pmatrix}=\begin{pmatrix}i\\-1\\-1\\i\end{pmatrix}.$$

Im dritten Schritt muß ein Vektor

$$\vec{c}_3=\vec{b}_3+\mu_1\vec{c}_1+\mu_2\vec{c}_2$$

berechnet werden mit

$$\vec{c}_3\cdot\vec{c}_1=\vec{c}_3\cdot\vec{c}_2=0\,.$$

Hier ist

$$\vec{c}_3\cdot\vec{c}_1=\vec{b}_3\cdot\vec{c}_1+\mu_1\vec{c}_1\cdot\vec{c}_1+\mu_2\vec{c}_2\cdot\vec{c}_1=\vec{b}_3\cdot\vec{c}_1+\mu_1\vec{c}_1\cdot\vec{c}_1\,,$$

da der zweite Schritt sicherstellte, daß $\vec{c}_2\cdot\vec{c}_1=0$ ist. Das HERMITEsche Skalarprodukt $\vec{c}_1\cdot\vec{c}_1=4$ kennen wir bereits, und eine kurze Rechnung zeigt, daß

$$\vec{b}_3\cdot\vec{c}_1=(1+2i)+(-1)+(1-4i)+(-1-2i)=-4i$$

ist. Also folgt

$$\mu_1=-\frac{\vec{b}_3\cdot\vec{c}_1}{\vec{c}_1\cdot\vec{c}_1}=-\frac{-4i}{4}=i$$

und entsprechend folgt auch

$$\mu_2=-\frac{\vec{b}_3\cdot\vec{c}_2}{\vec{c}_2\cdot\vec{c}_2}=-\frac{8}{4}=-2\,,$$

wobei diesesmal die genaue Berechnung der beiden HERMITEschen Skalarprodukte dem Leser als Übungsaufgabe überlassen sei. Der dritte Basisvektor $\vec{c}_3=\vec{b}_3+i\vec{c}_1-2\vec{c}_2$ ist daher gleich

$$\begin{pmatrix}1+2i\\-i\\-4-i\\1+2i\end{pmatrix}+i\begin{pmatrix}1\\i\\-i\\-1\end{pmatrix}-2\begin{pmatrix}i\\-1\\-1\\i\end{pmatrix}=\begin{pmatrix}1+i\\1-i\\-1-i\\1-i\end{pmatrix}.$$

Damit ist eine *Orthogonalbasis* gefunden; wenn wir eine *Orthonormalbasis* wollen, müssen wir die Vektoren noch durch ihre Längen dividieren: Wir wissen bereits, daß $\vec{c}_1\cdot\vec{c}_1=\vec{c}_2\cdot\vec{c}_2=4$ ist, so daß die ersten beiden Vektoren die Länge zwei haben; $\vec{c}_3$ hat vier Komponenten vom Betrag zwei, also ist

$$\vec{c}_3\cdot\vec{c}_3=4\cdot 2=8 \quad \text{und} \quad \|\vec{c}_2\|=\sqrt{8}=2\sqrt{2}\,.$$

Die gesuchte Orthonormalbasis besteht daher aus den Vektoren

$$\frac{1}{2}\begin{pmatrix}1\\i\\-i\\1\end{pmatrix},\quad \frac{1}{2}\begin{pmatrix}i\\-1\\-1\\i\end{pmatrix} \quad \text{und} \quad \frac{\sqrt{2}}{4}\begin{pmatrix}1+i\\1-i\\-1-i\\1-i\end{pmatrix}.$$

Die Nützlichkeit von Orthogonalbasen ergibt sich vor allem aus dem folgenden

Satz: $\{\vec{e}_1,\dots,\vec{e}_n\}$ sei eine Orthonormalbasis eines EUKLIDischen oder HERMITEschen Vektorraums V.
a) Sind $\vec{v}=a_1\vec{e}_1+\cdots+a_n\vec{e}_n$ und $\vec{w}=b_1\vec{e}_1+\cdots+b_n\vec{e}_n$ zwei Vektoren aus V, so ist

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

b) Für einen Vektor $\vec{v}\in V$ ist

$$\vec{v}=a_1\vec{e}_1+\cdots+a_n\vec{e}_n \quad \text{mit} \quad a_i=\vec{v}\cdot\vec{e}_i\,.$$

Der *Beweis* ist in beiden Fällen ein Einzeiler:

$$\vec{v}\cdot\vec{w}=\sum_{i=1}^n\sum_{j=1}^n a_i\overline{b_j}\;\vec{e}_i\cdot\vec{e}_j=\sum_{i=1}^n a_i\overline{b_i}\,, \quad \text{denn} \quad \vec{e}_i\cdot\vec{e}_j=\delta_{ij}$$

$$\vec{v} \cdot \vec{e}_i = (a_1\vec{e}_1 + \cdots + a_n\vec{e}_n) \cdot \vec{e}_i = \sum_{j=1}^{n} a_j \vec{e}_j \cdot \vec{e}_i = a_i \,.$$

Bezüglich einer Orthonormalbasis sieht also jedes (gewöhnliche oder HERMITEsche) Skalarprodukt aus wie das entsprechende Standardskalarprodukt – und das unabhängig von der gewählten Orthonormalbasis. Sobald man Vektoren bezüglich einer Orthonormalbasis dargestellt hat, lassen sich also alle Skalarprodukte sowie auch die Basisdarstellung eines beliebigen Vektors auf die einfachst denkbare Weise darstellen.

d) Normierte Vektorräume

Mit Hilfe von klassischen wie auch HERMITEschen Skalarprodukten konnten wir die Länge oder Norm

$$\|\vec{v}\| = \sqrt{\vec{v} \cdot \vec{v}}$$

eines Vektors definieren und damit beispielsweise auch Abstände zwischen Punkten eines affinen Raums –sofern der zugehörige Vektorraum EUKLIDisch oder HERMITEsch ist. Manchmal kommt es nur auf diese Längen an, nicht auf die Produkte; daher wollen wir diese hier für sich betrachten:

Sie haben folgende Eigenschaften:

a) $\|\lambda\vec{v}\| = |\lambda|\,\|\vec{v}\|$ für alle $\lambda \in \mathbb{R}$ und $\vec{v} \in V$

b) $\|\vec{v}\| \geq 0$ für alle $\vec{v} \in V$, und $\|\vec{v}\| = 0$ genau dann, wenn $\vec{v} = \vec{0}$

c) $\|\vec{v} + \vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|$.

Abgesehen von *c)*, der *Dreiecksungleichung*, sind diese Eigenschaften klar; wegen

$$\|\vec{v} + \vec{w}\|^2 = \|\vec{v}\|^2 + \|\vec{w}\|^2 + 2\vec{v}\vec{w} \qquad \text{und}$$

$$(\|\vec{v}\| + \|\vec{w}\|)^2 = \|\vec{v}\|^2 + \|\vec{w}\|^2 + 2\,\|\vec{v}\|\,\|\vec{w}\|$$

folgt letztere im Falle $\vec{v} \cdot \vec{w} \geq 0$ sofort aus der CAUCHY-SCHWARZschen Ungleichung und für $\vec{v} \cdot \vec{w} < 0$ aus der Nichtnegativität der Norm.

Definition: Ein normierter Vektorraum ist ein $\mathbb{R}$- oder $\mathbb{C}$-Vektorraum V zusammen mit einer Abbildung $\|\cdot\| : V \to \mathbb{R}$ mit den Eigenschaften *a)* bis *c)*.

Damit ist also jeder EUKLIDische oder HERMITEsche Vektorraum insbesondere auch ein normierter Vektorraum; man kann Normen aber auch anders definieren als über Skalarprodukte:

Beispielsweise erfüllt die *Maximumsnorm*

$$\|\cdot\|_\infty : \mathbb{R}^n \to \mathbb{R}; \qquad \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \mapsto \max_{i=1}^{n} |v_i|$$

offensichtlich die Bedingungen *a)* bis *c)*. Sie kommt aber nicht von einem Skalarprodukt, denn gäbe es ein Skalarprodukt $\star$ mit $\|\vec{v}\|_\infty = \sqrt{\vec{v} \star \vec{v}}$; so wäre etwa im $\mathbb{R}^2$

$$\begin{pmatrix}1\\0\end{pmatrix} \star \begin{pmatrix}1\\0\end{pmatrix} = \left\|\begin{pmatrix}1\\0\end{pmatrix}\right\|_\infty^2 = 1 \quad \text{und} \quad \begin{pmatrix}0\\1\end{pmatrix} \star \begin{pmatrix}0\\1\end{pmatrix} = \left\|\begin{pmatrix}0\\1\end{pmatrix}\right\|_\infty^2 = 1\,,$$

also

$$1 = \left\|\begin{pmatrix}1\\1\end{pmatrix}\right\|_\infty^2 = \left(\begin{pmatrix}1\\0\end{pmatrix} + \begin{pmatrix}0\\1\end{pmatrix}\right) \star \left(\begin{pmatrix}1\\0\end{pmatrix} + \begin{pmatrix}0\\1\end{pmatrix}\right) = 1{+}1{+}2\begin{pmatrix}1\\0\end{pmatrix} \star \begin{pmatrix}0\\1\end{pmatrix}$$

und somit $\begin{pmatrix}1\\0\end{pmatrix} \star \begin{pmatrix}0\\1\end{pmatrix} = -\dfrac{1}{2}$ und $\begin{pmatrix}2\\0\end{pmatrix} \star \begin{pmatrix}0\\1\end{pmatrix} = -1$. Mithin wäre

$$4 = \left\|\begin{pmatrix}2\\1\end{pmatrix}\right\|_\infty^2 = \left(\begin{pmatrix}2\\0\end{pmatrix} + \begin{pmatrix}0\\1\end{pmatrix}\right) \star \left(\begin{pmatrix}2\\0\end{pmatrix} + \begin{pmatrix}0\\1\end{pmatrix}\right) = 4 + 1 - 2 = 3\,,$$

ein offensichtlicher Widerspruch.

Maximumsnormen lassen sich nicht nur für $\mathbb{R}^n$ oder $\mathbb{C}^n$ definieren, sondern auch für Funktionenräume: Eine stetige Funktion $f : [a,\, b] \to \mathbb{R}$ auf einem *abgeschlossenen* Intervall $[a,\, b]$ nimmt ihr Maximum wirklich an, d.h die Abbildung

$$\|\cdot\|_\infty : \mathcal{C}^0\big([a,\, b],\, \mathbb{R}\big) \to \mathbb{R}; \qquad f \mapsto \max_{x \in [a,\, b]} |f(x)|$$

ist wohldefiniert (für eine stetige Funktion f ist auch $|f|$ eine stetige Funktion) und sie hat die Eigenschaften *a)* bis *c)*: *a)* und *b)* sind, wie in den meisten Fällen, trivial, und sind $f, g \in \mathcal{C}^0\big([a, b], \ \mathbb{R}\big)$ zwei Funktionen, von denen f ihr Maximum in x_1 annimmt, g in x_2 und $f+g$ in x^*, so ist

$$\begin{aligned} \|f+g\|_\infty &= \big|(f+g)(x^*)\big| = \big|f(x^*)+g(x^*)\big| \le \big|f(x^*)\big| + \big|g(x^*)\big| \\ &\le \max_{x\in[a,\,b]} |f(x)| + \max_{x\in[a,\,b]} |g(x)| = f(x_1)+g(x_2) \\ &= \|f\|_\infty + \|g\|_\infty \ . \end{aligned}$$

Maximumsnormen spielen unter anderem in der Numerik eine große Rolle; sie gestatten es beispielsweise, Fehlerschranken für numerische Rechnungen herzuleiten:

Führen wir beispielsweise auf dem Vektorraum $\mathbb{R}^{n\times m}$ aller $n \times m$-Matrizen die Maximumsnorm ein, so ist natürlich

$$\|A\|_\infty = \max_{i=1}^{n} \max_{j=1}^{m} \big|a_{ij}\big| \ .$$

Betrachten wir auch $\mathbb{R}^m$ mit der Maximumsnorm, so folgt für ein Produkt $A\vec{v} = \vec{b}$ sofort aus der Multiplikationsregel

$$b_i = \sum_{j=1}^{m} a_{ij} v_j$$

und der gewöhnlichen Dreiecksungleichung aus $\mathbb{R}$

$$|b_i| \le \sum_{j=1}^{m} \big|a_{ij}\big| \cdot \big|v_j\big| \ ,$$

daß $\big\|\vec{b}\big\|_\infty \le \|A\|_\infty \cdot \|\vec{v}\|_\infty$ ist.

Wird also der Vektor $\vec{v}$ durch einen Fehlervektor $\vec{\epsilon}$ gestört, so ist

$$A(\vec{v}+\vec{\epsilon}) = A\vec{v} + A\vec{\epsilon} = \vec{b} + A\vec{\epsilon}$$

mit einem Fehler behaftet, dessen Komponenten *mit Sicherheit* kleiner sind als

$$\|A\|_\infty \cdot \|\vec{\epsilon}\|_\infty \ .$$

(Tatsächlich wird diese Schranke in den meisten Fällen viel zu groß sein, aber realistische Schranken sind in der Numerik oft – wenn überhaupt – nur mit sehr großem Aufwand zu finden.)

Ebenfalls eine sehr wichtige Rolle spielen Normen bei Funktionenräumen; dafür werden wir im nächsten Semester zahlreiche Beispiele sehen. Der wesentliche Punkt ist, daß man bezüglich einer Norm in offensichtlicher Verallgemeinerung der klassischen Definitionen Begriffe wie Konvergenz und Stetigkeit definieren kann:

Definition: *a)* $\big(V, \|\cdot\|\big)$ sei ein normierter $\mathbb{R}$-Vektorraum. Eine Folge $\vec{v}_1, \vec{v}_2, \dots$ von Vektoren aus V konvergiert gegen den Vektor $\vec{v} \in V$, wenn es zu jedem $\varepsilon > 0$ eine natürliche Zahl $N \in \mathbb{N}$ gibt, so daß $\|\vec{v} - \vec{v}_n\| < \varepsilon$ für alle $n > N$.
b) Eine Abbildung $f\colon V \to W$ zwischen zwei normierten $\mathbb{R}$-Vektorräumen $\big(V, \|\cdot\|_1\big)$ und $\big(W, \|\cdot\|_2\big)$ heißt stetig in $\vec{v}_0 \in V$, wenn es für jedes $\varepsilon > 0$ ein $\delta > 0$ gibt, so daß für alle $\vec{v} \in V$ gilt: Ist $\|\vec{v} - \vec{v}_0\| < \delta$, so ist $\|f(\vec{v}) - f(\vec{v}_0)\| < \varepsilon$.

Somit kann man also in normierten Vektorräumen Analysis betreiben, wobei die Ergebnisse stark von der Norm abhängen können.

In diesem Semester werden uns allerdings auf die Analysis im $\mathbb{R}^n$ beschränken, und da dort alle Normen äquivalent sind, also auf denselben Konvergenz- und auch Stetigkeitsbegriff führen, werden wir im nächsten Kapitel auf normierte Vektorräume verzichten.