Wichtige Kenntnisse der Linearen Algebra

In Kapitel 3 der Vorlesung werden wir sehen (und in Kapitel 6 vertiefen), dass zur Beschreibung von Quantensystemen mathematische Begriffe aus dem Gebiet der Linearen Algebra erforderlich sind. Insbesondere Vektorräume und lineare Operatoren sind dabei sehr wichtig. Vieles davon kennen Sie aus früheren Vorlesungen, aber möglicherweise sind Ihnen die Zusammenhänge nicht immer bewusst. Die folgenden Ausführungen sollen Ihnen dabei helfen.

1 Vektorräume

In MaMe I haben wir in (5.11) die allgemeine Definition eines Vektorraums gegeben. Für einen Vektorraum benötigt man eine Menge $\mathcal{V}$ von Vektoren, einen Zahlenkörper K und zwei Rechenoperationen, nämlich die Addition $v + w$ von Vektoren und die Multiplikation cv von Vektoren mit Skalaren (also Elementen aus K). Die Rechenoperationen müssen dabei den Axiomen aus (5.11) genügen.

Das Paradebeispiel für einen Vektorraum ist der $\mathbb{R}^3$. Wir hatten in MaMe I aber auch schon gesehen, dass Vektorräume von Funktionen gebildet werden können. Solche **Funktionenräume** sind für die Quantenmechanik sehr wichtig. Als Beispiel hatten wir den Raum $K_n[x]$ betrachtet, der aus den Polynomen vom Grad $\leq n$ mit Koeffizienten in K besteht.

Jeder Vektorraum besitzt eine **Basis**. Dies ist eine Menge von Vektoren $\mathcal{B} = \{v_1, v_2, \ldots, v_N\}$ von minimaler Mächtigkeit mit der Eigenschaft, dass sich jedes $v \in \mathcal{V}$ als Linearkombination der v_i schreiben lässt, d.h.

$$v = \sum_{i=1}^{N} c_i v_i \tag{1}$$

mit geeigneten Koeffizienten $c_i \in K$. Man sagt daher, eine Basis sein ein „minimales Erzeugendensystem" von $\mathcal{V}$, da sich jeder Vektor gemäß (1) durch die Basisvektoren ausdrücken lässt. Aufgrund der Minimalität der Basis ist diese Darstellung – für gegebenes $\mathcal{B}$ – immer eindeutig.[1]

Auf vielen Vektorräumen kann man – zusätzlich zu den in der Definition (5.11) geforderten beiden Rechenoperationen – auch noch ein **Skalarprodukt** zwischen Vektoren definieren. Aus dem $\mathbb{R}^3$ ist uns das natürlich wohlbekannt. Dort schreiben wir dafür $\vec{v} \cdot \vec{w}$. In den abstrakten Vektorräumen der Quantenmechanik benutzt man stattdessen die Schreibweise $\langle v|w\rangle$. Das sieht schicker aus. Formal ist ein Skalarprodukt eine Abbildung $\mathcal{V} \times \mathcal{V} \to K$, die gewisse Eigenschaften besitzt. Zum Beispiel muss gelten

$$\langle u|v + w\rangle = \langle u|v\rangle + \langle u|w\rangle \quad \text{und} \quad \langle v|cw\rangle = c\langle v|w\rangle \ .$$

Insbesondere gilt für ein Skalarprodukt stets $\langle v|v\rangle \geq 0$. Und $\langle v|v\rangle = 0$ gilt dann und nur dann, wenn $v = 0$ (d.h., v ist der Nullvektor in $\mathcal{V}$). Bei Vorliegen eines Skalarprodukts kann man deshalb die **Norm** eines Vektors definieren als $||v|| := \sqrt{\langle v|v\rangle}$. (Im Fall des $\mathbb{R}^3$ mit seinem üblichen Skalarprodukt entspricht die Norm gerade der Länge von v.) Außerdem kann man den Begriff der Orthogonalität einführen: zwei Vektoren v und w mit $\langle v|w\rangle = 0$ heißen **orthogonal**.

Eine **Orthonormalbasis** $\mathcal{B} = \{v_1, v_2, \ldots, v_N\}$ ist eine Basis mit besonders schönen Eigenschaften. Sie besteht aus Vektoren der Norm 1, die paarweise orthogonal zueinander sind. D.h. es gilt $\langle v_i|v_j\rangle = \delta_{ij}$. Im Fall des $\mathbb{R}^3$ ist z.B. $\{\vec{e}_x, \vec{e}_y, \vec{e}_z\}$ eine Orthonormalbasis.

[1] Beachten Sie aber: Die Basis eines Vektorraums ist nicht eindeutig bestimmt. Es gibt sogar fast immer unendlich viele verschiedene Basen für einen Vektorraum.

Beispiel für einen Funktionenraum mit Orthonormalbasis: Die Menge

$$\mathcal{C}^1[-\pi,\pi] := \{f : [-\pi,\pi] \to \mathbb{R} \mid f \text{ ist stetig differenzierbar}\} \tag{2}$$

der stetig differenzierbaren Funktionen auf dem Intervall $[-\pi,\pi]$ bildet einen Vektorraum über $\mathbb{R}$. Die Vektoraddition und Multiplikation mit Skalaren ist punktweise definiert als[2]

$$(f+g)(x) := f(x) + g(x) \quad \text{und} \quad (cf)(x) := c\,f(x)\ . \tag{3}$$

Man sieht schnell ein, dass diese beiden Rechenoperationen alle Axiome aus (5.11) in MaMe I erfüllen. So gilt z.B. $f+g = g+f$, da für jedes $x \in [-\pi,\pi]$ gilt $(f+g)(x) = f(x)+g(x) = g(x)+f(x) = (g+f)(x)$. Das neutrale Element der Vektoraddition ist die Nullabbildung e mit $e(x) = 0$ für jedes $x \in [-\pi,\pi]$.

Gemäß der Theorie der Fourier-Reihen (vgl. MaMe II) kann man jedes Element aus $\mathcal{C}^1[-\pi,\pi]$ als (unendliche) Linearkombination von Sinus- und Kosinus-Funktionen schreiben, d.h.

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} [a_n \cos(nx) + b_n \sin(nx)]\ .$$

Dementsprechend bildet die Menge $\mathcal{B} = \{1, \sin(x), \cos(x), \sin(2x), \cos(2x), \ldots\}$ eine Basis für $\mathcal{C}^1[-\pi,\pi]$. Ihre Mächtigkeit – und damit die Dimension des Vektorraums – ist unendlich.[3]

Auf $\mathcal{C}^1[-\pi,\pi]$ kann man ein Skalarprodukt definieren als

$$\langle f|g\rangle := \int_{-\pi}^{\pi} f(x)g(x)\,\mathrm{d}x\ . \tag{4}$$

Bezüglich dieses Skalarprodukts sind die obigen Basisvektoren paarweise orthogonal, da z.B. gilt $\int_{-\pi}^{\pi} \cos(nx)\cos(mx)\,\mathrm{d}x = \pi\delta_{nm}$. Indem wir die Basisvektoren noch entsprechend normieren, erhalten wir mit

$$\mathcal{B}' = \left\{\frac{1}{\sqrt{2\pi}}, \frac{\sin(x)}{\sqrt{\pi}}, \frac{\cos(x)}{\sqrt{\pi}}, \frac{\sin(2x)}{\sqrt{\pi}}, \frac{\cos(2x)}{\sqrt{\pi}}, \ldots\right\} \tag{5}$$

eine Orthonormalbasis für $\mathcal{C}^1[-\pi,\pi]$.

2 Lineare Operatoren

Die Vektorräume in der Quantenmechanik werden von den Wellenfunktionen gebildet. Messprozesse an Quantensystemen werden wir durch bestimmte lineare Operatoren beschreiben. Deren mathematische Definition lautet folgendermaßen:

Definition: Sei $\mathcal{V}$ ein Vektorraum über K. Ein linearer Operator ist eine Abbildung $\varphi : \mathcal{V} \to \mathcal{V}$ mit der Eigenschaft, dass für alle Vektoren $v, w \in \mathcal{V}$ und Skalare $c \in K$ gilt

$$\varphi(v+w) = \varphi(v) + \varphi(w) \quad \text{und} \quad \varphi(cv) = c\varphi(v)\ . \tag{6}$$

[2] Um den mathematischen Gehalt der Zeile (3) zu verstehen, muss man sich den Unterschied zwischen f und $f(x)$ ins Gedächtnis rufen: f bezeichnet die Funktion, wohingegen $f(x)$ den Funktionswert von f an der Stelle x bezeichnet. Die Gleichung $(f+g)(x) := f(x)+g(x)$ bedeutet also, dass wir die Funktion $f+g$ dadurch definieren, dass wir ihr an der Stelle x den Funktionswert $f(x)+g(x)$ zuweisen. Da x dabei alle Werte aus $[-\pi,\pi]$ durchläuft, erhalten wir auf diese Weise eine Funktionsvorschrift auf dem gesamten Intervall.

[3] Bitte beachten Sie, dass wir hierbei wieder die in der Physik übliche Notation verwenden und z.B. mit $\sin(x)$ die Funktion (und nicht den Funktionswert) meinen.

Beispiele:

a) Streckung: Sei $\mathcal{V}$ ein beliebiger Vektorraum über dem Körper K und $\lambda \in K$ eine feste Zahl. Dann ist die Abbildung

$$\varphi : \mathcal{V} \to \mathcal{V} \quad \text{mit} \quad \varphi(v) = \lambda v$$

ein linearer Operator. Denn es gilt $\varphi(v + w) = \lambda(v + w) = \lambda v + \lambda w = \varphi(v) + \varphi(w)$ und $\varphi(cv) = \lambda(cv) = c(\lambda v) = c\varphi(v)$.

b) Drehung: Wir betrachten den Vektorraum $\mathcal{V} = \mathbb{R}^2$ mit Basis $\{\vec{e}_x, \vec{e}_y\}$. Auf $\mathcal{V}$ definieren wir eine Abbildung $\hat{D}$ durch die Vorschrift

$$\hat{D}(v) = (\cos\alpha\, v_x + \sin\alpha\, v_y)\vec{e}_x + (-\sin\alpha\, v_x + \cos\alpha\, v_y)\vec{e}_y \quad \text{für} \quad v = v_x\vec{e}_x + v_y\vec{e}_y \ . \tag{7}$$

Gemäß Beispiel II im Kapitel 5 von MaMe I stellt $\hat{D}$ eine Drehung um den Winkel α dar. Man kann wieder direkt nachrechnen, dass die Abbildungsvorschrift die Eigenschaften in (6) besitzt. Man kann sich diese auch geometrisch klarmachen: Es ist gleichwertig, ob man zwei gegebene Vektoren v und w zuerst addiert und anschließend den Summenvektor $v + w$ dreht oder zuerst v und w einzeln dreht und danach aufaddiert. Folglich ist $\hat{D}$ ein linearer Operator.

c) Ableitung: Als Beispiel für einen Funktionenraum betrachten wir noch den Raum $\mathcal{V} = K_n[x]$ der Polynome vom Grad $\leq n$. Auf $\mathcal{V}$ definieren wir die Abbildung $\hat{A}$ als Ableitung nach x, d.h.

$$\hat{A}(f) = \frac{\mathrm{d}f}{\mathrm{d}x} \quad \text{für} \quad f \in K_n[x] \ . \tag{8}$$

Aufgrund der Eigenschaften von Ableitungen [vgl. (3.2a) in MaMe I] ist $\hat{A}$ ein linearer Operator.

Lineare Operatoren als Matrizen: Einen linearen Operator kann man immer als eine Matrix auffassen – unabhängig davon, wie abstrakt der zugrundeliegende Vektorraum $\mathcal{V}$ ist. Die Wirkung des Operators wird dann einfach beschrieben durch die Multiplikation der Matrix mit Spaltenvektoren. Dazu muss man lediglich eine feste Basis $\mathcal{B}$ von $\mathcal{V}$ wählen. Da sich jeder Vektor v in eindeutiger Weise als Linearkombination der Basisvektoren schreiben lässt, kann man ihn als einen Spaltenvektor $\Phi(v)$ auffassen, dessen Einträge gerade die Entwicklungskoeffizienten c_i in (1) sind. Nach (5.12) in MaMe I nennt man $\Phi(v)$ den Koordinatenvektor zu v.

Die Berechnung der zu einem linearen Operator gehörigen Matrix gestaltet sich besonders einfach, wenn $\mathcal{B}$ eine Orthonormalbasis ist. Das ist im obigen Beispiel b) der Fall. Die Matrixeinträge erhält man hier über die Formel

$$D_{ij} = \langle \vec{e}_i | \hat{D}(\vec{e}_j) \rangle \ . \tag{9}$$

Um den Eintrag in der i-ten Zeile und j-ten Spalte zu finden, muss man also das Bild $\hat{D}(\vec{e}_j)$ des j-ten Basisvektors unter $\hat{D}$ bestimmen und dieses vermöge des Skalarprodukts auf den i-ten Basisvektor projizieren. Wegen $\hat{D}(\vec{e}_x) = \cos\alpha\,\vec{e}_x - \sin\alpha\,\vec{e}_y$ und $\hat{D}(\vec{e}_y) = \sin\alpha\,\vec{e}_x + \cos\alpha\,\vec{e}_y$ nach (7) erhalten wir die Matrix

$$(D_{ij}) = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \ .$$

Unter Verwendung von Koordinatenvektoren und dieser Matrix kann man die Vorschrift in (7) nun auch schreiben als

$$(D_{ij})\Phi(v) = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} v_x \\ v_y \end{pmatrix} = \begin{pmatrix} \cos\alpha\, v_x + \sin\alpha\, v_y \\ -\sin\alpha\, v_x + \cos\alpha\, v_y \end{pmatrix} = \Phi\left(\hat{D}(v)\right) \ . \tag{10}$$

In der Quantenmechanik werden wir stets mit Orthonormalbasen arbeiten. Deshalb genügt es, wenn Sie diese Art der Berechnung der zu einem linearen Operator gehörigen **Darstellungsmatrix** kennen. Der Vollständigkeit zuliebe betrachten wir aber auch noch das obige Beispiel c), in dem keine Orthonormalbasis vorliegt.

Als Basis des $K_n[x]$ hatten wir in MaMe I die Menge $\mathcal{B} = \{1, x, x^2, x^3, \ldots, x^n\}$ der Monome kennengelernt. Wir bezeichnen die Basisvektoren hier als f_0, f_1, f_2,. . . ,f_n. Um die Darstellungsmatrix des Ableitungsoperators bzgl. dieser Basis zu berechnen, müssen wir wieder zunächst die Bilder der Basisvektoren bestimmen:

$$\hat{A}(f_0) = 0\,,\ \ \hat{A}(f_1) = 1\,,\ \ \hat{A}(f_2) = 2x = 2f_1\,,\ \ \hat{A}(f_3) = 3x^2 = 3f_2\,,\ \ldots,\ \ \hat{A}(f_n) = nx^{n-1} = nf_{n-1}\,.$$

In die j-te Spalte der Matrix (A_{ij}) schreiben wir nun die Entwicklungskoeffizienten des Bildvektors $\hat{A}(f_j)$ bzgl. der Basis $\mathcal{B}$ (d.h., wir drücken $\hat{A}(f_j)$ als Linearkombination in der Basis $\mathcal{B}$ aus, was in den obigen Gleichungen bereits geschehen ist). Die Matrix lautet somit

$$(A_{ij}) = \begin{pmatrix} 0 & 1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 2 & 0 & \ldots & 0 \\ 0 & 0 & 0 & 3 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & n \\ 0 & 0 & 0 & 0 & \ldots & 0 \end{pmatrix}.$$

Wir können uns davon überzeugen, dass diese Matrix tatsächlich eine Ableitung darstellt. Dazu wählen wir ein allgemeines Element $f(x) = \sum_{k=0}^{n} a_k\, x^k$ aus $K_n[x]$ und bilden die Ableitung:

$$\hat{A}(f) = \sum_{k=1}^{n} k\, a_k\, x^{k-1} = a_1 + 2a_2\, x + 3a_3\, x^2 + \ldots na_n\, x^{n-1}\ .$$

Wenn wir stattdessen mit Koordinatenvektoren und der Matrix arbeiten, sieht das so aus:

$$(A_{ij})\Phi(f) = \begin{pmatrix} 0 & 1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 2 & 0 & \ldots & 0 \\ 0 & 0 & 0 & 3 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & n \\ 0 & 0 & 0 & 0 & \ldots & 0 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} a_1 \\ 2a_2 \\ 3a_3 \\ \vdots \\ na_n \\ 0 \end{pmatrix} = \Phi\left(\hat{A}(f)\right)\ . \tag{11}$$

Tatsächlich ist die Wirkung der Matrix (A_{ij}) auf den Koordinatenvektoren also äquivalent zur Wirkung des Operators $\hat{A}$ auf den Polynomfunktionen!

Spektralzerlegung von Matrizen: Ein wichtiges Merkmal einer Matrix sind ihre Eigenwerte und Eigenvektoren (vgl. Abschnitt 5.4 in MaMe I). Für reelle symmetrische Matrizen erlauben sie insbesondere eine sehr nützliche Darstellung, die als Spektralzerlegung bekannt ist. Sei M eine solche $n \times n$-Matrix mit Eigenwerten λ_i und normierten (!) Eigenvektoren $\vec{v}_i$. Dann gilt

$$M = \sum_{i=1}^{n} \lambda_i\, \vec{v}_i\, \vec{v}_i^{\,T}\ . \tag{12}$$

Hierbei bezeichnet $\vec{v}_i^{\,T}$ den zu $\vec{v}_i$ transponierten Vektor. Während $\vec{v}_i$ ein Spaltenvektor ist, ist $\vec{v}_i^{\,T}$ folglich ein Zeilenvektor. Das Produkt $\vec{v}_i\, \vec{v}_i^{\,T}$ ergibt deshalb gemäß der Regel „Zeile × Spalte" eine Matrix (im Gegensatz zum umgekehrten Produkt $\vec{v}_i^T \vec{v}_i$, das einen Skalar liefert).

Beispiel: Die Eigenwerte und zugehörigen normieren Eigenvektoren der Matrix

$$M = \begin{pmatrix} 3 & \sqrt{8} \\ \sqrt{8} & 5 \end{pmatrix} \quad \text{lauten} \quad \lambda_1 = 1 \ , \quad \lambda_2 = 7 \ , \quad \vec{v}_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} \sqrt{2} \\ -1 \end{pmatrix} \ , \quad \vec{v}_2 = \frac{1}{\sqrt{6}} \begin{pmatrix} \sqrt{2} \\ 2 \end{pmatrix} .$$

Eine einfache Rechnung ergibt (Prüfen Sie das nach!)

$$\lambda_1 \, \vec{v}_1 \, \vec{v}_1^T + \lambda_2 \, \vec{v}_2 \, \vec{v}_2^T = 1 \cdot \frac{1}{3} \begin{pmatrix} \sqrt{2} \\ -1 \end{pmatrix} \begin{pmatrix} \sqrt{2} & -1 \end{pmatrix} + 7 \cdot \frac{1}{6} \begin{pmatrix} \sqrt{2} \\ 2 \end{pmatrix} \begin{pmatrix} \sqrt{2} & 2 \end{pmatrix} = \begin{pmatrix} 3 & \sqrt{8} \\ \sqrt{8} & 5 \end{pmatrix} = M \ ,$$

in Übereinstimmung mit (12).

In der Quantenmechanik werden wir Spektralzerlegungen für bestimmte lineare Operatoren $\hat{A}$ benutzen. Es wird Sie nicht überraschen, dass das grundsätzlich möglich ist. Denn wir haben oben ja die enge Korrespondenz zwischen linearen Operatoren und Matrizen gesehen. Als komplexe Verallgemeinerung der Forderung, dass es sich bei (12) um reelle symmetrische Matrizen mit $M = M^T$ handeln muss, werden die Operatoren die Eigenschaft haben, **hermitesch** zu sein. Ihre Darstellungsmatrizen erfüllen also die Bedingung $A = (A^*)^T$ [vgl. auch (5.20) in MaMe I].

Oft sind die linearen Operatoren in der Quantenmechanik auf einem unendlich-dimensionalen Vektorraum definiert. Die zugehörigen Darstellungsmatrizen sind dann ebenfalls unendlich groß. Die obigen Konzepte lassen sich auch auf diesen Fall anwenden.