

Mathematisches Basiswissen für
Theoretische Physik T2

Helmut Neufeld
Fakultät für Physik
Universität Wien

Sommersemester 2012

Inhaltsverzeichnis

| | | |
|----------|---|-----------|
| 1 | Vektorräume und lineare Abbildungen | 1 |
| 1.1 | Vektorräume | 1 |
| 1.2 | Teilräume | 3 |
| 1.3 | Lineare Abhängigkeit und Unabhängigkeit | 4 |
| 1.4 | Basis eines Vektorraums | 4 |
| 1.5 | Matrizen | 5 |
| 1.6 | Lineare Abbildungen | 8 |
| 1.7 | Basiswechsel | 11 |
| 1.8 | Determinanten | 11 |
| 1.9 | Eigenwerte und Eigenvektoren | 13 |
| 1.10 | Charakteristisches Polynom, Spur | 16 |
| 2 | Euklidische Vektorräume | 17 |
| 2.1 | Skalarprodukt | 17 |
| 2.2 | Transponierte Abbildung | 19 |
| 2.3 | Norm | 20 |
| 2.4 | Orthonormalbasis | 22 |
| 2.5 | Orthogonale Transformationen | 25 |
| 2.6 | $SO(3)$ | 30 |
| 2.7 | Eulersche Winkel | 31 |

| | | |
|----------|---|-----------|
| 3 | Unitäre Vektorräume | 35 |
| 3.1 | Komplexes Skalarprodukt | 35 |
| 3.2 | Norm | 37 |
| 3.3 | Adjungierte Abbildung | 38 |
| 3.4 | Orthonormalbasis | 40 |
| 3.5 | Projektionsoperatoren | 43 |
| 3.6 | Hermitesche Operatoren | 44 |
| 3.7 | Unitäre Operatoren | 46 |
| 3.8 | Normale Operatoren | 47 |
| 3.9 | Spektralsatz für normale Operatoren | 48 |
| 3.10 | Gleichzeitige Diagonalisierbarkeit | 53 |
| 3.11 | Funktionen normaler Operatoren | 54 |
| 4 | Tensorprodukte von Vektorräumen | 57 |
| 4.1 | Tensorprodukte | 57 |
| 4.2 | Tensorprodukt von Zustandsräumen | 59 |
| 5 | Verallgemeinerte Funktionen | 61 |
| 5.1 | Motivation | 61 |
| 5.2 | Testfunktionen | 62 |
| 5.3 | Distributionen | 63 |
| 5.4 | Differentiation von Distributionen | 64 |
| 5.5 | Rechnen mit Distributionen | 65 |
| 5.6 | Deltafolgen | 66 |
| 5.7 | Greenfunktion des Laplaceoperators | 67 |

Kapitel 1

Vektorräume und lineare Abbildungen

Ein Vektorraum (oder linearer Raum) besteht aus Elementen (den so genannten Vektoren), für die sowohl eine Addition als auch eine Multiplikation mit Skalaren (den Elementen des Grundkörpers) erklärt ist. Hier werden die wichtigsten Definitionen und Sätze aus der Theorie der Vektorräume zusammengestellt.

1.1 Vektorräume

Definition: Ein **Vektorraum** \mathcal{V} über einem Körper \mathbb{K} (in der Physik sind nur $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ von Bedeutung) ist eine Menge von Elementen mit folgenden Eigenschaften:

(V1) $\forall v_1, v_2 \in \mathcal{V}$ ist eine Summe $v_1 + v_2 \in \mathcal{V}$ definiert

(V2) $v_1 + v_2 = v_2 + v_1$ (Kommutativität der Vektoraddition)

(V3) $v_1 + (v_2 + v_3) = (v_1 + v_2) + v_3$ (Assoziativität der Vektoraddition)

(V4) $\exists 0 \in \mathcal{V}$ mit $v + 0 = 0 + v = v \forall v \in \mathcal{V}$ (Nullvektor)

(V5) $\forall v \in \mathcal{V} \exists -v \in \mathcal{V}$ mit $v + (-v) = 0$

(V6) $\forall a \in \mathbb{K}$ und $\forall v \in \mathcal{V}$ ist $av \in \mathcal{V}$ definiert

(V7) $a_1(a_2v) = (a_1a_2)v$

(V8) $(a_1 + a_2)v = a_1v + a_2v$

$$(V9) \quad a(v_1 + v_2) = av_1 + av_2$$

$$(V10) \quad 1v = v \quad \forall v \in \mathcal{V} \quad (1 \text{ ist das Einheitsselement des Grundkörpers } \mathbb{K})$$

Beispiele für Vektorräume:

1. \mathbb{K}^n = Menge aller Spaltenvektoren

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad x_i \in \mathbb{K},$$

wobei die Addition zweier Vektoren $x, y \in \mathbb{K}^n$ durch

$$x + y = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}, \quad x_i, y_i \in \mathbb{K}$$

und die Multiplikation mit einem Skalar $a \in \mathbb{K}$ durch

$$ax = a \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} ax_1 \\ \vdots \\ ax_n \end{pmatrix}$$

definiert sind.

2. Die Menge aller Polynome

$$p(x) = \sum_{k=0}^{\infty} c_k x^k, \quad c_k \in \mathbb{C}$$

(nur endlich viele c_k von Null verschieden) auf der Zahlengeraden \mathbb{R} bildet einen Vektorraum über dem Grundkörper \mathbb{C} , wenn man die Addition der Polynome p und q ,

$$q(x) = \sum_{k=0}^{\infty} d_k x^k, \quad d_k \in \mathbb{C},$$

durch

$$(p + q)(x) = \sum_{k=0}^{\infty} (c_k + d_k) x^k$$

und die Multiplikation mit einem Skalar $a \in \mathbb{C}$ durch

$$(ap)(x) = \sum_{k=0}^{\infty} (ac_k) x^k$$

erklärt.

3. Die trigonometrischen Polynome der Form

$$\psi(x) = \sum_{n=-N}^N c_n e^{2\pi i n x}, \quad c_n \in \mathbb{C}, \quad x \in [0, 1]$$

auf dem Intervall $[0, 1]$ bilden einen Vektorraum über \mathbb{C} . Addition und Multiplikation mit Skalaren sind, wie im vorigen Beispiel, punktweise definiert.

4. S sei eine beliebige Menge. Dann bildet die Menge $\mathcal{F}(S)$ aller Funktionen $f : S \rightarrow \mathbb{K}$ einen Vektorraum über \mathbb{K} . Die algebraischen Operationen sind dabei punktweise definiert:

$$(f_1 + f_2)(s) = f_1(s) + f_2(s), \quad (af)(s) = af(s) \quad \forall s \in S.$$

5. Die Menge $C([a, b])$ aller reellwertigen stetigen Funktionen auf $[a, b] \subset \mathbb{R}$ bildet mit den punktweisen algebraischen Operationen einen Vektorraum über \mathbb{R} .

1.2 Teilräume

Definition: Unter einem (linearen) **Teilraum** \mathcal{W} eines Vektorraums \mathcal{V} versteht man eine Teilmenge von \mathcal{V} , die mit den in \mathcal{V} erklärten algebraischen Operationen selbst einen Vektorraum bildet.

Satz: Eine Teilmenge $\mathcal{W} \subseteq \mathcal{V}$ ist genau dann ein Teilraum von \mathcal{V} , wenn mit $w_1, w_2 \in \mathcal{W}$ auch $a_1 w_1 + a_2 w_2 \in \mathcal{W} \quad \forall a_1, a_2 \in \mathbb{K}$ gilt.

Beispiele:

1. $\{0\}$ und \mathcal{V} sind stets Teilräume von \mathcal{V} .
2. Ist $v \in \mathcal{V}$, dann ist $\{av \mid a \in \mathbb{K}\}$ ein Teilraum von \mathcal{V} .
3. Sind $v_1, v_2 \in \mathcal{V}$, dann ist die Menge aller Elemente der Gestalt $a_1 v_1 + a_2 v_2$ mit $a_1, a_2 \in \mathbb{K}$ ein Teilraum von \mathcal{V} .
4. Im \mathbb{R}^3 sind die durch den Ursprung gehenden Geraden und Ebenen Teilräume.

Bemerkung: Sei X eine beliebige Teilmenge von \mathcal{V} . Der **kleinste** Teilraum von \mathcal{V} , der X umfasst, wird dann mit $[X]$ bezeichnet. Man nennt ihn den von X **erzeugten** Teilraum. $[X]$ besteht aus allen **endlichen** Linearkombinationen der Gestalt $\sum_i a_i x_i$ mit $a_i \in \mathbb{K}$, $x_i \in X$.

1.3 Lineare Abhängigkeit und Unabhängigkeit

Definition:

1. Eine Teilmenge X von \mathcal{V} heißt **linear unabhängig**, wenn jede beliebige Linearkombination $\sum_i a_i x_i$ von Elementen $x_i \in X$ nur dann gleich dem Nullvektor sein kann ($\sum_i a_i x_i = 0$), wenn alle Koeffizienten $a_i = 0$ sind.
2. Eine Teilmenge X von \mathcal{V} heißt **linear abhängig**, wenn es eine **nicht triviale** Darstellung der Gestalt $\sum_i a_i x_i = 0$ gibt (d.h. wenigstens ein $a_i \neq 0$).
3. Ein Element $v \in \mathcal{V}$ heißt linear abhängig von X , falls es eine Darstellung $v = \sum_i a_i x_i$, $a_i \in \mathbb{K}$, $x_i \in X$ gibt.

Beispiele:

1. $X = \{0\}$ ist linear abhängig.
2. Im \mathbb{R}^3 sind die Vektoren

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

linear unabhängig.

3. Drei Vektoren in einer Ebene sind linear abhängig.

Satz: Sind X und Y endliche linear unabhängige Teilmengen eines Vektorraums \mathcal{V} , die denselben Teilraum erzeugen ($[X] = [Y]$), dann besitzen sie gleich viele Elemente.

1.4 Basis eines Vektorraums

Definition: Unter einer **Basis** eines Vektorraums \mathcal{V} versteht man eine linear unabhängige Teilmenge $B \subseteq \mathcal{V}$, die \mathcal{V} erzeugt.

Bemerkung: Man kann zeigen, dass jeder Vektorraum eine Basis (im obigen Sinn) besitzt. Dazu benötigt man im allgemeinen Fall allerdings das Auswahlaxiom (oder ein dazu äquivalentes Resultat). Wir werden uns hier aber nur mit dem wesentlich einfacheren Fall von endlichdimensionalen Vektorräumen beschäftigen.

Definition: Ein Vektorraum heißt **endlichdimensional**, wenn er von einer endlichen Teilmenge erzeugt wird.

Für einen endlichdimensionalen Vektorraum \mathcal{V} lässt sich die Existenz einer Basis aber leicht zeigen: Sei E eine endliche Teilmenge von \mathcal{V} mit $[E] = \mathcal{V}$. Sei B eine minimale Teilmenge von E , sodass $[B] = \mathcal{V}$. Man sieht sofort, dass B linear unabhängig ist (die Annahme der linearen Abhängigkeit führt auf einen Widerspruch) und somit ist gezeigt, dass B eine Basis von \mathcal{V} ist.

Aus dem letzten Satz des vorigen Abschnitts folgt, dass je zwei Basen eines endlichdimensionalen Vektorraums gleich viele Elemente besitzen. Man nennt diese Zahl die **Dimension** $\dim \mathcal{V}$ des Vektorraums \mathcal{V} . Ist also B eine Basis von \mathcal{V} , dann ist $\dim \mathcal{V}$ gleich der Anzahl der Elemente der Basis B .

Beispiele:

1. Der Vektorraum \mathbb{K}^n besitzt eine Basis, die aus den Elementen

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

besteht. Also ist \mathbb{K}^n ein n -dimensionaler Vektorraum. Man bezeichnet $\{e_1, e_2, \dots, e_n\}$ auch als **Standardbasis** des \mathbb{K}^n .

2. Der Vektorraum aller Polynome vom Grad $\leq n$ auf \mathbb{R} besitzt die Basis $\{1, x, x^2, \dots, x^n\}$.

Satz: Ist $\{f_1, \dots, f_n\}$ eine Basis des n -dimensionalen Vektorraums \mathcal{V} , dann lässt sich jedes $v \in \mathcal{V}$ **eindeutig** in der Form

$$v = \sum_{i=1}^n v_i f_i, \quad v_i \in \mathbb{K}$$

schreiben.

1.5 Matrizen

Eine $m \times n$ -Matrix $A = (A_{ik})$ über einem Körper \mathbb{K} ist ein Schema mit m Zeilen und n Spalten:

$$A = (A_{ik}) = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \vdots & \vdots & & \vdots \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{pmatrix}, \quad A_{ik} \in \mathbb{K}.$$

Ist A eine $m \times n$ -Matrix, so induziert A eine **lineare Abbildung** $A : \mathbb{K}^n \rightarrow \mathbb{K}^m$, welche das Element

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n$$

durch die Vorschrift

$$y_i = \sum_{k=1}^n A_{ik}x_k, \quad i = 1, \dots, m$$

in das Element

$$y = Ax = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} \in \mathbb{K}^m$$

überführt. Linearität bedeutet dabei, dass

$$A(ax + a'x') = aAx + a'Ax'$$

$\forall x, x' \in \mathbb{K}^n$ und $\forall a, a' \in \mathbb{K}$. \mathbb{K}^n ist der Definitionsbereich von A und \mathbb{K}^m der Wertebereich (range) von A . Man schreibt dafür $D(A) = \mathbb{K}^n$ und $R(A) = \mathbb{K}^m$.

Das **Produkt** $C = BA$ der Matrizen B und A ist genau dann definiert, falls $D(B) = R(A)$. Es ist definiert durch

$$C = \underbrace{B}_{m \times p} \underbrace{A}_{p \times n} = (C_{ik}) \quad \text{mit } C_{ik} = \sum_{l=1}^p B_{il}A_{lk},$$

wobei $D(BA) = D(A)$, $R(BA) = R(B)$:

$$\mathbb{K}^n \xrightarrow{A} \mathbb{K}^p \xrightarrow{B} \mathbb{K}^m.$$

Sind A, B, C Matrizen und die Produkte CB und BA definiert, dann sind auch die Produkte $(CB)A$ und $C(BA)$ definiert. Diese stimmen überein und man schreibt daher einfach CBA ohne Klammern.

Die Wirkung der Matrix A mit $D(A) = \mathbb{K}^n$ auf den Vektor $x \in \mathbb{K}^n$ ist ein Spezialfall des Matrixprodukts, wenn man x als $n \times 1$ -Matrix auffasst. Aus der Assoziativität ergibt sich $(BA)x = B(Ax)$, d.h. die Multiplikation zweier Matrizen entspricht der Hintereinanderausführung der entsprechenden Abbildungen.

Der **identischen Abbildung** auf \mathbb{K}^n entspricht die Einheitsmatrix

$$\mathbb{1}_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix},$$

wobei $\mathbb{1}_n x = x \forall x \in \mathbb{K}^n$. Ist die Dimension aus dem Zusammenhang klar, schreibt man einfach $\mathbb{1}$ statt $\mathbb{1}_n$.

Sind A und B $n \times n$ -Matrizen, dann sind AB und BA definiert, diese Produkte sind aber im Allgemeinen verschieden. Der **Kommutator** von A und B ,

$$[A, B] = AB - BA,$$

verschwindet also i.A. nicht.

Eine $n \times n$ -Matrix A heißt **invertierbar**, wenn es eine $n \times n$ -Matrix A^{-1} gibt mit $A^{-1}A = \mathbb{1}$. Daraus folgt, dass auch $AA^{-1} = \mathbb{1}$ ist. Man nennt A^{-1} die (zu A) **inverse Matrix**, diese ist eindeutig bestimmt. Sind zwei $n \times n$ -Matrizen A und B invertierbar, dann ist auch AB invertierbar und es gilt $(AB)^{-1} = B^{-1}A^{-1}$.

Je zwei $m \times n$ -Matrizen A, B kann man addieren und mit Elementen aus \mathbb{K} multiplizieren:

$$A + B = (A_{ik} + B_{ik}), \quad aA = (aA_{ik}).$$

Es gilt

$$(A + B)C = AC + BC, \quad A(B + C) = AB + AC, \quad (aA)B = A(aB) = a(AB),$$

sofern die auftretenden Produkte definiert sind.

Bemerkungen:

1. Die Menge aller $n \times n$ -Matrizen bildet einen **Ring** bezüglich der Addition und Matrixmultiplikation. Unter einem Ring versteht man eine Menge, in der alle Körperaxiome mit Ausnahme der Kommutativität der Multiplikation und der Existenz des inversen Elements erfüllt sind.
2. Von einer $n \times n$ -Matrix A kann man beliebige Potenzen A^k und Linearkombinationen davon, also Ausdrücke der Form

$$\sum_{k=0}^p a_k A^k, \quad a_k \in \mathbb{K}$$

bilden. Man kann diese Konstruktion auch auf unendliche Potenzreihen erweitern, indem man auf der Menge der quadratischen Matrizen eine geeignete Norm zur Klärung von Konvergenzfragen einführt.

Eine wichtige Matrixoperation ist die Bildung der **transponierten Matrix** A^T . Das ist jene Matrix, die man erhält, wenn man die Zeilen und Spalten der Matrix

A vertauscht:

$$A = \underbrace{\begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \vdots & \vdots & & \vdots \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{pmatrix}}_{m \times n\text{-Matrix}}, \quad A^T = \underbrace{\begin{pmatrix} A_{11} & A_{21} & \dots & A_{m1} \\ A_{12} & A_{22} & \dots & A_{m2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \dots & A_{mn} \end{pmatrix}}_{n \times m\text{-Matrix}}.$$

Es gilt:

$$(AB)^T = B^T A^T, \quad (A_1 + A_2)^T = A_1^T + A_2^T, \quad (aA)^T = aA^T.$$

1.6 Lineare Abbildungen

\mathcal{U} und \mathcal{V} seien Vektorräume über dem Grundkörper \mathbb{K} . Eine Abbildung $A : \mathcal{U} \rightarrow \mathcal{V}$ heißt **linear**, falls

$$A(a_1 u_1 + a_2 u_2) = a_1 A u_1 + a_2 A u_2 \quad \forall a_{1,2} \in \mathbb{K}, \forall u_{1,2} \in \mathcal{U}.$$

Die Menge aller **linearen** Abbildungen von \mathcal{U} nach \mathcal{V} bezeichnen wir mit $L(\mathcal{U}, \mathcal{V})$. Falls die Räume \mathcal{U} und \mathcal{V} übereinstimmen, schreiben wir $L(\mathcal{U}, \mathcal{U}) = L(\mathcal{U})$. Ist $\mathcal{V} = \mathbb{K}$ (also der Grundkörper), schreibt man $L(\mathcal{U}, \mathbb{K}) = \tilde{\mathcal{U}}$. $\tilde{\mathcal{U}}$ wird **Dualraum** von \mathcal{U} genannt, die Elemente von $\tilde{\mathcal{U}}$ heißen **lineare Funktionale**.

Beispiele von linearen Abbildungen:

1. Jede lineare Abbildung von \mathbb{K}^n nach \mathbb{K}^m hat die Gestalt $x \rightarrow Ax$ mit einer $m \times n$ -Matrix (A_{ik}) .
2. Jedes lineare Funktional $l \in \tilde{\mathbb{K}^n}$ lässt sich in der Form

$$l(x) = y^T x, \quad x \in \mathbb{R}^n$$

mit einem festen, eindeutig bestimmten $y \in \mathbb{K}^n$ schreiben. Da $l = y^T$ ein n -dimensionaler **Zeilenvektor** ist,

$$y^T = (y_1, \dots, y_n),$$

kann man die Menge der linearen Funktionale auf \mathbb{K}^n als Menge der n -dimensionalen Zeilenvektoren auffassen.

3. Sei \mathcal{V} ein beliebiger Vektorraum. Die durch

$$\mathbb{1}v = v \quad \forall v \in \mathcal{V}$$

definierte **identische Abbildung** ist eine lineare Abbildung $\mathbb{1} \in L(\mathcal{V})$.

4. \mathcal{P}_n sei der $(n + 1)$ -dimensionale Vektorraum aller reellwertigen Polynome vom Grad $\leq n$ auf \mathbb{R} . Durch die Vorschrift

$$(Dp)(x) = \frac{d}{dx}p(x)$$

wird eine lineare Abbildung $D : \mathcal{P}_n \rightarrow \mathcal{P}_n$ definiert.

5. Der durch $(T_h p)(x) = p(x-h)$ definierte Translationsoperator $T_h : \mathcal{P}_n \rightarrow \mathcal{P}_n$ ist ein weiteres Beispiel für eine lineare Abbildung.

6. Durch die Vorschrift

$$l(p) = \int_0^1 dx p(x)$$

wird auf \mathcal{P}_n ein lineares Funktional $l \in \tilde{\mathcal{P}}_n$ definiert.

7. Sei $x_0 \in \mathbb{R}$ eine feste reelle Zahl. Dann wird durch

$$\delta_{x_0}(p) = p(x_0)$$

ebenfalls ein lineares Funktional auf \mathcal{P}_n definiert.

Ist $\{f_1, \dots, f_n\}$ eine Basis von \mathcal{U} und $\{g_1, \dots, g_m\}$ eine Basis von \mathcal{V} , dann lässt sich die Wirkung einer linearen Abbildung $A \in L(\mathcal{U}, \mathcal{V})$ auf die Basisvektoren $\{f_1, \dots, f_n\}$ eindeutig bezüglich der Basis $\{g_1, \dots, g_m\}$ zerlegen:

$$Af_l = \sum_{k=1}^m g_k A_{kl}, \quad 1 \leq l \leq n, \quad A_{kl} \in \mathbb{K}.$$

Die $m \times n$ -Matrix (A_{kl}) bezeichnet man als **Matrixdarstellung** der linearen Abbildung A bezüglich der gewählten Basissysteme.

Zerlegt man einen Vektor $u \in \mathcal{U}$ bezüglich der Basis $\{f_1, \dots, f_n\}$,

$$u = \sum_{l=1}^n u_l f_l,$$

und den Vektor $v = Au \in \mathcal{V}$ bezüglich der Basis $\{g_1, \dots, g_m\}$,

$$v = \sum_{k=1}^m v_k g_k,$$

so besteht zwischen den Vektorkomponenten u_l und v_k der Zusammenhang

$$v_k = (Au)_k = \sum_{l=1}^n A_{kl} u_l,$$

d.h. der Vektor

$$\begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \in \mathbb{K}^n$$

wird durch die Matrixdarstellung (A_{kl}) von A auf den Vektor

$$\begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} = \begin{pmatrix} (Au)_1 \\ \vdots \\ (Au)_m \end{pmatrix} \in \mathbb{K}^m$$

abgebildet.

Bemerkung: Seien $\mathcal{A} \in L(\mathcal{U}, \mathcal{V})$, $\mathcal{B} \in L(\mathcal{V}, \mathcal{W})$ und $B_{\mathcal{U}}, B_{\mathcal{V}}, B_{\mathcal{W}}$ Basen von $\mathcal{U}, \mathcal{V}, \mathcal{W}$. Bezeichnet man die Matrixdarstellung von \mathcal{A} bezüglich der Basen $B_{\mathcal{U}}$ und $B_{\mathcal{V}}$ mit A und die Matrixdarstellung von \mathcal{B} bezüglich der Basen $B_{\mathcal{V}}$ und $B_{\mathcal{W}}$ mit B , so ist die Matrixdarstellung der zusammengesetzten Abbildung $\mathcal{B}\mathcal{A} \in L(\mathcal{U}, \mathcal{W})$ bezüglich $B_{\mathcal{U}}$ und $B_{\mathcal{W}}$ durch BA gegeben.

Aufgaben:

1. Ermitteln Sie die Matrixdarstellung der linearen Abbildung $D : \mathcal{P}_3 \rightarrow \mathcal{P}_3$ (Definition siehe oben), wenn sowohl im Ausgangs- als auch im Zielraum die Basis $B = \{1, x, x^2, x^3\}$ gewählt wird.
2. Gegeben sei die lineare Abbildung D des vorigen Beispiels. Wie wirken die linearen Abbildungen D^2, D^3, D^4 auf die Vektoren der Basis B ? Ermitteln Sie die Lösung sowohl durch Differenzieren als auch durch wiederholte Matrixmultiplikation der Matrixdarstellung von D .
3. Ermitteln Sie die Wirkung des Translationsoperators $T_h : \mathcal{P}_3 \rightarrow \mathcal{P}_3$ auf die Basis B und die dazugehörige Matrixdarstellung. Zeigen Sie, dass der Translationsoperator in der Form $T_h = \exp(-hD)$ geschrieben werden kann. Beachten Sie, dass die Potenzreihenentwicklung der Exponentialfunktion in diesem Fall abbricht, es sich also um eine endliche Summe handelt.
4. Betrachten Sie den von den Funktionen

$$e_n(x) = e^{2\pi i n x}, \quad x \in \mathbb{R}, \quad n = -N, -N + 1, \dots, N$$

erzeugten Funktionenraum über dem Grundkörper \mathbb{C} . (Diese Funktionen besitzen die Periode 1.) Zeigen Sie, dass durch die Operationen „Differentiation“ und „Translation“ lineare Abbildungen auf diesem Raum definiert werden. Ermitteln Sie sodann die dazugehörigen Matrixdarstellungen bezüglich der Basis $\{e_{-N}, \dots, e_N\}$. Verifizieren Sie wieder den Zusammenhang zwischen Translations- und Differentiationsoperator.

1.7 Basiswechsel

Sei \mathcal{V} ein n -dimensionaler Vektorraum mit einer Basis $\{f_1, \dots, f_n\}$. Ein Vektor $v \in \mathcal{V}$ hat dann die Zerlegung

$$v = \sum_{i=1}^n v_i f_i$$

bezüglich der gewählten Basis. Es sei $\{f'_1, \dots, f'_n\}$ eine andere Basis von \mathcal{V} . Bezüglich dieser besitzt der Vektor v die Zerlegung

$$v = \sum_{i=1}^n v'_i f'_i.$$

Zwischen den beiden Basissystemen besteht der Zusammenhang

$$f'_k = \sum_{i=1}^n f_i S_{ik}$$

mit einer invertierbaren Matrix

$$S = (S_{ik}) \in L(\mathbb{K}^n),$$

wobei

$$v'_i = (S^{-1})_{ik} v_k \Leftrightarrow v_i = S_{ik} v'_k.$$

Sei $\mathcal{A} \in L(\mathcal{V})$ mit der Matrixdarstellung $A = (A_{ik})$ bezüglich der Basis $\{f_1, \dots, f_n\}$ und der Matrixdarstellung $A' = (A'_{ik})$ bezüglich der Basis $\{f'_1, \dots, f'_n\}$. Dann besteht zwischen den beiden Matrixdarstellungen von \mathcal{A} der Zusammenhang

$$A = SA'S^{-1} \Leftrightarrow A' = S^{-1}AS.$$

1.8 Determinanten

Sei $A = (A_{ik})$ eine $n \times n$ -Matrix. Diese kann man sich aus n Spaltenvektoren $a_i \in \mathbb{K}^n$ aufgebaut denken,

$$A = (a_1, \dots, a_n),$$

wobei

$$a_i = Ae_i = \sum_{k=1}^n e_k A_{ki}.$$

Die **Determinante** von A ist durch

$$\det A = \det(a_1, \dots, a_n) = \sum_{\sigma \in \mathcal{S}_n} A_{\sigma(1)1} \dots A_{\sigma(n)n} \epsilon(\sigma)$$

definiert, wobei

$$\epsilon(\sigma) = \begin{cases} 1 & \text{falls } \sigma \text{ eine gerade Permutation} \\ -1 & \text{falls } \sigma \text{ eine ungerade Permutation} \end{cases}$$

ist. \mathcal{S}_n bezeichnet die Menge der Permutationen von $\{1, \dots, n\}$.

Die Determinante besitzt die folgenden Eigenschaften:

- (D1) $\det(a_1, \dots, a_n)$ ist multilinear (n -linear), d.h. $a_i \rightarrow \det(a_1, \dots, a_i, \dots, a_n)$ ist linear $\forall i$ bei festgehaltenen Argumenten $a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_n$.
- (D2) $\det(a_1, \dots, a_n) = 0$ falls $\{a_1, \dots, a_n\}$ linear unabhängig ist.
- (D3) $\det(e_1, \dots, e_n) = 1$, wobei $\{e_1, \dots, e_n\}$ die Standardbasis von \mathbb{K}^n ist.

Bemerkungen: Diese drei Eigenschaften legen die Determinante bereits eindeutig fest. Aus (D1) und (D2) folgt auch, dass $\det(a_1, \dots, a_n)$ **alternierend** ist, d.h. vertauscht man zwei Argumente, so ändert sich das Vorzeichen.

Satz: Die Vektoren $a_1, \dots, a_n \in \mathbb{K}^n$ sind genau dann linear unabhängig, wenn $\det(a_1, \dots, a_n) \neq 0$ ist.

Die Determinante liefert also ein Kriterium für die lineare Unabhängigkeit von n Vektoren im \mathbb{K}^n .

Weitere wichtige Eigenschaften der Determinante:

$$\det(AB) = \det A \det B, \quad \det A^{-1} = \frac{1}{\det A}, \quad \det A^T = \det A, \quad A, B \in L(\mathbb{K}^n).$$

Bemerkung: Sei \mathcal{V} ein beliebiger n -dimensionaler Vektorraum und $\mathcal{A} \in L(\mathcal{V})$. Dann kann man $\det \mathcal{A}$ durch

$$\det \mathcal{A} = \det A$$

definieren, wobei $A \in L(\mathbb{K}^n)$ eine beliebige Matrixdarstellung von \mathcal{A} ist. Diese Definition ist von der Wahl der Basis unabhängig, da aus $A' = S^{-1}AS$ und den Eigenschaften der Determinante folgt, dass $\det A' = \det A$.

1.9 Eigenwerte und Eigenvektoren

Definition: Sei $A \in L(\mathcal{V})$ eine lineare Abbildung auf einem Vektorraum \mathcal{V} . Angenommen es gibt eine Zahl $a \in \mathbb{K}$ und einen vom Nullvektor verschiedenen Vektor $v \in \mathcal{V}$, sodass

$$Av = av,$$

dann heißt a **Eigenwert** von A und v **Eigenvektor** von A (zum Eigenwert a).

$Av = av$, $v \neq 0$ ist äquivalent damit, dass $a\mathbb{1} - A$ nicht invertierbar ist und das ist genau dann der Fall, wenn

$$\det(a\mathbb{1} - A) = 0.$$

Die Eigenwerte von A sind also durch die Nullstellen des **charakteristischen Polynoms**

$$p_A(x) = \det(x\mathbb{1} - A)$$

bestimmt. Zu jedem Eigenwert gibt es mindestens einen Eigenvektor. Ist $\mathbb{K} = \mathbb{C}$, so existiert mindestens ein Eigenwert, da jedes Polynom in \mathbb{C} mindestens eine Nullstelle hat. Die Menge der Eigenwerte von A nennt man **Spektrum** von A , wir bezeichnen sie mit $\sigma(A)$.

Bemerkungen:

1. Die reelle 2×2 -Matrix

$$A = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}, \quad 0 \leq \theta < 2\pi$$

besitzt das charakteristische Polynom

$$p_A(x) = (x - \cos \theta)^2 + \sin^2 \theta.$$

Dieses hat die (i.A. komplexen) Nullstellen

$$x_{1,2} = \cos \theta \pm i \sin \theta = e^{\pm i\theta},$$

d.h. für $\theta \neq 0, \pi$ hat A keinen Eigenwert, wenn als Grundkörper $\mathbb{K} = \mathbb{R}$ gewählt wurde. Fasst man reelle Matrizen als Abbildungen auf dem \mathbb{R}^n auf, so müssen diese also keine Eigenwerte oder Eigenvektoren haben.

2. Die reelle 2×2 -Matrix

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

besitzt das charakteristische Polynom

$$p_A(x) = (x - 1)^2$$

mit der zweifach entarteten Nullstelle 1. Alle Eigenvektoren sind proportional zu

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Man sieht also, dass nicht notwendigerweise eine Basis von Eigenvektoren existieren muss, selbst wenn $\mathbb{K} = \mathbb{C}$ ist.

3. Sei \mathcal{V} ein n -dimensionaler Vektorraum und $A \in L(\mathcal{V})$. **Wenn** das charakteristische Polynom von A n **verschiedene** Nullstellen besitzt, dann bilden die dazugehörigen Eigenvektoren eine Basis.
4. Sei \mathcal{V} ein n -dimensionaler Vektorraum und $A \in L(\mathcal{V})$. Angenommen es gibt eine Basis $\{f_1, \dots, f_n\}$ von Eigenvektoren von A mit Eigenwerten a_1, \dots, a_n sind ($Af_i = a_i f_i$). Dann ist die Matrixdarstellung \widehat{A} von A bezüglich dieser Basis von Eigenvektoren die Diagonalmatrix

$$\widehat{A} = \begin{pmatrix} a_1 & & & \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot & \\ & & & & a_n \end{pmatrix},$$

d.h. durch die Wahl dieser Basis erhält die Matrixdarstellung von A eine besonders einfache Form. Lineare Abbildungen, die eine Basis von Eigenvektoren besitzen werden daher **diagonalisierbar** genannt.

Ein endliches Polynom der Form

$$p(A) = \sum_k c_k A^k, \quad c_k \in \mathbb{K}$$

ist wieder ein Element von $L(\mathcal{V})$. Die Wirkung von $p(A)$ auf die Basisvektoren ist durch

$$p(A)f_i = p(a_i)f_i, \quad 1 \leq i \leq n$$

gegeben, die Matrixdarstellung von $p(A)$ ist daher

$$\widehat{p(A)} = p(\widehat{A}) = \begin{pmatrix} p(a_1) & & & \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot & \\ & & & & p(a_n) \end{pmatrix}.$$

Die Anwendung von $p(A)$ auf einen beliebigen Vektor $v = \sum_{i=1}^n v_i f_i$ ergibt somit

$$p(A)v = \sum_{i=1}^n v_i p(a_i) f_i.$$

Sei nun $g : \sigma(A) \rightarrow \mathbb{K}$ eine beliebige, auf dem Spektrum von A definierte Funktion. Man definiert dann als Funktion $g(A)$ der linearen Abbildung A einfach

$$g(A)f_i = g(a_i)f_i, \quad 1 \leq i \leq n,$$

bzw.

$$g(A)v = \sum_{i=1}^n v_i g(a_i) f_i, \quad v = \sum_{i=1}^n v_i f_i.$$

5. Sei A eine $n \times n$ -Matrix. Aufgefasst als lineare Abbildung $A : \mathbb{K}^n \rightarrow \mathbb{K}^n$ ist A gerade die Matrixdarstellung bezüglich der Standardbasis $\{e_1, \dots, e_n\}$:

$$Ae_k = \sum_{i=1}^n e_i A_{ik}.$$

A möge eine Basis $\{f_1, \dots, f_n\}$ von Eigenvektoren ($Af_i = a_i f_i$) besitzen. Zwischen dieser Basis und der Standardbasis besteht die Beziehung

$$f_k = \sum_{i=1}^n e_i S_{ik}$$

und somit gilt

$$S_{ik} = (f_k)_i \Leftrightarrow S = (f_1, \dots, f_n),$$

sowie

$$\widehat{A} = S^{-1}AS = \text{diag}(a_1, \dots, a_n).$$

Beispiel: $A \in L(\mathbb{C}^2)$ sei gegeben durch

$$A = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

Die Eigenwerte sind $e^{\pm i\theta}$. Die Eigenvektoren zum Eigenwert $e^{+i\theta}$ haben die Form

$$x \begin{pmatrix} 1 \\ -i \end{pmatrix}, \quad x \neq 0 \text{ beliebig}$$

und die Eigenvektoren zum Eigenwert $e^{-i\theta}$ sind

$$x \begin{pmatrix} 1 \\ i \end{pmatrix}, \quad x \neq 0 \text{ beliebig.}$$

Wählt man als Basis von Eigenvektoren

$$f_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}, \quad f_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix},$$

so erhält man

$$S = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \Leftrightarrow S^{-1} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix}$$

mit

$$\widehat{A} = S^{-1}AS = \text{diag}(e^{i\theta}, e^{-i\theta}).$$

1.10 Charakteristisches Polynom, Spur

$A \in L(\mathcal{V})$ sei eine lineare Abbildung in einem n -dimensionalen Vektorraum \mathcal{V} . Für die Koeffizienten des **charakteristischen Polynoms**

$$p_A(x) = \det(x\mathbb{1} - A) = x^n + c_{n-1}x^{n-1} + \dots + c_1x + c_0, \quad c_k \in \mathbb{K}$$

gilt:

$$c_0 = \det(-A) = (-1)^n \det A, \quad c_{n-1} = -\operatorname{Tr}A.$$

Die **Spur** (trace) von A ist definiert durch

$$\operatorname{Tr}A = \sum_{i=1}^n A_{ii},$$

wobei (A_{ik}) eine beliebige Matrixdarstellung von A ist. Die Definition der Spur als Summe der Diagonalelemente A_{ii} ist selbstverständlich basisunabhängig.

Ist $\mathbb{K} = \mathbb{C}$, dann kann man das charakteristische Polynom in der Form

$$p_A(x) = \prod_{i=1}^n (x - a_i) = x^n - (a_1 + \dots + a_n)x^{n-1} + \dots + (-1)^n a_1 \dots a_n.$$

Durch Koeffizientenvergleich findet man in diesem Fall

$$\operatorname{Tr}A = a_1 + \dots + a_n = \sum_{i=1}^n a_i,$$

d.h. die Spur von A ist gleich der Summe der Eigenwerte von A , wobei entartete Eigenwerte entsprechend ihrer Vielfachheit zu zählen sind. Analog ist die Determinante von A gleich dem Produkt der Eigenwerte,

$$\det A = a_1 \dots a_n = \prod_{i=1}^n a_i.$$

Satz von Cayley-Hamilton: Sei $A \in L(\mathcal{V})$ und p_A das dazugehörige charakteristische Polynom. Dann gilt:

$$p_A(A) = 0.$$

Kapitel 2

Euklidische Vektorräume

In einem Vektorraum sind als Verknüpfungsvorschriften die Vektoraddition und die Multiplikation von Vektoren mit Skalaren (den Elementen des Grundkörpers) erklärt. Fügt man als weitere Struktur ein Skalarprodukt hinzu, so gelangt man im Fall eines reellen Vektorraums zur Definition des euklidischen und für komplexe Vektorräume zur Definition des unitären Vektorraums.

2.1 Skalarprodukt

Wie bereits aus dem Schulunterricht bekannt ist, kann man in der Ebene oder im dreidimensionalen Raum die Länge eines Vektors und den zwischen zwei Vektoren eingeschlossenen Winkel mit Hilfe des Skalarproduktes ausdrücken. Wir wollen hier Skalarprodukte axiomatisch, das heißt durch die Angabe ihrer charakteristischen Eigenschaften, einführen. Dadurch gewinnen wir einerseits mehr Einblick in die zugrundeliegenden mathematischen Strukturen und lösen uns andererseits von der Einengung auf elementargeometrische Anwendungen.

Definition: Ein (endlichdimensionaler) Vektorraum \mathcal{E} über dem Grundkörper \mathbb{R} heißt **euklidisch**, falls für je zwei Elemente $x, y \in \mathcal{E}$ ein **Skalarprodukt** (oder **inneres Produkt**) $(x|y) \in \mathbb{R}$ definiert ist, mit folgenden Eigenschaften:

$$(E1) \quad (a_1x_1 + a_2x_2|y) = a_1(x_1|y) + a_2(x_2|y), \quad x_{1,2} \in \mathcal{E}, \quad a_{1,2} \in \mathbb{R}$$

$$(E2) \quad (x|y) = (y|x)$$

$$(E3) \quad (x|x) \geq 0, \quad (x|x) = 0 \Leftrightarrow x = 0$$

Bemerkung: Die Eigenschaften (E1) (Linearität im ersten Argument) und (E2) (Symmetrie) implizieren Linearität auch im zweiten Argument, das heißt das Skalarprodukt ist **bilinear**.

Beispiele:

1. Das Paradebeispiel schlechthin ist natürlich der Raum \mathbb{R}^3 mit dem Skalarprodukt

$$(x|y) = \sum_{i=1}^3 x_i y_i.$$

Fasst man die Elemente des \mathbb{R}^3 als Spaltenvektoren (oder 3×1 -Matrizen) auf,

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},$$

so kann man mit Hilfe des dazugehörigen Zeilenvektors (bzw. der zu x transponierten 1×3 -Matrix)

$$x^T = (x_1, x_2, x_3)$$

das Skalarprodukt auch in Matrixnotation schreiben:

$$(x|y) = x^T y.$$

2. Im \mathbb{R}^n sei ein Skalarprodukt

$$(x|y) = x^T y = \sum_{i=1}^n x_i y_i$$

definiert. Diesen Raum nennen wir E^n .

3. Im \mathbb{R}^n sei ein inneres Produkt durch

$$(x|y) = \sum_{i=1}^n p_i x_i y_i$$

definiert, wobei $p_i > 0$ ($i = 1, \dots, n$) ist.

4. In dem Funktionenraum der auf einem endlichen Intervall $[a, b] \subset \mathbb{R}$ definierten reellen Polynome vom Grad $\leq n$,

$$p(x) = \sum_{k=0}^n c_k x^k, \quad c_k \in \mathbb{R}, \quad x \in [a, b],$$

kann man durch

$$(p|q) = \int_a^b dx p(x)q(x)$$

ein inneres Produkt definieren.

2.2 Transponierte Abbildung

Im Raum E^n (Beispiel 2) ist das Skalarprodukt durch $(x|y) = x^T y$ definiert, wobei x^T die „transponierte Matrix von x “ bedeutet. Diese sehr praktische Schreibweise kann für beliebige euklidische Räume verallgemeinert werden. Dazu geht man von der Beobachtung aus, dass im Falle des E^n die Abbildung $x \mapsto x^T$ eine „Identifizierung“ von E^n (dem Raum der n -dimensionalen Spaltenvektoren) mit seinem Dualraum \widetilde{E}^n (dem Raum der n -dimensionalen Zeilenvektoren) darstellt. Die Vorgabe eines Skalarproduktes erlaubt es nun, jeden euklidischen Vektorraum \mathcal{E} mit seinem Dualraum $\widetilde{\mathcal{E}}$ auf kanonische Weise zu identifizieren. Dazu ordnet man jedem $x \in \mathcal{E}$ das lineare Funktional $x^T \in \widetilde{\mathcal{E}}$ zu, welches durch

$$x^T(y) = x^T y = (x|y) \quad \forall y \in \mathcal{E}$$

definiert ist. Man überzeugt sich leicht davon, dass die Abbildung $x \mapsto x^T$ linear und injektiv ist. Da $\dim \mathcal{E} = \dim \widetilde{\mathcal{E}}$ ist, stimmt das Bild von \mathcal{E} unter der Abbildung T mit ganz $\widetilde{\mathcal{E}}$ überein. Bezeichnet man die inverse Abbildung ebenfalls mit T , so ist $x^{TT} = x \quad \forall x \in \mathcal{E}$ erfüllt.

Für eine lineare Abbildung $A : E^n \rightarrow E^n$ gilt

$$(x|Ay) = (A^T x|y),$$

da

$$(A^T x|y) = (A^T x)^T y = x^T A^{TT} y = x^T A y = (x|Ay).$$

Diese Konstruktion lässt sich ebenfalls auf den allgemeinen Fall übertragen: \mathcal{V} und \mathcal{W} seien euklidische Vektorräume mit dem Skalarprodukt $(\cdot|\cdot)_{\mathcal{V}}$, bzw. $(\cdot|\cdot)_{\mathcal{W}}$ und $A : \mathcal{V} \rightarrow \mathcal{W}$ eine lineare Abbildung. Dann ist die zu A **transponierte Abbildung** $A^T : \mathcal{W} \rightarrow \mathcal{V}$ durch

$$(A^T w|v)_{\mathcal{V}} = (w|Av)_{\mathcal{W}} \quad \forall v \in \mathcal{V}, \forall w \in \mathcal{W}$$

definiert. Wieder überzeugt man sich leicht davon, dass A^T wohldefiniert und linear ist.

Eigenschaften der transponierten Abbildung:

1. $(a_1 A_1 + a_2 A_2)^T = a_1 A_1^T + a_2 A_2^T, \quad a_{1,2} \in \mathbb{R}, \quad A_{1,2} \in L(\mathcal{V}, \mathcal{W}).$
2. $A^{TT} = A, \quad A \in L(\mathcal{V}, \mathcal{W}).$
3. $(BA)^T = A^T B^T, \quad A \in L(\mathcal{V}, \mathcal{W}), \quad B \in L(\mathcal{W}, \mathcal{X}).$

2.3 Norm

Im E^n kann die Länge eines Vektors durch

$$\|x\| = \sqrt{x_1^2 + \cdots + x_n^2} = \sqrt{(x|x)}$$

bestimmt werden. Im allgemeinen Fall eines beliebigen euklidischen Vektorraums \mathcal{E} definiert man die **Norm** $\|x\|$ eines Elementes $x \in \mathcal{E}$ durch

$$\|x\| = \sqrt{(x|x)}.$$

Für das Skalarprodukt gilt die **Ungleichung** von **Cauchy-Schwarz**:

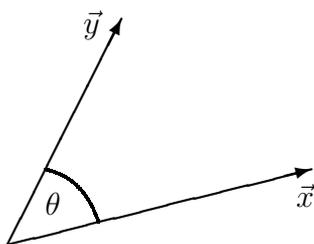
$$|(x|y)| \leq \|x\| \|y\|.$$

Beweis: Für $x = 0$ oder $y = 0$ ist die Ungleichung trivialerweise erfüllt. Man kann sich daher auf den Fall $x \neq 0 \wedge y \neq 0$ beschränken. Es sei $\epsilon = \pm 1$ so gewählt, dass

$$|(x|y)| = \epsilon(x|y).$$

$$\begin{aligned} \Rightarrow 0 &\leq \left(\frac{x}{\|x\|} - \epsilon \frac{y}{\|y\|} \mid \frac{x}{\|x\|} - \epsilon \frac{y}{\|y\|} \right) = \frac{(x|x)}{\|x\|^2} + \frac{(y|y)}{\|y\|^2} - 2 \frac{\epsilon(x|y)}{\|x\| \|y\|} \\ &= 2 \left(1 - \frac{|(x|y)|}{\|x\| \|y\|} \right) \Rightarrow |(x|y)| \leq \|x\| \|y\|. \end{aligned}$$

Anschauliche Bedeutung der Ungleichung von Cauchy-Schwarz in der Ebene oder im dreidimensionalen Raum:



$$|\vec{x} \cdot \vec{y}| = |\vec{x}| |\vec{y}| \underbrace{|\cos \theta|}_{\leq 1} \leq |\vec{x}| |\vec{y}|$$

Eigenschaften der Norm:

1. $\|x\| \geq 0, \quad \|x\| = 0 \Leftrightarrow x = 0$
2. $\|ax\| = |a| \|x\| \quad \forall a \in \mathbb{R}$
3. $\|x + y\| \leq \|x\| + \|y\|$

Beweis: Die ersten beiden Eigenschaften folgen unmittelbar aus der Definition der Norm. Die dritte Eigenschaft (**Dreiecksungleichung**) ergibt sich aus der Ungleichung von Cauchy-Schwarz:

$$\begin{aligned}\|x + y\|^2 &= (x + y|x + y) = (x|x) + (y|y) + 2(x|y) \\ &\leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| = (\|x\| + \|y\|)^2.\end{aligned}$$

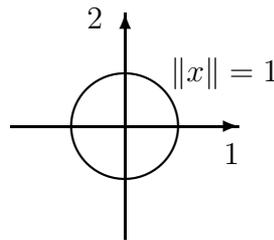
Definition: Ein Vektor x heißt **normiert** oder **Einheitsvektor**, wenn $\|x\| = 1$ ist. Zwei Vektoren x, y heißen **orthogonal**, wenn $(x|y) = 0$ ist.

Beispiele:

1. Im E^2 bildet die Menge aller Einheitsvektoren

$$\|x\|^2 = (x|x) = x_1^2 + x_2^2 = 1$$

einen Kreis mit Radius 1 (Einheitskreis).



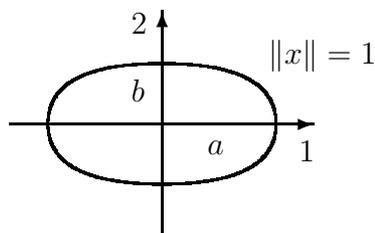
2. Im Fall des \mathbb{R}^2 mit dem inneren Produkt

$$(x|y) = \frac{x_1 y_1}{a^2} + \frac{x_2 y_2}{b^2}$$

liegen die Einheitsvektoren auf der Ellipse

$$\|x\|^2 = (x|x) = \frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1$$

(Ellipse als „Einheitskugel“). Die Richtungsvektoren zweier Durchmesser



sind genau dann „orthogonal“, wenn es sich um konjugierte Durchmesser handelt.

3. Wir betrachten in dem Vektorraum der Polynome vom Grad $\leq n$ auf $[0, 1]$ mit dem inneren Produkt

$$(p|q) = \int_0^1 dx p(x)q(x)$$

die Elemente $p_0(x) = 1$, $p_1(x) = x$ und $p_2(x) = x^2$. p_0 , $\sqrt{3}p_1$ und $\sqrt{5}p_2$ sind Einheitsvektoren, p_0 und p_1 sind nicht orthogonal, jedoch sind p_0 und $p_0 - 2p_1$ orthogonal.

2.4 Orthonormalbasis

Definition: Unter einem **Orthonormalsystem** (ONS) in einem euklidischen Vektorraum \mathcal{E} versteht man eine Menge von Einheitsvektoren, die paarweise aufeinander orthogonal stehen:

$$f_1, \dots, f_k \in \mathcal{E}, \quad (f_i|f_j) = \delta_{ij}.$$

Bildet $\{f_1, \dots, f_k\}$ ein ONS, dann sind diese Vektoren linear unabhängig, denn

$$\sum_{i=1}^k a_i f_i = 0 \quad \Rightarrow \quad 0 = (f_j | \sum_{i=1}^k a_i f_i) = \sum_{i=1}^k a_i \underbrace{(f_j | f_i)}_{\delta_{ij}} = a_j \quad \forall j, 1 \leq j \leq k.$$

Definition: Ein Orthonormalsystem heißt **vollständig** (VONS) oder **Orthonormalbasis** (ONB) von \mathcal{E} , wenn es eine Basis von \mathcal{E} bildet.

Beispiele:

1. Im E^n ist die Standardbasis

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \quad e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

eine ONB, weil $(e_i|e_j) = \delta_{ij}$ ist.

2. Im E^2 bilden die Vektoren

$$f_1 = \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}, \quad f_2 = \begin{pmatrix} -\sin \varphi \\ \cos \varphi \end{pmatrix}$$

eine ONB.

$\{f_1, \dots, f_n\}$ sei eine ONB eines n -dimensionalen euklidischen Vektorraumes \mathcal{E} . Dann lässt sich jedes $x \in \mathcal{E}$ in der Form

$$x = \sum_{i=1}^n f_i a_i$$

schreiben, wobei die Entwicklungskoeffizienten $a_i \in \mathbb{R}$ eindeutig durch

$$a_i = (f_i | x) = f_i^T x$$

bestimmt sind. Man hat also die Beziehung

$$x = \sum_{i=1}^n f_i f_i^T x \quad \forall x \in \mathcal{E}$$

und erhält somit die **Vollständigkeitsrelation**

$$\sum_{i=1}^n f_i f_i^T = \mathbb{1}.$$

Es ist instruktiv, die Vollständigkeitsrelation für die beiden oben angegebenen Beispiele zu verifizieren.

Eine ONB $\{f_1, \dots, f_n\}$ eines n -dimensionalen euklidischen Vektorraumes kann also auch durch die beiden Eigenschaften

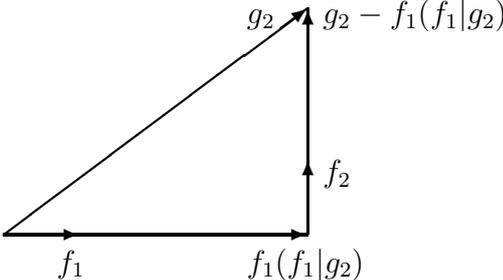
$$(f_i | f_j) = \delta_{ij}, \quad \sum_{i=1}^n f_i f_i^T = \mathbb{1}$$

charakterisiert werden.

Das **Gram-Schmidtsche Orthonormalisierungsverfahren** ist eine Methode, ausgehend von einer beliebigen Basis $\{g_1, \dots, g_n\}$ eines euklidischen Vektorraumes, eine ONB (ein VONS) $\{f_1, \dots, f_n\}$ zu konstruieren.

$$f_1 = \frac{g_1}{\|g_1\|}$$

$$g_2 = \underbrace{f_1(f_1 | g_2)}_{\|f_1\|} + \underbrace{g_2 - f_1(f_1 | g_2)}_{\perp f_1}$$

$$\Rightarrow f_2 = \frac{g_2 - f_1(f_1|g_2)}{\|g_2 - f_1(f_1|g_2)\|}$$


$$g_3 = \underbrace{f_1(f_1|g_3) + f_2(f_2|g_3)}_{\substack{\text{liegt in dem von } f_1, f_2 \\ \text{aufgespannten linearen Teilraum}}} + \underbrace{g_3 - f_1(f_1|g_3) - f_2(f_2|g_3)}_{\perp f_1, f_2}$$

$$\Rightarrow f_3 = \frac{g_3 - f_1(f_1|g_3) - f_2(f_2|g_3)}{\|g_3 - f_1(f_1|g_3) - f_2(f_2|g_3)\|}$$

usw. \rightarrow VONS $\{f_1, \dots, f_n\}$.

Bemerkung: Aus diesem Resultat folgt auch unmittelbar, dass ein VONS stets existiert.

Ein beliebiger n -dimensionaler euklidischer Vektorraum \mathcal{E} unterscheidet sich in gewisser Weise nicht von einem E^n . Wählt man nämlich in \mathcal{E} eine ONB $\{f_1, \dots, f_n\}$, so lässt sich jedes $x \in V$ eindeutig in der Form

$$x = \sum_{i=1}^n f_i x_i, \quad x_i = (f_i|x)$$

schreiben. Der Vektor x ist also (bei vorgegebener ONB) eindeutig durch die n Zahlen

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in E^n$$

bestimmt. Weiters gilt $(x+y)_i = x_i + y_i$, $(ax)_i = ax_i$ und

$$(x|y) = \left(\sum_{i=1}^n f_i x_i \mid \sum_{j=1}^n f_j y_j \right) = \sum_{i,j=1}^n x_i y_j \underbrace{(f_i|f_j)}_{\delta_{ij}} = \sum_{i=1}^n x_i y_i,$$

d.h. die Zuordnung

$$x \in \mathcal{E} \leftrightarrow \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in E^n$$

respektiert sowohl die Vektoraddition, die Multiplikation mit Skalaren als auch die Bildung des Skalarprodukts. Man sagt auch, dass jeder n -dimensionale euklidische Vektorraum **isomorph** zu E^n ist.

Bemerkung: Eine lineare Abbildung $T : \mathcal{V} \rightarrow \mathcal{W}$ heißt **Isomorphismus** der euklidischen Vektorräume \mathcal{V}, \mathcal{W} , falls T ein Vektorraumisomorphismus (d.h. eine bijektive lineare Abbildung) ist mit der zusätzlichen Eigenschaft

$$(Tx|Ty)_{\mathcal{W}} = (x|y)_{\mathcal{V}} \quad \forall x, y \in \mathcal{V}.$$

Die **Matrixdarstellung** (A_{ij}) einer linearen Abbildung $A \in L(\mathcal{E})$ (d.h. $A : \mathcal{E} \rightarrow \mathcal{E}$) in einem euklidischen Vektorraum \mathcal{E} hat **bezüglich einer ONB** $\{f_1, \dots, f_n\}$ eine besonders einfache Form. Bildet man nämlich das Skalarprodukt von f_i mit dem Ausdruck

$$Af_j = \sum_{k=1}^n f_k A_{kj},$$

so erhält man

$$(f_i|Af_j) = (f_i|\sum_{k=1}^n f_k A_{kj}) = \sum_{k=1}^n A_{kj} \underbrace{(f_i|f_k)}_{\delta_{ik}} = A_{ij}.$$

Das heißt man bildet einfach ein „Sandwich“ des Operators A zwischen den Vektoren f_i und f_j um das Matrixelement A_{ij} zu erhalten:

$$A_{ij} = (f_i|Af_j) = f_i^T Af_j.$$

Die Matrixdarstellung $(A^T)_{ij}$ der durch $(A^T x|y) = (x|Ay) \quad \forall x, y \in \mathcal{E}$ definierten, zu A transponieren Abbildung $A^T : \mathcal{E} \rightarrow \mathcal{E}$ ist dann **bezüglich einer ONB** durch

$$(A^T)_{ij} = (f_i|A^T f_j) = (Af_i|f_j) = (f_j|Af_i) = A_{ji}$$

gegeben, d.h. Zeilen und Spalten werden einfach vertauscht.

2.5 Orthogonale Transformationen

Wir wollen nun mit den sogenannten **orthogonalen Abbildungen** noch kurz eine wichtige Klasse von Operatoren $\in L(\mathcal{E})$ besprechen. Dazu gehen wir von der folgenden Beobachtung aus:

Satz: Sei $O \in L(\mathcal{E})$ (\mathcal{E} ein euklidischer Vektorraum). Dann sind die folgenden Aussagen äquivalent:

1. O ist invertierbar mit $O^{-1} = O^T$ (d.h. $O^T O = O O^T = \mathbb{1}$).
2. $(Ox|Oy) = (x|y) \quad \forall x, y \in \mathcal{E}$.
3. O ist **isometrisch**, d.h. $\|Ox\| = \|x\| \quad \forall x \in \mathcal{E}$.
4. O bildet jede ONB von \mathcal{E} wieder auf eine ONB von \mathcal{E} ab.

Ist eine dieser Bedingungen erfüllt, nennt man die Abbildung O **orthogonal**. In der früher eingeführten Sprechweise ist O dann ein Isomorphismus des euklidischen Vektorraumes \mathcal{E} mit sich, ein sogenannter **Automorphismus**.

Bei dem Beweis des obigen Satzes bereitet nur die Implikation $3 \Rightarrow 2$ etwas Mühe. Man benötigt dazu den folgenden **Hilfssatz (Polarisierungsidentität)**:

$$(x|y) = \frac{\|x + y\|^2 - \|x - y\|^2}{4}.$$

Beweis des Hilfssatzes: Man subtrahiere die Ausdrücke

$$\begin{aligned} (x + y|x + y) &= \|x\|^2 + \|y\|^2 + 2(x|y), \\ (x - y|x - y) &= \|x\|^2 + \|y\|^2 - 2(x|y). \end{aligned}$$

Nimmt man nun an, dass $\|Ox\| = \|x\| \quad \forall x \in \mathcal{E}$ erfüllt ist, so folgt unter Verwendung des Hilfssatzes tatsächlich, dass

$$\begin{aligned} (Ox|Oy) &= \frac{\|Ox + Oy\|^2 - \|Ox - Oy\|^2}{4} = \frac{\|O(x + y)\|^2 - \|O(x - y)\|^2}{4} \\ &= \frac{\|(x + y)\|^2 - \|(x - y)\|^2}{4} = (x|y) \quad \forall x, y \in \mathcal{E} \end{aligned}$$

erfüllt ist.

Bemerkungen:

1. Aus $O^T O = \mathbb{1}$ folgt für eine orthogonale Transformation $\det(O^T O) = \det O^T \det O = (\det O)^2 = 1$ und somit $\det O = \pm 1$.
2. Die orthogonalen Transformationen auf einem n -dimensionalen euklidischen Vektorraum bilden eine **Gruppe**, man nennt sie $O(n)$. Jene orthogonalen Transformationen R , welche die Zusatzbedingung $\det R = 1$ erfüllen, bilden ebenfalls eine Gruppe, diese wird als $SO(n)$ bezeichnet.

3. $\{f_1, \dots, f_n\}$ sei eine beliebige ONB eines euklidischen Vektorraumes \mathcal{E} . Wegen $\sum_{i=1}^n f_i f_i^T = \mathbb{1}$ lässt sich dann jede orthogonale Transformation $O \in L(\mathcal{E})$ in der Form

$$O = O\mathbb{1} = O \sum_{i=1}^n f_i f_i^T = \sum_{i=1}^n O f_i f_i^T = \sum g_i f_i^T$$

schreiben, wobei die Vektoren $g_i = O f_i$ ($i = 1, \dots, n$) wegen der Eigenschaft 4 wieder eine ONB von \mathcal{E} bilden. Für den Spezialfall $\mathcal{E} = E^n$ kann man daher jede orthogonale $n \times n$ -Matrix O in der Form

$$O = \sum_{i=1}^n g_i e_i^T = (g_1, \dots, g_n)$$

schreiben, das heißt in den Spalten von O steht die ONB $\{g_1, \dots, g_n\} = \{O e_1, \dots, O e_n\}$. Ebenso bilden die Zeilen einer orthogonalen Matrix ein VONS, da

$$\begin{aligned} O &= \mathbb{1}O = \sum_{i=1}^n e_i e_i^T O = \sum_{i=1}^n e_i (O^T e_i)^T = \sum_{i=1}^n e_i h_i^T \\ &= \begin{pmatrix} h_1^T \\ \vdots \\ h_n^T \end{pmatrix}, \quad h_i = O^T e_i. \end{aligned}$$

Die orthogonalen $n \times n$ -Matrizen sind also genau jene Elemente aus $L(E^n)$, deren Spalten bzw. Zeilen ein VONS von E^n bilden.

Beispiele:

1. Im E^2 lässt sich jeder Einheitsvektor f_1 in der Form

$$f_1 = \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix}$$

schreiben, für einen auf f_1 orthogonal stehenden Vektor f_2 gibt es dann nur mehr die zwei Möglichkeiten

$$f_2 = \varepsilon \begin{pmatrix} -\sin \alpha \\ \cos \alpha \end{pmatrix}, \quad \varepsilon = \pm 1.$$

Nach den obigen Überlegungen hat eine orthogonale Matrix O im E^2 dann die allgemeine Gestalt

$$O = \begin{pmatrix} \cos \alpha & -\varepsilon \sin \alpha \\ \sin \alpha & \varepsilon \cos \alpha \end{pmatrix}.$$

Für $\varepsilon = 1$ handelt es sich um eine **Drehung** mit Drehwinkel α . In diesem Fall ist $\det O = 1$. Im Fall $\varepsilon = -1$ hat man es mit einer **Drehspiegelung** zu tun, wobei $\det O = -1$ ist.

2. Eine reine **Spiegelung** im E^2 wäre z.B. die Transformation

$$S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

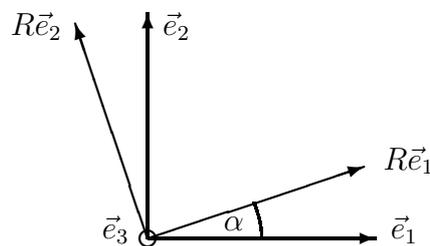
die einer Spiegelung an der 1-Achse entspricht.

3. Wir wollen im dreidimensionalen Raum eine Drehung um den Winkel α beschreiben. Die 3-Achse unseres Koordinatensystems legen wir dabei in Richtung der Drehachse. Als ONB haben wir (in der in der Physik für räumliche Vektoren üblichen Schreibweise) $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$, welche (in der angegebenen Reihenfolge) ein Rechtssystem bilden soll. Wegen der Linearität der Drehoperation R genügt es, ihre Wirkung auf die drei Basisvektoren zu kennen:

$$\begin{aligned} R\vec{e}_1 &= \vec{e}_1 \cos \alpha + \vec{e}_2 \sin \alpha, \\ R\vec{e}_2 &= -\vec{e}_1 \sin \alpha + \vec{e}_2 \cos \alpha, \\ R\vec{e}_3 &= \vec{e}_3. \end{aligned}$$

Daraus können wir unmittelbar die Matrixdarstellung (R_{ij}) von R bezüglich der gewählten ONB ablesen:

$$(R_{ij}) = \begin{pmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}$$



Bemerkungen:

- (a) Wir stellen fest, dass $\det R = 1$ ist. Reine Drehungen im dreidimensionalen Raum entsprechen den Elementen der **Drehgruppe** $SO(3)$.

- (b) Eine Drehung im dreidimensionalen Raum ist bereits festgelegt, wenn man das Bild von **zwei** orthonormierten Vektoren kennt, da dann auch das Bild des dritten Basisvektors bereits bestimmt ist. Tatsächlich folgt bei unserem Beispiel aus $R\vec{e}_2 = -\vec{e}_1 \sin \alpha + \vec{e}_2 \cos \alpha$ und $R\vec{e}_3 = \vec{e}_3$ mit $\vec{e}_1 = \vec{e}_2 \times \vec{e}_3$ das Transformationsverhalten von \vec{e}_1 :

$$\begin{aligned} R\vec{e}_1 &= R(\vec{e}_2 \times \vec{e}_3) = R\vec{e}_2 \times R\vec{e}_3 \\ &= (-\vec{e}_1 \sin \alpha + \vec{e}_2 \cos \alpha) \times \vec{e}_3 \\ &= \vec{e}_1 \cos \alpha + \vec{e}_2 \sin \alpha. \end{aligned}$$

- (c) Ein Vektor

$$\vec{x} = \sum_{i=1}^3 \vec{e}_i x_i$$

mit Komponenten x_i bezüglich der ONB $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ wird durch die Drehung $R \in SO(3)$ in den Vektor

$$R\vec{x} = R \sum_{j=1}^3 \vec{e}_j x_j = \sum_{j=1}^3 R\vec{e}_j x_j = \sum_{i,j=1}^3 \vec{e}_i R_{ij} x_j$$

mit den Komponenten

$$\sum_{j=1}^3 R_{ij} x_j$$

transformiert. Man spricht in diesem Fall von einer **aktiven** Transformation des Vektors. Bei einer **passiven** Transformation bleibt der Vektor \vec{x} selbst unverändert, man betrachtet jedoch seine Komponenten x'_i bezüglich einer transformierten ONB $\{\vec{e}'_1, \vec{e}'_2, \vec{e}'_3\}$. Im Fall unserer Drehung ist

$$\vec{e}'_j = R\vec{e}_j = \sum_{i=1}^3 \vec{e}_i R_{ij}.$$

Wegen

$$\vec{x} = \sum_{i=1}^3 \vec{e}_i x_i = \sum_{j=1}^3 \vec{e}'_j x'_j = \sum_{i,j=1}^3 \vec{e}_i R_{ij} x'_j,$$

ist der Zusammenhang zwischen den Komponenten x_i und den Komponenten x'_i durch

$$x_i = \sum_{j=1}^3 R_{ij} x'_j \quad \Leftrightarrow \quad x'_i = \sum_{j=1}^3 R_{ij}^{-1} x_j$$

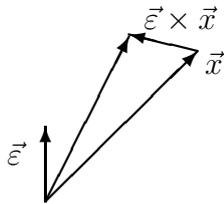
gegeben. Das Transformationsverhalten der Komponenten des Vektors ist also invers zu jenem der aktiven Transformation.

4. Das Standardbeispiel für eine orthogonale Transformation mit negativer Determinante ist in drei Raumdimensionen die **Paritätstransformation** $P = -\mathbb{1}$. Jedes Element $O \in O(3)$ mit $\det O = -1$ lässt sich in der Form $O = PR$, $R \in SO(3)$, schreiben.

2.6 $SO(3)$

Die Menge der Drehungen im dreidimensionalen Raum ist ein Beispiel für eine nichtabelsche Liegruppe.

Die Elemente der Gruppe $SO(3)$ können durch den sogenannten **Drehvektor** $\vec{\alpha}$ parametrisiert werden. Dieser zeigt (gemäß der Rechtsschraubenregel) in Richtung der Drehachse, sein Betrag ist gleich dem Drehwinkel $\alpha = |\vec{\alpha}|$ ($0 \leq \alpha < 2\pi$). Die entsprechende Transformation bezeichnen wir mit $R(\vec{\alpha})$. Es handelt sich bei der $SO(3)$ also um eine kontinuierliche Gruppe von Transformationen (eine sogenannte **Liegruppe**), deren wesentliche Eigenschaften durch die Gestalt ihrer Elemente in einer infinitesimalen Umgebung des Einheitselements bereits festgelegt sind. Für infinitesimales $\vec{\varepsilon}$ hat eine Drehung die einfache Form



$$R(\vec{\varepsilon}) \vec{x} = \vec{x} + \vec{\varepsilon} \times \vec{x} = (\mathbb{1} + \vec{\varepsilon} \cdot \vec{\Lambda}) \vec{x}.$$

Man kann nun eine beliebige **endliche** Drehung um den Drehwinkel $\vec{\alpha}$ durch die wiederholte Anwendung von N Drehungen um den Drehwinkel $\vec{\alpha}/N$ erhalten:

$$R(\vec{\alpha}) = R(\vec{\alpha}/N)^N.$$

Für genügend großes N (und somit einen entsprechend kleinen Wert von α/N) kann man schreiben:

$$R(\vec{\alpha}) \simeq \left(\mathbb{1} + \frac{\vec{\alpha} \cdot \vec{\Lambda}}{N} \right)^N$$

Im Limes $N \rightarrow \infty$ erhält man somit die Formel

$$R(\vec{\alpha}) = \exp(\vec{\alpha} \cdot \vec{\Lambda}).$$

Bemerkung: Wegen $(\vec{\varepsilon} \times \vec{x})_i = \epsilon_{ikj} \varepsilon_k x_j = -\varepsilon_k \epsilon_{kij} x_j$ (Summenkonvention!) sind die Matrixdarstellungen der drei in $\vec{\Lambda}$ zusammengefassten linearen Abbildungen

$\Lambda_1, \Lambda_2, \Lambda_3$ bezüglich der ONB $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ durch $(\Lambda_k)_{ij} = -\epsilon_{kij}$ gegeben. Diese Matrizen bilden eine dreidimensionale Darstellung der **Generatoren** (oder **Erzeuger**) der $SO(3)$.

Führt man zuerst die Drehung $R(\vec{\alpha})$ und anschließend die Drehung $R(\vec{\beta})$ aus, so wird sich das Resultat im Allgemeinen von jenem unterscheiden, das man erhält, wenn man die beiden Operationen in der umgekehrten Reihenfolge durchführt. Es handelt sich bei der $SO(3)$ nämlich um eine **nichtabelsche** (oder **nicht-kommutative**) Gruppe, das heißt ihre Elemente vertauschen (kommutieren) im Allgemeinen **nicht**:

$$R(\vec{\alpha})R(\vec{\beta}) \stackrel{i.A.}{\neq} R(\vec{\beta})R(\vec{\alpha}).$$

Konkret bedeutet das im Fall der $SO(3)$, dass

$$R(R(\vec{\beta})\vec{\alpha}) = R(\vec{\beta})R(\vec{\alpha})R(\vec{\beta})^{-1}.$$

Für infinitesimales $\vec{\alpha} = \vec{\varepsilon}$ ist diese Relation leicht zu sehen:

$$\begin{aligned} R(\vec{\beta})R(\vec{\varepsilon})R(\vec{\beta})^{-1}\vec{x} &= R(\vec{\beta})[R(\vec{\beta})^{-1}\vec{x} + \vec{\varepsilon} \times R(\vec{\beta})^{-1}\vec{x}] \\ &= \vec{x} + (R(\vec{\beta})\vec{\varepsilon}) \times \vec{x} \\ &= R(R(\vec{\beta})\vec{\varepsilon})\vec{x}. \end{aligned}$$

Im Fall eines endlichen Drehwinkels $\vec{\alpha}$ schreiben wir

$$R(\vec{\beta})R(\vec{\alpha})R(\vec{\beta})^{-1} = R(\vec{\beta})R(\vec{\alpha}/N)^N R(\vec{\beta})^{-1} = [R(\vec{\beta})R(\vec{\alpha}/N)R(\vec{\beta})^{-1}]^N.$$

Für genügend großes N können wir die für den infinitesimalen Drehwinkel $\vec{\alpha}/N$ bereits gezeigte Formel

$$R(\vec{\beta})R(\vec{\alpha}/N)R(\vec{\beta})^{-1} = R(R(\vec{\beta})\vec{\alpha}/N)$$

einsetzen und erhalten auf diese Weise das behauptete Ergebnis.

Bemerkung: Stimmen die durch $\vec{\alpha}$ und $\vec{\beta}$ beschriebenen Drehachsen überein, dann ist $R(\vec{\beta})\vec{\alpha} = \vec{\alpha}$ und die beiden Drehungen $R(\vec{\alpha}), R(\vec{\beta})$ kommutieren in diesem Fall.

2.7 Eulersche Winkel

Jedes Element der dreiparametrischen Liegruppe $SO(3)$ lässt sich eindeutig als Produkt von drei Drehungen um drei bestimmte Standardachsen beschreiben.

Eine andere Möglichkeit zur Parametrisierung der Elemente der $SO(3)$ stellen die drei **Eulerschen Winkel** dar. Sei $D \in SO(3)$ und $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ eine (rechtshändige) ONB. Ist $D\vec{e}_3 = \vec{e}_3$, dann handelt es sich um den bereits behandelten Fall

einer Drehung um die 3-Achse. Es sei also $D\vec{e}_3 \neq \vec{e}_3$. Die beiden durch den Ursprung gehenden und auf \vec{e}_3 , beziehungsweise auf $D\vec{e}_3$ normal stehenden Ebenen schneiden einander dann in einer Geraden, der sogenannten **Knotenlinie**. Wie auch bei Verwendung von Kugelkoordinaten üblich, bezeichnen wir den Winkel zwischen \vec{e}_3 und $D\vec{e}_3$ mit θ ($0 \leq \theta \leq \pi$) und den Winkel zwischen \vec{e}_1 und der Projektion von $D\vec{e}_3$ auf den von \vec{e}_1 und \vec{e}_2 aufgespannten Unterraum mit ϕ ($0 \leq \phi < 2\pi$).

Als ersten Schritt wenden wir eine Drehung

$$R_\phi = R(\phi\vec{e}_3)$$

mit der Drehachse \vec{e}_3 um den Winkel ϕ an. Diese bildet die ONB $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ auf die von den Vektoren

$$\vec{e}'_j = R_\phi \vec{e}_j = \sum_{i=1}^3 \vec{e}_i R_{\phi,ij}, \quad j = 1, \dots, 3,$$

gebildete ONB $\{\vec{e}'_1, \vec{e}'_2, \vec{e}'_3\}$ ab, wobei $(R_{\phi,ij})$ die Matrixdarstellung von R_ϕ bezüglich der ONB $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ bezeichnet. Die explizite Form dieser Transformation ist durch

$$\begin{aligned} \vec{e}'_1 &= R_\phi \vec{e}_1 = \vec{e}_1 \cos \phi + \vec{e}_2 \sin \phi, \\ \vec{e}'_2 &= R_\phi \vec{e}_2 = -\vec{e}_1 \sin \phi + \vec{e}_2 \cos \phi, \\ \vec{e}'_3 &= R_\phi \vec{e}_3 = \vec{e}_3 \end{aligned}$$

mit

$$(R_{\phi,ij}) = \begin{pmatrix} \cos \phi & -\sin \phi & 0 \\ \sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

gegeben. Der Vektor \vec{e}'_2 liegt in der Knotenlinie.

Im zweiten Schritt führen wir eine Drehung

$$R_\theta = R(\theta\vec{e}'_2)$$

mit der Drehachse \vec{e}'_2 um den Winkel θ durch. Die Wirkung dieser Drehung auf die gestrichene ONB $\{\vec{e}'_1, \vec{e}'_2, \vec{e}'_3\}$,

$$\vec{e}''_j = R_\theta \vec{e}'_j = \sum_{i=1}^3 \vec{e}'_i R'_{\theta,ij}, \quad j = 1, \dots, 3,$$

ist durch

$$\begin{aligned} \vec{e}''_1 &= R_\theta \vec{e}'_1 = \vec{e}'_1 \cos \theta - \vec{e}'_3 \sin \theta, \\ \vec{e}''_2 &= R_\theta \vec{e}'_2 = \vec{e}'_2 \\ \vec{e}''_3 &= R_\theta \vec{e}'_3 = \vec{e}'_1 \sin \theta + \vec{e}'_3 \cos \theta \end{aligned}$$

gegeben. Die Matrixdarstellung $(R'_{\theta,ij})$ der Drehung R_θ bezüglich der **gestrichelten** Basis liest man sofort als

$$(R'_{\theta,ij}) = \begin{pmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{pmatrix}$$

ab.

Da \vec{e}_3'' bereits mit $D\vec{e}_3$ übereinstimmt, benötigt man als dritte Drehung nur mehr eine Rotation um \vec{e}_3'' um den Winkel ψ ($0 \leq \psi < 2\pi$),

$$R_\psi = R(\psi\vec{e}_3''),$$

um die Wirkung der gesamten Drehung D als Produkt

$$D = R_\psi R_\theta R_\phi = R(\psi\vec{e}_3'') R(\theta\vec{e}_2') R(\phi\vec{e}_3)$$

darzustellen. Wir betrachten also die Wirkung der Transformation R_ψ auf die Orthonormalbasis $\{\vec{e}_1'', \vec{e}_2'', \vec{e}_3''\}$,

$$D\vec{e}_j = R_\psi \vec{e}_j'' = \sum_{i=1}^3 \vec{e}_i'' R_{\psi,ij}'', \quad j = 1, \dots, 3.$$

In expliziter Form hat man

$$\begin{aligned} D\vec{e}_1 &= R_\psi \vec{e}_1'' = \vec{e}_1'' \cos \psi + \vec{e}_2'' \sin \psi, \\ D\vec{e}_2 &= R_\psi \vec{e}_2'' = -\vec{e}_1'' \sin \psi + \vec{e}_2'' \cos \psi, \\ D\vec{e}_3 &= R_\psi \vec{e}_3'' = \vec{e}_3'', \end{aligned}$$

und die Matrixdarstellung $(R''_{\psi,ij})$ von R_ψ bezüglich der ONB $\{\vec{e}_1'', \vec{e}_2'', \vec{e}_3''\}$ lautet daher:

$$(R''_{\psi,ij}) = \begin{pmatrix} \cos \psi & -\sin \psi & 0 \\ \sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Um nun die Matrixdarstellung (D_{ij}) der Gesamtdrehung bezüglich der ursprünglichen ONB $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ zu erhalten, setzen wir diese aus den drei Einzeldrehungen zusammen:

$$\begin{aligned} D\vec{e}_j &= \sum_{i=1}^3 \vec{e}_i D_{ij} = R_\psi \vec{e}_j'' = \sum_{n=1}^3 \vec{e}_n'' R_{\psi,nj}'' \\ &= \sum_{n=1}^3 R_\theta \vec{e}_n' R_{\psi,nj}'' = \sum_{m,n=1}^3 \vec{e}_m' R'_{\theta,mn} R_{\psi,nj}'' \\ &= \sum_{m,n=1}^3 R_\phi \vec{e}_m R'_{\theta,mn} R_{\psi,nj}'' = \sum_{i,m,n=1}^3 \vec{e}_i R_{\phi,im} R'_{\theta,mn} R_{\psi,nj}'' \end{aligned}$$

Das heißt also, dass

$$D_{ij} = (R_\psi R_\theta R_\phi)_{ij} = \sum_{m,n=1}^3 R_{\phi,im} R'_{\theta,mn} R''_{\psi,nj},$$

bzw.

$$(D_{ij}) = \begin{pmatrix} \cos \phi & -\sin \phi & 0 \\ \sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{pmatrix} \begin{pmatrix} \cos \psi & -\sin \psi & 0 \\ \sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Wie aus diesem Ergebnis ersichtlich ist, lässt sich die Drehung D auch auf eine andere Art erhalten: Man dreht zunächst um \vec{e}_3 um den Winkel ψ , dann um die Drehachse \vec{e}_2 (**kein** Strich!) um den Winkel θ und schließlich wieder um \vec{e}_3 um den Winkel ϕ . Mit anderen Worten, es gilt die Beziehung

$$D = R(\psi \vec{e}_3'') R(\theta \vec{e}_2') R(\phi \vec{e}_3) = R(\phi \vec{e}_3) R(\theta \vec{e}_2) R(\psi \vec{e}_3).$$

Die zweite Form hat den Vorteil, dass nur um Achsen der **ursprünglichen** ONB gedreht wird.

Die hier gefundene Beziehung ist natürlich kein Zufall, sondern hat ihren tieferen Grund in der schon früher besprochenen Formel

$$R(R(\vec{\beta})\vec{\alpha}) = R(\vec{\beta})R(\vec{\alpha})R(\vec{\beta})^{-1}.$$

Um dies zu sehen, setzen wir $\vec{e}_2' = R(\phi \vec{e}_3)\vec{e}_2$ und

$$\begin{aligned} \vec{e}_3'' &= R(\theta \vec{e}_2')\vec{e}_3' = R(R(\phi \vec{e}_3)\theta \vec{e}_2)R(\phi \vec{e}_3)\vec{e}_3 = R(\phi \vec{e}_3)R(\theta \vec{e}_2) \underbrace{R(\phi \vec{e}_3)^{-1}R(\phi \vec{e}_3)}_{\mathbb{1}} \vec{e}_3 \\ &= R(\phi \vec{e}_3)R(\theta \vec{e}_2)\vec{e}_3 \end{aligned}$$

in $R(\psi \vec{e}_3'')R(\theta \vec{e}_2')R(\phi \vec{e}_3)$ ein:

$$\begin{aligned} & \underbrace{R(R(\phi \vec{e}_3)R(\theta \vec{e}_2)\psi \vec{e}_3)}_{R(\psi \vec{e}_3'')} \underbrace{R(R(\phi \vec{e}_3)\theta \vec{e}_2)}_{R(\theta \vec{e}_2')} R(\phi \vec{e}_3) \\ &= R(\phi \vec{e}_3)R(\theta \vec{e}_2)R(\psi \vec{e}_3)R(\theta \vec{e}_2)^{-1} \underbrace{R(\phi \vec{e}_3)^{-1}R(\phi \vec{e}_3)}_{\mathbb{1}} R(\theta \vec{e}_2) \underbrace{R(\phi \vec{e}_3)^{-1}R(\phi \vec{e}_3)}_{\mathbb{1}} \\ &= R(\phi \vec{e}_3)R(\theta \vec{e}_2)R(\psi \vec{e}_3) \underbrace{R(\theta \vec{e}_2)^{-1}R(\theta \vec{e}_2)}_{\mathbb{1}} \\ &= R(\phi \vec{e}_3)R(\theta \vec{e}_2)R(\psi \vec{e}_3). \end{aligned}$$

Bemerkung: Die hier gewählte Konvention zur Definition der Eulerschen Winkel hat in der Quantenmechanik gewisse Vorteile. In der klassischen Mechanik wird oft eine Konvention verwendet, bei der zunächst der Vektor \vec{e}_1 durch eine Rotation um \vec{e}_3 in die Knotenlinie gedreht wird, anschließend wird um \vec{e}_1' gedreht und schließlich um \vec{e}_3'' . Physikalische Resultate können natürlich von der Wahl einer bestimmten Konvention nicht abhängen.

Kapitel 3

Unitäre Vektorräume

Die Verwendung des Grundkörpers der komplexen Zahlen besitzt eine Reihe von rechentechnischen Vorteilen. Will man etwa die Eigenwerte einer $n \times n$ -Matrix bestimmen, so besitzt das dazugehörige charakteristische Polynom über \mathbb{C} stets n (eventuell teilweise zusammenfallende) Nullstellen, was über \mathbb{R} ja bekanntlich nicht der Fall zu sein braucht. Darüber hinaus ist die Theorie der unitären Vektorräume in der Physik von allergrößter Bedeutung, handelt es sich dabei doch um jene mathematischen Methoden, die in der Quantentheorie ihre natürliche Anwendung finden. Wir beschränken uns hier ausschließlich auf den endlichdimensionalen Fall. Es sei jedoch darauf hingewiesen, dass die hier dargestellte Theorie ihre volle Bedeutung erst im unendlichdimensionalen Fall erlangt, wo sie im Rahmen der Funktionalanalysis in die Theorie der (unendlichdimensionalen) Hilberträume übergeht.

3.1 Komplexes Skalarprodukt

Definition: Ein (endlichdimensionaler) Vektorraum \mathcal{H} über \mathbb{C} heißt **unitär** (oder endlichdimensionaler Hilbertraum), wenn $\forall \varphi, \psi \in \mathcal{H}$ ein **Skalarprodukt** (**inneres Produkt**) $\langle \varphi | \psi \rangle \in \mathbb{C}$ definiert ist mit den folgenden Eigenschaften:

$$(U1) \quad \langle \varphi | c_1\psi_1 + c_2\psi_2 \rangle = c_1\langle \varphi | \psi_1 \rangle + c_2\langle \varphi | \psi_2 \rangle, \quad \forall \varphi, \psi_{1,2} \in \mathcal{H}, \quad \forall c_{1,2} \in \mathbb{C}$$

$$(U2) \quad \langle \varphi | \psi \rangle = \langle \psi | \varphi \rangle^*$$

$$(U3) \quad \langle \psi | \psi \rangle \geq 0, \quad \langle \psi | \psi \rangle = 0 \Leftrightarrow \psi = 0$$

Bemerkungen: $\langle \varphi | \psi \rangle$ ist **sesquilinear**, d.h. linear in der zweiten Variablen

und antilinear in der ersten Variablen, da

$$\begin{aligned}\langle c_1\varphi_1 + c_2\varphi_2 \mid \psi \rangle &= \langle \psi \mid c_1\varphi_1 + c_2\varphi_2 \rangle^* \\ &= c_1^* \langle \psi \mid \varphi_1 \rangle^* + c_2^* \langle \psi \mid \varphi_2 \rangle^* \\ &= c_1^* \langle \varphi_1 \mid \psi \rangle + c_2^* \langle \varphi_2 \mid \psi \rangle.\end{aligned}$$

Wir verwenden hier die in der physikalischen Literatur üblichen Konventionen. In der mathematischen Literatur ist das innere Produkt oft in der ersten Variablen linear und in der zweiten antilinear. Weiters findet man dort häufig die Schreibweise \bar{z} statt z^* für die zur komplexen Zahl $z = a + ib$ ($a, b \in \mathbb{R}$) konjugiert komplexe Zahl $z^* = a - ib$.

Beispiele für unitäre Vektorräume:

1. $U^n = \mathbb{C}^n$ mit dem Skalarprodukt

$$\langle x \mid y \rangle = \sum_{k=1}^n x_k^* y_k.$$

2. Der Vektorraum, der auf $[a, b] \subset \mathbb{R}$ definierten komplexen Polynome vom Grad $\leq n$,

$$p(x) = \sum_{k=0}^n c_k x^k, \quad c_k \in \mathbb{C}, \quad x \in [a, b]$$

mit dem inneren Produkt

$$\langle p \mid q \rangle = \int_a^b dx p(x)^* q(x).$$

3. Der Vektorraum der komplexwertigen Funktionen der Form

$$\psi(x) = \sum_{n=-N}^N c_n e^{2\pi i n x}, \quad c_n \in \mathbb{C},$$

auf dem Intervall $[0, 1]$ mit dem Skalarprodukt

$$\langle \varphi \mid \psi \rangle = \int_0^1 dx \varphi(x)^* \psi(x).$$

4. Der Vektorraum der komplexwertigen Funktionen der Form

$$\varphi(x) = e^{-x^2/2}p(x)$$

auf \mathbb{R} , wobei $p(x)$ ein Polynom vom Grad $\leq n$ ist, mit dem Skalarprodukt

$$\begin{aligned}\langle \varphi | \psi \rangle &= \int_{-\infty}^{\infty} dx \varphi(x)^* \psi(x) \\ &= \int_{-\infty}^{\infty} dx e^{-x^2} p(x)^* q(x),\end{aligned}$$

wobei $\psi(x) = e^{-x^2/2}q(x)$.

3.2 Norm

In einem unitären Vektorraum \mathcal{H} definiert man die **Norm** $\|\psi\|$ eines Elementes $\psi \in \mathcal{H}$ durch

$$\|\psi\| = \sqrt{\langle \psi | \psi \rangle}.$$

Für das Skalarprodukt gilt die **Ungleichung** von **Cauchy-Schwarz**:

$$|\langle \varphi | \psi \rangle| \leq \|\varphi\| \|\psi\|.$$

Beweis: Für $\varphi = 0$ oder $\psi = 0$ ist die Ungleichung trivialerweise erfüllt. Man kann sich daher auf den Fall $\varphi \neq 0 \wedge \psi \neq 0$ beschränken. Es sei $\varepsilon \in \mathbb{C}$ ($|\varepsilon| = 1$) so gewählt, dass

$$\varepsilon \langle \varphi | \psi \rangle = |\langle \varphi | \psi \rangle| \quad \Rightarrow \quad \varepsilon^* \langle \varphi | \psi \rangle^* = |\langle \varphi | \psi \rangle|.$$

$$\begin{aligned}\Rightarrow 0 &\leq \left\langle \frac{\varphi}{\|\varphi\|} - \varepsilon \frac{\psi}{\|\psi\|} \middle| \frac{\varphi}{\|\varphi\|} - \varepsilon \frac{\psi}{\|\psi\|} \right\rangle = 2 - \frac{\varepsilon \langle \varphi | \psi \rangle}{\|\varphi\| \|\psi\|} - \frac{\varepsilon^* \langle \psi | \varphi \rangle}{\|\varphi\| \|\psi\|} \\ &= 2 \left(1 - \frac{|\langle \varphi | \psi \rangle|}{\|\varphi\| \|\psi\|} \right) \quad \Rightarrow \quad |\langle \varphi | \psi \rangle| \leq \|\varphi\| \|\psi\|.\end{aligned}$$

Eigenschaften der Norm:

$$1. \quad \|\psi\| \geq 0, \quad \|\psi\| = 0 \Leftrightarrow \psi = 0$$

2. $\|a\psi\| = |a|\|\psi\| \quad \forall a \in \mathbb{C}, \forall \psi \in \mathcal{H}$
3. $\|\varphi + \psi\| \leq \|\varphi\| + \|\psi\|$

Beweis: Die ersten beiden Eigenschaften folgen unmittelbar aus der Definition der Norm. Die dritte Eigenschaft (**Dreiecksungleichung**) ergibt sich aus der Ungleichung von Cauchy-Schwarz:

$$\begin{aligned} \|\varphi + \psi\|^2 &= \langle \varphi + \psi | \varphi + \psi \rangle = \langle \varphi | \varphi \rangle + \langle \psi | \psi \rangle + 2\operatorname{Re}\langle \varphi | \psi \rangle \\ &\leq \|\varphi\|^2 + \|\psi\|^2 + 2\|\varphi\|\|\psi\| = (\|\varphi\| + \|\psi\|)^2. \end{aligned}$$

Definition: Ein Vektor ψ heißt **normiert** oder **Einheitsvektor**, wenn $\|\psi\| = 1$ ist. Zwei Vektoren φ, ψ heißen **orthogonal**, wenn $\langle \varphi | \psi \rangle = 0$ ist.

Beispiel: Wir betrachten in dem Vektorraum der komplexen Polynome vom Grad $\leq n$ auf $[0, 1]$ mit dem inneren Produkt

$$\langle p | q \rangle = \int_0^1 dx p(x)^* q(x)$$

die Elemente $p_0(x) = 1$, $p_1(x) = x$ und $p_2(x) = x^2$. p_0 , $\sqrt{3}p_1$ und $\sqrt{5}p_2$ sind Einheitsvektoren, p_0 und p_1 sind nicht orthogonal, jedoch sind p_0 und $p_0 - 2p_1$ orthogonal.

In einem unitären Vektorraum gilt die sogenannte **Polarisierungsidentität**:

$$\langle \psi | \varphi \rangle = \frac{1}{4} \sum_{k=1}^4 i^k \|\varphi + i^k \psi\|^2$$

3.3 Adjungierte Abbildung

Die Rolle der transponierten Matrix A^T wird im komplexen Fall durch die **adjungierte Matrix** $A^\dagger = (A^T)^*$ übernommen. Die Operation $A \in L(U^n, U^m) \rightarrow A^\dagger \in L(U^m, U^n)$ besitzt die folgende Eigenschaften:

$$\begin{aligned} (c_1 A_1 + c_2 A_2)^\dagger &= c_1^* A_1^\dagger + c_2^* A_2^\dagger, \quad c_{1,2} \in \mathbb{C} \\ (AB)^\dagger &= B^\dagger A^\dagger, \quad A^{\dagger\dagger} = A \end{aligned}$$

Fasst man im Raum U^n ein Element

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

als $n \times 1$ -Matrix auf, dann ist

$$x^\dagger = (x_1^*, \dots, x_n^*)$$

und das Skalarprodukt kann in Matrixnotation

$$\langle x | y \rangle = x^\dagger y$$

geschrieben werden. Die Abbildung $x \mapsto x^\dagger$ stellt eine „Identifizierung“ des U^n mit seinem Dualraum \tilde{U}^n dar, allerdings ist diese „Identifizierung“ nicht linear:

$$(c_1 x_1 + c_2 x_2)^\dagger = c_1^* x_1^\dagger + c_2^* x_2^\dagger$$

Das Konzept der adjungierten Abbildung kann nun leicht auf den allgemeinen Fall eines beliebigen unitären Raumes \mathcal{H} erweitert werden. Da wir ein Skalarprodukt zur Verfügung haben, können wir \mathcal{H} mit seinem Dualraum $\tilde{\mathcal{H}}$ auf kanonische Weise „identifizieren“. Dazu ordnen wir jedem $\varphi \in \mathcal{H}$ das Funktional $\varphi^\dagger \in \tilde{\mathcal{H}}$ zu, welches durch

$$\varphi^\dagger \psi = \langle \varphi | \psi \rangle \quad \forall \psi \in \mathcal{H}$$

definiert ist. Diese Abbildung führt linear unabhängige Vektoren wieder in linear unabhängige Vektoren über und ist injektiv. Da die Dimension von \mathcal{H} mit der seines Dualraums übereinstimmt, ist sie auch surjektiv. Die inverse Abbildung bezeichnen wir ebenfalls mit † , sodass $\varphi^{\dagger\dagger} = \varphi$ gilt.

Seien \mathcal{U} und \mathcal{V} unitäre Vektorräume und $A : \mathcal{U} \rightarrow \mathcal{V}$ eine lineare Abbildung. Die zu A **adjungierte Abbildung** $A^\dagger : \mathcal{V} \rightarrow \mathcal{U}$ ist durch

$$\langle A^\dagger \varphi | \psi \rangle_{\mathcal{U}} = \langle \varphi | A \psi \rangle_{\mathcal{V}} \quad \forall \psi \in \mathcal{U}, \forall \varphi \in \mathcal{V}$$

definiert. Diese Definition ist äquivalent mit

$$A^\dagger \varphi = (\varphi^\dagger A)^\dagger \quad \forall \varphi \in \mathcal{V}.$$

Die so definierte Abbildung ist wohldefiniert und linear, da

$$\begin{aligned} A^\dagger(c_1 \varphi_1 + c_2 \varphi_2) &= ((c_1 \varphi_1 + c_2 \varphi_2)^\dagger A)^\dagger \\ &= ((c_1^* \varphi_1^\dagger + c_2^* \varphi_2^\dagger) A)^\dagger \\ &= c_1 (\varphi_1^\dagger A)^\dagger + c_2 (\varphi_2^\dagger A)^\dagger \\ &= c_1 A^\dagger \varphi_1 + c_2 A^\dagger \varphi_2 \quad \varphi_{1,2} \in \mathcal{V}, \quad c_{1,2} \in \mathbb{C}. \end{aligned}$$

Eigenschaften der adjungierten Abbildung:

1. $(c_1 A_1 + c_2 A_2)^\dagger = c_1^* A_1^\dagger + c_2^* A_2^\dagger, \quad c_{1,2} \in \mathbb{C}, \quad A_{1,2} \in L(\mathcal{U}, \mathcal{V}).$
2. $A^{\dagger\dagger} = A, \quad A \in L(\mathcal{U}, \mathcal{V}).$
3. $(BA)^\dagger = A^\dagger B^\dagger, \quad A \in L(\mathcal{U}, \mathcal{V}), \quad B \in L(\mathcal{V}, \mathcal{W}).$

3.4 Orthonormalbasis

Bezüglich einer ONB $\{\varphi_1, \dots, \varphi_n\}$ eines n -dimensionalen unitären Vektorraumes \mathcal{H} lässt sich jedes $\psi \in \mathcal{H}$ in der Form

$$\psi = \sum_{k=1}^n \varphi_k c_k$$

schreiben, wobei die Entwicklungskoeffizienten $c_k \in \mathbb{C}$ eindeutig durch

$$c_k = \langle \varphi_k | \psi \rangle = \varphi_k^\dagger \psi$$

gegeben sind. Das heißt, dass

$$\psi = \sum_{k=1}^n \varphi_k \varphi_k^\dagger \psi \quad \forall \psi \in \mathcal{H},$$

und man erhält somit die **Vollständigkeitsrelation** für den Fall eines unitären Vektorraumes:

$$\sum_{k=1}^n \varphi_k \varphi_k^\dagger = \mathbb{1}.$$

Das bedeutet, dass eine Menge von n Vektoren $\{\varphi_1, \dots, \varphi_n\}$ in einem n -dimensionalen unitären Vektorraum \mathcal{H} genau dann eine ONB von \mathcal{H} ist, falls die Bedingungen

$$\langle \varphi_k | \varphi_l \rangle = \delta_{kl}, \quad \sum_{k=1}^n \varphi_k \varphi_k^\dagger = \mathbb{1}$$

erfüllt sind.

Beispiel: Verifizieren Sie die oben angegebenen Eigenschaften für die folgende ONB von U^2 :

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}, \quad \frac{1}{\sqrt{2}} \begin{pmatrix} i \\ 1 \end{pmatrix}.$$

Bemerkung: Die Funktionen $e_n(x) = e^{2\pi i n x}$ ($-N \leq n \leq N$) bilden eine ONB des im Abschnitt 3.1 in Beispiel 3 angegebenen unitären Vektorraumes. Aus diesem Grund lassen sich die Entwicklungskoeffizienten c_n in der Form

$$c_n = \langle e_n | \psi \rangle = \int_0^1 dx e^{-2\pi i n x} \psi(x)$$

schreiben. Im (formalen) Limes $N \rightarrow \infty$ erhält man die **Fourierreihe**

$$\psi(x) = \sum_{n \in \mathbb{Z}} c_n e^{2\pi i n x}.$$

Das **Gram-Schmidtsche Orthonormalisierungsverfahren** ist eine Methode, ausgehend von einer beliebigen Basis $\{\chi_1, \dots, \chi_n\}$ eines unitären Vektorraumes, eine ONB (ein VONS) $\{\varphi_1, \dots, \varphi_n\}$ zu konstruieren.

$$\varphi_1 = \frac{\chi_1}{\|\chi_1\|}$$

$$\chi_2 = \underbrace{\varphi_1 \langle \varphi_1 | \chi_2 \rangle}_{\|\varphi_1\|} + \underbrace{\chi_2 - \varphi_1 \langle \varphi_1 | \chi_2 \rangle}_{\perp \varphi_1}$$

$$\Rightarrow \varphi_2 = \frac{\chi_2 - \varphi_1 \langle \varphi_1 | \chi_2 \rangle}{\|\chi_2 - \varphi_1 \langle \varphi_1 | \chi_2 \rangle\|}$$

$$\chi_3 = \underbrace{\varphi_1 \langle \varphi_1 | \chi_3 \rangle + \varphi_2 \langle \varphi_2 | \chi_3 \rangle}_{\text{liegt in dem von } \varphi_1, \varphi_2 \text{ aufgespannten linearen Teilraum}} + \underbrace{\chi_3 - \varphi_1 \langle \varphi_1 | \chi_3 \rangle - \varphi_2 \langle \varphi_2 | \chi_3 \rangle}_{\perp \varphi_1, \varphi_2}$$

$$\Rightarrow \varphi_3 = \frac{\chi_3 - \varphi_1 \langle \varphi_1 | \chi_3 \rangle - \varphi_2 \langle \varphi_2 | \chi_3 \rangle}{\|\chi_3 - \varphi_1 \langle \varphi_1 | \chi_3 \rangle - \varphi_2 \langle \varphi_2 | \chi_3 \rangle\|}$$

usw. \rightarrow VONS $\{\varphi_1, \dots, \varphi_n\}$.

Bemerkung: Aus diesem Resultat folgt auch unmittelbar, dass ein VONS stets existiert.

Ein beliebiger n-dimensionaler unitärer Vektorraum \mathcal{H} unterscheidet sich in gewisser Weise nicht von einem U^n . Wählt man nämlich in \mathcal{H} eine ONB $\{\varphi_1, \dots, \varphi_n\}$, so lässt sich jedes $\psi \in \mathcal{H}$ eindeutig in der Form

$$\psi = \sum_{k=1}^n \varphi_k c_k, \quad c_k = \langle \varphi_k | \psi \rangle$$

schreiben. Der Vektor $\psi \in \mathcal{H}$ ist also (bei vorgegebener ONB) eindeutig durch die n Zahlen

$$\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in U^n$$

bestimmt. Diese Zuordnung

$$\psi \in \mathcal{H} \leftrightarrow \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in U^n$$

respektiert sowohl die Vektoraddition, die Multiplikation mit Skalaren als auch die Bildung des Skalarprodukts. Man sagt auch, dass jeder n -dimensionale unitäre Vektorraum **isomorph** zu U^n ist.

Die **Matrixdarstellung** (A_{kl}) einer linearen Abbildung $A \in L(\mathcal{H})$ hat **bezüglich einer ONB** $\{\varphi_1, \dots, \varphi_n\}$ eine besonders einfache Form. Bildet man nämlich das Skalarprodukt von φ_k mit dem Ausdruck

$$A\varphi_l = \sum_{j=1}^n \varphi_j A_{jl},$$

so erhält man

$$\langle \varphi_k | A\varphi_l \rangle = \langle \varphi_k | \sum_{j=1}^n \varphi_j A_{jl} \rangle = \sum_{j=1}^n A_{jl} \underbrace{\langle \varphi_k | \varphi_j \rangle}_{\delta_{kj}} = A_{kl}.$$

Das heißt, man bildet einfach ein „Sandwich“ des Operators A zwischen den Vektoren φ_k und φ_l um das Matrixelement A_{kl} zu erhalten:

$$A_{kl} = \langle \varphi_k | A\varphi_l \rangle = \varphi_k^\dagger A\varphi_l.$$

Für die Matrixelemente des zu $A \in L(\mathcal{H})$ adjungierten Operators $A^\dagger \in L(\mathcal{H})$ erhält man daher **bezüglich einer ONB**

$$(A^\dagger)_{kl} = \langle \varphi_k | A^\dagger \varphi_l \rangle = \langle A\varphi_k | \varphi_l \rangle = \langle \varphi_l | A\varphi_k \rangle^* = A_{lk}^*.$$

Das heißt, die Zeilen werden mit den Spalten vertauscht und die Matrixelemente werden komplex konjugiert.

Bemerkung: In der Physik wird oft die sogenannte **Diracschreibweise** verwendet. Statt $\psi \in \mathcal{H}$ schreibt man dann $|\psi\rangle \in \mathcal{H}$ (**ket**-Vektor) und für ein Element des Dualraums $\psi^\dagger \in \tilde{\mathcal{H}}$ verwendet man die Notation $\langle\psi|$ (**bra**-Vektor). Die etwas seltsame Sprechweise kommt von dem englischen Wort **bracket** = Klammer = $\langle | \rangle$. In der Diracschreibweise schreibt man z.B. die Vollständigkeitsrelation als

$$\sum_{k=1}^n |\varphi_k\rangle\langle\varphi_k| = \mathbb{1},$$

oder oft in der Form

$$\sum_{k=1}^n |k\rangle\langle k| = \mathbb{1},$$

wobei $|k\rangle$ kurz für den Vektor $|\varphi_k\rangle$ der gewählten ONB steht. Einige weitere Beispiele mit „Übersetzung“:

$$\begin{aligned} |\varphi\rangle\langle\psi| &= \varphi\psi^\dagger \\ |k\rangle\langle l| &= |\varphi_k\rangle\langle\varphi_l| = \varphi_k\varphi_l^\dagger \\ \langle k|l\rangle &= \langle\varphi_k|\varphi_l\rangle = \varphi_k^\dagger\varphi_l \\ \langle k|A|l\rangle &= \langle\varphi_k|A|\varphi_l\rangle = \langle\varphi_k|A\varphi_l\rangle = A_{kl} \end{aligned}$$

3.5 Projektionsoperatoren

Sei \mathcal{M} ein Teilraum eines unitären Vektorraumes \mathcal{H} . Dann bezeichnen wir mit \mathcal{M}^\perp das **orthogonale Komplement** von \mathcal{M} , welches aus allen Elementen $\psi \in \mathcal{H}$ mit $\langle \varphi | \psi \rangle = 0 \quad \forall \varphi \in \mathcal{M}$ besteht. Dann ist \mathcal{M}^\perp wieder ein Teilraum von \mathcal{H} und $\mathcal{M} \cap \mathcal{M}^\perp = \{0\}$, denn

$$\begin{aligned} \psi_{1,2} \in \mathcal{M}^\perp &\Rightarrow \langle \varphi | c_1 \psi_1 + c_2 \psi_2 \rangle = c_1 \langle \varphi | \psi_1 \rangle + c_2 \langle \varphi | \psi_2 \rangle = 0 \\ &\Rightarrow c_1 \psi_1 + c_2 \psi_2 \in \mathcal{M}^\perp, \\ \psi \in \mathcal{M} \cap \mathcal{M}^\perp &\Rightarrow \underbrace{\langle \psi |}_{\in \mathcal{M}} \underbrace{|\psi\rangle}_{\in \mathcal{M}^\perp} = 0 \Rightarrow \psi = 0. \end{aligned}$$

Mit dem Skalarprodukt von \mathcal{H} ist \mathcal{M} selbst ein unitärer Raum. Daher besitzt \mathcal{M} eine ONB $\{\varphi_1, \dots, \varphi_m\}$. Für jeden Vektor $\psi \in \mathcal{H}$ sei

$$\psi_{\mathcal{M}} = \sum_{k=1}^m \varphi_k \langle \varphi_k | \psi \rangle = \sum_{k=1}^m \varphi_k \varphi_k^\dagger \psi$$

Dann ist $\psi_{\mathcal{M}} \in \mathcal{M}$ und $\psi - \psi_{\mathcal{M}} \in \mathcal{M}^\perp$, weil

$$\begin{aligned} \langle \varphi_k | \psi - \psi_{\mathcal{M}} \rangle &= \langle \varphi_k | \psi \rangle - \langle \varphi_k | \psi_{\mathcal{M}} \rangle \\ &= \langle \varphi_k | \psi \rangle - \left\langle \varphi_k \left| \sum_{l=1}^m \varphi_l \langle \varphi_l | \psi \rangle \right. \right\rangle \\ &= \langle \varphi_k | \psi \rangle - \sum_{l=1}^m \langle \varphi_l | \psi \rangle \underbrace{\langle \varphi_k | \varphi_l \rangle}_{\delta_{kl}} \\ &= \langle \varphi_k | \psi \rangle - \langle \varphi_k | \psi \rangle = 0 \quad \forall k = 1, \dots, m \end{aligned}$$

$\Rightarrow \psi = \psi_{\mathcal{M}} + (\psi - \psi_{\mathcal{M}})$ mit $\psi_{\mathcal{M}} \in \mathcal{M}$ und $\psi - \psi_{\mathcal{M}} = \psi_{\mathcal{M}^\perp} \in \mathcal{M}^\perp$. Diese Darstellung ist eindeutig, weil $\mathcal{M} \cap \mathcal{M}^\perp = \{0\}$ ist, d.h. $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^\perp$ und $\dim \mathcal{M}^\perp = \dim \mathcal{H} - \dim \mathcal{M}$. Weiters ist $\mathcal{M}^{\perp\perp} = \mathcal{M}$ für jeden Teilraum \mathcal{M} . Man bezeichnet $\psi_{\mathcal{M}}$ als die **orthogonale Projektion** von ψ auf den Teilraum \mathcal{M} .

Aus der Gleichung

$$\psi_{\mathcal{M}} = \sum_{k=1}^m \varphi_k \varphi_k^\dagger \psi$$

sieht man, dass der Operator

$$P_{\mathcal{M}} = \sum_{k=1}^m \varphi_k \varphi_k^\dagger = \sum_{k=1}^m |\varphi_k\rangle \langle \varphi_k| \in L(\mathcal{H}) \quad (m \leq \dim \mathcal{H})$$

die Projektion $\psi \rightarrow \psi_{\mathcal{M}}$ auf den m -dimensionalen Teilraum \mathcal{M} bewerkstelligt, wobei $P_{\mathcal{M}} \mathcal{M}^\perp = 0$. $P_{\mathcal{M}}$ heißt daher (**orthogonaler**) **Projektionsoperator** oder

(orthogonaler) Projektor auf \mathcal{M} . $P_{\mathcal{M}}$ ist natürlich von der Wahl der ONB $\{\varphi_1, \dots, \varphi_m\}$ in \mathcal{M} unabhängig.

Eigenschaften von $P_{\mathcal{M}}$:

1. $P_{\mathcal{M}}^2 = P_{\mathcal{M}}$ (**idempotent**) wegen $P_{\mathcal{M}}\psi = \psi_{\mathcal{M}}$, $P_{\mathcal{M}}\psi_{\mathcal{M}} = \psi_{\mathcal{M}}$.

Anderer Beweis:

$$P_{\mathcal{M}}^2 = \sum_{k=1}^m \varphi_k \varphi_k^\dagger \sum_{l=1}^m \varphi_l \varphi_l^\dagger = \sum_{k,l=1}^m \varphi_k \underbrace{\varphi_k^\dagger \varphi_l}_{\delta_{kl}} \varphi_l^\dagger = \sum_{k=1}^m \varphi_k \varphi_k^\dagger = P_{\mathcal{M}}$$

2. $P_{\mathcal{M}}^\dagger = P_{\mathcal{M}}$ (einen Operator mit dieser Eigenschaft nennt man **hermitesch** oder **selbstadjungiert**).

$$P_{\mathcal{M}}^\dagger = \left(\sum_{k=1}^m \varphi_k \varphi_k^\dagger \right)^\dagger = \sum_{k=1}^m \left(\varphi_k \varphi_k^\dagger \right)^\dagger = \sum_{k=1}^m \varphi_k^{\dagger\dagger} \varphi_k^\dagger = \sum_{k=1}^m \varphi_k \varphi_k^\dagger = P_{\mathcal{M}}$$

Umgekehrt definiert jeder Operator $P \in L(\mathcal{H})$, der die Eigenschaften $P^2 = P$ und $P^\dagger = P$ erfüllt, durch sein Bild $\mathcal{M} = P\mathcal{H}$ einen linearen Teilraum, wobei $P\mathcal{M}^\perp = 0$, da

$$\langle \chi | P\psi \rangle = \langle P^\dagger \chi | \psi \rangle = \underbrace{\langle P\chi | \psi \rangle}_{\in \mathcal{M}} = 0 \quad \forall \psi \in \mathcal{M}^\perp, \quad \forall \chi \in \mathcal{H}$$

wobei $P^\dagger = P$ verwendet wurde.

Bemerkung: $P_{\mathcal{M}^\perp} = \mathbb{1} - P_{\mathcal{M}}$, $P_{\mathcal{M}}P_{\mathcal{M}^\perp} = 0$.

$P_{\mathcal{M}}$ besitzt die Eigenwerte 0 (falls $\dim \mathcal{M} < \dim \mathcal{H} = n$) und 1, da $P_{\mathcal{M}}\varphi = \varphi$, $\forall \varphi \in \mathcal{M}$ und $P_{\mathcal{M}}\psi = 0$ $\forall \psi \in \mathcal{M}^\perp$ und $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^\perp$. Man kann die ONB $\{\varphi_1, \dots, \varphi_m\}$ von \mathcal{M} (die aus m linear unabhängigen Eigenvektoren von $P_{\mathcal{M}}$ zum Eigenwert 1 besteht) durch die Hinzunahme einer beliebigen ONB $\{\varphi_{m+1}, \dots, \varphi_n\}$ von \mathcal{M}^\perp (die aus $n - m$ linear unabhängigen Eigenvektoren von $P_{\mathcal{M}}$ zum Eigenwert 0 besteht) zu einem VONS $\{\varphi_1, \dots, \varphi_m, \varphi_{m+1}, \dots, \varphi_n\}$ des ganzen Raumes $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^\perp$ ergänzen.

3.6 Hermitesche Operatoren

Definition: Ein Operator $A \in L(\mathcal{H})$ (\mathcal{H} ist ein unitärer Vektorraum) heißt **hermitesch** oder **selbstadjungiert**, wenn $A^\dagger = A$ ist.

Die Eigenwerte eines hermiteschen Operators sind reell: Sei $\psi \in \mathcal{H}$ ein Eigenvektor von A mit Eigenwert a ($A\psi = a\psi$). Dann ist

$$\langle \psi | A\psi \rangle = \langle \psi | a\psi \rangle = a\langle \psi | \psi \rangle.$$

Andererseits ist wegen $A^\dagger = A$

$$\langle \psi | A\psi \rangle = \langle A^\dagger \psi | \psi \rangle = \langle A\psi | \psi \rangle = \langle a\psi | \psi \rangle = a^* \langle \psi | \psi \rangle,$$

und somit

$$(a - a^*)\langle \psi | \psi \rangle = 0.$$

Da ψ als Eigenvektor nicht der Nullvektor sein kann, ist auch $\langle \psi | \psi \rangle \neq 0$, woraus $a = a^* \in \mathbb{R}$ folgt.

Bemerkung: Die Umkehrung gilt nicht, das heißt ein Operator, der nur reelle Eigenwerte besitzt, ist nicht notwendigerweise hermitesch. Gegenbeispiel im U^2 :

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \neq A^\dagger = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix},$$

jedoch hat A den (einfach entarteten) Eigenwert 1.

Beispiele für hermitesche Operatoren:

1. Die allgemeine Form einer hermiteschen Matrix im U^2 ist

$$A = \begin{pmatrix} a & c + id \\ c - id & b \end{pmatrix}, \quad a, b, c, d \in \mathbb{R}.$$

Überzeugen Sie sich, dass die Eigenwerte von A tatsächlich reell sind.

2. Mit den orthogonalen Projektoren ($P^2 = P, P^\dagger = P$) haben wir bereits Spezialfälle von hermiteschen Operatoren kennengelernt. Man kann auch sagen, dass die Projektionsoperatoren genau jene hermiteschen Elemente aus $L(\mathcal{H})$ sind, deren Eigenwerte 0 oder 1 sind.
3. In einem unitären Vektorraum \mathcal{H} mögen die Vektoren $\{\varphi_1, \dots, \varphi_n\}$ ein VONS von \mathcal{H} bilden. Dann ist

$$A = \sum_{k=1}^n a_k \varphi_k \varphi_k^\dagger = \sum_{k=1}^n a_k |\varphi_k\rangle \langle \varphi_k|, \quad a_k \in \mathbb{R}$$

ein hermitescher Operator. Die reellen Zahlen a_k ($k = 1, \dots, n$) sind Eigenwerte von A und die ONB $\{\varphi_1, \dots, \varphi_n\}$ besteht aus Eigenvektoren von A . Wir werden später sehen, dass sich jeder hermitesche Operator in der oben angegebenen Form (**Spektraldarstellung**) schreiben lässt.

3.7 Unitäre Operatoren

Definition: Ein Operator $U \in L(\mathcal{H})$ (\mathcal{H} ein unitärer Vektorraum) heißt unitär, falls U invertierbar ist mit $U^{-1} = U^\dagger$ (d.h. $U^\dagger U = U U^\dagger = \mathbb{1}$).

Satz: Sei $U \in L(\mathcal{H})$ (\mathcal{H} ein unitärer Vektorraum). Dann sind die folgenden Aussagen äquivalent:

1. U ist unitär.
2. $\langle U\varphi | U\psi \rangle = \langle \varphi | \psi \rangle \quad \forall \varphi, \psi \in \mathcal{H}$.
3. $\|U\psi\| = \|\psi\| \quad \forall \psi \in \mathcal{H}$.
4. U bildet jede ONB von \mathcal{H} wieder auf eine ONB von \mathcal{H} ab.

Bei dem Beweis dieses Satzes bereitet nur die Implikation $3 \Rightarrow 2$ etwas Mühe. Nehmen wir an, dass $\|U\psi\| = \|\psi\| \quad \forall \psi \in \mathcal{H}$ erfüllt ist, so folgt unter Verwendung der Polarisierungsidentität, dass

$$\begin{aligned} \langle U\psi | U\varphi \rangle &= \frac{1}{4} \sum_{k=1}^4 i^k \|U\varphi + i^k U\psi\|^2 \\ &= \frac{1}{4} \sum_{k=1}^4 i^k \|U(\varphi + i^k \psi)\|^2 \\ &= \frac{1}{4} \sum_{k=1}^4 i^k \|(\varphi + i^k \psi)\|^2 \\ &= \langle \psi | \varphi \rangle. \end{aligned}$$

Sind $U_1, U_2 \in L(\mathcal{H})$ zwei unitäre Operatoren, dann ist auch das Produkt $U_1 U_2$ ein unitärer Operator, da

$$(U_1 U_2)^{-1} = U_2^{-1} U_1^{-1} = U_2^\dagger U_1^\dagger = (U_1 U_2)^\dagger$$

Die Eigenwerte eines unitären Operators haben den Betrag eins, da

$$\begin{aligned} U\psi &= u\psi \quad (\psi \neq 0, u \in \mathbb{C}) \quad \Rightarrow \\ \langle \psi | \psi \rangle &= \langle U\psi | U\psi \rangle = \langle u\psi | u\psi \rangle = |u|^2 \underbrace{\langle \psi | \psi \rangle}_{\neq 0} \quad \Rightarrow \quad |u| = 1 \\ \Rightarrow \quad u &= e^{i\alpha}, \quad \alpha \in \mathbb{R}. \end{aligned}$$

Das heißt, das Spektrum (die Menge der Eigenwerte) eines unitären Operators ist eine Teilmenge des Einheitskreises in der komplexen Zahlenebene.

Beispiele von unitären Abbildungen:

1. Sei U eine unitäre Matrix auf U^n . Dann wird die Standardbasis $\{e_1, \dots, e_n\}$ (VONS von U^n) in ein anderes VONS von U^n $\{Ue_1, \dots, Ue_n\} = \{f_1, \dots, f_n\}$ übergeführt. Wegen $\sum_{k=1}^n f_k f_k^\dagger = \mathbb{1}$ kann man schreiben:

$$U = U\mathbb{1} = U \sum_{k=1}^n e_k e_k^\dagger = \sum_{k=1}^n \underbrace{Ue_k}_{f_k} e_k^\dagger = \sum_{k=1}^n f_k e_k^\dagger = (f_1, \dots, f_n),$$

d.h. die Spalten einer unitären Matrix bilden ein VONS.

2. Die allgemeine Form einer unitären 2×2 Matrix ist daher:

$$U = \begin{pmatrix} a & -b^* \\ b & a^* \end{pmatrix} e^{i\alpha}, \quad a, b \in \mathbb{C}, |a|^2 + |b|^2 = 1, \alpha \in \mathbb{R}.$$

3. Die Vektoren $\{\varphi_1, \dots, \varphi_n\}$ mögen ein VONS eines unitären Vektorraums bilden. Dann ist

$$U = \sum_{k=1}^n e^{i\alpha_k} \varphi_k \varphi_k^\dagger = \sum_{k=1}^n e^{i\alpha_k} |\varphi_k\rangle \langle \varphi_k|, \quad \alpha_k \in \mathbb{R},$$

ein unitärer Operator. Die komplexe Zahl $e^{i\alpha_k}$ ist ein Eigenwert von U mit dazugehörigem Eigenvektor φ_k . Wieder kann man zeigen, dass sich jeder unitäre Operator in der obigen Form schreiben lässt.

3.8 Normale Operatoren

Definition: Ein Operator $A \in L(\mathcal{H})$ (\mathcal{H} ein unitärer Vektorraum) heißt **normal**, falls $AA^\dagger = A^\dagger A$, d.h. $[A, A^\dagger] = 0$.

Beispiele für normale Operatoren:

1. Hermitesche und unitäre Operatoren sind normal.
2. Sei $\{\varphi_1, \dots, \varphi_n\}$ ein VONS eines unitären Vektorraumes \mathcal{H} . Dann ist

$$A = \sum_{k=1}^n a_k \varphi_k \varphi_k^\dagger = \sum_{k=1}^n a_k |\varphi_k\rangle \langle \varphi_k|, \quad a_k \in \mathbb{C}$$

ein normaler Operator. Wir werden im nächsten Abschnitt zeigen, dass sich jeder normale Operator in dieser Form schreiben lässt. Das ist die Aussage des **Spektralsatzes für normale Operatoren**. Dieser behauptet nämlich, dass die normalen Operatoren genau diejenigen sind, die eine ONB von Eigenvektoren besitzen.

Bemerkungen:

1. Im Raum U^2 ist die Matrix

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

nicht normal ($AA^\dagger \neq A^\dagger A$). Diese Matrix hat nur **einen** Eigenvektor!

2. Für jedes $A \in L(\mathcal{H})$ sind $A^\dagger A$ und AA^\dagger hermitesch.
3. Für jedes $A \in L(\mathcal{H})$ sind

$$A_1 = \frac{1}{2}(A + A^\dagger) \quad \text{und} \quad A_2 = \frac{1}{2i}(A - A^\dagger)$$

hermitesche Operatoren und es gilt

$$A = A_1 + iA_2$$

Der Operator A ist genau dann normal, wenn $[A_1, A_2] = 0$ gilt. Daraus ergibt sich, dass man normale Operatoren in gewisser Hinsicht als Verallgemeinerung der komplexen Zahlen interpretieren kann, wobei die hermiteschen Operatoren den reellen Zahlen entsprechen. Die unitären Operatoren ($U^\dagger U = UU^\dagger = \mathbb{1}$) stellen das Analogon zu einer komplexen Zahl z mit $z^*z = 1$, d.h. $|z| = 1$, dar.

3.9 Spektralsatz für normale Operatoren

Normale Operatoren lassen sich dadurch charakterisieren, dass sie ein vollständiges Orthonormalsystem von Eigenvektoren besitzen. Die Spektraldarstellung normaler Operatoren ist eine unmittelbare Folgerung aus dieser Eigenschaft. Diese wird in den verschiedensten physikalischen Anwendungen benötigt.

Spektralsatz für normale Operatoren: Jeder normale Operator in einem unitären Vektorraum besitzt ein vollständiges Orthonormalsystem von Eigenvektoren.

Beweis: Sei $A \in L(\mathcal{H})$ ein normaler Operator ($[A, A^\dagger] = 0$) in einem n -dimensionalen unitären Vektorraum \mathcal{H} . Dann besitzt das charakteristische Polynom von A ,

$$p_A(a) = \det(a\mathbb{1} - A),$$

mindestens eine Nullstelle $a_1 \in \mathbb{C}$. Die Eigenwertgleichung

$$A\varphi = a_1\varphi$$

besitzt daher eine nichttriviale Lösung $\varphi_1 \neq 0$, wobei man $\langle \varphi_1 | \varphi_1 \rangle = 1$ wählen kann. Da A normal ist, ist φ_1 auch Eigenvektor des adjungierten Operators A^\dagger zum Eigenwert a_1^* :

$$A^\dagger \varphi_1 = a_1^* \varphi_1.$$

Das kann man folgendermaßen sehen:

$$\begin{aligned} 0 &= \langle (A - a_1)\varphi_1 | (A - a_1)\varphi_1 \rangle \\ &= \langle \varphi_1 | (A - a_1)^\dagger (A - a_1)\varphi_1 \rangle \\ &= \langle \varphi_1 | (A - a_1)(A - a_1)^\dagger \varphi_1 \rangle \\ &= \langle (A - a_1)^\dagger \varphi_1 | (A - a_1)^\dagger \varphi_1 \rangle \\ &= \langle (A^\dagger - a_1^*)\varphi_1 | (A^\dagger - a_1^*)\varphi_1 \rangle \\ &\Rightarrow A^\dagger \varphi_1 = a_1^* \varphi_1 \end{aligned}$$

Man betrachtet nun den $(n-1)$ -dimensionalen Teilraum

$$\mathcal{N} = [\varphi_1]^\perp = \{\psi \in \mathcal{H} \mid \langle \varphi_1 | \psi \rangle = 0\}.$$

Die Anwendung von A bzw. A^\dagger auf Elemente von \mathcal{N} führt aus diesem Teilraum nicht heraus, denn für $\psi \in \mathcal{N}$ gilt

$$\langle \varphi_1 | A\psi \rangle = \langle A^\dagger \varphi_1 | \psi \rangle = \langle a_1^* \varphi_1 | \psi \rangle = a_1 \langle \varphi_1 | \psi \rangle = 0$$

und

$$\langle \varphi_1 | A^\dagger \psi \rangle = \langle A\varphi_1 | \psi \rangle = \langle a_1 \varphi_1 | \psi \rangle = a_1^* \langle \varphi_1 | \psi \rangle = 0.$$

Man kann daher A als Operator auf \mathcal{N} auffassen, der natürlich auch normal ist. Man wiederholt nun das Verfahren von vorhin und erhält ein $\varphi_2 \in \mathcal{N}$ (d.h. $\langle \varphi_2 | \varphi_1 \rangle = 0$), $\|\varphi_2\| = 1$, mit $A\varphi_2 = a_2\varphi_2$.

Die weitere Fortsetzung dieser Prozedur liefert nach insgesamt n Schritten ein VONS $\{\varphi_1, \dots, \varphi_n\}$ von Eigenvektoren von A : $A\varphi_k = a_k\varphi_k$.

Zur **Spektraldarstellung** eines normalen Operators,

$$A = \sum_{k=1}^n a_k \varphi_k \varphi_k^\dagger = \sum_{k=1}^n a_k |\varphi_k\rangle \langle \varphi_k|,$$

gelangt man nun einfach durch die Verwendung der Vollständigkeitsrelation für die ONB $\{\varphi_1, \dots, \varphi_n\}$:

$$A = A\mathbb{1} = A \sum_{k=1}^n \varphi_k \varphi_k^\dagger = \sum_{k=1}^n A\varphi_k \varphi_k^\dagger = \sum_{k=1}^n a_k \varphi_k \varphi_k^\dagger$$

Bemerkung: Sei $A \in L(U^n)$ eine normale $n \times n$ -Matrix. Dann gibt es ein VONS $\{f_1, \dots, f_n\}$ von Eigenvektoren von A :

$$f_k^\dagger f_l = \delta_{kl}, \quad \sum_{k=1}^n f_k f_k^\dagger = \mathbb{1}, \quad A f_k = a_k f_k.$$

$$A = \sum_{k=1}^n f_k a_k f_k^\dagger = \underbrace{(f_1, \dots, f_n)}_U \underbrace{\begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix}}_{\hat{A}} \underbrace{\begin{pmatrix} f_1^\dagger \\ \vdots \\ f_n^\dagger \end{pmatrix}}_{U^\dagger}$$

Die Matrizen

$$U = \sum_{k=1}^n f_k e_k^\dagger = (f_1, \dots, f_n)$$

und

$$U^\dagger = \sum_{k=1}^n e_k f_k^\dagger = \begin{pmatrix} f_1^\dagger \\ \vdots \\ f_n^\dagger \end{pmatrix}$$

sind unitär ($UU^\dagger = U^\dagger U = \mathbb{1}$). Der Spektralsatz besagt also, dass man jede normale $n \times n$ -Matrix in der Form

$$A = U \hat{A} U^\dagger$$

schreiben kann, wobei U unitär ist und \hat{A} die aus den Eigenwerten a_1, \dots, a_n gebildete **Diagonalmatrix** $\hat{A} = \text{diag}(a_1, \dots, a_n)$ ist. Multipliziert man die obige Gleichung von links mit U^\dagger und von rechts mit U , so erhält man

$$U^\dagger A U = \hat{A}$$

d.h. eine $n \times n$ -Matrix lässt sich genau dann durch eine unitäre Transformation auf Diagonalgestalt bringen, wenn sie normal ist. An dieser Stelle sei an den allgemeinen Fall einer diagonalisierbaren $n \times n$ -Matrix erinnert, bei der eine Basis $\{f_1, \dots, f_n\}$ von Eigenvektoren $A f_k = a_k f_k$ existiert. Bildet man die Matrix $S = (f_1, \dots, f_n)$, so ist $S^{-1} A S = \hat{A} = \text{diag}(a_1, \dots, a_n)$ die gewünschte Diagonalisierung. In unserem Fall einer normalen Matrix haben wir also den Spezialfall, dass die Eigenvektoren eine ONB bilden und S daher unitär ist ($S^{-1} = S^\dagger$).

Kochrezept für das Diagonalisieren einer normalen $n \times n$ -Matrix A (d.h. einer Matrix mit $[A, A^\dagger] = 0$):

1. Man berechne das **charakteristische Polynom** von A :

$$p_A(a) = \det(a\mathbb{1} - A)$$

- Man berechne die **Eigenwerte** von A , indem man die Nullstellen des charakteristischen Polynoms bestimmt. Es kann natürlich vorkommen, dass manche der Eigenwerte entartet sind.
- Man ermittle die dazugehörigen **Eigenvektoren** f_k aus der **Eigenwertgleichung**

$$Af_k = a_k f_k$$

und normiere sie ($\|f_k\| = 1$). Eigenvektoren zu verschiedenen Eigenwerten ($a_k \neq a_l$) sind automatisch orthogonal, denn

$$\begin{aligned} f_k^\dagger A f_l &= a_l f_k^\dagger f_l = (A^\dagger f_k)^\dagger f_l = a_k f_k^\dagger f_l \\ \Rightarrow \underbrace{(a_k - a_l)}_{\neq 0} f_k^\dagger f_l &= 0 \quad \Rightarrow \quad f_k^\dagger f_l = 0. \end{aligned}$$

Ist ein Eigenwert d -fach entartet, so gibt es d linear unabhängige Eigenvektoren (d.h. der Eigenraum dieses Eigenwertes ist d -dimensional). Diese Eigenvektoren können aber stets orthonormal **gewählt** werden. Auf diese Weise erhält man ein VONS $\{f_1, \dots, f_n\}$ von U^n .

- Man bilde die unitäre Matrix $U = (f_1, \dots, f_n)$ mit den Eigenvektoren als Spalten $\Rightarrow U^\dagger A U = \tilde{A} = \text{diag}(a_1, \dots, a_n)$.

Beispiele:

- Die Matrix

$$R = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

ist in U^2 unitär, ihre Eigenwerte sind $e^{\pm i\theta}$, die dazugehörigen normierten Eigenvektoren sind

$$\begin{aligned} f_1 &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} && \text{(zum Eigenwert } e^{+i\theta}), \\ f_2 &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} && \text{(zum Eigenwert } e^{-i\theta}). \end{aligned}$$

somit ist

$$U = (f_1, f_2) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}$$

und

$$U^\dagger R U = \text{diag}(e^{i\theta}, e^{-i\theta}).$$

2. Die 3×3 -Matrix

$$A = \begin{pmatrix} 1 & -2 & 1 \\ -2 & 1 & 1 \\ 1 & 1 & -2 \end{pmatrix}$$

ist hermitesch, die Eigenwerte müssen daher reell sein. Für das charakteristische Polynom erhält man $p_A(a) = a(a^2 - 9)$ und somit die Eigenwerte

$$a_1 = 0, \quad a_2 = 3, \quad a_3 = -3.$$

Die Eigenwertgleichungen $Af_1 = 0$, $Af_2 = 3f_2$, $Af_3 = -3f_3$ ergeben die (normierten) Lösungen

$$f_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad f_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad f_3 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix},$$

Man überprüft leicht, dass $\{f_1, f_2, f_3\}$ tatsächlich eine ONB von U^3 ist. Ebenso kann man als Probe die Spektraldarstellung

$$A = \sum_{k=1}^3 a_k f_k f_k^\dagger = 3 \left(f_2 f_2^\dagger - f_3 f_3^\dagger \right)$$

„überprüfen“. Für die unitäre Matrix $U = (f_1, f_2, f_3)$, welche die Diagonalisierung von A in der Form $U^\dagger A U = \text{diag}(0, 3, -3)$ bewerkstelligt, erhält man daher

$$U = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \end{pmatrix}.$$

Bemerkung: Wir haben die Spektraldarstellung eines normalen Operators A in einem n -dimensionalen unitären Raum so formuliert, dass wir A als Linearkombination von **eindimensionalen** Projektionsoperatoren $|\varphi_k\rangle\langle\varphi_k|$ geschrieben haben,

$$A = \sum_{k=1}^n a_k |\varphi_k\rangle\langle\varphi_k|.$$

Wenn es entartete Eigenwerte gibt, dann sind einige der a_k **gleich** und es ist möglich die entsprechenden Terme zusammenzufassen. Dazu ändern wir unsere Notation etwas ab: $\{a_\alpha\}_{\alpha=1}^m$ sei das Spektrum von A (d.h. $a_\alpha \neq a_\beta$ für $\alpha \neq \beta$). a_1, \dots, a_m bezeichnen also jetzt die $m \leq n$ **verschiedenen** Eigenwerte von A mit dazugehörigen Eigenräumen \mathcal{M}_α ($\alpha = 1, \dots, m$), die paarweise aufeinander orthogonal stehen:

$$\mathcal{M}_\alpha \perp \mathcal{M}_\beta \quad \text{für} \quad \alpha \neq \beta, \quad \mathcal{H} = \mathcal{M}_1 \oplus \dots \oplus \mathcal{M}_m.$$

Durch eine geeignete Umbenennung der Vektoren der ONB $\{|\varphi_1\rangle, \dots, |\varphi_n\rangle\}$ in

$$\{|\alpha, r\rangle\} \quad \alpha = 1, \dots, m, \quad r = 1, \dots, d_\alpha = \dim \mathcal{M}_\alpha$$

kann man die Projektoren auf die Eigenräume \mathcal{M}_α ($\alpha = 1, \dots, m$) in der Form

$$P_{\mathcal{M}_\alpha} = \sum_{r=1}^{d_\alpha} |\alpha, r\rangle\langle\alpha, r|$$

schreiben. Die Tatsache, dass die $P_{\mathcal{M}_\alpha}$ orthogonale Projektionsoperatoren auf die paarweise orthogonal stehenden Teilräume $\mathcal{M}_1, \dots, \mathcal{M}_m$ sind, drückt sich durch die Relationen

$$P_{\mathcal{M}_\alpha} P_{\mathcal{M}_\beta} = \delta_{\alpha\beta} P_{\mathcal{M}_\alpha}, \quad P_{\mathcal{M}_\alpha}^\dagger = P_{\mathcal{M}_\alpha}$$

aus. Zusammen mit der Vollständigkeitsrelation

$$\sum_{\alpha=1}^m P_{\mathcal{M}_\alpha} = \mathbb{1}$$

gelangen wir schließlich zur gewünschten alternativen Formulierung der Spektraldarstellung von A :

$$A = \sum_{\alpha=1}^m a_\alpha P_{\mathcal{M}_\alpha}.$$

3.10 Gleichzeitige Diagonalisierbarkeit

Das folgende Kriterium für die gleichzeitige Diagonalisierbarkeit normaler Operatoren wird vor allem in der Quantenmechanik oft benötigt.

Satz: $A, B \in L(\mathcal{H})$ seien normale Operatoren, die miteinander **vertauschen** ($[A, B] = 0$). Dann gibt es eine **gemeinsame** ONB von Eigenvektoren für beide Operatoren.

Beweis: Seien a_α ($\alpha = 1, \dots, m$) die verschiedenen Eigenwerte von A und \mathcal{M}_α die entsprechenden Eigenräume. $\psi \in \mathcal{M}_\alpha \Rightarrow B\psi \in \mathcal{M}_\alpha$, denn

$$AB\psi = BA\psi = Ba_\alpha\psi = a_\alpha B\psi$$

$\Rightarrow B|_{\mathcal{M}_\alpha}$ ist ein normaler Operator auf $\mathcal{M}_\alpha \Rightarrow \exists$ ONB von \mathcal{M}_α $\{|\alpha, 1\rangle, \dots, |\alpha, d_\alpha\rangle\}$ von Eigenvektoren von B . Außerdem ist natürlich $A|\alpha, r\rangle = a_\alpha|\alpha, r\rangle$.

Die Vektoren $|\alpha, r\rangle$ ($\alpha = 1, \dots, m, r = 1, \dots, d_\alpha = \dim \mathcal{M}_\alpha$) bilden ein VONS von \mathcal{H} , das aus Eigenvektoren von A und B gebildet wird.

Bemerkung: Für normale $n \times n$ -Matrizen A, B mit $[A, B] = 0$ bedeutet das: \exists unitäre Matrix U , sodass $U^\dagger A U$ und $U^\dagger B U$ Diagonalmatrizen sind. (A und B sind gleichzeitig diagonalisierbar.)

3.11 Funktionen normaler Operatoren

Funktionen von linearen Operatoren treten etwa bei der Behandlung von Systemen linearer Differentialgleichungen, der Beschreibung von Symmetrietransformationen und der Zeitentwicklung in der Quantenmechanik auf. Die Berechnung der Funktion eines normalen Operators ist erstaunlich einfach. Sie kann auf die Bestimmung der Funktionswerte auf dem Spektrum des Operators zurückgeführt werden.

Um dies zu sehen, gehen wir von der Spektraldarstellung eines normalen Operators A in der Form

$$A = \sum_{k=1}^n a_k \varphi_k \varphi_k^\dagger$$

aus. Multipliziert man diesen Ausdruck mit sich selbst (was einer nochmaligen Anwendung der Abbildung A entspricht), so erhält man

$$\begin{aligned} A^2 &= AA = \sum_{k=1}^n a_k \varphi_k \varphi_k^\dagger \sum_{l=1}^n a_l \varphi_l \varphi_l^\dagger = \sum_{k,l=1}^n a_k a_l \varphi_k \underbrace{\varphi_k^\dagger \varphi_l}_{\delta_{kl}} \varphi_l^\dagger \\ &= \sum_{k=1}^n a_k^2 \varphi_k \varphi_k^\dagger \end{aligned}$$

bzw. für beliebiges $m \in \mathbb{N}$,

$$A^m = \underbrace{A \cdots A}_{m\text{-mal}} = \sum_{k=1}^n a_k^m \varphi_k \varphi_k^\dagger.$$

Ebenso erhält man für ein Polynom von A den Ausdruck

$$p(A) = \sum_m c_m A^m = \sum_{k=1}^n \left(\sum_m c_m a_k^m \right) \varphi_k \varphi_k^\dagger = \sum_{k=1}^n p(a_k) \varphi_k \varphi_k^\dagger.$$

Für eine beliebige komplexwertige Funktion f , die auf dem **Spektrum** des normalen Operators $A \in L(\mathcal{H})$ (das ist die Menge der Eigenwerte von A) definiert ist, kann man $f(A) \in L(\mathcal{H})$ daher durch

$$f(A) = \sum_{k=1}^n f(a_k) \varphi_k \varphi_k^\dagger$$

definieren, wobei $f(A)$ natürlich wieder ein normaler Operator ist.

Beispiele:

1. Sei $A \in L(U^n)$ eine normale $n \times n$ -Matrix. $A = U\hat{A}U^\dagger$, $\hat{A} = \text{diag}(a_1, \dots, a_n)$, U unitär $\Rightarrow f(A) = Uf(\hat{A})U^\dagger$, $f(\hat{A}) = \text{diag}(f(a_1), \dots, f(a_n))$.
2. Sei $U \in L(\mathcal{H})$ ein unitärer Operator. Seine Eigenwerte haben dann die Gestalt $e^{i\alpha_k}$ ($\alpha \in \mathbb{R}$) und die Spektraldarstellung lautet

$$U = \sum_{k=1}^n e^{i\alpha_k} \varphi_k \varphi_k^\dagger = e^{iA},$$

wobei $A = \sum_{k=1}^n \alpha_k \varphi_k \varphi_k^\dagger = A^\dagger$.

Das heißt, jeder unitäre Operator kann in der Form

$$U = e^{iA}, \quad A = A^\dagger,$$

geschrieben werden.

3. Die Funktion $f(A)$ eines normalen Operators A in einem **zweidimensionalen** unitären Raum kann als Linearkombination des Einheitsoperators und von A selbst geschrieben werden. Sind die zwei möglichen Eigenwerte entartet, so ist $A = a\mathbb{1}$ und daher $f(A) = f(a)\mathbb{1}$, womit die Behauptung offensichtlich erfüllt ist. Im Fall verschiedener Eigenwerte ($a_1 \neq a_2$) hat man

$$A = a_1 P_1 + a_2 P_2, \quad P_1 + P_2 = \mathbb{1},$$

wobei

$$P_1 = \varphi_1 \varphi_1^\dagger, \quad P_2 = \varphi_2 \varphi_2^\dagger$$

die beiden eindimensionalen Projektoren auf die von den Eigenvektoren $\varphi_{1,2}$ aufgespannten Eigenräume sind. Aus der Spektraldarstellung und der Vollständigkeitsrelation kann man die beiden Projektionsoperatoren als Linearkombinationen von $\mathbb{1}$ und A schreiben:

$$P_1 = \frac{A - a_2}{a_1 - a_2}, \quad P_2 = \frac{A - a_1}{a_2 - a_1}.$$

Somit erhält man

$$\begin{aligned} f(A) &= f(a_1)P_1 + f(a_2)P_2 = f(a_1)\frac{A - a_2}{a_1 - a_2} + f(a_2)\frac{A - a_1}{a_1 - a_2} \\ &= \frac{f(a_2)a_1 - f(a_1)a_2}{a_1 - a_2} \mathbb{1} + \frac{f(a_1) - f(a_2)}{a_1 - a_2} A. \end{aligned}$$

4. In der Quantenmechanik treten bei der Beschreibung des Spin 1/2-Systems die **Paulischen Spinmatrizen**

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

auf. Sie sind in U^2 hermitesch mit Eigenwerten ± 1 . Sie erfüllen die **Vertauschungsrelationen**

$$[\sigma_k, \sigma_l] = 2i \sum_{m=1}^3 \epsilon_{klm} \sigma_m$$

und die Antivertauschungsrelationen

$$\sigma_k \sigma_l + \sigma_l \sigma_k = 2\delta_{kl} \mathbb{1}.$$

Bildet man die Linearkombination

$$\vec{n} \cdot \vec{\sigma} = \sum_{k=1}^3 n_k \sigma_k$$

mit einem Einheitsvektor

$$\begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix} \in \mathbb{R}^3, \quad \sqrt{n_1^2 + n_2^2 + n_3^2} = 1,$$

so ist $\vec{n} \cdot \vec{\sigma}$ wieder hermitesch mit Eigenwerten ± 1 . Bei der Beschreibung der räumlichen Drehung eines Spins um den Winkel α mit Drehachse \vec{n} , tritt die Matrix

$$e^{-i\alpha \vec{n} \cdot \vec{\sigma} / 2}$$

auf. Überprüfen Sie, dass diese Matrix in der Form

$$e^{-i\alpha \vec{n} \cdot \vec{\sigma} / 2} = \mathbb{1} \cos \frac{\alpha}{2} - i \vec{n} \cdot \vec{\sigma} \sin \frac{\alpha}{2}$$

geschrieben werden kann. Hinweis: Wenden Sie die Formel von Beispiel 3 an.

Kapitel 4

Tensorprodukte von Vektorräumen

Tensorprodukte und Tensoren sind Standardkonstruktionen der linearen Algebra. In der Quantenmechanik benötigt man sie zur Beschreibung von Zustandsräumen und Zuständen zusammengesetzter Systeme.

4.1 Tensorprodukte

$\mathcal{V}^{(1)}, \dots, \mathcal{V}^{(r)}$ seien r (i.A. verschiedene) Vektorräume über einem Grundkörper K (\mathbb{R} oder \mathbb{C}). Unter dem **Tensorprodukt**

$$\mathcal{V}^{(1)} \otimes \dots \otimes \mathcal{V}^{(r)}$$

dieser Vektorräume versteht man die Menge aller **multilinearen Funktionale**

$$T : \tilde{\mathcal{V}}^{(1)} \times \dots \times \tilde{\mathcal{V}}^{(r)} \rightarrow K$$

auf dem kartesischen Produkt der Dualräume $\tilde{\mathcal{V}}^{(1)}, \dots, \tilde{\mathcal{V}}^{(r)}$.

Definiert man in $\mathcal{V}^{(1)} \otimes \dots \otimes \mathcal{V}^{(r)}$ eine Addition und eine Multiplikation mit Skalaren durch

$$(a_1 T_1 + a_2 T_2)(\tilde{v}^{(1)}, \dots, \tilde{v}^{(r)}) = a_1 T_1(\tilde{v}^{(1)}, \dots, \tilde{v}^{(r)}) + a_2 T_2(\tilde{v}^{(1)}, \dots, \tilde{v}^{(r)}),$$
$$a_{1,2} \in K, \quad T_{1,2} \in \mathcal{V}^{(1)} \otimes \dots \otimes \mathcal{V}^{(r)}, \quad \tilde{v}^{(k)} \in \tilde{\mathcal{V}}^{(k)},$$

so bildet $\mathcal{V}^{(1)} \otimes \dots \otimes \mathcal{V}^{(r)}$ seinerseits einen Vektorraum.

Das **Tensorprodukt**

$$v^{(1)} \otimes \dots \otimes v^{(r)} \in \mathcal{V}^{(1)} \otimes \dots \otimes \mathcal{V}^{(r)}$$

der Vektoren $v^{(1)}, \dots, v^{(r)}$ ($v^{(k)} \in \mathcal{V}^{(k)}$) wird durch

$$(v^{(1)} \otimes \dots \otimes v^{(r)})(\tilde{v}^{(1)}, \dots, \tilde{v}^{(r)}) = \tilde{v}^{(1)}(v^{(1)}) \dots \tilde{v}^{(r)}(v^{(r)})$$

definiert.

Sind in den Vektorräumen $\mathcal{V}^{(1)}, \dots, \mathcal{V}^{(r)}$ Basissysteme

$$\{e_1^{(k)}, \dots, e_{\dim \mathcal{V}^{(k)}}^{(k)}\}, \quad 1 \leq k \leq r,$$

gegeben, dann bilden die Elemente

$$\{e_{i_1}^{(1)} \otimes \dots \otimes e_{i_r}^{(r)}\}, \quad 1 \leq i_k \leq \dim \mathcal{V}^{(k)}, \quad 1 \leq k \leq r,$$

eine Basis von $\mathcal{V}^{(1)} \otimes \dots \otimes \mathcal{V}^{(r)}$. Somit gilt:

$$\dim(\mathcal{V}^{(1)} \otimes \dots \otimes \mathcal{V}^{(r)}) = \prod_{k=1}^r \dim \mathcal{V}^{(k)}.$$

Jedes Element von $\mathcal{V}^{(1)} \otimes \dots \otimes \mathcal{V}^{(r)}$ lässt sich daher in der Form

$$T = \sum_{i_1, \dots, i_r} e_{i_1}^{(1)} \otimes \dots \otimes e_{i_r}^{(r)} T_{i_1 \dots i_r}$$

schreiben. T bezeichnet man als **Tensor** r -ter Stufe mit Komponenten $T_{i_1 \dots i_r}$ bezüglich der gewählten Basis. Das Transformationsverhalten der Tensorkomponenten bei einem Basiswechsel ist damit offensichtlich.

Sind r lineare Abbildungen

$$A^{(k)} : \mathcal{V}^{(k)} \rightarrow \mathcal{V}^{(k)}, \quad 1 \leq k \leq r,$$

gegeben, dann definiert man die lineare Abbildung

$$A^{(1)} \otimes \dots \otimes A^{(r)} : \mathcal{V}^{(1)} \otimes \dots \otimes \mathcal{V}^{(r)} \rightarrow \mathcal{V}^{(1)} \otimes \dots \otimes \mathcal{V}^{(r)}$$

durch

$$(A^{(1)} \otimes \dots \otimes A^{(r)})(v^{(1)} \otimes \dots \otimes v^{(r)}) = (A^{(1)}v^{(1)}) \otimes \dots \otimes (A^{(r)}v^{(r)}).$$

Die Erweiterung auf beliebige Tensoren (d.h. solche, die sich nicht in Produktform schreiben lassen) erfolgt durch die Linearitätsforderung.

Bemerkung: Für die hier durchgeführten Konstruktionen müssen die Vektorräume $\mathcal{V}^{(1)}, \dots, \mathcal{V}^{(r)}$ nicht mit Skalarprodukten versehen sein.

4.2 Tensorprodukt von Zustandsräumen

Wird ein Teil A eines Quantensystems durch den Zustandsraum $\mathcal{H}^{(A)}$ und ein Teil B durch den Zustandsraum $\mathcal{H}^{(B)}$ beschrieben, dann ist das Tensorprodukt

$$\mathcal{H}^{(A)} \otimes \mathcal{H}^{(B)}$$

der geeignete Zustandsraum für das Gesamtsystem $A \cup B$. Wir wollen der Einfachheit halber annehmen, dass $\mathcal{H}^{(A)}$ und $\mathcal{H}^{(B)}$ **endlichdimensionale** unitäre Vektorräume mit Skalarprodukten

$$\langle \varphi^{(A)} | \psi^{(A)} \rangle_{\mathcal{H}^{(A)}}, \varphi^{(A)}, \psi^{(A)} \in \mathcal{H}^{(A)}, \quad \langle \varphi^{(B)} | \psi^{(B)} \rangle_{\mathcal{H}^{(B)}}, \varphi^{(B)}, \psi^{(B)} \in \mathcal{H}^{(B)}$$

sind. Das Skalarprodukt in $\mathcal{H}^{(A)} \otimes \mathcal{H}^{(B)}$ wird durch

$$\langle \varphi^{(A)} \otimes \varphi^{(B)} | \psi^{(A)} \otimes \psi^{(B)} \rangle_{\mathcal{H}^{(A)} \otimes \mathcal{H}^{(B)}} = \langle \varphi^{(A)} | \psi^{(A)} \rangle_{\mathcal{H}^{(A)}} \langle \varphi^{(B)} | \psi^{(B)} \rangle_{\mathcal{H}^{(B)}}$$

zusammen mit der Linearitätsforderung erklärt.

Ist

$$\{\varphi_1, \dots, \varphi_n\}, \quad \varphi_a \in \mathcal{H}^{(A)}$$

eine ONB von $\mathcal{H}^{(A)}$ und

$$\{\chi_1, \dots, \chi_\nu\}, \quad \chi_\alpha \in \mathcal{H}^{(B)}$$

eine ONB von $\mathcal{H}^{(B)}$, dann ist

$$\{\varphi_a \otimes \chi_\alpha\}, \quad 1 \leq a \leq n, 1 \leq \alpha \leq \nu$$

eine ONB von $\mathcal{H}^{(A)} \otimes \mathcal{H}^{(B)}$. Ein Vektor $\Psi \in \mathcal{H}^{(A)} \otimes \mathcal{H}^{(B)}$ besitzt dann bezüglich dieser ONB die Darstellung

$$\Psi = \sum_{a=1}^n \sum_{\alpha=1}^{\nu} \varphi_a \otimes \chi_\alpha c_{a\alpha}, \quad c_{a,\alpha} = \langle \varphi_a \otimes \chi_\alpha | \Psi \rangle \in \mathbb{C}.$$

Besteht zwischen den beiden Teilen des Systems keine Wechselwirkung, dann ist jener hermitesche Operator, welcher der Observablen „Gesamtenergie des Systems“ entspricht von der Form

$$H = H^{(A)} \otimes \mathbb{1} + \mathbb{1} \otimes H^{(B)}$$

(Hamiltonoperator des Gesamtsystems ohne Wechselwirkung), wobei die hermiteschen Operatoren

$$H^{(A)} = H^{(A)\dagger} \in L(\mathcal{H}^{(A)}), \quad H^{(B)} = H^{(B)\dagger} \in L(\mathcal{H}^{(B)})$$

die Hamiltonoperatoren der Teilsysteme sind. Bei Vorhandensein einer Wechselwirkung zwischen den beiden Teilsystemen hat man

$$H = H^{(A)} \otimes \mathbb{1} + \mathbb{1} \otimes H^{(B)} + W,$$

wobei der Wechselwirkungsterm $W = W^\dagger \in L(\mathcal{H}^{(A)} \otimes \mathcal{H}^{(B)})$ die Struktur

$$W = \sum_r (X_r^{(A)} \otimes Y_r^{(B)} + X_r^{(A)\dagger} \otimes Y_r^{(B)\dagger}), \quad X_r^{(A)} \in L(\mathcal{H}^{(A)}), \quad Y_r^{(B)} \in L(\mathcal{H}^{(B)})$$

besitzt.

Bemerkungen:

1. $(X^{(A)} \otimes Y^{(B)})^\dagger = X^{(A)\dagger} \otimes Y^{(B)\dagger}$, $X^{(A)} \in L(\mathcal{H}^{(A)})$, $Y^{(B)} \in L(\mathcal{H}^{(B)})$.
2. In der physikalischen Literatur findet man oft die Schreibweise $|a, \alpha\rangle$ statt $\varphi_a \otimes \chi_\alpha$. Statt $X^{(A)} \otimes \mathbb{1}$ und $\mathbb{1} \otimes Y^{(B)}$ schreibt man kurz $X^{(A)}$ und $Y^{(B)}$ wobei der Kommutator $[X^{(A)}, Y^{(B)}] = 0$ ist. Wegen $(X^{(A)} \otimes \mathbb{1})(\mathbb{1} \otimes Y^{(B)}) = X^{(A)} \otimes Y^{(B)}$ schreibt man dann einfach $X^{(A)} Y^{(B)}$ statt $X^{(A)} \otimes Y^{(B)}$.

Kapitel 5

Verallgemeinerte Funktionen

Wegen ihrer großen rechentechnischen Vorteile, werden in der theoretischen Physik häufig verallgemeinerte Funktionen (Distributionen) verwendet. So sind etwa viele im üblichen Sinn nicht differenzierbare Funktionen im Distributionssinn sehr wohl differenzierbar. Man handelt sich auf diese Weise allerdings recht singuläre Objekte ein. Prominentestes Beispiel ist die von P.A.M. Dirac in die physikalische Literatur eingeführte Deltafunktion, die den Anstoß für die Entwicklung der mathematischen Theorie der verallgemeinerten Funktionen gab.

5.1 Motivation

Eine ruhende, im Ursprung befindliche Punktladung q erzeugt das elektrische Feld

$$\vec{E}(\vec{x}) = \frac{q\vec{x}}{|\vec{x}|^3},$$

das dazugehörige skalare Potential ist wegen $\vec{E} = -\vec{\nabla}\phi$ durch

$$\phi(\vec{x}) = \frac{q}{|\vec{x}|}$$

gegeben. Sowohl $\vec{E}(\vec{x})$ als auch $\phi(\vec{x})$ sind auf $\mathbb{R}^3 \setminus \{\vec{0}\}$ beliebig oft differenzierbar. Da das elektrische Feld aber eine Lösung der (elektrostatistischen) Maxwellgleichungen

$$\vec{\nabla} \times \vec{E}(\vec{x}) = 0, \quad \vec{\nabla} \cdot \vec{E}(\vec{x}) = 4\pi\rho(\vec{x})$$

$\forall \vec{x} \in \mathbb{R}^3$ sein soll, benötigt man eine geeignete mathematische Beschreibung für die Ladungsdichte einer Punktladung an der Stelle $\vec{x} = \vec{0}$. Man schreibt dafür

$$\rho(\vec{x}) = q\delta^{(3)}(\vec{x}),$$

wobei die **Deltafunktion** in drei Raumdimensionen, $\delta^{(3)}(\vec{x})$, die folgenden Eigenschaften haben sollte:

$$\delta^{(3)}(\vec{x}) = \begin{cases} 0 & \text{für } \vec{x} \neq \vec{0} \\ \infty & \text{für } \vec{x} = \vec{0} \end{cases} \quad \text{und} \quad \int_{\mathbb{R}^3} d^3x \delta^{(3)}(\vec{x}) \varphi(\vec{x}) = \varphi(\vec{0})$$

für jede im Ursprung stetige Funktion φ . Das skalare Potential genügt der Poissongleichung

$$\Delta\phi(\vec{x}) = -4\pi\rho(\vec{x})$$

und so erhält man im Fall der Punktladung (mit $q = 1$) die Formel

$$\Delta \frac{1}{|\vec{x}|} = -4\pi\delta^{(3)}(\vec{x}).$$

Um die Diskussion zu vereinfachen, betrachten wir den Fall der Deltafunktion in **einer** Raumdimension, welche

$$\delta(x) = \begin{cases} 0 & \text{für } x \neq 0 \\ \infty & \text{für } x = 0 \end{cases} \quad \text{und} \quad \int_{-\infty}^{+\infty} dx \delta(x) \varphi(x) = \varphi(0)$$

für jede am Nullpunkt stetige Funktion φ erfüllen sollte. Es ist klar, dass es keine Funktion $\delta : \mathbb{R} \rightarrow \mathbb{R}$ mit diesen Eigenschaften geben kann. Man kann der Deltafunktion aber im Rahmen der Theorie der verallgemeinerten Funktionen (oder Distributionen) einen mathematisch wohldefinierten Sinn geben.

5.2 Testfunktionen

Zunächst wählt man einen Funktionenraum \mathcal{D} so genannter **Testfunktionen**. Das sind Funktionen mit besonders „schönen“ Eigenschaften. Eine mögliche Wahl ist der Raum aller komplexwertigen, beliebig oft differenzierbaren Funktionen auf \mathbb{R} , mit der zusätzlichen Eigenschaft, dass es zu jeder dieser Funktionen φ ein kompaktes Intervall $[a, b] \subset \mathbb{R}$ gibt mit $\varphi(x) = 0 \forall x \notin [a, b]$. Es handelt sich bei der hier gewählten Menge von Testfunktionen also um den Raum aller unendlich oft differenzierbaren Funktionen mit kompaktem Träger.

Auf \mathcal{D} kann man einen **Konvergenzbegriff** einführen:

Definition: Eine Folge von Testfunktionen $\varphi_n \in \mathcal{D}$ konvergiert in \mathcal{D} gegen die Testfunktion $\varphi \in \mathcal{D}$, wenn ein vom Index n unabhängiges kompaktes Intervall $[a, b] \subset \mathbb{R}$ existiert mit $\varphi_n(x) = 0 \forall x \notin [a, b]$ und

$$\varphi_n^{(k)} \xrightarrow{n \rightarrow \infty} \varphi^{(k)}, \quad \forall k = 0, 1, \dots$$

im Sinne der gleichmäßigen Konvergenz. Das heißt, dass

$$\forall \varepsilon > 0, \forall k \geq 0 \quad \exists N_k(\varepsilon) : \forall n \geq N_k : \sup_{x \in \mathbb{R}} |\varphi_n^{(k)}(x) - \varphi^{(k)}(x)| < \varepsilon.$$

5.3 Distributionen

Definition: Unter einer **Distribution** oder **verallgemeinerten Funktion** versteht man ein stetiges lineares Funktional auf \mathcal{D} , das ist eine Abbildung $l : \mathcal{D} \rightarrow \mathbb{C}$ mit den folgenden Eigenschaften:

1. $l(c_1\varphi_1 + c_2\varphi_2) = c_1l(\varphi_1) + c_2l(\varphi_2), \quad \forall \varphi_1, \varphi_2 \in \mathcal{D}, \forall c_1, c_2 \in \mathbb{C}$ (Linearität)
2. $\varphi_n \rightarrow \varphi$ in $\mathcal{D} \Rightarrow l(\varphi_n) \rightarrow l(\varphi)$ (Stetigkeit)

Bemerkung: Statt $l(\varphi)$ schreibt man oft auch $(l|\varphi)$.

Beispiele für Distributionen:

1. Sei $g : \mathbb{R} \rightarrow \mathbb{C}$ eine stetige Funktion. Dann wird durch

$$\varphi \rightarrow (g|\varphi) = \int_{-\infty}^{+\infty} dx g(x)\varphi(x), \quad \varphi \in \mathcal{D}$$

eine verallgemeinerte Funktion definiert.

Üblicherweise beschreibt man eine stetige Funktion $g : \mathbb{R} \rightarrow \mathbb{C}$ durch die Angabe ihrer Funktionswerte für alle $x \in \mathbb{R}$. Dieselbe Information ist aber auch in dem linearen Funktional $(g|\dots)$ enthalten, denn um den Funktionswert von g an der Stelle x_0 zu rekonstruieren, muss man ja nur eine Folge (φ_n) von Testfunktionen mit

$$\int_{-\infty}^{+\infty} dx \varphi_n(x) = 1$$

wählen, die für $n \rightarrow \infty$ immer mehr um den Punkt x_0 konzentriert ist $\Rightarrow (g|\varphi_n) \xrightarrow{n \rightarrow \infty} g(x_0)$.

2. Allgemein bezeichnet man eine Distribution, die man in der Form

$$\varphi \rightarrow (f|\varphi) = \int_{-\infty}^{+\infty} dx f(x)\varphi(x), \quad \varphi \in \mathcal{D},$$

mit einer geeigneten Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$, schreiben kann als **reguläre Distribution**.

3. Die δ -Distribution („Deltafunktion“) ist durch

$$\varphi \rightarrow (\delta|\varphi) = \varphi(0), \quad \varphi \in \mathcal{D}$$

definiert. Obwohl sie **keine** reguläre Distribution ist, schreibt man trotzdem **symbolisch**:

$$(\delta|\varphi) = \int_{-\infty}^{+\infty} dx \delta(x)\varphi(x).$$

5.4 Differentiation von Distributionen

Sei $g : \mathbb{R} \rightarrow \mathbb{C}$ eine einmal stetig differenzierbare Funktion. Dann ist

$$(g'|\varphi) = \int_{-\infty}^{+\infty} dx g'(x)\varphi(x)$$

eine reguläre Distribution, die man mit Hilfe partieller Integration so umformen kann:

$$(g'|\varphi) = \int_{-\infty}^{+\infty} dx g'(x)\varphi(x) = \underbrace{g(x)\varphi(x)}_0 \Big|_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} dx g(x)\varphi'(x) = -(g|\varphi').$$

Analog erhält man für eine k -mal stetige differenzierbare Funktion g die Formel

$$(g^{(k)}|\varphi) = (-1)^k (g|\varphi^{(k)}).$$

Für eine **beliebige** Distribution **definiert** man nun einfach

$$(l^{(k)}|\varphi) = (-1)^k (l|\varphi^{(k)}),$$

d.h. Distributionen sind **beliebig oft differenzierbar!**

Beispiel: Die Heavisidesche Stufenfunktion

$$\theta(x) = \begin{cases} 0 & \text{für } x < 0 \\ 1 & \text{für } x \geq 0 \end{cases}$$

ist im klassischen Sinn an der Stelle $x = 0$ nicht differenzierbar. Fasst man die θ -Funktion aber als Distribution auf, so erhält man

$$\begin{aligned} (\theta|\varphi) &= \int_{-\infty}^{+\infty} dx \theta(x)\varphi(x) = \int_0^{+\infty} dx \varphi(x), \\ (\theta'|\varphi) &= -(\theta|\varphi') = - \int_0^{+\infty} dx \varphi'(x) = \varphi(0) = (\delta|\varphi), \end{aligned}$$

d.h. $\theta' = \delta$, bzw. $\theta'(x) = \delta(x)$.

5.5 Rechnen mit Distributionen

Man kann Distributionen zwar i.A. **nicht** miteinander multiplizieren (der Ausdruck $\delta(x)\delta(x)$ ist nicht sinnvoll), aber man kann Distributionen mit ∞ -oft differenzierbaren Funktionen multiplizieren.

Definition: Sei $g \in C^\infty(\mathbb{R})$ und sei l eine beliebige Distribution. Dann ist die Distribution gl durch

$$\varphi \rightarrow (gl|\varphi) = (l|\underbrace{g\varphi}_{\in \mathcal{D}}), \quad \varphi \in \mathcal{D}$$

definiert.

Aufgabe: Zeigen Sie, dass die Produktregel $(gl)' = g'l + gl'$ erfüllt ist.

Für eine reguläre Distribution $g(x)$ ist $g(ax+b)$ ($a \neq 0, b \in \mathbb{R}$) durch

$$(g(ax+b)|\varphi(x)) = \int_{-\infty}^{+\infty} dx g(ax+b)\varphi(x)$$

definiert. Mit Hilfe der Variablentransformation $y = ax+b$ erhält man die Formel

$$(g(ax+b)|\varphi(x)) = \left(g(x) \left| \frac{1}{|a|} \varphi\left(\frac{x-b}{a}\right) \right. \right).$$

Nach dem üblichen Rezept **definiert** man nun für eine **beliebige** Distribution l :

$$(l(ax+b)|\varphi(x)) = \left(l(x) \left| \frac{1}{|a|} \varphi\left(\frac{x-b}{a}\right) \right. \right), \quad a \neq 0.$$

Beispiele:

1. An die Stelle x_0 verschobene Deltafunktion:

$$(\delta(x-x_0)|\varphi(x)) = (\delta(x)|\varphi(x+x_0)) = \varphi(x_0),$$

d.h.

$$\int_{-\infty}^{+\infty} dx \delta(x-x_0)\varphi(x) = \varphi(x_0).$$

2. $(\delta(ax)|\varphi(x)) = (\delta(x)|\frac{1}{|a|}\varphi(\frac{x}{a})) = \frac{1}{|a|}\varphi(0) = \frac{1}{|a|}(\delta(x)|\varphi(x)), \quad a \neq 0, \text{ d.h.}$

$$\delta(ax) = \frac{1}{|a|}\delta(x).$$

Bemerkung: Für eine ∞ -oft differenzierbare Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$, die \mathbb{R} eindeutig auf \mathbb{R} abbildet mit $f'(x) \neq 0 \forall x \in \mathbb{R}$, kann man in analoger Weise $l(f(x))$ definieren. Man kann dann für $l = \delta$ z.B. zeigen, dass

$$\delta(f(x)) = \frac{1}{|f'(x_0)|} \delta(x - x_0),$$

wobei $f(x_0) = 0$.

5.6 Deltafolgen

Die Deltafunktion kann man als Limes von regulären Distributionen erhalten.

Satz: Es sei $f : \mathbb{R} \rightarrow \mathbb{R}$ eine Funktion mit folgenden Eigenschaften:

$$\int_{-\infty}^{+\infty} dx |f(x)| < \infty \text{ (absolut integrierbar) und } \int_{-\infty}^{+\infty} dx f(x) = 1.$$

Dann gilt:

$$\lim_{\lambda \rightarrow \infty} \lambda f(\lambda x) = \delta(x) \quad \Leftrightarrow \quad \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} f\left(\frac{x}{\varepsilon}\right) = \delta(x),$$

d.h.

$$\lim_{\varepsilon \downarrow 0} \left(\frac{1}{\varepsilon} f\left(\frac{x}{\varepsilon}\right) \middle| \varphi(x) \right) = \lim_{\varepsilon \downarrow 0} \int_{-\infty}^{+\infty} dx \frac{1}{\varepsilon} f\left(\frac{x}{\varepsilon}\right) \varphi(x) = \varphi(0), \quad \varphi \in \mathcal{D}.$$

Beispiele:

1. $f(x) = c_{[-\frac{1}{2}, +\frac{1}{2}]}(x) \Rightarrow \delta_\varepsilon(x) = \frac{1}{\varepsilon} c_{[-\frac{\varepsilon}{2}, +\frac{\varepsilon}{2}]}(x) \xrightarrow{\varepsilon \downarrow 0} \delta(x)$, wobei $c_{[a,b]}(x)$ die charakteristische Funktion des Intervalls $[a, b]$ ist.
2. $f(x) = \frac{1}{\pi} \frac{1}{1+x^2}$, $\frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{1}{1+x^2} = \frac{1}{\pi} \arctan x \Big|_{-\infty}^{+\infty} = 1 \Rightarrow \frac{1}{\pi} \frac{\varepsilon}{x^2 + \varepsilon^2} \xrightarrow{\varepsilon \downarrow 0} \delta(x)$.
3. $f(x) = \frac{1}{\sqrt{\pi}} e^{-x^2} \Rightarrow \delta_\varepsilon(x) = \frac{1}{\varepsilon \sqrt{\pi}} e^{-x^2/\varepsilon^2} \xrightarrow{\varepsilon \downarrow 0} \delta(x)$.
4. Eine wichtige Deltafolge ist

$$\delta_\varepsilon(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dk e^{ikx} e^{-\varepsilon|k|}.$$

Die Berechnung des Integrals führt auf die Deltafolge des 2. Beispiels. Es gilt also

$$\lim_{\varepsilon \downarrow 0} \int_{-\infty}^{+\infty} dk e^{ikx} e^{-\varepsilon|k|} = 2\pi\delta(x),$$

wobei in der theoretischen Physik die Kurzschreibweise

$$\int_{-\infty}^{+\infty} dk e^{ikx} = 2\pi\delta(x).$$

üblich ist.

5.7 Greenfunktion des Laplaceoperators

Wir werden nun die am Beginn dieses Kapitels physikalisch motivierte Formel

$$\Delta \frac{1}{|\vec{x}|} = -4\pi\delta^{(3)}(\vec{x})$$

im Rahmen der Theorie der Distributionen beweisen. $\Delta(1/|\vec{x}|)$ ist durch

$$\left(\varphi(\vec{x}) \left| \Delta \frac{1}{|\vec{x}|} \right. \right) = \left(\Delta \varphi(\vec{x}) \left| \frac{1}{|\vec{x}|} \right. \right) = \int_{\mathbb{R}^3} d^3x \frac{1}{|\vec{x}|} \Delta \varphi(\vec{x}) = \lim_{\varepsilon \downarrow 0} \int_{|\vec{x}| \geq \varepsilon} d^3x \frac{1}{|\vec{x}|} \Delta \varphi(\vec{x})$$

gegeben, wobei φ eine Testfunktion ist. Diese ist laut unserer Definition eine Funktion mit kompaktem Träger und somit gibt es eine Zahl $a > 0$, sodass $\varphi(\vec{x})$ für alle \vec{x} mit $|\vec{x}| > a$ verschwindet. Man kann daher das erhaltene Integral als

$$\int_V d^3x \frac{1}{|\vec{x}|} \Delta \varphi(\vec{x}), \quad V := \{\vec{x} | \varepsilon \leq |\vec{x}| \leq a\}$$

schreiben und etwas umformen:

$$\int_V d^3x \frac{1}{|\vec{x}|} \nabla_i \nabla_i \varphi(\vec{x}) = \int_V d^3x \left[\nabla_i \left(\frac{1}{|\vec{x}|} \nabla_i \varphi(\vec{x}) \right) - \nabla_i \left(\varphi(\vec{x}) \nabla_i \frac{1}{|\vec{x}|} \right) + \varphi(\vec{x}) \Delta \frac{1}{|\vec{x}|} \right].$$

Der dritte Term des Integranden auf der rechten Seite verschwindet, da $1/|\vec{x}|$ in einem beliebigen Gebiet, das den Koordinatenursprung nicht enthält (also insbesondere in V) harmonisch ist,

$$\Delta \frac{1}{|\vec{x}|} = 0, \quad \text{für } \vec{x} \neq \vec{0}.$$

Auf die beiden verbleibenden Terme wendet man den Satz von Gauß an,

$$\int_V d^3x \frac{1}{|\vec{x}|} \Delta\varphi(\vec{x}) = \int_{\partial V} df_i \left(\frac{1}{|\vec{x}|} \nabla_i \varphi(\vec{x}) - \varphi(\vec{x}) \nabla_i \frac{1}{|\vec{x}|} \right).$$

Da der Teil des Randes mit $|\vec{x}| = a$ keinen Beitrag liefert, kann man für $|\vec{x}| = \varepsilon$ (Orientierung in Richtung des Ursprungs)

$$df_i = -\varepsilon^2 \frac{x_i}{|\vec{x}|} d\Omega, \quad |\vec{x}| = \varepsilon$$

setzen, wobei $d\Omega = d\cos\theta d\phi$ das Integrationsmaß auf der zweidimensionalen Sphäre S^2 bezeichnet. Man erhält daher

$$\begin{aligned} \int_V d^3x \frac{1}{|\vec{x}|} \Delta\varphi(\vec{x}) &= -\varepsilon^2 \int_{S^2} d\Omega \frac{x_i}{|\vec{x}|} \left(\frac{1}{|\vec{x}|} \nabla_i \varphi(\vec{x}) - \varphi(\vec{x}) \nabla_i \frac{1}{|\vec{x}|} \right) \\ &= -\varepsilon^2 \int_{S^2} d\Omega \left(\frac{1}{\varepsilon} \frac{\partial\varphi}{\partial r} - \varphi \frac{\partial}{\partial r} \frac{1}{r} \right) \Big|_{r=\varepsilon} \\ &= -\varepsilon \int_{S^2} d\Omega \frac{\partial\varphi}{\partial r} \Big|_{r=\varepsilon} + \varepsilon^2 \int_{S^2} d\Omega \varphi \underbrace{\frac{\partial}{\partial r} \frac{1}{r}}_{-\frac{1}{r^2}} \Big|_{r=\varepsilon} \\ &\xrightarrow{\varepsilon \downarrow 0} -4\pi\varphi(\vec{0}). \end{aligned}$$

Somit gilt

$$\left(\varphi(\vec{x}) \Big| \Delta \frac{1}{|\vec{x}|} \right) = -4\pi\varphi(\vec{0}) = -4\pi \left(\varphi(\vec{x}) \Big| \delta^{(3)}(\vec{x}) \right),$$

woraus

$$\Delta \frac{1}{|\vec{x}|} = -4\pi\delta^{(3)}(\vec{x})$$

folgt.

Aufgaben:

1. Zeigen Sie, dass für eine beliebige Dimension $n \geq 3$ gilt, dass

$$\Delta \frac{1}{r^{n-2}} = -(n-2)\Omega_{n-1}\delta^{(n)}(x)$$

gilt, wobei $\Omega_{n-1} = 2\pi^{n/2}/\Gamma(n/2)$ der Flächeninhalt von S^{n-1} ist.

2. Zeigen Sie, dass für $n = 2$ gilt:

$$\Delta \ln \frac{1}{r} = -2\pi\delta^{(2)}(x).$$

Literatur

J. Cigler: Einführung in die Lineare Algebra und Geometrie, 2. Teil, Manz, Wien 1977

H. Neufeld: Skriptum zur Vorlesung „Mathematische Methoden der Physik I“ (Sommersemester 2008)

J. Cigler: Skriptum zur Vorlesung „Verallgemeinerte Funktionen“ (Wintersemester 1975/76), verfasst von J. Hofbauer

I.M. Gelfand, G.E. Schilow: Verallgemeinerte Funktionen (Distributionen) I, Hochschulbücher für Mathematik, Bd. 47, VEB Deutscher Verlag der Wissenschaften, Berlin 1967