

Orthogonale Funktionenräume

Richard Küng*

February 27, 2014

Contents

1	Vektorräume mit Skalarprodukt	2
2	Lineare Abbildungen: Matrizen und Operatoren	8
2.1	Matrizen	8
2.1.1	Diagonalisieren von Matrizen	10
2.1.2	Selbstadjungierte Abbildungen/Symmetrische Matrizen	10
2.2	Operatoren	12
2.2.1	Selbstadjungierte Operatoren	13
2.2.2	Fallstudie: Der Operator $\hat{X} = x$ auf dem Funktionenraum $L^\infty([-1, 1])$	13
2.2.3	Fallstudie: Der Operator $\Delta_\phi := \frac{\partial^2}{\partial \phi^2}$ auf dem Vektorraum $L^2([0, 2\pi])$ der 2π -periodischen Funktionen	14
2.2.4	Fallstudie: Der sphärische Laplaceoperator Δ_{S^2} auf dem Vektorraum $L^2(S^2)$ der Funktionen auf der Einheitssphäre S^2	16
3	Lineare matrix- und operatorwertige Differentialgleichungen	18
3.1	Beispiel (matrixwertige DGL): Drei gekoppelte Oszillatoren	18
3.2	Beispiel (operatorwertige DGL): die Laplacegleichung	21

*e-mail: richard.kueng@physik.uni-freiburg.de

1 Vektorräume mit Skalarprodukt

In diesem Teilabschnitt rekapitulieren wir kurz das fundamentale Grundkonzept der linearen Algebra (Vektorräume) und führen das – in der Geometrie und Physik hochrelevante – Konzept eines Vektorraums mit Skalarprodukt ein. Dabei wollen wir uns vorläufig auf reelle Vektorräume beschränken. Die komplexe Verallgemeinerung ist jedoch straight-forward.

Definition (Vektorraum): Ein reeller Vektorraum (VR) V ist eine (meist unendliche) Menge von Elementen (Vektoren) $v \in V$, welche folgende Eigenschaften erfüllen:

$$\begin{aligned} v, w \in V &\Rightarrow v + w \in V, \\ v \in V, \lambda \in \mathbb{R} &\Rightarrow \lambda v \in V. \end{aligned}$$

Ein Vektorraum ist also abgeschlossen bezüglich der Addition zweier Elemente und der Streckung/Stauchung derselben. Das klassische Beispiel für einen reellen VR ist $\mathbb{R}^n$ mit $n \in \mathbb{N}$ beliebig. Interessanterweise haben auch Funktionen eine Vektorraumstruktur, wie das folgende Beispiel erläutert.

Beispiel: Es sei $I \in \mathbb{R}$ ein beliebiges Intervall und $L(I)$ die Menge aller Funktionen $f : I \rightarrow \mathbb{R}$. Definiert man für beliebige $f, g \in L(I)$ Addition und skalare Multiplikation mit $\lambda \in \mathbb{R}$ punktweise

$$\begin{aligned} (f + g)(x) &:= f(x) + g(x) \quad \forall x \in I, \\ (\lambda f)(x) &:= \lambda f(x) \quad \forall x \in I, \end{aligned}$$

so erhält auch $L(I)$ die Struktur eines Vektorraumes.

Diese scheinbar unschuldige Beobachtung eröffnet ungeahnte Möglichkeiten, da sie es uns erlauben wird, nützliche Vektorraumkonzepte wie Orthogonalität und Basiswahl auf Funktionenräume umzumünzen. Im Folgenden wollen wir daher diese Konzepte kurz rekapitulieren und auf Funktionenräume anwenden.

Definition (Basis): Sei V ein reeller Vektorraum. Eine Menge von n Vektoren $v_1, v_2, \dots, v_n \in V$ ist eine Basis von V falls alle v_i 's linear unabhängig sind (Minimalität) und sich jeder Vektor $v \in V$ als Linearkombination dieser Vektoren dargestellt werden kann (Vollständigkeit). Die Anzahl n der Basiselemente hängt nur von V ab und heißt dessen Dimension.

Vollständigkeit bedeutet formal, dass es für jedes $v \in V$ reelle Zahlen $\lambda_1(v), \lambda_2(v), \dots$ gibt, sodass:

$$v = \sum_{i=1}^n \lambda_i(v) v_i. \tag{1}$$

Informell kann man sich eine Basis als "Baukasten" vorstellen, mit dessen Hilfe man jedes einzelne Element im Vektorraum "zusammenbauen" (aufspannen) kann. Es ist leicht einzusehen, dass es im Allgemeinen unendlich viele verschiedene Basen eines Vektorraumes gibt.

Beispiel: $\mathbb{R}^n$ mit der Standardbasis $e_1 = (1, 0, \dots, 0)^T, e_2 = (0, 1, 0, \dots, 0)^T, \dots, e_n = (0, \dots, 0, 1)^T$.

Beispiel: $\mathbb{R}^3$ mit der Basis $v_1 = (1, 1, 1)^T, v_2 = (0, 1, 1)^T, v_3 = (0, 0, 1)^T$.

Beispiel: Der Vektorraum $L^\infty(\mathbb{R})$ aller analytischen, reellen Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}$ mit der Polynombasis $\{x^i\}_{i=1}^\infty = \{1, x, x^2, x^3, \dots\}$. Dieser Vektorraum ist offenbar unendlich-dimensional. Deshalb ist es prinzipiell subtil Minimalität und Vollständigkeit der Basis $\{x_i\}_{i=0}^\infty$ nachzuweisen. Glücklicherweise nimmt uns Taylor's Theorem diese Arbeit ab. Vollständigkeit folgt in der Tat direkt aus dem Konzept einer Taylorreihe. Sei $f \in L^\infty(\mathbb{R})$ beliebig, dann gilt gemäß Taylor

$$f(x) = \sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{d}{dy} f(y) \Big|_{y=0} \right) x^k = \sum_{k=0}^{\infty} \lambda_k(f) x^k \quad \forall x \in \mathbb{R},$$

wobei wir $\lambda_k(f) := \frac{1}{k!} \left(\frac{d}{dy} f(y) \Big|_{y=0} \right)$ definiert haben. Diese Reihendarstellung entspricht aber gerade einer Basisentwicklung (1) bezüglich $\{x^i\}_{i=0}^\infty$. Intuitiv ist auch klar, dass alle Monome linear unabhängig voneinander sind, und $\{x_i\}_{i=0}^\infty$ somit minimal ist. Auf einen rigorosen Beweis dieser Tatsache wollen wir hier allerdings verzichten.

Beispiel: Der Vektorraum $L^\infty([0, 2\pi])$ aller reellen, analytischen 2π -periodischen Funktionen $f : [0, 2\pi] \rightarrow \mathbb{R}$ mit der trigonometrischen Basis $\{1\} \cup \{\sin(kx), \cos(kx)\}_{k=1}^\infty = \{1, \cos(x), \sin(x), \cos(2x), \sin(2x), \dots\}$. Auch dieser Vektorraum ist unendlich-dimensional. Vollständigkeit folgt aus Fouriers Theorem, welches besagt, dass sich jede Funktion $f \in L^\infty([0, 2\pi])$ als Fourierreihe darstellen lässt:

$$\begin{aligned} f(x) &= \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos(kx) + b_k \sin(kx)) \quad \forall x \in [0, 2\pi], \quad (2) \\ \frac{a_0}{2} &= \frac{1}{2\pi} \int_0^{2\pi} f(x) dx, \\ a_k &= \frac{1}{2\pi} \int_0^{2\pi} f(x) \cos(kx) dx, \\ b_k &= \frac{1}{2\pi} \int_0^{2\pi} f(x) \sin(kx) dx. \end{aligned}$$

Obwohl die Fourierbasis auf den ersten Blick komplizierter erscheint, als die Polynombasis, ist sie doch in vielerlei Hinsicht "netter". Doch um dies präzise zu machen, brauchen wir noch ein weiteres geometrisches Konzept.

Definition (Skalarprodukt): Sei V ein Vektorraum. Eine bilinear¹ Abbildung $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$ ist ein Skalarprodukt, falls sie folgende Eigenschaften erfüllt:

$$\begin{aligned}\langle v, w \rangle &= \langle w, v \rangle \quad (\text{symmetrisch}), \\ \langle v, v \rangle &\geq 0 \quad (\text{positiv}), \\ \langle v, v \rangle = 0 &\Leftrightarrow v = 0 \quad (\text{positiv definit}).\end{aligned}$$

Jedes Skalarprodukt induziert eine zugehörige Norm² $\| \cdot \| : V \rightarrow \mathbb{R}$ via

$$\|v\| = \sqrt{\langle v, v \rangle} \quad \forall v \in V.$$

Skalarprodukt und Norm statten einen beliebigen Vektorraum V mit “Geometrie” aus. Zum Beispiel lassen sich mit ihrer Hilfe Konzepte wie Parallelität und Orthogonalität definieren:

Definition (Geometrie): Sei V ein Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$. Zwei Vektoren v, w sind

$$\begin{aligned}\text{parallel, falls} & \quad \langle v, w \rangle = \|v\| \|w\|, \\ \text{orthogonal, falls} & \quad \langle v, w \rangle = 0.\end{aligned}$$

Zusätzlich nennen wir einen Vektor $v \in V$ normiert, falls $\|v\| = 1$.

Zwei parallele Vektoren sind klarerweise linear abhängig, wohingegen zwei orthogonale Vektoren linear unabhängig sind. Diese beiden geometrischen Einsichten ermöglichen es, “besonders nette Basen” zu definieren:

Definition (orthonormale Basis): Sei V ein Vektorraum. Eine Basis $\{v_i\}_i$ von V ist eine orthonormale Basis (ONB), falls die v_i ’s normiert sind und zueinander orthogonal sind. Konkret bedeutet das:

$$\begin{aligned}\langle v_i, v_j \rangle &= 0 \quad \text{für } i \neq j \quad (\text{Orthogonalität}), \\ \langle v_i, v_i \rangle &= \|v_i\|^2 = 1 \quad (\text{Normiertheit}).\end{aligned}$$

Orthonormale Basen vereinfachen die Basisdarstellung von beliebigen Vektoren $v \in V$ beträchtlich, wie das folgende Theorem zeigt. Das macht sie zu besonders nützlichen Werkzeugen im Umgang mit beliebigen Vektorräumen.

¹Bilinear bedeutet linear in beiden Eingängen. Konkret ist $\langle \cdot, \cdot \rangle$ bilinear, falls für $u, v, w \in V$ und $\lambda, \mu \in \mathbb{R}$ beliebig gilt:

$$\langle \lambda u + \mu v, w \rangle = \lambda \langle u, w \rangle + \mu \langle v, w \rangle \quad \text{und} \quad \langle u, \lambda v + \mu w \rangle = \lambda \langle u, v \rangle + \mu \langle u, w \rangle.$$

²Zur Erinnerung: eine Norm ist eine positive definite Funktion, welche zusätzlich die Dreiecksungleichung erfüllt:

$$\|v + w\| \leq \|v\| + \|w\| \quad \forall v, w \in V.$$

Theorem: Sei V ein Vektorraum mit Orthonormalbasis $\{v_i\}_i$. Dann gilt für ein beliebiges Element $v \in V$:

$$v = \sum_i \lambda_i(v) v_i \quad \text{mit} \quad \lambda_j(v) = \langle v_j, v \rangle \quad \forall j = 1, 2, \dots$$

Diese Expansionsdarstellung ist zudem eindeutig.

Es ist also sehr einfach, die Expansionskoeffizienten λ_i einer ONB zu bestimmen.

Beweis: Da $\{v_i\}_i$ eine Basis von V ist, gilt für $v \in V$ beliebig per Definition

$$v = \sum_i \lambda_i(v) v_i.$$

Linearität des Skalarprodukts zusammen mit Orthogonalität und Normiertheit von $\{v_i\}_i$ garantieren nun zusätzlich

$$\langle v_j, w \rangle = \langle v_j, \sum_i \lambda_i(v) v_i \rangle = \sum_i \lambda_i(v) \langle v_j, v_i \rangle = \lambda_j(v) \langle v_j, v_j \rangle = \lambda_j(v)$$

und wir sind fertig.

Es ist nun höchste Zeit diese geometrischen Konzepte anhand einiger Beispiele zu erläutern.

Beispiel: $\mathbb{R}^n$ mit dem Standardskalarprodukt

$$\langle x, y \rangle = \left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle = \sum_{i=1}^n x_i y_i.$$

Die Standardbasis $e_1 = (1, 0, \dots, 0)^T, \dots, e_n = (0, \dots, 0, 1)^T$ ist das Paradebeispiel einer ONB für $\mathbb{R}^n$. Zusätzlich ermöglicht es das Gram-Schmidtsche Orthogonalisierungsverfahren eine beliebige Basis $\{v_1, \dots, v_n\}$ in eine ONB $\{\tilde{v}_1, \dots, \tilde{v}_n\}$ zu überführen.

Beispiel: Der Raum $L^\infty([-1, 1])$ der analytischen Funktionen $f : [-1, 1] \rightarrow \mathbb{R}$ mit dem Skalarprodukt

$$\langle f, g \rangle := \int_{-1}^1 f(x) g(x) dx. \quad (3)$$

Da dies unsere erste Begegnung mit einem Skalarprodukt auf einem unendlich-dimensionalen Vektorraum darstellt, wollen wir uns diesen Fall genauer ansehen und insbesondere nachweisen, dass (3) tatsächlich ein Skalarprodukt ist.

Symmetrie ist klar, da

$$\langle g, f \rangle = \int_{-1}^1 g(x) f(x) dx = \int_{-1}^1 f(x) g(x) dx = \langle f, g \rangle$$

für beliebige Funktionen f, g gilt. Um Positivität nachzuweisen, bemerken wir

$$\langle f, f \rangle = \int_{-1}^1 f(x) f(x) dx = \int_{-1}^1 f^2(x) dx$$

für f beliebig. Der letzte Ausdruck ist sicher nichtnegativ, da es sich um das Integral einer nicht-negativen Funktion handelt. Positive Definitheit können wir daraus ebenfalls ableiten, da $\int_{-1}^1 f^2(x) dx = 0$ notwendigerweise $f \equiv 0$ verlangt ("Cancellations" sind bei nicht-negativen Funktionen unmöglich). Somit haben wir gezeigt, dass (3) tatsächlich ein valides Skalarprodukt ist.

Wir wissen bereits, dass die Polynome $\{x^i\}_{i=0}^\infty$ eine Basis des Funktionenraumes $L^\infty(\mathbb{R})$ darstellen und Einschränken auf das Intervall $[-1, 1]$ ändert nichts an dieser Tatsache. Daher drängt sich die Frage auf, ob $\{x^i\}_{i=0}^\infty$ ein orthogonales Basissystem ist, oder nicht. Um das zu beantworten, wählen wir x^i, x^j beliebig und berechnen

$$\langle x^i, x^j \rangle = \int_{-1}^1 x^i x^j dx = \int_{-1}^1 x^{i+j} dx = \frac{1}{i+j+1} x^{i+j+1} \Big|_{-1}^1 = \begin{cases} \frac{2}{i+j+1} & \text{falls } i+j+1 \text{ ungerade,} \\ 0 & \text{falls } i+j+1 \text{ gerade.} \end{cases}$$

Folglich sind x^i und x^j im Allgemeinen NICHT orthogonal und $\{x^i\}_{i=0}^\infty$ ist dementsprechend KEINE Orthonormalbasis.

Beispiel: Der Vektorraum $L^\infty([0, 2\pi])$ aller reellen, analytischen 2π -periodischen Funktionen $f : [0, 2\pi] \rightarrow \mathbb{R}$ mit dem Skalarprodukt

$$\langle f, g \rangle := \frac{1}{2\pi} \int_0^{2\pi} f(x) g(x) dx \quad (4)$$

und der trigonometrischen Basis $\{1\} \cup \{\sin(kx), \cos(kx)\}_{k=1}^\infty$. Im Gegensatz zur Polynombasis ist die trigonometrische Basis tatsächlich orthogo-

nal. In der Tat, für $k, l \in \mathbb{N}_+$ beliebig, gilt

$$\begin{aligned}
2\pi \langle 1, \cos(kx) \rangle &= \int_0^{2\pi} 1 \cos(kx) \, dx = \int_0^{2\pi} \cos(kx) \, dx = \frac{1}{k} \sin(kx) \Big|_0^{2\pi} = 0 - 0 = 0, \\
2\pi \langle 1, \sin(kx) \rangle &= \int_0^{2\pi} 1 \sin(kx) \, dx = \int_0^{2\pi} \sin(kx) \, dx = -\frac{1}{k} \cos(kx) \Big|_0^{2\pi} = -\frac{1}{k} - \left(-\frac{1}{k}\right) = 0, \\
2\pi \langle \cos(kx), \cos(lx) \rangle &= \int_0^{2\pi} \cos(kx) \cos(lx) \, dx = \int_0^{2\pi} \frac{1}{2} (\cos((k-l)x) + \cos((k+l)x)) \, dx \\
&= \frac{1}{2} \int_0^{2\pi} \cos((k-l)x) \, dx + \frac{1}{2} \int_0^{2\pi} \cos((k+l)x) \, dx = \frac{1}{2} 0 + \frac{1}{2} 0 = 0, \\
2\pi \langle \sin(kx), \cos(lx) \rangle &= \int_0^{2\pi} \sin(kx) \cos(lx) \, dx = \int_0^{2\pi} \frac{1}{2} (\sin((k-l)x) + \sin((k+l)x)) \, dx = 0, \\
2\pi \langle \sin(kx), \sin(lx) \rangle &= \int_0^{2\pi} \sin(kx) \sin(lx) \, dx = \int_0^{2\pi} \frac{1}{2} (\cos((k-l)x) - \cos((k+l)x)) \, dx = 0.
\end{aligned}$$

Hier haben wir die trigonometrischen Additionstheoreme verwendet und in den späteren Rechnungen auf die vorhergehenden Resultate zurückgegriffen. Um Orthonormalität nachzuweisen, berechnen wir zusätzlich

$$\begin{aligned}
\|1\|^2 &= \langle 1, 1 \rangle = \frac{1}{2\pi} \int_0^{2\pi} 1 \, dx = 1, \\
\|\cos(kx)\|^2 &= \frac{1}{2\pi} \int_0^{2\pi} \cos^2(kx) \, dx = \frac{1}{2\pi} \int_0^{2\pi} \frac{1}{2} (1 + \cos(2kx)) \, dx = \frac{1}{2\pi} (2\pi + 0) = 1, \\
\|\sin(kx)\|^2 &= \frac{1}{2\pi} \int_0^{2\pi} \sin^2(kx) \, dx = \frac{1}{2\pi} \int_0^{2\pi} \frac{1}{2} (1 - \cos(2kx)) \, dx = \frac{1}{2\pi} (2\pi + 0) = 1.
\end{aligned}$$

Bezüglich dem Skalarprodukt (4) ist $\{1\} \cup \{\sin(kx), \cos(kx)\}_{k=1}^\infty$ tatsächlich eine Orthonormalbasis. Dies ist auch der Grund für den Vorfaktor $\frac{1}{2\pi}$ in (4), garantiert er doch Normiertheit der einzelnen Basisfunktionen. Mit diesem Wissen lässt sich die Fourierreihe (2) geometrisch verstehen. Tatsächlich ist sie nichts anderes als das Entwickeln eines Elements $f \in$

$L^\infty([0, 2\pi])$ in der Orthonormalbasis $\{1\} \cup \{\sin(kx), \cos(kx)\}_{k=1}^\infty$:

$$f = \sum_{i=1}^{\infty} \lambda_i(f) v_i = \lambda_0(f) + \sum_{k=1}^{\infty} (\lambda_k(f) \cos(kx) + \mu_k(f) \sin(kx)),$$

$$\lambda_0(f) = \langle 1, f(x) \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(x) dx = \frac{a_0}{2},$$

$$\lambda_k(f) = \langle \cos(kx), f(x) \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(x) \cos(kx) dx = a_k,$$

$$\mu_k(f) = \langle \sin(kx), f(x) \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(x) \sin(kx) dx = b_k.$$

Damit haben wir die Fourierreihe hoffentlich “entzaubert”. Sie ist kein komischer Voodoo-Trick für periodische Funktionen, sondern – im Gegenteil – lediglich eine weitere Anwendung eines euch wohlvertrauten Konzeptes: Das Entwickeln von Vektoren in einer ONB.

2 Lineare Abbildungen: Matrizen und Operatoren

Den Kern dieses Kapitels bildet die folgende Definition.

Definition (lineare Abbildung): Es seien V, W zwei Vektorräume. Eine Abbildung $f : V \rightarrow W$ heißt linear, falls sie die Vektorraumstruktur erhält. Konkret bedeutet dies, dass sie Vektoren $v \in V$ wieder auf Vektoren $f(v) = w \in W$ abbildet:

$$\begin{aligned} f(v_1 + v_2) &= f(v_1) + f(v_2) \quad \forall v_1, v_2 \in V, \\ f(\lambda v) &= \lambda f(v) \quad \forall v \in V, \forall \lambda \in \mathbb{R}. \end{aligned}$$

Für unsere Zwecke ist diese Definition ein bisschen zu allgemein. Wir wollen uns daher auf Selbstabbildungen $f : V \rightarrow V$ beschränken.

2.1 Matrizen

Es sei V ein endlich-dimensionaler Vektorraum mit Basis $\{v_i\}_{i=1}^n$. Eine lineare Abbildung $f : V \rightarrow V$ lässt sich durch ihre Wirkung auf diese Basisvektoren vollständig charakterisieren. Wegen der definierenden Eigenschaft einer Basis lässt sich nämlich jeder Vektor $v \in V$ als Linearkombination $v = \sum_{i=1}^n \lambda_i v_i$ darstellen. Zusammen mit den fundamentalen Eigenschaften einer linearen Abbildung gilt also:

$$f(v) = f\left(\sum_{i=1}^n \lambda_i v_i\right) = \sum_{i=1}^n \lambda_i f(v_i).$$

Somit ist die Wirkung der linearen Abbildung f durch ihr Transformationsverhalten bezüglich der n Basiselement $w_i = f(v_i) \in V$ vollständig charakterisiert! Entwickeln wir diese $w_i \in V$ wieder bezüglich der Basis $\{v_i\}_{i=1}^n$, lässt sich dieses Transformationsverhalten besonders ökonomisch niederschreiben:

1. Bezüglich der Basis $\{v_i\}_{i=1}^n$ lässt sich jeder Vektor $v \in V$ als "Zahlenarray" $(\lambda_1, \dots, \lambda_n)^T$ darstellen.
2. Die lineare Abbildung f überträgt jeden Basisvektor $v_i \in V$ auf einen Vektor $w_i = f(v_i) \in V$.
3. Erneut lässt sich jeder solche Vektor $w_i \in V$ als Linearkombination $w_i = \sum_{j=1}^n w_{ji} v_j$ darstellen und entspricht somit dem "Zahlenarray" $(w_{i1}, \dots, w_{in})^T$.
4. Kombiniert man diese Basisentwicklungen erhalten wir folgende Matrix und das vertraute Vektor-Matrix-Zusammenspiel:

$$f(v) = \begin{pmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & & w_{2n} \\ \vdots & & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n \lambda_i w_{1i} \\ \vdots \\ \sum_{i=1}^n \lambda_i w_{ni} \end{pmatrix}.$$

Unsere Matrix gibt also tatsächlich das exakte Verhalten der linearen Abbildung f wieder. Allerdings hängt sie essentiell von der gewählten Basis ab! Die lineare Abbildung selbst ist natürlich basisunabhängig.

Nota Bene: Das Konzept einer linearen Abbildung ist eindeutig. Ihre Darstellung als Matrix ist es jedoch nicht (es gibt i.A. unendlich viele verschiedene Matrizen, welche dieselbe lineare Abbildung darstellen).

Es ist leicht einzusehen, dass Matrizen von reellen Selbstabbildungen $f: V \rightarrow V$ (mit $\dim V = n$) einen Vektorraum bilden. Diesen bezeichnen wir als $\mathcal{B}(V)$ (den Raum der linearen Abbildungen von V nach V).

Definition (Äquivalenz von Matrizen): Zwei Matrizen $X, Y \in M_{n \times n}(\mathbb{R})$ sind äquivalent, falls eine invertierbare Matrix B existiert, sodass

$$X = BYB^{-1} \quad (5)$$

gilt. Zwei äquivalente Matrizen beschreiben dieselbe Selbstabbildung $f: V \rightarrow V$.

Der Grund für diese Definition sollte einleuchten. Eine beliebige invertierbare Matrix B beschreibt nämlich (eindeutig) einen Basiswechsel in V . Bedingung (5) verlangt somit, dass sich Y durch einen Basiswechsel in X überführen lässt.

2.1.1 Diagonalisieren von Matrizen

Es sei V ein n -dimensionaler (endlicher) Vektorraum. Eine $n \times n$ -Matrix M ist diagonalisierbar, falls sie äquivalent zu einer Diagonalmatrix $D = \text{diag}(\lambda_1, \dots, \lambda_n)$ ist, d.h. falls es eine invertierbare $n \times n$ -Matrix gibt, welche

$$M = BDB^{-1}$$

garantiert. Bis auf Vertauschen der λ_i 's ist diese Matrix eindeutig. Nicht jede Matrix ist diagonalisierbar, ein Gegenbeispiel ist durch $M = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ gegeben.

Mit unserer Interpretation von Matrizen als Manifestierung von linearen Abbildung lässt sich die Diagonalisierung einer Matrix auf anschauliche Art interpretieren: Die invertierbare Matrix B entspricht einer Basiswechselmatrix $\{e_i\}_{i=1}^n \rightarrow \{b_i\}_{i=1}^n$. Bezüglich dieser neuen Basis ist die zugrunde liegende lineare Abbildung f durch die Diagonalmatrix $D(f) = \text{diag}(\lambda_1, \dots, \lambda_n)$ vollständig charakterisiert. Insbesondere bedeutet dies, dass

$$\begin{aligned} f(b_1) &= \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \lambda_1 b_1, \\ &\vdots \\ f(b_n) &= \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = \lambda_n b_n. \end{aligned}$$

Die lineare Abbildung f bildet somit die Vektoren $b_i \in V$ auf ein λ_i -faches von sich selbst ab. Aber das ist genau die formale Definition von Eigenwert und zugehörigem Eigenvektor! Darüber hinaus ist f eine lineare Abbildung und damit basisunabhängig. Die Eigenvektoren $b_i \in V$ sind somit direkt über die lineare Abbildung f – und nicht über eine explizite Matrixdarstellung – definiert! Somit sind sie genau wie die lineare Abbildung basisunabhängig (das sollten sie auch sein, da sie fundamental Elementen des Vektorraums V entsprechen und solche immer basisunabhängig sind).

Nota Bene: Eigenvektoren und Eigenwerte einer Matrix sind basisunabhängige Entitäten der der Matrix zugrundeliegenden linearen Abbildung (vgl. Fixpunkte).

2.1.2 Selbstadjungierte Abbildungen/Symmetrische Matrizen

Es sei nun V ein (endlicher) n -dimensionaler Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$.

Definition (selbstadjungierte Abbildung): Eine lineare Abbildung $f \in \mathcal{B}(V)$ ist selbstadjungiert (s.a.), falls

$$\langle v, f(w) \rangle = \langle f(v), w \rangle \quad \forall v, w \in V \quad (6)$$

gilt.

Hat f bezüglich einer beliebigen Basis die Matrixdarstellung $A \in M_{n \times n}(\mathbb{R})$, so ist die definierende Eigenschaft (6) äquivalent zu

$$A^T = A.$$

Somit liegen den euch wohlvertrauten symmetrischen Matrizen selbstadjungierte Abbildungen zugrunde. Auf einen Beweis dieser Äquivalenz wollen wir hier verzichten (er ist ziemlich “straight forward”).

Selbstadjungierte Abbildungen sind in vielerlei Hinsicht besonders “nett”, wie das folgende Theorem zeigt.

Theorem: Es sei $f \in \mathcal{B}(V)$ eine selbstadjungierte Abbildung. Dann sind die Eigenwerte von $\lambda_1, \dots, \lambda_n$ reell und die zugehörigen Eigenvektoren können orthonormal zueinander gewählt werden.

Dass die Eigenwerte einer reellen Matrix reell sind mag banal klingen, ist es aber nicht. Ihr wisst, dass die Eigenwerte einer Matrix A den Nullstellen des charakteristischen Polynoms $p_A(x) = \det(A - x\mathbb{I})$ entsprechen. Allerdings gibt es durchaus reelle Polynome welche komplexe Nullstellen haben, z.B. $p(x) = x^2 + 1$. Folgerichtig hat eine reelle Matrix nicht zwingend reelle Eigenwerte!

Beweis: Für den Beweis, dass die Eigenwerte reell sind, brauchen wir einen komplexen Vektorraum V mit sesquilinearem Skalarprodukt. Das ist weiters kein Problem, da sich jede reelle lineare Abbildung in einen komplexen Vektorraum einbetten lässt. Für ein komplexes Skalarprodukt gilt $\langle v, w \rangle = \overline{\langle w, v \rangle}$ für $v, w \in V$ beliebig. Es sei nun $b_i \in V$ ein normierter Eigenvektor zum Eigenwert λ_i . Dann gilt

$$\begin{aligned} \lambda_i &= \lambda_i \langle b_i, b_i \rangle = \langle b_i, \lambda_i b_i \rangle = \langle b_i, f(b_i) \rangle = \langle f(b_i), b_i \rangle = \langle \lambda_i b_i, b_i \rangle \\ &= \overline{\langle b_i, \lambda_i b_i \rangle} = \bar{\lambda}_i \langle b_i, b_i \rangle = \bar{\lambda}_i. \end{aligned}$$

Somit muss $\bar{\lambda}_i = \lambda_i$ gelten, was nur für reelle Zahlen möglich ist.

Um Orthonormalität der Eigenvektoren zu zeigen, nehmen wir zunächst an, dass b_i und b_j Eigenvektoren von verschiedenen Eigenwerten $\lambda_i \neq \lambda_j$ sind. Da die Eigenwerte zudem reell sind, folgern wir

$$\lambda_i \langle b_i, b_j \rangle = \langle \lambda_i b_i, b_j \rangle = \langle f(b_i), b_j \rangle = \langle b_i, f(b_j) \rangle = \langle b_i, \lambda_j b_j \rangle = \lambda_j \langle b_i, b_j \rangle.$$

Das impliziert

$$(\lambda_i - \lambda_j) \langle b_i, b_j \rangle = 0$$

und wegen $\lambda_i \neq \lambda_j$ muss somit $\langle b_i, b_j \rangle = 0$ gelten. Nun bleibt nur noch der Spezialfall von degenierten Eigenwerten (d.h. Eigenwerte mit mehr

als einem zugehörigen Eigenvektor) zu behandeln. Es sei λ_i ein s -fach degenerierter Eigenwert (falls es ihn gibt) und $b'_1, \dots, b'_s$ alle zugehörigen Eigenvektoren. Dann gilt offenbar für beliebige Linearkombinationen $v' = \sum_{i=1}^s \nu_i b'_i$:

$$f(v') = f\left(\sum_{i=1}^s \nu_i b'_i\right) = \sum_{i=1}^s \nu_i f(b'_i) = \lambda_i \sum_{i=1}^s \nu_i b'_i = \lambda_i v'.$$

Beliebige Linearkombinationen von Eigenvektoren mit demselben Eigenwert sind also wieder Eigenvektoren zum selben Eigenwert. Diese Freiheit erlaubt es uns nun aber das Gram-Schmitt-Verfahren anzuwenden und somit die Eigenvektoren $b'_1, \dots, b'_s$ zu "orthonormalisieren".

2.2 Operatoren

Bislang haben wir uns mit linearen Abbildungen auf endlichen Vektorräumen beschäftigt und ihre Darstellung als Matrizen rekapituliert. In diesem Kapitel wollen wir uns auf lineare Abbildungen auf unendlich-dimensionalen Vektorräumen stürzen. Diese werden gemeinhin als Operatoren bezeichnet und mit einem "Dach" $\hat{}$ über dem Ausdruck gekennzeichnet. Das Grundkonzept – lineare Abbildungen – bleibt dabei unverändert (die Definition am Beginn dieses Kapitels ist agnostisch bezüglich der Dimension des Vektorraumes V). Da wir jedoch mit dem Operatorkonzept das Regime der linearen Algebra verlassen, wollen wir zunächst einige – für die Physik hochrelevant – Beispiele bringen:

- Der Operator $\hat{X} : f \mapsto xf(x)$ (punktweise) auf dem Funktionenraum $L^0(\mathbb{R})$. Dieser Operator ist offensichtlich linear, da für $f(x), g(x) \in L^0(\mathbb{R})$ und $\lambda \in \mathbb{R}$ beliebig gilt:

$$\begin{aligned} \left(\hat{X}(f+g)\right)(x) &= x(f(x)+g(x)) = xf(x) + xg(x) = \left(\hat{X}f\right)(x) + \left(\hat{X}g\right)(x), \\ \left(\hat{X}(\lambda f)\right)(x) &= x\lambda f(x) = \lambda xf(x) = \lambda \left(\hat{X}(f)\right)(x) \quad \forall x \in \mathbb{R}. \end{aligned}$$

- Der Operator $\hat{D} := \frac{d}{dx}$ auf dem Funktionenraum $L^\infty(\mathbb{R})$. Auch dieser Operator ist linear, da für $f(x), g(x) \in L^0(\mathbb{R})$ und $\lambda, \mu \in \mathbb{R}$ beliebig gilt:

$$\begin{aligned} \hat{D}(\lambda f + \mu g)(x) &= \frac{d}{dx}(\lambda f(x) + \mu g(x)) = \lambda \frac{d}{dx}f(x) + \mu \frac{d}{dx}g(x) \\ &= \lambda \hat{D}(f)(x) + \mu \hat{D}(g)(x) \quad \forall x \in \mathbb{R}. \end{aligned}$$

Die Notation mag hier zunächst verwirrend sein, konzeptionell gilt jedoch einfach: Der Operator $\hat{X}$ multipliziert eine beliebige Funktion f mit x und bildet sie somit wieder auf eine Funktion $xf(x) \forall x \in \mathbb{R}$ ab. Zudem erhält er die Vektorraumstruktur, da Multiplikation mit x linear ist. Ähnliches gilt für den Ableitungsoperator $\hat{D}$, da ableiten ebenfalls linear ist $((f+g)' = f' + g')$.

Zudem machen Konzepte wie Eigenwerte und Eigenvektoren (Eigenfunktionen) für Operatoren Sinn. Das liegt daran, dass sie fundamentale Eigenschaften dieser linearen Abbildung sind – und Operatoren sind nichts anderes als lineare Abbildungen. Zum Beispiel ist es leicht einzusehen, dass $e^{\lambda x} \in L^\infty(\mathbb{R})$ eine Eigenfunktion des Operators $\hat{D}$ mit zugehörigem Eigenwert λ ist:

$$\hat{D}(e^x) = \frac{d}{dx} e^{\lambda x} = \lambda e^x.$$

Diesen “Linear-Algebra-Approach” für Operatoren wollen wir im Folgenden beträchtlich vertiefen.

Vorschau: In der Quantenmechanik (Schrödinger-approach) sind Operatoren von fundamentaler Bedeutung. Die auch aus der Hamiltonischen Mechanik wohlvertrauten Orts- und Impulskoordinaten q und p werden in der Quantenmechanik durch lineare Operatoren $\hat{X} = x$ und $\hat{P} = -i\hbar\nabla$ auf Hilberträumen (komplexer Vektorraum mit Skalarprodukt) ersetzt.

2.2.1 Selbstadjungierte Operatoren

Es sei nun V ein unendlich-dimensionaler Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$. Wir bezeichnen einen Operator $\hat{O} \in \mathcal{B}(V)$ als selbstadjungiert, falls er Bedingung (6) erfüllt, also:

$$\langle f, \hat{O}(g) \rangle = \langle \hat{O}(f), g \rangle \quad \forall f, g \in V.$$

Ähnlich wie symmetrische Matrizen, haben auch selbst-adjungierte Operatoren “tolle Eigenschaften”. Tatsächlich gilt obiges Theorem auch für Operatoren. Das kann man zum Beispiel daran erkennen, dass wir im Beweis nie Endlichdimensionalität des zugrundeliegenden Vektorraumes gebraucht haben. Allerdings kann diese Unendlichkeit durchaus zu subtilen Problemen führen. Ich möchte hier allerdings nicht näher darauf eingehen und verweise auf Standardvorlesungen in Funktionalanalysis. Nichtsdestotrotz gilt:

Theorem: Es sei $\hat{O} \in \mathcal{B}(V)$ eine selbstadjungierte Abbildung. Dann sind die Eigenwerte von $\lambda_1, \dots, \lambda_n$ reell und die zugehörigen Eigenfunktionen können orthonormal zueinander gewählt werden.

Wir wollen uns nun mit ein paar besonders wichtige selbst-adjungierte Operatoren beschäftigen

2.2.2 Fallstudie: Der Operator $\hat{X} = x$ auf dem Funktionenraum $L^\infty([-1, 1])$

Im vorigen Kapitel haben wir bereits das Skalarprodukt

$$\langle f, g \rangle = \int_{-1}^1 f(x) g(x) dx$$

auf $L^\infty([-1, 1])$ eingeführt. Bezüglich diesem Skalarprodukt ist $\hat{X}$ selbstadjungiert, da

$$\langle \hat{X}(f), g \rangle = \int_{-1}^1 x f(x) g(x) dx = \int_{-1}^1 f(x) x g(x) dx = \langle f, \hat{X}(g) \rangle.$$

Somit wissen wir, dass $\hat{X}$ nur reelle Eigenwerte hat und seine Eigenfunktionen orthonormal zueinander gewählt werden können. Allerdings wollen wir uns hier nicht mit dem expliziten Suchen nach unendlich (!!!) vielen Eigenwerten und Eigenfunktionen aufhalten. Anstatt dessen bemerken wir, dass für ein beliebiges Monom x^k gilt:

$$\hat{X}(x^k) = x x^k = x^{k+1}.$$

Somit bildet $\hat{X}$ Monome auf Monome ab und bezüglich der Standardpolynombasis $\{x^i\}_{i=0}^\infty$ können wir $\hat{X}$ formal als “unendlich-dimensionale” Matrix darstellen:

$$\hat{X} = \begin{pmatrix} 0 & \dots & & \\ 1 & 0 & \dots & \\ \vdots & 1 & 0 & \\ & \vdots & \ddots & \ddots \end{pmatrix}.$$

Klarerweise ist diese Matrix nicht diagonal (die Monome sind ja keine Eigenfunktionen von $\hat{X}$). Diese Äquivalenz von Operatoren und unendlich-dimensionalen Matrizen kann präzise gemacht werden und schlägt eine Brücke zwischen der Schrödingerschen und der Heisenbergschen Quantenmechanik. Heisenbergs Approach zur Quantenmechanik beruht nämlich auf derartigen unendlich-großen Matrizen!

2.2.3 Fallstudie: Der Operator $\Delta_\phi := \frac{\partial^2}{\partial \phi^2}$ auf dem Vektorraum $L^2([0, 2\pi])$ der 2π -periodischen Funktionen

Zu diesem Vektorraum haben wir bereits das Skalarprodukt

$$\langle f, g \rangle = \int_0^{2\pi} f(\phi) g(\phi) d\phi$$

eingeführt. Bezüglich dieses Skalarprodukts ist der Operator $\hat{\Delta}_\phi$ selbstadjungiert. Um das zu sehen, wählen wir $f, g \in L^2([0, 2\pi])$ beliebig und integrieren zweimal

partiell:

$$\begin{aligned}
\langle f, \Delta_\phi g \rangle &= \int_0^{2\pi} f(\phi) \left(\frac{\partial^2}{\partial \phi^2} g(\phi) \right) d\phi = \int_0^{2\pi} f(\phi) g''(\phi) d\phi \\
&= f(\phi) g'(\phi) \Big|_0^{2\pi} - \int_0^{2\pi} f'(\phi) g'(\phi) d\phi \\
&= 0 - f'(\phi) g(\phi) \Big|_0^{2\pi} + \int_0^{2\pi} f''(\phi) g(\phi) d\phi \\
&= 0 + \int_0^{2\pi} \left(\frac{\partial^2}{\partial \phi^2} f(\phi) \right) g(\phi) d\phi = \langle \Delta_\phi f, g \rangle.
\end{aligned}$$

Hier haben wir verwendet, dass sich aufgrund der Periodizität von f und g die jeweiligen Randterme gegenseitig aufheben. Aus dem Wissen, dass Δ_ϕ können wir sofort folgern, dass seine Eigenwerte reell sind und die Eigenfunktionen orthonormal zueinander gewählt werden können. Dies hat weitreichende Konsequenzen, da wir einige Eigenfunktionen von Δ_ϕ bereits kennen! Einmal mehr handelt es sich dabei um die trigonometrische Basis $\{1\} \cup \{\cos(k\phi), \sin(k\phi)\}_{k=1}^\infty$:

$$\begin{aligned}
\Delta_\phi 1 &= 0 \quad \Rightarrow \quad 1 \text{ ist eine EF mit EW } 0, \\
\Delta_\phi \cos(k\phi) &= -k^2 \cos(k\phi) \quad \Rightarrow \quad \cos(k\phi) \text{ ist eine EF mit EW } -k^2, \\
\Delta_\phi \sin(k\phi) &= -k^2 \sin(k\phi) \quad \Rightarrow \quad \sin(k\phi) \text{ ist eine EF mit EW } -k^2.
\end{aligned}$$

Tatsächlich sind dies alle Eigenfunktionen von Δ_ϕ , da die trigonometrischen Funktionen eine Basis $\{1\} \cup \{\cos(k\phi), \sin(k\phi)\}_{k=1}^\infty$ darstellen und somit den gesamten $L^2([0, 2\pi])$ aufspannen.

Zudem können wir Orthonormalität der trigonometrischen Basis mit unseren “advanced tools” viel cleverer beweisen. Schließlich wissen wir ja bereits, dass Eigenfunktionen selbstadjungierter Operatoren automatisch orthogonal zueinander sind, wenn sie verschiedene Eigenwerte haben. Daraus können wir folgende Orthogonalitätsrelationen bequem ableiten:

$$\begin{aligned}
\langle 1, \cos(k\phi) \rangle &= 0 \quad \forall k \geq 1, \\
\langle 1, \sin(k\phi) \rangle &= 0 \quad \forall k \geq 1, \\
\langle \cos(k\phi), \cos(l\phi) \rangle &= 0 \quad \forall k \neq l, \\
\langle \cos(k\phi), \sin(l\phi) \rangle &= 0 \quad \forall k \neq l, \\
\langle \sin(k\phi), \sin(l\phi) \rangle &= 0 \quad \forall k \neq l.
\end{aligned}$$

Somit müssen wir einzig und alleine $\langle \cos(k\phi), \sin(k\phi) \rangle = 0$ noch herleiten, da $\cos(k\phi)$ und $\sin(k\phi)$ Eigenfunktionen mit gleichem Eigenwert (nämlich $-k^2$) sind. Abschließend wollen wir noch erwähnen, dass sich Δ_ϕ bezüglich der

trigonometrischen Basis als folgende unendlichdimensionale Matrix schreiben lässt:

$$\Delta_\phi = - \begin{pmatrix} 0 & & & & & 0 \\ & 1 & & & & \\ & & 1 & & & \\ & & & 4 & & \\ & & & & 4 & \\ & & & & & 9 \\ & & & & & & 9 \\ 0 & & & & & & & \ddots \end{pmatrix}.$$

Im Gegensatz zur vorigen Fallstudie, ist diese Matrix nun diagonal. Das verwundert nicht weiter, da wir für die Darstellung die zugehörige Eigenbasis verwendet haben.

2.2.4 Fallstudie: Der sphärische Laplaceoperator Δ_{S^2} auf dem Vektorraum $L^2(S^2)$ der Funktionen auf der Einheitssphäre S^2

Hier wollen wir uns dem sphärischen Laplaceoperator

$$\Delta_{S^2} = \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2}$$

auf der Einheitssphäre $S^2 = \{x \in \mathbb{R}^3 : \|x\|_2^2 = 1\}$ widmen. Wählen wir die Kugelkoordinaten θ, ϕ für das Parametrisieren von Funktionen auf der Sphäre, so macht folgende Definition eines Skalarproduktes Sinn:

$$\langle f, g \rangle := \int_0^\pi \int_0^{2\pi} f(\theta, \phi) g(\theta, \phi) \sin \theta d\theta d\phi.$$

Es sollte offensichtlich sein, dass es sich dabei um ein valides Skalarprodukt handelt. Zusätzlich ist Δ_{S^2} bezüglich dieses Skalarproduktes selbstadjungiert. In der Tat gilt für $f, g \in L^2(S^2)$ beliebig:

$$\begin{aligned} \langle f, \Delta_{S^2} g \rangle &= \int_0^\pi \int_0^{2\pi} f(\theta, \phi) \left(\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial g(\theta, \phi)}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2 g(\theta, \phi)}{\partial \phi^2} \right) \sin \theta d\phi d\theta \\ &= \int_0^{2\pi} \left[\int_0^\pi f(\theta, \phi) \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial g(\theta, \phi)}{\partial \theta} \right) d\theta \right] d\phi + \int_0^\pi \frac{1}{\sin \theta} \left[\int_0^{2\pi} f(\theta, \phi) \frac{\partial^2 g(\theta, \phi)}{\partial \phi^2} d\phi \right] d\theta. \end{aligned}$$

Im zweiten Integral können wir nun zweimal partiell integrieren und die ϕ -Ableitung von g auf f schieben. Dabei verschwinden die auftretenden Randterme wegen der 2π -Periodizität der Funktionen in ϕ . Im ersten Teil vernachlässigen wir fürs erste die ϕ -Integration und integrieren ebenfalls zweimal partiell über

θ :

$$\begin{aligned}
& \int_0^\pi f(\theta, \phi) \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial g(\theta, \phi)}{\partial \theta} \right) d\theta \\
&= f(\theta, \phi) \sin \theta \frac{\partial g(\theta, \phi)}{\partial \theta} \Big|_{\theta=0}^{\theta=\pi} - \int_0^\pi \sin(\theta) \frac{\partial f(\theta, \phi)}{\partial \theta} \frac{\partial g(\theta, \phi)}{\partial \theta} d\theta \\
&= 0 - \sin(\theta) \frac{\partial f(\theta, \phi)}{\partial \theta} g(\theta, \phi) \Big|_{\theta=0}^{\theta=\pi} + \int_0^\pi \frac{\partial}{\partial \theta} \left(\sin(\theta) \frac{\partial f(\theta, \phi)}{\partial \theta} \right) g(\theta, \phi) d\theta \\
&= \int_0^\pi \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin(\theta) \frac{\partial f(\theta, \phi)}{\partial \theta} \right) g(\theta, \phi) d\theta.
\end{aligned}$$

Hier verschwinden die Randterme wegen $\sin(0) = \sin(\pi) = 0$. Fügen wir diese Resultate wieder zusammen, erhalten wir somit tatsächlich:

$$\begin{aligned}
\langle f, \Delta_{S^2} g \rangle &= \int_0^{2\pi} \left[\int_0^\pi f(\theta, \phi) \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial g(\theta, \phi)}{\partial \theta} \right) d\theta \right] d\phi + \int_0^\pi \frac{1}{\sin \theta} \left[\int_0^{2\pi} f(\theta, \phi) \frac{\partial^2 g(\theta, \phi)}{\partial \phi^2} d\phi \right] d\theta \\
&= \int_0^{2\pi} \left[\int_0^\pi \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin(\theta) \frac{\partial f(\theta, \phi)}{\partial \theta} \right) g(\theta, \phi) d\theta \right] d\phi + \int_0^\pi \frac{1}{\sin \theta} \left[\int_0^{2\pi} \frac{\partial^2 f(\theta, \phi)}{\partial \phi^2} g(\theta, \phi) d\phi \right] d\theta \\
&= \langle \Delta_{S^2} f, g \rangle.
\end{aligned}$$

Da wir soeben bewiesen haben, dass Δ_{S^2} selbstadjungiert ist, können wir sofort auf reelle Eigenwerte und orthogonale Eigenfunktionen schließen. Da dieser Operator so wichtig ist, sind diese selbstverständlich bekannt. Allerdings ist der "Diagonalisierungsaufwand" beträchtlich und wir wollen uns daher damit begnügen, die Resultate aufzuzählen:

- Die Eigenwerte von Δ_{S^2} entsprechen $\{-l(l+1)\}_{l=0}^\infty = \{0, -2, -6, \dots\}$. Dabei ist der Eigenwert $-l(l+1)$ $(2l+1)$ -fach entartet. Für jedes $-l(l+1)$ gibt es somit $(2l+1)$ verschiedene Eigenfunktionen, welche orthogonal zueinander gewählt werden können.
- Die Kugelflächenfunktionen $\{Y_l^m(\theta, \phi) : l \in \mathbb{N}, -l \leq m \leq l\}$ sind die Eigenfunktionen von Δ_{S^2} . Konkret gilt für $l \in \mathbb{N}$ beliebig:

$$\Delta_{S^2} Y_l^m(\theta, \phi) = -l(l+1) Y_l^m(\theta, \phi) \quad \forall -l \leq m \leq l.$$

Somit codiert der Suffix l den zugehörigen Eigenwert $-l(l+1)$. Das Superscript m rangiert zwischen $-l$ und l und labelt $(2l+1)$ verschiedene Eigenfunktionen zu betreffendem Eigenwert (das entspricht genau der $(2l+1)$ -fachen Entartung des Eigenwertes $-l(l+1)$). Diese Eigenfunktionen sind zudem so konstruiert, dass sie orthogonal zueinander stehen

(vgl. Gram-Schmitt). Konkret gilt also für $l \in \mathbb{N}$ beliebig:

$$\langle Y_l^m, Y_l^{m'} \rangle = 0 \quad \text{für } m \neq m'.$$

Mit unserem Theorem über selbstadjungierte Operatoren können wir zudem sofort folgern, dass

$$\langle Y_l^m, Y_{l'}^{m'} \rangle = 0 \quad \text{für } l \neq l'$$

gelten muss. Schließlich sind für $l' \neq l$ Y_l^m und $Y_{l'}^{m'}$ Eigenfunktionen zu verschiedenen Eigenwerten $(-l(l+1) \neq -l'(l'+1))$.

Nota Bene: Bei den Kugelflächenfunktionen ist Orthogonalität für verschiedene m 's konstruiert (vgl. Gram-Schmitt), Orthogonalität für verschiedene l 's jedoch eine fundamentale Eigenschaft des Operators Δ_{S^2} !

Daraus schließen wir, dass die Kugelflächenfunktionen eine orthogonale Funktionenfamilie darstellen. Tatsächlich sind sie sogar eine orthogonale Funktionenbasis. Minimalität folgt aus der Orthogonalität, wohingegen Vollständigkeit subtiler ist. Auf einen Beweis dieser Tatsache wollen wir hier allerdings nicht näher eingehen.

3 Lineare matrix- und operatorwertige Differentialgleichungen

Dieses Kapitel bildet den Abschluss und Höhepunkt unserer Abhandlung über unendlich-dimensionale Vektorräume und lineare Operatoren. Dabei wollen wir uns zunächst auf matrixwertige Differentialgleichungen fokussieren und Lösungsmethoden rekapitulieren. Anschließend benützen wir die “take-home message”

Funktionenräume sind unendlichdimensionale Vektorräume und lineare Operatoren unendlichdimensionale Matrizen

um unser Wissen über Matrixdifferentialgleichungen dahingehend zu verwenden, mehrdimensionale Differentialgleichungen zu lösen.

3.1 Beispiel (matrixwertige DGL): Drei gekoppelte Oszillatoren

Hier wollen wir uns auf ein konkretes Beispiel einer gewöhnlichen linearen matrixwertigen Differentialgleichung beschränken. Für eine genauere Studie dieser Konzepte verweise ich auf meine Notiz “Differentialgleichungen und Matrizen”, welches auf meiner Homepage verfügbar ist. Das folgende Beispiel findet sich ebenfalls dort.

Wir betrachten 3 gekoppelte harmonische Oszillatoren mit gleicher Masse m und identischer Federkonstante k (vgl. Beispiel 9.7.2 im Script von Wulf Fetscher, ETH Zürich 2007). Wir bezeichnen die zeitabhängige Auslenkung

des jeweiligen Oszillators aus der Ruhelage mit $x_i(t)$ ($i = 1, 2, 3$) und betrachten alle drei Auslenkungen zusammen als einen Auslenkungsvektor $\vec{x}(t)^T = (x_1(t), x_2(t), x_3(t))^T \in \mathbb{R}^3$. Mit dem Hooke'schen Gesetz ergeben sich folgende Bewegungsgleichungen:

$$\begin{aligned} m\ddot{x}_1 &= k(x_2 - x_1), \\ m\ddot{x}_2 &= -k(x_2 - x_1) + k(x_3 - x_2), \\ m\ddot{x}_3 &= -k(x_3 - x_1). \end{aligned}$$

Mithilfe der Definition $\omega := \sqrt{k/m}$ (Eigenfrequenz) ist dieses Gleichungssystem äquivalent zu folgender Matrixgleichung:

$$\ddot{\vec{x}}(t) + \omega^2 A \vec{x}(t) = 0, \quad \text{mit} \quad A = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix} \quad (7)$$

Abgesehen von der Matrix sieht diese Differentialgleichung wie die gewöhnliche Gleichung des harmonischen Oszillators aus. Tatsächlich lässt sie sich in eine solche überführen, wenn wir die Struktur der Matrix ausnützen. Diese ist nämlich offenbar symmetrisch und ist somit bezüglich einer Orthogonalbasis aus Eigenvektoren diagonalisierbar. Bezeichnen wir diese (orthonormalen) Eigenvektoren mit $\vec{v}_j$ und die assoziierten (reellen) Eigenwerte mit λ_j ($j = 1, 2, 3$), so macht es aus strukturellen Gründen Sinn folgenden Ansatz zu wählen:

$$\vec{x}_j(t) := f_j(t) \vec{v}_j,$$

wobei $f(t)$ eine beliebige (2mal stetig differenzierbare) skalare Funktion bezeichnet. Einsetzen dieses Ansatzes in (7) zeigt sofort die Kraft dieses Ansatzes. Tatsächlich gilt

$$\begin{aligned} 0 &= \ddot{\vec{x}}_j(t) + \omega^2 A \vec{x}_j(t) \\ &= \ddot{f}_j(t) \vec{v}_j + \omega^2 f_j(t) A \vec{v}_j \\ &= \ddot{f}_j(t) \vec{v}_j + \omega^2 f_j(t) \lambda_j \vec{v}_j \\ &= \left(\ddot{f}_j(t) + \omega^2 \lambda_j f_j(t) \right) \vec{v}_j. \end{aligned}$$

Aufgrund der letzten Zeile ist diese Gleichung äquivalent zu

$$\ddot{f}_j(t) + \omega_j^2 f_j(t) = 0,$$

wobei wir $\omega_j := \sqrt{\lambda_j} \omega$ definiert haben. Das ist nun aber tatsächlich die Standardgleichung eines harmonischen Oszillators und folglich gilt:

$$f_j(t) = a_j \cos(\omega_j t) + b_j \sin(\omega_j t),$$

sowie

$$\vec{x}_j(t) = f_j(t) \vec{v}_j = (a_j \cos(\omega_j t) + b_j \sin(\omega_j t)) \vec{v}_j \quad i = 1, 2, 3.$$

Nun wissen wir aber auch, dass die allgemeine Lösung einer beliebigen linearen Differentialgleichung einer Linearkombination aller möglichen Lösungen entspricht. In unserem Fall bedeutet das:

$$\vec{x}_{\text{allg}}(t) = \sum_{j=1}^3 (a_j \cos(\omega_j t) + b_j \sin(\omega_j t)) \vec{v}_j.$$

Doch damit noch nicht genug! Die Tatsache, dass die $\vec{v}_j$'s orthonormal zueinander gewählt werden können, vereinfacht das Anpassen dieser allgemeinen Lösung an Randbedingungen ungemein. Um das zu illustrieren, nehmen wir folgende Randbedingung an:

$$\begin{aligned} \vec{x}(0) &= \vec{c}, \\ \dot{\vec{x}}(0) &= 0. \end{aligned}$$

Daraus folgt sofort, dass $b_j = 0$ für alle $j = 1, 2, 3$ gelten muss und somit verlangt die erste Gleichung

$$\vec{x}(0) = \sum_{j=1}^3 a_j \cos(\omega_j 0) \vec{v}_j = \sum_{j=1}^3 a_j \vec{v}_j = \vec{c}.$$

Allerdings können wir nun benützen, dass $\{\vec{v}_j\}_{j=1}^3$ eine Orthonormalbasis von $\mathbb{R}^3$ ist und somit insbesondere gilt:

$$\vec{c} = \sum_{j=1}^3 \langle \vec{v}_j, \vec{c} \rangle \vec{v}_j.$$

Ein Koeffizientenvergleich zeigt somit $a_j = \langle \vec{v}_j, \vec{c} \rangle$ und folglich gilt

$$\vec{x}(t) = \sum_{j=1}^3 \langle \vec{v}_j, \vec{c} \rangle \cos(\omega_j t) \vec{v}_j.$$

Um das Ganze etwas greifbarer zu machen, erwähnen wir noch, dass die Matrix A folgende Eigenwerte und Eigenvektoren hat:

$$\begin{aligned} \lambda_1 &= 0, & \vec{v}_1 &= \frac{1}{\sqrt{3}} (1, 1, 1)^T, \\ \lambda_2 &= 1, & \vec{v}_2 &= \frac{1}{\sqrt{2}} (1, 0, -1)^T, \\ \lambda_3 &= 3, & \vec{v}_3 &= \frac{1}{\sqrt{4}} (1, -2, 1)^T. \end{aligned}$$

Somit gilt für die konkrete allgemeine Lösung:

$$\begin{aligned} \vec{x}_{\text{allg}}(t) &= (a_1 + b_1 t) \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + (a_2 \cos(\omega t) + b_2 \sin(\omega t)) \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \\ &+ (a_3 \cos(\sqrt{3}\omega t) + b_3 \sin(\sqrt{3}\omega t)) \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}. \end{aligned}$$

In physikalischem Sinne hat jeder dieser Eigenvektoren eine Interpretation als jeweilige “Eigenmode” – wir wollen hier aber nicht weiter auf diese Interpretation eingehen.

3.2 Beispiel (operatorwertige DGL): die Laplacegleichung

Dieses Beispiel bildet den Höhepunkt dieses Manuskripts. Wir wollen uns dabei endlich auf das Lösen der Laplacegleichung

$$\Delta_{x,y,z} f(x, y, z) = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) f(x, y, z) = 0$$

stürzen, welche in der Elektrostatik von entscheidender Bedeutung ist. Weit weg von einer beliebigen Ladungsverteilung muss diese Gleichung stets gelten. Zudem sieht jede Ladungsverteilung von großer Distanz wie eine Punktladung aus. Es macht daher Sinn zusätzlich eine Kugelsymmetrie anzunehmen und die sphärische Laplacegleichung

$$\Delta_{r,\theta,\phi} f(r, \theta, \phi) = 0 \quad \text{mit} \quad \Delta_{r,\theta,\phi} = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2} \Delta_{S^2}$$

zu betrachten. Hier bezeichnet Δ_{S^2} den sphärischen Laplaceoperator $\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} (\sin \theta \frac{\partial}{\partial \theta}) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2}$ – einen selbstadjungierten Operator auf dem Vektorraum aller Funktionen auf der Sphäre S^2 , den wir bereits aus unserem letzten Kapitel kennen. Der oben angedeuteten Aufteilung in Radialkoordinate r und Winkelkoordinaten θ, ϕ wollen wir durch einen Separationsansatz von der Form

$$f(r, \theta, \phi) = R(r) \Theta(\theta, \phi)$$

Folge leisten, wobei $R :]0, \infty[\rightarrow \mathbb{R}$ und $\Theta : S^2 \rightarrow \mathbb{R}$ beliebige 2mal differenzierbare Funktionen darstellen. Setzen wir diesen Ansatz in die Laplacegleichung ein, erhalten wir

$$\Delta_{r,\theta,\phi} R(r) \Theta(\theta, \phi) = \frac{\Theta(\theta, \phi)}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} R(r) \right) + \frac{R(r)}{r^2} \Delta_{S^2} \Theta(\theta, \phi) = 0.$$

Multiplizieren mit r^2 liefert dann

$$\left(\frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} R(r) \right) \right) \Theta(\theta, \phi) + R(r) \Delta_{S^2} \Theta(\theta, \phi) = 0. \quad (8)$$

Hier lohnt es sich kurz innezuhalten und sich der Ähnlichkeiten dieser Gleichung zur Matrixdifferentialgleichung im letzten Beispiel gewahr zu werden. Wir wissen, dass wir “moralisch” $\Theta(\theta, \phi)$ als unendlich-dimensionalen Vektor und Δ_{S^2} als “unendlichdimensionale” Matrix auffassen können. Der Funktion $R(r)$ kommt folglich die Rolle der skalaren Funktion $f(t)$ im vorigen Beispiel zu. Konsequenterweise haben wir es hier “quasi” mit einer gewöhnlichen linearen DGL in $R(r)$ und einer Matrixgleichung in $\Theta(\theta, \phi)$ zu tun. Aus dem

vorigen Beispiel wissen wir, dass sich letztere durch einen schlaun Ansatzvektor – nämlich einen Eigenvektor der betreffenden Matrix – gut in den Griff bekommen lässt. Wir machen von dieser Intuition Gebrauch und setzen

$$R(r) \Theta(\theta, \phi) := R_l(r) Y_l^m(\theta, \phi) \quad l \in \mathbb{N}, \quad m \in [-l, l],$$

da wir bereits aus dem vorigen Kapitel wissen, dass $Y_l^m(\theta, \phi)$ eine Eigenfunktion des Operators Δ_{S^2} ist. Konkret gilt

$$\Delta_{S^2} Y_l^m(\theta, \phi) = -l(l+1) Y_l^m(\theta, \phi).$$

Einsetzen dieser Tatsache in (8) liefert somit

$$\begin{aligned} 0 &= \left(\frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} R_l(r) \right) Y_l^m(\theta, \phi) + R_l(r) \Delta_{S^2} Y_l^m(\theta, \phi) \right) \\ &= \left(\frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} R_l(r) \right) - l(l+1) R_l(r) \right) Y_l^m(\theta, \phi). \end{aligned}$$

Genau wie im Matrixfall ist diese DGL somit äquivalent zu folgender homogener linearen und vor allem skalaren Differentialgleichung:

$$0 = \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} R_l(r) \right) - l(l+1) R_l(r).$$

Wie ihr bereits wisst, lässt sich eine derartige Differentialgleichung oft durch einen Ansatz der Form

$$R_l(r) = cr^k$$

mit $k \in \mathbb{N}$ lösen. Das ist auch hier der Fall. Tatsächlich folgt aus

$$\begin{aligned} 0 &= \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} cr^k \right) - l(l+1) cr^k = c \frac{\partial}{\partial r} (kr^{k+1}) - l(l+1) cr^k \\ &= c(k+1)kr^k - cl(l+1)r^k = c((k+1)k - l(l+1))r^k, \end{aligned}$$

dass die DGL von Interesse genau dann von cr^k gelöst wird, wenn

$$(k+1)k = l(l+1)$$

erfüllt ist. Das ist aber genau für $k = l$ und $k = -(l+1)$ der Fall. Somit erhalten wir mit diesem Ansatz

$$R_l(r) = A_l r^l + \frac{B_l}{r^{l+1}}$$

und folglich

$$R_l(r) Y_l^m(\theta, \phi) = \left(A_l r^l + \frac{B_l}{r^{l+1}} \right) Y_l^m(\theta, \phi).$$

Nun gilt es hervorzuheben, dass jede Wahl von $l \in \mathbb{N}$ und $m \in [-l, l]$ eine derartige valide Lösung liefert. Folglich entspricht die allgemeine Lösung der

Laplacegleichung einer Linearkombination a priori aller (das sind unendlich viele!) solcher Lösungen:

$$f(r, \theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l \left(A_l r^l + \frac{B_l}{r^{l+1}} \right) Y_l^m(\theta, \phi).$$

Abschließend wollen wir diese allgemeine Lösung noch etwas spezifizieren. Dazu nehmen wir an, dass ein Potential U auf der Einheitssphäre ($U : S^2 \rightarrow \mathbb{R}$) gegeben ist. Wir wollen dieses Potential auf den gesamten Außenraum der Kugel fortsetzen. Für große Distanzen muss das Potential verschwinden und wir können folglich A_l gleich 0 setzen. Die allgemeine Lösung $\Phi(r, \theta, \phi)$ der Potentialgleichung mit der Randbedingung

$$\Phi(1, \theta, \phi) \stackrel{!}{=} U(\theta, \phi)$$

vereinfacht sich für den Außenraum folglich zu

$$\Phi(r, \theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l \frac{B_{l,m}}{r^{l+1}} Y_l^m(\theta, \phi). \quad (9)$$

Darüber hinaus können wir wie im Matrixbeispiel oben die Tatsache, dass die Kugelflächenfunktionen eine Orthonormalbasis für den Funktionenraum $L^2(S^2)$ bildet, nützen. Klarerweise gilt nämlich $U \in L^2(S^2)$ und folglich lässt sich das Anfangspotential schreiben als

$$U(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l \langle Y_l^m(\theta, \phi), U(\theta, \phi) \rangle Y_l^m(\theta, \phi).$$

Wie im Matrixfall liefert ein einfacher Koeffizientenvergleich nun

$$B_{l,m} = \langle Y_l^m(\theta, \phi), U(\theta, \phi) \rangle$$

und Einsetzen dieser Tatsache in (9) liefert

$$\begin{aligned} \Phi(r, \theta, \phi) &= \sum_{l=0}^{\infty} \sum_{m=-l}^l \frac{1}{r^{l+1}} \langle Y_l^m(\theta, \phi), U(\theta, \phi) \rangle Y_l^m(\theta, \phi) \\ &= \sum_{l=0}^{\infty} \sum_{m=-l}^l \frac{1}{r^{l+1}} \left(\int_0^{\pi} \int_0^{2\pi} \bar{Y}_l^m(\theta, \phi) U(\theta, \phi) \sin \theta d\theta d\phi \right) Y_l^m(\theta, \phi), \end{aligned} \quad (10)$$

wobei wir in der zweiten Zeile lediglich die Definition des (komplexen) Standardskalarproduktes $\langle \cdot, \cdot \rangle_{S^2}$ eingesetzt haben. Dies ist nun die Formel, welche ihr aus der Vorlesung bereits kennt.

Abschließend wollen wir noch ein einfaches, konkretes Beispiel behandeln, nämlich eine homogen geladene Kugeloberfläche mit Gesamtladung Q . In Tabellen

lässt sich das konkrete Aussehen der Kugelflächenfunktionen $Y_l^m(\theta, \phi)$ nachschlagen und es gilt insbesondere

$$Y_0^0(\theta, \phi) = \frac{1}{\sqrt{4\pi}} = \text{const.}$$

Somit können wir insbesondere

$$U(\theta, \phi) = \frac{Q}{4\pi} = \frac{Q}{\sqrt{4\pi}} Y_0^0(\theta, \phi)$$

schreiben. Einsetzen dieser Ladungsverteilung in Formel (10) und Ausnützen der Orthonormalität von Kugelflächenfunktionen ($\langle Y_l^m, Y_{l'}^{m'} \rangle = \delta_{l,l'} \delta_{m,m'}$) liefert

$$\begin{aligned} \Phi(r, \theta, \phi) &= \sum_{l=0}^{\infty} \sum_{m=-l}^l \frac{4\pi}{r^{l+1}} \langle Y_l^m(\theta, \phi), \frac{Q}{\sqrt{4\pi}} Y_0^0(\theta, \phi) \rangle Y_l^m(\theta, \phi) \\ &= \frac{Q}{\sqrt{4\pi}} \sum_{l=0}^{\infty} \sum_{m=-l}^l \frac{4\pi}{r^{l+1}} \langle Y_l^m(\theta, \phi), Y_0^0(\theta, \phi) \rangle Y_l^m(\theta, \phi) \\ &= \frac{Q}{\sqrt{4\pi}} \sum_{l=0}^{\infty} \sum_{m=-l}^l \frac{4\pi}{r^{l+1}} \delta_{l,0} \delta_{m,0} Y_l^m(\theta, \phi) \\ &= \frac{4\pi Q}{\sqrt{4\pi}} \frac{1}{r} Y_0^0(\theta, \phi) = \frac{Q}{r}, \end{aligned}$$

wie erwartet. Denn schließlich ist das genau das bekannte Potential einer homogen geladenen Kugel, oder eben einer Punktladung im Ursprung.