2 Mathematische Grundlagen. Banach- und Hilberträume

Häufig beschränkt man sich bei der mathematischen Beschreibung von physikalischen Problemen auf Hilberträume. Das hat Vor- und Nachteile. Der Vorteil, der Fülle von Bearbeitungsmethoden im Hilbertraum kommt erst zur Geltung, wenn man für sein spezielles Problem den richtigen mathematischen Rahmen gefunden hat. Dazu ist es sinnvoll, physikalische Größen verschiedener Typen auch unterschiedlich zu beschreiben. Dazu stellt es sich als besonders zweckmäßig heraus, intensive und extensive physikalische Größen in Bachachräume und ihre dualen zu platzieren.
Hat man das getan, kann man überlegen, welcher der vielen möglichen Hilberträume der richtige für das betrachtete Problem ist.

2.1 Banachräume

2.1.1 Banachräume und Operatoren

Wir bezeichnen Banachräume mit $\mathcal{X}$ und $\mathcal{Y}$, ihre dualen (die Banachräume der Funktionale) mit $\mathcal{X}^*$ und $\mathcal{Y}^*$. Elemente daraus bezeichnen wir mit x, y bzw. x^*, y^*.
Den zu $\mathcal{X}$ bidualen Raum bezeichnen wir mit $\mathcal{X}^{**}$. Ist $\mathcal{X} = \mathcal{X}^{**}$, d.h. wenn sich die Elemente von $\mathcal{X}$ und $\mathcal{X}^{**}$ geeignet identifizieren lassen, heißt $\mathcal{X}$ reflexiv. Im allgemeinen ist $\mathcal{X} \subset \mathcal{X}^{**}$ dank der Dualitätsabbildung stetig eingebettet.
In der Mathematik werden verschiedene Banachräume benutzt. Der Ausgangspunkt einer Überlegung in der klassischen Physik sollte aber stets $\mathcal{C}(\mathcal{Z})$, der Raum der stetigen reellwertigen Funktionen auf einem kompakten topologischen Raum $\mathcal{Z}$ sein. Die Elemente aus $\mathcal{C}(\mathcal{Z})$ sind – im Gegensatz zu Elementen aus L_2-Räumen – tatsächlich Funktionen, also eindeutige Abbildungen $\mathcal{Z} \to \mathbb{R}$. $\mathcal{C}(\mathcal{Z})$ ist mit der sup-Norm ein Banachraum (das muß man beweisen, da $\mathcal{C}(\mathcal{Z})$ – im Gegensatz zu den Lebesgueräumen – nicht durch Vervollständigung entsteht), mit der punktweisen Ordnung ein Banachverband und mit der punktweisen Multiplikation eine Banachalgebra.
Der zu $\mathcal{C}(\mathcal{Z})$ duale Raum $\mathcal{C}^*(\mathcal{Z})$ ist nach dem fundamentalen Darstellungssatz von Riesz-Markow der Raum der Radonmaße.
Der Fall, daß $\mathcal{X}$ endlich dimensional ist, entsteht bei der Betrachtung einer endlichen Menge $\mathcal{Z} = \{z_1, ..., z_n\} =: \mathcal{Z}_n$. $\mathcal{C}(\mathcal{Z}_)$ ist dann der $\mathbb{R}_n$ mit der Maximum-Norm. Das ist nicht der euklidische $\mathbb{R}^n$! $\mathcal{C}^*(\mathcal{Z}_n)$ ist dann ebenfalls der $\mathbb{R}_n$ aber mit der l_1-Norm. Wir schreiben für ihn $\mathbb{R}_n^*$. Wir betrachten in ihm stets die kanonische Basis e_i bzw. e_i^*.
Es sei $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ die Menge der linearen beschränkten Abbildungen zwischen zwei Banachräumen $\mathcal{X}$ und $\mathcal{Y}$. Sie bildet einen linearen Raum und mit der üblichen Operatornorm einen Banachraum.
Betrachtet man einen Banachraum $\mathcal{X}$ und seinen dualen $\mathcal{X}^*$, so gibt es prinzipiell zwei Typen von Operatoren, solche aus $\mathcal{L}(\mathcal{X}, \mathcal{X})$ und solche $\mathcal{L}(\mathcal{X}, \mathcal{X}^*)$.
1) Es sei $\mathbf{A} : \mathcal{X} \to \mathcal{X}$, dann bildet sein adjungierter $\mathbf{A}^* : \mathcal{X}^* \to \mathcal{X}^*$ den dualen Raum auf sich selbst ab. Es ist klar, daß das Gleichsetzen von $\mathbf{A}$ und $\mathbf{A}^*$ Unsinn ist, weil die Operatoren in völlig verschiedenen Räumen wirken.
Für Operatoren aus $\mathcal{L}(\mathcal{X}, \mathcal{X})$ kann es einen inversen Operator, einen Fixpunkt oder Potenzen geben.
Im $\mathcal{Z}_n$-Fall entspricht (in der kanonischen Basis) einem Operator $\mathbf{A}$ eine $n \times n$-Matrix und $\mathbf{A}^*$ ihre transponierte Matrix. Formal kann man symmetrische Matrizen betrachten, was aber nicht

bedeutet, daß $\mathbf{A}$ symmetrisch ist, was wie erläutert keinen Sinn hat. Der identischen Abbildung entspricht die Einheitsmatrix $\mathbf{I}$.

2) Es sei $\mathbf{A} : \mathcal{X} \to \mathcal{X}^*$, dann wirkt sein adjungierter $\mathbf{A}^* : \mathcal{X}^{**} \to \mathcal{X}^*$ im bidualen Raum. Da $\mathcal{X}$ in $\mathcal{X}^{**}$ stetig eingebettet ist, kann man die Einschränkung von $\mathbf{A}^*$ auf $\mathcal{X}$ betrachten (wir nennen sie ebenfalls $\mathbf{A}^*$, was nicht zu Verwechslungen führt): $\mathbf{A}^* : \mathcal{X} \longrightarrow \mathcal{X}^*$. Für so einen Operator kann man die Frage stelle, ob er symmetrisch ist. Wir nennen einen Operator $\mathbf{A}$ symmetrisch, falls er mit der Einschränkung seines adjungierten übereinstimmt. Äquivalent ist $\mathbf{A} : \mathcal{X} \to \mathcal{X}^*$ symmetrisch, wenn

$$\langle x, \mathbf{A}y \rangle = \langle y, \mathbf{A}x \rangle, \ \ x, y \in \mathcal{X}$$

Offensichtlich gibt es in $\mathcal{L}(\mathcal{X}, \mathcal{X}^*)$ keine Identität.

Für Operatoren aus $\mathcal{L}(\mathcal{X}, \mathcal{X}^*)$ sind Begriffe wie “Fixpunkt” oder “Potenzen von Operatoren” sinnlos.

Es kann zu einem Operator $\mathbf{A} : \mathcal{X} \longrightarrow \mathcal{X}^*$ einen inversen geben. Davon kann man sprechen, wenn er bijektiv $\mathcal{X}$ auf $R(\mathcal{X}) \subset \mathcal{X}^*$ abbildet. Im allgemeinen (insbesondere in ∞-dimensionalen Raum, wenn $\mathcal{X}$ nicht reflexiv ist) gibt es aber keine bijektiven beschränkten Operatoren, die $\mathcal{X}$ auf ganz $\mathcal{X}^*$ abbilden. Es gilt der Satz: $\mathcal{X}$ ist reflexiv genau dann, wenn es einen stetigen und stetig invertierbaren Operator in $\mathbf{A} : \mathcal{X} \to \mathcal{X}^*$ gibt.

Im $\mathcal{Z}_n$-Fall entspricht (in der kanonischen Basis) einem Operator $\mathbf{A}$ eine $n \times n$-Matrix und $\mathbf{A}^*$ ihre transponierte Matrix. Ein Operator ist symmetrisch, wenn seine Matrix symmetrisch ist.

Formal kann man eine Einheitsmatrix $\mathbf{I}$ betrachten, die aber nicht die Matrix der Identität ist, welche es in $\mathcal{L}(\mathcal{X}, \mathcal{X}^*)$ nicht gibt.

2.1.2 Symmetrische und formenpositive Operatoren im Banachraum

Es sei $\mathcal{X}$ ein reeller Banachraum, $\mathcal{X}^*$ sein dualer, $\mathcal{X}^{**}$ sein bidualer und $\langle x, x^* \rangle$ die duale Paarung zwischen $\mathcal{X}$ und $\mathcal{X}^*$. Es sei $\mathbf{A} : \mathcal{X} \to \mathcal{X}^*$ ein linearer beschränkter Operator und $\mathbf{A}^* : \mathcal{X}^{**} \to \mathcal{X}^*$ sein adjungierter.

Ein linearer Operator $\mathbf{Q} : \mathcal{X} \longrightarrow \mathcal{X}^*$ definiert auf $\mathcal{X}$ durch $\langle x, \mathbf{Q}y \rangle$ eine Bilinearform (auch quadratisches Funktional genannt).

Ein Operator $\mathbf{Q}$ und die entsprechende Bilinearform heißt **symmetrisch**, wenn $\langle x, \mathbf{Q}y \rangle = \langle y, \mathbf{Q}x \rangle$ für alle $x, y \in \mathcal{X}$ gilt. Diese Eigenschaft bedeutet im reflexiven Raum $\mathbf{Q} = \mathbf{Q}^*$ und im allgemeinen Fall $\mathbf{Q} = \mathbf{Q}^*\big|_{\mathcal{X}}$.

Ein symmetrischer Operator $\mathbf{Q}$ heißt **positiv** (im Formensinne)[3], wenn $\langle x, \mathbf{Q}x \rangle \geq 0$ für alle $x \in \mathcal{X}$. Er heißt **strikt positiv**, wenn $\langle x, \mathbf{Q}x \rangle > 0$ für alle $\mathcal{X} \ni x \neq 0$. Er heißt **positiv definit**, wenn es ein $k > 0$ gibt mit $\langle x, \mathbf{Q}x \rangle \geq k\|x\|^2$ für alle $\mathcal{X} \ni x \neq 0$.

[3] In der Mathematik werden zwei Arten von positiven Operatoren betrachtet. Im Banachverband spielen Operatoren, die die Positivtät erhalten eine wichtige Rolle. Sie werden **positive Operatoren im Kegelsinn** genannt. Operatoren, die eine auf der Diagonale positive Bilinearform definieren, heißen **positive Operatoren im Formensinn**. Wir betrachten in diesem Kurs vorläufig nur positive Operatoren im Formensinn und nennen sie einfach “positiv”.

2.1.3 Die Hauptidentität für invertierbare positive Operatoren

Es sei $\mathbf{A} > 0$ und invertierbar. Dann gilt

$$\begin{aligned}\frac{1}{2}\langle x, \mathbf{A}x\rangle - \langle x, x^*\rangle + \frac{1}{2}\langle \mathbf{A}^{-1}x^*, x^*\rangle &= \frac{1}{2}\Big\langle x - \mathbf{A}^{-1}x^*, \mathbf{A}x - x^*)\Big\rangle = \\ &= \frac{1}{2}\Big\langle (x - \mathbf{A}^{-1}x^*), \mathbf{A}(x - \mathbf{A}^{-1}x^*)\Big\rangle = \\ &= \frac{1}{2}\Big\langle \mathbf{A}^{-1}(\mathbf{A}x - x^*), (\mathbf{A}x - x^*)\Big\rangle \geq 0\end{aligned}$$

Diese Identität wird auch Formel der quadratische Ergänzung für symmetrische Operatoren genannt. Aus ihr folgt die Ungleichung

$$\frac{1}{2}\langle x, \mathbf{A}x\rangle - \langle x, x^*\rangle \geq -\frac{1}{2}\langle \mathbf{A}^{-1}x^*, x^*\rangle$$

oder

$$\frac{1}{2}\langle \mathbf{A}^{-1}x^*, x^*\rangle \geq \langle x, x^*\rangle - \frac{1}{2}\langle x, \mathbf{A}x\rangle$$

die man als Variante der Youngschen Ungleichung bezeichnen kann.
Diese Ungleichungen sind scharf, es gilt Gleichheit für $\mathbf{A}x = x^*$ oder $x = \mathbf{A}^{-1}x^*$.

2.1.4 Beschränkte Operatoren in vier Banachräumen

Typischerweise enstehen physikalische Probleme in zwei Banachräumen und ihren dualen.
Es sei $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ die Menge der linearen beschränkten Abbildungen zwischen $\mathcal{X}$ und $\mathcal{Y}$. Sie bildet einen linearen Raum und mit der üblichen Operatornorm einen Banachraum.
Es sei $\mathbf{A} : \mathcal{X} \to \mathcal{Y}$ ein linearer beschränkter Operator und $\mathcal{X}^*$ und $\mathcal{Y}^*$ die zu $\mathcal{X}$ bzw. $\mathcal{Y}$ dualen Räume. $\langle \mathbf{A}x, y^*\rangle$ ist für festes x eine lineare beschränkte Abbildung nach $\mathbb{R}$. D.h. ein Funktional auf $\mathcal{X}$. Es existiert also ein x^* sodaß $\langle \mathbf{A}x, y^*\rangle = \langle x, x^*\rangle$. Wir nennen die Abbildung $y^* \longrightarrow x^*$ adjungierten Operator und schreiben $x^* = \mathbf{A}^* y^*$. Die Operatoren $\mathbf{Q}$ und $\mathbf{R}$ werden später eine Rolle spielen. Sie verbinden die 4 Räume zu einem kommutativen Diagramm.

$$\begin{tikzcd} \mathcal{X} \arrow[r, "\mathbf{A}"'] \arrow[d, "\mathbf{Q}"] & \mathcal{Y} \arrow[d, "\mathbf{R}"'] \\ \mathcal{X}^* & \mathcal{Y}^* \arrow[l, "\mathbf{A}^*"'] \end{tikzcd}$$

Der adjungierte Operator hat folgende Eigenschaften:

- Eindeutig definiert, linear, beschränkt
- $\mathbf{A}^* \in \mathcal{L}(\mathcal{Y}^*, \mathcal{X}^*)$, d.h., $\mathbf{A}^* : \mathcal{Y}^* \to \mathcal{X}^*$.

  Sehr wichtig: Der adjungierte Operator bildet die dualen Räume in umgekehrter Richtung ab.
- Wir führen hier die möglichen Opertoren zwischen den Räumen und die die adjungierten Operatoren (wir schränken sie ein, falls sie auf dem bidualen Raum definiert sind) definierenden dualen Paarungen an:

$$\begin{array}{llll} \mathbf{A}: \ \mathcal{X} \to \mathcal{Y}, & \mathbf{A}^*: \ \mathcal{Y}^* \to \mathcal{X}^*, & \langle y^*, \mathbf{A}x\rangle_{\mathcal{Y}} = & \langle \mathbf{A}^* y^*, x\rangle_{\mathcal{X}} \\ \mathbf{A}: \ \mathcal{X} \to \mathcal{Y}^*, & \mathbf{A}^*: \ \mathcal{Y} \to \mathcal{X}^*, & \langle y, \mathbf{A}x\rangle_{\mathcal{Y}} = & \langle \mathbf{A}^* y, x\rangle_{\mathcal{X}} \end{array}$$

- In einem reflexiven B-Raum hat jeder Operator einen prädualen, d.h. zu einem Operator $\mathbf{B} \in \mathcal{L}(\mathcal{Y}^*, \mathcal{X}^*)$ existiert ein Operator $\mathbf{A} : \mathcal{X} \to \mathcal{Y}$ mit $\mathbf{A}^* = \mathbf{B}$.

  Im nichtreflexiven B-Raum muß das nicht der Fall sein. Wir betrachten im weiteren stets Operatoren in $\mathcal{L}(\mathcal{Y}^*, \mathcal{X}^*)$, die einen prädualen besitzen.

  ÜA Konstruiere einen Operator $\mathbf{B} : \mathcal{Y}^* \to \mathcal{X}^*$, der keinen prädualen hat, d.h. für den es keinen beschränkten Operator $\mathbf{A} : \mathcal{X} \to \mathcal{Y}$ mit $\mathbf{B} = \mathbf{A}^*$ gibt.

- Im endlich dimensionalen Raum kann man einen Operator durch eine Matrix darstellen. Der Übergang zum adjungierten Operator entspricht dem Transponieren der Matrix. Man sieht, daß die transponierte Matrix in umgekehrter Richtung abbildet als die Matrix selbst, wenn man rechteckige Matrizen betrachtet.

Für Operatoren, die zwischen verschiedenen Räumen wirken sind Begriffe wie "inverser Operator", "Fixpunkt" oder "Potenzen von Operatoren" sinnlos.
Im allgemeinen endlichdimensionalen Fall haben die Räume $\mathcal{X}$ und $\mathcal{Y}$ verschiedene Dimensionen. Die Matrizen von Operatoren aus $\mathcal{L}(\mathcal{X}, \mathcal{Y})$, $\mathcal{L}(\mathcal{X}, \mathcal{Y}^*)$ oder $\mathcal{L}(\mathcal{X}^*, \mathcal{Y}^*)$ sind dann rechteckig.

2.2 Hilberträume

Ein Hilbertraum $\mathcal{H}$ ist ein linearer Raum, indem ein Skalarprodukt $(\cdot, \cdot)$ definiert ist, das die Norm als $\|x\|^2 = (x, x)$ bestimmt, in der der Raum vollständig ist.
In einem reellen Hilbertraum $\mathcal{H}$ wird häufig von symmetrischen oder selbstadjungierten Operatoren $\mathbf{A} \in \mathcal{L}(\mathcal{H}, \mathcal{H})$ gesprochen, wenn $\mathbf{A} = \mathbf{A}^*$ oder äquivalent $(\mathbf{A}x, y) = (x, \mathbf{A}y)$, $x, y \in \mathcal{H}$ gilt.
Ein symmetrischer Operator $\mathbf{A}$ heißt **positiv**, wenn $(x, \mathbf{A}x) \geq 0$ für alle $x \in \mathcal{H}$. Er heißt **strikt positiv**, wenn $(x, \mathbf{A}x) > 0$ für alle $\mathcal{H} \ni x \neq 0$. Er heißt **positiv definit**, wenn es ein $k > 0$ gibt mit $(x, \mathbf{A}x) \geq k\|x\|^2$ für alle $x \in \mathcal{H}$.
Die kanonische Form, einen Hilbertraum zu definieren ist die Definition über einen strikt positiven Operator in einem Banachraum, z.B. in $\mathcal{C}$ (siehe weiter unten).

2.2.1 Die kanonische Definition von Hilberträumen

Ein strikt positiver Operator $\mathbf{Q} : \mathcal{X} \to \mathcal{X}^*$ definiert einen reellen Hilbertraum $L_2(\mathbf{Q})$ durch das Skalarprodukt $(x, y)_\mathbf{Q} = \langle x, \mathbf{Q}y \rangle$ mit der Norm $\|x\|^2_{L_2(\mathbf{Q})} = \langle x, \mathbf{Q}x \rangle$. $L_2(\mathbf{Q})$ ist die Vervollständigung von $\mathcal{X}$ in dieser Norm.
Ein typisches Beispiel hierfür ist der $L_2(\mu)$: Es sei $\mathcal{Z}$ ein gegebener kompakter toplogischer Raum und $\mathcal{X} = \mathcal{C}(\mathcal{Z})$. Es sei $\mu \in \mathcal{C}^*$ ein positives Radonmaß. Da $\mathcal{X}$ mit der punktweisen Multiplikation eine Banachalgebra ist, ist

$$(g, f)_\mu := \langle g \cdot f, \mu \rangle = \int_\mathcal{Z} g(z) f(z) \mu(dz)$$

ein auf $\mathcal{X}$ definiertes Skalarprodukt. $L_2(\mu)$ ist dann der durch Vervollständigung von $\mathcal{X}$ in der Norm $\|g\|_\mu = \sqrt{(g, g)_\mu}$ definierte Banachraum. Zu beachten ist, daß $L_2(\mu)$ ein Raum von Grenzwerten ist, dessen Elemente im allgemeinen nicht als Funktionen $\mathcal{Z} \longrightarrow \mathbb{R}$ betrachtet werden können (siehe auch die Bemerkungen in Abschnitt 2.3 auf S.35).
Einen gegebenen beschränkten Operator $\mathbf{A} : \mathcal{X} \to \mathcal{X}$ kann man unter bestimmten Bedingungen nach $L_2(\mathbf{Q})$ stetig fortsetzen (wir nennen ihn ebenfalls $\mathbf{A}$). Ist das der Fall, dann kann man

fragen, was der zu $\mathbf{A}$ in $L_2(\mathbf{Q})$ adjungierte Operator ist. Wir nennen ihn $\mathbf{A}^\star$ (im Gegensatz zu $\mathbf{A}^*$). $\mathbf{A}^\star$ ist der Operator, für den die Gleichung

$$(x, \mathbf{A}y)_{\mathcal{H}} = (y, \mathbf{A}^\star x)_{\mathcal{H}}$$

gilt. In $\mathcal{X}$ (da $\mathcal{X}$ in $\mathcal{H}$ dicht ist, reicht das zur Betrachtung aus) bedeutet das

$$\langle x, \mathbf{QA}y\rangle = \langle y, \mathbf{QA}^\star x\rangle$$

oder

$$\langle y, \mathbf{A}^*\mathbf{Q}y\rangle = \langle y, \mathbf{QA}^\star x\rangle$$

$\mathbf{A}^\star$ erfüllt also die Gleichung

$$\mathbf{QA}^\star = \mathbf{A}^*\mathbf{Q}$$

Sollte $\mathbf{Q}^{-1}$ existieren, dann gilt

$$\mathbf{A}^\star = \mathbf{Q}^{-1}\mathbf{A}^*\mathbf{Q}$$

Eine genauere Beschreibung der Zusammenhänge dieser Operatoren siehe in Punkt 2.2.4 auf S.34.
Der selbstadjungierte Fall $\mathbf{A}^\star = \mathbf{A}$ tritt für Operatoren $\mathbf{A}$ ein, die die Gleichung $\mathbf{QA} = \mathbf{A}^*\mathbf{Q}$ erfüllen.
Im endlich dimensionalen Raum ist $\mathbf{A}^*$ stets die zu $\mathbf{A}$ transponierte Matrix, wogegen $\mathbf{A}^\star$ genau dann die zu $\mathbf{A}$ transponierte Matrix ist, wenn $\mathbf{Q} = \mathbf{I}$. Dann ist offensichtlich $\mathbf{A}^\star = \mathbf{A}^*$.

2.2.2 Der Abstand eines Vektors zu einem Unterraum

Eine häufige Aufgabe in einem Hilbertraum $\mathcal{H}$ mit Skalarprodukt $(\cdot,\cdot)$ und Norm $\|\cdot\|$ ist es, den Abstand eines gegebenen Vektors $x \in \mathcal{H}$ zu einem Unterraum $U \subset \mathcal{H}$ zu bestimmen. Das ist

$$\text{dist}\,(x, U) = \min_{y\in U} \|x - y\|$$

oder

$$\text{dist}^2(x, U) = \min_{y\in U} \|x - y\|^2 = \frac{1}{2}\min_{y\in U}(x - y, x - y)$$

Angenommen, U ist der Wertebereich eines gegebenen Operators $\mathbf{G}$, $U = R(\mathbf{G})$, dann läßt sich dist (x, U) auch als

$$\begin{aligned}\text{dist}^2(x, R(\mathbf{G})) &= \frac{1}{2}\min_{y\in\mathcal{H}}(x - \mathbf{G}y, x - \mathbf{G}y) = \\ &= \min_{y\in\mathcal{H}}\left(\frac{1}{2}(\mathbf{G}^\star\mathbf{G}y, y) - (\mathbf{G}^\star x, y)\right) + \|x\|^2\end{aligned}$$

Ableiten bezüglich y führt auf die Gleichung

$$\mathbf{G}^\star\mathbf{G}y_0 = \mathbf{G}^\star x$$

zur Bestimmung von

$$y_0 = (\mathbf{G}^\star\mathbf{G})^{-1}\mathbf{G}^\star x$$

Der Vektor $\mathbf{G}y_0$ ist die orthogonale Projektion des Vektors x auf $R(\mathbf{G})$ und wird vermittelt durch den Projektor $\mathbf{P} : \mathcal{H} \to \mathcal{H}$

$$\mathbf{P} = \mathbf{G}(\mathbf{G}^\star\mathbf{G})^{-1}\mathbf{G}^\star$$

Die nötigen Eigenschaften $\mathbf{P}^2 = \mathbf{P}$ und $\mathbf{P}^\star = \mathbf{P}$ sind leicht zu überprüfen.
Ist $\mathcal{H} = L_2(\mathbf{Q})$, dann läßt sich $\mathbf{P} : \mathcal{X} \to \mathcal{X}$ wegen $\mathbf{G}^\star = \mathbf{Q}^{-1}\mathbf{G}^*\mathbf{Q}$ bestimmen als

$$\mathbf{P} = \mathbf{G}(\mathbf{Q}^{-1}\mathbf{G}^*\mathbf{Q}\mathbf{G})^{-1}\mathbf{Q}^{-1}\mathbf{G}^*\mathbf{Q} = \mathbf{G}(\mathbf{G}^*\mathbf{Q}\mathbf{G})^{-1}\mathbf{G}^*\mathbf{Q}$$

Berechnung der Größe des Minimums
Eine besonders wichtige Anwendung für diese Projektion ist die Methode der kleinsten Quadrate (siehe 1.3.2 auf Seite 21).

2.2.3 Orthogonale Zerlegung im Hilbertraum

ES sei $\mathcal{H}$ ein gegebener Hilbertraum und $\mathbf{P} : \mathcal{H} \to \mathcal{H}$ ein orthogonaler Projektor (ein Operator mit den Eigenschaften $\mathbf{P}^2 = \mathbf{P}$ und $\mathbf{P}^\star = \mathbf{P}$) auf den Unterraum $\mathcal{H}_1 = \mathbf{P}\mathcal{H}$. Dann läßt sich $\mathcal{H}$ in die orthogonale Summe $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$ zerlegen, wobei $\mathcal{H}_2$ das Bild des Projektors $\mathbf{I} - \mathbf{P}$ ist. Es gilt also für alle $h \in \mathcal{H}$

$$h = \mathbf{P}h + (\mathbf{I} - \mathbf{P})h$$

Die Orthogonalität von $\mathbf{P}h$ und $(\mathbf{I} - \mathbf{P})h$ folgt aus

$$(\mathbf{P}h, (\mathbf{I} - \mathbf{P})h) = (h, \mathbf{P}(\mathbf{I} - \mathbf{P})h) = (h, (\mathbf{P} - \mathbf{P}^2)h) = 0$$

Bemerkung: Die Orthogonalität folgt nicht allein aus der Projektoreigenschaft $\mathbf{P}^2 = \mathbf{P}$. Auch $\mathbf{P}^\star = \mathbf{P}$ wird benötigt.
Ein gegebener Operator $\mathbf{G} : \mathcal{H} \to \mathcal{H}$ zerlegt $\mathcal{H}$ in die orthogonale Summe $\mathcal{H} = R(\mathbf{G}) \oplus K(\mathbf{G}^\star)$.
Der Projektor auf $R(\mathbf{G})$ ist (siehe 2.2.2 auf Seite 33)

$$\mathbf{P} = \mathbf{G}(\mathbf{G}^\star\mathbf{G})^{-1}\mathbf{G}^\star$$

Es gilt der Satz des Pythagoras

$$\|h\|^2 = \mathrm{dist}^2(h, R(\mathbf{G})) + \mathrm{dist}^2(h, K(\mathbf{G}^\star))$$

2.2.4 Operatoren zwischen verschiedenen Hilberträumen

Da wir es typischerweise mit 4 Banachräumen zu tun haben, entstehen kanonisch auch zwei Hilberträume $\mathcal{H}_\mathcal{X}$ und $\mathcal{H}_\mathcal{Y}$ durch Operatoren $\mathbf{Q}_\mathcal{X}$ bzw. $\mathbf{Q}_\mathcal{Y}$. Wir wollen den adjungierten Operator $\mathbf{G}^\star$ zu einem gegebenen Operator $\mathbf{G} : \mathcal{H}_\mathcal{X} \to \mathcal{H}_\mathcal{Y}$ ermitteln. $\mathbf{G}^\star$ erfüllt die Gleichung

$$(y, \mathbf{G}x)_{\mathbf{Q}_\mathcal{Y}} = (x, \mathbf{G}y)_{\mathbf{Q}_\mathcal{X}}$$

Übergang zu den Banachräumen ergibt

$$\begin{aligned}(y, \mathbf{G}x)_{\mathbf{Q}_\mathcal{Y}} &= \langle \mathbf{Q}_\mathcal{Y} y, \mathbf{G}x\rangle_\mathcal{Y} = \langle \mathbf{G}^*\mathbf{Q}_\mathcal{Y} y, x\rangle_\mathcal{Y} \\ (x, \mathbf{G}y)_{\mathbf{Q}_\mathcal{X}} &= \langle \mathbf{Q}_\mathcal{X} x, \mathbf{G}^\star y\rangle_\mathcal{X} = \langle x, \mathbf{Q}_\mathcal{X}\mathbf{G}^\star y\rangle_\mathcal{X}\end{aligned}$$

Jetzt können wir die Operatoren gleichsetzen und erhalten

$$\mathbf{Q}_\mathcal{X}\mathbf{G}^\star = \mathbf{G}^*\mathbf{Q}_\mathcal{Y} \tag{7}$$

oder, falls $\mathbf{Q}_\mathcal{X}^{-1}$ existiert

$$\mathbf{G}^\star = \mathbf{Q}_\mathcal{X}^{-1}\mathbf{G}^*\mathbf{Q}_\mathcal{Y} \tag{8}$$

$$\begin{tikzcd}
\mathcal{H}_\mathcal{X} \arrow[d, shift left, "\mathbf{G}"] & \mathcal{X} \arrow[l, hook'] \arrow[r, "\mathbf{Q}_\mathcal{X}"] \arrow[d, shift left, "\mathbf{A}"] & \mathcal{X}^* \\
\mathcal{H}_\mathcal{Y} \arrow[u, shift left, "\mathbf{G}^\star"] & \mathcal{Y} \arrow[l, hook'] \arrow[r, "\mathbf{Q}_\mathcal{Y}"] \arrow[u, shift left, "\mathbf{A}^+"] & \mathcal{Y}^* \arrow[u, "\mathbf{A}^*"']
\end{tikzcd}$$

Genauer ist diese Konstruktion folgendermaßen zu verstehen: Wir unterscheiden hier vorübergehend den Operator $\mathbf{A} : \mathcal{X} \to \mathcal{Y}$ und seine Erweiterung $\mathbf{G} : \mathcal{H}_\mathcal{X} \to \mathcal{H}_\mathcal{Y}$. Es gelte also $\mathbf{A} = \mathbf{G}$ auf $\mathcal{X}$.
Es sei $\mathbf{A}^+ : \mathcal{Y} \to \mathcal{X}$ ein Operator, der die Gleichung

$$\mathbf{Q}_\mathcal{X}\mathbf{A}^+ = \mathbf{A}^*\mathbf{Q}_\mathcal{Y} \qquad : \mathcal{Y} \to \mathcal{X}^* \tag{9}$$

erfüllt. Unter generischen Bedingungen existiert stets wenigstens ein solcher Operator. (Im endlich dimensionalen Fall ist die Existenz gesichert, im allgemeinen noch nicht vollständig untersucht.) Es kann aber sein, daß er nicht einzig ist.
Es sei $\mathbf{G}$ die Erweiterung von $\mathbf{A}$. Wir nehmen an, daß $\mathbf{G}$ ebenfalls stetig ist. $\mathbf{G}^\star$ sei sein adjungierter im Hilbertraum-Sinne. $\mathbf{G}^\star$ existiert und ist einzig (wenn $\mathbf{G}$ dicht definiert ist, was der Fall ist, wenn $\mathbf{G}$ stetig ist). Dann ist $\mathbf{G}^\star$ die Erweiterung von allen $\mathbf{A}^+$, die (9) erfüllen.
Falls $\mathbf{Q}_\mathcal{X}^{-1}$ existiert, ist Gleichung (8) eine Möglichkeit $\mathbf{G}^\star$ explizit zu berechnen. Schon der endlich dimensionalen Fall, wenn $\mathbf{Q}_\mathcal{X}$ keine Diagonalmatrix ist, zeigt, das die Berechnung von $\mathbf{G}^\star$ mühevoll sein kann.
Zum Zusammenhang zwischen den Gleichungen (7) und (8): Es ist klar, daß der Übergang von Gleichung (7) zu (8) nur dann möglich ist, wenn $\mathbf{Q}_\mathcal{X}^{-1}$ in irgendeinem Sinn existiert. Das ist z.B. der Fall, wenn $\mathbf{Q}_\mathcal{X}$ bijektiv ist. Dann läßt sich $\mathbf{Q}_\mathcal{X}^{-1}$ auf $R(\mathbf{Q}_\mathcal{X})$ eindeutig definieren. Um Gleichung (8) dann verwenden zu können, muß gezeit werden, daß $\mathbf{G}^*\mathbf{Q}_\mathcal{Y}$ in den Wertebereich von $\mathbf{Q}_\mathcal{X}$ abbildet.
Ein typischer Fall, daß $\mathbf{Q}_\mathcal{X}^{-1}$ nicht existiert und damit Gleichung (8) sinnlos ist, ist der, $K(\mathbf{Q}_\mathcal{X}) \neq \{0\}$. In diesem Fall besitzt Gleichung (7) *mehrere* Lösungen. Aber jede Lösung besitzt dieselbe Erweiterung $\mathbf{G}^\star$.

In unserem speziellen Fall ist

$$\begin{aligned}
\mathbf{Q}_\mathcal{X} &= \mathbf{A} \\
\mathbf{Q}_\mathcal{Y} &= \mathbf{C} \\
\mathbf{G} &= \mathbf{B} = \mathbf{C}^{-1}\mathbf{D}
\end{aligned}$$

und wir erhalten wegen $\mathbf{B}^* = \mathbf{D}^*\mathbf{C}^{-1}$

$$\mathbf{B}^\star = (\mathbf{C}^{-1}\mathbf{D})^\star = \mathbf{A}^{-1}(\mathbf{C}^{-1}\mathbf{D})^*\mathbf{C} = \mathbf{A}^{-1}\mathbf{D}^*\mathbf{C}^{-1}\mathbf{C} = \mathbf{A}^{-1}\mathbf{D}^*$$

2.3 Zum Unterschied zwischen Banach- und Hilberträumen

Ein Banachraum ist das ursprüngliche, der Startpunkt für die Untersuchung eines Problems. Zwischen Banach- und Hilberträumen gibt es prinzipielle Unterschiede, die bereits bei ihren Definitionen klar werden.
Die Theorie der **Banachräume** läuft so ab:

- Wir starten mit einem linearen Raum $\mathcal{X}$ und definieren eine Norm.
- Wir beweisen, daß $\mathcal{X}$ vollständig ist.
- Wir versuchen den dualen Raum $\mathcal{X}^*$ und die duale Paarung zwischen beiden Räumen zu beschreiben. Das ist so zu verstehen: $\mathcal{X}^*$ ist abstrakt als Raum der stetigen Funktionale definiert. Um mit ihm arbeiten zu können, ist es sinnvoll, ihn genauer zu kennen. Manchmal stellt sich heraus, daß er isomorph zu einem bereits gut bekannten Raum ist. So stellt sich z.B.heraus, daß der abstrakte Raum L_p^* isomorph zum gut bekannten Raum L_q^* mit $q = p/(p-1)$ ist. Solche Beweise sind nicht trivial. Für viele interessante Räume hat man noch keine befriedigende Beschreibung des dualen gefunden.

Wir betrachten diese Schritte am Beispiel des klassischen Banachraumes der stetigen Funktionen.

- Es sei $\mathcal{Z}$ ein kompakter, Hausdorffscher, erstabzählbarer topologischer Raum, wir betrachten die Menge der stetigen reellwertigen Funktionen g darauf. Es stellt sich heraus, daß sie einen linearer Raum bilden. Wir normieren ihn mit der Maximumnorm. Wir nennen den Raum $\mathcal{X} = \mathcal{C}(\mathcal{Z})$
- Wir beweisen, daß $\mathcal{X}$ vollständig ist. Das ist relativ einfach zu beweisen, es liegt aber nicht in der Definition von $\mathcal{X}$.
- Wir beweisen (Darstellungssatz von Riesz-Markow), daß der duale Raum $\mathcal{X}^*$ der Raum der Radonmaße auf der σ-Algebra der Borelmengen von $\mathcal{Z}$ (induziert von den abgeschlossenen Mengen aus $\mathcal{Z}$) ist und daß sich die duale Paarung zwischen beiden Räumen als Lebesgueintergral

  $$\langle g, p \rangle = \int_{\mathcal{Z}} g(z)\mu(dz)$$

  schreiben läßt.

Der endlichdimensionale Raum entsteht, wenn $\mathcal{Z} = \{z_1, ..., z_n\}$ (endliche Menge) mit der diskreten Toplogie ist. Die enstehenden Räume $\mathcal{X}$ und $\mathcal{X}^*$ sind nicht euklidische Räume, da sie andere Normen haben.

Ein **Hilbertraum** wird folgendermaßen definiert:

- Wir starten mit einem Banachraum $\mathcal{X}$ und definieren ein Skalarprodukt und mit seiner Hilfe eine Norm.
- Wir vervollständigen $\mathcal{X}$ in dieser Norm und nennen den entstehenden Hilbertraum $\mathcal{H} = \overline{\mathcal{X}}$. Per definition ist dieser Raum vollständig. Da ist nichts zu beweisen. Offensichtlich ist auch $\mathcal{X} \subset \mathcal{H}$. (Stichworte: Gelfand-Tripel, energetische Räume)
- Der duale eines Hilbertraumes ist isomorph zu sich selbst (Satz von Fréchet-Riesz), sodaß $\mathcal{H}^*$ mit $\mathcal{H}$ identifiziert werden kamm. Diese Identifizierung wird mit der Dualitätsabbildung erreicht. Es muß also – im Gegensatz zum Banachraum – nicht bestimmt werden, ob der abstrakte duale Raum vielleicht zu einem bekannten Raum isomorph ist. Mit dem Satz von Fréchet-Riesz ist das ein für allemal erledigt.

Die Frage ist jetzt, was gibt es für Gründe, ein spezielles Skalarprodukt auszuwählen. Die kanonische Methode ist hier, einen symmetrischen, positiven Operator $\mathbf{Q} : \mathcal{X} \to \mathcal{X}^*$ festzulegen und nach der beschriebenen Methode das Skalarprodukt zu definieren. Der Operator $\mathbf{Q}$ muß sich aus der Aufgabe ergeben. Jede spezielle Aufgabe erfordert ein spezielles $\mathbf{Q}$.
Der endlichdimensionale euklidische Hilbertraum entsteht, wenn die duale Paarung mit dem homogenen Maß μ betracht wird. Der erzeugende Operator $\mathbf{Q}$ ist hier die "Einheitsmatrix" $\mathbf{I}$. Sie sieht zwar aus wie die Einheitsmatrix im endlichdimensionalen euklidischen (Hilbert)Raum, sie ist aber keine Identität, denn sie bildet $\mathbf{I} : \mathcal{X} \to \mathcal{X}^*$ ab.
Grob gesagt ist der Banachraum das richtige Objekt, wenn es darum geht, das Problem formal zu verstehen. Die Betrachtung des Problems in vier verschiedenen Banachräumen sichert, daß man jede physikalische Größe ihrem Typ (extensiv oder intensiv) entsprechend verwendet. Man verwechselt nicht die Operatoren. Vor allem im endlich dimensionalen Raum sieht man einem Vektor nicht unbedingt an, ob er als Element eines Raumes oder seines dualen auftritt.
Der Hilbertraum ist dagegen das Mittel der Wahl, wenn es darum geht, ein konkretes Problem explizit zu lösen. Hat man aus der Fülle der möglichen Hilberträume den richtigen ausgewählt, steht das ganze Spektrum an Methoden zur Verfügung, die es nur im Hilbertraum gibt, z.B.

- Selbstadjungierte und normale Operatoren
- Eigenwerte, Eigenvektoren, Diagonalisierung
- Orthogonalität, Projektoren

Allerdings kann die Benutzung dieser Methoden im allgemeinen schwerer (und vor allem ungewöhnlicher) sein, als man es vom euklidischen Raum gewöhnt ist.
Wie häufig ist die historische Entwicklung den umgekehrten Weg der kanonischen Entwicklung gegangen (die kanonische Entwicklung wird erst deutlich, nachdem der historische Weg beschritten wurde). So steht z.B. (April 2023) in der *wikipedia* die historische Sicht: Die duale Paarung ist in der Mathematik eine Abbildung, die einem Vektor und einem linearen Funktional eine Zahl zuweist. Sie stellt eine Verallgemeinerung des Skalarproduktes dar.
Die kanonische Sicht ist dagegen: Das Skalarprodukt ist eine spezielle Bilinearform, die mit Hilfe der dualen Paarung definiert wird.
Wir schließen mit ein paar **Merksätzen:**

- Gewichte (Operatoren $\mathbf{Q}$ in Diagonalform) sind im Hilbertraum allgegenwärtig. Im Banachraum gibt es prinzipiell keine Gewichte.
- Im endlichdimensionalen ist der adjungierte Operator im Banachraum stets die transponierte Matrix, im Hilbertraum im allgemeinen nicht.
- Im Banachraum gibt es (meistens) eine kanonische Basis. Im Hilbertraum wäre eine kanonische Basis eine orthonormale Basis. Die muß man – falls man sie braucht – erst berechnen. Sie hängt natürlich von $\mathbf{Q}$ ab.
- Der Banachraum ist bei der Modellierung der ursprünglich Raum. Unter bestimmten Bedingungen, wenn z.B. Materialgesetze linear sind, ist es sinnvoll zu einem Hilbertraum überzugehen. D.h., es ist durchaus sinnvoll, auch nichtlineare Operatoren im (linearen) Banachraum zu betrachten. Im Hilbertraum eher nicht.

2.4 Differentialrechnung

Es gibt verschiedene Begriffe im Zusammenhang mit Ableitungen, die vom betrachteten Raum abhängen.

2.4.1 Die Richtungsableitung

Es sei $\mathcal{X}$ eine Teilmenge eines linearen Raumes. und $F : \mathcal{X} \to \mathbb{R}$ eine genug reguläre reellwertige Abbildung und $x \in \mathcal{X}$ ein fixierter Punkt. Desweiteren sei $y \in \mathcal{X}$ ebenfalls ein gegebebener Punkt, den wir als Vektor betrachten (siehe Identifizierung von Punkten und Vektoren).
Die **Richtungsableitung** von F am Punkt x entlang y ist definiert durch den Limes

$$D_y F(x) = \lim_{t \to +0} \frac{F(x+ty) - F(x)}{t}$$

falls dieser existiert. $D_y F(x)$ ist eine reelle Zahl.

2.4.2 Die Gateauxableitung

Im Banachraum kann man untersuchen, wie sich die Richtungsableitung $D_y F(x)$ in Abhängigkeit von y verhält. Es kann z.B. sein, daß $D_y F(x)$ (für festes x) als Funktion von y linear und beschränkt ist. In diesem Fall existiert ein Funktional $\xi^* \in \mathcal{X}^*$ (das von x abhängt), sodaß

$$D_y F(x) = \langle \xi^*, y \rangle, \ y \in \mathcal{X}$$

dann heißt $\xi^*(x) = F'(x) \in \mathcal{X}^*$ **Gateauxableitung** von F in x.

2.4.3 Die Frechetableitung

Definition: Eine Abbildung $G : \mathcal{Y} \longrightarrow \mathcal{X}$ heißt im Punkt y **Frechet**-differenzierbar, wenn ein linearer Operator $G'(y) : \mathcal{Y} \to \mathcal{X}$ existiert und für alle $h \in \mathcal{Y}$ gilt

$$G(y+h) = G(y) + G'(y)h + \omega(x,h)$$

mit $\|\omega\|_{\mathcal{X}} \to 0$ für $\|h\|_{\mathcal{Y}} \to 0$.
Ist $G(y) = \mathbf{G}y$ eine lineare Abbildung, so gilt offensichtlich

$$G(y+h) = \mathbf{G}(y+h) = \mathbf{G}y + \mathbf{G}h + 0$$

Folglich ist $\mathbf{G}$ die Frechetableitung von $G(y) = \mathbf{G}y$.

2.4.4 Beispiele

Beispiel 1: Es sei $x^* \in \mathcal{X}^*$ gegeben. Wir betrachten das lineare Funktional $F(x) = \langle x, x^* \rangle$. Es sei $y \in \mathcal{X}$ gegeben. Dann ist offenbar

$$D_y F(x) = \lim_{t \to +0} \frac{F(x+ty) - F(x)}{t} = \lim_{t \to +0} \frac{\langle x+ty, x^* \rangle - \langle x, x^* \rangle}{t} = \lim_{t \to +0} \frac{1}{t} \langle ty, x^* \rangle = \langle y, x^* \rangle$$

Offenbar existiert für $D_y F(x) = \langle y, x^* \rangle$ ein derartiges ξ^*, nämlich $\xi^* = x^*$ (hängt nicht von x ab, wie für lineare Funktionen üblich). Es ist also $F'(x) = x^*$.
Beispiel 2: Es sei $\mathbf{A} : \mathcal{X} \to \mathcal{X}^*$ mit $\mathbf{A} > 0$ gegeben. Wir betrachten das quadratische Funktional $F(x) = \frac{1}{2}\langle x, \mathbf{A}x \rangle$. Es sei $y \in \mathcal{X}$ gegeben. Dann ist offenbar wegen der Symmetrie von $\mathbf{A}$

$$D_y F(x) = \lim_{t \to +0} \frac{\langle x+ty, \mathbf{A}x+ty \rangle - \langle x, \mathbf{A}x \rangle}{2t} = \lim_{t \to +0} \frac{1}{t} \left(t\langle y, \mathbf{A}x \rangle + \frac{1}{2} t^2 \langle y, \mathbf{A}y \rangle \right) = \langle y, \mathbf{A}x \rangle$$

Wegen $D_y F(x) = \langle y, \mathbf{A}x \rangle$ existiert offenbar ein geeignetes $\xi^* \in \mathcal{X}^*$, nämlich $\xi^* = \mathbf{A}x$. Es ist also $F'(x) = \mathbf{A}x$. Die Gateauxableitung eines quadratischen Funktionals hängt linear von x ab.

$G(x) = \mathbf{A}x$ ist eine lineare Abbildung $\mathcal{X} \to \mathcal{X}^*$. Ihre Frechetableitung ist $\mathbf{G}$. Man kann sie als zweite Gateauxableitung bezeichnen, was in folgendem Beispiel deutlich wird:
Es gilt die Identität

$$\frac{1}{2}\langle x', \mathbf{A}x'\rangle = \frac{1}{2}\langle x, \mathbf{A}x\rangle + \langle x' - x, \mathbf{A}x\rangle + \frac{1}{2}\big\langle (x' - x), \mathbf{A}(x' - x)\big\rangle$$

Diese Identität entspricht der Taylorreihe

$$F(x') = F(x) + \langle x' - x, F'(x)\rangle + \frac{1}{2}\big\langle (x' - x), F''(x)(x' - x)\big\rangle + ...$$

Hier ist $F''(x)$ die zweite Gateaux-Ableitung an der Stelle x, ein linearer Operator $F''(x) : \mathcal{X} \to \mathcal{X}^*$.
Beispiel 3: Es sei $F(x) = \frac{1}{2}\langle x, \mathbf{A}x\rangle - \langle x, x^*\rangle$. Dann gilt $F'(x) = \mathbf{A}x - x^*$. Die Gateauxableitung diese Funktionales verschwindet im Punkt x_0, wenn die Gleichung $\mathbf{A}x_0 = x^*$ erfüllt ist.

2.4.5 Der Gradient

Gateauxableitungen werden typischerweise in Banachräumen betrachtet. In Hilberträumen gibt es den Begriff des **Gradient**en.
Definiert man mit Hilfe eines linearen, beschränkten symmetrischen positiven Operators $\mathbf{Q} : \mathcal{X} \to \mathcal{X}^*$ einen Hilbertraum $L_2(\mathbf{Q})$ als Erweiterung von $\mathcal{X}$ durch das Skalarprodukt

$$(x, y)_{\mathbf{Q}} = \langle x, \mathbf{Q}y\rangle$$

dann läßt sich der **Gradient** $\nabla F(x)$ einer glatten Funktion $F : \mathcal{X} \to \mathbb{R}$ an der Stelle x definieren. Er ist dasjenge Element aus $L_2(\mathbf{Q})$, dessen Skalarprodukt mit einem beliebigen $y \in \mathcal{X}$ gerade die Richtungsableitung von F an der Stelle x in Richtung y ist, also (falls $\nabla F(x) \in \mathcal{X}$)

$$(\nabla F(x), y)_{\mathbf{Q}} = D_y F(x)$$

Als Zusammenhang mit der Gateauxableitung ergibt sich aus den Definitionen im Hilbert- und im Banachraum

$$\begin{aligned} D_y F(x) &= (\nabla F(x), y)_{\mathbf{Q}} = \langle y, \mathbf{Q}\nabla F(x)\rangle = \langle \nabla F(x), \mathbf{Q}y\rangle \\ D_y F(x) &= \langle y, F'(x)\rangle \end{aligned}$$

die Beziehung

$$F'(x) = \mathbf{Q}\nabla F(x)$$

Beispiel: Wie bekannt gilt für den Gradienten der Norm in jedem Hilbertraum $\nabla \frac{1}{2}\|x\|_{\mathcal{H}}^2 = x$. Gilt $\mathcal{H} = L_2(\mathbf{Q})$, mit einem positiven Operator $\mathbf{Q} : \mathcal{X} \to \mathcal{X}^*$, dann kann andererseits die Norm als Funktional auf $\mathcal{X}$ betrachtet werden: $F(x) = \frac{1}{2}\|x\|_{\mathcal{H}}^2 = \frac{1}{2}\langle x, \mathbf{Q}x\rangle$. Seine Gateauxableitung ist $F'(x) = \mathbf{Q}x$. Es gilt also $F'(x) = \mathbf{Q}\nabla F(x)$.

2.5 Das unbedingte Minimumproblem im Banachraum

Wir betrachten jetzt den Zusammenhang zwischen linearen Gleichungen und Minimumsproblemen quadratischer Funktionale.
Es sei $x^* \in \mathcal{X}^*$ und $\mathbf{A} > 0$ ein strikt positiver Operator. Wir definieren das quadratische Funktional $P : \mathcal{X} \to \mathbb{R}$ durch

$$P(x) = \frac{1}{2}\langle x, \mathbf{A}x\rangle - \langle x, x^*\rangle$$

Analog zu quadratischen Funktionen mit positivem Koeffizienten bei x^2, hat so ein Funktional ein Minimum. Es gilt folgender
Satz: $P(x)$ hat ein Minimum in x_0 genau dann, wenn x_0 Lösung der Gleichung $\mathbf{A}x_0 = x^*$ ist.
Beweis: $\Longrightarrow$ Es sei $P(x_0) = \min_{x\in\mathcal{X}} P(x)$. Da $P(x)$ differenzierbar ist, gilt $\nabla P(x)\big|_{x=x_0} = 0$. Das bedeutet $\mathbf{A}x_0 - x^* = 0$.
$\Longleftarrow$ Es sei x_0 Lösung der Gleichung $\mathbf{A}x_0 = x^*$. Analog zur Hauptidentität folgt

$$\begin{aligned}
P(x) - P(x_0) &= \frac{1}{2}\langle x, \mathbf{A}x\rangle - \langle x, x^*\rangle - \frac{1}{2}\langle x_0, \mathbf{A}x_0\rangle + \langle x_0, x^*\rangle = \\
&= \frac{1}{2}\langle x, \mathbf{A}x\rangle - \langle x, \mathbf{A}x_0\rangle - \frac{1}{2}\langle x_0, \mathbf{A}x_0\rangle + \langle x_0, \mathbf{A}x_0\rangle = \\
&= \frac{1}{2}\langle x, \mathbf{A}x\rangle - \langle x, \mathbf{A}x_0\rangle + \frac{1}{2}\langle x_0, \mathbf{A}x_0\rangle = \\
&= \frac{1}{2}\langle (x - x_0), \mathbf{A}(x - x_0)\rangle \geq 0
\end{aligned}$$

Wobei Gleichheit für $x = x_0$ gilt. □
Wegen $\mathbf{A} > 0$ ist x_0 einzig und man kann für die Lösung $x_0 = \mathbf{A}^{-1}x^*$ schreiben. Für das Minimum

$$P(x_0) = P(\mathbf{A}^{-1}x^*) = -\frac{1}{2}\langle \mathbf{A}^{-1}x^*, x^*\rangle$$

erhält man. Es gilt die Ungleichung

$$P(x) = \frac{1}{2}\langle x, \mathbf{A}x\rangle - \langle x, x^*\rangle \geq -\frac{1}{2}\langle \mathbf{A}^{-1}x^*, x^*\rangle = P(x_0)$$

2.6 Das Minimumproblem im allgemeinen Hilbertraum für unbeschränkte Operatoren

$\mathcal{H}$ sei Hilbertraum, $\mathbf{A}$ sei ein linearer Operator, der auf $\mathcal{D}(\mathbf{A}) \subset \mathcal{H}$ dicht definiert ist.

$$\overline{\mathcal{D}(\mathbf{A})} = \mathcal{H} \tag{10}$$

$$(\mathbf{A}g, g) \geq c^2\|g\|^2 \quad , \quad \forall\, g \in \mathcal{D}(\mathbf{A}) \tag{11}$$

Gefragt ist nach der Lösbarkeit der Gleichung

$$\mathbf{A}g = f \tag{12}$$

für beliebige $f \in \mathcal{H}$. Wenn (12) eine Lösung hat, so folgt aus (11) die Einzigkeit.
Daneben wird das Funktional

$$F(g) = (\mathbf{A}g, g) - (g, f) - (f, g) \tag{13}$$

betrachtet und nach seinem Minimum gefragt. Die Frage hat Sinn, da $F(g)$ wegen

$$F(g) = (\mathbf{A}g, g) - 2\Re\mathrm{e}\,(g, f)$$

und (11) nur reelle Werte annimmt.

2.6.1 Äquivalenz von Gleichung und Minimumproblem

Theorem 1

$$\mathbf{A}g_0 = f \quad \Longleftrightarrow \quad F(g_0) = \min_{g \in \mathcal{D}(\mathbf{A})} F(g)$$

Proof: Es sei $\mathbf{A}g_0 = f$. Dann gilt wegen $(\mathbf{A}g_0, g) = (g_0, \mathbf{A}g)$

$$\begin{aligned} F(g) &= (\mathbf{A}g, g) - (g, \mathbf{A}g_0) - (\mathbf{A}g_0, g) = \\ &= (\mathbf{A}g, g) - (\mathbf{A}g, g_0) - (\mathbf{A}g_0, g) = \\ &= \Big(\mathbf{A}(g - g_0), (g - g_0)\Big) - (\mathbf{A}g_0, g_0) \end{aligned}$$

Hieraus folgt, daß $F(g)$ sein Minimum bei $g = g_0$ annimmt und es gilt

$$\min_{g \in \mathcal{D}(\mathbf{A})} F(g) = -(\mathbf{A}g_0, g_0) = -(f, g_0) = -(g_0, f)$$

Umgekehrt sei

$$F(g_0) = \min_{g \in \mathcal{D}(\mathbf{A})} F(g)$$

Es sei $h \in \mathcal{D}(\mathbf{A})$ und $t \in \mathbb{R}$, dann nimmt die Funktion $f(t) = F(g_0 + th)$ für $t = 0$ ihr Minimum an. Das heißt

$$0 \;=\; \frac{d}{dt} F(g_0 + th) = \Re\mathrm{e}\, (\mathbf{A}g_0 - f, h)$$

Für $h := ih$ folgt $\Im\mathrm{m}\, (\mathbf{A}g_0 - f, h) = 0$ und damit $(\mathbf{A}g_0 - f, h) = 0$. Wegen (10) folgt $\mathbf{A}g_0 = f$. ■

2.6.2 Lösung des Minimumproblems im Hilbertraum (Existenzbeweis)

Neues Skalarprodukt auf $\mathcal{D}(\mathbf{A})$

$$[g, h] = (\mathbf{A}g, h)\ , \quad g, h \in \mathcal{D}(\mathbf{A}) \tag{14}$$

und Norm

$$|g|^2 = [g, g] \tag{15}$$

Es sei $\mathcal{H}_A$ der Abschluß von $\mathcal{D}(\mathbf{A})$ in der Norm $|\cdot|$. Es gilt wegen (11)

$$\|g\| \le \frac{1}{c}|g|$$

und — weil deshalb eine Fundamentalfolge in $\mathcal{H}_A$ auch fundamental in $\mathcal{H}$ ist und $\mathcal{H}$ ist vollständig:

$$\mathcal{H}_A \subset \mathcal{H}$$

Wegen

$$|(g, f)| \le \|g\| \|f\| \le \frac{\|f\|}{c} \cdot |g|$$

ist (g, f) ein beschränktes, lineares Funktional auf $\mathcal{H}_A$. Aus dem Satz von Riesz folgt die Existenz eines $g_0 \in \mathcal{H}_A$ mit

$$(g, f) = [g, g_0]$$

und damit

$$\begin{aligned} F(g) &= [g, g] - [g, g_0] - [g_0, g] = \\ &= [g - g_0, g - g_0] - [g_0, g_0] = |g - g_0|^2 - |g_0|^2 \end{aligned}$$

also gilt

$$\min_{g \in \mathcal{H}_A} F(g) = F(g_0) = -\|g_0\|^2$$

Lösungsmethode: Householdertransformation. Das ist die Transformation $x \to \mathbf{H}x$, wobei $\mathbf{H} = \mathbf{I} - n^* \otimes n$ eine um $\mathbf{I}$ verschobene Rank-1 Matrix ist. $\mathbf{H}x$ liefert den Vektor, der sich aus x bei Spieglung an der Ebene mit Normalenvektor n ergibt. Wird verwendet zur $\mathbf{A} = \mathbf{QR}$-Zerlegung mit $\mathbf{Q} = (\mathbf{H}_n \cdots \mathbf{H}_1)^*$.