MATHEMATISCHE BILDVERARBEITUNG

Vorlesungsskript, Wintersemester 2020/21

Christian Clason

Stand vom 9. Februar 2021

Fakultät für Mathematik
Universität Duisburg-Essen

INHALTSVERZEICHNIS

ÜBERBLICK

Inhalt dieser Vorlesung sind moderne mathematische Methoden für Aufgaben in der Bildverarbeitung und Bildgebung; darunter fallen Entrauschen, Schärfen oder Interpolation von Bildern sowie die Rekonstruktion von Bildern aus (unvollständigen) Messungen wie in der Computer-Tomographie oder Magnetresonanztomographie. Bilder werden dabei aufgefasst als Funktionen $u : \Omega \to M$ für ein Gebiet $\Omega \subset \mathbb{R}^n$ und eine Wertemenge $M \subset \mathbb{R}^d$. Im einfachsten Fall eines Schwarz-Weiss-Bildes (auf den wir uns hier konzentrieren wollen) bildet u einen Punkt $x \in [0,1]^2$ auf die Helligkeitswert $u(x) \in [0,1]$ ab (wobei 0 schwarz und 1 weiss bedeutet); dieses Modell umfasst aber ebenso Farbbilder ($u(x) = (r, g, b) \in \mathbb{R}^3$ entspricht dann dem Farbwert) bis hin zu dreidimensionalen und zeitabhängigen Datensätzen in der medizinischen Bildgebung. Durch diesen abstrakten Zugang werden Begriffe und Methoden aus der Funktionalanalysis und der nichtlinearen Optimierung anwendbar; man spricht von *Variationsmethoden*. Ein Kernaspekt dabei ist es, Funktionenräume zu finden, die die Struktur von Bildern möglichst gut berücksichtigen.

Ein Beispiel soll dies verdeutlichen. Nimmt man ein Photo bei schlechten Lichtverhältnissen auf, muss das eingefangene Licht stark verstärkt werden um ein sichtbares Bild zu erhalten. Dabei wird neben dem gewünschten Bildinhalt auch thermisches *Rauschen* verstärkt; dessen nachträgliches Entfernen wird als *Entrauschen* bezeichnet. Wir nehmen also an, das aufgenommene Bild $f : [0,1]^2 \to \mathbb{R}$ setzt sich zusammen aus dem gewünschten, rauschfreien, Bildinhalt $u^0 : [0,1]^2 \to \mathbb{R}$ und einer unerwünschten Störung $\eta : [0,1]^2 \to \mathbb{R}$, d. h.

$$f = u^0 + \eta.$$

Die Aufgabe ist also, bei gegebenem f das unbekannte u^0 zu finden. Natürlich gibt es unendlich viele Möglichkeiten, eine Funktion als Summe zweier Funktionen zu schreiben. Wir müssen also diese Möglichkeiten einschränken, indem wir Annahmen an die Struktur von u^0 und f machen:

- Thermisches Rauschen entsteht durch Zufallsprozesse, die unabhängig in jedem Punkt $x \in [0,1]^2$ wirken, deren Stärke aber (zum Beispiel) normalverteilt ist. Solche Funktionen sind in dem Raum der quadrat-integrierbaren Funktionen enthalten, d. h. $\eta \in L^2(\Omega)$.

- Bilder haben dagegen eine räumliche Struktur, benachbarte Bildpunkte sind also (in der Regel) ähnlich. Eine erste Möglichkeit, dies zu beschreiben, ist durch Ableitungen – je ähnlicher benachbarte Helligkeitswerte, desto kleiner die Ableitung. Wir setzen also u^0 als (schwach) differenzierbar an: $u^0 \in H^1(\Omega)$, d. h. der (schwache) Gradient $\nabla u^0 \in L^2(\Omega)^2$. (Eine mathematisch rigorose Definition folgt in einem späteren Kapitel.)

Wollen wir jeweils aus dieser Funktionenklasse das beste Element finden, so führt dies auf das Minimierungsproblem

$$\min_{\substack{u^0 \in H^1(\Omega), \eta \in L^2(\Omega) \\ u^0 + \eta = f}} \int_\Omega |\eta(x)|^2 \, dx + \int_\Omega |\nabla u^0(x)|^2 \, dx.$$

Eliminieren wir die letzte Bedingung durch Setzen von $\eta = f - u^0$, so erhalten wir

$$\min_{u \in H^1(\Omega)} \int_\Omega |u(x) - f(x)|^2 \, dx + \int_\Omega |\nabla u(x)|^2 \, dx.$$

Wir werden sehen, dass viele Probleme in der Bildverarbeitung in dieser Form geschrieben werden können:

$$\min_{u \in U} F(u) + \alpha R(u),$$

wobei der *Diskrepanzerm* F die Struktur der Störung und der *Regularisierungsterm* R die Struktur des gesuchten Bildes beschreibt; blicherweise verwendet man hier (Halb-)Normen in einem geeigneten Banachraum $X \supset U$. Der *Regularisierungsparameter* α gewichtet dabei, was uns wichtiger ist: Die Struktur des Bildes oder die Nähe zu den gegebenen Daten.

Die Fragen, die wir uns nun zu stellen haben, sind

(i) Hat dieses Problem eine Lösung, d. h. existiert ein $\bar{u} \in U$ mit

$$F(\bar{u}) + \alpha R(\bar{u}) \leq F(u) + \alpha R(u) \qquad \text{für alle } u \in U?$$

(ii) Gibt es eine intrinsische Charakterisierung von $\bar{u}$, d. h. ohne Vergleich mit allen anderen $u \in U$?

(iii) Wie kann dieses $\bar{u}$ (effizient) berechnet werden?

Das Vorgehen lässt sich anhand der einfachsten Situation, nämlich $U \subset \mathbb{R}$, skizzieren:

(i) Ist U kompakt und sind F und G stetig, so nimmt die Funktion $J := F + G$ nach dem Satz von Weierstraß ihr Minimum in $\bar{u} \in U$ an.

(ii) Sind F und G differenzierbar, so gilt das *Fermat-Prinzip*

$$0 = J'(\bar{u}) = F'(\bar{u}) + \alpha R'(\bar{u}).$$

(iii) Sind F und G stetig differenzierbar, so kann man das Verfahren des steilsten Abstiegs verwenden: Wähle u^0 und setze für $k = 1, \ldots$

$$u^{k+1} = u^k - t_k J'(u^k)$$

mit geeigneter Schrittweite t_k, dann gilt $u^k \to \bar{u}$ für $k \to \infty$.

Ziel der Vorlesung ist es, diese Ansätze so zu verallgemeinern, dass sie auf Bildverarbeitungsprobleme angewendet werden können. Wir werden sehen, dass wir dabei schnell gezwungen sind, den Standardrahmen zu verlassen, denn Bilder haben eine besondere Struktur, die zu berücksichtigen ist: Sie bestehen typischerweise aus glatten Regionen, die durch *Kanten* getrennt sind (und können dazu *Texturen*, d. h. fein aufgelöste, regelmässige, Details enthalten). Allerdings sind (auch schwach) differenzierbare Funktionen stetig, und können daher keine Sprünge enthalten; Sobolevräume wie $H^1(\Omega)$ sind also ungeeignet, Bilder zu beschreiben. Andererseits enthalten Lebesgueräume wie $L^2(\Omega)$ auch Rauschen, erlauben also keine Trennung von Bild und Rauschen. Wir brauchen also Räume „dazwischen"; insbesondere der Raum der Funktionen mit *beschränkter Variation* hat sich als nützlich herausgestellt. Diese Räume sind aber in der Regel keine Hilberträume und führen zu nichtdifferenzierbaren Diskrepanz- bzw. Regularisierungstermen, so dass die oben genannten Schritte schwierig werden. Dies macht aber auch den Reiz der Bildverbeitung aus: tiefe Methoden der Funktionalanalysis, Maßtheorie und konvexen Analysis werden notwendig für eine praktische Anwendung!

In dieser Vorlesung konzentrieren wir uns jedoch auf algorithmische Aspekte und verzichten dafür auf die maßtheoretischen Details bei der Konstruktion der auftretenden Funktionenräume.

Dieses Skriptum basiert vor allem auf den folgenden Werken:

[1] K. Bredies & D. A. Lorenz (2011), *Mathematische Bildverarbeitung, Einführung in Grundlagen und moderne Theorie*, Vieweg+Teubner, doi: 10.1007/978-3-8348-9814-2

[2] W. Schirotzek (2007), *Nonsmooth Analysis*, Universitext, Springer, Berlin, doi: 10.1007/978-3-540-71333-3

[3] H. H. Bauschke & P. L. Combettes (2017), *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, 2. Aufl., Springer, New York, doi: 10.1007/978-3-319-48311-5

[4] M. Brokate (2014), Konvexe Analysis und Evolutionsprobleme, Vorlesungsskript, Zentrum Mathematik, TU München, url: http://www-m6.ma.tum.de/~brokate/cev_ss14.pdf

[5] H. Attouch, G. Buttazzo & G. Michaille (2014), *Variational Analysis in Sobolev and BV Spaces*, 2. Aufl., Society for Industrial & Applied Mathematics (SIAM), Philadelphia, PA, doi: 10.1137/1.9781611973488

[6] O. Scherzer u. a. (2009), *Variational Methods in Imaging*, Springer, New York, doi: 10.1007/978-0-387-69277-7

Teil I

GRUNDLAGEN

1 GRUNDLAGEN DER FUNKTIONALANALYSIS

In diesem Kapitel stellen wir die für diese Vorlesung wesentlichen Begriffe, Notationen und Resultate zusammen. Für Beweise wird auf die Standardliteratur verwiesen, z. B. auf [Werner 2011], sowie auf [Clason 2015].

1.1 NORMIERTE RÄUME

Im Folgenden bezeichne X einen Vektorraum über dem Körper $\mathbb{K}$, wobei wir uns hier stets auf den Fall $\mathbb{K} = \mathbb{R}$ beschränken. Eine Abbildung $\|\cdot\| : X \to \mathbb{R}^+ := [0, \infty)$ heißt *Norm* (auf X), falls für alle $x \in X$ gilt

(i) $\|\lambda x\| = |\lambda|\|x\|$ für alle $\lambda \in \mathbb{K}$,

(ii) $\|x + y\| \leq \|x\| + \|y\|$ für alle $y \in X$,

(iii) $\|x\| = 0$ genau dann, wenn $x = 0 \in X$.

Beispiel 1.1. (i) Auf $X = \mathbb{R}^N$ werden Normen definiert durch

$$\|x\|_p = \left(\sum_{i=1}^{N} |x_i|^p\right)^{1/p} \qquad 1 \leq p < \infty,$$
$$\|x\|_\infty = \max_{i=1,\dots,N} |x_i|.$$

(ii) Auf $X = \ell^p$ (dem Raum der reellen Folgen, auf dem folgende Ausdrücke endlich sind) sind Normen definiert durch

$$\|x\|_p = \left(\sum_{i=1}^{\infty} |x_i|^p\right)^{1/p} \qquad 1 \leq p < \infty,$$
$$\|x\|_\infty = \sup_{i=1,\dots,\infty} |x_i|.$$

(iii) Auf $X = L^p(\Omega)$ (dem Raum der messbaren reellen Funktionen auf dem Gebiet $\Omega \subset \mathbb{R}^n$, auf dem folgende Ausdrücke endlich sind) sind Normen definiert durch

$$\|u\|_p = \left(\int_\Omega |u(x)|^p\right)^{1/p} \qquad 1 \leq p < \infty,$$
$$\|u\|_\infty = \operatorname*{ess\ sup}_{x\in\Omega} |u(x)|.$$

(iv) Auf $X = C(\overline{\Omega})$ (dem Raum der stetigen Funktionen auf $\overline{\Omega}$) ist eine Norm definiert durch

$$\|u\|_C = \sup_{x\in\overline{\Omega}} |u(x)|.$$

Eine analoge Norm ist auf $X = C_0(\Omega)$ (dem Raum der stetigen Funktionen auf Ω mit kompaktem Träger) definiert, wenn das Supremum nur über $x \in \Omega$ genommen wird.

Ist $\|\cdot\|$ eine Norm auf X, so bezeichnet man das Paar $(X, \|\cdot\|)$ als *normierten Raum*, und schreibt in diesem Fall oft $\|\cdot\|_X$. Ist die Norm kanonisch (etwa in Beispiel 1.1 (ii)–(iv)), so wird sie oft weggelassen.

Zwei Normen $\|\cdot\|_1$, $\|\cdot\|_2$ heißen *äquivalent* auf X, falls $c_1, c_2 > 0$ existieren mit

$$c_1\|x\|_2 \leq \|x\|_1 \leq c_2\|x\|_2 \qquad \text{für alle } x \in X.$$

Ist X endlichdimensional, so sind alle Normen auf X äquivalent. Die Konstanten c_1, c_2 hängen dann jedoch von der Dimension N von X ab; die Vermeidung solcher dimensionsabhängiger Konstanten ist einer der Gründe, warum wir Bilder – unabhängig von ihrer eigentlich diskreten Pixelstruktur – als Elemente in einem unendlichdimensionalen Funktionenraum betrachten wollen.

Sind $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ normierte Räume mit $X \subset Y$, so heißt X *stetig eingebettet* in Y, geschrieben $X \hookrightarrow Y$, falls ein $C > 0$ existiert mit

$$\|x\|_Y \leq C\|x\|_X \qquad \text{für alle } x \in X.$$

Wir betrachten nun Abbildungen zwischen normierten Räumen. Seien im Folgenden stets $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ normierte Räume, $U \subset X$, und $F : U \to Y$ eine Abbildung. Wir bezeichnen mit

- $\operatorname{dom} F := U$ den *Definitionsbereich* (englisch „domain“) von F;
- $\ker F := \{x \in U : F(x) = 0\}$ den *Kern* (englisch „kernel“ oder „null space“) von F;
- $\operatorname{ran} F := \{F(x) \in Y : x \in U\}$ das *Bild* (englisch „range“) von F;
- $\operatorname{graph} F := \{(x, y) \in X \times Y : y = F(x)\}$ den *Graph* von F.

Wir sagen, F ist

- *stetig* in $x \in U$, wenn für alle $\varepsilon > 0$ ein $\delta > 0$ existiert mit

$$\|F(x) - F(z)\|_Y \leq \varepsilon \qquad \text{für alle } z \in U \text{ mit } \|x - z\|_X \leq \delta;$$

- *Lipschitz-stetig*, wenn ein $L > 0$ existiert (genannt *Lipschitz-Konstante*) mit

$$\|F(x_1) - F(x_2)\|_Y \leq L\|x_1 - x_2\|_X \qquad \text{für alle } x_1, x_2 \in U.$$

Ist $T : X \to Y$ linear, so ist die Stetigkeit äquivalent zu der Bedingung, dass eine Konstante $C > 0$ existiert mit

$$\|Tx\|_Y \leq C\|x\|_X \qquad \text{für alle } x \in X.$$

Stetige lineare Abbildungen nennt man daher auch *beschränkt*; man spricht auch von einem beschränkten linearen *Operator*. Der Raum $L(X, Y)$ der beschränkten linearen Operatoren ist ein normierter Raum versehen mit der *Operatornorm*

$$\|T\|_{L(X,Y)} = \sup_{x \in X \setminus \{0\}} \frac{\|Tx\|_Y}{\|x\|_X} = \sup_{\|x\|_X = 1} \|Tx\|_Y = \sup_{\|x\|_X \leq 1} \|Tx\|_Y$$

(die gleich der kleinstmöglichen Konstante C in der Definition der Stetigkeit ist). Ist $T \in L(X, Y)$ bijektiv, dann ist die Inverse $T^{-1} : Y \to X$ stetig genau dann, wenn ein $c > 0$ existiert mit

$$c\|x\|_X \leq \|Tx\|_Y \qquad \text{für alle } x \in X.$$

In diesem Fall ist $\|T^{-1}\|_{L(Y,X)} = c^{-1}$ für die größtmögliche Wahl von c.

1.2 STARKE UND SCHWACHE KONVERGENZ

Eine Norm vermittelt auf direkte Weise einen Konvergenzbegriff, die sogenannte *starke Konvergenz*: Eine Folge $\{x_n\}_{n\in\mathbb{N}} \subset X$ konvergiert (stark in X) gegen ein $x \in X$, geschrieben $x_n \to x$, wenn gilt

$$\lim_{n \to \infty} \|x_n - x\|_X = 0.$$

Eine Teilmenge $U \subset X$ nennen wir

- *abgeschlossen*, falls für jede konvergente Folge $\{x_n\}_{n\in\mathbb{N}} \subset U$ auch der Grenzwert $x \in U$ liegt;
- *kompakt*, falls jede Folge $\{x_n\}_{n\in\mathbb{N}} \subset U$ eine konvergente Teilfolge $\{x_{n_k}\}_{k\in\mathbb{N}}$ besitzt, deren Grenzwert $x \in U$ liegt.

Eine Abbildung $F : X \to Y$ ist genau dann stetig, wenn aus $x_n \to x$ auch $F(x_n) \to F(x)$ folgt, und *abgeschlossen*, wenn für $x_n \to x$ und $F(x_n) \to y$ folgt, dass $F(x) = y$ ist (also $\operatorname{graph} F$ eine abgeschlossene Menge ist).

Weiterhin definieren wir für späteren Gebrauch für $x \in X$ und $r > 0$

- die *offene Kugel* $O_r(x) := \{z \in X : \|x - z\|_X < r\}$ und
- die *abgeschlossene Kugel* $K_r(x) := \{z \in X : \|x - z\|_X \leq r\}$.

Die abgeschlossene Kugel um 0 mit Radius 1 bezeichnet man auch als *Einheitskugel* B_X (englisch „unit ball"). Eine Menge $U \subset X$ heißt

- *offen*, falls für alle $x \in U$ ein $r > 0$ existiert mit $O_r(x) \subset U$ (d. h. alle $x \in U$ *innere Punkte* von U sind, deren Menge wir als *Inneres* U^o bezeichnen);
- *beschränkt*, falls sie in einer abgeschlossenen Kugel $K_r(0)$ für ein $r > 0$ enthalten ist;
- *konvex*, falls für $x, y \in U$ auch $\lambda x + (1 - \lambda)y \in U$ für alle $\lambda \in [0, 1]$ gilt.

In normierten Räumen gilt, dass das Komplement einer offenen Menge abgeschlossen ist und umgekehrt (d. h. die abgeschlossenen Mengen im Sinne der Topologie sind genau die (Folgen-)abgeschlossenen Mengen im Sinne unserer Definition). Sowohl offene als auch abgeschlossene Kugeln sind wegen der Norm-Axiome konvex.

Ein normierter Raum X heißt *vollständig*, wenn jede Cauchy-Folge in X konvergiert; man nennt X dann auch *Banachraum*. Alle Räume in Beispiel 1.1 sind Banachräume. Ebenso ist $L(X, Y)$, versehen mit der Operatornorm, ein Banachraum, wenn Y ein Banachraum ist. Für konvexe Teilmengen von Banachräumen gilt folgende nützliche Eigenschaft, die auf dem Satz von Baire beruht.

Lemma 1.2. *Sei X ein Banachraum und $U \subset X$ abgeschlossen und konvex. Dann gilt*

$$U^o = \{x \in U : \text{für alle } h \in X \text{ existiert } \delta > 0 \text{ mit } x + th \in U \text{ für alle } t \in [0, \delta]\}.$$

Die Menge auf der rechten Seite wird auch als *algebraisches Inneres* oder englisch „*core*" bezeichnet, weshalb Lemma 1.2 auch manchmal „core-int-Lemma" genannt wird.

Von wesentlicher Bedeutung wird für uns der Spezialfall $Y = \mathbb{R}$ sein, das heißt der Raum $L(X, \mathbb{R})$ der *linearen stetigen Funktionale* auf X. In diesem Fall bezeichnet man $X^* := L(X, \mathbb{R})$ als *Dualraum* von X. Ist $x^* \in X^*$, so schreibt man auch

$$\langle x^*, x\rangle_X := x^*(x) \in \mathbb{R}.$$

Diese *duale Paarung* soll andeuten, dass man auch x auf x^* wirkend auffassen kann, was später wichtig sein wird. Aus der Definition der Operatornorm folgt sofort, dass gilt

$$|\langle x^*, x\rangle_X| \leq \|x^*\|_{X^*}\|x\|_X \qquad \text{für alle } x \in X, x^* \in X^*. \tag{1.1}$$

In vielen Fällen kann der Dualraum eines Banachraums mit einem bekannten Banachraum identifiziert werden.

Beispiel 1.3. (i) $(\mathbb{R}^N, \|\cdot\|_p)^* \cong (\mathbb{R}^N, \|\cdot\|_q)$ mit $p^{-1}+q^{-1} = 1$, wobei $0^{-1} = \infty$ und $\infty^{-1} = 0$ gesetzt wird. Die duale Paarung ist gegeben durch

$$\langle x^*, x\rangle_p = \sum_{i=1}^{N} x_i^* x_i.$$

(ii) $(\ell^p)^* \cong (\ell^q)$ für $1 < p < \infty$. Die duale Paarung ist gegeben durch

$$\langle x^*, x\rangle_p = \sum_{i=1}^{\infty} x_i^* x_i.$$

Darüber hinaus ist $(\ell^1)^* = \ell^\infty$, aber $(\ell^\infty)^*$ ist selber kein Folgenraum.

(iii) Ebenso ist $L^p(\Omega)^* \cong L^q(\Omega)$ für $1 < p < \infty$. Die duale Paarung ist gegeben durch

$$\langle u^*, u\rangle_p = \int_\Omega u^*(x)u(x)\,dx.$$

Es gilt auch $L^1(\Omega)^* \cong L^\infty(\Omega)$, aber $L^\infty(\Omega)^*$ ist selber kein Funktionenraum.

(iv) $C_0(\Omega)^* \cong \mathcal{M}(\Omega)$, dem Raum der *Radon-Maße*; er enthält unter anderem das Lebesgue-Maß, aber auch Dirac-Maße δ_x für $x \in \Omega$, definiert durch $\delta_x(u) = u(x)$ für $u \in C_0(\Omega)$. Die duale Paarung ist gegeben durch

$$\langle u^*, u\rangle_C = \int_\Omega u(x)\,du^*.$$

Ein zentrales Resultat über Dualräume ist der Satz von Hahn–Banach, auf dem viele der folgenden Aussagen beruhen. Es gibt von ihm eine algebraische und eine geometrische Version.

Satz 1.4 (Hahn–Banach, algebraisch). *Sei X ein normierter Raum. Zu jedem $x \in X$ existiert ein $x^* \in X^*$ mit*

$$\|x^*\|_{X^*} = 1 \qquad \text{und} \qquad \langle x^*, x\rangle_X = \|x\|_X.$$

Satz 1.5 (Hahn–Banach, geometrisch). *Seien X ein normierter Raum und $A, B \subset X$ konvex, nichtleer und disjunkt. Ist A offen, dann existiert ein $x^* \in X^*$ und ein $\lambda \in \mathbb{R}$ mit*

$$\langle x^*, x_1\rangle_X < \lambda \leq \langle x^*, x_2\rangle_X \qquad \text{für alle } x_1 \in A, x_2 \in B.$$

Insbesondere die geometrische Version – auch als *Trennungssatz* bekannt – ist von zentraler Bedeutung für die konvexe Analysis; wir werden sie in der folgenden Variante benötigen, die als *Satz von Eidelheit* bekannt ist.

Folgerung 1.6. *Seien X ein normierter Raum und $A, B \subset X$ konvex und nichtleer. Ist die Menge A^o der inneren Punkte von A nichtleer und disjunkt zu B, dann existiert ein $x^* \in X^* \setminus \{0\}$ und ein $\lambda \in \mathbb{R}$ mit*

$$\langle x^*, x_1\rangle_X \leq \lambda \leq \langle x^*, x_2\rangle_X \qquad \text{für alle } x_1 \in A, x_2 \in B. \tag{1.2}$$

Beweis. Satz 1.5 liefert die Existenz von x^* und λ, so dass die Aussage gilt für alle $x_1 \in A^o$ (sogar mit strikter Ungleichung, woraus auch $x^* \neq 0$ folgt). Es bleibt also zu zeigen, dass $\langle x^*, x\rangle_X \leq \lambda$ auch für $x \in A \setminus A^o$ gilt. Da A^o nichtleer ist, existiert ein $x_0 \in A^o$, d. h. es existiert $r > 0$ mit $O_r(x_0) \subset A$. Aus der Konvexität von A folgt dann, dass für alle $t \in [0, 1]$ und $\tilde{x} \in O_r(x_0)$ auch $t\tilde{x} + (1 - t)x \in A$ ist. Damit ist

$$tO_r(x_0) + (1 - t)x = O_{tr}(tx_0 + (1 - t)x) \subset A, \tag{1.3}$$

und insbesondere gilt $x(t) := tx_0 + (1 - t)x \in A^o$ für alle $t \in (0, 1)$.

Wir finden also eine Folge $\{x_n\}_{n\in\mathbb{N}} \subset A^o$ (zum Beispiel $x_n = x(n^{-1})$) mit $x_n \to x$. Aus der Stetigkeit von $x^* \in X = L(X, \mathbb{R})$ folgt mit Grenzübergang dann

$$\langle x^*, x\rangle_X = \lim_{n\to\infty} \langle x^*, x_n\rangle_X \leq \lambda. \qquad \square$$

Ein normierter Raum wird also in gewisser Weise durch seinen Dualraum charakterisiert. Als direkte Folgerung von Satz 1.4 erhalten wir, dass die Norm in einem Banachraum als Operatornorm dargestellt werden kann.

Folgerung 1.7. *Sei X ein Banachraum. Dann gilt für alle $x \in X$*

$$\|x\|_X = \sup_{\|x^*\|_{X^*} \leq 1} |\langle x^*, x\rangle_X|,$$

und das Supremum wird angenommen.

Ein $x \in X$ können wir also auch als lineares und wegen (1.1) stetiges Funktional auf X^* auffassen, also als Element im *Bidualraum* $X^{**} := (X^*)^*$. Die Einbettung $X \hookrightarrow X^{**}$ wird dabei vermittelt durch die *kanonische Injektion*

$$J : X \to X^{**}, \qquad \langle Jx, x^*\rangle_{X^*} := \langle x^*, x\rangle_X \quad \text{für alle } x^* \in X^*.$$

Offensichtlich ist J linear; aus Satz 1.4 folgt weiterhin $\|Jx\|_{X^{**}} = \|x\|_X$. Ist die kanonische Injektion surjektiv – können wir also X^{**} mit X identifizieren – so nennt man X *reflexiv*.

Endlichdimensionale Räume sind reflexiv, sowie Beispiel 1.1 (ii) und (iii) für $1 < p < \infty$, nicht aber ℓ^1, ℓ^∞, und $L^1(\Omega), L^\infty(\Omega)$ sowie $C(\overline{\Omega})$.

Durch die duale Paarung werden weitere Konvergenzbegriffe erzeugt: die *schwache* Konvergenz auf X sowie die *schwach-∗* Konvergenz auf X^*.

(i) Eine Folge $\{x_n\}_{n\in\mathbb{N}} \subset X$ konvergiert schwach (in X) gegen $x \in X$, geschrieben $x_n \rightharpoonup x$, falls
$$\langle x^*, x_n\rangle_X \to \langle x^*, x\rangle_X \qquad \text{für alle } x^* \in X^*.$$

(ii) Eine Folge $\{x_n^*\}_{n\in\mathbb{N}} \subset X^*$ konvergiert schwach-∗ (in X^*) gegen $x^* \in X^*$, geschrieben $x_n^* \rightharpoonup^* x^*$, falls
$$\langle x_n^*, x\rangle_X \to \langle x^*, x\rangle_X \qquad \text{für alle } x \in X.$$

Die schwache Konvergenz verallgemeinert den Begriff der komponentenweisen Konvergenz in $\mathbb{R}^n$, der – wie aus dem Beweis des Satzes von Heine–Borel ersichtlich ist – im Kontext der Kompaktheit der wesentliche ist. Aus starker Konvergenz folgt schwache Konvergenz; ebenso folgt aus Konvergenz in der Operatornorm (auch *punktweise Konvergenz* genannt) die schwach-∗ Konvergenz. Ist X reflexiv, so stimmen schwache und schwach-∗ Konvergenz (beide in $X = X^{**}$!) überein. In endlichdimensionalen Räumen stimmen alle Konvergenzbegriffe überein.

Konvergiert $x_n \to x$ und $x_n^* \rightharpoonup^* x^*$ oder $x_n \rightharpoonup x$ und $x_n^* \to x^*$, so gilt $\langle x_n^*, x_n\rangle_X \to \langle x^*, x\rangle_X$. Die duale Paarung aus schwach(-∗) konvergenten Folgen konvergiert aber in der Regel nicht! Analog zur starken Konvergenz definiert man nun schwache(-∗) Stetigkeit und Abgeschlossenheit von Abbildungen sowie schwache(-∗) Abgeschlossenheit und Kompaktheit von Mengen. Letztere Eigenschaft wird für uns wesentlich sein; ihre Charakterisierung ist daher ein zentrales Resultat dieses Kapitels.

Satz 1.8 (Eberlein–Šmulyan). *Sei X ein normierter Raum. Dann ist B_X schwach kompakt genau dann, wenn X reflexiv ist.*

In einem reflexiven Raum enthalten also insbesondere alle beschränkten Folgen eine schwach (aber im allgemeinen nicht stark) konvergente Teilfolge. Beachten Sie, dass schwache Abgeschlossenheit eine *stärkere* Forderung ist als Abgeschlossenheit. Für konvexe Mengen stimmen die Begriffe aber überein.

Lemma 1.9. *Eine konvexe Menge $U \subset X$ ist genau dann abgeschlossen, wenn sie schwach abgeschlossen ist.*

Ist X nicht reflexiv (wie z. B. $L^\infty(\Omega)$), so müssen wir auf die schwach-∗ Konvergenz ausweichen.

Satz 1.10 (Banach–Alaoglu). *Ist der normierte Raum X separabel (d. h. es gibt eine abzählbare dichte Teilmenge), so ist B_{X^*} schwach-$*$ kompakt.*

Nach dem Weierstraßschen Approximationssatz sind $C(\overline{\Omega})$ und $L^p(\Omega)$ für $1 \leq p < \infty$ separabel; auch ℓ^p ist separabel für $1 \leq p < \infty$. Also sind beschränkte und schwach-$*$ abgeschlossene Kugeln in ℓ^∞, $L^\infty(\Omega)$ und $\mathcal{M}(\Omega)$ schwach-$*$ kompakt; diese Räume sind aber selber nicht separabel. Beachten Sie aber, dass abgeschlossene konvexe Mengen in nichtreflexiven Räumen *nicht* schwach-$*$ abgeschlossen sein müssen.

Da ein Dualraum den ursprünglichen Raum charakterisiert, ist dies auch der Fall für lineare Operatoren auf diesem Raum. Für $T \in L(X, Y)$ ist durch $T^* : Y^* \to X^*$,

$$\langle T^* y^*, x\rangle_X = \langle y^*, Tx\rangle_Y \qquad \text{für alle } x \in X, y^* \in Y^*$$

der *adjungierte Operator* $T^* \in L(Y^*, X^*)$ definiert. Es gilt stets $\|T^*\|_{L(Y^*,X^*)} = \|T\|_{L(X,Y)}$. Außerdem folgt aus der Stetigkeit von T, dass T^* schwach-$*$ stetig (und T natürlich schwach stetig) ist.

1.3 HILBERTRÄUME

Besonders weitgehende Dualitäts-Aussagen gelten in Hilberträumen. Eine Abbildung $(\cdot,\cdot) : X \times X \to \mathbb{R}$ auf dem Vektorraum X über $\mathbb{R}$ heißt *Skalarprodukt*, falls gilt

(i) $(\alpha x + \beta y, z) = \alpha\,(x, z) + \beta\,(y, z)$ für alle $x, y, z \in X$ und $\alpha, \beta \in \mathbb{R}$;

(ii) $(x, y) = (y, x)$ für alle $x, y \in X$;

(iii) $(x, x) \geq 0$ für alle $x \in X$ mit Gleichheit genau dann, wenn $x = 0$.

Ein Banachraum mit Skalarprodukt $(X, (\cdot,\cdot)_X)$ wird *Hilbertraum* genannt; ist das Skalarprodukt kanonisch, lässt man es weg. Durch das Skalarprodukt wird eine Norm

$$\|x\|_X := \sqrt{(x, x)_X}$$

induziert, die der *Cauchy–Schwarz-Ungleichung* gehorcht:

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

Beispiel 1.3 (i–iii) für $p = 2(= q)$ ist jeweils ein Hilbertraum, wobei das Skalarprodukt der dualen Paarung entspricht und die kanonischen Normen induziert.

Ein Skalarprodukt vermittelt den Begriff der *Orthogonalität*: Ist X ein Hilbertraum, so nennt man $x, y \in X$ *orthogonal*, falls $(x, y)_X = 0$ gilt. Eine Menge $U \subset X$, deren Elemente paarweise orthogonal sind, heißt *Orthogonalsystem*. Gilt sogar für alle $x, y \in U$, dass

$$(x, y)_X = \begin{cases} 1 & \text{falls } x = y, \\ 0 & \text{sonst,} \end{cases}$$

so heißt U *Orthonormalsystem*. Ein Orthonormalsystem ist *vollständig*, falls kein Orthonormalsystem $V \subset X$ mit $U \subsetneq V$ existiert. Ein höchstens abzählbares vollständiges Orthonormalsystem nennt man auch *Orthonormalbasis*.

Jedes Orthonormalsystem $U \subset X$ erfüllt die *Besselsche Ungleichung*:

$$\sum_{y \in U} |\,(x, y)_X\,|^2 \leq \|x\|_X^2 \qquad \text{für alle } x \in X, \tag{1.4}$$

wobei höchstens abzählbar viele Summanden von Null verschieden sind. Ist U vollständig und X ein Hilbertraum, so gilt sogar Gleichheit; dies folgt aus der *Parseval-Identität*

$$(x_1, x_2)_X = \sum_{y \in U} (x_1, y)_X\,(x_2, y)_X \qquad \text{für alle } x_1, x_2 \in X \tag{1.5}$$

(man spricht daher im Fall der Gleichheit in (1.4) auch von der *Parseval-Relation*).

Der für uns wesentliche Punkt ist, dass der Dualraum eines Hilbertraums X mit X identifiziert werden kann.

Satz 1.11 (Fréchet–Riesz). *Sei X ein Hilbertraum. Dann existiert zu jedem $x^* \in X^*$ genau ein $z_{x^*} \in X$ mit $\|x^*\|_{X^*} = \|z_{x^*}\|_X$ und*

$$\langle x^*, x\rangle_X = (x, z_{x^*})_X \qquad \textit{für alle } x \in X.$$

Man bezeichnet z_{x^*} als *Riesz-Repräsentant* von x^*. Die (lineare) Abbildung $J_X : X^* \to X$, $x^* \mapsto z_{x^*}$ wird *Riesz-Isomorphismus* genannt. Mit ihrer Hilfe zeigt man zum Beispiel, dass jeder Hilbertraum reflexiv ist.

Satz 1.11 erlaubt, anstelle der dualen Paarung das Skalarprodukt zu verwenden. So konvergiert $\{x_n\}_{n\in\mathbb{N}}$ schwach gegen x genau dann, wenn gilt

$$(x_n, z)_X \to (x, z)_X \qquad \text{für alle } z \in X. \tag{1.6}$$

Ähnliches gilt für lineare Operatoren auf Hilberträumen. Für Hilberträume X, Y wird zu $T \in L(X, Y)$ der *Hilbertraum-adjungierte Operator* $T^\star \in L(Y, X)$ definiert durch

$$\left(T^\star y, x\right)_X = (Tx, y)_Y \qquad \text{für alle } x \in X, y \in Y.$$

Ist $T^\star = T$, so nennt man T *selbstadjungiert*. Zwischen den beiden Definitionen einer Adjungierten besteht die Beziehung $T^\star = J_X T^* J_Y^{-1}$. Ist der Kontext klar, werden wir nicht in der Notation unterscheiden.

2 GRUNDLAGEN DER VARIATIONSRECHNUNG

Wir betrachten zuerst die Frage nach der Existenz von Minimierern eines (nichtlinearen) Funktionals $F : U \to \mathbb{R}$ für eine Teilmenge U eines Banachraums X. Mit solchen Problemen beschäftigt sich die *Variationsrechnung*.

2.1 DIREKTE METHODE DER VARIATIONSRECHNUNG

Es ist hilfreich, die Beschränkung $\bar{x} \in U$ in das Funktional aufzunehmen, indem wir F auf X erweitern, dafür aber den Wert ∞ zulassen. Wir betrachten also

$$\bar{F} : X \to \overline{\mathbb{R}} := \mathbb{R} \cup \{\infty\}, \qquad \bar{F}(x) = \begin{cases} F(x) & x \in U, \\ \infty & x \in X \setminus U. \end{cases}$$

Dabei wird $\overline{\mathbb{R}}$ mit der üblichen Arithmetik versehen, d. h. $t < \infty$ und $t + \infty = \infty$ für alle $t \in \mathbb{R}$; Subtraktion und Multiplikation von negativen Zahlen mit ∞ und insbesondere $F(x) = -\infty$ sind nicht zugelassen. Existiert überhaupt ein $x \in U$, so kann ein Minimierer $\bar{x}$ also nur in U liegen.

Wir betrachten also in Folge Funktionale $F : X \to \overline{\mathbb{R}}$. Die Menge, auf der F endlich ist, bezeichnet man als (*effektiven*) *Definitionsbereich*

$$\operatorname{dom} F := \{x \in X : F(x) < \infty\}.$$

Ist $\operatorname{dom} F \neq \emptyset$, so nennt man F *eigentlich* (englisch: „proper“).

Wir verallgemeinern nun den Satz von Weierstraß (jede reellwertige stetige Funktion auf kompakten Mengen nimmt ihr Minimum und Maximum an) auf Banachräume und insbesondere auf Funktionen der Form $\bar{F}$. Da wir nur an Minima interessiert sind, reicht dafür eine „einseitige“ Stetigkeit: Man nennt F *unterhalbstetig* in $x \in X$ (englisch: „lower semicontinuous“), falls gilt

$$F(x) \leq \liminf_{n\to\infty} F(x_n) \quad \text{für alle Folgen } \{x_n\}_{n\in\mathbb{N}} \subset X \text{ mit } x_n \to x.$$

Analog definiert man *schwach(-∗) unterhalbstetige* Funktionen über schwach(-∗) konvergente Folgen. Gilt schließlich für jede Folge $\{x_n\}_{n\in\mathbb{N}} \subset X$ mit $\|x_n\|_X \to \infty$ auch $F(x_n) \to \infty$, so heißt F *koerziv*.

Damit haben wir alle Begriffe zur Hand, um das zentrale Resultat der Variationsrechnung zu beweisen.[1]

Satz 2.1. *Sei X ein reflexiver Banachraum und $F : X \to \overline{\mathbb{R}}$ eigentlich, koerziv und schwach unterhalbstetig. Dann hat das Minimierungsproblem*

$$\min_{x\in X} F(x)$$

eine Lösung $\bar{x} \in \operatorname{dom} F$.

Beweis. Der Beweis kann in drei Schritte aufgeteilt werden.

(i) *Zeige, dass eine Minimalfolge existiert.*

Da F eigentlich ist, ist $M := \inf_{x\in X} F(x) < \infty$ (wobei $M = -\infty$ noch nicht ausgeschlossen ist). Wir können also eine Folge $\{y_n\}_{n\in\mathbb{N}} \subset \operatorname{ran} F \setminus \{\infty\} \subset \mathbb{R}$ finden mit $y_n \to M$, d. h. es existiert eine Folge $\{x_n\}_{n\in\mathbb{N}} \subset X$ mit

$$F(x_n) \to M = \inf_{x\in X} F(x).$$

Eine solche Folge wird *Minimalfolge* genannt. Beachten Sie, dass wir aus der Konvergenz von $\{F(x_n)\}_{n\in\mathbb{N}}$ (noch) nicht auf die Konvergenz von $\{x_n\}_{n\in\mathbb{N}}$ schließen können.

(ii) *Zeige, dass die Minimalfolge eine konvergente Teilfolge besitzt.*

Wir zeigen zuerst, dass $\{x_n\}_{n\in\mathbb{N}}$ beschränkt ist. Angenommen, das ist nicht der Fall, d. h. $\|x_n\|_X \to \infty$ für $n \to \infty$. Aus der Koerzivität von F folgt dann auch $F(x_n) \to \infty$, im Widerspruch zu $F(x_n) \to M < \infty$ nach Definition der Minimalfolge. Also ist $\{x_n\}_{n\in\mathbb{N}}$ beschränkt und enthält daher nach dem Satz von Eberlein–Šmulyan eine schwach konvergente Teilfolge $\{x_{n_k}\}_{k\in\mathbb{N}}$ mit Grenzwert $\bar{x} \in X$. Dieser Grenzwert ist Kandidat für einen Minimierer.

(iii) *Zeige, dass dieser Grenzwert ein Minimierer ist.*

Aus der Definition der Minimalfolge folgt, dass auch für die Teilfolge $F(x_{n_k}) \to M$ gilt. Mit der schwachen Unterhalbstetigkeit von F und der Definition des Infimums erhalten wir daher

$$\inf_{x\in X} F(x) \leq F(\bar{x}) \leq \liminf_{k\to\infty} F(x_{n_k}) = M = \inf_{x\in X} F(x) < \infty.$$

[1] Diese Beweis-Strategie, bekannt unter dem Namen *direkte Methode der Variationsrechnung*, wird so häufig angewendet, dass in Forschungsarbeiten üblicherweise nur geschrieben wird: „Die Existenz eines Minimierers folgt aus Standard-Argumenten." Das Grundprinzip geht auf Hilbert zurück; die hier verwendete Formulierung für unterhalbstetige Funktionen stammt von Leonida Tonelli (1885–1946), der damit die moderne Variationsrechnung nachhaltig geprägt hat.

Daraus folgt $\bar{x} \in \operatorname{dom} F$ sowie $\inf_{x\in X} F(x) = F(\bar{x}) > -\infty$ (da F eigentlich ist). Das Infimum wird also in $\bar{x} \in \operatorname{dom} F$ angenommen, und damit ist $\bar{x}$ der gesuchte Minimierer. □

Ist X nicht reflexiv, aber Dualraum eines separablen Banachraums, so zeigt man analog die Existenz von Minimierern schwach-$*$ unterhalbstetiger Funktionale mit dem Satz von Banach–Alaoglu.

Beachten Sie, wie im Beweis die zu verwendende Topologie auf X durch Schritt (ii) und (iii) eingeschränkt wird: Schritt (ii) profitiert von einer groben Topologie (in der mehr Folgen konvergieren), Schritt (iii) von einer feinen (je weniger Folgen konvergieren, desto einfacher ist die $\liminf$-Bedingung zu erfüllen). Da wir in unseren Fällen nicht mehr als die Beschränktheit einer Minimalfolge erwarten können, können wir keine feinere als die schwache Topologie verwenden. Es bleibt daher die Frage, ob genügend (interessante) Funktionale schwach unterhalbstetig sind.

Ein erstes Beispiel sind beschränkte lineare Funktionale: Ist $x^* \in X^*$, so ist

$$F : X \to \mathbb{R}, \qquad x \mapsto \langle x^*, x\rangle_X,$$

schwach stetig (nach Definition der schwachen Konvergenz) und damit insbesondere schwach unterhalbstetig. Ein weiterer Vorzug der (schwachen) Unterhalbstetigkeit ist, dass sie unter bestimmten Operationen erhalten bleibt.

Lemma 2.2. *Seien X, Y Banachräume und sei $F : X \to \overline{\mathbb{R}}$ schwach unterhalbstetig. Dann sind schwach unterhalbstetig*

(i) αF für alle $\alpha \geq 0$;

(ii) $F + G$ für $G : X \to \overline{\mathbb{R}}$ schwach unterhalbstetig;

(iii) $\varphi \circ F$ für $\varphi : \overline{\mathbb{R}} \to \overline{\mathbb{R}}$ unterhalbstetig und monoton steigend;

(iv) $F \circ \Phi$ für $\Phi : Y \to X$ schwach stetig, d. h. aus $y_n \rightharpoonup y$ folgt $\Phi(y_n) \rightharpoonup \Phi(y)$;

(v) $x \mapsto \sup_{i\in I} F_i(x)$ mit $F_i : X \to \overline{\mathbb{R}}$ schwach unterhalbstetig für eine beliebige Menge I.

Beachte, dass Aussage (v) für stetige Funktionen *nicht* gilt!

Beweis. Die Aussagen (i) und (ii) folgen direkt aus den Rechenregeln für den $\liminf$.

Aussage (iii) folgt aus der Monotonie und der schwachen Unterhalbstetigkeit von φ, denn für $x_n \rightharpoonup x$ gilt

$$\varphi(F(x)) \leq \varphi(\liminf_{n\in\mathbb{N}} F(x_n)) \leq \liminf_{n\in\mathbb{N}} \varphi(F(x_n)).$$

Aussage (iv) folgt direkt aus der schwachen Stetigkeit von Φ: Gilt $y_n \rightharpoonup y$, so gilt auch $x_n := \Phi(y_n) \rightharpoonup \Phi(y) =: x$, und aus der Unterhalbstetigkeit von F folgt

$$F(\Phi(y)) \leq \liminf_{n\to\infty} F(\Phi(y_n)).$$

Sei schließlich $\{x_n\}_{n\in\mathbb{N}}$ eine schwach konvergente Folge mit Grenzwert $x \in X$. Dann gilt nach Definition des Supremums

$$F_j(x) \leq \liminf_{n\to\infty} F_j(x_n) \leq \liminf_{n\to\infty} \sup_{i\in I} F_i(x_n) \qquad \text{für alle } j \in I. \tag{2.1}$$

Nehmen wir auf beiden Seiten das Supremum über alle $j \in I$, so folgt die Aussage (v). □

Folgerung 2.3. *Sei X ein Banachraum. Dann ist $\|\cdot\|_X$ eigentlich, koerziv und schwach unterhalbstetig.*

Beweis. Koerzivität und $\operatorname{dom} \|\cdot\|_X = X$ folgen direkt aus der Definition; schwache Unterhalbstetigkeit folgt aus Lemma 2.2 (v) und Folgerung 1.7, denn

$$\|x\|_X = \sup_{\|x^*\|_{X^*}\leq 1} |\langle x^*, x\rangle_X|. \tag{2.2}$$

□

Ein weiteres häufig auftretendes Funktional ist die *Indikator-Funktion*[2] einer Menge $U \subset X$, definiert als

$$\delta_U(x) = \begin{cases} 0 & x \in U, \\ \infty & x \in X \setminus U. \end{cases}$$

Der Zweck dieser Definition ist natürlich, die Minimierung eines Funktionals $F : X \to \mathbb{R}$ unter der Nebenbedingung $x \in U$ auf die unbeschränkte Minimierung von $\bar{F} := F + \delta_U$ zurückführen. Für die Existenz von Minimierern ist daher das folgende Resultat wichtig.

Lemma 2.4. *Sei X ein Banachraum und $U \subset X$. Dann ist $\delta_U : X \to \overline{\mathbb{R}}$*

(i) eigentlich, wenn U nichtleer ist;

(ii) schwach unterhalbstetig, wenn U konvex und abgeschlossen ist;

(iii) koerziv, wenn U beschränkt ist.

[2] nicht zu verwechseln mit der *charakteristischen* Funktion $\mathbb{1}_U$ mit $\mathbb{1}_U(x) = 1$ für $x \in U$ und 0 sonst!

Beweis. Aussage (i) ist klar. Für (ii) betrachte eine schwach konvergente Folge $\{x_n\}_{n\in\mathbb{N}} \subset X$ mit Grenzwert $x \in X$. Ist $x \in U$, dann ist wegen $\delta_U \geq 0$ natürlich

$$\delta_U(x) = 0 \leq \liminf_{n\to\infty} \delta_U(x_n).$$

Sei nun $x \notin U$. Da U konvex und abgeschlossen und daher nach Lemma 1.9 auch schwach abgeschlossen ist, muss ein $N \in \mathbb{N}$ existieren mit $x_n \notin U$ für alle $n \geq N$ (sonst könnten wir – durch Übergang zu einer Teilfolge – eine Folge mit $x_n \rightharpoonup x \in U$ konstruieren, im Widerspruch zur Annahme). Also gilt $\delta_U(x_n) = \infty$ für alle $n \geq N$ und damit

$$\delta_U(x) = \infty = \liminf_{n\to\infty} \delta_U(x_n).$$

Für (iii) sei U beschränkt, d. h. es gebe ein $M > 0$ mit $U \subset K_M(0)$. Gilt $\|x_n\|_X \to \infty$, so existiert ein $N \in \mathbb{N}$ mit $\|x_n\|_X > M$ für alle $n \geq N$, und damit $x_n \notin K_M(0) \supset U$ für alle $n \geq N$. Also gilt auch $\delta_U(x_n) \to \infty$. □

2.2 DIFFERENZIERBARKEIT IN BANACHRÄUMEN

Auch im Banachraum möchte man Minimierer intrinsisch mit Hilfe des Fermatschen Prinzip charakterisieren. Dafür übertragen wir den klassischen Ableitungsbegriff in den Banachraum. Dies geschieht in mehreren Schritten.

Seien X, Y Banachräume, $F : X \to Y$ eine Abbildung, und $x, h \in X$.

- Existiert der einseitige Grenzwert

  $$F'(x; h) := \lim_{t\to 0^+} \frac{F(x + th) - F(x)}{t} \in Y,$$

  so nennen wir diesen *Richtungsableitung* in x in Richtung h.

- Falls $F'(x; h)$ für alle $h \in X$ existiert und durch

  $$DF(x) : X \to Y, h \mapsto F'(x; h)$$

  ein linearer beschränkter(!) Operator definiert wird, so heißt F *Gâteaux-differenzierbar* (in x) und $DF \in L(X, Y)$ *Gâteaux-Ableitung*.

- Gilt zusätzlich

  $$\lim_{\|h\|_X\to 0} \frac{\|F(x + h) - F(x) - DF(x)h\|_Y}{\|h\|_X} = 0,$$

  so heißt F *Fréchet-differenzierbar* (in x) und $F'(x) := DF(x) \in L(X, Y)$ *Fréchet-Ableitung*.

- Ist die Abbildung $x \mapsto F'(x)$ stetig, so heißt F *stetig differenzierbar*.

Der Unterschied zwischen Gâteaux- und Fréchet-Differenzierbarkeit liegt also im Approximationsfehler von F in der Nähe von x durch $F(x) + DF(x)h$: Während für Gâteaux-differenzierbare Funktionen dieser nur beschränkt durch $\|h\|_X$ – also linear in $\|h\|_X$ – sein muss, ist er für Fréchet-differenzierbare Funktionen sogar superlinear in $\|h\|_X$. (Für eine feste Richtung h ist dies natürlich auch für Gâteaux-differenzierbare Funktionen der Fall; für Fréchet-differenzierbare Funktionen ist zusätzlich also Gleichmäßigkeit in h gefordert.)

Ist F Gâteaux-differenzierbar, kann man die Gâteaux-Ableitung berechnen via

$$DF(x)h = \left(\tfrac{d}{dt}F(x+th)\right)\Big|_{t=0}.$$

Offensichtlich sind lineare beschränkte Operatoren $F \in L(X,Y)$ überall Fréchet-differenzierbar mit Ableitung $F'(x) = F \in L(X,Y)$ für alle $x \in X$. Weitere Ableitungen erhält man durch die üblichen Rechenregeln, die genau wie in $\mathbb{R}^n$ gezeigt werden. Beispielhaft sei die folgende Kettenregel angegeben.

Satz 2.5. *Seien X, Y, Z Banachräume und $F : X \to Y$ Fréchet-differenzierbar in $x \in X$ und $G : Y \to Z$ Fréchet-differenzierbar in $y := F(x) \in Y$. Dann ist $G \circ F$ Fréchet-differenzierbar in x und*

$$(G \circ F)'(x) = G'(F(x)) \circ F'(x).$$

Eine analoge Regel für Gâteaux-Ableitungen gilt dagegen nicht!

Wir betrachten nun die Charakterisierung von Minimierern eines differenzierbaren Funktionals $F : X \to \mathbb{R}$.[3]

Satz 2.6 (Fermat-Prinzip). *Sei $F : X \to \mathbb{R}$ Gâteaux-differenzierbar und $\bar{x} \in X$ ein lokaler Minimierer von F. Dann gilt $DF(\bar{x}) = 0 \in X^*$, d. h.*

$$\langle DF(\bar{x}), h\rangle_X = F'(x; h) = 0 \qquad \text{für alle } h \in X.$$

Beweis. Sei $h \in X$ beliebig. Da $\bar{x}$ ein lokaler Minimierer ist, existiert nach Lemma 1.2 ein $\varepsilon > 0$ mit $F(\bar{x}) \leq F(\bar{x} + th)$ für alle $t \in (0, \varepsilon)$ und damit

$$0 \leq \frac{F(\bar{x}+th) - F(\bar{x})}{t} \to F'(\bar{x}; h) = \langle DF(\bar{x}), h\rangle_X \quad \text{für } t \to 0,$$

wobei wir die Gâteaux- und damit Richtungsdifferenzierbarkeit von F verwendet haben. Da die rechte Seite linear in h ist, können wir analog für $-h$ argumentieren und erhalten $\langle DF(\bar{x}), h\rangle_X \leq 0$ und damit die Behauptung. □

[3] In der *indirekten Methode* der Variationsreichnung zeigt man darüber auch die Existenz von Minimierern, etwa als Lösung einer partiellen Differentialgleichung.

Beachten Sie, dass Gâteaux-Ableitungen eines Funktionals $F : X \to \mathbb{R}$ Elemente des *Dualraums* $X^* = L(X, \mathbb{R})$ sind und daher nicht zu Vektoren in X addiert werden können. In Hilberträumen (und insbesondere $\mathbb{R}^n$) kann man aber $DF(x) \in X^*$ mit Hilfe des Satz von Fréchet–Riesz kanonisch mit einem Element $\nabla F(x) \in X$, genannt *Gradient* von F, identifizieren über

$$\langle DF(x), h\rangle_X = (\nabla F(x), h)_X \qquad \text{für alle } h \in X.$$

Als Beispiel betrachten wir für die durch das Skalarprodukt induzierte Norm in einem Hilbertraum das Funktional $F(x) = \frac{1}{2}\|x\|_X^2$. Dann gilt für alle $x, h \in X$

$$F'(x; h) = \lim_{t\to 0^+} \frac{\frac{1}{2}\,(x + th, x + th)_X - \frac{1}{2}\,(x, x)_X}{t} = (x, h)_X = \langle DF(x), h\rangle_X,$$

da das Skalarprodukt für festes x linear in h ist. Die quadrierte Norm ist also Gâteaux-differenzierbar in x mit Ableitung $DF(x) = h \mapsto (x, h)_X \in X^*$; wegen

$$\lim_{\|h\|_X\to 0} \frac{\left|\frac{1}{2}\|x + h\|_X^2 - \frac{1}{2}\|x\|_X^2 - (x, h)_X\right|}{\|h\|_X} = \lim_{\|h\|_X\to 0} \frac{1}{2}\|h\|_X = 0$$

ist sie sogar Fréchet-differenzierbar. Der Gradient $\nabla F(x) \in X$ erfüllt nach Definition

$$(\nabla F(x), h)_X = \langle DF(x), h\rangle_X = (x, h)_X \qquad \text{für alle } h \in X,$$

d. h. $\nabla F(x) = x$. Um zu illustrieren, wie der Gradient (im Gegensatz zur Ableitung) von dem Skalarprodukt abhängt, sei $M \in L(X, X)$ selbstadjungiert und positiv definit (und damit stetig invertierbar). Dann definiert $(x, y)_Z := (Mx, y)_X$ ebenfalls ein Skalarprodukt auf dem Vektorraum X und induziert daher eine (äquivalente) Norm $\|x\|_Z := (x, x)_Z^{1/2}$. Damit ist $(X, (\cdot, \cdot)_Z)$ ebenfalls ein Hilbertraum, den wir mit Z bezeichnen. Betrachte nun das Funktional $\tilde F : Z \to \mathbb{R}$ mit $\tilde F(x) = \frac{1}{2}\|x\|_X^2$ (welches wohldefiniert ist, da $\|\cdot\|_X$ auch eine äquivalente Norm auf Z ist). Dann ist $D\tilde F(x) \in Z^*$ immer noch gegeben durch $\langle D\tilde F(x), h\rangle_Z = (x, h)_X$ für alle $h \in Z$ (bzw. alle $h \in X$, da wir Z über den selben Vektorraum definiert haben). Aber $\nabla \tilde F(x) \in Z^*$ ist nun charakterisiert durch

$$(x, h)_X = \langle D\tilde F(x), h\rangle_Z = (\nabla \tilde F(x), h)_Z = (M\nabla \tilde F(x), h)_X \qquad \text{für alle } h \in Z,$$

d. h. $\nabla \tilde F(x) = M^{-1}x \neq \nabla F(x)$. (Die Situation wird noch komplizierter, wenn M nur auf einem Unterraum positiv definit ist, etwa im Fall $X = L^2(\Omega)$ und $Z = H^1(\Omega)$.)

3 GRUNDLAGEN DER KONVEXEN ANALYSIS

Die klassischen Ableitungsbegriffe des letzten Kapitels sind für unsere Zwecke nicht ausreichend, denn viele interessante Funktionale sind in diesem Sinne nicht differenzierbar; auch Funktionale mit Werten in $\overline{\mathbb{R}}$ können damit nicht behandelt werden. Wir brauchen daher einen Ableitungsbegriff, der allgemeiner als Gâteaux- und Fréchet-Ableitungen ist, aber immer noch ein Fermatsches Prinzip und praktische Rechenregeln erlaubt.

3.1 KONVEXE FUNKTIONEN

Wir betrachten zuerst die Klasse der Funktionale, die solch eine verallgemeinerte Ableitung zulassen. Ein eigentliches Funktional $F : X \to \overline{\mathbb{R}}$ heißt *konvex*, wenn für alle $x, y \in X$ und $\lambda \in [0,1]$ gilt

$$F(\lambda x + (1-\lambda)y) \leq \lambda F(x) + (1-\lambda)F(y) \tag{3.1}$$

(dabei ist der Funktionswert ∞ auf beiden Seiten zugelassen). Gilt für $x \neq y$, $x, y \in \operatorname{dom} F$, und $\lambda \in (0,1)$ sogar

$$F(\lambda x + (1-\lambda)y) < \lambda F(x) + (1-\lambda)F(y),$$

so heißt F *strikt konvex*.

Eine alternative Charakterisierung der Konvexität eines Funktionals $F : X \to \overline{\mathbb{R}}$ basiert auf ihrem *Epigraph*

$$\operatorname{epi} F := \{(x,t) \in X \times \mathbb{R} : F(x) \leq t\}.$$

Lemma 3.1. *Für $F : X \to \overline{\mathbb{R}}$ ist* $\operatorname{epi} F$

(i) nichtleer genau dann, wenn F eigentlich ist;

(ii) konvex genau dann, wenn F konvex ist;

(iii) (schwach) abgeschlossen genau dann, wenn F (schwach) unterhalbstetig ist.

Beweis. Aussage (i) folgt direkt aus der Definition: F ist eigentlich genau dann, wenn ein $x \in X$ und ein $t \in \mathbb{R}$ existiert mit $F(x) \leq t < \infty$, d. h. $(x, t) \in \operatorname{epi} F$.

Für (ii) sei F konvex und seien $(x, r), (y, s) \in \operatorname{epi} F$ gegeben. Für beliebige $\lambda \in [0, 1]$ folgt dann aus (3.1), dass gilt

$$F(\lambda x + (1 - \lambda)y) \leq \lambda F(x) + (1 - \lambda)F(y) \leq \lambda r + (1 - \lambda)s,$$

d. h. es ist

$$\lambda(x, r) + (1 - \lambda)(y, s) = (\lambda x + (1 - \lambda)y, \lambda r + (1 - \lambda)s) \in \operatorname{epi} F$$

und damit $\operatorname{epi} F$ konvex. Sei umgekehrt $\operatorname{epi} F$ konvex und $x, y \in X$ beliebig, wobei wir $F(x) < \infty$ und $F(y) < \infty$ annehmen können (ansonsten ist (3.1) trivialerweise erfüllt). Offensichtlich sind $(x, F(x)), (y, F(y)) \in \operatorname{epi} F$. Aus der Konvexität von $\operatorname{epi} F$ folgt dann, dass für alle $\lambda \in [0, 1]$ gilt

$$(\lambda x + (1 - \lambda)y, \lambda F(x) + (1 - \lambda)F(y)) = \lambda(x, F(x)) + (1 - \lambda)(y, F(y)) \in \operatorname{epi} F$$

und damit nach Definition von $\operatorname{epi} F$ auch (3.1).

Nun zu (iii): Sei zuerst F unterhalbstetig und $\{(x_n, t_n)\}_{n\in\mathbb{N}} \subset \operatorname{epi} F$ eine beliebige Folge mit $(x_n, t_n) \to (x, t) \in X \times \mathbb{R}$. Dann gilt

$$F(x) \leq \liminf_{n\to\infty} F(x_n) \leq \limsup_{n\to\infty} t_n = t,$$

d. h $(x, t) \in \operatorname{epi} F$. Sei umgekehrt $\operatorname{epi} F$ abgeschlossen und angenommen, F ist eigentlich (sonst ist die Aussage trivial) und nicht unterhalbstetig. Dann existiert eine Folge $\{x_n\}_{n\in\mathbb{N}} \subset X$ mit $x_n \to x \in X$ und

$$F(x) > \liminf_{n\to\infty} F(x_n) =: M \in [-\infty, \infty).$$

Wir machen nun eine Fallunterscheidung:

a) $x \in \operatorname{dom} F$: Dann können wir eine Teilfolge auswählen, die wir wieder mit $\{x_n\}_{n\in\mathbb{N}}$ bezeichnen, so dass ein $\varepsilon > 0$ existiert mit $F(x_n) \leq F(x) - \varepsilon$ und damit $(x_n, F(x) - \varepsilon) \in \operatorname{epi} F$ für alle $n \in \mathbb{N}$. Wegen $x_n \to x$ folgt aus der Abgeschlossenheit von $\operatorname{epi} F$ auch $(x, F(x) - \varepsilon) \in \operatorname{epi} F$ und damit $F(x) \leq F(x) - \varepsilon$, im Widerspruch zu $\varepsilon > 0$.

b) $x \notin \operatorname{dom} F$: In diesem Fall argumentiert man analog mit $F(x_n) \leq M + \varepsilon$ für $M > -\infty$ bzw. $F(x_n) \leq \varepsilon$ für $M = -\infty$, um einen Widerspruch zu $F(x) = \infty$ zu erhalten.

Genauso zeigt man die Äquivalenz von schwacher Unterhalbstetigkeit und schwacher Abgeschlossenheit. □

Zusammen mit Lemma 1.9 erhalten wir daraus sofort

Folgerung 3.2. *Sei $F : X \to \overline{\mathbb{R}}$ konvex. Dann ist F schwach unterhalbstetig genau dann, wenn F unterhalbstetig ist.*

Direkt aus der Definition folgt die Konvexität

- der Norm $\|\cdot\|_X$ in einem normierten Raum X;
- der Indikatorfunktion δ_C für eine konvexe Menge C.

Ist X ein Hilbertraum, so ist $F(x) = \|x\|_X^2$ sogar strikt konvex: Für $x, y \in X$ mit $x \neq y$ und $\lambda \in (0,1)$ beliebig gilt

$$\begin{aligned}
\|\lambda x + (1-\lambda)y\|_X^2 &= (\lambda x + (1-\lambda)y, \lambda x + (1-\lambda)y)_X \\
&= \lambda^2 (x,x)_X + 2\lambda(1-\lambda)(x,y)_X + (1-\lambda)^2 (y,y)_X \\
&= \lambda\Big(\lambda (x,x) - (1-\lambda)(x-y,x)_X + (1-\lambda)(y,y)_X\Big) \\
&\quad + (1-\lambda)\Big(\lambda (x,x) + \lambda (x-y,y)_X + (1-\lambda)(y,y)_X\Big) \\
&= (\lambda + (1-\lambda))\Big(\lambda (x,x)_X + (1-\lambda)(y,y)_X\Big) - \lambda(1-\lambda)(x-y,x-y)_X \\
&= \lambda\|x\|_X^2 + (1-\lambda)\|y\|_X^2 - \lambda(1-\lambda)\|x-y\|_X^2 \\
&< \lambda\|x\|_X^2 + (1-\lambda)\|y\|_X^2.
\end{aligned}$$

Weitere Beispiele lassen sich durch folgende Operationen erzeugen.

Lemma 3.3. *Seien X, Y normierte Räume und sei $F : X \to \overline{\mathbb{R}}$ konvex. Dann sind konvex*

(i) *αF für alle $\alpha \geq 0$;*

(ii) *$F + G$ für $G : X \to \overline{\mathbb{R}}$ konvex (ist F oder G strikt konvex, so auch $F + G$);*

(iii) *$\varphi \circ F$ für $\varphi : \overline{\mathbb{R}} \to \overline{\mathbb{R}}$ konvex und monoton steigend;*

(iv) *$F \circ A$ für $A \in L(Y, X)$;*

(v) *$x \mapsto \sup_{i\in I} F_i(x)$ für eine beliebige Indexmenge I und F_i konvex für alle $i \in I$.*

Nach all der Vorarbeit können wir nun schnell das Hauptresultat über die Existenz von Lösungen konvexer Minimierungsaufgaben beweisen.

Satz 3.4. *Sei X ein reflexiver Banachraum und seien*

(i) *$U \subset X$ nichtleer, abgeschlossen, und konvex,*

(ii) *$F : X \to \overline{\mathbb{R}}$ eigentlich, konvex, und unterhalbstetig mit $\operatorname{dom} F \cap U \neq \emptyset$,*

(iii) *U beschränkt oder F koerziv.*

Dann hat das Problem

$$\min_{x \in U} F(x)$$

eine Lösung $\bar{x} \in U \cap \operatorname{dom} F$. *Ist* F *strikt konvex, so ist die Lösung eindeutig.*

Beweis. Wir betrachten das Funktional $\bar{F} = F + \delta_U$. Aus Voraussetzung (i) folgt mit Lemma 2.2, dass δ_U eigentlich, konvex und schwach unterhalbstetig ist. Wegen (ii) existiert ein Punkt $x_0 \in U$ mit $\bar{F}(x_0) < \infty$, daher ist auch $\bar{F}$ eigentlich, konvex und damit sogar unterhalbstetig. Aus (iii) folgt schließlich die Koerzivität von $\bar{F}$ (da die Summe koerziv ist, sobald es einer dieser Summanden ist). Da X reflexiv ist, können wir nun Satz 2.1 anwenden und erhalten die Existenz eines Minimierers $\bar{x} \in \operatorname{dom} \bar{F} = U \cap \operatorname{dom} F$ von $\bar{F}$ mit

$$F(\bar{x}) = \bar{F}(\bar{x}) \leq \bar{F}(x) = F(x) \qquad \text{für alle } x \in U,$$

d. h. $\bar{x}$ die gesuchte Lösung.

Sei nun F strikt konvex, und seien $\bar{x}, \bar{x}' \in U$ zwei verschiedene Minimierer, d. h. $\bar{x} \neq \bar{x}'$ aber $F(\bar{x}) = F(\bar{x}') = \min_{x \in U} F(x)$. Dann ist für alle $\lambda \in (0,1)$ wegen der Konvexität von U auch

$$x_\lambda := \lambda \bar{x} + (1-\lambda)\bar{x}' \in U,$$

aber wegen der strikten Konvexität von F gilt

$$F(x_\lambda) < \lambda F(\bar{x}) + (1-\lambda)F(\bar{x}') = F(\bar{x}),$$

im Widerspruch zu $F(\bar{x}) \leq F(x)$ für alle $x \in U$. □

3.2 KONVEXE SUBDIFFERENTIALE

Wir wenden uns nun der Charakterisierung von Minimierern konvexer Funktionen durch ein Fermatsches Prinzip zu. Ein erster Kandidat für den nötigen Ableitungsbegriff ist die Richtungsableitung, denn diese existiert (zumindest in den erweiterten reellen Zahlen) für jede konvexe Funktion.

Lemma 3.5. *Sei* $F : X \to \overline{\mathbb{R}}$ *konvex und seien* $x \in \operatorname{dom} F$ *und* $h \in X$ *gegeben. Dann gilt:*

(i) *Die Funktion*

$$\varphi : (0,\infty) \to \overline{\mathbb{R}}, \qquad t \mapsto \frac{F(x+th) - F(x)}{t},$$

ist monoton steigend.

(ii) *Der Grenzwert* $F'(x;h) = \lim_{t \to 0^+} \varphi(t) \in [-\infty, \infty]$ *existiert und erfüllt*

$$F'(x;h) \leq F(x+h) - F(x). \tag{3.2}$$

Beweis. Zu (i): Durch Einsetzen und Umformen sieht man, dass für alle $0 < s \leq t$ die Bedingung $\varphi(s) \leq \varphi(t)$ äquivalent ist zu

$$F(x + sh) \leq \frac{s}{t} F(x + th) + \left(1 - \frac{s}{t}\right) F(x).$$

Dies folgt aber wegen $x + sh = (1 - \frac{s}{t})x + \frac{s}{t}(x + th)$ aus der Konvexität von F.

Aussage (ii) folgt sofort nun aus (i) wegen

$$F'(x; h) = \lim_{t \to 0^+} \varphi(t) = \inf_{t>0} \varphi(t) \leq \varphi(1) = F(x + h) - F(x). \qquad \square$$

Leider liefert dieser Begriff noch nicht das Gewünschte, denn die konvexe Funktion $f : \mathbb{R} \to \mathbb{R}$, $f(t) = |t|$ hat ein Minimum in $t = 0$, dort aber nur die Richtungsableitung $f'(0; h) = |h| > 0$ für $h \in \mathbb{R} \setminus \{0\}$. Es gilt also nicht $f'(0; h) = 0$ für irgendein $h \neq 0$, aber es gilt zumindest $0 \leq f'(0; h)$ für alle $h \in \mathbb{R}$. Diese Bedingung wollen wir auf allgemeine normierte Räume verallgemeinern. Dafür betrachten wir für $F : X \to \overline{\mathbb{R}}$ und $x \in \operatorname{dom} F$ die Menge

$$\{x^* \in X^* : \langle x^*, h \rangle_X \leq F'(x; h) \text{ für alle } h \in X\}.$$

Diese Menge (die auch leer sein kann!) ist mit Hilfe von Lemma 3.5 auch ohne Richtungsableitung charakterisierbar.

Lemma 3.6. *Seien $F : X \to \overline{\mathbb{R}}$ konvex und $x \in \operatorname{dom} F$. Dann sind für $x^* \in X^*$ äquivalent*

(i) $\langle x^*, h \rangle_X \leq F'(x; h) \qquad$ *für alle* $h \in X$;

(ii) $\langle x^*, h \rangle_X \leq F(x + h) - F(x) \qquad$ *für alle* $h \in X$.

Beweis. Gilt (i), so folgt direkt aus Lemma 3.5 (ii)

$$\langle x^*, h \rangle_X \leq F'(x; h) \leq F(x + h) - F(x) \quad \text{für alle } h \in X.$$

Gilt (ii) für alle $h \in X$, so auch für th für alle $h \in X$ und $t > 0$. Division durch t und Grenzübergang liefert dann

$$\langle x^*, h \rangle_X \leq \lim_{t \to 0^+} \frac{F(x + th) - F(x)}{t} = F'(x; h). \qquad \square$$

Führen wir $\tilde{x} = x + h \in X$ ein, so führt die zweite Bedingung auf die folgende Definition. Für $F : X \to \overline{\mathbb{R}}$ und $x \in \operatorname{dom} F$ definieren wir das *(konvexe) Subdifferential* als

$$\partial F(x) := \{x^* \in X^* : \langle x^*, \tilde{x} - x \rangle_X \leq F(\tilde{x}) - F(x) \quad \text{für alle } \tilde{x} \in X\}. \tag{3.3}$$

(Beachten Sie, dass $\tilde{x} \notin \operatorname{dom} F$ zugelassen ist, da dann die Ungleichung trivialerweise erfüllt ist.) Für $x \notin \operatorname{dom} F$ setzen wir $\partial F(x) = \emptyset$. Ein Element $\xi \in \partial F(x)$ heißt *Subgradient.*

Diese Definition liefert nun das Gewünschte.

Satz 3.7. *Seien* $F : X \to \overline{\mathbb{R}}$ *und* $\bar{x} \in \operatorname{dom} F$. *Dann sind äquivalent:*

(i) $0 \in \partial F(\bar{x})$.

(ii) $F(\bar{x}) = \min_{x \in X} F(x)$;

Beweis. Dies folgt direkt aus den Definitionen: Es ist $0 \in \partial F(\bar{x})$ genau dann, wenn gilt

$$0 = \langle 0, \tilde{x} - \bar{x} \rangle_X \leq F(\tilde{x}) - F(\bar{x}) \qquad \text{für alle } \tilde{x} \in X,$$

d. h. $F(\bar{x}) \leq F(\tilde{x})$ für alle $\tilde{x} \in X$. □

Dies entspricht auch der geometrischen Anschauung: Für $X = \mathbb{R} = X^*$ beschreibt $\tilde{f}(\tilde{x}) = f(x) + \xi(\tilde{x} - x)$ mit $\xi \in \partial f(x)$ eine Tangente an $(x, f(x))$ mit Steigung ξ; die Bedingung $\xi = 0 \in \partial f(\bar{x})$ bedeutet also, dass f in $\bar{x}$ eine waagerechte Tangente hat.[1]

Wir betrachten einige Beispiele. Zunächst ist aus der Konstruktion ersichtlich, dass das Subdifferential die Gâteaux-Ableitung verallgemeinert.

Satz 3.8. *Sei* $F : X \to \overline{\mathbb{R}}$ *konvex und Gâteaux-differenzierbar in* x. *Dann ist* $\partial F(x) = \{DF(x)\}$.

Beweis. Nach Definition der Gâteaux-Ableitung gilt

$$\langle DF(x), h \rangle_X = DF(x)h = F'(x; h) \quad \text{für alle } h \in X.$$

Aus Lemma 3.6 folgt nun mit $\tilde{x} = x + h$ sofort $DF(x) \in \partial F(x)$.

Umgekehrt folgt aus $\xi \in \partial F(x)$ mit $h := \tilde{x} - x \in X$, dass gilt

$$\langle \xi, h \rangle_X \leq F'(x; h) = \langle DF(x), h \rangle_X.$$

Da $\tilde{x} \in X$ beliebig war, gilt dies für alle $h \in X$. Supremum über alle h mit $\|h\|_X \leq 1$ liefert dann $\|\xi - DF(x)\|_{X^*} \leq 0$, d. h. $\xi = DF(x)$. □

Natürlich möchten wir auch Subdifferentiale von Funktionen berechnen, die nicht differenzierbar sind. Das kanonische Beispiel ist die Norm $\|x\|_X$ in einem normierten Raum, die ja in $x = 0$ nicht differenzierbar ist.

[1] Beachten Sie, dass in Satz 3.7 nirgendwo die Konvexität von F eingeht! Tatsächlich charakterisiert $0 \in \partial F(\bar{x})$ die *globalen* Minimierer jeder Funktion. Nichtkonvexe Funktionen können aber auch lokale Minimierer haben, für die die Subdifferentialinklusion nicht erfüllt ist. Tatsächlich sind (konvexe) Subdifferentiale nichtkonvexer Funktionen in der Regel leer. Dies führt insbesondere für den Beweis einiger Rechenregeln zu Problemen, für die wir in der Tat Konvexität voraussetzen müssen.

Satz 3.9. *Für $x \in X$ ist*

$$\partial(\|\cdot\|_X)(x) = \begin{cases} \{x^* \in X^* : \langle x^*, x\rangle_X = \|x\|_X \text{ und } \|x^*\|_{X^*} = 1\} & \text{falls } x \neq 0, \\ B_{X^*} & \text{falls } x = 0. \end{cases} \tag{3.4}$$

Beweis. Für $x = 0$ ist nach Definition $\xi \in \partial(\|\cdot\|_X)(x)$ genau dann, wenn gilt

$$\langle \xi, \tilde{x}\rangle_X \leq \|\tilde{x}\|_X \qquad \text{für alle } \tilde{x} \in X \setminus \{0\}$$

(für $\tilde{x} = 0$ ist die Ungleichung trivial). Dies ist aber wegen der Definition der Operatornorm äquivalent mit $\|\xi\|_{X^*} \leq 1$.

Sei nun $x \neq 0$ und betrachte $\xi \in \partial(\|\cdot\|_X)(x)$. Indem wir nacheinander $\tilde{x} = 0$ und $\tilde{x} = 2x$ in die Definition (3.3) einsetzen, erhalten wir

$$\|x\|_X \leq \langle \xi, x\rangle_X = \langle \xi, 2x - x\rangle \leq \|2x\|_X - \|x\|_X = \|x\|_X,$$

d. h. $\langle \xi, x\rangle_X = \|x\|_X$. Analog haben wir für alle $\tilde{x} \in X$, dass gilt

$$\langle \xi, \tilde{x}\rangle_X = \langle \xi, (\tilde{x} + x) - x\rangle_X \leq \|\tilde{x} + x\|_X - \|x\|_X \leq \|\tilde{x}\|_X,$$

woraus wie im Fall $x = 0$ folgt $\|\xi\|_{X^*} \leq 1$. Für $\tilde{x} = x/\|x\|_X$ gilt nun

$$\langle \xi, \tilde{x}\rangle_X = \|x\|_X^{-1}\langle \xi, x\rangle_X = \|x\|_X^{-1}\|x\|_X = 1.$$

Also ist tatsächlich $\|\xi\|_{X^*} = 1$.

Es sei umgekehrt $x^* \in X^*$ mit $\langle x^*, x\rangle_X = \|x\|_X$ und $\|x^*\|_{X^*} = 1$. Dann gilt mit (1.1) für alle $\tilde{x} \in X$ die Relation

$$\langle x^*, \tilde{x} - x\rangle_X = \langle x^*, \tilde{x}\rangle_X - \langle x^*, x\rangle_X \leq \|\tilde{x}\|_X - \|x\|_X,$$

und daher nach Definition $x^* \in \partial(\|\cdot\|_X)(x)$ □

Für den Fall $X = \mathbb{R}$ erhalten wir daraus das Subdifferential der Betragsfunktion als

$$\partial(|\cdot|)(t) = \operatorname{sign}(t) := \begin{cases} \{1\} & \text{falls } t > 0, \\ \{-1\} & \text{falls } t < 0, \\ [-1, 1] & \text{falls } t = 0. \end{cases} \tag{3.5}$$

Durch komponentenweise Betrachtung erhält man daraus die Charakterisierung des Subdifferentials der Norm in $(\mathbb{R}^N, \|\cdot\|_1)$.

Schließlich haben wir auch eine einfache Darstellung für das Subdifferential der Indikatorfunktion einer konvexen Menge $C \subset X$. Für $x \in C = \operatorname{dom} \delta_C$ gilt nämlich

$$\begin{aligned} x^* \in \partial\delta_C(x) &\Leftrightarrow \langle x^*, \tilde{x} - x\rangle_X \leq \delta_C(\tilde{x}) \quad \text{für alle } \tilde{x} \in X \\ &\Leftrightarrow \langle x^*, \tilde{x} - x\rangle_X \leq 0 \quad \text{für alle } \tilde{x} \in C, \end{aligned}$$

da die Ungleichung für alle $\tilde{x} \notin C$ trivialerweise erfüllt ist. Die Menge $\partial\delta_C(x)$ nennt man auch *Normalenkegel* an C in x. Wieder ist das Beispiel $X = \mathbb{R}$ erhellend: Betrachte $C = [-1, 1]$ und $t \in C$. Dann ist $\xi \in \partial\delta_{[-1,1]}(t)$ genau dann, wenn $\xi(\tilde{t} - t) \leq 0$ für alle $\tilde{t} \in [-1, 1]$ gilt. Wir machen nun eine Fallunterscheidung:

1. Fall: $t = 1$. Dann ist $\tilde{t} - t \in [-2, 0]$ und damit gilt die Bedingung genau dann, wenn $\xi \geq 0$ ist.

2. Fall: $t = -1$. Dann ist $\tilde{t} - t \in [0, 2]$ und damit gilt die Bedingung genau dann, wenn $\xi \leq 0$ ist.

3. Fall: $t \in (-1, 1)$. Dann kann $\tilde{t} - t$ sowohl positive als auch negative Werte annehmen, und damit muss $\xi = 0$ sein.

Also ist

$$\partial\delta_{[-1,1]}(t) = \begin{cases} [0, \infty) & \text{falls } t = 1, \\ (-\infty, 0] & \text{falls } t = -1, \\ \{0\} & \text{falls } t \in (-1, 1), \\ \emptyset & \text{falls } t \in \mathbb{R} \setminus [-1, 1]. \end{cases}$$

Dies entspricht den aus der linearen Optimierung bekannten *Komplementaritätsbedingungen* für Lagrange-Multiplikatoren zu den Ungleichungen $-1 \leq t \leq 1$. Analoges gilt für komponentenweise Schranken in $X = \mathbb{R}^N$.

Subdifferentiale weiterer Funktionale erhält man durch Rechenregeln. Es ist naheliegend, dass diese umso aufwändiger zu beweisen sind, je schwächer der Differenzierbarkeitsbegriff ist (d. h. je mehr Funktionen in diesem Sinne differenzierbar sind). Die ersten beiden Regeln folgen noch direkt aus der Definition.

Lemma 3.10. *Für $F : X \to \overline{\mathbb{R}}$ konvex und $x \in \operatorname{dom} F$ gilt*

(i) $\partial(\lambda F)(x) = \lambda(\partial F(x)) := \{\lambda\xi : \xi \in \partial F(x)\}$ *für* $\lambda > 0$;

(ii) $\partial F(\cdot + x_0)(x) = \partial F(x + x_0)$ *für* $x_0 \in X$ *mit* $x + x_0 \in \operatorname{dom} F$.

Schon die Summenregel ist deutlich aufwändiger.

Satz 3.11 (Summenregel). *Seien $F, G : X \to \overline{\mathbb{R}}$ konvex. Dann gilt für alle $x \in \operatorname{dom} F \cap \operatorname{dom} G$*

$$\partial F(x) + \partial G(x) \subset \partial(F + G)(x).$$

Existiert ein $x_0 \in \operatorname{dom} F \cap \operatorname{dom} G$ mit F stetig in x_0, so gilt Gleichheit.

Beweis. Die Inklusion folgt durch Addition der Definitionen der Subdifferentiale. Seien daher $x \in \operatorname{dom} F \cap \operatorname{dom} G$ und $\xi \in \partial(F+G)(x)$, erfüllen also

$$\langle \xi, \tilde{x} - x\rangle_X \leq (F(\tilde{x}) + G(\tilde{x})) - (F(x) + G(x)) \quad \text{für alle } \tilde{x} \in X. \tag{3.6}$$

Unser Ziel ist nun, mit Hilfe der Charakterisierung konvexer Funktionale durch ihren Epigraphen und des Trennungssatzes ein lineares Funktional $\zeta \in \partial G(x) \subset X^*$ zu finden mit $\xi - \zeta \in \partial F(x)$, d. h.

$$\begin{aligned} F(\tilde{x}) - F(x) - \langle \xi, \tilde{x} - x\rangle_X &\geq \langle \zeta, x - \tilde{x}\rangle_X \quad \text{für alle } \tilde{x} \in \operatorname{dom} F,\\ G(x) - G(\tilde{x}) &\leq \langle \zeta, x - \tilde{x}\rangle_X \quad \text{für alle } \tilde{x} \in \operatorname{dom} G. \end{aligned}$$

Wir definieren dafür die Mengen

$$\begin{aligned} C_1 &:= \{(\tilde{x}, t - (F(x) - \langle \xi, x\rangle_X)) : \tilde{x} \in \operatorname{dom} F, t \geq F(\tilde{x}) - \langle \xi, \tilde{x}\rangle_X\},\\ C_2 &:= \{(\tilde{x}, G(x) - t) : \tilde{x} \in \operatorname{dom} G, t \geq G(\tilde{x})\}, \end{aligned}$$

d. h.

$$C_1 = \operatorname{epi}(F - \xi) - (0, F(x) - \langle \xi, x\rangle_X), \qquad C_2 = -(\operatorname{epi} G - (0, G(x))).$$

Wegen $x \in \operatorname{dom} F \cap \operatorname{dom} G$ sind C_1 und C_2 nichtleer. Weiter kann man direkt nachrechnen, dass C_1 und C_2 konvex sind. Weiter ist x_0 ein innerer Punkt von $\operatorname{dom} F = \operatorname{dom}(F - \xi)$ (sonst wäre F nicht stetig in x_0) und damit (x_0, α) für α groß genug ein innerer Punkt von C_1 (denn F ist stetig und damit in einer Umgebung $U \subset (\operatorname{dom} F)^o$ von x_0 beschränkt; wir finden also ein offenes Intervall $I \subset \mathbb{R}$ mit $U \times I \subset C_1$, wobei $U \times I$ nach Definition der Produkttopologie offen in $X \times \mathbb{R}$ ist). Das Innere $(C_1)^o$ von C_1 ist also nichtleer. Bleibt zu zeigen, dass $(C_1)^o$ und C_2 disjunkt sind. Dafür sei $(\tilde{x}, \alpha) \in (C_1)^o \cap C_2$, für das nach Definition von C_1 und C_2 (und einigem Umformen) gilt

$$F(\tilde{x}) - F(x) - \langle \xi, \tilde{x} - x\rangle_X < \alpha \leq G(x) - G(\tilde{x}), \tag{3.7}$$

im Widerspruch zu (3.6). Folgerung 1.6 liefert also ein $(x^*, s) \in (X^* \times \mathbb{R}) \setminus \{(0,0)\} = (X \times \mathbb{R})^* \setminus \{(0,0)\}$ und ein $\lambda \in \mathbb{R}$ mit

$$\langle x^*, \tilde{x}\rangle_X + s(t - (F(x) - \langle \xi, x\rangle_X)) \leq \lambda, \quad \tilde{x} \in \operatorname{dom} F, t \geq F(\tilde{x}) - \langle \xi, \tilde{x}\rangle_X, \tag{3.8}$$

$$\langle x^*, \tilde{x}\rangle_X + s(G(x) - t) \geq \lambda, \quad \tilde{x} \in \operatorname{dom} G, t \geq G(\tilde{x}). \tag{3.9}$$

Wir zeigen nun, dass $s < 0$ ist. Für $s = 0$ folgt mit $\tilde{x} = x_0 \in \operatorname{dom} F \cap \operatorname{dom} G$ sofort der Widerspruch

$$\langle x^*, x_0\rangle_X < \lambda \leq \langle x^*, x_0\rangle_X,$$

da (x_0, α) für α groß genug innerer Punkt von C_1 ist und daher nach Satz 1.5 die Trennung sogar mit strikter Ungleichung erfüllt ist. Gilt $s > 0$, so ist für $t > F(x) - \langle \xi, x\rangle_X$ die Klammer in (3.8) positiv, und $t \to \infty$ mit $\tilde{x}$ fest führt zum Widerspruch zur Beschränktheit durch λ.

Also ist $s < 0$, und aus (3.8) mit $t = F(\tilde{x}) - \langle \xi, \tilde{x} \rangle_X$ und aus (3.9) mit $t = G(\tilde{x})$ folgt

$$F(\tilde{x}) - F(x) + \langle \xi, \tilde{x} - x \rangle_X \geq s^{-1}(\lambda - \langle x^*, \tilde{x} \rangle_X), \quad \text{für alle } \tilde{x} \in \operatorname{dom} F,$$
$$G(x) - G(\tilde{x}) \leq s^{-1}(\lambda - \langle x^*, \tilde{x} \rangle_X), \quad \text{für alle } \tilde{x} \in \operatorname{dom} G.$$

Setzen wir $\tilde{x} = x \in \operatorname{dom} F \cap \operatorname{dom} G$ in beiden Ungleichungen, so folgt sofort $\lambda = \langle x^*, x \rangle_X$. Damit ist $\zeta = s^{-1}x^* \in X^*$ das gewünschte Funktional mit $(\xi - \zeta) \in \partial F(x)$ und $\zeta \in \partial G(x)$, d. h. $\xi \in \partial F(x) + \partial G(x)$. ◻

Daraus erhält man per Induktion Summenregeln für beliebig viele Summanden (wobei in x_0 alle bis auf ein Summand stetig sein müssen). Analog beweist man eine Kettenregel für lineare Operatoren.

Satz 3.12 (Kettenregel). *Seien X, Y normierte Räume, $A \in L(X, Y)$, und $F : Y \to \overline{\mathbb{R}}$ konvex. Existiert ein $x_0 \in X$ so, dass F stetig in Ax_0 ist, dann gilt für alle $x \in \operatorname{dom}(F \circ A)$*

$$\partial(F \circ A)(x) = A^* \partial F(Ax) := \{A^* y^* : y^* \in \partial F(Ax)\}.$$

Beweis. Die rechte Inklusion folgt wieder direkt aus der Definition: Für $\eta \in \partial F(Ax) \subset Y^*$ gilt insbesondere für alle $A\tilde{x} \in Y$, $\tilde{x} \in X$, dass

$$F(A\tilde{x}) - F(Ax) \geq \langle \eta, A\tilde{x} - Ax \rangle_Y = \langle A^*\eta, \tilde{x} - x \rangle_X,$$

d. h. $\xi := A^*\eta \in \partial(F \circ A) \subset X^*$.

Sei nun $x \in \operatorname{dom}(F \circ A)$ und $\xi \in \partial(F \circ A)(x)$, d. h.

$$F(Ax) + \langle \xi, \tilde{x} - x \rangle_X \leq F(A\tilde{x}) \quad \text{für alle } \tilde{x} \in X. \tag{3.10}$$

Ähnlich wie im Beweis von Satz 3.11 zeigen wir durch Trennung die Existenz eines $\eta \in \partial F(Ax)$ mit $\xi = A^*\eta$. Wir definieren dafür

$$C_1 := \operatorname{epi} F, \qquad C_2 := \{(A\tilde{x}, F(Ax) + \langle \xi, \tilde{x} - x \rangle_X) : \tilde{x} \in X\}.$$

Man vergewissert sich leicht, dass C_2 nichtleer und konvex ist. Ebenso ist C_1 nichtleer und konvex und besitzt nach Voraussetzung den inneren Punkt (Ax_0, α) für α groß genug. Schließlich gilt für $(\tilde{y}, t) \in (C_1)^o \cap C_2$ sowohl $F(\tilde{y}) < t$ als auch $\tilde{y} = A\tilde{x}$ und $t = F(Ax) + \langle \xi, \tilde{x} - x \rangle_X$ für ein $\tilde{x} \in X$, woraus ein Widerspruch zu (3.10) folgt.

Wir können also wieder Folgerung 1.6 anwenden und erhalten $(y^*, s) \in (Y * \times \mathbb{R}) \setminus \{(0, 0)\}$ und ein $\lambda \in \mathbb{R}$ mit

$$\langle y^*, \tilde{y} \rangle_Y + st \leq \lambda, \quad \tilde{y} \in \operatorname{dom} F, t \geq F(\tilde{y}), \tag{3.11}$$
$$\langle y^*, A\tilde{x} \rangle_Y + s(F(Ax) + \langle \xi, \tilde{x} - x \rangle_X) \geq \lambda, \quad \tilde{x} \in X. \tag{3.12}$$

Wie zuvor können wir den Fall $s = 0$ durch Wahl $\tilde{x} = x_0$, $\tilde{y} = Ax_0$, und den Fall $s > 0$ durch Betrachtung von $t \to \infty$ für $\tilde{y}$ fest ausschließen. Also ist $s < 0$, und die Wahl von $\tilde{x} = x \in \operatorname{dom}(F \circ A)$, $\tilde{y} = Ax \in \operatorname{dom} F$ und $t = F(Ax)$ ergibt sofort $\lambda = \langle y^*, Ax\rangle_Y + sF(Ax)$.

Aus (3.11) folgt nun wegen $s < 0$ mit $t = F(\tilde{y})$, dass

$$F(\tilde{y}) - F(Ax) \geq \langle s^{-1}y^*, Ax - \tilde{y}\rangle_Y \quad \text{für alle } \tilde{y} \in \operatorname{dom} F,$$

d. h. $\eta := -s^{-1}y^* \in \partial F(Ax)$. Damit kann man (3.12) vereinfachen zu

$$\langle \xi - A^*\eta, \tilde{x} - x\rangle_X \leq 0 \qquad \text{für alle } \tilde{x} \in X.$$

Da an $\tilde{x}$ keine Einschränkungen gelten, kann die Ungleichung nur gelten, falls $\xi = A^*\eta$ ist, was zu zeigen war. □

Damit erhalten wir eine Charakterisierung von Minimierern konvexer Funktionen unter (konvexen) Nebenbedingungen.

Folgerung 3.13. *Sei $U \subset X$ nichtleer, konvex und abgeschlossen, und sei $F : X \to \overline{\mathbb{R}}$ eigentlich, konvex und unterhalbstetig. Existiert ein $x_0 \in U^o \cap \operatorname{dom} F$, so ist $\bar{x} \in U$ Lösung von*

$$\min_{x \in U} F(x)$$

genau dann, wenn ein $\xi \in X^*$ existiert mit

$$\begin{cases} -\xi \in \partial F(\bar{x}), \\ \langle \xi, \tilde{x} - \bar{x}\rangle \leq 0 \quad \textit{für alle } \tilde{x} \in U. \end{cases} \tag{3.13}$$

Beweis. Da F und U konvex sind, können wir Satz 3.7 auf $J := F + \delta_U$ anwenden; und da δ_U in $x_0 \in U^o$ stetig ist, können wir auch die Summenregel anwenden. Also hat F in $\bar{x}$ ein Minimum genau dann, wenn gilt

$$0 \in \partial J(\bar{x}) = \partial F(\bar{x}) + \partial \delta_U(\bar{x}).$$

Zusammen mit der Charakterisierung des Subdifferentials der Indikatorfunktion als Normalenkegel erhält man damit (3.13). □

Für eine Gâteaux-differenzierbare (und damit stetige) Funktion $F : X \to \mathbb{R}$ ergibt (3.13) die klassischen *Karush–Kuhn–Tucker-Bedingungen*; die Existenz des inneren Punktes $x_0 \in U^o$ entspricht dabei genau der *Slater-Bedingung*.

3.3 KONVEXE KONJUGIERTE UND DUALITÄT

Ein wesentliches Werkzeug in der konvexen Optimierung ist die *Dualität*: Jedem Optimierungsproblem kann ein *duales* Problem zugeordnet werden, und die gleichzeitige Betrachtung beider Probleme liefert zusätzliche Informationen. Die Verbindung wird über das folgende Konzept hergestellt.

Sei X ein normierter Raum und $F : X \to \overline{\mathbb{R}}$ eigentlich. Dann ist die *Fenchel-Konjugierte* zu F definiert als

$$F^* : X^* \to \overline{\mathbb{R}}, \qquad F^*(x^*) = \sup_{x\in X} \langle x^*, x\rangle_X - F(x).$$

(Da $\operatorname{dom} F \neq \emptyset$ angenommen ist, gilt $F^*(x^*) > -\infty$ für alle $x^* \in X^*$, also ist die Definition sinnvoll.) Anschaulich ist $F^*(x^*)$ der (negative) affine Anteil der Tangente an F (im Punkt x, in dem das Supremum angenommen wird) mit der Steigung x^*. Aus Lemma 3.3 (v) und Lemma 2.2 (v) folgt sofort, dass F^* für eigentliche F stets konvex und unterhalbstetig ist. Ist F nach unten durch ein affin-lineares Funktional beschränkt, so ist auch F^* eigentlich. Aus der Definition folgt außerdem sofort die *Fenchel–Young-Ungleichung*

$$\langle x^*, x\rangle_X \leq F(x) + F^*(x^*) \qquad \text{für alle } x \in X, x^* \in X^*. \tag{3.14}$$

Falls X nicht reflexiv ist, definieren wir analog für $F : X^* \to \mathbb{R}$ die *Fenchel-Präkonjugierte*

$$F_* : X \to \overline{\mathbb{R}}, \qquad F_*(x) = \sup_{x^*\in X^*} \langle x^*, x\rangle_X - F(x^*).$$

Durch diese Konvention ist die *Bikonjugierte*

$$F^{**} : X \to \mathbb{R}, \qquad F^{**}(x) = (F^*)_*(x),$$

selbst für nicht-reflexive Räume wieder auf X definiert (anstatt auf $X^{**} \supset X$); für reflexive X gilt $F_* = F^*$. Anschaulich ist F^{**} die untere konvexe Hülle von F, die für konvexe Funktionen ja mit F übereinstimmt. Um dies zu zeigen, benötigen wir das folgende Resultat, das wir hier nicht beweisen wollen.

Lemma 3.14. *Sei $F : X \to \overline{\mathbb{R}}$ eigentlich. Dann ist F genau dann konvex und unterhalbstetig, wenn gilt*

$$F = F^\Gamma := \sup \{G : X \to \mathbb{R} : G(x) = \langle x^*, x\rangle_X - \alpha,\ x^* \in X^*, \alpha \in \mathbb{R} \quad \textit{mit} \quad G \leq F\}.$$

Der etwas längliche Beweis beruht ähnlich wie der von Satz 3.11 auf dem Trennungssatz von Hahn–Banach und findet sich z. B. in [Schirotzek 2007, Theorem 2.1.2].

Satz 3.15 (Fenchel–Moreau–Rockafellar). *Sei $F : X \to \overline{\mathbb{R}}$ eigentlich. Dann gilt*

(i) $F^{**} \leq F$;

(ii) *$F^{**} = F$ genau dann, wenn F konvex und unterhalbstetig ist.*

Beweis. Für Aussage (i) nehmen wir in der Fenchel–Young-Ungleichung (3.14) das Supremum über alle $x^* \in X^*$ und erhalten

$$F(x) \geq \sup_{x^* \in X^*} \langle x^*, x \rangle_X - F^*(x^*) = F^{**}(x).$$

Für (ii) zeigen wir, dass $F^{**} = F^\Gamma$ gilt, woraus mit Lemma 3.14 die Behauptung folgt. Nach Definition der Fenchel-Konjugierten ist F^{**} konvex und unterhalbstetig, und wegen (i) ist F^{**} auch eigentlich. Also gilt nach Lemma 3.14

$$F^{**} = \sup \left\{ G : X \to \mathbb{R} : G(x) = \langle x^*, x \rangle_X - \alpha,\ x^* \in X^*, \alpha \in \mathbb{R} \quad \text{mit} \quad G \leq F^{**} \right\}.$$

Wir zeigen nun, dass wir in der Definition F^{**} durch F ersetzen können. Sei dafür $G(x) = \langle x^*, x \rangle_X - \alpha$ für $x^* \in X^*$ und $\alpha \in \mathbb{R}$ beliebig. Gilt $G \leq F^{**}$, so folgt aus (i) sofort $G \leq F$. Gilt umgekehrt $G \leq F$, so ist $\langle x^*, x \rangle_X - F(x) \leq \alpha$ für alle $x \in X$ und durch Supremum über alle $x \in X$ erhalten wir $\alpha \geq F^*(x^*)$. Daraus folgt nun

$$G(x) = \langle x^*, x \rangle_X - \alpha \leq \langle x^*, x \rangle_X - F^*(x^*) \leq F^{**}(x) \quad \text{für alle } x \in X,$$

d. h. $G \leq F^{**}$. □

Wir betrachten wieder relevante Beispiele.

Beispiel 3.16.

(i) Sei X ein Hilbertraum und $F(x) = \frac{1}{2}\|x\|_X^2$. Wir identifizieren X über den Satz von Fréchet–Riesz mit seinem Dualraum X^*, so dass die duale Paarung durch das Skalarprodukt ausgedrückt werden kann. Da F Fréchet-differenzierbar ist mit Gradient $\nabla F(x) = x$, muss die Lösung $\bar{x}$ von

$$\sup_{x \in X} (x^*, x)_X - \tfrac{1}{2}(x, x)_X$$

das Fermat-Prinzip erfüllen, d. h. $x^* = \bar{x}$. Einsetzen und vereinfachen liefert die Konjugierte

$$F^* : X \to \mathbb{R}, \qquad F^*(x^*) = \tfrac{1}{2}\|x^*\|_X^2.$$

(ii) Sei B_X die Einheitskugel im normierten Raum X und setze $F = \delta_{B_X}$. Dann ist für $x^* \in X^*$

$$(\delta_{B_X})^*(x^*) = \sup_{x \in X} \langle x^*, x \rangle_X - \delta_{B_X}(x) = \sup_{\|x\|_X \leq 1} \langle x^*, x \rangle_X = \|x^*\|_{X^*}.$$

Analog zeigt man unter Verwendung der Definition der Präkonjugierten im Dualraum und Folgerung 1.7, dass $(\delta_{B_{X^*}})_*(x) = \|x\|_X$ ist.

(iii) Sei X ein normierter Raum und setze $F(x) = \|x\|_X$. Für $x^* \in X^*$ unterscheiden wir nun zwei Fälle:

1. Fall: $\|x^*\|_{X^*} \leq 1$. Dann gilt mit (1.1) für alle $x \in X$ dass $\langle x^*, x\rangle_X - \|x\|_X \leq 0$ ist. Weiterhin ist $\langle x^*, 0\rangle = 0 = \|0\|_X$. Also gilt

$$F^*(x^*) = \sup_{x^* \in X^*} \langle x^*, x\rangle_X - \|x\|_X = 0.$$

2. Fall: $\|x^*\|_{X^*} > 1$. Nach Definition der Norm im Dualraum existiert dann ein $x_0 \in X$ mit $\langle x^*, x_0\rangle_X > \|x_0\|_X$. Lassen wir daher $t \to \infty$ gehen in

$$0 < t(\langle x^*, x_0\rangle_X - \|x_0\|_X) = \langle x^*, tx_0\rangle_X - \|tx_0\|_X \leq F^*(x^*),$$

so erhalten wir $F^*(x^*) = \infty$.

Zusammen ergibt dies $F^* = \delta_{B_{X^*}}$.

Wie oben zeigt man analog, dass $(\|\cdot\|_{X^*})_* = \delta_{B_X}$ ist.

Wir notieren noch einige nützliche Rechenregeln.

Lemma 3.17. *Sei $F : X \to \overline{\mathbb{R}}$ eigentlich. Dann ist*

(i) $(\alpha F)^* = \alpha F^* \circ (\alpha^{-1}\,\mathrm{Id})$ *für* $\alpha > 0$*;*

(ii) $(F(\cdot + x_0) + \langle x_0^*, \cdot\rangle_X)^* = F^*(\cdot - x_0^*) - \langle \cdot - x_0^*, x_0\rangle_X$ *für alle* $x_0 \in X$, $x_0^* \in X^*$*;*

(iii) $(F \circ A)^* = F^* \circ A^{-*}$ *für* $A \in L(Y, X)$ *stetig invertierbar und* $A^{-*} := (A^{-1})^*$.

Beweis. Die Regeln folgen direkt aus den Eigenschaften des Supremums.

Aussage (i) gilt wegen $\alpha > 0$ und

$$(\alpha F)^*(x^*) = \sup_{x\in X} \big(\alpha\langle \alpha^{-1}x^*, x\rangle_X - \alpha F(x)\big) = \alpha \sup_{x\in X} \big(\langle \alpha^{-1}x^*, x\rangle_X - F(x)\big) = \alpha F^*(\alpha^{-1}x^*).$$

Aussage (ii) gilt wegen $\{x + x_0 : x \in X\} = X$ und

$$\begin{aligned}
(F(\cdot + x_0) + \langle x_0^*, \cdot\rangle_X)^*(x^*) &= \sup_{x\in X} \langle x^*, x\rangle_X - F(x + x_0) - \langle x_0^*, x\rangle_X \\
&= \sup_{x\in X} \big(\langle x^* - x_0^*, x + x_0\rangle_X - F(x + x_0)\big) - \langle x^* - x_0^*, x_0\rangle_X \\
&= \sup_{\tilde{x} = x + x_0, x\in X} \big(\langle x^* - x_0^*, \tilde{x}\rangle_X - F(\tilde{x})\big) - \langle x^* - x_0^*, x_0\rangle_X \\
&= F^*(x^* - x_0^*) - \langle x^* - x_0^*, x_0\rangle_X.
\end{aligned}$$

Aussage (iii) gilt wegen $X = \operatorname{ran} A$ und

$$
\begin{aligned}
(F \circ A)^*(y^*) &= \sup_{y \in Y} \langle y^*, A^{-1}Ay\rangle_Y - F(Ay) \\
&= \sup_{x=Ay, y \in Y} \langle A^{-*}y^*, x\rangle_X - F(x) = F^*(A^{-*}y^*).
\end{aligned}
$$

□

Die Definition der Fenchel-Konjugierten ist auf besondere Weise verträglich mit der des Subdifferentials.

Lemma 3.18. *Sei $F : X \to \overline{\mathbb{R}}$ eigentlich, konvex und unterhalbstetig. Dann sind äquivalent für $x \in X$ und $x^* \in X^*$:*

(i) $\langle x^*, x\rangle_X = F(x) + F^*(x^*)$;

(ii) $x^* \in \partial F(x)$;

(iii) $x \in \partial F^*(x^*)$.

Beweis. Gilt (i), so folgt aus der Definition von F^* als Supremum, dass für alle $\tilde{x} \in X$ gilt

$$
\langle x^*, x\rangle_X - F(x) = F^*(x^*) \geq \langle x^*, \tilde{x}\rangle_X - F(\tilde{x}), \tag{3.15}
$$

was nach Definition äquivalent ist zu $x^* \in \partial F(x)$. Umgekehrt ergibt Supremum über alle $\tilde{x} \in X$ auf beiden Seiten von (3.15)

$$
\langle x^*, x\rangle_X \geq F(x) + F^*(x^*),
$$

und zusammen mit der Fenchel–Young-Ungleichung (3.14) folgt (i).

Analog erhält man aus (i) zusammen mit Satz 3.15 für alle $\tilde{x}^* \in X^*$ die Ungleichung

$$
\langle x^*, x\rangle_X - F^*(x^*) = F(x) = F^{**}(x) \geq \langle \tilde{x}^*, x\rangle_X - F^*(\tilde{x}^*),
$$

woraus wie oben die Äquivalenz von (i) und (iii) folgt. □

Bemerkung. Ist X nicht reflexiv, so ist dabei $x \in \partial F^*(x^*) \subset X^{**}$ über die kanonische Injektion aufzufassen, d. h. via

$$
\langle J(x), \tilde{x}^* - x^*\rangle_{X^*} = \langle \tilde{x}^* - x^*, x\rangle_X \leq F^*(\tilde{x}^*) - F^*(x^*) \quad \text{für alle } \tilde{x}^* \in X.
$$

Gleichheit in der Fenchel–Young-Ungleichung (oder äquivalent die Subdifferentialinklusion (ii)) garantiert also in (iii) die Existenz eines Subgradienten in $X \subset X^{**}$ und nicht nur in X^{**}; umgekehrt ist in (iii) die Existenz eines Subgradienten in $X \subset X^{**}$ notwendig für die Gleichheit (und damit (ii)). (Analoges gilt für $F : X^* \to \overline{\mathbb{R}}$ und $F_* : X \to \overline{\mathbb{R}}$.)

Lemma 3.18 spielt die Rolle des „Satzes von der konvexen Umkehrfunktion". Damit kann man insbesondere das Subdifferential einer komplizierten Norm durch das (einfachere) der konjugierten Indikatorfunktion ersetzen. Hat man zum Beispiel ein Problem der Form

$$\inf_{x\in X} F(x) + G(Ax) \tag{3.16}$$

für $F : X \to \overline{\mathbb{R}}$ und $G : Y \to \overline{\mathbb{R}}$ eigentlich, konvex, und unterhalbstetig sowie $A \in L(X,Y)$, so können wir G mit Hilfe von Satz 3.15 ersetzen durch die Definition von G^{**} und erhalten

$$\inf_{x\in X} \sup_{Y^*\in Y^*} F(x) + \langle y^*, Ax\rangle_Y - G^*(y^*). \tag{3.17}$$

Dürften wir nun inf und sup vertauschen, so könnten wir schreiben (mit $\inf F = -\sup(-F)$)

$$\begin{aligned}\inf_{x\in X} \sup_{y^*\in y^*} F(x) + \langle y^*, Ax\rangle_Y - G^*(y^*) &= \sup_{y^*\in Y^*} \inf_{x\in X} F(x) + \langle y^*, Ax\rangle_Y - G^*(y^*)\\ &= \sup_{y^*\in Y^*} -\left(\sup_{x\in X} -F(x) + \langle -A^*y^*, x\rangle_X\right) - G^*(y^*).\end{aligned}$$

Einsetzen der Definition von F^* ergibt dann das *duale Problem*

$$\sup_{y^*\in Y^*} -F^*(-A^*y^*) - G^*(y^*). \tag{3.18}$$

Als Nebeneffekt haben wir den Operator A zwischen den Funktionalen verschoben.

Der folgende Satz nutzt auf elegante Weise das Fermat-Prinzip, Summen- und Kettenregel sowie die Fenchel–Young-Gleichung, um hinreichende Bedingungen für die Vertauschbarkeit zu geben.

Satz 3.19 (Fenchel–Rockafellar). *Seien X, Y normierte Räume, $F : X \to \overline{\mathbb{R}}$ und $G : Y \to \overline{\mathbb{R}}$ eigentlich, konvex und unterhalbstetig, und $A \in L(X,Y)$. Gelte weiterhin:*

(i) das primale Problem (3.16) *hat eine Lösung $\bar x \in X$;*

(ii) es existiert ein $x_0 \in \operatorname{dom} F \cap \operatorname{dom}(G\circ A)$ so dass G stetig ist in Ax_0.

Dann hat das duale Problem (3.18) eine Lösung $\bar y^* \in Y^*$ und es gilt

$$\min_{x\in X} F(x) + G(Ax) = \max_{y^*\in Y^*} -F^*(-A^*y^*) - G^*(y^*). \tag{3.19}$$

Weiterhin sind $\bar x$ und $\bar y^*$ Lösungen von (3.16) bzw. (3.18) genau dann, wenn gilt

$$\begin{cases} -A^*\bar y^* \in \partial F(\bar x),\\ \quad\;\; \bar y^* \in \partial G(A\bar x).\end{cases} \tag{3.20}$$

Beweis. Nach Satz 3.7 ist $\bar{x} \in X$ genau dann eine Lösung von (3.16), wenn $0 \in \partial(F+G\circ A)(\bar{x})$ gilt. Wegen Voraussetzung (ii) sind Satz 3.11 und Satz 3.12 anwendbar; wir erhalten also

$$0 \in \partial(F + G \circ A)(\bar{x}) = \partial F(\bar{x}) + A^*\partial G(A\bar{x})$$

und damit die Existenz eines $\bar{y}^* \in \partial G(A\bar{x})$ mit $-A^*\bar{y}^* \in \partial F(\bar{x})$, d. h. (3.20) ist erfüllt. Aus (3.20) folgt nun mit Lemma 3.18 die Gleichheit in den Fenchel–Youngschen Ungleichungen für F und G, d. h.

$$\begin{cases} \langle -A^*\bar{y}^*, \bar{x}\rangle_X = F(\bar{x}) + F^*(-A^*\bar{y}^*), \\ \quad \langle \bar{y}^*, A\bar{x}\rangle_Y = G(A\bar{x}) + G^*(\bar{y}^*). \end{cases} \tag{3.21}$$

Durch Summieren beider Gleichungen erhalten wir

$$F(\bar{x}) + G(A\bar{x}) = -F^*(-A^*\bar{y}^*) - G^*(\bar{y}^*). \tag{3.22}$$

Es bleibt zu zeigen, dass $\bar{y}^*$ Lösung von (3.18) ist. Setze dafür

$$L : X \times Y^* \to \overline{\mathbb{R}}, \qquad L(x, y^*) = F(x) + \langle y^*, Ax\rangle_Y - G^*(y^*).$$

Dann gilt für alle $\tilde{x} \in X$ und $\tilde{y}^* \in Y^*$ stets

$$\sup_{y^*\in Y^*} L(\tilde{x}, y^*) \geq L(\tilde{x}, \tilde{y}^*) \geq \inf_{x\in X} L(x, \tilde{y}^*),$$

und damit (Infimum über alle $\tilde{x}$ in der ersten und Supremum über alle $\tilde{y}^*$ in der zweiten Ungleichung)

$$\inf_{x\in X} \sup_{y^*\in Y^*} L(x, y^*) \geq \sup_{y^*\in Y^*} \inf_{x\in X} L(x, y^*).$$

Also ist wegen $G = G^{**} = (G^*)_*$

$$\begin{aligned} F(\bar{x}) + G(A\bar{x}) &= \inf_{x\in X} \sup_{Y^*\in Y^*} F(x) + \langle y^*, Ax\rangle_Y - G^*(y^*) \\ &\geq \sup_{Y^*\in Y^*} \inf_{x\in X} F(x) + \langle y^*, Ax\rangle_Y - G^*(y^*) \\ &= \sup_{y^*\in Y^*} -F^*(-A^*y^*) - G^*(y^*). \end{aligned} \tag{3.23}$$

Zusammen mit der Gleichung (3.22) erhalten wir damit

$$-F^*(-A^*\bar{y}^*) - G(\bar{y}^*) = F(\bar{x}) + G(A\bar{x}) \geq \sup_{y^*\in Y^*} -F^*(-A^*y^*) - G^*(y^*),$$

d. h. $\bar{y}^*$ ist Lösung von (3.18). Damit folgt aus (3.22) auch (3.19).

Sind schließlich $\bar{x} \in X$ und $\bar{y}^* \in Y^*$ Lösungen von (3.16) bzw. (3.18), so folgt aus (3.19) auch (3.22). Zusammen mit der produktiven Null erhalten wir daraus

$$0 = [F(\bar{x}) + F^*(-A^*\bar{y}^*) - \langle -A^*\bar{y}^*, \bar{x}\rangle_X] + [G(A\bar{x}) + G^*(\bar{y}^*) - \langle \bar{y}^*, A\bar{x}\rangle_Y] .$$

Da beide Klammern wegen der Fenchel–Young-Ungleichung nicht-negativ sind, müssen sie einzeln verschwinden. Also gilt (3.21), und aus Lemma 3.18 folgt (3.20). □

Man nennt (3.20) auch *Fenchel-Extremalitätsbedingungen*; mit Hilfe von Lemma 3.18 können wir aus ihnen weitere äquivalente Optimalitätsbedingungen erzeugen, indem wir eine oder beide Subdifferentialinklusionen invertieren. Dies werden wir später nutzen, um praktisch durchführbare Algorithmen für die Lösung von Optimierungsproblemen dieser Form herzuleiten.

Teil II

ALGORITHMEN

4 PROXIMALPUNKT-VERFAHREN

Wir beschäftigen uns jetzt mit Verfahren zur numerischen Berechnung von Minimierern von Funktionalen $J : X \to \overline{\mathbb{R}}$ der Form

$$J(x) := F(x) + G(x)$$

für F, G konvex aber nicht differenzierbar. Die Schwierigkeit dabei ist, dass im Gegensatz zu differenzierbaren Funktionen das Äquivalent zum Gradientenabstieg, die Iteration

$$x^{k+1} \in x^k - \tau_k \partial J(x^k),$$

nicht wohldefiniert ist, da ein beliebiges Element im Subdifferential keine Abstiegsrichtung sein muss; dies kann man nur für den Subgradienten minimaler Norm garantieren – und für J erhält man diesen auch nicht aus den entsprechenden Subgradienten von F und G. Dies macht einen Subgradientenabstieg in der Regel nicht praktikabel. Wir ändern daher unseren Blickwinkel und suchen eine Nullstelle der mengenwertigen Abbildung $x \mapsto \partial J(x)$ durch eine geeignete Fixpunktiteration.

Um technischen Schwierigkeiten aus dem Weg zu gehen – und da wir später die Algorithmen auf diskretisierte Bilder, d. h. im $\mathbb{R}^N$, anwenden werden – beschränken wir uns in diesem und dem nächsten Kapitel auf den Fall, dass X ein Hilbertraum ist. Dies erlaubt, X^* mit X identifizieren; insbesondere identifizieren wir $\partial J(x) \subset X^*$ mit der Menge der entsprechenden Riesz-Repräsentanten in X. Damit können wir $\bar{x}$ mit $0 \in \partial J(\bar{x})$ durch Fixpunktiteration auf der äquivalenten Inklusion $\bar{x} \in \{\bar{x}\} + \partial J(\bar{x})$ suchen.

4.1 MONOTONE OPERATOREN

Zuerst betrachten wir die Klasse von Abbildungen der Form $x \mapsto \partial J(x)$. Für zwei normierte Räume X, Y betrachten wir eine *mengenwertige Abbildung* $A : X \to \mathcal{P}(Y)$, geschrieben auch $A : X \rightrightarrows Y$, und definieren weiter

- den *Definitionsbereich* $\operatorname{dom} A = \{x \in X : Ax \neq \emptyset\}$;
- das *Bild* $\operatorname{ran} A = \bigcup_{x \in X} Ax$;

- die *Inverse* $A^{-1} : Y \rightrightarrows X$ durch $A^{-1}(y) = \{x \in X : y \in Ax\}$ für alle $y \in Y$. (Beachten Sie, dass $A^{-1}(y) = \emptyset$ möglich ist.)

Für $A, B : X \rightrightarrows Y$, $C : Y \rightrightarrows Z$, und $\lambda \in \mathbb{R}$ definieren wir

- $\lambda A : X \rightrightarrows Y$ durch $(\lambda A)(x) = \{\lambda y : y \in Ax\}$;
- $A + B : X \rightrightarrows Y$ durch $(A + B)(x) = \{y + z : y \in Ax, z \in Bx\}$;
- $C \circ A : X \rightrightarrows Z$ durch $(C \circ A)(x) = \{z : \text{es gibt } y \in Y \text{ mit } y \in Ax, z \in Cy\}$.

Eine mengenwertige Abbildung $A : X \rightrightarrows X$ heißt *monoton*, falls gilt

$$(y_1 - y_2, x_1 - x_2)_X \geq 0 \quad \text{für alle } x_1, x_2 \in X, y_1 \in Ax_1, y_2 \in Ax_2.$$

Gilt Gleichheit nur für $x_1 = x_2$, so heißt A *streng monoton*.

Klarerweise ist die Identität $\mathrm{Id} : X \rightrightarrows X$, $x \mapsto \{x\}$, monoton. Für eine konvexe Funktion $F : X \to \overline{\mathbb{R}}$ ist auch der Operator $\partial F : X \rightrightarrows X$, $x \mapsto \partial F(x)$ monoton: Sind $x_1, x_2 \in X$ und $x_1^* \in \partial F(x_1)$ und $x_2^* \in \partial F(x_2)$, so gilt nach Definition

$$\begin{aligned} \left(x_1^*, \tilde{x} - x_1\right) &\leq F(\tilde{x}) - F(x_1) \qquad \text{für alle } \tilde{x} \in X, \\ \left(x_2^*, \tilde{x} - x_2\right) &\leq F(\tilde{x}) - F(x_2) \qquad \text{für alle } \tilde{x} \in X, \end{aligned}$$

woraus für $\tilde{x} = x_2$ in der ersten und $\tilde{x} = x_1$ in der zweiten Ungleichung und Addition die Monotonie folgt. Ebenso sind für A, B monoton und $\lambda \geq 0$ auch λA und $A + B$ monoton. Wir werden aber noch eine stärkere Eigenschaft benötigen, die die Abgeschlossenheit von A garantiert.

Definition 4.1. Ein monotoner Operator $A : X \rightrightarrows X$ ist maximal monoton genau dann, wenn für $x \in X$ und $x^* \in X$ die Bedingung

$$(x^* - \tilde{x}^*, x - \tilde{x})_X \geq 0 \qquad \text{für alle } \tilde{x} \in X, \tilde{x}^* \in A\tilde{x} \tag{4.1}$$

impliziert, dass $x^* \in Ax$ ist.

Für feste $x \in X$ und $x^* \in X$ besagt die Bedingung (4.1) gerade, dass die Erweiterung $\tilde{A} \supset A$, definiert durch

$$\tilde{A}\tilde{x} := \begin{cases} Ax \cup \{x^*\} & \tilde{x} = x, \\ A\tilde{x} & \tilde{x} \neq x, \end{cases}$$

monoton ist. Die Behauptung ist dann genau, dass dies keine echte Erweiterung darstellt. Beispielsweise ist $A : \mathbb{R} \rightrightarrows \mathbb{R}$ definiert durch

$$At = \begin{cases} 1 & t \geq 0, \\ -1 & t < 0, \end{cases}$$

monoton aber nicht maximal monoton, da A eine Teilmenge des durch $\tilde{A}t = \mathrm{sign}(t) = \partial(|\cdot|)(t)$ definierten monotonen Operators ist.

Aus der Definition folgen sofort einige nützliche Aussagen.

Lemma 4.2. *Ist $A : X \rightrightarrows X$ maximal monoton, dann ist es auch λA für alle $\lambda > 0$.*

Beweis. Seien $x, x^* \in X$ und

$$0 \le (x^* - \tilde{x}^*, x - \tilde{x})_X = \lambda \left(\lambda^{-1}x^* - \lambda^{-1}\tilde{x}^*, x - \tilde{x}\right)_X \qquad \text{für alle } \tilde{x} \in X, \tilde{x}^* \in \lambda A x.$$

Da für $\tilde{x}^* \in \lambda A x$ gilt $\lambda^{-1}\tilde{x}^* \in Ax$ und A maximal monoton ist, folgt daraus $\lambda^{-1}\bar{x}^* \in A\bar{x}$, d. h. $\bar{x}^* \in \lambda A\bar{x}$. Damit ist λA maximal monoton. □

Lemma 4.3. *Ist $A : X \rightrightarrows X$ maximal monoton, dann ist A schwach–stark abgeschlossen, d. h. aus $x_n \rightharpoonup x$ und $Ax_n \ni x_n^* \to x^*$ folgt $x^* \in Ax$.*

Beweis. Für $\tilde{x} \in X$ und $\tilde{x}^* \in A\tilde{x}$ beliebig gilt

$$0 \le \left(x_n^* - \tilde{x}^*, x_n - \tilde{x}\right)_X \to (x^* - \tilde{x}^*, x - \tilde{x})_X$$

da x_n schwach und x_n^* stark in X konvergieren. Da A maximal monoton ist, folgt daraus $x^* \in Ax$. □

Von zentraler Bedeutung für die Theorie der monotonen Operatoren ist der *Satz von Minty*.

Satz 4.4 (Minty). *Sei $A : X \rightrightarrows X$ monoton. Dann ist A maximal monoton genau dann, wenn* $\operatorname{ran}(\operatorname{Id} + A) = X$.

Beweis. Wir zeigen nur, dass diese Bedingung hinreichend ist; der Beweis der Notwendigkeit ist deutlich aufwendiger,[1] weshalb hierfür auf die Literatur verwiesen wird; siehe z. B. [Bauschke & Combettes 2017, Theorem 21.1].

Sei $\operatorname{Id} + A$ surjektiv, und seien $x \in X$ und $x^* \in X$ mit

$$(x^* - \tilde{x}^*, x - \tilde{x})_X \ge 0 \qquad \text{für alle } \tilde{x} \in X, \tilde{x}^* \in Ax. \tag{4.2}$$

Nach Annahme existiert nun für $x + x^* \in X$ ein $z \in X$ und ein $z^* \in Az$ mit

$$x + x^* = z + z^* \in (\operatorname{Id} + A)z. \tag{4.3}$$

Einsetzen von $(\tilde{x}, \tilde{x}^*) = (z, z^*)$ in (4.2) ergibt dann

$$0 \le (x^* - z^*, x - z)_X = (z - x, x - z)_X = -\|x - z\|_X^2 \le 0,$$

d. h. $x = z$. Aus (4.3) folgt dann auch $x^* = z^* \in Az = Ax$, und damit ist A maximal monoton. □

[1] Er verläuft analog zu dem Beweis von Folgerung 4.5 über die Konstruktion eines Funktionals F_A – der *Fitzpatrick-Funktion* – welches für A die selbe Rolle spielt wie F für ∂F.

Damit lässt sich nun zeigen, dass Subdifferentiale maximal monoton sind.

Folgerung 4.5. *Für $F : X \to \overline{\mathbb{R}}$ eigentlich, konvex und unterhalbstetig ist $\partial F : X \rightrightarrows X$ maximal monoton.*

Beweis. Da ∂F monoton ist, genügt es zu zeigen, dass $\operatorname{Id} + \partial F$ surjektiv ist. Für gegebenes $z \in X$ betrachten wir das Funktional $J : X \to \overline{\mathbb{R}}$,

$$J(x) = \frac{1}{2}\|x - z\|_X^2 + F(x).$$

Da sowohl die quadrierte Norm als auch F eigentlich, konvex und unterhalbstetig sind, ist J eigentlich, konvex und unterhalbstetig. Weiter ist J koerziv, da F für $\|x\|_X \to \infty$ schlimmstenfalls linear fällt. Nach Satz 3.4 existiert also ein $\bar{x} \in X$ mit $J(\bar{x}) = \min_{x \in X} J(x)$, für welches nach Sätze 3.7 und 3.11 gilt

$$0 \in \partial J(\bar{x}) = \{\bar{x} - z\} + \partial F(\bar{x}),$$

d. h. $z \in \{\bar{x}\} + \partial F(\bar{x})$. □

4.2 RESOLVENTEN UND PROXIMALPUNKTE

Wie wir im Beweis von Folgerung 4.5 gesehen haben, ist $\operatorname{Id} + \partial F$ surjektiv; da die quadrierte Norm im Hilbertraum und damit J strikt konvex ist, ist der Minimierer und damit das Urbild sogar eindeutig. Also ist $(\operatorname{Id} + \partial F)^{-1}$ einwertig, auch wenn ∂F dies nicht ist. Es besteht daher die Hoffnung, dieses Urbild anstelle von Subgradienten – oder allgemeiner, Elementen aus dem Bild eines monotonen Operators – für Algorithmen nutzen zu können.

Wir definieren daher für $A : X \rightrightarrows X$ maximal monoton die *Resolvente*

$$\mathcal{R}_A : X \rightrightarrows X, \qquad \mathcal{R}_A(x) = (\operatorname{Id} + A)^{-1} x,$$

und für $F : X \to \overline{\mathbb{R}}$ eigentlich, konvex und unterhalbstetig die *Proximalpunktabbildung*

$$\operatorname{prox}_F : X \to X, \qquad \operatorname{prox}_F(x) = \arg\min_{w \in X} \frac{1}{2}\|w - x\|_X^2 + F(w).$$

Da wie im Beweis von Folgerung 4.5 $w = \mathcal{R}_{\partial F}(x)$ die notwendige und hinreichende Bedingung dafür ist, dass w der eindeutige Minimierer (man sagt auch *Proximalpunkt*) dieser strikt konvexen Funktion ist, gilt

$$\operatorname{prox}_F = (\operatorname{Id} + \partial F)^{-1} = \mathcal{R}_{\partial F}.$$

Wir müssen noch zeigen, dass auch die Resolvente von beliebigen maximal monotonen Operators einwertig ist, d. h. die Schreibweise $\mathcal{R}_A : X \to X$ gerechtfertigt ist.

Lemma 4.6. *Sei $A : X \rightrightarrows X$ maximal monoton. Dann ist $\mathcal{R}_A$ auf ganz X definiert und einwertig.*

Beweis. Da A maximal monoton ist, ist $\operatorname{Id} + A$ surjektiv, woraus $\operatorname{dom} \mathcal{R}_A = X$ folgt. Seien nun $x^*, z^* \in X$ beliebig und $x \in \mathcal{R}_A(x^*)$ und $z \in \mathcal{R}_A(z^*)$, d. h. es gilt $x^* \in \{x\} + Ax$ und $z^* \in \{z\} + Az$. Aus $x^* - x \in Ax$ und $z^* - z \in Az$ folgt mit der Monotonie von A, dass

$$\|x - z\|_X^2 \leq (x^* - z^*, x - z)_X$$

gilt. Ist $x^* = z^*$, so muss daher auch $x = z$ sein. Also ist $\mathcal{R}_A$ einwertig. □

Tatsächlich lassen sich Minimierer konvexer Funktionen auch als Proximalpunkte charakterisieren, indem man den folgende nützlichen Zusammenhang verwendet.

Lemma 4.7. *Sei $F : X \to \overline{\mathbb{R}}$ eigentlich, konvex und unterhalbstetig, und seien $x, x^* \in X$. Dann ist für beliebige $\gamma > 0$*

$$x^* \in \partial F(x) \quad \Leftrightarrow \quad x = \operatorname{prox}_{\gamma F}(x + \gamma x^*).$$

Beweis. Mit Hilfe der Rechenregeln für mengenwertige Abbildungen sieht man, dass

$$\begin{aligned} x^* \in \partial F(x) &\Leftrightarrow x + \gamma x^* \in (\operatorname{Id} + \gamma \partial F)(x) \\ &\Leftrightarrow x \in (\operatorname{Id} + \gamma \partial F)^{-1}(x + \gamma x^*) \\ &\Leftrightarrow x = \operatorname{prox}_{\gamma F}(x + \gamma x^*). \end{aligned}$$

Im letzten Schritt haben wir dabei $\operatorname{prox}_{\gamma F} = \mathcal{R}_{\partial(\gamma F)} = \mathcal{R}_{\gamma \partial F}$ nach Lemma 3.10 (i) und die Einwertigkeit der Proximalpunktabbildung verwendet. □

Wendet man diese Umformulierung auf das Fermat-Prinzip $0 \in \partial F(\bar{x})$ an, erhält man sofort das folgende Resultat.

Folgerung 4.8. *Sei $F : X \to \overline{\mathbb{R}}$ eigentlich, konvex und unterhalbstetig und $\gamma > 0$ beliebig. Dann ist $\bar{x} \in \operatorname{dom} F$ ein Minimierer von F genau dann, wenn gilt*

$$\bar{x} = \operatorname{prox}_{\gamma F}(\bar{x}).$$

Damit liefert uns die Proximalpunkt-Abbildung eine Möglichkeit, Minimierer konvexer Funktionen statt durch (explizite) Mengeninklusion durch eine (implizite) Gleichung zu charakterisieren.

Weiter erhält man daraus eine Verallgemeinerung der orthogonalen Zerlegung von Vektorräumen.

Satz 4.9 (Moreau-Zerlegung). *Sei $F : X \to \overline{\mathbb{R}}$ eigentlich, konvex und unterhalbstetig. Dann gilt für alle $x \in X$*

$$x = \operatorname{prox}_F(x) + \operatorname{prox}_{F^*}(x).$$

Beweis. Setze $w = \operatorname{prox}_F(x)$. Dann folgt aus Lemmata 3.18 und 4.7 für $\gamma = 1$, dass gilt

$$\begin{aligned} w = \operatorname{prox}_F(x) = \operatorname{prox}_F(w + (x - w)) &\Leftrightarrow x - w \in \partial F(w) \\ &\Leftrightarrow w \in \partial F^*(x - w) \\ &\Leftrightarrow x - w = \operatorname{prox}_{F^*}((x - w) + w) = \operatorname{prox}_{F^*}(x). \quad \square \end{aligned}$$

Folgende Rechenregeln werden nützlich sein.

Lemma 4.10. *Sei $F : X \to \overline{\mathbb{R}}$ eigentlich, konvex und unterhalbstetig.*

(i) Für $\lambda \neq 0$ und $z \in X$ gilt mit $H(x) = F(\lambda x + z)$

$$\operatorname{prox}_H(x) = \lambda^{-1}(\operatorname{prox}_{\lambda^2 F}(\lambda x + z) - z).$$

(ii) Für $\gamma > 0$ gilt

$$\operatorname{prox}_{\gamma F^*}(x) = x - \gamma \operatorname{prox}_{\gamma^{-1} F}(\gamma^{-1} x).$$

(iii) Für $G : Y \to \overline{\mathbb{R}}$ gilt mit $H : X \times Y \to \overline{\mathbb{R}}$, $H(x, y) = F(x) + G(y)$

$$\operatorname{prox}_{\gamma H}(x, y) = \begin{pmatrix} \operatorname{prox}_{\gamma F}(x) \\ \operatorname{prox}_{\gamma G}(y) \end{pmatrix}.$$

Beweis. Zu (i): Nach Definition ist

$$\bar{w} := \operatorname{prox}_H(x) = \arg\min_{w \in X} \frac{1}{2}\|w - x\|_X^2 + F(\lambda w + z).$$

Da X ein Vektorraum ist, gilt weiter

$$\min_{w \in X} \frac{1}{2}\|w - x\|_X^2 + F(\lambda w + z) = \min_{v \in X} \frac{1}{2}\|\lambda^{-1}(v - z) - x\|_X^2 + F(v),$$

und die entsprechenden Minimierer erfüllen $\bar{v} = \lambda \bar{w} + z$. Die Aussage folgt also aus

$$\begin{aligned} \bar{v} &= \arg\min_{v \in X} \frac{1}{2}\|\lambda^{-1}(v - z) - x\|_X^2 + F(v) \\ &= \arg\min_{v \in X} \frac{1}{2\lambda^2}\|v - (\lambda x + z)\|_X^2 + F(v) \\ &= \arg\min_{v \in X} \frac{1}{2}\|v - (\lambda x + z)\|_X^2 + \lambda^2 F(v) \\ &= \operatorname{prox}_{\lambda^2 F}(\lambda x + z). \end{aligned}$$

Zu (ii): Aus der Moreau-Zerlegung, den Rechenregeln für die Fenchel-Konjugierte, und (i) für $\lambda = \gamma^{-1}$, $z = 0$, folgt

$$\begin{aligned}\operatorname{prox}_{\gamma F}(x) &= x - \operatorname{prox}_{(\gamma F)^*}(x)\\ &= x - \operatorname{prox}_{\gamma F^* \circ (\gamma^{-1} \operatorname{Id})}(x)\\ &= x - \gamma \operatorname{prox}_{\gamma(\gamma^{-2} F^*)}(\gamma^{-1} x).\end{aligned}$$

Anwenden auf F^* unter Verwendung von $F^{**} = F$ nach Satz 3.15 liefert die Aussage.

Zu (iii): Nach Definition der Norm auf $X \times Y$ gilt

$$\begin{aligned}\operatorname{prox}_{\gamma H}(x, y) &= \arg \min_{(u,v) \in X \times Y} \frac{1}{2} \|(u, v) - (x, y)\|_{X \times Y}^2 + \gamma H(u, v)\\ &= \arg \min_{u \in X, v \in Y} \left(\frac{1}{2}\|u - x\|_X^2 + \gamma F(u)\right) + \left(\frac{1}{2}\|v - y\|_Y^2 + \gamma G(v)\right).\end{aligned}$$

Da keine gemischten Terme in u und v auftreten, können die Klammern separat minimiert werden. Also ist $\operatorname{prox}_{\gamma H}(x, y) = (\bar{u}, \bar{v})$ mit

$$\begin{aligned}\bar{u} &= \arg \min_{u \in X} \frac{1}{2}\|u - x\|_X^2 + \gamma F(u) = \operatorname{prox}_{\gamma F(x)},\\ \bar{v} &= \arg \min_{v \in Y} \frac{1}{2}\|v - y\|_Y^2 + \gamma G(v) = \operatorname{prox}_{\gamma G(x)}.\end{aligned}$$

□

Das Ausrechnen von Proximalpunkten ist in der Regel schwierig, denn die Auswertung von prox_F enthält ja bereits die Minimierung von F. In manchen Fällen kann man aber eine geschlossene Formel angeben.

Beispiel 4.11. Wir betrachten zuerst $X = \mathbb{R}$.

(i) $F(t) = \frac{1}{2}|t|^2$, d. h. $\partial F(t) = \{t\}$ und damit ist $\operatorname{prox}_{\gamma F}(t) = (1 + \gamma)^{-1} t$.

(ii) $F(t) = |t|$, d. h. $\partial F(t) = \operatorname{sign}(t)$; siehe (3.5). Also ist $s = \operatorname{prox}_{\gamma F}(t) = (\operatorname{Id} + \gamma \operatorname{sign})^{-1}(t)$ genau dann, wenn $t \in \{s\} + \gamma \operatorname{sign}(s)$ ist. Angenommen, für gegebenes t gelte diese Inklusion für ein $\bar{s}$. Wir machen nun eine Fallunterscheidung.

1. Fall $\bar{s} = 0$. Dies impliziert $t \in \gamma[-1, 1] = [-\gamma, \gamma]$.

2. Fall $\bar{s} > 0$. Dies impliziert $t = \bar{s} + \gamma$, d. h. $\bar{s} = t - \gamma$ und damit $t > \gamma$.

3. Fall $\bar{s} < 0$. Dies impliziert $t = \bar{s} - \gamma$, d. h. $\bar{s} = t + \gamma$ und damit $t < -\gamma$.

Da dies auch eine vollständige und disjunkte Fallunterscheidung für $t \in \mathbb{R}$ ist, erhalten wir

$$\operatorname{prox}_{\gamma F}(t) = \begin{cases} t - \gamma & t > \gamma,\\ 0 & t \in [-\gamma, \gamma],\\ t + \gamma & t < -\gamma.\end{cases}$$

Diese Abbildung wird auch als *soft-shrinkage*-Operator bezeichnet.

(iii) $F(t) = \delta_{[-1,1]}(t)$. Nach Beispiel 3.16 (iii) gilt $F^*(t) = |t|$, und daraus erhalten wir mit Lemma 4.10 (ii)

$$\begin{aligned}
\operatorname{prox}_{\gamma F}(t) &= t - \gamma \operatorname{prox}_{\gamma^{-1}F^*}(\gamma^{-1}t) \\
&= \begin{cases} t - \gamma(\gamma^{-1}t - \gamma^{-1}) & \gamma^{-1}t > \gamma^{-1}, \\ t - 0 & \gamma^{-1}t \in [-\gamma^{-1}, \gamma^{-1}], \\ t - \gamma(\gamma^{-1}t + \gamma^{-1}) & \gamma^{-1}t < -\gamma^{-1} \end{cases} \\
&= \begin{cases} 1 & t > 1, \\ t & t \in [-1, 1], \\ -1 & t < -1. \end{cases}
\end{aligned}$$

Die Proximalpunktabbildung ist also – unabhängig von γ – die Projektion von t auf $[-1, 1]$.

Beispiel 4.12. Beispiel 4.11 kann auf $X = \mathbb{R}^N$ (versehen mit dem Euklidischen Skalarprodukt) verallgemeinert werden, indem man N mal Lemma 4.10 (iii) anwendet. Man erhält dann komponentenweise

(i) für $F(x) = \frac{1}{2}\|x\|_2^2$

$$[\operatorname{prox}_{\gamma F}(x)]_i = \frac{1}{1+\gamma} x_i, \quad 1 \le i \le N;$$

(ii) für $F(x) = \|x\|_1$

$$[\operatorname{prox}_{\gamma F}(x)]_i = (|x_i| - \gamma)^+ \operatorname{sign}(x_i), \quad 1 \le i \le N;$$

(iii) für $F(x) = \delta_{B_\infty}(x)$

$$\begin{aligned}
[\operatorname{prox}_{\gamma F}(x)]_i &= x_i - (x_i - 1)^+ - (x_i + 1)^- \\
&= \frac{x_i}{\max\{1, |x_i|\}}, \qquad 1 \le i \le N
\end{aligned}$$

mit der Schreibweise $(t)^+ = \max\{t, 0\}$ und $(t)^- := \min\{t, 0\}$.

Viele weitere Beispiele findet man in [Parikh & Boyd 2014, § 6.5]. Allgemein gilt in einem Hilbertraum X:

(i) Ist $F = \frac{1}{2}\|\cdot\|_X^2 = \frac{1}{2}(\cdot, \cdot)_X$, so ist $\operatorname{prox}_{\gamma F}(x) = (1+\gamma)^{-1}x$.

(ii) Ist $F = \|\cdot\|_X$, so zeigt man analog mit Fallunterscheidung nach Satz 3.9

$$\operatorname{prox}_{\gamma F}(x) = \left(1 - \frac{\gamma}{\|x\|_X}\right)^+ x.$$

(iii) Ist $F = \delta_C$ für $C \subset X$ nichtleer, konvex und abgeschlossen, so ist

$$\operatorname{prox}_{\gamma F}(x) = \operatorname{proj}_C(x),$$

die *Projektion* von x auf C; die Proximalpunktabbildung verallgemeinert also die metrische Projektion auf (konvexe) Mengen. Projektionsformeln bzw. -verfahren für verschiedene Klassen von Mengen sind in [Cegielski 2012, Kapitel 4.1] aufgeführt.

4.3 PROXIMALPUNKT-VERFAHREN

Wir haben gesehen, dass ein Minimierer $\bar{x}$ von $F : X \to \overline{\mathbb{R}}$ charakterisiert werden kann als Fixpunkt von $\operatorname{prox}_{\gamma F}$ für beliebiges $\gamma > 0$. Dies legt nahe, diesen durch eine Fixpunktiteration zu berechnen: Wähle $x^0 \in X$ und setze für eine geeignete Folge $\{\gamma_k\}_{k \in \mathbb{N}}$

$$x^{k+1} = \operatorname{prox}_{\gamma_k F}(x^k). \tag{4.4}$$

Wie bei der klassischen Fixpunktiteration müssen wir nun nachweisen, dass die Fixpunktabbildung in einem passenden Sinne kontrahierend wirkt. Eine Abbildung $T : X \to X$ heißt *nichtexpansiv*, wenn für alle $x, z \in X$ gilt $\|Tx - Tz\|_X \leq \|x - z\|_X$. Sie heißt *fest nichtexpansiv* (englisch: „firmly nonexpansive"), wenn sogar gilt

$$\|Tx - Tz\|_X^2 \leq (Tx - Tz, x - z)_X \qquad \text{für alle } x, z \in X.$$

Resolventen maximal monotoner Operatoren sind stets fest nichtexpansiv.

Lemma 4.13. *Sei $A : X \rightrightarrows X$ maximal monoton. Dann ist $\mathcal{R}_A : X \to X$ fest nichtexpansiv, und es gilt*

$$\|\mathcal{R}_A x - \mathcal{R}_A z\|_X^2 + \|(\operatorname{Id} - \mathcal{R}_A)x - (\operatorname{Id} - \mathcal{R}_A)z\|_X^2 \leq \|x - z\|_X^2 \quad \textit{für alle } x, z \in X.$$

Beweis. Wie im Beweis von Lemma 4.6 erhalten wir aus der Monotonie von A, dass für $x^* = \mathcal{R}_A(x)$ und $z^* = \mathcal{R}_A(z)$ gilt

$$\|x^* - z^*\|_X^2 \leq (x^* - z^*, x - z)_X,$$

also $\mathcal{R}_A$ fest nichtexpansiv ist. Daraus folgt nun, dass gilt

$$\begin{aligned} \|(\operatorname{Id} - \mathcal{R}_A)x - (\operatorname{Id} - \mathcal{R}_A)z\|_X^2 &= \|x - z\|_X^2 - 2\,(x - z, \mathcal{R}_A x - \mathcal{R}_A z)_X + \|\mathcal{R}_A x - \mathcal{R}_A z\|_X^2 \\ &\leq \|x - z\|_X^2 - \|\mathcal{R}_A x - \mathcal{R}_A z\|_X^2 \end{aligned}$$ □

Damit können wir nun die Konvergenz der Fixpunktiteration für Resolventen maximal monotoner Operatoren zeigen, woraus sofort die Konvergenz des Proximalpunktverfahrens (4.4) folgt.

Satz 4.14. *Sei $A : X \rightrightarrows X$ maximal monoton mit Nullstelle $x^* \in A^{-1}(0)$, und sei $\{\gamma_k\}_{k\in\mathbb{N}} \subset (0,\infty)$ mit $\sum_{k=0}^\infty \gamma_k^2 = \infty$. Erfüllt $\{x^k\}_{k\in\mathbb{N}} \subset X$ die Rekursion*

$$x^{k+1} = \mathcal{R}_{\gamma_k A} x^k,$$

dann konvergiert $x^k \rightharpoonup \bar{x}$ mit $0 \in A\bar{x}$.

Beweis. Aus der Rekursion $x^{k+1} = \mathcal{R}_{\gamma_k A} x^k = (\operatorname{Id} + \gamma_k A)^{-1} x^k$ folgt

$$w^k := \gamma_k^{-1}(x^k - x^{k+1}) \in Ax^{k+1}$$

und damit auch $x^{k+1} - x^{k+2} = \gamma_{k+1} w^{k+1}$. (Der Vektor w^k spielt die Rolle des Residuums in der verallgemeinerten Gleichung $0 \in Ax$.) Da A monoton ist, gilt für $\gamma_{k+1} > 0$

$$\begin{aligned} 0 &\leq \gamma_{k+1}^{-1} \left(w^k - w^{k+1}, x^{k+1} - x^{k+2} \right)_X \\ &= \left(w^k - w^{k+1}, w^{k+1} \right)_X \\ &= \left(w^k, w^{k+1} \right)_X - \|w^{k+1}\|_X^2 \\ &\leq \|w^{k+1}\|_X \left(\|w^k\|_X - \|w^{k+1}\|_X \right). \end{aligned}$$

Also ist die Folge $\{\|w^k\|_X\}_{k\in\mathbb{N}}$ der Residuumsnormen nicht-negativ und monoton fallend und daher konvergent. (Solange $w^{k+1} \neq 0$, aber in diesem Fall ist wegen $w^{k+1} \in Ax^{k+2}$ bereits x^{k+2} die gesuchte Nullstelle.)

Sei nun $x^* \in X$ eine Nullstelle von A, d. h. $0 \in Ax^*$. Dann ist $x^* = \mathcal{R}_{\gamma A} x^*$ für alle $\gamma > 0$. Aus Lemma 4.13 folgt

$$\begin{aligned} \|x^{k+1} - x^*\|_X^2 &= \|\mathcal{R}_{\gamma_k A} x^k - \mathcal{R}_{\gamma_k A} x^*\|_X^2 \\ &\leq \|x^k - x^*\|_X^2 - \|(\operatorname{Id} - \mathcal{R}_{\gamma_k A})x^k - (\operatorname{Id} - \mathcal{R}_{\gamma_k A})x^*\|_X^2 \\ &= \|x^k - x^*\|_X^2 - \gamma_k^2 \|w^k\|_X^2. \end{aligned} \tag{4.5}$$

Also ist $\{\|x^k - x^*\|_X\}_{k\in\mathbb{N}}$ für jede Nullstelle x^* monoton fallend (man nennt solche Folgen *Féjer-monoton*) und damit beschränkt. Also ist auch $\{x^k\}_{k\in\mathbb{N}}$ beschränkt und hat daher eine schwach konvergente Teilfolge $x^{k_l} \rightharpoonup \bar{x}$. Aus (4.5) folgt außerdem

$$\|x^{k+1} - x^*\|_X^2 \leq \|x^0 - x^*\|_X^2 - \sum_{j=0}^{k} \gamma_j^2 \|w^j\|_X^2. \tag{4.6}$$

Die Partialsummenfolge ist also monoton wachsend und beschränkt, und deshalb konvergiert $\sum_{k=0}^\infty \gamma_k^2 \|w^k\|_X^2$. Da die Folge $\{\gamma_k^2\}_{k\in\mathbb{N}}$ nach Annahme nicht summierbar ist, muss $\liminf_{k\to\infty} \|w^k\|_X^2 = 0$ gelten. Aus der Konvergenz von $\{\|w^k\|_X\}_{k\in\mathbb{N}}$ folgt damit $w^k \to 0$.

Lemma 4.3 angewendet auf $\{x^{k_l}\}_{l\in\mathbb{N}}$ mit $x^{k_l} \rightharpoonup \bar{x}$ und $\{w^{k_l}\}_{l\in\mathbb{N}}$ mit $w^{k_l} \to 0$, ergibt dann $0 \in A\bar{x}$, d. h. jeder schwache Häufungspunkt von $\{x^k\}_{k\in\mathbb{N}}$ ist eine Nullstelle von A.

Wir zeigen nun, dass die gesamte Folge $\{x^k\}_{k\in\mathbb{N}}$ konvergiert. Seien $\bar{x}$ und $\hat{x}$ schwache Häufungspunkte und damit Nullstellen von A. Dann sind wegen der Féjer-Monotonie von $\{x^k\}_{k\in\mathbb{N}}$ sowohl $\{\|x^k - \bar{x}\|_X\}_{k\in\mathbb{N}}$ als auch $\{\|x^k - \hat{x}\|_X\}_{k\in\mathbb{N}}$ monoton fallend und beschränkt und daher konvergent. Also konvergiert auch

$$\left(x^k, \bar{x} - \hat{x}\right)_X = \frac{1}{2}\left(\|x^k - \hat{x}\|_X^2 - \|x^k - \bar{x}\|_X^2 + \|\bar{x}\|_X^2 - \|\hat{x}\|_X^2\right)$$

gegen ein $c \in \mathbb{R}$. Da $\bar{x}$ schwacher Häufungspunkt ist, existiert eine Teilfolge $\{x^{k_n}\}_{n\in\mathbb{N}}$ mit $x^{k_n} \rightharpoonup \bar{x}$; ebenso existiert eine Teilfolge $\{x^{k_m}\}_{m\in\mathbb{N}}$ mit $x^{k_m} \rightharpoonup \hat{x}$. Also ist

$$(\bar{x}, \bar{x} - \hat{x})_X = \lim_{n\to\infty}\left(x^{k_n}, \bar{x} - \hat{x}\right)_X = c = \lim_{m\to\infty}\left(x^{k_m}, \bar{x} - \hat{x}\right)_X = (\hat{x}, \bar{x} - \hat{x})_X\,,$$

und damit

$$0 = \|\bar{x} - \hat{x}\|_X^2,$$

d. h. $\bar{x} = \hat{x}$. Jede Teilfolge hat daher den selben Grenzwert, und deshalb konvergiert die gesamte Folge $\{x^k\}_{k\in\mathbb{N}}$ gegen diesen. □

5 SPLITTING-VERFAHREN

Für Funktionale aus der Bildverarbeitung, die üblicherweise die Form $J(x) = F(x) + G(x)$ haben, ist das Proximalpunkt-Verfahren in der Regel nicht praktikabel, da die Auswertung prox_J nicht wesentlich einfacher ist als das ursprüngliche Problem. Anstatt die Proximalpunkt-Formulierung auf $0 \in \partial J(\bar{x})$ anzuwenden, verwendet man zuerst die Summenregel und erhält dadurch ein $\bar{p} \in X$ mit

$$\begin{cases} \bar{p} \in \partial F(\bar{x}), \\ -\bar{p} \in \partial G(\bar{x}). \end{cases} \tag{5.1}$$

Man kann nun eine oder beide der Subdifferentialinklusionen durch eine Proximalpunkt-Abbildung ersetzen.

5.1 EXPLIZITES SPLITTING

In expliziten Splitting-Verfahren – auch *forward–backward splitting* genannt – wendet man Lemma 4.7 nur auf die zweite Inklusion in (5.1) an und erhält

$$\begin{cases} \bar{p} \in \partial F(\bar{x}), \\ \bar{x} = \operatorname{prox}_{\gamma G}(\bar{x} - \gamma \bar{p}). \end{cases} \tag{5.2}$$

Die zugehörige Fixpunkt-Iteration ist dann

(i) Wähle $p^k \in \partial F(x^k)$ (mit minimaler Norm).

(ii) Setze $x^{k+1} = \operatorname{prox}_{\gamma_k G}(x^k - \gamma_k p^k)$.

Dies ist im Allgemeinen kein praktikables Verfahren, da die Bestimmung eines Subgradienten minimaler Norm aufwändig ist. Eine Ausnahme ist jedoch, wenn F zusätzlich differenzierbar ist: Dann gilt (im Hilbertraum) $\partial F(x) = \{\nabla F(x)\}$. In diesem Fall erhalten wir das *Proximal-Gradientenverfahren*

$$x^{k+1} = \operatorname{prox}_{\gamma_k G}(x^k - \gamma_k \nabla F(x^k)) \tag{5.3}$$

(im Spezialfall $G = \delta_C$ – d. h. $\operatorname{prox}_{\gamma G}(x) = \operatorname{proj}_C(x)$ – auch *projiziertes Gradientenverfahren* genannt).

Um nun analog zum Proximalpunkt-Verfahren Konvergenz zu zeigen, müssen wir die Lipschitz-Stetigkeit des Gradienten voraussetzen (da wir für F ja keine Resolvente verwenden, die automatisch fest nichtexpansiv und damit Lipschitz-stetig ist).

Lemma 5.1. *Sei $F : X \to \mathbb{R}$ Gâteaux-differenzierbar mit Lipschitz-stetigem Gradienten, d. h. $\|\nabla F(x) - \nabla F(y)\|_X \leq L\|x-y\|_X$ für alle $x, y \in X$. Dann gilt*

$$F(y) \leq F(x) + (\nabla F(x), y-x)_X + \frac{L}{2}\|x-y\|_X^2 \quad \text{für alle } x, y \in X.$$

Beweis. Aus der Gâteaux-Differenzierbarkeit von F folgt

$$\frac{d}{dt}F(x+t(y-x)) = (\nabla F(x+t(y-x)), y-x)_X \quad \text{für alle } x, y \in X.$$

Integration über t auf $[0,1]$ liefert damit

$$\int_0^1 (\nabla F(x+t(y-x)), y-x)_X \, dt = F(y) - F(x).$$

Zusammen mit der produktiven Null erhalten wir dann unter Verwendung der Cauchy–Schwarz-Ungleichung und Lipschitz-Stetigkeit

$$\begin{aligned} F(y) &= F(x) + (\nabla F(x), y-x)_X + \int_0^1 (\nabla F(x+t(y-x)) - \nabla F(x), y-x)_X \, dt \\ &\leq F(x) + (\nabla F(x), y-x)_X + \int_0^1 \|\nabla F(x+t(y-x)) - \nabla F(x)\|_X \|x-y\|_X \, dt \\ &\leq F(x) + (\nabla F(x), y-x)_X + \int_0^1 Lt\|x-y\|_X^2 \, dt \\ &= F(x) + (\nabla F(x), y-x)_X + \frac{L}{2}\|x-y\|_X^2. \end{aligned}$$ □

Für genügend kleine Schrittweiten kann man damit die Konvergenz des Proximal-Gradientenverfahrens zeigen.

Satz 5.2. *Seien $F : X \to \mathbb{R}$ und $G : X \to \overline{\mathbb{R}}$ eigentlich, konvex, und unterhalbstetig. Sei zusätzlich F Gâteaux-differenzierbar mit Lipschitz-stetigem Gradienten. Gilt $0 < \gamma_{\min} \leq \gamma_k \leq L^{-1}$, dann konvergiert die durch (5.3) erzeugte Folge $\{x^k\}_{k\in\mathbb{N}}$ schwach gegen einen Minimierer $\bar{x}$ von J.*

Beweis. Wir argumentieren ähnlich wie im Beweis von Satz 4.14, wobei wir die Monotonie der Folge der Residuumsnormen $w^k \in Ax^{k+1}$ durch die der Funktionswerte $J(x^k)$ ersetzen. Wir definieren dafür

$$T_\gamma : X \to X, \qquad x \mapsto \gamma^{-1}(x - \mathrm{prox}_{\gamma G}(x - \gamma\nabla F(x))), \tag{5.4}$$

womit wir die Iteration (5.3) umschreiben können als

$$x^{k+1} = \operatorname{prox}_{\gamma_k G}(x^k - \gamma_k \nabla F(x^k)) = x^k - \gamma_k T_{\gamma_k}(x^k). \tag{5.5}$$

Nach Lemma 4.7 gilt daher

$$T_{\gamma_k}(x^k) - \nabla F(x^k) \in \partial G(x^k - \gamma_k T_{\gamma_k}(x^k)). \tag{5.6}$$

Aus Lemma 5.1 folgt weiter mit $x = x^k$, $y = x^{k+1} = x^k - \gamma_k T_{\gamma_k}(x^k)$, und $\gamma_k \leq L^{-1}$

$$\begin{aligned} F(x^k - \gamma_k T_{\gamma_k}(x^k)) &\leq F(x^k) - \gamma_k \left(\nabla F(x^k), T_{\gamma_k}(x^k)\right)_X + \frac{\gamma_k^2 L}{2} \|T_{\gamma_k}(x^k)\|_X^2 \\ &\leq F(x^k) - \gamma_k \left(\nabla F(x^k), T_{\gamma_k}(x^k)\right)_X + \frac{\gamma_k}{2} \|T_{\gamma_k}(x^k)\|_X^2. \end{aligned} \tag{5.7}$$

Damit erhalten wir aus der Definition von Subgradienten unter Verwendung von (5.6) und $\nabla F(x) \in \partial F(x)$ für beliebige $z \in X$, dass gilt

$$\begin{aligned} J(x^{k+1}) &= F(x^k - \gamma_k T_{\gamma_k}(x^k)) + G(x^k - \gamma_k T_{\gamma_k}(x^k)) \\ &\leq F(x^k) - \gamma_k \left(\nabla F(x^k), T_{\gamma_k}(x^k)\right)_X + \frac{\gamma_k}{2} \|T_{\gamma_k}(x^k)\|_X^2 \\ &\quad + G(z) + \left(T_{\gamma_k}(x^k) - \nabla F(x^k), x^k - \gamma_k T_{\gamma_k}(x^k) - z\right)_X \\ &\leq F(z) + \left(\nabla F(x^k), x^k - z\right)_X - \gamma_k \left(\nabla F(x^k), T_{\gamma_k}(x^k)\right)_X + \frac{\gamma_k}{2} \|T_{\gamma_k}(x^k)\|_X^2 \\ &\quad + G(z) + \left(T_{\gamma_k}(x^k) - \nabla F(x^k), x^k - z - \gamma_k T_{\gamma_k}(x^k)\right)_X \\ &= J(z) + \left(T_{\gamma_k}(x^k), x^k - z\right)_X - \frac{\gamma_k}{2} \|T_{\gamma_k}(x^k)\|_X^2. \end{aligned} \tag{5.8}$$

Für $z = x^k$ folgt daraus

$$J(x^{k+1}) \leq J(x^k) - \frac{\gamma_k}{2} \|T_{\gamma_k}(x^k)\|_X^2, \tag{5.9}$$

d. h. $\{J(x^k)\}_{k\in\mathbb{N}}$ ist monoton fallend. (Das Proximal-Gradientenverfahren ist also ein *Abstiegsverfahren.*) Für $z = x^*$ mit $J(x^*) = \min_{x\in X} J(x)$ folgt weiter durch quadratisches Ergänzen

$$\begin{aligned} 0 \leq J(x^{k+1}) - J(x^*) &\leq \left(T_{\gamma_k}(x^k), x^k - x^*\right)_X - \frac{\gamma_k}{2} \|T_{\gamma_k}(x^k)\|_X^2 \\ &= \frac{1}{2\gamma_k} \left(\|x^k - x^*\|_X^2 - \|x^k - x^* - \gamma_k T_{\gamma_k}(x^k)\|_X^2\right) \\ &= \frac{1}{2\gamma_k} \left(\|x^k - x^*\|_X^2 - \|x^{k+1} - x^*\|_X^2\right). \end{aligned} \tag{5.10}$$

Insbesondere ist $\{\|x^k - x^*\|_X\}_{k\in\mathbb{N}}$ monoton fallend und damit $\{x^k\}_{k\in\mathbb{N}}$ Féjer-monoton und beschränkt. Es existiert also eine schwach konvergente Teilfolge $\{x^{k_l}\}_{l\in\mathbb{N}}$ mit $x^{k_l} \rightharpoonup \bar{x}$.

Summieren über $k = 1, \dots, n$ liefert nun

$$\begin{aligned}\sum_{k=1}^{n}(J(x^k) - J(x^*)) &\leq \frac{1}{2\gamma_{\min}} \sum_{k=1}^{n} \left(\|x^{k-1} - x^*\|_X^2 - \|x^k - x^*\|_X^2\right) \\ &= \frac{1}{2\gamma_{\min}} \left(\|x^0 - x^*\|_X^2 - \|x^n - x^*\|_X^2\right) \\ &\leq \frac{1}{2\gamma_{\min}}\|x^0 - x^*\|_X^2.\end{aligned}$$

Da $\{J(x^k)\}_{k\in\mathbb{N}}$ monoton fallend ist, folgt

$$J(x^n) - J(x^*) \leq \frac{1}{n}\sum_{k=1}^{n}(J(x^k) - J(x^*)) \leq \frac{1}{2n\gamma_{\min}}\|x^0 - x^*\|_X^2 \tag{5.11}$$

und damit $J(x^n) \to J(x^*)$ für $n \to \infty$. Also gilt auch wegen der schwachen Unterhalbstetigkeit von F und G, dass

$$J(\bar{x}) \leq \liminf_{l\to\infty} J(x^{k_l}) = J(x^*).$$

Wie im Beweis von Satz 4.14 folgt nun aus der Féjer-Monotonie von $\{x^k\}_{k\in\mathbb{N}}$, dass $x^k \rightharpoonup \bar{x}$ für die gesamte Folge gilt. □

Aus (5.11) folgt insbesondere $J(x^k) = J(x^*) + O(k^{-1})$. Um $J(x^k) \leq J(x^*) + \varepsilon$ zu erreichen, sind also $O(\varepsilon^{-1})$ Iterationen notwendig. Durch eine geschickte Extrapolation lässt sich dies auf $O(\varepsilon^{-1/2})$ reduzieren, was beweisbar optimal ist; siehe [Nesterov 1983], [Nesterov 2004, Theorem 2.1.7]. (Dafür ist die Folge der Funktionswerte allerdings nicht mehr monoton fallend.) Die Iterationsvorschrift dafür ist

$$\begin{cases} x^{k+1} = \operatorname{prox}_{\gamma_k G}(\bar{x}^k - \gamma_k \nabla F(\bar{x}^k)), \\ \bar{x}^{k+1} = \left(1 - \dfrac{1-\tau_k}{\tau_{k+1}}\right) x^{k+1} + \dfrac{1-\tau_k}{\tau_{k+1}} x^k, \end{cases} \tag{5.12}$$

für die (schwer zu motivierende) Folge

$$\tau_0 = 1, \qquad \tau_k = \frac{1 + \sqrt{1 + 4\tau_{k-1}^2}}{2} \quad (\to \infty),$$

siehe [Beck & Teboulle 2009, § 4].

Ein Nachteil des expliziten Splittings ist die Notwendigkeit, für die Wahl einer zulässigen Schrittweite γ_k die Lipschitzkonstante L von ∇F zu kennen. Im Beweis von Satz 5.2 erkennt man, dass dies lediglich dazu dient, die Abschätzung (5.7) zu garantieren. Ist die Lipschitzkonstante nicht bekannt, kann man versuchen, die Gültigkeit von (5.7) durch Liniensuche zu erfüllen: Man startet mit $\gamma^0 > 0$ und verkleinert γ_k (etwa durch Halbieren) so lange, bis

$$F(x^k - \gamma_k T_{\gamma_k}(x^k)) \le F(x^k) - \gamma_k \left(\nabla F(x^k), T_{\gamma_k}(x^k)\right)_X + \frac{\gamma_k}{2}\|T_{\gamma_k}(x^k)\|_X^2$$

erfüllt ist (was spätestens für $\gamma_k < L^{-1}$ der Fall ist). Dafür ist in jedem Schritt der Liniensuche jeweils eine Auswertung von F und $\operatorname{prox}_{\gamma_k G}$ erforderlich (wobei Letzteres durch Vertauschen von Proximal- und Gradientenschritt – *backward–forward splitting* – vermieden werden kann).

5.2 PRIMAL-DUALES SPLITTING

Diese Verfahren wurden speziell für Probleme der Form

$$\min_{x\in X} F(x) + G(Ax)$$

für F, G eigentlich, konvex und unterhalbstetig und $A \in L(X, Y)$ entwickelt. Aus Satz 3.19 und Lemma 3.18 erhält man dafür die Fenchel-Extremalitätsbedingungen

$$\begin{cases} -A^*\bar y \in \partial F(\bar x), \\ \quad\ \bar y \in \partial G(A\bar x), \end{cases} \quad\Leftrightarrow\quad \begin{cases} -A^*\bar y \in \partial F(\bar x), \\ \quad A\bar x \in \partial G^*(\bar y). \end{cases} \tag{5.13}$$

Die zugehörige Proximalpunktformulierung nach Lemma 4.7 ist

$$\begin{cases} \bar x = \operatorname{prox}_{\tau F}(\bar x - \tau A^*\bar y), \\ \bar y = \operatorname{prox}_{\sigma G^*}(\bar y + \sigma A\bar x), \end{cases}$$

für beliebige $\sigma, \tau > 0$. Daraus lässt sich direkt eine Fixpunktiteration gewinnen. Es ist aber sowohl aus praktischer als auch aus theoretischer Sicht hilfreich, einen Extrapolationsschritt einzuschieben:

$$\begin{cases} x^{k+1} = \operatorname{prox}_{\tau F}(x^k - \tau A^* y^k), \\ \bar x^{k+1} = 2x^{k+1} - x^k, \\ y^{k+1} = \operatorname{prox}_{\sigma G^*}(y^k + \sigma A\bar x^{k+1}). \end{cases} \tag{5.14}$$

Dieses *primal-duale Extragradienten-Verfahren* kann man nun – dank des zusätzlichen Extrapolationsschritts! – als Proximalpunkt-Verfahren auffassen.[1] Dazu formen wir (5.14)

[1] Eingeführt wurde das Verfahren in [Chambolle & Pock 2011], weshalb es in der Literatur auch oft als *Chambolle–Pock-Verfahren* bezeichnet wird; der Zusammenhang mit Proximalpunkt-Verfahren wurde in [He & Yuan 2012] beschrieben.

so um, dass (x^k, y^k) und (x^{k+1}, y^{k+1}) auf jeweils einer Seite stehen. Verwenden von $\operatorname{prox}_{\tau F} = (\operatorname{Id} + \tau\partial F)^{-1}$ ergibt für die erste Gleichung:

$$
\begin{aligned}
x^{k+1} = \operatorname{prox}_{\tau F}(x^k - \tau A^* y^k) &\Leftrightarrow x^k - \tau A^* y^k \in \{x^{k+1}\} + \tau\partial F(x^{k+1})\\
&\Leftrightarrow \tau^{-1}x^k - A^* y^k \in \{\tau^{-1}x^{k+1}\} + \partial F(x^{k+1}).
\end{aligned}
$$

Analog haben wir für die zweite Gleichung (nach Elimination von $\bar{x}^{k+1}$)

$$
y^{k+1} = \operatorname{prox}_{\sigma G^*}(y^k + \sigma A(2x^{k+1} - x^k)) \Leftrightarrow \sigma^{-1}y^k - Ax^k \in \{\sigma^{-1}y^{k+1} - 2Ax^{k+1}\} + \partial G^*(y^{k+1}).
$$

Setzen wir $Z = X \times Y$, $z = (x, y)$,

$$
M = \begin{pmatrix} \tau^{-1}\operatorname{Id} & -A^* \\ -A & \sigma^{-1}\operatorname{Id} \end{pmatrix}, \qquad T = \begin{pmatrix} \partial F & A^* \\ -A & \partial G^* \end{pmatrix},
$$

so ist (5.14) äquivalent zu

$$
Mz^k \in (M+T)z^{k+1} \qquad \Leftrightarrow \qquad z^{k+1} \in (M+T)^{-1}Mz^k.
$$

Ist M invertierbar, so gilt $M = (M^{-1})^{-1}$ und damit $(M+T)^{-1}Mz^k = (\operatorname{Id} + M^{-1}T)^{-1}z^k$; es handelt sich dann also in der Tat um ein Proximalpunkt-Verfahren für den Operator $M^{-1}T$. Dazu zeigen wir, dass – unter Voraussetzungen an σ, τ – M selbstadjungiert und positiv definit ist bezüglich des Skalarproduktes

$$
(z_1, z_2)_Z = (x_1, x_2)_X + (y_1, y_2)_Y \quad \text{für alle } z_1 = (x_1, y_1) \in Z, z_2 = (x_2, y_2) \in Z.
$$

Lemma 5.3. *Der Operator M ist beschränkt und selbstadjungiert. Gilt $\sigma\tau\|A\|^2_{L(X,Y)} < 1$, so ist M positiv definit.*

Beweis. Direkt aus der Definition von M folgt die Beschränktheit (da $A \in L(X, Y)$ beschränkt ist) und Selbstadjungiertheit. Sei nun $z = (x, y) \in Z$ beliebig. Dann ist

$$
\begin{aligned}
(Mz, z)_Z &= \left(\tau^{-1}x - A^* y, x\right)_X + \left(\sigma^{-1}y - Ax, y\right)_Y\\
&= \tau^{-1}\|x\|_X^2 - 2\,(x, A^* y)_X + \sigma^{-1}\|y\|_Y^2\\
&\geq \tau^{-1}\|x\|_X^2 - 2\|A\|_{L(X,Y)}\|x\|_X\|y\|_Y + \sigma^{-1}\|y\|_Y^2\\
&\geq \tau^{-1}\|x\|_X^2 - \|A\|_{L(X,Y)}\sqrt{\sigma\tau}(\tau^{-1}\|x\|_X^2 + \sigma^{-1}\|y\|_Y^2) + \sigma^{-1}\|y\|_Y^2\\
&= (1 - \|A\|_{L(X,Y)}\sqrt{\sigma\tau})(\tau^{-1}\|x\|_X^2 + \sigma^{-1}\|y\|_Y^2)\\
&\geq C(\|x\|_X^2 + \|y\|_Y^2)
\end{aligned}
$$

für $C := (1 - \|A\|_{L(X,Y)}\sqrt{\sigma\tau})\min\{\tau^{-1}, \sigma^{-1}\} > 0$. Also ist $(Mz, z)_Z \geq C\|z\|_Z^2$ für alle $z \in Z$ und damit M positiv definit. □

Der Operator M induziert also ein Skalarprodukt $(z_1, z_2)_M := (Mz_1, z_2)_Z$ und dadurch eine Norm $\|z\|_M^2 = (z, z)_M$, für die gilt

$$c_1\|z\|_Z \leq \|z\|_M \leq c_2\|z\|_Z \qquad \text{für alle } z \in Z. \tag{5.15}$$

Folgerung 5.4. *Gilt $\sigma\tau\|A\|^2_{L(X,Y)} < 1$, so ist M stetig invertierbar, d. h. es gilt $M^{-1} \in L(Z, Z)$.*

Beweis. Wegen (5.15) ist für beliebige $z \in Z$ die Abbildung $v \mapsto (z, v)_Z$ ein (bezüglich $\|\cdot\|_M$) beschränktes Funktional. Nach dem Satz von Fréchet–Riesz, angewendet auf den Hilbertraum $(Z, (\cdot,\cdot)_M)$, existiert also ein eindeutiges $z^* \in Z$ mit

$$(Mz^*, v)_Z = (z^*, v)_M = (z, v)_Z \qquad \text{für alle } v \in Z,$$

und die Abbildung $M^{-1} : z \mapsto z^*$ ist linear. Damit gilt

$$c_1^2\|z^*\|_Z^2 \leq \|z^*\|_M^2 = (Mz^*, z^*)_Z = (z, z^*)_Z \leq \|z\|_Z\|z^*\|_Z,$$

und nach Division durch $c_1^2\|z^*\|_Z$ folgt die behauptete Beschränktheit von M^{-1}. □

Also ist $M^{-1}T$ wohldefiniert, d. h. $\operatorname{graph} M^{-1}T \neq \emptyset$. Wir zeigen nun, dass $M^{-1}T$ maximal monoton ist bezüglich des Skalarproduktes $(\cdot,\cdot)_M$

Lemma 5.5. *Gilt $\sigma\tau\|A\|^2_{L(X,Y)} < 1$, so ist $M^{-1}T$ maximal monoton in $(X, (\cdot,\cdot)_M)$.*

Beweis. Wir zeigen zunächst die Monotonie. Sei $z \in Z$ und $z^* \in M^{-1}Tz$, d. h. $Mz^* \in Tz$. Nach Definition von T existieren daher für $z = (x, y)$ ein $\xi \in \partial F(x)$ und ein $\eta \in \partial G^*(y)$ mit $Mz^* = (\xi + A^*y, \eta - Ax)$. Analog können wir für $\bar z = (\bar x, \bar y) \in Z$ und $\bar z^* \in M^{-1}T\bar z$ schreiben $M\bar z^* = (\bar\xi + A^*\bar y, \bar\eta - A\bar x)$ für $\bar\xi \in \partial F(\bar x)$ und $\bar\eta \in \partial G^*(\bar y)$. Dann gilt

$$\begin{aligned}
(\bar z^* - z^*, \bar z - z)_M = (M\bar z^* - Mz^*, \bar z - z)_Z &= \big((\bar\xi + A^*\bar y) - (\xi + A^*y), \bar x - x\big)_X \\
&\quad + ((\bar\eta - A\bar x) - (\eta - Ax), \bar y - y)_Y \\
&= \big(\bar\xi - \xi, \bar x - x\big)_X + (A^*(\bar y - y), \bar x - x)_X \\
&\quad - (A(\bar x - x), \bar y - y)_Y + (\bar\eta - \eta, \bar y - y)_Y \\
&= \big(\bar\xi - \xi, \bar x - x\big)_X + (\bar\eta - \eta, \bar y - y)_Y \geq 0
\end{aligned}$$

wegen der Monotonie der Subdifferentiale.

Für die maximale Monotonie seien nun $\bar z^*, \bar z \in Z$ und es gelte

$$(\bar z^* - z^*, \bar z - z)_M \geq 0 \quad \text{für alle } (z^*, z) \in \operatorname{graph} M^{-1}T.$$

Wie oben ist nun $Mz^* = (\xi + A^*y, \eta - Ax)$ für ein $\xi \in \partial F(x)$ und ein $\eta \in \partial G^*(y)$. Also gilt

$$(M\bar z^* - Mz^*, \bar z - z)_Z = (\bar z^* - z^*, \bar z - z)_M \geq 0 \quad \text{für alle } z \in Z, Mz^* \in Tz. \tag{5.16}$$

Setze nun $\bar{\xi} := \tilde{x}^* - A^*\bar{y}$ und $\bar{\eta} := \tilde{y}^* + A\bar{x}$ für $M\bar{z}^* = (\tilde{x}^*, \tilde{y}^*)$ und $\bar{z} = (\bar{x}, \bar{y})$. Dann ist $M\bar{z}^* = (\bar{\xi} + A^*\bar{y}, \bar{\eta} - A\bar{x})$, und aus (5.16) folgt daher für alle $(x, y) \in Z$:

$$
\begin{aligned}
0 &\le \big((\bar{\xi} + A^*\bar{y}) - (\xi + A^*y), \bar{x} - x\big)_X + ((\bar{\eta} - A\bar{x}) - (\eta - Ax), \bar{y} - y)_Y \\
&= \big(\bar{\xi} - \xi, \bar{x} - x\big)_X + (\bar{\eta} - \eta, \bar{y} - y)_Y\,.
\end{aligned}
$$

Setzen wir speziell Paare der Form $(x, \bar{y})$ für $x \in X$ beliebig und $(\bar{x}, y)$ für $y \in Y$ beliebig ein, erhalten wir, dass beide Skalarprodukte auf der rechten Seiten nichtnegativ sind. Aus der maximalen Monotonie von Subdifferentialen folgt nun $\bar{\xi} \in \partial F(\bar{x})$ und $\bar{\eta} \in \partial G^*(\bar{y})$. Also ist

$$
M\bar{z}^* = (\bar{\xi} + A^*\bar{y}, \bar{\eta} - A\bar{x}) \in T\bar{z}
$$

und damit $\bar{z}^* \in M^{-1}T\bar{z}$. Dies zeigt die maximale Monotonie von $M^{-1}T$. □

Das primal-duale Extragradienten-Verfahren (5.14) ist also äquivalent mit der Proximalpunktiteration $z^{k+1} = R_{M^{-1}T}z^k$ für $M^{-1}T$ maximal monoton, so dass aus Satz 4.14 zusammen mit der positiven Definitheit von M die schwache Konvergenz folgt.

Satz 5.6. *Erfüllen F, G, A die Voraussetzungen von Satz 3.19 und gilt $\sigma\tau\|A\|^2_{L(X,Y)} < 1$, so konvergiert $\{(x^k, y^k)\}_{k\in\mathbb{N}}$ schwach in Z gegen eine Lösung $(\bar{x}, \bar{y})$ von (5.13).*

Beweis. Aus Satz 4.14 erhalten wir wegen Lemma 5.5 zunächst die schwache Konvergenz von $z_k := (x^k, y^k)$ bezüglich des Skalarprodukts $(\cdot, \cdot)_M$. Da aber M symmetrisch und positiv definit und damit invertierbar ist, folgt daraus

$$
(z_k, Mw)_X = (z_k, w)_M \to (z, Mw)_X \quad \text{für alle } w \in X
$$

und damit schwache Konvergenz bezüglich $(\cdot, \cdot)_X$, da mit w auch Mw ganz X durchläuft. □

Beachten Sie, dass das Verfahren durch die Proximalpunktformulierung der Fenchel-Extremalitätsbedingungen zwar implizit in F und G, nicht jedoch in A ist; es ist also nicht verwunderlich, dass dadurch eine Schrittweitenrestriktion an τ und σ durch A entsteht.[2]

[2]Die Proximalpunktabbildung zu $G \circ A$ würde zu einem voll impliziten Verfahren führen, dafür aber aufgrund der Rechenregeln die Anwendung von A^{-1} benötigen. Es ist gerade Sinn und Zweck des primal-dualen Extragradientenverfahrens, die Invertierung von A – was in der Praxis oft schlecht konditioniert oder sogar unmöglich ist – zu vermeiden.

Teil III

BILDMODELLE

6 KONTINUIERLICHE BILDMODELLE

Um die Werkzeuge der Variationsrechnung auf Probleme der Bildverarbeitung anwenden zu können, fassen wir nun (Schwarz-Weiss-)Bilder auf als Funktionen $u : \Omega \to \mathbb{R}$ mit $\Omega \subset \mathbb{R}^d$ offen und beschränkt (z. B. $\Omega = (0,1)^2$). Dabei ist es wichtig, einen korrekten Funktionenraum zu finden, in dem wir Bilder suchen: Er sollte genug Regularität erfordern, um zufälliges Rauschen auszuschliessen, aber nicht so viel Regularität, dass keine Sprünge (d. h. Kanten im Bild) erlaubt sind.

6.1 LEBESGUE-RÄUME

Die grundlegende Klasse von Funktionenräumen (die wir später einschränken wollen) sind die Räume der Lebesgue-integrierbaren Funktionen. Dabei handelt es sich um Äquivalenzklassen von Funktionen, die bezüglich des Lebesgue-Maßes messbar sind; identifiziert werden Funktionen, die sich nur auf einer Menge vom Lebesgue-Maß Null unterscheiden. Für $1 \leq p \leq \infty$ sind die *Lebesgue-Räume* definiert als

$$L^p(\Omega) := \{u \text{ messbar} : \|u\|_{L^p} < \infty\}$$

mit den in Beispiel 1.1 (iii) definierten Normen. Wir erinnern, dass es sich dabei um Banachräume und für $p = 2$ um einen Hilbertraum handelt, und dass $L^p(\Omega)^* \cong L^q(\Omega)$ für $1 \leq p < \infty$ und $p^{-1} + q^{-1} = 1$ ist. (Diese Identifizierung werden wir im Folgenden stillschweigend verwenden und $L^q(\Omega)$ als Dualraum bezeichnen.) Die Ungleichung (1.1) wird dann zur *Hölder-Ungleichung*

$$\int_\Omega u(x)v(x)\,dx \leq \|u\|_{L^p}\|v\|_{L^q} \quad \text{für alle } u \in L^p(\Omega), v \in L^q(\Omega).$$

Daraus folgt sofort für beschränkte Gebiete Ω die Einbettung $L^p(\Omega) \hookrightarrow L^{p'}(\Omega)$ für $p \geq p'$. Da dies für Bilder keine Einschränkung darstellt, betrachten wir in Folge stets beschränkte Gebiete.

Eine Sonderrolle (als der "größtmögliche" Lebesgue-Raum) spielt

$$L^1_{\text{loc}}(\Omega) := \left\{u \text{ messbar} : u|_K \in L^1(K) \text{ für alle } K \subset \Omega \text{ kompakt}\right\}.$$

Ein Grund, mit Lebesgue-integrierbaren Funktionen zu arbeiten, ist die Verträglichkeit dieses Integralbegriffs mit der Unterhalbstetigkeit.

Lemma 6.1. *Sei $\varphi : \mathbb{R} \to \overline{\mathbb{R}}$ eigentlich, konvex und unterhalbstetig mit $\varphi \geq 0$. Dann gilt dies auch für*

$$\Phi : L^p(\Omega) \to \overline{\mathbb{R}}, \qquad \Phi(u) = \begin{cases} \int_\Omega \varphi(u(x))\,dx & \varphi \circ u \in L^1(\Omega), \\ \infty & \text{sonst.} \end{cases}$$

Beweis. Da φ eigentlich ist, existiert ein $t_0 \in \operatorname{dom}\varphi$. Also ist die konstante Funktion $u \equiv t_0 \in \operatorname{dom}\Phi$, da $\varphi(u) \equiv \varphi(t_0) \in L^\infty(\Omega) \subset L^1(\Omega)$.

Für $u, v \in \operatorname{dom}\Phi$ und $\lambda \in (0,1)$ folgt für fast alle $x \in \Omega$ aus der Konvexität von φ, dass

$$\varphi(\lambda u(x) + (1-\lambda)v(x)) \leq \lambda\varphi(u(x)) + (1-\lambda)\varphi(v(x)).$$

Da für $f, g \in L^1(\Omega)$ und $\alpha, \beta \in \mathbb{R}$ auch $\alpha f + \beta g \in L^1(\Omega)$ gilt, ist also $\lambda u + (1-\lambda)v \in \operatorname{dom}\Phi$, und durch Integration der Ungleichung über Ω folgt die Konvexität von Φ.

Für die Unterhalbstetigkeit verwenden wir Lemma 3.1. Sei dafür $\{(u_n, t_n)\}_{n\in\mathbb{N}} \subset \operatorname{epi} F$ mit $u_n \to u$ in $L^p(\Omega)$ und $t_n \to t$ in $\mathbb{R}$. Dann existiert eine Teilfolge $\{u_{n_k}\}_{k\in\mathbb{N}}$ mit $u_{n_k}(x) \to u(x)$ fast überall. Wegen der Unterhalbstetigkeit und Nichtnegativität von f folgt mit dem Lemma von Fatou

$$\begin{aligned} F(u) &= \int_\Omega f(u(x))\,dx \leq \int_\Omega \liminf_{k\to\infty} f(u_{n_k}(x))\,dx \leq \liminf_{k\to\infty} \int_\Omega f(u_{n_k}(x))\,dx \\ &= \liminf_{k\to\infty} F(u_{n_k}) \leq \limsup_{k\to\infty} t_{n_k} = t, \end{aligned}$$

d. h. $(u, t) \in \operatorname{epi} F$. Also ist $\operatorname{epi} F$ abgeschlossen und damit F unterhalbstetig nach Lemma 3.1 (iii). $\square$

Für solche Funktionale lässt sich das Subdifferential punktweise charakterisieren. Im Folgenden sei stets $p^{-1} + q^{-1} = 1$, d. h. $L^q(\Omega) = L^p(\Omega)^*$ für $1 \leq p < \infty$, angenommen.

Lemma 6.2. *Sei $\varphi : \mathbb{R} \to \overline{\mathbb{R}}$ eigentlich, konvex und unterhalbstetig und $\Phi : L^p(\Omega) \to \overline{\mathbb{R}}$ mit $1 \leq p < \infty$ wie in Lemma 6.1. Dann ist für alle $u \in \operatorname{dom}\Phi$*

$$\partial\Phi(u) = \{u^* \in L^q(\Omega) : u^*(x) \in \partial\varphi(u(x)) \quad \text{für fast alle } x \in \Omega\}.$$

Beweis. Sei $u \in \operatorname{dom}\Phi$, d. h. $\varphi \circ u \in L^1(\Omega)$. Gilt für $u^* \in L^q(\Omega)$ dass $u^*(x) \in \partial\varphi(u(x))$ fast überall, so folgt daraus durch Integration

$$\Phi(\tilde{u}) - \Phi(u) = \int_\Omega \varphi(\tilde{u}(x)) - \varphi(u(x))\,dx \geq \int_\Omega u^*(x)(\tilde{u}(x) - u(x))\,dx = \langle u^*, \tilde{u} - u\rangle_{L^p}$$

für alle $\tilde{u} \in L^p(\Omega)$.

Sei umgekehrt $u^* \in \partial\Phi(u)$. Dann gilt nach Definition

$$\int_\Omega u^*(x)(\tilde{u}(x) - u(x))\,dx \le \int_\Omega \varphi(\tilde{u}(x)) - \varphi(u(x))\,dx \quad \text{für alle } \tilde{u} \in L^p(\Omega).$$

Sei nun $t \in \mathbb{R}$ beliebig und sei $A \subset \Omega$ eine beliebige messbare Menge. Für die Wahl

$$\tilde{u}(x) := \begin{cases} t & x \in A, \\ u(x) & x \notin A, \end{cases}$$

folgt wegen $\tilde{u} \in L^p(\Omega)$ aus der obigen Ungleichung

$$\int_A u^*(x)(t - u(x))\,dx \le \int_A \varphi(t) - \varphi(u(x))\,dx.$$

Da A beliebig war, muss gelten

$$u^*(x)(t - u(x)) \le \varphi(t) - \varphi(u(x)) \qquad \text{für fast alle } x \in \Omega.$$

Da $t \in \mathbb{R}$ beliebig war, folgt daraus $u^*(x) \in \partial u(x)$ für fast alle $x \in \Omega$. □

Auf ähnliche Weise zeigt man, dass sich auch die Fenchel-Konjugierte punktweise berechnen lässt.[1]

Lemma 6.3. *Sei $\varphi : \mathbb{R} \to \overline{\mathbb{R}}$ eigentlich, konvex und unterhalbstetig, und sei $\Phi : L^p(\Omega) \to \overline{\mathbb{R}}$ mit $1 \le p < \infty$ wie in Lemma 6.1. Dann ist die Fenchel-Konjugierte von Φ gegeben durch*

$$\Phi^* : L^q(\Omega) \to \overline{\mathbb{R}}, \quad \Phi^*(u^*) = \int_\Omega \varphi^*(u^*(x))\,dx.$$

Auf dem anderen Ende der Glattheitsskala liegen die unendlich oft differenzierbaren *Testfunktionen* in

$$C_0^\infty(\Omega) := \{\varphi : \Omega \to \mathbb{R} : \varphi \text{ unendlich oft stetig differenzierbar, } \operatorname{supp}\varphi \subset \Omega \text{ kompakt}\},$$

wobei $\operatorname{supp}\varphi = \overline{\{x \in \Omega : \varphi(x) \neq 0\}}$ den *Träger* von φ bezeichnet. Ihr Wert liegt darin, dass sich Lebesgue-integrierbare Funktionen beliebig gut durch Testfunktionen approximieren lassen; man sagt, $C_0^\infty(\Omega)$ *liegt dicht* in $L^p(\Omega)$.[2]

Satz 6.4. *Für alle $u \in L^p(\Omega)$, $1 \le p < \infty$, existiert eine Folge $\{\varphi_n\}_{n\in\mathbb{N}} \subset C_0^\infty(\Omega)$ mit $\varphi_n \to u$ in $L^p(\Omega)$.*

[1] siehe z. B. [Ekeland & Témam 1999, Proposition IV.1.2]
[2] siehe z. B. [Dobrowolski 2010, Satz 4.23]

Dieses Resultat erlaubt die klassische Beweistechnik des *Dichtheitsarguments*: Will man eine Eigenschaft für alle $u \in L^p(\Omega)$ nachweisen, so zeigt man sie zuerst für alle $\varphi \in C_0^\infty(\Omega)$ und argumentiert, dass sie beim Grenzübergang in der approximierenden Folge erhalten bleibt. So zeigt man beispielsweise das folgende wichtige Resultat.

Satz 6.5 (Fundamentallemma der Variationsrechnung[3]). *Sei $u \in L^1_{\text{loc}}(\Omega)$ mit*

$$\int_\Omega u(x)\varphi(x)\,dx \geq 0 \qquad \textit{für alle } \varphi \in C_0^\infty(\Omega).$$

Dann ist $u(x) = 0$ für fast alle $x \in \Omega$.

6.2 SOBOLEV-RÄUME

Es ist klar, dass Testfunktionen kein gutes Modell für Bilder darstellen. Statt Funktionen, die im klassischen Sinn differenzierbar sind, betrachten wir daher als nächstes *schwach differenzierbare* Funktionen. Diese Funktionen sind die Grundlage der modernen Theorie partieller Differentialgleichungen und eignen sich besonders für Funktionale mit Integraldarstellung. Ausgangspunkt ist, die klassische punktweise Definition der Ableitung durch eine „Definition im Mittel" zu ersetzen; als grundlegende Eigenschaft fordert man dabei, dass der neue Ableitungsbegriff verträglich ist mit der partiellen Integration. Dafür ist es nützlich, einen *Multiindex*

$$\alpha := (\alpha_1, \dots, \alpha_d) \in \mathbb{N}_0^d$$

mit Länge $|\alpha| := \sum_{i=1}^d \alpha_i$ zu definieren, um für f definiert auf $\Omega \subset \mathbb{R}^d$ eine (klassische) partielle Ableitung der Ordnung $|\alpha|$ kompakt schreiben zu können als

$$D^\alpha f(x_1, \dots, x_d) := \frac{\partial^{|\alpha|} f(x_1, \dots, x_d)}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}}.$$

(Wir schreiben auch kurz $\partial_i := \frac{\partial}{\partial x_i}$.) Eine Funktion $u \in L^1_{\text{loc}}(\Omega)$ heißt *schwach differenzierbar*, falls ein $v \in L^1_{\text{loc}}(\Omega)$ existiert mit

$$\int_\Omega v(x)\varphi(x)\,dx = (-1)^{|\alpha|} \int_\Omega u(x) D^\alpha \varphi(x)\,dx \qquad \text{für alle } \varphi \in C_0^\infty(\Omega). \tag{6.1}$$

Nach dem Fundamentallemma der Variationsrechnung ist dann die *schwache Ableitung* eindeutig definiert als $D^\alpha u := v$; besitzt u eine entsprechende klassische Ableitung, so stimmt die schwache Ableitung mit dieser überein (und wir unterscheiden daher nicht in der Notation). Die schwache Ableitung existiert aber für ein grössere Klasse von Funktionen;

[3]siehe z. B. [Dobrowolski 2010, Satz 5.1]

zum Beispiel hat $u : (-1,1) \to \mathbb{R}$, $u(x) = |x|$, die schwache Ableitung $Du : (-1,1) \to \mathbb{R}$, $Du(x) = 1$ für $x \geq 0$ und $Du(x) = -1$ für $x < 0$.

Wir können nun die *Sobolev-Räume* $W^{k,p}(\Omega)$ für $k \in \mathbb{N}_0$ und $1 \leq p \leq \infty$ definieren als

$$W^{k,p}(\Omega) = \left\{ u \in L^p(\Omega) : D^\alpha u \in L^p(\Omega) \text{ für alle } |\alpha| \leq k \right\}.$$

Diese sind Banachräume[4] versehen mit der Norm

$$\begin{aligned} \|u\|_{W^{k,p}} &= \left(\sum_{|\alpha| \leq k} \|D^\alpha u\|_{L^p}^p \right)^{\frac{1}{p}} \qquad \text{für } 1 \leq p < \infty, \\ \|u\|_{W^{k,\infty}} &= \max_{|\alpha| \leq k} \|D^\alpha u\|_{L^\infty}. \end{aligned}$$

Für $p = 2$ handelt es sich um Hilberträume mit Skalarprodukt

$$(u, v)_{W^{k,2}} = \sum_{|\alpha| \leq k} (D^\alpha u, D^\alpha v)_{L^2}.$$

Man schreibt daher üblicherweise $H^k(\Omega) := W^{k,2}(\Omega)$.

Auch Sobolev-Funktionen lassen sich beliebig gut durch Testfunktionen (ohne kompakten Träger) approximieren.[5]

Satz 6.6. *Für alle $k \in \mathbb{N}$ und $1 \leq p < \infty$ liegt $C^\infty(\Omega) \cap W^{k,p}(\Omega)$ dicht in $W^{k,p}(\Omega)$.*

Damit kann man etliche Eigenschaften von Sobolev-Funktionen mit Dichtheitsargumenten beweisen. Ein nützliches Resultat ist

Lemma 6.7. *Gilt für ein $u \in W^{k,p}(\Omega)$, dass $D^\alpha u = 0$ für alle $|\alpha| = k$ ist, dann stimmt u fast überall mit einem Polynom vom Grad $k - 1$ überein.*

Die Norm im Sobolev-Raum misst die Regularität von Funktionen auf unterschiedliche Weise, bestimmt durch die beiden Zahlen k und p: Je grösser k, desto stärker fallen *Oszillationen* ins Gewicht; je grösser m, desto stärker fällt eine *Konzentration* ins Gewicht. Ist Ω glatt genug, so kann man diese beiden Eigenschaften in einem bestimmten Rahmen in Beziehung setzen.[6] Ist $\Omega \subset \mathbb{R}^d$ offen, zusammenhängend und beschränkt, und kann der Rand $\partial\Omega$ von Ω durch eine Lipschitz-stetige Funktion parametrisiert werden, so nennt man Ω *Lipschitz-Gebiet.* Insbesondere ist das Einheitsquadrat $(0,1)^2$ ein Lipschitzgebiet.

[4] siehe z. B. [Dobrowolski 2010, Satz 5.10]

[5] Dies wurde bewiesen von Meyers und Serrin in einer Arbeit, die zu Recht berühmt ist sowohl für die Bedeuting ihres Inhalts als auch für die Kürze ihres Titels, „$H = W$“. Für den Beweis siehe z. B. [Dobrowolski 2010, Satz 5.16].

[6] Für eine schöne Darstellung dieser Intuition, siehe [Tao 2010b] sowie [Tao 2010a].

Satz 6.8 (Einbettungssatz von Sobolev, Morrey[7]). *Sei $1 \leq p, q < \infty$ und $\Omega \subset \mathbb{R}^d$ ein Lipschitz-Gebiet. Dann sind die folgenden Einbettungen stetig:*

$$W^{k,p}(\Omega) \hookrightarrow \begin{cases} L^q(\Omega) & p < \frac{d}{k} \text{ und } p \leq q \leq \frac{dp}{d-kp}, \\ L^q(\Omega) & p = \frac{d}{k} \text{ und } p \leq q < \infty, \\ C(\overline{\Omega}) & p > \frac{d}{k}. \end{cases}$$

Auf solchen Gebieten lassen sich Sobolev-Funktionen auch stetig (bzgl. der Norm) auf größere Gebiete fortsetzen. Je glatter das Gebiet, desto mehr Regularität bleibt dabei erhalten. Wir definieren $C^{m,1}$-Gebiete analog zu Lipschitz-Gebieten (die für $m = 0$ in dieser Definition enthalten sind).

Satz 6.9 (Fortsetzung[8]). *Sei $m \in \mathbb{N}$ und Ω ein $C^{m-1,1}$-Gebiet. Zu jedem Gebiet Ω', das $\overline{\Omega}$ als kompakte Teilmenge enthält, existiert für alle $0 \leq k \leq m$ ein $E \in L(W^{k,p}(\Omega), W^{k,p}(\Omega'))$ mit $(Eu)|_\Omega = u$.*

Wir beschränken uns in Folge auf den Fall $k = 1$, und verwenden für $u \in W^{1,p}(\Omega)$ die übliche Notation $\nabla u := (\partial_1 u, \dots, \partial_d u)^T \in L^p(\Omega)^d$ für den *Gradienten* von u sowie

$$|u|_{W^{1,p}} := \|\nabla u\|_{(L^p)^d} = \left(\int_\Omega |\nabla u(x)|_p^p \, dx \right)^{1/p} \tag{6.2}$$

für die *Seminorm* auf $W^{1,p}(\Omega)$. Damit ist

$$\|u\|_{W^{1,p}} = (\|u\|_{L^p}^p + |u|_{W^{1,p}}^p)^{1/p},$$

sowie für $p = 2$

$$(u, v)_{H^1} = (u, v)_{L^2} + (\nabla u, \nabla v)_{(L^2)^d}\,.$$

Durch partielle Integration definieren wir nun für eine Vektor-Funktion $u = (u_1, \dots, u_d)^T \in L^1_{\mathrm{loc}}(\Omega)^d$ die schwache *Divergenz*

$$\operatorname{div} u = \sum_{i=1}^d \partial_i u_i \in L^1_{\mathrm{loc}}(\Omega),$$

falls gilt

$$\int_\Omega \operatorname{div} u \varphi \, dx = -\int_\Omega u \cdot \nabla \varphi \, dx \qquad \text{für alle } \varphi \in C_0^\infty(\Omega).$$

Sucht man nun für Bildverarbeitungsaufgaben Minimierer in $W^{1,p}(\Omega)$, so ist es naheliegend, die entsprechende Seminorm als Regularisierungsterm zu verwenden. Für die direkte Methode müssen wir Konvexität, Unterhalbstetigkeit und Koerzivität untersuchen. Dafür zeigen wir zuerst die Abgeschlossenheit von ∇.

[7] siehe z. B. [Adams & Fournier 2003, Theorem 4.12], [Dobrowolski 2010, Satz 6.24]
[8] siehe z. B. [Dobrowolski 2010, Satz 6.10]

Lemma 6.10. *Sei $\Omega \subset \mathbb{R}^d$ und $1 \leq p, q < \infty$. Dann ist*

$$\nabla : \left\{u \in L^q(\Omega) : \nabla u \in L^p(\Omega)^d\right\} \to L^p(\Omega)^d$$

schwach abgeschlossen, d. h. $u_n \rightharpoonup u$ in $L^p(\Omega)$ und $\nabla u_n \rightharpoonup v$ in $L^q(\Omega)^d$ impliziert $v = \nabla u$. Die Aussage gilt für $p = \infty$ oder $q = \infty$, wenn die entsprechende Konvergenz durch schwach-$*$ Konvergenz ersetzt wird.

Beweis. Sei $\{u_n\}_{n\in\mathbb{N}} \subset L^q(\Omega)$ eine Folge mit $u_n \rightharpoonup u$ in $L^q(\Omega)$ und $L^p(\Omega)^d \ni \nabla u_n \rightharpoonup v$ in $L^p(\Omega)^d$). Für alle $\varphi \in C_0^\infty(\Omega)$ gilt insbesondere $\varphi \in L^p(\Omega)^*$ und $\partial_k \varphi \in L^q(\Omega)^*$ für alle $1 \leq k \leq d$. Nach Definition der schwachen Konvergenz und der schwachen Ableitung gilt dann

$$\begin{aligned}
\int_\Omega u \partial_k \varphi \, dx &= \langle \partial_k \varphi, u\rangle_{L^q} = \lim_{n\to\infty} \langle \partial_k \varphi, u_n\rangle_{L^q} \\
&= \lim_{n\to\infty} \int_\Omega u_n \partial_k \varphi \, dx = -\lim_{n\to\infty} \int_\Omega \partial_k u_n \varphi \, dx \\
&= -\lim_{n\to\infty} \langle \varphi, \partial_k u_n\rangle_{L^p} = -\langle \varphi, \partial_k u\rangle_{L^p} \\
&= -\int_\Omega v_k \varphi \, dx,
\end{aligned} \tag{6.3}$$

d. h. $v_k = \partial_k u$ für alle $1 \leq k \leq d$ nach Definition der schwachen Ableitung. Damit ist $\nabla u = v$. □

Aus der Gleichheit des zweiten und des vorletzten Terms in (6.3) folgt, dass formal $\nabla^* = -\operatorname{div}$ gilt. Dies kann durch korrekte Wahl der Bild- und Definitionsbereiche auch rigoros bewiesen werden; siehe z. B. [Kurula & Zwart 2012].

Es ist klar, dass die Seminorm nicht koerziv ist, da $|u|_{W^{1,p}} = 0$ gilt für $u \equiv c$ konstant. Eine Möglichkeit, diesen Fall auszuschliessen, ist die Beschränkung auf Funktionen mit Mittelwert 0; dann folgt die Koerzivität aus der Poincaré-Ungleichung. Sei dafür Ω beschränkt mit Lebesgue-Maß $|\Omega|$ und $u \in L^1(\Omega)$, und setze

$$\mu(u) := \frac{1}{|\Omega|} \int_\Omega u(x) \, dx$$

sowie für alle $1 \leq p \leq \infty$

$$\Pi_0 : L^p(\Omega) \to L^p(\Omega), \qquad \Pi_0(u) = u - \mu(u).$$

Lemma 6.11 (Poincaré–Wirtinger-Ungleichung[9]). *Sei Ω ein konvexes Lipschitz-Gebiet und $1 \leq p < \infty$. Dann existiert ein $c > 0$ mit*

$$\|\Pi_0 u\|_{L^p} \leq c|u|_{W^{1,p}} \quad \text{für alle } u \in W^{1,p}(\Omega).$$

[9] siehe z. B. [Dobrowolski 2010, Satz 6.23]

Damit können wir nun das Gewünschte zeigen.

Satz 6.12. *Sei Ω ein konvexes Lipschitz-Gebiet und $1 < p < \infty$. Dann ist*

$$R : L^p(\Omega) \to \overline{\mathbb{R}}, \qquad R(u) := \begin{cases} \frac{1}{p}|u|^p_{W^{1,p}} & u \in W^{1,p}(\Omega), \\ \infty & \text{sonst} \end{cases}$$

eigentlich, konvex und unterhalbstetig, und für $\{u_n\}_{n\in\mathbb{N}} \subset L^p(\Omega)$ mit $\|\Pi_0 u_n\|_{L^p} \to \infty$ gilt $R(u_n) \to \infty$.

Beweis. Da $R(u) = 0$ für $u = 0 \in W^{1,p}(\Omega)$, ist R eigentlich. Weiter ist die Vektornorm $|\cdot|_p$ konvex auf $\mathbb{R}^d$, und zusammen mit der Konvexität und Monotonie von $t \mapsto t^p$ auf $[0,\infty)$ folgt die Konvexität von R analog zum Beweis von Lemma 6.1.

Für die Unterhalbstetigkeit betrachten wir eine Folge $u_n \rightharpoonup u$ in $L^p(\Omega)$. Gilt $|u_n|_{W^{1,p}} \to \infty$, so ist nichts zu zeigen. Ansonsten ist $\{\nabla u_n\}_{n\in\mathbb{N}}$ beschränkt in $L^p(\Omega)^d$ und hat daher wegen dessen Reflexivität für $1 < p < \infty$ nach dem Satz von Eberlein–Šmulyan eine konvergente Teilfolge $\nabla u_{n_k} \rightharpoonup v \in L^p(\Omega)^d$. Da nach Voraussetzung auch $u_{n_k} \rightharpoonup u$ gilt, folgt aus Lemma 6.10 also $\nabla u_{n_k} \rightharpoonup \nabla u$. Da wir dieses Argument auf jede Teilfolge von $\{u_n\}_{n\in\mathbb{N}}$ anwenden können, gilt dies für die gesamte Folge. Die Unterhalbstetigkeit der Norm in $L^p(\Omega)^d$ ergibt dann

$$\|\nabla u\|_{(L^p)^d} \leq \liminf_{n\to\infty} \|\nabla u_n\|_{(L^p)^d},$$

und aus der Monotonie von $t \mapsto \frac{1}{p}|t|^p$ folgt die Unterhalbstetigkeit von R.

Sei nun $\{u_n\}_{n\in\mathbb{N}} \subset L^p(\Omega)$ mit $\|\Pi_0 u_n\|_{L^p} \to \infty$. Angenommen, es gäbe ein $M > 0$ mit $R(u_n) \leq M$ für unendlich viele $n \in \mathbb{N}$. Nach Definition von R und $W^{1,p}(\Omega)$ muss dann auch $|u_n|_{W^{1,p}} \leq \tilde{M}$ und damit insbesondere $u_n \in W^{1,p}(\Omega)$ für diese n gelten. Aus Lemma 6.11 folgt dann $\|\Pi_0 u_n\|_{L^p} \leq c|u_n|_{W^{1,p}} \leq cM$, ein Widerspruch. □

Wir berechnen schließlich die Richtungsableitungen von R.

Lemma 6.13. *Seien $u, h \in \operatorname{dom} R = W^{1,p}(\Omega)$ für $1 < p < \infty$. Dann gilt*

$$R'(u;h) = \int_\Omega |\nabla u|_p^{p-2} \nabla u \cdot \nabla h \, dx.$$

Beweis. Die Abbildung $t \mapsto \frac{1}{p}t^p$ ist konvex und differenzierbar auf $[0,\infty)$ mit Ableitung $|t|^{p-2}t$. Wie im Beweis von Lemma 3.5 folgt daraus, dass für fast alle $x \in \Omega$ die Abbildung

$$\varphi_x : (0,\infty) \to \mathbb{R}, \qquad \varphi_x(t) := \frac{1}{t}\left(\tfrac{1}{p}|\nabla u(x) + t\nabla h(x)|_p^p - \tfrac{1}{p}|\nabla u(x)|_p^p\right),$$

monoton steigend ist mit $|\varphi_x(t)| \leq |\varphi_x(1)| \leq \frac{1}{p}|\nabla h(x)|^p$ für alle $t \in (0,1]$. Wegen $h \in W^{1,p}(\Omega)$ ist daher $\varphi_x(t) \in L^1(\Omega)$ (als Funktion auf Ω für festes t) für alle $t \in (0,1]$. Nach dem Satz von Lebesgue[10] können wir daher den Grenzübergang $t \to 0^+$ unter dem Integral durchführen und erhalten

$$R'(u;h) = \int_\Omega \lim_{t\to 0^+} \varphi_x(t)\,dx = \int_\Omega |\nabla u|_p^{p-2} \nabla u \cdot \nabla h\,dx. \qquad \square$$

Da diese Richtungsableitung linear in h ist, ist R Gâteaux-differenzierbar, und wir erhalten mit Satz 3.8 eine Charakterisierung des Subdifferentials.

Folgerung 6.14. *Für $u \in W^{1,p}(\Omega)$, $1 < p < \infty$, gilt $\xi \in \partial R(u)$ genau dann, wenn*

$$\int_\Omega \xi\, h\,dx = \int_\Omega |\nabla u|_p^{p-2} \nabla u \cdot \nabla h\,dx \quad \textit{für alle } h \in W^{1,p}(\Omega).$$

Auf der rechten Seite steht die schwache Formulierung eines partiellen Differentialoperators (der für $p \neq 2$ nichtlinear ist; für $p = 2$ erhält man den negativen Laplace-Operator) mit homogenen Neumann-Randbedingungen, siehe z. B. [Attouch, Buttazzo & Michaille 2014, Theorem 6.6.1, Remark 6.6.1].

6.3 FUNKTIONEN MIT BESCHRÄNKTER VARIATION

Zwar lässt $W^{1,p}(\Omega)$ für $p \to 1$ immer stärkere Konzentration von ∇u zu, echte Kanten (d. h. Unstetigkeiten entlang Linien) sind dennoch auch für $p = 1$ nicht zugelassen. Das sieht man bereits an einem einfachen Beispiel: Betrachte für $\Omega = (-1, 1)$ die Funktion definiert durch $u(x) = 1$ für $x \geq 0$ und $u(x) = -1$ für $x < 0$. Falls eine schwache Ableitung v existiert, müsste diese für beliebige $\varphi \in C_0^\infty(\Omega)$ erfüllen

$$\int_{-1}^1 v(x)\varphi(x)\,dx = -\int_{-1}^1 u(x)\varphi'(x)\,dx = \int_{-1}^0 \varphi'(x)\,dx - \int_0^1 \varphi'(x)\,dx = 2\varphi(0).$$

Es gibt aber keine Funktion $v \in L^1(\Omega)$ mit $\int_\Omega v\varphi\,dx = \varphi(0)$ für alle $\varphi \in C_0^\infty(\Omega)$ (denn diese müsste fast überall gleich Null sein). Hingegen gilt dies für $\delta_0 \in \mathcal{M}(\Omega) = C_0(\Omega)^*$; siehe Beispiel 1.3 (iv). Um Kanten zu erhalten, müssen wir also auch Funktionen zulassen, deren schwacher Gradient nur ein Radon-Maß ist.

Analog zur schwachen Ableitung definieren wir daher für $u \in L^1_{\mathrm{loc}}(\Omega)$ die *sehr schwache Ableitung* $D^\alpha u \in \mathcal{M}(\Omega)$ via (6.1). Der entsprechende *sehr schwache Gradient* ist dann ein

[10] siehe z. B. [Dobrowolski 2010, Satz 4.11]

vektorwertiges Maß, d. h. ein Element im Dualraum $\mathcal{M}(\Omega;\mathbb{R}^d)$ des Raums $C_0(\Omega;\mathbb{R}^d)$ der stetigen Funktionen von Ω nach $\mathbb{R}^d$, versehen mit der Norm[11]

$$\|\varphi\|_{(C)^d} := \sup_{x\in\overline{\Omega}} |\varphi(x)|_2.$$

Dann ist $\nabla u := v \in \mathcal{M}(\Omega;\mathbb{R}^d)$ der sehr schwache Gradient von u genau dann, wenn gilt

$$\int_\Omega u \operatorname{div}\varphi\, dx = -\int_\Omega \varphi \cdot dv = -\sum_{i=1}^d \int_\Omega \varphi_i\, dv_i \qquad \text{für alle } \varphi \in C_0^\infty(\Omega;\mathbb{R}^d),$$

und wir können für sehr schwach differenzierbare u wegen der Dichtheit von $C_0^\infty(\Omega)$ in $C_0(\Omega)$ schreiben

$$TV(u) := \|\nabla u\|_{(\mathcal{M})^d} = \sup\left\{\int_\Omega u \operatorname{div}\varphi\, dx : \varphi \in C_0^\infty(\Omega;\mathbb{R}^d) \text{ mit } \|\varphi\|_{(C)^d} \le 1\right\}.$$

Hat u keinen sehr schwachen Gradienten, setzen wir $TV(u) = \infty$. Man nennt $TV(u)$ die *totale Variation* von u.[12] Für die eingangs definierte Funktion ist zum Beispiel

$$TV(u) = \sup_{\|\varphi\|_C\le 1} \int_{-1}^1 u(x)\varphi'(x)\, dx = \sup_{\|\varphi\|_C\le 1} 2\varphi(0) = 2,$$

was genau der Höhe des Sprungs von u in $x = 0$ entspricht. Weiter folgt für $u \in W^{1,1}(\Omega)$ durch partielle Integration und Dichtheit von $C_0^\infty(\Omega)$ in $L^\infty(\Omega)$ bezüglich der punktweisen fast überall (aber nicht gleichmäßigen!) Konvergenz,[13] dass $TV(u) = \|\nabla u\|_{(L^1)^d}$ gilt. Die totale Variation stellt also eine Erweiterung der Seminorm in $W^{1,1}(\Omega)$ dar.

Wir möchten nun TV als Regularisierungsterm verwenden; dafür müssen wir nachweisen, dass TV eigentlich (klar wegen $TV(0) = 0$), konvex und unterhalbstetig ist. Dafür zeigt man zuerst die Abgeschlossenheit des sehr schwachen Gradienten; der Beweis läuft genau wie in Lemma 6.10.

Lemma 6.15. *Sei $\Omega \subset \mathbb{R}^d$ und $1 \le p < \infty$. Dann ist der sehr schwache Gradient*

$$\nabla : \left\{u \in L^p(\Omega) : \nabla u \in \mathcal{M}(\Omega;\mathbb{R}^d)\right\} \to \mathcal{M}(\Omega;\mathbb{R}^d)$$

schwach-∗ abgeschlossen, d. h. $u_n \rightharpoonup u$ in $L^p(\Omega)$ und $\nabla u_n \rightharpoonup^* v$ in $\mathcal{M}(\Omega;\mathbb{R}^d)$ impliziert $v = \nabla u$.

[11] Die Wahl der Vektornorm für $\varphi(x) \in \mathbb{R}^d$ ist für das Folgende unwesentlich, hat aber Auswirkungen auf die geometrische Struktur von Minimierern der unten definierten totalen Variation. Die Euklidische Norm hat den Vorteil der Rotationsinvarianz, d. h. die totale Variation eines Bildes hängt nicht davon ab, ob Kanten mit Koordinatenrichtungen zusammenfallen. Man spricht daher in diesem Fall auch von der *isotropen* totalen Variation.

[12] Im Sinne der Maßtheorie ist $TV(u)$ eigentlich die totale Variation $|\nabla u|(\Omega)$ des Maßes ∇u; die Bezeichnung hat sich in diesem Kontext trotzdem eingebürgert.

[13] siehe z. B. [Brezis 2010, Exercise 4.25]

Analog zu Satz 6.12 (mit dem Satz von Banach–Alaoglu anstelle von Eberlein–Šmulyan) folgt daraus Konvexität und Unterhalbstetigkeit von TV.

Satz 6.16. *Sei $\Omega \subset \mathbb{R}^d$ und $1 \leq p < \infty$. Dann ist $TV : L^p(\Omega) \to \overline{\mathbb{R}}$ eigentlich, konvex und unterhalbstetig.*

Dagegen ist die $W^{1,1}$-Seminorm *nicht* unterhalbstetig in $L^1(\Omega)$: Für die eingangs betrachtete stückweise konstante Funktion $u \in L^\infty(-1,1)$ konvergiert die Folge $\{u_n\}_{n\in\mathbb{N}} \subset L^\infty(-1,1)$, definiert durch

$$u_n(x) := \begin{cases} 1 & x \geq 1/n, \\ nx & |x| < 1/n, \\ -1 & x \leq -1/n, \end{cases}$$

stark (und damit auch schwach) in $L^1(-1,1)$ gegen u und erfüllt

$$|u_n|_{W^{1,1}} = \int_{-1/n}^{1/n} n\,dx = 2,$$

aber $u \notin W^{1,1}(-1,1)$ und daher $|u|_{W^{1,1}} = \infty$. Dies ist eine weitere Begründung für die Notwendigkeit, auf sehr schwache Ableitungen und damit die totale Variation auszuweichen.

Funktionen, deren totale Variation endlich ist, nennt man *Funktionen mit beschränkter Variation*. Der zugehörige Funktionenraum

$$BV(\Omega) := \left\{u \in L^1(\Omega) : TV(u) < \infty\right\}$$

ist ein Banachraum versehen mit der Norm[14]

$$\|u\|_{BV} := \|u\|_{L^1} + TV(u).$$

Weitere Eigenschaften von BV erhält man wie für Sobolevräume über Dichtheitsargumente. Allerdings liegen die Testfunktionen nicht dicht in BV bezüglich der starken Konvergenz (laut Satz 6.6 ist der Abschluss von $C^\infty(\overline{\Omega})$ bezüglich dieser Norm gerade $W^{1,1}(\Omega)$). Wir brauchen dafür einen schwächeren Konvergenzbegriff: Wir sagen, u_n *konvergiert strikt* in BV gegen u, wenn $u_n \to u$ in L^1 und zusätzlich $TV(u_n) \to TV(u)$ gilt.

Satz 6.17 (Dichtheit[15]). *Sei Ω ein beschränktes Gebiet. Dann existiert für alle $u \in BV(\Omega)$ eine Folge $\{\varphi_n\}_{n\in\mathbb{N}} \subset C^\infty(\overline{\Omega})$, die strikt in BV gegen u konvergiert.*

Damit lassen sich jetzt die üblichen Resultate zeigen. Wir setzen für alle Aussagen voraus, dass Ω ein beschränktes Lipschitz-Gebiet ist.

[14] siehe z. B. [Attouch, Buttazzo & Michaille 2014, Theorem 10.1.1]
[15] siehe z. B. [Attouch, Buttazzo & Michaille 2014, Theorem 10.1.2]

Satz 6.18 (Einbettung[16]). *Für alle* $1 \leq p \leq \frac{d}{d-1}$ *ist die Einbettung* $BV(\Omega) \hookrightarrow L^p(\Omega)$ *stetig.*

Satz 6.19 (Fortsetzung[17]). *Es gibt einen beschränkten linearen Operator* $E : BV(\Omega) \to BV(\mathbb{R}^d)$ *mit* $Eu|_\Omega = u$, $Eu|_{\mathbb{R}^d\setminus\Omega} = 0$, *und* $\|\nabla(Eu)\|_{\mathcal{M}(\partial\Omega)} = 0$ *für alle* $u \in BV(\Omega)$.

Satz 6.20 (Poincaré–Wirtinger[18]). *Sei* $1 \leq p \leq d/(d-1)$. *Dann existiert ein* $c > 0$ *so dass für alle* $u \in BV(\Omega)$ *gilt*

$$\|\Pi_0 u\|_{L^p} \leq c\, TV(u).$$

Insbesondere gilt für $\{u_n\}_{n\in\mathbb{N}} \subset L^p(\Omega)$ *mit* $\|\Pi_0 u_n\|_{L^p} \to \infty$ *auch* $TV(u_n) \to \infty$.

Die Untersuchung der Sprungstellen von Funktionen $u \in BV(\Omega)$ und deren Verbindung mit der totalen Variation $TV(u)$ ist Inhalt der *geometrischen Maßtheorie*, deren Behandlung den Rahmen dieser Vorlesung sprengen würde. Hierfür sei auf [Ambrosio, Fusco & Pallara 2000], [Ziemer 1989], [Valkonen 2014] oder [Attouch, Buttazzo & Michaille 2014, Kapitel 10.3] verwiesen; wir betrachten lediglich einige beispielhafte Resultate.

Wir betrachten zuerst die charakteristische Funktion $\mathbb{1}_E$ für ein Lipschitz-Gebiet $E \subset \Omega$. Es gilt nun für alle $\varphi \in C_0^\infty(\Omega; \mathbb{R}^d)$ nach dem Satz von Gauß, dass

$$\int_\Omega \mathbb{1}_E \operatorname{div}\varphi \, dx = \int_E \operatorname{div}\varphi\, dx = \int_{\partial E} \varphi \cdot \nu \, d\mathcal{H}^{d-1},$$

wobei ν die äußere Einheitsnormale an $\partial\Omega$ und $\mathcal{H}^{d-1}$ das $(d-1)$-dimensionale Hausdorff-Maß ist. Also ist $\nabla \mathbb{1}_E = -\nu\mathcal{H}^{d-1}(\partial E \cap \Omega)$. Man kann nun zeigen,[19] dass

$$TV(\mathbb{1}_E) = \mathcal{H}^{d-1}(\partial E \cap \Omega) < \infty$$

und damit $\mathbb{1}_E \in BV(\Omega)$ ist. Die totale Variation der charakteristischen Funktion eines genügend regulären Gebietes E entspricht also seinem *Umfang* (oder *Perimeter*) $\operatorname{Per}(E)$. Für stückweise glatte Funktionen, d. h. $u \in BV(\Omega)$ mit $u_k := u|_{\Omega_k} \in W^{1,1}(\Omega_k)$ für eine disjunkte Zerlegung von Ω in Lipschitzgebiete $\Omega_k \subset \Omega$, zeigt man mit ein wenig mehr Aufwand

$$TV(u) = \sum_k \|\nabla u_k\|_{L^1(\Omega_k)} + \sum_{l\neq k} \int_{\partial\Omega_k \cap \partial\Omega_l} |u_k - u_l| \, d\mathcal{H}^{d-1}.$$

Die totale Variation misst also den schwachen Gradienten in den glatten Bereichen plus den Gesamtsprung über alle Kanten. (Vergleiche auch das Eingangsbeispiel.) Dies zeigt wieder die Nützlichkeit der totalen Variation im Zusammenhang mit Bildern.

Kernstück der geometrischen Maßtheorie ist das folgende Resultat, das eine Funktion $u \in BV(\Omega)$ mit ihren *Niveaumengen* $E_t := \{x \in \Omega : u(x) > t\}$ in Beziehung setzt.

[16] siehe z. B. [Attouch, Buttazzo & Michaille 2014, Theorem 10.1.3]

[17] siehe z. B. [Ambrosio, Fusco & Pallara 2000, Proposition 3.21]

[18] siehe z. B. [Ambrosio, Fusco & Pallara 2000, Theorem 3.44] zusammen mit der Einbettung von $BV(\Omega)$ in $L^p(\Omega)$

[19] siehe z. B. [Ziemer 1989, Remark 5.4.2]

Satz 6.21 (Koflächenformel[20]). *Für $u \in BV(\Omega)$ gilt*

$$TV(u) = \int_{\mathbb{R}} \operatorname{Per}(E_t)\, dt.$$

Dieses Resultat erlaubt, die (analytische) Struktur von BV-Funktionen über die (geometrischen) Eigenschaften ihrer Niveaumengen zu beschreiben; zum Beispiel lässt sich mit Hilfe der Koflächenformel zeigen, dass $u \in BV(\Omega)$ fast überall (bezüglich des Lebesgue-Maßes) differenzierbar ist und dass die Sprungstellen von u Kurven von einer bestimmten Regularität darstellen; siehe z. B. [Ambrosio, Fusco & Pallara 2000, Kapitel 3.7] und [Valkonen 2014]. Diese tiefe Verbindung von analytischen und geometrischen Begriffen macht den Reiz der geometrischen Maßtheorie aus.

Ähnlich subtil ist die Charakterisierung des Subdifferentials $\partial TV \subset BV(\Omega)^*$. Formal lässt sich für $TV(u) = \|\nabla u\|_{\mathcal{M}^d}$ mit der Kettenregel (Satz 3.12) schreiben al:w s

$$\partial TV(u) = \nabla^* \partial(\|\cdot\|_{\mathcal{M}^d})(\nabla u),$$

für $\xi \in \partial TV(u)$ gibt es also ein $\eta \in \partial(\|\cdot\|_{\mathcal{M}^d})(\nabla u)$ mit $\xi = -\operatorname{div}\eta$. Einer mathematisch sauberen Ausführung stehen nun im Wege, dass $\partial(\|\cdot\|_{\mathcal{M}^d}) \subset \mathcal{M}(\Omega;\mathbb{R}^d)^*$ ist – ein Raum, der sehr schwierig zu charakterisieren ist – und weder Definitions- noch Bildbereich von ∇^* offensichtlich sind. Interessierte seien verwiesen auf [Bredies & Lorenz 2011, Seite 346–355]; wir machen uns das Leben leichter und betrachten stattdessen eine Diskretisierung des Problems, für die der Unterschied zwischen Prädual- und Dualraum entfällt und daher die Verfahren aus Abschnitt 5.2 angewendet werden können.

[20] siehe z. B. [Ziemer 1989, Theorem 5.4.4]

7 DISKRETE BILDMODELLE

Für die konkrete Berechnung von Minimierern mit Hilfe der in Teil II vorgestellten Algorithmen muss man eine endlichdimensionale Approximation der zu minimierenden Funktionale betrachten. Hier verwenden wir eine *pixelweise* Diskretisierung, die verträglich ist mit der üblichen Darstellung in der *digitalen* Bildverarbeitung.

Ein kontinuierliches Bild $u : [0,1]^2 \to \mathbb{R}$ wird dabei angenähert durch ein diskretes Bild $u_h \in \mathbb{R}^n$, indem wir das Gebiet $[0,1]^2$ in $n = N^2$ quadratische Teilgebiete – genannt *Pixel* – mit Fläche $h^2 = 1/N^2$ unterteilen, auf denen u_h stückweise konstant ist. Da der Trend zum Breitbild geht, lassen wir auch nichtquadratische Bilder mit $n = NM$ Pixeln (die dann nicht mehr quadratisch sein müssen) und entsprechend $h^2 = 1/(NM)$ zu. Dem kontinuierlichen Grauwert $u(ih, jh)$ entspricht also der diskrete Grauwert $(u_h)_{ij}$ für $1 \leq i \leq N$ und $1 \leq j \leq M$.[1] Um die Notation übersichtlich zu halten, verzichten wir im Rest dieses Kapitels auf den Index h.

Wir versehen den Vektorraum $\mathbb{R}^{NM}$ mit dem Skalarprodukt

$$(u, v) = h^2 \sum_{i=1}^{N} \sum_{j=1}^{M} u_{ij} v_{ij}$$

und der dadurch induzierten Norm

$$\|u\|_2 = \sqrt{(u, u)} = h \left(\sum_{i=1}^{N} \sum_{j=1}^{M} |u_{ij}|^2 \right)^{1/2}.$$

Analog definieren auch die diskreten p-Normen für $p \neq 2$ als

$$\begin{aligned} \|u\|_p &= \left(h^2 \sum_{i=1}^{N} \sum_{j=1}^{M} |u_{ij}|^p \right)^{1/p}, \quad 1 \leq p < \infty, \\ \|u\|_\infty &= \max_{\substack{1 \leq i \leq N \\ 1 \leq j \leq M}} |u_{ij}|, \quad p = \infty. \end{aligned}$$

[1] Fasst man in verbreiteter Matlab-Konvention u_h als Matrix in $\mathbb{R}^{N \times M}$ auf, so zeigt die diskrete x_1-Koordinate nach *unten*.

Durch diese Definition wird sichergestellt, dass die Norm eines konstanten Bildes nicht von der Auflösung abhängt. Für vektorwertige Funktionen (wie etwa den Gradienten eines Bildes) in $\mathbb{R}^{2NM}$ ersetzen wir $|u_{ij}|$ durch die Euklidische Norm

$$|\xi_{ij}|_2 = \sqrt{\xi_{ij} \cdot \xi_{ij}} = \left(\sum_{k=1}^{2} (\xi_{ijk})^2 \right)^{1/2}.$$

Insbesondere erhalten wir das Skalarprodukt von ξ und η in $\mathbb{R}^{2NM}$ (bzw. die duale Paarung in $(\mathbb{R}^{2NM}, \|\cdot\|_p)$ und $(\mathbb{R}^{2NM}, \|\cdot\|_{p'})$)

$$(\xi, \eta) = h^2 \sum_{i=1}^{N} \sum_{j=1}^{M} \sum_{k=1}^{2} \xi_{ijk} \eta_{ijk}.$$

Für die Approximation der Normen in Sobolevräumen oder BV brauchen wir noch eine Diskretisierung des Gradienten. Dafür interpretieren wir den Grauwert eines Pixels (i, j) wieder mit dem Wert $u(ih, jh)$ und approximieren die partiellen Ableitungen $\partial_k u(ih, jh)$ durch Vorwärtsdifferenzienquoten mit konstanter Fortsetzung am Rand, d. h. wir setzen $u(ih, (M + 1)h) = u(ih, Mh)$.[2] Die partiellen Ableitungen, ausgedrückt durch den Vektor $u \in \mathbb{R}^{NM}$, sind dann

$$(\partial_1^{h,+} u)_{ij} = \begin{cases} \frac{u_{i+1,j} - u_{ij}}{h} & i < N, \\ 0 & i = N, \end{cases}$$

$$(\partial_2^{h,+} u)_{ij} = \begin{cases} \frac{u_{i,j+1} - u_{ij}}{h} & j < M, \\ 0 & J = M. \end{cases}$$

Der diskrete Gradient $\nabla_h u \in \mathbb{R}^{2NM}$ eines diskreten Bildes $u \in \mathbb{R}^{NM}$ ist dann komponentenweise gegeben durch

$$(\nabla_h u)_{ijk} = (\partial_k^{h,+} u)_{ij}.$$

Damit kann die diskrete Sobolev-Seminorm approximiert werden durch

$$|u|_{W_h^{1,p}} := \|\nabla_h u\|_p = \left(h^2 \sum_{i=1}^{N} \sum_{j=1}^{M} |(\nabla_h u)_{ij}|_p^p \right)^{1/p}$$

und die (isotrope) totale Variation durch

$$TV_h(u) = h^2 \sum_{i=1}^{N} \sum_{j=1}^{M} |(\nabla_h u)_{ij}|_2 = \| |\nabla_h u|_2 \|_1.$$

[2] Alternativen wären zentrale Differenzenquotienten bzw. periodische oder Null-Fortsetzung.

Aus der Definition ist ersichtlich, dass aus $\nabla_h u = 0$ folgt $u \equiv \lambda$ für ein $\lambda \in \mathbb{R}$, denn damit ist $u_{i-1,j} = u_{ij}$ für alle $i < N$ und $u_{i,j-1} = u_{ij}$ für alle $j < M$. Per Induktion erhalten wir daraus $u_{ij} = u_{NM}$ für alle i, j und damit ist u ein konstanter Vektor.

Da es sich bei dem diskreten Gradienten um einen linearen Operator von $\mathbb{R}^{NM}$ nach $\mathbb{R}^{2NM}$ handelt, gibt es natürlich eine Darstellung als Matrix in $\mathbb{R}^{2NM\times NM}$; es ist jedoch in der Regel weder notwendig noch sinnvoll, diese zu aufzustellen, da für die hier betrachteten Algorithmen nur die Anwendung des Gradienten mit Hilfe der obigen Definition notwendig ist. Für die Berechnung von zulässigen Schrittweiten im primal-dualen Extragradienten-Verfahren brauchen wir dennoch die Operatornorm dieser Matrix.

Lemma 7.1. *Für $\nabla_h : \mathbb{R}^{NM} \to \mathbb{R}^{2NM}$ und die durch $\|\cdot\|_2$ induzierte Operatornorm gilt*

$$\|\nabla_h\|_2 < \frac{\sqrt{8}}{h}.$$

Beweis. Wir zeigen zuerst eine Hilfsabschätzung: Für $\|u\|_2 \leq 1$ ist

$$-h^2 \sum_{i=1}^{N-1} \sum_{j=1}^{M} u_{i+1,j} u_{ij} < 1 \qquad \text{und} \qquad -h^2 \sum_{i=1}^{N} \sum_{j=1}^{M-1} u_{i,j+1} u_{ij} < 1.$$

Führen wir $v_{ij} = -u_{i+1,j}$ für $i < N$ und $v_{Nj} = 0$ ein, so ist die erste Ungleichung äquivalent zu $(v, u) < 1$. Sei nun $u \neq 0$ (sonst ist die Abschätzung trivialerweise erfüllt) und angenommen, die Ungleichung gilt nicht, d. h. $(u, v) \geq 1$. Wegen $\|v\|_2 \leq \|u\|_2 \leq 1$ ist dann

$$1 \leq (u, v) \leq \|u\|_2 \|v\|_2 \leq 1,$$

es gilt also Gleichheit in der Cauchy–Schwarz-Ungleichung. Dies ist aber dann und nur dann der Fall, wenn $v = \lambda u$ für ein $\lambda \in \mathbb{R} \setminus \{0\}$ ist. Nach Konstruktion von v ist gilt daher für beliebige $j \in \{1, \ldots, M\}$

$$0 = |v_{Nj}| = |\lambda u_{Nj}| = |\lambda v_{N-1,j}| = |\lambda^2 u_{N-1,j}| = \cdots = |\lambda^N u_{1j}|,$$

und damit $u = 0$ im Widerspruch zur Annahme. Die zweite Abschätzung zeigt man analog.

Mit dieser Abschätzung erhalten wir für alle $\|u\|_2 \leq 1$, dass gilt

$$\begin{aligned}
\|\nabla_h u\|_2^2 &= h^2 \left(\sum_{i=1}^{N-1} \sum_{j=1}^{M} \left(\frac{u_{i+1,j} - u_{ij}}{h} \right)^2 \right) + h^2 \left(\sum_{i=1}^{N} \sum_{j=1}^{M-1} \left(\frac{u_{i,j+1} - u_{ij}}{h} \right)^2 \right) \\
&= \left(\sum_{i=1}^{N-1} \sum_{j=1}^{M} u_{i+1,j}^2 + u_{ij}^2 - 2u_{i+1,j} u_{ij} \right) + \left(\sum_{i=1}^{N} \sum_{j=1}^{M-1} u_{i,j+1}^2 + u_{ij}^2 - 2u_{i,j+1} u_{ij} \right) \\
&\leq 4 \left(\sum_{i=1}^{N} \sum_{j=1}^{M} u_{ij}^2 \right) - 2 \left(\sum_{i=1}^{N-1} \sum_{j=1}^{M} u_{i+1,j} u_{ij} \right) - 2 \left(\sum_{i=1}^{N} \sum_{j=1}^{M-1} u_{i,j+1} u_{ij} \right) \\
&< \frac{4}{h^2} \|u\|_2^2 + \frac{2}{h^2} + \frac{2}{h^2} \leq \frac{8}{h^2},
\end{aligned}$$

woraus die Aussage folgt. □

Für die primal-dualen Algorithmen brauchen wir auch den adjungierten Operator, die *diskrete Divergenz* $\operatorname{div}_h : \mathbb{R}^{2NM} \to \mathbb{R}^{NM}$. Da es sich um den adjungierten Operator bezüglich des *diskreten* Skalarprodukts handeln muss, können wir nicht einfach den (bezüglich $(\cdot,\cdot)_{L^2}$) adjungierten Operator ∇^* analog zum Gradienten diskretisieren. Stattdessen verwenden wir hier Rückwärtsdifferenzenquotienten und Nullfortsetzung, und definieren für $u \in \mathbb{R}^{NM}$ komponentenweise

$$(\partial_1^{h,-} u)_{ij} = \begin{cases} \frac{u_{1j}}{h} & i = 1, \\ \frac{u_{ij}-u_{i-1,j}}{h} & 1 < i < N, \\ \frac{-u_{N-1,j}}{h} & i = N, \end{cases}$$

$$(\partial_2^{h,-} u)_{ij} = \begin{cases} \frac{u_{i1}}{h} & j = 1, \\ \frac{u_{ij}-u_{i,j-1}}{h} & 1 < j < M, \\ \frac{-u_{i,M-1}}{h} & j = M, \end{cases}$$

sowie für $v \in \mathbb{R}^{2NM}$

$$(\operatorname{div}_h v)_{ij} = \sum_{k=1}^{2} \left(\partial_k^{h,-} v_{\cdot,\cdot,k}\right)_{ij}.$$

(Beachten Sie die Sonderbehandlung der Randterme $i = N$ und $j = M$!)

Lemma 7.2. *Für alle $u \in \mathbb{R}^{NM}$ und $v \in \mathbb{R}^{2NM}$ gilt*

$$(\nabla_h u, v) = (u, -\operatorname{div}_h v).$$

Beweis. Wir verwenden „diskrete partielle Integration“:

$$\begin{aligned}
(\nabla_h u, v) &= h\left(\sum_{i=1}^{N-1}\sum_{j=1}^{M}(u_{i+1,j}-u_{ij})v_{ij,1}\right) + h\left(\sum_{i=1}^{N}\sum_{j=1}^{M-1}(u_{i,j+1}-u_{ij})v_{ij,2}\right) \\
&= h\sum_{j=1}^{M}\left(\left(\sum_{i=2}^{N} u_{ij}v_{i-1,j,1}\right) - \left(\sum_{i=1}^{N-1} u_{ij}v_{ij,1}\right)\right) \\
&\quad + h\sum_{i=1}^{N}\left(\left(\sum_{j=2}^{M} u_{ij}v_{i,j-1,2}\right) - \left(\sum_{j=1}^{M-1} u_{ij}v_{ij,2}\right)\right) \\
&= h\sum_{j=1}^{M}\left(-v_{1j,1}u_{1j} + \left(\sum_{i=2}^{N-1}(v_{i-1,j,1}-v_{ij,1})u_{ij}\right) + v_{N-1,j,1}u_{N,j}\right) \\
&\quad + h\sum_{i=1}^{N}\left(-v_{i1,2}u_{i1} + \left(\sum_{j=2}^{M-1}(v_{i,j-1,2}-v_{ij,2})u_{ij}\right) + v_{i,M-1,2}u_{i,M}\right) \\
&= h^2\sum_{i=1}^{N}\sum_{j=1}^{M}\left(-(\partial_1^{h,-}v_{\cdot,\cdot,1})_{ij} - (\partial_2^{h,-}v_{\cdot,\cdot,2})_{ij}\right)u_{ij} \\
&= (u, -\operatorname{div}_h v).
\end{aligned}$$

□

Teil IV

REKONSTRUKTIONSMODELLE

8 ENTRAUSCHEN

Wir kehren nun zurück zu dem Entrausch-Problem, das uns am Anfang als Motivation diente: Ein gegebenes verrauschtes Bild $f : [0,1]^2 \to \mathbb{R}$ soll additiv zerlegt werden in einen gewünschten Bildinhalt $u^0 : [0,1]^2 \to \mathbb{R}$ und eine unerwünschte Störung $\eta : [0,1]^2 \to \mathbb{R}$. Je nach Struktur der Störung und des erwarteten Bildinhalts wird man nun die Summe geeigneter (Semi-)Normen von η und u^0 minimieren, so dass $f = u^0 + \eta$ gilt. Wir betrachten also

$$\min_{u \in X} F(u) + \alpha R(u)$$

und untersuchen Existenz, Optimalitätsbedingungen und numerische Lösung für drei konkrete Beispiele.

8.1 L²-H¹-ENTRAUSCHEN

Ist die Störung η normalverteilt mit Erwartungswert 0 und Varianz σ, so bietet sich für F die quadrierte L^2-Norm von $f - u$ an. Für R setzen wir zunächst der Einfachheit halber die H^1-Seminorm an, betrachten also für $\Omega = (0,1)^2$, $f \in L^2(\Omega)$, und $\alpha > 0$

$$\min_{u \in L^2(\Omega)} \frac{1}{2}\|f - u\|^2_{L^2(\Omega)} + \frac{\alpha}{2}|u|^2_{H^1(\Omega)}.$$

Wir kommen schnell zur Sache, denn wir sind gut vorbereitet. Für die Existenz wenden wir Satz 3.4 an. Wegen $\operatorname{dom} R = H^1(\Omega) \subset L^2(\Omega) = \operatorname{dom} F$ ist das Funktional eigentlich. Da $u \mapsto f - u$ stetig und $t \mapsto \frac{1}{2}t^2$ stetig und monoton steigend ist, ist F nach Lemma 2.2 schwach unterhalbstetig; und wegen der strikten Konvexität von $\|\cdot\|_X^2$ im Hilbertraum ist F strikt konvex. Die entsprechenden Eigenschaften von αR folgen aus Satz 6.12. Somit ist das gesamte Funktional strikt konvex und unterhalbstetig. Schließlich gilt offensichtlich $F(u_n) \to \infty$ für $\|u_n\|_{L^2} \to \infty$, d. h. das Funktional ist wegen $R \geq 0$ koerziv. Nach Satz 3.4 hat (8.1) also eine eindeutige Lösung $\bar{u} \in \operatorname{dom} R = H^1(\Omega)$.

Nach Satz 3.7 ist $\bar{u}$ genau dann Lösung, wenn $0 \in \partial(F + \alpha R)(\bar{u})$ gilt. Nun ist F Fréchet-differenzierbar mit $\nabla F(u) = u - f$ (Fréchet-Differenzierbarkeit der quadrierten Norm im Hilbertraum und Kettenregel, Satz 2.5). Also ist F insbesondere stetig auf ganz $L^2(\Omega)$, und Anwendung der Summenregel aus Satz 3.11 ergibt

$$f - \bar{u} \in \partial(\alpha R)(\bar{u}).$$

Nach Lemma 6.13 ist also $u \in H^1(\Omega)$ Lösung von (8.1) genau dann, wenn gilt

$$\alpha(\nabla\bar{u}, \nabla v)_{(L^2)^d} + (\bar{u}, v)_{L^2} = (f, v)_{L^2} \qquad \text{für alle } v \in H^1(\Omega).$$

Ist nun mit $f_h \in \mathbb{R}^{NM}$ eine Pixelbild, so erhalten wir für das diskretisierte Problem die Optimalitätsbedingung

$$(f_h, v_h) = \alpha(\nabla_h u_h, \nabla_h v_h) + (u_h, v_h) = (\alpha(-\operatorname{div}_h)\nabla_h u_h + u_h, v_h) \quad \text{für alle } v_h \in \mathbb{R}^{NM},$$

d. h. (nach Kürzen von h^2 auf beiden Seiten) $\bar{u}_h$ als Lösung des linearen Gleichungssystems

$$K_h u_h = f_h$$

mit der Matrix $K_h := -\alpha \operatorname{div}_h \nabla_h + \operatorname{Id} \in \mathbb{R}^{NM\times NM}$, zum Beispiel mit dem CG-Verfahren.[1]

Ganz analog geht man für $F(u) = \frac{1}{q}\|f-u\|_{L^q}^q$ und $R(u) = \frac{1}{p}|u|_{W^{1,p}}^p$ für $1 < q \leq p < \infty$ vor (wobei dann die Optimalitätsbedingungen auf eine nichtlineare Gleichung führen).

8.2 L^2-TV-ENTRAUSCHEN

Da L^2-H^1-Entrauschen Kanten nicht erhalten kann, ist es für Bilder in der Regel nicht geeignet. Mittlerweile hat sich daher das L^2-TV-Entrauschen[2] durchgesetzt. Hier betrachtet man

$$\min_{u\in L^2(\Omega)} \frac{1}{2}\|f-u\|_{L^2(\Omega)}^2 + \alpha TV(u).$$

Auch hier sind wir dank unserer Vorarbeiten schnell am Ziel. Wie zuvor ist der Diskrepanzterm F eigentlich, strikt konvex, unterhalbstetig, und koerziv auf L^2. Ebenso ist $R = \alpha TV$ nach Satz 6.16 eigentlich, konvex und unterhalbstetig mit $\operatorname{dom} TV = BV(\Omega) \hookrightarrow L^2(\Omega) = \operatorname{dom} F$ wegen Satz 6.18 und $d = 2$. Auch dieses Funktional ist wegen $TV(u) \geq 0$ koerziv, und aus Satz 3.4 folgt daher wieder die Existenz eines eindeutigen Minimierers $\bar{u} \in \operatorname{dom} TV = BV(\Omega)$.

Aus Satz 3.7 und der Summenregel erhalten wir wie zuvor die Optimalitätsbedingungen

$$f - \bar{u} \in \partial(\alpha TV)(\bar{u}),$$

die wir mangels einer expliziten Darstellung des Subdifferentials nicht weiter ausreizen können. Für die Diskretisierung

$$\min_{u_h\in\mathbb{R}^{NM}} \frac{1}{2}\|u_h - f_h\|_2^2 + \alpha\||\nabla_h u_h|_2\|_1$$

[1] Dies entspricht der Finite-Differenzen-Approximation von $-\alpha\Delta u + u = f$ mit homogenen Neumann-Randbedingungen.

[2] nach den Autoren, die diesen Zugang in [Rudin, Osher & Fatemi 1992] vorgeschlagen haben, auch als *ROF-Modell* bekannt

können wir jedoch den Satz von Fenchel–Rockafellar für $A = \nabla_h$ sowie Lemma 3.18 anwenden und erhalten die Existenz eines $\bar{y}_h \in \mathbb{R}^{2NM}$ mit

$$\begin{cases} \operatorname{div}_h \bar{y}_h = \bar{u}_h - f_h, \\ \quad \nabla_h \bar{u}_h \in \partial\delta_{\{v_h : \| |v_h|_2 \|_\infty \le \alpha\}}(\bar{y}_h), \end{cases}$$

wobei wir die nicht offensichtliche, aber direkt nachprüfbare, Tatsache verwendet haben, dass $(\mathbb{R}^{2NM}, \| |\cdot|_q \|_p)^* = (\mathbb{R}^{2NM}, \| |\cdot|_{q'} \|_{p'})$ für $1 \le p, p' \le \infty$ mit $\frac{1}{p} + \frac{1}{p'} = 1$ ist.

Auf (8.2) können wir nun das primal-duale Extragradienten-Verfahren (5.14) anwenden. Unter Verwendung von Beispiel 4.12 (i) und Lemma 4.10 (i) erhält man für $F_h(u_h) = \frac{1}{2}\|u_h - f_h\|_2^2$ komponentenweise

$$[\operatorname{prox}_{\tau F_h}(v_h)]_{ij} = \frac{1}{1+\tau}(v_{ij} + \tau f_{ij}).$$

Für $(\alpha G_h)^*$ mit $G_h(u_h) = \| |u_h|_2 \|_1$ entspricht die Proximalpunkt-Abbildung der Projektion auf $\{v_h : \| |v_h|_2 \|_\infty \le \alpha\}$, ist also – analog zu Beispiel 4.12 (iii) – komponentenweise gegeben durch

$$[\operatorname{prox}_{\sigma(\alpha G_h)^*}(v_h)]_{ijk} = \frac{\alpha v_{ijk}}{\max\{\alpha, |v_{ij}|_2\}}.$$

Das komplette Verfahren ist im folgenden Algorithmus zusammengefasst, der nach Satz 5.6 für $\sigma\tau < \|\nabla_h\|^{-2} = \frac{h^2}{8}$ (Lemma 7.1) gegen eine Lösung von (8.2) konvergiert.[3]

Algorithmus 8.1 : Primal-duales L^2-TV-Entrauschen

Input : $f \in \mathbb{R}^{NM}$, $u^0 \in \mathbb{R}^{NM}$, $y^0 \in \mathbb{R}^{2NM}$, $\sigma\tau < h^2/8$

for $k = 1 \dots$ **do**
- $u^k = (u^{k-1} + \tau(f + \operatorname{div}_h y^{k-1}))/(1+\tau)$
- $\bar{u}^k = 2u^k - u^{k-1}$
- $w = y^{k-1} + \sigma\nabla_h \bar{u}^k$
- $y^k = \alpha w / \max\{\alpha, |w|_2\}$

Dabei sollten ∇_h und div_h als Prozeduren implementiert werden, die für gegebene $\bar{x}^k$ und y^{k-1} die entsprechenden Vektoren komponentenweise durch Differenzenquotienten ausrechnen.

8.3 L^1-TV-ENTRAUSCHEN

In der Praxis sind nicht alle Störungen normalverteilt: In der digitalen Bildgebung tritt oft *impulsives Rauschen* auf, bei dem nur einzelne Pixel selektiv gestört sind. Zum Beispiel

[3] Da der Algorithmus für Bilder in $BV(\Omega)$ nicht wohldefiniert ist, überrascht es nicht, dass es im Grenzübergang $h \to 0$ ebenfalls versagt.

können einzelne Datenleitungen defekt sein und nur Rauschen übermitteln. Ein Modell für solche Störungen η ist punktweise gegeben durch

$$\eta(x) = \begin{cases} \xi & \text{mit Wahrscheinlichkeit } r, \\ 0 & \text{mit Wahrscheinlichkeit } 1-r, \end{cases}$$

wobei $r \in [0,1]$ der Anteil der defekten Leitungen ist und für jede betroffene Leitung ξ unabhängig normalverteilt ist mit Mittelwert 0 und Varianz σ. Noch extremer ist *Salz-und-Pfeffer-Rauschen*, bei dem die Daten in den betroffenen Pixeln durch 1 (bzw. den Maximalwert) oder 0 ersetzt werden (z. B. wenn CCD-Sensoren komplett ausfallen oder durch kosmische Strahlung gesättigt werden). Solch eine Störung ist nicht mehr additiv; für gegebenes $u^0 : [0,1]^2 \to [0,1]$ ist das verrauschte Bild punktweise gegeben durch

$$f(x) = \begin{cases} 0 & \text{mit Wahrscheinlichkeit } r/2, \\ 1 & \text{mit Wahrscheinlichkeit } r/2, \\ u^0(x) & \text{mit Wahrscheinlichkeit } 1-r. \end{cases}$$

In beiden Fällen ist die Störung charakterisiert durch *Ausreisser*, d. h. große Störungen sind wahrscheinlicher, als sie es bei normalverteiltem Rauschen wären. Hier kann man mit statistischen Überlegungen begründen, dass die L^1-Norm als Diskrepanzterm robuster ist als die L^2-Norm. Wir betrachten daher für $f \in L^1(\Omega)$ das Problem[4]

$$\min_{u\in L^1(\Omega)} \|f-u\|_{L^1(\Omega)} + \alpha TV(u).$$

Ganz analog wie für L^2-TV-Entrauschen liefert die direkte Methode der Variationsrechnung die Existenz eines Minimierers $\bar u \in BV(\Omega) \subset L^1(\Omega)$; allerdings ist nun weder $\|f-u\|_{L^1}$ noch $TV(u)$ strikt konvex, so dass die Lösung nicht mehr eindeutig sein muss.

Hier sind beide Summanden nicht-differenzierbar; da F stetig auf $L^1(\Omega)$ ist, liefert aber die Summenregel und Lemma 3.10 die Existenz eines $\bar p \in L^\infty(\Omega)$ mit

$$\begin{cases} \bar p \in \operatorname{sign}(\bar u - f), \\ -\bar p \in \partial(\alpha TV)(\bar u). \end{cases}$$

Im Diskreten kann man stattdessen wieder den Satz von Fenchel–Rockafellar anwenden und erhält ein $\bar y_h \in \mathbb{R}^{2NM}$ mit

$$\begin{cases} \operatorname{div}_h \bar y_h \in \operatorname{sign}(\bar u_h - f_h), \\ \nabla_h \bar u_h \in \partial\delta_{\{v_h:\||v_h|_2\|_\infty\le\alpha\}}(\bar y_h). \end{cases}$$

[4] zuerst untersucht in [Chan & Esedoglu 2005], weshalb es manchmal auch *Chan–Esedoğlu-Modell* genannt wird

In diesem Fall ist nach Beispiel 4.12 (ii) und Lemma 4.10 (i) die Proximalpunkt-Abbildung für $F_h(u_h) = \|u_h - f_h\|_1$ komponentenweise gegeben durch

$$[\operatorname{prox}_{\tau F_h}(v_h)]_{ij} = (|v_{ij} - f_{ij}| - \tau)^+ \operatorname{sign}(v_{ij} - f_{ij}) + f_{ij} = \begin{cases} f_{ij} & |v_{ij} - f_{ij}| \leq \tau, \\ v_{ij} - \tau & v_{ij} - f_{ij} > \tau, \\ v_{ij} + \tau & v_{ij} - f_{ij} < -\tau. \end{cases}$$

(Der Proximalpunkt-Schritt wirkt also als „Ausreisser-Detektor“: Ist das Residuum $v_h - f_h$ in einem Pixel kleiner als τ, so werden dort die Daten als exakt angesehen; ansonsten liegt ein Ausreisser vor und die Daten werden verworfen.) Das primal-duale Extragradienten-Verfahren lautet nun wie folgt.

Algorithmus 8.2 : Primal-duales L^1-TV-Entrauschen

Input : $f \in \mathbb{R}^{NM}$, $u^0 \in \mathbb{R}^{NM}$, $y^0 \in \mathbb{R}^{2NM}$, $\sigma\tau < h^2/8$

for $k = 1\dots$ **do**

- $r = u^{k-1} - f + \tau \operatorname{div}_h y^{k-1}$
- $u^k = (|r| - \tau)^+ \operatorname{sign}(r) + f$
- $\bar{u}^k = 2u^k - u^{k-1}$
- $w = y^{k-1} + \sigma \nabla_h \bar{u}^k$
- $y^k = \alpha w / \max\{\alpha, |w|_2\}$

9 ENTFALTEN

Die zweite Standard-Aufgabe in der Bildverarbeitung ist das *Schärfen* von Bildern. Das optische System einer Kamera bildet jeden Punkt in der Objektebene auf einen Punkt in der Bildebene (in der sich der Bildsensor befindet) ab. Punkte, die sich näher zu oder weiter weg von der Kamera befinden, werden dagegen auf einen Kreis abgebildet, dessen Radius von der Entfernung des Punktes von der Objektebene abhängt. Umgekehrt überlagern sich an jedem Bildpunkt alle Kreise von Punkten, die innerhalb dieses Radius liegen. Dies lässt sich mathematisch wie folgt modellieren: Angenommen, wir sind interessiert an einem Bild u^0 eines Gegenstandes, der sich in einer Ebene mit festen Abstand zur Objektebene befindet, und alle Punkte in dieser Ebene werden auf einen Kreis mit Radius r in der Bildebene abgebildet. Definiere für jeden Punkt x in der Bildebene den Kreis

$$B_r(x) := \left\{ y \in \mathbb{R}^2 : |x - y|_2 < r \right\}.$$

Das von einem Sensor aufgenommene Bild f ist dann punktweise gegeben durch

$$\begin{aligned} f(x) = \int_{B_r(x)} u^0(y)\,dy = \int_{\mathbb{R}^2} u^0(y)\mathbb{1}_{B_r(x)}(y)\,dy &= \int_{\mathbb{R}^2} u^0(y)\mathbb{1}_{B_r(0)}(x-y)\,dy \\ &= \int_{\mathbb{R}^2} u^0(x-y)\mathbb{1}_{B_r(0)}(y)\,dy, \end{aligned}$$

wobei wir verwendet haben, dass $y \in B_r(x)$ genau dann gilt, wenn $x - y \in B_r(0)$ ist, sowie im letzten Schritt die Variablentransformation $y \mapsto x - y$. Solch eine Operation bezeichnet man als *Faltung* (von u^0 mit dem *Faltungskern* $\mathbb{1}_{B_r}$); das nachträgliche Schärfen von Bildern entspricht also einer *Entfaltung*. Andere Faltungskerne tauchen unter dem Namen *Punktantwort* (englisch: *point spread function*) in der optischen Mikroskopie auf; dort kann man durch Entfalten die effektive Auflösung erhöhen.

Um für solche Probleme einen Diskrepanzterm zu formulieren, ist zunächst etwas Notation notwendig. Sei $\Omega \subset \mathbb{R}^d$ das Gebiet, auf dem das gesuchte Bild u^0 definiert ist. Weiterhin sei $\omega \subset \mathbb{R}^d$ das Gebiet, auf dem der Faltungskern k definiert ist (wobei wir ohne Beschränkung annehmen können, dass $\operatorname{supp} k = \omega$ ist). Schließlich sei $\Omega' \subset \mathbb{R}^d$ das Gebiet, auf dem das gefaltete Bild f definiert ist. Damit die Faltung für alle $x \in \Omega'$ wohldefiniert ist, nehmen wir in Folge stets an, dass gilt

$$\Omega' - \omega = \left\{ x - z \in \mathbb{R}^d : x \in \Omega', z \in \omega \right\} \subset \Omega.$$

Die Faltung von u mit k, geschrieben $k \star u$, ist dann definiert für alle $x \in \Omega'$ durch

$$(k \star u)(x) := \int_\omega k(y)u(x-y)\,dy.$$

Verwenden wir die Substitution $z := x - y \in \Omega' - \omega \subset \Omega$ und setzen $k'(z) := k(x-z)$ auf $\Omega \setminus (\Omega' - \omega)$ durch Null fort, so können wir auch schreiben

$$(k \star u)(x) = \int_\Omega k(x-z)u(z)\,dz = (u \star k)(x),$$

d. h. die Faltung ist symmetrisch.

Aus der Definition ist weiterhin sofort ersichtlich, dass durch $u \mapsto k \star u$ eine lineare Abbildung definiert wird; diese ist sogar stetig.

Lemma 9.1. *Für $k \in L^1(\omega)$ und $u \in L^p(\Omega)$ mit $1 \le p \le \infty$ ist $k \star u \in L^p(\Omega')$, und es gilt*

$$\|k \star u\|_{L^p(\Omega')} \le \|k\|_{L^1(\omega)} \|u\|_{L^p(\Omega)}.$$

Beweis. Für $p = \infty$ können wir das Supremum über $|u(x)|$ vor das Integral ziehen und erhalten die gewünschte Abschätzung. Für $p = 1$ gilt nach dem Satz von Fubini (denn der Integrand ist nicht-negativ) und wegen $\Omega' - \{y\} \subset \Omega$ für alle $y \in \omega$ nach Annahme

$$\begin{aligned}\int_{\Omega'} |(k \star u)(x)|\,dx &= \int_\omega \int_{\Omega'} |u(x-y)|\,dx\,|k(y)|\,dy \\ &\le \int_\omega \int_\Omega |u(z)|\,dz\,|k(y)|\,dy \\ &= \|k\|_{L^1(\omega)} \|u\|_{L^1(\Omega)}.\end{aligned}$$

Insbesondere ist damit $k \star u \in L^1(\Omega')$.

Für $1 < p < \infty$ wenden wir zuerst die Höldersche Ungleichung an (mit $1/p + 1/p' = 1$) und erhalten wie zuvor

$$\begin{aligned}\int_{\Omega'} |(k \star u)(x)|^p\,dx &\le \int_{\Omega'} \left(\int_\omega |u(x-y)|^p |k(y)|\,dy\right)\left(\int_\omega |k(y)|\,dy\right)^{p/p'} dx \\ &= \int_\omega \int_{\Omega'} |u(x-y)|^p\,dx\,|k(y)|\,dy\, \|k\|_{L^1(\omega)}^{p/p'} \\ &\le \|u\|_{L^p(\Omega)}^p \|k\|_{L^1(\omega)}^{p/p'+1},\end{aligned}$$

woraus wegen $p/p' + 1 = p$ nach Ziehen der p-ten Wurzel die Aussage folgt. □

Aus der Symmetrie der Faltung folgt also, dass $k \star u$ mindestens so glatt ist wie die glattere der beiden Funktionen. Ist insbesondere $k \in C^1(\bar\omega)$, so ist auch $k \star u \in C^1(\bar\Omega')$.

Für einen gegebenen Faltungskern $k \in L^1(\omega)$ wird also durch

$$K : L^p(\Omega) \to L^p(\Omega'), \qquad u \mapsto k \star u,$$

ein linearer und beschränkter Operator definiert, dessen Norm gleich $\|k\|_{L^1(\omega)}$ ist. Der zugehörige adjungierte Operator ist wieder eine Faltung.

Lemma 9.2. *Für $K : L^p(\Omega) \to L^p(\Omega')$, $u \mapsto k \star u$, ist $K^* : L^{p'}(\Omega') \to L^{p'}(\Omega)$ gegeben durch $K^*u = \bar k \star u$ für $\bar k(y) = k(-y)$ für alle $y \in \omega$.*

Beweis. Für $w \in L^{p'}(\Omega')$ gilt nach Fubini

$$\begin{aligned}\langle w, k \star u\rangle_{L^p(\Omega')} &= \langle w, u \star k\rangle_{L^p(\Omega')}\\ &= \int_{\Omega'} w(x) \int_\Omega k(x-y)u(y)\,dy\,dx\\ &= \int_\Omega u(y) \int_{\Omega'} \bar k(y-x)w(x)\,dx\,dy\\ &= \langle u, \bar k \star w\rangle_{L^{p'}(\Omega)},\end{aligned}$$

wobei wieder k für $x - y \notin \omega$ durch Null fortgesetzt sein soll. □

Für ein gegebenes unscharfes Bild $f \in L^p(\Omega')$ mit $1 \le p < \infty$ und Faltungskern $k \in L^1(\omega)$ betrachten wir nun den Diskrepanzterm

$$F : L^p(\Omega) \to \mathbb{R}, \qquad F(u) = \frac{1}{p}\|k \star u - f\|^p_{L^p(\Omega')}.$$

Da $u \mapsto k \star u$ ein linearer beschränkter Operator ist, ist F nach Lemmata 2.2 und 3.3 eigentlich, konvex und unterhalbstetig. Für die Anwendung der direkten Methode benötigen wir nur noch die Koerzivität, wofür bei Regularisierung mit Seminormen ausreicht, diese Eigenschaft nur für konstante Funktionen nachzuweisen. Ist $u \equiv c$ konstant, so ist

$$(k \star u)(x) = \int_\omega k(y)u(x-y)\,dy = c\int_\omega k(y)\,dy,$$

und damit gilt $F(u) \to \infty$ für $c \to \infty$, solange $\int_\omega k(y)\,dy \neq 0$ ist.

Nun haben wir alles zur Hand, um die Existenz von Lösungen von Entfaltungsproblemen zu beweisen. Wir betrachten exemplarisch den Fall $p = 2$ mit der totalen Variation als Regularisierungsterm, d. h. das Problem

$$\min_{u \in L^2} \frac{1}{2}\|k \star u - f\|^2_{L^2} + \alpha TV(u). \tag{9.1}$$

Satz 9.3. *Für $k \in L^1(\omega)$ mit $\int_\omega k(y)\,dy \neq 0$ hat das Problem (9.1) eine Lösung $\bar u \in BV(\Omega)$. Ist $u \mapsto k \star u$ injektiv, so ist die Lösung eindeutig.*

Beweis. Wir wissen bereits, dass das Funktional eigentlich, konvex und unterhalbstetig ist, es bleibt also nur noch die Koerzivität zu zeigen. Sei dafür $\{u_n\}_{n\in\mathbb{N}} \subset L^2(\Omega)$ eine Folge mit $\|u_n\|_{L^2(\Omega)} \to \infty$. Wir zerlegen nun $u_n \in L^2(\Omega)$ in einen Anteil mit Mittelwert 0 und einen konstanten Anteil (der gleich dem Mittelwert ist), d. h. $u_n = \Pi_0 u_n + \mu(u_n)$. Die Dreiecksungleichung ergibt dann

$$\|\Pi_0 u_n\|_{L^2(\Omega)} + \|\mu(u_n)\|_{L^2(\Omega)} \geq \|u_n\|_{L^2(\Omega)} \to \infty.$$

Also muss wenigstens einer der beiden Terme auf der linken Seite auch unbeschränkt sein. Ist der erste Term unbeschränkt, so existiert eine Teilfolge von $\{\Pi_0 u_n\}_{n\in\mathbb{N}}$ mit $\|\Pi_0 u_{n_m}\|_{L^2(\Omega)} \to \infty$. Wegen $1 \leq 2 = d/(d-1)$ können wir die Poincaré–Wirtinger-Ungleichung (Satz 6.20) anwenden und erhalten mit $\nabla\mu(u) = 0$

$$TV(u_{n_m}) = TV(\Pi_0(u_{n_m})) \geq c^{-1}\|\Pi_0 u_{n_m}\|_{L^2(\Omega)} \to \infty.$$

Ansonsten muss der zweite Term unbeschränkt sein. Wir finden also wieder eine Teilfolge mit $\|\mu(u_{n_m})\|_{L^2(\Omega)} \to \infty$. Da $\mu(u_{n_m})$ konstant ist, gilt $k \star \mu(u_{n_m}) = C(k)\mu(u_{n_m})$ mit $C(k) = \int_\omega k(y)\,dy$. Aus der Linearität der Faltung folgt dann mit Hilfe der umgekehrten Dreiecksungleichung und Lemma 9.1

$$\|k \star u_{n_m} - f\|_{L^2(\Omega')} \geq |C(k)|\|\mu(u_{n_m})\|_{L^2(\Omega')} - \|k\|_{L^1(\omega)}\|\Pi_0(u_{n_m})\|_{L^2(\Omega)} - \|f\|_{L^2(\Omega')} \to \infty,$$

da der zweite Term auf der rechten Seite nach Annahme beschränkt ist.

Da wir diese Argumention auf jede Teilfolge von $\{u_n\}_{n\in\mathbb{N}}$ anwenden können, muss also

$$\|k \star u_n - f\|_{L^2(\Omega')} + \alpha TV(u_n) \to \infty$$

für die gesamte Folge gelten. Daraus folgt die Koerzivität des Funktionals und nach Satz 3.4 die Existenz einer Lösung $\bar u \in BV(\Omega)$.

Ist die Faltung injektiv, so ist für $u_1 \neq u_2$ auch $k \star u_1 \neq k \star u_2$. Aus der strikten Konvexität von $v \mapsto \frac{1}{2}\|v\|^2_{L^2(\Omega')}$ folgt dann für alle $\lambda \in (0,1)$

$$\begin{aligned}\frac{1}{2}\|k \star (\lambda u_1 + (1-\lambda)u_2) - f\|^2_{L^2(\Omega')} &= \frac{1}{2}\|\lambda(k \star u_1 - f) + (1-\lambda)(k \star u_2 - f)\|^2_{L^2(\Omega')} \\ &< \frac{\lambda}{2}\|k \star u_1 - f\|^2_{L^2(\Omega')} + \frac{(1-\lambda)}{2}\|k \star u_2 - f\|^2_{L^2(\Omega')}.\end{aligned}$$

Der Diskrepanzterm und damit das gesamte Funktional ist also strikt konvex, was die Eindeutigkeit des Minimierers garantiert. □

Aus der Linearität von K und der Fréchet-Differenzierbarkeit von $v \mapsto \frac{1}{2}\|v-f\|^2_{L^2(\Omega')}$ folgt mit der Kettenregel aus Satz 2.5 auch die Fréchet-Differenzierbarkeit von F mit Gradient

$$\nabla F(u) = K^*(Ku-f) = \bar{k} \star (k \star u - f).$$

Da F stetig ist auf $L^2(\Omega)$, folgt aus der Summenregel (Satz 3.11), dass $\bar{u} \in BV(\Omega)$ Lösung von (9.1) ist genau dann, wenn gilt

$$\bar{k} \star (f - k \star \bar{u}) \in \partial(\alpha TV(\bar{u})).$$

Für die numerische Lösung müssen wir zuerst das diskretisierte Problem formulieren, wofür wir eine diskrete Faltung brauchen. Auch hier müssen wir aufpassen, dass die Definitionsbereiche der verschiedenen diskreten Bilder bzw. Faltungskerne zusammenpassen. Sei wieder $u_h \in \mathbb{R}^{NM}$ ein diskretes Bild, und $k_h \in \mathbb{R}^{KL}$ ein diskreter Faltungskern. In der Praxis sind üblicherweise K und L ungerade, d. h. $K = 2r+1$ und $L = 2s+1$, und die Elemente von k_h symmetrisch indiziert: $k_h = (k_{ij})_{i=-r,\dots,r,j=-s,\dots,s}$. Die *diskrete Faltung* $k_h \star u_h \in \mathbb{R}^{(N-2r)(M-2s)}$ (wir schenken uns die Notation $\star_h$) ist dann definiert durch

$$[k_h \star u_h]_{ij} = h \sum_{n=-r}^{r} \sum_{m=-s}^{s} k_{nm} u_{i-n,j-m} = h \sum_{n=i-r}^{i+r} \sum_{m=j-s}^{j+s} u_{nm} k_{i-n,j-m}.$$

Für festes k_h wird dadurch ein linearer Operator $K_h : \mathbb{R}^{NM} \to \mathbb{R}^{(N-2r)(M-2s)}$ mit Norm $\|K_h\| = \|k_h\|_1$ definiert. Analog zur diskreten Divergenz rechnet man nun nach (durch Vertauschung der Summation und Indexverschiebung), dass $K_h^* : \mathbb{R}^{(N-2r)(M-2s)} \to \mathbb{R}^{NM}$ gegeben ist durch

$$[K_h^* w_h]_{ij} = h \sum_{n=i-r}^{i+r} \sum_{m=j-s}^{j+s} w_{nm} k_{n-i,m-j} = \bar{k}_h \star w_h,$$

d. h. durch die diskrete Faltung mit dem gespiegelten Kern $[\bar{k}_h]_{ij} = [k_h]_{-i,-j}$, wobei auch hier w_h durch Null fortgesetzt wird.

Das diskrete Problem ist also

$$\min_{u_h \in \mathbb{R}^{NM}} \frac{1}{2}\|K_h u_h - f_h\|_2^2 + \alpha \| |\nabla_h u_h|_2 \|_1.$$

Hier kommen nun in beiden Termen lineare Operatoren vor. Um zu vermeiden, für den ersten Term eine Proximalabbildung ausrechnen zu müssen (in der K_h^{-1} auftauchen würde, und K_h muss ja genau wie K nicht injektiv sein), wenden wir den Satz von Fenchel–Rockafellar an auf

$$\begin{aligned}
&F_h : \mathbb{R}^{NM} \to \mathbb{R}, && F_h(u_h) = 0,\\
&G_h : (\mathbb{R}^{(N-2r)(M-2s)} \times \mathbb{R}^{2NM}) \to \mathbb{R}, && G_h(y_h, z_h) = \frac{1}{2}\|y_h - f_h\|_2^2 + \alpha \| |z_h|_2 \|_1,\\
&A_h : \mathbb{R}^{NM} \to (\mathbb{R}^{(N-2r)(M-2s)} \times \mathbb{R}^{2NM}), && A_h u_h = \begin{pmatrix} K_h u_h \\ \nabla_h u_h \end{pmatrix}.
\end{aligned}$$

Nun ist A_h ein beschränkter linearer Operator mit Norm $\|A_h\|^2 \le \|k_h\|_1^2 + \frac{8}{h^2}$ und Adjungierter

$$A_h^* : (\mathbb{R}^{(N-2r)(M-2s)} \times \mathbb{R}^{2NM}) \to \mathbb{R}^{NM}, \qquad A_h^*(y_h, z_h) = K_h^* y_h - \operatorname{div}_h z_h.$$

Da G stetig ist in $A_h 0$, erhalten wir die Existenz von $(\bar{y}_h, \bar{z}_h) \in (\mathbb{R}^{(N-2r)(M-2s)} \times \mathbb{R}^{2NM})$ mit

$$\begin{cases} -A_h^*(\bar{y}_h, \bar{z}_h) \in \partial F_h(\bar{u}_h), \\ \qquad A_h \bar{u}_h \in \partial G_h^*(\bar{y}_h, \bar{z}_h). \end{cases}$$

Um darauf das primal-duale Extragradienten-Verfahren anwenden zu können, brauchen wir nur noch die entsprechenden Proximalpunktabbildungen. Für $F_h \equiv 0$ haben wir $\partial F_h(u_h) = \{0\}$ und damit

$$\operatorname{prox}_{\tau F_h}(u_h) = (\operatorname{Id} + \tau \partial F_h)^{-1}(u_h) = (\operatorname{Id})^{-1} u_h = u_h.$$

Weiter ist

$$\begin{aligned} &G_h^* : (\mathbb{R}^{(N-2r)(M-2s)} \times \mathbb{R}^{2NM}) \to \overline{\mathbb{R}}, \\ &G_h^*(y_h^*, z_h^*) = \frac{1}{2}\|y_h^*\|_2^2 + (y_h^*, f_h) + \delta_{\{v_h : \||v_h|_2\|_\infty \le \alpha\}}(z_h^*), \end{aligned}$$

da das Supremum in y_h und z_h separat berechnet werden kann. Auch die zugehörige Proximalpunktabbildung können wir nach Lemma 4.10 (iii) separat bezüglich jedem Argument berechnen. Unter Verwendung der Rechnung in Kapitel 8.2 folgt damit

$$\operatorname{prox}_{\sigma G^*}(y_h, z_h) = \begin{pmatrix} \frac{1}{1+\sigma}(y_h - \sigma f_h) \\ \dfrac{\alpha z_h}{\max\{\alpha, |z_h|_2\}} \end{pmatrix}.$$

Damit können wir das Verfahren (5.14) für das Entfaltungsproblem (und analog für beliebige lineare Operatoren K_h) als konkreten Algorithmus formulieren:

Algorithmus 9.1 : Primal-duales L^2-TV-Entfalten

Input : $f \in \mathbb{R}^{NM}$, $k \in \mathbb{R}^{(2r+1)(2s+1)}$, $u^0 \in \mathbb{R}^{NM}$, $y^0 \in \mathbb{R}^{(N-2r)(M-2s)}$, $z^0 \in \mathbb{R}^{2NM}$, $\sigma\tau < (\|k_h\|_1^2 + h^2/8)^{-1}$

for $k = 1 \dots$ **do**
- $u^k = u^{k-1} + \tau(\operatorname{div}_h z^{k-1} - K_h^* y^{k-1})$
- $\bar{u}^k = 2u^k - u^{k-1}$
- $y^k = (y^{k-1} + \sigma(K_h \bar{u}^k - f))/(1+\sigma)$
- $w = z^{k-1} + \sigma \nabla_h \bar{u}^k$
- $z^k = \alpha w / (\max\{\alpha, |w|_2\})$

Natürlich gibt es nichts geschenkt: Durch diesen Trick wird die zulässige Wahl der Schrittweiten $\sigma\tau < (\|k_h\|_1^2 + 8/h^2)^{-1} < h^2/8$ weiter eingeschränkt.

LITERATUR

R. A. Adams & J. J. F. Fournier (2003), *Sobolev Spaces*, 2. Aufl., Academic Press, Amsterdam.

L. Ambrosio, N. Fusco & D. Pallara (2000), *Functions of Bounded Variation and Free Discontinuity Problems*, Oxford Mathematical Monographs, The Clarendon Press, Oxford University Press, New York.

H. Attouch, G. Buttazzo & G. Michaille (2014), *Variational Analysis in Sobolev and BV Spaces*, 2. Aufl., Society for Industrial & Applied Mathematics (SIAM), Philadelphia, PA, doi: 10.1137/1.9781611973488.

H. H. Bauschke & P. L. Combettes (2017), *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, 2. Aufl., Springer, New York, doi: 10.1007/978-3-319-48311-5.

A. Beck & M. Teboulle (2009), A fast iterative shrinkage-thresholding algorithm for linear inverse problems, *SIAM J. Imaging Sci.* 2(1), 183–202, doi: 10.1137/080716542.

K. Bredies & D. A. Lorenz (2011), *Mathematische Bildverarbeitung, Einführung in Grundlagen und moderne Theorie*, Vieweg+Teubner, doi: 10.1007/978-3-8348-9814-2.

H. Brezis (2010), *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Springer, New York, doi: 10.1007/978-0-387-70914-7.

M. Brokate (2014), Konvexe Analysis und Evolutionsprobleme, Vorlesungsskript, Zentrum Mathematik, TU München, url: http://www-m6.ma.tum.de/~brokate/cev_ss14.pdf.

A. Cegielski (2012), *Iterative methods for fixed point problems in Hilbert spaces*, Bd. 2057, Lecture Notes in Mathematics, Springer, Heidelberg, doi: 10.1007/978-3-642-30901-4.

A. Chambolle & T. Pock (2011), A first-order primal-dual algorithm for convex problems with applications to imaging, *J Math Imaging Vis* 40(1), 120–145, doi: 10.1007/s10851-010-0251-1.

T. Chan & S. Esedoglu (2005), Aspects of total variation regularized L1 function approximation, *SIAM Journal on Applied Mathematics* 65(5), 1817–1837, doi: 10.1137/040604297.

C. Clason (2015), Funktionalanalysis I, Vorlesungsskript, Fakultät für Mathematik, Universität Duisburg-Essen, url: https://www.uni-due.de/~adf040p/skripte/FunktAnSkript15.pdf.

M. Dobrowolski (2010), *Angewandte Funktionalanalysis*, Springer-Verlag, Berlin, doi: 10.1007/978-3-642-15269-6.

I. Ekeland & R. Témam (1999), *Convex Analysis and Variational Problems*, Bd. 28, Classics Appl. Math. SIAM, Philadelphia, doi: 10.1137/1.9781611971088.

B. He & X. Yuan (2012), Convergence analysis of primal-dual algorithms for a saddle-point problem: from contraction perspective, *SIAM Journal on Imaging Sciences* 5(1), 119–149, doi: 10.1137/100814494.

M. Kurula & H. J. Zwart (2012), The duality between the gradient and divergence operators on bounded Lipschitz domains, Techn. Ber. 1994, Enschede: Department of Applied Mathematics, University of Twente, URL: http://eprints.eemcs.utwente.nl/22373/.

Y. E. Nesterov (1983), A method for solving the convex programming problem with convergence rate $O(1/k^2)$, *Soviet Math. Doklad.* 27(2), 372–376.

Y. Nesterov (2004), *Introductory Lectures on Convex Optimization*, Bd. 87, Applied Optimization, A basic course, Kluwer Academic Publishers, Boston, MA, xviii+236, DOI: 10.1007/978-1-4419-8853-9.

N. Parikh & S. Boyd (2014), Proximal algorithms, *Foundations and Trends in Optimization* 1(3), 123–231, DOI: 10.1561/2400000003.

L. I. Rudin, S. Osher & E. Fatemi (1992), Nonlinear total variation based noise removal algorithms, *Physica D: Nonlinear Phenomena* 60(1), 259–268, DOI: 10.1016/0167-2789(92)90242-f.

O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier & F. Lenzen (2009), *Variational Methods in Imaging*, Springer, New York, DOI: 10.1007/978-0-387-69277-7.

W. Schirotzek (2007), *Nonsmooth Analysis*, Universitext, Springer, Berlin, DOI: 10.1007/978-3-540-71333-3.

T. Tao (11. März 2010a), A type diagram for function spaces, (blog post), What's new, URL: https://terrytao.wordpress.com/2010/03/11 (besucht am 05. 12. 2018).

T. Tao (12. März 2010b), Way to memorize relations between the Sobolev spaces?, (answer), MathOverflow, URL: http://mathoverflow.net/a/17906/30516 (besucht am 05. 12. 2018).

T. Valkonen (2014), Measure and Image, Vorlesungsskript, University of Cambridge, URL: http://tuomov.iki.fi/mathematics/mandi.pdf.

D. Werner (2011), *Funktionalanalysis*, 7. Aufl., Springer-Verlag, Berlin, DOI: 10.1007/978-3-642-21017-4.

W. P. Ziemer (1989), *Weakly Differentiable Functions*, Bd. 120, Graduate Texts in Mathematics, Sobolev spaces and functions of bounded variation, Springer-Verlag, New York, DOI: 10.1007/978-1-4612-1015-3.