MATHEMATISCHE BILDVERARBEITUNG

Vorlesungsskript, Sommersemester 2014

Christian Clason

Stand vom 6. Oktober 2014

Fakultät für Mathematik
Universität Duisburg-Essen

INHALTSVERZEICHNIS

I GRUNDLAGEN

1 GRUNDLAGEN DER FUNKTIONALANALYSIS 5
- 1.1 Normierte Räume 5
- 1.2 Starke und schwache Topologie 7
- 1.3 Hilberträume 11

2 GRUNDLAGEN DER VARIATIONSRECHNUNG 14
- 2.1 Direkte Methode der Variationsrechnung 14
- 2.2 Differenzierbarkeit in Banachräumen 18

3 GRUNDLAGEN DER KONVEXEN ANALYSIS 21
- 3.1 Konvexe Funktionen 21
- 3.2 Konvexe Subdifferentiale 23
- 3.3 Konvexe Konjugierte und Dualität 31

II ALGORITHMEN

4 PROXIMALPUNKT-VERFAHREN 37
- 4.1 Monotone Operatoren 37
- 4.2 Resolventen und Proximalpunkte 40
- 4.3 Proximalpunkt-Verfahren 45

5 SPLITTING-VERFAHREN 47
- 5.1 Explizites Splitting 47
- 5.2 Implizites Splitting 51
- 5.3 Primal-duale Verfahren 51

III BILDMODELLE

6 KONTINUIERLICHE BILDMODELLE 57
6.1 Lebesgue-Räume 57
6.2 Sobolev-Räume 60
6.3 Funktionen mit beschränkter Variation 65

7 DISKRETE BILDMODELLE 70

IV REKONSTRUKTIONSMODELLE

8 ENTRAUSCHEN 75
8.1 L^2-H^1-Entrauschen 75
8.2 L^2-TV-Entrauschen 76
8.3 L^1-TV-Entrauschen 77

9 ENTFALTEN 80

10 DEKOMPRESSION 86

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, d. h. $u^0 \in H^1(\Omega)$. (Dieser Raum wird später eingeführt.)

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

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

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

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

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

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

1. 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.

2. Sind F und G differenzierbar, so gilt das *Fermat-Prinzip*

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

3. 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 auf algorithmische Aspekte und verzichten auf die maßtheoretischen Details bei der Konstruktion der auftretenden Funktionenräume. Nichtsdestotrotz sind elementare Kenntnisse aus Funktionalanalysis sowie Numerischer Mathematik sinnvoll für das Verständnis der Inhalte.

Dieses Skriptum basiert vor allem auf den folgenden Werken:

[1] K. Bredies und 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 und P. L. Combettes (2011). *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC. Springer, New York. DOI: 10.1007/978-1-4419-9467-7

[4] H. Attouch u. a. (2006). *Variational Analysis in Sobolev and BV Spaces*. Bd. 6. MPS/SIAM Series on Optimization. Society for Industrial und Applied Mathematics (SIAM), Philadelphia, PA. DOI: 10.1137/1.9780898718782

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].

1.1 NORMIERTE RÄUME

Im Folgenden bezeichne X ein 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) $\|x + y\| \leqslant \|x\| + \|y\|$ für alle $y \in X$,

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

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

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

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 als Elemente in einem unendlichdimensionalen Funktionenraum betrachten wollen.

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

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

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

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

(iii) Auf $X = L^p(\Omega)$ (dem Raum der messbaren reellen Funktionen, auf dem folgende Ausdrücke endlich sind) sind Normen definiert durch

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

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

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

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.

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

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

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;
- $\ker F := \{x \in U : Fx = 0\}$ den *Kern* (englisch „kernel“ oder „null space“) von F;
- $\operatorname{rg} F := \{Fx \in Y : x \in U\}$ 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

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

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

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

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

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

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*

$$\|F\|_{L(X,Y)} = \sup_{x \in X \setminus \{0\}} \frac{\|Fx\|_Y}{\|x\|_X} = \sup_{\|x\|_X \leqslant 1} \|Fx\|_Y$$

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

$$c\|x\|_X \leqslant \|Fx\|_Y \qquad \text{für alle } x \in X,$$

in diesem Fall ist $\|F^{-1}\|_{L(Y,X)} = c^{-1}$.

1.2 STARKE UND SCHWACHE TOPOLOGIE

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 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.

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

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

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);
- *beschränkt*, falls sie in einer abgeschlossenen Kugel $K_r(0)$ für ein $r > 0$ enthalten ist;
- *konvex*, falls für $u, v \in U$ auch $\lambda u + (1 - \lambda)v \in U$ 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 dann X 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.

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 wird. Aus der Definition der Operatornorm folgt sofort, dass

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

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

Beispiel 1.2. (i) $(\mathbb{R}^N, \|\cdot\|_p)^* = (\mathbb{R}^N, \|\cdot\|_q)$ mit $p^{-1} + q^{-1} = 1$, wobei $0^{-1} = \infty$ und $\infty^{-1} = 0$ gesetzt wird. Die duale Paarung ist gegeben durch

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

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

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

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

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

$$\langle u^*, u \rangle_{p,q} = \int_\Omega u^*(x) u(x)\, dx.$$

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

(iv) $C(\overline{\Omega})^*$ ist der Raum $\mathcal{M}(\Omega)$ 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 $u \in C(\overline{\Omega})$. Die duale Paarung ist gegeben durch

$$\langle u^*, u \rangle_{\mathcal{M},C} = \int_\Omega u(x)\, du^*.$$

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

Satz 1.3 (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*

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

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

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

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

Folgerung 1.5. *Sei X ein Banachraum. Dann gilt*

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

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

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

Aus Satz 1.4 folgt dann $\|J(x)\|_{X^{**}} = \|x\|_X$. Ist die kanonische Injektion surjektiv – können wir also X^{**} mit X identifizieren – so nennt man X *reflexiv*. Endlichdimensionale Räume sind reflexiv, sowie Beispiele 1.1 (ii) und (iii) für $1 < p < \infty$, nicht aber ℓ^1 und $L^1(\Omega)$ sowie $C(\overline{\Omega})$.

Durch die duale Paarung werden weitere Topologien erzeugt: die *schwache* Topologie auf X sowie die *schwach-$*$* Topologie 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

$$\langle x_n^*, x \rangle_X \to \langle x^*, x \rangle_X \qquad \text{für alle } x \in X.$$

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!) überein. In endlichdimensionalen Räumen stimmen alle Konvergenzbegriffe überein.

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

In einem reflexiven Raum sind also insbesondere alle beschränkten und schwach abgeschlossenen Mengen schwach kompakt. Beachten Sie, dass schwache Abgeschlossenheit eine *stärkere* Forderung ist als Abgeschlossenheit. Für konvexe Mengen stimmen die Begriffe aber überein.

Lemma 1.7. *Eine konvexe Menge $U \subset X$ ist genau dann abgeschlossen, wenn sie schwach abgeschlossen ist.*

Beweis. Da eine konvergente Folge stets auch schwach konvergiert, ist jede schwach abgeschlossene Menge auch abgeschlossen. Sei daher $U \subset X$ konvex und abgeschlossen, und betrachte eine Folge $\{x_n\}_{n\in\mathbb{N}} \subset U$ mit $x_n \rightharpoonup x \in X$. Angenommen, $x \in X \setminus U$. Da U abgeschlossen ist, ist $X \setminus U$ offen, und es existiert daher ein $z \in X$ und $r > 0$ mit $x \in B_r(z) \subset X \setminus U$. Die Mengen U und $B_r(z)$ erfüllen also alle Bedingungen des Hahn–Banach-Trennungssatzes (Satz 1.3); wir finden daher ein $x^* \in X^*$ mit

$$\langle x^*, x\rangle_X < \lambda \leqslant \langle x^*, x_n\rangle_X \quad \text{für alle } n \in \mathbb{N}.$$

Grenzübergang auf beiden Seiten ergibt dann aber den Widerspruch

$$\langle x^*, x\rangle_X < \langle x^*, x\rangle_X. \qquad \square$$

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

Satz 1.8 (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 $1 \leqslant p < \infty$ separabel; auch ℓ^p ist separabel für $1 \leqslant p < \infty$. 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. Allerdings sind abgeschlossene konvexe Mengen in nichtreflexiven Räumen nicht unbedingt schwach-$*$ abgeschlossen.

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

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

der *adjungierte Operator* $F^* \in L(Y^*, X^*)$ definiert. Es gilt stets $\|F^*\|_{L(Y^*,X^*)} = \|F\|_{L(X,Y)}$. Ausserdem folgt aus der Stetigkeit von F, dass F^* schwach-$*$ stetig (und F 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) $(x, x) \geqslant 0$ 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

$$\|x\|_X := ((x,x)_X)^{1/2}$$

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

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

Die Beispiele 1.2 (i–iii) für $p = 2$ $(= q)$ sind Hilberträume, 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*:

$$\sum_{y \in U} |(x,y)_X|^2 \leqslant \|x\|_X^2 \qquad \text{für alle } x \in X, \tag{1.2}$$

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*

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

(man spricht daher im Fall der Gleichheit in (1.2) auch von der *Parseval-Relation*).

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

Satz 1.9 (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*

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

Die Abbildung $J_X : x^* \mapsto z(x^*)$ wird *Riesz-Isomorphismus* genannt.

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

$$(F^\star y, x)_X = (Fx, y)_Y \qquad \text{für alle } x \in X, y \in Y.$$

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

GRUNDLAGEN DER VARIATIONSRECHNUNG

2

Variationsmethoden führen auf die Aufgabe, ein (nichtlineares) Funktional $F : X \to \mathbb{R}$ über eine Teilmenge U eines Banachraums X zu minimieren. Mit solchen Problemen beschäftigt sich die *Variationsrechnung*.

2.1 DIREKTE METHODE DER VARIATIONSRECHNUNG

Wir befassen uns zuerst mit der Frage nach der Existenz eines Minimierers $\bar{x}$ von F. Dabei ist es 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

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

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

Wir betrachten also in Folge Funktionen $F : X \to \overline{\mathbb{R}}$. 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 wollen nun den Satz von Weierstrass (jede reelle stetige Funktion auf kompakten Mengen nimmt ihr Minimum und Maximum an) auf Banachräume und insbesondere Funktionen der Form $\tilde{F}$ erweitern. Da wir nur an Minima interessiert sind, reicht dafür eine „einseitige" Stetigkeit: Man nennt F *unterhalbstetig* (englisch: „lower semicontinuous"), falls für alle konvergenten Folgen $\{x_n\}_{n\in\mathbb{N}} \subset X$ mit Grenzwert $x \in X$ gilt

$$F(x) \leqslant \liminf_{n\to\infty} F(x_n).$$

Analog definiert man *schwach*(-$*$) unterhalbstetige (bzw. oberhalbstetige) 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, 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 F nach unten beschränkt ist.*

Angenommen, F ist nicht nach unten beschränkt. Dann existiert eine Folge $\{x_n\}_{n\in\mathbb{N}} \subset X$ mit $F(x_n) \to -\infty$. Aus der Koerzivität von F folgt dann, dass diese Folge beschränkt ist (sonst müsste ja $F(x_n) \to \infty$ gelten), d. h. es gibt ein $M > 0$ mit $\|x_n\| \leqslant M$ für alle $n \in \mathbb{N}$. Insbesondere ist also $\{x_n\}_{n\in\mathbb{N}} \subset B_M(0)$. Aus dem Satz von Eberlein–Šmulyan folgt dann die Existenz einer schwach konvergenten Teilfolge $\{x_{n_k}\}_{k\in\mathbb{N}}$, für deren Grenzwert $\tilde{x} \in X$ wegen der schwachen Unterhalbstetigkeit gilt

$$F(\tilde{x}) \leqslant \liminf_{k\to\infty} F(x_{n_k}) = -\infty.$$

Dies ist aber ein Widerspruch zu $F(x) \in (-\infty, \infty]$.

Da F eigentlich und nach unten beschränkt ist, muss ein $M := \inf_{x\in X} F(x) \in \mathbb{R}$ existieren. Aus der Definition des Infimums folgt dann, dass eine Folge $\{y_n\}_{n\in\mathbb{N}} \subset \operatorname{rg} F \subset \mathbb{R}$ existiert 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 $\{F(x_n)\}_{n\in\mathbb{N}}$ nicht auf die Konvergenz von $\{x_n\}_{n\in\mathbb{N}}$ schliessen können.

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

Aus der Koerzivität von F folgt wieder, dass die Minimalfolge beschränkt ist und daher nach dem Satz von Eberlein–Šmulyan eine schwach konvergente Teilfolge $\{x_{n_k}\}_{k\in\mathbb{N}}$ mit Grenzwert $\bar{x} \in X$ besitzt. Dieser Grenzwert ist ein Kandidat für einen Minimierer.

[1] Diese Beweis-Strategie, eingeführt unter dem Namen *direkte Methode der Variationsrechnung* von Leonida Tonelli (1885–1946), wird so häufig angewendet, dass in Forschungsarbeiten üblicherweise nur geschrieben wird: „Die Existenz eines Minimierers folgt aus Standard-Argumenten."

(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

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

Das Infimum wird also in $\bar{x}$ angenommen, d. h. $F(\bar{x}) = \min_{x\in X} F(x)$.

□

Ist X nicht reflexiv, aber Dualraum eines separablen Banachraums, so zeigt man analog die Existenz von Minimierern 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 schwachen Topologie (je schwächer die Topologie, desto mehr Folgen konvergieren und desto mehr Mengen sind daher kompakt), Schritt (iii) von einer starken (je weniger Folgen konvergieren, desto einfacher ist die lim inf-Bedingung zu erfüllen). Da wir in unseren Fällen nicht mehr als die Beschränktheit einer Minimalfolge erwarten können, müssen wir wenigstens 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

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

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.

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

(i) αF *für alle* $\alpha \geqslant 0$*;*

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

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

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

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

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

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

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

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

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

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

Sei schliesslich $\{x_n\}_{n\in\mathbb{N}}$ eine schwach konvergente Folge mit Grenzwert $x \in X$. Dann gilt nach Definition des Supremums für alle $j \in I$, dass

$$F_j(x) \leqslant \liminf_{n\to\infty} F_j(x_n) \leqslant \liminf_{n\to\infty} \sup_{i\in I} F_i(x_n).$$

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 $\|\cdot\|_X$ schwach unterhalbstetig und koerziv.*

Beweis. Koerzivität folgt direkt aus der Definition, und schwache Unterhalbstetigkeit aus Lemma 2.2 (v) und Folgerung 1.5, denn

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

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

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

Lemma 2.4. *Sei $U \subset X$. Dann ist δ_U*

(i) eigentlich, wenn U nichtleer ist;

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

(iii) koerziv, wenn U beschränkt ist.

Beweis. Aussage (i) ist klar. Nach Lemma 1.7 ist eine konvexe Menge abgeschlossen genau dann, wenn sie schwach abgeschlossen ist. Sei daher U schwach abgeschlossen und $\{x_n\}_{n\in\mathbb{N}} \subset X$ eine schwach konvergente Folge mit Grenzwert $x \in X$. Angenommen, $x \in U$, dann ist wegen $\delta_U \geqslant 0$ natürlich

$$0 = \delta_U(x) \leqslant \lim \inf_{n\to\infty} \delta_U(x_n).$$

Ist $x \notin U$, dann muss ein $N \in \mathbb{N}$ existieren mit $x_n \notin U$ für alle $n \geqslant N$ (sonst wäre U nicht abgeschlossen). Also gilt $\delta_U(x_n) = \infty$ für alle $n \geqslant N$ und damit

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

Für (iii) sei U beschränkt, d. h. es gibt ein $M > 0$ mit $U \subset B_M(0)$. Gilt $\|x_n\|_X \to \infty$, so existiert ein $N \in \mathbb{N}$ mit $\|x_n\|_X > M$ für alle $n \geqslant N$, d. h. $x_n \notin B_M(0) \supset U$ für alle $n \geqslant M$. 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

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

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

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

  $$DF(x)h := F'(x; h)$$

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

- Gilt zusätzlich

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

  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 $F \in L(X, Y)$ Fréchet-differenzierbar mit Ableitung $DF = F \in L(X, Y)$. Weitere Ableitungen erhält man durch die üblichen Rechenregeln, die genau wie in $\mathbb{R}^n$ gezeigt werden. Beispielhaft beweisen wir eine Kettenregel.

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* $y = F(x) \in X$. *Dann ist* $G \circ F$ *Fréchet-differenzierbar in* x *und*

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

Beweis. Für $h \in X$ mit $x + h \in \operatorname{dom} F$ gilt

$$(G \circ F)(x+h) - (G \circ F)(x) = G(F(x+h)) - G(F(x)) = G(y+g) - G(y)$$

für $g = F(x+h) - F(x)$. Aus der Fréchet-Differenzierbarkeit von G folgt daher, dass

$$\|(G \circ F)(x+h) - (G \circ F)(x) - G'(y)g\|_Z = r_1(\|g\|_Y)$$

gilt mit $r_1(t)/t \to 0$ für $t \to 0$. Aus der Fréchet-Differenzierbarkeit von F folgt aber, dass

$$\|g - F'(x)h\|_Y = r_2(\|h\|_X)$$

gilt mit $r_2(t)/t \to 0$ für $t \to 0$. Insbesondere ist

$$\|g\|_Y \leqslant \|F'(x)h\|_Y + r_2(\|h\|_X). \tag{2.1}$$

Also gilt, dass

$$\|(G \circ F)(x+h) - (G \circ F)(x) - G'(F(x))F'(x)h\|_Z \leqslant r_1(\|g\|_Y) + r_2(\|h\|_X).$$

Für $\|h\| \to 0$ folgt aus (2.1) und $F'(x) \in L(X, Y)$ auch $\|g\|_Y \to 0$ und damit die gewünschte Aussage. □

Eine analoge Regel für Gâteaux-Ableitungen gilt dagegen nicht!

Wir betrachten nun die Charakterisierung von Minimierern eines differenzierbaren Funktionals $F : X \to \mathbb{R}$.[2]

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

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

Beweis. Wenn $\bar{x}$ ein Minimierer von F ist, so muss insbesondere für alle $h \in X$ die Funktion $f : t \mapsto F(\bar{x} + th)$ ein Minimum in $t = 0$ haben. Da F Gâteaux-differenzierbar ist, existiert die Ableitung $f'(t)$ in $t = 0$ und damit muss gelten

$$0 = f'(0) = \lim_{t \to 0^+} \frac{f(t) - f(0)}{t} = F'(x; h). \qquad \square$$

Beachten Sie, dass 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 Hilbert-Räumen (darunter auch $\mathbb{R}^n$) 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

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

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 $x, h \in X$, dass

$$DF(x) = \frac{1}{2}\left(\frac{d}{dt}(x + th, x + th)_X\right)\bigg|_{t=0} = (x, h)_X\,.$$

Die quadrierte Norm ist also Gâteaux-differenzierbar in x mit Ableitung $DF(x) : h \mapsto (x, h)_X$ und Gradient $\nabla F(x) = x \in X$; wegen

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

ist sie sogar Fréchet-differenzierbar. Fasst man nun dieselbe Abbildungsvorschrift auf als definiert auf einem kleineren Hilbertraum $X' \hookrightarrow X$ (zum Beispiel $X = L^2(\Omega)$, $X' = H^1(\Omega)$), so ist immer noch $DF(x)h = (x, h)_X \in (X')^*$, aber $\nabla F \in X'$ ist nun charakterisiert durch

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

Unterschiedliche Skalarprodukte werden nun zu unterschiedlichen Gradienten führen.

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

GRUNDLAGEN DER KONVEXEN ANALYSIS

3

Die klassischen Ableitungsbegriffe des letzten Kapitels sind für unsere Zwecke aber 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 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 $F : X \to \bar{\mathbb{R}}$ heisst *konvex*, wenn für alle $x, y \in X$ und $\lambda \in [-1, 1]$ gilt

$$F(\lambda x + (1 - \lambda)y) \leqslant \lambda F(x) + (1 - \lambda)F(y)$$

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

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

so heisst F *strikt konvex*.

Ein Funktional F ist genau dann konvex, wenn ihr *Epigraph*

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

konvex (als Teilmenge von $X \times \mathbb{R}$) ist. Weiterhin ist F genau dann konvex und unterhalbstetig, wenn $\operatorname{epi} F$ konvex und abgeschlossen ist. Daraus folgt auch, dass ein konvexes Funktional genau dann unterhalbstetig ist, wenn es schwach 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.

Weitere Beispiele lassen sich durch folgende Operationen erzeugen.

Lemma 3.1. *Seien* X, Y *normierte Räume und sei* $F : X \to \bar{\mathbb{R}}$ *konvex. Dann sind konvex*

(i) αF *für alle* $\alpha \geqslant 0$*;*

(ii) $F + G$ *für* $G : X \to \bar{\mathbb{R}}$ *konvex;*

(iii) $\varphi \circ F$ *für* $\varphi : \bar{\mathbb{R}} \to \bar{\mathbb{R}}$ *konvex und monoton steigend (ist* φ *strikt konvex, so auch* $\varphi \circ F$*);*

(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 3.2. *Seien*

(i) U *eine nichtleere, abgeschlossene, konvexe Teilmenge eines reflexiven Banachraums* X,

(ii) $F : U \to \overline{\mathbb{R}}$ *eigentlich, konvex, 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 $\bar{x} \in U$. *Ist* F *strikt konvex, so ist die Lösung eindeutig.*

Beweis. Wir betrachten das Funktional $\tilde{F} = F + \delta_U$. 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 $\tilde{F}(x_0) < \infty$, daher ist auch $\tilde{F}$ eigentlich, konvex und damit sogar unterhalbstetig. Aus (iii) folgt schliesslich die Koerzivität von $\tilde{F}$ (da die Summe koerziv ist, sobald ein Summand es ist). Wir können daher Theorem 2.1 anwenden und erhalten die Existenz eines Minimierers $\bar{x} \in U \subset \operatorname{dom} \tilde{F}$. Da für alle $x \in U$ gilt

$$F(\bar{x}) = \tilde{F}(\bar{x}) \leqslant \tilde{F}(x) = F(x),$$

ist $\bar{x}$ die gesuchte Lösung.

Sei nun F strikt konvex, und seien $\bar{x}, \bar{x}' \in U$ zwei verschiedene Minimierer, $\eth F(\bar{x}) = F(\bar{x}') = \min_{x \in U} F(x)$. 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

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

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

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 für jede konvexe Funktion.

Lemma 3.3. *Sei* $F : X \to \overline{\mathbb{R}}$ *konvex und* $x \in \operatorname{dom} F$. *Dann existiert* $F'(x; h)$ *für jedes* $h \in X$, *und es gilt*

$$F'(x; h) \leqslant F(x + h) - F(x). \tag{3.1}$$

Beweis. Wir zeigen zuerst, dass für $x \in \operatorname{dom} F$ und $h \in X$ die Abbildung

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

monoton steigend ist. Durch Einsetzen und Umformen sieht man, dass für alle $0 < s < t$ die Bedingung $\varphi(s) \leqslant \varphi(t)$ äquivalent ist zu

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

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

$$F'(x; h) = \lim_{t \to 0^+} \varphi(t) = \inf_{t > 0} \varphi(t),$$

und damit

$$F'(x; h) \leqslant \varphi(1) = F(x + h) - F(x). \qquad \square$$

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 $0 \leqslant f'(0; h)$ für alle $h \in \mathbb{R}$. Diese Bedingung wollen wir auf allgemeine normierte Räume verallgemeinern. Dafür betrachten wir für $F : X \to \overline{\mathbb{R}}$ konvex und $x \in \operatorname{dom} F$ die Menge

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

Diese Menge kann man mit Hilfe von Lemma 3.3 auch ohne Richtungsableitung charakterisieren.

Lemma 3.4. *Seien* $F : X \to \overline{\mathbb{R}}$ *konvex und* $x \in \operatorname{dom} F$. *Dann sind für* $x^* \in X^*$ *äquivalent*

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

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

Beweis. Gilt (i), so folgt direkt aus (3.1), dass für alle $h \in X$ gilt

$$\langle x^*, h\rangle_X \leqslant F'(x;h) \leqslant F(x+h) - F(x).$$

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

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

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

$$\partial F(x) := \{x^* \in X^* : \langle x^*, \tilde{x} - x\rangle_X \leqslant F(\tilde{x}) - F(x) \quad \text{für alle } \tilde{x} \in X\}. \tag{3.2}$$

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

Diese Definition liefert nun das Gewünschte.

Satz 3.5. *Seien* $F : X \to \overline{\mathbb{R}}$ *konvex und* $\bar{x} \in \operatorname{dom} F$. *Dann sind äquivalent:*

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

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

Beweis. Dies folgt direkt aus den Definitionen: $\bar{x}$ minimiert F genau dann, wenn

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

d. h. wenn $0 \in \partial F(\bar{x})$ gilt. $\square$

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

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

Satz 3.6. *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 3.4 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

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

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

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

Satz 3.7. *Für $x \in X$ ist*

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

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

$$\langle \xi, \tilde{x}\rangle_X \leqslant \|\tilde{x}\|_X \qquad \text{für alle } \tilde{x} \in X.$$

Dies ist aber wegen der Definition der Operatornorm äquivalent mit $\|\xi\| \leqslant 1$.

Sei nun $x \neq 0$ und betrachte $\xi \in \partial(\|\cdot\|_X)(x)$. Indem wir nacheinander $\tilde{x} = 0$ und $\tilde{x} = 2x$ in (3.2) einsetzen, erhalten wir

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

Analog haben wir für alle $\tilde{x} \in X$, dass

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

woraus wie im Fall $x = 0$ folgt $\|\xi\| \leqslant 1$. Für $\tilde{x} = x/\|x\|$ 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

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

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

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

$$\partial(|\cdot|)(t) = \operatorname{sign}(t) := \begin{cases} \{1\} & t > 0, \\ \{-1\} & t < 0, \\ [-1, 1] & t = 0. \end{cases} \tag{3.3}$$

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

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

da die Ungleichung für alle $\tilde{x} \notin C$ trivialerweise gilt. 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 $\xi(\tilde{t} - t) \leqslant 0$ 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 $\xi \geqslant 0$ ist.

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

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

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

Subdifferentiale weiterer Funktionale erhält man durch Rechenregeln. Es ist naheliegend, dass diese umso aufwendiger 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 3.8. *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 aufwendiger. Wir benötigen die folgende Variante des Hahn–Banach-Trennsatzes.

Lemma 3.9. *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*

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

Beweis. Satz 1.3 liefert die Existenz von x^* und λ, so dass die Aussage gilt für alle $x_1 \in A^o$ (sogar mit strikter Ungleichung). Es bleibt also zu zeigen, dass $\langle x^*, x\rangle_X \leqslant \lambda$ 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

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

und es gilt sogar $x(t) := tx_0 + (1-t)x \in A^o$ 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 $x_n \to x$. Grenzübergang liefert dann

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

Satz 3.10 (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 aus der Definition des Subdifferentials. Seien daher $x \in \operatorname{dom} F \cap G$ und $\xi \in \partial(F + G)(x)$, erfüllen also

$$\langle \xi, \tilde{x} - x\rangle_X \leqslant (F(\tilde{x}) + G(\tilde{x})) - (F(x) + G(x)) \quad \text{für alle } \tilde{x} \in X. \tag{3.4}$$

Unser Ziel ist nun, mit Hilfe der Charakterisierung konvexer Funktionen durch ihren Epigraphen und des Trennsatzes ein lineares Funktional $\zeta \in X^*$ zu finden mit

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

d. h. $\xi = (\xi - \zeta) + \zeta \in \partial F(x) + \partial G(x)$.

Wir definieren dafür die Mengen

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

d. h.

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

Da es sich um verschobene (und, für C_2, gespiegelte) Epigraphen eigentlicher konvexer Funktionen (lineare Funktionale sind konvex) handelt, sind C_1 und C_2 nichtleer und konvex. Weiter ist x_0 ein innerer Punkt von $\operatorname{dom} F$ (sonst wäre $F - \xi$ und damit F nicht stetig in x_0) und damit (x_0, α) für α groß genug ein innerer Punkt von C_1. 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 gilt

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

im Widerspruch zu (3.4). Lemma 3.9 liefert also ein $(x^*, s) \in (X \times \mathbb{R})^* \setminus \{0\}$ und ein $\lambda \in \mathbb{R}$ mit

$$\langle x^*, \tilde{x}\rangle_X + s(t - (F(x) - \langle \xi, x\rangle_X)) \leqslant \lambda, \quad \tilde{x} \in \operatorname{dom} F, t \geqslant F(\tilde{x}) - \langle \xi, \tilde{x}\rangle_X, \tag{3.5}$$

$$\langle x^*, \tilde{x}\rangle_X + s(G(x) - t) \geqslant \lambda, \quad \tilde{x} \in \operatorname{dom} G, t \geqslant G(\tilde{x}). \tag{3.6}$$

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

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

da (x_0, α) für α groß genug innerer Punkt von C_1 ist und daher nach Satz 1.3 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 (3.5) positiv, und $t \to \infty$ mit $\tilde{x}$ fest führt zum Widerspruch zur Beschränktheit durch λ.

Also ist $s < 0$, und aus (3.5) mit $t = F(\tilde{x}) - \langle \xi, \tilde{x}\rangle_X$ und aus (3.6) mit $t = G(\tilde{x})$ folgt

$$\begin{aligned} F(\tilde{x}) - F(x) + \langle \xi, \tilde{x} - x\rangle_X &\geqslant s^{-1}(\lambda - \langle x^*, \tilde{x}\rangle_X), \quad &\tilde{x} \in \operatorname{dom} F,\\ G(x) - G(\tilde{x}) &\leqslant s^{-1}(\lambda - \langle x^*, \tilde{x}\rangle_X), \quad &\tilde{x} \in \operatorname{dom} G. \end{aligned}$$

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 $\zeta = s^{-1}x^*$ das gewünschte Funktional, für das gilt $(\xi - \zeta) \in \partial F$ und $\zeta \in \partial G$, d. h. $\xi \in \partial F + \partial G$. □

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

Satz 3.11 (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

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

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.

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

Wir konstruieren nun $\eta \in \partial F(Ax)$ durch Trennung von $\operatorname{epi} F$ und

$$\operatorname{graph} A = \{(x, Ax) : x \in A\},$$

und zwar durch Anwenden der Summenregel auf $H : X \times Y \to \overline{\mathbb{R}}$,

$$H(x, y) := F(y) + \delta_{\operatorname{graph} A}(x, y).$$

Da A linear ist, ist $\operatorname{graph} A$ und damit $\delta_{\operatorname{graph} A}$ konvex. Weiter ist Ax in $\operatorname{dom} F$ und damit $(x, Ax) \in \operatorname{dom} H$.

Zuerst zeigen wir, dass $\xi \in \partial(F \circ A)(x)$ genau dann gilt, wenn $(\xi, 0) \in \partial H(x, Ax)$ ist. Sei zuerst $(\xi, 0) \in \partial H(x, Ax)$. Dann gilt für alle $\tilde{x} \in X$, $\tilde{y} \in Y$

$$\langle \xi, \tilde{x} - x \rangle_X + \langle 0, \tilde{y} - Ax \rangle_Y \leqslant F(\tilde{y}) - F(Ax) + \delta_{\operatorname{graph} A}(\tilde{x}, \tilde{y}) - \delta_{\operatorname{graph} A}(x, Ax).$$

Insbesondere gilt dies für alle $\tilde{y} \in \operatorname{rg}(A) = \{A\tilde{x} : \tilde{x} \in X\}$. Also ist wegen $\delta_{\operatorname{graph} A}(\tilde{x}, A\tilde{x}) = 0$

$$\langle \xi, \tilde{x} - x \rangle_X \leqslant F(A\tilde{x}) - F(Ax) \quad \text{für alle } \tilde{x} \in X,$$

d. h. $\xi \in \partial(F \circ A)(x)$. Sei umgekehrt $\xi \in \partial(F \circ A)(x)$. Dann ist für alle $\tilde{x} \in X$ und $\tilde{y} \in Y$ wegen $\delta_{\operatorname{graph} A}(x, Ax) = 0$ und $\delta_{\operatorname{graph} A}(\tilde{x}, \tilde{y}) \geqslant 0$

$$\begin{aligned}
\langle \xi, \tilde{x} - x \rangle_X + \langle 0, \tilde{y} - Ax \rangle_Y &= \langle \xi, \tilde{x} - x \rangle_X \\
&\leqslant F(A\tilde{x}) - F(Ax) + \delta_{\operatorname{graph} A}(\tilde{x}, \tilde{y}) - \delta_{\operatorname{graph} A}(x, Ax) \\
&= F(\tilde{y}) - F(Ax) + \delta_{\operatorname{graph} A}(\tilde{x}, \tilde{y}) - \delta_{\operatorname{graph} A}(x, Ax).
\end{aligned}$$

da für $\tilde{y} \neq A\tilde{x}$ beide Seiten der letzten Gleichung unendlich sind. Also ist $(\xi, 0) \in \partial H(x, Ax)$.

Da F stetig ist in Ax_0, ist $(x, y) \mapsto F(y)$ stetig in $(x_0, Ax_0) \in \operatorname{graph} A = \operatorname{dom} \delta_{\operatorname{graph} A}$. Aus der Summenregel folgt nun

$$(\xi, 0) \in \partial H(x, Ax) = \partial F(Ax) + \partial \delta_{\operatorname{graph} A}(x, Ax).$$

Nun ist $(x^*, y^*) \in \partial F(Ax)$ (aufgefasst als Abbildung $(x, y) \mapsto F(y)$ genau dann, wenn

$$\langle x^*, \tilde{x} - x \rangle_X + \langle y^*, \tilde{y} - Ax \rangle_Y \leqslant F(\tilde{y}) - F(Ax) \quad \text{für alle } \tilde{x} \in X, \tilde{y} \in Y$$

ist. Festhalten von $\tilde{x} = x$ bzw. $\tilde{y} = Ax$ liefert $y^* \in \partial F(Ax)$ (als Abbildung auf Y) und $x^* = 0$. Weiter ist $(w^*, z^*) \in \partial \delta_{\operatorname{graph} A}(x, Ax)$ genau dann, wenn

$$\langle w^*, \tilde{x} - x\rangle_X + \langle z^*, \tilde{y} - Ax\rangle_Y \leqslant 0 \quad \text{für alle } (\tilde{x}, \tilde{y}) \in \operatorname{graph} A,$$

d. h. für alle $\tilde{x} \in X$ und $\tilde{y} = A\tilde{x}$. Also ist

$$\langle w^* + A^*z^*, \tilde{x} - x\rangle_X \leqslant 0 \quad \text{für alle } \tilde{x} \in X$$

und damit $w^* = -A^*z^*$. Zusammen erhalten wir

$$(\xi, 0) = (0, y^*) + (-A^*z^*, z^*),$$

woraus $\xi = -A^*z^* = A^*y^*$ mit $y^* \in \partial F(Ax)$ folgt, d. h. $\xi \in A^*\partial F(Ax)$. □

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

Folgerung 3.12. *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

$$\begin{cases} -\xi \in \partial F(\bar{x}), \\ \langle \xi, \tilde{x} - x\rangle \leqslant 0 \quad \text{für alle } \tilde{x} \in U. \end{cases} \tag{3.7}$$

Beweis. Da F und U konvex sind, können wir Satz 3.5 auf $J := F + \delta_U$ anwenden. Da δ_U in $x_0 \in U^o$ stetig ist, folgt aus der Summenregel, dass F in $\bar{x}$ ein Minimum hat 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 (3.7). □

Für eine Gâteaux-differenzierbare (und damit stetige) Funktion $F : X \to \mathbb{R}$ ergibt (3.7) die klassischen *Karush–Kuhn–Tucker-Bedingungen*.

3.3 KONVEXE KONJUGIERTE UND DUALITÄT

Ein Grund für die Nützlichkeit des konvexen Subdifferentials ist, wie wir sehen werden, seine Verbindung mit der Fenchel–Legendre-Transformation. 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.) Aus Lemma 3.1 (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 F^* eigentlich. Aus der Definition folgt auch sofort die *Fenchel–Young-Ungleichung*

$$\langle x^*, x \rangle_X \leqslant F(x) + F^*(x^*) \qquad \text{für alle } x \in X, x^* \in X^*. \tag{3.8}$$

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^*. Analog definieren wir für $F : X^* \to \overline{\mathbb{R}}$ die Fenchel-Konjugierte als

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

Durch diese Konvention ist die *Bikonjugierte* $F^{**} = (F^*)^*$ selbst für nicht-reflexive Räume wieder auf X definiert (anstatt auf X^{**}). Anschaulich ist F^{**} die untere konvexe Hülle von F, die für konvexe Funktionen ja mit F übereinstimmt. Der Satz von Moreau – den wir hier nicht beweisen wollen[1] – besagt, dass das tatsächlich der Fall ist.

Satz 3.13 (Fenchel–Moreau–Rockafellar)**.** *Sei $F : X \to \overline{\mathbb{R}}$ eigentlich. Dann gilt*

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

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

Wir betrachten wieder relevante Beispiele.

Beispiel 3.14. (i) Sei $X = L^2(\Omega)$ und $F(x) = \frac{1}{2}\|x\|_{L^2}^2$. Wir identifizieren den Hilbertraum 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

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

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

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

[1] siehe z. B. [Schirotzek 2007, Theorem 2.2.4]

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

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

Analog zeigt man unter Verwendung der Definition der Konjugierten im Dualraum und Folgerung 1.5, 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: $\|x^*\|_{X^*} \leqslant 1$. Dann gilt mit (1.1) für alle $x \in X$ dass $\langle x^*, x\rangle_X - \|x\|_X \leqslant 0$ 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

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

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 3.15. *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

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

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

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

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

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

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

Lemma 3.16. *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

$$
\langle x^*, x\rangle_X - F(x) = F^*(x^*) \geqslant \langle x^*, \tilde{x}\rangle_X - F(\tilde{x}), \tag{3.9}
$$

was nach Definition äquivalent ist zu $x^* \in \partial F(x)$. Umgekehrt ergibt Supremum über alle $\tilde{x} \in X$ auf beiden Seiten von (3.9)

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

und zusammen mit (3.8) folgt (i).

Analog zeigt man unter Zuhilfenahme von Satz 3.13 die Äquivalenz von (i) und (iii). □

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

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

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 3.16 spielt die Rolle des „Satz 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

$$\inf_{x\in X} F(x) + G(Ax) \tag{3.10}$$

mit $G = \|\cdot\|_Y$ (wir werden später mehrere davon sehen), so können wir G ersetzen durch die Definition von $G^{**} = (\delta_{B_{Y^*}})^* = \|\cdot\|_Y = G$ und erhalten

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

(Nach dem Satz von Moreau geht das sogar für jedes konvexe unterhalbstetige Funktional G). Dürften wir nun inf und sup vertauschen, so könnten wir (wegen $\inf F = -\sup(-F)$) schreiben

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

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

$$\sup_{y^*\in Y^*} -F^*(-A^*y^*) - G^*(y^*). \tag{3.11}$$

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 3.17 (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 (3.10) *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 (3.11) eine Lösung $\bar{y}^* \in Y^*$ und es gilt

$$\min_{x\in X} F(x) + G(Ax) = \max_{y^*\in Y^*} -F^*(-A^*y^*) - G^*(y^*). \tag{3.12}$$

Weiterhin sind $\bar{x}$ und $\bar{y}^*$ Lösungen von (3.10) bzw. (3.11) genau dann, wenn gilt

$$\begin{cases} -A^*\bar{y}^* \in \partial F(\bar{x}),\\ \qquad \bar{y}^* \in \partial G(A\bar{x}).\end{cases} \tag{3.13}$$

Beweis. Nach Satz 3.5 ist $\bar{x} \in X$ genau dann eine Lösung von (3.10), wenn $0 \in \partial(F(\bar{x})+G(A\bar{x}))$ gilt. Wegen Voraussetzung (ii) sind Satz 3.10 und Satz 3.11 anwendbar; wir erhalten also

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

und damit die Existenz eines $\bar{y}^* \in \partial G(A\bar{x})$ mit $0 - A^*\bar{y}^* \in \partial F(\bar{x})$, d. h. (3.13) ist erfüllt.

Aus (3.13) folgt nun mit Lemma 3.16 die Gleichheit in den Fenchel–Youngschen Ungleichungen für F und G, d. h.

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

Durch Summieren beider Gleichungen erhalten wir (3.12).

Es bleibt zu zeigen, dass (3.13) auch die Lösungen von (3.11) charakterisiert. 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

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

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

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

Also ist

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

Zusammen mit der Gleichung (3.12) erhalten wir damit

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

d. h. $\bar{y}^*$ ist Lösung von (3.11). $\square$

Man nennt (3.13) auch *Fenchel-Extremalitätsbedingungen*; mit Hilfe von Lemma 3.16 können wir aus ihnen weitere äquivalente Optimalitätsbedingungen erzeugen, indem wir eine oder beide Subdifferentialinklusionen invertieren. Dies werden wir später nutzen, um numerisch zugängliche Optimalitätsbedingungen für Bildverarbeitungsprobleme herzuleiten (wobei F und G Diskrepanz- bzw. Regularisierungsterme sein werden).

Teil II

ALGORITHMEN

PROXIMALPUNKT-VERFAHREN

4

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 ein Gradientenabstieg

$$x^{k+1} = x^k - \tau_k \nabla J(x^k)$$

nicht wohldefiniert ist, da das konvexe Subdifferential eine Menge ist. Ein beliebiges Element im Subdifferential muss aber keine Abstiegsrichtung sein; 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 $x \mapsto \partial J(x)$.

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. Insbesondere werden wir stets X^* mit X identifizieren.

4.1 MONOTONE OPERATOREN

Zuerst betrachten wir die Klasse von Abbildungen der Form $x \mapsto \partial J(x)$. Für zwei normierte Räume X, Y wird durch eine Relation $R \subset X \times Y$ eine *mengenwertige Abbildung* $A : X \to \mathcal{P}(Y)$, geschrieben auch $A : X \rightrightarrows Y$, definiert durch

$$Ax = \{y \in Y : (x, y) \in R\}.$$

Wir definieren weiter für eine mengenwertige Abbildung $A : X \rightrightarrows Y$

- den *Definitionsbereich* $\operatorname{dom} A = \{x \in X : Ax \neq \emptyset\}$;
- das *Bild* $\operatorname{rg} A = \bigcup_{x \in X} Ax$;
- 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 $\lambda \in \mathbb{R}$ setzen wir

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

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

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

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. Aus der Definition folgt auch sofort, dass für eine konvexe Funktion $F : X \to \overline{\mathbb{R}}$ der Operator $x \mapsto \partial F(x)$ monoton ist. Ebenso sind für A, B monoton und $\lambda \geqslant 0$ auch λA und $A + B$ monoton. Wir werden aber noch eine stärkere Eigenschaft benötigen, die die Abgeschlossenheit von A garantiert. Ist ein monotoner Operator $A : X \rightrightarrows X$ (als Teilmenge von $X \times X$) nicht enthalten in einem weiteren monotonen Operator, so heisst A *maximal monoton*. Beispielsweise ist $A : \mathbb{R} \rightrightarrows \mathbb{R}$ definiert durch

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

monoton aber nicht maximal monoton, da A eine Teilmenge des durch $\tilde{A}t = \operatorname{sign}(t) = \partial(|\cdot|)(t)$ definierten monotonen Operators ist. Die folgende äquivalente Charakterisierung kann auch als formale Definition der maximalen Monotonie aufgefasst werden.

Lemma 4.1. $A : X \rightrightarrows X$ *ist maximal monoton genau dann, wenn für alle* $\bar{x} \in X$ *und* $\bar{x}^* \in X$ *die Bedingung*

$$(\bar{x}^* - x^*, \bar{x} - x)_X \geqslant 0 \qquad \textit{für alle } x \in X, x^* \in Ax$$

impliziert, dass $\bar{x}^* \in A\bar{x}$ *ist.*

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

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

monoton ist. Die Behauptung ist dann genau, dass dies keine echte Erweiterung darstellt. $\square$

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

Beweis. Anwenden von Lemma 4.1 auf $\lambda^{-1}\bar{x}^*$ liefert $\lambda^{-1}\bar{x}^* \in A\bar{x}$, d. h. $\bar{x}^* \in \lambda A\bar{x}$, und damit ist λA maximal monoton. □

Die gewünschte Abgeschlossenheit von maximal monotonen Operatoren folgt jetzt direkt aus der Charakterisierung in Lemma 4.1.

Folgerung 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

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

da x_n schwach und x_n^* stark in X konvergieren. Aus Lemma 4.1 folgt dann $x^* \in Ax$. □

Aus Lemma 4.1 erhalten wir auch eine Bedingung, wann ein monotoner Operator A maximal monoton ist.

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

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 und Combettes 2011, Theorem 21.1].)

Angenommen, $\operatorname{Id} + A$ wäre surjektiv, aber A nicht maximal monoton. Dann existieren $z \in X$ und $z^* \notin Az$ mit

$$(z^* - \tilde{x}^*, z - \tilde{x})_X \geqslant 0 \qquad \text{für alle } \tilde{x} \in X, \tilde{x}^* \in Ax. \tag{4.1}$$

Nach Annahme existiert nun für $z + z^* \in X$ ein $\bar{x} \in X$ und ein $\bar{x}^* \in A\bar{x}$ mit

$$\bar{x} + \bar{x}^* = z + z^*. \tag{4.2}$$

Einsetzen von $(\tilde{x}, \tilde{x}^*) = (\bar{x}, \bar{x}^*)$ in (4.1) ergibt dann

$$0 \leqslant (z^* - \bar{x}^*, z - \bar{x})_X = (\bar{x} - z, z - \bar{x})_X = -\|z - \bar{x}\|_X^2 \leqslant 0,$$

d. h. $\bar{x} = z$. Aus (4.2) folgt dann auch $z^* = \bar{x}^* \in A\bar{x}$, im Widerspruch zu $z^* \notin Az = A\bar{x}$. □

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

[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.

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

Beweis. Wir zeigen, dass $\mathrm{Id} + \partial F$ surjektiv ist. Für gegebenes $z \in X$ betrachten wir das Funktional $G : X \to \overline{\mathbb{R}}$,

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

Da die quadrierte Norm strikt konvex ist, ist G auch eigentlich, konvex und unterhalbstetig. Nach Satz 3.2 existiert also ein eindeutiges $\bar{x} \in X$ mit $G(\bar{x}) = \min_{x \in X} G(x)$, für welches nach Satz 3.5 und Satz 3.10 gilt

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

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

4.2 RESOLVENTEN UND PROXIMALPUNKTE

Wie wir im Beweis von Lemma 4.5 gesehen haben, ist $\mathrm{Id} + \partial F$ surjektiv und hat ein eindeutiges Urbild. Also ist $(\mathrm{Id} + \partial F)^{-1}$ einwertig, auch wenn ∂F dies nicht ist. Es besteht daher die Hoffnung, dieses Objekt anstelle des Subdifferentials – oder allgemeiner, eines monotonen Operators – für Algorithmen nutzen zu können.

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

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

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

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

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

$$\mathrm{prox}_F = \mathcal{R}_{\partial F}.$$

Wir müssen noch zeigen, dass auch die Resolvente 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 $\mathrm{Id} + A$ surjektiv, woraus $\mathrm{dom}\, \mathcal{R}_A = X$ folgt. Sei nun $x \in \mathcal{R}_A(x^*)$ und $z \in \mathcal{R}_A(z^*)$, d. h. $x^* \in x + Ax$ und $z^* \in z + Az$. Aus $x^* - x \in Ax$ und $z^* - z \in Az$ folgt mit der Monotonie von A, dass

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

gilt. Ist $x^* = z^*$, so muss 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 verwendet, dass die Proximalpunktabbildung einwertig ist, also Gleichheit statt Mengeninklusion gilt. □

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

Satz 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 und $\gamma > 0$ beliebig. 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 Lemma 4.7 und Lemma 3.16, dass

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

□

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

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

Setzen wir $v = \lambda w + z$, so ist $\bar{w} = \lambda^{-1}(\bar{v} - z)$ mit

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

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

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

Anwenden auf F^* unter Verwendung von $F^{**} = F$ 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

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

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 $X = \mathbb{R}$.

1. $F(t) = \frac{1}{2}|t|^2$, d. h. $\partial F(t) = \{t\}$ und damit ist $\operatorname{prox}_{\gamma F}(t) = (1+\gamma)^{-1}t$.
2. $F(t) = |t|$, d. h. $\partial F(t) = \operatorname{sign}(t)$; siehe (3.3). Also ist $w = \operatorname{prox}_{\gamma F}(t) = (\mathrm{Id} + \gamma \operatorname{sign})^{-1}(t)$ genau dann, wenn $t \in w + \gamma \operatorname{sign}(w)$ ist. Angenommen, für gegebenes t gelte diese Inklusion für ein $\bar w$. Wir machen nun eine Fallunterscheidung.

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

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

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

   Wir erhalten also

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

   Dieser Operator wird auch als *soft-shrinkage* bezeichnet.

3. $F(t) = \delta_{[-1,1]}(t)$. Nach Beispiel 3.14 (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. Die Beispiele 4.11 können 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 $F(x) = \frac{1}{2}\|x\|_2^2$

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

(ii) für $F(x) = \|x\|_1$

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

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

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

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

Viele weitere Beispiele findet man in [Parikh und 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 3.7, dass

$$\operatorname{prox}_{\gamma F}(x) = \begin{cases} \left(1 - \frac{\gamma}{\|x\|_X}\right)x & \|x\|_X > \gamma, \\ 0 & \|x\|_X \leqslant \gamma. \end{cases}$$

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

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

die *Projektion* von x auf C; die Proximalpunktabbildung verallgemeinert also die orthogonale Projektion auf Unterräume. 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 beliebige $\gamma > 0$. Dies legt nahe, diesen durch eine Fixpunktiteration zu berechnen: Wähle $x^0 \in X$ und setze für geeignete $\{\gamma_k\}_{k\in\mathbb{N}}$

$$x^{k+1} = \operatorname{prox}_{\gamma_k F}(x^k). \tag{4.3}$$

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 $\|Tx - Tz\|_X \leqslant \|x - z\|_X$. Sie heisst *fest nichtexpansiv* (englisch: „firmly nonexpansive“, wenn sogar gilt

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

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*

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

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

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

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

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

gilt. $\square$

Damit können wir nun die Konvergenz der Fixpunktiteration für Resolventen maximal monotoner Operatoren zeigen, woraus sofort die Konvergenz des Proximalpunktverfahrens (4.3) folgt.

Satz 4.14. *Sei $A : X \rightrightarrows X$ maximal monoton, und sei $\{\gamma_n\}_{n\in\mathbb{N}}$ mit $\sum_{n=0}^{\infty} \gamma_n^2 = \infty$. Erfüllt $\{x^n\}_{n\in\mathbb{N}} \subset \mathbb{R}^+$ die Rekursion*

$$x^{n+1} = \mathcal{R}_{\gamma_n A} x^n,$$

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

Beweis. Aus der Rekursion $x^{n+1} = \mathcal{R}_{\gamma_n A} x^n = (\mathrm{Id} + \gamma_n A)^{-1} x^n$ folgt $w^n := \gamma_n^{-1}(x^n - x^{n+1}) \in Ax^{n+1}$ und damit auch $x^{n+1} - x^{n+2} = \gamma_{n+1} w^{n+1}$. Da A monoton ist, gilt für $\gamma_{n+1} > 0$

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

Also ist $\{\|w^n\|_X\}_{n\in\mathbb{N}}$ nicht-negativ und monoton fallend und daher konvergent.

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

$$
\begin{aligned}
\|x^{n+1} - x^*\|_X^2 &= \|\mathcal{R}_{\gamma_n A} x^n - \mathcal{R}_{\gamma_n A} x^*\|_X^2 \\
&\leqslant \|x^n - x^*\|_X^2 - \|(\mathrm{Id} - \mathcal{R}_{\gamma_n A})x^n - (\mathrm{Id} - \mathcal{R}_{\gamma_n A})x^*\|_X^2 \\
&= \|x^n - x^*\|_X^2 - \gamma_n^2 \|w^n\|_X^2,
\end{aligned}
$$

und damit

$$
\|x^{n+1} - x^*\|_X^2 \leqslant \|x^0 - x^*\|_X^2 - \sum_{k=0}^{n} \gamma_k^2 \|w^k\|_X^2. \tag{4.4}
$$

Die Partialsummenfolge ist also monoton wachsend und beschränkt, und deshalb konvergiert $\sum_{n=0}^\infty \gamma_n^2 \|w^n\|_X^2$. Da die Folge $\{\gamma_n^2\}_{n\in\mathbb{N}}$ nach Annahme nicht summierbar ist, muss $\liminf_{n\to\infty} \|w^n\|_X^2 = 0$ gelten. Aus der Konvergenz von $\{\|w^n\|_X\}_{n\in\mathbb{N}}$ folgt damit $w^n \to 0$. Aus (4.4) erhalten wir ausserdem, dass $\{\|x^n - x^*\|\}_{n\in\mathbb{N}}$ beschränkt ist. Also ist auch $\{x^n\}_{n\in\mathbb{N}}$ beschränkt und hat daher eine schwach konvergente Teilfolge $x^{n_k} \rightharpoonup \bar{x}$. Folgerung 4.3 angewendet auf $\{x^{n_k}\}_{k\in\mathbb{N}}$, $x^{n_k} \rightharpoonup \bar{x}$, und $\{w^{n_k}\}_{k\in\mathbb{N}}$, $w^{n_k} \to 0$, ergibt dann $0 \in A\bar{x}$, d. h. jeder schwache Häufungspunkt von $\{x^n\}_{n\in\mathbb{N}}$ ist eine Nullstelle von A.

Wir zeigen nun, dass die gesamte Folge $\{x^n\}_{n\in\mathbb{N}}$ konvergiert. Seien $\bar{x}$ und $\hat{x}$ schwache Häufungspunkte und damit Nullstellen von A. Dann sind wegen (4.4) sowohl $\{\|x^n - \bar{x}\|_X\}_{n\in\mathbb{N}}$ als auch $\{\|x^n - \hat{x}\|_X\}_{n\in\mathbb{N}}$ monoton fallend und beschränkt und daher konvergent. (Man sagt, die Folge $\{x^n\}_{n\in\mathbb{N}}$ ist *Féjer-monoton*.) Also konvergiert auch

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

gegen ein $c \in \mathbb{R}$. Da $\bar{x}$ schwacher Häufungspunkt ist, existiert eine Teilfolge $\{x^{n_k}\}_{k\in\mathbb{N}}$ mit $x^{n_k} \rightharpoonup \bar{x}$; ebenso existiert eine Teilfolge $\{x^{n_l}\}_{l\in\mathbb{N}}$ mit $x^{n_l} \rightharpoonup \hat{x}$. Also ist

$$
(\bar{x}, \bar{x} - \hat{x})_X = \lim_{k\to\infty} (x^{n_k}, \bar{x} - \hat{x})_X = c = \lim_{l\to\infty} (x^{n_l}, \bar{x} - \hat{x})_X = (\hat{x}, \bar{x} - \hat{x})_X,
$$

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 $\{x^n\}_{n\in\mathbb{N}}$ gegen diesen. $\square$

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:

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

Die zugehörige Fixpunkt-Iteration ist dann

1. Wähle $p^k \in \partial F(x^k)$ (mit minimaler Norm).
2. 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 aufwendig ist. Eine Ausnahme ist jedoch, wenn F zusätzlich differenzierbar ist: Dann gilt $\partial F(x) = \{DF(x)\}$. In diesem Fall erhalten wir das *Proximal-Gradientenverfahren*

$$x^{k+1} = \operatorname{prox}_{\gamma_k G}(x^k - \gamma_k \nabla F(x^k)) \tag{5.2}$$

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

Um nun analog zum Proximalpunktverfahren Konvergenz zu zeigen, müssen wir die Nichtexpansivität des Gradienten voraussetzen (da wir für F ja keine Resolvente verwenden, die automatisch nichtexpansiv ist).

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

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

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

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

□

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* $0 < \gamma_{\min} \leqslant \gamma_k \leqslant L^{-1}$, *dann konvergiert* $x^k \rightharpoonup \bar{x}$ *mit* $0 \in J(\bar{x})$.

Beweis. Wir zeigen die Konvergenz, indem wir die Iteration (5.2) als Abstiegsverfahren auffassen. Wir definieren dafür $T : X \to X$,

$$T_\gamma(x) = \gamma^{-1}(x - \operatorname{prox}_{\gamma G}(x - \gamma\nabla F(x))),$$

und erhalten dadurch

$$x^{n+1} = \mathrm{prox}_{\gamma_n G}(x^n - \gamma_n \nabla F(x^n)) = x^n - \gamma_n T_{\gamma_n}(x^n)$$

und, nach Lemma 4.7,

$$T_{\gamma_n}(x^n) - \nabla F(x^n) \in \partial G(x^n - \gamma_n T_{\gamma_n}(x^n)). \tag{5.3}$$

Aus Lemma 5.1 folgt dann mit $x = x^n$, $y = x^{n+1} = x^n - \gamma_n T_{\gamma_n}(x^n)$, und $\gamma_n \leqslant L^{-1}$

$$\begin{aligned} F(x^n - \gamma_n T_{\gamma_n}(x^n)) &\leqslant F(x^n) - \gamma_n \left(\nabla F(x^n), T_{\gamma_n}(x^n)\right)_X + \frac{\gamma_n^2 L}{2}\|T_{\gamma_n}(x^n)\|_X^2 \\ &\leqslant F(x^n) - \gamma_n \left(\nabla F(x^n), T_{\gamma_n}(x^n)\right)_X + \frac{\gamma_n}{2}\|T_{\gamma_n}(x^n)\|_X^2. \end{aligned} \tag{5.4}$$

Damit gilt auch für alle $z \in X$ (unter Verwendung von (5.3) und $\nabla F(x) \in \partial F(x)$), dass

$$\begin{aligned} J(x^{n+1}) &= F(x^n - \gamma_n T_{\gamma_n}(x^n)) + G(x^n - \gamma_n T_{\gamma_n}(x^n)) \\ &\leqslant F(x^n) - \gamma_n \left(\nabla F(x^n), T_{\gamma_n}(x^n)\right)_X + \frac{\gamma_n}{2}\|T_{\gamma_n}(x^n)\|_X^2 \\ &\quad + G(z) + (T_{\gamma_n}(x^n) - \nabla F(x^n), x^n - \gamma_n T_{\gamma_n}(x^n) - z)_X \\ &\leqslant F(z) + (\nabla F(x^n), x^n - z)_X - \gamma_n \left(\nabla F(x^n), T_{\gamma_n}(x^n)\right)_X + \frac{\gamma_n}{2}\|T_{\gamma_n}(x^n)\|_X^2 \\ &\quad + G(z) + (T_{\gamma_n}(x^n) - \nabla F(x^n), x^n - z - \gamma_n T_{\gamma_n}(x^n))_X \\ &= J(z) + (T_{\gamma_n}, x^n - z)_X - \frac{\gamma_n}{2}\|T_{\gamma_n}(x^n)\|_X^2. \end{aligned} \tag{5.5}$$

Für $z = x^n$ folgt daraus mit (5.5)

$$J(x^{n+1}) \leqslant J(x^n) - \frac{\gamma_n}{2}\|T_{\gamma_n}(x^n)\|_X^2,$$

d. h. das Proximal-Gradientenverfahren ist ein Abstiegsverfahren. Für $z = x^*$ mit $J(x^*) = \min_{x \in X} J(x)$ folgt weiter

$$\begin{aligned} 0 \leqslant J(x^{n+1}) - J(x^*) &\leqslant (T_{\gamma_n}, x^n - x^*)_X - \frac{\gamma_n}{2}\|T_{\gamma_n}(x^n)\|_X^2 \\ &= \frac{1}{2\gamma_n}\left(\|x^n - x^*\|_X^2 - \|x^n - x^* - \gamma_n T_{\gamma_n}(x^n)\|_X^2\right) \\ &= \frac{1}{2\gamma_n}\left(\|x^n - x^*\|_X^2 - \|x^{n+1} - x^*\|_X^2\right). \end{aligned}$$

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

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

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

Da $J(x^n)$ monoton fallend ist, folgt

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

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

$$J(\bar{x}) \leqslant \liminf_{l \to \infty} J(x^{n_l}) = J(x^*). \tag{5.6}$$

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

Aus (5.6) folgt insbesondere $J(x^k) = \mathcal{O}(k^{-1})$. Um $J(x^k) \leqslant J(x^*) - \varepsilon$ zu erreichen, sind also $\mathcal{O}(\varepsilon^{-1})$ Iterationen notwendig. Durch eine geschickte Extrapolation lässt sich dies auf $\mathcal{O}(\varepsilon^{-1/2})$ reduzieren, was beweisbar optimal ist; siehe [Nesterov 1983]. (Dafür ist die Folge der Funktionswerte allerdings nicht mehr monoton fallend.) Die Iterationsvorschrift dafür ist

$$\begin{cases} x^{k+1} = \operatorname{prox}_{\gamma_k G}(\bar{x}^k - \gamma_k \nabla F(\bar{x}^k)) \\ \bar{x}^{k+1} = \left(1 - \dfrac{1-\tau_k}{\tau_{k+1}}\right) x^{k+1} + \dfrac{1-\tau_k}{\tau_{k+1}} x^k \end{cases} \tag{5.7}$$

für die (schwer zu motivierende) Folge

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

Ein Nachteil des expliziten Splittings ist die Notwendigkeit, für die Wahl einer zulässigen Schrittweite γ_n die Lipschitzkonstante L von ∇F zu kennen. Im Beweis von Satz 5.2 erkennt man, dass dies dazu dient, die Abschätzung (5.4) zu garantieren. Ist die Lipschitzkonstante nicht bekannt, kann man versuchen, die Gültigkeit von (5.4) durch Liniensuche zu erfüllen: Man startet mit $\gamma^0 > 0$ und verkleinert γ^k (etwa durch Halbieren) so lange, bis

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

erfüllt ist (was spätestens für $\gamma^k < L^{-1}$ der Fall ist). Dafür ist in jedem Schritt der Liniensuche eine Auswertung von F und eine von $\operatorname{prox}_{\gamma^k G}$ erforderlich.

5.2 IMPLIZITES SPLITTING

Auch mit Liniensuche bleiben die Anforderungen an die Schrittweite γ_n in expliziten Verfahren unbefriedigend. Dies kann man mit einem impliziten Splitting umgehen. Hier wenden wir die Proximalpunkt-Formulierung auf beide Subdifferential-Inklusionen in (5.1) an und erhalten

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

Um $\bar{p}$ aus den Gleichungen zu eliminieren, setzen wir $\bar{z} := \bar{x} + \gamma\bar{p}$ und $\bar{w} := \bar{x} - \gamma\bar{p}$. Dann ist $\bar{z} + \bar{w} = 2\bar{x}$, d. h.

$$\bar{w} = 2\bar{x} - \bar{z}.$$

Damit brauchen wir nur noch eine Rekursion für $\bar{z}$; diese gewinnen wir aus der produktiven Null

$$\bar{z} = \bar{z} + (\bar{x} - \bar{x}).$$

Durch Fixpunktiteration auf diesen Gleichungen erhalten wir das *Douglas–Rachford-Verfahren*

$$\begin{cases} x^{k+1} = \operatorname{prox}_{\gamma F}(z^k), \\ y^{k+1} = \operatorname{prox}_{\gamma G}(2x^{k+1} - z^k), \\ z^{k+1} = z^k + y^{k+1} - x^{k+1}. \end{cases} \tag{5.8}$$

Diese Iteration kann man durch Einführen von geeigneten Operatoren als Proximalpunkt-Verfahren schreiben, woraus mit etwas Aufwand (nach Nachweisen der starken Nichtexpansivität) die Konvergenz mit Satz 4.14 folgt; siehe etwa [Eckstein und Bertsekas 1992]. Wir betrachten hier gleich eine allgemeinere Variante, die sich in der Bildverarbeitung als sehr erfolgreich erwiesen hat.

5.3 PRIMAL-DUALE VERFAHREN

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 3.17 und Lemma 3.16 erhält man dafür die Fenchel-Extremalitätsbedingungen

$$\begin{cases} -A^*\bar{y} \in \partial F(\bar{x}), \\ \quad\;\; \bar{y} \in \partial G(A\bar{x}), \end{cases} \Leftrightarrow \begin{cases} -A^*\bar{y} \in \partial F(\bar{x}), \\ \quad\; A\bar{x} \in \partial G^*(\bar{y}). \end{cases} \tag{5.9}$$

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 (vergleiche (5.7)) einzuschieben:

$$\begin{cases} x^{k+1} = \operatorname{prox}_{\tau F}(x^k - \tau A^* y^k), \\ \bar{x}^{k+1} = 2x^{k+1} - x^k, \\ y^{k+1} = \operatorname{prox}_{\sigma G^*}(y^k + \sigma A \bar{x}^{k+1}). \end{cases} \tag{5.10}$$

Dieses *primal-duale Extragradienten-Verfahren* kann man nun als Proximalpunkt-Verfahren auffassen.[1] Dazu formen wir (5.10) so um, dass (x^n, y^n) und (x^{n+1}, y^{n+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

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

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 (5.10) äquivalent mit

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

Wäre M invertierbar, so verifiziert man leicht, dass $(M + T)^{-1} M z^k = (\operatorname{Id} + M^{-1} T)^{-1} z^k$ ist, es sich also um ein Proximalpunktverfahren für den Operator $M^{-1} T$ handelt. Dazu zeigen wir, dass M – unter Voraussetzungen an σ, τ – selbstadjungiert und positiv definit bezüglich des Skalarproduktes

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

ist.

[1] Eingeführt wurde das Verfahren in [Chambolle und Pock 2011]; der Zusammenhang mit Proximalpunktverfahren wurde in [He und Yuan 2012] beschrieben.

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 $z = (x,y) \in Z \setminus \{0\}$ beliebig. Dann ist

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

für $C := (1 - \|A\|_{L(X,Y)}\sqrt{\sigma\tau})\min\{\tau^{-1}, \sigma^{-1}\} > 0$. Also ist $(Mz,z)_Z > C\|z\|_Z^2$ für alle $z \neq 0$ 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

$$c_1\|z\|_Z \leqslant \|z\|_M \leqslant c_2\|z\|_Z \qquad \text{für alle } z \in Z. \tag{5.11}$$

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

Beweis. Wegen (5.11) ist für beliebige $z \in Z$ die Abbildung $f : Z \to \mathbb{R}$, $f(v) = (z,v)_Z$ ein (bezüglich $\|\cdot\|_M$) beschränktes Funktional. Nach dem Satz von Fréchet–Riesz existiert also ein eindeutiges $z^* \in Z$ mit

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

und

$$\|z^*\|_Z \leqslant c_2^{-1}\|z^*\|_M = c_2^{-1}\sup_{v\neq 0}\frac{(z,v)_Z}{\|v\|_M} \leqslant (c_1c_2)^{-1}\sup_{v\neq 0}\frac{(z,v)_Z}{\|v\|_Z} = (c_1c_2)^{-1}\|z\|_Z,$$

wieder mit Fréchet–Riesz für den Hilbertraum $(Z, (\cdot,\cdot)_Z)$. Die Abbildung $M^{-1} : Z \to Z$, $z \mapsto z^*$, definiert also einen beschränkten linearen Operator. □

Also ist $M^{-1}T$ wohldefiniert. Um die maximale Monotonie von $M^{-1}T$ zu beweisen, benötigen wir auch noch, dass die Skalarprodukte $(\cdot,\cdot)_M$ und $(\cdot,\cdot)_Z$ äquivalent sind.

Folgerung 5.5. *Es existiert ein $C > 0$, so dass gilt*

$$(z,z')_Z \leqslant C\,(z,z')_M \qquad \textit{für alle } z, z' \in Z.$$

Beweis. Aus der Polarisationsformel und (5.11) folgt

$$\begin{aligned}(z, z')_Z &= \frac{1}{4}\left(\|z+z'\|_Z^2 - \|z-z'\|_Z^2\right)\\ &\leqslant \frac{1}{4}\left(c_1^{-2}\|z+z'\|_M^2 - c_2^{-2}\|z-z'\|_M^2\right)\\ &\leqslant \max\{c_1^{-2}, c_2^{-2}\}\,(z, z')_M\,.\end{aligned}$$

□

Lemma 5.6. *Gilt* $\sigma\tau\|A\|_{L(X,Y)}^2 < 1$*, so ist* $M^{-1}T$ *maximal monoton.*

Beweis. Wir verwenden Lemma 4.1. Seien dafür $\bar z^*, \bar z \in Z$ und es gelte

$$(\bar z^* - z^*, \bar z - z)_Z \geqslant 0 \quad \text{für alle } z \in Z, z^* \in M^{-1}Tz.$$

Für $z^* \in M^{-1}Tz$ ist nun $Mz^* \in Tz$ und daher existieren $\xi \in \partial F(x)$ und $\eta \in \partial G^*(y)$ mit $z = (x, y)$ und $Mz^* = (\xi + A^*y, \eta - Ax)$. Weiterhin ergibt Folgerung 5.5

$$\begin{aligned}(Mz̄^* - Mz^*, \bar z - z)_Z &= (\bar z^* - z^*, \bar z - z)_M\\ &\geqslant C^{-1}\,(\bar z^* - z^*, \bar z - z)_Z \geqslant 0 \quad \text{für alle } z \in Z, Mz^* \in Tz.\end{aligned} \tag{5.12}$$

Setze nun $\bar\xi = \bar x^* - A^*\bar y$ und $\bar\eta = \bar y^* + A\bar x$ für $M\bar z^* = (\bar x^*, \bar y^*)$. Aus (5.12) folgt daher für alle $(x, y) \in Z$:

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

Da der erste Term nur von x und der zweite nur von y abhängt und beide unabhängig gewählt werden können, müssen beide nichtnegativ sein. Aus der maximalen Monotonie von Subdifferentialen folgt nun mit Lemma 4.1 $\bar\xi \in \partial F(\bar x)$ 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 $\bar z^* \in M^{-1}T\bar z$. Lemma 4.1 liefert nun die maximale Monotonie von $M^{-1}T$. □

Das primal-duale Extragradienten-Verfahren (5.10) ist also äquivalent mit der Proximalpunktiteration $z^{k+1} = \mathcal{R}_{M^{-1}T} z^k$ für $M^{-1}T$ maximal monoton, so dass aus Satz 4.14 die Konvergenz folgt.

Satz 5.7. *Erfüllen* F, G, A *die Voraussetzungen von Satz 3.17 und gilt* $\sigma\tau\|A\|_{L(X,Y)}^2 < 1$*, so konvergiert* $\{(x^k, y^k)\}_{k\in\mathbb{N}}$ *schwach in* Z *gegen eine Lösung* $(\bar x, \bar y)$ *von (5.9).*

Setzt man $A = \mathrm{Id}$, $\tau = \gamma$, $\sigma = \gamma^{-1}$ und $z^k = x^k - \gamma y^k$ und wendet Lemma 4.10 (ii) an, so erkennt man, dass die Iteration (5.10) äquivalent ist mit (5.8). Die Konvergenz des Douglas–Rachford-Verfahrens folgt aber wegen $\sigma\tau\|A\|_{L(X,Y)}^2 = 1$ nicht aus Satz 5.7.

Eine Schwierigkeit in der konkreten Durchführung des Verfahrens ist das Abbruchkriterium. Insbesondere steht für nichtdifferenzierbare Funktionen kein bequemes Äquivalent zur Residuumsnorm in den Optimalitätsbedingungen zur Verfügung. Aus Satz 3.17 folgt aber, dass in der Lösung $(\bar{x}, \bar{y})$ die Funktionswerte des primalen und des dualen Problems übereinstimmen. Das legt nahe, als Kriterium die *Dualitätslücke*

$$g(x, y) = F(x) + G(Ax) + F^*(-Ay) + G^*(y)$$

zu verwenden. Im Beweis von Satz 3.17 haben wir gesehen, dass $(\bar{x}, \bar{y})$ Sattelpunkte der Funktion

$$L(x, y) = F(x) + (y, Ax)_Y - G^*(y)$$

ist. Es gilt also für alle $(x, y) \in Z$

$$\inf_{x \in X} L(x, y) \leqslant L(\bar{x}, y) \leqslant L(\bar{x}, \bar{y}) \leqslant L(x, \bar{y}) \leqslant \sup_{y \in Y} L(x, y).$$

Daher ist wegen

$$g(x, y) \geqslant F(x) + G(Ax) - F(\bar{x}) - G(A\bar{x}) \geqslant 0$$

die Dualitätslücke eine obere Schranke für den Abstand zum minimalen Funktionswert. Wir können also abbrechen, falls $g(x^k, y^k) \leqslant \varepsilon$ ist für eine vorgegebene Schranke $\varepsilon > 0$. Allerdings ist nicht garantiert, dass $g(x^k, y^k) \to 0$ gilt (aus der Unterhalbstetigkeit folgt lediglich $0 \leqslant \liminf_{k \to \infty} g(x^k, y^k)$); auch kann die Dualitätslücke unendlich sein. Dies muss in konkreten Problemen (z. B. durch geeignete Modifikation des Kriteriums) berücksichtigt werden.

Teil III

BILDMODELLE

KONTINUIERLICHE BILDMODELLE

6

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 $\Omega \subset \mathbb{R}^d$ (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 $1 \leqslant p \leqslant \infty$ 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 $L^p(\Omega)^* = L^q(\Omega)$ für $1 \leqslant p < \infty$ und $p^{-1} + q^{-1} = 1$ ist. Die Ungleichung (1.1) wird dann zur *Hölder-Ungleichung*

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

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

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

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

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. Dann gilt dies auch für*

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

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

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

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.

Sei nun $\{u_n\}_{n\in\mathbb{N}}$ mit $u_n \to u$ in $L^p(\Omega)$. Dann existiert eine Teilfolge $\{u_{n_k}\}_{k\in\mathbb{N}}$ mit $u_{n_k}(x) \to u(x)$ fast überall.[1] Wegen der Unterhalbstetigkeit von φ und mit dem Lemma von Fatou gilt

$$\begin{aligned} \Phi(u) &= \int_\Omega \varphi(u(x))\,dx \\ &\leqslant \int_\Omega \liminf_{k\to\infty} \varphi(u_{n_k}(x))\,dx \\ &\leqslant \liminf_{k\to\infty} \int_\Omega \varphi(u_{n_k}(x))\,dx = \liminf_{k\to\infty} \Phi(u_{n_k}). \end{aligned}$$ □

Für solche Funktionale lässt sich das Subdifferential punktweise charakterisieren.

Lemma 6.2. *Sei* $\varphi : \mathbb{R} \to \overline{\mathbb{R}}$ *eigentlich, konvex und unterhalbstetig und* $\Phi : L^p(\Omega) \to \overline{\mathbb{R}}$ *wie in Lemma 6.1. Dann ist*

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

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

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

für alle $\tilde{u} \in L^p(\Omega)$.

[1] siehe z. B. [Dobrowolski 2010, Satz 4.25]

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

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

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 aus der obigen Ungleichung

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

Da A beliebig war, folgt daraus

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

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.[2]

Lemma 6.3. *Sei* $\varphi : \mathbb{R} \to \overline{\mathbb{R}}$ *eigentlich, konvex und unterhalbstetig, und sei* $\Phi : L^p(\Omega) \to \overline{\mathbb{R}}$ *wie in Lemma 6.1. Dann ist*

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

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 $L^p(\Omega)$.[3]

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

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.

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

Satz 6.5 (Fundamentallemma der Variationsrechnung[4]). *Sei* $u \in L^1_{loc}(\Omega)$ *mit*

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

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 $u \in L^1_{loc}(\Omega)$ heisst *schwach differenzierbar*, falls ein $v \in L^1_{loc}(\Omega)$ 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 Variationsrechnng ist dann die *schwache Ableitung* eindeutig definiert als $D^\alpha u := g$; falls u eine entsprechende klassische Ableitung besitzt, 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; 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 $x \geqslant 0$ und $Du(x) = -1$ für $x < 0$.

Wir können nun die *Sobolev-Räume* $W^{k,p}(\Omega)$ für $k \in \mathbb{N}_0$ und $1 \leqslant p \leqslant \infty$ definieren als

$$W^{k,p}(\Omega) = \{u \in L^p(\Omega) : D^\alpha u \in L^p(\Omega) \text{ für alle } |\alpha| \leqslant k\}.$$

[4] siehe z. B. [Dobrowolski 2010, Satz 5.1]

Diese sind Banachräume[5] versehen mit der Norm

$$\|u\|_{W^{k,p}} = \left(\sum_{|\alpha| \leqslant k} \|D^\alpha u\|_{L^p}^p \right)^{\frac{1}{p}} \qquad \text{für } 1 \leqslant p < \infty,$$
$$\|u\|_{W^{k,\infty}} = \max_{|\alpha| \leqslant k} \|D^\alpha u\|_{L^\infty}.$$

Für $p = 2$ handelt es sich um Hilberträume mit Skalarprodukt

$$(u, v)_{W^{k,2}} = \sum_{|\alpha| \leqslant k} (D^\alpha u, D^\alpha v)_{L^2}.$$

Man schreibt daher üblicherweise $H^k(\Omega) := W^{k,2}(\Omega)$.

Auch Sobolev-Funktionen lassen sich beliebig gut durch Testfunktionen (ohne kompakten Träger) approximieren.[6]

Satz 6.6. *Für alle $k \in \mathbb{N}$ und $1 \leqslant p < \infty$ liegt $C^\infty(\overline{\Omega}) \cap W^{k,p}(\Omega)$ 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. 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.

Satz 6.8 (Einbettungssatz von Sobolev, Morrey[7]). *Sei $1 \leqslant p, q < \infty$ und $\Omega \subset \mathbb{R}^d$ ein Lipschitz-Gebiet. Dann sind die folgenden Einbettungen stetig:*

$$W^{k,p}(\Omega) \hookrightarrow \begin{cases} L^q(\Omega) & p < \frac{d}{k} \textit{ und } p \leqslant q \leqslant \frac{dp}{d-p}, \\ L^q(\Omega) & p = \frac{d}{k} \textit{ und } p \leqslant q < \infty, \\ C(\overline{\Omega}) & p > \frac{d}{k}. \end{cases}$$

[5] siehe z. B. [Dobrowolski 2010, Satz 5.10]

[6] Die wurde bewiesen von Meyers und Serrin, in einer Arbeit die zu Recht bekannt ist sowohl für ihren Inhalt als auch für die Kürze ihres Titels, „$H = W$“. Für den Beweis siehe z. B. [Dobrowolski 2010, Satz 5.16].

[7] siehe z. B. [Adams und Fournier 2003, Theorem 4.12],[Dobrowolski 2010, Satz 6.24]

Auf solchen Gebieten lassen sich Sobolev-Funktionen auch stetig 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 Ω als kompakte Teilmenge enthält, existiert für alle $0 \leqslant k \leqslant m$ 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

$$|u|_{H^1} := \|\nabla u\|_{(L^p)^d} = \left(\int_\Omega |\nabla u(x)|_p^p \, dx\right)^{1/p}$$

für die *Seminorm* auf $H^1(\Omega)$. Damit ist

$$\|u\|_{W^{1,p}} = (\|u\|_{L^p}^p + |u|_{H^1}^p)^{1/p},$$

sowie

$$(u, v)_{W^{1,p}} = (u, v)_{L^2} + (\nabla u, \nabla v)_{(L^2)^d}.$$

Durch partielle Integration definieren wir nun für Eine Vektor-Funktion $u = (u_1, \dots, u_d)^T \in L^1_{loc}(\Omega)^d$ hat die schwache *Divergenz*

$$\operatorname{div} u = \sum_{i=1}^d \partial_i u_i \in L^1_{loc}(\Omega),$$

falls gilt

$$\int_\Omega \operatorname{div} u(x) \varphi \, dx = -\int_\Omega u \cdot \nabla \varphi \, dx \qquad \text{für alle } \varphi \in C_0^\infty(\Omega).$$

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 ∇.

Lemma 6.10. *Sei $\Omega \subset \mathbb{R}^d$ und $1 \leqslant p, q < \infty$. 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.

[8]siehe z. B. [Dobrowolski 2010, Satz 6.10]

Beweis. Sei eine Folge u_n mit den geforderten Eigenschaften gegeben. 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 $1 \leqslant k \leqslant d$. Nach Definition der schwachen Konvergenz und der schwachen Ableitung gilt dann

$$
\begin{aligned}
\int_\Omega u\partial_k\varphi\,dx &= \langle\partial_k\varphi, u\rangle_{L^p} = \lim_{n\to\infty}\langle\partial_k\varphi, u_n\rangle_{L^p} \\
&= \lim_{n\to\infty}\int_\Omega u_n\partial_k\varphi\,dx = \lim_{n\to\infty} -\int_\Omega \partial_k u_n\varphi\,dx \\
&= \lim_{n\to\infty} -\langle\varphi, \partial_k u_n\rangle_{L^q} = -\langle\varphi, \partial_k u\rangle_{L^q} \\
&= -\int_\Omega v_k\varphi\,dx,
\end{aligned} \tag{6.2}
$$

d. h. $v_k = \partial_k u$ für alle $1 \leqslant k \leqslant d$ nach Definition der schwachen Ableitung und damit $\nabla u = v$. □

Aus der Gleichheit des zweiten und des vorletzten Terms in (6.2) folgt, dass formal $\nabla^* = -\operatorname{div}$ gilt. Dies kann durch korrekte Wahl der Bild- und Definitionsbereiche auch rigoros bewiesen werden; siehe z. B. [Kurula und 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 $1 \leqslant p \leqslant \infty$

$$\Pi_0 := L^p(\Omega) \to L^p(\Omega), \qquad \Pi_0(u) = u - \mu(u).$$

Lemma 6.11 (Poincaré–Wirtinger-Ungleichung[9]). *Sei Ω ein konvexes Lipschitz-Gebiet und $1 \leqslant p < \infty$. Dann existiert ein $c > 0$ so dass für alle $u \in W^{1,p}(\Omega)$ gilt*

$$\|\Pi_0 u\|_{L^p} \leqslant c|u|_{W^{1,p}}.$$

Damit können wir nun das Gewünschte zeigen.

Satz 6.12. *Sei Ω ein Lipschitz-Gebiet und $1 < p < \infty$. Dann ist*

$$\mathcal{R} : L^p(\Omega) \to \overline{\mathbb{R}}, \qquad \mathcal{R}(u) := \begin{cases} \frac{1}{p}|u|_{W^{1,p}}^p & u \in W^{1,p}(\Omega), \\ \infty & \textit{sonst} \end{cases}$$

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 $\mathcal{R}(u_n) \to \infty$.

[9]siehe z. B. [Dobrowolski 2010, Satz 6.23]

Beweis. Da $\mathcal{R}(u) = 0$ für $u \equiv 0 \in W^{1,p}(\Omega)$, ist $\mathcal{R}$ 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 $\mathcal{R}$ 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 wegen $u \notin W^{1,p}(\Omega)$ nichts zu zeigen. Ansonsten ist $\{\nabla u_n\}_{n\in\mathbb{N}}$ beschränkt in $L^p(\Omega)$ und hat daher ($L^p(\Omega)$ ist reflexiv 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 $u_{n_k} \rightharpoonup u$, folgt aus Lemma 6.10 also $\nabla u_{n_k} \rightharpoonup \nabla u$. Die Unterhalbstetigkeit der Norm in $L^p(\Omega)^d$ ergibt dann

$$\|\nabla u\|_{(L^p)^d} \leqslant \liminf_{k\to\infty} \|\nabla u_{n_k}\|_{(L^p)^d}.$$

Da wir dieses Argument auf jede Teilfolge von $\{u_n\}_{n\in\mathbb{N}}$ anwenden können, gilt dies für die gesamte Folge.

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 $|u_n|_{W^{1,p}} \leqslant M$ für unendlich viele $n \in \mathbb{N}$. Dies bedeutet insbesondere $u_n \in W^{1,p}(\Omega)$ für diese n. Aus der Poincaré–Wirtinger-Ungleichung[10] folgt dann $\|\Pi_0 u_n\|_{L^p} \leqslant c|u_n|_{W^{1,p}} \leqslant cM$, ein Widerspruch. □

Wir berechnen schliesslich die Richtungsableitungen von $\mathcal{R}$.

Lemma 6.13. *Seien* $u, h \in \operatorname{dom}\mathcal{R} = W^{1,p}(\Omega)$ *für* $1 < p < \infty$. *Dann gilt*

$$\mathcal{R}'(u;h) = \int_\Omega |\nabla u|_p^{p-2} \nabla u \cdot \nabla h \, dx.$$

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 3.3 folgt daraus, dass für fast alle $x \in \Omega$

$$\varphi_x(t) := \frac{1}{t}\left(\frac{1}{p}|\nabla u(x) + t\nabla h(x)|_p^p - \frac{1}{p}|\nabla u(x)|_p^p\right)$$

monoton steigend ist mit $|\varphi_x(t)| \leqslant |\varphi_x(1)| \leqslant \frac{1}{p}|\nabla h(x)|^p \in L^1(\Omega)$ (wegen $h \in W^{1,p}(\Omega)$) für $t \in (0,1]$. Nach dem Satz von Lebesgue[11] können wir also den Grenzübergang $t \to 0^+$ unter dem Integral durchführen und erhalten

$$\mathcal{R}'(u;h) = \int_\Omega \lim_{t\to 0^+} \varphi_x(t)\, dx = \int_\Omega |\nabla u|_p^{p-2}\nabla u \cdot \nabla h \, dx. \qquad \square$$

Da diese Richtungsableitung linear in h ist, ist $\mathcal{R}$ Gâteaux-differenzierbar, und wir erhalten mit Satz 3.6 eine Charakterisierung des Subdifferentials.

[10] siehe z. B. [Dobrowolski 2010, Satz 6.23]
[11] siehe z. B. [Dobrowolski 2010, Satz 4.11]

Folgerung 6.14. *Für* $u \in W^{1,p}(\Omega)$, $1 < p < \infty$, *gilt* $\xi \in \partial\mathcal{R}(u)$ *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 u. a. 2006, 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 $x \geqslant 0$ 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 $\delta_0 \in \mathcal{M}(\Omega) = C(\overline{\Omega})^*$; siehe Beispiel 1.2 (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 vektorwertiges Maß, d. h. ein Element im Dualraum $\mathcal{M}(\Omega; \mathbb{R}^d)$ des Raums $C(\overline{\Omega}; \mathbb{R}^d)$ der stetigen Funktionen von Ω nach $\mathbb{R}^d$, versehen mit der Norm[12]

$$\|\varphi\|_{(C)^d} := \sup_{x \in \overline{\Omega}} |\varphi(x)|_2.$$

Dann ist $\nabla u := v \in \mathcal{M}(\Omega; \mathbb{R}^d)$ der sehr schwache Gradient von u genau dann, wenn gilt

$$\int_\Omega u \operatorname{div} \varphi\, dx = -\int_\Omega \varphi \cdot dv = \sum_{i=1}^{d} \int_\Omega \varphi_i\, dv_i \qquad \text{für alle } \varphi \in C_0^\infty(\Omega; \mathbb{R}^d),$$

[12]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.

und wir können für sehr schwach differenzierbare u wegen der Dichtheit von $C_0^\infty(\Omega)$ in $C(\overline{\Omega})$ schreiben

$$\mathrm{TV}(u) := \|\nabla u\|_{(\mathcal{M})^d} = \sup\left\{\int_\Omega u \operatorname{div} \varphi \, dx : \varphi \in C_0^\infty(\Omega; \mathbb{R}^d) \text{ mit } \|\varphi\|_{(C)^d} \leqslant 1\right\}.$$

Hat u keinen sehr schwachen Gradienten, setzen wir $\mathrm{TV}(u) = \infty$. Man nennt $\mathrm{TV}(u)$ die *totale Variation* von u.[13] Für die eingangs definierte Funktion ist zum Beispiel

$$\mathrm{TV}(u) = \sup_{\|\varphi\|_C \leqslant 1} \int_{-1}^{1} u(x)\varphi'(x) \, dx = \sup_{\|\varphi\|_C \leqslant 1} 2\varphi(0) = 2,$$

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,[14] dass $\mathrm{TV}(u) = \|\nabla u\|_{(L^1)^d}$ 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 $\mathrm{TV}(0) = 0$), 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 $\Omega \subset \mathbb{R}^d$ ein Lipschitz-Gebiet und $1 \leqslant p < \infty$. 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 $v = \nabla u$.

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 $1 \leqslant p < \infty$. Dann ist* $\mathrm{TV} : L^p(\Omega) \to \overline{\mathbb{R}}$ *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

$$u_n(x) := \begin{cases} 1 & x \geqslant 1/n, \\ nx & |x| < 1/n, \\ -1 & x < -1/n, \end{cases}$$

[13] Im Sinne der Maßtheorie ist $\mathrm{TV}(u)$ eigentlich die totale Variation $|\nabla u|(\Omega)$ des Maßes ∇u; die Bezeichnung hat sich in diesem Kontext trotzdem eingebürgert.

[14] siehe z. B. [Brezis 2010, Exercise 4.25]

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[15]

$$\|u\|_{BV} := \|u\|_{L^1} + TV(u).$$

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[16]). *Sei Ω ein Lipschitz-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 Lipschitz-Gebiet ist.

Satz 6.18 (Einbettung[17]). *Für alle $1 \leqslant p \leqslant d/(d-1)$ ist die Einbettung $BV(\Omega) \hookrightarrow L^p(\Omega)$ stetig.*

Satz 6.19 (Fortsetzung[18]). *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[19]). *Sei $1 \leqslant p \leqslant d/(d-1)$. Dann existiert ein $c > 0$ so dass für alle $u \in BV(\Omega)$ gilt*

$$\|\Pi_0 u\|_{L^p} \leqslant c\,TV(u).$$

Insbesondere gilt für $\{u_n\}_{n\in\mathbb{N}} \subset L^p(\Omega)$ mit $\|\Pi_0 u_n\|_{L^p} \to \infty$ auch $TV(u_n) \to \infty$.

[15] siehe z. B. [Attouch u. a. 2006, Theorem 10.1.1]
[16] siehe z. B. [Attouch u. a. 2006, Theorem 10.1.2]
[17] siehe z. B. [Attouch u. a. 2006, Theorem 10.1.3]
[18] siehe z. B. [Ambrosio u. a. 2000, Proposition 3.21]
[19] siehe z. B. [Ambrosio u. a. 2000, Theorem 3.44] zusammen mit der Einbettung von $BV(\Omega)$ in $L^p(\Omega)$

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 u. a. 2000], [Ziemer 1989] oder [Attouch u. a. 2006, Kapitel 10.3] verwiesen; wir betrachten lediglich einige beispielhafte Resultate.

Wir betrachten zuerst die charakteristische Funktion χ_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

$$\int_\Omega \chi_E \operatorname{div} \varphi \, dx = \int_E \operatorname{div} \varphi \, dx = \int_{\partial E} \varphi \cdot \nu \, d\mathcal{H}^{d-1},$$

d. h. $\nabla \chi_E = -\nu \mathcal{H}^{d-1}(\partial E \cap \Omega)$, wobei ν die äussere Einheitsnormale an $\partial\Omega$ und $\mathcal{H}^{d-1}$ das $(d-1)$-dimensionale Hausdorff-Maß ist. Man kann nun zeigen,[20] dass

$$TV(\chi_E) = \mathcal{H}^{d-1}(\partial E \cap \Omega) < \infty$$

und damit $\chi_E \in BV(\Omega)$ 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.

Satz 6.21 (Koflächenformel[21]). *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 u. a. 2000, Kapitel 3.7]. Diese tiefe Verbindung von

[20] siehe z. B. [Ziemer 1989, Remark 5.4.2]
[21] siehe z. B. [Ziemer 1989, Theorem 5.4.4]

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 3.11) schreiben

$$\partial TV(u) = \nabla^* \partial(\|\cdot\|_{\mathcal{M}^d})(\nabla u),$$

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 und Lorenz 2011, Seite 346–355]; wir nehmen in Folge an, dass η regulär genug und insbesondere $\eta \in C(\overline{\Omega};\mathbb{R}^d)$ ist, so dass wir Lemma 3.16 anwenden können. Wir erhalten dann

$$\nabla u \in \partial\delta_{B_{C(\overline{\Omega};\mathbb{R}^d)}}(\eta),$$

was eine deutlich einfachere Darstellung erlaubt. Insbesondere entfällt nach Diskretisierung der Unterschied zwischen Prädual- und Dualraum, sodass die Verfahren aus Abschnitt 5.3 auf diese Darstellung angewendet werden können.

DISKRETE BILDMODELLE

7

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 $u_h : \mathbb{R}^n \to \mathbb{R}$, 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 sind) zu. Dem kontinuierlichen Grauwert $u(ih, jh)$ entspricht also der diskrete Grauwert $(u_h)_{ij}$ für $1 \leqslant i \leqslant N$ und $1 \leqslant j \leqslant m$.[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

$$\|u\|_2 = \sqrt{(u, v)} = h \left(\sum_{i=1}^{N} \sum_{j=1}^{M} |u_{ij}|^2 \right)^{1/2}.$$

Analog definieren auch die diskreten p-Normen für $p \neq 2$ als

$$\|u\|_p = \left(h^2 \sum_{i=1}^{N} \sum_{j=1}^{M} |u_{ij}|^p \right)^{1/p}, \quad 1 \leqslant p < \infty,$$
$$\|u\|_\infty = \max_{\substack{1 \leqslant i \leqslant N \\ 1 \leqslant j \leqslant M}} |u_{ij}|, \quad p = \infty.$$

[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

$$(\nabla_h u)_{ijk} = (\partial_k^{h,+})_{ij}.$$

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

$$\mathrm{TV}_h(u) = h^2 \sum_{i=1}^{N} \sum_{j=1}^{M} |(\nabla_h u)_{ij}|_2 = \||\nabla_h u|_2\|_1.$$

[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 $\|u\|_2 \leqslant 1$ 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. $(u, v) \geqslant 1$. Wegen $\|v\|_2 \leqslant \|u\|_2 \leqslant 1$ ist dann

$$1 \leqslant (u, v) \leqslant \|u\|_2\|v\|_2 \leqslant 1,$$

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 $j \in \{1, \dots, M\}$

$$0 = v_{Nj} = \lambda u_{Nj} = \lambda v_{N-1,j} = \lambda^2 u_{N-1,j} = \cdots = \lambda^N u_{1j},$$

und damit $u = 0$ im Widerspruch zur Annahme. Die zweite Abschätzung folgt analog.

Mit dieser Abschätzung erhalten wir für alle $\|u\|_2 \leqslant 1$, dass

$$\begin{aligned}
\|\nabla_h u\|_2^2 &= h^2 \left(\sum_{i=1}^{N-1}\sum_{j=1}^{M}\left(\frac{u_{i+1,j}-u_{ij}}{h}\right)^2\right) + h^2\left(\sum_{i=1}^{N}\sum_{j=1}^{M-1}\left(\frac{u_{i,j+1}-u_{ij}}{h}\right)^2\right)\\
&= \left(\sum_{i=1}^{N-1}\sum_{j=1}^{M} u_{i+1,j}^2 + u_{ij}^2 - 2u_{i+1,j}u_{ij}\right) + \left(\sum_{i=1}^{N}\sum_{j=1}^{M-1} u_{i,j+1}^2 + u_{ij}^2 - 2u_{i,j+1}u_{ij}\right)\\
&\leqslant 4\left(\sum_{i=1}^{N}\sum_{j=1}^{M} u_{ij}^2\right) - 2\left(\sum_{i=1}^{N-1}\sum_{j=1}^{M} u_{i+1,j}u_{ij}\right) - 2\left(\sum_{i=1}^{N}\sum_{j=1}^{M-1} u_{i,j+1}u_{ij}\right)\\
&< \frac{4}{h^2}\|u\|_2^2 + \frac{2}{h^2} + \frac{2}{h^2} \leqslant \frac{8}{h^2},
\end{aligned}$$

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

ENTRAUSCHEN

8

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 $\eta : [0,1]^{\to}\mathbb{R}$. 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

$$\min_{u \in X} \mathcal{F}(u) + \alpha\mathcal{R}(u)$$

und untersuchen Existenz, Optimalitätsbedingungen und numerische Lösung für drei konkrete Beispiele.

8.1 L^2-H^1-ENTRAUSCHEN

Ist die Störung η normalverteilt mit Erwartungswert 0 und Varianz σ, so bietet sich für $\mathcal{F}$ die quadrierte L^2-Norm von $f - u$ an. Für $\mathcal{R}$ setzen wir zunächst der Einfachheit halber die H^1-Seminorm an, betrachten also $\Omega = (0,1)^2$ und

$$\min_{u \in L^2(\Omega)} \frac{1}{2}\|f - u\|_{L^2(\Omega)}^2 + \frac{\alpha}{2}|u|_{H^1(\Omega)}^2. \tag{8.1}$$

Wir kommen schnell zur Sache, denn wir sind gut vorbereitet. Wir wenden Satz 3.2 an. Wegen $\operatorname{dom}\mathcal{R} = H^1(\Omega) \subset L^2(\Omega) = \operatorname{dom}\mathcal{F}$ ist das Funktional eigentlich. Da $u \mapsto f - u$ stetig und $t \mapsto \frac{1}{2}t^2$ stetig und monoton steigend ist, ist $\mathcal{F}$ nach Lemma 2.2 schwach unterhalbstetig; und wegen der Konvexität der Norm und strikten Konvexität von $t \mapsto \frac{1}{2}t^2$ ist $\mathcal{F}$ strikt konvex. Die entsprechenden Eigenschaften von $\alpha\mathcal{R}$ folgen aus Satz 6.12. Somit ist das gesamte Funktional strikt konvex und unterhalbstetig. Schließlich gilt offensichtlich $\mathcal{F}(u_n) \to \infty$ für $\|u_n\|_{L^2} \to \infty$, d. h. das Funktional ist koerziv. Nach Satz 3.2 hat (8.1) also eine eindeutige Lösung $\bar{u} \in \operatorname{dom}\mathcal{R} = H^1(\Omega)$.

Nach Satz 3.5 ist das genau dann der Fall, wenn $0 \in \partial(\mathcal{F} + \alpha\mathcal{R})(\bar{u})$ gilt. Nun ist $\mathcal{F}$ Fréchet-differenzierbar mit $\mathcal{F}'(u) = u - f$ (Fréchet-differenzierbarkeit der quadrierten Norm im

Hilbertraum und Kettenregel, Satz 2.5). Also ist $\mathcal{F}$ insbesondere stetig auf ganz $L^2(\Omega)$, und Anwendung der Summenregel 3.10 ergibt

$$f - \bar{u} \in \partial(\alpha\mathcal{R})(\bar{u}).$$

Nach Folgerung 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

$$(f_h, v_h) = \alpha(\nabla_h u_h, \nabla_h v_h) + (u_h, v_h) = (\alpha\nabla_h^T\nabla_h u_h + u_h, v_h) \quad \text{für alle } v_h \in \mathbb{R}^{NM},$$

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 $K_h := \alpha\nabla_h^T\nabla_h + \mathrm{Id} \in \mathbb{R}^{NM\times NM}$, zum Beispiel mit dem CG-Verfahren.[1]

Ganz analog geht man für $\mathcal{F}(u) = \frac{1}{q}\|f-u\|_{L^q}^q$ und $\mathcal{R}(u) = \frac{1}{p}|u|_{W^{1,p}}^p$ für $1 < q \leqslant p < \infty$ 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 $\mathcal{F}$ strikt konvex, unterhalbstetig und koerziv auf L^2. Ebenso ist $\mathcal{R} = \alpha TV$ nach Satz 6.16 eigentlich, konvex und unterhalbstetig mit $\operatorname{dom} TV = BV(\Omega) \hookrightarrow L^2(\Omega) = \operatorname{dom}\mathcal{F}$ wegen Satz 6.18 und $d = 2$. Aus Satz 3.2 folgt nun wieder die Existenz eines eindeutigen Minimierers $\bar{u} \in \operatorname{dom} TV = BV(\Omega)$.

Aus Satz 3.5 und der Summenregel erhalten wir wie zuvor die Optimalitätsbedingungen

$$f - \bar{u} \in \partial(\alpha TV)(\bar{u}),$$

[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 u. a. 1992] vorgeschlagen haben, auch als *ROF-Modell* bekannt

die wir mangels einer expliziten Darstellung des Subdifferentials nicht weiter ausreizen können. Für die Diskretisierung

$$\min_{u_h \in \mathbb{R}^{NM}} \frac{1}{2}\|u_h - f_h\|_2^2 + \alpha\||\nabla_h u_h|_2\|_1 \tag{8.2}$$

können wir jedoch den Satz von Fenchel–Rockafellar für $A = \nabla_h$ sowie Lemma 3.16 anwenden und erhalten die Existenz eines $\bar{y}_h \in \mathbb{R}^{2NM}$ mit

$$\begin{cases} \operatorname{div}_h \bar{y}_h = \bar{u}_h - f_h, \\ \nabla_h \bar{u}_h \in \partial\delta_{\{v_h : \||v_h|_2\|_\infty \leqslant \alpha\}}(\bar{y}_h). \end{cases} \tag{8.3}$$

Auf (8.3) können wir nun das primal-duale Extragradienten-Verfahren (5.10) anwenden. Unter Verwendung von Beispiel 4.12 (i) und Lemma 4.10 (i) erhält man für $\mathcal{F}$ komponentenweise

$$[\operatorname{prox}_{\tau\mathcal{F}}(v_h)]_{ij} = \frac{1}{1+\tau}(v_{ij} + \tau f_{ij}).$$

Für $(\alpha\mathcal{R})^*$ entspricht die Proximalpunkt-Abbildung der Projektion auf $\{v_h : \||v_h|_2\|_\infty \leqslant \alpha\}$, ist also – analog zu Beispiel 4.12 (iii) – komponentenweise gegeben durch

$$[\operatorname{prox}_{\sigma(\alpha\mathcal{R})^*}(v_h)]_{ijk} = \frac{\alpha v_{ijk}}{\max\{\alpha, |v_{ij}|_2\}}.$$

Das komplette Verfahren ist im folgenden Algorithmus zusammengefasst, der nach Satz 5.7 für $\sigma\tau < \|\nabla_h\|^{-2} = h^2/8$ (Lemma 7.1) gegen eine Lösung von (8.2) konvergiert.

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**

- $u^k = (u^{k-1} + \tau(f + \operatorname{div}_h y^{k-1}))/(1+\tau)$
- $\bar{u}^k = 2u^k - u^{k-1}$
- $w = y^{k-1} + \sigma\nabla_h \bar{u}^k$
- $y^k = \alpha w / \max\{\alpha, |w|\}$

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 können einzelne Datenleitungen defekt sein und nur Rauschen übermitteln. Ein Modell für solch eine Störung 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 das Problem[3]

$$\min_{u \in L^1(\Omega)} \|f-u\|_{L^1(\Omega)} + \alpha TV(u).$$

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.

Nun sind beide Summanden nicht-differenzierbar; die Summenregel und Lemma 3.8 liefert also 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

$$\begin{cases} \operatorname{div}_h \bar{y}_h \in \operatorname{sign}(\bar{u}_h - f_h), \\ \nabla_h \bar{u}_h \in \partial\delta_{\{v_h : \||v_h|_2\|_\infty \leqslant \alpha\}}(\bar{y}_h). \end{cases}$$

[3] zuerst untersucht in [Chan und 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 $\mathcal{F}$ komponentenweise gegeben durch

$$[\mathrm{prox}_{\tau\mathcal{F}}(v)]_{ij} = (|v_{ij} - f_{ij}| - \tau)^{+}\,\mathrm{sign}(v_{ij} - f_{ij}) + f_{ij} = \begin{cases} f_{ij} & |v_{ij} - f_{ij}| \leqslant \tau, \\ v_{ij} - \tau & v_{ij} - f_{ij} > \tau, \\ v_{ij} + \tau & v_{ij} - f_{ij} < -\tau. \end{cases}$$

(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 also

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**

- $r = u^{k-1} - f + \tau \operatorname{div}_h y^{k-1}$
- $u^k = (|r| - \tau)^{+} \operatorname{sign}(r) + f$
- $\bar{u}^k = 2u^k - u^{k-1}$
- $w = y^{k-1} + \sigma\nabla_h \bar{u}^k$
- $y^k = \alpha w / \max\{\alpha, |w|\}$

ENTFALTEN

9

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

$$\begin{aligned} f(x) = \int_{B_r(x)} u^0(y)\,dy = \int_{\mathbb{R}^2} u^0(y)\chi_{B_r(x)}(y)\,dy &= \int_{\mathbb{R}^2} u^0(y)\chi_{B_r(0)}(x-y)\,dy \\ &= \int_{\mathbb{R}^2} u^0(x-y)\chi_{B_r(0)}(y)\,dy, \end{aligned}$$

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* χ_{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 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* $1 \leqslant p \leqslant \infty$ *ist* $k \star u \in L^p(\Omega')$, *und es gilt*

$$\|k \star u\|_{L^p(\Omega')} \leqslant \|k\|_{L^1(\omega)}\|u\|_{L^p(\Omega)}.$$

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 und der obigen Substitution

$$\int_{\Omega'} |(k \star u)(x)|\,dx \leqslant \int_\omega \int_{\Omega'} |u(x-y)|\,dx\,|k(y)|\,dy \leqslant \|k\|_{L^1(\omega)}\|u\|_{L^1(\Omega)}.$$

Damit ist insbesondere $k \star u \in L^1(\Omega')$, und wegen der Existenz dieses Integrals war die Anwendung des Satzes von Fubini gerechtfertigt.

Für $1 < p < \infty$ wenden wir zuerst die Höldersche Ungleichung an (mit $1/p + 1/p' = 1$) und erhalten wie zuvor

$$\begin{aligned}
\int_{\Omega'} |(k \star u)(x)|^p\,dx &\leqslant \int_{\Omega'} \left(\int_\omega |u(x-y)|^p |k(y)|\,dy\right) \left(\int_\omega |k(y)|\,dy\right)^{p/p'} dx \\
&= \int_\omega \int_{\Omega'} |u(x-y)|^p\,dx\,|k(y)|\,dy\,\|k\|_{L^1(\omega)}^{p/p'} \\
&\leqslant \|u\|_{L^p(\Omega)}^p \|k\|_{L^1(\omega)}^{p/p'+1},
\end{aligned}$$

woraus 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

$$A : L^p(\Omega) \to L^p(\Omega'), \qquad Au = k \star u,$$

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* $A : L^p(\Omega) \to L^p(\Omega')$, $Au = k \star u$ *ist* $A^* : L^{p'}(\Omega') \to L^{p'}(\Omega)$ *gegeben durch* $A^*u = \bar{k} \star u$, *mit* $\bar{k}(y) = k(-y)$ *für alle* $y \in \omega$.

Beweis. Für $w \in L^{p'}(\Omega')$ gilt nach Fubini

$$\begin{aligned}
\langle w, k \star u\rangle_{L^p(\Omega')} &= \langle w, u \star k\rangle_{L^p(\Omega')} \\
&= \int_{\Omega'} w(x) \int_{\Omega} k(x-y)u(y)\, dy\, dx \\
&= \int_{\Omega} u(y) \int_{\Omega'} k(x-y)w(x)\, dx\, dy \\
&= \langle u, \bar{k} \star w\rangle_{L^{p'}(\Omega)},
\end{aligned}$$

wobei wieder k für $x - y \notin \omega$ durch Null fortgesetzt sein soll. □

Für ein gegebenes unscharfes Bild $f \in L^p(\Omega')$ mit $1 \leqslant p < \infty$ und Faltungskern $k \in L^1(\omega)$ betrachten wir nun den Diskrepanzterm

$$\mathcal{F} : L^p(\Omega) \to \mathbb{R}, \qquad \mathcal{F}(u) = \frac{1}{p}\|k \star u - f\|^p_{L^p(\Omega')}.$$

Da $u \mapsto k \star u$ ein linearer beschränkter Operator ist, ist $\mathcal{F}$ nach Lemma 2.2 und Lemma 3.1 eigentlich, konvex und unterhalbstetig. Für die Anwendung der direkten Methode benötigen wir nur noch die Koerzivität, wofür dank des Regularisierungsterms 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 $\mathcal{F}(u) \to \infty$ für $c \to \infty$.

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

$$\min_{u \in L^2} \frac{1}{2}\|k \star u - f\|^2_{L^2} + \alpha \mathrm{TV}(u). \tag{9.1}$$

Satz 9.3. *Das Problem* (9.1) *hat eine Lösung* $\bar{u} \in BV(\Omega)$. *Ist* $A : u \mapsto k \star u$ *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. Dafür zerlegen wir $u \in L^2(\Omega)$ in einen Anteil mit Mittelwert 0 und einen konstanten Anteil (der gleich dem Mittelwert ist), d. h. $u = \Pi_0 u + \mu(u)$. Sei nun $\{u_n\}_{n\in\mathbb{N}} \subset L^2(\Omega)$ eine Folge mit

$$\|\Pi_0 u_n\|_{L^2(\Omega)} + \|\mu(u_n)\|_{L^2(\Omega)} \geqslant \|u_n\|_{L^2(\Omega)} \to \infty.$$

Dann muss wenigstens einer der beiden Terme auf der linken Seite auch unbeschränkt sein. Ist der erste Term unbeschränkt, so können wir wegen $1 \leqslant 2 = d/d-1$ die Poincaré–Wirtinger-Ungleichung (Satz 6.20) anwenden und erhalten wegen $\nabla\mu(u) = 0$

$$TV(u_n) = TV(\Pi_0(u_n)) \geqslant \|\Pi_0 u_n\|_{L^2(\Omega)} \to \infty.$$

Ansonsten muss der zweite Term unbeschränkt sein. Da $\mu(u)$ konstant ist, gilt $k \star \mu(u) = C(k)\mu(u)$ mit $C(k) = \int_\omega k(y)\,dy$. Aus der Linearität der Faltung folgt dann mit Hilfe der umgekehrten Dreiecksungleichung und Lemma 9.1, dass

$$\|k \star u_n - f\|_{L^2(\Omega')} \geqslant |C(k)|\|\mu(u_n)\|_{L^2(\Omega')} - \|k\|_{L^1(\omega)}\|\Pi_0(u_n)\|_{L^2(\Omega)} - \|f\|_{L^2(\Omega')} \to \infty$$

da der zweite Term auf der rechten Seite nach Annahme beschränkt ist. Daraus folgt die Koerzivität in beiden Fällen des Funktionals und nach Satz 3.2 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 $\lambda \in (0, 1)$, dass

$$\begin{aligned}\frac{1}{2}\|k \star (\lambda u_1 + (1-\lambda)u_2) - f\|^2_{L^2(\Omega')} &= \frac{1}{2}\|\lambda(k \star u_1 - f) + (1-\lambda)(k \star u_2 - f)\|^2_{L^2(\Omega')} \\ &< \frac{\lambda}{2}\|k \star u_1 - f\|^2_{L^2(\Omega')} + \frac{(1-\lambda)}{2}\|k \star u_2 - f\|^2_{L^2(\Omega')},\end{aligned}$$

d. h. der Diskrepanzterm und damit das gesamte Funktional ist strikt konvex, was die Eindeutigkeit des Minimierers garantiert. □

Aus der Linearität von A und der Fréchet-differenzierbarkeit von $v \mapsto \frac{1}{2}\|v - f\|^2_{L^2(\Omega')}$ folgt mit der Kettenregel 2.5 auch die Fréchet-differenzierbarkeit von $\mathcal{F}$ mit

$$\mathcal{F}(u) = A^*(Au - f) = \bar{k} \star (k \star u - f).$$

Damit ist $\mathcal{F}$ stetig auf $L^2(\Omega)$ und aus der Summenregel 3.10 folgt, 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 $A_h : \mathbb{R}^{NM} \to \mathbb{R}^{(N-2r)(M-2s)}$ definiert. Analog zur diskreten Divergenz rechnet man nun nach (durch Vertauschung der Summation und Indexverschiebung), dass $A_h^* : \mathbb{R}^{(N-2r)(M-2s)} \to \mathbb{R}^{NM}$ gegeben ist durch

$$(A_h^* w_h)_{ij} = h \sum_{n=i-r}^{i+r} \sum_{m=j-s}^{j+s} w_{nm} k_{n-i,m-j} = \bar{k}_h \star w_h,$$

d. h. durch die diskrete Faltung mit dem gespiegelten Kern $(\bar{k}_h)_{ij} = k_{-i,-j}$.

Das diskrete Problem ist also

$$\min_{u_h \in \mathbb{R}^{NM}} \frac{1}{2} \|A_h u_h - f_h\|_2^2 + \alpha \| |\nabla_h u_h|_2 \|_1.$$

Hier kommen nun in beiden Termen lineare Operatoren vor. Um zu vermeiden, für den ersten Term eine Proximalabbildung ausrechnen zu müssen (in der A_h^{-1} auftauchen würde, und A_h muss ja genau wie A nicht injektiv sein), wenden wir den Satz von Fenchel–Rockafellar an auf

$$\begin{aligned}
&F : \mathbb{R}^{NM} \to \mathbb{R}, && F(u_h) = 0,\\
&G : (\mathbb{R}^{(N-2r)(M-2s)} \times \mathbb{R}^{2NM}) \to \overline{\mathbb{R}}, && G(y_h, z_h) = \frac{1}{2}\|y_h - f_h\|_2^2 + \alpha \| |z_h|_2 \|_1,\\
&B_h : \mathbb{R}^{NM} \to (\mathbb{R}^{(N-2r)(M-2s)} \times \mathbb{R}^{2NM}), && B_h u_h = \begin{pmatrix} A_h u_h \\ \nabla_h u_h \end{pmatrix}.
\end{aligned}$$

Nun ist B_h ein beschränkter linearer Operator mit Norm $\|B_h\|^2 \leqslant \|k_h\|_1^2 + \frac{8}{h^2}$ und Adjungierter

$$B_h^* : (\mathbb{R}^{(N-2r)(M-2s)} \times \mathbb{R}^{2NM}) \to \mathbb{R}^{NM}, \qquad B_h^*(y_h, z_h) = A_h^* y_h - \operatorname{div}_h z_h.$$

Da G stetig ist in $B_h 0$, erhalten wir die Existenz von $(\bar{y}_h, \bar{z}_h) \in (\mathbb{R}^{(N-2r)(M-2s)} \times \mathbb{R}^{2NM})$ mit

$$\begin{cases} -B_h^*(\bar{y}_h, \bar{z}_h) \in \partial F(\bar{u}_h), \\ \qquad B_h \bar{u}_h \in \partial G^*(\bar{y}_h, \bar{z}_h). \end{cases}$$

Um darauf das primal-duale Extragradienten-Verfahren anwenden zu können, brauchen wir nur noch die entsprechenden Proximalpunktabbildungen. Für $F \equiv 0$ haben wir $\partial F(u_h) = \{0\}$ und damit

$$\operatorname{prox}_{\tau F}(u_h) = (\mathrm{Id} + \tau\partial F)^{-1}(u_h) = (\mathrm{Id})^{-1} u_h = u_h.$$

Weiter ist

$$\begin{aligned} G^* &: (\mathbb{R}^{(N-2r)(M-2s)} \times \mathbb{R}^{2NM}) \to \overline{\mathbb{R}}, \\ G^*(y_h^*, z_h^*) &= \frac{1}{2}\|y_h^*\|_2^2 + (y_h^*, f_h) + \delta_{\{v_h : \||v_h|_2\|_\infty \leqslant \alpha\}}(z_h^*), \end{aligned}$$

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 (5.10) für das Entfaltungsproblem (und analog für beliebige lineare Operatoren A_h) als konkreten Algorithmus formulieren:

Algorithmus 9.1 : Primal-duales L^2-TV-Entfalten

Input : $f \in \mathbb{R}^{NM}$, $k \in \mathbb{R}^{(2r+1)(2s+1)}$, $u^0 \in \mathbb{R}^{NM}$, $y^0 \in \mathbb{R}^{(N-2r)(M-2s)}$, $z^0 \in \mathbb{R}^{2NM}$, $\sigma\tau < (\|k_h\|_1^2 + h^2/8)$

for $k = 1 \dots$ **do**

- $u^k = u^{k-1} + \tau(\operatorname{div}_h z^{k-1} - A_h^* y^{k-1})$
- $\bar{u}^k = 2u^k - u^{k-1}$
- $y^k = (y^{k-1} + \sigma(A_h \bar{u}^k - f))/(1+\sigma)$
- $w = z^{k-1} + \sigma\nabla_h \bar{u}^k$
- $z^k = \alpha w/(\max\{\alpha, |w|_2\})$

10 DEKOMPRESSION

Der große Vorteil des Variationsansatzes in der Bildverarbeitung ist die Flexibilität in der Wahl der Bild- und Rekonstruktionsmodelle. Um dies zu illustrieren, betrachten wir zum Abschluss die Reduktion von Artefakten in der JPEG-Dekompression nach [Bredies und Holler 2012].

Bei JPEG handelt es sich um einen Standard für die verlustbehaftete Datenreduktion von digitalen Bilddaten. Die Kompression eines gegebenen Pixelbildes $u_h : \mathbb{R}^{NM} \to [0, 255]$ verläuft in groben Zügen wie folgt (wir ignorieren Vorverabeitungsschritte, die spezifisch für Farbbilder sind):

1. Das Bild wird in Blöcke von 8×8 Pixeln unterteilt.
2. Jeder Block wird durch eine *diskrete Kosinus-Transformation* in eine Linearkombination von 8×8 Kosinusfunktionen unterschiedlicher Frequenz zerlegt.
3. Jede 8×8-Matrix der Koeffizienten wird elementweise durch eine vorgegebene *Quantisierungsmatrix* dividiert (die von der gewünschten Datenreduktion abhängt: je größer die Einträge der Quantisierungsmatrix, desto höher die Kompression, wobei die Einträge für höhere – und damit für den visuellen Eindruck weniger relevante – Frequenzen größer sind als die für niedrige).
4. Die Ergebnisse der Division werden auf ganze Zahlen gerundet. (Hier tritt der Datenverlust auf.)
5. Die ganzzahligen Einträge aller Blöcke werden verlustfrei komprimiert und als Bitstrom gespeichert.

Für die Dekompression werden diese Schritte umgekehrt, was mit Ausnahme von Schritt 4 eindeutig möglich ist. Hier wird in der Standard-Implementierung Schritt 4 einfach übersprungen, d. h. die Umkehrung von Schritt 3 erhält die ganzzahligen Einträge als exakte Koeffizienten. Dies führt (bei starker Kompression) zu den bekannten Block-Artefakten. Alternativ kann man die Dekompression als Variationsproblem formulieren: Gesucht ist dasjenige Bild u mit (zum Beispiel) minimaler totaler Variation, dessen JPEG-Kompression den gegebenen Daten entspricht.

Um einen passenden Diskrepanzterm zu definieren, ist etwas Vorarbeit nötig. Wir beginnen mit der Block-Kosinus-Transformation, und nehmen dafür an, dass $\Omega = (0, 8k) \times (0, 8l) \subset \mathbb{R}^2$ für $k, l \in \mathbb{N}$ ist. Für $0 \leqslant i < k$ und $0 \leqslant j < l$ sei

$$E_{ij} = ([8i, 8i+8) \times [8j, 8j+8)) \cap \Omega$$

mit charakteristischer Funktion

$$\chi_{ij}(x) = \begin{cases} 1 & x \in E_{ij}, \\ 0 & \text{sonst.} \end{cases}$$

Sei weiterhin $\{b_{nm}\}_{nm \geqslant 0} \subset L^2(\Omega)$ definiert durch

$$b_{nm}(x_1, x_2) = c_n c_m \cos(n\pi x_1) \cos(m\pi x_2)$$

mit

$$c_s = \begin{cases} 1 & s = 0, \\ \sqrt{2} & s \neq 0. \end{cases}$$

Man rechnet nun nach, dass dies ein vollständiges Orthonormalsystem auf $L^2((0,1)^2)$ ergibt. Damit ist auch $\{a^{ij}_{nm}\}_{ij,nm}$, definiert durch

$$a^{ij}_{nm}(x_1, x_2) = \frac{1}{8} b_{nm}\left(\frac{x_1 - 8i}{8}, \frac{x_2 - 8j}{8}\right) \chi_{ij}(x_1, x_2),$$

ein vollständiges Orthonormalsystem auf $L^2(\Omega)$. Wir bezeichnen den zugehörigen *Transformations-Operator* mit

$$A : L^2(\Omega) \to \ell^2, \qquad [Au]^{ij}_{nm} = (u, a^{ij}_{nm})_{L^2(\Omega)},$$

der wegen der Parseval-Relation $\|Au\|_{\ell^2} = \|u\|_{L^2(\Omega)}$ erfüllt. Mit etwas Aufwand verifiziert man, dass A unitär ist, d. h. $A^{-1} = A^*$ gilt mit

$$A^* : \ell^2 \to L^2(\Omega), \qquad [A^*v](x) = \sum_{ij,m,n} v^{ij}_{nm} a^{ij}_{nm}(x).$$

Seien schließlich $\{q_{nm}\}_{m,n \geqslant 0}$ die gegebenen Quantisierungskoeffizienten und $\{z^{ij}_{nm}\}_{ij,nm}$ die zugehörigen quantisierten Daten des gesuchten Bildes $u^0 \in L^2(\Omega)$, d. h. es gilt

$$[Au^0]^{ij}_{nm} \in J^{ij}_{nm} := \left[q_{nm} z^{ij}_{nm} - \frac{q_{nm}}{2}, q_{nm} z^{ij}_{nm} + \frac{q_{nm}}{2}\right].$$

Wir setzen also

$$\mathcal{F}(u) := \delta_U(u), \qquad U := \left\{u \in L^2(\Omega) : [Au]^{ij}_{nm} \in J^{ij}_{nm}\right\}. \tag{10.1}$$

Wir nehmen an, dass $1 \leqslant q_{nm} < \infty$ gilt (da sonst keine Kompression stattfindet oder Koeffizienten komplett ignoriert werden) und damit alle J^{ij}_{nm} nichtleer und beschränkt sind. Also ist wegen der Linearität und Beschränktheit von A auch U nichtleer, konvex und abgeschlossen.

Wir betrachten nun

$$\min_{u \in L^2(\Omega)} \mathcal{F}(u) + TV(u) = \min_{u \in U} TV(u). \tag{10.2}$$

Für die Existenz einer Lösung ist also nur noch zu zeigen, dass dieses Funktional eigentlich ist, d. h. ein $u \in U \cap BV(\Omega)$ existiert. Für die spätere Verwendung zeigen wir gleich, dass dieses u im Inneren U^o von U liegt. Da nach Annahme alle J^{ij}_{nm} nichtleeres Inneres haben. hat U ebenfalls ein nichtleeres Inneres. Sei nun $\tilde{u} \in U^o$. Aus der Dichtheit von $C_0^\infty(\Omega) \subset BV(\Omega)$ in $L^2(\Omega)$ (Satz 6.4) erhalten wir die Existenz einer Folge $\{u_n\}_{n\in\mathbb{N}} \subset BV(\Omega)$ mit

$$\|Au_n - A\tilde{u}\|_{\ell^2} = \|u_n - \tilde{u}\|_{L^2(\Omega)} \to 0 \qquad \text{für } n \to \infty.$$

Deshalb existiert eine Teilfolge $\{Au_{n_k}\}_{k\in\mathbb{N}}$, die punktweise gegen Au_0 konvergiert. Es gibt also einen Index n_{k^*} groß genug, so dass $u_0 := u_{n_{k^*}} \in U^o \cap BV(\Omega)$ liegt. Aus Satz 3.2 folgt nun die Existenz einer Lösung $\bar{u} \in U \cap BV(\Omega)$ von (10.2).

Die Existenz von u_0 erlaubt auch die Anwendung der Summenregel sowie der Kettenregel; $\bar{u}$ ist also Lösung genau dann, wenn gilt

$$0 \in \partial TV(\bar{u}) + A^*\partial\delta_J(A\bar{u})$$

mit

$$J := \left\{ v \in \ell^2 : v^{ij}_{nm} \in J^{ij}_{nm} \right\}. \tag{10.3}$$

In anderen Worten: Es existiert ein $\bar{p} \in \ell^2$ mit

$$\begin{cases} -A^*\bar{p} \in \partial TV(\bar{u}), \\ \quad\ \bar{p} \in \partial\delta_J(A\bar{u}) \end{cases}$$

Im Diskreten ist nun $u_h \in \mathbb{R}^{8k\cdot 8l}$ gesucht, und statt der Block-Kosinus-Transformation $A : L^2(\Omega) \to \ell^2$ verwenden wir die diskrete Block-Kosinus-Transformation $A_h : \mathbb{R}^{8k\cdot 8l} \to \mathbb{R}^{8k\cdot 8l}$, d. h. $0 \leqslant m, n \leqslant 7$. Die zulässige Mengen U_h bzw. J_h sind dann analog zu (10.1) bzw. (10.3) definiert. Der Satz von Fenchel–Rockafellar zusammen mit der Summenregel liefert dann die Extremalitätsbedingungen

$$\begin{cases} \operatorname{div}_h \bar{y}_h \in \partial\delta_U(u_h) = A_h^*\partial\delta_{J_h}(A_h\bar{u}_h), \\ \ \nabla_h\bar{u}_h \in \partial\delta_{\{v_h:\||v_h|_2\|_\infty \leqslant 1\}}(\bar{y}_h). \end{cases}$$

Um hierfür das primal-dualen-Extragradienten-Verfahren anwenden zu können, benötigen wir noch die Proximalpunkt-Abbildung der ersten Relation. Da A_h ein unitärer linearer Operator ist (wobei A_h^* durch die inverse Block-Kosinus-Transformation gegeben ist), gilt

$$\begin{aligned} v_h \in A_h^* \partial \delta_{J_h}(A_h u_h) &\Leftrightarrow A_h v_h \in \partial \delta_{J_h}(A_h u_h) \\ &\Leftrightarrow A_h u_h = \operatorname{prox}_{\tau \delta_{J_h}}(A_h u_h + \tau A_h v_h) \\ &\Leftrightarrow u_h = A_h^* \operatorname{prox}_{\tau \delta_{J_h}}(A_h(u_h + \tau v_h)). \end{aligned}$$

Schliesslich haben wir analog zu Beispiel 4.12 (iii)

$$\begin{aligned} [\operatorname{prox}_{\tau\delta_{J_h}}(v_h)]_{nm}^{ij} &= \begin{cases} [v_h]_{nm}^{ij} & [v_h]_{nm}^{ij} \in J_{nm}^{ij} \\ q_{nm}[z_h]_{nm}^{ij} + \frac{q_{nm}}{2} & [v_h]_{nm}^{ij} > q_{nm}[z_h]_{nm}^{ij} + \frac{q_{nm}}{2} \\ q_{nm}[z_h]_{nm}^{ij} - \frac{q_{nm}}{2} & [v_h]_{nm}^{ij} < q_{nm}[z_h]_{nm}^{ij} - \frac{q_{nm}}{2} \end{cases} \\ &=: [\operatorname{proj}_{J_h}(v_h)]_{nm}^{ij}. \end{aligned}$$

Der gesamte Algorithmus ist also wie unten angegeben; als Startvektor $u^0 \in U_h$ kann man zum Beispiel die Standard-JPEG-Dekompression der Daten verwenden.

Algorithmus 10.1 : Primal-duales JPEG-TV-Entrauschen

Input : $q \in \mathbb{R}^{8\cdot 8}$, $z \in \mathbb{R}^{8k\cdot 8l}$, $u^0 \in U_h$, $y^0 \in \mathbb{R}^{8k\cdot 8l}$, $\sigma\tau < 1/8$

for $k = 1 \dots$ **do**

- $u^k = A_h^* \operatorname{proj}_{J_h}(A_h(u^{k-1} + \tau \operatorname{div}_h y^{k-1}))$
- $\bar{u}^k = 2u^k - u^{k-1}$
- $w = y^k + \sigma \nabla_h \bar{u}^k$
- $y^k = w/(\max\{1, |w|_2\})$

LITERATUR

R. A. Adams und J. J. F. Fournier (2003). *Sobolev Spaces*. 2. Aufl. Academic Press, Amsterdam.

L. Ambrosio, N. Fusco und 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 und G. Michaille (2006). *Variational Analysis in Sobolev and BV Spaces*. Bd. 6. MPS/SIAM Series on Optimization. Society for Industrial und Applied Mathematics (SIAM), Philadelphia, PA. DOI: 10.1137/1.9780898718782.

H. H. Bauschke und P. L. Combettes (2011). *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC. Springer, New York. DOI: 10.1007/978-1-4419-9467-7.

K. Bredies und M. Holler (2012). *A total variation-based JPEG decompression model*. SIAM Journal on Imaging Sciences 5.1, S. 366–393. DOI: 10.1137/110833531.

K. Bredies und 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.

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 und T. Pock (2011). *A first-order primal-dual algorithm for convex problems with applications to imaging*. J Math Imaging Vis 40.1, S. 120–145. DOI: 10.1007/s10851-010-0251-1.

T. Chan und S. Esedoglu (2005). *Aspects of total variation regularized L1 function approximation*. SIAM Journal on Applied Mathematics 65.5, S. 1817–1837. DOI: 10.1137/040604297.

M. Dobrowolski (2010). *Angewandte Funktionalanalysis*. Springer-Verlag, Berlin. DOI: 10.1007/978-3-642-15269-6.

J. Eckstein und D. P. Bertsekas (1992). *On the Douglas–Rachford splitting method and the proximal point algorithm for maximal monotone operators*. Mathematical Programming 55.1-3, S. 293–318. DOI: 10.1007/BF01581204.

I. Ekeland und R. Témam (1999). *Convex Analysis and Variational Problems*. Bd. 28. Classics Appl. Math. SIAM, Philadelphia. DOI: 10.1137/1.9781611971088.

B. He und X. Yuan (2012). *Convergence analysis of primal-dual algorithms for a saddle-point problem: from contraction perspective*. SIAM Journal on Imaging Sciences 5.1, S. 119–149. DOI: 10.1137/100814494.

M. Kurula und 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, S. 372–376.

N. Parikh und S. Boyd (2014). *Proximal algorithms*. Foundations and Trends in Optimization 1.3, S. 123–231. DOI: 10.1561/2400000003.

L. I. Rudin, S. Osher und E. Fatemi (1992). *Nonlinear total variation based noise removal algorithms*. Physica D: Nonlinear Phenomena 60.1, S. 259–268. DOI: 10.1016/0167-2789(92)90242-F.

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

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.