

# Institut für Numerische und Angewandte Mathematik

Masterarbeit

## Konvexe Krümmungsregularisierung für Linienmaße in drei Dimensionen

Bernd Mekes

Matr. 373772

Betreuer: Prof. Dr. Benedikt Wirth  
Betreuender Assistent: Ulrich Hartleif

# Zusammenfassung

## Zusammenfassung

In dieser Arbeit übertragen wir eine Methode aus [2] zur konvexen Krümmungsregularisierung von Niveaulinien in zweidimensionalen Bildern auf einen dreidimensionalen Fall, in dem das Bild dünne, linienartige Strukturen enthält. Indem wir die Ecken dieser Strukturen in geeigneter Weise bestrafen erhalten wir Informationen über deren Krümmung. Mit Hilfe einer konvexen Relaxierung eines Funktionals zur Minimierung der Krümmung erhalten wir dann ein konvexes Optimierungsproblem. Zur Lösung eines Entrauschungsproblems, das auf dieser Krümmungsminimierung basiert, geben wir zum Abschluss eine geeignete Diskretisierung an und lösen das dabei entstehende Sattelpunktproblem numerisch.

## Abstract

In this thesis we aim at transferring an approach from [2] for convex curvature regularization of level lines in two dimensional images to a three dimensional case with images that contain thin, line-like structures. By penalizing the vertices of these structures in an appropriate way, we gain information about their curvature. We introduce a functional which minimizes the curvature of these structures. By using a convex relaxation, we derive a convex optimization problem for curvature minimization. At the end of this thesis we will make use of this curvature minimization in order to solve a denoising problem. We will derive a suitable discretization in form of a saddle point optimization problem which will then be solved numerically.

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
<b>2</b>	<b>Einführendes Kapitel</b>	<b>3</b>
2.1	Konvexe Optimierung . . . . .	3
2.2	Weitere Grundlagen . . . . .	8
<b>3</b>	<b>Ein Funktional zum Bestrafen von Knoten in Linienmaßen</b>	<b>12</b>
3.1	Dreidimensionale Linienmaße . . . . .	12
3.2	Eine Norm auf $\mathcal{M}(S^2)$ , die Richtungsänderungen bestraft . . . . .	15
3.3	Eine Funktional, das Richtungsänderungen von Linienmaßen bestraft . . . . .	18
<b>4</b>	<b>Konvexe Krümmungsregularisierung</b>	<b>24</b>
4.1	Konvexität des Minimierungsproblems . . . . .	24
4.2	Existenz eines Minimums . . . . .	28
4.2.1	Existenz eines Minimums für $L^2$ Daten . . . . .	30
4.2.2	Existenz eines Minimums für einen Wasserstein Datenterm . . . . .	32
<b>5</b>	<b>Numerische Umsetzung</b>	<b>35</b>
5.1	Diskretisierung des Optimierungsproblems . . . . .	35
5.2	Primal-dualer Algorithmus für das diskrete Optimierungsproblem . . . . .	39
5.2.1	Primal-Dualer Algorithmus für $\text{TVX}_0^{\alpha,\beta}$ . . . . .	40
5.2.2	Primal-Dualer Algorithmus für $\text{TVX}_1^{\alpha,\beta}$ . . . . .	42
5.3	Resultate und Auswertung . . . . .	43
<b>6</b>	<b>Fazit</b>	<b>50</b>
	<b>Literaturverzeichnis</b>	<b>51</b>

---



---

## KAPITEL 1

---

# Einleitung

Ein typisches Problem der Bildverarbeitung ist das Entrauschen von Bildern, bei dem aus einem gegebenen verrauschten Bild  $u_0$  ein Bild  $u^*$  konstruiert werden soll, welches dieses Rauschen nicht aufweist. Dies geschieht häufig mittels Minimierungsproblemen der Form

$$u^* = \min_v G_{u_0}(v) + R(v)$$

bei dem einerseits ein von  $u_0$  abhängiger Datenterm  $G_{u_0}$  sicherstellt, dass das Ergebnis möglichst stark den verrauschten Daten  $u_0$  ähnelt, während ein Strafterm  $R$  typische Eigenschaften des Rauschens bestraft.

Häufig wird beispielsweise die Länge von Kanten in Bildern bestraft. Da in Bildern auf benachbarte Pixel mit gleichen Farbwerten unterschiedlich starkes Rauschen wirken kann und daher durch Rauschen künstliche Kanten entstehen, ist diese Methode zur Entrauschung erfolgreich. Motiviert durch psychologische Erkenntnisse über die Rekonstruktion von Kanten teilweise verdeckter Objekte im Gehirn (siehe [7]), schlägt Mumford in [12] vor auch die Krümmung dieser Kanten zu berücksichtigen. Mumford tut dies, indem er die Eulersche Elastica in die Bildverarbeitung einführt. Euler untersuchte 1744 die Verformung dünner Drähte. Dabei wurden die Endpunkte eines Drahtes fixiert und die Tangentialrichtungen in den Endpunkten vorgegeben. Euler beschrieb die daraus resultierenden Kurven (sinngemäß) als Minimum des Energiefunktional

$$E[\kappa(s)] = \int_0^l \kappa(s)^2 ds$$

über die Bogenlänge  $0 \leq s \leq l$ , wobei  $\kappa(s)$  die Krümmung der Kurve in  $s$  sei [9]. In [12] wird die Elastica allgemeiner definiert als diejenige Kurve  $\gamma$ , die das Funktional

$$\int_{\gamma} \alpha + \beta \kappa^p d\mathcal{H}^1 \tag{1.1}$$

unter Nebenbedingungen in den Endpunkten minimiert. Hierbei sind  $\alpha$  und  $\beta$  nichtnegative Parameter,  $p \geq 1$  ein Exponent und  $\kappa$  die Krümmung der Kurve  $\gamma$ .

In [2] führen Bredies, Pock und Wirth Variationsmodelle ein, mit denen sie die Krümmung von Niveaulinien in Bildern minimieren können. Dazu definieren sie zunächst Metriken, mit denen Sie Richtungsänderungen in Ecken von Niveaulinien messen und bestrafen können. Mit Hilfe eines Funktional, das diese Strafen für alle Ecken aufsummiert erhalten sie Informationen über die gesamte Krümmung in einem Bild. Diese nutzen sie dann um ein durch die Eulersche Elastica motiviertes Funktional zur Minimierung der Krümmung in Bildern zu definieren. Mit Hilfe einer konvexen Relaxierung erhalten sie dann ein konvexes Sattelpunktproblem, welches diskretisiert und numerisch gelöst wird.

---

In dieser Arbeit möchten wir diese Methode auf einen dreidimensionalen Fall übertragen, in dem die Bilder aus dünnen, linienartigen Strukturen bestehen. Wir werden dazu in Kapitel 2 einige Grundlagen einführen, die im Verlauf der Arbeit benötigt werden. Anschließend werden wir zu Beginn von Kapitel 3 den Begriff des Linienmaßes definieren. Im Rest des Kapitels 3 werden wir die Richtungsänderungen bestrafenden Metriken und das Funktional welches die Krümmung von Linienmaßen mit Hilfe dieser Metriken ermittelt einführen. In Kapitel 4 werden wir ein Optimierungsproblem zur Minimierung der Krümmung in Linienmaßen erhalten, welches wir anschließend konvex relaxieren. Dabei entsteht ein Sattelpunktproblem, welches in Kapitel 5 zunächst diskretisiert und anschließend numerisch gelöst wird.

---



---

## KAPITEL 2

---

# Einführendes Kapitel

Zu Beginn dieser Arbeit führen wir einige mathematische Grundlagen ein, die wir im Laufe der Arbeit benötigen werden.

Zunächst werden wir im Abschnitt 2.1 auf einige Grundlagen der Optimierung, insbesondere der konvexen Optimierung und im Speziellen der primal-dualen Optimierungsprobleme eingehen. Wir werden zudem einen Algorithmus zur Lösung solcher Probleme vorstellen.

Im zweiten Abschnitt 2.2 dieses Kapitels werden wir einige weitere, insbesondere maßtheoretische Begriffe definieren.

### 2.1 Konvexe Optimierung

Wir führen nun also zunächst die benötigten Begriffe aus der Optimierung ein. Grundlegend ist dabei der Begriff des Optimierungsproblems:

**Definition 1 (Optimierungsproblem)** *Ein Optimierungsproblem ist ein Problem bei dem ein Kostenfunktional  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  unter gegebenen Nebenbedingungen minimiert werden soll. Wir schreiben:*

$$\min_{x \in \mathbb{R}^n} f(x), \text{ so dass } c_E(x) = b_E, c_I(x) \leq b_I.$$

Hierbei sind

- $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\pm\infty\}$  das Kostenfunktional,
- $x \in \mathbb{R}^n$  die Optimierungsvariable,
- $c_E \equiv (g_1, \dots, g_p)^T : \mathbb{R}^n \rightarrow (\mathbb{R} \cup \{\infty\})^p$  die Funktionen für die Gleichungsnebenbedingungen,
- $c_I \equiv (h_1, \dots, h_q)^T : \mathbb{R}^n \rightarrow (\mathbb{R} \cup \{\infty\})^q$  die Funktionen für die Ungleichungsnebenbedingungen,
- $b_E$  ein Vektor für die rechte Seite der Gleichungsnebenbedingung und
- $b_I$  ein Vektor für die rechte Seite der Ungleichungsnebenbedingung.

*Ein Minimierungsproblem heißt konvex, falls  $f$  konvex,  $c_E$  affin und  $c_I$  komponentenweise konvex ist. Falls wir die Definitionsbereiche des Kostenfunktionals  $f$  und der Nebenbedingungen  $c_E$  und  $c_I$  auf Teilmengen des  $\mathbb{R}^n$  einschränken, so fordern wir zudem, dass die Definitionsbereiche der Funktionen konvexe Mengen sind. Der Definitionsbereich des Optimierungsproblems  $\mathcal{D} = (\text{dom}f) \cap (\bigcap_{i=1}^p \text{dom}g_i) \cap (\bigcap_{i=1}^q \text{dom}h_i)$  ist in dem Fall ebenfalls konvex.*

Nachdem wir zunächst allgemeine konvexe Optimierungsprobleme eingeführt haben, möchten wir nun ein spezielles konvexes Optimierungsproblem, welches wir im Verlauf der Arbeit noch benötigen werden, einführen. Dazu benötigen wir die folgende Definition.

**Definition 2 (Fenchel-Konjugierte, [14], Sec. 12)** Gegeben sei ein reeller, normierter Vektorraum  $X$ , sein Dualraum  $X^*$  sowie die duale Paarung

$$\langle \cdot, \cdot \rangle : X \times X^* \rightarrow \mathbb{R}, (x, y) \mapsto \langle x, y \rangle.$$

Für ein gegebenes Funktional  $f : X \rightarrow \mathbb{R} \cup \{\infty\}$  definieren wir die Fenchel-Konjugierte  $f^* : X^* \rightarrow \mathbb{R} \cup \{\infty\}$  durch

$$f^*(y) = \sup_{x \in X} \{\langle x, y \rangle - f(x)\}. \quad (2.1)$$

Zur Herleitung unseres Optimierungsproblems möchten wir die Bikonjugierte  $f^{**}$  verwenden. Da allerdings im Allgemeinen nur die Ungleichung  $f^{**} \leq f$  gilt, brauchen wir eine Aussage darüber, wann wir Gleichheit erwarten können. Dazu müssen wir zunächst sicherstellen, dass das Minimum einer konvexen Funktion endlich ist. Wir definieren daher:

**Definition 3 (Zulässige konvexe Funktion, [14], Sec. 4)** Eine konvexe Funktion  $f$  heißt zulässig, falls

- $f(x) < \infty$  für mindestens ein  $x$  und
- $-\infty < f(x)$  für alle  $x$

gilt.

Eine Aussage über die Gleichheit einer Funktion und ihrer Bikonjugierten liefert nun folgendes Theorem:

**Theorem 1 ([14], Thm. 12.2)** Es sei  $X$  ein reeller, normierter Vektorraum. Für jede unterhalbstetige, zulässig konvexe Funktion  $f : X \rightarrow \mathbb{R} \cup \{\infty\}$  gilt  $f = f^{**}$ .

Mit den Definitionen und dem Theorem sind wir nun in der Lage das primal-duale Optimierungsproblem einzuführen:

**Definition 4 (Primal-duales Optimierungsproblem, [3], Sec. 2)** Gegeben seien

- zwei endlichdimensionale, reelle Vektorräume  $X$  und  $Y$  mit Skalarprodukten  $\langle \cdot, \cdot \rangle_X$  bzw.  $\langle \cdot, \cdot \rangle_Y$  und Normen  $\|\cdot\|_X$  bzw.  $\|\cdot\|_Y$ ,
- ein stetiger und linearer Operator  $\mathcal{K} : X \rightarrow Y$  mit induzierter Norm  $\|\mathcal{K}\|$  und
- zwei unterhalbstetige, zulässige konvexe Funktionale  $g : X \rightarrow [0, \infty]$  und  $f : Y \rightarrow [0, \infty]$ .

Ein primales Optimierungsproblem ist ein Minimierungsproblem der Form

$$\min_{x \in X} f(\mathcal{K}x) + g(x). \quad (2.2)$$

Mit Hilfe des Theorems 1 schreiben wir

$$f(\mathcal{K}x) = \max_{y \in Y} \{\langle \mathcal{K}x, y \rangle - f^*(y)\}.$$

Setzen wir diese Gleichung in das primale Optimierungsproblem ein, so erhalten wir das primal-duale Optimierungsproblem

$$\min_{x \in X} \max_{y \in Y} \{\langle \mathcal{K}x, y \rangle + g(x) - f^*(y)\}. \quad (2.3)$$

Wir nennen  $x$  die primale Variable und  $y$  die duale Variable des Optimierungsproblems.

Wir werden nun im verbleibenden Teil dieses Abschnitts analog zu [2] einen Algorithmus zum Lösen primal-dualer Optimierungsprobleme einführen. Der zugrunde liegende Algorithmus wurde von A. Chambolle und T. Pock in [3] entwickelt. Um diesen Algorithmus einzuführen benötigen wir noch einige Definitionen. Zunächst müssen wir mittels Subdifferentialen den Begriff des Gradienten auf nicht differenzierbare konvexe Funktionen verallgemeinern.

**Definition 5 (Subdifferential, [14], Sec. 23)** Gegeben sei eine konvexe Funktion  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ . Ein Vektor  $x^* \in \mathbb{R}^n$  heißt Subgradient von  $f$ , falls

$$f(z) \geq f(x) + \langle x^*, z - x \rangle$$

für alle  $z \in \mathbb{R}^n$  gilt. Das Subdifferential  $\partial f(x)$  ist die Menge aller Subgradienten der Funktion  $f$  im Punkt  $x$ .

Stark mit dem Subdifferential verknüpft ist der Proximaloperator.

**Definition 6 (Proximaloperator, [14], Sec. 31)** Der Proximaloperator einer unterhalbstetigen, zulässig konvexen Funktion  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$  zu einem gegebenen Parameter  $\lambda > 0$  ist definiert durch

$$\text{prox}_{\lambda, f}(y) = \arg \min_x \left( f(x) + \frac{1}{2\lambda} \|x - y\|_2^2 \right). \quad (2.4)$$

Ein Wert  $z = \text{prox}_{\lambda, f}(y)$  heißt Proximum.

Wie bereits erwähnt hängen der Proximaloperator und das Subdifferential eng zusammen, denn für  $z = \text{prox}_{\lambda, f}(y)$  lautet die Optimalitätsbedingung erster Ordnung:

$$\begin{aligned} 0 &\in \partial_z \left( f(z) + \frac{1}{2\lambda} \|z - y\|_2^2 \right) \\ \iff 0 &\in \partial_z f(z) + (z - y) \\ \iff y - z &\in \lambda \partial_z f(z). \end{aligned}$$

Aufgrund der strikten Konvexität von  $\|\cdot\|_2$  ist das Proximum eindeutig. Man formuliert die letzte Zeile daher häufig noch um zu  $z = (I + \lambda \partial f)^{-1}(y)$ . Der Operator  $(I + \lambda \partial f)^{-1}$  wird dann Resolventen-Operator genannt.

Wir werden diesen Operator zur Formulierung des Algorithmus von Chambolle und Pock benötigen. Daher müssen wir sicherstellen, dass Proxima einfach zu berechnen sind. Im Allgemeinen kann man allerdings nicht erwarten, dass Proximaloperatoren eine Darstellung in geschlossener Form besitzen. Wir definieren daher:

**Definition 7 (Einfache Funktion)** Eine Funktion  $f$  heißt einfach, falls ihr Proximaloperator eine Darstellung in geschlossener Form besitzt.

Falls eine Funktion  $f$  einfach ist, so ist auch ihre Fenchel-Konjugierte  $f^*$  einfach, denn es gilt:

**Theorem 2 (Moreau's Identität, [14], Thm. 31.5.)** Für die Proximaloperatoren einer Funktion  $f$  und ihrer Fenchel-Konjugierten  $f^*$  gilt

$$x = \text{prox}_{\lambda, f}(x) + \text{prox}_{\frac{1}{\lambda}, f^*}\left(\frac{x}{\lambda}\right).$$

Wir sind mit diesen Definitionen nun in der Lage den Algorithmus von Chambolle und Pock zur Lösung primal-dualer Optimierungsprobleme anzugeben. Dazu seien die Räume  $X, Y$ , der Operator  $\mathcal{K}$  und die Funktionen  $f$  und  $g$  wie in Definition 4 gegeben. Wir setzen zudem im Folgenden voraus, dass die auftretenden Funktionen  $f$  und  $g$  einfach im Sinne der Definition 7 sind.

<p><b>Algorithmus 1 :</b> Algorithmus von Chambolle und Pock aus [3] zur Lösung primal-dualer Probleme der Form (2.3)</p>
<p>Initialisierung: Wähle <math>\sigma, \tau &gt; 0, \theta \in [0, 1]</math>, einen Startwert <math>(x^0, y^0) \in X \times Y</math> und setze <math>\hat{x}^0 = x^0</math>.</p> <p>Iterationen (<math>n \geq 0</math>):</p> $\begin{cases} x^{n+1} = \text{prox}_{\tau, g}(x^n - \tau \mathcal{K}^T y^n) \\ \hat{x}^{n+1} = x^{n+1} + \theta(x^{n+1} - x^n) \\ y^{n+1} = \text{prox}_{\sigma, f^*}(y^n + \sigma \mathcal{K} \hat{x}^{n+1}) \end{cases} \quad (2.5)$

Zur Veranschaulichung schreiben wir die Gleichung für  $x$  um:

$$\begin{aligned} x^{n+1} &= \text{prox}_{\tau, g}(x^n - \tau \mathcal{K}^T y^n) \\ &= (I + \tau \partial g)^{-1}(x^n - \tau \mathcal{K}^T y^n) \\ &= x^n - \tau \mathcal{K}^T y^n - \tau \partial g(x^{n+1}) \end{aligned}$$

Falls wir nun voraussetzen, dass  $g$  differenzierbar ist, so steht in der letzten Zeile ein Gradientenabstiegsverfahren in der primalen Variable  $x$ , bei dem im linearen Anteil ein expliziter und im Rest ein impliziter Gradientenschritt gemacht wird. Analog erhält man für die dritte Gleichung in (2.5) ein entsprechendes Gradientenaufstiegsverfahren für die duale Variable  $y$ . Die zweite Gleichung in (2.5) ist ein Relaxationsschritt in der primalen Variable.

Wir schränken uns im Folgenden auf den Fall  $\theta = 1$  ein und zitieren zunächst ein Konvergenzresultat aus [3]:

**Theorem 3 (Konvergenz des Algorithmus 1, [3], Thm. 1)** Wir nehmen an, dass das primal-duale Problem (2.3) eine Lösung  $(\tilde{x}, \tilde{y})$  besitze. Weiter wählen wir  $\theta = 1$  und die Schrittweiten  $\sigma$  und  $\tau$  so, dass  $\sigma\tau\|\mathcal{K}\|^2 < 1$  gilt. Es seien  $(x^n, y^n, \hat{x}^n)$  wie im Algorithmus 1 gegeben. Dann gilt:

- Für jedes  $n$  ist  $(x^n, y^n)$  beschränkt durch

$$\frac{\|y^n - \tilde{y}\|^2}{2\sigma} + \frac{\|x^n - \tilde{x}\|^2}{2\tau} \leq (1 - \sigma\tau\|\mathcal{K}\|^2) \left( \frac{\|y^0 - \tilde{y}\|^2}{2\sigma} + \frac{\|x^0 - \tilde{x}\|^2}{2\tau} \right).$$

- Es seien  $x_N = \frac{1}{N} \left( \sum_{n=1}^N x^n \right)$  und  $y_N = \frac{1}{N} \left( \sum_{n=1}^N y^n \right)$  gegeben, dann sind die Häufungspunkte von  $(x_N, y_N)_N$  Lösungen des primal-dualen Problems (2.3) und für beschränkte Teilmengen  $B_x \subset X$  und  $B_y \subset Y$  gilt die Abschätzung

$$\max_{y \in B_y} \langle y, Kx_N \rangle - f^*(y) + g(x_N) - \min_{x \in B_x} \langle y_N, Kx \rangle - f^*(y_N) + g(x) \leq \frac{D(B_x, B_y)}{N}$$

mit

$$D(B_x, B_y) := \sup_{(x,y) \in B_x \times B_y} \frac{\|x - x_0\|^2}{2\tau} + \frac{\|y - y_0\|^2}{2\sigma}.$$

- Es existiert eine Lösung  $(x^*, y^*)$  des primal-dualen Problems, so dass  $x^n \rightarrow x^*$  und  $y^n \rightarrow y^*$  gelten.

Falls  $g$  oder  $f^*$  strikt konvex sind, so lassen sich bessere Konvergenzresultate für die primale Variable  $x$  ([3], Thm. 2) und eine primal-duale Lücke ([4], Sec. 5) zeigen. In diesen Fällen lässt sich eine Konvergenzrate  $\mathcal{O}(1/N^2)$  zeigen, wobei  $N$  die Anzahl der Iterationen sei. Allerdings ist für diese Resultate eine spezielle Schrittweitenwahl nötig. Falls der Operator  $K$  keine einfache Struktur hat, kann sich jedoch insbesondere die Schrittweitenwahl als schwierig gestalten, da die Norm  $\|\mathcal{K}\|$  bestimmt werden muss und eventuell groß ist, was zu kleinen Schrittweiten führt. In [13] nehmen Chambolle und Pock an, dass  $\mathcal{K} \in \mathbb{R}^{m \times n}$  gilt, wobei  $n = \dim(X)$  und  $m = \dim(Y)$  seien. Sie schlagen daher vor die Schrittweiten  $\sigma$  und  $\tau$  durch eine Vorkonditionierung mit symmetrisch positiv definiten Matrizen zu ersetzen.

**Algorithmus 2** : Algorithmus von Chambolle und Pock mit Vorkonditionierung

Initialisierung: Wähle symmetrische und positiv definite Matrizen  $\mathcal{T}, \mathcal{S}$  und einen Startwert  $(x^0, y^0) \in X \times Y$ .

Iterationen ( $n \geq 0$ ):

$$\begin{cases} x^{n+1} = \text{prox}_{\mathcal{T},g} \left( x^n - \mathcal{T}\mathcal{K}^T y^n \right) \\ y^{n+1} = \text{prox}_{\mathcal{S},f^*} \left( y^n + \mathcal{S}\mathcal{K} (2x^{n+1} - x^n) \right) \end{cases} \quad (2.6)$$

Die Proximaloperatoren sind hierbei definiert durch

$$\text{prox}_{\mathcal{T},g}(x^*) = \arg \min_x \frac{1}{2} \langle \mathcal{T}^{-1}(x - x^*), (x - x^*) \rangle_2 + g(x) \quad (2.7)$$

und

$$\text{prox}_{\mathcal{S},f^*}(y^*) = \arg \min_y \frac{1}{2} \langle \mathcal{S}^{-1}(y - y^*), (y - y^*) \rangle_2 + f^*(y). \quad (2.8)$$

Für diesen Algorithmus erhält man für spezielle Wahlen von  $\mathcal{S}$  und  $\mathcal{T}$  wieder ein Konvergenzresultat.

**Theorem 4 (Konvergenz des Algorithmus 2, [13], Thm. 1 & Lemma 2)** Für  $\mathcal{K} \in \mathbb{R}^{m \times n}$  definieren wir Diagonalmatrizen  $\mathcal{S} \in \mathbb{R}^{m \times m}$  und  $\mathcal{T} \in \mathbb{R}^{n \times n}$  mit Diagonaleinträgen

$$\mathcal{S}_{i,i} = \frac{1}{\sum_{j=1}^n |\mathcal{K}_{i,j}|} \quad \text{beziehungsweise} \quad \mathcal{T}_{j,j} = \frac{1}{\sum_{i=1}^m |\mathcal{K}_{i,j}|}.$$

Dann konvergiert die durch (2.6) generierte Folge  $(x_n, y_n)$  gegen eine optimale Lösung des primal-dualen Problems (2.3).

Um den finalen Algorithmus aus [2] zu erhalten, führen wir einen zusätzlichen Überrelaxierungsschritt mit Parameter  $\gamma$  ein:

**Algorithmus 3** : Algorithmus von Chambolle und Pock mit Vorkonditionierung und Überrelaxierung, [2]

Initialisierung: Wähle  $\mathcal{T}, \mathcal{S}$  wie in Theorem 4,  $\gamma \in [0, 1[$  und einen Startwert  $(x^0, y^0) \in X \times Y$ .

Iterationen ( $n \geq 0$ ):

$$\begin{cases} \hat{x}^{n+1} = \text{prox}_{\mathcal{T}, g}(x^n - \mathcal{T}\mathcal{K}^T y^n) \\ \hat{y}^{n+1} = \text{prox}_{\mathcal{S}, f^*}(y^n + \mathcal{S}\mathcal{K}(2\hat{x}^{n+1} - x^n)) \\ (x^{n+1}, y^{n+1}) = (\hat{x}^{n+1}, \hat{y}^{n+1}) + \gamma(\hat{x}^{n+1} - x^n, \hat{y}^{n+1} - y^n) \end{cases}$$

In einem sehr aktuellen Paper zeigen Chambolle und Pock, dass für  $\mathcal{T} = \tau I_n$  und  $\mathcal{S} = \sigma I_m$  und spezielle Schrittweiten  $\tau$  und  $\sigma$  die Konvergenz des Algorithmus 3 für  $\gamma \rightarrow 1$  beschleunigt wird ([4], Thm. 2).

## 2.2 Weitere Grundlagen

Wir führen nun noch einige weitere, insbesondere maßtheoretische Begriffe ein. Zunächst definieren wir den Begriff des Radonmaßes, da dieser in der Literatur nicht einheitlich verwendet wird. Maßtheoretische Grundlagen setzen wir hierbei voraus.

**Definition 8 (Radonmaß, [2])** Gegeben sei ein topologischer Raum  $\Omega$  und die Borelsche  $\sigma$ -Algebra  $\mathcal{B}(\Omega)$  auf  $\Omega$ . Ein (endliches,  $\mathbb{R}$ -wertiges) Radonmaß  $\mu$  ist eine reguläre, abzählbar additive Mengenfunktion  $\mu : \mathcal{B}(\Omega) \rightarrow \mathbb{R}$  mit  $\mu(\emptyset) = 0$ . Für ein gegebenes Radonmaß  $\mu$  definieren wir das Variationsmaß  $|\mu|$  durch

$$|\mu|(E) = \sup \left\{ \sum |\mu(E_n)| \mid E_n \in \mathcal{B}(\Omega) \text{ paarweise disjunkt, } E = \cup_{n=0}^{\infty} E_n \right\}.$$

Die Variationsnorm  $\|\mu\|_{\mathcal{M}}$  eines Radonmaßes  $\mu$  ist dann definiert als  $\|\mu\|_{\mathcal{M}} = |\mu|(\Omega)$ . Es sei  $\mathcal{M}(\Omega) = \{\mu : \mathcal{B}(\Omega) \rightarrow \mathbb{R} \mid \mu \text{ endliches Radonmaß}\}$  der Raum der endlichen Radonmaße ausgestattet mit  $\|\cdot\|_{\mathcal{M}}$ . Dann ist  $\mathcal{M}(\Omega)$  ein Banachraum.

Für uns werden im Folgenden nur endliche Radonmaße relevant sein. Der Riesz'sche Darstellungssatz liefert eine Beschreibung von  $\mathcal{M}(\Omega)$  als Dualraum von  $\mathcal{C}_0(\Omega)$ , dem Raum der im Unendlichen verschwindenden stetigen Funktionen. Der Raum  $\mathcal{C}_0(\Omega)$  sei hierbei ausgestattet mit der Supremumsnorm  $\|f\|_{\infty} = \sup_{x \in \Omega} |f(x)|$ . Der (topologische) Dualraum  $\mathcal{C}_0^*(\Omega)$  kann dann in natürlicher Weise mit der Dualnorm  $\|\mu\| = \sup_{\|f\|_{\infty} \leq 1} |\langle \mu, f \rangle|$  ausgestattet werden.

**Theorem 5 (Riesz'scher Darstellungssatz für  $C_0^*$ , [6], Kap. VIII, 2.23, 2.26)** *Es sei  $\Omega$  ein lokal-kompakter Hausdorff-Raum. Dann ist durch*

$$\Theta : \mathcal{M}(\Omega) \rightarrow C_0^*(\Omega), \quad \Theta(\nu)(f) = \int_{\Omega} f \, d\nu, \quad (\nu \in \mathcal{M}(\Omega), f \in C_0(\Omega))$$

*ein linearer, ordnungstreuer Norm-Isomorphismus mit*

$$\|\Theta(\nu)\| = \|\nu\|_{\mathcal{M}}$$

*definiert.*

Der Darstellungssatz von Riesz versetzt uns also in die Lage die Räume  $\mathcal{M}(\Omega)$  und  $C_0^*(\Omega)$  miteinander zu identifizieren.

Wir benötigen im Folgenden einen Konvergenzbegriff, der schwächer ist, als der durch die Dualnorm definierte und führen daher die schwach-\* Konvergenz ein.

**Definition 9 (Schwach-\* Konvergenz, [5], Chap. VI, Sec 15)** *Für einen normierten Vektorraum  $X$  und seinen (topologischen) Dualraum  $X^*$  ist die schwach-\* Topologie die größte Topologie auf  $X^*$  bezüglich derer die Funktionale*

$$T_x : X^* \rightarrow \mathbb{R}, \quad T_x(x') = x'(x) \quad (2.9)$$

*stetig sind.*

*Eine Folge  $(x'_n)_n$  in  $X^*$  konvergiert in der schwach-\* Topologie gegen  $x' \in X^*$ , falls*

$$\lim_{n \rightarrow \infty} x'_n(x) = x'(x)$$

*für alle  $x \in X$  gilt. Wir schreiben dann  $x'_n \xrightarrow{*} x'$ .*

Wir benötigen in der Arbeit den Begriff der schwach-\* Konvergenz von Radonmaßen. Eine Folge von Radonmaßen  $\mu_n \in \mathcal{M}(\Omega)$  heißt nach obiger Definition schwach-\* konvergent gegen ein Radonmaß  $\mu \in \mathcal{M}(\Omega)$ , falls

$$\lim_{n \rightarrow \infty} \langle \mu_n, \varphi \rangle = \langle \mu, \varphi \rangle$$

für alle  $\varphi \in C_0(\Omega)$  gilt. Hierbei ist  $\langle \cdot, \cdot \rangle : \mathcal{M}(\Omega) \times C_0(\Omega) \rightarrow \mathbb{R}$ ,  $\langle \mu, \varphi \rangle := \int_{\Omega} \varphi(x) \, d\mu(x)$  die duale Paarung.

Im Verlauf der Arbeit brauchen wir eine Aussage darüber, wann eine beschränkte Folge von Radonmaßen eine schwach-\* konvergente Teilfolge besitzt. Hilfreich sind dabei die folgenden Sätze.

**Theorem 6 (Satz von Banach-Alaoglu für separable Räume, [1], Satz 6.5)** *Gegeben sei ein separabler Raum  $X$ . Dann besitzt jede Folge in der abgeschlossenen Einheitskugel  $\overline{B_1(0)}$  von  $X^*$  eine schwach-\* konvergente Teilfolge.*

**Bemerkung 1** *Dieser Satz gilt auch für jede andere Kugel  $\overline{B_R(0)}$  in  $X^*$  mit Radius  $R > 0$  [1]. Für jede beschränkte Folge  $(x'_n)_n$  in  $X^*$  existieren ein  $R > 0$  und eine Kugel  $\overline{B_R(0)}$  in  $X^*$ , sodass  $x'_n \in \overline{B_R(0)}$  für alle  $n$  gilt. Daher besitzt in einem separablen Raum jede beschränkte Folge eine schwach-\* konvergente Teilfolge.*

Um den Satz von Banach-Alaoglu für  $\mathcal{M}(\Omega)$  anwenden zu können, müssen wir sicherstellen, dass  $\mathcal{C}_0(\Omega)$  separabel ist. Ein entsprechendes Resultat für  $\mathcal{C}(\Omega)$  ist in [5] zu finden:

**Theorem 7 ([5], Corollary 16.2)** *Für jede kompakte Menge  $\Omega \subset \mathbb{R}^n$  ist der Raum der stetigen Funktionen  $\mathcal{C}(\Omega)$ , ausgestattet mit der Topologie der uniformen Konvergenz, separabel.*

Da der Raum  $\mathcal{C}_0(\Omega)$  ein Teilraum von  $\mathcal{C}(\Omega)$  ist, folgt, dass  $\mathcal{C}_0(\Omega)$  ebenfalls separabel ist, falls  $\Omega$  kompakt ist. Es gilt daher:

**Theorem 8 (Satz von Banach-Alaoglu für  $\mathcal{M}(\Omega)$ )** *Für jede kompakte Menge  $\Omega \subset \mathbb{R}^n$  besitzt jede beschränkte Folge in  $\mathcal{M}(\Omega)$  eine konvergente Teilfolge.*

*Beweis:* Da  $\Omega$  kompakt ist, folgt aus Theorem 7 dass  $\mathcal{C}(\Omega)$  und daher auch  $\mathcal{C}_0(\Omega)$  separabel sind. Durch Anwenden des Satzes von Banach-Alaoglu (Theorem 6) erhalten wir mithilfe der Bemerkung 1 die Aussage. ■

Wir möchten nun einen Abstands begriff für Wahrscheinlichkeitsmaße auf  $\Omega$ , also Maße  $\mu : \mathcal{B}(\Omega) \rightarrow [0, 1]$  mit  $\mu(\Omega) = 1$  definieren. Den Raum der Wahrscheinlichkeitsmaße bezeichnen wir mit  $\mathcal{P}(\Omega)$ .

Gegeben seien zunächst zwei vollständige, separable metrische Räume  $\Omega_1$  und  $\Omega_2$ , die jeweils mit der Borelschen  $\sigma$ -Algebra ausgestattet seien, sowie Maße  $\mu \in \mathcal{P}(\Omega_1)$  und  $\nu \in \mathcal{P}(\Omega_2)$ . Es sei weiterhin  $\pi$  ein Maß, sodass

$$\int_{\Omega_1 \times \Omega_2} (\varphi(x) + \psi(y)) \, d\pi(x, y) = \int_{\Omega_1} \varphi \, d\mu + \int_{\Omega_2} \psi \, d\nu$$

für alle messbaren Funktionen  $\varphi \in L^1(\Omega_1, \mu)$  und  $\psi \in L^1(\Omega_2, \nu)$  gelte. Dann heißen  $\mu$  und  $\nu$  Marginale von  $\pi$  und wir bezeichnen mit  $\Pi(\mu, \nu)$  die Menge aller Maße auf  $(\Omega_1 \times \Omega_2, \mathcal{B}(\Omega_1 \times \Omega_2))$  deren Marginale  $\mu$  und  $\nu$  sind. Für ein gegebenes Kostenfunktional  $c : \Omega_1 \times \Omega_2 \rightarrow \mathbb{R}$  ist dann das Monge-Kantorovich Minimierungsproblem definiert durch

$$\inf_{\pi \in \Pi(\mu, \nu)} \int_{\Omega_1 \times \Omega_2} c(x, y) \, d\pi(x, y).$$

Anschaulich minimiert dieses Funktional die Kosten eines Transports der Masse von  $\mu$  nach  $\nu$ . Für den Spezialfall, dass  $\Omega = \Omega_1 = \Omega_2$  gilt und das Kostenfunktional durch eine auf  $\Omega$  definierte Metrik gegeben ist, erhalten wir die Wasserstein-1 Metrik:

**Definition 10 (Wasserstein-1 Metrik, [15], Def. 6.1)** *Gegeben seien ein vollständiger, separabler, mit der Borelschen  $\sigma$ -Algebra ausgestatteter, metrischer Raum  $(\Omega, d_\Omega)$  und zwei Maße  $\mu, \nu \in \mathcal{P}(\Omega)$ . Dann ist die Wasserstein-1 Metrik auf  $\mathcal{P}(\Omega)$  durch*

$$W_1(\mu, \nu) = \inf_{\pi \in \Pi(\mu, \nu)} \int_{\Omega \times \Omega} d_\Omega(x, y) \, d\pi(x, y) \tag{2.10}$$

definiert.

Für die Wasserstein-1 Metrik existiert eine duale Formulierung, denn es gilt für je zwei Maße  $\mu, \nu \in \mathcal{P}_1(\Omega)$ , die Gleichung

$$W_1(\mu, \nu) = \sup_{\text{Lip}(\psi) \leq 1} \left\{ \int_{\Omega} \psi \, d\mu - \int_{\Omega} \psi \, d\nu \right\}, \tag{2.11}$$

wobei das Supremum über alle Lipschitz-stetigen Funktionen  $\psi$  mit Lipschitzkonstante  $\text{Lip}(\psi) \leq 1$  läuft ([15] *Remark 6.5*).

Wir werden im Verlauf dieser Arbeit Maße benötigen, die Linien im  $\mathbb{R}^3$  messen. Dies gelingt beispielsweise nicht mit Lebesgue-Maßen, da jede eindimensionale Linie im  $\mathbb{R}^3$  eine Lebesgue-Nullmenge ist. Wir führen daher jetzt geeignete Maße ein:

**Definition 11 (Hausdorff-Maße, [10], Def. 4.3)** *Gegeben seien ein separabler Raum  $\Omega$  und  $0 \leq s < \infty$ . Für jede Teilmenge  $E \subseteq \Omega$  ist durch  $\text{diam}(E) = \sup\{|x - y| \mid x, y \in E\}$  der Durchmesser von  $E$  definiert.*

*Für  $\delta \in [0, \infty)$  und  $n \in \mathbb{N}$  sei*

$$\mathcal{H}_\delta^n(E) = \inf \left\{ \sum_i \text{diam}(E_i)^n \mid E \subset \cup_{i=1}^\infty E_i, E_i \subset X, \text{diam}(E_i) < \delta \right\}.$$

*Das  $n$ -dimensionale Hausdorff-Maß  $\mathcal{H}^n$  definieren wir dann durch  $\mathcal{H}^n(E) = \lim_{\delta \rightarrow 0} \mathcal{H}_\delta^n(E)$ .*

Wir benötigen im Folgenden nur das 1-dimensionale Hausdorff-Maß. Für eine rektifizierbare Kurve  $\Gamma \subset \mathbb{R}^n$  entspricht  $\mathcal{H}^1(\Gamma)$  der Länge von  $\Gamma$ . Falls  $A \subset \mathbb{R}^n$  für  $n > 1$  eine nichtleere offene Menge ist, so gilt  $\mathcal{H}^1(A) = \infty$ . Das Hausdorff-Maß  $\mathcal{H}^1$  ist daher im Allgemeinen kein Radonmaß. Allerdings ist für eine  $\mathcal{H}^1$ -messbare Menge  $B \subset \mathbb{R}^n$  mit  $\mathcal{H}^1(B) < \infty$  die Einschränkung  $\mathcal{H}^1 \llcorner B$  des Hausdorff-Maßes auf  $B$  ein Radonmaß [10].

Wir benötigen zum Abschluss dieses Kapitels noch eine Aussage darüber, wann ein Teilraum von  $\mathcal{C}(\Omega)$  dicht in  $\mathcal{C}(\Omega)$  liegt. Eine Aussage darüber wann eine Unteralgebra dicht in der Algebra  $\mathcal{C}(\Omega)$  liegt, liefert der Satz von Stone-Weierstraß. Eine Unteralgebra ist ein Teilraum einer Algebra, der bezüglich Addition, Skalarmultiplikation und bezüglich der auf  $\mathcal{C}(\Omega)$  definierten Multiplikation abgeschlossen ist.

**Theorem 9 (Satz von Stone-Weierstraß, [5] Theorem 17.1)** *Gegeben sei ein kompakter Hausdorffraum  $\Omega$  und eine Unteralgebra  $\mathcal{F} \subset \mathcal{C}(\Omega)$ , die*

- *die konstanten Funktionen enthält und*
- *die für je zwei Punkte  $x, y \in \Omega$  mit  $x \neq y$  eine Funktion  $f \in \mathcal{F}$  enthält, so dass  $f(x) \neq f(y)$  gilt.*

*Dann liegt  $\mathcal{F}$  dicht in  $\mathcal{C}(\Omega)$ .*

---



---

## KAPITEL 3

---

# Ein Funktional zum Bestrafen von Knoten in Linienmaßen

Dieses Kapitel dient dazu ein Funktional einzuführen, welches in der Lage ist Knoten in Linienmaßen entsprechend der Krümmung, die in diesen Knoten auftritt, zu bestrafen.

Wir werden zunächst in Abschnitt 3.1 den Begriff des Linienmaßes einführen, wie wir ihn in dieser Arbeit verwenden, anschließend einige Resultate zu unterhalbstetigen Normen auf der  $S^2$  zeigen und zum Abschluss das Funktional einführen.

Das Vorgehen in diesem Kapitel orientiert sich an [2]. Dort wurde Krümmung von Niveaulinien in zweidimensionalen Bildern mit der hier vorgestellten Methode regularisiert.

### 3.1 Dreidimensionale Linienmaße

Wir möchten zunächst klären, was wir unter einem Linienmaß verstehen. Dazu führen wir zunächst einige graphentheoretische Begriffe ein, da wir mithilfe von Graphen die Träger der Linienmaße beschreiben wollen.

**Definition 12 (Ungerichteter und gerichteter Graph, [8], Kap. 2.1.)** *Ein geordnetes Paar  $G = (V, E)$ , welches aus einer Menge von Knoten  $V$  und einer Menge von Kanten  $E$  besteht, heißt Graph.*

*Wir nennen einen Graphen ungerichtet, falls die Menge aller Kanten  $E$  eine Teilmenge aller zwei-elementigen Teilmengen von  $V$  ist.*

*Ein Graph heißt gerichtet, falls die Menge aller Kanten  $E$  aus geordneten Paaren von Knoten  $(i, j)$ ,  $i, j \in V$  besteht.*

*Ein Graph heißt endlich, falls die Knotenmenge, und damit auch die Kantenmenge, endlich ist.*

*Eine Kante der Form  $(i, i) \in E$  für ein  $i \in V$  heißt Schleife. Wir nennen einen Graphen schleifenlos, falls keine seiner Kanten eine Schleife ist.*

Um Graphen zur Beschreibung von Trägern von Linienmaßen nutzen zu können, müssen wir den Begriff des Graphen auf reelle Vektorräume, im Speziellen auf den  $\mathbb{R}^3$  übertragen. Das geschieht mit folgender Definition:

**Definition 13 (Euklidischer Graph)** *Ein Graph heißt euklidisch, falls die Knoten des Graphen durch Punkte in  $\Omega \subseteq \mathbb{R}^3$  repräsentiert werden.*

*Es seien zwei Knoten  $i, j \in V$  und die zugehörigen Punkte  $v_i, v_j \in \Omega$  eines euklidischen Graphen gegeben. Dann wird jede Kante  $e = (i, j) \in E$  eines ungerichteten euklidischen*

Graphen durch eine Strecke  $|e| = \{\lambda v_i + (1 - \lambda)v_j | v_i, v_j \in \Omega, \lambda \in [0, 1]\}$  repräsentiert. Die Kanten  $e = (i, j) \in E$  eines gerichteten Graphen werden hingegen durch Pfeile, also gerichtete Strecken, bei denen einer der beiden Knoten  $v_i, v_j$  als Anfang und der andere als Ende identifiziert werden, repräsentiert. Mit  $\vec{e}$  bezeichnen wir den zu einer solchen gerichteten Strecke gehörigen normierten Richtungsvektor.

Wir möchten zudem die Kanten von Graphen mit zusätzlicher Information versehen. Dazu definieren wir:

**Definition 14 (Kantengewichteter Graph, [8], Kap. 2.1.)** Gegeben seien ein Graph  $G = (V, E)$  und eine Funktion  $w : E \rightarrow \mathbb{R}$ , die jeder Kante  $e \in E$  ein Gewicht  $w(e)$  zuordnet. Das geordnete Tripel  $(V, E, w)$  heißt dann kantengewichteter Graph.

Wir müssen zudem im Folgenden die Menge der betrachteten Graphen weiter einschränken. Sei dazu  $G = (V, E, w)$  ein gerichteter, kantengewichteter Graph und  $j \in V$  ein Knoten. Wir bezeichnen mit  $e_i^{\text{in}} = (i, j) \in E$  die Kanten mit Endknoten  $j$  und mit  $e_i^{\text{out}} = (j, i) \in E$  die Kanten mit Anfangsknoten  $j$ . Wir fordern dann, dass die im Folgenden betrachteten gerichteten Graphen der Ein- und Ausflussbedingung

$$\sum_{\substack{i \in V \text{ s.d.} \\ (i, j) \in E}} w(e_i^{\text{in}}) - \sum_{\substack{i \in V \text{ s.d.} \\ (j, i) \in E}} w(e_i^{\text{out}}) = 0 \quad (3.1)$$

für jeden Knoten  $j \in V$  genügen.

Wir sind mit diesen Definitionen nun in der Lage Linienmaße zu definieren. Dazu sei im Folgenden  $\Omega \subset \mathbb{R}^3$  ein beschränktes Gebiet. Wir führen zunächst Linienmaße auf  $\Omega$  ein

**Definition 15 (Linienmaße auf  $\Omega$ )** Gegeben sei ein ungerichteter, kantengewichteter, schleifenloser, endlicher und euklidischer Graph  $G = (V, E, w)$  mit Punkten in  $\Omega$ . Die Kantengewichte  $w(e)$  seien zudem für alle Kanten  $e \in E$  positiv. Dann definieren wir das Linienmaß  $u \in \mathcal{M}(\Omega)$  durch

$$\int_{\Omega} \varphi(x) \, du(x) = \sum_{e \in E} w(e) \int_{|e|} \varphi(x) \, d\mathcal{H}^1 \llcorner |e|$$

für alle  $\varphi \in \mathcal{C}_0(\Omega)$ .

Als nächstes führen wir, analog zum *functional lifting* in [2] ein Linienmaß auf  $\Omega \times S^2$  ein.

**Definition 16 (Linienmaße auf  $\Omega \times S^2$ )** Gegeben sei ein gerichteter, kantengewichteter, schleifenloser, endlicher und euklidischer Graph  $G = (V, E, w)$  mit Punkten in  $\Omega$ , welcher der Ein- und Ausflussbedingung (3.1) genügt. Die Kantengewichte  $w(e)$  seien zudem für alle Kanten  $e \in E$  positiv. Dann definieren wir das Linienmaß  $\mu \in \mathcal{M}(\Omega \times S^2)$  durch

$$\int_{\Omega \times S^2} \varphi(x, v) \, d\mu(x, v) = \sum_{e \in E} w(e) \int_{|e|} \varphi(x, \vec{e}) \, d\mathcal{H}^1 \llcorner |e|$$

für alle  $\varphi \in \mathcal{C}_0(\Omega \times S^2)$ .

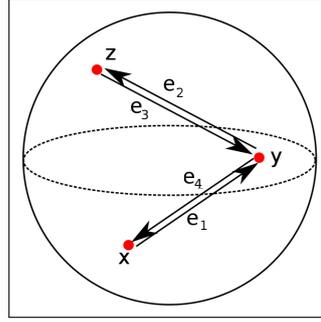


Abbildung 3.1: Graph zum motivierenden Beispiel

Wir bemerken, dass  $\mathcal{H}^1 \llcorner |e|$  Einschränkungen von  $\mathcal{H}^1$  auf  $\mathcal{H}^1$ -messbare und bezüglich  $\mathcal{H}^1$  endliche Mengen sind. Für die Einschränkungen gilt daher  $\mathcal{H}^1 \llcorner |e| \in \mathcal{M}(\Omega)$ . Die Maße  $u$  und  $\mu$  sind daher ebenso endliche Radonmaße.

Aus einem ungerichteten kantengewichteten Graphen  $G = (V, E, w)$  erhalten wir einen zugehörigen gerichteten kantengewichteten Graphen  $G^* = (V, E^*, w^*)$ , indem wir fordern dass für jede Kante  $e = \{i, j\} \in E$  mit  $i, j \in V$  gilt, dass beide Kanten  $e_1^* = (i, j)$  und  $e_2^* = (j, i)$  in  $E^*$  sind und, dass  $w(e) = \frac{1}{2}w^*(e_1^*) = \frac{1}{2}w^*(e_2^*)$  gilt. Für die oben definierten Linienmaße  $u$  und  $\mu$  gilt dann

$$\int_{\Omega} \varphi \, du = \int_{\Omega \times S^2} \varphi(x) \, d\mu(x, v) \quad \forall \varphi \in \mathcal{C}_0(\Omega). \quad (3.2)$$

Für solche  $u$  und  $\mu$  möchten wir im Folgenden eine konvexe Krümmungsregularisierung einführen und nennen (3.2) Projektionsbedingung.

Die Idee dabei ist es für ein gegebenes Linienmaß  $u \in \mathcal{M}(\Omega)$  das Linienmaß  $\mu \in \mathcal{M}(\Omega \times S^2)$  entsprechend der Bedingung (3.2) zu konstruieren und mit Hilfe von  $\mu$  dann analog zu [2] die Krümmung in Knoten zu bestrafen.

Zum Abschluss dieses Abschnitts möchten wir noch das weitere Vorgehen mit Hilfe eines Beispiels motivieren. Dazu nehmen wir an, dass wir  $\mu$  mit der Projektionsbedingung (3.2) aus  $u$  konstruiert haben. Der  $\mu$  zu Grunde liegende euklidische Graph habe drei Punkte  $x, y, z \in \Omega = B_1(0)$  und Kanten  $e_1, e_2, e_3$  und  $e_4$  wie in Abbildung 3.1. Für die Kantengewichte gelte  $w(e_1) = w(e_2) = w(e_3) = w(e_4) = 1$ . Dann gilt für  $\varphi \in \mathcal{C}_0(\Omega \times S^2)$

$$\begin{aligned} \int_{\Omega \times S^2} \varphi \, d\mu &= w(e_1) \int_{|e_1|} \varphi(x, \vec{e}_1) \, d\mathcal{H}^1 \llcorner |e_1| + w(e_2) \int_{|e_2|} \varphi(x, \vec{e}_2) \, d\mathcal{H}^1 \llcorner |e_2| \\ &\quad + w(e_3) \int_{|e_3|} \varphi(x, \vec{e}_3) \, d\mathcal{H}^1 \llcorner |e_3| + w(e_4) \int_{|e_4|} \varphi(x, \vec{e}_4) \, d\mathcal{H}^1 \llcorner |e_4|. \end{aligned}$$

Mit  $\varphi(x, v) = \nabla_x \psi(x, v) \cdot v$  für  $\nabla_x \psi \in \mathcal{C}_0(\Omega \times S^2, \mathbb{R}^2)$  gilt weiter

$$\begin{aligned}
 \int_{\Omega \times S^2} \varphi \, d\mu &= \int_0^1 |y-x| \nabla_x \psi((1-t)x + ty, \vec{e}_1) \cdot \vec{e}_1 \, dt \\
 &\quad + \int_0^1 |z-y| \nabla_x \psi((1-t)y + tz, \vec{e}_2) \cdot \vec{e}_2 \, dt \\
 &\quad + \int_0^1 |y-z| \nabla_x \psi((1-t)z + ty, \vec{e}_3) \cdot \vec{e}_3 \, dt \\
 &\quad + \int_0^1 |x-y| \nabla_x \psi((1-t)y + tx, \vec{e}_4) \cdot \vec{e}_4 \, dt \\
 &= \int_0^1 \frac{\partial}{\partial t} \psi((1-t)x + ty, \vec{e}_1) \, dt + \int_0^1 \frac{\partial}{\partial t} \psi((1-t)y + tz, \vec{e}_2) \, dt \\
 &\quad + \int_0^1 \frac{\partial}{\partial t} \psi((1-t)z + ty, \vec{e}_3) \, dt + \int_0^1 \frac{\partial}{\partial t} \psi((1-t)y + tx, \vec{e}_4) \, dt \\
 &= \psi(y, \vec{e}_1) - \psi(x, \vec{e}_1) + \psi(z, \vec{e}_2) - \psi(y, \vec{e}_2) \\
 &\quad + \psi(y, \vec{e}_3) - \psi(z, \vec{e}_3) + \psi(x, \vec{e}_4) - \psi(y, \vec{e}_4) \\
 &= \langle \delta_{\vec{e}_4} - \delta_{\vec{e}_1}, \psi(x, \cdot) \rangle + \langle \delta_{\vec{e}_1} - \delta_{\vec{e}_2} + \delta_{\vec{e}_3} - \delta_{\vec{e}_4}, \psi(y, \cdot) \rangle + \langle \delta_{\vec{e}_2} - \delta_{\vec{e}_3}, \psi(z, \cdot) \rangle.
 \end{aligned}$$

Wir erhalten also für jeden Knoten von  $\mu$  eine Differenz von Diracmaßen  $\delta_{\vec{e}}$ , die eine Richtungsänderung im Knoten beschreibt. Unser Ziel ist es mit Hilfe der auftretenden Differenzen von Diracmaßen  $\delta_{\vec{e}_i} - \delta_{\vec{e}_j}$  die Richtungsänderungen in den Knoten von Linienmaßen entsprechend des Abstandes der Richtungen  $\vec{e}_i$  und  $\vec{e}_j$  auf der  $S^2$  zu bestrafen.

### 3.2 Eine Norm auf $\mathcal{M}(S^2)$ , die Richtungsänderungen bestraft

In diesem Abschnitt möchten wir eine Norm auf  $\mathcal{M}(S^2)$  definieren, welche Differenzen von Diracmaßen wie vorhin beschrieben bestraft. Unser Vorgehen ist dabei analog zu [2]. Dort wurden solche Normen auf  $\mathcal{M}(S^1)$  eingeführt.

Im Folgenden sei  $\rho : S^2 \times S^2 \rightarrow [0, \infty)$  immer eine unterhalbstetige Metrik auf der  $S^2$ . Es existieren also stetige Metriken  $\rho_i \in \mathcal{C}(S^2 \times S^2)$  für  $i \in I \neq \emptyset$  sodass  $\rho = \sup_{i \in I} \rho_i$  gilt. Für gegebenes  $\rho$  definieren wir die Menge

$$C_\rho = \left\{ \varphi \in \mathcal{C}(S^2) \mid \varphi(\eta_1) - \varphi(\eta_2) \leq \rho(\eta_1, \eta_2) \text{ für } (\eta_1, \eta_2) \in S^2 \times S^2 \right\} \quad (3.3)$$

und definieren anschließend ein Funktional auf  $\mathcal{M}(S^2)$  durch

$$\|\mu\|_\rho = \sup_{\varphi \in C_\rho} \langle \mu, \varphi \rangle. \quad (3.4)$$

Im Rest des Abschnitts werden wir zeigen, dass  $\|\cdot\|_\rho$  das Verlangte leistet. Die vorgestellten Resultate dazu werden aus [2] entnommen und auf unseren Fall übertragen.

Wir beginnen indem wir zeigen, dass  $\|\cdot\|_\rho$  schwach-\* folgenunterhalbstetig ist.

**Proposition 1** *Das Funktional  $\|\cdot\|_\rho : \mathcal{M}(S^2) \rightarrow \mathbb{R}$  ist folgenunterhalbstetig bezüglich schwach-\* konvergenter Folgen, d.h. es gilt*

$$\liminf_{n \rightarrow \infty} \|\mu_n\|_\rho \geq \|\mu\|_\rho \quad (3.5)$$

für alle schwach-\* konvergenten Folgen  $\mu_n \xrightarrow{*} \mu$  in  $\mathcal{M}(S^2)$ .

*Beweis:* Per Definition der schwach-\* Topologie ist die Abbildung  $\langle \cdot, \varphi \rangle : \mathcal{M}(\Omega) \rightarrow \mathbb{R}$  für gegebenes  $\varphi$  folgenstetig bezüglich schwach-\* konvergenter Folgen. Dann ist  $\|\cdot\|_\rho$  schwach-\* folgenunterhalbstetig als Supremum schwach-\* folgenstetiger Funktionale. ■

Im Rest des Abschnitts zeigen wir, dass  $\|\cdot\|_\rho$  das am Ende von Abschnitt 3.1 Verlangte leistet. Zunächst zeigen wir dazu, dass Differenzen von Diracmaßen  $\delta_{\vec{e}_i} - \delta_{\vec{e}_j}$  entsprechend des Abstandes von  $\vec{e}_i$  und  $\vec{e}_j$  auf der  $S^2$  bestraft werden:

**Proposition 2** *Gegeben sei eine unterhalbstetige Metrik  $\rho : S^2 \times S^2 \rightarrow [0, \infty)$ . Dann gilt*

$$\|\delta_{v_1} - \delta_{v_2}\|_\rho = \rho(v_1, v_2)$$

für alle  $v_1, v_2 \in S^2$ .

*Beweis:* Es gilt  $\|\delta_{v_1} - \delta_{v_2}\|_\rho = \sup_{\varphi \in C_\rho} \varphi(v_1) - \varphi(v_2)$  per Definition von  $\|\cdot\|_\rho$ . Die Abschätzung  $\|\delta_{v_1} - \delta_{v_2}\|_\rho \leq \rho(v_1, v_2)$  ergibt sich daher aus der Definition von  $C_\rho$ , da für alle  $\varphi \in C_\rho$  die Ungleichung  $\varphi(v_1) - \varphi(v_2) \leq \rho(v_1, v_2)$  gilt.

Zu zeigen bleibt also die Ungleichung  $\|\delta_{v_1} - \delta_{v_2}\|_\rho \geq \rho(v_1, v_2)$ . Dazu bemerken wir, dass

$$\sup_{\varphi \in C_\rho} \varphi(v_1) - \varphi(v_2) \geq \sup_{i \in I} \sup_{\varphi \in C_{\rho_i}} \varphi(v_1) - \varphi(v_2)$$

gilt, da  $C_{\rho_i} \subseteq C_\rho$  gilt. Wir müssen nun  $\sup_{\varphi \in C_{\rho_i}} \varphi(v_1) - \varphi(v_2)$  nach unten durch  $\rho_i(v_1, v_2)$  abschätzen. Es sei dazu  $\varphi_i^* \in C(S^2)$ ,  $\varphi_i^*(v) = \rho_i(v, v_2)$  gegeben. Dann ist  $\varphi_i^* \in C_{\rho_i}$ , da

$$\varphi_i^*(\eta_1) - \varphi_i^*(\eta_2) = \rho_i(\eta_1, v_2) - \rho_i(\eta_2, v_2) \leq \rho_i(\eta_1, \eta_2)$$

gilt. Wählen wir  $\eta_1 = v_1$  und  $\eta_2 = v_2$ , so gilt

$$\sup_{\varphi_i \in C_{\rho_i}} \varphi_i(v_1) - \varphi_i(v_2) \geq \varphi_i^*(v_1) - \varphi_i^*(v_2) = \rho_i(v_1, v_2) - \rho_i(v_2, v_2) = \rho_i(v_1, v_2).$$

Mithilfe der Unterhalbstetigkeit von  $\rho$  erhalten wir weiter

$$\begin{aligned} \|\delta_{v_1} - \delta_{v_2}\|_\rho &\geq \sup_{\varphi \in C_\rho} \varphi(v_1) - \varphi(v_2) \\ &\geq \sup_{i \in I} \sup_{\varphi \in C_{\rho_i}} \varphi(v_1) - \varphi(v_2) \\ &\geq \sup_{i \in I} \rho_i(v_1, v_2) \\ &= \rho(v_1, v_2). \end{aligned}$$

Somit haben wir beide Ungleichungen gezeigt und es gilt die behauptete Gleichheit. ■

Wir müssen nun noch zeigen, dass durch  $\|\cdot\|_\rho$  tatsächlich eine Norm auf  $\mathcal{M}(S^2)$  definiert wird.

**Theorem 10** *Gegeben sei eine unterhalbstetige Metrik  $\rho : S^2 \times S^2 \rightarrow [0, \infty)$ . Dann ist das durch  $\|\mu\|_\rho = \sup_{\varphi \in C_\rho} \langle \mu, \varphi \rangle$  für  $\mu \in \mathcal{M}(S^2)$  definierte Funktional  $\|\cdot\|_\rho : \mathcal{M}(S^2) \rightarrow \mathbb{R} \cup \{\infty\}$  eine Norm auf  $\mathcal{M}(S^2)$ .*

*Beweis:* Wir weisen die Eigenschaften einer Norm nach und beginnen mit der Nichtnegativität. Da die Metrik  $\rho$  symmetrisch ist, gelten die Ungleichungen  $\rho(\eta_1, \eta_2) \geq \varphi(\eta_1) - \varphi(\eta_2)$  und  $\rho(\eta_1, \eta_2) \geq \varphi(\eta_2) - \varphi(\eta_1) = -(\varphi(\eta_1) - \varphi(\eta_2))$  für  $\varphi \in C_\rho$ . Deshalb gilt  $C_\rho = -C_\rho$  und  $\|\mu\|_\rho = \sup_{\varphi \in C_\rho} \langle \mu, \varphi \rangle$  ist daher nichtnegativ für alle  $\mu \in \mathcal{M}(S^2)$ .

Die positive Homogenität folgt mit  $C_\rho = -C_\rho$  und der Bilinearität der dualen Paarung  $\langle \cdot, \cdot \rangle$  aus

$$\|\lambda\mu\|_\rho = \sup_{\varphi \in C_\rho} \langle \lambda\mu, \varphi \rangle = \sup_{\varphi \in C_\rho} \langle |\lambda|\mu, \operatorname{sgn}(\lambda)\varphi \rangle = |\lambda| \sup_{\varphi \in C_\rho} \langle \mu, \varphi \rangle = |\lambda| \|\mu\|_\rho.$$

Die Dreiecksungleichung ergibt sich direkt aus

$$\langle \mu_1 + \mu_2, \varphi \rangle = \langle \mu_1, \varphi \rangle + \langle \mu_2, \varphi \rangle \leq \|\mu_1\|_\rho + \|\mu_2\|_\rho$$

durch Übergang zum Supremum.

Zu zeigen bleibt also noch die positive Definitheit. Sei dazu  $\mu \in \mathcal{M}(S^2)$  so gegeben, dass  $\|\mu\|_\rho = 0$  gilt. Dies bedeutet aufgrund der Definition von  $\|\cdot\|_\rho$ , dass für alle  $\varphi \in C_\rho$  auch  $\langle \mu, \varphi \rangle \leq 0$  gilt. Deshalb gilt auch  $\langle \mu, \varphi \rangle \leq 0$  für alle  $\varphi \in V := \bigcup_{\lambda \geq 0} \lambda C_\rho$ . Wir möchten nun zeigen, dass  $V$  dicht in  $\mathcal{C}(S^2)$  liegt. Dazu nutzen wir den Satz von Stone-Weierstraß (Theorem 9) und weisen zunächst nach, dass  $V$  eine Unteralgebra von  $\mathcal{C}(S^2)$  ist.

Wir beginnen mit der Abgeschlossenheit der Skalarmultiplikation auf  $V$ . Es sei dazu ein Skalar  $\alpha \in \mathbb{R}$  gegeben. Dann ist wegen  $|\alpha|\lambda \geq 0$  für  $\lambda \geq 0$  die Menge  $|\alpha|\lambda C_\rho$  in  $V$  enthalten. Für eine Funktion  $\varphi_0 \in C_\rho$  ist wegen  $C_\rho = -C_\rho$  auch  $\operatorname{sgn}(\alpha)\varphi_0 \in C_\rho$ . Da die Funktion  $\varphi := \lambda\varphi_0$  für  $\lambda \geq 0$  in  $V$  ist, gilt

$$\alpha\varphi = |\alpha|\lambda \operatorname{sgn}(\alpha)\varphi_0 \in |\alpha|\lambda C_\rho \subset V$$

und  $V$  ist somit bezüglich Skalarmultiplikation abgeschlossen.

Es sind nun noch die Abgeschlossenheit von  $V$  bezüglich Addition und Multiplikation zu zeigen. Für  $\varphi, \psi \in V$  existieren Funktionen  $\varphi_0, \psi_0 \in C_\rho$  und Konstanten  $\lambda_1 \geq 0$  und  $\lambda_2 \geq 0$ , sodass  $\varphi = \lambda_1\varphi_0$  und  $\psi = \lambda_2\psi_0$  gelten.

Die Abgeschlossenheit von  $V$  bezüglich der Addition erhalten wir aus der Abschätzung

$$\begin{aligned} (\varphi + \psi)(v_1) - (\varphi + \psi)(v_2) &= \lambda_1(\varphi_0(v_1) - \varphi_0(v_2)) + \lambda_2(\psi_0(v_1) - \psi_0(v_2)) \\ &\leq (\lambda_1 + \lambda_2)\rho(v_1, v_2), \end{aligned}$$

denn es gilt  $(\varphi + \psi) \in (\lambda_1 + \lambda_2)C_\rho \subset V$ .

Aus der Abschätzung

$$\begin{aligned} (\varphi\psi)(v_1) - (\varphi\psi)(v_2) &= \lambda_1\lambda_2(\psi_0(v_1)(\varphi_0(v_1) - \varphi_0(v_2))) \\ &\leq \lambda_1\lambda_2(\|\psi_0\|_\infty + \|\varphi_0\|_\infty)\rho(v_1, v_2) \end{aligned}$$

erhalten wir  $(\varphi\psi) \in \lambda_1\lambda_2(\|\psi_0\|_\infty + \|\varphi_0\|_\infty)C_\rho \subset V$  und damit die Abgeschlossenheit von  $V$  bezüglich der Multiplikation.

Damit ist gezeigt, dass  $V$  eine  $\mathbb{R}$ -Unteralgebra von  $\mathcal{C}(S^2)$  ist. Um den Satz von Stone-Weierstraß anwenden zu können muss nun noch gezeigt werden, dass  $V$  Punkte separiert und, dass die konstanten Funktionen in  $V$  enthalten sind.

Dass die konstanten Funktionen in  $V$  enthalten sind, ist offensichtlich, denn für eine konstante Funktion  $\varphi \in \mathcal{C}(S^2)$  gilt

$$\varphi(v_1) - \varphi(v_2) = 0 \leq \rho(v_1, v_2)$$

für alle  $v_1, v_2 \in S^2$  und damit ist  $\varphi \in \lambda C_\rho \subset V$  für alle  $\lambda \geq 0$ .

Die Eigenschaft, dass  $V$  Punkte separiert folgt aus Proposition 2, denn für  $v_1, v_2 \in S^2$  mit  $v_1 \neq v_2$  existiert ein  $\varphi \in C_\rho$ , sodass  $\varphi(v_1) - \varphi(v_2) = \rho_i(v_1, v_2) > 0$  für ein  $i \in I$  gilt.

Die Unteralgebra  $V$  erfüllt also alle Voraussetzungen um den Satz von Stone-Weierstraß anwenden zu können. Es gilt daher  $\bar{V} = \mathcal{C}(S^2)$ . Da die Abbildung  $\varphi \mapsto \langle \varphi, \mu \rangle$  stetig ist und per Voraussetzung die Ungleichung  $\langle \varphi, \mu \rangle \leq 0$  für  $\varphi \in V$  gilt, muss aufgrund der Dichtheit von  $V$  in  $\mathcal{C}(S^2)$  diese Ungleichung für alle  $\varphi \in \mathcal{C}(S^2)$  gelten. Für  $\mu \neq 0$  widerspricht dies allerdings der Bilinearität der dualen Paarung  $\langle \cdot, \cdot \rangle : \mathcal{M}(S^2) \times \mathcal{C}(S^2) \rightarrow \mathbb{R}$ . Es muss also  $\mu = 0$  sein. Das Funktional  $\|\cdot\|_\rho$  ist also positiv definit.

Damit ist nun nachgewiesen, dass  $\|\cdot\|_\rho$  eine Norm auf  $\mathcal{M}(S^2)$  definiert. ■

Wir haben in diesem Abschnitt die Resultate aus [2] auf unseren Fall übertragen können und so eine Norm auf  $\mathcal{M}(S^2)$  definiert, die Richtungsänderungen entsprechend der Distanz der Richtungen auf der  $S^2$  bezüglich einer gegebenen Metrik  $\rho$  bestraft.

### 3.3 Eine Funktional, das Richtungsänderungen von Linienmaßen bestraft

Wir übertragen nun die Resultate aus dem letzten Abschnitt auf die in Definition von Linienmaßen  $\mu$  auf  $\Omega \times S^2$  für ein beschränktes Gebiet  $\Omega \subset \mathbb{R}^3$ . Dazu nutzen wir die im motivierenden Beispiel in Abschnitt 3.1 verdeutlichte Möglichkeit aus dem Integral

$$\int_{\Omega \times S^2} \varphi \, d\mu \quad \text{für } \varphi \in \mathcal{C}_0(\Omega \times S^2)$$

durch Übergang zu  $\varphi(x, v) = \nabla_x \psi(x, v) \cdot v$  für  $\psi \in \mathcal{C}_0(\Omega \times S^2, \mathbb{R}^3)$  Differenzen von Diracmaßen zu erhalten, welche Richtungsänderungen in den Knoten von  $\mu$  beschreiben. Diese Differenzen möchten wir anschließend mit Hilfe der in Abschnitt 3.2 definierten Norm bestrafen. Das geschieht mit folgender Definition, die sich analog zum zweidimensionalen Fall in [2] ergibt:

**Definition 17** Gegeben seien eine unterhalbstetige Metrik  $\rho : S^2 \times S^2 \rightarrow [0, \infty)$ . Dann definieren wir das Funktional  $T_\rho : \mathcal{M}(\Omega \times S^2) \rightarrow \mathbb{R}$  durch

$$T_\rho(\mu) = \sup_{\psi \in M_\rho(\Omega)} \int_{\Omega \times S^2} \nabla_x \psi(x, v) \cdot v \, d\mu(x, v) \quad (3.6)$$

wobei die Menge  $M_\rho(\Omega)$  durch

$$M_\rho(\Omega) = \left\{ \psi \in \mathcal{C}_0(\Omega \times S^2) \mid \nabla_x \psi \in \mathcal{C}_0(\Omega \times S^2, \mathbb{R}^3), \psi(x, \cdot) \in C_\rho \text{ für alle } x \in \Omega \right\}$$

definiert sei.

Wir möchten nun zeigen, dass  $T_\rho$  das Verlangte leistet. Es sei dazu ein Graph  $G = (V, E, w)$  wie in Definition 16 gegeben. Für jede Kante  $e \in E$  bezeichnen wir mit  $x_e \in \Omega$  den Anfangspunkt und mit  $y_e$  den Endpunkt der zur Kante  $e$  gehörigen gerichteten Strecke in  $\Omega$ .

Weiterhin seien  $s_e(t) := (1-t)x_e + ty_e$  und  $v_e = \frac{y_e - x_e}{|y_e - x_e|}$  gegeben. Dann gilt:

$$\begin{aligned}
 T_\rho(\mu) &= \sup_{\psi \in M_\rho(\Omega)} \int_{\Omega \times S^2} \nabla_x \psi(x, v) \cdot v \, d\mu(x, v) \\
 &= \sup_{\psi \in M_\rho(\Omega)} \sum_{e \in E} w(e) \int_{|e|} \nabla_x \psi(x, \vec{e}) \cdot \vec{e} \, d\mathcal{H}^1 \llcorner |e| \\
 &= \sup_{\psi \in M_\rho(\Omega)} \sum_{e \in E} w(e) \int_0^1 |y_e - x_e| \nabla_x \psi(s_e(t), v_e) \cdot v_e \, dt \\
 &= \sup_{\psi \in M_\rho(\Omega)} \sum_{e \in E} w(e) \int_0^1 \frac{\partial}{\partial t} \psi(s_e(t), v_e) \, dt \\
 &= \sup_{\psi \in M_\rho(\Omega)} \sum_{e \in E} w(e) (\psi(y_e, v_e) - \psi(x_e, v_e)) \\
 &= \sup_{\psi \in M_\rho(\Omega)} \sum_{y \in V} \left( \sum_{\substack{e=(x,y) \in E \\ \text{für } x \in V}} w(e) \psi(y, v_e) - \sum_{\substack{d=(y,x) \in E \\ \text{für } x \in V}} w(d) \psi(y, v_d) \right).
 \end{aligned}$$

In der letzten Gleichung haben wir hierbei nur die Summanden so umsortiert, dass alle Summanden zum gleichen Knoten  $y \in V$  in einer Summe stehen. Wir formulieren diese Summe mit Hilfe von Diracmaßen nun weiter um zu

$$T_\rho(\mu) = \sup_{\psi \in M_\rho(\Omega)} \sum_{y \in V} \left( \sum_{\substack{e=(x,y) \in E \\ \text{für } x \in V}} w(e) \langle \delta_{v_e}, \psi(y, \cdot) \rangle - \sum_{\substack{d=(y,x) \in E \\ \text{für } x \in V}} w(d) \langle \delta_{v_d}, \psi(y, \cdot) \rangle \right).$$

Wir halten kurz fest, dass  $T_\rho$  einige Eigenschaften von  $\|\cdot\|_\rho$  erbt:

**Proposition 3** *Das Funktional  $T_\rho$  ist nichtnegativ, positiv homogen, erfüllt die Dreiecksungleichung  $T_\rho(\mu + \nu) \leq T_\rho(\mu) + T_\rho(\nu)$  und ist folgenunterhalbstetig bezüglich schwach-\* konvergenter Folgen.*

*Beweis:* Die Aussagen ergeben sich analog zu den entsprechenden Aussagen für  $\|\cdot\|_\rho$  in Abschnitt 3.2. Für  $\varphi \in M_\rho$  ist auch  $-\varphi \in M_\rho$ , da zum Einen für  $\nabla_x \varphi \in \mathcal{C}_0(\Omega \times S^2, \mathbb{R}^3)$  auch  $\nabla_x(-\varphi) \in \mathcal{C}_0(\Omega \times S^2, \mathbb{R}^3)$  gilt und zum Anderen  $C_\rho = -C_\rho$  ist (Theorem 10). Es gilt daher  $M_\rho = -M_\rho$  und da in  $T_\rho$  das Supremum über die Menge  $M_\rho$  gebildet wird, folgt die Nichtnegativität von  $T_\rho$ .

Die positive Homogenität erhalten wir wieder mit Hilfe der Mengengleichheit  $M_\rho = -M_\rho$ , denn es gilt

$$\begin{aligned}
 T_\rho(\lambda\mu) &= \sup_{\varphi \in M_\rho} \int_{\Omega \times S^2} |\lambda| \varphi \, d(\text{sgn}(\lambda)\mu) \\
 &= |\lambda| \sup_{\varphi \in \text{sgn}(\lambda)M_\rho} \int_{\Omega \times S^2} \varphi \, d\mu \\
 &= |\lambda| T_\rho(\mu).
 \end{aligned}$$

Die Dreiecksungleichung ist ebenfalls erfüllt, da

$$\begin{aligned} T_\rho(\mu_1 + \mu_2) &= \sup_{\varphi \in M_\rho} \int_{\Omega \times S^2} \varphi \, d(\mu_1 + \mu_2) \\ &\leq \sup_{\varphi \in M_\rho} \int_{\Omega \times S^2} \varphi \, d\mu_1 + \sup_{\varphi \in M_\rho} \int_{\Omega \times S^2} \varphi \, d\mu_2 \\ &= T_\rho(\mu_1) + T_\rho(\mu_2) \end{aligned}$$

gilt.

Die schwach-\* Folgenunterhalbstetigkeit ergibt sich wie im Beweis zu Proposition 1.  $\blacksquare$

**Bemerkung 2** Das Funktional  $T_\rho$  ist konvex. Dies ergibt sich direkt aus der positiven Homogenität und der Dreiecksungleichung, denn für  $\mu_1, \mu_2 \in \mathcal{M}(\Omega \times S^2)$  und  $t \in [0, 1]$  gilt:

$$T_\rho((1-t)\mu_1 + t\mu_2) \leq T_\rho((1-t)\mu_1) + T_\rho(t\mu_2) = (1-t)T_\rho(\mu_1) + tT_\rho(\mu_2).$$

Wir werden nun anhand von zwei Beispielen verdeutlichen, wie dieses Funktional Richtungsänderungen in Linienmaßen bestraft. Die Beispiele sind analog zu [2] gewählt und werden im Verlauf der Arbeit noch benötigt.

**Beispiel 1** Wir wählen als Metrik die diskrete Metrik  $\rho_0 : S^2 \times S^2 \rightarrow [0, \infty)$ , die durch

$$\rho_0(v_1, v_2) = \begin{cases} 0, & \text{falls } v_1 = v_2 \\ 1, & \text{sonst} \end{cases}$$

definiert ist. Diese Metrik ist eine geeignete Wahl als Metrik in Abschnitt 3.2, da sie wegen  $\rho_0(v_1, v_2) = \sup_{\lambda > 0} \min \left\{ 1, \frac{|v_1 - v_2|}{\lambda} \right\}$  unterhalbstetig ist.

Unser Ziel ist es nun  $T_{\rho_0}$  zu bestimmen. Dazu benötigen wir zunächst die Menge

$$\begin{aligned} C_{\rho_0} &= \left\{ \varphi \in \mathcal{C}(S^2) \mid \varphi(v_1) - \varphi(v_2) \leq \rho(v_1, v_2) \forall v_1, v_2 \in S^2 \right\} \\ &= \left\{ \varphi \in \mathcal{C}(S^2) \mid \max_{v \in S^2} \varphi(v) - \min_{v \in S^2} \varphi(v) \leq 1 \right\} \\ &= \left\{ \varphi \in \mathcal{C}(S^2) \mid \varphi = \psi + c \text{ für } \|\psi\|_\infty \leq \frac{1}{2} \text{ und } c \in \mathbb{R} \right\}. \end{aligned}$$

Es gilt daher

$$\|\mu\|_{\rho_0} = \sup_{\varphi \in C_{\rho_0}} \langle \mu, \varphi \rangle = \frac{1}{2} \sup_{\|\varphi\|_\infty \leq 1} \langle \mu, \varphi \rangle = \frac{1}{2} \|\mu\|_{\mathcal{M}}.$$

Wir möchten nun eine Interpretation von  $T_{\rho_0}$  finden. Es sei dazu ein Linienmaß  $\mu \in (\Omega \times S^2)$  und der zugehörige Graph  $G = (V, E, w)$  gegeben. Wir führen zur Erleichterung der Notation zunächst einige Mengen ein. Es seien für jedes  $y \in V$  die Mengen

$$\begin{aligned} E_y^{in} &= \{d \in E \mid d = (x, y) \text{ für ein } x \in V\} \text{ und} \\ E_y^{out} &= \{e \in E \mid e = (y, x) \text{ für ein } x \in V\} \end{aligned}$$

gegeben. Dies sind genau diejenigen Kanten des (gerichteten) Graphen  $G$ , für die  $y$  entweder der Anfangs- oder der Endknoten ist. Weiterhin seien die Menge

$$R_y = \{(d, e) \in E_y^{in} \times E_y^{out} \mid \vec{d} = \vec{e}\}$$

sowie die Mengen

$$\begin{aligned} S_y^{in} &= \{d \in E_y^{in} \mid \vec{d} \neq \vec{e} \forall e \in E_y^{out}\} \text{ und} \\ S_y^{out} &= \{e \in E_y^{out} \mid \vec{e} \neq \vec{d} \forall d \in E_y^{in}\} \end{aligned}$$

gegeben. Anschaulich enthält die Menge  $R_y$  genau jene Paare  $(d, e)$  von Kanten entlang derer eine Masse ohne Richtungsänderung im Knoten  $y$  fließen kann. Die Menge  $S_y^{in} \cup S_y^{out}$  enthält genau diejenigen Kanten für die das nicht möglich ist. Für  $T_{\rho_0}$  gilt dann

$$\begin{aligned} T_{\rho_0}(\mu) &= \sup_{\psi \in M_{\rho_0}(\Omega)} \int_{\Omega \times S^2} \nabla_x \psi(x, v) \cdot v \, d\mu(x, v) \\ &= \sup_{\psi \in M_{\rho_0}(\Omega)} \sum_{y \in V} \left( \sum_{d \in E_y^{in}} w(d) \langle \delta_{\vec{d}}, \psi(y, \cdot) \rangle - \sum_{e \in E_y^{out}} w(e) \langle \delta_{\vec{e}}, \psi(y, \cdot) \rangle \right) \\ &= \sup_{\psi \in M_{\rho_0}(\Omega)} \sum_{y \in V} \left( \sum_{(d,e) \in R_y} \left( w(d) \langle \delta_{\vec{d}}, \psi(y, \cdot) \rangle - w(e) \langle \delta_{\vec{e}}, \psi(y, \cdot) \rangle \right) \right. \\ &\quad \left. + \sum_{d \in S_y^{in}} w(d) \langle \delta_{\vec{d}}, \psi(y, \cdot) \rangle - \sum_{e \in S_y^{out}} w(e) \langle \delta_{\vec{e}}, \psi(y, \cdot) \rangle \right) \\ &= \sum_{y \in V} \sup_{\psi \in M_{\rho_0}(\Omega)} \left( \sum_{(d,e) \in R_y} \left( w(d) \langle \delta_{\vec{d}}, \psi(y, \cdot) \rangle - w(e) \langle \delta_{\vec{e}}, \psi(y, \cdot) \rangle \right) \right. \\ &\quad \left. + \sum_{d \in S_y^{in}} w(d) \langle \delta_{\vec{d}}, \psi(y, \cdot) \rangle - \sum_{e \in S_y^{out}} w(e) \langle \delta_{\vec{e}}, \psi(y, \cdot) \rangle \right), \end{aligned}$$

wobei wir in der letzten Gleichung ausgenutzt haben, dass die Summanden der äußeren Summe voneinander unabhängig sind. Wir schränken unsere Betrachtung nun zunächst auf einen Knoten  $y_0 \in V$  ein und bemerken zunächst, dass wegen  $\vec{d} = \vec{e}$  in der Summe über  $R_{y_0}$  der Anteil

$$\min \{w(d), w(e)\} \left( \langle \delta_{\vec{d}}, \psi(y_0, \cdot) \rangle - \langle \delta_{\vec{e}}, \psi(y_0, \cdot) \rangle \right)$$

Null wird. In dieser Summe bleibt als Rest also

$$\left( \max \{w(d), w(e)\} - \min \{w(d), w(e)\} \right) \cdot \begin{cases} \langle \delta_{\vec{d}}, \psi(y_0, \cdot) \rangle & , \text{ falls } w(d) > w(e) \\ -\langle \delta_{\vec{e}}, \psi(y_0, \cdot) \rangle & , \text{ falls } w(e) > w(d) \\ 0 & , \text{ sonst} \end{cases}$$

stehen. Nutzen wir nun aus, dass

$$\frac{1}{2} \|\delta_{\vec{d}}\|_{\mathcal{M}} = \sup_{\psi(y_0, \cdot) \in C_{\rho_0}} \langle \delta_{\vec{d}}, \psi(y_0, \cdot) \rangle = \sup_{\psi(y_0, \cdot) \in C_{\rho_0}} \langle \delta_{\vec{e}}, -\psi(y_0, \cdot) \rangle = \frac{1}{2} \|\delta_{\vec{e}}\|_{\mathcal{M}}$$

wegen  $C_{\rho_0} = -C_{\rho_0}$  gilt und, dass für die Totalvariationsnorm des Diracmaßes  $\|\delta_{\vec{d}}\|_{\mathcal{M}} = 1$  gilt, so erhalten wir

$$\begin{aligned} & \sup_{\psi(y_0, \cdot) \in C_{\rho_0}} \sum_{(d,e) \in R_{y_0}} \left( w(e) \langle \delta_{\vec{d}}, \psi(y_0, \cdot) \rangle - w(d) \langle \delta_{\vec{d}}, \psi(y_0, \cdot) \rangle \right) \\ &= \frac{1}{2} \cdot \sum_{(d,e) \in R_{y_0}} (\max \{w(d), w(e)\} - \min \{w(d), w(e)\}). \end{aligned}$$

Entsprechend gilt für die Summen über  $S_{y_0}^{in}$  und  $S_{y_0}^{out}$

$$\sup_{\psi(y_0, \cdot) \in C_{\rho_0}} \sum_{d \in S_{y_0}^{in}} \left( w(d) \langle \delta_{\vec{d}}, \psi(y_0, \cdot) \rangle \right) - \sum_{e \in S_{y_0}^{out}} \left( w(e) \langle \delta_{\vec{e}}, \psi(y_0, \cdot) \rangle \right) = \frac{1}{2} \cdot \sum_{d \in S_{y_0}^{in}} w(d) + \frac{1}{2} \cdot \sum_{e \in S_{y_0}^{out}} w(e).$$

Fasst man die Gewichte als Massen auf, die über die gerichteten Kanten durch den Knoten  $y_0$  fließen, so wird durch das Funktional  $T_{\rho_0}$  genau derjenige Anteil der Masse bestraft, der nicht ohne Richtungsänderung durch den Knoten fließen kann. Insgesamt erhalten wir dann

$$T_{\rho_0}(\mu) = \frac{1}{2} \sum_{y \in V} \left( \sum_{(d,e) \in R_y} (\max \{w(d), w(e)\} - \min \{w(d), w(e)\}) + \sum_{d \in S_y^{in}} w(d) + \sum_{e \in S_y^{out}} w(e) \right).$$

Das Funktional  $T_{\rho_0}$  bestraft alle Richtungsänderungen, die in einem Knoten auftreten in gleicher Weise. Im folgenden Beispiel wählen wir eine Metrik, die eine Richtungsänderung entsprechend der Distanz der beiden Richtungen auf der  $S^2$  bestraft.

**Beispiel 2** Gegeben seien zwei Richtungen  $v_1, v_2 \in S^2$ . Den Winkel  $\lambda$  zwischen  $v_1$  und  $v_2$  können wir dann mit Hilfe des Skalarprodukts  $v_1 \cdot v_2 = \cos(\lambda)$  berechnen und durch

$$\text{dist}_{S^2}(v_1, v_2) = \lambda = \cos^{-1}(v_1 \cdot v_2)$$

eine Metrik auf der  $S^2$  definieren. Diese ist stetig, da  $\cos^{-1} : [-1, 1] \rightarrow [0, \pi]$  stetig ist und  $v_1 \cdot v_2 \in [-1, 1]$  ist. Sie ist also insbesondere unterhalbstetig und ist daher eine geeignete Wahl in Abschnitt 3.2. Wir setzen daher

$$\rho_1(v_1, v_2) = \text{dist}_{S^2}(v_1, v_2) \tag{3.7}$$

und erhalten

$$\begin{aligned} C_{\rho_1} &= \left\{ \varphi \in \mathcal{C}(S^2) \mid \varphi(v_1) - \varphi(v_2) \leq \rho_1(v_1, v_2) \forall v_1, v_2 \in S^2 \right\} \\ &= \left\{ \varphi \in \mathcal{C}(S^2) \mid \varphi \text{ Lipschitz-stetig mit Lipschitzkonstante } \text{Lip}(\varphi) \leq 1 \right\}. \end{aligned}$$

Wir möchten nun wieder eine Interpretation dessen erhalten, was das Funktional  $T_{\rho_1}$  bewirkt. Es sei dazu wieder ein Linienmaß  $\mu \in \mathcal{M}(\Omega \times S^2)$  und der zugehörige Graph  $G = (V, E, w)$  gegeben. Dann gilt

$$\begin{aligned} T_{\rho_1}(\mu) &= \sup_{\psi \in M_{\rho_1}(\Omega)} \int_{\Omega \times S^2} \nabla_x \psi(x, v) \cdot v \, d\mu(x, v) \\ &= \sup_{\psi \in M_{\rho_1}(\Omega)} \sum_{y \in V} \left( \sum_{\substack{d=(x,y) \in E \\ \text{für } x \in V}} w(d) \langle \delta_{\vec{d}}, \psi(y, \cdot) \rangle - \sum_{\substack{e=(y,x) \in E \\ \text{für } x \in V}} w(e) \langle \delta_{\vec{e}}, \psi(y, \cdot) \rangle \right). \end{aligned}$$

Zur Vereinfachung der Notation definieren wir nun die Maße

$$\nu_y^{in} = \sum_{\substack{d=(x,y) \in E \\ \text{für } x \in V}} w(d) \delta_{\vec{d}} \text{ und } \nu_y^{out} = \sum_{\substack{e=(x,y) \in E \\ \text{für } x \in V}} w(e) \delta_{\vec{e}}$$

und erhalten

$$T_{\rho_1}(\mu) = \sup_{\psi \in M_{\rho_1}(\Omega)} \sum_{y \in Y} \langle \nu_y^{in}, \psi(y, \cdot) \rangle - \langle \nu_y^{out}, \psi(y, \cdot) \rangle.$$

Wegen der Unabhängigkeit der Summanden folgt weiterhin

$$T_{\rho_1}(\mu) = \sum_{y \in Y} \sup_{\psi \in M_{\rho_1}(\Omega)} \left( \langle \nu_y^{in}, \psi(y, \cdot) \rangle - \langle \nu_y^{out}, \psi(y, \cdot) \rangle \right).$$

Wir bemerken zunächst, dass wegen der Ein- und Ausflussbedingung 3.1 die Maße  $\nu_y^{in}$  und  $\nu_y^{out}$  die gleiche Masse haben, es gilt also  $\nu_y^{in}(S^2) = \nu_y^{out}(S^2)$ . Schränken wir unsere Betrachtung nun wieder auf einen Knoten  $y_0 \in Y$  ein, so erhalten wir daher

$$\begin{aligned} \sup_{\psi(y_0, \cdot) \in C_{\rho_1}} \langle \nu_{y_0}^{in}, \psi(y_0, \cdot) \rangle - \langle \nu_{y_0}^{out}, \psi(y_0, \cdot) \rangle &= \sup_{\text{Lip}(\psi(y_0, \cdot)) \leq 1} \langle \nu_{y_0}^{in}, \psi(y_0, \cdot) \rangle - \langle \nu_{y_0}^{out}, \psi(y_0, \cdot) \rangle \\ &= W_1(\nu_{y_0}^{in}, \nu_{y_0}^{out}). \end{aligned}$$

Hierbei ist  $W_1$  die in Definition 10 eingeführte Wasserstein-Metrik. Insgesamt ergibt sich somit:

$$T_{\rho_1}(\mu) = \sum_{y \in Y} W_1(\nu_y^{in}, \nu_y^{out}).$$

---



---

## KAPITEL 4

---

# Konvexe Krümmungsregularisierung

Unser Ziel ist es nun die Resultate aus den vergangenen Kapiteln zusammenzutragen und ein konvexes Minimierungsproblem zur Entrauschung von Linienmaßen

$$\min_{(u, \mu)} G(u) + R(u)$$

einzuführen.

In der Einleitung haben wir als Wahl für den Strafterm  $R$  die Eulersche Elastica (1.1) eingeführt, die sowohl die Länge als auch die Krümmung von Kurven bestraft. Für Linienmaße erhalten wir eine Bestrafung der Länge durch  $\|\mu\|_{\mathcal{M}} = \sum_{e \in E} w(e) \mathcal{H}^1(|e|)$ , wobei  $\|\cdot\|_{\mathcal{M}}$  die Totalvariationsnorm aus Definition 8 ist. Mit Hilfe des Funktionals  $T_\rho$  aus Definition 17 können wir zudem die Krümmung von  $\mu$  in Knoten bestrafen. Für Linienmaße  $u \in \mathcal{M}(\Omega)$  und  $\mu \in \mathcal{M}(\Omega \times S^2)$ , welche die Projektionsbedingung (3.2) erfüllen, wählen wir daher entsprechend der Eulerschen Elastica das Funktional

$$R_\rho^{\alpha, \beta} : \mathcal{M}(\Omega) \rightarrow [0, \infty], \quad R_\rho^{\alpha, \beta}(u) = \alpha \|\mu\|_{\mathcal{M}} + \beta T_\rho(\mu).$$

als Strafterm. Hierbei sind  $\alpha$  und  $\beta$  positive Parameter und  $\rho$  eine unterhalbstetige Metrik auf  $S^2$ .

Um ein konvexes Minimierungsproblem zu erhalten, müssen wir zum einen sicherstellen, dass  $R_\rho^{\alpha, \beta}$  ein konvexes Funktional ist und zum anderen, dass der Definitionsbereich des Minimierungsproblems eine konvexe Menge ist. Dies geschieht in Abschnitt 4.1. Falls wir dann für den Datenterm  $G$  ebenfalls ein konvexes Funktional wählen, so erhalten wir ein konvexes Optimierungsproblem. Im Abschnitt 4.2 werden wir dann für spezielle Wahlen von  $G$  zeigen, dass ein Minimum des Optimierungsproblems existiert.

In [2] wurden entsprechende Resultate zur Konvexität des Minimierungsproblems und der Existenz von Minima für zweidimensionale Bilder hergeleitet. Wir übertragen die dortigen Resultate auf unser Problem.

### 4.1 Konvexität des Minimierungsproblems

Der Definitionsbereich des Minimierungsproblems ist gegeben durch die Menge

$$N = \left\{ (u, \mu) \mid u \in \mathcal{M}(\Omega), \mu \in \mathcal{M}(\Omega \times S^2) \text{ Linienmaße, } \int_{\Omega} \varphi \, du = \int_{\Omega \times S^2} \varphi(x) \, d\mu \, \forall \varphi \in \mathcal{C}_0(\Omega) \right\}.$$

Analog zu Proposition 4.1. in [2] definieren wir eine konvexe Obermenge, über die wir stattdessen optimieren können.

**Proposition 4** *Die Menge*

$$M = \left\{ (u, \mu) \mid u \in \mathcal{M}(\Omega), \mu \in \mathcal{M}(\Omega \times S^2), \mu \geq 0 \right. \\ \left. \int_{\Omega} \varphi \, du = \int_{\Omega \times S^2} \varphi(x) \, d\mu \text{ für alle } \varphi \in \mathcal{C}_0(\Omega) \text{ und} \right. \\ \left. \int_{\Omega \times S^2} \nabla_x \varphi(x) \cdot v \, d\mu(x, v) = 0 \text{ für alle } \varphi \in \mathcal{C}_0(\Omega) \text{ mit } \nabla_x \varphi \in \mathcal{C}_0(\Omega, \mathbb{R}^3) \right\}$$

ist nicht leer, konvex und folgenabgeschlossen bezüglich schwach-\* konvergenter Folgen.

*Beweis:* Wir beginnen indem wir zeigen, dass  $N \subset M$  gilt. Gegeben sei dazu  $(u, \mu) \in N$ . Wegen der Definitionen der Linienmaße  $u$  und  $\mu$  gelten  $u \in \mathcal{M}(\Omega)$  und  $\mu \in \mathcal{M}(\Omega \times S^2)$ . Da zudem die Gewichte der Linienmaße nicht negativ sind, gilt weiterhin  $\mu \geq 0$ . Wir müssen noch zeigen, dass die Gleichung

$$\int_{\Omega \times S^2} \nabla_x \varphi(x) \cdot v \, d\mu(x, v) = 0 \quad (4.1)$$

für alle  $\varphi \in \mathcal{C}_0(\Omega)$  mit  $\nabla_x \varphi \in \mathcal{C}_0(\Omega, \mathbb{R}^3)$  gilt. Da  $\mu$  ein Linienmaß ist, existiert ein zu Grunde liegender Graph  $G = (V, E, w)$ . Für eine Kante  $e \in E$  sei  $x_e$  der Anfangs- und  $y_e$  der Endknoten. Es gilt dann

$$\begin{aligned} \int_{\Omega \times S^2} \nabla_x \varphi(x) \cdot v \, d\mu(x, v) &= \sum_{e \in E} w(e) \int_{|e|} \nabla_x \varphi(x) \cdot \vec{e} \, d\mathcal{H}^1 \llcorner |e| \\ &= \sum_{e \in E} w(e) \int_0^1 |y_e - x_e| \nabla_x \varphi((1-t)x_e + ty_e) \cdot \vec{e} \, dt \\ &= \sum_{e \in E} w(e) \int_0^1 \partial_t \varphi((1-t)x_e + ty_e) \, dt \\ &= \sum_{e \in E} w(e) (\varphi(y_e) - \varphi(x_e)). \end{aligned}$$

Fassen wir nun Summanden zum gleichen Knoten zusammen, so erhalten wir

$$\int_{\Omega \times S^2} \nabla_x \varphi(x) \cdot v \, d\mu(x, v) = \sum_{y \in V} \varphi(y) \left( \sum_{\substack{d=(x,y) \in E \\ x \in V}} w(d) - \sum_{\substack{e=(y,x) \in E \\ x \in V}} w(e) \right)$$

für alle  $\varphi \in \mathcal{C}_0(\Omega)$  mit  $\nabla_x \varphi \in \mathcal{C}_0(\Omega, \mathbb{R}^3)$ . Wegen der Ein- und Ausflussbedingung (3.1) wird der Ausdruck in der Klammer 0 und daher gilt die Gleichung (4.1). Damit ist gezeigt, dass  $N \subset M$  gilt.

Wir möchten nun die Konvexität der Menge  $M$  zeigen. Es seien dazu  $(u_1, \mu_1) \in M$  und  $(u_2, \mu_2) \in M$  gegeben. Dann werden durch  $u_t = (1-t)u_1 + tu_2$  und  $\mu_t = (1-t)\mu_1 + t\mu_2$  für  $t \in [0, 1]$  endliche, nichtnegative Radonmaße auf  $\Omega$  beziehungsweise  $\Omega \times S^2$  definiert. Für diese gilt

$$\begin{aligned} \int_{\Omega} \varphi \, du_t &= (1-t) \int_{\Omega} \varphi \, du_1 + t \int_{\Omega} \varphi \, du_2 \\ &= (1-t) \int_{\Omega} \varphi(x) \, d\mu_1(x, v) + t \int_{\Omega} \varphi(x) \, d\mu_2(x, v) \\ &= \int_{\Omega} \varphi(x) \, d\mu_t \end{aligned}$$

für alle  $\varphi \in \mathcal{C}_0(\Omega)$ . Die erste und dritte Gleichung gelten aufgrund der Linearitätsaussage im Riesz'schen Darstellungssatz (Theorem 5). Die zweite Gleichung gilt, da die Paare  $(u_1, \mu_1)$  und  $(u_2, \mu_2)$  Elemente aus  $M$  sind.

Ebenso gilt deshalb auch für alle  $\varphi \in \mathcal{C}_0(\Omega)$  mit  $\nabla_x \varphi(x) \in \mathcal{C}_0(\Omega, \mathbb{R}^3)$  die Gleichung

$$\begin{aligned} & \int_{\Omega \times S^2} \nabla_x \varphi(x) \cdot v \, d\mu_t \\ &= (1-t) \int_{\Omega \times S^2} \nabla_x \varphi(x) \cdot v \, d\mu_1(x, v) + t \int_{\Omega \times S^2} \nabla_x \varphi(x) \cdot v \, d\mu_2(x, v) \\ &= (1-t) \cdot 0 + t \cdot 0. \end{aligned}$$

Damit ist für  $(u_1, \mu_1) \in M$  und  $(u_2, \mu_2) \in M$  auch  $(u_t, \mu_t) \in M$  für  $t \in [0, 1]$ . Die Menge  $M$  ist also konvex.

Zu zeigen bleibt also noch die Folgenabgeschlossenheit bezüglich schwach-\* konvergenter Folgen. Es sei dazu  $\{(u_n, \mu_n)\}_n$  eine Folge in  $M$ , so dass  $u_n \xrightarrow{*} u$  und  $\mu_n \xrightarrow{*} \mu$  für ein  $u \in \mathcal{M}(\Omega)$  und ein  $\mu \in \mathcal{M}(\Omega \times S^2)$  gilt. Per Definition der schwach-\* gelten dann die Gleichungen

$$\int_{\Omega} \varphi(x) \, du(x) - \int_{\Omega \times S^2} \varphi(x) \, d\mu(x, v) = \lim_{n \rightarrow \infty} \int_{\Omega} \varphi(x) \, du_n(x) - \int_{\Omega \times S^2} \varphi(x) \, d\mu_n(x, v) = 0$$

für alle  $\varphi \in \mathcal{C}_0(\Omega)$  und

$$\int_{\Omega \times S^2} \nabla_x \varphi(x) \cdot v \, d\mu(x, v) = \lim_{n \rightarrow \infty} \int_{\Omega \times S^2} \nabla_x \varphi(x) \cdot v \, d\mu_n(x, v) = 0$$

für alle  $\varphi \in \mathcal{C}_0(\Omega)$  mit  $\nabla_x \varphi \in \mathcal{C}_0(\Omega, \mathbb{R}^3)$ . Die Menge  $M$  ist daher folgenabgeschlossen.  $\blacksquare$

Mit Hilfe der Menge  $M$  sind wir nun in der Lage analog zu [2] ein konvexes Funktional  $\overline{R}_\rho^{\alpha, \beta}$  zu definieren, welches wir im Anschluss als Strafterm des Minimierungsproblems nutzen können.

**Definition 18** Für ein Maß  $u \in \mathcal{M}(\Omega)$  und Parameter  $a > 0$  und  $b > 0$  sei das Funktional  $\overline{R}_\rho^{\alpha, \beta} : \mathcal{M}(\Omega) \rightarrow [0, \infty]$  definiert durch

$$\overline{R}_\rho^{\alpha, \beta}(u) = \inf_{\substack{\mu \in \mathcal{M}(\Omega \times S^2) \\ (u, \mu) \in M}} \alpha \|\mu\|_{\mathcal{M}} + \beta T_\rho(\mu).$$

Hierbei sei das Infimum über die leere Menge  $\infty$ .

Wir weisen nun analog zu Proposition 4.4. in [2] einige Eigenschaften von  $\overline{R}_\rho^{\alpha, \beta}$  nach. Diese werden wir im nächsten Abschnitt benötigen um die Existenz von Minima des Optimierungsproblems sicherstellen zu können.

**Proposition 5** Für  $\alpha > 0$ ,  $\beta > 0$  und eine unterhalbstetige Metrik  $\rho$  auf  $S^2$  ist das Funktional  $\overline{R}_\rho^{\alpha, \beta} : \mathcal{M}(\Omega) \rightarrow [0, \infty]$  für jede kompakte Teilmenge  $\Omega \subset \mathbb{R}^3$  konvex und folgenunterhalbstetig bezüglich schwach-\* konvergenter Folgen.

*Beweis:* Die Nichtnegativität von  $\bar{R}_\rho^{\alpha,\beta}$  ist klar, da die Totalvariationsnorm per Definition und das Funktional  $T_\rho$  nach Proposition 3 nichtnegativ sind.

Wir beginnen daher mit der schwach-\* Folgenunterhalbstetigkeit. Dazu sei eine schwach-\* konvergente Folge  $(u_n)_n$  in  $\mathcal{M}(\Omega)$  mit Grenzwert  $u \in \mathcal{M}(\Omega)$  gegeben. Für jedes  $n$  sei eine Minimalfolge  $((\mu_n)_m)_m$  in  $\mathcal{M}(\Omega \times S^2)$  so gegeben, dass  $(u_n, (\mu_n)_m) \in M$  und

$$\lim_{m \rightarrow \infty} \alpha \|(\mu_n)_m\|_{\mathcal{M}} + \beta T_\rho((\mu_n)_m) = \bar{R}_\rho^{\alpha,\beta}(u_n)$$

gilt. Wir möchten nun zeigen, dass eine Teilfolge von  $((\mu_n)_m)_m$  beschränkt ist. Dazu nehmen wir zunächst an, dass die Folge  $\bar{R}_\rho^{\alpha,\beta}(u_n)$  Werte in  $[0, \infty)$  annehme und gegen einen endlichen Wert konvergiert. Falls  $\bar{R}_\rho^{\alpha,\beta}(u_n) \rightarrow \infty$  gilt, so ist  $\bar{R}_\rho^{\alpha,\beta}(u) \leq \lim_{n \rightarrow \infty} \bar{R}_\rho^{\alpha,\beta}(u_n)$ . In diesem Fall ist also nichts zu zeigen. Falls  $\bar{R}_\rho^{\alpha,\beta}(u_n) = \infty$  für ein  $n \in \mathbb{N}$  gilt, so ist ebenfalls  $\bar{R}_\rho^{\alpha,\beta}(u) \leq \bar{R}_\rho^{\alpha,\beta}(u_n)$ .

Wir nehmen deshalb nun an, dass  $\bar{R}_\rho^{\alpha,\beta}(u_n) \in [0, \infty)$  für alle  $n$  und  $\bar{R}_\rho^{\alpha,\beta}(u) < \infty$  für  $u_n \xrightarrow{*} u$  gilt. Dann existiert eine Diagonalfolge  $(\tilde{\mu}_n)$  von  $((\mu_n)_m)$ , für die

$$\alpha \|\tilde{\mu}_n\|_{\mathcal{M}} + \beta T_\rho(\tilde{\mu}_n) \leq \bar{R}_\rho^{\alpha,\beta}(u_n) + \frac{1}{n}$$

für alle  $n \geq 1$  gilt. Wir erhalten daher für  $\|\tilde{\mu}_n\|_{\mathcal{M}}$  die Abschätzung

$$\begin{aligned} \|\tilde{\mu}_n\|_{\mathcal{M}} &\leq \alpha^{-1} (\alpha \|\tilde{\mu}_n\|_{\mathcal{M}} + \beta T_\rho(\tilde{\mu}_n)) \\ &\leq \alpha^{-1} \left( \sup_{n \geq 1} \bar{R}_\rho^{\alpha,\beta}(u_n) + 1 \right) \\ &< \infty, \end{aligned}$$

da  $\bar{R}_\rho^{\alpha,\beta}(u_n)$  per Annahme endlich ist. Die Diagonalfolge  $(\tilde{\mu}_n)_n$  in  $\mathcal{M}(\Omega \times S^2)$  ist also beschränkt. Da  $\Omega \times S^2$  per Annahme kompakt ist, können wir den Satz von Banach-Alaoglu (Theorem 8) anwenden. Es existiert also eine Teilfolge  $(\tilde{\mu}_{n_k})_k$ , so dass  $\tilde{\mu}_{n_k} \xrightarrow{*} \mu$  für ein  $\mu \in \mathcal{M}(\Omega \times S^2)$  und für  $k \rightarrow \infty$  gilt. Da  $M \subset \mathcal{M}(\Omega) \times \mathcal{M}(\Omega \times S^2)$  nach Proposition 4 folgenabgeschlossen bezüglich schwach-\* konvergenter Folgen ist, gilt auch  $(u, \mu) \in M$ .

Wir möchten nun zeigen, dass  $\bar{R}_\rho^{\alpha,\beta}$  schwach-\* folgenunterhalbstetig ist. Dazu bemerken wir zunächst, dass die Variationsnorm  $\|\cdot\|_{\mathcal{M}}$  schwach-\* folgenunterhalbstetig ist ([1], Bemerkungen 6.3 (3)). In Proposition 3 haben wir gezeigt, dass das Funktional  $T_\rho$  ebenfalls schwach-\* folgenunterhalbstetig ist. Es gilt daher

$$\begin{aligned} \bar{R}_\rho^{\alpha,\beta}(u) &\leq \alpha \|u\|_{\mathcal{M}} + \beta T_\rho(u) \\ &\leq \liminf_{k \rightarrow \infty} \alpha \|\tilde{\mu}_{n_k}\|_{\mathcal{M}} + \beta T_\rho(\tilde{\mu}_{n_k}) \\ &= \lim_{n \rightarrow \infty} \bar{R}_\rho^{\alpha,\beta}(u_n). \end{aligned}$$

Die erste Ungleichung gilt, da  $(u, \mu) \in M$  ist und die zweite Ungleichung gilt aufgrund der schwach-\* Unterhalbstetigkeiten von  $\|\cdot\|_{\mathcal{M}}$  und  $T_\rho$ .

Damit gilt  $\bar{R}_\rho^{\alpha,\beta}(u) \leq \lim_{n \rightarrow \infty} \bar{R}_\rho^{\alpha,\beta}(u_n)$  für alle schwach-\* konvergenten Folgen  $u_n \xrightarrow{*} u$  in  $\mathcal{M}(\Omega)$ . Das Funktional  $\bar{R}_\rho^{\alpha,\beta}$  ist also schwach-\* folgenunterhalbstetig.

Es bleibt noch die Konvexität von  $\overline{R}_\rho^{\alpha,\beta}$  zu zeigen. Diese ergibt sich genau wie in [2]. Dazu seien Maße  $u^1, u^2 \in \mathcal{M}(\Omega)$  gegeben. Wir müssen zeigen, dass

$$\overline{R}_\rho^{\alpha,\beta} \left( (1-t)u^1 + tu^2 \right) \leq (1-t)\overline{R}_\rho^{\alpha,\beta}(u^1) + t\overline{R}_\rho^{\alpha,\beta}(u^2)$$

für  $t \in [0, 1]$  gilt. Falls  $\overline{R}_\rho^{\alpha,\beta}(u^1) = \infty$  oder  $\overline{R}_\rho^{\alpha,\beta}(u^2) = \infty$  gilt, so ist diese Ungleichung erfüllt. Wir nehmen daher an, dass  $\overline{R}_\rho^{\alpha,\beta}(u^1) < \infty$  und  $\overline{R}_\rho^{\alpha,\beta}(u^2) < \infty$  gilt. Dann existieren Minimalfolgen  $(\mu_n^1)_n$  und  $(\mu_n^2)_n$ . Für diese Minimalfolgen sind  $(u^1, (\mu_n^1))$  und  $(u^2, (\mu_n^2))$  für alle  $n$  in  $M$ . Weiterhin gelten die Gleichungen

$$\overline{R}_\rho^{\alpha,\beta}(u^1) = \lim_{n \rightarrow \infty} \alpha \|\mu_n^1\|_{\mathcal{M}} + \beta T_\rho(\mu_n^1) \text{ und } \overline{R}_\rho^{\alpha,\beta}(u^2) = \lim_{n \rightarrow \infty} \alpha \|\mu_n^2\|_{\mathcal{M}} + \beta T_\rho(\mu_n^2).$$

Wegen der Konvexität von  $M$  ist für jedes  $n$  auch  $((1-t)u^1 + tu^2, (1-t)\mu_n^1 + t\mu_n^2)$  für alle  $t \in [0, 1]$  in  $M$  enthalten. Die Variationsnorm  $\|\cdot\|_{\mathcal{M}}$  ist positiv homogen und subadditiv und daher konvex. Das Funktional  $T_\rho$  ist wegen Bemerkung 2 ebenfalls konvex. Es gilt deshalb

$$\begin{aligned} \overline{R}_\rho^{\alpha,\beta}((1-t)u^1 + tu^2) &\leq \liminf_{n \rightarrow \infty} \left( \alpha \|(1-t)\mu_n^1 + t\mu_n^2\|_{\mathcal{M}} + \beta T_\rho \left( (1-t)\mu_n^1 + t\mu_n^2 \right) \right) \\ &\leq (1-t) \lim_{n \rightarrow \infty} \left( \alpha \|\mu_n^1\|_{\mathcal{M}} + \beta T_\rho(\mu_n^1) \right) \\ &\quad + t \lim_{n \rightarrow \infty} \left( \alpha \|\mu_n^2\|_{\mathcal{M}} + \beta T_\rho(\mu_n^2) \right) \\ &= (1-t)\overline{R}_\rho^{\alpha,\beta}(u^1) + t\overline{R}_\rho^{\alpha,\beta}(u^2). \end{aligned}$$

■

Falls wir nun ein konvexes Funktional  $G : \mathcal{M}(\Omega) \rightarrow (-\infty, \infty]$  als Datenterm wählen, so erhalten wir durch

$$\min_{u \in \mathcal{M}(\Omega)} G(u) + \overline{R}_\rho^{\alpha,\beta}(u) \tag{4.2}$$

ein konvexes Optimierungsproblem.

## 4.2 Existenz eines Minimums

Wir möchten nun sicherstellen, dass Minima des Optimierungsproblem (4.2) existieren. Dies geschieht mit Hilfe einer weiteren Annahme an den Datenterm  $G$ :

**Theorem 11** *Gegeben seien positive Parameter  $\alpha > 0$  und  $\beta > 0$ , eine unterhalbstetige Metrik  $\rho$  auf  $S^2$  und eine kompakte Teilmenge  $\Omega \subset \mathbb{R}^3$ . Weiterhin sei ein nach unten beschränktes und schwach-\* folgenunterhalbstetiges Funktional  $G : \mathcal{M}(\Omega) \rightarrow (-\infty, \infty]$  mit*

$$G(u_n) \rightarrow \infty, \text{ falls } \int_{\Omega} \varphi du_n \rightarrow \infty \text{ für ein } \varphi \in C_0(\Omega) \text{ mit } \|\varphi\|_{\infty} \leq 1$$

*gegeben. Dann existiert eine Lösung des Optimierungsproblems 4.2. Diese Lösung ist eindeutig, falls  $G$  strikt konvex ist.*

*Beweis:* Wir nutzen zum Beweis des Theorems die direkte Methode der Variationsrechnung. Dazu sei das Funktional

$$F := G + \overline{R}_\rho^{\alpha,\beta} : \mathcal{M}(\Omega) \rightarrow (-\infty, \infty]$$

gegeben. Dieses Funktional ist nach unten beschränkt, da  $\overline{R}_\rho^{\alpha,\beta}$  nach Proposition 5 und  $G$  per Annahme nach unten beschränkt sind. Nehmen wir an, dass  $F(u) = \infty$  für alle  $u$  gilt, so ist die Aussage klar. Wir nehmen daher an, dass  $F$  für mindestens ein  $u \in \mathcal{M}(\Omega)$  endlich ist. Da  $F$  nach unten beschränkt ist, existiert dann eine Minimalfolge  $(u_n)_n$  für die

$$\lim_{n \rightarrow \infty} F(u_n) = \inf_u F(u) < \infty$$

gilt. Wir möchten nun zeigen, dass die Folge  $(u_n)_n$  beschränkt ist. Dazu bemerken wir, dass  $\int_\Omega \varphi du_n$  für alle  $\varphi \in C_0(\Omega)$  mit  $\|\varphi\|_\infty \leq 1$  endlich sein muss, da sonst  $G(u_n) \rightarrow \infty$  und deshalb auch  $F(u_n) \rightarrow \infty$  gelten würde, was der Annahme widerspricht. Es gilt also

$$\|u_n\|_{\mathcal{M}} = \sup_{\substack{\varphi \in C_0(\Omega) \\ \|\varphi\|_\infty \leq 1}} \int_\Omega \varphi du_n < \infty$$

und damit ist die Folge  $(u_n)_n$  beschränkt. Da  $\Omega$  per Annahme kompakt ist, können wir den Satz von Banach-Alaoglu (Theorem 8) anwenden. Es existiert also eine Teilfolge  $(u_{n_k})_k$  von  $(u_n)_n$ , sodass  $u_{n_k} \xrightarrow{*} u^*$  für ein  $u^* \in \mathcal{M}(\Omega)$  gilt.

Weiterhin ist  $F$  schwach-\* folgenunterhalbstetig, denn  $G$  ist es per Annahme und  $\overline{R}_\rho^{\alpha,\beta}$  ist nach Proposition 5 schwach-\* folgenunterhalbstetig.

Wir erhalten damit die Abschätzung

$$\inf_{u \in \mathcal{M}(\Omega)} F(u) \stackrel{1.}{=} \lim_{n \rightarrow \infty} F(u_n) \stackrel{2.}{=} \lim_{k \rightarrow \infty} F(u_{n_k}) \stackrel{3.}{\geq} F(u^*) \stackrel{4.}{\geq} \inf_{u \in \mathcal{M}(\Omega)} F(u),$$

wobei wir

- 1. die Existenz einer Minimalfolge,
- 2. die schwach-\* Konvergenz einer Teilfolge,
- 3. die Unterhalbstetigkeit von  $F$  und
- 4. die Definition des Infimums

ausgenutzt haben. Es gilt also  $F(u^*) = \inf_{u \in \mathcal{M}} F(u) > -\infty$  wegen der Beschränktheit von  $F$ . Damit ist die Existenz einer Lösung des Optimierungsproblems (4.2) gezeigt.

Falls  $G$  nun strikt konvex folgt aufgrund der Konvexität von  $\overline{R}_\rho^{\alpha,\beta}$  (Prop. 5), dass auch  $F = G + \overline{R}_\rho^{\alpha,\beta}$  strikt konvex ist. Für  $u_1, u_2 \in \mathcal{M}(\Omega)$  mit  $F(u_1) = F(u_2) = \inf_{u \in \mathcal{M}(\Omega)} F(u)$  gilt  $F((1-t)u_1 + tu_2) < F(u_1) = \inf_{u \in \mathcal{M}(\Omega)} F(u)$ . Dies ist ein Widerspruch, da auch  $(1-t)u_1 + tu_2 \in \mathcal{M}(\Omega)$  ist. Es muss daher  $u_1 = u_2$  gelten und die Lösung des Optimierungsproblems ist eindeutig. ■

Zum Abschluss dieses Kapitels möchten wir nun zwei Beispiele für Datenterme  $G$  geben, die den Voraussetzungen im Theorem 11 genügen.

### 4.2.1 Existenz eines Minimums für $L^2$ Daten

In der Bildverarbeitung werden häufig  $L^2$ -Datenterme zur Lösung von Entrauschungsproblemen genutzt, da diese insbesondere bei additivem Gausschen Rauschen, das in der Digitalfotographie beispielsweise bei zu geringer Beleuchtung auftritt, gute Ergebnisse liefern. Einen solchen Datenterm möchten wir auch für unser Optimierungsproblem nutzen können. Wir benötigen dazu noch die folgende Definition:

**Definition 19 (Schwache Konvergenz in  $L^2$ , [5], Chap. V, Sec. 9)** Eine Folge  $(f_n)_n$  in  $L^2(\Omega)$  heißt schwach konvergent gegen ein  $f \in L^2(\Omega)$ , falls

$$\lim_{n \rightarrow \infty} \int_{\Omega} f_n g \, d\lambda = \int_{\Omega} f g \, d\lambda$$

für alle  $g \in L^2(\Omega)$  gilt. Hierbei ist  $\lambda \in \mathcal{M}(\Omega)$  das dreidimensionale Lebesgue-Maß. Wir schreiben dann  $f_n \rightharpoonup f$ .

Ein Funktional  $G : L^2(\Omega) \rightarrow [0, \infty]$  heißt schwach folgenunterhalbstetig, falls für jede Folge  $(f_n)_n$  in  $L^2(\Omega)$  mit  $f_n \rightharpoonup f$  die Abschätzung:

$$\liminf_{n \rightarrow \infty} G(f_n) \geq G(f)$$

gilt.

Gegeben seien nun verrauschte Daten  $f \in L^2(\Omega)$ . Dann definieren wir den  $L^2$ -Datenterm durch

$$G : L^2(\Omega) \rightarrow [0, \infty), \quad G(v) = \frac{1}{2} \int_{\Omega} |v - f|^2 \, d\lambda(x)$$

wobei  $\lambda \in \mathcal{M}(\Omega)$  das Lebesgue-Maß auf  $\Omega \subset \mathbb{R}^3$  sei.

Wir nutzen nun aus, dass  $L^2(\Omega) \subset L^1(\Omega)$  für Kompakta  $\Omega \subset \mathbb{R}^3$  gilt ([6], Kapitel VI, Satz 2.10) und daher durch

$$T : L^2(\Omega) \rightarrow \mathcal{M}(\Omega), \quad v \mapsto \left( u : \mathcal{B}(\Omega) \rightarrow \mathbb{R}, \quad u(A) = \int_A v(x) \, d\lambda(x) \right)$$

eine Einbettung von  $L^2(\Omega)$  in  $\mathcal{M}(\Omega)$  definiert wird.

Für verrauschte Daten  $f \in L^2(\Omega)$  können wir deshalb durch

$$G_{\mathcal{M}} : \mathcal{M}(\Omega) \rightarrow [0, \infty], \quad G_{\mathcal{M}}(u) = \frac{1}{2} \int_{\Omega} \left| \frac{du}{d\lambda} - f \right|^2 \, d\lambda(x) + \begin{cases} 0 & , \quad u \in T(L^2(\Omega)) \\ \infty & , \quad \text{sonst} \end{cases}$$

einen Datenterm definieren. Hierbei sei  $\frac{du}{d\lambda}$  die Radon-Nikodym Dichte von  $u$  bezüglich  $\lambda$ . Wir weisen nun nach, dass  $G_{\mathcal{M}}$  die Eigenschaften aus Theorem 11 erfüllt.

Wir beginnen mit der schwach-\* Folgenunterhalbstetigkeit von  $G_{\mathcal{M}}$ . Gegeben sei dazu eine schwach-\* konvergente Folge  $(u_n)_n$ , sodass  $u_n \xrightarrow{*} u$  in  $\mathcal{M}(\Omega)$  gilt. Wir müssen nun zeigen, dass  $G_{\mathcal{M}}(u) \leq \liminf_{n \rightarrow \infty} G_{\mathcal{M}}(u_n)$  gilt.

Falls  $\liminf_{n \rightarrow \infty} G_{\mathcal{M}}(u_n) = \infty$  gilt, so ist die Aussage klar. Es sei daher

$$\liminf_{n \rightarrow \infty} G_{\mathcal{M}}(u_n) \leq c$$

für ein  $c \in \mathbb{R}$ . Es existiert dann eine Teilfolge, die wir ebenfalls mit  $(u_n)_n$  bezeichnen, für die  $G_{\mathcal{M}}(u_n) \leq 2c$  für alle  $n \in \mathbb{N}$  gilt. Die Maße  $u_n$  sind also in  $T(L^2(\Omega))$ . Deshalb existieren Dichtefunktionen  $v_n \in L^2(\Omega)$ , so dass

$$u_n : \mathcal{B}(\Omega) \rightarrow [0, \infty), u_n(A) = \int_A v_n(x) d\lambda(x)$$

gilt. Weil  $\frac{du_n}{d\lambda} = v_n$   $\lambda$ -fast überall gilt ([6], Kapitel VII, Satz 2.3), folgt weiterhin

$$G_{\mathcal{M}}(u_n) = \frac{1}{2} \int_{\Omega} \left| \frac{du_n}{d\lambda} - f \right|^2 d\lambda = \frac{1}{2} \int_{\Omega} |v_n - f|^2 d\lambda = G(v_n)$$

und insbesondere  $\|v_n\|_{L^2(\Omega)} < \infty$ . Es existiert daher eine in  $L^2(\Omega)$  schwach konvergente Teilfolge, die wir ebenfalls mit  $(v_n)_n$  bezeichnen, so dass  $v_n \rightharpoonup v$  für ein  $v \in L^2(\Omega)$  gilt. Da für alle  $n \in \mathbb{N}$  die Gleichung  $\frac{du_n}{d\lambda} = v_n$   $\lambda$ -fast überall gilt, gilt weiterhin

$$\int_{\Omega} \varphi du_n = \int_{\Omega} \varphi \frac{du_n}{d\lambda} d\lambda = \int_{\Omega} \varphi v_n d\lambda \quad (4.3)$$

für alle  $\varphi \in \mathcal{C}^0(\Omega) \subset L^2(\Omega)$  und alle  $n \in \mathbb{N}$ . Da zudem wegen  $\mathcal{C}^0(\Omega) \subset L^2(\Omega)$  auch

$$\lim_{n \rightarrow \infty} \int_{\Omega} \varphi v_n d\lambda = \int_{\Omega} \varphi v d\lambda$$

für alle  $\varphi \in \mathcal{C}^0(\Omega)$  gilt, muss durch Übergang zum Grenzwert in (4.3) auch

$$\int_{\Omega} \varphi du = \lim_{n \rightarrow \infty} \int_{\Omega} \varphi du_n = \lim_{n \rightarrow \infty} \int_{\Omega} \varphi v_n d\lambda = \int_{\Omega} \varphi v d\lambda$$

für alle  $\varphi \in \mathcal{C}^0(\Omega)$  gelten. Die Gleichung  $\frac{du}{d\lambda} = v$  gilt daher ebenfalls  $\lambda$ -fast überall. Deshalb erhalten wir analog zu oben  $G_{\mathcal{M}}(u) = G(v)$ .

Da  $G$  (strikt) konvex und stetig ist, folgt, dass  $G$  schwach folgenunterhalbstetig in  $L^2(\Omega)$  ist ([16], Lemma III.5.9). Dies bedeutet, dass für die Folge  $(v_n)_n$  die Ungleichung

$$\liminf_{n \rightarrow \infty} G(v_n) \geq G(v)$$

gilt. Da  $G(v_n) = G_{\mathcal{M}}(u_n)$  für alle  $n \in \mathbb{N}$  und  $G(v) = G_{\mathcal{M}}(u)$  gelten, muss ebenfalls

$$\liminf_{n \rightarrow \infty} G_{\mathcal{M}}(u_n) \geq G_{\mathcal{M}}(u)$$

gelten. Damit ist gezeigt, dass das Funktional  $G_{\mathcal{M}}$  schwach-\* folgenunterhalbstetig in  $\mathcal{M}(\Omega)$  ist.

Es bleibt zu zeigen, dass

$$G_{\mathcal{M}}(u_n) \rightarrow \infty, \text{ falls } \int_{\Omega} \varphi du_n \rightarrow \infty \text{ für ein } \varphi \in \mathcal{C}_0(\Omega) \text{ mit } \|\varphi\|_{\infty} \leq 1$$

gilt. Wir zeigen dazu die Ungleichungskette:

$$\begin{aligned}
 G_{\mathcal{M}}(u_n) &= \int_{\Omega} \left| \frac{du_n}{d\lambda} - f \right|^2 d\lambda \\
 &\stackrel{1.}{\geq} \frac{1}{2} \int_{\Omega} \left| \frac{du_n}{d\lambda} \right|^2 d\lambda - \int_{\Omega} |f|^2 d\lambda \\
 &\stackrel{2.}{\geq} \frac{1}{2|\Omega|} \left( \int_{\Omega} \left| \frac{du_n}{d\lambda} \right| d\lambda \right)^2 - \int_{\Omega} |f|^2 d\lambda \\
 &\stackrel{3.}{\geq} \frac{1}{2|\Omega|} |u_n|(\Omega)^2 - \int_{\Omega} |f|^2 d\lambda \\
 &\stackrel{4.}{\geq} \frac{1}{2|\Omega|} (\|u_n\|_{\mathcal{M}} \|\varphi\|_{\infty})^2 - \int_{\Omega} |f|^2 d\lambda \\
 &\stackrel{5.}{\geq} \frac{1}{2|\Omega|} \left| \int_{\Omega} \varphi du_n \right|^2 - \int_{\Omega} |f|^2 d\lambda.
 \end{aligned}$$

Wir gehen dazu die Ungleichungen entsprechend ihrer Nummerierung durch: die Ungleichung

1. ist eine umgekehrte Jensensche Ungleichung, die beispielsweise in [11], Chapter I, Section 7 zu finden ist.
2. ist eine Anwendung der Jensenschen Ungleichung ([5], Chap. IV, Prop. 14.1).
3. ist eigentlich eine Gleichung, da für ein Maß  $u_n$  mit Radon-Nikodym Dichte  $\frac{du_n}{d\lambda}$  das Variationsmaß  $|u_n|$  die Radon-Nikodym Dichte  $|\frac{du_n}{d\lambda}|$  besitzt.
4. gilt, da per Definition der Variationsnorm  $\|u_n\|_{\mathcal{M}} = |u_n|(\Omega)$  und per Annahme  $\|\varphi\|_{\infty} \leq 1$  gilt.
5. ist beispielsweise in [6], Kapitel VIII, Prop. 2.22 zu finden.

Da  $f \in L^2(\Omega)$  ist, muss  $\int_{\Omega} |f|^2 d\lambda$  endlich sein. Es sei nun ein  $\varphi \in \mathcal{C}_0(\Omega)$  mit  $\|\varphi\|_{\infty} \leq 1$  gegeben, sodass  $\int_{\Omega} \varphi du_n \rightarrow \infty$  gilt. Dann muss wegen  $G_{\mathcal{M}}(u_n) \geq \frac{1}{2|\Omega|} \left| \int_{\Omega} \varphi du_n \right|^2 - \int_{\Omega} |f|^2 d\lambda$  auch  $G_{\mathcal{M}}(u_n) \rightarrow \infty$  gelten.

Damit sind die Voraussetzungen von Theorem 11 erfüllt und eine Lösung von

$$\min_{u \in \mathcal{M}(\Omega)} G_{\mathcal{M}}(u) + \bar{R}_{\rho}^{\alpha, \beta}(u)$$

existiert für Kompakta  $\Omega \subset \mathbb{R}^3$ , positive Parameter  $\alpha, \beta$  und unterhalbstetige Metriken  $\rho$  auf  $S^2$ . Da zudem  $G_{\mathcal{M}}$  strikt konvex ist, ist diese Lösung eindeutig.

#### 4.2.2 Existenz eines Minimums für einen Wasserstein Datenterm

Zum Abschluss dieses Kapitels führen wir einen weiteren Datenterm ein, für den wir die bereits in Definition 10 eingeführte Wasserstein-1 Metrik nutzen. Die Wasserstein-1 Metrik misst nur die Distanz zwischen Wahrscheinlichkeitsmaßen. Allerdings lassen sich die Definition und die im Folgenden benötigten Resultate auch auf Radonmaße  $\mu \in \mathcal{M}(\Omega)$  mit endlicher Masse  $\mu(\Omega) < \infty$  übertragen. Gegeben seien dazu verrauschte Daten  $f \in \mathcal{M}(\Omega)$  mit Masse  $f(\Omega) < \infty$ . Wir definieren zunächst die Menge

$$\mathcal{M}^+(\Omega) := \{u \in \mathcal{M}(\Omega) \mid u(A) = |u|(A) \forall A \in \mathcal{B}(\Omega)\}.$$

Mit Hilfe dieser Menge definieren wir dann durch

$$G_{\mathcal{W}} : \mathcal{M}^+(\Omega) \rightarrow \mathbb{R} \cup \{\infty\}, \quad G_{\mathcal{W}}(u) = \sup_{\text{Lip}(\varphi) \leq 1} \left\{ \int_{\Omega} \varphi \, du - \int_{\Omega} \varphi \, df \right\}$$

einen Wasserstein-Datenterm. Wir zeigen nun, dass dieser Datenterm die Voraussetzungen von Theorem 11 erfüllt und beginnen mit der schwach-\* Folgenunterhalbstetigkeit.

Gegeben sei eine Folge von Maßen  $(u_n)_n$  in  $\mathcal{M}^+(\Omega)$ , die schwach-\* gegen ein Maß  $u \in \mathcal{M}^+(\Omega)$  konvergiere.

Wir nehmen zunächst an, dass  $u_n(\Omega) \neq f(\Omega)$  gelte. Falls  $u_n(\Omega) > f(\Omega)$  ist, so erhalten wir für  $k \in \mathbb{N}$  und  $\varphi = k$ :

$$\int_{\Omega} \varphi \, d(u - f) = \int_{\Omega} k \, d(u - f) = k (\|u\|_{\mathcal{M}} - \|f\|_{\mathcal{M}}) \xrightarrow{k \rightarrow \infty} \infty.$$

Falls  $u_n(\Omega) < f(\Omega)$  ist, so wählen wir entsprechend  $\varphi = -k$  für  $k \in \mathbb{N}$ . Da die konstanten Funktionen 1-Lipschitz stetig sind, gilt somit auch  $G_{\mathcal{W}}(u_n) = \infty$  und die Abschätzung

$$G_{\mathcal{W}}(u) \leq \liminf_{n \rightarrow \infty} G_{\mathcal{W}}(u_n) = \infty.$$

Es sei daher eine schwach-\* konvergente Folge  $u_n \xrightarrow{*} u$  mit  $u_n(\Omega) = f(\Omega)$  gegeben. Die schwach-\* Folgenunterhalbstetigkeit ergibt sich dann aus folgendem Lemma:

**Lemma 1 ([15], Lemma 4.3)** *Gegeben seien ein polnischer Raum  $X$  und ein nichtnegatives, unterhalbstetiges Kostenfunktional  $d^X : X \times X \rightarrow \mathbb{R} \cup \{\infty\}$  auf  $X$ . Weiterhin sei  $(\pi_k)_k$  eine Folge von Maßen in  $\mathcal{M}(X \times X)$ , so dass  $\pi_k \rightarrow \pi$  für ein Maß  $\pi \in \mathcal{M}(X \times X)$  und  $\pi_k(X \times X) = \pi(X \times X)$  für alle  $k \in \mathbb{N}$  gilt. Dann gilt*

$$\int_{X \times X} d^X(x, y) \, d\pi(x, y) \leq \liminf_{k \rightarrow \infty} \int_{X \times X} d^X(x, y) \, d\pi_k(x, y).$$

*Beweis:* Das Funktional  $d^X$  ist punktweiser Grenzwert einer nichtfallenden Folge stetiger, reellwertiger Funktionen  $(d_i^X)_{i \in \mathbb{N}}$ . Deshalb gilt mit Hilfe des Satzes von Beppo-Levi die Abschätzung

$$\int_{X \times X} d^X \, d\pi = \lim_{l \rightarrow \infty} \int_{X \times X} d_l^X \, d\pi = \lim_{l \rightarrow \infty} \lim_{k \rightarrow \infty} \int_{X \times X} d_l^X \, d\pi_k \leq \liminf_{k \rightarrow \infty} \int_{X \times X} d^X \, d\pi_k. \quad \blacksquare$$

Die Menge  $\Omega$  ist als kompakte Teilmenge des  $\mathbb{R}^3$  polnisch. Zudem ist die Standardmetrik auf  $\mathbb{R}^3$  nichtnegativ und unterhalbstetig. Nutzen wir nun aus, dass die duale Formulierung (2.11) der Wasserstein-1 Metrik äquivalent zu (2.10) ist, so ergibt sich aus dem Lemma sofort die schwach-\* Folgenunterhalbstetigkeit von  $G_{\mathcal{W}}$ .

Wir müssen nun noch zeigen, dass

$$G_{\mathcal{W}}(u_n) \rightarrow \infty, \quad \text{falls} \quad \int_{\Omega} \varphi \, du^n \rightarrow \infty \quad \text{für ein} \quad \varphi \in \mathcal{C}_0(\Omega) \quad \text{mit} \quad \|\varphi\|_{\infty} \leq 1$$

gilt.

Falls  $\int_{\Omega} \varphi \, du^n \rightarrow \infty$  gilt, so auch  $\|\varphi\|_{\infty} \|u_n\|_{\mathcal{M}} \rightarrow \infty$ . Wegen  $\|u_n\|_{\mathcal{M}} = |u_n|(\Omega) = u_n(\Omega)$  gilt

dann auch  $u_n(\Omega) \rightarrow \infty$  und damit ergibt sich

$$\begin{aligned} G_{\mathcal{W}}(u_n) &= \sup_{\text{Lip}(\varphi) \leq 1} \int_{\Omega} \varphi \, du^n - \int_{\Omega} \varphi \, df \\ &\geq \int_{\Omega} 1 \, du_n - \|f\|_{\mathcal{M}} \\ &= u_n(\Omega) - \|f\|_{\mathcal{M}} \rightarrow \infty \end{aligned}$$

da  $\|f\|_{\mathcal{M}}$  nach Voraussetzung endlich ist.

Damit folgt mit Theorem 11 die Existenz von Lösungen des Minimierungsproblems (4.2) für den Datenterm  $G_{\mathcal{W}}$ . Diese Lösungen sind im Allgemeinen aber nicht eindeutig, da  $G_{\mathcal{W}}$  nicht strikt konvex sondern nur linear ist.

---



---

## KAPITEL 5

---

# Numerische Umsetzung

Nachdem wir nun zur Regularisierung der Krümmung dreidimensionaler Linienmaße ein Minimierungsproblem der Form

$$\min_{u \in \mathcal{M}(\Omega)} G(u) + \overline{R}_\rho^{\alpha, \beta}(u)$$

aufgestellt haben und die Existenz von Lösungen für spezielle Datenterme  $G_{\mathcal{M}}$  und  $G_{\mathcal{W}}$  gezeigt haben, widmen wir uns zum Abschluss dieser Arbeit der numerischen Umsetzung. Wir gehen dabei analog zu [2] vor und diskretisieren dazu in Abschnitt 5.1 das Problem mit finiten Differenzen. Dabei erhalten wir ein primal-duales Optimierungsproblem, welches wir in Abschnitt 5.2 mit Hilfe des Algorithmus von Chambolle und Pock lösen. Im abschließenden Abschnitt werden wir den Algorithmus auf das Entrauschungsproblem anwenden und einige Resultate zeigen.

### 5.1 Diskretisierung des Optimierungsproblems

Wir diskretisieren zunächst die kompakten Gebiete  $\Omega$  und  $\Omega \times S^2$ . Das Gebiet  $\Omega$  diskretisieren wir durch ein reguläres kartesisches Gitter der Größe  $I \times J \times K$  mit Gitterweite  $h > 0$ . Die Gitterpunkte sind dann gegeben durch

$$\{(ih, jh, kh) | 1 \leq i \leq I, 1 \leq j \leq J, 1 \leq k \leq K\}.$$

Da sich im Folgenden aus dem Kontext ergibt, ob die Menge  $\Omega$  oder ihre Diskretisierung gemeint ist, bezeichnen wir diese Menge ebenfalls mit  $\Omega$ . Wir benötigen noch eine Diskretisierung der  $S^2$ . Dazu wählen wir  $L$  Gitterpunkte  $g_l$ , die äquidistant auf der  $S^2$  verteilt seien. Für  $L = 4, 6, 8, 12, 20$  können wir das beispielsweise erreichen indem wir der  $S^2$  einen der fünf platonischen Körper so einbeschreiben, dass der Mittelpunkt des Körpers mit dem der  $S^2$  übereinstimmt. Die Ecken können wir dann auf die  $S^2$  projizieren und die projizierten Punkte  $g_l$ ,  $1 \leq l \leq L$ , auf der  $S^2$  dann als Gitterpunkte wählen. Die Gitterpunkte auf  $\Omega \times S^2$  sind dann gegeben durch die Menge

$$\{(ih, jh, kh, g_l) | 1 \leq i \leq I, 1 \leq j \leq J, 1 \leq k \leq K, 1 \leq l \leq L\},$$

die wir ebenfalls mit  $\Omega \times S^2$  bezeichnen.

Die Linienmaße  $u \in \mathcal{M}(\Omega)$  und  $\mu \in \mathcal{M}(\Omega \times S^2)$  sind im diskreten Fall durch Felder in  $\mathbb{R}^{I \times J \times K}$  beziehungsweise  $\mathbb{R}^{I \times J \times K \times L}$  gegeben. Zur Vereinfachung der Notation bezeichnen wir diese Felder ebenfalls mit  $u$  und  $\mu$ . Wie üblich speichern wir die Felder durch eine lineare Indizierung in Spaltenvektoren  $u \in \mathbb{R}^{IJK}$  und  $\mu \in \mathbb{R}^{IJKL}$ . Diese Indizierung erlaubt

es die linearen diskreten Operatoren des Minimierungsproblems durch Matrizen anzugeben. Wir führen nun zunächst diese Operatoren ein und werden im Anschluss mit Hilfe dieser Operatoren das Minimierungsproblem diskretisieren.

Wir beginnen mit der Diskretisierung des Operators  $T_\rho$ . Dazu bemerken wir, dass das Integral  $\int_{\Omega \times S^2} \nabla_x \psi(x, v) \cdot v \, d\mu(x, v)$  als distributionelle Richtungsableitung von  $\mu$  aufgefasst werden kann ([2], Remark 3.11). Wir benötigen daher zunächst eine Diskretisierung der Richtungsableitung. Wie in [2] nutzen wir dazu nichtlokale finite Differenzen, da die Punkte  $(ih, jh, kh)^T + g_l$  im Allgemeinen keine Gitterpunkte sind. Wir definieren daher für jedes  $g_l = (g_l^1, g_l^2, g_l^3)^T$  einen Reskalierungsfaktor

$$\delta_l = \begin{cases} 1 & , \text{ falls } g_l \in \left\{ \pm (1, 0, 0)^T, \pm (0, 1, 0)^T, \pm (0, 0, 1)^T \right\} \\ \min \left\{ |g_l^b|, |g_l^c| \right\} & , g_l^a = 0 \text{ für genau ein } a \in \{1, 2, 3\}, b \neq a \neq c \\ \min \left\{ |g_l^1|, |g_l^2|, |g_l^3| \right\} & , \text{ sonst} \end{cases}$$

und die reskalierten Richtungen  $v_l = \frac{g_l}{\delta_l}$ . Der Vektor  $(ih, jh, kh)^T + v_l$  ist dann in mindestens einem Eintrag ganzzahlig. Den Wert  $\mu_{i+v_l^1, j+v_l^2, k+v_l^3, g_l}$  von  $\mu$  im Punkt  $(ih, jh, kh)^T + v_l$  berechnen wir dann durch eine bilineare Interpolation mit Nullrandwerten in den nicht ganzzahligen Einträgen. Den diskreten Richtungsableitungsoperator  $B \in \mathbb{R}^{IJKL \times IJKL}$  definieren wir dann durch

$$(B\mu)_{i,j,k,g_l} = \frac{\mu_{i+v_l^1, j+v_l^2, k+v_l^3, g_l} - \mu_{i,j,k,g_l}}{|v_l|}.$$

Wir benötigen nun noch eine Diskretisierung der Mengen

$$M_{\rho_0} = \left\{ \psi \in \mathcal{C}_0(\Omega \times S^2) \mid \nabla_x \psi \in \mathcal{C}_0(\Omega \times S^2, \mathbb{R}^3), |\psi(x, \cdot)| \leq \frac{1}{2} \forall x \in \Omega \right\}$$

und

$$M_{\rho_1} = \left\{ \psi \in \mathcal{C}_0(\Omega \times S^2) \mid \nabla_x \psi \in \mathcal{C}_0(\Omega \times S^2, \mathbb{R}^3), \text{Lip}(\psi(x, \cdot)) \leq 1 \forall x \in \Omega \right\}.$$

Analog zu [2] erhalten wir durch Faltung der Funktionen mit Glättungskernen und mit Hilfe glatter Abschneidefunktionen (siehe [5], Chap. V, Prop 21.1) dichte Teilmengen

$$\tilde{M}_{\rho_0} = \left\{ (x, v) \mapsto \psi(x, v) + \varphi(x) \mid \psi \in \mathcal{C}_c^\infty(\Omega \times S^2), \varphi \in \mathcal{C}_c^\infty(\Omega), \|\psi\|_\infty \leq \frac{1}{2} \right\}$$

und

$$\tilde{M}_{\rho_1} = \left\{ (x, v) \mapsto \psi(x, v) + \varphi(x) \mid \psi \in \mathcal{C}_c^\infty(\Omega \times S^2), \varphi \in \mathcal{C}_c^\infty(\Omega), \|\partial_v \psi\|_\infty \leq 1, \int_{S^2} \psi(x, v) \, dv = 0 \right\}$$

über die wir stattdessen optimieren können. Die Bedingung  $\int_{S^2} \psi(x, v) \, dv = 0$  verändert die Menge  $\tilde{M}_{\rho_1}$  nicht, da für  $\int_{S^2} \psi(x, v) \, dv = a \neq 0$  auch  $(\psi + a) \in \mathcal{C}_c^\infty(\Omega)$  ist. Im Algorithmus verkleinert diese Bedingung allerdings die Menge der Testfunktionen  $\psi$ , was zu höherer numerischer Stabilität führen kann.

Wir möchten diese Mengen nun diskretisieren. Da im Folgenden aus dem Kontext klar sein wird, ob die Menge  $M_{\rho_i}$ ,  $i = 0, 1$  oder deren Diskretisierung genutzt werden, bezeichnen

wir die diskretisierten Mengen ebenfalls mit  $M_{\rho_i}$ ,  $i = 0, 1$ . Die Diskretisierung  $M_{\rho_0}$  ist dann gegeben durch

$$M_{\rho_0} = \left\{ \psi \in \mathbb{R}^{IJKL} \mid |\psi_{i,j,k,g_l}| \leq \frac{1}{2} \forall i, j, k, l \right\}.$$

Um die Menge  $M_{\rho_1}$  zu diskretisieren müssen wir zunächst die Lipschitzbedingung  $\|\partial_v \psi\|_\infty \leq 1$  diskretisieren. Dazu nutzen wir, dass die  $S^2$  äquidistant diskretisiert wurde. Daher existieren für jeden Gitterpunkt  $g_l \in S^2$ ,  $1 \leq l \leq L$ , ein  $M \in \mathbb{N}$  und benachbarte Gitterpunkte  $g_l^m$ ,  $1 \leq m \leq M$  die alle minimalen Abstand  $g > 0$  zu  $g_l$  auf der  $S^2$  haben. Eine Diskretisierung des Operators  $\|\partial_v \psi\|_\infty \leq 1$  erhalten wir mit Hilfe einer Matrix  $C \in \mathbb{R}^{IJKLM \times IJKL}$ , die durch

$$(C\psi)_{i,j,k,g_l^m} = \frac{1}{g} \left( \psi_{i,j,k,g_l^m} - \psi_{i,j,k,g_l} \right)$$

definiert ist. Die Lipschitzbedingung kann dann durch  $|(C\psi)_{i,j,k,g_l^m}| \leq 1$  diskretisiert werden. Daher kann bei der Implementierung von  $C$  auf die Hälfte der Bedingungen verzichtet werden. Eine Diskretisierung der Menge  $M_{\rho_1}$  erhalten wir dann mit Hilfe von  $C$  durch

$$M_{\rho_1} = \left\{ \psi \in \mathbb{R}^{IJKL} \mid |(C\psi)_{i,j,k,g_l^m}| \leq 1 \forall i, j, k, l, m \text{ und } \sum_l \psi_{i,j,k,g_l} = 0 \forall i, j, k \right\}.$$

Wir müssen nun noch die Menge  $M$  aus Proposition 4 diskretisieren. Wir beginnen mit der Diskretisierung von  $\int_{\Omega \times S^2} \nabla_x \varphi(x) \cdot v \, d\mu(x, v) = 0$  und definieren dazu einen Operator  $D \in \mathbb{R}^{IJLK \times IJK}$  durch

$$(D\varphi)_{i,j,k,l} = \frac{1}{h} \begin{pmatrix} \varphi_{i+1,j,k} - \varphi_{i,j,k} \\ \varphi_{i,j+1,k} - \varphi_{i,j,k} \\ \varphi_{i,j,k+1} - \varphi_{i,j,k} \end{pmatrix} \cdot g_l.$$

Die Diskretisierung der Projektionsbedingung  $\int_\Omega \varphi \, du = \int_{\Omega \times S^2} \varphi(x) \, d\mu$  ist aufwendiger. Wir diskretisieren dazu  $u$  auf den Halbgitterpunkten und bezeichnen daher im Folgenden mit  $\tilde{\Omega}$  die Menge der Gitterpunkte und mit

$$\tilde{\Omega} = \left\{ \begin{pmatrix} bh \\ ch \\ dh \end{pmatrix} \mid \begin{pmatrix} bh \\ ch \\ dh \end{pmatrix} = \begin{pmatrix} ih \\ jh \\ kh \end{pmatrix} + \begin{pmatrix} \frac{1}{2}h \\ \frac{1}{2}h \\ \frac{1}{2}h \end{pmatrix} \text{ für ein } \begin{pmatrix} ih \\ jh \\ kh \end{pmatrix} \in \Omega \right\}$$

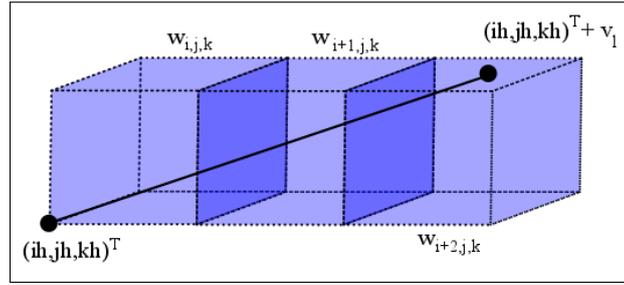
die Menge der Halbgitterpunkte. Für einen Halbgitterpunkt  $a = (bh, ch, dh)^T \in \tilde{\Omega}$  bezeichnen wir mit

$$m_{b,c,d} = \left\{ a \pm \begin{pmatrix} \frac{1}{2}h \\ \frac{1}{2}h \\ \frac{1}{2}h \end{pmatrix}, a \pm \begin{pmatrix} -\frac{1}{2}h \\ \frac{1}{2}h \\ \frac{1}{2}h \end{pmatrix}, a \pm \begin{pmatrix} \frac{1}{2}h \\ -\frac{1}{2}h \\ \frac{1}{2}h \end{pmatrix}, a \pm \begin{pmatrix} \frac{1}{2}h \\ \frac{1}{2}h \\ -\frac{1}{2}h \end{pmatrix} \mid a = \begin{pmatrix} bh \\ ch \\ dh \end{pmatrix} \right\}$$

die Menge der benachbarten Gitterpunkte und mit  $w_{b,c,d}$  den abgeschlossenen Würfel, dessen Mittelpunkt der Halbgitterpunkt  $a = (bh, ch, dh)^T$  und dessen Eckpunkte die Gitterpunkte  $e \in m_{b,c,d}$  sind.

Für einen Punkt  $(ih, jh, kh)^T \in \Omega \times S^2$  sei

$$s^{i,j,k,l} = \left\{ t(ih, jh, kh)^T + (1-t) \left( (ih, jh, kh)^T + v_l \right) \mid t \in [0, 1] \right\}$$


 Abbildung 5.1: Projektion von  $\mu$  auf  $u$ :

Die Ecken der Würfel seien Gitterpunkte in  $\Omega$  und die Mittelpunkte der Würfel Halbgridpunkte in  $\tilde{\Omega}$ . Dann projizieren wir eine Masse  $\mu_{i,j,k,l}$  auf  $u$  indem wir die Masse von  $\mu_{i,j,k,l}$  auf die Würfel, die von der Strecke  $s^{i,j,k,l}$  geschnitten werden, verteilen. Auf jeden Halbgridpunkt  $u_{i,j,k}$  wird dann genau der Anteil von  $\mu_{i,j,k,l}$  projiziert, der dem Anteil der Strecke  $s^{i,j,k,l}$  im zu  $u_{i,j,k}$  gehörigen Würfel  $w_{i,j,k}$  entspricht.

gegeben. Dann definieren wir für jeden Halbgridpunkt  $a = (bh, ch, dh)^T$  die Menge

$$n_{b,c,d} = \left\{ (i, j, k, l) \mid s^{i,j,k,l} \cap w_{b,c,d} \neq \emptyset \text{ für } (ih, jh, kh, gl)^T \in \Omega \times S^2 \right\}.$$

Weiterhin seien für alle  $(bh, ch, dh)^T \in \tilde{\Omega}$  und alle  $(ih, jh, kh, gl) \in \Omega \times S^2$  Faktoren

$$\lambda_{b,c,d}^{i,j,k,l} = \frac{\mathcal{H}^1(s^{i,j,k,l} \cap w_{b,c,d})}{\mathcal{H}^1(s^{i,j,k,l})}$$

gegeben. Mit Hilfe dieser Faktoren bestimmen wir den Anteil von  $\mu_{i,j,k,l}$ , der auf  $u_{b,c,d}$  projiziert wird.

Wir müssen allerdings noch Effekte berücksichtigen, die auftreten, falls die Strecken  $s^{i,j,k,l}$  Teilmengen der Seitenflächen oder Kanten der Würfel sind. Da die Strecke dann in mehreren benachbarten Würfeln mit der gleichen Seitenfläche oder Kante liegt, müssen wir sicherstellen, dass die Masse von  $\mu_{i,j,k,l}$  nicht vollständig auf alle Würfel projiziert wird, sondern auf diese aufgeteilt wird. Dies geschehe durch einen weiteren Faktor  $c^{i,j,k,l}$ . Die Projektionsbedingung können wir dann für eine Matrix  $P \in \mathbb{R}^{IJK \times IJKL}$  angeben durch

$$(P\mu)_{b,c,d} = (Iu)_{b,c,d} = \sum_{(i,j,k,l) \in n_{b,c,d}} c^{i,j,k,l} \lambda_{b,c,d}^{i,j,k,l} \mu_{i,j,k,l},$$

wobei  $I \in \mathbb{R}^{IJK \times IJK}$  die Einheitsmatrix sei. Für eine gegebene Anzahl an Richtungen  $L$  können die Faktoren  $c^{i,j,k,l}$  und  $\lambda_{b,c,d}^{i,j,k,l}$  sowie die Mengen  $n_{b,c,d}$  im Voraus rechnerisch bestimmt werden, so dass eine Implementation des Operators  $P$  einfach möglich ist.

Zur Diskretisierung des Wasserstein-1 Datenterms  $G_{\mathcal{W}}$  aus Abschnitt 4.2.2 benötigen wir noch einen weiteren Operator für die Lipschitz-Bedingung in der dualen Variable. Wir führen dazu den Operator  $W \in \mathbb{R}^{3IJK \times IJL}$  definiert durch

$$(W\varphi)_{i,j,k} = \frac{1}{h} \begin{pmatrix} \varphi_{i+1,j,k} - \varphi_{i,j,k} \\ \varphi_{i,j+1,k} - \varphi_{i,j,k} \\ \varphi_{i,j,k+1} - \varphi_{i,j,k} \end{pmatrix} \quad (5.1)$$

ein, wobei wir am Rand Nullrandwerte nutzen.

Wir möchten nun analog zu [2] die Strafterme  $\bar{R}_\rho^{\alpha,\beta}$  der Optimierungsprobleme

$$\begin{aligned}\text{TVX}_0^{\alpha,\beta}(u) &= G(u) + \bar{R}_{\rho_0}^{\alpha,\beta}(u) \text{ und} \\ \text{TVX}_1^{\alpha,\beta}(u) &= G(u) + \bar{R}_{\rho_1}^{\alpha,\beta}(u)\end{aligned}$$

diskretisieren.

Wir beginnen mit  $\text{TVX}_0$  und diskretisieren zunächst das Funktional  $T_{\rho_0}$  durch

$$T_{\rho_0}(u) = \max_{\psi \in M_{\rho_0}} \langle B\mu, \psi \rangle$$

und anschließend den Strafterm durch

$$\bar{R}_{\rho_0}^{\alpha,\beta}(u) = \min_{(u,\mu) \in M} \max_{\psi \in M_{\rho_0}} \alpha \sum_{i,j,k,l} \mu_{i,j,k,l} + \beta \langle B\mu, \psi \rangle.$$

Führen wir nun Lagrange-Multiplikatoren  $\varphi \in \mathbb{R}^{IJK}$  und  $\xi \in \mathbb{R}^{IJK}$  für die Nebenbedingungen in  $M$  ein, so erhalten wir als Diskretisierung des Strafterms in  $\text{TVX}_0^{\alpha,\beta}$

$$\begin{aligned}\bar{R}_{\rho_0}^{\alpha,\beta}(u) &= \min_{u,\mu} \max_{\psi,\phi,\xi} \alpha \sum_{i,j,k,l} \mu_{i,j,k,l} + \beta \langle B\mu, \psi \rangle + \langle \mu, D\varphi \rangle + \langle P\mu - Iu, \xi \rangle \\ &\text{so dass } \mu_{i,j,k,l} \geq 0, |\psi_{i,j,k,l}| \leq \frac{1}{2}.\end{aligned}$$

Um  $\text{TVX}_1^{\alpha,\beta}$  diskretisieren zu können benötigen wir zunächst eine Diskretisierung von  $\bar{R}_{\rho_1}^{\alpha,\beta}(u)$ . Dazu brauchen wir einen weiteren Lagrange-Multiplikator  $\nu \in \mathbb{R}^{IJKLM}$  und eine weitere duale Variable  $\chi \in \mathbb{R}^{IJKLM}$ . Wir erhalten analog zu  $\text{TVX}_0^{\alpha,\beta}$  eine Diskretisierung des Strafterms von  $\text{TVX}_1^{\alpha,\beta}$  durch

$$\begin{aligned}\bar{R}_{\rho_1}^{\alpha,\beta}(u) &= \min_{u,\mu} \max_{\psi,\phi,\xi} \alpha \sum_{i,j,k,l} \mu_{i,j,k,l} + \beta \langle B\mu, \psi \rangle \\ &\quad + \langle \mu, D\varphi \rangle + \langle P\mu - Iu, \xi \rangle + \langle \nu, C\psi - I\chi \rangle \\ &\text{so dass } \mu_{i,j,k,l} \geq 0, \sum_l \psi_{i,j,k,l} = 0, |\chi_{i,j,k,l,m}| \leq 1.\end{aligned}$$

## 5.2 Primal-dualer Algorithmus für das diskrete Optimierungsproblem

In Abschnitt 5.1 haben wir die Strafterme der Optimierungsprobleme  $\text{TVX}_0^{\alpha,\beta}$  und  $\text{TVX}_1^{\alpha,\beta}$  diskretisiert und dabei Sattelpunktprobleme erhalten. Zur Lösung primal-dualer Probleme, die spezielle Sattelpunktprobleme sind, haben wir bereits in Abschnitt 2.1 Algorithmen eingeführt. Wir formen daher zunächst in den beiden folgenden Unterkapiteln die Sattelpunktprobleme  $\text{TVX}_0^{\alpha,\beta}$  und  $\text{TVX}_1^{\alpha,\beta}$  zu primal-dualen Problemen

$$\min_{x \in X} \max_{y \in Y} \{ \langle \mathcal{K}x, y \rangle + g(x) - f^*(y) \} \quad (5.2)$$

für einen stetigen linearen Operator  $\mathcal{K} : X \rightarrow Y$  und unterhalbstetige, zulässig konvexe und einfache Funktionale  $g : X \rightarrow [0, \infty]$  und  $f : Y \rightarrow [0, \infty]$  um.

### 5.2.1 Primal-Dualer Algorithmus für $\text{TVX}_0^{\alpha,\beta}$

Wir beginnen mit der Umformulierung von  $\text{TVX}_0^{\alpha,\beta}$ . Dazu seien zunächst

$$x = (\mu, u)^T \in \mathbb{R}^{IJKL+IJK} \text{ und } y = (\psi, \varphi, \xi)^T \in \mathbb{R}^{IJKL+2IJK}$$

sowie der Operator

$$\mathcal{K} = \begin{pmatrix} \beta B & 0 \\ D^T & 0 \\ P & -I \end{pmatrix} \in \mathbb{R}^{(IJKL+2IJK) \times (IJKL+IJK)}$$

gegeben. Weiterhin seien die Funktionale

$$g(x) = \alpha \sum_{i,j,k,l} \mu_{i,j,k,l} + \iota_{[0,\infty)}^{IJKL}(\mu) + G(u) \text{ sowie } f^*(y) = \iota_{[-\frac{1}{2}, \frac{1}{2}]}^{IJKL}(\psi)$$

gegeben wobei  $\iota_S$  die Indikatorfunktion der Menge  $S$  sei. Diese Funktionale sind unterhalbstetig und zulässig konvex falls es der Datenterm  $G(u)$  ist. Da wir im Folgenden nur die unterhalbstetigen und konvexen Datenterme  $G_{\mathcal{M}}$  und  $G_{\mathcal{W}}$  aus Abschnitt 4.2 nutzen werden, bleibt nur noch zu zeigen, dass die Funktionale  $g$  und  $f$  einfach im Sinne der Definition 7 sind. Die zugehörigen Proximaloperatoren (2.7) und (2.8) müssen also eine Darstellung in geschlossener Form besitzen. Wir bezeichnen mit  $\mathcal{T}^\mu \in \mathbb{R}^{IJKL \times IJKL}$  die zu  $\mu$  gehörige Untermatrix und mit  $\mathcal{T}_{i,j,k,l}^\mu$  den zum Eintrag  $\mu_{i,j,k,l}$  gehörigen Diagonaleintrag von  $\mathcal{T}$ . Für die Proximaloperatoren ergibt sich dann analog zu [2] mit Hilfe der Optimalitätsbedingung erster Ordnung

$$\begin{aligned} x = \text{prox}_{\mathcal{T},g}(\hat{x}) &\Leftrightarrow \mu_{i,j,k,l} = \max \left\{ 0, \hat{\mu}_{i,j,k,l} - \alpha \mathcal{T}_{i,j,k,l}^\mu \right\} \\ u &= \text{prox}_{\mathcal{T}^u,g}(\hat{u}) \end{aligned}$$

und

$$\begin{aligned} y = \text{prox}_{\mathcal{S},f^*}(\hat{y}) &\Leftrightarrow \psi_{i,j,k,l} = \max \left\{ -\frac{1}{2}, \min \left\{ \frac{1}{2}, \hat{\psi}_{i,j,k,l} \right\} \right\} \\ \varphi_{i,j,k} &= \hat{\varphi}_{i,j,k} \\ \xi_{i,j,k} &= \hat{\xi}_{i,j,k}, \end{aligned}$$

wobei  $\text{prox}_{\mathcal{T}^u,g}(\hat{u})$  abhängig vom Datenterm zu bestimmen ist.

Den  $L^2$ -Datenterm  $\lambda G_{\mathcal{M}}$  mit einem Parameter  $\lambda > 0$  diskretisieren wir wie in [2] mit Hilfe der euklidischen Norm durch  $\lambda \|u - f\|_2^2$  und erhalten

$$u = \text{prox}_{\mathcal{T}^u,g}(\hat{u}) \Leftrightarrow u_{i,j,k} = \frac{\hat{u}_{i,j,k} + \lambda \mathcal{T}_{i,j,k}^u f_{i,j,k}}{1 + \lambda \mathcal{T}_{i,j,k}^u}.$$

Durch Einsetzen in den Algorithmus 3 erhalten wir dann einen primal-dualen Algorithmus zur Lösung des Optimierungsproblems  $\text{TVX}_0^{\alpha,\beta}$  mit einem  $L^2$ -Datenterm.

Als zweiten Datenterm möchten wir noch den Wasserstein-1 Datenterm  $G_{\mathcal{W}}$  aus Abschnitt 4.2.2 nutzen. Zur Diskretisierung führen wir zunächst weitere Variablen  $\tau \in \mathbb{R}^{IJK}$  und

$\kappa \in \mathbb{R}^{3IJK}$  sowie einen Lagrange Multiplikator  $\omega \in \mathbb{R}^{3IJK}$  ein. Eine Diskretisierung des Wasserstein-1 Datenterms  $\lambda G_{\mathcal{W}}$  für einen Parameter  $\lambda \geq 0$  ist dann gegeben durch

$$\lambda \max_{\tau} \langle \tau, u \rangle - \langle \tau, f \rangle + \langle \omega, W\tau - I\kappa \rangle \text{ so dass } |\kappa| \leq 1,$$

wobei  $W$  der in (5.1) eingeführte Operator ist. Die Skalarprodukte  $\langle \tau, u \rangle$  und  $\langle \omega, W\tau - I\kappa \rangle$  können wir in den Operator  $\mathcal{K}$  und  $-\langle \tau, f \rangle$  in das Funktional  $f^*$  einfügen. Für Variablen

$$x = (\mu, u, \omega)^T \in \mathbb{R}^{IJKL+4IJK} \text{ und } y = (\psi, \varphi, \xi, \tau, \kappa) \in \mathbb{R}^{IJKL+6IJK}$$

erhalten wir daher nun einen Operator

$$\mathcal{K} = \begin{pmatrix} \beta B & 0 & 0 \\ D^T & 0 & 0 \\ P & -I & 0 \\ 0 & \lambda I & W^T \\ 0 & 0 & -I \end{pmatrix} \in \mathbb{R}^{(IJKL+6IJK) \times (IJKL+4IJK)}$$

und Funktionale

$$g(x) = \alpha \sum_{i,j,k,l} \mu_{i,j,k,l} + \iota_{[0,\infty)}^{IJKL}(\mu)$$

sowie

$$f^*(y) = \iota_{[-\frac{1}{2}, \frac{1}{2}]}^{IJKL}(\psi) - \langle \tau, f \rangle + \iota_{[-1,1]}^{3IJK}(\kappa).$$

Die Proximaloperatoren erhalten wir wieder mit Hilfe der Optimalitätsbedingung erster Ordnung:

$$\begin{aligned} x = \text{prox}_{\mathcal{T},g}(\hat{x}) &\Leftrightarrow \mu_{i,j,k,l} = \max \left\{ 0, \hat{\mu}_{i,j,k,l} - \alpha \mathcal{T}_{i,j,k,l}^{\mu} \right\} \\ u_{i,j,k} &= \hat{u}_{i,j,k} \\ \omega_{i,j,k}^a &= \hat{\omega}_{i,j,k}^a, \quad a = 1, 2, 3 \end{aligned}$$

und

$$\begin{aligned} y = \text{prox}_{\mathcal{S},f^*}(\hat{y}) &\Leftrightarrow \psi_{i,j,k,l} = \max \left\{ -\frac{1}{2}, \min \left\{ \frac{1}{2}, \hat{\psi}_{i,j,k,l} \right\} \right\} \\ \varphi_{i,j,k} &= \hat{\varphi}_{i,j,k} \\ \xi_{i,j,k} &= \hat{\xi}_{i,j,k} \\ \tau_{i,j,k} &= -\lambda \mathcal{S}_{i,j,k}^{\tau} f_{i,j,k} + \hat{\tau}_{i,j,k} \\ \kappa_{i,j,k}^a &= \max \left\{ -1, \min \left\{ 1, \hat{\kappa}_{i,j,k}^a \right\} \right\}, \quad a = 1, 2, 3. \end{aligned}$$

Einen primal-dualen Algorithmus zur Lösung des Optimierungsproblems  $\text{TVX}_0^{\alpha,\beta}$  mit einem Wasserstein-1 Datenterm erhalten wir nun durch Einsetzen in den Algorithmus 3.

### 5.2.2 Primal-Dualer Algorithmus für $\text{TVX}_1^{\alpha,\beta}$

Nachdem wir das Optimierungsproblem  $\text{TVX}_0^{\alpha,\beta}$  in die Form eines primal-dualen Algorithmus gebracht haben, möchten wir dasselbe nun auch für  $\text{TVX}_\rho^{\alpha,\beta}$  erreichen. Dazu seien die Variablen

$$x = (\mu, u, \nu)^T \text{ und } y = (\psi, \varphi, \xi, \chi)^T,$$

der Operator

$$\mathcal{K} = \begin{pmatrix} \beta B & 0 & C^T \\ D^T & 0 & 0 \\ P & -I & 0 \\ 0 & 0 & -I \end{pmatrix} \in \mathbb{R}^{(IJKL+IJK+IJKLM) \times (IJKL+2IJK+IJKLM)}$$

sowie die Funktionale

$$g(x) = \iota_{[0,\infty)}^{IJKL}(\mu) + \alpha \sum_{i,j,k,l} \mu_{i,j,k,l} + G(u)$$

und

$$f^*(y) = \iota_{[-1,1]}^{IJKLM}(\chi) + \iota_{\{\sum_l \psi_{i,j,k,l} = 0 \forall i,j,k\}}(\psi)$$

gegeben. Für den  $L^2$  Datenterm  $G_{\mathcal{M}}$  ergeben sich die Proximaloperatoren analog zu den vorherigen Fällen:

$$\begin{aligned} x = \text{prox}_{\mathcal{T},g}(\hat{x}) &\Leftrightarrow \mu_{i,j,k,l} = \max \left\{ 0, \hat{\mu}_{i,j,k,l} - \alpha \mathcal{T}_{i,j,k,l}^\mu \right\} \\ u_{i,j,k} &= \frac{\hat{u}_{i,j,k} + \lambda \mathcal{T}_{i,j,k}^u f_{i,j,k}}{1 + \lambda \mathcal{T}_{i,j,k}^u} \\ \nu_{i,j,k,l,m} &= \hat{\nu}_{i,j,k,l,m}, \end{aligned}$$

und

$$\begin{aligned} y = \text{prox}_{\mathcal{S},f^*}(\hat{y}) &\Leftrightarrow \psi_{i,j,k,l} = \hat{\psi}_{i,j,k,l} - \frac{\sum_l \psi_{i,j,k,l}}{L} \\ \varphi_{i,j,k} &= \hat{\varphi}_{i,j,k} \\ \xi_{i,j,k} &= \hat{\xi}_{i,j,k} \\ \chi_{i,j,k,l,m} &= \max \{-1, \min \{1, \hat{\chi}_{i,j,k,l,m}\}\}. \end{aligned}$$

Zum Abschluss dieses Kapitels müssen wir nun noch das Optimierungsproblem  $\text{TVX}_1^{\alpha,\beta}$  für den Wasserstein-1 Datenterm  $G_{\mathcal{W}}$  zu einem primal-dualen Algorithmus umformulieren. Wir benötigen dazu wieder die bereits eingeführten Variablen  $\tau, \kappa$  und  $\omega$  und erhalten als primale Variable  $x = (\mu, u, \nu, \omega)^T$  und als duale Variable  $y = (\psi, \varphi, \xi, \chi, \tau, \kappa)^T$ . Der Operator  $\mathcal{K}$  ist gegeben durch

$$\mathcal{K} = \begin{pmatrix} \beta B & 0 & C^T & 0 \\ D^T & 0 & 0 & 0 \\ P & -I & 0 & 0 \\ 0 & 0 & -I & 0 \\ 0 & \lambda I & 0 & W^T \\ 0 & 0 & 0 & -I \end{pmatrix} \in \mathbb{R}^{(IJKL+6IJK+IJKLM) \times (IJKL+4IJK+IJKLM)}$$

und die Funktionale sind gegeben durch

$$g(x) = \alpha \sum_{i,j,k,l} \mu_{i,j,k,l} + \iota_{[0,\infty)}^{IJKL}(\mu)$$

und

$$f^*(y) = \iota_{\{\sum_l \psi_{i,j,k,l}=0 \forall i,j,k\}}(\psi) + \iota_{[-1,1]}^{IJKLM}(\chi) - \lambda \tau^T f + \iota_{[-1,1]}^{3IJK}(\kappa).$$

Als Proximaloperatoren erhalten wir daher nun wieder:

$$\begin{aligned} x = \text{prox}_{\mathcal{T},g}(\hat{x}) &\Leftrightarrow \mu_{i,j,k,l} = \max \left\{ 0, \hat{\mu}_{i,j,k,l} - \alpha \mathcal{T}_{i,j,k,l}^\mu \right\} \\ u_{i,j,k} &= \hat{u}_{i,j,k} \\ \nu_{i,j,k,l,m} &= \hat{\nu}_{i,j,k,l,m}, \\ \omega_{i,j,k}^a &= \hat{\omega}_{i,j,k}^a, \quad a = 1, 2, 3 \end{aligned}$$

und

$$\begin{aligned} y = \text{prox}_{\mathcal{S},f^*}(\hat{y}) &\Leftrightarrow \psi_{i,j,k,l} = \hat{\psi}_{i,j,k,l} - \frac{\sum_l \psi_{i,j,k,l}}{L} \\ \varphi_{i,j,k} &= \hat{\varphi}_{i,j,k} \\ \xi_{i,j,k} &= \hat{\xi}_{i,j,k} \\ \chi_{i,j,k,l,m} &= \max \left\{ -1, \min \left\{ 1, \hat{\chi}_{i,j,k,l,m} \right\} \right\} \\ \tau_{i,j,k} &= -\lambda \mathcal{S}_{i,j,k}^\tau f_{i,j,k} + \hat{\tau}_{i,j,k} \\ \kappa_{i,j,k}^a &= \max \left\{ -1, \min \left\{ 1, \hat{\kappa}_{i,j,k}^a \right\} \right\}, \quad a = 1, 2, 3. \end{aligned}$$

Durch Einsetzen in den Algorithmus 3 erhalten wir für beide Datenterme primal-duale Algorithmen zur Lösung der entsprechenden Optimierungsprobleme für  $\text{TVX}_1^{\alpha,\beta}$ .

## 5.3 Resultate und Auswertung

Im Folgenden werden wir Resultate zur Implementierung der im letzten Kapitel eingeführten Algorithmen vorstellen. Die Implementierung kann der beigefügten CD entnommen werden. Zur Visualisierung der Resultate ist im Ordner "imagine" das Programm *imagine* (siehe [17]) aus dem Matlab Central File Exchange beigefügt, mit dem dreidimensionale Bilder durch Schnittbilder und Projektionen visualisiert werden können. Da das Programm die Grauwerte der Bilder allerdings reskaliert, werden wir im Folgenden stattdessen Schnittbilder mit der Matlab internen Funktion *imshow* darstellen und explizit angeben, falls wir eine Reskalierung vornehmen.

Aufgrund der Laufzeiten des Algorithmus ist ein Test nur mit sehr kleinen, einfachen Bildern möglich. Die im Folgenden genutzten Bilder sind in der Regel  $40 \times 40 \times 20$  Pixel groß. Wir möchten im Folgenden zum einen testen, ob die verschiedenen Optimierungsprobleme zur Entrauschung von Bildern geeignet sind und zum anderen herausfinden, wie sich die Krümmungsregularisierung in Bildern bemerkbar macht.

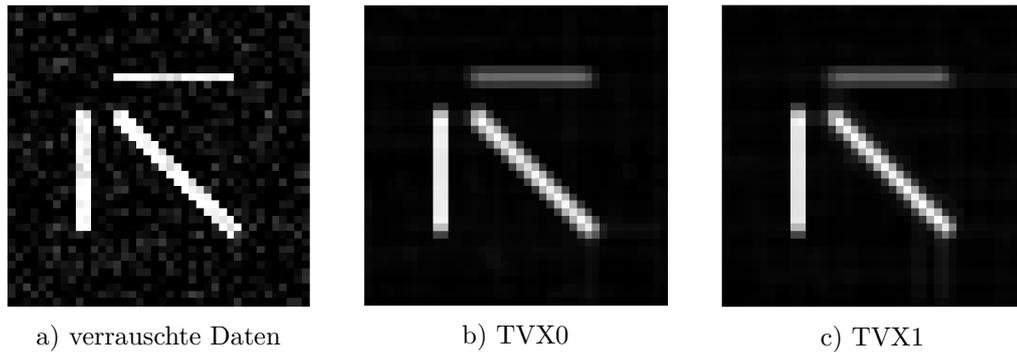


Abbildung 5.2: Entrauschung mit dem  $L^2$ -Datenterm und Parametern  $\alpha = 0.1$ ,  $\beta = 1$ ,  $\lambda = 4$  für 6 Richtungen, Gauss-Rauschen mit Varianz  $\sigma^2 = 0.01$

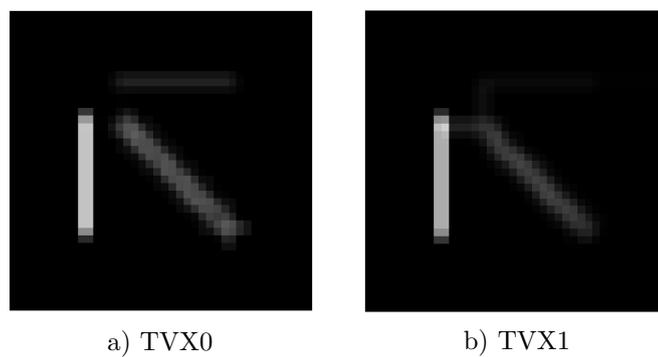


Abbildung 5.3: Entrauschung mit dem  $L^2$ -Datenterm und Parametern  $\alpha = 0.5$ ,  $\beta = 5$ ,  $\lambda = 4$  für 6 Richtungen, Gauss-Rauschen mit Varianz  $\sigma^2 = 0.01$

Die Entrauschung funktioniert, allerdings zeigen sich bereits in Abbildung 5.2 erste Probleme. Aufgrund der gewählten Projektion kann die horizontale Linie nicht exakt rekonstruiert werden. Außerdem sind in den Abbildung 5.2b und 5.2c bei genauer Betrachtung im Hintergrund noch linienförmige Strukturen erkennbar, deren Richtung den diskretisierten Richtungen entspricht. Diese verschwinden mit zunehmend stärkerer Regularisierung jedoch und sind in Abbildung 5.3 nicht mehr erkennbar.

Eine starke Krümmungsregularisierung wirkt sich insbesondere auf die schräge Linie aus. Während die horizontale und die vertikale Linie jeweils diskretisierten Richtungen auf der  $S^2$  entsprechen, muss die schräge Linie aus eben diesen Richtungen zusammengesetzt werden und weist eine dementsprechend starke Krümmung auf, was sie anfällig für die Regularisierung macht. Dies ist besonders deutlich in Abbildung 5.3 zu sehen, bei der wir eine stärkere Regularisierung gewählt haben. Zudem ist in dieser Abbildung der für TV-Regularisierungen typische Kontrastverlust bei höherer Regularisierung deutlich bemerkbar. In Abbildung 5.3b entsteht zudem eine Verbindung der Linien am oberen Ende, um so die Krümmung zu minimieren. Bei der Regularisierung mit TVX0 entsteht diese nicht, da sich die Anzahl der Kanten dadurch nicht verringert.

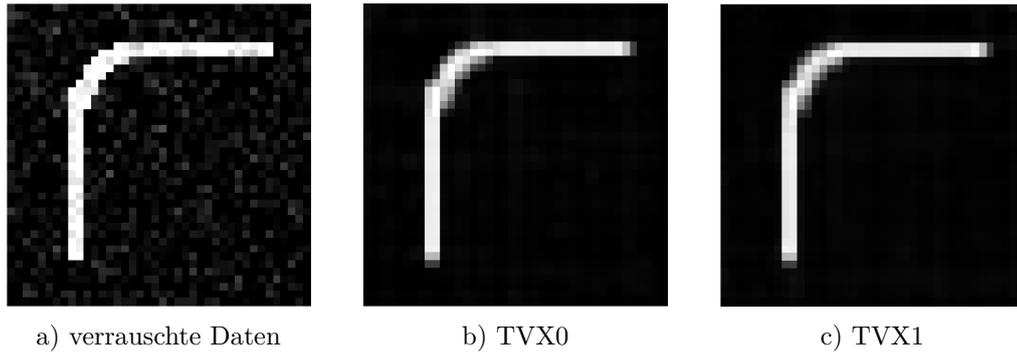


Abbildung 5.4: Entrauschung mit dem  $L^2$ -Datenterm und Parametern  $\alpha = 0.1$ ,  $\beta = 1$ ,  $\lambda = 4$  für 6 Richtungen, Gauss-Rauschen mit Varianz  $\sigma^2 = 0.01$

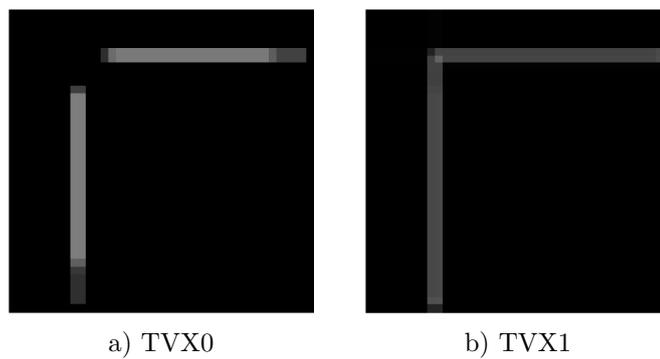


Abbildung 5.5: Entrauschung mit dem  $L^2$ -Datenterm und Parametern  $\alpha = 0.1$ ,  $\beta = 3$ ,  $\lambda = 0.5$  für 6 Richtungen, Gauss-Rauschen mit Varianz  $\sigma^2 = 0.01$

In den Abbildungen 5.4 und 5.5 zeigt sich deutlich, wie sich eine starke Krümmungsregularisierung auf Bilder auswirkt. Während die Krümmungen bei geringer Krümmungsregularisierung in Abbildung 5.4 noch rekonstruiert werden können, ist dieses bei starker Krümmungsregularisierung in Abbildung 5.5 nicht mehr der Fall. Der gekrümmte Teil der Linie verschwindet. Zudem zeigen sich hier auch Unterschiede zwischen TVX0 und TVX1. Für TVX0 ist es nicht günstiger die beiden Linien zu verbinden, da die entstehende Ecke von  $\mu$  aufgrund der Ein- Ausflussbedingungen in den Endpunkten in beiden Richtungen durchlaufen und damit doppelt zählen würde. Für TVX1 verringert das Verbinden der Linien allerdings den Winkel und wird daher bei genügend starker Krümmungsregularisierung bevorzugt. In Abbildung 5.5 zeigt sich zudem eine Tendenz die Linien bis zum Rand des Bildes fortzusetzen, was vermutlich an den gewählten Randbedingungen liegt.

Während der  $L^2$ -Datenterm also zur Entrauschung Gauss-verrauschter Daten geeignet ist, können wir dies beim Wasserstein-1 Datenterm, aufgrund der Voraussetzung, dass die Masse des verrauschten Bildes und die der optimalen Lösung gleich sein müssen, nicht erwarten. In Abbildung 5.6 zeigt sich, dass statt einer Entrauschung im Hintergrund neue linienförmige Strukturen entstehen.

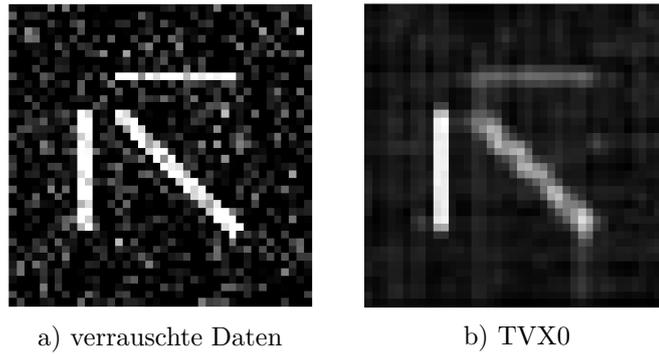


Abbildung 5.6: Entrauschung mit dem  $W^1$ -Datenterm und Parametern  $\alpha = 0.1$ ,  $\beta = 1$ ,  $\lambda = 0.3$  für 6 Richtungen, Gauss-Rauschen mit Varianz  $\sigma^2 = 0.05$

Wir testen das Entrauschungsproblem mit Wasserstein-1 Datenterm daher nicht mit einem Gaussrauschen sondern mit einem geometrischen Rauschen. Dazu verrauschen wir gegebene Daten  $u$  indem wir auf jede Koordinate der Pixel  $u_{i,j,k}$  mit  $u_{i,j,k} \neq 0$  eine zufällige ganze Zahl  $i$  mit  $-n \leq i \leq n$  für ein vorgegebenes  $n \in \mathbb{N}$  aufaddieren. In den Abbildungen 5.7 und 5.8 sind die Originaldaten und verrauschte Daten für  $n = 1$  dargestellt. Hierbei ist zu beachten, dass wir die Darstellung der verrauschten Daten reskaliert haben, da aufgrund der Verrauschung in einigen Pixeln Grauwerte angenommen werden, die doppelt so hoch sind wie der maximale Grauwert in den Originaldaten.

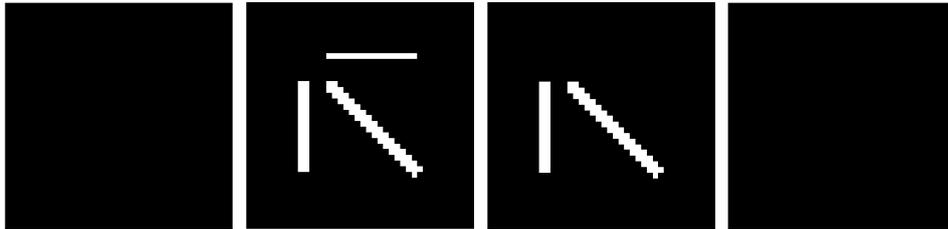
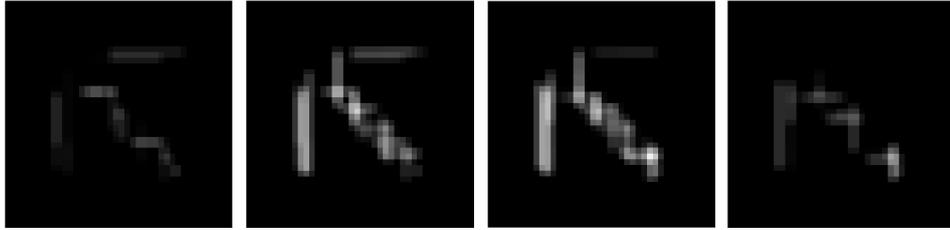
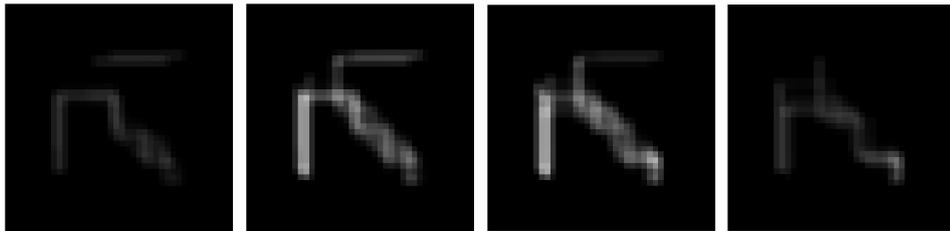


Abbildung 5.7: relevante Schnittbilder der unverrauschten Daten



Abbildung 5.8: geometrisch verrauschte Daten, Darstellung reskaliert mit Faktor  $\frac{1}{2}$

Abbildung 5.9: Entrauschung mit TVX0 für Parameter  $\alpha = 0.1$ ,  $\beta = 1$ ,  $\lambda = 0.3$  für 6 RichtungenAbbildung 5.10: Entrauschung mit TVX1 für Parameter  $\alpha = 0.1$ ,  $\beta = 1$ ,  $\lambda = 0.3$  für 6 Richtungen

In den Abbildungen 5.9 und 5.10 wurden die Daten aus Abbildung 5.8 mit den Optimierungsproblemen TVX0 und TVX1 mit dem Wasserstein-1 Datenterm entrauscht. Deutlich erkennbar ist in beiden Fällen eine Tendenz die Strukturen mit horizontalen und vertikalen Linien zu rekonstruieren, da die zugehörigen Richtungen diskretisierten Richtungen auf der  $S^2$  entsprechen. Weiterhin zeigt sich auch in diesem Fall der aus Abbildung 5.5 bekannte Unterschied zwischen TVX0 und TVX1. Während es für TVX1 günstiger ist die horizontalen und vertikalen Linien zu verbinden, um so die Krümmung zu minimieren, geschieht dies für TVX0 nicht. Der Kontrastverlust gegenüber den unverrauschten Daten ist mit der geometrischen Verrauschung zu erklären, da sich in der anschließenden Entrauschung die insgesamt gleichgroße Masse auf eine größere Anzahl an Pixeln verteilt.

Wir möchten nun kurz auf die Konvergenz der Algorithmen eingehen. Wir wählen im Folgenden als Parameter  $\alpha = 0.1$ ,  $\beta = 1$ ,  $\lambda = 1$  und lösen für die  $L^2$ -Datenterme ein Entrauschungsproblem für Gauss-verrauschte Daten und für die Wasserstein-1-Datenterme ein Entrauschungsproblem für geometrisch verrauschte Daten. Wir führen für jeden Algorithmus 100000 Iterationen durch und speichern dabei Zwischenergebnisse  $(x_n, y_n)$ , wobei  $n$  die Anzahl der Iterationen sei. Die 100000. Iterierte nutzen wir als Approximation an die optimale Lösung  $(x^*, y^*)$  der Entrauschungsprobleme und berechnen zunächst die Abstände der Zwischenergebnisse zur optimalen Lösung  $\|x^* - x_n\|_1$  und  $\|y^* - y_n\|_1$  für die verschiedenen Algorithmen. Die Ergebnisse sind in Abbildung 5.11 durch doppeltlogarithmische Plots dargestellt. Die gespeicherten und ausgewerteten Zwischenergebnisse sind hierbei durch Marker dargestellt.

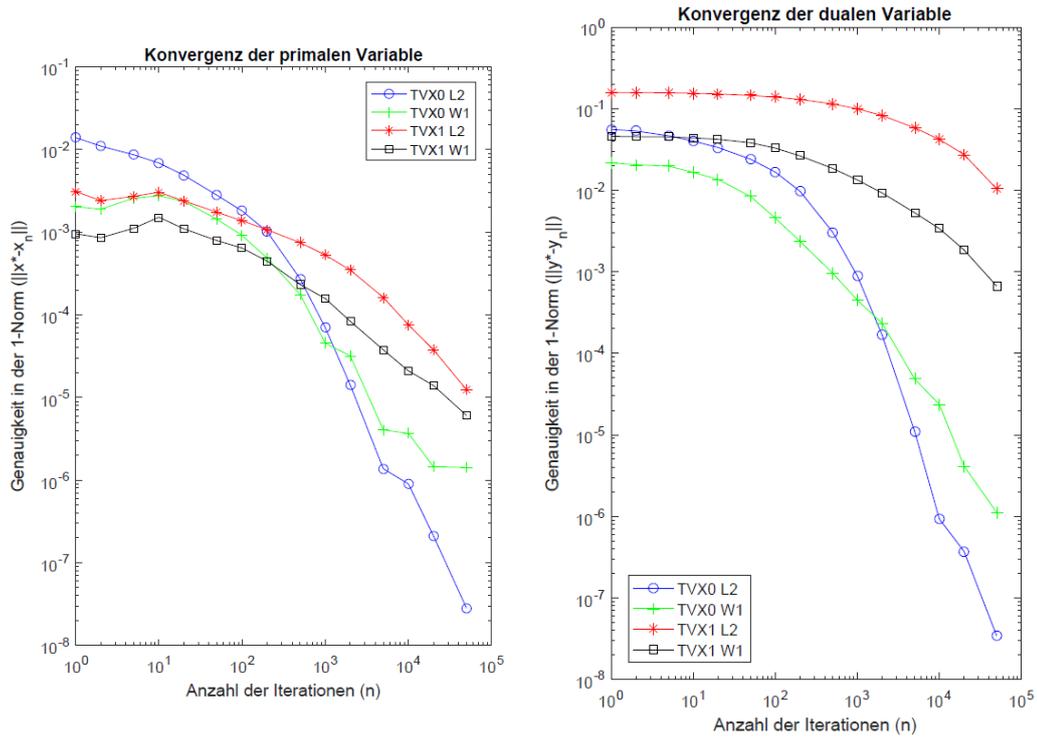
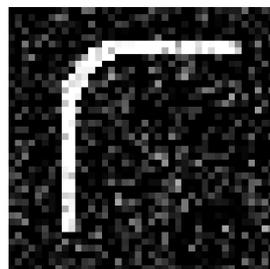
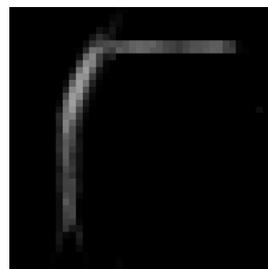


Abbildung 5.11: Durchschnittlicher Abstand der Einträge der Zwischenergebnisse  $(x_n, y_n)$  zur optimalen Lösung  $(x^*, y^*)$

Auffällig ist die in beiden Variablen schlechtere Konvergenz der Optimierungsprobleme für TVX1. Dies ist vermutlich darauf zurückzuführen, dass der Operator  $K$  für diese Optimierungsprobleme deutlich größer ist und daher die Vorkonditionierungsmatrizen  $\mathcal{T}$  und  $\mathcal{S}$  aus Theorem 4 kleinere Einträge haben, was zu kleineren Schrittweiten im Algorithmus führt. Der Algorithmus für TVX1 braucht dann deutlich mehr Iterationen um ein mathematisch genaues Ergebnis zu liefern. Dies ist allerdings aufgrund der Laufzeiten des Algorithmus auf unserem Testsystem (8 GB RAM, Intel i7-4510U CPU) problematisch. Während der Algorithmus für das Optimierungsproblem TVX0 mit  $L^2$  Datenterm etwa 4,2 Stunden (15136 Sekunden) für 100000 Iterationen benötigte, wurden für das Optimierungsproblem TVX1 mit  $L^2$  Datenterm etwa 10,5 Stunden (37907 Sekunden) benötigt.



a) Verrauschte Daten



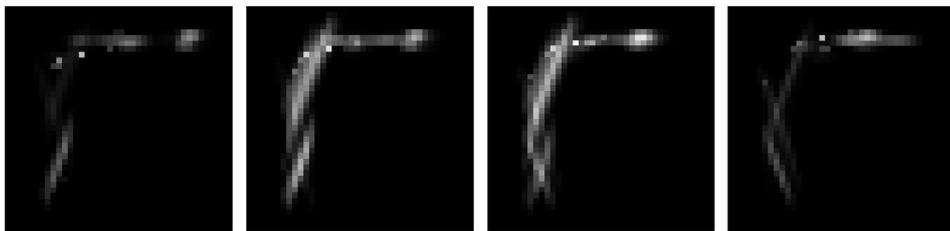
b) Entrausung mit TVX0

Abbildung 5.12: Entrausung mit dem  $L^2$  Datenterm für Parameter  $\alpha = 0.1$ ,  $\beta = 1$ ,  $\lambda = 0.5$  für 20 Richtungen, Gauss-Rauschen mit Varianz  $\sigma^2 = 0.05$

Zum Abschluss möchten wir noch kurz auf die Algorithmen eingehen, bei denen die  $S^2$  durch 20 Gitterpunkte diskretisiert wurde. Die Laufzeiten für diese Algorithmen sind deutlich länger. Tests sind daher insbesondere für TVX1 nur eingeschränkt möglich. Es zeigen sich jedoch auch für 20 Richtungen die bereits beobachteten Effekte. In Abbildung 5.12 wird der gekrümmte Teil der Linie gut durch eine der 20 diskretisierten Richtungen rekonstruiert. Probleme treten in diesem Fall bei der Rekonstruktion der horizontalen und vertikalen Anteile der Linie auf, da diese nicht diskretisierten Richtungen entsprechen. Dies wird vor allem bei der geometrischen Entrauschung mit dem Wasserstein-1 Datenterm in Abbildung 5.15 deutlich, in der diese Linien nicht gut rekonstruiert werden können. Es wäre daher sinnvoll die  $S^2$  durch noch mehr als 20 Gitterpunkte zu diskretisieren, was allerdings aufgrund der Laufzeiten problematisch ist und deshalb ein Ziel zukünftiger Arbeit bleibt.



Abbildung 5.13: relevante Schnittbilder der unverrauschten Daten

Abbildung 5.14: Verrauschte Daten, Darstellung reskaliert mit Faktor  $\frac{1}{3}$ Abbildung 5.15: Entrauschung mit TVX0 für Parameter  $\alpha = 0.3$ ,  $\beta = 1$ ,  $\lambda = 0.075$  für 20 Richtungen

---

---

## KAPITEL 6

---

### Fazit

Das Ziel dieser Arbeit war es die Resultate aus [2] auf einen dreidimensionalen Fall zu übertragen, in dem die Bilder dünne, linienartige Strukturen enthalten.

Wir haben dazu zunächst einige Grundlage der Optimierung, Maßtheorie und Funktionalanalysis eingeführt, die wir im Verlauf der Arbeit benötigt haben. Im Hauptteil der Arbeit haben wir dann zunächst den Begriff des Linienmaßes eingeführt und veranschaulicht, wie wir Informationen über Richtungsänderungen in Knoten von Linienmaßen gewinnen können. Wir haben anschließend Normen auf  $\mathcal{M}(S^2)$  definiert, die diese Richtungsänderungen, die durch Differenzen von Diracmaßen beschrieben werden, bestrafen. Mit Hilfe des Funktionals  $T_\rho$ , welches alle Strafen für Richtungsänderungen in einem Linienmaß aufsummiert, sind wir in der Lage die gesamten Richtungsänderungen in einem Linienmaß zu bestrafen und so Informationen über dessen Krümmung zu gewinnen. Zum Einen haben wir dabei ein Funktional  $T_{\rho_0}$  eingeführt, mit dem wir in der Lage sind die Anzahl der Richtungsänderungen in einem Linienmaß zu minimieren. Zum Anderen haben wir ein Funktional  $T_{\rho_1}$  eingeführt, welches die Winkel der Richtungsänderungen minimiert. Mit den Informationen über die Krümmung waren wir in der Lage von der Eulerschen Elastica inspirierte Funktionale zur Minimierung der Krümmung in Linienmaßen aufzustellen. Nach einer konvexen Relaxierung haben wir Optimierungsprobleme zur Entrauschung von Linienmaßen aufgestellt und für einen Wasserstein- und einen  $L^2$  Datenterm Existenzresultate gezeigt. Im abschließenden Kapitel dieser Arbeit haben wir zunächst die Entrauschungsprobleme diskretisiert und dabei Optimierungsprobleme in Sattelpunktform erhalten, die wir zu primal-dualen Optimierungsproblemen umformen konnten. Anhand einer Implementation in Matlab haben wir gesehen, dass der  $L^2$  erwartungsgemäß gut zur Entrauschung Gaussverrauschter Daten geeignet ist. Der Wasserstein-Datenterm ist hingegen zur Entrauschung geometrisch verrauschter Daten geeignet.

Wir haben angenommen, dass die Linien in den Linienmaßen durch Strecken gegeben sind, und daher Krümmungsänderungen nur in Knoten der Linienmaße auftreten. Ein Ziel zukünftiger Forschung könnte daher das Übertragen der Resultate auf  $\mathcal{C}^1$  Kurven sein. Weiterhin ist insbesondere die Laufzeit des Algorithmus problematisch. Für den Fall, dass die  $S^2$  mit 20 Richtungen diskretisiert ist, waren ausführliche Tests aufgrund der Laufzeit nicht möglich. Deshalb sollte ein weiteres Ziel die Verbesserung der Laufzeit sein. Dann könnten beispielsweise auch Tests mit mehr als 20 diskretisierten Richtungen auf der  $S^2$  und Tests mit realistischen Daten, etwa dreidimensionale Aufnahmen von Blutgefäßen, durchgeführt werden. Für das Entrauschungsproblem mit Wasserstein-Datenterm bieten sich zudem alternativ zum Algorithmus von Chambolle und Pock Verfahren der linearen Optimierung, beispielsweise Innere-Punkte Methoden oder Simplex-Verfahren, an.

---

---

## Literaturverzeichnis

- [1] H. ALT, *Lineare Funktionalanalysis*, Springer Verlag, 1985.
- [2] K. BREDIES, T. POCK, AND B. WIRTH, *Convex relaxation of a class of vertex penalizing functionals*, Journal of Mathematical Imaging and Vision, 47 (2013), pp. 278–302.
- [3] A. CHAMBOLLE AND T. POCK, *A first-order primal-dual algorithm for convex problems with applications to imaging*, Journal of Mathematical Imaging and Vision, 40 (2011), pp. 120–145.
- [4] ———, *On the ergodic convergence of a first-order primal-dual algorithm*, Mathematical Programming, 159 (2016), pp. 253–287.
- [5] E. DIBENEDETTO, *Real Analysis*, Birkhäuser Advanced Texts, Birkhäuser Basel, 2002.
- [6] J. ELSTRODT, *Maß- und Integrationstheorie*, Springer Verlag, 1996.
- [7] G. KANISZA, *Organization in Vision: Essays on Gestalt Perception*, Praeger, 1979.
- [8] B. KORTE AND J. VYGEN, *Kombinatorische Optimierung*, Springer Verlag, 2008.
- [9] R. LEVIEN, *The elastica: a mathematical history*, Tech. Report UCB/EECS-2008-103, EECS Department, University of California, Berkeley, 2008.
- [10] P. MATTILA, *Geometry of Sets and Measures in Euclidean Spaces: Fractals and Rectifiability*, vol. 44 of Cambridge studies in advanced mathematics, Cambridge University Press, 1995.
- [11] D. S. MITRINOVIC, J. PECARIC, AND A. FINK, *Classical and New Inequalities in Analysis*, Springer Verlag, 1993.
- [12] D. MUMFORD, *Elastica and computer vision*, Algebraic Geometry and Its Applications, (1994), pp. 491–506.
- [13] T. POCK AND A. CHAMBOLLE, *Diagonal preconditioning for first order primal-dual algorithms in convex optimization*, International Conference of Computer Vision, ICCV 2011, (2011), pp. 1762–1769.
- [14] R. T. ROCKAFELLAR, *Convex Analysis*, Princeton University Press, 1970.
- [15] C. VILLANI, *Optimal transport, old and new*, Grundlehren der mathematischen Wissenschaften, Springer Verlag, 2009.
- [16] D. WERNER, *Funktionalanalysis*, Springer Verlag, 1995.
- [17] C. WUERSLIN, *Imagine*. <http://de.mathworks.com/matlabcentral/fileexchange/40440-imagine>. Abgerufen: 11.09.2016.

Hiermit versichere ich, dass die vorliegende Arbeit mit dem Titel *Konvexe Krümmungsregularisierung für Linienmaße in drei Dimensionen* selbstständig verfasst worden ist, dass keine anderen Quellen und Hilfsmittel als die angegebenen benutzt worden sind und dass die Stellen der Arbeit, die anderen Werken – auch elektronischen Medien – dem Wortlaut oder Sinn nach entnommenen wurden, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht worden sind.

Münster, 26. September 2016

---

(Bernd Mekes)

Ich erkläre mich mit einem Abgleich der Arbeit mit anderen Texten zwecks Auffindung von Übereinstimmungen sowie mit einer zu diesem Zweck vorzunehmenden Speicherung der Arbeit in eine Datenbank einverstanden.

Münster, 26. September 2016

---

(Bernd Mekes)