

WESTFÄLISCHE WILHELMS-UNIVERSITÄT MÜNSTER

MASTERARBEIT

---

# Dimensionsreduktion von dynamischer Super-Resolution

---

*Autor:*

Alexander SCHLÜTER  
alx.schlueter@gmail.com  
Matrikelnr. 409649

*Betreuer:*

Prof. Dr. Benedikt WIRTH  
Benedikt.Wirth@uni-muenster.de

eingereicht am 31. Oktober 2019



FACHBEREICH 10  
MATHEMATIK UND  
INFORMATIK

# Inhaltsverzeichnis

<b>Einleitung</b>	<b>3</b>
<b>1. Theorie komplexwertiger Maße</b>	<b>6</b>
1.1. Allgemeine Maßräume . . . . .	6
1.2. Borelmaße und der Darstellungssatz von Riesz . . . . .	13
<b>2. Super-Resolution über konvexe Optimierung</b>	<b>18</b>
2.1. Der Messoperator . . . . .	19
2.2. Duale Struktur und Existenz von Lösungen . . . . .	25
2.3. Exakte Rekonstruktion . . . . .	30
2.4. Konstruktion dualer Zertifikate . . . . .	33
<b>3. Dynamische Super-Resolution</b>	<b>39</b>
3.1. Vorstellung des Modells . . . . .	39
3.2. Der dynamische Messoperator . . . . .	40
3.3. Dynamische duale Zertifikate . . . . .	41
<b>4. Dimensionsreduktion</b>	<b>47</b>
4.1. Radon-Transformation für Maße . . . . .	47
4.2. Problemformulierung . . . . .	49
4.3. Träger der neuen Variablen . . . . .	55
4.4. Exakte Rekonstruktion . . . . .	56
<b>5. Numerik</b>	<b>69</b>
5.1. Gitter-Diskretisierung . . . . .	69
5.2. ADCG-Algorithmus . . . . .	71
5.3. Dimensionsreduziertes Gitterverfahren in 2D . . . . .	76
5.4. Vergleich der Verfahren . . . . .	78
<b>Schlusswort</b>	<b>83</b>
<b>Literatur</b>	<b>85</b>
<b>A. Anhang</b>	<b>87</b>
A.1. Beweise aus Kapitel 2 . . . . .	87
A.2. Beweise aus Abschnitt 4.3 . . . . .	89

# Einleitung

Wir beschäftigen uns mit dem Problem der Super-Resolution (auf Deutsch auch Super-Auflösung), also mit der Fragestellung, wie feinskalige Information aus der Messung eines grobskaligen Signals rekonstruiert werden können. Konkret stellen wir uns eine Sammlung von Punktquellen vor, bei denen es sich zum Beispiel um fluoreszierende Moleküle oder Sterne handeln kann. Mathematisch wird eine solche Punkteverteilung repräsentiert durch eine Linearkombination von Dirac-Maßen

$$\hat{u} = \sum_{i=1}^N \alpha_i \delta_{z_i}$$

mit Intensitäten  $\alpha_i \in \mathbb{C}$  an Positionen  $z_i$  innerhalb eines Gebietes  $\Omega \subset \mathbb{R}^d$ .

Als Messung kommen Bildgebungsverfahren wie zum Beispiel die Mikroskopie, PET-Scanner in der Medizin oder die Teleskopie in der Astronomie in Frage. Wir stellen uns vor, dass die Messung insofern unvollständig ist, als feinskalige Informationen im Laufe des Messprozesses verloren gehen. Im Fall der optischen Bildgebungsverfahren führt zum Beispiel die Beugung des Lichtes an der Blende dazu, dass Punktquellen nicht als scharfe Punkte, sondern als verschwommene Scheiben abgebildet werden. Dies stellt eine physikalische Auflösungsgrenze dar, welche auch als *Abbe-Limit* bekannt ist.

Im Artikel „Towards a Mathematical Theory of Super-resolution“ [9] wurde von Candès et al. gezeigt, dass sich die Positionen der Punktquellen durch ein konvexes Minimierungsproblem der Form

$$\min_{u \in \mathcal{M}(\Omega)} \|u\|_{\text{TV}} \quad \text{s.t.} \quad \mathcal{G}u = \hat{y} \quad (\text{ER})$$

über dem Raum der Maße  $\mathcal{M}(\Omega)$  exakt rekonstruieren lassen. Hierbei modelliert ein Operator  $\mathcal{G}$  den Messprozess,  $\hat{y} = \mathcal{G}\hat{u}$  sind die gemessenen Daten und  $\|\cdot\|_{\text{TV}}$  ist die sogenannte Totale-Variations-Norm eines Maßes. Dies ist möglich, solange die aufzulösenden Punkte einen Mindestabstand nicht unterschreiten, welcher von dem Frequenzbereich abhängt, der durch den Messprozess abgeschnitten wird. Der Beweis nutzt auf entscheidende Weise die Dualität zwischen dem Raum der Maße und dem Raum der stetigen Funktionen, welche als der *Darstellungssatz von Riesz* bekannt ist. Zu einer gegebenen Punkteverteilung werden sogenannte *dual-Zertifikate* konstruiert. Hierbei handelt es sich um trigonometrische Polynome, welche eine von den Positionen der Punktquellen abhängige Interpolationsaufgabe lösen müssen. Die Existenz eines dualen Zertifikates garantiert, dass die tatsächliche Punkteverteilung die eindeutige Lösung des Problems (ER) ist.

In der Realität stehen die Dinge selten still: Sterne wandern am Himmel, Zellen bewegen sich in der Blutbahn, der Patient zuckt im Computertomographen. Der Experimentator befindet sich in einem Dilemma: Misst er über einen langen Zeitraum, so können viele Daten für eine Rekonstruktion gesammelt werden. Währenddessen wird sich die Probe jedoch über eine weite Strecke bewegt haben und das Bild verschwimmen.

Verkürzt er die Messdauer, so ist die Bewegung im Vergleich zur Größe der abzubildenden Struktur möglicherweise vernachlässigbar. Allerdings können im kurzen Zeitraum nur wenige Daten gesammelt werden, was eine Rekonstruktion erschwert.

Die Lösung des Dilemmas ist eine Berücksichtigung der Bewegung in der Modellierung, sodass diese während der Rekonstruktion wieder herausgerechnet werden kann. Die entsprechende Verallgemeinerung von (ER) für lineare Dynamiken wurde von Alberti et al. unter dem Begriff „Dynamic Spike Super-resolution“ [2] veröffentlicht und führt auf ein sehr ähnliches TV-Minimierungsproblem. Allerdings treten statt der Orte  $x_i \in \mathbb{R}^d$  nun Orte und Geschwindigkeiten  $(x_i, v_i) \in \mathbb{R}^d \times \mathbb{R}^d$  als Variablen auf. Das zu rekonstruierende Maß hat also die Form

$$\hat{\lambda} = \sum_{i=1}^N \alpha_i \delta_{(x_i, v_i)}.$$

Dieses *Phasenraum-Lifting* des Problems erlaubt es, vieles von der Theorie zu (ER) auf das dynamische Problem zu übertragen.

Doch das Lifting ist nicht ohne Nachteil: Das resultierende Problem hat Dimension  $2d$  statt  $d$ . Ein üblicher Diskretisierungsansatz besteht darin, ein Gitter über den Parameterraum zu legen. Für  $d = 3$  wäre hierzu ein sechsdimensionales Gitter notwendig, was schnell zu numerischen Problemen aufgrund der Vielzahl an Variablen führt.

In dieser Arbeit wird eine Formulierung untersucht, welche in gewisser Weise „zwischen“ dem statischen und dem dynamischen Problem angesiedelt ist: Für jeden Richtungsvektor  $\theta \in \mathbb{S}^{d-1}$  wird die gemeinsame Projektion der Orts- und Geschwindigkeitsvariablen auf  $\theta$  betrachtet, was auf ein Maß der Form

$$\hat{\mu}_\theta = \sum_{i=1}^N \alpha_i \delta_{(\theta \cdot x_i, \theta \cdot v_i)}$$

führt. Es wird ein dimensionsreduziertes Problem formuliert, welches die Dynamik der Teilchen berücksichtigt, aber nur noch Variablen der Dimension  $d + 1$  statt  $2d$  enthält.

Die Arbeit ist wie folgt organisiert: Zu Beginn werden die mathematischen Werkzeuge eingeführt, welche später benötigt werden. Dies ist die Theorie komplexwertiger Maße, darunter der Zerlegungssatz von Lebesgue und der Darstellungssatz von Riesz.

Anschließend wird die Methode der Super-Resolution mittels konvexer Optimierung in Kapitel 2 vorgestellt. Hier wird ein duales Problem hergeleitet und der

---

Begriff der dualen Zertifikate definiert, was es erlaubt, Aussagen über Existenz und Eindeutigkeit von Lösungen zu treffen.

Der Super-Resolution von bewegten Teilchen wenden wir uns in Kapitel 3 zu. Nach einer Vorstellung des Modells werden Situationen untersucht, in welchen dynamische duale Zertifikate existieren. Hierfür wird der Begriff des Durchschnittszertifikates sowie der Begriff der Geisterteilchen zentral sein.

Es folgt die Vorstellung einer neuen Methode zur Dimensionsreduktion des dynamischen Super-Resolution Problems in Kapitel 4. Verschiedene mögliche Formulierungen werden diskutiert. Es wird bewiesen, dass die Methode trotz reduzierter Dimension des Problems das Zielmaß unter bestimmten Voraussetzungen noch exakt rekonstruieren kann.

Die Numerik ist Thema des letzten Kapitels. Hier werden zunächst zwei Verfahren aus der Literatur zur numerischen Lösung des Minimierungsproblems beschrieben: Eine einfache Gitter-Diskretisierung sowie der sogenannte ADCG-Algorithmus. Der Schwerpunkt liegt jedoch in der Entwicklung einer geeigneten Diskretisierung für die dimensionsreduzierte Methode, welche schließlich anhand eines hierfür generierten Datensatzes mit mehreren anderen Methoden verglichen wird.

# 1. Theorie komplexwertiger Maße

Zur mathematischen Modellierung der Bilder im Kontinuierlichen sollen Maße herangezogen werden. Damit beliebige, auch negative und sogar komplexwertige Intensitäten modelliert werden können, müssen die Maße auch Werte in  $\mathbb{C}$  annehmen dürfen. Die Theorie komplexwertiger Maße soll in diesem Kapitel kurz eingeführt werden, wobei auf Beweise größtenteils verzichtet wird; die Hauptquelle ist Kapitel 6 aus dem Buch von Rudin [20]. Hierbei werden die Grundlagen der Theorie von positiven, das heißt  $[0, \infty]$ -wertigen Maßen, als bekannt vorausgesetzt.

## 1.1. Allgemeine Maßräume

Wir beginnen mit dem Teil der Theorie für allgemeine Maßräume, der auf der Grundmenge  $X$  keine besondere Struktur außer einer  $\sigma$ -Algebra  $\mathcal{A}$  verlangt. Im Folgenden seien alle Maße auf dem messbaren Raum  $(X, \mathcal{A})$  definiert, sofern nicht anders angegeben.

**1.1 Definition** ([20, Kap. 6.1]) Ein komplexwertiges Maß auf  $(X, \mathcal{A})$  ist eine Abbildung  $\mu : \mathcal{A} \rightarrow \mathbb{C}$ , sodass für alle Folgen  $(E_i)_{i \in \mathbb{N}} \subset \mathcal{A}$  von paarweise disjunkten messbaren Mengen gilt: Die Reihe  $\sum_{i=1}^{\infty} \mu(E_i)$  konvergiert und

$$\mu \left( \bigcup_{i=1}^{\infty} E_i \right) = \sum_{i=1}^{\infty} \mu(E_i). \quad (\sigma\text{-Additivität})$$

Ein reellwertiges Maß ist eine Abbildung  $\mathcal{A} \rightarrow \mathbb{R}$ , welche dieselbe Eigenschaft erfüllt.

Da der Term auf der linken Seite der Gleichheit unabhängig von der Reihenfolge der Mengen  $E_i$  ist, muss dies auch für den Grenzwert der Reihe gelten. Die Reihe konvergiert somit absolut in  $\mathbb{C}$  bzw.  $\mathbb{R}$ . Außerdem bilden die komplexwertigen Maße einen Vektorraum über  $\mathbb{C}$  und die reellwertigen einen Vektorraum über  $\mathbb{R}$ .

Das komplex konjugierte Maß  $\bar{\mu}$  zu einem komplexwertigen Maß  $\mu$  ist definiert durch

$$\bar{\mu}(E) := \overline{\mu(E)}, \quad E \in \mathcal{A}.$$

Auf die gleiche Weise werden Real- und Imaginärteil  $\operatorname{Re} \mu$  bzw.  $\operatorname{Im} \mu$  eines komplexwertigen Maßes definiert.

Eine *Partition* einer Menge  $E$  ist eine Folge  $(E_i)_{i \in \mathbb{N}}$  von paarweise disjunkten Mengen, sodass  $E = \bigcup_{i=1}^{\infty} E_i$ .

---

**1.2 Definition und Satz** (Totale Variation) Es sei  $\mu$  ein komplexwertiges Maß. Definiere eine Abbildung  $|\mu| : \mathcal{A} \rightarrow [0, \infty]$  durch

$$|\mu|(E) := \sup \left\{ \sum_{i=1}^{\infty} |\mu(E_i)| \mid (E_i)_{i \in \mathbb{N}} \subset \mathcal{A} \text{ ist Partition von } E \right\}.$$

Dann ist  $|\mu|$  ein endliches positives Maß auf  $\mathcal{A}$  und wird das *Totale-Variations-Maß* von  $\mu$  genannt. Die Zahl

$$\|\mu\|_{\text{TV}} := |\mu|(X) \in [0, \infty)$$

heißt *Totale Variation* von  $\mu$ . Durch die Vorschrift  $\mu \mapsto \|\mu\|_{\text{TV}}$  wird eine Norm auf dem Vektorraum der komplexwertigen Maße definiert.

*Beweis.* Zu den Maßeigenschaften von  $|\mu|$  siehe [20, Thm. 6.2].

Nach [20, Thm 6.4] gilt  $|\mu|(X) < \infty$ , sodass  $|\mu|$  tatsächlich ein endliches Maß ist und auch  $\|\cdot\|_{\text{TV}}$  nur endliche Werte annimmt. Aus der Definition der Totalen Variation und den Eigenschaften des Betrages und des Supremums folgt sofort die Dreiecksungleichung für  $\|\cdot\|_{\text{TV}}$ , sowie  $\|c\mu\|_{\text{TV}} = |c|\|\mu\|_{\text{TV}}$  für alle  $c \in \mathbb{C}$ .

Sei  $\mu \neq 0$  ein komplexwertiges Maß, dann existiert  $E \in \mathcal{A}$  mit  $\mu(E) \neq 0$ . Die Familie  $E_1 := E$ ,  $E_i := \emptyset$  für  $i > 1$  ist trivialerweise eine Partition von  $E$ , sodass folgt

$$\|\mu\|_{\text{TV}} \geq \sum_{i=1}^{\infty} |\mu(E_i)| = |\mu(E)| > 0.$$

Also ist  $\|\cdot\|_{\text{TV}}$  auch positiv definit und insgesamt eine Norm. □

Das *Dirac-Maß*  $\delta_x$  an einem Punkt  $x \in X$  ist das Maß definiert durch

$$\delta_x(E) := \begin{cases} 1 & \text{falls } x \in E, \\ 0 & \text{sonst.} \end{cases}$$

Die Modellierung von Punktquellen führt in natürlicher Weise auf Linearkombinationen von Dirac-Maßen

$$\sum_{j=1}^N \alpha_j \delta_{x_j}$$

für Gewichte  $\alpha_j \in \mathbb{C}$  und paarweise verschiedene Punkte  $x_j \in X$ ,  $1 \leq j \leq N$ . Für solche Maße ist die Totale-Variations-Norm besonders leicht zu berechnen: Wählt man in der Definition die Partition  $E_i := \{x_i\}$ ,  $1 \leq i \leq N$ ,  $E_{N+1} := X \setminus \{x_1, \dots, x_N\}$  (vorausgesetzt, diese Mengen sind in der  $\sigma$ -Algebra  $\mathcal{A}$  enthalten), so bekommt man

$$\left\| \sum_{j=1}^N \alpha_j \delta_{x_j} \right\|_{\text{TV}} \geq \sum_{i=1}^{N+1} \left| \sum_{j=1}^N \alpha_j \delta_{x_j}(E_i) \right| = \sum_{j=1}^N |\alpha_j|.$$

Umgekehrt ist auch die TV-Norm durch die rechte Seite beschränkt. Es gilt daher

$$\left\| \sum_{j=1}^N \alpha_j \delta_{x_j} \right\|_{\text{TV}} = \|\alpha\|_1,$$

d.h. die TV-Norm ist gleich der 1-Norm des Gewichtsvektors  $\alpha = (\alpha_1, \dots, \alpha_N) \in \mathbb{C}^N$ .

**1.3 Definition** (vgl. [20, Def. 6.7]) Es seien  $\mu, \mu_1, \mu_2$  komplexwertige Maße.

1.  $\mu$  heißt *konzentriert* auf  $A \in \mathcal{A}$ , falls  $|\mu|(A^c) = 0$ .
2.  $\mu_1$  heißt *singulär* bezüglich  $\mu_2$ , falls  $A \in \mathcal{A}$  existiert, sodass  $\mu_1$  konzentriert auf  $A$  und  $\mu_2$  konzentriert auf  $A^c$  ist. Schreibe in diesem Fall  $\mu_1 \perp \mu_2$ .
3.  $\mu_1$  heißt *absolutstetig* bezüglich  $\mu_2$ , falls für alle  $A \in \mathcal{A}$  gilt:

$$|\mu_2|(A) = 0 \implies |\mu_1|(A) = 0.$$

Nutze hierfür die Notation  $\mu_1 \ll \mu_2$ .

Aus der Definition der Totalen Variation über das Supremum folgt sofort die Ungleichung  $|\mu(E)| \leq |\mu|(E)$  für alle  $E \in \mathcal{A}$ . Insbesondere folgt  $\mu \ll |\mu|$  für alle komplexwertigen Maße  $\mu$ .

Für ein positives Maß  $\lambda$  und  $1 \leq p < \infty$  ist die  $L^p$ -Norm einer messbaren Funktion  $f : X \rightarrow \mathbb{C}$  definiert durch

$$\|f\|_{L^p(\lambda)} := \left( \int_X |f|^p d\lambda \right)^{1/p}.$$

Die Menge

$$\mathcal{L}^p(\lambda) := \{f : X \rightarrow \mathbb{C} \mid f \text{ messbar, } \|f\|_{L^p(\lambda)} < \infty\}$$

ist ein  $\mathbb{C}$ -Vektorraum. Identifiziert man alle Funktionen in  $\mathcal{L}^p(\lambda)$ , welche  $\lambda$ -fast überall übereinstimmen, so erhält man die Lebesgue-Räume  $L^p(\lambda)$ , auf welchen  $\|\cdot\|_{L^p(\lambda)}$  tatsächlich die Eigenschaften einer Norm erfüllt. Ist  $\lambda$  das Lebesgue-Maß auf einer messbaren Menge  $X = \Omega \subset \mathbb{R}^d$ , so schreiben wir auch  $L^p(\Omega)$ .

Ein positives Maß  $\lambda$  heist  *$\sigma$ -endlich*, falls eine Folge  $(E_i)_{i \in \mathbb{N}}$  von messbaren Mengen mit  $\lambda(E_i) < \infty$  existiert, sodass

$$\bigcup_{i=1}^{\infty} E_i = X.$$

Der folgende Satz ist zentral für die späteren Beweise zur exakten Rekonstruktion, da er die Zerlegung von Lösungen des konvexen Minimierungsproblems in einen Anteil auf den tatsächlichen Punktquellen und einen Fehleranteil erlaubt.

---

**1.4 Satz** (Satz von Lebesgue-Radon-Nikodym [20, Thm. 6.10]) Es sei  $\lambda$  ein positives,  $\sigma$ -endliches Maß und  $\mu$  ein komplexwertiges Maß.

1. Es existiert ein eindeutiges Paar von komplexwertigen Maßen  $\mu_a$  und  $\mu_s$ , sodass

$$\mu = \mu_a + \mu_s, \quad \mu_a \ll \lambda, \quad \mu_s \perp \lambda.$$

Falls  $\mu$  ein positives endliches Maß ist, dann gilt dies auch für  $\mu_a$  und  $\mu_s$ .

2. Es existiert ein  $h \in \mathcal{L}^1(\lambda)$ , sodass

$$\mu_a(E) = \int_E h \, d\lambda$$

für alle messbaren Mengen  $E \in \mathcal{A}$  und die Funktion  $h$  mit dieser Eigenschaft ist  $\lambda$ -fast überall eindeutig bestimmt.

Wie schon bemerkt, gilt stets der Zusammenhang  $\mu \ll |\mu|$ . Punkt 2 des obigen Satzes lässt sich für den Fall  $\lambda = |\mu|$  jedoch noch verschärfen:

**1.5 Satz** (Polarzerlegung [20, Thm. 6.12]) Es sei  $\mu$  ein komplexwertiges Maß. Dann existiert eine messbare Funktion  $h : X \rightarrow \mathbb{C}$  mit  $|h(x)| = 1$  für alle  $x \in X$  und

$$\mu(E) = \int_E h \, d|\mu|$$

für alle messbaren Mengen  $E \in \mathcal{A}$ .

Die obige Aussage erlaubt es uns nun, die Integration einer messbaren Funktion  $f : X \rightarrow \mathbb{C}$  bezüglich eines komplexwertigen Maßes  $\mu$  zu definieren als

$$\int_X f \, d\mu := \int_X fh \, d|\mu|,$$

solange das rechte Integral wohldefiniert ist. Dementsprechend definiert man

$$\mathcal{L}^1(\mu) := \mathcal{L}^1(|\mu|).$$

Für ein  $f \in \mathcal{L}^1(\mu)$  wird durch  $\nu : \mathcal{A} \rightarrow \mathbb{C}$ ,

$$\nu(E) := \int_E f \, d\mu$$

ein weiteres komplexwertiges Maß definiert und wir schreiben hierfür kurz  $\nu = f \, d\mu$ .

Einige Eigenschaften des Integrals bezüglich eines komplexwertigen Maßes sind in folgendem Lemma zusammengefasst:

**1.6 Lemma** Es sei  $\mu$  ein komplexwertiges Maß.

1. Die Abbildung

$$T_\mu : \mathcal{L}^1(\mu) \rightarrow \mathbb{C}, \quad f \mapsto \int_X f \, d\mu$$

ist  $\mathbb{C}$ -linear.

2. Ist  $f : X \rightarrow \mathbb{C}$  messbar und beschränkt, so ist das Integral von  $f$   $\mathbb{C}$ -linear im Maßargument, d.h. es gilt

$$\int_X f \, d(c_1\mu_1 + c_2\mu_2) = c_1 \int_X f \, d\mu_1 + c_2 \int_X f \, d\mu_2$$

für  $c_i \in \mathbb{C}$  und komplexwertige Maße  $\mu_i$ ,  $i = 1, 2$ .

3. Für  $f \in \mathcal{L}^1(\mu)$  gilt die Abschätzung

$$\left| \int_X f \, d\mu \right| \leq \int_X |f| \, d|\mu|.$$

4. Für die komplexe Konjugation des Integrals von  $f \in \mathcal{L}^1(\mu)$  gilt die Rechenregel

$$\overline{\int_X f \, d\mu} = \int_X \bar{f} \, d\bar{\mu}.$$

*Beweis.* Punkt 1 folgt durch einfaches Nachrechnen und wegen der Linearität des Integrals bezüglich  $|\mu|$ . Für Punkt 2 verweisen wir auf [20, Abschnitt 6.18]. Aussage 3 ergibt sich anhand der Ungleichungskette

$$\left| \int_X f \, d\mu \right| = \left| \int_X fh \, d|\mu| \right| \leq \int_X |fh| \, d|\mu| = \int_X |f| \, d|\mu|.$$

Zum Nachrechnen von Punkt 4 können Funktion und Maß jeweils in ihre Real- und Imaginärteile aufgespalten werden und die Aussage folgt nach Anwendung der Linearität.  $\square$

Die Begriffe des Bildmaßes sowie die zugehörige Transformationsformel übertragen sich fast analog auf komplexwertige Maße, wobei jedoch in der Transformationsformel die Integrierbarkeit der Verknüpfung der Funktionen gefordert werden muss (vgl. [4, Ex. 3.10.68]).

**1.7 Definition** (Bildmaß) Es seien  $(X_1, \mathcal{A}_1)$ ,  $(X_2, \mathcal{A}_2)$  messbare Räume,  $\mu$  ein komplexwertiges Maß auf  $X_1$  und  $f : X_1 \rightarrow X_2$  messbar. Das *Bildmaß* bzw. der *Pushforward* von  $\mu$  unter  $f$  ist das Maß  $f_{\#}\mu$  auf  $X_2$  definiert durch

$$f_{\#}\mu(E) := \mu(f^{-1}(E)), \quad E \in \mathcal{A}_2.$$

---

**1.8 Satz** ([4, Abschnitt 3.6]) Es seien  $(X_1, \mathcal{A}_1)$ ,  $(X_2, \mathcal{A}_2)$  messbare Räume,  $\mu$  ein komplexwertiges Maß auf  $X_1$ ,  $f : X_1 \rightarrow X_2$  messbar und  $g \circ f \in \mathcal{L}(\mu)$ . Dann ist  $g \in \mathcal{L}^1(f_{\#}\mu)$  und es gilt

$$\int_{X_2} g \, d f_{\#}\mu = \int_{X_1} g \circ f \, d\mu.$$

**1.9 Lemma** ([20, Thm. 1.39]) Es sei  $\lambda$  ein positives Maß,  $A \in \mathcal{A}$  und  $f : X \rightarrow [0, \infty]$  messbar mit  $\int_A f \, d\lambda = 0$ . Dann ist  $f = 0$   $\lambda$ -fast überall auf  $A$ .

*Beweis.* Sei  $E_n := \{x \in A \mid f(x) > 1/n\}$  für  $n \in \mathbb{N}$ , dann gilt

$$\frac{1}{n} \lambda(E_n) \leq \int_{E_n} f \, d\lambda \leq \int_A f \, d\lambda = 0,$$

d.h.  $\lambda(E_n) = 0$  für alle  $n \in \mathbb{N}$ . Nun ist

$$\lambda(\{x \in A \mid f(x) > 0\}) = \lambda\left(\bigcup_{n \in \mathbb{N}} E_n\right) \leq \sum_{n=1}^{\infty} \lambda(E_n) = 0$$

und somit gilt  $f = 0$   $\lambda$ -fast überall auf  $A$ . □

Aus diesem klassischen Lemma über positive Maße lässt sich nun eine Aussage über komplexwertige Maße folgern, welche später zum Beweis der Eindeutigkeit bei der Rekonstruktion von Maßen nützlich sein wird.

**1.10 Korollar** (vgl. [10, Lem. A.1]) Es sei  $\mu$  ein komplexwertiges Maß und  $f \in \mathcal{L}^1(\mu)$  mit  $|f(x)| \leq 1$  für alle  $x \in X$ . Es gelte

$$\int_X f \, d\mu = \|\mu\|_{\text{TV}}.$$

Dann ist  $\mu$  konzentriert auf  $\{x \in X \mid |f(x)| = 1\}$ .

*Beweis.* Nach Polarzerlegung von  $\mu$  existiert eine messbare Funktion  $h : X \rightarrow \mathbb{C}$  mit Betrag 1, sodass gilt

$$\int_X 1 - fh \, d|\mu| = \|\mu\|_{\text{TV}} - \int_X f \, d\mu = 0.$$

Insbesondere muss auch das Integral über den Realteil verschwinden:

$$\int_X 1 - \operatorname{Re}(fh) \, d|\mu| = 0$$

Es gilt für alle  $x \in X$

$$1 - \operatorname{Re}(f(x)h(x)) \geq 1 - |f(x)h(x)| = 1 - |f(x)| \geq 0.$$

Nach Lemma 1.9 folgt  $1 - \operatorname{Re}(fh) = 0$   $|\mu|$ -fast überall, also nach obiger Ungleichungskette auch  $1 = |f(x)|$  für  $|\mu|$ -fast alle  $x \in X$ .  $\square$

Schließlich sollen noch zwei Rechenregeln für die Totale Variation bewiesen werden.

**1.11 Lemma** Sind  $\mu, \nu$  komplexwertige Maße, so gilt:

1.  $\mu \perp \nu \implies |\mu + \nu| = |\mu| + |\nu|$
2.  $\sqrt{\|\operatorname{Re} \mu\|_{\text{TV}}^2 + \|\operatorname{Im} \mu\|_{\text{TV}}^2} \leq \|\mu\|_{\text{TV}}$

*Beweis.* Zu Punkt 1: Es sei  $E \in \mathcal{A}$  beliebig. Die Ungleichung  $|\mu + \nu|(E) \leq |\mu|(E) + |\nu|(E)$  folgt direkt aus der Dreiecksungleichung für den Betrag. Für die umgekehrte Ungleichung sei  $\varepsilon > 0$  beliebig und  $(E_i)_{i \in \mathbb{N}}$  eine Partition von  $E$  mit

$$|\mu|(E) \leq \sum_{i=1}^{\infty} |\mu(E_i)| + \varepsilon \quad \text{und} \quad |\nu|(E) \leq \sum_{i=1}^{\infty} |\nu(E_i)| + \varepsilon.$$

Weiter sei  $A \in \mathcal{A}$  mit  $|\mu|(A^c) = 0 = |\nu|(A)$ . Definiere eine feinere Partition  $(F_i)_{i \in \mathbb{N}}$  von  $E$  über

$$F_i := \begin{cases} E_{i/2} \cap A & \text{falls } i \text{ gerade,} \\ E_{(i+1)/2} \cap A^c & \text{falls } i \text{ ungerade} \end{cases}$$

für  $i \in \mathbb{N}$ . Dann gilt aufgrund der Definition der Totalen Variation als das Supremum über alle Partitionen die Ungleichung

$$|\mu + \nu|(E) \geq \sum_{i=1}^{\infty} |(\mu + \nu)(F_i)|.$$

Für jedes gerade  $i \in \mathbb{N}$  ist  $|\nu(F_i)| \leq |\nu|(F_i) \leq |\nu|(A) = 0$  und analog ist  $\mu(F_i) = 0$  für alle ungeraden  $i$ , sodass

$$\sum_{i=1}^{\infty} |(\mu + \nu)(F_i)| = \sum_{i=1}^{\infty} |\mu(E_i)| + \sum_{i=1}^{\infty} |\nu(E_i)| \geq |\mu|(E) + |\nu|(E) - 2\varepsilon.$$

Da  $\varepsilon$  beliebig war, gilt insgesamt  $|\mu + \nu|(E) \geq |\mu|(E) + |\nu|(E)$ .

Zu Punkt 2: Es sei  $\varepsilon > 0$  beliebig. Wähle eine Partition  $(E_i)_{i \in \mathbb{N}}$  von  $X$  mit

$$\|\operatorname{Re} \mu\|_{\text{TV}} \leq \sum_{i=1}^{\infty} |\operatorname{Re} \mu(E_i)| + \varepsilon, \quad \|\operatorname{Im} \mu\|_{\text{TV}} \leq \sum_{i=1}^{\infty} |\operatorname{Im} \mu(E_i)| + \varepsilon$$

und schreibe kurz

$$\begin{aligned} a_i &:= |\operatorname{Re} \mu(E_i)|, & b_i &:= |\operatorname{Im} \mu(E_i)|, \\ s_1 &:= \sum_{i=1}^{\infty} a_i, & s_2 &:= \sum_{i=1}^{\infty} b_i. \end{aligned}$$

Es gilt

$$\sqrt{\|\operatorname{Re} \mu\|_{\text{TV}}^2 + \|\operatorname{Im} \mu\|_{\text{TV}}^2} \leq \sqrt{(s_1 + \varepsilon)^2 + (s_2 + \varepsilon)^2}$$

und die Dreiecksungleichung für die 2-Norm angewendet auf die Vektoren  $(s_1, s_2)$  und  $(\varepsilon, \varepsilon)$  in  $\mathbb{R}^2$  liefert

$$\sqrt{(s_1 + \varepsilon)^2 + (s_2 + \varepsilon)^2} \leq \|(s_1, s_2)\|_2 + \|(\varepsilon, \varepsilon)\|_2 = \|(s_1, s_2)\|_2 + \sqrt{2}\varepsilon$$

Weiter gilt mithilfe der Dreiecksungleichung für Reihen

$$\|(s_1, s_2)\| = \left\| \sum_{i=1}^{\infty} (a_i, b_i) \right\|_2 \leq \sum_{i=1}^{\infty} \|(a_i, b_i)\|_2.$$

Nun ist aber

$$\sum_{i=1}^{\infty} \|(a_i, b_i)\|_2 = \sum_{i=1}^{\infty} |\mu(E_i)| \leq \|\mu\|_{\text{TV}}$$

und wir gelangen insgesamt zu der Ungleichung

$$\sqrt{\|\operatorname{Re} \mu\|_{\text{TV}}^2 + \|\operatorname{Im} \mu\|_{\text{TV}}^2} \leq \|\mu\|_{\text{TV}} + \sqrt{2}\varepsilon.$$

Die Behauptung folgt, da  $\varepsilon$  beliebig gewählt war.  $\square$

## 1.2. Borelmaße und der Darstellungssatz von Riesz

Bisher wurden Maße auf einem allgemeinen messbaren Raum betrachtet. Nun spezialisieren wir uns auf den Fall, dass  $X$  ein topologischer Hausdorffraum mit einer lokalkompakten Topologie ist. Es sei  $\mathcal{A} = \mathcal{B}(X)$  die zugehörige Borel- $\sigma$ -Algebra, d.h. die kleinste  $\sigma$ -Algebra auf  $X$ , die alle offenen Mengen enthält. Maße auf  $\mathcal{B}(X)$  werden *Borelmaße* genannt.

**1.12 Definition** Es sei  $\mu$  ein komplexwertiges Borelmaß auf  $X$ . Definiere den *Träger* von  $\mu$  als

$$\operatorname{supp}(\mu) := \{x \in X \mid \text{für alle offenen Umgebungen } U \text{ von } x \text{ gilt } |\mu|(U) > 0\}.$$

**1.13 Lemma** (vgl. [19, Ch. 2, Thm. 2.1])  $X$  habe eine abzählbare Basis der Topologie. Dann ist  $\operatorname{supp}(\mu) \subset X$  abgeschlossen und  $\mu$  ist konzentriert auf  $\operatorname{supp}(\mu)$ .

*Beweis.* Wähle eine abzählbare Basis  $B$  der Topologie auf  $X$ . Für jedes  $x \in \text{supp}(\mu)^c$  existiert eine offene Umgebung  $V$ , sodass  $|\mu|(V) = 0$ . Also existiert auch  $U \in B$  mit  $x \in U \subset V$  und es gilt  $|\mu|(U) \leq |\mu|(V) = 0$ .

Umgekehrt ist jede offene Menge  $U \in B$  mit  $|\mu|(U) = 0$  in  $\text{supp}(\mu)^c$  enthalten. Es gilt also die Darstellung

$$\text{supp}(\mu)^c = \bigcup \{U \in B \mid |\mu|(U) = 0\}$$

und  $\text{supp}(\mu)^c$  ist als Vereinigung von offenen Mengen offen. Da  $B$  abzählbar ist, handelt es sich hierbei außerdem um eine abzählbare Vereinigung von Nullmengen von  $|\mu|$ , sodass auch  $\text{supp}(\mu)^c$  eine  $|\mu|$ -Nullmenge sein muss.  $\square$

**1.14 Definition** ([20, Def. 3.16]) Eine stetige Funktion  $f : X \rightarrow \mathbb{C}$  *verschwindet im Unendlichen*, falls für jedes  $\varepsilon > 0$  eine kompakte Menge  $K \subset X$  existiert, sodass  $|f(x)| < \varepsilon$  für alle  $x \in X \setminus K$ . Bezeichne den komplexen Vektorraum dieser Funktionen mit  $C_0(X)$ .

**1.15 Satz** (vgl. [20, Thm. 3.17]) Der Raum  $C_0(X)$  wird durch die Supremumsnorm

$$\|f\|_\infty := \sup_{x \in X} |f(x)|$$

zu einem Banachraum.

*Beweis.* Zeige nur die Vollständigkeit: Es sei  $(f_n)_{n \in \mathbb{N}}$  eine Cauchy-Folge in  $C_0(X)$ . Dann ist insbesondere für jedes  $x \in X$  die Folge  $(f_n(x))_{n \in \mathbb{N}}$  eine Cauchy-Folge in  $\mathbb{C}$  und da  $\mathbb{C}$  vollständig ist, konvergiert  $(f_n)$  punktweise gegen eine Grenzfunktion  $f : X \rightarrow \mathbb{C}$ .

Für beliebiges  $\varepsilon > 0$  existiert  $N \in \mathbb{N}$ , sodass  $\|f_n - f_m\|_\infty < \varepsilon$  für alle  $n, m \geq N$ . Es folgt für alle  $n \geq N$  und alle  $x \in X$

$$|f_n(x) - f(x)| = \lim_{m \rightarrow \infty} |f_n(x) - f_m(x)| \leq \varepsilon,$$

d.h.  $\|f_n - f\|_\infty \leq \varepsilon$  für alle  $n \geq N$ . Die Konvergenz gegen  $f$  ist also sogar gleichmäßig.

Nach einem klassischen Resultat aus der Analysis („ $\frac{\varepsilon}{3}$ -Beweis“) ist die Grenzfunktion aufgrund der gleichmäßigen Konvergenz wieder stetig. Noch zu zeigen ist, dass  $f$  im Unendlichen verschwindet.

Sei wieder  $\varepsilon > 0$  beliebig. Nach dem vorherigen Schritt existiert ein  $n \in \mathbb{N}$ , sodass  $\|f_n - f\|_\infty < \varepsilon/2$ . Da  $f_n$  ein Element von  $C_0(X)$  ist, existiert außerdem eine kompakte Menge  $K \subset X$ , sodass  $|f_n(x)| < \varepsilon/2$  für alle  $x \in X \setminus K$ . Insgesamt folgt

$$|f(x)| \leq |f_n(x)| + |f_n(x) - f(x)| < \varepsilon$$

für alle  $x \in X \setminus K$ .  $\square$

Die Funktionen in  $C_0(X)$  sind stetig, also Borel-messbar, und außerdem be-

---

schränkt, sodass ihr Integral bezüglich eines beliebigen komplexwertigen Borelmaßes  $\mu$  wohldefiniert ist. Die Vorschrift

$$T_\mu : C_0(X) \rightarrow \mathbb{C}, \quad f \mapsto \int_X f \, d\mu$$

definiert eine lineare Abbildung, welche aufgrund der Abschätzung

$$|T_\mu f| \leq \int_X |f| \, d|\mu| \leq \|f\|_\infty \|\mu\|_{\text{TV}} \quad (1.1)$$

beschränkt, also Element des Dualraumes

$$C_0(X)' := \{\Phi : C_0(X) \rightarrow \mathbb{C} \mid \Phi \text{ ist stetig}\}$$

ist.

Interessant ist nun die Umkehrung dieser Beobachtung: Lässt sich *jedes* Funktional in  $C_0(X)'$  durch ein Borelmaß auf obige Weise ausdrücken? Diese Frage wird vom Rieszschen Darstellungssatz beantwortet. Damit sogar eine *eindeutige* Zuordnung zwischen Maßen und Funktionalen aufgeschrieben werden kann, muss die Menge der betrachteten Maße verkleinert werden, was durch folgende Definition geschieht:

**1.16 Definition** ([20, Def. 2.15]) Ein positives Borelmaß  $\lambda$  heißt *regulär*, falls für alle  $E \in \mathcal{B}(X)$  gilt

$$\lambda(E) = \inf\{\lambda(V) \mid E \subset V, V \text{ offen}\} \quad (\text{Regularität von außen})$$

und

$$\lambda(E) = \sup\{\lambda(K) \mid K \subset E, K \text{ kompakt}\}. \quad (\text{Regularität von innen})$$

Ein komplexwertiges Borelmaß  $\mu$  heißt *regulär*, wenn  $|\mu|$  regulär ist.

Bezeichne die Menge der regulären komplexwertigen Borelmaße auf  $X$  mit  $\mathcal{M}(X, \mathbb{C})$  bzw. kürzer mit  $\mathcal{M}(X)$ . Die Menge der regulären reellwertigen Borelmaße wird mit  $\mathcal{M}(X, \mathbb{R})$  notiert. Die Teilmenge  $\mathcal{M}^+(X) \subset \mathcal{M}(X, \mathbb{R})$  enthält diejenigen Maße, welche zusätzlich nichtnegativ sind.

Die Menge  $\mathcal{M}(X)$  wird auch die Menge der (endlichen) *Radonmaße* genannt. Es lässt sich zeigen, dass  $\mathcal{M}(X)$  ein komplexer Vektorraum ist [15, Prop. 7.16].

Nun kommen wir zum zentralen Satz dieses Abschnittes, welcher den Zusammenhang zwischen Maßen und stetigen Funktionen herstellt.

**1.17 Satz** (Darstellungssatz von Riesz, [20, Thm. 6.19]) Es sei  $X$  ein lokalkompakter Hausdorffraum und  $\Phi \in C_0(X)'$  ein stetiges lineares Funktional. Dann existiert ein eindeutiges komplexwertiges reguläres Borelmaß  $\mu$ , sodass

$$\Phi(f) = \int_X f \, d\mu$$

für alle  $f \in C_0(X)$ . Außerdem gilt dann für die Operatornorm von  $\Phi$

$$\|\Phi\|_{C_0(X)'} = \|\mu\|_{\text{TV}}.$$

**1.18 Korollar** Die Abbildung

$$\mathcal{M}(X) \rightarrow C_0(X)', \quad \mu \mapsto T_\mu \text{ mit } T_\mu f = \int_X f \, d\mu$$

ist ein isometrischer Isomorphismus von normierten Räumen. Außerdem ist der Raum  $(\mathcal{M}(X), \|\cdot\|_{\text{TV}})$  ein Banachraum. Die TV-Norm eines komplexwertigen Borelmaßes  $\mu$  lässt sich ausdrücken als

$$\|\mu\|_{\text{TV}} = \sup \left\{ \left| \int_X f \, d\mu \right| \mid f \in C_0(X), \|f\|_\infty \leq 1 \right\}.$$

*Beweis.* Die Abbildung ist wohldefiniert nach der Abschätzung in Gleichung (1.1), linear aufgrund der Linearität des Integraloperators als Funktion des Maßes und bijektiv und isometrisch nach dem Darstellungssatz von Riesz.

Wie aus der Funktionalanalysis bekannt ist, ist der Dualraum eines normierten Raumes vollständig, sodass  $C_0(X)'$  ein Banachraum ist. Da  $(\mathcal{M}(X), \|\cdot\|_{\text{TV}})$  isometrisch isomorph zu  $C_0(X)'$  ist, muss auch  $(\mathcal{M}(X), \|\cdot\|_{\text{TV}})$  vollständig sein.

Die Charakterisierung der TV-Norm folgt direkt aus der Isometrie, da der Ausdruck auf der rechten Seite der Gleichung lediglich die Operatornorm des Funktionals  $T_\mu$  ist.  $\square$

Die Regularitätsbedingungen aus Definition 1.16 sind sehr technisch, was der Allgemeinheit des Raumes  $X$  geschuldet ist. Unter einer zusätzlichen Annahme an die Topologie auf  $X$  ist die Situation deutlich einfacher: Falls jede offene Menge in  $X$   $\sigma$ -kompakt ist, d.h. sich als abzählbare Vereinigung kompakter Mengen schreiben lässt, so gilt folgende Vereinfachung:

**1.19 Satz** ([20, Thm. 2.18]) Es sei  $X$  ein lokalkompakter Hausdorffraum, in welchem jede offene Menge  $\sigma$ -kompakt ist. Es sei  $\lambda$  ein positives Borelmaß auf  $X$ , sodass  $\lambda(K) < \infty$  für jede kompakte Menge  $K$ .

Dann ist  $\lambda$  regulär.

Mitunter ist in dieser Situation jedes komplexwertige Borelmaß  $\mu$  regulär, da  $|\mu|$  immer ein endliches Maß ist. Im Rest dieser Arbeit werden wir fast ausschließlich Grundräume  $X = \mathbb{R}^d$  bzw.  $X = \Omega$  für kompakte Teilmengen  $\Omega \subset \mathbb{R}^d$  betrachten. Jede offene Teilmenge lässt sich als abzählbare Vereinigung von kompakten Bällen schreiben (nutze die Dichtheit von  $\mathbb{Q}^d$ ), sodass die obige Vereinfachung greift. Außerdem gilt im Fall eines kompakten Grundraums  $\Omega$ , dass  $C_0(\Omega) = C(\Omega)$ , d.h. jede stetige Funktion verschwindet im Unendlichen.

Schließlich wollen wir noch die Fourier-Transformation für Maße einführen. Wir werden diese später sowohl für Maße auf dem Würfel  $[0, 1]^d$  als auch für Maße auf  $\mathbb{R}^d$  benötigen, sodass wir zur Unterscheidung letzere mit  $\tilde{\mathcal{F}}$  bezeichnen.

---

**1.20 Definition** Definiere die Fourier-Transformation eines Maßes  $u \in \mathcal{M}([0, 1]^d)$  durch

$$\mathcal{F}u := \left( \int_{[0,1]^d} e^{-2\pi i(l \cdot z)} \, d u(z) \right)_{l \in \mathbb{Z}^d} .$$

Die Fourier-Transformation eines Maßes  $\mu \in \mathcal{M}(\mathbb{R}^d)$  sei durch

$$\tilde{\mathcal{F}}\mu : \mathbb{R}^d \rightarrow \mathbb{C}, \quad (\tilde{\mathcal{F}}\mu)(y) := \int_{\mathbb{R}^d} e^{-i(x \cdot y)} \, d\mu(x)$$

definiert. Die Fourier-Transformation einer Funktion  $f \in \mathcal{L}^1([0, 1]^d)$  bzw.  $f \in \mathcal{L}^1(\mathbb{R}^d)$  sei die Fourier-Transformation des Maßes  $f \, d\ell$  mit Dichte  $f$  bezüglich des Lebesgue-Maßes.

**1.21 Proposition** Die Fourier-Transformation auf  $[0, 1]^d$  definiert eine lineare Kontraktion  $\mathcal{F} : \mathcal{M}([0, 1]^d) \rightarrow \ell^\infty(\mathbb{Z}^d)$  in den Raum der beschränkten Folgen auf  $\mathbb{Z}^d$ , versehen mit der Supremumsnorm.

*Beweis.* Die Linearität ist klar. Für alle  $l \in \mathbb{Z}^d$  gilt die Abschätzung

$$|\mathcal{F}u(l)| \leq \int_{[0,1]^d} |e^{-2\pi i(l \cdot z)}| \, d|u|(z) = \|u\|_{\text{TV}},$$

sodass  $\mathcal{F}$  tatsächlich eine Kontraktion ist. □

**1.22 Proposition** Die Fourier-Transformation auf  $\mathbb{R}^d$  definiert eine injektive Abbildung  $\tilde{\mathcal{F}} : \mathcal{M}(\mathbb{R}^d) \rightarrow C(\mathbb{R}^d)$ .

*Beweis.* Das Ergebnis  $\tilde{\mathcal{F}}\mu$  der Fourier-Transformation eines Maßes  $\mu \in \mathcal{M}(\mathbb{R}^d)$  ist stetig nach dem Satz über die majorisierte Konvergenz, da  $|e^{-ix \cdot y}| = 1$  und somit eine Majorante durch die konstante Einsfunktion gegeben ist.

Der Beweis der Injektivität nutzt die Dichtheit der trigonometrischen Polynome nach dem Satz von Weierstrass. Für Details siehe [4, Prop. 3.8.6]. □

## 2. Super-Resolution über konvexe Optimierung

Super-Resolution meint grundsätzlich eine Technik, welche eine bessere Auflösung erlaubt, als nach klassischer Auffassung möglich wäre. Um dies zu erreichen, ist man im Allgemeinen auf *a priori* Informationen angewiesen, d.h. es müssen gewisse Vorinformationen über die gesuchte Lösung bekannt sein. In vielen Anwendungen ist die Annahme erfüllt, dass den Bildern eine Sammlung von Punktquellen zugrundeliegt. Man spricht dann auch von einer dünnbesetzten Repräsentation (engl. *sparsity*).

**2.1 Beispiel (STORM)** In der Fluoreszenzmikroskopie werden interessante Strukturen mit einem fluoreszierenden Stoff gefärbt. Die Fluoreszenz wird durch Laserlicht einer bestimmten Wellenlänge angeregt, sodass der Stoff Licht einer anderen Wellenlänge zurücksendet. Dieses wird schließlich durch ein Mikroskop abgebildet (vgl. [18]).

Die *stochastic optical reconstruction microscopy* (STORM, [21]) ist eine Technik, welche die Abbildung von kleinsten Strukturen mit Größen unterhalb von ca. 200 nm erlaubt, obwohl eine Auflösung durch klassische Lichtmikroskopie aufgrund der Lichtbeugung nicht mehr möglich wäre. Dazu werden Moleküle genutzt, deren Fluoreszenz sich mit zwei Lasern verschiedener Wellenlänge (typischerweise ein grüner und ein roter) an bzw. abschalten lässt. Nun werden viele Bilder der eingefärbten Struktur aufgenommen, sodass bei jedem Bild nur ein zufälliger Bruchteil der Moleküle fluoresziert. Ist die Anzahl der fluoreszierenden Moleküle pro Bild gering genug, so ist mit hoher Wahrscheinlichkeit der Abstand zwischen den Molekülen im Bild groß, sodass diese trotz der Lichtbeugung einzeln aufgelöst werden können. Die Moleküle werden lokalisiert und schließlich werden die Lokalisierungen in den Einzelbildern zu einem Gesamtbild zusammengefügt.

Wir wählen eine kompakte Menge  $\Omega \subset \mathbb{R}^d$ , welche das Gebiet repräsentiert, auf dem das Bild aufgenommen wird. Weiter gehen wir davon aus, dass das Stück Realität, welches abgebildet werden soll, durch ein Radonmaß  $\hat{u} \in \mathcal{M}(\Omega)$  modelliert werden kann. Der Raum der Radonmaße bietet viel Flexibilität in der Modellierung, da sowohl Punktquellen als auch kontinuierliche Intensitätsverteilungen dargestellt werden können.

Die Bildaufnahme (oder allgemeiner: der Messprozess) wird nun durch einen Operator

$$\mathcal{G} : \mathcal{M}(\Omega) \rightarrow H$$

---

formalisiert, wobei  $H$  ein komplexer Hilbertraum mit Skalarprodukt  $\langle \cdot, \cdot \rangle$  sei. Später werden wir uns auf den endlichdimensionalen Fall  $H = \mathbb{C}^m$  spezialisieren, welcher sich z.B. nach Diskretisierung eines allgemeineren Messoperators ergibt.

Vom Zielmaß  $\hat{u}$  wird nun eine Messung

$$\hat{y} = \mathcal{G}\hat{u}$$

aufgenommen und die Aufgabe lautet,  $\hat{u}$  aus dieser Messung zu rekonstruieren. Hierbei handelt es sich um ein sogenanntes *Inverses Problem* und der Operator  $\mathcal{G}$  wird in diesem Zusammenhang auch der *Vorwärtsoperator* genannt.

Die Annahme der Dünnbesetztheit führt nun auf eine Linearkombination von  $N \in \mathbb{N}$  Dirac-Maßen

$$\hat{u} = \sum_{i=1}^N \alpha_i \delta_{z_i} \quad (2.1)$$

mit  $z_1, \dots, z_N \in \Omega$  und einem Gewichtsvektor  $\alpha \in \mathbb{K}^N$  für  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{C}$ .

Wie erhält man ein effizientes Verfahren zur Rekonstruktion von  $\hat{u}$  aus den Daten  $\hat{y}$ ? Im Artikel „Towards a Mathematical Theory of Super-resolution“ [9] schlagen Candès et al. vor, das Minimierungsproblem

$$\min_{u \in \mathcal{M}(\Omega, \mathbb{K})} \|u\|_{\text{TV}} \quad \text{s.t.} \quad \mathcal{G}u = \hat{y} \quad (\text{ER}_{\mathbb{K}})$$

zu betrachten. Hier ist  $\|\cdot\|_{\text{TV}}$  die Totale-Variations-Norm aus Kapitel 1 und die Nebenbedingung  $\mathcal{G}u = \hat{y}$  erzwingt Konsistenz der Lösung mit den Daten. Dieses Problem hat eine wichtige Eigenschaft: Da  $\|\cdot\|_{\text{TV}}$  eine Norm und die Nebenbedingung linear ist, ist  $(\text{ER}_{\mathbb{K}})$  ein konvexes Minimierungsproblem. Dies ist wichtig, da sich aus konvexen Minimierungsproblemen häufig effiziente numerische Lösungsverfahren gewinnen lassen, wie auch das letzte Kapitel dieser Arbeit zeigt.

Der Rest dieses Kapitels ist wie folgt strukturiert: Zunächst werden wir genauer auf den Messoperator  $\mathcal{G}$  eingehen und einige Beispiele diskutieren. Anschließend wird ein duales Problem zu  $(\text{ER}_{\mathbb{K}})$  hergeleitet und gleichzeitig die Existenz von Lösungen bewiesen. In Abschnitt 2.3 beschäftigen wir uns mit Bedingungen, welche garantieren, dass  $\hat{u}$  die eindeutige Lösung von  $(\text{ER}_{\mathbb{K}})$  ist. Die hierfür benötigten dualen Zertifikate werden schließlich in Abschnitt 2.4 konstruiert.

## 2.1. Der Messoperator

Der Operator  $\mathcal{G}$  modelliert die Messung der Daten  $\hat{y}$  ausgehend von dem Zielmaß  $\hat{u}$ . Die genaue Form hängt in der Anwendung von der Beschaffenheit des konkreten Bildgebungsprozesses ab. Einige mögliche Messoperatoren sollen im Folgenden kurz untersucht werden.

Wir treffen jedoch zunächst eine Annahme an die Regularität von  $\mathcal{G}$ , welche später für die mathematische Behandlung benötigt wird: Wir nehmen an, dass zu

$\mathcal{G} : \mathcal{M}(\Omega) \rightarrow H$  ein Operator

$$\mathcal{G}^* : H \rightarrow C(\Omega)$$

existiert, sodass für alle  $c \in H$  und  $u \in \mathcal{M}(\Omega)$  die Gleichung

$$\langle c, \mathcal{G}u \rangle = \int_{\Omega} \mathcal{G}^* c \, d\bar{u} \quad (2.2)$$

erfüllt ist. Der Operator  $\mathcal{G}^*$  wird in diesem Zusammenhang der *prädual* Operator zu  $\mathcal{G}$  genannt (vgl. [8]).

### 2.1.1. Faltung

Für einen stetigen, beschränkten Faltungskern  $\varphi : \mathbb{R}^d \rightarrow \mathbb{C}$  und ein Maß  $u \in \mathcal{M}(\Omega)$  sei die Faltung  $\varphi * u : \Omega \rightarrow \mathbb{C}$  definiert durch

$$(\varphi * u)(t) := \int_{\Omega} \varphi(t - z) \, du(z).$$

Für die Anwendung in der Mikroskopie wäre  $\varphi$  typischerweise die sog. *point spread function* (PSF) des Mikroskops.

Nach dem Satz von Fubini ist  $\varphi * u$  messbar. Es gilt für alle  $t \in \Omega$

$$|(\varphi * u)(t)| = \left| \int_{\Omega} \varphi(t - z) \, du(z) \right| \leq \int_{\Omega} |\varphi(t - z)| \, d|u|(z) \leq \|\varphi\|_{\infty} \|u\|_{\text{TV}},$$

d.h. das Ergebnis der Faltung ist beschränkt. Aufgrund der Kompaktheit von  $\Omega$  ist das Lebesgue-Maß  $\ell(\Omega)$  endlich und wir können einfach abschätzen

$$\|\varphi * u\|_{L^2(\Omega)} = \left( \int_{\Omega} |\varphi * u|^2 \, dt \right)^{1/2} \leq \sqrt{\ell(\Omega)} \|\varphi\|_{\infty} \|u\|_{\text{TV}}, \quad (2.3)$$

sodass  $\varphi * u \in L^2(\Omega)$ .

Der Messoperator sei nun gegeben durch

$$\mathcal{G} : \mathcal{M}(\Omega) \rightarrow L^2(\Omega), \quad \mathcal{G}u = \varphi * u.$$

Dieser ist beschränkt nach der Rechnung in Gleichung (2.3). Für alle  $c \in L^2(\Omega)$  und  $u \in \mathcal{M}(\Omega)$  gilt nach dem Satz von Fubini

$$\begin{aligned} \langle c, \mathcal{G}u \rangle_{L^2} &= \int_{\Omega} c(t) \overline{(\varphi * u)(t)} \, dt = \int_{\Omega} c(t) \int_{\Omega} \overline{\varphi(t - z)} \, d\bar{u}(z) \, dt \\ &= \int_{\Omega} \int_{\Omega} c(t) \overline{\varphi(t - z)} \, dt \, d\bar{u}(z), \end{aligned}$$

sodass der präduale Operator zu  $\mathcal{G}$  gegeben ist durch  $\mathcal{G}^* : L^2(\Omega) \rightarrow C(\Omega)$ ,

$$(\mathcal{G}^*c)(z) = \int_{\Omega} c(t) \overline{\varphi(t-z)} dt, \quad z \in \Omega.$$

Hierbei ist  $\mathcal{G}^*c$  tatsächlich stetig nach dem Satz über die majorisierte Konvergenz, da  $\varphi$  stetig ist und eine integrierbare Majorante von  $(t, z) \mapsto c(t) \overline{\varphi(t-z)}$  gegeben ist durch  $t \mapsto |c(t)| \cdot \|\varphi\|_{\infty}$ .

### 2.1.2. Faltung nach Radon-Transformation

Medizinische Bildgebungsverfahren wie die Computertomographie und die Positronen-Emissions-Tomographie lassen sich durch die Radon-Transformation modellieren. Hierfür nutzen wir einen Faltungskern  $\psi \in C_c(\mathbb{R})$ , also eine stetige Abbildung  $\mathbb{R} \rightarrow \mathbb{C}$  mit kompaktem Träger, und betrachten den Operator  $\mathcal{G}$  auf  $\mathcal{M}(\Omega)$  definiert durch

$$(\mathcal{G}u)(\theta, s) := \int_{\Omega} \psi(\theta \cdot z - s) du(z), \quad \theta \in \mathbb{S}^{d-1}, s \in \mathbb{R}.$$

Wegen der Kompaktheit von  $\Omega$  und des Trägers von  $\psi$  existiert  $R > 0$  mit  $\Omega \cup \text{supp } \psi \subset B_{R/2}(0)$ . Ist  $s \in \mathbb{R}$  mit  $|s| > R$ , so gilt für alle  $z \in \Omega$  und alle  $\theta \in \mathbb{S}^{d-1}$

$$|\theta \cdot z - s| \geq |s| - |\theta \cdot z| > R - \|z\| > \frac{R}{2}.$$

Hieraus folgt  $\text{supp}(\mathcal{G}u) \subset \mathbb{S}^{d-1} \times [-R, R]$ . Wir nutzen die Notation  $\chi_A$  für die Indikatorfunktion einer Menge  $A$  und erhalten eine Majorante von  $\mathcal{G}u$  über die Abschätzung

$$|\mathcal{G}u(\theta, s)| \leq \int_{\Omega} |\psi(\theta \cdot z - s)| d|u|(z) \leq \|\psi\|_{\infty} \|u\|_{\text{TV}} \chi_{[-R, R]}(s).$$

Bezeichne mit  $L^2(\mathbb{S}^{d-1} \times \mathbb{R})$  den Lebesgue-Raum des Produktmaßes  $\omega_{d-1} \otimes \ell$  aus dem normierten Oberflächenmaß  $\omega_{d-1}$  auf  $\mathbb{S}^{d-1}$  und dem Lebesgue-Maß auf  $\mathbb{R}$ . Wir behaupten, dass  $\mathcal{G}$  nach  $L^2(\mathbb{S}^{d-1} \times \mathbb{R})$  abbildet. Es gilt nämlich

$$\begin{aligned} \|\mathcal{G}u\|_{L^2(\mathbb{S}^{d-1} \times \mathbb{R})} &\leq \|\psi\|_{\infty} \|u\|_{\text{TV}} \left( \int_{\mathbb{S}^{d-1} \times \mathbb{R}} \chi_{[-R, R]}(s) d\omega_{d-1} \otimes \ell(\theta, s) \right)^{1/2} \\ &= \sqrt{2R} \|\psi\|_{\infty} \|u\|_{\text{TV}}. \end{aligned}$$

Dies zeigt auch, dass  $\mathcal{G}$  ein beschränkter Operator ist.

Wie sieht der präduale Operator zu  $\mathcal{G}$  aus? Für  $c \in L^2(\mathbb{S}^{d-1} \times \mathbb{R})$ ,  $u \in \mathcal{M}(\Omega)$

können wir rechnen

$$\begin{aligned} \langle c, \mathcal{G}u \rangle_{L^2(\mathbb{S}^{d-1} \times \mathbb{R})} &= \int_{\mathbb{S}^{d-1} \times \mathbb{R}} c(\theta, s) \overline{\mathcal{G}u(\theta, s)} \, d\omega_{d-1} \otimes \ell(\theta, s) \\ &= \int_{\mathbb{S}^{d-1} \times \mathbb{R}} c(\theta, s) \int_{\Omega} \overline{\psi(\theta \cdot z - s)} \, d\bar{u}(z) \, d\omega_{d-1} \otimes \ell(\theta, s). \end{aligned} \quad (2.4)$$

Es gilt

$$\begin{aligned} \int_{\mathbb{S}^{d-1} \times \mathbb{R}} \int_{\Omega} |c(\theta, s) \psi(\theta \cdot z - s)| \, d|u|(z) \, d\omega_{d-1} \otimes \ell(\theta, s) \\ \leq \|\psi\|_{\infty} \|u\|_{\text{TV}} \int_{\mathbb{S}^{d-1} \times [-R, R]} |c(\theta, s)| \, d\omega_{d-1} \otimes \ell(\theta, s) < \infty, \end{aligned}$$

sodass wir in Gleichung (2.4) den Satz von Fubini anwenden dürfen. Dies bringt uns auf den Ausdruck

$$\langle c, \mathcal{G}u \rangle_{L^2(\mathbb{S}^{d-1} \times \mathbb{R})} = \int_{\Omega} \int_{\mathbb{S}^{d-1} \times \mathbb{R}} c(\theta, s) \overline{\psi(\theta \cdot z - s)} \, d\omega_{d-1} \otimes \ell(\theta, s) \, d\bar{u}(z).$$

Der präduale Operator  $\mathcal{G}^* : L^2(\mathbb{S}^{d-1} \times \mathbb{R}) \rightarrow C(\Omega)$ ,

$$(\mathcal{G}^* c)(z) = \int_{\mathbb{S}^{d-1} \times \mathbb{R}} c(\theta, s) \overline{\psi(\theta \cdot z - s)} \, d\omega_{d-1} \otimes \ell(\theta, s), \quad z \in \Omega$$

ist wohldefiniert nach dem Satz über die majorisierte Konvergenz: Die Abbildung  $\psi$  ist stetig und eine von  $z \in \Omega$  unabhängige Majorante des Integranden ist gegeben durch  $(\theta, s) \mapsto \|\psi\|_{\infty} |c(\theta, s)| \chi_{[-R, R]}(s)$ .

### 2.1.3. Der ideale Frequenzfilter

Die Aussagen in diesem Abschnitt sind angelehnt an [17, Ch. 3]. Es sei  $\mathbb{T}^d = \mathbb{R}^d / \mathbb{Z}^d$  der  $d$ -dimensionale Torus und  $C(\mathbb{T}^d)$  die Menge der stetigen, komplexwertigen Funktionen hierauf. Wir identifizieren diese mit den stetigen Abbildungen  $\mathbb{R}^d \rightarrow \mathbb{C}$ , welche in jeder Komponente Periode 1 haben und nutzen die Fourier-Transformation  $\mathcal{F}$  auf  $[0, 1]^d$  gemäß Definition 1.20.

**2.2 Lemma** Es sei  $\varphi \in C(\mathbb{T}^d)$  und  $u \in \mathcal{M}([0, 1]^d)$ . Dann gilt

$$(\mathcal{F}(\varphi * u))(l) = (\mathcal{F}\varphi)(l) \cdot (\mathcal{F}u)(l)$$

für alle  $l \in \mathbb{Z}^d$ .

*Beweis.* Es sei  $l \in \mathbb{Z}^d$  beliebig. Einsetzen der Definitionen von Faltung und Fourier-Transformation führt zu

$$(\mathcal{F}(\varphi * u))(l) = \int_{[0, 1]^d} e^{-2\pi i(l \cdot t)} \int_{[0, 1]^d} \varphi(t - z) \, du(z) \, dt.$$

---

Die Integrale dürfen nach dem Satz von Fubini vertauscht werden, denn

$$\int_{[0,1]^d} \int_{[0,1]^d} |e^{-2\pi i(l \cdot t)} \varphi(t-z)| \, d|u|(z) \, dt < \infty$$

aufgrund der Beschränktheit der Integranden und der Endlichkeit des Maßes. Wir erhalten

$$(\mathcal{F}(\varphi * u))(l) = \int_{[0,1]^d} \int_{[0,1]^d} e^{-2\pi i l \cdot (t-z)} \varphi(t-z) \, dt \, e^{-2\pi i(l \cdot z)} \, du(z).$$

Da die Abbildung  $t \mapsto e^{-2\pi i(l \cdot t)} \varphi(t)$  Periode 1 hat, gilt für alle  $z \in [0,1]^d$

$$\int_{[0,1]^d} e^{-2\pi i l \cdot (t-z)} \varphi(t-z) \, dt = \int_{[0,1]^d} e^{-2\pi i l \cdot t} \varphi(t) \, dt$$

und es folgt

$$\begin{aligned} (\mathcal{F}(\varphi * u))(l) &= \left( \int_{[0,1]^d} e^{-2\pi i(l \cdot t)} \varphi(t) \, dt \right) \left( \int_{[0,1]^d} e^{-2\pi i(l \cdot z)} \, du(z) \right) \\ &= (\mathcal{F}\varphi)(l) \cdot (\mathcal{F}u)(l). \end{aligned}$$

□

Für Fouriermessungen betrachten wir zu einer Grenzfrequenz  $f_c \in \mathbb{N}$  die Menge der erlaubten ganzzahligen Frequenzen gegeben durch

$$F := \{l \in \mathbb{Z}^d \mid \|l\|_\infty \leq f_c\}.$$

Es sei  $m = |F|$  und  $\Omega = [0,1]^d$ . Wir identifizieren  $\mathbb{C}^F$  mit  $\mathbb{C}^m$  und können so den idealen Frequenzfilter zur Grenzfrequenz  $f_c$  definieren durch  $\mathcal{F}_{f_c} : \mathcal{M}(\Omega) \rightarrow \mathbb{C}^m$ ,

$$\mathcal{F}_{f_c} u := (\mathcal{F}u(l))_{l \in F}.$$

Der Operator ist stetig nach Proposition 1.21.

Der Zusammenhang zur Faltung ist folgender: Sei  $D_{f_c} \in C(\mathbb{T}^d)$ ,

$$D_{f_c}(z) = \sum_{l \in F} e^{2\pi i(l \cdot z)}$$

der sogenannte *Dirichlet-Kern*. Wir berechnen die Fourier-Koeffizienten von  $D_{f_c}$ :

$$(\mathcal{F}D_{f_c})(l') = \int_{[0,1]^d} e^{-2\pi i(l' \cdot z)} \sum_{l \in F} e^{2\pi i(l \cdot z)} \, dz = \sum_{l \in F} \int_{[0,1]^d} e^{2\pi i(l-l') \cdot z} \, dz$$

Durch Aufspaltung in das Produkt der Integrale über die einzelnen Komponenten

kann nachgerechnet werden, dass

$$\int_{[0,1]^d} e^{2\pi i(l-l') \cdot z} dz = \begin{cases} 1 & l = l', \\ 0 & \text{sonst,} \end{cases}$$

sodass

$$(\mathcal{F}D_{f_c})(l') = \begin{cases} 1 & l' \in F, \\ 0 & \text{sonst.} \end{cases}$$

Nach obigem Lemma gilt  $\mathcal{F}(D_{f_c} * u) = (\mathcal{F}D_{f_c}) \cdot (\mathcal{F}u)$ , d.h. nach Faltung mit dem Dirichlet-Kern bleiben im Fourier-Raum genau die niedrigsten Frequenzen von  $u$  bis zur Grenzfrequenz  $f_c$  erhalten.

Der präduale Operator ist gegeben durch  $\mathcal{F}_{f_c}^* : \mathbb{C}^m \cong \mathbb{C}^F \rightarrow C([0,1]^d)$ ,

$$(\mathcal{F}_{f_c}^* c)(z) := \sum_{l \in F} c_l e^{2\pi i(l \cdot z)}.$$

Falls a priori bekannt ist, dass das Zielmaß  $\hat{u}$  reellwertig ist, so ist es natürlich, das Minimierungsproblem für  $\mathbb{K} = \mathbb{R}$  zu betrachten. Allerdings stellt sich die Frage, ob auch Lösungen des Minimierungsproblems mit allgemeinen, komplexwertigen Maßen zwangsläufig reellwertig sind, wenn die Daten  $\hat{y}$  von einem reellwertigen Maß gemessen wurden. Im Fall des idealen Frequenzfilters ist dies positiv zu beantworten.

**2.3 Lemma** Der ideale Frequenzfilter hat die Eigenschaft

$$(\mathcal{F}\bar{u})(l) = \overline{(\mathcal{F}u)(-l)} \quad \forall l \in \mathbb{Z}^d.$$

*Beweis.* Mit der Rechenregel für die komplexe Konjugation des Integrals (Lemma 1.6) können wir dies einfach nachrechnen:

$$\begin{aligned} \overline{(\mathcal{F}u)(-l)} &= \overline{\int_{[0,1]^d} e^{2\pi i(l \cdot z)} du(z)} = \int_{[0,1]^d} \overline{e^{2\pi i(l \cdot z)}} d\bar{u}(z) \\ &= \int_{[0,1]^d} e^{-2\pi i(l \cdot z)} d\bar{u}(z) = (\mathcal{F}\bar{u})(l) \end{aligned}$$

□

**2.4 Satz** (Äquivalenz der Minimierungsprobleme für den idealen Frequenzfilter) Angenommen, das Zielmaß ist reellwertig, d.h.  $\hat{u} \in \mathcal{M}(\Omega, \mathbb{R})$ , und  $\hat{y} = \mathcal{F}_{f_c} \hat{u}$  sind die gemessenen Daten. Dann sind die Lösungsmengen der Minimierungsprobleme

$$\min_{u \in \mathcal{M}(\Omega, \mathbb{R})} \|u\|_{\text{TV}} \quad \text{s.t.} \quad \mathcal{F}_{f_c} u = \hat{y} \quad (\text{ER}_{\mathbb{R}})$$

und

$$\min_{u \in \mathcal{M}(\Omega, \mathbb{C})} \|u\|_{\text{TV}} \quad \text{s.t.} \quad \mathcal{F}_{f_c} u = \hat{y} \quad (\text{ER}_{\mathbb{C}})$$

gleich.

*Beweis.* Da das Zielmaß  $\hat{u}$  reellwertig ist, gilt für die Daten für alle  $l \in F$

$$\hat{y}_l = (\mathcal{F}_{f_c} \hat{u})(l) = (\mathcal{F}_{f_c} \bar{\hat{u}})(l) = \overline{(\mathcal{F}_{f_c} \hat{u})(-l)} = \overline{\hat{y}_{-l}}.$$

Es sei  $u \in \mathcal{M}(\Omega, \mathbb{C})$  eine Lösung von  $\text{ER}_{\mathbb{C}}$ . Da  $u$  die Nebenbedingung  $\mathcal{F}_{f_c} u = \hat{y}$  erfüllt, muss für alle  $l \in F$  gelten

$$(\mathcal{F}_{f_c} \bar{u})(l) = \overline{(\mathcal{F}_{f_c} u)(-l)} = \overline{\hat{y}_{-l}} = \hat{y}_l,$$

d.h.  $\mathcal{F}_{f_c} \bar{u} = \hat{y}$ . Nun wissen wir, dass auch der Realteil  $\text{Re } u$  die Nebenbedingung erfüllt, denn

$$\mathcal{F}_{f_c}(\text{Re } u) = \mathcal{F}_{f_c} \left( \frac{1}{2}(u + \bar{u}) \right) = \frac{1}{2}(\mathcal{F}_{f_c} u + \mathcal{F}_{f_c} \bar{u}) = \hat{y}.$$

Nach Lemma 1.11 gilt die Abschätzung

$$\|\text{Re } u\|_{\text{TV}} \leq \sqrt{\|\text{Re } u\|_{\text{TV}}^2 + \|\text{Im } u\|_{\text{TV}}^2} \leq \|u\|_{\text{TV}}.$$

Wäre  $\text{Im } u \neq 0$ , so wäre die erste Ungleichung strikt, was ein Widerspruch zur Optimalität von  $u$  für  $\text{ER}_{\mathbb{C}}$  wäre. Also ist jede Lösung von  $\text{ER}_{\mathbb{C}}$  reellwertig und somit auch eine Lösung von  $\text{ER}_{\mathbb{R}}$ .

Sei umgekehrt  $u \in \mathcal{M}(\Omega, \mathbb{R})$  eine Lösung von  $\text{ER}_{\mathbb{R}}$ , dann ist  $u$  offensichtlich auch zulässig für  $\text{ER}_{\mathbb{C}}$ . Ist  $u'$  zulässig für  $\text{ER}_{\mathbb{C}}$ , so ist nach obiger Überlegung  $\text{Re } u'$  zulässig für  $\text{ER}_{\mathbb{R}}$  und es gilt  $\|u'\|_{\text{TV}} \geq \|\text{Re } u'\|_{\text{TV}} \geq \|u\|_{\text{TV}}$ . Also ist  $u$  auch eine Lösung des Minimierungsproblems über  $\mathcal{M}(\Omega, \mathbb{C})$ .  $\square$

## 2.2. Duale Struktur und Existenz von Lösungen

Das Problem

$$\min_{u \in \mathcal{M}(\Omega)} \|u\|_{\text{TV}} \quad \text{s.t.} \quad \mathcal{G}u = \hat{y} \tag{ER}$$

ist ein konvexes Minimierungsproblem mit einer Variablen in  $\mathcal{M}(\Omega)$  und einer Gleichheitsnebenbedingung im Hilbertraum  $H$ . Um die Beweisideen für die anschließenden Aussagen zur exakten Rekonstruktion zu motivieren, wollen wir ein duales Problem zu (ER) herleiten. Dieser Abschnitt ist angelehnt an die Überlegungen in [13], wo jedoch nur reellwertige Maße behandelt wurden. Wir wollen uns dagegen mit dem komplexwertigen Fall  $\mathbb{K} = \mathbb{C}$  befassen.

Es bezeichne  $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$  die erweiterten reellen Zahlen.

**2.5 Definition** (Fenchel-Transformation) Es sei  $V$  ein reeller topologischer Vektorraum und  $V'$  der Dualraum. Für eine Abbildung  $f : V \rightarrow \bar{\mathbb{R}}$  sei die Fenchel-

Transformation von  $f$  gegeben durch  $f^* : V' \rightarrow \overline{\mathbb{R}}$ ,

$$f^*(v') := \sup_{v \in V} (v'(v) - f(v)).$$

Folgender Satz ist eine Zusammenfassung der Ergebnisse aus [14, Abschnitt III.4(I)]:

**2.6 Satz** Es seien  $V, W$  reelle lokalkonvexe topologische Vektorräume und  $V', W'$  die zugehörigen Dualräume. Weiter sei  $\Lambda : V \rightarrow W$  eine stetige lineare Abbildung und  $\Lambda^T : W' \rightarrow V'$  die transponierte Abbildung.

Es sei eine Abbildung  $J : V \times W \rightarrow \overline{\mathbb{R}}$  gegeben. Betrachte die Optimierungsprobleme

$$\inf_{v \in V} J(v, \Lambda v) =: p_1 \tag{P_1}$$

und

$$\sup_{w' \in W'} -J^*(\Lambda^T w', -w') =: p_2. \tag{P_2}$$

Es gelte

- 1)  $J$  ist konvex,
- 2)  $p_1$  ist endlich,
- 3) es existiert  $v_0 \in V$ , sodass  $J(v_0, \Lambda v_0) < \infty$  und die Funktion  $w \mapsto J(v_0, w)$  ist stetig im Punkt  $w = \Lambda v_0$ .

Dann folgt

$$p_1 = p_2$$

und das Supremum in  $(P_2)$  wird angenommen.

**2.7 Bemerkung** Der Produktraum  $V \times W$  ist wieder ein topologischer Vektorraum, dessen Dualraum  $(V \times W)'$  mit  $V' \times W'$  identifiziert werden kann über  $f \mapsto (f(\cdot, 0), f(0, \cdot))$ . Also ist  $J^*$  eine Abbildung  $V' \times W' \rightarrow \overline{\mathbb{R}}$ .

**2.8 Bemerkung** Problem  $(P_1)$  wird in diesem Kontext das *primale* und  $(P_2)$  das *duale* Problem genannt. Da  $\mathcal{M}(\Omega)$  nach dem Darstellungssatz von Riesz der Dualraum zu  $C(\Omega)$  ist, würde dies dazu führen, dass unser Ausgangsproblem (ER) die Rolle des dualen Problems einnimmt. Diese Bezeichnung ist aber genau umgekehrt zu der in der Literatur üblichen Bezeichnung. Um die Verwirrung zu vermeiden, verzichten wir deshalb hier zunächst auf die Vokabeln primal und dual und halten uns später an die Konvention aus der Literatur.

Wir wenden den obigen Satz an auf  $V = H, W = C(\Omega)$ . Dazu wollen wir in diesem Abschnitt beide Vektorräume als Vektorräume über  $\mathbb{R}$  auffassen. Der Hilbertraum  $H$  wird zu einem Hilbertraum über  $\mathbb{R}$  mit Skalarprodukt  $\operatorname{Re}\langle \cdot, \cdot \rangle$  und der reelle

Dualraum  $H'^{\mathbb{R}}$  kann hierüber wieder mit  $H$  identifiziert werden. Den Dualraum von  $C(\Omega)$ , aufgefasst als reeller Vektorraum, charakterisieren wir mit folgendem Lemma:

**2.9 Lemma** Fasse  $\mathcal{M}(\Omega)$  und  $C(\Omega)$  als Vektorräume über  $\mathbb{R}$  auf. Die Abbildung

$$\Gamma : \mathcal{M}(\Omega) \rightarrow C(\Omega)'^{\mathbb{R}}, \quad \mu \mapsto \Gamma\mu \text{ mit } (\Gamma\mu)(f) = \operatorname{Re} \int_{\Omega} f \, d\bar{\mu}$$

ist ein isometrischer Isomorphismus von normierten Räumen.

*Beweis.* Siehe Anhang A.1.1. □

Wir können also auch den reellen Dualraum von  $C(\Omega)$  mit  $\mathcal{M}(\Omega)$  identifizieren. Wir sind nun bereit, Satz 2.6 auf unser Minimierungsproblem (ER) anzuwenden. Es sei  $\mathcal{I}_{\bar{B}_{\infty}}$  die Indikatorfunktion des abgeschlossenen  $\|\cdot\|_{\infty}$ -Balls in  $C(\Omega)$ , also

$$\mathcal{I}_{\bar{B}_{\infty}}(f) := \begin{cases} 0 & \|f\|_{\infty} \leq 1, \\ \infty & \text{sonst.} \end{cases}$$

Wir definieren  $J : H \times C(\Omega) \rightarrow \bar{\mathbb{R}}$ ,

$$J(c, f) := -\operatorname{Re}\langle c, \hat{y} \rangle + \mathcal{I}_{\bar{B}_{\infty}}(f),$$

wobei  $\hat{y} = \mathcal{G}\hat{u}$  wieder die von einem Zielmaß  $\hat{u} \in \mathcal{M}(\Omega)$  gemessenen Daten seien. Die Funktion  $J$  ist konvex, da der erste Summand  $\mathbb{R}$ -linear in  $c$  ist und Normbälle konvex sind. Weiter wählen wir  $\Lambda = \mathcal{G}^*$ . Wenden wir den Realteil auf die definierende Gleichung des prädualen Operators (vgl. Gleichung (2.2)) an, so erhalten wir

$$\operatorname{Re}\langle c, \mathcal{G}u \rangle = \operatorname{Re} \int_{\Omega} \mathcal{G}^* c \, d\bar{u}, \quad \forall c \in H, \forall u \in \mathcal{M}(\Omega).$$

Hieran lesen wir ab, dass tatsächlich  $\mathcal{G} = \Lambda^T$  die transponierte Abbildung zu  $\mathcal{G}^*$  ist.

Das Problem  $(P_1)$  ist gegeben durch

$$\begin{aligned} p_1 &= \inf_{c \in H} J(c, \mathcal{G}^*c) = \inf_{c \in H} -\operatorname{Re}\langle c, \hat{y} \rangle + \mathcal{I}_{\bar{B}_{\infty}}(\mathcal{G}^*c) \\ &= - \sup_{\|\mathcal{G}^*c\|_{\infty} \leq 1} \operatorname{Re}\langle c, \hat{y} \rangle. \end{aligned} \tag{2.5}$$

Das Supremum ist endlich, da für zulässige  $c$  gilt

$$\operatorname{Re}\langle c, \hat{y} \rangle = \operatorname{Re}\langle c, \mathcal{G}\hat{u} \rangle = \operatorname{Re} \int_{\Omega} \mathcal{G}^* c \, d\bar{\hat{u}} \leq \left| \int_{\Omega} \mathcal{G}^* c \, d\bar{\hat{u}} \right| \leq \|\hat{u}\|_{\operatorname{TV}}.$$

Bevor wir das zweite Problem aufschreiben können, müssen wir die Fenchel-Transformierte von  $J$  berechnen. Nach Definition dieser gilt für alle  $c \in H$  und alle  $u \in \mathcal{M}(\Omega)$

$$\begin{aligned} J^*(c, u) &= \sup_{\substack{c' \in H \\ f \in C(\Omega)}} \operatorname{Re}\langle c', c \rangle + \operatorname{Re} \int_{\Omega} f \, d\bar{u} + \operatorname{Re}\langle c', \hat{y} \rangle - \mathcal{I}_{\overline{B_{\infty}}}(f) \\ &= \sup_{c' \in H} \operatorname{Re}\langle c', c + \hat{y} \rangle + \sup_{\|f\|_{\infty} \leq 1} \operatorname{Re} \int_{\Omega} f \, d\bar{u} \\ &= \mathcal{I}_{\{\hat{y}\}}(-c) + \|u\|_{\text{TV}}. \end{aligned}$$

Für Problem  $(P_2)$  erhalten wir

$$\begin{aligned} p_2 &= \sup_{u \in \mathcal{M}(\Omega)} -J^*(\mathcal{G}u, -u) = \sup_{u \in \mathcal{M}(\Omega)} -\mathcal{I}_{\{\hat{y}\}}(-\mathcal{G}u) - \|-u\|_{\text{TV}} \\ &= \sup_{u \in \mathcal{M}(\Omega)} -\mathcal{I}_{\{\hat{y}\}}(\mathcal{G}u) - \|u\|_{\text{TV}} \\ &= - \inf_{\substack{u \in \mathcal{M}(\Omega) \\ \mathcal{G}u = \hat{y}}} \|u\|_{\text{TV}}. \end{aligned} \tag{2.6}$$

Bis auf das geänderte Vorzeichen entspricht dies dem Ausgangsproblem (ER).

Noch zu prüfen ist die letzte Bedingung in Satz 2.6. Im Punkt  $c = 0$  ist  $J(c, \mathcal{G}^*c) = 0$ . Die Funktion  $f \mapsto J(0, f)$  ist konstant gleich 0 im offenen  $\|\cdot\|_{\infty}$ -Einheitsball und dieser ist eine offene Umgebung von  $0 = \mathcal{G}^*c$  in  $C(\Omega)$ .

Nach Satz 2.6 folgt starke Dualität der beiden Probleme, d.h.  $p_1 = p_2$ . Außerdem wird das Infimum in Gleichung (2.6) angenommen.

Lässt sich auch etwas über das Supremum im ersten Problem aussagen? Wir können Gleichung (2.5) (hier ohne das Vorzeichen) umschreiben zu

$$\begin{aligned} \sup_{\|\mathcal{G}^*c\|_{\infty} \leq 1} \operatorname{Re}\langle c, \hat{y} \rangle &= \sup_{\|\mathcal{G}^*c\|_{\infty} \leq 1} \operatorname{Re}\langle c, \mathcal{G}\hat{u} \rangle \\ &= \sup_{\|\mathcal{G}^*c\|_{\infty} \leq 1} \operatorname{Re} \int_{\Omega} \mathcal{G}^*c \, d\bar{u} = \sup_{f \in \operatorname{Im}(\mathcal{G}^*) \cap \overline{B_{\infty}}} \operatorname{Re} \int_{\Omega} f \, d\bar{u}. \end{aligned}$$

Wie schon bekannt ist, ist der Integraloperator stetig in der  $\|\cdot\|_{\infty}$ -Topologie (vgl. Satz 1.17). Nehmen wir nun zusätzlich an, dass das Bild  $\operatorname{Im}(\mathcal{G}^*)$  des prädualen Operators endlichdimensional ist (z.B. wenn  $H = \mathbb{C}^m$  gilt), so ist  $\operatorname{Im}(\mathcal{G}^*) \cap \overline{B_{\infty}}$  nach dem Satz von Heine-Borel eine kompakte Teilmenge von  $C(\Omega)$ . Unter dieser Annahme wird also auch das Supremum in Gleichung (2.5) angenommen.

Die obigen Ergebnisse werden in folgendem Satz zusammengefasst:

**2.10 Satz** (Starke Dualität) Es sei  $\hat{u} \in \mathcal{M}(\Omega)$  und  $\hat{y} = \mathcal{G}\hat{u}$ . Nehme an, dass  $\operatorname{Im}(\mathcal{G}^*)$  endlichdimensional ist.

---

Dann sind die Lösungsmengen der Optimierungsprobleme

$$\min_{u \in \mathcal{M}(\Omega)} \|u\|_{\text{TV}} \quad \text{s.t.} \quad \mathcal{G}u = \hat{y} \quad (\text{ER})$$

und

$$\max_{c \in H} \text{Re}\langle c, \hat{y} \rangle \quad \text{s.t.} \quad \|\mathcal{G}^*c\|_{\infty} \leq 1 \quad (\text{D(ER)})$$

nichtleer und ihr Optimalwert ist gleich.

Natürlich ist es im Hinblick auf eine Rekonstruktion wünschenswert, dass die eindeutige Lösung von (ER) durch das Zielmaß  $\hat{u}$  gegeben ist. In Bezug auf Optimalität kann folgende Beobachtung gemacht werden: Ist  $\hat{c} \in H$  ein Vektor, welcher die Bedingungen

$$\begin{aligned} \|\mathcal{G}^*\hat{c}\|_{\infty} &\leq 1, \\ \int_{\Omega} \mathcal{G}^*\hat{c} \, d\bar{u} &= \|\hat{u}\|_{\text{TV}} \end{aligned} \quad (2.7)$$

erfüllt, so muss  $\hat{u}$  zwangsläufig optimal für (ER) sein. Es gilt in diesem Fall nämlich

$$\begin{aligned} \|\hat{u}\|_{\text{TV}} &= \int_{\Omega} \mathcal{G}^*\hat{c} \, d\bar{u} = \text{Re}\langle \hat{c}, \mathcal{G}\hat{u} \rangle = \text{Re}\langle \hat{c}, \hat{y} \rangle \\ &\leq \sup_{\|\mathcal{G}^*c\|_{\infty} \leq 1} \text{Re}\langle c, \hat{y} \rangle = \min_{\mathcal{G}u = \hat{y}} \|u\|_{\text{TV}} \leq \|\hat{u}\|_{\text{TV}}. \end{aligned}$$

Hier wurde für die zweite Gleichheit die Definition des prädualen Operators genutzt. Der Realteil kann ohne Änderung genommen werden, da die TV-Norm am Anfang der Gleichung reell ist. Ein Vektor  $\hat{c} \in H$ , welcher diese Bedingungen erfüllt, bescheinigt also die Optimalität von  $\hat{u}$  und wird deshalb *duales Zertifikat für Optimalität* genannt.

**2.11 Bemerkung** Sind wir in der Situation von Satz 2.10, dann wissen wir sogar, dass für *jede* Lösung  $u \in \mathcal{M}(\Omega)$  von (ER) ein solches duales Zertifikat für Optimalität existiert. Denn es gilt dann

$$\max_{\|\mathcal{G}^*c\|_{\infty} \leq 1} \text{Re}\langle c, \hat{y} \rangle = \min_{\mathcal{G}u = \hat{y}} \|u\|_{\text{TV}} = \|u\|_{\text{TV}},$$

d.h. es existiert  $c \in H$  mit  $\|\mathcal{G}^*c\|_{\infty} \leq 1$  und

$$\|u\|_{\text{TV}} = \text{Re}\langle c, \hat{y} \rangle = \text{Re}\langle c, \mathcal{G}u \rangle = \text{Re} \int_{\Omega} \mathcal{G}^*c \, d\bar{u}. \quad (2.8)$$

Tatsächlich muss der Imaginärteil des letzten Integrals verschwinden, denn sonst wäre

$$\left| \text{Re} \int_{\Omega} \mathcal{G}^*c \, d\bar{u} \right| < \left| \int_{\Omega} \mathcal{G}^*c \, d\bar{u} \right| \leq \|\mathcal{G}^*c\|_{\infty} \|u\|_{\text{TV}} \leq \|u\|_{\text{TV}}$$

im Widerspruch zu Gleichung (2.8).

Bis jetzt gelten die Betrachtungen dieses Abschnittes für allgemeine Zielmaße  $\hat{u} \in \mathcal{M}(\Omega)$ . Eigentlich sind wir jedoch an diskreten Maßen der Form

$$\hat{u} = \sum_{i=1}^N \alpha_i \delta_{z_i}$$

mit paarweise verschiedenen  $z_1, \dots, z_N \in \Omega$  und einem Gewichtsvektor  $\alpha \in (\mathbb{C} \setminus \{0\})^N$  interessiert. Für ein solches Maß gilt  $\|\hat{u}\|_{\text{TV}} = \|\alpha\|_1$  und die zweite Bedingung in Gleichung (2.7) kann umgeformt werden zu

$$\int_{\Omega} \mathcal{G}^* \hat{c} d\bar{u} = \sum_{i=1}^N \bar{\alpha}_i \mathcal{G}^* \hat{c}(z_i) \stackrel{!}{=} \sum_{i=1}^N |\alpha_i|.$$

Somit sind die Bedingungen aus Gleichung (2.7) für den Fall eines diskreten Zielmaßes äquivalent zu

$$\begin{aligned} \|\mathcal{G}^* \hat{c}\|_{\infty} &\leq 1, \\ (\mathcal{G}^* \hat{c})(z_i) &= \frac{\alpha_i}{|\alpha_i|}, \quad i = 1, \dots, N. \end{aligned} \tag{2.9}$$

**2.12 Bemerkung** In der Literatur wird häufig im Fall des idealen Frequenzfilters nicht der Koeffizientenvektor  $\hat{c} \in \mathbb{C}^m$ , sondern das frequenzbeschränkte trigonometrische Polynom

$$(\mathcal{F}_{f_c}^* \hat{c})(z) = \sum_{l \in F} \hat{c}_l e^{2\pi i l \cdot z}$$

als „duales Zertifikat“ bezeichnet. Zur besseren Unterscheidung nutzen wir für obigen Ausdruck den Begriff „duales Polynom“.

### 2.3. Exakte Rekonstruktion

Die Bedingungen aus Gleichung (2.9) garantieren Optimalität von  $\hat{u}$ . Um mithilfe des TV-Minimierungsproblems das Maß  $\hat{u}$  anhand der Daten  $\hat{y}$  rekonstruieren zu können, würden wir jedoch auch gerne Bedingungen für die Eindeutigkeit der Lösung formulieren.

In diesem Abschnitt behandeln wir wieder die Probleme

$$\min_{u \in \mathcal{M}(\Omega, \mathbb{K})} \|u\|_{\text{TV}} \quad \text{s.t.} \quad \mathcal{G}u = \hat{y} \tag{ER_{\mathbb{K}}}$$

für  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{C}$ .

**2.13 Definition** Ein Maß  $\hat{u} \in \mathcal{M}(\Omega, \mathbb{K})$  heißt *exakt rekonstruierbar*, wenn  $\hat{u}$  eindeutige Lösung des Minimierungsproblems (ER $_{\mathbb{K}}$ ) zu den Daten  $\hat{y} = \mathcal{G}\hat{u}$  ist. In

diesem Fall sagen wir auch „ $(\text{ER}_{\mathbb{K}})$  rekonstruiert  $\hat{u}$  exakt“.

**2.14 Bemerkung** Ist  $\hat{c}$  ein duales Zertifikat für Optimalität zu einem Zielmaß  $\hat{u} \in \mathcal{M}(\Omega, \mathbb{K})$ , dann ist  $\hat{u}$  optimal für  $(\text{ER}_{\mathbb{K}})$ . Für jede weitere Lösung  $u$  von  $(\text{ER}_{\mathbb{K}})$  gilt dann

$$\begin{aligned} \|u\|_{\text{TV}} = \|\hat{u}\|_{\text{TV}} &= \int_{\Omega} \mathcal{G}^* \hat{c} \, d\bar{u} = \langle \hat{c}, \mathcal{G}\hat{u} \rangle = \langle \hat{c}, \hat{y} \rangle = \langle \hat{c}, \mathcal{G}u \rangle \\ &= \int_{\Omega} \mathcal{G}^* \hat{c} \, d\bar{u}. \end{aligned}$$

Erfüllt das Zertifikat zusätzlich die *strikte* Ungleichung außerhalb des Trägers von  $\hat{u}$ , d.h.

$$|(\mathcal{G}^* \hat{c})(z)| < 1 \quad \forall z \in \Omega \setminus \text{supp}(\hat{u}), \quad (2.10)$$

so folgt mit Korollar 1.10, dass auch  $u$  konzentriert ist auf  $\text{supp}(\hat{u})$ .

Falls also ein Zertifikat  $\hat{c}$  für  $\hat{u}$  existiert, sodass die stetige Funktion  $\mathcal{G}^* \hat{c}$  sogar eine strikte Ungleichung außerhalb des Trägers von  $\hat{u}$  erfüllt, dann ist *jede* Lösung von  $(\text{ER}_{\mathbb{K}})$  ein diskretes Maß auf den Punkten  $z_1, \dots, z_N$ .

Es sei

$$\text{sgn } \mathbb{K}^N := \{\eta \in \mathbb{K}^N \mid |\eta_i| = 1, i = 1, \dots, N\}.$$

Die folgende Verschärfung der Bedingungen an ein Zertifikat liefert auch Eindeutigkeit der Lösung von  $(\text{ER}_{\mathbb{K}})$ . Neben der strikten Ungleichung wird zusätzlich gefordert, dass zu jedem Vektor von (komplexen) Vorzeichen  $\eta \in \text{sgn } \mathbb{K}^N$  ein Zertifikat existiert, welches an den Punkten  $\{z_1, \dots, z_N\}$  diese Vorzeichen annimmt.

**2.15 Satz** ([9, Prop. A.1]) Es sei  $\hat{u} = \sum_{i=1}^N \alpha_i \delta_{z_i}$  mit paarweise verschiedenen  $z_1, \dots, z_N \in \Omega$  und einem Gewichtsvektor  $\alpha \in (\mathbb{K} \setminus \{0\})^N$ . Angenommen, für jeden Vektor  $\eta \in \text{sgn } \mathbb{K}^N$  existiert ein  $c \in H$ , sodass gilt

$$\begin{aligned} (\mathcal{G}^* c)(z_i) &= \eta_i, \quad i = 1, \dots, N, \\ |(\mathcal{G}^* c)(z)| &< 1 \quad \forall z \in \Omega \setminus \text{supp}(\hat{u}). \end{aligned}$$

Dann ist  $\hat{u}$  die eindeutige Lösung von  $(\text{ER}_{\mathbb{K}})$ .

*Beweis.* Es sei  $u$  eine beliebige Lösung von  $(\text{ER}_{\mathbb{K}})$ . Definiere die Differenz aus Rekonstruktion und Zielmaß als

$$h := u - \hat{u}.$$

Mittels Lebesguezerlegung (Satz 1.4) kann  $h$  bezüglich  $|\hat{u}|$  zerlegt werden in zwei Maße  $h_a, h_s \in \mathcal{M}(\Omega, \mathbb{K})$ , sodass gilt

$$h = h_a + h_s, \quad h_a \ll |\hat{u}|, \quad h_s \perp |\hat{u}|.$$

Das Maß  $h_a$  ist konzentriert auf  $\text{supp}(\hat{u}) = \{z_1, \dots, z_N\}$  und lässt sich deshalb schreiben als

$$h_a = \sum_{i=1}^N \beta_i \delta_{z_i}$$

mit Gewichten  $\beta_i \in \mathbb{K}$ .

Es sei nun  $c \in H$  das Zertifikat aus den Voraussetzungen zu  $\eta \in \text{sgn } \mathbb{K}^N$  gegeben durch

$$\eta_i := \begin{cases} \frac{\beta_i}{|\beta_i|} & \text{falls } \beta_i \neq 0, \\ 1 & \text{sonst,} \end{cases} \quad i = 1, \dots, N.$$

Da  $u$  und  $\hat{u}$  die Nebenbedingung von  $(\text{ER}_{\mathbb{K}})$  erfüllen, gilt  $\mathcal{G}h = \mathcal{G}u - \mathcal{G}\hat{u} = 0$ , also

$$0 = \langle \mathcal{G}h, c \rangle = \int_{\Omega} \overline{\mathcal{G}^*c} dh = \int_{\Omega} \overline{\mathcal{G}^*c} dh_a + \int_{\Omega} \overline{\mathcal{G}^*c} dh_s. \quad (2.11)$$

Das Maß  $h_a$  ist diskret, darum lässt sich das erste Integral leicht berechnen:

$$\int_{\Omega} \overline{\mathcal{G}^*c} dh_a = \sum_{i=1}^N \beta_i \overline{(\mathcal{G}^*c)(z_i)} = \sum_{i=1}^N \beta_i \overline{\eta_i} = \sum_{i=1}^N |\beta_i| = \|h_a\|_{\text{TV}}.$$

Wir stellen Gleichung (2.11) um und nutzen die Tatsache, dass  $\|\mathcal{G}^*c\|_{\infty} \leq 1$ :

$$\|h_a\|_{\text{TV}} = \left| \int_{\Omega} \overline{\mathcal{G}^*c} dh_s \right| \leq \|h_s\|_{\text{TV}} \quad (2.12)$$

Diese Ungleichung zwischen den TV-Normen von  $h_a$  und  $h_s$  ist zentral für den Beweis. Aufgrund der Optimalität von  $u$  gilt

$$\|\hat{u}\|_{\text{TV}} \geq \|u\|_{\text{TV}} = \|\hat{u} + h\|_{\text{TV}}. \quad (2.13)$$

Es gilt  $(\hat{u} + h_a) \perp h_s$  und Lemma 1.11 erlaubt uns, die TV-Norm aufzuspalten:

$$\|\hat{u} + h\|_{\text{TV}} = \|\hat{u} + h_a\|_{\text{TV}} + \|h_s\|_{\text{TV}} \geq \|\hat{u}\|_{\text{TV}} - \|h_a\|_{\text{TV}} + \|h_s\|_{\text{TV}}$$

Schließlich gilt wegen Gleichung (2.12)

$$\|\hat{u}\|_{\text{TV}} - \|h_a\|_{\text{TV}} + \|h_s\|_{\text{TV}} \geq \|\hat{u}\|_{\text{TV}}$$

und mit Blick auf den Start der Ungleichungskette in (2.13) muss überall Gleichheit gelten. Insbesondere folgt Gleichheit in (2.12), d.h.

$$\left| \int_{\Omega} \overline{\mathcal{G}^*c} dh_s \right| = \|h_s\|_{\text{TV}}.$$

Da jedoch  $\mathcal{G}^*c$  die Bedingung

$$|(\mathcal{G}^*c)(z)| < 1 \quad \forall z \in \Omega \setminus \text{supp}(\hat{u})$$

erfüllt und  $h_s$  konzentriert auf  $\Omega \setminus \text{supp}(\hat{u})$  ist, impliziert dies nach Korollar 1.10

$$h_s = 0.$$

Aus Gleichung (2.12) erhalten wir sofort auch  $h_a = 0$ , also  $u = \hat{u}$ .  $\square$

Wir wollen noch einen alternativen Beweis für obige Aussage präsentieren, um den Zusammenhang zur dualen Struktur deutlicher auszuarbeiten: Ist  $u$  eine Lösung von  $(\text{ER}_{\mathbb{K}})$ , dann folgt aus der strikten Ungleichung wie in Bemerkung 2.14, dass  $u$  konzentriert auf  $\text{supp}(\hat{u})$  ist. Betrachte das Differenz-Minimierungsproblem

$$\min_{h \in \mathcal{M}(\Omega, \mathbb{K})} \|h\|_{\text{TV}} \quad \text{s.t.} \quad \mathcal{G}h = 0.$$

Das Zertifikat mit der Wahl von  $\eta$  wie im obigen Beweis ist nun ein duales Zertifikat für Optimalität von  $h = u - \hat{u}$  für das Differenz-Minimierungsproblem. Die eindeutige Lösung hierfür ist jedoch offensichtlich das Nullmaß, d.h.  $h = 0$ .

## 2.4. Konstruktion dualer Zertifikate

Unter welchen Voraussetzungen kann ein Maß  $\hat{u} \in \mathcal{M}(\Omega, \mathbb{K})$  durch  $(\text{ER}_{\mathbb{K}})$  exakt rekonstruiert werden? Eine Strategie zur Beantwortung dieser Frage ist es, Situationen zu untersuchen, in denen Zertifikate für  $\hat{u}$  konstruiert werden können, welche die Voraussetzungen aus Satz 2.15 erfüllen. In [9] wird diese Strategie in der Situation des idealen Frequenzfilters verfolgt. Dort wird gezeigt, dass geeignete Zertifikate existieren, solange die Punkte  $\{z_1, \dots, z_N\}$  nicht zu nah beieinander liegen.

Dies ist intuitiv einleuchtend: Ein duales Zertifikat entspricht hier einem frequenzbeschränkten trigonometrischen Polynom

$$(\mathcal{F}_{f_c}^*c)(z) = \sum_{l \in F} c_l e^{2\pi i(l \cdot z)}, \quad z \in [0, 1]^d, \quad (2.14)$$

wobei

$$F = \{l \in \mathbb{Z}^d \mid \|l\|_{\infty} \leq f_c\}$$

für die Menge der ganzzahligen Frequenzen bis zur Grenzfrequenz  $f_c$  stand. Dieses Polynom muss zwischen den Vorzeichen in  $\eta \in \text{sgn } \mathbb{K}^N$  an den Punkten  $z_1, \dots, z_N$  interpolieren. Liegen die Punkte zu nah beieinander, so ist eine Interpolation aufgrund der Frequenzbeschränkung nicht mehr möglich.

Die trigonometrischen Polynome aus Gleichung (2.14) sind periodisch, was bei der Formulierung eines Abstandsbegriffs für die Punkte  $\{z_1, \dots, z_N\}$  berücksichtigt

werden muss. Dies führt auf die sogenannte Torus-Metrik auf dem Intervall  $(0, 1)$ :

$$d_{\mathbb{T}}(x, y) := \min\{|x - y|, 1 - |x - y|\}, \quad x, y \in (0, 1).$$

Anschaulich nimmt diese Metrik den kürzesten Abstand auch über die Intervallgrenzen hinweg, sodass z.B.  $d_{\mathbb{T}}(0.1, 0.9) = 0.2$ . Die  $d$ -dimensionale Torus-Metrik ist dann die Maximumsmetrik zu  $d_{\mathbb{T}}$  auf  $(0, 1)^d$ :

$$d_{\mathbb{T}^d}(v, w) := \max_{i=1}^d d_{\mathbb{T}}(v_i, w_i), \quad v, w \in (0, 1)^d. \quad (2.15)$$

**2.16 Definition** Die minimale Separation einer Familie von Punkten  $T \subset (0, 1)^d$  ist definiert durch

$$\Delta(T) := \inf_{\substack{z, z' \in T \\ z \neq z'}} d_{\mathbb{T}^d}(z, z').$$

Nun können wir das Hauptresultat von Candès et al. zur exakten Rekonstruktion aus Messungen des idealen Frequenzfilters formulieren:

**2.17 Satz** ([9, Thm. 1.2]) Es sei  $d = 1$  und  $\hat{u} = \sum_{i=1}^N \alpha_i \delta_{z_i}$  mit Gewichten  $\alpha \in \mathbb{K}^N$  und Träger  $T := \{z_1, \dots, z_N\} \subset (0, 1)$ . Weiter sei  $\mathcal{G} = \mathcal{F}_{f_c}$  der ideale Frequenzfilter für eine Grenzfrequenz  $f_c \in 2\mathbb{N}$  mit  $f_c \geq 128$ .

Ist  $\mathbb{K} = \mathbb{C}$  und ist die Ungleichung

$$\Delta(T) \geq 2/f_c$$

erfüllt, dann ist  $\hat{u}$  die eindeutige Lösung von  $(\text{ER}_{\mathbb{K}})$ .

Ist  $\mathbb{K} = \mathbb{R}$ , so genügt die Bedingung

$$\Delta(T) \geq 1.87/f_c.$$

In zwei Dimensionen gilt ein analoges Resultat:

**2.18 Satz** ([9, Thm. 1.3]) Es sei  $d = 2$  und  $\hat{u} = \sum_{i=1}^N \alpha_i \delta_{z_i}$  mit Gewichten  $\alpha \in \mathbb{R}^N$  und Träger  $T := \{z_1, \dots, z_N\} \subset (0, 1)^2$ . Weiter sei  $\mathcal{G} = \mathcal{F}_{f_c}$  der ideale Frequenzfilter für eine Grenzfrequenz  $f_c \in 2\mathbb{N}$  mit  $f_c \geq 512$ .

Falls gilt

$$\Delta(T) \geq 2.38/f_c,$$

dann ist  $\hat{u}$  die eindeutige Lösung von  $(\text{ER}_{\mathbb{K}})$  für  $\mathbb{K} = \mathbb{R}$ .

**2.19 Bemerkung** Da der ideale Frequenzfilter periodisch ist, würde eine Punktquelle bei  $z = 0$  dieselbe Messung verursachen wie eine Punktquelle bei  $z = 1$ . Die Maße  $\hat{u} = \delta_0$  und  $\hat{u} = \delta_1$  sind somit anhand der Messung nicht voneinander zu unterscheiden. Candès et al. formulieren das Problem  $(\text{ER}_{\mathbb{K}})$  über dem Raum der Maße  $\mathcal{M}(\mathbb{T})$  auf dem Torus  $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ . Hier sind die Intervallgrenzen 0 und 1

miteinander identifiziert und man kann auch eine Punktquelle auf dem Rand exakt rekonstruieren (ohne jedoch unterscheiden zu können, ob diese nun bei 0 oder bei 1 lag).

Wie angekündigt, funktioniert der Beweis über die Konstruktion von dualen Zertifikaten. Unter den Voraussetzungen der obigen Sätze ist es also zu jedem Vektor von Vorzeichen  $\eta \in \text{sgn } \mathbb{K}^N$  möglich, ein trigonometrisches Polynom zu konstruieren, welches keine höheren Frequenzen als  $f_c$  enthält und zwischen den Vorzeichen in  $\eta$  interpoliert. Der Beweis für  $d = 1$  soll im Folgenden skizziert werden.

Wir befinden uns also in der Situation von Satz 2.17 und es sei  $\eta \in \text{sgn } \mathbb{C}^N$  ein Vektor von komplexen Zahlen mit Betrag 1. Um Satz 2.15 zur exakten Rekonstruktion anwenden zu können, muss ein trigonometrisches Polynom  $q \in \text{Im}(\mathcal{F}_{f_c}^*)$  gefunden werden, welches die Bedingungen

$$\begin{aligned} q(z_i) &= \eta_i, & i &= 1, \dots, N, \\ |q(z)| &< 1 & \forall z &\in [0, 1] \setminus T \end{aligned}$$

erfüllt.

Candès et al. betrachten den Faltungskern  $K \in C(\mathbb{T})$  definiert durch

$$K(z) := \left( \frac{\sin((f_c/2 + 1)\pi z)}{(f_c/2 + 1) \sin(\pi z)} \right)^4. \quad (2.16)$$

Dieser Kern hat bei  $z = 0$  den Wert 1 und fällt zu beiden Seiten schnell ab. Da wir uns auf gerade Grenzfrequenzen  $f_c \in 2\mathbb{N}$  beschränken, handelt es sich bei dieser Funktion um das (normalisierte) Quadrat des sog. *Fejér-Kerns*

$$F_n(z) := \frac{1}{n+1} \sum_{k=0}^n D_k(z)$$

mit der Wahl  $n = f_c/2$ . Hierbei ist der *Dirichlet-Kern* wie schon in Abschnitt 2.1.3 durch

$$D_k(z) := \sum_{l=-k}^k e^{i2\pi lz}$$

definiert. Die Zusammenhänge werden von folgender Proposition erfasst:

**2.20 Proposition** (vgl. [17, Abschnitt 3.1.3]) Der Dirichlet-Kern hat folgende einfache Darstellung durch Sinus-Funktionen:

$$D_k(z) = \frac{\sin((2k+1)\pi z)}{\sin(\pi z)}$$

Für den Fejér-Kern gilt analog

$$F_n(z) = \frac{1}{n+1} \frac{\sin((n+1)\pi z)^2}{\sin(\pi z)^2}.$$

Die Koeffizienten des Fejér-Kerns als trigonometrisches Polynom können abgelesen werden an der Gleichung

$$F_n(z) = \sum_{l=-n}^n \left(1 - \frac{|l|}{n+1}\right) e^{i2\pi lz}.$$

Schließlich geben wir auch eine konkrete Form der Koeffizienten von  $K$  an (vgl. [23]), wobei wir zu Abkürzung  $n = f_c/2$  schreiben:

$$K(z) = \sum_{r=-2n}^{2n} g(r) e^{i2\pi r z}$$

$$\text{mit } g(r) = \frac{1}{(n+1)^2} \sum_{l=\max\{r-n, -n\}}^{\min\{r+n, n\}} \left(1 - \frac{|l|}{n+1}\right) \left(1 - \frac{|r-l|}{n+1}\right)$$

*Beweis.* Siehe Anhang A.1.2. □

**2.21 Bemerkung** Der Fall ungerader Grenzfrequenzen  $f_c = 2w + 1$ ,  $w \in \mathbb{N}$  wird von den Autoren in [9] nicht explizit behandelt. Der Kern aus Gleichung (2.16) kann nicht mehr genutzt werden, da

$$\lim_{z \nearrow 1} |K(z)| = \lim_{z \nearrow 1} \left| \frac{\sin((w+1+1/2)\pi z)}{(w+1+1/2)\sin(\pi z)} \right|^4 = \infty.$$

Der Sinus über dem Bruchstrich konvergiert gegen  $\sin((z+1+1/2)\pi) \in \{-1, 1\}$ , während der Sinus unter dem Bruchstrich gegen 0 konvergiert. Im Fall einer ungeraden Grenzfrequenz wird also durch diesen Ausdruck kein trigonometrisches Polynom mehr definiert.

Allerdings kann man sich natürlich auf den Fall  $f_c - 1$  zurückziehen (möglicherweise müssen hierzu die Konstanten in den obigen Sätzen leicht angepasst werden).

Die letzte Aussage der obigen Proposition zeigt, dass der Kern aus Gleichung (2.16) tatsächlich nur Frequenzen in  $F = \{-f_c, \dots, f_c\}$  enthält. Natürlich gilt dies dann auch für die Ableitung  $K'$ , sodass jede Linearkombination der beiden Funktionen im Bild von  $\mathcal{F}_{f_c}^*$  liegt und somit zur Konstruktion eines dualen Polynoms genutzt werden kann.

Es wird nun für Koeffizientenvektoren  $a, b \in \mathbb{C}^N$  die Linearkombination

$$q(z) = \sum_{j=1}^N a_j K(z - z_j) + b_j K'(z - z_j) \tag{2.17}$$

---

aus dem verschobenen Kern und seiner Ableitung  $K'$  gebildet. Die Idee ist es, Koeffizienten  $a, b$  zu wählen, sodass an den Punkten  $z_i$  die Werte  $\eta_i$  angenommen werden:

$$q(z_i) = \sum_{j=1}^N a_j K(z_i - z_j) + b_j K'(z_i - z_j) \stackrel{!}{=} \eta_i, \quad i = 1, \dots, N \quad (2.18)$$

Zusätzlich soll das duale Polynom an allen anderen Punkten  $z \in [0, 1] \setminus T$  betragslich strikt kleiner als 1 bleiben. Dazu wird gefordert, dass an den Punkten  $z_i$  die Ableitung gleich 0 ist, d.h.

$$q'(z_i) = \sum_{j=1}^N a_j K'(z_i - z_j) + b_j K''(z_i - z_j) \stackrel{!}{=} 0, \quad i = 1, \dots, N. \quad (2.19)$$

Diese Forderung soll lokale Extremstellen an den Punkten  $z_i$  erzwingen. Zu zeigen ist noch, dass

1. Koeffizientenvektoren  $a, b$  existieren, sodass Gleichungen (2.18) und (2.19) erfüllt sind und
2. das resultierende trigonometrische Polynom  $q$  außerhalb von  $T$  tatsächlich betragslich strikt kleiner als 1 bleibt.

Definiere Matrizen  $M_0, M_1, M_2 \in \mathbb{R}^{N \times N}$  durch

$$(M_0)_{ij} := K(z_i - z_j), \quad (M_1)_{ij} := K'(z_i - z_j), \quad (M_2)_{ij} := K''(z_i - z_j).$$

Gleichungen (2.18) und (2.19) sind äquivalent zu dem linearen Gleichungssystem

$$\begin{pmatrix} M_0 & M_1 \\ M_1 & M_2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \eta \\ 0 \end{pmatrix} \quad (2.20)$$

für die Koeffizientenvektoren  $a, b \in \mathbb{C}^N$ . Bezeichne die obige Blockmatrix mit  $M$ .

**2.22 Lemma** ([9, vgl. Lem. 2.2]) Unter den Annahmen von Satz 2.17 ist die Matrix in Gleichung (2.20) invertierbar.

Zum Beweis wir der Abstand  $\|I_{2N} - M\|$  zur Einheitsmatrix in einer geeigneten Matrixnorm gegen 1 abgeschätzt. Hierzu werden allgemeine Abschätzungen über den Kern  $K$  und seine Ableitungen genutzt. Wir verzichten hier auf die Details und verweisen auf [9].

**2.23 Bemerkung** Ist  $\mathbb{K} = \mathbb{R}$ , dann werden nur Vorzeichenvektoren  $\eta \in \text{sgn } \mathbb{R}^N$  interpoliert, deshalb sind in diesem Fall auch die Koeffizienten  $(a, b)$  sowie das duale Polynom  $q$  reellwertig.

Zur Vervollständigung zitieren wir noch das Lemma, welches strikte Ungleichung außerhalb des Trägers des Zielmaßes garantiert. Da der Beweis jedoch auf einer Reihe von technischen Abschätzungen des Kerns  $K$  und seiner Ableitungen beruht, wird dieser hier nicht angeführt.

**2.24 Lemma** ([9, Lem 2.3 und 2.4]) Es gelten die Annahmen von Satz 2.17. Für ein  $\eta \in \text{sgn } \mathbb{K}^N$  sei  $(a, b)$  die Lösung des Gleichungssystems (2.20) und  $q$  das zugehörige duale Polynom definiert durch Gleichung (2.17).

Dann gilt  $|q(z)| < 1$  für alle  $z \in [0, 1] \setminus T$ .

### 3. Dynamische Super-Resolution

Häufig möchte man in Anwendungen nicht statische, sondern bewegte Motive abbilden. Wird diese Bewegung bei der Rekonstruktion nicht berücksichtigt, so kommt es zu einem Phänomen, was auch Hobbyfotografen als *Bewegungsunschärfe* bekannt ist. Hierdurch verschlechtert sich die Auflösbarkeit der abgebildeten Objekte. Eine schnelle Reparatur besteht in einer Verringerung der Belichtungszeit. Dies ist jedoch in Anwendungen problematisch, in denen das vom Motiv gemessene Signal schon so gering sind, dass es sich kaum vom Hintergrundrauschen abhebt. Stattdessen wäre es vorteilhaft, Daten über einen längeren Zeitraum zu messen und den Einfluss der Bewegung in der Rekonstruktionphase wieder herauszurechnen.

Ein zweiter Aspekt, welcher für eine Modellierung der Dynamik spricht, ist der Wunsch nach einer Rekonstruktion der Geschwindigkeiten der Motive. Eine denkbare Methode hierzu wäre, zunächst mehrere statische Rekonstruktionen (d.h. ohne Berücksichtigung der Dynamik) zu mehreren Zeitschritten durchzuführen und anschließend anhand der rekonstruierten Positionen die Geschwindigkeiten zu schätzen. Dieser Ansatz hätte zum Nachteil, dass für die einzelnen statischen Rekonstruktionen jeweils nur die Daten aus einem Zeitschritt zur Verfügung ständen.

Stattdessen soll in diesem Kapitel der Ansatz einer gleichzeitigen Rekonstruktion von Positionen und Geschwindigkeiten verfolgt werden. Für den Spezialfall der bewegten Punktquellen mit linearer Dynamik wurde dieser Ansatz von Alberti et al. in [2] behandelt.

#### 3.1. Vorstellung des Modells

Das zugrundeliegende Modell soll zunächst eingeführt werden: Für ein  $\delta > 0$  sei  $[-\delta, \delta]$  das Zeitfenster der Beobachtung. Weiter sei  $K \in \mathbb{N}_{>0}$  und  $\tau := \delta/K$  der Abstand zwischen zwei Zeitschritten. Die Annahmen des Modells lauten

- 1) Zu jeder Zeit in  $[-\delta, \delta]$  befinden sich dieselben  $N \in \mathbb{N}$  Teilchen im Gebiet  $[0, 1]^d$
- 2) Die Änderung der Geschwindigkeiten ist während der Messung hinreichend klein, sodass von einer linearen Dynamik ausgegangen werden kann
- 3) Es werden  $2K + 1$  Messungen zu den Zeitpunkten  $t_k = \tau k$ ,  $k \in \{-K, \dots, K\}$  aufgenommen
- 4) Der Messprozess ist zu jedem Zeitpunkt der Gleiche und lässt sich modellieren durch einen Operator  $B : \mathcal{M}([0, 1]^d) \rightarrow \tilde{H}$  für einen Hilbertraum  $\tilde{H}$ .

Um die Rekonstruktion durchzuführen, schlagen Alberti et al. einen Übergang in den Phasenraum vor, d.h. in den Parameterraum, der aus der Kombination von Position und Geschwindigkeit besteht. Unter Berücksichtigung von Annahme 1) sei der kombinierte Parameterraum gegeben durch

$$\Omega = \{(x, v) \in \mathbb{R}^d \times \mathbb{R}^d \mid \forall t \in [-\delta, \delta] : x + tv \in [0, 1]^d\}.$$

Annahmen 1) und 2) führen auf eine Darstellung durch ein zeitabhängiges Maß der Form

$$\hat{\rho}_t = \sum_{i=1}^N \alpha_i \delta_{x_i + tv_i}, \quad t \in [-\delta, \delta]$$

mit Gewichten  $\alpha_i \in \mathbb{K}$  für  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{C}$  und Positionen und Geschwindigkeiten  $(x_i, v_i) \in \Omega$ . Wir gehen davon aus, dass die  $(x_i, v_i)$  paarweise verschieden sind (sonst könnten sie zusammengefasst werden) und schreiben  $T := \{(x_i, v_i) \mid i = 1, \dots, N\}$ .

Wir wollen nun zum Phasenraum übergehen, um die gleichzeitige Rekonstruktion an allen Zeitpunkten zu formulieren. Die Repräsentation im Phasenraum ist gegeben durch

$$\hat{\lambda} := \sum_{i=1}^N \alpha_i \delta_{(x_i, v_i)}.$$

Wir schreiben  $\mathcal{K} := \{-K, \dots, K\}$  für die Menge der betrachteten Zeitschritte und  $|\mathcal{K}| = 2K + 1$  für die Anzahl dieser. Definieren wir für  $k \in \mathcal{K}$  und eine Dimension  $l \in \mathbb{N}$  die Abbildung

$$D_k^l : \mathbb{R}^l \times \mathbb{R}^l \rightarrow \mathbb{R}^l, \quad (x, v) \mapsto x + k\tau v$$

so erhalten wir den Zusammenhang  $(D_k^d)_\# \hat{\lambda} = \hat{\rho}_{k\tau}$ . Also repräsentiert das Maß

$$\hat{u}_k := (D_k^d)_\# \hat{\lambda} \in \mathcal{M}([0, 1]^d, \mathbb{K})$$

die Konfiguration der Teilchen zum  $k$ -ten Zeitschritt und entsprechend Annahme 4) messen wir zu diesem Zeitschritt die Daten

$$\hat{y}_k := B\hat{u}_k \in \tilde{H}.$$

## 3.2. Der dynamische Messoperator

Das Ziel ist es nun, die Rekonstruktion von  $\hat{\lambda}$  aus den Daten  $\hat{y} := (\hat{y}_k)_{k \in \mathcal{K}}$  wieder durch ein TV-Minimierungsproblem wie in Kapitel 2 zu formulieren. Hierfür benötigen wir noch einen geeigneten Messoperator  $\mathcal{G}$ , der die Messung für alle Zeitschritte zusammengenommen repräsentiert. Zu diesem Zweck definieren wir  $H := \tilde{H}^{\mathcal{K}}$  und den Operator

$$\mathcal{G} : \mathcal{M}(\Omega) \rightarrow H \quad \lambda \mapsto (B(D_k^d)_\# \lambda)_{k \in \mathcal{K}},$$

welcher einfach die Messungen zu allen Zeitschritten zu einem Vektor konkateniert.

Analog zu Kapitel 2 wollen wir die Annahme treffen, dass zu  $B$  ein prädualer Operator  $B^* : \tilde{H} \rightarrow C([0, 1]^d)$  existiert, welcher

$$\langle \tilde{c}, Bu \rangle_{\tilde{H}} = \int_{[0,1]^d} B^* \tilde{c} d\bar{u}$$

für alle  $\tilde{c} \in \tilde{H}$  und alle  $u \in \mathcal{M}([0, 1]^d)$  erfüllt. Ist dies der Fall, so existiert auch zum zusammengesetzten Messoperator  $\mathcal{G}$  ein prädualer Operator  $\mathcal{G}^*$ , welcher sich aus  $B^*$  berechnen lässt: Für alle  $c = (c)_{k \in \mathcal{K}} \in H$  und alle  $\lambda \in \mathcal{M}(\Omega)$  gilt nämlich

$$\begin{aligned} \langle c, \mathcal{G}\lambda \rangle_H &= \sum_{k \in \mathcal{K}} \langle c_k, B(D_k^d)_{\#} \lambda \rangle_{\tilde{H}} = \sum_{k \in \mathcal{K}} \int_{[0,1]^d} B^* c_k d(D_k^d)_{\#} \bar{\lambda} \\ &= \int_{[0,1]^d} \sum_{k \in \mathcal{K}} (B^* c_k)(x + k\tau v) d\bar{\lambda}. \end{aligned}$$

Der präduale Operator von  $\mathcal{G}$  ist also gegeben durch

$$\mathcal{G}^* : H \rightarrow C(\Omega), \quad c \mapsto \mathcal{G}^* c$$

mit

$$(\mathcal{G}^* c)(x, v) = \sum_{k \in \mathcal{K}} (B^* c_k)(x + k\tau v), \quad (x, v) \in \Omega.$$

Nun können wir das dynamische Minimierungsproblem

$$\min_{\lambda \in \mathcal{M}(\Omega, \mathbb{K})} \|\lambda\|_{\text{TV}} \quad \text{s.t.} \quad \mathcal{G}\lambda = \hat{y} \quad (\text{ERdyn})$$

aufstellen. Wir stellen fest, dass dieses Minimierungsproblem genau die Form aus Kapitel 2 für einen allgemeinen Messoperator  $\mathcal{G}$  hat.

### 3.3. Dynamische duale Zertifikate

Es ist also ein Zielmaß

$$\hat{\lambda} = \sum_{i=1}^N \alpha_i \delta_{(x_i, v_i)}$$

gegeben und wir wollen dieses durch (ERdyn) aus den Daten  $\hat{y} = \mathcal{G}\hat{\lambda}$  rekonstruieren. Um exakte Rekonstruktion für das dynamische Problem zu zeigen, soll wieder Satz 2.15 genutzt werden. Doch wie sehen duale Zertifikate für den neuen Messoperator  $\mathcal{G}$  aus? Um Satz 2.15 anwenden zu können, muss für jeden Vektor  $\eta \in \text{sgn } \mathbb{K}^N$  ein  $c \in H$  existieren, sodass die stetige Funktion

$$(\mathcal{G}^* c)(x, v) = \sum_{k \in \mathcal{K}} (B^* c_k)(x + k\tau v)$$

die Bedingungen

$$\begin{aligned} (\mathcal{G}^*c)(x_i, v_i) &= \eta_i, & i &= 1, \dots, N, \\ |(\mathcal{G}^*c)(x, v)| &< 1, & (x, v) &\in \Omega \setminus T \end{aligned} \quad (3.1)$$

erfüllt.

**3.1 Definition** Einen Vektor  $c \in H$ , welcher die Bedingungen (3.1) für ein  $\eta \in \text{sgn } \mathbb{K}^N$  erfüllt, nennen wir *dynamisches duales Zertifikat* (zu  $\eta$ ).

Im Folgenden soll ein Bezug zu den dualen Zertifikaten für das statische Rekonstruktionsproblem zum Zeitschritt  $k$  hergestellt werden, welches gegeben sei durch

$$\min_{u \in \mathcal{M}([0,1]^d, \mathbb{K})} \|u\|_{\text{TV}} \quad \text{s.t.} \quad Bu = \hat{y}_k. \quad (\text{ER}k)$$

Zu diesem Zeitschritt haben sich die Teilchen auf die Positionen

$$T_k = \{x_i + k\tau v_i \mid i = 1, \dots, N\}$$

bewegt und das Problem (ER $k$ ) entspricht dem aus Kapitel 2 mit dem Zielmaß

$$\hat{u}_k = \sum_{i=1}^N \alpha_i \delta_{x_i + k\tau v_i},$$

dem Messoperator  $B$  und den Daten  $\hat{y}_k$ . Entsprechend Satz 2.15 ist ein Vektor  $\tilde{c}_k \in \tilde{H}$  ein Zertifikat zu  $\eta \in \text{sgn } \mathbb{K}^N$  für dieses statische Problem, falls gilt

$$\begin{aligned} (B^*\tilde{c}_k)(x_i + k\tau v_i) &= \eta_i, & i &= 1, \dots, N, \\ |(B^*\tilde{c}_k)(y)| &< 1, & y &\in [0, 1]^d \setminus T_k. \end{aligned} \quad (3.2)$$

**3.2 Definition** Einen Vektor  $\tilde{c}_k \in \tilde{H}$ , welcher die Bedingungen (3.2) für ein  $\eta \in \text{sgn } \mathbb{K}^N$  erfüllt, nennen wir *statisches duales Zertifikat* (zu  $\eta$ ).

**3.3 Bemerkung** Es fällt auf, dass die erste Bedingung in (3.2) nicht für beliebige Vorzeichen  $\eta$  erfüllbar ist, wenn zwei unterschiedliche Teilchen  $(x_i, v_i)$ ,  $(x_j, v_j)$  am Zeitschritt  $k$  dieselbe Position einnehmen. Für Vorzeichen  $\eta_i \neq \eta_j$  würde in der ersten Bedingung von der Funktion  $(B^*\tilde{c}_k)$  gefordert, zwei unterschiedliche Werte an einem Punkt  $x_i + k\tau v_i = x_j + k\tau v_j$  anzunehmen.

**3.4 Definition** (Koinzidenz) Es sei  $\{(x_i, v_i) \mid i = 1, \dots, N\}$  eine Konfiguration von Teilchen. Wir sagen, die Konfiguration hat eine *Koinzidenz* zu einer Zeit  $k \in \mathcal{K}$ , wenn am  $k$ -ten Zeitschritt zwei Teilchen dieselbe Position einnehmen, d.h. wenn gilt

$$\exists i, j \in \{1, \dots, N\}, i \neq j : \quad x_i + k\tau v_i = x_j + k\tau v_j.$$

Eine wichtige Beobachtung von Alberti et al. ist es, dass dynamische duale Zer-

tifikate als Mittelung aus statischen dualen Zertifikaten gewonnen werden können.

**3.5 Definition** (Durchschnittszertifikat, vgl. [2, Def. 3]) Es sei  $\eta \in \text{sgn } \mathbb{K}^N$  und  $\mathcal{K}' \subset \mathcal{K}$  mit  $|\mathcal{K}'| \geq 3$ . Angenommen, für jedes  $k \in \mathcal{K}'$  existiert ein statisches duales Zertifikat, d.h. es existiert  $\tilde{c}_k \in \tilde{H}$  mit

$$\begin{aligned} (B^* \tilde{c}_k)(x_i + k\tau v_i) &= \eta_i, & i &= 1, \dots, N, \\ |(B^* \tilde{c}_k)(y)| &< 1, & y &\in [0, 1]^d \setminus T_k. \end{aligned}$$

Dann nennen wir den Vektor  $c \in H$  definiert durch

$$c_k := \begin{cases} \frac{1}{|\mathcal{K}'|} \tilde{c}_k, & k \in \mathcal{K}', \\ 0 & \text{sonst} \end{cases}$$

ein  $\mathcal{K}'$ -Durchschnittszertifikat.

Für solch ein Durchschnittszertifikat  $c$  gilt per Definition

$$(\mathcal{G}^* c)(x, v) = \sum_{k \in \mathcal{K}} (B^* c_k)(x + k\tau v) = \frac{1}{|\mathcal{K}'|} \sum_{k \in \mathcal{K}'} (B^* \tilde{c}_k)(x + k\tau v)$$

für alle  $(x, v) \in \Omega$ . Offensichtlich sind die Bedingungen

$$\begin{aligned} (\mathcal{G}^* c)(x_i, v_i) &= \eta_i, & i &= 1, \dots, N, \\ |(\mathcal{G}^* c)(x, v)| &\leq 1, & (x, v) &\in \Omega \setminus T \end{aligned} \tag{3.3}$$

sofort erfüllt. Allerdings kann eine strikte Ungleichung in der zweiten Bedingung im Allgemeinen nicht mehr garantiert werden.

**3.6 Beispiel** Wir wollen in  $d = 1$  drei statische Teilchen mit Abstand  $\Delta x \in (0, 1/2)$  zueinander betrachten. Diese seien im Phasenraum gegeben durch  $(0, 0)$ ,  $(\Delta x, 0)$  und  $(2\Delta x, 0)$ . Für Gewichte  $\alpha_i = 1$  ist das Zielmaß demnach

$$\hat{\lambda} = \delta_{(0,0)} + \delta_{(\Delta x,0)} + \delta_{(2\Delta x,0)}.$$

Von dieser Konfiguration sollen Messungen zu drei Zeitpunkten aufgenommen werden, d.h.  $K = 1$ ,  $\mathcal{K} = \{-1, 0, 1\}$  und wir setzen der Einfachheit halber  $\tau = 1$ .

Angenommen, wir haben für jedes  $k \in \mathcal{K}$  ein statisches Zertifikat  $\tilde{c}_k \in \tilde{H}$  zu den Vorzeichen  $\eta = (1, 1, 1)$ . Dann gilt

$$\begin{aligned} \frac{1}{3} \sum_{k=-1}^1 (B^* \tilde{c}_k)(\Delta x + k\tau \Delta x) &= \frac{1}{3} ((B^* \tilde{c}_{-1})(0) + (B^* \tilde{c}_0)(\Delta x) + (B^* \tilde{c}_1)(2\Delta x)) \\ &= 1, \end{aligned}$$

denn zu jedem der Zeitpunkte  $k = -1, 0, 1$  befindet sich ein Teilchen auf den Positionen  $0, \Delta x, 2\Delta x$  und nach den Bedingungen an die statischen Zertifikate  $\tilde{c}_k$  müssen diese hier das Vorzeichen 1 annehmen. Die obige Rechnung zeigt, dass das Durchschnittszertifikat auch für den Punkt  $(\Delta x, \Delta x)$  im Phasenraum, welcher keinem Teilchen der eigentlichen Konfiguration entspricht, den Wert 1 annimmt. Eine analoge Rechnung gilt für den Punkt  $(\Delta x, -\Delta x)$  im Phasenraum.

**3.7 Definition** (Geisterteilchen, [2, Def. 4]) Es sei  $T = \{(x_i, v_i) \mid i = 1, \dots, N\}$  eine Konfiguration von Teilchen und  $\mathcal{K}' \subset \mathcal{K}$  eine Menge von Zeitschritten. Ein Tupel  $(g, w) \in \mathbb{R}^d \times \mathbb{R}^d$  heißt *Geisterteilchen* der Konfiguration (bzgl.  $\mathcal{K}'$ ), falls paarweise verschiedene Indizes  $(i_k)_{k \in \mathcal{K}'} \subset \{1, \dots, N\}$  existieren, sodass

$$\forall k \in \mathcal{K}' : \quad x_{i_k} + k\tau v_{i_k} = g + k\tau w.$$

Bezeichne die Menge der Geisterteilchen bzgl.  $\mathcal{K}'$  mit  $G(T, \mathcal{K}')$ .

Ein Geisterteilchen ist also ein potentielles Teilchen, welches zu jedem Zeitschritt in  $\mathcal{K}'$  von einem anderen Teilchen der Konfiguration überdeckt würde. In der Situation von Beispiel 3.6 sind zwei Geisterteilchen vorhanden, nämlich genau die Teilchen  $(\Delta x, \Delta x)$  und  $(\Delta x, -\Delta x)$ .

Definiere wie in [2] die  $d$ -dimensionalen affinen Unterräume

$$L_{i,k} := \{(x, v) \in \mathbb{R}^{2d} \mid x + k\tau v = x_i + k\tau v_i\}$$

für  $i \in \{1, \dots, N\}, k \in \mathcal{K}$ . Dann ist  $L_{i,k}$  die Menge aller Punkte im Phasenraum, deren zugehörige Teilchen am Zeitschritt  $k$  die gleiche Position einnehmen wie das  $i$ -te Teilchen der Konfiguration  $T$ . Offensichtlich lässt sich die Menge der Geisterteilchen ausdrücken als

$$G(T, \mathcal{K}') = \bigcup_{(i_k)_k} \bigcap_{k \in \mathcal{K}'} L_{i_k, k}, \quad (3.4)$$

wobei die Vereinigung über alle Folgen  $(i_k)_{k \in \mathcal{K}'} \subset \{1, \dots, N\}$  von paarweise verschiedenen Indizes läuft.

Es kann nun das Hauptresultat von Alberti et al. zur Rekonstruktion eines Phasenraummaßes durch das dynamische TV-Minimierungsproblem

$$\min_{\lambda \in \mathcal{M}(\Omega, \mathbb{K})} \|\lambda\|_{\text{TV}} \quad \text{s.t.} \quad \mathcal{G}\lambda = \hat{y} \quad (\text{ERdyn})$$

angegeben werden:

**3.8 Satz** ([2, Thm. 5]) Es sei  $T = \{(x_i, v_i) \mid i = 1, \dots, N\} \subset \Omega$  eine Konfiguration von  $N$  Teilchen,  $\alpha \in \mathbb{K}^N$  ein Gewichtsvektor und  $\mathcal{K}' \subset \mathcal{K}$  mit  $|\mathcal{K}'| \geq 3$ . Betrachte das Zielmaß

$$\hat{\lambda} = \sum_{i=1}^N \alpha_i \delta_{(x_i, v_i)} \in \mathcal{M}(\Omega, \mathbb{K}).$$

---

Es gelte

- 1) Für jedes  $k \in \mathcal{K}'$  und jedes  $\eta \in \text{sgn } \mathbb{K}^N$  existiere ein statisches duales Zertifikat  $\tilde{c}_k \in \tilde{H}$ , welches die Bedingungen (3.2) erfülle
- 2) Die Konfiguration lasse keine Geisterteilchen bzgl.  $\mathcal{K}'$  zu.

Dann ist  $\hat{\lambda}$  die eindeutige Lösung von (ERdyn) zu den Daten  $\hat{y} = \mathcal{G}\hat{\lambda}$ .

*Beweis.* Es ist zu zeigen, dass für einen beliebigen Vorzeichenvektor  $\eta \in \text{sgn } \mathbb{K}^N$  das  $\mathcal{K}'$ -Durchschnittszertifikat ein gültiges duales Zertifikat für das dynamische Rekonstruktionsproblem ist. Wie schon beobachtet, gelten für das Durchschnittszertifikat die Bedingungen (3.3) und es muss nur noch gezeigt werden, dass in der zweiten Bedingung auch die strikte Ungleichung gilt, um Satz 2.15 anwenden zu können.

Es sei also  $\eta \in \text{sgn } \mathbb{K}^N$  beliebig und  $c \in H$  das zugehörige  $\mathcal{K}'$ -Durchschnittszertifikat. Wegen

$$(\mathcal{G}^*c)(x, v) = \frac{1}{|\mathcal{K}'|} \sum_{k \in \mathcal{K}'} (B^*\tilde{c}_k)(x + k\tau v), \quad (x, v) \in \Omega$$

ist  $\mathcal{G}^*c$  der Mittelwert von Summanden mit Betrag kleiner 1. Ist also  $(x, v) \in \Omega$  ein Punkt mit  $|(\mathcal{G}^*c)(x, v)| = 1$ , so muss auch

$$|(B^*\tilde{c}_k)(x + k\tau v)| = 1 \quad \forall k \in \mathcal{K}'$$

gelten. Die statischen dualen Zertifikate erfüllen die strikte Ungleichung außerhalb des Trägers von  $\hat{u}_k$ , also existiert für jedes  $k \in \mathcal{K}'$  ein  $i_k \in \{1, \dots, N\}$  mit

$$x + k\tau v = x_{i_k} + k\tau v_{i_k}. \quad (3.5)$$

Es können nun zwei Fälle eintreten:

**Fall 1:** Ein Index  $i_k$  tritt mehrfach auf. Die Variablen  $(x, v)$  werden durch die lineare Gleichung (3.5) zu zwei Zeitpunkten eindeutig festgelegt und es muss  $(x, v) = (x_{i_k}, v_{i_k}) \in T$  gelten.

**Fall 2:** Die Indizes  $(i_k)_{k \in \mathcal{K}'}$  sind paarweise verschieden. Dann wäre  $(x, v)$  ein Geisterteilchen für die Teilchenkonfiguration bezüglich  $\mathcal{K}'$ , was im Widerspruch zur Annahme stünde.

Wir haben gezeigt, dass  $|(\mathcal{G}^*c)(x, v)| < 1$  für alle  $(x, v) \in \Omega \setminus T$ . Die exakte Rekonstruktion für (ERdyn) folgt mit Satz 2.15.  $\square$

**3.9 Bemerkung** Die Voraussetzungen des Satzes fordern, dass zu mindestens drei Zeitschritten statische duale Zertifikate existieren. Dies bedeutet, dass zu diesen Zeitschritten auch die statischen Probleme (ER $k$ ) die Maße  $\hat{u}_k$  exakt rekonstruieren können. Eigentlich erwartet man jedoch, dass die Rekonstruktion über das dynamische Problem, welches alle Zeitschritte gleichzeitig betrachtet, auch in Situationen rekonstruieren kann, in denen die statischen Rekonstruktionen scheitern.

Alberti et al. zeigen hierzu nur ein numerisches Beispiel, welches letztlich auf der Konstruktion aus Abschnitt 2.4 basiert.

Ein Vorteil gegenüber einzelnen statischen Rekonstruktionen besteht jedoch auch unter diesen Voraussetzungen: Die gleichzeitige Rekonstruktion der Geschwindigkeiten. Außerdem muss bei einer Aufnahme von mehr als drei Bildern ( $|\mathcal{K}| > 3$ ) nicht bekannt sein, *welche* Teilmenge  $\mathcal{K}' \subset \mathcal{K}$  die Voraussetzungen erfüllt.

Wie auch in Beispiel 3.6 zu sehen war, ist für die Existenz von Geisterteilchen eine bestimmte Symmetrie in der Teilchenkonfiguration erforderlich, damit ein Geisterteilchen in jedem Zeitschritt von einer Position zur nächsten springen kann. Es ist deshalb intuitiv ersichtlich, dass Teilchenkonfigurationen, welche ein Geisterteilchen zulassen, eher selten sind.

Dazu zitieren wir abschließend noch folgende Aussage, welche diesen Gedanken formalisiert:

**3.10 Proposition** ([2, Prop. 7]) Es sei  $|\mathcal{K}'| \geq 3$ . Die Variablen  $(X_i, V_i), i = 1, \dots, N$  seien unabhängige Zufallsvariablen, welche aus einer absolutstetigen Verteilung bezüglich des Lebesgue-Maßes auf  $\Omega$  gezogen seien. Dann lässt die Konfiguration  $\{(X_i, Y_i) \mid i = 1, \dots, N\}$  fast sicher keine Geisterteilchen bezüglich  $\mathcal{K}'$  zu.

## 4. Dimensionsreduktion

Durch das Phasenraum-Lifting hat sich die Dimension des dynamischen Problems im Vergleich zum statischen verdoppelt. Dies kann zu numerischen Problemen führen, da z.B. bei einer Gitter-Diskretisierung die Anzahl der Gitterpunkte nun mit  $s^{2d}$  skaliert, wenn die Anzahl der Gitterpunkte in einer Dimension mit  $s$  bezeichnet wird. Es stellt sich die Frage, ob die Dimension des dynamischen Problems reduziert werden kann, ohne alle Vorteile der gemeinsamen Rekonstruktion an mehreren Zeitschritten zu verlieren. In diesem Kapitel soll ein geeignetes Verfahren entwickelt werden.

### 4.1. Radon-Transformation für Maße

Eine typische Idee im Zusammenhang mit Dimensionsreduktion ist die Projektion von hochdimensionalen Daten auf einen Unterraum von niedrigerer Dimension. Eng mit dieser Idee verbunden ist die Radon-Transformation: Sei für  $\theta \in \mathbb{S}^{d-1}$  und  $s \in \mathbb{R}$  mit  $H_{\theta,s}$  die affine Hyperebene  $\{x \in \mathbb{R}^d \mid \theta \cdot x = s\}$  bezeichnet. Für solch eine Hyperebene sei  $\sigma$  das zugehörige Oberflächenmaß. Dann ist die Radon-Transformation einer Funktion  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  definiert durch

$$\mathcal{R}f : \mathbb{S}^{d-1} \times \mathbb{R} \rightarrow \mathbb{R}, \quad (\mathcal{R}f)(\theta, s) = \int_{H_{\theta,s}} f(x) d\sigma(x),$$

solange diese Integrale wohldefiniert sind. Die klassische Radon-Transformation repräsentiert also eine Funktion durch die Sammlung ihrer Integrale über alle affinen Hyperebenen.

Es sei  $\omega_{d-1}$  wieder das normierte Oberflächenmaß auf der Einheitssphäre in  $\mathbb{R}^d$ . Die nun folgende Definition erlaubt uns, die Radon-Transformation auch auf Maße in unserem Setting anzuwenden.

**4.1 Definition** (Radon-Transformation) Es sei für  $\theta \in \mathbb{S}^{d-1}$  die Linienprojektion  $R_\theta : \mathbb{R}^d \rightarrow \mathbb{R}$  definiert durch  $R_\theta x := \theta \cdot x$ . Die *Radon-Transformation zur Richtung*  $\theta$  ist dann die Abbildung

$$(R_\theta)_\# : \mathcal{M}(\mathbb{R}^d) \rightarrow \mathcal{M}(\mathbb{R}), \quad \xi \mapsto (R_\theta)_\# \xi,$$

welche ein Maß auf sein Bildmaß unter  $R_\theta$  schickt.

Ist ein Maß  $\xi \in \mathcal{M}(\mathbb{R}^d)$  gegeben, so definiert die Abbildung  $\theta \mapsto (R_\theta)_\# \xi$  einen Übergangskern  $\mathbb{S}^{d-1} \rightarrow \mathbb{R}$ . Die *Radon-Transformation* von  $\xi$  ist definiert als das

Maß

$$\mathcal{R}\xi := \omega_{d-1} \otimes (R_\theta)_\# \xi \in \mathcal{M}(\mathbb{S}^{d-1} \times \mathbb{R}).$$

Der Zusammenhang zur klassischen Radon-Transformation für Funktionen soll kurz skizziert werden: In der Verallgemeinerung von Funktionen auf Maße entspricht eine Funktion  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  dem Maß  $f \, d\ell$  mit Dichte  $f$  bezüglich des  $d$ -dimensionalen Lebesgue-Maßes  $\ell$ . Für eine beliebige Borelmenge  $A \in \mathcal{B}(\mathbb{R})$  gilt

$$\begin{aligned} (R_\theta)_\# f \, d\ell(A) &= f \, d\ell(\{x \in \mathbb{R}^d \mid x \cdot \theta \in A\}) \\ &= \int_{x: x \cdot \theta \in A} f \, d\ell = \int_A \int_{H_{\theta,s}} f(x) \, d\sigma(x) \, ds. \end{aligned}$$

Hierbei folgen die ersten Gleichungen aus der Definition des Bildmaßes bzw. eines Maßes mit Dichte. Die letzte Gleichheit entspricht dem Satz von Fubini, falls  $\theta$  Element der Standardbasis von  $\mathbb{R}^d$  ist. Der Fall für allgemeine  $\theta$  kann mithilfe der Rotationsinvarianz des Lebesgue-Maßes hierauf zurückgeführt werden. An dieser Rechnung ist abzulesen, dass die Radon-Transformation des Maßes  $f \, d\ell$  gerade die Dichte  $(\mathcal{R}f)(\theta, \cdot)$  bezüglich des eindimensionalen Lebesgue-Maßes hat, was die Konsistenz der obigen Definition mit der klassischen Radon-Transformation für Funktionen belegt.

Nimmt man von der Radon-Transformation einer Funktion zur Richtung  $\theta$  die eindimensionale Fourier-Transformation bezüglich des zweiten Argumentes, so erhält man dasselbe Ergebnis, als wenn man die  $d$ -dimensionale Fourier-Transformation der Funktion entlang der Richtung  $\theta$  auswertet. Dieser Zusammenhang ist bekannt als Fourier-Slice Theorem und gilt auch für die Radon-Transformation für Maße:

**4.2 Satz** (Fourier-Slice, [5, Lem. 2]) Es sei  $\xi \in \mathcal{M}(\mathbb{R}^d)$ . Dann gilt für alle  $p \in \mathbb{R}, \theta \in \mathbb{S}^{d-1}$ :

$$(\tilde{\mathcal{F}}(R_\theta)_\# \xi)(p) = (\tilde{\mathcal{F}}\xi)(p\theta).$$

*Beweis.* Dies folgt durch Nachrechnen und der Transformationsformel für das Bildmaß (Satz 1.8):

$$\begin{aligned} (\tilde{\mathcal{F}}(R_\theta)_\# \xi)(p) &= \int_{\mathbb{R}} e^{-ips} \, d(R_\theta)_\# \xi(s) \\ &= \int_{\mathbb{R}^d} e^{-ip\theta \cdot x} \, d\xi(x) = (\tilde{\mathcal{F}}\xi)(p\theta). \end{aligned}$$

□

Die obige Aussage hilft uns dabei, die Injektivität der Radon-Transformation zu beweisen.

**4.3 Satz** (vgl. [5, Thm. A]) Es sei  $\xi \in \mathcal{M}(\mathbb{R}^d)$  mit  $(R_\theta)_\# \xi = 0$  für  $\omega_{d-1}$ -fast alle  $\theta \in \mathbb{S}^{d-1}$ . Dann ist  $\xi = 0$ .

---

*Beweis.* Für jedes  $\theta \in \mathbb{S}^{d-1}$  mit  $(R_\theta)_\# \xi = 0$  gilt auch  $\tilde{\mathcal{F}}(R_\theta)_\# \xi = 0$ . Mit dem Fourier-Slice Theorem folgt  $(\tilde{\mathcal{F}}\xi)(p\theta) = 0$  für solch ein  $\theta$  und alle  $p \in \mathbb{R}$ .

Da die Radon-Transformation zu fast allen Richtungen  $\theta$  verschwindet, ist die Menge  $\{p\theta \mid p \in \mathbb{R}, (R_\theta)_\# \xi = 0\}$  dicht in  $\mathbb{R}^d$ . Aufgrund der Stetigkeit von  $\tilde{\mathcal{F}}\xi$  folgt somit  $\tilde{\mathcal{F}}\xi = 0$ . Da die Fourier-Transformation injektiv ist, muss schließlich  $\xi = 0$  gelten.  $\square$

Durch die so definierte Radon-Transformation lässt sich noch keine Dimensionsreduktion erreichen: Würde man eine Gitter-Diskretisierung auf einer Teilmenge von  $\mathbb{S}^{d-1} \times \mathbb{R}$  ansetzen, so würde die Anzahl der Gitterpunkte mit  $s^{d-1} \cdot s = s^d$  skalieren – genauso schnell wie bei einem Gitter auf  $\mathbb{R}^d$ . Für die Reduktion wird deshalb eine leichte Abwandlung der Radon-Transformation benötigt.

**4.4 Definition** Es sei für  $\theta \in \mathbb{S}^{d-1}$  die gemeinsame Linienprojektion definiert durch  $\hat{R}_\theta : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^2, \hat{R}_\theta(x, v) := (\theta \cdot x, \theta \cdot v)$ . Die *gemeinsame Radon-Transformation zur Richtung  $\theta$*  ist dann die Abbildung

$$(R_\theta)_\# : \mathcal{M}(\mathbb{R}^d \times \mathbb{R}^d) \rightarrow \mathcal{M}(\mathbb{R}^2), \quad \xi \mapsto (\hat{R}_\theta)_\# \xi.$$

Ist ein Maß  $\xi \in \mathcal{M}(\mathbb{R}^d \times \mathbb{R}^d)$  gegeben, so definiert die Abbildung  $\theta \mapsto (\hat{R}_\theta)_\# \xi$  einen Übergangskern  $\mathbb{S}^{d-1} \rightarrow \mathbb{R}^2$  und die *gemeinsame Radon-Transformation* von  $\xi$  ist analog definiert als das Maß

$$\hat{\mathcal{R}}\xi := \omega_{d-1} \otimes (\hat{R}_\theta)_\# \xi \in \mathcal{M}(\mathbb{S}^{d-1} \times \mathbb{R}^2).$$

Eine naive Zählung der Dimensionen zeigt eine Reduktion der Dimension von  $2d$  auf  $(d-1) + 2 = d+1$  durch die gemeinsame Radon-Transformation. Diese Reduktion soll nun auf das dynamische Problem umgesetzt werden.

## 4.2. Problemformulierung

Bevor wir ein dimensionsreduziertes Problem erarbeiten können, benötigen wir noch einige Vorbereitungen. Definiere für  $n \in \mathbb{S}^1$  und eine Dimension  $l \in \mathbb{N}$  die Abbildung

$$Q_n^l : \mathbb{R}^l \times \mathbb{R}^l \rightarrow \mathbb{R}^l, \quad (a, b) \mapsto n_1 a + n_2 b.$$

Weiter seien für jeden Zeitschritt  $k \in \mathcal{K}$  der Skalierungsfaktor  $s_k := \sqrt{1 + (k\tau)^2}$  und der Richtungsvektor

$$n_k := \frac{1}{s_k} \begin{pmatrix} 1 \\ k\tau \end{pmatrix} \in \mathbb{S}^1$$

definiert. Zur Erinnerung geben wir auch die Abbildung

$$D_k^l : \mathbb{R}^l \times \mathbb{R}^l \rightarrow \mathbb{R}^l, \quad (x, v) \mapsto x + k\tau v$$

aus dem vorherigen Kapitel nocheinmal an, welche von der Phasenraumrepräsentation zu der Position am  $k$ -ten Zeitschritt wechselt. Im folgenden Lemma sind einige

Rechenregeln aufgelistet, welche später die Kommutation verschiedener Bildmaße erlauben werden:

**4.5 Lemma** Für alle  $k \in \mathcal{K}, \theta \in \mathbb{S}^{d-1}$  und  $n \in \mathbb{S}^1$  gilt:

1.  $D_k^l = [z \mapsto z s_k] \circ Q_{n_k}^l$
2.  $R_\theta \circ Q_n^d = Q_n^1 \circ \hat{R}_\theta$
3.  $R_\theta \circ D_k^d = D_k^1 \circ \hat{R}_\theta$

*Beweis.* Zu Punkt 1: Für alle  $(a, b) \in \mathbb{R}^l \times \mathbb{R}^l$  gilt

$$D_k^l(a, b) = a + k\tau b = s_k \cdot \frac{1}{s_k}(a + k\tau b) = [z \mapsto z s_k] \circ Q_{n_k}^l(a, b).$$

Zu Punkt 2: Für alle  $(x, v) \in \mathbb{R}^d \times \mathbb{R}^d$  gilt

$$R_\theta \circ Q_n^d(x, v) = R_\theta(n_1 x + n_2 v) = n_1 \theta \cdot x + n_2 \theta \cdot v = Q_n^1 \circ \hat{R}_\theta(x, v).$$

Zu Punkt 3: Dies folgt direkt aus den vorherigen beiden Aussagen, da die Skalierung um  $s_k$  offensichtlich mit allen Abbildungen kommutiert.  $\square$

Um die Notation zu vereinfachen, lassen wir in diesem Abschnitt die explizite Abhängigkeit von  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{C}$  fallen, auch wenn alle Probleme natürlich auch für reellwertige Maße formuliert werden könnten. Zur Erinnerung wird hier das volldimensionale dynamische Problem noch einmal angegeben: Für ein Zielmaß

$$\hat{\lambda} = \sum_{i=1}^N \alpha_i \delta_{(x_i, v_i)}$$

im Phasenraum wird das Problem

$$\min_{\lambda \in \mathcal{M}(\Omega)} \|\lambda\|_{\text{TV}} \quad \text{s.t.} \quad \mathcal{G}\lambda = \hat{y} \quad (\text{ERdyn})$$

zu den Daten  $\hat{y} = \mathcal{G}\hat{\lambda}$  betrachtet.

Die Idee ist es nun, in obigem Problem stattdessen die gemeinsame Radon-Transformation eines Phasenraummaßes als Variable zu nutzen. Hierzu könnte  $\lambda \in \mathcal{M}(\Omega)$  durch eine Variable  $\mu \in \mathcal{M}(\mathbb{S}^{d-1} \times \mathbb{R}^2)$  ersetzt werden. Somit wäre die Anwendung der TV-Norm auf die neue Variable wohldefiniert und im dimensionsreduzierten Problem könnte stattdessen über  $\|\mu\|_{\text{TV}}$  minimiert werden. Der Nachteil dieses Ansatzes ist, dass Auswertungen  $\mu(\theta, \cdot)$  für einzelne Richtungen  $\theta \in \mathbb{S}^{d-1}$  dann im Allgemeinen keinen Sinn mehr ergeben.

Stattdessen wird im Folgenden eine Familie von Maßen  $(\mu_\theta)_{\theta \in \mathbb{S}^{d-1}} \subset \mathcal{M}(\mathbb{R}^2)$  eingeführt, wobei wir uns vorstellen, dass diese als  $\mu_\theta = (\hat{R}_\theta)_\# \lambda$  aus einem Phasenraummaß  $\lambda \in \mathcal{M}(\Omega)$  hervorgehen.

Um das volldimensionale Maß aus dem Problem zu eliminieren, muss die Nebenbedingung in (ERdyn) angepasst werden. Als erster Schritt werden die Konfigurationen  $(u_k)_{k \in \mathcal{K}} \subset \mathcal{M}([0, 1]^d)$  zu den einzelnen Zeitschritten wieder als Variablen eingeführt und wir erhalten statt  $\mathcal{G}\lambda = \hat{y}$  die Gleichungen

$$\begin{aligned} Bu_k &= \hat{y}_k & \forall k \in \mathcal{K}, \\ u_k &= (D_k^d)_{\#} \lambda & \forall k \in \mathcal{K}. \end{aligned}$$

Um in der zweiten Gleichung das volldimensionale Maß  $\lambda$  durch die gemeinsame Radon-Transformation ersetzen zu können, wird auf beiden Seiten zunächst die einfache Radon-Transformation zu einer Richtung  $\theta \in \mathbb{S}^{d-1}$  angewandt:

$$(R_\theta)_{\#} u_k = (R_\theta)_{\#} (D_k^d)_{\#} \lambda = (D_k^1)_{\#} (\hat{R}_\theta)_{\#} \lambda$$

Hier wurde die Rechenregel  $g_{\#} f_{\#} = (g \circ f)_{\#}$  für die Hintereinanderausführung des Pushforwards in Verbindung mit Lemma 4.5 genutzt. Bis jetzt ist durch die Ersetzung der Nebenbedingung noch keine Information verloren gegangen, da die (einfache) Radon-Transformation injektiv ist.

Ein erster Ansatz für ein dimensionsreduziertes Problem ist nun

$$\begin{aligned} \min_{\substack{(\mu_\theta)_{\theta \in \mathbb{S}^{d-1}} \subset \mathcal{M}(\mathbb{R}^2) \\ (u_k)_{k \in \mathcal{K}} \subset \mathcal{M}([0, 1]^d)}} \sup_{\theta \in \mathbb{S}^{d-1}} \|\mu_\theta\|_{\text{TV}} \quad \text{s.t.} \\ Bu_k &= \hat{y}_k & \forall k \in \mathcal{K}, \\ (R_\theta)_{\#} u_k &= (D_k^1)_{\#} \mu_\theta & \forall k \in \mathcal{K} \forall \theta \in \mathbb{S}^{d-1}, \\ \exists \lambda \in \mathcal{M}(\Omega) : \forall \theta \in \mathbb{S}^{d-1} : \mu_\theta &= (\hat{R}_\theta)_{\#} \lambda. \end{aligned} \tag{4.1}$$

Die letzte Bedingung soll sicherstellen, dass  $\mu$  tatsächlich aus einem Phasenraum  $\lambda$  hervorgeht. Allerdings ist nicht klar, wie diese Bedingung numerisch zu prüfen ist.

Bevor wir fortfahren, benötigen wir noch ein Lemma:

**4.6 Lemma** Es sei  $\lambda \in \mathcal{M}(\Omega)$  gegeben und  $\mu_\theta := (\hat{R}_\theta)_{\#} \lambda$  die gemeinsame Radon-Transformation von  $\lambda$  zu  $\theta \in \mathbb{S}^{d-1}$ . Dann existiert zu jedem  $n \in \mathbb{S}^1$  ein Maß  $\nu_n \in \mathcal{M}(\mathbb{R}^d)$  mit

$$(R_\theta)_{\#} \nu_n = (Q_n^1)_{\#} \mu_\theta \quad \forall \theta \in \mathbb{S}^{d-1}. \tag{4.2}$$

*Beweis.* Wähle  $\nu_n := (Q_n^d)_{\#} \lambda$ ,  $n \in \mathbb{S}^1$ . Dann gilt für alle  $\theta \in \mathbb{S}^{d-1}$ ,  $n \in \mathbb{S}^1$  nach Lemma 4.5

$$(R_\theta)_{\#} \nu_n = (R_\theta)_{\#} (Q_n^d)_{\#} \lambda = (Q_n^1)_{\#} (\hat{R}_\theta)_{\#} \lambda = (Q_n^1)_{\#} \mu_\theta.$$

□

Es bieten sich nun zwei Alternativen, um die letzte Bedingung in Gleichung (4.1) zu umgehen:

1. **Verzicht auf die letzte Bedingung:** Das resultierende Problem lautet

$$\begin{aligned} & \min_{\substack{(\mu_\theta)_{\theta \in \mathbb{S}^{d-1}} \subset \mathcal{M}(\mathbb{R}^2) \\ (u_k)_{k \in \mathcal{K}} \subset \mathcal{M}([0,1]^d)}} \sup_{\theta \in \mathbb{S}^{d-1}} \|\mu_\theta\|_{\text{TV}} \quad \text{s.t.} \\ & Bu_k = \hat{y}_k \quad \forall k \in \mathcal{K}, \\ & (R_\theta)_\# u_k = (D_k^1)_\# \mu_\theta \quad \forall k \in \mathcal{K} \forall \theta \in \mathbb{S}^{d-1}. \end{aligned}$$

Die Existenz eines zu  $\mu$  passenden Phasenraummaßes wird nicht mehr erzwungen. Allerdings wird in Abschnitt 4.4 gezeigt, dass unter bestimmten Voraussetzungen die gemeinsame Radon-Transformation des Zielmaßes  $\hat{\lambda}$  trotzdem exakt rekonstruiert wird.

2. **Relaxierung der Bedingung:** Die Bedingung wird abgeschwächt, indem die Beobachtung aus Lemma 4.6 genutzt wird. Statt der Existenz eines zu  $\mu$  passenden Phasenraummaßes  $\lambda$  fordern wir also die Existenz einer Familie von Maßen  $(\nu_n)_{n \in \mathbb{S}^1} \subset \mathcal{M}(\mathbb{R}^d)$ , sodass Bedingung (4.2) erfüllt ist. Der Vorteil gegenüber der ursprünglichen Bedingung liegt darin, dass die Dimension der Familie von Variablen  $(\nu_n)_{n \in \mathbb{S}^1}$  nur noch  $d + 1$  beträgt. Das Problem lautet dann

$$\begin{aligned} & \min_{\substack{(\mu_\theta)_{\theta \in \mathbb{S}^{d-1}} \subset \mathcal{M}(\mathbb{R}^2) \\ (u_k)_{k \in \mathcal{K}} \subset \mathcal{M}([0,1]^d) \\ (\nu_n)_{n \in \mathbb{S}^1} \subset \mathcal{M}(\mathbb{R}^d)}} \sup_{\theta \in \mathbb{S}^{d-1}} \|\mu_\theta\|_{\text{TV}} \quad \text{s.t.} \\ & Bu_k = \hat{y}_k \quad \forall k \in \mathcal{K}, \\ & (R_\theta)_\# u_k = (D_k^1)_\# \mu_\theta \quad \forall k \in \mathcal{K}, \forall \theta \in \mathbb{S}^{d-1}, \\ & (R_\theta)_\# \nu_n = (Q_n^1)_\# \mu_\theta \quad \forall n \in \mathbb{S}^1, \forall \theta \in \mathbb{S}^{d-1}. \end{aligned} \tag{ERdyn_u^{\text{rel}}}$$

#### 4.2.1. Elimination von $u$

In der Version des Problems mit relaxierter Nebenbedingung lassen sich die Variablen  $u_k$  noch aus dem Problem eliminieren.

**4.7 Lemma** Es seien  $(\mu_\theta)_{\theta \in \mathbb{S}^{d-1}}, (u_k)_{k \in \mathcal{K}}, (\nu_n)_{n \in \mathbb{S}^1}$  zulässig für  $(\text{ERdyn}_u^{\text{rel}})$ . Dann gilt für alle  $k \in \mathcal{K}$

$$u_k = [x \mapsto s_k x]_\# \nu_{n_k}.$$

*Beweis.* Es sei  $k \in \mathcal{K}$  beliebig. Aus der dritten Nebenbedingung mit  $n = n_k$  folgt für alle  $\theta \in \mathbb{S}^{d-1}$

$$(R_\theta)_\# \nu_{n_k} = (Q_{n_k}^1)_\# \mu_\theta = [z \mapsto z/s_k]_\# (D_k^1)_\# \mu_\theta,$$

wobei für die zweiten Gleichheit Lemma 4.5 genutzt wurde. Aus der zweiten Ne-

benbedingung folgt weiter

$$\begin{aligned} [z \mapsto z/s_k]_{\#}(D_k^1)_{\#}\mu_{\theta} &= [z \mapsto z/s_k]_{\#}(R_{\theta})_{\#}u_k \\ &= (R_{\theta})_{\#}[x \mapsto \frac{1}{s_k}x]_{\#}u_k. \end{aligned}$$

Wir haben also die Gleichung

$$(R_{\theta})_{\#}\nu_{n_k} = (R_{\theta})_{\#}[x \mapsto \frac{1}{s_k}x]_{\#}u_k \quad \forall \theta \in \mathbb{S}^{d-1}.$$

Aus der Injektivität der Radon-Transformation (Satz 4.3) folgt die Behauptung.  $\square$

Die Variablen  $(u_k)_{k \in \mathcal{K}}$  können somit aus dem Minimierungsproblem eliminiert werden, indem in der ersten Nebenbedingung statt  $Bu_k = \hat{y}_k$  die Gleichheit

$$B[x \mapsto s_k x]_{\#}\nu_{n_k} = \hat{y}_k \quad \forall k \in \mathcal{K}$$

gefordert wird und die zweite Nebenbedingung fallen gelassen wird.

Eine letzte Anpassung muss noch vorgenommen werden: Da der Messoperator  $B$  nur für Maße auf  $[0, 1]^d$  definiert ist, muss erzwungen werden, dass die Variablen  $\nu_n$  geeignete Träger haben. Mit Blick auf den Beweis von Lemma 4.6 ist klar, dass die Variablen als

$$\nu_n = (Q_n^d)_{\#}\lambda \tag{4.3}$$

aus dem Phasenraummaß  $\lambda \in \mathcal{M}(\Omega)$  hervorgehen sollen. Für  $n \in \mathbb{S}^1$  ist  $\nu_n$  definiert durch Gleichung (4.3) konzentriert auf der Menge  $Q_n^d(\Omega)$ . Eine natürlicher Raum für die Variablen  $(\nu_n)$  ist daher

$$(\nu_n)_{n \in \mathbb{S}^1} \in \prod_{n \in \mathbb{S}^1} \mathcal{M}(Q_n^d(\Omega)) =: \mathcal{Z}.$$

Das resultierende Problem lautet

$$\begin{aligned} \min_{\substack{(\mu_{\theta})_{\theta \in \mathbb{S}^{d-1}} \subset \mathcal{M}(\mathbb{R}^2) \\ (\nu_n)_{n \in \mathbb{S}^1} \in \mathcal{Z}}} \sup_{\theta \in \mathbb{S}^{d-1}} \|\mu_{\theta}\|_{\text{TV}} \quad \text{s.t.} \\ B[x \mapsto s_k x]_{\#}\nu_{n_k} = \hat{y}_k \quad \forall k \in \mathcal{K}, \\ (R_{\theta})_{\#}\nu_n = (Q_n^1)_{\#}\mu_{\theta} \quad \forall n \in \mathbb{S}^1 \forall \theta \in \mathbb{S}^{d-1}. \end{aligned} \tag{ERdyn<sup>rel</sup>}$$

#### 4.2.2. Geschwindigkeitsproblem

Die Gewichte der Teilchen sowie ihre Positionen zu allen Zeitschritten lassen sich an den Lösungen für die Variablen  $u_k$  (bzw.  $[z \mapsto z s_k]_{\#}\nu_{n_k}$  im Problem (ERdyn<sup>rel</sup>)) direkt ablesen, vorausgesetzt, diese werden durch das jeweilige Minimierungsproblem exakt rekonstruiert. Es ist jedoch wünschenswert, auch im dimensionsreduzierten Problem noch die Geschwindigkeiten rekonstruieren zu können.

Für das Zielmaß im Phasenraum  $\hat{\lambda} = \sum_{i=1}^N \alpha_i \delta_{(x_i, v_i)}$  ist die gemeinsame Radon-Transformation gegeben durch

$$\hat{\mu}_\theta := (\hat{R}_\theta)_\# \hat{\lambda} = \sum_{i=1}^N \alpha_i \delta_{(\theta \cdot x_i, \theta \cdot v_i)}.$$

Es sei

$$\mathcal{M}_1(\mathbb{R}^2) := \{\xi \in \mathcal{M}(\mathbb{R}^2) \mid \int_{\mathbb{R}^2} |w| d|\xi|(y, w) < \infty\}$$

die Menge der komplexwertigen Maße auf  $\mathbb{R}^2$ , welche endlichen ersten Moment in der zweiten Komponente haben. Die Abbildung

$$\begin{aligned} \mathcal{W} : \mathcal{M}_1(\mathbb{R}^2) &\rightarrow \mathcal{M}(\mathbb{R}^2), \quad \xi \mapsto \mathcal{W}\xi \\ \text{mit } \mathcal{W}\xi(E) &= \int_E w d\xi(y, w) \text{ für } E \in \mathcal{B}(\mathbb{R}^2) \end{aligned}$$

angewendet auf  $\hat{\mu}_\theta$  ergibt

$$\mathcal{W}\hat{\mu}_\theta = \sum_{i=1}^N \alpha_i (\theta \cdot v_i) \delta_{(\theta \cdot x_i, \theta \cdot v_i)}.$$

Schließlich gilt

$$[(y, w) \mapsto y]_\# \mathcal{W}\hat{\mu}_\theta = \sum_{i=1}^N \alpha_i (\theta \cdot v_i) \delta_{(\theta \cdot x_i)}.$$

Ein Ansatz zur Rekonstruktion der Geschwindigkeiten ist nun die Lösung des Problems

$$(R_\theta)_\# \sum_{j=1}^d \theta_j V_j = [(y, w) \mapsto y]_\# \mathcal{W}\mu_\theta \quad (\text{ERvel})$$

für Maße  $V_1, \dots, V_d \in \mathcal{M}(\Omega)$ , wobei in die rechte Seite eine Lösung  $\mu_\theta$  des dimensionsreduzierten Problems eingesetzt werden kann. Die Idee ist folgende: Setzt man in die rechte Seite  $\mu_\theta = \hat{\mu}_\theta$  ein, so ist eine Lösung gegeben durch

$$V_j = \sum_{i=1}^N \alpha_i v_i^{(j)} \delta_{x_i}, \quad j = 1, \dots, d,$$

wobei  $v_i^{(j)}$  für die  $j$ -te Komponente der wahren Geschwindigkeit des  $i$ -ten Teilchens stehe. An den Maßen  $V_j$  könnten die Geschwindigkeiten des  $i$ -ten Teilchens abgelesen werden, falls Gewichte  $\alpha_i$  und Positionen  $x_i$  schon bekannt sind.

Im Rest dieser Arbeit wollen wir uns jedoch auf die dimensionsreduzierte Rekonstruktion von Gewichten und Positionen konzentrieren und werden das Geschwindigkeitsproblem nicht näher untersuchen.

### 4.3. Träger der neuen Variablen

Im letzten Abschnitt wurden zum Zweck der Dimensionsreduktion die Variablen  $(\mu_\theta)_{\theta \in \mathbb{S}^{d-1}}$  sowie  $(\nu_n)_{n \in \mathbb{S}^1}$  neu eingeführt. Hierbei ist stets der Gedanke, dass ein volldimensionales Maß  $\lambda \in \mathcal{M}(\Omega)$  existiert, sodass

$$\mu_\theta = (\hat{R}_\theta)_\# \lambda \quad \text{und} \quad \nu_n = (Q_n^d)_\# \lambda$$

für alle  $\theta \in \mathbb{S}^{d-1}$  bzw. alle  $n \in \mathbb{S}^1$  gilt. Aus der Definition des Bildmaßes folgt sofort, dass die Variablen  $\mu_\theta$  konzentriert auf  $\hat{R}_\theta(\Omega) \subset \mathbb{R}^2$  und  $\nu_n$  konzentriert auf  $Q_n^d(\Omega) \subset \mathbb{R}^d$  sind. Diese Mengen sind jeweils kompakt, da die Abbildungen  $\hat{R}_\theta$  bzw.  $Q_n^d$  stetig sind und  $\Omega$  kompakt ist. Für die numerische Umsetzung der Verfahren ist es nützlich, konkrete Ausdrücke für diese Mengen herzuleiten, um später ein geeignetes Gitter zur Diskretisierung der Maße generieren zu können.

Zur Erinnerung wird hier der Parameterraum für das dynamische Problem noch einmal angegeben: Dieser war definiert durch

$$\Omega = \{(x, v) \in \mathbb{R}^d \times \mathbb{R}^d \mid \forall t \in [-\delta, \delta] : x + tv \in [0, 1]^d\}.$$

Es sei

$$\Omega_1 := \{(x, v) \in \mathbb{R}^2 \mid \forall t \in [-\delta, \delta] : x + tv \in [0, 1]\}$$

die entsprechende Menge in einer Dimension. Offensichtlich ist  $(x, v) \in \Omega$  genau dann, wenn für jede Komponente  $j = 1, \dots, d$  gilt, dass  $(x^{(j)}, v^{(j)}) \in \Omega_1$ . Bis auf Umordnung der Komponenten ist  $\Omega$  also das  $d$ -fache kartesische Produkt aus  $\Omega_1$ .

Durch Umformung der Ungleichung aus der Definition von  $\Omega_1$  für  $t = -\delta$  und  $t = \delta$  ist abzulesen, dass  $(x, v) \in \Omega_1$  genau dann, wenn die Bedingungen

$$\begin{aligned} 0 &\leq x \leq 1, \\ \max\left\{\frac{-x}{\delta}, \frac{x-1}{\delta}\right\} &\leq v \leq \min\left\{\frac{x}{\delta}, \frac{1-x}{\delta}\right\} \end{aligned} \tag{4.4}$$

erfüllt sind. Da  $\Omega_1$  durch die linearen Ungleichungen (4.4) begrenzt ist, ist  $\Omega_1$  ein Schnitt von vier affinen Halbebenen und somit ein Polytop in  $\mathbb{R}^2$ . Genauer ist  $\Omega_1$  ein Parallelogramm mit Eckpunkten

$$S := \left\{ (0, 0), (1, 0), \left(\frac{1}{2}, \frac{1}{2\delta}\right), \left(\frac{1}{2}, -\frac{1}{2\delta}\right) \right\},$$

d.h. es gilt  $\Omega_1 = \text{conv}(S)$ .

Nun kann eine konkrete Darstellung von  $\hat{R}_\theta(\Omega)$  angegeben werden.

**4.8 Lemma** Es sei  $\theta \in \mathbb{S}^{d-1}$  beliebig und  $s_+ := \sum_{j:\theta_j > 0} \theta_j$ ,  $s_- := \sum_{j:\theta_j < 0} \theta_j$ .

Die Menge  $\hat{R}_\theta(\Omega)$  ist ein Parallelogramm in  $\mathbb{R}^2$ , dessen Eckpunkte abzulesen sind

an der Gleichung

$$\hat{R}_\theta(\Omega) = \text{conv} \left\{ (s_-, 0), (s_+, 0), \left( \frac{s_+ + s_-}{2}, \frac{s_+ - s_-}{2\delta} \right), \left( \frac{s_+ + s_-}{2}, \frac{s_- - s_+}{2\delta} \right) \right\}$$

*Beweis.* Siehe Anhang A.2. □

Ebenso kann ein konkreter Ausdruck für die Träger von  $\nu_n$  gefunden werden.

**4.9 Lemma** Es sei  $n \in \mathbb{S}^1$  beliebig.

1. Falls  $n_1 \geq \frac{|n_2|}{\delta}$ , so ist  $Q_n^d(\Omega) = [0, n_1]^d$ .
2. Falls  $n_1 \leq -\frac{|n_2|}{\delta}$ , so ist  $Q_n^d(\Omega) = [n_1, 0]^d$ .
3. Falls  $|n_1| < \frac{|n_2|}{\delta}$ , so ist  $Q_n^d(\Omega) = [\frac{n_1}{2} - \frac{|n_2|}{2\delta}, \frac{n_1}{2} + \frac{|n_2|}{2\delta}]^d$ .

*Beweis.* Siehe Anhang A.2. □

## 4.4. Exakte Rekonstruktion

In diesem Abschnitt sollen die dimensionsreduzierten Probleme mathematisch analysiert und ihr Zusammenhang zum volldimensionalen dynamischen Problem sowie zu den einzelnen statischen Problemen untersucht werden. Unser Fokus liegt dabei auf der Frage der exakten Rekonstruktion, also unter welchen Voraussetzungen die eindeutige Lösung der Minimierungsprobleme durch die „wahre“ Lösung gegeben ist. Was hierunter genau zu verstehen ist, wird im Folgenden definiert.

Es sei ein Zielmaß

$$\hat{\lambda} = \sum_{i=1}^N \alpha_i \delta_{(x_i, v_i)}$$

und Daten  $\hat{y} = \mathcal{G}\hat{\lambda}$  gegeben. Wie gehabt seien die Konfigurationen zu den Zeitschritten  $k \in \mathcal{K}$  gegeben durch

$$\hat{u}_k := (D_k^d)_\# \hat{\lambda} = \sum_{i=1}^N \alpha_i \delta_{x_i + k\tau v_i}. \quad (4.5)$$

Außerdem haben wir die gemeinsame Radon-Transformation des Zielmaßes

$$\hat{\mu}_\theta := (\hat{R}_\theta)_\# \hat{\lambda} = \sum_{i=1}^N \alpha_i \delta_{(\theta \cdot x_i, \theta \cdot v_i)}, \quad \theta \in \mathbb{S}^{d-1}. \quad (4.6)$$

Schließlich können wir auch die erwarteten Lösungen für die Variablen  $\nu_n$  in den Problemen mit relaxierter Nebenbedingung berechnen:

$$\hat{\nu}_n := (Q_n^d)_\# \hat{\lambda} = \sum_{i=1}^N \alpha_i \delta_{n_1 x_i + n_2 v_i}, \quad n \in \mathbb{S}^1. \quad (4.7)$$

Wir geben die in diesem Abschnitt relevanten Probleme noch einmal an. Dabei wollen wir uns zur Vereinfachung auf den Fall der nichtnegativen reellen Maße konzentrieren. Wir nehmen also an, dass  $\alpha_i \in [0, \infty)$  und schränken die Räume für die Variablen der Minimierungsprobleme dementsprechend ein.

Zunächst wird zu jedem Zeitschritt  $k \in \mathcal{K}$  das statische Problem

$$\min_{u \in \mathcal{M}^+([0,1]^d)} \|u\|_{\text{TV}} \quad \text{s.t.} \quad Bu = \hat{y}_k \quad (\text{ER}k)$$

betrachtet. Das volldimensionale dynamische Problem lautet

$$\min_{\lambda \in \mathcal{M}^+(\Omega)} \|\lambda\|_{\text{TV}} \quad \text{s.t.} \quad \mathcal{G}\lambda = \hat{y}. \quad (\text{ERdyn})$$

Das dimensionsreduzierte Problem bei Verzicht auf die letzte Nebenbedingung hat die Form

$$\begin{aligned} & \min_{\substack{(\mu_\theta)_{\theta \in \mathbb{S}^{d-1}} \subset \mathcal{M}^+(\mathbb{R}^2) \\ (u_k)_{k \in \mathcal{K}} \subset \mathcal{M}^+([0,1]^d)}} \sup_{\theta \in \mathbb{S}^{d-1}} \|\mu_\theta\|_{\text{TV}} \quad \text{s.t.} \\ & Bu_k = \hat{y}_k \quad \forall k \in \mathcal{K}, \\ & (R_\theta)_\# u_k = (D_k^1)_\# \mu_\theta \quad \forall k \in \mathcal{K} \forall \theta \in \mathbb{S}^{d-1}. \end{aligned} \quad (\text{ERdyn}^-)$$

Zuletzt betrachten wir das Problem mit relaxierter Nebenbedingung in der Version nach Elimination von  $u$ :

$$\begin{aligned} & \min_{\substack{(\mu_\theta)_{\theta \in \mathbb{S}^{d-1}} \subset \mathcal{M}^+(\mathbb{R}^2) \\ (\nu_n)_{n \in \mathbb{S}^1} \in \mathcal{Z}^+}} \sup_{\theta \in \mathbb{S}^{d-1}} \|\mu_\theta\|_{\text{TV}} \quad \text{s.t.} \\ & B[x \mapsto s_k x]_\# \nu_{n_k} = \hat{y}_k \quad \forall k \in \mathcal{K}, \\ & (R_\theta)_\# \nu_n = (Q_n^1)_\# \mu_\theta \quad \forall n \in \mathbb{S}^1 \forall \theta \in \mathbb{S}^{d-1}. \end{aligned} \quad (\text{ERdyn}^{\text{rel}})$$

Der Raum für die Variable  $\nu$  ist dabei entsprechend der Annahme der Nichtnegativität definiert als

$$\mathcal{Z}^+ := \prod_{n \in \mathbb{S}^1} \mathcal{M}^+(Q_n^d(\Omega)).$$

**4.10 Bemerkung** Da wir uns auf nichtnegative Maße beschränken, ist die TV-Norm einfach die Gesamtmasse, d.h. das Maß des gesamten Raumes. Für Bildmaße bleibt diese offensichtlich erhalten, also erzwingen die Nebenbedingungen der obigen Probleme, dass alle auftretenden Maße die gleiche TV-Norm haben. Insbesondere sind die Suprema in den Zielfunktionen trivial.

**4.11 Definition** (Exakte Rekonstruktion) 1. Das Problem (ER $k$ ) rekonstruiert exakt, falls eine eindeutige Lösung durch  $u = \hat{u}_k$  gegeben ist.

2. Das Problem (ERdyn) rekonstruiert exakt, falls eine eindeutige Lösung durch  $\lambda = \hat{\lambda}$  gegeben ist.

3. Das Problem  $(\text{ERdyn}^-)$  rekonstruiert exakt, falls für jede Lösung  $(\mu, u)$  gilt

$$\begin{aligned} u_k &= \hat{u}_k \quad \forall k \in \mathcal{K}, \\ \mu_\theta &= \hat{\mu}_\theta \quad \text{für } \omega_{d-1}\text{-fast alle } \theta \in \mathbb{S}^{d-1}. \end{aligned}$$

4. Das Problem  $(\text{ERdyn}^{\text{rel}})$  rekonstruiert exakt, falls für jede Lösung  $(\mu, \nu)$  gilt

$$\begin{aligned} \mu_\theta &= \hat{\mu}_\theta \quad \text{für } \omega_{d-1}\text{-fast alle } \theta \in \mathbb{S}^{d-1}, \\ \nu_n &= \hat{\nu}_n \quad \text{für } \omega_1\text{-fast alle } n \in \mathbb{S}^1. \end{aligned}$$

Intuitiv liegt das Problem mit relaxierter Nebenbedingung näher am volldimensionalen Problem, als das Problem, bei welchem auf die Nebenbedingung völlig verzichtet wurde. Wir erwarten also eine Hierarchie

$$">(\text{ERdyn}) \geq (\text{ERdyn}^{\text{rel}}) \geq (\text{ERdyn}^-)".$$

Hierauf zielen die nächsten beiden Propositionen ab.

**4.12 Proposition** Angenommen,  $(\text{ERdyn}^-)$  rekonstruiert exakt. Dann rekonstruiert auch  $(\text{ERdyn}^{\text{rel}})$  exakt.

*Beweis.* Zunächst zeigen wir, dass die Variablen  $(\hat{\mu}, \hat{\nu})$  definiert durch Gleichungen (4.6) und (4.7) zulässig für  $(\text{ERdyn}^{\text{rel}})$  sind. Die erste Nebenbedingung ist erfüllt, da für alle  $k \in \mathcal{K}$  unter Verwendung von Lemma 4.5 gilt

$$\begin{aligned} B[x \mapsto s_k x]_{\#} \hat{\nu}_{n_k} &= B[x \mapsto s_k x]_{\#} (Q_{n_k}^d)_{\#} \hat{\lambda} \\ &= B(D_k^d)_{\#} \hat{\lambda} = \hat{y}_k. \end{aligned}$$

Die Rechnung

$$\begin{aligned} (R_\theta)_{\#} \hat{\nu}_n &= (R_\theta)_{\#} (Q_n^d)_{\#} \hat{\lambda} \\ &= (Q_n^1)_{\#} (\hat{R}_\theta)_{\#} \hat{\lambda} = (Q_n^1)_{\#} \hat{\mu}_\theta \end{aligned}$$

beweist, dass auch die zweite Nebenbedingung von  $(\text{ERdyn}^{\text{rel}})$  für alle  $\theta \in \mathbb{S}^{d-1}, n \in \mathbb{S}^1$  erfüllt ist.

Sei nun  $(\mu, \nu)$  eine weitere Lösung von  $(\text{ERdyn}^{\text{rel}})$  mit

$$\sup_{\theta \in \mathbb{S}^{d-1}} \|\mu_\theta\|_{\text{TV}} = \sup_{\theta \in \mathbb{S}^{d-1}} \|\hat{\mu}_\theta\|_{\text{TV}}.$$

Definiere

$$u_k := [x \mapsto s_k x]_{\#} \nu_{n_k}, \quad k \in \mathcal{K},$$

dann ist  $(\mu, u)$  zulässig für  $(\text{ERdyn}^-)$ : Die erste Nebenbedingung ist nämlich erfüllt, da  $\nu$  die erste Nebenbedingung von  $(\text{ERdyn}^{\text{rel}})$  erfüllen muss. Die zweite Nebenbe-

dingung ist erfüllt nach der Rechnung

$$\begin{aligned}
(R_\theta)_\# u_k &= (R_\theta)_\# [x \mapsto s_k x]_\# \nu_{n_k} \\
&= [z \mapsto s_k z]_\# (R_\theta)_\# \nu_{n_k} \\
&= [z \mapsto s_k z]_\# (Q_{n_k}^1)_\# \mu_\theta \\
&= (D_k^1)_\# \mu_\theta,
\end{aligned}$$

wobei im vorletzten Schritt genutzt wurde, dass  $(\mu, \nu)$  die zweite Nebenbedingung von  $(\text{ERdyn}^{\text{rel}})$  erfüllen.

Da nach Annahme  $(\text{ERdyn}^-)$  exakt rekonstruiert, folgt

$$\begin{aligned}
u_k &= \hat{u}_k \quad \forall k \in \mathcal{K}, \\
\mu_\theta &= \hat{\mu}_\theta \quad \text{für } \omega_{d-1}\text{-fast alle } \theta \in \mathbb{S}^{d-1}.
\end{aligned}$$

Es ist noch die exakte Rekonstruktion von  $\hat{\nu}$  zu zeigen. Dies folgt jedoch aus Satz 4.3, da für alle  $n \in \mathbb{S}^1$  und fast alle  $\theta \in \mathbb{S}^{d-1}$

$$(R_\theta)_\# \nu_n = (Q_n^1)_\# \mu_\theta = (Q_n^1)_\# \hat{\mu}_\theta = (R_\theta)_\# \hat{\nu}_n$$

gilt. □

Es sei mit  $\mathcal{M}_{\text{disc}}(\mathbb{R}^{2d}) \subset \mathcal{M}(\mathbb{R}^{2d})$  die Menge der endlichen Radonmaße auf  $\mathbb{R}^{2d}$  bezeichnet, welche zusätzlich diskret sind. Hierbei verstehen wir unter einem diskreten Maß ein Maß, welches konzentriert auf einer abzählbaren Menge ist.

**4.13 Lemma** Die gemeinsame Radon-Transformation

$$\hat{\mathcal{R}} : \mathcal{M}_{\text{disc}}(\mathbb{R}^{2d}) \rightarrow \mathcal{M}(\mathbb{S}^{d-1} \times \mathbb{R}^2),$$

eingeschränkt auf den Raum der diskreten Maße, ist injektiv.

*Beweis.* Es sei  $\xi \in \mathcal{M}_{\text{disc}}(\mathbb{R}^{2d})$  mit  $\hat{\mathcal{R}}\xi = 0$ . Nach Definition von  $\hat{\mathcal{R}}$  gilt

$$(\hat{R}_\theta)_\# \xi = 0 \quad \text{für } \omega_{d-1}\text{-fast alle } \theta \in \mathbb{S}^{d-1}.$$

Da es sich um ein diskretes Maße handelt, existiert eine abzählbare Menge  $A \subset \mathbb{R}^{2d}$  mit  $|\xi|(\mathbb{R}^{2d} \setminus A) = 0$ .

Es sei  $(x, v) \in A$  beliebig. Für jedes weitere  $(x', v') \in A$ ,  $(x', v') \neq (x, v)$  ist die Menge

$$\begin{aligned}
&\{\theta \in \mathbb{S}^{d-1} \mid \hat{R}_\theta(x, v) = \hat{R}_\theta(x', v')\} \\
&= \{\theta \in \mathbb{S}^{d-1} \mid \theta \cdot (x - x') = 0\} \cap \{\theta \in \mathbb{S}^{d-1} \mid \theta \cdot (v - v') = 0\}
\end{aligned}$$

enthalten im Schnitt von  $\mathbb{S}^{d-1}$  mit einer Hyperebene, also eine  $\omega_{d-1}$ -Nullmenge. Da

$A$  abzählbar ist, gilt insgesamt für  $\omega_{d-1}$ -fast alle  $\theta \in \mathbb{S}^{d-1}$

$$(\hat{R}_\theta)_\# \xi = 0$$

und

$$\forall (x', v') \in A, (x', v') \neq (x, v) : \hat{R}_\theta(x, v) \neq \hat{R}_\theta(x', v').$$

Wähle irgendein solches  $\theta$  aus, dann folgt

$$\xi(\{(x, v)\}) = (\hat{R}_\theta)_\# \xi(\{\hat{R}_\theta(x, v)\}) = 0$$

und da  $(x, v) \in A$  beliebig war, folgt  $\xi = 0$ . □

**4.14 Proposition** Angenommen,  $(\text{ERdyn}^{\text{rel}})$  rekonstruiert exakt. Dann rekonstruiert auch  $(\text{ERdyn})$  exakt.

*Beweis.* Zunächst ist  $\hat{\lambda}$  offensichtlich zulässig für  $(\text{ERdyn})$ . Es sei also  $\lambda \in \mathcal{M}^+(\Omega)$  eine beliebige Lösung von  $(\text{ERdyn})$  mit

$$\|\lambda\|_{\text{TV}} = \|\hat{\lambda}\|_{\text{TV}}$$

und wir müssen zeigen, dass  $\lambda = \hat{\lambda}$ .

Definiere für  $\theta \in \mathbb{S}^{d-1}, n \in \mathbb{S}^1$

$$\mu_\theta := (\hat{R}_\theta)_\# \lambda, \quad \nu_n := (Q_n^d)_\# \hat{\lambda}.$$

Dann ist  $(\mu, \nu)$  zulässig für  $(\text{ERdyn}^{\text{rel}})$  nach analogen Rechnungen wie in Lemma 4.6 bzw. Abschnitt 4.2.1. Es gilt nach Bemerkung 4.10 und der Optimalität von  $\lambda$

$$\sup_{\theta \in \mathbb{S}^{d-1}} \|\mu_\theta\|_{\text{TV}} = \|\lambda\|_{\text{TV}} = \|\hat{\lambda}\|_{\text{TV}} = \sup_{\theta \in \mathbb{S}^{d-1}} \|\hat{\mu}_\theta\|_{\text{TV}},$$

sodass  $(\mu, \nu)$  auch optimal für  $(\text{ERdyn}^{\text{rel}})$  ist. Da wir exakte Rekonstruktion für dieses Problem angenommen haben, folgt

$$\mu_\theta = \hat{\mu}_\theta \quad \text{für } \omega_{d-1}\text{-fast alle } \theta \in \mathbb{S}^{d-1}.$$

Übersetzt auf die Variable  $\lambda$  bedeutet dies

$$(\hat{R}_\theta)_\# \lambda = (\hat{R}_\theta)_\# \hat{\lambda} = \sum_{i=1}^N \alpha_i \delta_{(\theta \cdot x_i, \theta \cdot v_i)} \quad (4.8)$$

für fast alle  $\theta \in \mathbb{S}^{d-1}$ . Hieraus wollen wir folgern, dass  $\lambda$  ein diskretes Maß sein muss:

Es sei wieder  $T := \{(x_i, v_i) \mid i = 1, \dots, N\}$  der Träger von  $\hat{\lambda}$ . Aus Gleichung (4.8) können wir schließen, dass

$$(\hat{R}_\theta)_\# \lambda(\mathbb{R}^2 \setminus \hat{R}_\theta(T)) = 0 \quad \text{für fast alle } \theta \in \mathbb{S}^{d-1},$$

d.h.  $\lambda$  muss konzentriert auf den Mengen  $\hat{R}_\theta^{-1}(\hat{R}_\theta(T))$  sein. Wir wählen nun eine Basis  $\{\theta_1, \dots, \theta_d\} \subset \mathbb{S}^{d-1}$  von  $\mathbb{R}^d$ , sodass Gleichung (4.8) für alle  $\theta_j, 1 \leq j \leq d$  gilt. Nun ist  $\lambda$  auch konzentriert auf

$$\bigcap_{j=1}^d \hat{R}_{\theta_j}^{-1}(\hat{R}_{\theta_j}(T)).$$

Da aber die Mengen  $\hat{R}_{\theta_j}(T)$  endlich sind, ist auch obiger Schnitt endlich und somit muss  $\lambda$  diskret sein. Mit Lemma 4.13 folgt schließlich  $\lambda = \hat{\lambda}$  wie gewünscht.  $\square$

Der Rest dieses Kapitels soll dazu dienen, ein analoges Resultat zu Satz 3.8 für das dimensionsreduzierte Problem (ERdyn<sup>-</sup>) zu beweisen. Dazu wollen wir annehmen, dass für eine Teilmenge  $\mathcal{K}' \subset \mathcal{K}$  von Zeitschritten das statische Problem (ERk) exakt rekonstruiert.

Es bezeichne  $T = \{(x_i, v_i) \mid i = 1, \dots, N\}$  wieder den Träger des Zielmaßes  $\hat{\lambda}$ . Wir nutzen die Begriffe der Koinzidenz und der Geisterteilchen aus Abschnitt 3.3.

**4.15 Satz** Es existiere eine Teilmenge  $\mathcal{K}' \subset \mathcal{K}$  von Zeitschritten ( $|\mathcal{K}'| \geq 3$ ), sodass folgende Annahmen gelten:

- 1) Die Teilchenkonfiguration  $T$  habe zu den Zeiten in  $\mathcal{K}'$  keine Koinzidenz,
- 2) es existiere kein Geisterteilchen für die Konfiguration  $T$  (bezüglich  $\mathcal{K}'$ ),
- 3) für jedes  $k \in \mathcal{K}'$  sei exakte Rekonstruktion für das statische Problem ERk gegeben, d.h.  $\hat{u}_k = (D_k^d)_\# \hat{\lambda}$  sei die eindeutige Lösung von ERk.

Dann rekonstruiert (ERdyn<sup>-</sup>) exakt.

**4.16 Bemerkung** In Satz 3.8 wurde die Existenz eines statischen dualen Zertifikates für jeden Zeitschritt  $k \in \mathcal{K}'$  (und zu jedem  $\eta \in \text{sgn } \mathbb{K}^N$ ) gefordert. Da hieraus folgt, dass die statischen Probleme exakt rekonstruieren, ist die Annahme des obigen Satzes nicht stärker.

Außerdem hatten wir schon bemerkt (vgl. Bemerkung 3.3), dass für die Existenz geeigneter Zertifikate für Satz 3.8 keine Koinzidenzen vorliegen dürfen, sodass auch diese Annahme nicht stärker ist.

Allerdings beschränken wir uns in diesem Abschnitt auf nichtnegative Maße, während Satz 3.8 sogar für komplexwertige Maße gilt.

Um diesen Satz zu beweisen, werden wir einige Hilfsbehauptungen benötigen. Konkret wollen wir zu den Richtungen  $\theta \in \mathbb{S}^{d-1}$  die projizierten Teilchenkonfigurationen

$$\hat{R}_\theta(T) = \{(\theta \cdot x_i, \theta \cdot v_i) \mid i = 1, \dots, N\} \subset \mathbb{R}^2$$

untersuchen, welche also aus eindimensionalen Teilchen mit Positionen  $\theta \cdot x_i$  und Geschwindigkeiten  $\theta \cdot v_i$  bestehen. In Abschnitt 3.3 hatten wir die Mengen

$$L_{i,k} = \{(x, v) \in \mathbb{R}^{2d} \mid x + k\tau v = x_i + k\tau v_i\}$$

für  $i = 1, \dots, N, k \in \mathcal{K}$  eingeführt. Auch für die projizierten Konfigurationen betrachten wir die Mengen

$$L_{i,k}(\theta) = \{(y, w) \in \mathbb{R}^2 \mid y + k\tau w = \theta \cdot x_i + k\tau\theta \cdot v_i\}$$

und untersuchen den Zusammenhang im nächsten Lemma.

**4.17 Lemma** Es seien  $i, j \in \{1, \dots, N\}$  und  $\theta \in \mathbb{S}^{d-1}$  gegeben.

1. Es gilt für alle  $k \in \mathcal{K}$ :

$$L_{i,k}(\theta) = \hat{R}_\theta(L_{i,k}).$$

2. Für zwei Zeitschritte  $k_1, k_2 \in \mathcal{K}, k_1 \neq k_2$  sind die Schnitte

$$L_{i,k_1} \cap L_{j,k_2} \quad \text{sowie} \quad L_{i,k_1}(\theta) \cap L_{j,k_2}(\theta)$$

einelementig und es gilt

$$L_{i,k_1}(\theta) \cap L_{j,k_2}(\theta) = \hat{R}_\theta(L_{i,k_1} \cap L_{j,k_2}).$$

*Beweis.* Zu Punkt 1: Ein beliebiges  $(x, v) \in L_{i,k}$  erfüllt die Gleichung

$$x + k\tau v = x_i + k\tau v_i.$$

Bildung des Skalarproduktes mit  $\theta$  ergibt

$$\theta \cdot x + k\tau\theta \cdot v = \theta \cdot x_i + k\tau\theta \cdot v_i,$$

also ist  $\hat{R}_\theta(x, v) = (\theta \cdot x, \theta \cdot v) \in L_{i,k}(\theta)$ .

Sei nun  $(y, w) \in L_{i,k}(\theta)$ . Definiere  $x := x_i + k\tau(v_i - w\theta)$  und  $v := w\theta$ , dann gilt  $\theta \cdot v = w$  und

$$\theta \cdot x = \theta \cdot (x_i + k\tau v_i) - k\tau w = y,$$

d.h.  $\hat{R}_\theta(x, v) = (y, w)$ . Des Weiteren ist  $(x, v) \in L_{i,k}$ , denn

$$x + k\tau v = x_i + k\tau(v_i - w\theta) + k\tau w\theta = x_i + k\tau v_i.$$

Somit sind beide Inklusionen gezeigt.

Zu Punkt 2: Definiere

$$p_1 := x_i + k_1\tau v_i \quad \text{und} \quad p_2 := x_j + k_2\tau v_j,$$

dann ist  $(x, v) \in L_{i,k_1} \cap L_{j,k_2}$  genau dann, wenn das lineare Gleichungssystem

$$\begin{aligned} x + k_1\tau v &= p_1 \\ x + k_2\tau v &= p_2 \end{aligned}$$

erfüllt ist. Dieses wird eindeutig gelöst durch

$$x = \frac{1}{k_2 - k_1}(k_2 p_1 - k_1 p_2) \quad \text{und} \quad v = \frac{1}{(k_2 - k_1)\tau}(p_2 - p_1).$$

Die Gleichungen

$$\begin{aligned} y + k_1 \tau w &= \theta \cdot p_1 \\ y + k_2 \tau w &= \theta \cdot p_2 \end{aligned}$$

werden eindeutig gelöst durch

$$y = \frac{1}{k_2 - k_1}(k_2 \theta \cdot p_1 - k_1 \theta \cdot p_2) \quad \text{und} \quad w = \frac{1}{(k_2 - k_1)\tau}(\theta \cdot p_2 - \theta \cdot p_1).$$

Aufgrund der Eindeutigkeit der Lösungen sind die Schnitte in der Behauptung einelementig und wir lesen ab, dass  $(y, w) = (\theta \cdot x, \theta \cdot v) = \hat{R}_\theta(x, v)$  gilt.  $\square$

Auch in der projizierten Konfiguration  $\hat{R}_\theta(T)$  können Koinzidenzen auftreten. Das nächste Lemma besagt, dass durch die gemeinsame Projektion fast sicher keine neuen Koinzidenzen entstehen.

**4.18 Lemma** Angenommen, die Teilchenkonfiguration  $T$  ist zu den Zeitschritten in  $\mathcal{K}' \subset \mathcal{K}$  frei von Koinzidenzen. Dann ist

$$M_1 := \{\theta \in \mathbb{S}^{d-1} \mid \hat{R}_\theta(T) \text{ hat eine Koinzidenz zu einer Zeit } k \in \mathcal{K}'\}$$

eine  $\omega_{d-1}$ -Nullmenge.

*Beweis.* Für ein  $\theta \in \mathbb{S}^{d-1}$  und Indizes  $i, j \in \{1, \dots, N\}, i \neq j$  liegen die projizierten Teilchen  $(\theta \cdot x_i, \theta \cdot v_i)$  und  $(\theta \cdot x_j, \theta \cdot v_j)$  genau dann im  $k$ -ten Zeitschritt auf derselben Position, wenn gilt

$$\theta \cdot x_i + k\tau\theta \cdot v_i = \theta \cdot x_j + k\tau\theta \cdot v_j.$$

Umstellen führt zu der Gleichung

$$\theta \cdot (x_i - x_j + k\tau(v_i - v_j)) = 0.$$

in der Variable  $\theta$ . Da nach Voraussetzung  $x_i - x_j + k\tau(v_i - v_j) \neq 0$  für alle  $k \in \mathcal{K}'$  gilt, ist die Lösungsmenge der obigen Gleichung eine Hyperebene in  $\mathbb{R}^d$ . Der Schnitt einer Hyperebene mit  $\mathbb{S}^{d-1}$  ist jedoch eine  $\omega_{d-1}$ -Nullmenge. Insgesamt ist

$$M_1 = \bigcup_{k \in \mathcal{K}'} \bigcup_{i \neq j} \{\theta \in \mathbb{S}^{d-1} \mid \theta \cdot (x_i - x_j + k\tau(v_i - v_j)) = 0\}$$

eine endliche Vereinigung von  $\omega_{d-1}$ -Nullmengen.  $\square$

Die Menge der Geisterteilchen der volldimensionalen Konfiguration  $T$  bezüglich  $\mathcal{K}'$  hatten wir in Abschnitt 3.3 mit  $G(T, \mathcal{K}')$  bezeichnet und festgestellt, dass sich

diese ausdrücken lässt als

$$G(T, \mathcal{K}') = \bigcup_{(i_k)_k} \bigcap_{k \in \mathcal{K}'} L_{i_k, k},$$

wobei die Vereinigung über alle Folgen  $(i_k)_{k \in \mathcal{K}'} \subset \{1, \dots, N\}$  von paarweise verschiedenen Indizes läuft. Analog betrachten wir nun die Geisterteilchen  $G(\hat{R}_\theta(T), \mathcal{K}')$  der projizierten Konfiguration. Auch diese lässt sich berechnen aus

$$G(\hat{R}_\theta(T), \mathcal{K}') = \bigcup_{(i_k)_k} \bigcap_{k \in \mathcal{K}'} L_{i_k, k}(\theta). \quad (4.9)$$

**4.19 Lemma** Es gelten folgende Zusammenhänge:

1. Für alle  $\theta \in \mathbb{S}^{d-1}$  gilt  $G(\hat{R}_\theta(T), \mathcal{K}') \supset \hat{R}_\theta(G(T, \mathcal{K}'))$
2. Für  $\omega_{d-1}$ -fast alle  $\theta \in \mathbb{S}^{d-1}$  gilt sogar  $G(\hat{R}_\theta(T), \mathcal{K}') = \hat{R}_\theta(G(T, \mathcal{K}'))$ .

Aussage 1 des Lemmas besagt, dass jedes Geisterteilchen der volldimensionalen Konfiguration ein Geisterteilchen der projizierten Konfiguration induziert. Aus Aussage 2 folgt sofort das Korollar:

**4.20 Korollar** Angenommen, die Teilchenkonfiguration  $T$  lässt kein Geisterteilchen bezüglich  $\mathcal{K}'$  zu.

Dann lässt für  $\omega_{d-1}$ -fast alle  $\theta \in \mathbb{S}^{d-1}$  auch die projizierte Konfiguration  $\hat{R}_\theta(T)$  kein Geisterteilchen bezüglich  $\mathcal{K}'$  zu.

*Beweis von Lemma 4.19.* Es sei  $\mathcal{K}' = \{k_1, \dots, k_{m'}\}$ .

Zu Punkt 1: Es sei  $\theta \in \mathbb{S}^{d-1}$  beliebig und  $(x, v) \in G(T, \mathcal{K}')$  ein Geisterteilchen. Die Definition eines Geisterteilchens besagt, dass eine Folge  $(i_p)_{p \in \{1, \dots, m'\}}$  von paarweise verschiedene Indizes existiert, sodass

$$(x, v) \in \bigcap_{p=1}^{m'} L_{i_p, k_p}.$$

Es folgt, dass  $\hat{R}_\theta(x, v) \in \hat{R}_\theta(\bigcap_{p=1}^{m'} L_{i_p, k_p})$ . Nun ist aber das Bild des Schnittes eine Teilmenge des Schnittes über die Bilder und mit Punkt 1 von Lemma 4.17 folgt

$$\hat{R}_\theta \left( \bigcap_{p=1}^{m'} L_{i_p, k_p} \right) \subset \bigcap_{p=1}^{m'} \hat{R}_\theta(L_{i_p, k_p}) = \bigcap_{p=1}^{m'} L_{i_p, k_p}(\theta)$$

Also liegt  $\hat{R}_\theta(x, v)$  auch in Schnitt auf der rechten Seite und folglich ist  $\hat{R}_\theta(x, v) \in G(\hat{R}_\theta(T), \mathcal{K}')$ .

Zu Punkt 2: Wir definieren die Menge an Projektionsrichtungen, für die eine

strikte Inklusion der Geisterteilchen gilt:

$$M := \{\theta \in \mathbb{S}^{d-1} \mid G(\hat{R}_\theta(T), \mathcal{K}') \setminus \hat{R}_\theta(G(T, \mathcal{K}')) \neq \emptyset\}$$

Zu zeigen ist, dass es sich hierbei um eine  $\omega_{d-1}$ -Nullmenge handelt. Nach Gleichung (4.9) lässt sich die Menge ausdrücken als endliche Vereinigung

$$\begin{aligned} M &= \left\{ \theta \in \mathbb{S}^{d-1} \mid \bigcup_{(i_p)_p} \bigcap_{p=1}^{m'} L_{i_p, k_p}(\theta) \setminus \hat{R}_\theta(G(T, \mathcal{K}')) \neq \emptyset \right\} \\ &= \bigcup_{(i_p)_p} \left\{ \theta \in \mathbb{S}^{d-1} \mid \bigcap_{p=1}^{m'} L_{i_p, k_p}(\theta) \setminus \hat{R}_\theta(G(T, \mathcal{K}')) \neq \emptyset \right\} \end{aligned} \quad (4.10)$$

und es genügt zu zeigen, dass die Mengen innerhalb der Vereinigung Nullmengen sind.

Es sei also  $(i_p)_{p \in \{1, \dots, m'\}} \subset \{1, \dots, N\}$  eine Folge von paarweise verschiedenen Indizes und  $\theta \in \mathbb{S}^{d-1}$  beliebig.

Laut Lemma 4.17 sind die Schnitte  $L_{i_p, k_p} \cap L_{i_{p+1}, k_{p+1}}$  einelementig. Schreibe also

$$\{(a_p, b_p)\} = L_{i_p, k_p} \cap L_{i_{p+1}, k_{p+1}}, \quad p = 1, \dots, m' - 1$$

und rechne mittels Lemma 4.17

$$\begin{aligned} \bigcap_{p=1}^{m'} L_{i_p, k_p}(\theta) &= \bigcap_{p=1}^{m'-1} L_{i_p, k_p}(\theta) \cap L_{i_{p+1}, k_{p+1}}(\theta) \\ &= \bigcap_{p=1}^{m'-1} \hat{R}_\theta(L_{i_p, k_p} \cap L_{i_{p+1}, k_{p+1}}) = \bigcap_{p=1}^{m'-1} \{\hat{R}_\theta(a_p, b_p)\}. \end{aligned} \quad (4.11)$$

Nun unterscheiden wir zwischen zwei Fällen.

**Fall 1:** Alle Punkte  $(a_p, b_p)$  stimmen überein, d.h.

$$\bigcap_{p=1}^{m'} L_{i_p, k_p} = \{(a_1, b_1)\}.$$

In diesem Fall ist  $(a_1, b_1) \in G(T, \mathcal{K}')$  ein Geisterteilchen der volldimensionalen Konfiguration. Aus Gleichung (4.11) folgt

$$\bigcap_{p=1}^{m'} L_{i_p, k_p}(\theta) \setminus \hat{R}_\theta(G(T, \mathcal{K}')) = \emptyset.$$

**Fall 2:** Es existiert ein  $q \in \{1, \dots, m' - 1\}$ , sodass  $(a_q, b_q) \neq (a_{q+1}, b_{q+1})$ . Es sei o.B.d.A  $a_q \neq a_{q+1}$ . Angenommen, der letzte Schnitt in Gleichung (4.11) ist nicht

leer, dann muss insbesondere

$$(\theta \cdot a_q, \theta \cdot b_q) = \hat{R}_\theta(a_q, b_q) = \hat{R}_\theta(a_{q+1}, b_{q+1}) = (\theta \cdot a_{q+1}, \theta \cdot b_{q+1})$$

gelten, d.h.  $\theta$  ist in der Hyperebene

$$\{\theta \in \mathbb{R}^d \mid \theta \cdot (a_q - a_{q+1}) = 0\}$$

enthalten.

Nach der obigen Fallunterscheidung sind die Mengen in der Vereinigung aus Gleichung (4.10) jeweils entweder leer oder in einer Hyperebene enthalten, d.h.  $M$  ist enthalten in einer endlichen Vereinigung von  $\omega_{d-1}$ -Nullmengen und somit selbst eine Nullmenge.  $\square$

Nun sind wir bereit, Satz 4.15 zu beweisen.

*Beweis von Satz 4.15.* Es sei  $(\mu, u)$  eine Lösung von  $\text{ERdyn}^-$ .

**Schritt 1:** Exakte Rekonstruktion der  $\hat{u}_k$  für  $k \in \mathcal{K}'$ .

Es sei  $k \in \mathcal{K}'$  beliebig. Die erste Nebenbedingung von  $(\text{ERdyn}^-)$  stimmt mit der Nebenbedingung von  $(\text{ER}k)$  überein, sodass  $u_k$  auch zulässig für  $(\text{ER}k)$  ist. Nach Annahme ist  $(\mu, u)$  eine Lösung von  $(\text{ERdyn}^-)$ , sodass die Ungleichung

$$\sup_{\theta \in \mathbb{S}^{d-1}} \|\mu_\theta\|_{\text{TV}} \leq \sup_{\theta \in \mathbb{S}^{d-1}} \|\hat{\mu}_\theta\|_{\text{TV}}$$

gilt. Da wir nur nichtnegative Maße betrachten, überträgt sich diese Ungleichung sofort auch auf die anderen Variablen (vgl. Bemerkung 4.10):

$$\|u_k\|_{\text{TV}} = \sup_{\theta \in \mathbb{S}^{d-1}} \|\mu_\theta\|_{\text{TV}} \leq \sup_{\theta \in \mathbb{S}^{d-1}} \|\hat{\mu}_\theta\|_{\text{TV}} = \|\hat{u}_k\|_{\text{TV}}$$

Nach Annahme 3) von Satz 4.15 rekonstruiert  $(\text{ER}k)$  für  $k \in \mathcal{K}'$  exakt und es folgt  $u_k = \hat{u}_k$ .

**Schritt 2:** Exakte Rekonstruktion von  $\hat{\mu}_\theta$  für fast alle  $\theta$ .

Nach Annahmen 1) und 2) von Satz 4.15 zusammen mit Lemma 4.18 und Korollar 4.20 hat die projizierte Teilchenkonfiguration  $\hat{R}_\theta(T)$  für  $\omega_{d-1}$ -fast alle  $\theta \in \mathbb{S}^{d-1}$  weder Koinzidenzen noch Geisterteilchen bezüglich der Zeiten in  $\mathcal{K}'$ . Fixiere solch ein  $\theta$ .

Es sei  $h := \mu_\theta - \hat{\mu}_\theta \in \mathcal{M}(\mathbb{R}^2, \mathbb{R})$  das signierte Maß, welches sich aus der Differenz von Rekonstruktion und Zielmaß ergibt. Die Lebesguezerlegung von  $h$  bzgl.  $|\hat{\mu}_\theta|$  erlaubt eine Darstellung

$$h = h_a + h_s, \quad h_a = \sum_{i=1}^N \beta_i \delta_{(\theta \cdot x_i, \theta \cdot v_i)}, \quad h_s \perp \hat{\mu}_\theta$$

mit Gewichten  $\beta_i \in \mathbb{R}$ . Definiere für  $k \in \mathcal{K}'$  Funktionen  $\tilde{q}_k : \mathbb{R} \rightarrow \mathbb{R}$ ,

$$\tilde{q}_k(s) = \begin{cases} \operatorname{sgn} \beta_i & \text{falls } s = \theta \cdot (x_i + k\tau v_i), \\ 0 & \text{sonst.} \end{cases}$$

Diese sind wohldefiniert, da die projizierte Konfiguration nach Wahl von  $\theta$  keine Koinzidenzen zu den Zeitschritten in  $\mathcal{K}'$  hat. Definiere weiter  $q : \mathbb{R}^2 \rightarrow \mathbb{R}$ ,

$$q(y, w) := \frac{1}{|\mathcal{K}'|} \sum_{k \in \mathcal{K}'} \tilde{q}_k(y + k\tau w),$$

dann gilt offensichtlich  $q(\theta \cdot x_i, \theta \cdot v_i) = \operatorname{sgn} \beta_i$  für alle  $i \in \{1, \dots, N\}$ .

Die Funktion  $q$  ist definiert als der Mittelwert von Summanden mit Betrag kleiner 1. Ist also  $(y, w) \in \mathbb{R}^2$  mit  $|q(y, w)| = 1$ , so muss  $|\tilde{q}_k(y + k\tau w)| = 1$  für alle  $k \in \mathcal{K}'$  gelten. Hieraus folgt

$$\forall k \in \mathcal{K}' : \exists i_k \in \{1, \dots, N\} : y + k\tau w = \theta \cdot (x_{i_k} + k\tau v_{i_k}) \quad (4.12)$$

und es lassen sich wieder zwei Fälle unterscheiden:

**Fall 1:** Ein Index  $i_k$  tritt mehrfach auf. Dann wird der Punkt  $(y, w)$  durch die Gleichung (4.12) zu zwei Zeitschritten eindeutig festgelegt und es muss  $(y, w) = (\theta \cdot x_{i_k}, \theta \cdot v_{i_k})$  gelten.

**Fall 2:** Die Indizes  $(i_k)_{k \in \mathcal{K}'}$  sind paarweise verschieden. Dann wäre  $(y, w)$  ein Geisterteilchen der projizierten Konfiguration, was nach Wahl von  $\theta$  nicht möglich ist.

Wir haben also gezeigt, dass

$$|q(y, w)| < 1 \quad \forall (y, w) \in \mathbb{R}^2 \setminus \hat{R}_\theta(T). \quad (4.13)$$

Nach Schritt 1 des Beweises sind für alle  $k \in \mathcal{K}'$  die Variablen  $u_k = \hat{u}_k$  exakt rekonstruiert. Die zweite Nebenbedingung von  $(\text{ERdyn}^-)$  ergibt

$$(D_k^1)_\# \mu_\theta = (R_\theta)_\# u_k = (R_\theta)_\# \hat{u}_k = (D_k^1)_\# \hat{\mu}_\theta,$$

folglich gilt  $0 = (D_k^1)_\# h$  für alle  $k \in \mathcal{K}'$ . Wir können also rechnen

$$\begin{aligned} 0 &= \frac{1}{|\mathcal{K}'|} \sum_{k \in \mathcal{K}'} \int_{\mathbb{R}} \tilde{q}_k d(D_k^1)_\# h = \int_{\mathbb{R}^2} q dh \\ &= \int_{\mathbb{R}^2} q dh_a + \int_{\mathbb{R}^2} q dh_s \\ &= \sum_{i=1}^N |\beta_i| + \int_{\mathbb{R}^2} q dh_s. \end{aligned}$$

Wir bekommen wegen  $\|q\|_\infty \leq 1$  die Ungleichung

$$\|h_a\|_{\text{TV}} = \left| \int_{\mathbb{R}^2} q \, dh_s \right| \leq \|h_s\|_{\text{TV}}. \quad (4.14)$$

Wegen der Optimalität von  $\mu$  und Bemerkung 4.10 gilt

$$\|\hat{\mu}_\theta\|_{\text{TV}} \geq \|\mu_\theta\|_{\text{TV}} = \|\hat{\mu}_\theta + h\|_{\text{TV}} \quad (4.15)$$

und nach Lemma 1.11 folgt

$$\|\hat{\mu}_\theta + h\|_{\text{TV}} = \|\hat{\mu}_\theta + h_a\|_{\text{TV}} + \|h_s\|_{\text{TV}} \geq \|\hat{\mu}_\theta\|_{\text{TV}} - \|h_a\|_{\text{TV}} + \|h_s\|_{\text{TV}}.$$

Schließlich gilt nach Gleichung (4.14)

$$\|\hat{\mu}_\theta\|_{\text{TV}} - \|h_a\|_{\text{TV}} + \|h_s\|_{\text{TV}} \geq \|\hat{\mu}_\theta\|_{\text{TV}}$$

und mit Blick auf den Start der Ungleichungskette in (4.15) muss überall Gleichheit gelten. Insbesondere gilt

$$\|h_a\|_{\text{TV}} = \left| \int_{\mathbb{R}^2} q \, dh_s \right| = \|h_s\|_{\text{TV}}. \quad (4.16)$$

Nun ist jedoch  $h_s \perp \hat{\mu}_\theta$  und  $q$  erfüllt die strikte Ungleichung aus Gleichung (4.13), sodass nach Korollar 1.10 schon  $h_s = 0$  gelten muss. Schließlich folgt aus Gleichung (4.16) auch  $h_a = 0$  und insgesamt  $\mu_\theta = \hat{\mu}_\theta$ .

**Schritt 3:** Exakte Rekonstruktion von  $\hat{u}_k$  für  $k \in \mathcal{K} \setminus \mathcal{K}'$ .

Tatsächlich besagt die zweite Nebenbedingung von  $(\text{ERdyn}^-)$  kombiniert mit dem soeben gezeigten Rekonstruktionsresultat für  $\mu$ , dass für jedes  $k \in \mathcal{K}$  und  $\omega_{d-1}$ -fast alle  $\theta \in \mathbb{S}^{d-1}$  gilt

$$(R_\theta)_\# u_k = (D_k^1)_\# \mu_\theta = (D_k^1)_\# \hat{\mu}_\theta = (R_\theta)_\# \hat{u}_k.$$

Aufgrund der Injektivität der Radon-Transformation (Satz 4.3) folgt  $u_k = \hat{u}_k$ .  $\square$

## 5. Numerik

Der Raum der Radonmaße ist unendlichdimensional, sodass die bisher diskutierten Optimierungsprobleme nicht direkt numerisch lösbar sind. Im diesem Kapitel wollen wir Verfahren vorstellen, um die Optimierungsprobleme in eine Form zu bringen, welche tatsächlich implementiert werden kann. Dazu stellen wir zunächst zwei Verfahren zur Lösung des ursprünglichen Problems

$$\min_{u \in \mathcal{M}(\Omega, \mathbb{R})} \|u\|_{\text{TV}} \quad \text{s.t.} \quad \mathcal{G}u = \hat{y} \quad (\text{ER})$$

vor: Die Gitter-Diskretisierung sowie den ADCG-Algorithmus. Anschließend entwickeln wir eine geeignete Diskretisierung für die dimensionsreduzierten Probleme aus Kapitel 4, welche auf der Gitter-Diskretisierung basiert. Am Ende des Kapitels werden die vorgestellten Methoden anhand eines Datensatzes von zufällig generierten Teilchenkonfigurationen verglichen.

In diesem Kapitel gehen wir davon aus, dass der Messoperator  $\mathcal{G}$  (bzw.  $B$  im dynamischen Fall) nach  $\mathbb{C}^m$  abbildet. Außerdem betrachten wir nur reellwertige Maße, d.h. den Fall  $\mathbb{K} = \mathbb{R}$ .

### 5.1. Gitter-Diskretisierung

Das unendlichdimensionale Problem (ER) kann nicht direkt gelöst werden. Stattdessen kann eine endliche Teilmenge  $\Omega_{\#} = \{g_1, \dots, g_L\} \subset \Omega$  gewählt werden. Stellvertretend für das kontinuierliche Optimierungsproblem (ER) betrachtet man nun das diskrete Problem

$$\min_{u \in \mathcal{M}(\Omega_{\#}, \mathbb{R})} \|u\|_{\text{TV}} \quad \text{s.t.} \quad \mathcal{G}u = \hat{y}. \quad (\text{ER}_{\#})$$

Da  $\Omega_{\#}$  endlich ist, sind die Maße  $u \in \mathcal{M}(\Omega_{\#}, \mathbb{R})$  genau die diskreten Maße, welche sich als Linearkombination von Dirac-Maßen

$$u = \sum_{j=1}^L \xi_j \delta_{g_j}$$

mit einem Gewichtsvektor  $\xi \in \mathbb{R}^L$  schreiben lassen. Die TV-Norm von  $u$  entspricht einfach der 1-Norm des Gewichtsvektors und die Nebenbedingung lässt sich auf-

grund der Linearität von  $\mathcal{G}$  umschreiben zu einer Matrix-Vektor-Multiplikation:

$$\mathcal{G}u = \sum_{j=1}^L \xi_j \mathcal{G}\delta_{g_j} = (\mathcal{G}\delta_{g_1}, \dots, \mathcal{G}\delta_{g_L}) \cdot \xi$$

Mit der Bezeichnung

$$M := (\mathcal{G}\delta_{g_1}, \dots, \mathcal{G}\delta_{g_L}) \in \mathbb{C}^{m \times L}$$

ist das diskrete Optimierungsproblem äquivalent zu

$$\min_{\xi \in \mathbb{R}^N} \|\xi\|_1 \quad \text{s.t.} \quad M\xi = \hat{y}.$$

Probleme dieser Art werden im Teilgebiet des *compressed sensing* untersucht und die Matrix  $M$  ist die zum Operator  $\mathcal{G}$  zugehörige *Messmatrix*. Hierbei handelt es sich um ein endlichdimensionales konvexes Optimierungsproblem mit einer linearen Gleichheitsnebenbedingung.

Betrachte den Fall  $\Omega = [0, 1]^d$  und das Gitter gegeben durch

$$\Omega_{\#} = \{j/s \mid j = 0, \dots, s-1\}^d$$

für  $s \in \mathbb{N}$ . Die Anzahl der Variablen im obigen diskreten Problem beträgt  $s^d$ . Im Falle der statischen Super-Resolution ist dies in Raumdimensionen  $d \leq 3$  häufig noch gut lösbar. Für die volldimensionale dynamische Super-Resolution kommen jedoch die Geschwindigkeits-Variablen hinzu, sodass insgesamt für Raumdimension  $d = 2$  ein Problem auf einem vierdimensionalen Gitter, für  $d = 3$  sogar auf einem sechsdimensionalen Gitter zu lösen wäre.

Eine weitere Schwierigkeit ergibt sich aus der obigen Diskretisierung: Es ist im Allgemeinen nicht zu erwarten, dass die Gleichheitsnebenbedingung noch exakt erfüllt ist. Die wahre Teilchenkonfiguration  $T$  ist nämlich im Allgemeinen nicht in den Punkten der Diskretisierungsmenge enthalten, d.h.  $T \not\subset \Omega_{\#}$ , sodass das tatsächliche Zielmaß  $\hat{u}$  in dieser Diskretisierung nicht exakt repräsentiert werden kann. Dieses Phänomen ist bekannt als *basis mismatch* [11]. Zusätzlich ist in der Realität damit zu rechnen, dass die Messdaten  $\hat{y}$  fehlerbehaftet sind. Deshalb bietet es sich an, auf eine Formulierung der Form

$$\min_{\xi \in \mathbb{R}^N} \|\xi\|_1 \quad \text{s.t.} \quad \|M\xi - \hat{y}\|_2 \leq \epsilon. \quad (5.1)$$

für geeignetes  $\epsilon > 0$  auszuweichen.

**5.1 Beispiel** Im Artikel „Faster STORM using compressed sensing“ [24] verwenden Zhu et al. die Formulierung aus Gleichung (5.1), um die Moleküle in den STORM Einzelbildern aus Beispiel 2.1 zu lokalisieren.

---

## 5.2. ADCG-Algorithmus

In diesem Abschnitt soll der Alternating Descent Conditional Gradient (ADCG) Algorithmus von Boyd et al. [6] vorgestellt werden. Für eine differenzierbare Verlustfunktion  $\ell : \mathbb{C}^m \rightarrow \mathbb{R}$  und eine Massenschranke  $\Sigma > 0$  soll das allgemeine Problem

$$\min_{u \in \mathcal{M}^+(\Omega)} \ell(\mathcal{G}u - \hat{y}) \quad \text{s.t.} \quad u(\Omega) \leq \Sigma \quad (5.2)$$

gelöst werden. Der Operator  $\mathcal{G} : \mathcal{M}(\Omega, \mathbb{R}) \rightarrow \mathbb{C}^m$  sowie der Vektor  $\hat{y} \in \mathbb{C}^m$  seien wie in Kapitel 2.

Da es sich hierbei um eine Erweiterung des klassischen Frank-Wolfe-Algorithmus (bzw. Conditional Gradient Method, CGM) [16] zur Lösung von differenzierbaren konvexen Optimierungsproblemen handelt, wird dieser zunächst eingeführt.

### 5.2.1. Frank-Wolfe-Algorithmus

Der klassische Frank-Wolfe-Algorithmus löst Optimierungsprobleme der Form

$$\min_{x \in C} f(x), \quad (5.3)$$

wobei  $C \subset \mathbb{R}^d$  eine kompakte, konvexe Menge und  $f : C \rightarrow \mathbb{R}$  differenzierbar und konvex ist.

Der Algorithmus geht von einem Startpunkt  $x_0 \in \mathbb{R}^d$  aus und verfährt iterativ. In Iteration  $k$  wird ein neuer Punkt  $x_k$  aus dem vorhergehenden Punkt  $x_{k-1}$  wie folgt berechnet:

Zunächst wird die Linearisierung  $\hat{f}_k : C \rightarrow \mathbb{R}$  von  $f$  am Punkt  $x_{k-1}$  betrachtet, welche gegeben ist durch

$$\hat{f}_k(x) := f(x_{k-1}) + \langle \nabla f(x_{k-1}), x - x_{k-1} \rangle. \quad (5.4)$$

Die Linearisierung  $\hat{f}_k$  wird nun über die Menge  $C$  minimiert:

$$s_k \in \operatorname{argmin}_{x \in C} \hat{f}_k(x) \quad (5.5)$$

Da die anderen Terme konstant sind, ist dies äquivalent zur Minimierung von  $\langle \nabla f(x_{k-1}), \cdot \rangle$  über  $C$ .

An dieser Stelle kann die neue Iterationsvariable  $x_k$  auf verschiedene Arten berechnet werden, was auf verschiedene Versionen des Frank-Wolfe-Algorithmus führt:

- a) **Vorgegebener Schritt:** Es wird eine Konvexkombination aus aktuellem Punkt  $x_{k-1}$  und Minimum  $s_k$  der Linearisierung genommen:

$$x_k = \frac{k}{k+2} x_{k-1} + \frac{2}{k+2} s_k$$

- b) **Linienuche:** Es wird ein Punkt  $x_k$  gewählt, welcher die Funktion  $f$  auf der Linie zwischen  $x_{k-1}$  und  $s_k$  minimiert:

$$x_k \in \operatorname{argmin}_{x \in \operatorname{conv}\{x_{k-1}, s_k\}} f(x)$$

- c) **Fully-corrective:** Der Punkt  $x_k$  wird so gewählt, dass er die Funktion  $f$  über die konvexe Hülle aller zuvor betrachteten Minimierer der Linearisierung minimiert:

$$x_k \in \operatorname{argmin}_{x \in \operatorname{conv}\{x_0, s_1, \dots, s_k\}} f(x)$$

### 5.2.2. Frank-Wolfe im Raum der Maße

Als Zwischenschritt zum ADCG-Algorithmus wird von Boyd et al. der klassische Frank-Wolfe-Algorithmus auf das Problem (5.2) im Vektorraum  $\mathcal{M}(\Omega, \mathbb{R})$  der reellwertigen Maße übertragen. Die zugehörige Iterationsvariable für Iteration  $k$  wird im Folgenden mit  $\gamma_k$  bezeichnet.

Da  $\ell$  differenzierbar und  $\mathcal{G}$  linear ist, ist die Abbildung

$$f : \mathcal{M}(\Omega, \mathbb{R}) \rightarrow \mathbb{R}, \quad \gamma \mapsto \ell(\mathcal{G}\gamma - \hat{y})$$

Gateaux-differenzierbar. Genauer gilt für die Gateaux-Ableitung von  $f$  an einem Punkt  $\gamma \in \mathcal{M}(\Omega, \mathbb{R})$  in Richtung eines Maßes  $h \in \mathcal{M}(\Omega, \mathbb{R})$

$$\begin{aligned} D_h f(\gamma) &= \lim_{\varepsilon \rightarrow 0} \frac{f(\gamma + \varepsilon h) - f(\gamma)}{\varepsilon} = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} (\ell(\mathcal{G}\gamma + \varepsilon \mathcal{G}h - \hat{y}) - \ell(\mathcal{G}\gamma - \hat{y})) \\ &= \langle \nabla_{\operatorname{Re}} \ell(\mathcal{G}\gamma - \hat{y}), \operatorname{Re} \mathcal{G}h \rangle + \langle \nabla_{\operatorname{Im}} \ell(\mathcal{G}\gamma - \hat{y}), \operatorname{Im} \mathcal{G}h \rangle, \end{aligned}$$

wobei  $\nabla_{\operatorname{Re}}$  bzw.  $\nabla_{\operatorname{Im}}$  für die Gradienten bezüglich der reellen bzw. imaginären Variablen stehen. Schreibe  $\tilde{\nabla} = \nabla_{\operatorname{Re}} + i \nabla_{\operatorname{Im}}$  und rechne weiter

$$\begin{aligned} D_h f(\gamma) &= \operatorname{Re} \langle \tilde{\nabla} \ell(\mathcal{G}\gamma - \hat{y}), \mathcal{G}h \rangle = \operatorname{Re} \int_{\Omega} \mathcal{G}^* (\tilde{\nabla} \ell(\mathcal{G}\gamma - \hat{y})) \, d\bar{h} \\ &= \int_{\Omega} \operatorname{Re} \mathcal{G}^* (\tilde{\nabla} \ell(\mathcal{G}\gamma - \hat{y})) \, dh. \end{aligned}$$

Die letzte Gleichheit gilt, da  $h$  ein reellwertiges Maß ist.

#### Lineares Programm

Die Gateaux-Ableitung  $D_h f(\gamma)$  in Richtung  $h$  tritt an die Stelle der Richtungsableitung  $\langle \nabla f(x), h \rangle$  im Frank-Wolfe-Algorithmus. Somit muss im ersten Schritt das Minimierungsproblem

$$s_k \in \operatorname{argmin}_{h \in \mathcal{M}^+(\Omega), h(\Omega) \leq \Sigma} \int_{\Omega} \operatorname{Re} \mathcal{G}^* (\tilde{\nabla} \ell(\mathcal{G}\gamma_k - \hat{y})) \, dh \quad (5.6)$$

gelöst werden. Definiere  $F : \Omega \rightarrow \mathbb{R}$ ,

$$F(z) := \operatorname{Re} \mathcal{G}^*(\tilde{\nabla} \ell(\mathcal{G}\gamma_k - \hat{y}))(z).$$

Der präduale Operator  $\mathcal{G}^*$  bildet in die stetigen Funktionen auf  $\Omega$  ab und  $\Omega$  ist kompakt, sodass  $F$  auf  $\Omega$  ein Minimum annimmt. Es sei  $z_k \in \operatorname{argmin}_\Omega F$ .

**Fall 1:**  $F(z_k) \geq 0$ . Dann ist  $s_k = 0$  optimal für (5.6), da  $F \geq 0$  und für alle zulässigen  $h \in \mathcal{M}^+(\Omega)$  gilt  $\int_\Omega F dh \geq 0$ .

**Fall 2:**  $F(z_k) < 0$ . In diesem Fall wird (5.6) minimal für  $s_k = \Sigma \delta_{z_k}$ , d.h. wenn die gesamte zulässige Masse auf dem Minimalpunkt liegt. Es gilt nämlich für zulässige  $h$

$$\int_\Omega F dh \geq \int_\Omega F(z_k) dh = F(z_k)h(\Omega) \geq F(z_k)\Sigma = \int_\Omega F ds_k.$$

### Gewichtsprblem

Nun soll die „fully-corrective“ Variante des Frank-Wolfe-Algorithmus auf das Setting der Maße übertragen werden. Gesucht ist also

$$\gamma_k \in \operatorname{argmin}_{\gamma \in \operatorname{conv}\{\gamma_0, s_1, \dots, s_k\}} \ell(\mathcal{G}\gamma - \hat{y}). \quad (5.7)$$

Da die Maße  $s_k$  aber nach obiger Rechnung nur aus Punktmassen bestehen, sind die vom Algorithmus produzierten Maße  $\gamma_k$  diskrete Maße mit endlichem Träger (solange das Startmaß  $\gamma_0$  dies auch erfüllt). Diese Eigenschaft macht den Frank-Wolfe-Algorithmus im Raum der Maße attraktiv: Aus dem ursprünglichen, unendlichdimensionalen Problem (5.2) wird somit ein Verfahren, welches tatsächlich im Computer implementiert werden kann.

Es sei

$$S_k = \operatorname{supp}(\gamma_0) \cup \operatorname{supp}(s_1) \cup \dots \cup \operatorname{supp}(s_k),$$

dann ist  $S_k \subset \Omega$  endlich und jedes Maß  $\gamma \in \operatorname{conv}\{\gamma_0, s_1, \dots, s_k\}$  lässt sich schreiben als

$$\gamma = \sum_{z \in S_k} \gamma(\{z\}) \delta_z.$$

Mit der Linearität des Messoperators  $\mathcal{G}$  ist der Minimierungsschritt (5.7) der „fully-corrective“ Methode äquivalent zu

$$\gamma_k \in \operatorname{argmin}_{\substack{\gamma \geq 0, \gamma(S_k) \leq \Sigma \\ \gamma(S_k^c) = 0}} \ell\left(\sum_{z \in S_k} \gamma(\{z\}) \mathcal{G}\delta_z - \hat{y}\right)$$

und hierbei handelt es sich um ein Optimierungsproblem bezüglich der Gewichte  $\gamma(\{z\}), z \in S_k$ .

Das volle Verfahren ist in Algorithmus 1 aufgelistet. Zwischen jeder Iteration müssen nur der Träger  $S_k$  und die Gewichte  $\gamma_k(\{z\}), z \in S_k$  gespeichert werden. Bei der Optimierung der Gewichte kann es passieren, dass einige Gewichte auf 0

gesetzt werden. Deshalb ist es sinnvoll, die zu diesen Gewichten gehörenden Punkte im letzten Schritt jeder Iteration aus dem Träger zu entfernen.

---

**Algorithmus 1** Frank-Wolfe für Maße
 

---

```

1: for  $k = 1, \dots, k_{\max}$  do
2:    $g_k \leftarrow \tilde{\nabla} \ell(\mathcal{G}\gamma_{k-1} - \hat{y})$  ▷ Gradient der Verlustfunktion
3:   Wähle  $z_k \in \operatorname{argmin}_{\Omega} \operatorname{Re} \mathcal{G}^*(g_k)$  ▷ Neuer Quellpunkt
4:    $S_k \leftarrow S_{k-1} \cup \{z_k\}$  ▷ Update des Trägers
5:    $\gamma_k \leftarrow \operatorname{argmin}_{\substack{\gamma \geq 0, \gamma(S_k) \leq \Sigma \\ \gamma(S_k^c) = 0}} \ell(\sum \gamma(\{z\}) \mathcal{G}\delta_z - \hat{y})$  ▷ Optimierung der Gewichte
6:    $S_k = \operatorname{supp}(\gamma_k)$  ▷ Löschung massfreier Punkte
7: end for

```

---

### 5.2.3. Ausnutzung der Differenzierbarkeit

Boyd et al. machen die Beobachtung, dass Algorithmus 1 in der Praxis oft langsam konvergiert und Maße mit unnötig vielen Trägerpunkten generiert. Dies wird darauf zurückgeführt, dass Algorithmus 1 nur Trägerpunkte hinzufügen und löschen kann, diese aber nicht glatt innerhalb des Parameterraumes verschoben werden können.

Häufig besitzt der Messoperator  $\mathcal{G}$  jedoch zusätzliche differenzierbare Struktur: Für das Bild eines Dirac-Maßes am Punkt  $z \in \Omega$  gilt in der  $j$ -ten Komponente

$$(\mathcal{G}\delta_z)_j = \overline{\langle e_j, \mathcal{G}\delta_z \rangle} = \int_{\Omega} \overline{\mathcal{G}^* e_j} d\delta_z = \overline{(\mathcal{G}^* e_j)(z)},$$

wobei  $e_j$  wieder der  $j$ -te Einheitsvektor sei. Nehmen wir nun zusätzlich an, dass das Bild von  $\mathcal{G}^*$  nicht nur aus stetigen, sondern sogar aus differenzierbaren Abbildungen  $U \rightarrow \mathbb{C}$  für eine offene Menge  $U \supset \Omega$  besteht, so ist insbesondere die Abbildung  $z \mapsto \mathcal{G}\delta_z$  differenzierbar.

**5.2 Beispiel** Ist  $\mathcal{G}$  zum Beispiel der ideale Frequenzfilter, so besteht das Bild von  $\mathcal{G}^* = \mathcal{F}_{f_c}^*$  aus den trigonometrischen Polynomen  $z \mapsto e^{2\pi i l \cdot z}$  und die Annahme der Differenzierbarkeit ist offensichtlich erfüllt.

Die Idee ist nun, dem Verfahren einen neuen Schritt **local\_descent** hinzuzufügen, in welchem die Gewichte festgehalten, die Trägerpunkte jedoch glatt variiert werden: Für eine Menge  $A = \{(z_1, \alpha_1), \dots, (z_r, \alpha_r)\} \subset \Omega \times [0, \infty)$  von Paaren aus Trägerpunkten und Gewichten nutzt die Subroutine **local\_descent**(A) die Differenzierbarkeit der Funktion

$$(z_1, \dots, z_r) \mapsto \ell\left(\sum_{j=1}^r \alpha_j \mathcal{G}\delta_{z_j} - \hat{y}\right)$$

aus, um die Positionen  $z_j$  lokal zu verbessern.

Dieser Schritt wird in einer inneren Schleife abwechselnd mit der Gewichtsoptimierung ausgeführt, bis keine Verbesserung mehr erreicht wird oder eine gewisse Zahl von Iterationen überschritten ist.

---

**Algorithmus 2** Alternating descent conditional gradient method (ADCG)

---

```

1: for  $k = 1, \dots, k_{\max}$  do
2:    $g_k \leftarrow \tilde{\nabla} \ell(\mathcal{G}\gamma_{k-1} - \hat{y})$  ▷ Gradient der Verlustfunktion
3:   Wähle  $z_k \in \operatorname{argmin}_{\Omega} \operatorname{Re} \mathcal{G}^*(g_k)$  ▷ Neuer Quellpunkt
4:    $S_k \leftarrow S_{k-1} \cup \{z_k\}$  ▷ Update des Trägers
5:   repeat
6:      $\gamma_k \leftarrow \operatorname{argmin}_{\substack{\gamma \geq 0, \gamma(S_k) \leq \Sigma \\ \gamma(S_k^c) = 0}} \ell(\sum \gamma(\{z\}) \mathcal{G}\delta_z - \hat{y})$  ▷ Optimierung der Gewichte
7:      $S_k = \operatorname{supp}(\gamma_k)$  ▷ Löschung massfreier Punkte
8:      $S_k \leftarrow \mathbf{local\_descent}(\{(z, \gamma_k(\{z\})) \mid z \in S_k\})$  ▷ Lokaler Abstieg
9:   until convergence
10: end for

```

---

#### 5.2.4. Implementation der Teilschritte

Für die Implementation der Gewichtsoptimierung muss ein geeignetes Verfahren abhängig von der Verlustfunktion  $\ell$  gewählt werden. Handelt es sich z.B. um die quadrierte 2-Norm, d.h.  $\ell = \|\cdot\|_2^2$ , so ist das entsprechende Problem als LASSO-Regression bekannt [7]. Hierbei ist es von Vorteil, ein Verfahren zu wählen, welches dünnbesetzte Darstellungen bevorzugt, damit so wenig Punkte wie möglich im Träger  $S_k$  behalten werden müssen.

In der Subroutine **local\_descent** können die Positionen im Träger mittels eines Gradientenabstiegsverfahrens lokal optimiert werden, wobei die durch  $\Omega$  vorgegebenen Parametergrenzen eingehalten werden müssen.

Problematischer ist der Schritt

$$z_k \in \operatorname{argmin}_{z \in \Omega} (\operatorname{Re} \mathcal{G}^* g_k)(z) \quad (5.8)$$

zur Wahl eines neuen Trägerpunktes. Die genaue Form hängt natürlich vom konkreten Messoperator  $\mathcal{G}$  ab. Im Allgemeinen handelt es sich hierbei um ein nicht-konvexes Optimierungsproblem, welches aufgrund der entsprechenden Annahme an  $\mathcal{G}$  allerdings differenzierbar ist. Boyd et al. schlagen zur Lösung von (5.8) vor, zunächst ein grobes Gitter über den Parameterraum  $\Omega$  zu legen und anschließend die Differenzierbarkeit zu einer lokalen Optimierung auszunutzen.

Dieses Vorgehen wird jedoch schnell ineffizient, sobald die Dimension des Parameterraumes  $\Omega$  steigt. Insofern vermeidet der ADCG-Algorithmus zwar zunächst eine feine Gitter-Diskretisierung wie in Abschnitt 5.1, ist aber grundsätzlich von den gleichen Nachteilen betroffen. Gerade in der Anwendung der dynamischen Super-

Resolution, wo  $\Omega \subset \mathbb{R}^{2d}$  die Teilchen im Phasenraum parametrisiert, kann dies zu Problemen führen.

### 5.3. Dimensionsreduziertes Gitterverfahren in 2D

Das Ziel dieses Abschnittes ist es, eine Diskretisierung der dimensionsreduzierten Formulierung

$$\begin{aligned} \min_{\substack{(\mu_\theta)_{\theta \in \mathbb{S}^{d-1}} \subset \mathcal{M}(\mathbb{R}^2) \\ (\nu_n)_{n \in \mathbb{S}^1} \in \mathcal{Z}}} \sup_{\theta \in \mathbb{S}^{d-1}} \|\mu_\theta\|_{\text{TV}} \quad \text{s.t.} \\ B[x \mapsto s_k x]_{\#} \nu_{n_k} = \hat{y}_k \quad \forall k \in \mathcal{K}, \\ (R_\theta)_{\#} \nu_n = (Q_n^1)_{\#} \mu_\theta \quad \forall n \in \mathbb{S}^1 \forall \theta \in \mathbb{S}^{d-1} \end{aligned} \quad (\text{ERdyn}^{\text{rel}})$$

in  $d = 2$  zu entwickeln, bei der sich die Dimensionsreduktion in eine tatsächliche Ersparnis bezüglich der Anzahl der benötigten Variablen umsetzt. Den Ausgangspunkt hierfür bietet die Gitter-Diskretisierung aus Abschnitt 5.1.

#### 5.3.1. Diskretisierung der Variablen

Hierzu wählen wir zunächst endliche Mengen von Richtungen  $\Theta \subset \mathbb{S}^1$  sowie  $\mathcal{N} \subset \mathbb{S}^1$ . Da in der ersten Nebenbedingung von (ERdyn<sup>rel</sup>) die Variable  $\nu_n$  für  $n = n_k$  ausgewertet wird, müssen die Richtungen  $n_k$  auf jeden Fall in  $\mathcal{N}$  enthalten sein. Wir definieren also  $\mathcal{N}_{\mathcal{K}} := \{n_k \mid k \in \mathcal{K}\}$  und stellen die Bedingung

$$\mathcal{N} = \mathcal{N}_{\mathcal{K}} \cup \mathcal{N}_+$$

für eine Menge  $\mathcal{N}_+ \subset \mathbb{S}^1$  von zusätzlichen Richtungen.

Für jedes  $\theta \in \Theta$  wird die Variable  $\mu_\theta$  wie folgt diskretisiert: Nach Lemma 4.8 ist  $\mu_\theta$  auf einem Parallelogramm in  $\mathbb{R}^2$  konzentriert. Es wird nun ein Gitter über dieses Parallelogramm gelegt, indem zunächst ein reguläres Rechtecks-Gitter mit  $s^2$  Gitterpunkten auf dem Quadrat  $[0, 1]^2$  generiert wird und dieses anschließend mittels einer geeigneten affinen Abbildung auf das Parallelogramm abgebildet wird. Das so konstruierte Gitter wird mit  $G_\mu(\theta)$  bezeichnet und besteht aus  $s^2$  Punkten. Die Variable  $\mu_\theta$  wird nun repräsentiert durch einen Vektor  $\vec{\mu}_\theta \in \mathbb{R}^{s^2}$ , dessen Einträge die Gewichte des Maßes  $\mu_\theta$  an den Gitterpunkten aus  $G_\mu(\theta)$  sind.

Analog werden die Variablen  $\nu_n$  für  $n \in \mathcal{N}$  diskretisiert, wobei hierfür ein Gitter auf den Träger aus Lemma 4.9 gelegt wird. Die Menge  $G_\nu(n)$  sei das entsprechende Gitter und  $\vec{\nu}_n \in \mathbb{R}^{s^2}$  der Vektor der Gewichte auf den Gitterpunkten.

#### 5.3.2. Diskretisierung der Operatoren

Für die erste Nebenbedingung in (ERdyn<sup>rel</sup>) ist für jedes  $k \in \mathcal{K}$  der Operator  $B[x \mapsto s_k x]_{\#}$  zu diskretisieren. Dieser Operator wird in der Nebenbedingung auf

die Variable  $\nu_{n_k}$  angewendet. Da jedoch  $\nu_{n_k}$  auf unterschiedlichen Gittern abhängig von  $n_k$  diskretisiert ist, sind insgesamt  $|\mathcal{K}|$  Matrizen erforderlich:

$$M_B(k) := (B[x \mapsto s_k x]_{\#} \delta_g)_{g \in G_{\nu}(n_k)} \in \mathbb{R}^{m \times s^2}, \quad k \in \mathcal{K}$$

Für die zweite Nebenbedingung sind die Operatoren  $(R_{\theta})_{\#}, (Q_n^1)_{\#}$  für  $\theta \in \Theta, n \in \mathcal{N}$  zu diskretisieren. Da wir in  $d = 2$  arbeiten, stimmen diese im Kontinuierlichen tatsächlich überein, d.h.  $(R_{\theta})_{\#} = (Q_n^1)_{\#}$  für  $\theta = n$ . Allerdings werden die Operatoren in der zweiten Nebenbedingung auf die Variablen  $\mu_{\theta}, \nu_n$  für unterschiedliche  $\theta, n$  angewendet, sodass aufgrund der verschiedenen Gitter auch unterschiedliche Matrizen benötigt werden.

Die Diskretisierung der Radon-Transformation wird am Beispiel von  $\mu$  erläutert. Die Konstruktion für  $\nu$  funktioniert analog, jedoch mit vertauschten Rollen von  $\theta$  und  $n$ .

Es sei nun ein  $\theta \in \Theta$  sowie ein  $n \in \mathcal{N}$  gegeben. Es ist eine Matrix zu konstruieren, welche die Projektion der Variable  $\mu_{\theta}$  in Richtung  $n$ , d.h. die Operation  $(Q_n^1)_{\#} \mu_{\theta}$  realisiert.

Hierzu wird für jeden Gitterpunkt  $g \in G_{\mu}(\theta)$  die Projektion  $n \cdot g \in \mathbb{R}$  in Richtung  $n$  betrachtet. Es werden Radien  $r_1 < r_2 < \dots < r_q \in \mathbb{R}$  gewählt, sodass die Projektionen aller Gitterpunkte aus  $G_{\mu}(\theta)$  im Intervall  $[r_1, r_q]$  liegen.

Bezeichne  $G_{\mu}(\theta) = \{g_1, \dots, g_{s^2}\}$ . Die Matrix  $M = M_{\mu}(\theta, n) \in \mathbb{R}^{q \times s^2}$  ist nun definiert durch

$$M_{i,j} := \begin{cases} \frac{r_{i+1} - n \cdot g_j}{r_{i+1} - r_i} & \text{falls } n \cdot g_j \in [r_i, r_{i+1}), \\ \frac{n \cdot g_j - r_{i-1}}{r_i - r_{i-1}} & \text{falls } n \cdot g_j \in [r_{i-1}, r_i), \\ 0 & \text{sonst,} \end{cases} \quad 2 \leq i \leq q-1, \quad 1 \leq j \leq s^2$$

$$M_{1,j} := \begin{cases} \frac{r_2 - n \cdot g_j}{r_2 - r_i} & \text{falls } n \cdot g_j \in [r_1, r_2), \\ 0 & \text{sonst,} \end{cases} \quad 1 \leq j \leq s^2$$

$$M_{q,j} := \begin{cases} \frac{n \cdot g_j - r_{q-1}}{r_q - r_{q-1}} & \text{falls } n \cdot g_j \in [r_{q-1}, r_q], \\ 0 & \text{sonst,} \end{cases} \quad 1 \leq j \leq s^2.$$

Wie schon erwähnt verläuft die Konstruktion der entsprechenden Matrizen für  $\nu$  analog und führt auf Matrizen  $M_{\nu}(n, \theta) \in \mathbb{R}^{q \times s^2}$ .

### 5.3.3. Diskretes Minimierungsproblem

Die kontinuierliche Formulierung in (ERdyn<sup>rel</sup>) fordert exakte Gleichheit in den Nebenbedingungen. Die Diskretisierung der Variablen und Operatoren sowie numerische Fehler führen jedoch dazu, dass exakte Gleichheit nicht mehr zu erwarten ist. Stattdessen wird im diskreten Problem die Abweichung von exakter Gleichheit in der 2-Norm minimiert. Analog zum ADCG-Algorithmus wird die Annahme gemacht, dass die Gesamtmasse  $\Sigma > 0$  des Zielmaßes bekannt ist. Die diskrete

Formulierung lautet

$$\begin{aligned}
 \min_{\substack{\vec{\mu}_\theta \in \mathbb{R}^{s^2}, \theta \in \Theta \\ \vec{v}_n \in \mathbb{R}^{s^2}, n \in \mathcal{N}}} & \sum_{k \in \mathcal{K}} \|M_B(k) \vec{v}_{n_k} - \hat{y}_k\|_2^2 + \sum_{\theta \in \Theta, n \in \mathcal{N}} \|M_\mu(\theta, n) \vec{\mu}_\theta - M_\nu(n, \theta) \vec{v}_n\|_2^2 \\
 & \text{subject to} \\
 & \|\vec{\mu}_\theta\|_1 \leq \Sigma \quad \forall \theta \in \Theta, \\
 & \|\vec{v}_n\|_1 \leq \Sigma \quad \forall n \in \mathcal{N}.
 \end{aligned} \tag{ERdyn\#}$$

Die einzelnen Matrizen werden zu einer Systemmatrix zusammengefügt und das entstehende konische Programm mit der Pythonbibliothek *CVXOPT* [12] gelöst. Als Backend wird der Löser *MOSEK* [3] genutzt.

### 5.3.4. Nachbearbeitung

Das Ergebnis des diskreten Problems (ERdyn<sub>#</sub><sup>-</sup>) sind Vektoren  $\vec{\mu}_\theta, \vec{v}_n$  für  $\theta \in \Theta, n \in \mathcal{N}$ . Insbesondere enthalten die Vektoren  $\vec{v}_n$  für  $n = n_k, k \in \mathcal{K}$  diejenigen Gewichte, welche die rekonstruierten Teilchen zum  $k$ -ten Zeitschritt repräsentieren.

Um einen Vergleich zu gitterfreien Verfahren zu ermöglichen, ist stattdessen eine Liste der erkannten Teilchen vorteilhaft. Hierzu wird, Tang et al. [22] folgend, ein Clustering-Verfahren auf die Gewichtsvektoren angewandt. Zunächst werden dabei alle Gewichte, deren Beträge unterhalb eines Schwellwertes  $\alpha_{\min}$  liegen, auf 0 gesetzt. Anschließend werden aneinander angrenzende Gewichte zu Clustern zusammengefasst. Pro Cluster wird anhand der Gewichte und Positionen der Schwerpunkt des Clusters berechnet, welcher schließlich zur Liste der erkannten Teilchenpositionen hinzugefügt wird.

## 5.4. Vergleich der Verfahren

Verglichen werden folgende Verfahren:

1. **static:** Die Gitter-Diskretisierung des statischen Verfahrens, welches nur einzelne Zeitschritte betrachtet. Hierzu wird einfach die Implementation von (ERdyn<sub>#</sub><sup>-</sup>) mit  $\Theta = \mathcal{N}_+ = \emptyset$  verwendet.
2. **ADCG:** Das volldimensionale dynamische Problem wird mit dem ADCG-Algorithmus gelöst.
3. **dir5n0:** Das dimensionsreduzierte Verfahren mit  $|\Theta| = 5$  und  $\mathcal{N}_+ = \emptyset$ .
4. **dir30n0:** Das dimensionsreduzierte Verfahren mit  $|\Theta| = 30$  und  $\mathcal{N}_+ = \emptyset$ .
5. **dir30n5:** Das dimensionsreduzierte Verfahren mit  $|\Theta| = 30$  und  $|\mathcal{N}_+| = 5$ .

---

### 5.4.1. Datensatz

Zum Vergleich der Methoden wird ein Datensatz von 1000 Teilchenkonfigurationen in  $d = 2$  generiert. Jede Konfiguration besteht aus  $4 \leq N \leq 20$  Teilchen mit zufälligen Positionen und Geschwindigkeiten  $(x_i, v_i) \in \Omega$ . Die Gewichte  $\alpha_i$  werden gleichverteilt zwischen 0.9 und 1.1 gewählt. Es wird  $K = 1$  und  $\tau = 1$  gewählt, d.h. es werden drei Zeitschritte  $\mathcal{K} = \{-1, 0, 1\}$  betrachtet.

Die dynamische Separation einer Teilchenkonfiguration  $T = \{(x_i, v_i)\}_{i=1}^N$  sei definiert als (vgl. [2, eq. (25)])

$$\Delta_{\text{dyn}}(T) := \min_{k \in \{-1, 0, 1\}} \min_{i \neq j} d_{\mathbb{T}^2}(x_i + k\tau v_i, x_j + k\tau v_j)$$

mit der Torus-Metrik  $d_{\mathbb{T}^2}$  aus Gleichung (2.15). Die dynamische Separation ist also der minimale Abstand zweier Teilchen der Konfiguration über die betrachteten Zeitschritte.

Um die Eigenschaften der verschiedenen Rekonstruktionsmethoden zuverlässig für verschiedene Werte von  $\Delta_{\text{dyn}}$  untersuchen zu können, ist es wünschenswert, dass im Datensatz genügend Beispielkonfigurationen in allen Wertebereichen von  $\Delta_{\text{dyn}}$  enthalten sind. Bei einer gleichverteilten Ziehung der Teilchen auf  $\Omega$  sind jedoch niedrige und hohe Werte von  $\Delta_{\text{dyn}}$  unterrepräsentiert.

Um dies zu beheben, wird ein *rejection sampling*-Algorithmus aus [1] übernommen. Kurz gesagt wird zunächst aus der  $N$ -fachen Gleichverteilung auf  $\Omega$  ein Vorschlag für eine neue Teilchenkonfiguration generiert. Diese wird mit einer bestimmten Wahrscheinlichkeit angenommen, d.h. dem Datensatz hinzugefügt, ansonsten wird sie abgelehnt. Die Akzeptanzwahrscheinlichkeit wird so gewählt, dass die Ungleichverteilung bezüglich  $\Delta_{\text{dyn}}$  ausgeglichen wird.

Der im Folgenden genutzte Datensatz enthält Konfigurationen  $T$  mit Werten  $\Delta_{\text{dyn}}(T) \in [0, 0.1]$ .

### 5.4.2. Parameterwahl

#### Modellparameter

Wie schon erwähnt, werden Zeitschritte  $\mathcal{K} = \{-1, 0, 1\}$  betrachtet. Als Messoperator  $B$  soll der ideale Frequenzfilter auf dem Gebiet  $[0, 1]^2$  dienen.

Zur Wahl der Grenzfrequenz  $f_c$  ist folgende Überlegung anzustellen: Es ist zu erwarten, dass die gitterbasierten Verfahren Teilchen nicht mehr auflösen können, deren Separation unterhalb der Gitterfeinheit liegt. Gleichzeitig lassen die Aussagen zum kontinuierlichen Problem aus Satz 2.18 vermuten, dass es eine fundamentalere Limitierung der Auflösung abhängig von der Grenzfrequenz gibt.

Um diese Effekte voneinander trennen zu können, sollte die Grenzfrequenz so niedrig gewählt werden, dass auch Teilchenkonfigurationen mit Separation über der Gitterfeinheit nicht in allen Fällen exakt rekonstruiert werden. Gleichzeitig ist klar, dass der Rechenaufwand der gitterbasierten Verfahren stark von der Anzahl

der Gitterpunkte abhängt. Es wird daher eine relativ niedrige Grenzfrequenz  $f_c = 3$  gewählt.

### Diskretisierungsparameter

Die Parameter für das diskrete Problem werden wie folgt gewählt:

- Anzahl Gitterpunkte entlang einer Dimension:  $s = 100$
- Schwellwert für das Clustering:  $\alpha_{\min} = 0.1$
- Anzahl Radian für die Diskretisierung der Radon-Transformation:  $q = 100$

### Parameter für ADCG

Es wird die Implementation aus [1] genutzt und für den ADCG-Algorithmus werden folgende Parameter gewählt:

- Maximale Iterationen der äußeren Schleife: 100
- Gitter im Schritt zur Suche eines neuen Quellpunktes: 20x20 Gitter für Positionen, 20x20 Gitter für Geschwindigkeiten

### 5.4.3. Evaluation

Zur Evaluation der Methoden werden nur die Lokalisierungen im Zeitschritt  $k = 0$  herangezogen. Die Rekonstruktion einer Teilchenkonfiguration durch eine der Methoden wird als Erfolg betrachtet, wenn

1. es zu jedem tatsächlichen Teilchen  $(x_i, v_i)$  ein rekonstruiertes Teilchen  $\tilde{x}$  gibt, sodass

$$\|\tilde{x} - x_i\|_2 < \Delta x,$$

2. keine anderen Teilchen erkannt wurden.

Es ist klar, dass die Auflösung der gitterbasierten Verfahren durch die Feinheit des Gitters begrenzt ist. Dies muss bei der Wahl der Toleranz berücksichtigt werden und wir wählen deshalb  $\Delta x = 0.01$ .

Schließlich wird der Wertebereich  $[0, 0.1]$  für die im Datensatz enthaltenen dynamischen Separationen in gleichmäßige Abschnitte unterteilt. In jedem Abschnitt wird für die Methoden die *correct reconstruction rate* (CRR), d.h. das Verhältnis von erfolgreich rekonstruierten Konfigurationen zur Gesamtanzahl an Konfigurationen berechnet.

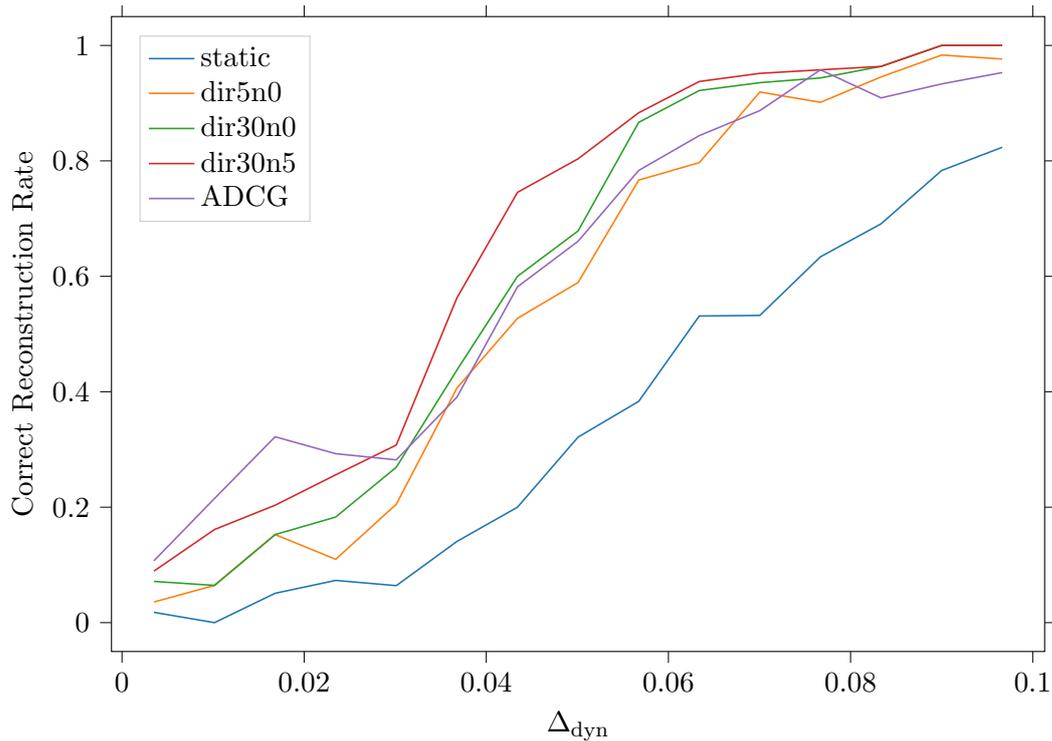


Abbildung 5.1.: Vergleich der Verfahren

#### 5.4.4. Ergebnis und Diskussion

Das Ergebnis des Vergleichs ist in Abb. 5.1 dargestellt. Wie erwartet, schneidet die *static* Methode, welche nur die Daten aus einem Zeitschritt betrachtet, am schlechtesten ab. Besonders für größere Werte von  $\Delta_{\text{dyn}}$  erreichen die Verfahren, welche die Dynamik des Problems mit einbeziehen, eine deutlich höhere Rekonstruktionsrate.

Die dimensionsreduzierten Methoden rekonstruieren besser, je mehr Richtungen  $\theta$  bzw.  $n$  in das Problem aufgenommen werden. Dies geht natürlich auch mit einem höheren Rechenaufwand einher. Im Bereich  $\Delta_{\text{dyn}} \leq 0.03$  rekonstruiert der ADCG-Algorithmus am besten. Dies ist konsistent mit der Erwartung, da die gitterbasierten Verfahren hier an die durch die Gitterfeinheit vorgegebene Grenze stoßen. Für größere Werte von  $\Delta_{\text{dyn}}$  liegen die Rekonstruktionsraten der ADCG-Methode ungefähr zwischen denen der *dir5n0* und der *dir30n0* Methode.

Auffällig ist ein Absinken der Rekonstruktionsrate des ADCG-Algorithmus im Bereich um  $\Delta_{\text{dyn}} = 0.03$ . Dies widerspricht zunächst der Erwartung, da größere Teilchenabstände die Rekonstruktion bei allen Methoden erleichtern sollten. Eigentlich wird also bei allen Methoden ein monotoner Verlauf erwartet. Jedoch sind auch an anderen Stellen Verstöße gegen diese Erwartung festzustellen, wie z.B. bei der *dir5n0* Methode um  $\Delta_{\text{dyn}} = 0.08$ . Diese sind deshalb wahrscheinlich auf statis-

tische Schwankungen aufgrund eines zu kleinen Datensatzes zurückzuführen. Eine Wiederholung des Experiments mit einem größeren Datensatz könnte diese Theorie bestätigen.

Überraschend ist auch, dass der ADCG-Algorithmus, welcher auf dem volldimensionalen kontinuierlichen Modell basiert, von den dimensionsreduzierten Verfahren *dir30n0* und *dir30n5* übertroffen wird. Offensichtlich scheint sich die höhere Dimension des kontinuierlichen Modells nicht in bessere Rekonstruktionsraten (außer im Bereich  $\Delta_{\text{dyn}} \leq 0.03$ ) umzusetzen. Möglicherweise würden andere Parameter für den ADCG-Algorithmus, wie z.B. ein feineres Gitter im Schritt zur Suche eines neuen Quellpunktes, zu einem besseren Ergebnis führen.

Insgesamt stellen wir fest, dass die dimensionsreduzierten Verfahren die Informationen aus allen Zeitschritten erfolgreich zu einer besseren Rekonstruktion nutzen. Sogar die Methode *dir5n0*, welche lediglich 5 Richtungen  $\theta$  in das Problem aufnimmt, übertrifft die *static* Methode weit.

# Schlusswort

In dieser Arbeit wurde ein neues Verfahren zur Dimensionsreduktion des dynamischen Super-Resolution Problems vorgestellt. Hierzu wurde zunächst ein Überblick über die Theorie der Super-Resolution von Punktquellen gegeben, wobei der Fokus auf der exakten Rekonstruktion lag. Anschließend wurde das dynamische Modell eingeführt und es wurden die Resultate zur exakten Rekonstruktion in diesem Setting besprochen. Dabei wurde angenommen, dass die statischen Probleme die Teilchen zu mindestens drei Zeitschritten rekonstruieren können.

Die Dimensionsreduktion des Problems konnte durch eine Abwandlung der Radon-Transformation erreicht werden, wodurch für eine Raumdimension  $d$  eine Reduktion von  $2d$  auf  $d+1$  möglich ist. Es wurden verschiedene mögliche Problemformulierungen diskutiert. Anschließend wurde das dimensionsreduzierte Problem unter dem Aspekt der exakten Rekonstruktion untersucht und die verschiedenen Formulierungen wurden untereinander und mit dem volldimensionalen dynamischen Problem in Beziehung gesetzt. Unser Hauptresultat ist eine Aussage, welche die exakte Rekonstruktion des dimensionsreduzierten Problems garantiert, solange zu mindestens drei Zeitschritten eine statische Rekonstruktion möglich ist. Dieses ist somit analog zu dem für das volldimensionale Problem bekannten Resultat.

Im letzten Kapitel wurden die theoretischen Überlegungen numerisch untermauert. Hierzu musste zunächst eine geeignete Diskretisierung des dimensionsreduzierten Problems entwickelt werden. Die Experimente haben gezeigt, dass sogar eine geringe Kopplung der einzelnen Zeitschritte durch nur wenige hinzugenommene Projektionsrichtungen zu einer besseren Rekonstruktion führt. Das dimensionsreduzierte Verfahren ist konkurrenzfähig mit dem volldimensionalen Modell, implementiert über den ADCG-Algorithmus, und übertrifft diesen für eine höhere Anzahl an Projektionsrichtungen sogar.

Einige Dinge wurden in dieser Arbeit nicht behandelt und stehen für zukünftige Forschung offen: Die Resultate zur exakten Rekonstruktion knüpfen an die Theorie zum volldimensionalen dynamischen Problem an. Hier wären jedoch Aussagen wünschenswert, welche nicht die exakte Rekonstruktion der statischen Probleme annehmen müssen. Diese müssten jedoch der Einfachheit halber zunächst für das volldimensionale Problem gezeigt werden.

Nicht untersucht wurde die Stabilität des Verfahrens, d.h. der Einfluss von fehlerbehafteten Daten auf die Rekonstruktion. Hierzu sind für das volldimensionale dynamische Problem einige Aussagen von Alberti et al. bekannt, welche möglicherweise auf das dimensionsreduzierte Problem übertragen werden könnten.

Die Rekonstruktion der Geschwindigkeiten aus der Lösung des dimensionsreduzierten Problems ist möglich (vgl. Abschnitt 4.2.2), wurde jedoch in dieser Arbeit

nicht weiter behandelt. Auch hier wäre ein Vergleich zum volldimensionalen Verfahren interessant.

# Literatur

- [1] Giovanni S. Alberti et al. *Code for „Dynamic Spike Super-resolution and Applications to Ultrafast Ultrasound Imaging“*. 2018. URL: [https://github.com/panchoop/dynamic\\_spike\\_super\\_resolution](https://github.com/panchoop/dynamic_spike_super_resolution).
- [2] Giovanni S. Alberti et al. „Dynamic Spike Super-resolution and Applications to Ultrafast Ultrasound Imaging“. In: *CoRR* abs/1803.03251 (2018). arXiv: 1803.03251. URL: <http://arxiv.org/abs/1803.03251>.
- [3] MOSEK ApS. *MOSEK Optimizer API for Python 9.1.5*. 2019. URL: <https://docs.mosek.com/9.1/pythonapi/index.html>.
- [4] Vladimir I. Bogachev. *Measure Theory*. Springer Berlin Heidelberg, 2007. ISBN: 978-3-540-34513-8. DOI: 10.1007/978-3-540-34514-5.
- [5] Jan Boman und Filip Lindskog. „Support Theorems for the Radon Transform and Cramér-Wold Theorems“. In: *Journal of Theoretical Probability* 22.3 (Sep. 2009), S. 683–710. ISSN: 1572-9230. DOI: 10.1007/s10959-008-0151-0.
- [6] N. Boyd, G. Schiebinger und B. Recht. „The alternating descent conditional gradient method for sparse inverse problems“. In: *Proc. IEEE 6th Int. Workshop Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*. Dez. 2015, S. 57–60. DOI: 10.1109/CAMSAP.2015.7383735.
- [7] Stephen Boyd und Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004. DOI: 10.1017/CB09780511804441.
- [8] Kristian Bredies und Hanna Katriina Pikkarainen. „Inverse problems in spaces of measures“. In: *ESAIM: Control, Optimisation and Calculus of Variations* 19.1 (2013), S. 190–218. DOI: 10.1051/cocv/2011205.
- [9] Emmanuel J. Candès und Carlos Fernandez-Granda. „Towards a Mathematical Theory of Super-resolution“. In: *Communications on Pure and Applied Mathematics* 67.6 (2014), S. 906–956. DOI: 10.1002/cpa.21455.
- [10] Yohann de Castro und Fabrice Gamboa. „Exact reconstruction using Beurling minimal extrapolation“. In: *Journal of Mathematical Analysis and Applications* 395.1 (2012), S. 336–354. ISSN: 0022-247X. DOI: 10.1016/j.jmaa.2012.05.011.
- [11] Y. Chi et al. „Sensitivity to Basis Mismatch in Compressed Sensing“. In: *IEEE Transactions on Signal Processing* 59.5 (Mai 2011), S. 2182–2195. DOI: 10.1109/TSP.2011.2112650.
- [12] CVXOPT. *CVXOPT User’s Guide*. 2019. URL: <http://cvxopt.org/userguide/index.html>.

- [13] Vincent Duval und Gabriel Peyré. „Exact Support Recovery for Sparse Spikes Deconvolution“. In: *Foundations of Computational Mathematics* 15.5 (Okt. 2015), S. 1315–1355. ISSN: 1615-3383. DOI: 10.1007/s10208-014-9228-6.
- [14] Ivar Ekeland und Roger Témam. *Convex Analysis and Variational Problems*. Society for Industrial und Applied Mathematics, 1999. DOI: 10.1137/1.9781611971088.
- [15] Gerald B. Folland. *Real Analysis: Modern Techniques and Their Applications*. 2. Aufl. Pure and Applied Mathematics. John Wiley & Sons, 1999. ISBN: 978-0-471-31716-6.
- [16] Marguerite Frank und Philip Wolfe. „An algorithm for quadratic programming“. In: *Naval Research Logistics Quarterly* 3.1-2 (1956), S. 95–110. DOI: 10.1002/nav.3800030109.
- [17] Loukas Grafakos. *Classical Fourier Analysis*. 2. Aufl. Springer New York, 2008. ISBN: 978-0-387-09432-8. DOI: 10.1007/978-0-387-09432-8.
- [18] Jeff W. Lichtman und José-Angel Conchello. „Fluorescence microscopy“. In: *Nature Methods* 2.12 (Dez. 2005), S. 910–919. ISSN: 1548-7105. DOI: doi : 10.1038/nmeth817.
- [19] K.R. Parthasarathy. *Probability Measures on Metric Spaces*. Probability and Mathematical Statistics: A Series of Monographs and Textbooks. Academic Press, 1967. ISBN: 978-0-8218-3889-1.
- [20] Walter Rudin. *Real and complex analysis*. 3. Aufl. McGraw-Hill, 1987. ISBN: 978-0-07-054234-1.
- [21] Michael J. Rust, Mark Bates und Xiaowei Zhuang. „Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)“. In: *Nature Methods* 3.10 (Okt. 2006), S. 793–796. ISSN: 1548-7105. DOI: 10.1038/nmeth929.
- [22] G. Tang, B. N. Bhaskar und B. Recht. „Sparse recovery over continuous dictionaries-just discretize“. In: *2013 Asilomar Conference on Signals, Systems and Computers*. Nov. 2013, S. 1043–1047. DOI: 10.1109/ACSSC.2013.6810450.
- [23] G. Tang et al. „Compressed Sensing Off the Grid“. In: *IEEE Transactions on Information Theory* 59.11 (Nov. 2013), S. 7465–7490. DOI: 10.1109/TIT.2013.2277451.
- [24] Lei Zhu et al. „Faster STORM using compressed sensing“. In: *Nature Methods* 9.7 (Juli 2012), S. 721–723. ISSN: 1548-7105. DOI: 10.1038/nmeth.1978.

# A. Anhang

## A.1. Beweise aus Kapitel 2

### A.1.1. Beweis von Lemma 2.9

Zunächst zitieren wir eine Aussage aus [20] über den Zusammenhang zwischen komplexen und reellen Dualräumen:

**A.1 Proposition** ([20, Prop. 5.17]) Es sei  $V$  ein komplexer Vektorraum.

1. Ist  $\rho$  der Realteil eines komplex-linearen Funktionals  $\gamma$  auf  $V$ , so ist  $\rho$  ein reell-lineares Funktional und es gilt

$$\gamma(v) = \rho(v) - i\rho(iv) \quad \forall v \in V. \quad (\text{A.1})$$

2. Ist  $\rho$  ein reell-lineares Funktional auf  $V$  und sei  $\gamma$  das Funktional definiert durch Gleichung (A.1), dann ist  $\gamma$  ein komplex-lineares Funktional auf  $V$ .
3. Ist  $V$  ein normierter Vektorraum und erfüllen  $\rho$  und  $\gamma$  die Beziehung in (A.1), dann gilt  $\|\rho\|_{V'} = \|\gamma\|_{V'}$ .

Nun zum Beweis des Lemmas:

*Beweis von Lemma 2.9.* Zunächst ist die Abbildung  $f \mapsto \operatorname{Re} \int_{\Omega} f \, d\bar{\mu}$  für jedes  $\mu \in \mathcal{M}(\Omega)$  reell-linear und stetig, also im Dualraum  $C(\Omega)'^{\mathbb{R}}$  enthalten und  $\Gamma$  ist wohldefiniert. Klar ist auch, dass  $\Gamma$  selbst eine reell-lineare Abbildung ist, denn das Integral ist auch linear im Maß.

Zur Surjektivität: Sei  $\rho : C(\Omega) \rightarrow \mathbb{R}$  ein stetiges, reell-lineares Funktional. Nach Proposition A.1 ist  $\gamma$  definiert über Gleichung (A.1) ein komplex-lineares, stetiges Funktional auf  $C(\Omega)$  mit  $\rho = \operatorname{Re} \gamma$ . Definiere  $\nu : C(\Omega) \rightarrow \mathbb{C}$  durch

$$\nu(f) := \overline{\gamma(f)}.$$

Auch  $\nu$  ist komplex-linear und stetig und darum ein Element von  $C(\Omega)'$ .

Der Darstellungssatz von Riesz (Satz 1.17) liefert nun ein Maß  $\mu \in \mathcal{M}(\Omega)$ , sodass für alle  $f \in C(\Omega)$  gilt

$$\nu(f) = \int_{\Omega} f \, d\mu.$$

Dieses Maß erfüllt

$$\gamma(f) = \overline{\nu(f)} = \overline{\int_{\Omega} f \, d\mu} = \int_{\Omega} f \, d\bar{\mu}$$

für alle  $f \in C(\Omega)$ , also ist  $\rho = \operatorname{Re} \gamma = \Gamma\mu$  im Bild von  $\Gamma$ .

Zur Isometrie: Sei  $\mu \in \mathcal{M}(\Omega)$  beliebig und nutze die Bezeichnung

$$T_{\bar{\mu}}f = \int_{\Omega} f \, d\bar{\mu}, \quad f \in C(\Omega).$$

Nach Proposition A.1 gilt

$$\|\operatorname{Re} T_{\bar{\mu}}\|_{C(\Omega)'} = \|T_{\bar{\mu}}\|_{C(\Omega)'}$$

und nach dem Darstellungssatz von Riesz gilt auch  $\|T_{\bar{\mu}}\|_{C(\Omega)'} = \|\bar{\mu}\|_{\operatorname{TV}}$ . Schließlich ist offensichtlich  $\|\bar{\mu}\|_{\operatorname{TV}} = \|\mu\|_{\operatorname{TV}}$ . Die Injektivität folgt aus der Isometrie.  $\square$

### A.1.2. Beweis von Proposition 2.20

*Beweis.* Die Darstellung des Dirichlet-Kerns über Sinus-Funktionen ergibt sich aus folgender Rechnung über eine geometrische Summe von komplexen Exponentialen:

$$\begin{aligned} D_k(z) &= \sum_{l=-k}^k e^{i2\pi lz} = e^{-i2\pi kz} \sum_{l=0}^{2k} e^{i2\pi lz} = \frac{e^{-i\pi(2k+1)z} e^{i2\pi(2k+1)z} - 1}{e^{-i\pi z} - e^{i2\pi z} - 1} \\ &= \frac{e^{i\pi(2k+1)z} - e^{-i\pi(2k+1)z}}{e^{i\pi z} - e^{-i\pi z}} = \frac{\sin((2k+1)\pi z)}{\sin(\pi z)} \end{aligned}$$

Der Ausdruck für den Fejér-Kern kann hierauf zurückgeführt werden:

$$\begin{aligned} F_n(z) &= \frac{1}{n+1} \sum_{k=0}^n D_k(z) = \frac{1}{(n+1) \sin(\pi z)} \sum_{k=0}^n \sin((2k+1)\pi z) \\ &= \frac{1}{2(n+1) \sin(\pi z)^2} \sum_{k=0}^n 2 \sin((2k+1)\pi z) \sin(\pi z) \\ &= \frac{1}{2(n+1) \sin(\pi z)^2} \sum_{k=0}^n (\cos(k2\pi z) - \cos((k+1)2\pi z)) \\ &= \frac{1 - \cos((n+1)2\pi z)}{2(n+1) \sin(\pi z)^2} = \frac{1}{n+1} \frac{\sin((n+1)\pi z)^2}{\sin(\pi z)^2} \end{aligned} \tag{A.2}$$

Im letzten Schritt wurde die Winkelverdopplungsformel des Kosinus genutzt.

Die Koeffizienten des Fejér-Kerns berechnen sich auf folgende Weise:

$$\begin{aligned}
F_n(z) &= \frac{1}{n+1} \sum_{k=0}^n \sum_{l=-k}^k e^{i2\pi lz} \\
&= \frac{1}{n+1} \sum_{l=-n}^n e^{i2\pi lz} \#\{k \in \{0, \dots, n\} \mid |l| \leq k\} \\
&= \frac{1}{n+1} \sum_{l=-n}^n e^{i2\pi lz} (n+1 - |l|) = \sum_{l=-n}^n \left(1 - \frac{|l|}{n+1}\right) e^{i2\pi lz} \quad (\text{A.3})
\end{aligned}$$

Aus Definition von  $K$  zusammen mit Gleichung (A.2) folgt, dass  $K$  tatsächlich der normalisierte quadratische Fejér-Kern zu  $n = f_c/2$  ist:

$$K(z) = \frac{1}{(f_c/2 + 1)^2} \left( \frac{\sin((f_c/2 + 1)\pi z)^2}{(f_c/2 + 1) \sin(\pi z)^2} \right)^2 = \frac{1}{(f_c/2 + 1)^2} F_n(z)^2$$

Schreibe kurz  $n := f_c/2$ . Einsetzen von Gleichung (A.3) führt zu folgender Rechnung:

$$\begin{aligned}
K(z) &= \frac{1}{(n+1)^2} \sum_{l_1=-n}^n \sum_{l_2=-n}^n \left(1 - \frac{|l_1|}{n+1}\right) \left(1 - \frac{|l_2|}{n+1}\right) e^{i2\pi(l_1+l_2)z} \\
&= \frac{1}{(n+1)^2} \sum_{r=-2n}^{2n} \sum_{\substack{l \in \{-n, \dots, n\} \\ |r-l| \leq n}} \left(1 - \frac{|l|}{n+1}\right) \left(1 - \frac{|r-l|}{n+1}\right) e^{i2\pi rz}
\end{aligned}$$

Da die Bedingung  $|r-l| \leq n$  umgeformt werden kann zu  $r-n \leq l \leq r+n$ , entsprechen die Koeffizienten des obigen trigonometrischen Polynoms genau dem Ausdruck  $g(r)$  aus der Aussage der Proposition.  $\square$

## A.2. Beweise aus Abschnitt 4.3

*Beweis von Lemma 4.8.* Es ist  $(y, w) \in \hat{R}_\theta(\Omega)$  genau dann, wenn  $(x^{(j)}, v^{(j)}) \in \Omega_1$ ,  $j = 1, \dots, d$  existieren, sodass

$$(y, w) = \theta_1(x^{(1)}, v^{(1)}) + \dots + \theta_d(x^{(d)}, v^{(d)}).$$

Also gilt unter Verwendung der Minkowski-Summe von Mengen

$$\hat{R}_\theta(\Omega) = \theta_1\Omega_1 + \dots + \theta_d\Omega_1 = \sum_{j:\theta_j>0} \theta_j\Omega_1 + \sum_{j:\theta_j<0} (-\theta_j)(-\Omega_1).$$

Die Mengen  $\Omega_1$  und  $-\Omega_1$  sind konvex und für die Minkowski-Summe gilt die Formel

$$\begin{aligned} \sum_{j:\theta_j>0} \theta_j \Omega_1 + \sum_{j:\theta_j<0} (-\theta_j)(-\Omega_1) &= \left( \sum_{j:\theta_j>0} \theta_j \right) \Omega_1 + \left( \sum_{j:\theta_j<0} (-\theta_j) \right) (-\Omega_1) \\ &= s_+ \Omega_1 + s_- (-\Omega_1) \end{aligned}$$

Es bleibt nur, die Minkowski-Summe von

$$s_+ \Omega_1 = \text{conv} \left\{ (0, 0), (s_+, 0), \left( \frac{s_+}{2}, \frac{s_+}{2\delta} \right), \left( \frac{s_+}{2}, -\frac{s_+}{2\delta} \right) \right\}$$

und

$$s_- \Omega_1 = \text{conv} \left\{ (0, 0), (s_-, 0), \left( \frac{s_-}{2}, -\frac{s_-}{2\delta} \right), \left( \frac{s_-}{2}, \frac{s_-}{2\delta} \right) \right\}$$

zu berechnen. Nach der Formel  $\text{conv}(A) + \text{conv}(B) = \text{conv}(A + B)$  führt dies auf eine konvexe Hülle von 16 Punkten, wobei jedoch leicht zu sehen ist, dass sich diese allesamt als Konvexkombinationen der vier Eckpunkte in der Aussage des Lemmas schreiben lassen.  $\square$

*Beweis von Lemma 4.9.* Es ist  $z \in Q_n^d(\Omega)$  genau dann, wenn ein  $(x, v) \in \Omega$  existiert, sodass  $n_1 x + n_2 v = z$ . Da  $\Omega$  bis auf Umordnung der Variablen das  $d$ -fache kartesische Produkt von  $\Omega_1$  ist, ist dies genau dann erfüllt, wenn für jedes  $j = 1, \dots, d$  ein  $(x^{(j)}, v^{(j)}) \in \Omega_1$  existiert mit  $z_j = n_1 x^{(j)} + n_2 v^{(j)}$ . Es gilt also

$$Q_n^d(\Omega) = \prod_{j=1}^d Q_n^1(\Omega_1)$$

und es genügt, die Menge  $Q_n^1(\Omega_1)$  zu verstehen.

Da  $Q_n^1$  linear ist, gilt

$$Q_n^1(\Omega_1) = Q_n^1(\text{conv } S) = \text{conv}(Q_n^1(S)).$$

Nun ist

$$\begin{aligned} \text{conv}(Q_n^1(S)) &= \text{conv} \left\{ 0, n_1, \frac{n_1}{2} + \frac{n_2}{2\delta}, \frac{n_1}{2} - \frac{n_2}{2\delta} \right\} \\ &= \text{conv} \left\{ 0, n_1, \frac{n_1}{2} + \frac{|n_2|}{2\delta}, \frac{n_1}{2} - \frac{|n_2|}{2\delta} \right\} \end{aligned}$$

und hierbei handelt es sich um eine konvexe Teilmenge von  $\mathbb{R}$ , also ein Intervall. Um die Intervallgrenzen zu bestimmen, ist eine Fallunterscheidung notwendig:

**Fall 1:**  $n_1 \geq \frac{|n_2|}{\delta}$ . Hieraus folgt

$$0 \leq \frac{n_1}{2} - \frac{|n_2|}{2\delta} \leq \frac{n_1}{2} + \frac{|n_2|}{2\delta} \leq \frac{n_1}{2} + \frac{n_1}{2} = n_1.$$

---

Insgesamt ist  $\text{conv}(Q_n^1(S)) = [0, n_1]$ , d.h.  $Q_n^d(\Omega) = [0, n_1]^d$ .

**Fall 2:**  $n_1 \leq -\frac{|n_2|}{\delta}$ . Hieraus folgt

$$n_1 = \frac{n_1}{2} + \frac{n_1}{2} \leq \frac{n_1}{2} - \frac{|n_2|}{2\delta} \leq \frac{n_1}{2} + \frac{|n_2|}{2\delta} \leq 0$$

und somit  $Q_n^d(\Omega) = [n_1, 0]^d$ .

**Fall 3:**  $|n_1| < \frac{|n_2|}{\delta}$ . In diesem Fall ist

$$\frac{n_1}{2} - \frac{|n_2|}{2\delta} \leq 0 \leq \frac{n_1}{2} + \frac{|n_2|}{2\delta}$$

und

$$\frac{n_1}{2} - \frac{|n_2|}{2\delta} \leq \frac{n_1}{2} - \frac{|n_1|}{2} \leq \frac{n_1}{2} + \frac{n_1}{2} = n_1.$$

Umgekehrt gilt auch

$$n_1 = \frac{n_1}{2} + \frac{n_1}{2} \leq \frac{n_1}{2} + \frac{|n_1|}{2} \leq \frac{n_1}{2} + \frac{|n_2|}{2\delta},$$

sodass insgesamt  $Q_n^d(\Omega) = [\frac{n_1}{2} - \frac{|n_2|}{2\delta}, \frac{n_1}{2} + \frac{|n_2|}{2\delta}]^d$ .

□



# Eigenständigkeitserklärung

Hiermit versichere ich, dass die vorliegende Arbeit

## **Dimensionsreduktion von dynamischer Super-Resolution**

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 entnommen wurden, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht worden sind.

---

(Ort, Datum)

---

(Unterschrift)

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.

---

(Ort, Datum)

---

(Unterschrift)