



Fachbereich VI Informatik und Medien

**Masterarbeit**

von

Radmir Gesler

Zur Erlangung  
des akademischen Grades  
Master of Engineering (M.Eng.)

Im Studiengang  
Technische Informatik

Thema:

Physics Informed Neural Networks (PINN) zur Lösung eines inversen Problems der  
Wärmeleitungsgleichung

Betreuer: Prof. Dr. Frank Haußer  
Gutachterin: Prof. Dr. Yin Amy Siu

Eingereicht am: 07. Juli 2025

## Eidesstattliche Erklärung

Hiermit erkläre ich, dass ich die vorliegende Masterarbeit selbstständig verfasst und nur unter Zuhilfenahme der angegebenen Quellen und Hilfsmitteln angefertigt habe. Sämtliche Stellen der Arbeit, die im Wortlaut oder dem Sinn nach anderen gedruckten oder im Internet verfügbaren Werken entnommen sind, habe ich durch genaue Quellenangaben kenntlich gemacht.

Berlin, 28. Februar 2025

---

Radmir Gelser

## Abstract

The present work contains a comprehensive discussion about the current methods of machine learning methods to solve partial differential equations and their comparison to traditional methods like FEM.

....

## Zusammenfassung

Die vorliegende Arbeit stellt eine zusammenfassende Diskussion über die Methoden des maschinellen Lernens zur Lösung der partiellen Differentialgleichungen und deren Vergleich zu numerischen Methoden. Als speziell Fall wird das Problem der Wärmeleitungsgleichung betrachtet.

...

# Inhaltsverzeichnis

<b>1. Einführung</b>	<b>1</b>
1.1. Was ist Wärmeübertragung . . . . .	1
1.1.1. Wärmeleitung . . . . .	1
1.1.2. Wärmekonvektion . . . . .	1
1.1.3. Wärmestrahlung . . . . .	1
1.2. Definition des Vorhabens . . . . .	2
1.2.1. Das Problem . . . . .	2
<b>2. Wärmeleitungsproblem</b>	<b>3</b>
2.1. Herleitung der Wärmeleitungsgleichung . . . . .	3
2.1.1. Analytische Untersuchung der Wärmeleitungsgleichung . . . . .	6
2.2. Vorwärtsproblem der Wärmeleitungsgleichung . . . . .	6
2.2.1. Herleitung der Variationsformulierung der Wärmegleichung . . . . .	6
2.2.2. Geometrie und Gitter definieren . . . . .	6
2.2.3. Funktionsräume definieren . . . . .	7
2.2.4. Variationsformulierung herleiten . . . . .	7
2.2.5. Zeitdiskretisierung . . . . .	8
2.2.6. Backward-Euler-Verfahren . . . . .	8
2.2.7. Schlüsselkonzepte . . . . .	9
2.3. Inverses Problem der Wärmeleitung . . . . .	9
2.3.1. Existenz und Eindeutigkeit . . . . .	10
2.3.2. Stabilität . . . . .	11
2.3.3. Regularisierung des inversen Problems . . . . .	11
2.4. Ansatz des maschinellen Lernens . . . . .	12
2.5. Statistisches Lernen: Grundlagen . . . . .	12
2.5.1. Definition des statistischen Lernens . . . . .	12
2.5.2. Lineare Modelle und Regression . . . . .	12
2.5.3. Nichtlineare Modelle und neuronale Netze . . . . .	13
2.5.4. Das Prinzip der neuronalen Netze (NN) . . . . .	13
2.5.5. Optimierte Gradienten-Methoden . . . . .	15
2.6. Einführung in Physics-Informed Neural Networks . . . . .	15
2.6.1. Motivation für PINNs . . . . .	15
2.6.2. Theorie für das inverse Problem der Wärmeleitung . . . . .	15
2.6.3. Formulierung als Variationsproblem . . . . .	16
2.6.4. Existenz und Eindeutigkeit . . . . .	16
2.6.5. PINNs für das inverse Problem . . . . .	16
2.6.6. Eigenschaften und Herausforderungen . . . . .	17
<b>3. Auswertung der praktischen Arbeit</b>	<b>18</b>
3.1. Architektur der Implementierung . . . . .	18
3.2. Diskussion der Ergebnisse . . . . .	18

3.3. Ausblick . . . . .	18
<b>A. Implementierung</b>	<b>55</b>
<b>B. Inhalte der vollen Masterarbeit</b>	<b>56</b>
B.1. PDF-Dateien . . . . .	56
B.2. Python Code . . . . .	56
<b>Quellenverzeichnis</b>	<b>57</b>

# 1. Einführung

## 1.1. Was ist Wärmeübertragung

Die Beobachtung der Naturvorgänge, die mit dem Transport der Wärme zu tun haben, hat zu einer Feststellung geführt. Und zwar, falls ein Medium zwei Gebiete mit unterschiedlichen Temperaturen besitzt, bewegt sich die Wärme immer aus dem Gebiet mit der höheren Temperatur in das, wo eine niedrigere Temperatur herrscht. Dieser Vorgang der Wärmeübertragung kann auf drei Arten geschehen, durch sogenannte Wärmeleitung, durch Wärmekonvektion und durch Wärmestrahlung. Im folgendem wollen wir diese Begrifflichkeiten zuerst klären.

### 1.1.1. Wärmeleitung

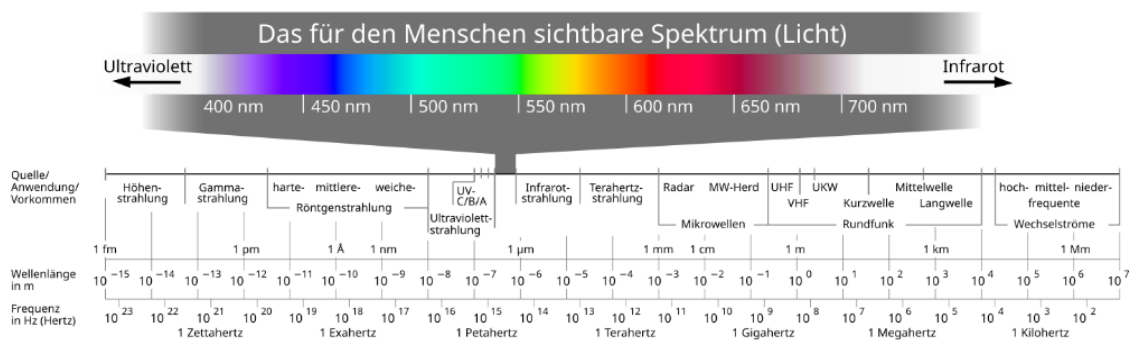
Man stelle sich vor, wie ein Metallstab an einem Ende aufgeheizt wird. Nach einiger Zeit würde man an dem anderen Ende eine Temperaturerhöhung messen können. Im Umkehrschluss hieße es, dass die Wärme von dem Stabende, wo eine Wärmequelle angelegt wurde, zu dem anderem gewandert ist. Eine physikalische Erklärung dafür könnte so lauten, die Atome, wo eine Wärmequelle angelegt ist, fangen an stärker zu schwingen, also bekommen mehr Energie. Durch die Gitterstruktur im Metall regen die schon stärker schwingende Atome ihre Nachbarn an und so weiter, bis die Atome an dem anderen Ende auch stärker schwingen. Somit wird die Wärmeenergie in Form schwingender Atome durch einen Festkörper transportiert, in solchen Fällen spricht man von *Wärmeleitung*.

### 1.1.2. Wärmekonvektion

Bei der *Konvektion* handelt es sich um Transport der Stoffmengen. Dieser Prozess findet in festen, flüssigen und gasförmigen Materialzuständen statt. Zum Beispiel bei Flüssigkeiten und Gasen tritt eine Ausgleichsströmung durch den Dichteunterschied zwischen warmen und kühleren Gebieten auf. In solchen Fällen sind die wärmeren Gebiete spezifisch leichter und steigen in die Höhe, wo sie sich dann abkühlen. Der beschriebene Verlauf ist ein Fall für sogenannte *freie Konvektion*. Eine *erzwungene Konvektion* liegt dann vor, wenn eine äußere Einwirkung dem System hinzugefügt wird. Hierfür könnte eine Vulkaneruption als Beispiel dienen.

### 1.1.3. Wärmestrahlung

Im Vakuum, da wo keine Materie vorhanden ist, kann die Wärmeübertragung weder durch Wärmeleitung noch durch Konvektion stattfinden. Nehmen wir ein Beispiel der Wärmeübertragung von der Sonne zur Erde an. Trotz der Tatsache, dass der



**Abbildung 1.1.:** Das Spektrum der elektromagnetischen Strahlung. Quelle: Wikipedia.

Raum zwischen der Erde und Sonne (fast) leer von Materie ist, findet die Wärmeübertragung von der Sonne auf die Erde statt. Es ist die *Wärmestrahlung*, die für die Energieübertragung sorgt. Im Allgemeinen werden solche Prozesse durch das Konzept der elektromagnetischen Strahlung erklärt. Fakt ist, dass die elektromagnetische Strahlung kein Medium braucht um sich im Raum auszubreiten. Strahlung an sich kann mittels Wellen mathematisch beschrieben werden. Wellen schwingen, alles was schwingt hat eine Frequenz und kann daher in einem Frequenzspektrum eingeordnet werden. So ist die Wärmestrahlung nur ein Teil des gesamten elektromagnetischen Spektrums, wobei die UV-Strahlung, das sichtbare Spektrum und die Infrarot-Strahlung die Wärmestrahlung ausmachen (Abb.1.1). Tritt die Wärmestrahlung auf Materie, so wird sie absorbiert, wodurch sich die Materie aufwärmt oder gar erhitzt.

## 1.2. Definition des Vorhabens

### 1.2.1. Das Problem

Die vorangegangenen Überlegungen werden unten im Text herangezogen, um die theoretische Herangehensweise zum Problem der Wärmeausbreitung in Stoffen beizustellen ... bla bla ... was wird gemacht:

1. Überlegungen, die aus physikalischen Gegebenheiten auf eine Gleichung Führen, im Kapitel 2.
2. Gleichung benennen, als Wärmegleichung, ev. Kapitel so und so.
3. Mögliche Problemstellungen benennen, direkte und inverse ..., ev. Kapitel so und so
4. Das Vorhaben benennen. Mit NN inverses Problem lösen, bla bla
5. Die Struktur des Vorhabens schildern: FEM -> Lösung -> einige Werte der Lösung als Input für NN, , ev. Kapitel so und so
6. Vergleich der Ergebnisse, ev. Kapitel so und so

Bla Bla ... somit ist das Ziel der vorliegenden Arbeit festgelegt worden

## 2. Wärmeleitungsproblem

### 2.1. Herleitung der Wärmeleitungsgleichung

Im Allgemeinen beschreibt die Wärmegleichung die Ausbreitung der Wärme in einem Medium. Das Medium kann beispielsweise ein räumlicher Gegenstand sein, zum Beispiel eine rechteckige Metallplatte. Man stelle sich vor, die gegebene Platte wird in der Mitte kurz erhitzt. Diesen Prozess kann man mithilfe einer Wärmebildkamera sichtbar machen.

Würde man diese Situation über einen Zeitraum beobachten, so würde man feststellen, dass die Temperatur in der Mitte der Platte sich über die gesamte Fläche verteilt. Und irgendwann kommt es zum Abkühlen, also zur Gleichverteilung der Temperatur in der Platte.

Der beschriebene Prozess kann mithilfe sogenannter Wärmeleitungsgleichung modelliert werden, die wir zunächst herleiten möchten.

#### Wärmeausbreitung

Sei  $\Omega \subset \mathbb{R}^3$  das Gebiet auf dem die Verteilung der Temperatur bestimmt werden soll.<sup>1</sup> Die gesuchte Temperaturverteilung auf  $\Omega$  in einem Zeitintervall  $T = [t_0, t_1] \subset \mathbb{R}$  bezeichnen wir mit  $u$ .  $u$  ordnet jedem Punkt  $x \in \Omega$  zu jedem Zeitpunkt  $t \in T$  eine Zahl  $u(x, t) \in \mathbb{R}$  zu, also  $u : \Omega \times T \rightarrow \mathbb{R}$ . Zusätzlich wollen wir die Bewegung der Wärme in  $\Omega$  auffassen, es ist der Wärmefluss  $j$ . Für jeden Punkt  $x \in \Omega$  zu jedem Zeitpunkt  $t \in T$  gibt der Wärmefluss die Richtung an, in der die Wärmeenergie transportiert wird an, somit  $j : \Omega \times T \rightarrow \mathbb{R}^3$ . Damit sind nötigen *Variablen* für die Modellierung festgelegt.

Für die Herleitung der Gleichungen der Modellierung wird ein beliebiges Volumen  $V \subset \Omega$  betrachtet. In einem homogenen Fall, also es gibt keine Wärmequellen, die in das Volumen  $V$  die Wärmeenergie hinein oder heraus transportieren, würde man keinen Wärmestrom am Rand  $dV$  detektieren, es gilt also

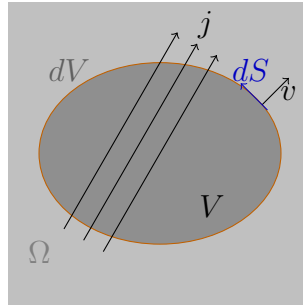
$$\int_{\partial V} \langle j, v \rangle dS = 0. \quad (2.1)$$

Mit der Gleichung 2.1 wird der Strom  $j$  in Richtung jeder Normalen  $v$  am Rand  $dV$  von  $V$  aufsummiert (Abb.2.1). Und mit der Annahme, dass es keine Quellen gibt, erwarten wir eine 0 Summe.

---

<sup>1</sup> Im Allgemeinen gilt  $\Omega \subset \mathbb{R}^n$ .





**Abbildung 2.1.:** Schematische Darstellung der Gleichung 2.1, mit  $V$  das Volumen eines Mediums,  $dV$  Rand (orange) von  $V$ ,  $j$  Der Wärmestrom durch  $V$ ,  $v \in \mathbb{R}^3$  Normalenvektor an  $dV$  und  $dS$  Tangente an  $dV$ .

An dieser Stelle benötigen wir den *Gauß'schen Satz* um das Oberflächenintegral 2.1 in ein Volumenintegral umzuschreiben.

**Definition 1. Der Gauß'sche Satz**

Sei  $V \subset \mathbb{R}^n$  kompakt mit abschnittsweise glattem Rand  $S = \partial V$ , der Rand ist orientiert durch äußeres Normaleneinheitsvektorfeld  $n$ . Das Vektorfeld  $F$  sei stetig differenzierbar auf einer offenen Menge  $\Omega$  mit  $V \in \Omega$ . Dann gilt

$$\int_V \langle \nabla, F \rangle dV = \int_{\partial V} \langle F, n \rangle dS. \quad (2.2)$$

**Temporäres DenkZettel**

**Definition 2. Kompakte Menge**

$\mathbb{R}^n$  ist metrischer Raum. Wenn es für jede **offene Überdeckung** von  $A \subset \mathbb{R}^n$ , eine endliche Überdeckung gibt, also

$$A \subset \bigcup_{i \in I} U_i, \quad \text{mit endlichem } I$$

**Definition 3. Glatter Rand**

Sei  $A \in \mathbb{R}^n$  kompakt,  $A$  hat einen glatten Rand, wenn es zu jedem Punkt  $x \in \partial A$  eine offene Umgebung  $U \subseteq \mathbb{R}^n$  und  $\psi : U \rightarrow \mathbb{R}$ ,  $\psi \in C^n$  so gibt, dass

- (a)  $A \cap U = \{x \in U : \psi(x) \leq 0\}$
- (b)  $\forall x \in U : \psi'(x) \neq 0$

gilt.

Nun, der Gauß'sche Satz angewandt auf Gleichung 2.1 liefert folgendes Ergebnis

$$\int_V \langle \nabla, j \rangle dx = \int_{\partial V} \langle j, v \rangle dS = 0 \quad (2.3)$$

Da, das Volumen  $V$  nicht näher spezifiziert war, können wir aus 2.3 folgern, dass **Dahmen/Reusken: S.456, sinnvolle Argumentation nach dem Gaußschen Satz**

$$\langle \nabla, j \rangle = 0 \quad (2.4)$$

auf ganz  $\Omega$  gilt.

Als Nächstes, wollen wir herausfinden, wie  $j$  und  $u$  von einander abhängen. Eine einfache Modellannahme gibt uns das *Fourier'sche Gesetz*:<sup>2</sup>

*Die Wärmeenergie strömt immer von Wärmeren Regionen zu den kälteren. Die Flussrate des Stroms ist proportional zum Temperaturgradienten.*

Eine weitere Annahme für das Modell ist, dass jedes Material eine Eigenschaft hat, die Wärme auf seine Art und Weise zu leiten. Diese Eigenschaft nennt man die *Wärmeleitfähigkeit*  $a$ . Sie ordnet jedem Punkt im Körper eine positive Zahl zu, also  $a : \Omega \rightarrow \mathbb{R}_+$ . Führen wir beide Überlegungen zusammen und erhalten

$$j = -a \nabla u. \quad (2.5)$$

Einsetzen der Gleichung 2.5 in 2.4 führt uns zum folgenden Ausdruck

$$-\langle \nabla, a \nabla u \rangle = 0 \quad (2.6)$$

Falls auf  $\Omega$  noch zusätzliche Wärmequellen  $f : \Omega \rightarrow \mathbb{R}$  vorhanden sind und die Wärmeverteilung *instationär* ist, erhalten wir insgesamt folgende Gleichung [0, S.5]

$$c\rho \partial_t u - \langle \nabla, a \nabla u \rangle = f \quad (2.7)$$

$c$  spezifische Wärmekapazität [ $J/Kkg$ ]  
 $\rho$  - Einheit [ $kg/m^3$ ]

### Einschub über die Klassifikation der DGL

Nie Gleichung 2.7 beschreibt ein abstraktes System. Die sogenannten Anfangs- und Randbedingungen würden eine mehr der Realität entsprechende Beschreibung des Wärmeleitungsproblems liefern. Jedes physikalische System besitzt zum Anfang der Messung oder der Beobachtung einen gewissen Anfangszustand  $u_0$ . Meistens wird die Zeit  $t$  zu diesem Anfang auf den Wert 0 gesetzt, somit erhalten wir

$$u(x, 0) = u_0(x), \quad x \in \Omega \quad (2.8)$$

---

<sup>2</sup> J. B. Fourier, 1768-1830

Die Randbedingungen geben die Temperaturverteilung  $u$  am Rand  $\partial\Omega$  des Gebiets  $\Omega$  und können im Allgemeinen mit Sturm'schen Randbedingungen angegeben werden (oder gemischte Randbedingung?):

$$u(x, t) = g(x, t), \quad x \in \partial\Omega \quad (2.9)$$

Nun kann man die Gleichungen 2.7, 2.8 und 2.9 zusammen in einer Rand- und Anfangswertaufgabe (RAWP) der Wärmeleitung zusammenfassen, also:

**Definition 4. Rand- und Anfangswertaufgabe der Wärmeleitung**

Sei  $\Omega \subset \mathbb{R}^3$  ein Gebiet mit glattem Rand  $\partial\Omega$ , der Rand ist orientiert durch äußeres Normaleneinheitsvektorfeld  $n$ . So beschreibt das Modell

$$\begin{aligned} \partial_t u - \langle \nabla, a \nabla u \rangle &= f \\ u(x, 0) &= u_0(x), \quad x \in \Omega \\ u(x, t) &= g(x, t), \quad x \in \partial\Omega \end{aligned} \quad (2.10)$$

die Rand- und Anfangswertaufgabe der Wärmeleitung für alle  $t \in T \subset \mathbb{R}$ .

Das Ergebnis dieses Kapitels ist in der obigen Definition 4 angegeben und wird als Grundmodell in folgenden Kapiteln dieser Arbeit vorausgesetzt.

### 2.1.1. Analytische Untersuchung der Wärmeleitungsgleichung

Eventuell ...

## 2.2. Vorwärtsproblem der Wärmeleitungsgleichung

FEM als Grundlage des numerischen Ansatzes, warum und wieso ...

### 2.2.1. Herleitung der Variationsformulierung der Wärmegleichung

Die Wärmegleichung lautet:

$$\partial_t u - \langle \nabla, a \nabla u \rangle = f, \quad (2.11)$$

und soll in die Variationsformulierung in schwacher Form umgeformt werden:

$$\int_{\Omega} \frac{\partial u}{\partial t} v \, dx + \int_{\Omega} a(u) \nabla u \cdot \nabla v \, dx = \int_{\Omega} f v \, dx. \quad (2.12)$$

Im Folgenden werden die Schritte zur Herleitung detailliert erläutert.

### 2.2.2. Geometrie und Gitter definieren

Die Geometrie und das Gitter definieren den Bereich  $\Omega$ , in dem die Gleichung gelöst wird. Das Gitter teilt diesen Bereich in kleine Unterbereiche (Finite Elemente) auf.

### Warum Geometrie und Gitter definieren?

- Die Methode der Finiten Elemente (FEM) basiert auf der Approximation der Lösung  $u(x, y, t)$  auf einer diskreten Darstellung des Gebiets.
- Ein Gitter zerlegt  $\Omega$  in einfachere Formen (z. B. Dreiecke oder Vierecke). Innerhalb jedes Elements wird die Lösung lokal approximiert.

### Beispiel für ein rechteckiges Gebiet

$$\Omega = \{(x, y) \mid 0 \leq x \leq L_x, 0 \leq y \leq L_y\}. \quad (2.13)$$

Dieses Gebiet könnte mit Dreiecken oder Rechtecken diskretisiert werden.

### 2.2.3. Funktionsräume definieren

Ein Funktionsraum beschreibt, wie die Lösung  $u$  und die Testfunktionen  $v$  innerhalb der Finiten Elemente dargestellt werden.

#### Was sind $P_1$ - und $P_2$ -Elemente?

- $P_1$ : Lineare Basisfunktionen –  $u$  wird innerhalb eines Elements als lineares Polynom approximiert.
- $P_2$ : Quadratische Basisfunktionen –  $u$  wird innerhalb eines Elements als quadratisches Polynom approximiert.

Diese Basisfunktionen gewährleisten, dass  $u$  über das gesamte Gebiet hinweg stetig ist.

#### Warum Funktionsräume definieren?

Die Funktionsräume sorgen dafür, dass die approximierte Lösung  $u$  und die Testfunktionen  $v$  die notwendigen Kontinuitäts- und Randbedingungen der PDE erfüllen.

### 2.2.4. Variationsformulierung herleiten

Die Variationsformulierung (schwache Form) wird durch Multiplikation der PDE mit einer Testfunktion  $v$  und Integration über das Gebiet  $\Omega$  hergeleitet. Dieser Schritt erlaubt es, auch Lösungen  $u$  zu berücksichtigen, die möglicherweise nicht überall differenzierbar sind.

#### Herleitung der schwachen Form

##### 1. Starte mit der PDE:

$$\frac{\partial u}{\partial t} - \nabla \cdot (a(u) \nabla u) = f. \quad (2.14)$$

2. **Multipliziere mit einer Testfunktion  $v$ :** Multipliziere beide Seiten mit einer glatten Testfunktion  $v \in V$ , wobei  $V$  der Raum der zulässigen Testfunktionen ist:

$$v \left( \frac{\partial u}{\partial t} - \nabla \cdot (a(u) \nabla u) \right) = v f. \quad (2.15)$$

3. Integriere über  $\Omega$ :

$$\int_{\Omega} v \frac{\partial u}{\partial t} dx - \int_{\Omega} v \nabla \cdot (a(u) \nabla u) dx = \int_{\Omega} v f dx. \quad (2.16)$$

4. **Wende den Divergenzsatz an:** Für den zweiten Term wende den Divergenzsatz an, um das Volumenintegral in ein Oberflächenintegral umzuwandeln:

$$\int_{\Omega} v \nabla \cdot (a(u) \nabla u) dx = \int_{\partial\Omega} v (a(u) \nabla u \cdot \mathbf{n}) dS - \int_{\Omega} \nabla v \cdot (a(u) \nabla u) dx. \quad (2.17)$$

Hier ist  $\partial\Omega$  der Rand des Gebiets und  $\mathbf{n}$  der nach außen gerichtete Normalenvektor.

5. **Vereinfachung bei homogenen Randbedingungen:** Unter der Annahme homogener Dirichlet-Randbedingungen ( $u = 0$  auf dem Rand) verschwindet der Oberflächenterm:

$$\int_{\Omega} v \nabla \cdot (a(u) \nabla u) dx = - \int_{\Omega} \nabla v \cdot (a(u) \nabla u) dx. \quad (2.18)$$

6. **Kombiniere die Terme:** Substituiere zurück in die Gleichung. Die schwache Form lautet:

$$\int_{\Omega} v \frac{\partial u}{\partial t} dx + \int_{\Omega} a(u) \nabla u \cdot \nabla v dx = \int_{\Omega} v f dx. \quad (2.19)$$

## 2.2.5. Zeitdiskretisierung

Die Zeitdiskretisierung ist notwendig, um die Lösung schrittweise für diskrete Zeitpunkte zu berechnen.

**Warum Zeitdiskretisierung?**

- Der Term  $\frac{\partial u}{\partial t}$  ist eine partielle Ableitung nach der Zeit. FEM-Frameworks lösen jedoch räumliche Ableitungen.
- Die Zeitdiskretisierung ersetzt  $\frac{\partial u}{\partial t}$  durch eine algebraische Differenz.

## 2.2.6. Backward-Euler-Verfahren

Das Backward-Euler-Verfahren approximiert  $\frac{\partial u}{\partial t}$  zum Zeitpunkt  $t_{n+1}$ :

$$\frac{\partial u}{\partial t} \approx \frac{u^{n+1} - u^n}{\Delta t}. \quad (2.20)$$

Setze dies in die schwache Form ein:

$$\int_{\Omega} \frac{u^{n+1} - u^n}{\Delta t} v dx + \int_{\Omega} a(u^{n+1}) \nabla u^{n+1} \cdot \nabla v dx = \int_{\Omega} v f^{n+1} dx. \quad (2.21)$$

Dies führt zu einem System von Gleichungen, das für jeden Zeitschritt numerisch gelöst werden kann.

### 2.2.7. Schlüsselkonzepte

- **Testfunktion  $v$ :** Eine Testfunktion  $v$  gehört demselben Funktionsraum wie  $u$  und stellt sicher, dass die schwache Form für alle zulässigen Lösungen  $u$  gilt.
- **Warum schwache Form?**
  - Die schwache Form ermöglicht es, die PDE für Funktionen  $u$  zu lösen, die möglicherweise nicht überall differenzierbar sind (z. B. wenn  $\nabla u$  nicht existiert).
  - Die Integration durch Teile verschiebt die Ableitung von  $u$  auf die Testfunktion  $v$ , die als glatt angenommen wird.

....

## 2.3. Inverses Problem der Wärmeleitung

Inverses Problem aus der Praktischen Sicht, also  $u(x, y, t)$  wird gemessen und  $a(x, y)$  dann bestimmt ...

Betrachten wir das Problem aus der Definition 4, insbesondere die Wärmeleitfähigkeit  $a$ . In vielen praktischen Problemstellungen hängt die Wärmeleitfähigkeit direkt von der Temperatur  $u(x, t)$  ab ([0, S. 506]), also gilt  $a : \mathbb{R} \rightarrow \mathbb{R}$ .

Die DGL aus der Definition 4 bleibt bei dieser Betrachtung dieselbe, jedoch wird  $a$  jetzt explizit als Funktion von  $u$  betrachtet, also:

$$\frac{\partial u}{\partial t} - \nabla \cdot (a(u) \nabla u) = f \quad (2.22)$$

Um  $a(u)$  zu isolieren, können wir folgendermaßen vorgehen:

1. Schreibe die Gleichung um, sodass der diffusive Term isoliert wird:

$$-\nabla \cdot (a(u) \nabla u) = f - \frac{\partial u}{\partial t}.$$

2. Verwende die Produktregel für den diffusen Term:

$$-\nabla \cdot (a(u) \nabla u) = -a(u) \nabla^2 u - \nabla a(u) \cdot \nabla u = -a(u) \Delta u - \nabla a(u) \cdot \nabla u.$$

3. Wenn wir annehmen, dass  $a(u)$  nur von  $u$  abhängt (keine explizite Abhängigkeit von  $x, y$ ), vereinfacht sich der Gradienten-Term:

$$\nabla a(u) = \frac{da(u)}{du} \nabla u.$$

4. Setze dies in die Gleichung ein:

$$-a(u) \nabla^2 u - \frac{da(u)}{du} (\nabla u \cdot \nabla u) = f - \frac{\partial u}{\partial t}.$$

5. Gruppiere die Terme, um  $a(u)$  zu isolieren:

$$a(u) = -\frac{1}{\nabla^2 u} \left( f - \frac{\partial u}{\partial t} + \frac{da(u)}{du} (\nabla u \cdot \nabla u) \right).$$

die Wärmeleitungsgleichung im Gebiet  $\Omega \subset \mathbb{R}^2$  mit zeitlicher Dimension  $t \in T \subset \mathbb{R}$ :

$$\frac{\partial u}{\partial t} - \nabla \cdot (a \nabla u) = f,$$

ergänzt durch Anfangsbedingungen

$$u(x, 0) = u_0(x, y) \quad \text{für } x \in \Omega,$$

und Dirichlet-Randbedingungen

$$u(x, t) = g(x, t) \quad \text{für } x \in \partial\Omega, t \in (0, T].$$

Das inverse Problem besteht darin,  $a(u)$  zu finden, sodass die Differentialgleichung und die Randbedingungen erfüllt werden, wenn  $u(x, t)$  und  $f(x, t)$  bekannt sind.

### 2.3.1. Existenz und Eindeutigkeit

#### Satz 1: Existenz der Lösung

Sei  $u(x, y, t)$  eine hinreichend glatte Lösung der Wärmeleitungsgleichung und  $f(x, y, t)$  sowie  $u_0(x, y)$  bekannt. Dann existiert unter geeigneten Regularitätsvoraussetzungen an die Daten eine Funktion  $a(x, y)$ , die das inverse Problem löst.

#### Beweisidee:

1. Schreibe die Wärmeleitungsgleichung in schwacher Form:

$$\int_{\Omega} \left( \frac{\partial u}{\partial t} v + a(x, y) \nabla u \cdot \nabla v \right) d\Omega = \int_{\Omega} f(x, y, t) v d\Omega,$$

für alle Testfunktionen  $v \in H^1(\Omega)$ .

2. Die Lösung  $a(x, y)$  ergibt sich aus der schwachen Form, indem  $u$  und  $\nabla u$  durch die Messdaten approximiert werden. Standardresultate aus der Variationsrechnung garantieren die Existenz von  $a(x, y)$ .

#### Satz 2: Eindeutigkeit der Lösung

Seien  $u(x, y, t)$  und  $f(x, y, t)$  eindeutig gegeben, und sei  $u$  hinreichend differenzierbar. Dann ist  $a(x, y)$  eindeutig bestimmt, wenn  $\nabla u \neq 0$  in  $\Omega$ .

**Beweis:** Nehmen wir an, es existieren zwei Funktionen  $a_1(x, y)$  und  $a_2(x, y)$ , die das inverse Problem lösen. Dann gilt:

$$\frac{\partial u}{\partial t} - \nabla \cdot (a_1 \nabla u) = f = \frac{\partial u}{\partial t} - \nabla \cdot (a_2 \nabla u).$$

Durch Subtraktion erhalten wir:

$$\nabla \cdot ((a_1 - a_2) \nabla u) = 0.$$

Da  $\nabla u \neq 0$ , folgt aus der Divergenzfreiheit, dass  $a_1 = a_2$  fast überall in  $\Omega$ . Dies zeigt die Eindeutigkeit.

### 2.3.2. Stabilität

#### Satz 3: Stabilität der Lösung

Sei  $u_{\text{obs}}(x, y, t)$  eine Rauschbehaftung der exakten Lösung  $u(x, y, t)$  mit einem maximalen Fehler  $\epsilon > 0$ , d. h.

$$\|u_{\text{obs}} - u\|_{L^2} \leq \epsilon.$$

Dann existiert eine Konstante  $C > 0$ , sodass der Fehler in  $a(x, y)$  beschränkt ist durch

$$\|a_{\text{obs}} - a\|_{L^2} \leq C\epsilon.$$

#### Beweisidee:

1. Schreibe die Wärmeleitungsgleichung als Operatorgleichung:

$$\mathcal{A}(a) = f,$$

wobei  $\mathcal{A}(a) = \frac{\partial u}{\partial t} - \nabla \cdot (a \nabla u)$ .

2. Zeige, dass der Operator  $\mathcal{A}$  stetig und Lipschitz-stetig bezüglich  $a$  ist.
3. Wende den Satz von Lax-Milgram oder perturbative Stabilitätsresultate an, um die beschränkte Fehlerfortpflanzung von  $u$  auf  $a$  zu zeigen.

### 2.3.3. Regularisierung des inversen Problems

Da inverse Probleme oft schlecht gestellt sind, d. h. kleine Fehler in den Daten können zu großen Abweichungen in der Lösung führen, wird eine Regularisierung benötigt.

#### Definition (Regularisierte Zielfunktion)

Das inverse Problem wird als Minimierung einer modifizierten Zielfunktion formuliert:

$$J(a) = \int_{\Omega} \int_0^T \left| \frac{\partial u}{\partial t} - \nabla \cdot (a \nabla u) - f \right|^2 dt d\Omega + \lambda R(a),$$

wobei  $R(a)$  eine Regularisierungsfunktion ist, typischerweise:

- $R(a) = \|a\|_{H^1(\Omega)}^2$  (Tikhonov-Regularisierung) oder
- $R(a) = \int_{\Omega} |\nabla a|^2 d\Omega$  (Glattheitsregularisierung).

#### Satz 4: Existenz der regularisierten Lösung

Sei  $J(a)$  schwach unterer-halbstetig und sei  $R(a)$  strikt konvex. Dann existiert mindestens eine minimierende Funktion  $a \in H^1(\Omega)$ , die das reguläre inverse Problem löst.

#### Beweis:

1. Schwache Halbstetigkeit und kompakte Einbettungen von Sobolev-Räumen garantieren die Existenz einer minimierenden Funktion.
2. Die Konvexität von  $R(a)$  sorgt dafür, dass die Lösung eindeutig ist.



## 2.4. Ansatz des maschinellen Lernens

Traditionelle numerische Methoden wie Finite-Differenzen oder Finite-Elemente sind häufig mit hohen Rechenkosten verbunden. Eine vielversprechende Methode zur Lösung dieser Probleme stellen Physics-Informed Neural Networks (PINNs) dar, die maschinelles Lernen mit physikalischen Modellen kombinieren.

Dieses Kapitel beginnt mit einer detaillierten Einführung in die Grundlagen des statistischen Lernens, das eine mathematische Grundlage für maschinelles Lernen bildet. Anschließend wird die Theorie von PINNs eingeführt und detailliert ausgebaut.

## 2.5. Statistisches Lernen: Grundlagen

### 2.5.1. Definition des statistischen Lernens

Das Ziel des statistischen Lernens ist es, eine Abbildung  $f : X \rightarrow Y$  zu finden, die die zugrunde liegende Verteilung der Daten möglichst gut approximiert. Hierbei sind  $X$  der Raum der Eingabedaten und  $Y$  der Raum der Zielgrößen (z. B. Labels). Formal lässt sich das Problem des Lernens als Minimierung eines Risikos beschreiben.

Sei  $(x_i, y_i)$  eine i.i.d. Stichprobe aus einer unbekannten gemeinsamen Verteilung  $P(x, y)$ . Ziel ist es, ein Modell  $f$  aus einem Hypothesenraum  $\mathcal{F}$  zu finden, das die zugehörige Verlustfunktion minimiert. Das Risiko-Funktional ist gegeben durch:

$$R(f) = \mathbb{E}_{(x,y) \sim P}[L(y, f(x))],$$

wobei  $L : Y \times Y \rightarrow \mathbb{R}^+$  eine Verlustfunktion ist. Die Erwartung wird hierbei über die wahre Datenverteilung  $P(x, y)$  genommen.

Da jedoch die wahre Verteilung  $P(x, y)$  nicht bekannt ist, wird das Risiko durch den empirischen Risikofunktional approximiert:

$$R_n(f) = \frac{1}{n} \sum_{i=1}^n L(y_i, f(x_i)),$$

wobei  $(x_i, y_i)$  die beobachteten Trainingsdaten sind.

Die Zielsetzung des Lernens besteht also darin, den Schätzer  $f$  zu finden, der das Risiko minimiert:

$$\hat{f} = \arg \min_{f \in \mathcal{F}} R_n(f).$$

Eine zentrale Annahme des statistischen Lernens ist, dass das Modell  $f$  generalisierbar ist, d. h., es kann auf neue, unbekannte Daten angewendet werden.

### 2.5.2. Lineare Modelle und Regression

Eine der einfachsten und am häufigsten verwendeten Klassen von Modellen im statistischen Lernen sind lineare Modelle. Das grundlegende lineare Modell wird durch

$$f(x) = w^\top x + b$$

gegeben, wobei  $w \in \mathbb{R}^d$  der Gewichtsvektor und  $b \in \mathbb{R}$  der Bias-Term sind. In der linearen Regression wird  $f(x)$  so gewählt, dass der Fehler zwischen den Vorhersagen  $f(x_i)$  und den tatsächlichen Werten  $y_i$  minimiert wird.

Die Verlustfunktion ist hierbei oft die quadratische Verlustfunktion:

$$L(y, f(x)) = \|y - f(x)\|^2.$$

Um die Parameter  $w$  und  $b$  zu optimieren, minimiert man das empirische Risiko:

$$R_n(w, b) = \frac{1}{n} \sum_{i=1}^n \|y_i - (w^\top x_i + b)\|^2.$$

Durch Anwendung der Methode der kleinsten Quadrate ergibt sich eine Lösung für  $w$  und  $b$ , die durch die Normalengleichung beschrieben wird:

$$\hat{w}, \hat{b} = \arg \min_{w, b} R_n(w, b).$$

Die Lösung dieser Gleichung lässt sich durch:

$$\hat{w} = (X^\top X)^{-1} X^\top y$$

für  $w$  (wobei  $X$  die Designmatrix ist) und  $\hat{b}$  durch den Mittelwert der Zielwerte berechnen.

### 2.5.3. Nichtlineare Modelle und neuronale Netze

Da lineare Modelle für viele Probleme nicht ausreichend sind, hat sich die Nutzung nichtlinearer Modelle als Standard etabliert. Ein prominentes nichtlineares Modell ist das neuronale Netz. Ein neuronales Netz ist eine Verknüpfung von linearen Transformationen und nichtlinearen Aktivierungsfunktionen. Ein einfaches Feedforward-Netzwerk besteht aus mehreren Schichten, wobei die Funktion eines neuronalen Netzes durch

$$f(x) = \sigma(W^{(L)} \sigma(W^{(L-1)} \dots \sigma(W^{(1)} x + b^{(1)}) \dots + b^{(L-1)}) + b^{(L)})$$

gegeben ist. Hierbei sind  $W^{(l)}$  die Gewichtsmatrizen und  $b^{(l)}$  die Bias-Vektoren der  $l$ -ten Schicht, und  $\sigma$  ist eine nichtlineare Aktivierungsfunktion wie ReLU oder Sigmoid.

Neuronale Netze werden häufig mit Gradientenabstieg optimiert, indem die Kreuzentropie-Verlustfunktion oder die Mean Squared Error (MSE)-Verlustfunktion minimiert wird:

$$L(y, f(x)) = \|y - f(x)\|^2.$$

Die Optimierung erfolgt durch Rückpropagation und den Gradientenabstiegsalgorithmus.

Hier ein kurzer Exkurs zum Prinzip der Werteapproximation mit NN.

### 2.5.4. Das Prinzip der neuronalen Netze (NN)

#### Vorwärtsthroughlauf (Forward Pass)

Ein neuronales Netz besteht aus **Schichten** von Neuronen, die jeweils eine affine Transformation gefolgt von einer nichtlinearen Aktivierungsfunktion anwenden.

- **Eingangsvektor:**  $x \in \mathbb{R}^d$  (mit Eingabedimension  $d$ )
- **Gewichte:**  $W^{(l)} \in \mathbb{R}^{n_l \times n_{l-1}}$  (Anzahl der Neuronen pro Schicht  $n_l$ )
- **Bias:**  $b^{(l)} \in \mathbb{R}^{n_l}$
- **Aktivierungsfunktion:**  $\sigma : \mathbb{R} \rightarrow \mathbb{R}$

Die Transformation einer Schicht:

$$z^{(l)} = W^{(l)} a^{(l-1)} + b^{(l)} \quad (2.23)$$

Die Aktivierung:

$$a^{(l)} = \sigma(z^{(l)}) \quad (2.24)$$

Gängige Aktivierungsfunktionen:

- **ReLU:**  $\sigma(x) = \max(0, x)$
- **Sigmoid:**  $\sigma(x) = \frac{1}{1+e^{-x}}$
- **Tanh:**  $\sigma(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$

Der endgültige Output nach  $L$  Schichten:

$$\hat{y} = a^{(L)} \quad (2.25)$$

Nach der Berechnung des Outputs erfolgt dessen Fehleranalyse.

### Fehlermessung (Loss-Funktion)

Die Loss-Funktion misst den Unterschied zwischen Vorhersage und wahrem Wert. Hier ein Paar Ansätze für die Loss-Funktionen:

#### Mittlerer quadratischer Fehler (MSE)

$$\mathcal{L}(\theta) = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (2.26)$$

### Backpropagation (Gradientenberechnung)

Backpropagation nutzt die Kettenregel der Ableitung, um die Änderung des Fehlers bezüglich der Gewichte zu berechnen.

$$\frac{\partial \mathcal{L}}{\partial a^{(L)}} = -\frac{2}{N} (y - \hat{y}) \quad (2.27)$$

Für das letzte Layer gilt:

$$\frac{\partial \mathcal{L}}{\partial W^{(L)}} = \delta^{(L)} a^{(L-1)T} \quad (2.28)$$

mit

$$\delta^{(L)} = \frac{\partial \mathcal{L}}{\partial a^{(L)}} \cdot \sigma'(z^{(L)}) \quad (2.29)$$

Rekursiv für vorherige Schichten:

$$\delta^{(l)} = (W^{(l+1)})^T \delta^{(l+1)} \cdot \sigma'(z^{(l)}) \quad (2.30)$$

**Gradientenabstieg (Gewichtsaktualisierung)**

$$W^{(l)} = W^{(l)} - \eta \frac{\partial \mathcal{L}}{\partial W^{(l)}} \quad (2.31)$$

mit der Lernrate  $\eta$ .

**2.5.5. Optimierte Gradienten-Methoden****Stochastischer Gradientenabstieg (SGD)**

$$W = W - \eta \nabla \mathcal{L}(W, x_i) \quad (2.32)$$

**Adam Optimizer**

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \quad (2.33)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \quad (2.34)$$

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t}, \quad \hat{v}_t = \frac{v_t}{1 - \beta_2^t} \quad (2.35)$$

$$W = W - \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon} \quad (2.36)$$

**2.6. Einführung in Physics-Informed Neural Networks****2.6.1. Motivation für PINNs**

In vielen physikalischen Problemen, wie der Wärmeleitungsgleichung, sind die zugrunde liegenden physikalischen Gesetze bekannt und können als Randbedingungen und PDEs formuliert werden. In vielen praktischen Anwendungen ist jedoch das inverse Problem von Interesse: Man möchte aus Messdaten, die an den Randbedingungen oder an einem Teil des Systems erfasst werden, die vollständige Lösung der PDE rekonstruieren.

**2.6.2. Theorie für das inverse Problem der Wärmeleitung****Das zugrunde liegende Problem**

Die modifizierte Wärmeleitungsgleichung lautet:

$$\frac{\partial u}{\partial t} - \nabla \cdot (a(u) \nabla u) = f(x, y, t),$$

wobei:

- $u(x, y, t)$  die Temperaturverteilung ist,
- $a(u)$  eine unbekannte temperaturabhängige Wärmeleitfähigkeit darstellt,
- $f(x, y, t)$  eine unbekannte Wärmequelle ist.

Das inverse Problem besteht darin, die Funktionen  $a(u)$  und  $f(x, y, t)$  basierend auf Messungen der Temperatur  $u_{\text{obs}}(x, y, t)$  zu rekonstruieren.

### 2.6.3. Formulierung als Variationsproblem

Für das inverse Problem wird das Modell in ein Optimierungsproblem überführt. Die Zielfunktion besteht aus einem Datenanpassungsterm und physikalischen Residuen:

$$\mathcal{J}(u, a, f) = \mathcal{J}_{\text{Data}}(u) + \mathcal{J}_{\text{PDE}}(u, a, f),$$

wobei:

- **Datenanpassung:**

$$\mathcal{J}_{\text{Data}}(u) = \frac{1}{2} \int_0^T \int_{\Omega} (u(x, y, t) - u_{\text{obs}}(x, y, t))^2 \, d\Omega \, dt,$$

- **PDE-Residuum:**

$$\mathcal{J}_{\text{PDE}}(u, a, f) = \frac{1}{2} \int_0^T \int_{\Omega} \left( \frac{\partial u}{\partial t} - \nabla \cdot (a(u) \nabla u) - f(x, y, t) \right)^2 \, d\Omega \, dt.$$

Zur Stabilisierung wird eine Regularisierung hinzugefügt:

$$\mathcal{J}_{\text{Reg}}(a, f) = \alpha \|a(u)\|_{H^1}^2 + \beta \|f(x, y, t)\|_{L^2}^2,$$

wobei  $\alpha, \beta > 0$  Regularisierungsparameter sind.

Die kombinierte Zielfunktion lautet:

$$\mathcal{J}(u, a, f) = \mathcal{J}_{\text{Data}}(u) + \mathcal{J}_{\text{PDE}}(u, a, f) + \mathcal{J}_{\text{Reg}}(a, f).$$

### 2.6.4. Existenz und Eindeutigkeit

**Satz 1 (Existenz):** Unter der Annahme, dass die Beobachtungsdaten  $u_{\text{obs}}$  ausreichend regelmäßig sind ( $u_{\text{obs}} \in H^1(\Omega \times (0, T))$ ), existiert eine Lösung  $(u, a, f)$  des Optimierungsproblems.

*Beweis (Skizze):*

1. Zeige, dass die Zielfunktion  $\mathcal{J}$  nach unten beschränkt ist.
2. Zeige, dass die Menge der zulässigen Parameter  $(u, a, f)$  schwach kompakt ist.
3. Wende den Satz von Weierstraß auf das minimierende Funktional an.

**Bemerkung:** Die Eindeutigkeit ist im Allgemeinen nicht garantiert, da das Problem schlecht gestellt ist (mehrere Kombinationen von  $a(u)$  und  $f(x, y, t)$  können dieselben Messdaten erklären). Regularisierungen wie  $H^1$ - und  $L^2$ -Normen tragen dazu bei, die Lösung zu stabilisieren.

### 2.6.5. PINNs für das inverse Problem

Ein Physics-Informed Neural Network (PINN) wird verwendet, um sowohl die Temperatur  $u(x, y, t)$  als auch die unbekannten Funktionen  $a(u)$  und  $f(x, y, t)$  zu approximieren. Die neuronalen Netze sind wie folgt definiert:

- $u(x, y, t; \theta_u)$ : Neuronales Netz zur Approximation der Temperatur,
- $a(u; \theta_a)$ : Neuronales Netz für die Wärmeleitfähigkeit,
- $f(x, y, t; \theta_f)$ : Neuronales Netz für die Wärmequelle.

Die Verlustfunktion des PINN-Modells kombiniert mehrere Terme:

$$\mathcal{L} = \mathcal{L}_{\text{Data}} + (\mathcal{L}_{\text{PDE}} + \mathcal{L}_{\text{BC}} + \mathcal{L}_{\text{IC}}) + \mathcal{L}_{\text{Reg}}.$$

- **Datenanpassung:**

$$\mathcal{L}_{\text{Data}} = \frac{1}{N_{\text{obs}}} \sum_{i=1}^{N_{\text{obs}}} \left( u(x_i, y_i, t_i; \theta_u) - u_{\text{obs}}(x_i, y_i, t_i) \right)^2,$$

- **PDE-Residuum:**

$$\mathcal{L}_{\text{PDE}} = \frac{1}{N_{\text{PDE}}} \sum_{i=1}^{N_{\text{PDE}}} \left( \frac{\partial u}{\partial t} - \nabla \cdot (a(u; \theta_a) \nabla u) - f(x, y, t; \theta_f) \right)^2,$$

- **Regularisierung:**

$$\mathcal{L}_{\text{Reg}} = \alpha \|\nabla_{\theta} a(u; \theta_a)\|^2 + \beta \|\nabla_{\theta} f(x, y, t; \theta_f)\|^2.$$

Das PINN wird durch Minimierung der Verlustfunktion trainiert:

$$\theta_u^*, \theta_a^*, \theta_f^* = \arg \min_{\theta_u, \theta_a, \theta_f} \mathcal{L}.$$

### 2.6.6. Eigenschaften und Herausforderungen

- **Kopplung von  $a(u)$  und  $f(x, y, t)$ :** Die Rekonstruktion von  $a(u)$  ist nichtlinear gekoppelt mit der Lösung  $u(x, y, t)$ , da  $a$  explizit von  $u$  abhängt. Diese Nichtlinearität führt zu Herausforderungen in der Konvergenz der Optimierung.
- **Ill-posedness:** Das inverse Problem bleibt empfindlich gegenüber Rauschen in den Beobachtungsdaten. Regularisierungen und sorgfältig gewählte Netzarchitekturen sind entscheidend.
- **Einschränkungen:** Die Lösungsgüte hängt stark von der Verfügbarkeit hochauflösender Messdaten  $u_{\text{obs}}$  ab, insbesondere für stark heterogene Materialien oder schnell variierende Wärmequellen.

## **3. Auswertung der praktischen Arbeit**

**3.1. Architektur der Implementierung**

**3.2. Diskussion der Ergebnisse**

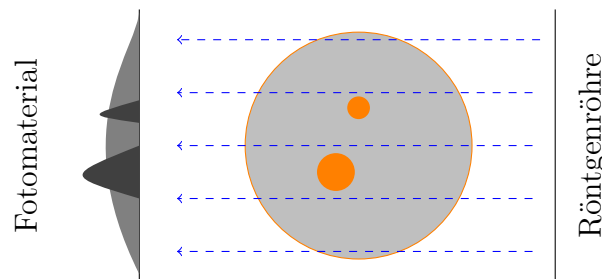
**3.3. Ausblick**

# Vorlage der BA

## Das des Computertomographen

Das Wort Tomographie setzt sich aus zwei dem Altgriechischen entstammten Wörtern zusammen.  $\tau\omicron\mu\eta$  [*tome*], bedeutet Schnitt und  $\gamma\rho\alpha\varphi\epsilon\iota\nu$  [*graphien*] bezeichnet Schreiben. Zusammengefasst zu einem Wort *Tomographie* kann es als Schnittbild eines Objekts verstanden werden.

Die CT ist ein bildgebendes Verfahren, das in der medizinischen Radiologie und in der Materialforschung eingesetzt wird. Die Technologie des Verfahrens erlaubt eine zerstörungsfreie Materialuntersuchung. Das ausschlaggebende Wort ist hier *zerstörungsfrei*, damit ist die Untersuchung des inneren Aufbaus eines Objekts, ohne es zu öffnen, gemeint.



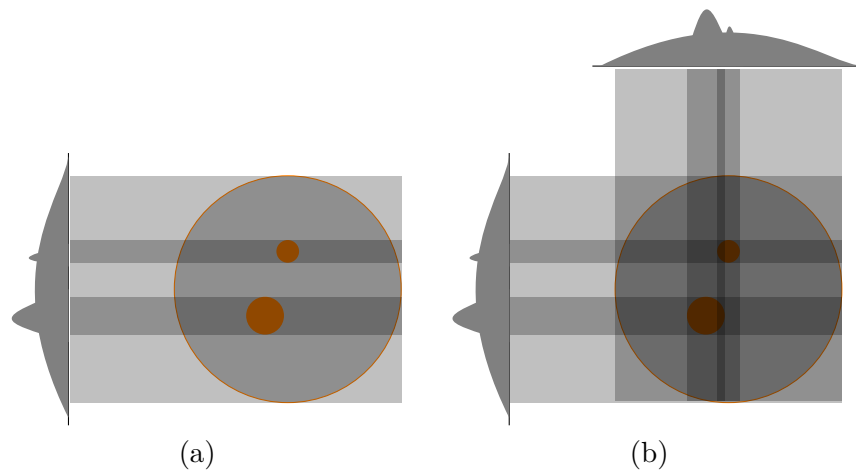
**Abbildung 3.1.:** Schematischer Aufbau einer Röntgenaufnahme im Draufsicht. Links ist das Schwächungsprofil nach Durchgang der Röntgenstrahlen durch ein Medium mit verschiedenen Dichten zu sehen (mit der Annahme, dass orangene Kreise höhere Dichte haben, als die graue Fläche).

Sicherlich wäre die Computertomographie ohne die Röntgenstrahlung nicht denkbar. Diese verdankt ihren Namen dem deutschen Physiker Wilhelm Conrad Röntgen (1845-1923). Im Jahre 1895 gelang es ihm zum ersten Mal eine elektromagnetische Strahlung zu erzeugen, deren Wellenlänge zwischen  $10\text{nm}$  und  $1\text{pm}$  lag. Solche kurze Wellenlängen haben nur energiereiche Strahlen. Das erlaubt ihnen die Durchdringung der Materie. Anzumerken ist, dass die Intensität hochenergetischer Strahlen beim Durchgehen der Materie mit der Eindringtiefe abfällt. Dies bedeutet, dass die Röntgenstrahlung beim Auftreffen auf ein Probestück eine höhere Intensität besitzt als beim Austreten. Bei den herkömmlichen Röntgenaufnahmen wurde diese Eigenschaft direkt ausgenutzt und man hat überlagerte Schatten von Objekten mit verschiedenen Dichten auf das Aufnahmematerial projiziert bekommen. Damit erschienen die dichteren Stellen des Objekts dunkler, weiche entsprechend heller. Für gewisse Situationen geben solche Projektionen genug Information her, jedoch nicht, wenn die räumliche Anordnung des Inneren eines Objekts diskutiert werden soll



(Abb.??).

Um das Problem der räumlichen Anordnung der Objekte zu lösen, ist folgendes Vorgehen sehr nützlich. Würde man das Schattenbild aus der Abbildung ?? zurück in die Schnittebene des Objekts projizieren, so wie es in der Abbildung 3.2 (a) gezeigt ist, erhält man etwas Information über die gegebene Objekte. Projiziert man mehrere Schattenbilder, die aus verschiedenen Winkeln aufgenommen worden sind, so erhält man als Summe der Rückprojektionen, die sogenannte Rückprojektion (Abb. 3.2(b)). Später werden wir sehen, welcher mathematischer Aufwand sich hinter dieser Idee verbirgt, bevor ein annehmbares Resultat zustande kommt.



**Abbildung 3.2.:** Das Prinzip der Bildrekonstruktion aus CT-Daten mit einer Rückprojektion (a) und zwei Rückprojektionen (b).

Ein Pionier der Tomographie, A. M. Cormack<sup>3</sup>, veröffentlichte 1964 eine Arbeit über die mathematische Bestimmung der Dichten eines Objekts. Aus den Projektionsdaten konnte er die innere Dichteverteilung des Objekts bestimmen.

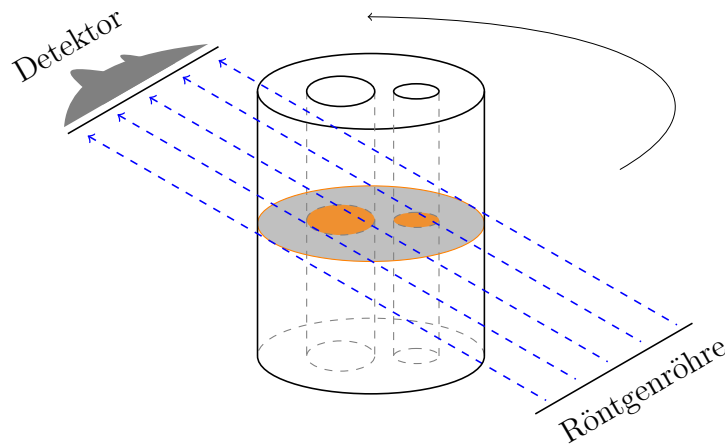
Der erste Prototyp eines CT-Scanners wurde vom G. N. Hounsfield<sup>4</sup> im Jahr 1968 entwickelt. Seine Arbeit beruhte auf den Erkenntnissen von M. Cormack. Die Idee, der um das Objekt rotierenden Scanneinheit (Abb. 3.3) findet man hier wieder. Vorerst konnten nur anatomische Präparate vermessen werden. Aber schon im Jahre 1972 entstehen erste klinische Untersuchungen mit einem Schädelscanner durch Hounsfield und Ambrose. 1973 werden Berichte über die Weichteildiagnostik des Gehirns mithilfe eines CT-Scanners veröffentlicht. Hinsichtlich der Nützlichkeit des CT-Verfahrens in der Medizin, wurden A. M. Cormack und G. N. Hounsfield im Jahre 1979 mit dem Nobelpreis für Medizin ausgezeichnet.

In darauf folgenden Jahren entstand eine ganze Reihe von Generationen des CT-Scanners. Die Hauptunterschiede bestehen in der Komposition der Scanneinheit und dessen Technik, siehe auch [0, Kap. 3]. Heutzutage hat sich die dritte Generation am meisten durchgesetzt. Der prinzipielle Grundaufbau (Abb. 3.3) dieses Scanners besteht darin, dass die Röntgenröhre und Detektor fest auf einem Ringkörper, einander gegenüber installiert sind. In der Mitte des Rings ist ein fest installierter Probetisch, sodass die Scanneinheit um diesen rotieren kann. Mittels dieser Konstruktion ist es

<sup>3</sup> Allan M. Cormack (1924-1998) südafrikanischer Physiker.

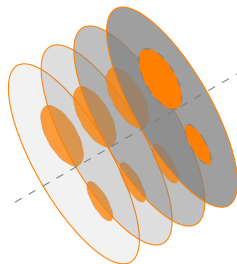
<sup>4</sup> Godfrey N. Hounsfield (1919 - 2004) englischer Ingenieur.

möglich *tomographische* Bilddaten aus allen Positionen um die Längsachse des Objekts aufzunehmen. Die oben beschriebene Rückprojektion aller tomographischen



**Abbildung 3.3.:** Prinzipieller Aufbau eines CT-Scanners der dritten Generation.

Daten (Projektionen) erzeugt ein Schnittbild. Erzeugt man mehrere Schnittbilder entlang der Längsachse des Objekts, so kann man diese der Reihe nach aufstellen und mithilfe programmierter technischer Werkzeuge ein dreidimensionales Bild dieses Objekts gewinnen (Abb. 3.4) .



**Abbildung 3.4.:** Das Schichtmodell eines Objekts als Grundlage zur dreidimensionalen Rekonstruktion.

Im Folgenden gehen wir mehr auf die mathematisch-physikalischen Gegebenheiten der oben beschriebenen Situation ein.

## Die Radon Transformation

Das Verständnis des mathematischen Problems der Bildrekonstruktion in der Computertomographie verlangt zunächst das Verständnis der Entstehung der tomographischen Bilddaten. Die Entstehung solcher Daten kann mithilfe eines mathematischen Modells nachvollzogen werden. Bekanntlich baut man das mathematische Modell auf physikalischen Annahmen auf. Die Abbildung 3.5 soll in diesem Fall als physikalisches Ausgangsmodell dienen. Wir nehmen an, dass der graue Kreis eine homogene Dichteverteilung besitzt. Unter weiteren Annahme sollen die Röntgenstrahlen sich geradlinig ohne Streuung in Richtung des Detektors ausbreiten.

Sei nun  $f(x, y)$  die gesuchte Dichtefunktion, die auf einem Einheitskreis  $\Omega = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$  definiert ist. Dass  $f$  auf einem Kreis liegen soll, folgern wir aus der physikalischen Situation. Das Objekt wird aus einem bestimmten Radius um den Mittelpunkt der Scanneinheit gescannt, daher möchten wir, dass ein Objekt in die CT-Scanneinheit passt, gegebenenfalls muss die Scanneinheit skaliert werden.

Die blau gestrichelte Linie in Abbildung 3.5 stelle einen Röntgenstrahl dar. Wir möchten das Verhalten der Intensität entlang der Strahllinie an den Punkten  $L_0$  und  $L_0 + \Delta L$  beobachten. Wie schon erwähnt wurde, nimmt die Intensität  $I$  der Röntgenstrahlen mit der Eindringtiefe in das Objekt ab. Das Verhalten ist in der Physik unter dem Namen *Lambert-Beer'sches Gesetz* bekannt und besagt:

*Die Intensitätsabnahme der Strahlung ist proportional zur Intensität. Der Proportionalitätsfaktor hängt von der Art der durchdrungenen Materie ab.*

$$I(L_0 + \Delta L) - I(L_0) = -\alpha I(L_0). \quad (3.1)$$

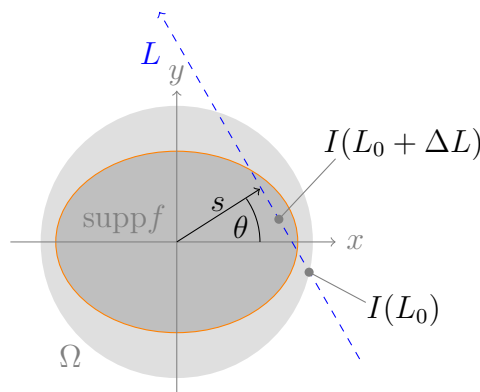
Der Proportionalitätsfaktor  $\alpha$  ist durch das Produkt der Dichteverteilung entlang des Strahls, also  $f(L)$  und seiner zurückgelegten Strecke  $\|\Delta L\|$  gegeben. Hier wird angenommen, dass  $f(L)$  die Eigenschaften des Materials in sich trägt. Insgesamt ergibt sich

$$I(L_0 + \Delta L) - I(L_0) = -f(L) \|\Delta L\| I(L_0). \quad (3.2)$$

Nimmt man an, dass die Gleichung (3.2) asymptotisch im Sinne  $\|\Delta L\| \rightarrow 0$  verstanden werden kann, so kann die linke Seite von (3.2) als Ableitung von  $I$  nach  $L$  interpretiert werden und man schreibt:

$$\lim_{\Delta L \rightarrow 0} \frac{I(L_0 + \Delta L) - I(L_0)}{\|\Delta L\|} = \frac{\partial I}{\partial L} = -f(L)I(L_0). \quad (3.3)$$

Mit 3.3 liegt eine lineare Differentialgleichung (DGL) erster Ordnung vor. Diese DGL kann mithilfe einfacher analytischer Werkzeuge, wie Variablenseparation gelöst werden. Nach der Separation der Variablen in (3.3) und anschließender Integration



**Abbildung 3.5.:** Intensitätsänderung der Röntgenstrahlung beim Passieren eines homogenen Objekts.

bekommt man nun folgenden Ausdruck

$$\ln\left(\frac{I_1}{I_0}\right) = - \int_{L_0}^{L_1} f(L) dL. \quad (3.4)$$

Durch das Exponentieren von (3.4) gelangt man zu dem *Lambert-Beer'schen Gesetz*  $I(L_1) = I_0 e^{-\int_{L_0}^{L_1} f(L) dL}$ . Jedoch für die Zwecke der Computertomographie interessiert man sich für etwas von (3.4) abgewandelte Darstellung

$$\ln\left(\frac{I_0}{I_1}\right) = \int_{L_0}^{L_1} f(L) dL. \quad (3.5)$$

Die rechte Seite in (3.5) liefert einen Summenwert von  $f$  entlang der Geraden  $L$  (Abb. 3.6), die wie folgt parametrisiert werden kann:

$$\begin{aligned} L(s, \theta, t) &= s \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} + t \begin{pmatrix} -\sin(\theta) \\ \cos(\theta) \end{pmatrix}, \quad s, t \in \mathbb{R}; \quad \theta \in [0, \pi]. \\ &= s\omega(\theta) + t\omega^\perp(\theta) \end{aligned} \quad (3.6)$$

Betrachtet man alle Geraden, die den Einheitskreis schneiden, dann ergibt sich folgende Menge

$$G = \{L(s, \theta, t) \mid |s| \leq 1, \theta \in [0, \pi], t \in [-\gamma(s), \gamma(s)]\}. \quad (3.7)$$

Wobei  $\gamma(s)$  den Parameter  $t$  für (3.6) festlegt und folgendermaßen definiert werden kann

$$\gamma(s) = \begin{cases} \sqrt{1-s^2} & : |s| \leq 1 \\ 0 & : \text{sonst} \end{cases}. \quad (3.8)$$

Die Gleichung (3.8) liefert nun die Integrationsgrenzen für (3.5). Da die Werte von  $\pm\gamma(s)$  genau die Schnittpunkte von  $L$  mit  $\Omega$  liefern und  $\text{supp } f^5 \subseteq \Omega$ , heißt es insgesamt, dass nur die Werte von  $L \subset \text{supp } f$  bei der Integration von (3.5) berücksichtigt werden.

Kehren wir zur Gleichung (3.5) zurück und betrachten die Abbildung 3.6, werden wir feststellen, dass die Funktion  $f$  über den Parameter  $t$  integriert werden muss. Hierzu bedarf es einer Variablentransformation:

$$|(J(L))| = \left| \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \right| = 1$$

und

$$(3.9)$$

$$dL = \det(J(L))dt.$$

Mit den obigen Abmachungen haben wir nun alles Nötige in der Hand, um die Radon<sup>6</sup> Transformation angeben zu können.

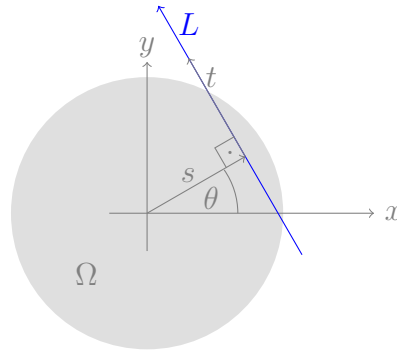
**Definition 5.** Sei nun  $f : \Omega \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$  eine integrierbare Funktion, dann ist ihre Radon Transformation durch

$$\mathcal{R}f(s, \theta) := \int_{-\gamma(s)}^{\gamma(s)} f(s\omega(\theta) + t\omega^\perp(\theta)) dt \quad (3.10)$$

definiert. Hier ist  $\Omega = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$ .

<sup>5</sup> Mit  $\text{supp } f$  (Support) ist der Träger einer Funktion  $f$  gemeint. Def.:  $\text{supp } f := \overline{\{x \in \Omega \mid f(x) \neq 0\}}$

<sup>6</sup> Johann Karl August Radon (1887-1956) österreichischer Mathematiker.



**Abbildung 3.6.:** Parametrisierung des Strahls  $L$  auf dem Gebiet  $\Omega$ .

**Bemerkung 1.**  $\mathcal{R}$  wird als Integraloperator oder einfach als Operator bezeichnet.

Die Gleichung (3.10) trägt das Wesen der Computertomographie in sich und gehört in den Bereich der Integraltransformationen. In der Tat ist man aber an der Inversion von  $\mathcal{R}$  interessiert. Das wird klarer, wenn man (3.10) etwas anders auffasst:

$$p(s, \theta) = \mathcal{R}f(s, \theta). \quad (3.11)$$

Somit stellt  $p(s, \theta)$  in (3.11) genau die Projektion einer unbekannten Dichteverteilung entlang eines Röntgenstrahls  $L(s, \theta)$  dar. Da die Projektionen als Messwerte beim CT-Verfahren entstehen, möchte man  $f$  folgendermaßen wiedergewinnen

$$f(x) = \mathcal{R}^{-1}p(x). \quad (3.12)$$

Jedoch offenbaren sich einige mathematische Hürden bei der Inversion von (3.10), was die Rekonstruktion von  $f$  aus den Projektionsdaten schwierig macht. Um die dahinterstehende Problematik besser verstehen zu können, untersuchen wir zunächst die Eigenschaften von (3.10).

### Kleiner Einschub über $L^2$ -Raum

Um den Operator  $\mathcal{R}$  untersuchen zu können, sollte man wissen auf welchen Räumen er operiert. Dieses Wissen erlaubt uns Aussagen über die Stetigkeit, Kompaktheit und auch über die Singulärwertzerlegung (SWZ) von  $\mathcal{R}$  zu treffen. Besonders wichtig ist die SWZ, denn anhand dieser ergibt sich eine Möglichkeit den Operator  $\mathcal{R}$  zu invertieren. Zudem wollen wir im nächsten Kapitel die Schlechtgestelltheit des computertomographischen Problems diskutieren, welche die Eigenschaften der SWZ benutzt.

In Bezug auf das hier vorgestellte Rekonstruktionsproblem der Computertomographie betrachten wir Funktionen auf dem Gebiet  $\Omega \subset \mathbb{R}^2$ . Wir setzen voraus, dass die Funktionen auf  $\Omega$  integrierbar sind. Wie dieser Sachverhalt zu verstehen ist, wollen wir kurz und grob andeuten.

Zunächst klären wir den Begriff der Integrierbarkeit einer Funktion im Sinne von Lebesgue<sup>7</sup>. Das Lebesgue-Integral wird für Funktionen auf einem geeigneten Maßraum definiert. Ein Maßraum ist durch ein Tripel  $(\Omega, \mathcal{A}, \lambda)$  gegeben.

<sup>7</sup>Henri Léon Lebesgue (1875-1941) französischer Mathematiker.

Eine sogenannte  $\sigma$ -Algebra  $\mathcal{A} \subset \mathcal{P}(\Omega)$  auf  $\Omega$  ist ein Mengensystem bestehend aus Teilmengen von  $\Omega$ .  $\mathcal{A}$  besitzt folgende Eigenschaften: (i). Jede endliche oder abzählbar unendliche Vereinigung oder jeder Durchschnitt der Elemente aus  $\mathcal{A}$  liegen wieder in  $\mathcal{A}$ , sowie  $\Omega$  und  $\emptyset$ . (ii). Die Komplemente der Elemente aus  $\mathcal{A}$ , also  $\forall A \in \mathcal{A} : \Omega \setminus A = A^C \in \mathcal{A}$ , liegen auch in  $\mathcal{A}$ .

Nun suchen wir eine Funktion, die allen Elementen  $A$  aus  $\mathcal{A}$  eine positive Zahl zuordnet, also  $\lambda(A) \in [0, \infty]$ . Somit beschreibt  $\lambda$  das Maß einer Menge in  $\mathcal{A}$ . Im Falle  $\Omega \subset \mathbb{R}^2$  ordnet  $\lambda$  den Flächeninhalt einer Teilmenge von  $\Omega$  zu. Des Weiteren hat  $\lambda$  folgende Eigenschaften: (i):  $\lambda(\emptyset) = 0$  und (ii): die  $\sigma$ -Additivität, das heißt für paarweise disjunkte Mengen  $(A_n)_{n \in \mathbb{N}}$  aus  $\mathcal{A}$  gilt

$$\lambda\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \lambda(A_n).$$

Nun gilt per Definition *messbarer Abbildungen* zwischen zwei Maßräumen, dass die Urbilder messbarer Mengen messbar sind. Das heißt  $f : (X_1, \mathcal{A}_1) \rightarrow (X_2, \mathcal{A}_2)$  ist messbar, wenn für alle  $A_2 \in \mathcal{A}_2$  gilt  $f^{-1}(A_2) \in \mathcal{A}_1$ .

Um den Begriff der Integrierbarkeit im Sinne von Lebesgue einführen zu können, fehlt uns noch der Begriff der Borel<sup>8</sup>- $\sigma$ -Algebra. Dazu benutzen wir für uns relevante Mengen  $\mathbb{R}$  und  $\mathbb{R}^2$ . Das ist insofern klar, weil die gesuchten Dichtefunktionen in einem Punkt  $x \in \Omega \subset \mathbb{R}^2$  einen skalaren Wert  $f(x) \in \mathbb{R}$  liefern. Noch legen wir fest, dass die Werte der Funktionen auf  $\Omega$  immer positiv sind, also  $f(x) \in \mathbb{R}^+$ , denn es gibt keine negativen Dichten. Betrachten wir den topologischen Raum  $(\mathbb{R}, \mathcal{T})$ , hier stellt  $\mathcal{T}$ <sup>9</sup> ein System von offenen Mengen von  $\mathbb{R}$  dar, dann heißt  $\mathcal{B}(\mathbb{R})$  die von  $\mathcal{T}$  erzeugte  $\sigma$ -Algebra auf  $\mathbb{R}$ , die Borel- $\sigma$ -Algebra.  $\mathcal{B}(\mathbb{R})$  ist die kleinste Algebra, die  $\mathcal{T}$  enthält. Das bringt uns ein Vorteil bezüglich der Messbarkeit, denn somit erhalten wir die größtmögliche Menge messbarer Mengen von  $\mathbb{R}$ . Für  $\mathbb{R}^2$  gilt das Obere auch.

Im letzten Schritt definieren wir die sogenannten positiven Treppenfunktionen  $h^+ : (\Omega, \mathcal{B}(\Omega)) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ , die endlich viele Werte  $c_j \in \mathbb{R}_0^+$ ,  $j \in [1, n]$ ,  $n \in \mathbb{N}$  auf  $\Omega$  annehmen. Auf diese Weise ist es möglich das Integral über  $h^+$  folgendermaßen anzugeben:

$$\mathcal{I}(h^+) = \sum_{j=1}^n c_j \lambda(A_j), \quad \text{für } A_j \in \mathcal{B}(\Omega).$$

Jetzt können wir das Lebesgue-Integral einer messbaren Funktion  $f : \Omega \rightarrow \mathbb{R}^+$  definieren, indem wir die Funktion  $f$  ‘von unten’ mit Treppenfunktionen approximieren:

**Definition 6.** Sei  $f : \Omega \rightarrow \mathbb{R}^+$  messbar, dann ist sie integrierbar im Sinne von

<sup>8</sup>Émile Borel (1871-1956) französischer Mathematiker.

<sup>9</sup>Sei  $X$  eine Menge, dann ist  $\mathcal{T} \subset 2^X$  eine Topologie auf  $X$ , falls folgendes gilt:

- (i)  $\emptyset, X \in \mathcal{T}$ .
- (ii)  $A, B \in \mathcal{T} \Rightarrow A \cap B \in \mathcal{T}$ .
- (iii)  $\{A_i\}_{i \in I} \in \mathcal{T} \Rightarrow \bigcup_{i \in I} A_i \in \mathcal{T}$

So nennt man das Paar  $(X, \mathcal{T})$  ein topologischer Raum. Alle Elemente von  $\mathcal{T}$  sind offene Mengen.

Lebesgue, wenn gilt

$$\int_{\Omega} f(x) = \sup \{ \mathcal{I}(h^+) \mid \forall x \in \Omega : h^+(x) \leq f(x) \} < \infty.$$

**Bemerkung 2.** Im Falle einer Lebesgue-Integrierbarkeit spricht man meistens davon, dass eine Funktion integrierbar ist. Streng genommen müsste man das Lebesgue-Integral für beliebige Funktionen  $f : \Omega \rightarrow \mathbb{R}$  mit einer geeigneten Fortsetzung einführen, um dem Begriff der Integrierbarkeit gerecht zu werden. Aber die Definition 6 reicht uns, um das Prinzip der Integration verstanden zu haben.

Nun sind wir im Stande den Raum der quadratintegrierbarer Funktionen folgendermaßen zu definieren:

$$\mathcal{L}^2(\Omega, \mathcal{A}, \lambda) = \{ f : \Omega \rightarrow \mathbb{R} \mid f \text{ ist integrierbar und } \int_{\Omega} |f(x)|^2 dx \leq \infty \}. \quad (3.13)$$

Versehen wir den Raum  $\mathcal{L}^2(\Omega, \mathcal{A}, \lambda)$  mit einer Norm

$$\|f\|_2 := \langle f, f \rangle^{\frac{1}{2}} = \left( \int_{\Omega} |f(x)|^2 dx \right)^{\frac{1}{2}}, \quad x \in \Omega, \quad (3.14)$$

so erhalten wir den Raum  $L^2(\Omega, \mathcal{A}, \lambda)$  (einfach  $L^2(\Omega)$ ), also den *Hilbertraum der quadratintegrierbaren Funktionen*.

Im Allgemeinen ist  $L^2(\Omega)$  ein unendlichdimensionaler Vektorraum über dem Körper der komplexen Zahlen, der bezüglich der durch das Skalarprodukt induzierten Norm vollständig ist.

Kehren wir zurück, zu den gesuchten Funktionen auf einer Kreisscheibe  $\Omega$ , dann sei das Urbild des Operators  $\mathcal{R}$  genau die Menge  $L^2(\Omega)$ . Das Bild von  $\mathcal{R}$  liegt jedoch auf einem halben Zylinder  $Z := [0, \pi] \times [-1, 1]$ . Der Grund hierfür ist ein technischer, denn bei Projektionen über den ganzen Durchmesser einer Kreisfläche auf einem Intervall von 0 bis  $2\pi$ , entstehen redundante Daten, also reichen die Projektionen von 0 bis  $\pi$ .

Des Weiteren wissen wir an dieser Stelle noch nicht wohin  $\mathcal{R}$  den Raum  $L^2(\Omega)$  abbildet. Aber wie oben schon erwähnt wurde, wollen wir später die SWZ von  $\mathcal{R}$  angeben, dafür muss ein *adjungierter* Operator  $\mathcal{R}^*$  zu  $\mathcal{R}$  existieren<sup>10</sup>. Um das zu bewerkstelligen, muss man fordern, dass der Bildraum sowie der Kern von  $\mathcal{R}$  Orthogonalbasen besitzen sollen, was auch einen Skalarprodukt impliziert. Hieraus kann eine Vermutung aufgestellt werden, dass  $\mathcal{R} : L^2(\Omega) \rightarrow L^2(Z)$  abbilden soll, weil eben  $L^2$  Räume der Forderung gerecht sind. Somit erhalten wir insgesamt

$$\mathcal{R} : L^2(\Omega) \rightarrow L^2(Z). \quad (3.15)$$

Mit (3.15) kann man die Eigenschaften des Operators  $\mathcal{R}$  gut untersuchen.

<sup>10</sup> Der Grund weshalb  $\mathcal{R}^*$  existieren muss ist folgender: Es kann nicht vorausgesetzt werden, dass  $\mathcal{R}$  selbstadjungiert ist, wonach der *Spektralsatz für selbstadjungierte Operatoren* für  $\mathcal{R}$  nicht gelten würde. Aber für  $\mathcal{R}\mathcal{R}^*$  funktioniert der Satz schon, da  $\mathcal{R}\mathcal{R}^* = (\mathcal{R}\mathcal{R}^*)^* = \mathcal{R}^*\mathcal{R}$  gelte.

## Eigenschaften der Radon Transformation

In diesem Kapitel sollen einige Eigenschaften des Operators  $\mathcal{R}$  erarbeitet werden. Anhand der Eigenschaften von  $\mathcal{R}$  kann die Problematik der Gleichung (3.12) besser verstanden werden. Zunächst zeigen wir die Linearität von  $\mathcal{R}$ .

### Linearität von $\mathcal{R}$

$\mathcal{R}$  ist linear, denn für  $f, g \in L^2(\Omega)$  beliebig, ergibt sich

$$\begin{aligned} \mathcal{R}(\alpha f + \beta g) &\stackrel{\text{Def.5}}{=} \int_{-\gamma(s)}^{\gamma(s)} (\alpha f + \beta g) dt \\ &\stackrel{\text{Lin.}}{=} \alpha \int_{-\gamma(s)}^{\gamma(s)} f dt + \beta \int_{-\gamma(s)}^{\gamma(s)} g dt \\ &= \alpha \mathcal{R}f + \beta \mathcal{R}g. \end{aligned} \tag{3.16}$$

Der Übersichtlichkeit halber wurden in (3.16) die Argumente weggelassen, also für alle  $f \in \Omega$  gilt  $f = f(s\omega(\theta) + t\omega^\perp(\theta))$ .

### Stetigkeit von $\mathcal{R}$

Mit dem obigen Ergebnis kann man die Theorie der linearen Operatoren in normierten Räumen auf  $\mathcal{R}$  anwenden. Nun wollen wir die Stetigkeit von  $\mathcal{R}$  überprüfen.

**Lemma 1.** *Der Operator  $R : L^2(\Omega) \rightarrow L^2(Z)$  ist stetig.*

*Beweis.* Das Ziel ist zu zeigen, dass das Bild von  $\mathcal{R}$  beschränkt ist. Demnach ist jeder linearer beschränkter Operator stetig (s.a. [0, S.82]). Wir führen den Beweis in zwei Schritten durch. Zuerst geben wir eine Abschätzung für  $|\mathcal{R}f(s, \theta)|^2$  und dann für  $\|\mathcal{R}\|_2$ .

- i. Sei nun  $f \in L^2(\Omega)$  beliebig, dann gilt mit Benutzung der Cauchy-Schwarzschen Ungleichung für Integrale

$$\begin{aligned} |\mathcal{R}f(s, \theta)|^2 &= \left| \int_{-\gamma(s)}^{\gamma(s)} f(s\omega(\theta) + t\omega^\perp(\theta)) dt \right|^2 \\ &\leq \int_{-\gamma(s)}^{\gamma(s)} 1 dt \int_{-\gamma(s)}^{\gamma(s)} |f(s\omega(\theta) + t\omega^\perp(\theta))|^2 dt \\ &= 2\gamma(s) \int_{-\gamma(s)}^{\gamma(s)} |f(s\omega(\theta) + t\omega^\perp(\theta))|^2 dt. \end{aligned} \tag{3.17}$$



Aus der Gleichung (3.8) wissen wir, dass  $\gamma(s) \leq 1$  ist, damit können wir (3.18) weiter abschätzen

$$\begin{aligned} |\mathcal{R}f(s, \theta)|^2 &\leq 2\gamma(s) \int_{-\gamma(s)}^{\gamma(s)} |f(s\omega(\theta) + t\omega^\perp(\theta))|^2 dt \\ &\stackrel{(3.8)}{\leq} 2 \int_{-\gamma(s)}^{\gamma(s)} |f(s\omega(\theta) + t\omega^\perp(\theta))|^2 dt. \end{aligned} \quad (3.18)$$

ii. Mit (3.18) lässt sich das Skalarprodukt  $\|\mathcal{R}f(s, \theta)\|_{L^2(Z)}^2$  abschätzen:

$$\begin{aligned} \|\mathcal{R}f(s, \theta)\|_{L^2(Z)}^2 &= \int_0^\pi \int_{-1}^1 |\mathcal{R}f(s, \theta)|^2 ds d\theta \\ &\stackrel{(3.18)}{\leq} 2 \int_0^\pi \int_{-1}^1 \left( \int_{-\gamma(s)}^{\gamma(s)} |f(s\omega(\theta) + t\omega^\perp(\theta))|^2 dt \right) ds d\theta \\ &= 2 \int_0^\pi \left( \int_{-1}^1 \int_{-\gamma(s)}^{\gamma(s)} |f(s\omega(\theta) + t\omega^\perp(\theta))|^2 dt ds \right) d\theta \\ &= 2 \int_0^\pi \left( \int_\Omega |f(x)|^2 dx \right) d\theta \\ &= 2\pi \|f\|_{L^2(\Omega)}^2. \end{aligned} \quad (3.19)$$

In der vorletzten Zeile von (3.19) wurde  $s\omega(\theta) + t\omega^\perp(\theta) = x$  für ein festes  $\theta$  gesetzt. Damit wurde der Integrationsbereich über  $t$  und  $s$  zu einem über  $x$  gefasst, um insgesamt über ganz  $\Omega$  integrieren zu können.

Nach dem Wurzelziehen aus (3.19), kommt man zur Erkenntnis, dass  $\mathcal{R}$  beschränkt ist, nämlich  $\|\mathcal{R}\|_2 \leq \sqrt{2\pi}$ . Da  $\mathcal{R}$  linear und beschränkt ist, dürfen wir auf seine Stetigkeit schließen.  $\square$

Ist  $\mathcal{R}$  linear und stetig so schreiben wir  $\mathcal{R} \in \mathcal{L}(L^2(\Omega), L^2(Z))^{11}$ .

### Adjungierter Operator $\mathcal{R}$

Da  $\mathcal{R}$  stetig von  $L^2(\Omega)$  nach  $L^2(Z)$  abbildet, existiert zu  $\mathcal{R}$  ein stetiger adjungierter Operator  $\mathcal{R}^*$ , d.h.

$$\forall f \in L^2(\Omega) \wedge g \in L^2(Z) : \langle \mathcal{R}f, g \rangle_{L^2(Z)} = \langle f, \mathcal{R}^*g \rangle_{L^2(\Omega)}. \quad (3.20)$$

**Lemma 2.** Der adjungierte Operator  $\mathcal{R}^*$  zu  $\mathcal{R}$  ist gegeben durch

$$\begin{aligned} \mathcal{R}^* : L^2(Z) &\rightarrow L^2(\Omega) \\ \mathcal{R}^*g(x) &= \int_0^\pi g(x^T \omega(\theta), \theta) d\theta. \end{aligned} \quad (3.21)$$

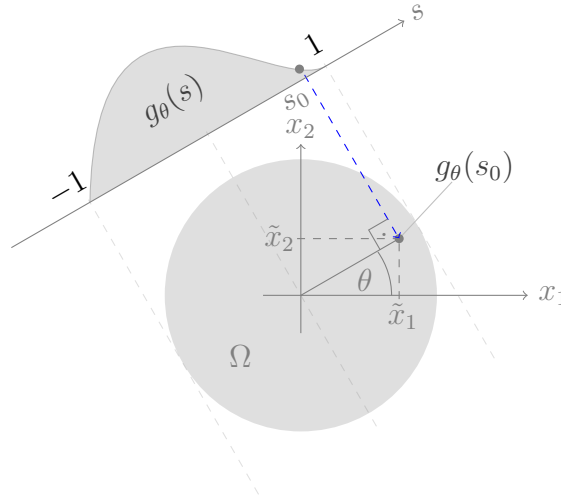
<sup>11</sup>  $\mathcal{L}(X, Y)$  : Raum der stetigen linearen Abbildungen zwischen normierten Räumen  $X$  und  $Y$ .

**Bemerkung 3.** Bevor wir (3.20) für (3.21) zeigen, wollen wir die letztere Gleichung etwas besser in unsere Vorstellung einprägen. Sei dazu die Abb.3.7 betrachtet. Wir wählen einen beliebigen Punkt  $x = (\tilde{x}_1, \tilde{x}_2)^T \in \Omega$  und einen Wert  $s \in [-1, 1]$ . So existiert folgende Beziehung zwischen  $x$  und  $s$

$$x^T \omega(\theta) = (s \cos(\theta), s \sin(\theta)) \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} = s(\cos^2(\theta) + \sin^2(\theta)) = s. \quad (3.22)$$

Wobei wir gemäß der Abb.3.7 für  $x = (s \cos(\theta), s \sin(\theta))^T$  schreiben dürfen.

(3.21) stellt eine Ungefilterte Rückprojektion dar. Man kann sich das folgendermaßen vorstellen. Für ein festes  $x$  in der Ebene werden entsprechende Projektionen an der Stelle  $s$  und Winkel  $\theta$  zurück auf den Punkt  $x$  projiziert. Nach der Addition der Projektionen auf  $x$  über alle Winkel, bekommt man den Wert der ungefilterten Rückprojektion an dem Punkt  $x$ .



**Abbildung 3.7.:** Ein Beispiel der Rückprojektion eines Projektionswerts  $g_\theta(s_0)$  auf einen Punkt  $x \in \Omega$ . Mit  $g_\theta(s)$  ist hier eine Projektion bei einem festen Winkel  $\theta$  bezeichnet, also  $g_\theta(s) = g(s, \theta)$ .

*Beweis von Lemma 2.* Seien  $f \in L^2(\Omega)$  und  $g \in L^2(Z)$  beliebig, dann gilt

$$\begin{aligned} \langle \mathcal{R}f, g \rangle_{L^2(Z)} &= \int_Z (\mathcal{R}f(s, \theta)) g(x^T \omega(\theta), \theta) \, ds d\theta \\ &\stackrel{(3.10)}{=} \int_0^\pi \int_{-1}^1 \left( \int_{-\gamma(s)}^{\gamma(s)} f(s\omega(\theta) + t\omega^\perp(\theta)) dt \right) g(x^T \omega(\theta), \theta) \, ds d\theta. \end{aligned} \quad (3.23)$$

An dieser Stelle machen wir eine Substitution in der rechten Seite von (3.23). Hier nutzen wir (3.22) aus und schreiben  $x^T \omega(\theta) = \omega^T(\theta)x = s$ . Dann bekommt man

$g(x^T \omega(\theta), \theta) = g(\omega^T(\theta)x, \theta)$  und wir wollen die Funktion  $f$  im Punkt  $x$  betrachten, also  $f(s\omega(\theta) + t\omega^\perp(\theta)) = f(x)$ . Somit erhalten wir einen Integrationsbereich über  $\Omega$ . Nun führen wir (3.24) weiter:

$$\begin{aligned}
\langle \mathcal{R}f, g \rangle_{L^2(Z)} &= \int_0^\pi \int_{-1}^1 \left( \int_{-\gamma(s)}^{\gamma(s)} f(s\omega(\theta) + t\omega^\perp(\theta)) dt \right) g(x^T \omega(\theta), \theta) ds d\theta \\
&\stackrel{\text{Subst.}}{=} \int_0^\pi \left( \int_{-1}^1 \int_{-\gamma(s)}^{\gamma(s)} f(x) g(\omega^T(\theta)x, \theta) dt ds \right) d\theta \\
&= \int_0^\pi \int_\Omega f(x) g(\omega^T(\theta)x, \theta) dx d\theta \\
&= \int_\Omega f(x) \left( \int_0^\pi g(\omega^T(\theta)x, \theta) d\theta \right) dx \\
&= \langle f, \mathcal{R}^* g \rangle_{L^2(\Omega)}.
\end{aligned} \tag{3.24}$$

Somit ist  $\mathcal{R}^*$  adjungierter Operator der Radon Transformation und ist nach Satz von Schauder auch kompakt, wenn  $\mathcal{R}$  kompakt ist. ([0, S. 185]).  $\square$

### Kompaktheit von $\mathcal{R}$

Um die Kompaktheit von  $\mathcal{R}$  zeigen zu können, werden wir an dieser Stelle einen Umweg einschlagen. Dafür ziehen wir ein Ergebnis aus [0, S. 45] heran. Es handelt sich um einen zu  $\mathcal{R}$  adjungierten Operator  $\tilde{\mathcal{R}}^*$  folgender Gestalt:

$$\tilde{\mathcal{R}}^* : L^2(Z, \frac{1}{\gamma(s)}) \rightarrow L^2(\Omega). \tag{3.25}$$

Hier wurde  $\gamma(s)$  im Sinne der Gleichung (3.8) benutzt. Und der Raum  $L^2(Z, \frac{1}{\gamma(s)})$  ist mit dem folgenden Skalarprodukt versehen

$$\langle g, f \rangle = \int_Z g \bar{f} \frac{1}{\gamma(s)} ds d\theta. \tag{3.26}$$

Wir können festhalten dass der Operator  $\mathcal{R}' : L^2(\Omega) \rightarrow L^2(Z, \frac{1}{\gamma(s)})$  kompakt ist<sup>12</sup>. Die Kompaktheit folgt eben aus der SWZ von  $\mathcal{R}'$ . Denn die Singulärwerte von  $\mathcal{R}'$

$$\sigma_j = 2\sqrt{\frac{\pi}{j+1}}, \quad j \in \mathbb{N}$$

bilden eine Nullfolge, was nach einer längeren Rechnung aus [0, S. 43] folgt. Also schreiben wir  $\mathcal{R}' \in \mathcal{K}(L^2(\Omega), L^2(Z, \frac{1}{\gamma(s)}))$ <sup>13</sup>.

<sup>12</sup>Prinzipiell sind  $\mathcal{R}$  und  $\mathcal{R}'$  gleich, es ist derselbe Operator, der aber  $L^2(\Omega)$  in einen anderen Raum abbildet. Die Bezeichnung  $\mathcal{R}'$  führen wir ein, um uns Verwechslungen zu ersparen.

<sup>13</sup> $\mathcal{K}(X, Y) = \{A \in \mathcal{L}(X, Y) \mid A \text{ ist kompakt} \}$

Wir veranschaulichen die entstandene Situation in einem kommutativen Diagramm

$$\begin{array}{ccc}
 L^2(\Omega) & \xrightarrow{\mathcal{R}} & L^2(Z) \\
 \nwarrow \mathcal{R}' & & \nearrow \text{Id} \\
 & L^2(Z, \frac{1}{\gamma(s)}) &
 \end{array}
 \quad (3.27)$$

und stellen fest, dass  $\mathcal{R}$  als eine Komposition zweier Abbildungen darstellbar ist. Hier bedeutet Id die Identität auf  $L^2(Z)$  dar. Der Unterschied zwischen  $L^2(Z)$  und  $L^2(Z, \frac{1}{\gamma(s)})$  besteht nur in der Definition des Skalarprodukts auf diesen beiden Räumen.

Insgesamt würde es heißen, dass wir die Komposition  $\text{Id} \circ \mathcal{R}' = \mathcal{R}$  bezüglich der Kompaktheit beurteilen müssen. Dafür wird uns ein nützlicher Satz aus [0, S. 27] gute Dienste leisten (ohne Beweis) :

**Satz 1.** *Seien  $X, Y$  und  $Z$  normierte Räume. Falls  $A \in \mathcal{L}(X, Y)$  und  $B \in \mathcal{K}(Y, Z)$  oder  $A \in \mathcal{K}(X, Y)$  und  $B \in \mathcal{L}(Y, Z)$ , dann gilt  $BA \in \mathcal{K}(X, Z)$ .*

Somit wollen wir zeigen, dass der Operator  $\text{Id} : L^2(Z, \frac{1}{\gamma(s)}) \rightarrow L^2(Z)$  beschränkt ist, und somit auch stetig ist. Danach könnten wir den Satz 1 benutzen, um die Kompaktheit von  $\mathcal{R}$  zu zeigen.

Sei nun  $g(s, \theta) \in L^2(Z)$  beliebig, dann können wir folgende Abschätzung durchführen:

$$\begin{aligned}
 \| \text{Id } g(s, \theta) \|_{L^2(Z)}^2 &= \int_Z |\mathcal{R}f(s, \theta)|^2 ds d\theta \\
 &\stackrel{(3.26)}{\leq} \int_Z |\mathcal{R}f(s, \theta)|^2 \frac{1}{\gamma(s)} ds d\theta \\
 &= \| g(s, \theta) \|_{L^2(Z, \frac{1}{\gamma(s)})}^2 .
 \end{aligned}
 \quad (3.28)$$

Die letzte Abschätzung beruht auf der Kenntnis, dass  $\gamma(s) \leq 1$  für alle  $s \in (-1, 1)$ , somit aber  $\frac{1}{\gamma(s)} \geq 1$  gilt. Nun haben wir gezeigt, dass  $\| \text{Id} \|_2 \leq 1$  beschränkt und demnach stetig ist. Also gilt  $\text{Id} \in \mathcal{L}(L^2(Z, \frac{1}{\gamma(s)}), L^2(Z))$ . Dieses Ergebnis ist für uns ausreichend um den Satz 1 benutzen zu können, d.h

$$\mathcal{R} \stackrel{(3.27)}{=} \text{Id} \circ \mathcal{R}' \stackrel{\text{Satz 1}}{\implies} \mathcal{R} \in \mathcal{K}(L^2(\Omega), L^2(Z)). \quad (3.29)$$

Nun sind wir am Ende dieses Kapitels angelangt und nach einigen Untersuchungen können wir festhalten, dass der lineare stetige Operator  $\mathcal{R} : L^2(\Omega) \rightarrow L^2(Z)$  kompakt ist und einen kompakten adjungierten Operator  $\mathcal{R}^*$  hat. Diese Ergebnisse spielen eine zentrale Rolle bei der Singulärwertzerlegung (SWZ) eines kompakten Operators und beleuchten das Problem der Invertierung solcher Operatoren sehr gut.

### Die Singulärwertzerlegung (SWZ) von $\mathcal{R}$

Im Folgenden wird die allgemeine Form der SWZ des Operators  $\mathcal{R}$  angegeben, eine konkrete wird in [0, Kap. 2.5] hergeleitet, jedoch aber nur für  $\mathcal{R}'$ . Um zu SWZ

von  $\mathcal{R}$  zu gelangen, ermittelt man zuerst die Eigenwerte und Eigenvektoren von  $\mathcal{R}\mathcal{R}^* : L^2(Z) \rightarrow L^2(Z)$ , die als Komposition kompakter Operatoren kompakt ist. Demnach gilt der *Spektralsatz für selbstadjungierte kompakte Operatoren* [0, S. 30] und man bekommt

$$\mathcal{R}\mathcal{R}^*g = \sum_{j=1}^{\infty} \lambda_j \langle g, v_j \rangle_{L^2(Z)} v_j, \quad \text{für alle } g \in L^2(Z). \quad (3.30)$$

Hier bildet  $\{\lambda_j\}_{j \in \mathbb{N}} \in \mathbb{R}$  eine Folge mit der absteigenden Anordnung der Eigenwerte  $|\lambda_1| \geq |\lambda_2| \geq \dots > 0$  und  $\{v_j\}_{j \in \mathbb{N}} \in L^2(Z)$  ein Orthonormalsystem in  $\text{Kern}(\mathcal{R})^\perp$ . Eine wichtige Tatsache ist, dass die Folge  $\{\lambda_j\}_{j \in \mathbb{N}}$  entweder abbricht oder gegen Null konvergiert.

Die Singulärwerte und das Orthonormalsystem  $\{u_j\}_{j \in \mathbb{N}}$  des Bildes von  $\mathcal{R}$  bekommen wir aus

$$\begin{aligned} (i). \quad \sigma_j &:= \sqrt{\lambda_j} \\ (ii). \quad u_j &:= \sigma_j^{-1} \mathcal{R}v_j, \quad j \in \mathbb{N}. \end{aligned} \quad (3.31)$$

Mit den Ergebnissen aus (3.30) und (3.31) kann nun die SWZ der Radon Transformation angegeben werden

$$\mathcal{R}f = \sum_{j=1}^{\infty} \sigma_j \langle f, v_j \rangle_{L^2(\Omega)} u_j. \quad (3.32)$$

Das Tripel  $\{(\sigma_j, v_j, u_j)\}$  nennt man *Singulärsystem* eines kompakten Operators und die Folgen  $\{u_j\}_{j \in \mathbb{N}}$  und  $\{v_j\}_{j \in \mathbb{N}}$  bilden folgende Orthonormalsysteme

$$\overline{\text{Span}\{u_j | j \in \mathbb{N}\}} = \overline{\text{Bild}(\mathcal{R})} \quad \text{und} \quad \overline{\text{Span}\{v_j | j \in \mathbb{N}\}} = \text{Kern}(\mathcal{R})^\perp. \quad (3.33)$$

Die Gleichung (3.32) wird im nächsten Kapitel bei der Klärung des Begriffs *schlecht gestellter Probleme* sehr hilfreich sein.

# Das inverse Problem

In diesem Abschnitt soll der Bezug der Bildrekonstruktion aus CT-Daten zum Begriff *inverser Probleme* hergestellt und der Frage nachgegangen werden, warum inverse Probleme meist schlecht gestellt seien.

Zunächst können wir festhalten, dass in allen naturwissenschaftlichen Fragestellungen, die Wirkungen von den Ursachen herbei geführt werden. Dies ist eine Kausalität, welche sich in einem mathematischen Modell beschreiben lässt. Anlehnend an [0, S. 14] definieren wir ein mathematisches Modell, wie folgt

**Definition 7.** Ein *mathematisches Modell* ist eine Abbildung

$$A : X \rightarrow Y$$

von der Menge der Ursachen (Parameter)  $X$  in die Menge der Wirkungen (Daten)  $Y$ . Im Falle des **inversen Problems** wird zu einer Wirkung  $y \in Y$  die Ursache  $x \in X$  gesucht, sodass  $Ax = y$  erfüllt ist. Umgekehrt liegt ein **direktes Problem** vor.

Jetzt ist es klar, warum Bildrekonstruktion aus CT-Daten zu den inversen Problemen gehört. Man möchte aus der Menge der Projektionsdaten ( $Y$ ) die unbekannte Dichteverteilung ( $X$ ) des bestrahlten Materials berechnen.

Im nächsten Schritt soll geklärt werden, was ein *schlecht gestelltes* Problem ist. Dazu betrachten wir die Definition aus [0, S. 24]

**Definition 8.** Sei  $A : X \rightarrow Y$  eine Abbildung zwischen Hilberträumen  $X$  und  $Y$ . Das Problem  $(A, X, Y)$  heißt **schlecht gestellt nach Nashed[0]** (*ill-posed*), wenn das Bild von  $A$  nicht abgeschlossen ist. Ansonsten heißt das Problem  $(A, X, Y)$  **gut gestellt nach Nashed**.

Nun ist es an der Zeit, die erarbeiteten Ergebnisse aus dem Kapitel (??) heranzuziehen. Entsprechend der Definition 8. schreiben wir den Ausdruck (3.10) als  $(\mathcal{R}, L^2(\Omega), L^2(Z))$  um. Somit ist ein schlecht gestelltes Problem gegeben, welches durch die SWZ (3.32) des Operators  $\mathcal{R}$

$$\mathcal{R}f = \sum_{j=1}^{\infty} \sigma_j \langle f, v_j \rangle_{L^2(\Omega)} u_j.$$

einzusehen ist. Man erkennt, dass das Bild des Operators  $\mathcal{R}$  nicht abgeschlossen ist.

Wir wollen nun den Operator  $\mathcal{R}$  invertieren. Aus dem *Spektralsatz für selbstadjungierte kompakte Operatoren* [0, S. 30] folgt, dass die Eigenwerte von  $\mathcal{R}\mathcal{R}^*$  der Bedingung

$$\forall j \in \mathbb{N} : \lambda_j > 0 \quad \text{und} \quad \lambda_j \xrightarrow{j \rightarrow \infty} 0$$

genügen. Entsprechend (3.31, (i)) sind die Singulärwerte von  $\mathcal{R}$  alle größer Null. Dies erlaubt die Invertierung der Gleichung (3.32) und wir schreiben

$$\mathcal{R}^+g = \sum_{j=1}^{\infty} \sigma_j^{-1} \langle g, u_j \rangle_{L^2(Z)} v_j \quad \text{für } g \in \text{Bild}(\mathcal{R}). \quad (3.34)$$

Aus der Gleichung (3.34) ist zu erkennen, dass der Ausdruck  $\sigma_j^{-1}$  für  $j \rightarrow \infty$  gegen Unendlich läuft und somit ist  $\mathcal{R}^+$  nicht beschränkt, was also die Nichtstetigkeit bedeutet.

Die Bezeichnung  $\mathcal{R}^+$  nennt man *verallgemeinerte Inverse* des kompakten Operators  $\mathcal{R}$  und wir bestätigen die obige Überlegung mit folgendem Satz, dessen Beweis in [0, S. 32] nachgelesen werden kann.

**Satz 2.** *Sei  $A : X \rightarrow Y$  ein kompakter Operator mit einem Singulärsystem  $\{(\sigma_j, v_j, u_j)\}$ . Dann ist die verallgemeinerte Inverse durch*

$$A^+g = \sum_{j=1}^{\infty} \sigma_j^{-1} \langle g, u_j \rangle_Y v_j \quad \text{für } g \in \mathcal{D}_A^{14}$$

*gegeben. Hat  $A$  ein endlichdimensionales Bild, dann ist  $A^+$  stetig.*

Im Falle des Operators  $\mathcal{R} : L^2(\Omega) \rightarrow L^2(Z)$  ist das Bild *unendlichdimensional*, was auch dem Ausdruck (3.32) anzusehen ist.

Gemäß obigen Überlegungen ist das Problem rein theoretisch. In der Praxis hängt die Radon Transformation nur von endlich vielen Argumenten ab, also von  $s_i$  mit  $i \in [1, k]$ ,  $k \in \mathbb{N}$  und von  $\theta_l$  mit  $l \in [1, q]$ ,  $q \in \mathbb{N}$ . Damit ist gemeint, dass man gerade  $n = kq$  Projektionen erzeugen kann. Somit bekommen wir

$$\mathcal{R}f = \sum_{j=1}^n \sigma_j \langle f, v_j \rangle_{L^2(\Omega)} u_j, \quad n \in \mathbb{N}. \quad (3.35)$$

**Bemerkung 4.** *An dieser Stelle lassen wir die Diskretisierungstheorie unendlichdimensionaler Operatoren bei Seite. Wir verweisen hier auf [0, S. 153]. Zu erwähnen ist, dass  $n \in \mathbb{N}$  in (3.35) nicht die ersten  $n$  Eigenvektoren bezeichnet, sondern, dass die unendliche Dimension von  $\mathcal{R}$  auf  $n$ -Dimensionalität interpoliert wurde.*

**Bemerkung 5.** *Die Elemente  $\mathcal{R}$ ,  $f$ ,  $\sigma_j$ ,  $v_j$  und  $u_j$  in (3.35) sind nicht mit denen aus (3.32) zu verwechseln. In diesem Falle sind es die diskretisierte Elemente.*

Wir bilden zuerst die Inverse  $\mathcal{R}^+$  von  $\mathcal{R}$  und diskutieren das Verhalten einzelner Summenglieder von

$$\mathcal{R}^+g = \sum_{j=1}^n \sigma_j^{-1} \langle g, u_j \rangle_{L^2(Z)} v_j. \quad (3.36)$$

Hier stellt  $g$  eine reale Messung der Projektionen dar, bei der gewisse Messungenauigkeiten anzunehmen sind. Diese Messungenauigkeiten werden mit steigendem  $j$  um dem Faktor  $\sigma_j^{-1}$  in Richtung des Eigenvektors  $u_j$  verstärkt. Anderes ausgedrückt,

<sup>14</sup> $\mathcal{D}_A$  : Definitionsbereich der Abbildung  $A : X \rightarrow Y$ .

wird auf diese Weise rekonstruierte Lösung  $\mathcal{R}^+g$  immer verschwommener, was auf die Oszillation der Eigenvektoren zurückzuführen ist.

Im Allgemeinen bedarf es einer Bedingung an die Koeffizienten  $\langle g, u_j \rangle_{L^2(Z)}$ , die besagt, dass sie schnell genug abfallen müssen, um nicht von  $\sigma_j^{-1}$  ins Unendliche geschickt zu werden. Genau das besagt die *Picard-Bedingung* (s.a. [0, S. 31]).

**Satz 3 (*Picard-Bedingung*).** *Seien  $X, Y$  Hilberträume und  $A : X \rightarrow Y$  ein kompakter Operator mit einem singulärem System  $\{(\sigma_j, v_j, u_j)\}$ . Die Reihe*

$$\sum_{j=1}^{\infty} \frac{|\langle g, u_j \rangle_Y|^2}{\sigma_j^2} \quad (3.37)$$

*konvergiert genau dann, wenn  $g \in \text{Bild}(A)$ .*

Die verrauschten Messdaten liegen nie in dem Bild der Radon Transformation. Deshalb werden die Anteile der Lösung mit kleinen Singulärwerten (oder hochfrequente Anteile) mit speziellen Techniken *gedämpft*, so dass die Reihe (3.37) konvergiert. Die Dämpfung hochfrequenter Anteile nennt man im Allgemeinen Regularisierung eines schlecht gestellten Problems. Die Art der Regularisierung unterliegt dem Rekonstruktionsverfahren. Das aber, wird das Hauptthema des nächsten Kapitels sein.



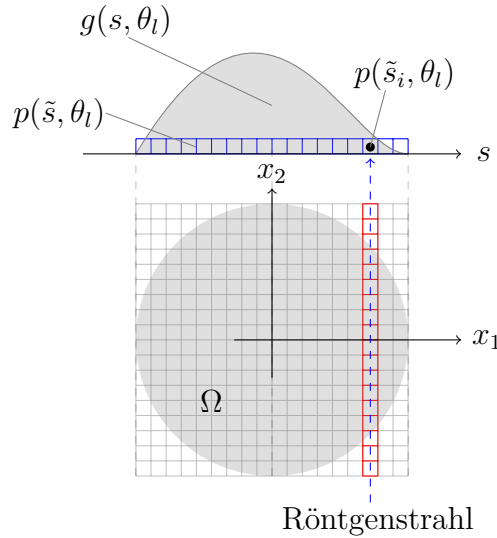
# Rekonstruktionsverfahren und ihre Implementierung

In diesem Kapitel werden wir die praktische Vorgehensweise einiger Bildrekonstruktionstechniken in der Computertomographie diskutieren. Den Stützpunkt folgender Diskussion werden wir zuerst hier festigen. Dazu schauen wir uns nochmal das diskrete Modell (3.35) an:

$$\mathcal{R}f = \sum_{j=1}^n \sigma_j \langle f, v_j \rangle_{L^2(\Omega)} u_j, \quad n \in \mathbb{N}.$$

Die obige Gleichung stellt genau  $n$  Projektionen  $p \in \mathbb{R}^n$  einer unbekannten Dichtefunktion  $f \in \mathbb{R}^m$  dar. Die Anzahl der Projektionen  $n$  setzt sich aus der Anzahl der Detektorzellen  $k \in \mathbb{N}$  im Detektorband und der Anzahl der Winkel  $q \in \mathbb{N}$  zu denen die Projektionen erzeugt worden sind zusammen. Insgesamt ergibt sich

$$n = kq \quad \text{für } n, k, q \in \mathbb{N}. \quad (3.38)$$



**Abbildung 3.8.:** Anschauliche Darstellung der Diskretisierung der Dichtefunktion  $f$ . Entlang des Röntgenstrahls (alle rot umrandeten Pixel) an der Stelle  $s_i$  und zu einem festen Winkel  $\theta_l = 0$ , werden die Werte von  $f$  zu dem Wert  $p(\tilde{s}_i, \theta_l)$  aufsummiert.  $p(\tilde{s}, \theta_l)$  ist der Vektor zu allen Strahlen an den Stellen  $\tilde{s}_i$  und somit die Diskretisierung von  $g(s, \theta_l)$ .

Die Diskretisierung von  $f$  hängt hingegen nur von der Anzahl der Detektorzellen  $k \in \mathbb{N}$  im Detektorband ab. Betrachtet man die Abbildung 3.8, dann wird man erkennen, dass die  $k$  Detektorzellen die Breite und die Höhe des zu rekonstruierenden Bildes  $f$  festlegen, demnach ergibt sich

$$m = k^2 \quad \text{für } k, m \in \mathbb{N}. \quad (3.39)$$

**Bemerkung 6.** *Praktisch gesehen, stellt  $k$  die Anzahl der digitalen Pixel in der Breite und Tiefe eines digitalen Bildes  $f$  dar.*

Bevor wir im nächsten Schritt zur Bildrekonstruktion übergehen, schauen wir uns zuerst die Entstehung der Projektionsdaten aus der Radon Transformation etwas genauer an. Betrachten wir dazu die Gleichung (3.10) und die Abbildung 3.6.

$$\mathcal{R}f(s, \theta) = \int_{-\gamma(s)}^{\gamma(s)} f(s\omega(\theta) + t\omega^\perp(\theta)) dt.$$

Im diskreten Fall  $\tilde{f} \in \mathbb{R}^{k \times k}$  wird man feststellen, dass (3.10) die Summe aller Werte von  $\tilde{f}$  entlang einer Geraden  $G$ , den Projektionswert  $p(\tilde{s}_i, \theta_l)$  an der Stelle  $\tilde{s}_i, i \in [1, k]$ , entsprechend dem Winkel  $\theta_l, l \in [1, q]$ , liefert. Damit kann (3.10) für die Transformation eines digitalen Bildes folgendermaßen interpretiert werden

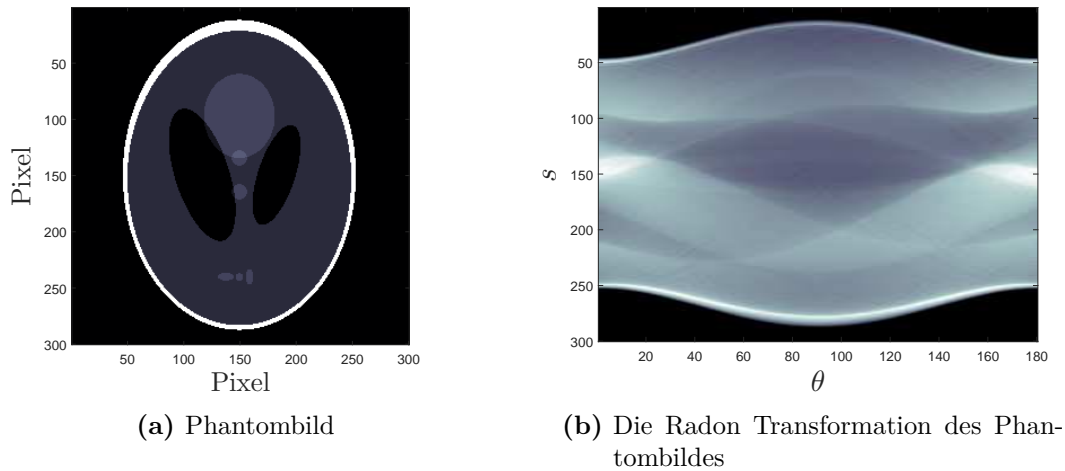
$$p(\tilde{s}_i, \theta_l) = \sum_{j=1}^m \mu_j \Delta \tilde{t}, \quad \text{mit } m = k^2, \quad (3.40)$$

$$\text{für } \mu_j = \begin{cases} \tilde{f}_j & : \text{ falls } \tilde{f}_j \cap G \neq \emptyset \\ 0 & : \text{ sonst} \end{cases}.$$

Fasst man auf diese Weise zu jedem  $\tilde{s}_i$  (wobei  $\Delta \tilde{t} = 1$ ) die entsprechenden Werte von  $\tilde{f}$  zusammen, so entsteht der Projektionsvektor  $p(\tilde{s}, \theta_l)$  (Abb. 3.8). Stellt man alle Projektionen  $p(\tilde{s}, \theta_l)$  entsprechend der Winkelreihenfolge  $\theta_l, l \in [1, q]$  auf, so entsteht ein digitales Bild der Projektionsdaten der diskreten Radon Transformation.

**Bemerkung 7.** *Die Gleichung (3.40) entspricht der Parallelstrahlengeometrie. Das bedeutet, dass alle Geraden  $G$ , die das Objekt durchdringen, parallel zueinander sind. Um (3.40) in die Fächerstrahlengeometrie zu überführen, müssen entsprechende geometrische Transformationen durchgeführt werden.*

Schauen wir uns ein konkretes Beispiel einer Radon Transformation an. Sei dazu die Abbildung 3.9 betrachtet. In Abb. 3.9b ist die Radontransformierte oder *das Sinogramm* von Abb. 3.9a gemäß (3.40) zu sehen. Der Aufbau des Sinogramms entspricht dem oben beschriebenen Prinzip. Fasst man also das Sinogramm in Abb. 3.9b als eine Matrix auf, so ist jede Spalte der Matrix gleich einer Projektion  $g(\tilde{s}, \theta_l)$  zu einem festem Winkel  $\theta_l$ . Die Implementierung der (diskreten) Radon Transformation ist im Abschnitt ?? kurz erläutert.



**Abbildung 3.9.:** Links (a) ist ein Phantombild (Quelle: MATLAB, `phantom(300)`) zu sehen, welches wir für alle Testzwecke in diesem Kapitel benutzen werden. Rechts (b) ist die Radontransformierte von (a) abgebildet. Die Transformation wurde über  $\theta \in [0^\circ, 180^\circ]$  mit einer Schrittweite von  $1^\circ$  durchgeführt. Das Detektorband hat eine Länge von  $k = 300$ , wobei  $k$  gleich der Anzahl der Tiefenpixel von (a) ist. Somit bestimmt  $k$  auch die Anzahl der Strahlen, die durch das Phantombild geschickt worden sind.

Sinogramme sind in der Praxis der computertomographischen Bildrekonstruktion genau die gemessenen Daten. Somit ist die Aufgabe das Bild des Ursprungsobjekts aus einem Sinogramm zu rekonstruieren. In folgenden Abschnitten werden wir uns vier solcher Rekonstruktionsverfahren anschauen. Dabei unterscheiden wir zwischen *analytischen* und *algebraischen* Verfahren. Als Nächstes wollen wir den Begriff der analytischen Rekonstruktion klären.

## Analytische Rekonstruktionsverfahren

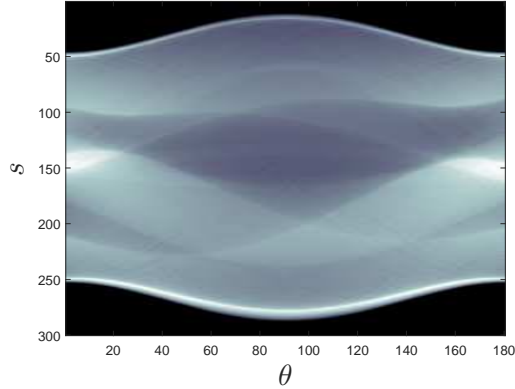
Der Begriff der *analytischen Rekonstruktion* bedeutet in diesem Fall, dass die Algorithmen der CT-Bildrekonstruktion direkt aus den analytischen Formeln hergeleitet werden. Solche Verfahren nutzen die Projektionsdaten zur Rekonstruktion des Bildes direkt aus. Zuerst schauen wir uns das einfachste davon an.

### Ungefilterte Rückprojektion

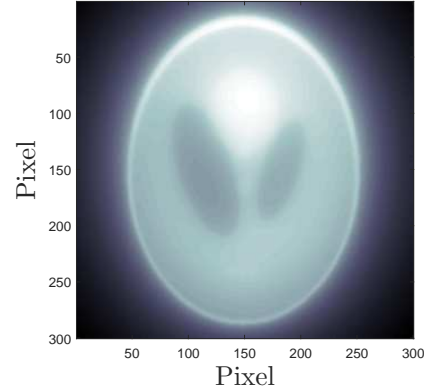
Den Ausgangspunkt dieses Abschnitts liefert uns der Adjungierte Operator (3.21). Schauen wir uns die Gleichung des Operators  $\mathcal{R}^*$  noch einmal an:

$$\mathcal{R}^*g(x) = \int_0^\pi g(x^T \omega(\theta), \theta) d\theta.$$

Aus der Bemerkung 3 wissen wir bereits, dass (3.21) die ungefilterte Rückprojektion darstellt. Das Argument des Integranden in der obigen Gleichung ist ein fester Punkt  $x$  der Kreisscheibe  $\Omega$ . Auf diesen Punkt  $x$  wird aus allen Winkeln zurück projiziert. Führt man das für alle  $x \in \Omega$  durch, so entsteht ein verschwommenes Bild des Originals, wie es die Abbildung 3.10b zeigt.



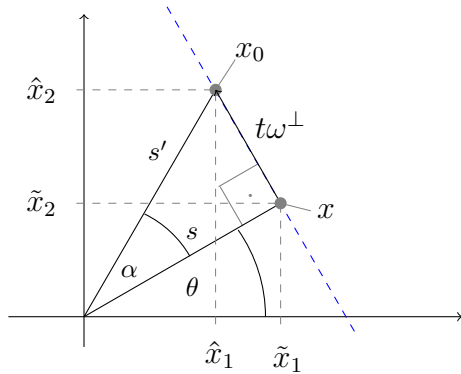
(a) Die Radon Transformation des Phantombildes aus Abb. 3.9a.



(b) Ungefilterte Rückprojektion

**Abbildung 3.10.:** Links (a) ist die Radontransformierte wie in der Abb. 3.9b. Rechts (b) ist die ungefilterte Rückprojektion von (b).

Jetzt wollen wir nachvollziehen, warum der Effekt der Verschwommenheit bei der ungefilterten Rückprojektion auftritt. Dazu sei zunächst die Abbildung 3.11 betrachtet, denn mit ihrer Hilfe können wir den Zusammenhang zweier Punkte in der Ebene  $\Omega$  nachvollziehen.



$$\beta = \alpha + \theta \quad (3.41)$$

$$x = \begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_2 \end{pmatrix} = s \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} = s\omega(\theta) \quad (3.42)$$

$$x_0 = \begin{pmatrix} \hat{x}_1 \\ \hat{x}_2 \end{pmatrix} = s' \begin{pmatrix} \cos(\beta) \\ \sin(\beta) \end{pmatrix} = s'\omega(\beta) \quad (3.43)$$

$$s'\omega(\beta) = s\omega(\theta) + t\omega^\perp(\theta) \quad (3.44)$$

**Abbildung 3.11.:** Eine Skizze (links) zur Verdeutlichung der Beziehung zweier Punkte  $x, x_0 \in \Omega$  bezüglich ihrer kartesischen- und Polar-Koordinaten. Die dazugehörige Nebenrechnung (rechts) zeigt deren formelmäßigen Zusammenhang auf.

Hier ist das Ziel zu zeigen dass, dass die ungefilterte Rückprojektion eine Faltung von  $f(x)$ ,  $x \in \Omega$  mit einer noch näher zu bestimmenden Funktion  $h(x)$ ,  $x \in \Omega$  darstellt. Dazu formulieren wir folgendes Lemma:

**Lemma 3.** Sei  $f(x) \in L^2(\Omega)$  beliebig und  $h(x) \in L^2(\Omega)$  speziell, dann kann die ungefilterte Rückprojektion als Faltung zweier Funktionen als

$$\mathcal{R}^*g(x_0) = \int_{\mathbb{R}^2} f(x)h(x - x_0)dx = f(x) * h(x).$$

geschrieben werden. Da  $\text{supp}f \subset \Omega$  betrachten wir das obige Integral nur auf  $\Omega$ .

**Bemerkung 8.** Bevor wir zum Beweis übergehen, führen wir die dafür benötigte  $\delta$ -Distribution und einige für uns nützliche Eigenschaften ein. Seien  $x, x_0 \in \mathbb{R}^2$ , dann gilt:

$$\delta(x - x_0) = \begin{cases} 0 & : x \neq x_0 \\ \infty & : x = x_0 \end{cases} \quad (3.45)$$

Eigenschaften :

$$(i). \quad \langle \delta(x - x_0), f(x) \rangle = \int_{-\infty}^{\infty} f(x)\delta(x - x_0)dx = f(x_0) \quad (3.46)$$

$$(ii). \quad \delta(ax) = \frac{1}{|a|}\delta(x) \quad (3.47)$$

$$(iii). \quad \int_{-\infty}^{\infty} \delta(x - x_0)dx = 1 \quad (3.48)$$

*Beweis.* Wir Beginnen mit der Definition der ungefilterten Rückprojektion (3.21) und benutzen dabei die Nebenrechnung zu Abb. 3.11 sowie die eingeführten Eigenschaften der  $\delta$ -Distribution aus der Bemerkung 8.

$$\begin{aligned} \mathcal{R}^*g(x_0) &= \int_0^\pi g(x_0^T\omega(\theta), \theta) d\theta \stackrel{(3.10, 3.44)}{=} \int_0^\pi \left( \int_{-\gamma(s)}^{\gamma(s)} f(s'\omega(\beta)) dt \right) d\theta \\ &\stackrel{(3.46)}{=} \int_0^\pi \left( \int_{-\gamma(s)}^{\gamma(s)} \int_{-1}^1 f(s\omega(\theta))\delta(s\omega(\theta) - s'\omega(\beta)) ds dt \right) d\theta \\ &\stackrel{(3.44)}{=} \int_0^\pi \left( \int_{\Omega} f(x)\delta(t\omega^\perp(\theta)) dx \right) d\theta \stackrel{(3.47)}{=} \int_{\Omega} f(x) \frac{\int_0^\pi \delta(\beta - \alpha) d\beta}{|t \frac{\partial \omega^\perp(\theta)}{\partial \theta}|} dx \\ &\stackrel{(3.48)}{=} \int_{\Omega} f(x) \frac{1}{|x - x_0|} dx = f(x) * h(x). \end{aligned} \quad (3.49)$$

In der letzten Zeile wurde  $|t \frac{\partial \omega^\perp(\theta)}{\partial \theta}| = |t| \left| \begin{pmatrix} -\cos(\theta) \\ -\sin(\theta) \end{pmatrix} \right| = |t| = |x - x_0|$  benutzt.  $\square$

Somit haben wir auch die spezielle Funktion  $h(x) = \frac{1}{|x|}$  gefunden. Die Funktion  $h$  heißt *Point-Spread-Function*. Anschaulich kann  $f$  in jedem Punkt als  $\delta$ -Distribution verstanden werden. Das heißt, setzt man  $f$  auf ganz  $\Omega$  Null, außer im Punkt  $x_0 \in \Omega$  und faltet mit  $h(x)$ , so wird das Ergebnis  $\tilde{f}$  um den Punkt  $x_0$  nach außen radial abfallen. Führt man die Faltung von  $f$  für jeden Punkt in  $\Omega$  und setzt auf diese Weise entstandene Faltungsbilder zusammen führt das zu dem Effekt der Verschwommenheit.

## Gefilterte Rückprojektion

Das in dem letzten Abschnitt erreichte Ergebnis ist für die Praxis unbrauchbar. Man kann an einem verschwommenen Bild keine qualitative Analyse durchführen. Deshalb werden wir in diesem Abschnitt eine Verbesserung des ersten Verfahrens suchen. Wie der Name schon verrät, bekommt man ein ungefiltertes Bild zurück. Wir wollen nun verstehen, wie man die Projektionsdaten filtert, bevor man sie zurück projiziert.

Um den Ausdruck der *gefilterten Rückprojektion* zu bekommen, werden wir uns des sogenannten *Fourier<sup>15</sup>-Slice-Theorem* (FST) bedienen (siehe z.B. [0, S. 120]). Das Schema des Theorems kann nun folgendermaßen wiedergegeben werden:

1. Berechne für einen festen Winkel  $\theta$  die Fouriertransformierte<sup>16</sup> (FT) einer Projektion  $\mathcal{R}f(s, \theta) = g(s, \theta)$

$$(\mathcal{F}g)(s, \theta) = G(q, \theta).$$

2. Konstruiere die Fouriertransformierte  $F(u, v)$  von  $f(x)$  aus  $G(q, \theta)$

$$G(q, \theta) \rightsquigarrow F(u).$$

3. Berechne die inverse Fouriertransformierte von  $F$

$$(\mathcal{F}^{-1}F)(u) = \tilde{f}(x).$$

Wir werden den Punkt 2 des FST vom Punkt 1 und dann vom Punkt 3 aus schrittweise annähern. Bevor wir die Herleitung der gefilterten Rückprojektion angehen, machen wir uns die vorliegende Situation an einer Skizze (Abb. 3.12) klar.

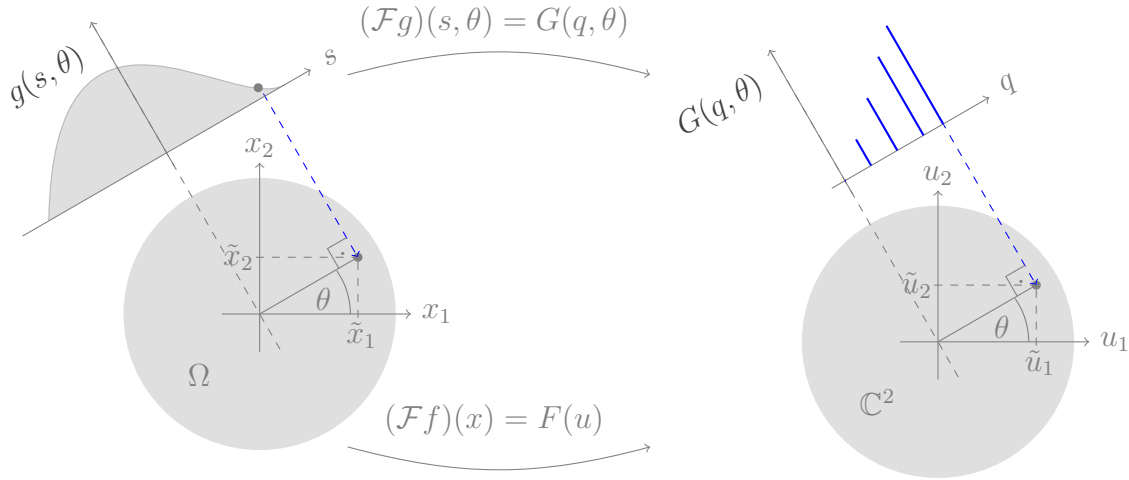
In der Abbildung 3.12 ist schematisch wiedergegeben, dass die Fourier-Transformation an der Natur der Koordinatensysteme nichts ändert. Diese Tatsache wird in dem *Plancherel-Theorem* bewiesen, welches unter anderem besagt, dass die die FT ein unitärer Operator für Funktionen aus  $L^2$  ist. Das heißt, die FT ist linear, isometrisch und bijektiv. Das weiteren beweist der Satz die Invertierbarkeit von  $\mathcal{F}$  für  $L^2$  Funktionen.

<sup>15</sup>Jean Baptiste Joseph Fourier (1768 - 1830) ein französischer Mathematiker und Physiker.

<sup>16</sup>**Definition** *Fourier-Transformation*. Sei  $f(x) \in L^2(\Omega)$ , dann ist ihre Fourier-Transformation durch

$$(\mathcal{F}f)(x) = \int_{\Omega} f(x) e^{-2\pi i \langle x^T, u \rangle} dx = F(u), \quad u \in \mathbb{C}^2 \quad (3.50)$$

gegeben.



**Abbildung 3.12.:** Eine Veranschaulichung des Fourier-Slice-Theorems. Links ist der Vorgang der Radon Transformation abgebildet, rechts seine Fouriertransformierte.

Jetzt haben wir genug Argumente gesammelt, um die gefilterte Rückprojektion herzuleiten. Dafür beginnen wir mit dem Punkt 1 des FST: wir bilden die Fouriertransformierte (Fußnote 16) einer Projektion  $p(s, \theta)$ :

$$\begin{aligned}
 P(q, \theta) &= \int_{-1}^1 p(s, \theta) e^{-2\pi i q s} ds \\
 &\stackrel{(3.10)}{=} \int_{-1}^1 \left( \int_{-\gamma(s)}^{\gamma(s)} f(s\omega(\theta) + t\omega^\perp(\theta)) dt \right) e^{-2\pi i q s} ds \\
 &\stackrel{(3.22)}{=} \int_{\Omega} f(x) e^{-2\pi i q x^T \omega(\theta)} dx.
 \end{aligned} \tag{3.51}$$

An dem Punkt 3 des FST ist nicht viel zu tun, deshalb geben wir hier direkt die Fouriertransformierte von  $f$  an. Bei der Wahl der Koordinaten für  $\mathcal{F}f$  und ihre Umrechnung stützen wir uns auf die rechte Seite der Abbildung 3.12. Dann gilt

$$F(q \cos(\theta), q \sin(\theta)) = \int_{\Omega} f(x) e^{-2\pi i q x^T \omega(\theta)} dx. \tag{3.52}$$

Aus den Gleichungen (3.51) und (3.52) können wir festhalten, dass die Fouriertransformierten von  $f$  und  $p$  im folgenden Bezug zueinander stehen

$$F(q \cos(\theta), q \sin(\theta)) = P(q, \theta). \tag{3.53}$$

Das heißt (3.53) soll eine Koordinatentransformation zwischen  $\mathcal{F}f$  und  $\mathcal{F}p$  darstellen. Somit ist der Punkt 2 des FST auch erledigt.

Im zweiten Schritt invertieren wir die Gleichung (3.52) und führen eine Variablensubstitution<sup>17</sup> durch, sodass man zu der Gleichung

$$f(x) = \int_0^\pi \int_{-1}^1 F(q \cos(\theta), q \sin(\theta)) e^{2\pi i q x^T \omega(\theta)} q \, dq \, d\theta \quad (3.54)$$

gelangt. Die Symmetriebetrachtungen, wie in [0, S. 135] und die Tatsache, dass man nur reelle Funktionen auf  $\Omega$  betrachtet, führen zu folgendem Ausdruck

$$\begin{aligned} f(x) &= \int_0^\pi \int_{-1}^1 F(q \cos(\theta), q \sin(\theta)) e^{2\pi i q x^T \omega(\theta)} |q| \, dq \, d\theta \\ &\stackrel{(3.53)}{=} \int_0^\pi \left( \int_{-1}^1 (P(q, \theta) |q|) e^{2\pi i q x^T \omega(\theta)} \, dq \right) d\theta. \end{aligned} \quad (3.55)$$

Mit der Gleichung (3.55) haben wir die gefilterte Rückprojektion erhalten. Es bleibt noch das Gefilterte darin zu entdecken. Beginnen wir mit der inneren Klammerung von (3.55).  $P(q, \theta)$  ist die Fouriertransformierte  $\mathcal{F}p$  von  $p(s, \theta)$ , die punktweise mit einer Betragsfunktion  $b(q) = |q|$  multipliziert wird. Da wir jetzt  $p(s, \theta)$  im Frequenzraum betrachten und die Funktion  $b(q)$  über das gleiche Frequenzband, wie  $P(q, \theta)$  läuft, werden durch punktweise Multiplikation kleinfrequente Anteile von  $P(q, \theta)$  von  $b(q)$  unterdrückt und die hochfrequente im Gegensatz sehr verstärkt. Dieses Vorgehen wird im Allgemeinen *Hochpassfilterung* genannt. Die Funktion  $b(q) = |q|$  bezeichnet man als Rampenfilter. Im Folgenden bezeichnen wir einen Filter mit  $h_\xi(q)$  und wir schreiben  $h_b(q) = b(q) = |q|$ . Für die Filterung der Projektionen führen wir die Bezeichnung  $P_{h_\xi}(q, \theta) = P(q, \theta) h_\xi(q)$  ein.

Nach dem die Filterung der Projektion durchgeführt wurde, wird diese mittels der inversen FT zurück in den Ortsraum der Projektionen überführt. Was durch das innere Integral von (3.55) dargestellt wird. Anschließend projiziert man die gefilterte Projektion in den Ortsraum von  $f$  zurück, dies ist das äußere Integral von (3.55).

Wir fassen die nötigen Schritte der gefilterten Rückprojektion zu einem Algorithmus zusammen

---

**Algorithmus 3.1:** Gefilterte Rückprojektion (Filtered Back Projection, FBP)

---

- 1: Berechne die FT von  $(\mathcal{F}p)(s, \theta) = P(s, \theta)$ .
  - 2: Führe die Filterung  $P_{h_\xi}(q, \theta) = P(q, \theta) h_\xi(q)$  durch.
  - 3: Berechne die inverse FT  $(\mathcal{F}^{-1} P_{h_\xi})(q, \theta) = \tilde{p}(s, \theta)$ .
  - 4: Projiziere  $\tilde{p}(s, \theta)$  auf  $\Omega$  zurück.
- 

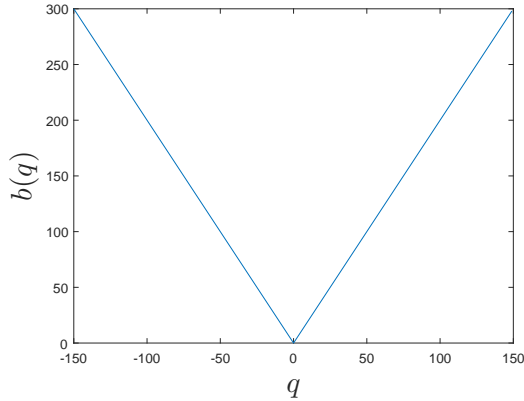
In der Abbildung 3.13b ist das erste Ergebnis der gefilterten Rückprojektion, gefiltert mit einem Rampenfilter zu sehen.

---

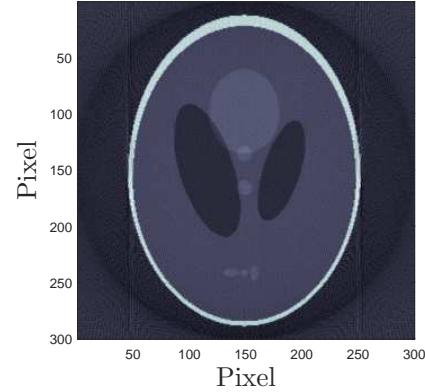
<sup>17</sup>Für die Substitution von  $du_1 du_2$  berechnen wir die Jacobi-Determinante von  $|J(\begin{pmatrix} u_1 \\ u_2 \end{pmatrix})| = \left| \begin{pmatrix} c \cos(\theta) & \sin(\theta) \\ -q \sin(\theta) & q \cos(\theta) \end{pmatrix} \right| = q$ , somit erhalten wir  $du_1 du_2 = q dq d\theta$ .



Wenn man das Sinogramm (Abb. 3.9b) betrachtet, so stellt man fest, dass das Sinogramm keine Störungen aufweist. Das würde eine perfekte Messung bedeuten. Dies ist aber bei den realen Messungen nie der Fall. Deshalb wollen wir uns die Rekonstruktion von gestörten Daten anschauen. Dafür soll das Phantom Bild (Abb. 3.9a) mit einem Salzstreuer-Effekt (Quelle: MATLAB) beaufschlagt werden. Das soll die Homogenität einiger Pixelbereiche von der Abbildung 3.9a aufheben, was eine kleine Annäherung zur Realität widerspiegelt (Abb. 3.14a).



(a) Das Rampenfilter.



(b) Eine mit Rampenfilter gefilterte Rückprojektion.

**Abbildung 3.13.:** Links (a) das Rampenfilter, konstruiert passend zu dem Sinogramm aus 3.9b. Das Frequenzband des Filters ist durch die isometrische Eigenschaften der FT gleich der Detektorbreite, also ist auch gleich der Tiefe des Sinogramms. Rechts (b) ist die gefilterte Rückprojektion von 3.9b, durchgeführt unter der Einwirkung des Rampenfilters aus (a).

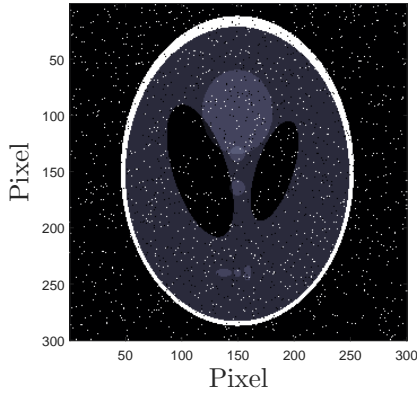
Die Radon Transformation von der Abbildung 3.14a ohne zusätzliche Störung würde nicht die Messfehler repräsentieren, höchstens Rundungsfehler, die bei der Integration entstehen, hier aber nicht wesentlich sind. Deshalb stören wir die Projektionsdaten zusätzlich mit dem Salzstreuer-Effekt (Abb. 3.14b).

Wir erinnern uns an die Ergebnisse des zweiten Kapitels. Insbesondere daran, dass der Operator  $\mathcal{R}$  schlecht gestellt ist. Dazu sei die Gleichung (3.36) nochmal betrachtet

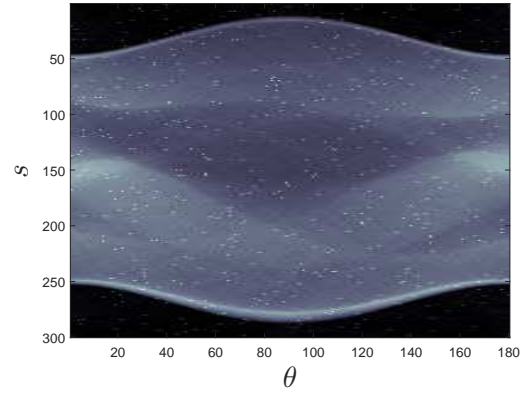
$$\mathcal{R}^+ g = \sum_{j=1}^n \sigma_j^{-1} \langle g, u_j \rangle_{L^2(Z)} v_j.$$

Wir wissen, dass  $g$  die Projektionen in (3.36) darstellt. Sei nun  $p_\epsilon$  unsere gestörte Projektion. Des Weiteren betrachten wir den Schritt 2,3 des Algorithmus 3.1 und schreiben ihn, wie folgt auf

$$\tilde{p}_\epsilon = \int_{-1}^1 P(q, \theta) h_\xi(q) e^{2\pi i q x^T \omega(\theta)} dq. \quad (3.56)$$



(a) Ein verrauschtes Phantombild.



(b) Die Radontransformierte von (a) und einem zusätzlichem Rauschen.

**Abbildung 3.14.:** Links (a) ist ein Phantombild verrauscht mit einem Salzstreuer-Effekt (Quelle: MATLAB) dargestellt. Rechts (b) ist Die Radontransformierte von (a) und zusätzlich mit einem Salzstreuer-Effekt verrauscht.

Wir wollen nun  $\tilde{p}_\epsilon$  als eine Linearkombination oder auch als verallgemeinerte Fourierreihe angeben. Das tun wir direkt für den diskreten Fall unter Beachtung der Bemerkung 5. Somit bekommen wir

$$\tilde{p}_\epsilon \approx \sum_{i=1}^n c_i h_\xi(q_i) u_i, \quad \text{mit } c_i = \langle \tilde{P}, u_i \rangle. \quad (3.57)$$

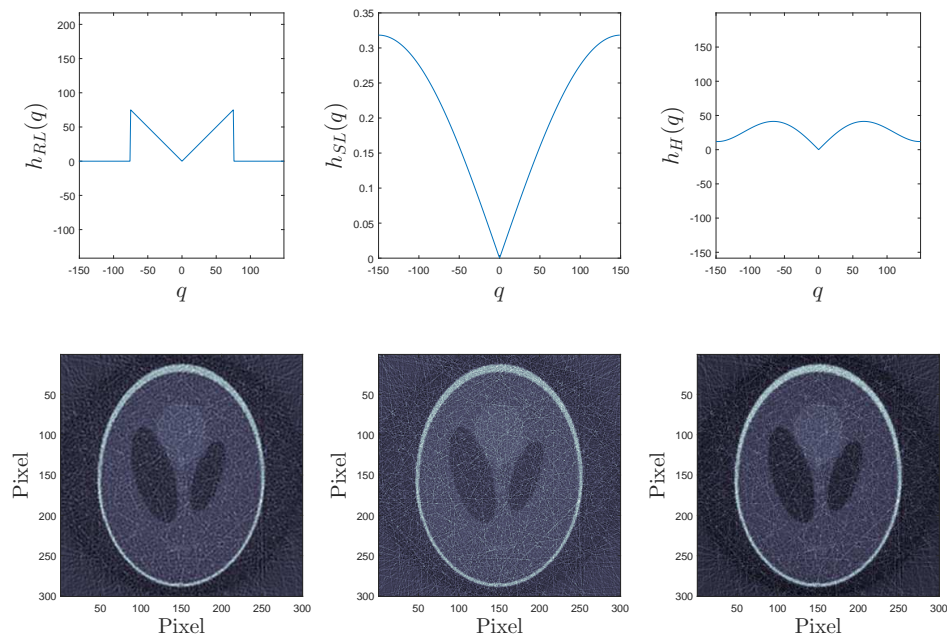
Diese Gleichung werden wir nicht beweisen, aber praktisch bedeutet das, wenn man  $h_\xi(q) = 1$  setzt und es dem FBP Algorithmus übergibt, kommt genau die ungefilterte Projektion dabei raus.

Nun setzen wir unsere gestört-gefilterte Projektion (3.57) in (3.36) ein und bekommen einen Ausdruck

$$\begin{aligned} \mathcal{R}^+ g &= \sum_{j=1}^n \sigma_j^{-1} \left\langle \left( \sum_{i=1}^n c_i h_\xi(q_i) u_i \right), u_j \right\rangle_{L^2(Z)} v_j \\ &= \sum_{j,i=1}^n (\sigma_j^{-1} h_\xi(q_j)) c_i \delta_{ij} v_j = \sum_{j=1}^n (\sigma_j^{-1} h_\xi(q_j)) c_j v_j, \end{aligned} \quad (3.58)$$

der uns prinzipiell keine neue Erkenntnis liefert, aber offensichtlich macht, wie die hochfrequenten Anteile bei der Rekonstruktion von dem Filter beeinflusst werden können. Das heißt, das Design des Filters kann die Rückprojektion in verschiedenen Frequenzbereichen dämpfen und somit ihre Oszillation bis zu einem gewissen Grade beschränken. Anders gesagt, wir beschleunigen somit die Konvergenz der Reihe (3.58) zu unseren Gunsten.

In der Abbildung 3.15 sind drei Ergebnisse mit drei verschiedenen Filtern dargestellt. Zum Aufbau verschiedener Filterkerne sei man hier auf [0, S. 191] verwie-



**Abbildung 3.15.:** Eine Rekonstruktion aus den Daten 3.14b unter der Einwirkung von drei verschiedenen Filtern. Unten links ist die Filterung mit dem *Ramachandran-Lakshminarayanan-Filter*, mittig mit dem *Shepp-Logan-Filter* und rechts mit dem *Hamming-Fiter* dargestellt.

sen. In der Abbildung 3.15 links unten, ist zu erkennen, dass ein *Ramachandran-Lakshminarayanan-Filter* die hohen Frequenzen abschneidet, wodurch das Bild insgesamt an Kontrast verliert. Das untere Ergebnis in der Mitte zeigt, dass das *Shepp-Logan-Filter* die mittleren Frequenzen anhebt, was die Helligkeit des Ergebnisses etwas erhöht. Das untere rechte Bild hat durch das *Hamming-Fiter* volles Frequenzband erhalten, allerdings sind die hohen Frequenzen stark gedämpft worden.

Insgesamt existiert ein Dutzend Filter, die bei der CT-Bildrekonstruktion eingesetzt werden können. Dieses Vorgehen nennt man *Regularisierung* des Rückprojektionsproblems mit Filtern.

## Algebraische Rekonstruktionsverfahren

Algebraische Verfahren basieren auf der Auffassung des Operators  $\mathcal{R}$  als eine Matrix. Das hat zu Folge, dass man die Wirkung von  $\mathcal{R}$  auf  $f$  in Form eines linearen Gleichungssystems auffassen kann. Somit bekommen wir

$$\mathcal{R}f = Af = p. \quad (3.59)$$

**Bemerkung 9.** Für den Rest dieses Kapitels vereinbaren wir nur diskretisierte Systeme zu betrachten.

Mit der Bemerkung 9 und den vereinbarten Größen  $f \in \mathbb{R}^m$  (3.39) und  $p \in \mathbb{R}^n$  (3.38) ist  $A$  eine endliche Matrix.  $A$  wird *Systemmatrix* des Abbildungssystems von

$\mathcal{R}$  genannt. Wir werden zunächst die Entstehung und den Aufbau von  $A$  etwas genauer anschauen.

### Aufbau der Systemmatrix $A$ von $\mathcal{R}$

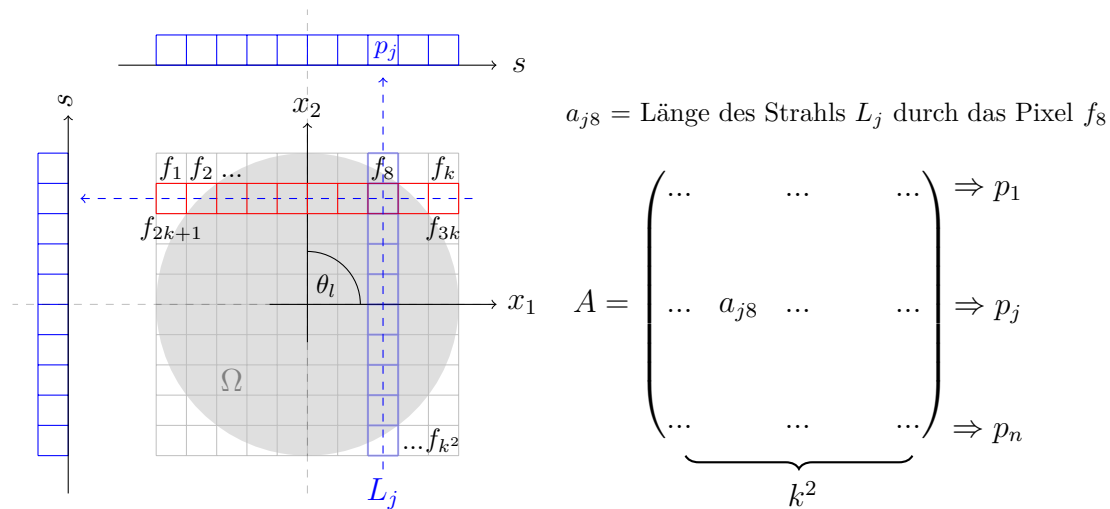
Da  $f \in \mathbb{R}^m$  und  $p \in \mathbb{R}^n$  sind bedeutet es, dass  $A \in \mathbb{R}^{n \times m}$  ist. Das heißt, die Zeilen von  $A$  laufen über die Indizierung der Projektionen  $p$ , also die Zeile  $j$  ist die  $j$ -te Projektion. Die Spalten von  $A$  laufen über die Indizierung von  $f$ .

$f$  wird durch die Anzahl der Detektoren  $k$  im Detektorband diskretisiert (Abb. 3.16). Man kann also hypothetisch  $f \in \mathbb{R}^{k \times k}$  annehmen. Lassen wir die Indizierung von  $f$  Zeilenweise durchlaufen, so kommen wir genau auf  $f \in \mathbb{R}^{k^2}$ . Wobei nach (3.39)  $k^2 = m$  gilt.

Aus der Gleichung (3.59) können wir folgern, dass für  $j$ -te Projektion

$$\sum_{i=1}^m a_{ji} f_i = p_j \quad (3.60)$$

gelten muss. Beachtet man, dass nur die  $f_j$  die vom Strahl getroffen worden sind zur Summe (3.60) beitragen, ergibt sich die Gleichung (3.40).



**Abbildung 3.16.:** Eine Skizze zur Veranschaulichung des Aufbauprinzips der Abbildungsmatrix  $A$  von  $\mathcal{R}$ .

Jetzt bleibt noch zu klären, wie das Systemmatrix  $A$  initialisiert wird. In der Abbildung 3.16 sehen wir den Strahl  $L_j$ , der zu den Projektionen unter dem Winkel  $\theta_1 = 0^\circ$  gehört. Der Strahl trifft unter anderen auch das Bildpixel  $f_8$ . Berechnet man die Länge des Strahls durch dieses Pixel, so kann man die errechnete Größe an der Stelle des Gewichts  $a_{j8}$  in  $A$  eintragen. Dementsprechend gilt es für alle betroffenen Pixel pro Strahl.

Für jeden Winkel werden immer  $k \in \mathbb{N}$  Projektionen erzeugt, so ergibt sich aus der Anzahl der Winkel  $q \in \mathbb{N}$  die Anzahl der Projektionen  $n = kq$ , was dem Ausdruck (3.38) entspricht.

Zu bemerken ist noch, dass die Matrix  $A$  eine *dünnbesetzte* Matrix ist. Mann kann leicht einsehen, dass auf der Diagonalen von  $f$  maximal  $\lfloor \sqrt{2} \rfloor k$  Pixel liegen.

Wenn der Strahl etwas neben der Diagonalen parallel verläuft, trifft er höchstens  $2\lfloor\sqrt{2}\rfloor k - 1$  Pixel von  $f$ ; dies ist auch die maximale Anzahl der Einträge von  $A$  in einer Zeile.

**Bemerkung 10.** *Die Längen des Strahls pro Pixel zu berechnen, ist nicht die einzige Möglichkeit  $A$  zu modellieren. Die Gewichte  $a_{ji}$  können auch als Verhältnis der Fläche des Strahls zu der Pixelfläche aufgestellt werden. Das modelliert zwar das Abbildungssystem von  $\mathcal{R}$  besser, ist aber viel rechenaufwendiger, als die hier betrachtete Vorgehensweise [0, S. 141].*

Die Implementierung der Funktion zur Erstellung der Systemmatrix ist in ?? kommentiert und ist auf der beigelegten CD-ROM einzusehen.

Nachdem die Konstruktion der Systemmatrix etwas erhellt wurde, wollen wir zu der Diskussion der darauf aufbauenden Rekonstruktionsverfahren übergehen. Die sogenannten *algebraische Rekonstruktionstechniken* (eng. *algebraic reconstruction technique, ART*) beruhen auf der Kenntnis von  $A$ . Als erstes würde man wahrscheinlich auf die Idee kommen, die Gleichung  $Af = p$  zu invertieren. Nun in der Regel ist  $A$  sehr groß und hat keine einfache Struktur, sodass keine schnelle Inversion gefunden werden kann und wenn, dann ist es sehr zeit- und speicherintensiv (vergl. [0, S. 157]). Zudem kann es vorkommen, dass  $m > n$  ist, was heißen würde, dass das System unterbesetzt ist und dadurch keine eindeutige oder gar keine Lösung haben kann. Deshalb sucht man nach anderen Techniken  $f$  zu rekonstruieren.

## Iterative Rekonstruktion nach Kaczmarz

In diesem Abschnitt wollen wir ein iteratives Verfahren zur Lösung des Gleichungssystems (3.59) nach Kaczmarz<sup>18</sup> anschauen. Das Verfahren kann folgendermaßen hergeleitet werden. Dafür betrachten wir (3.59) in expandierter Form

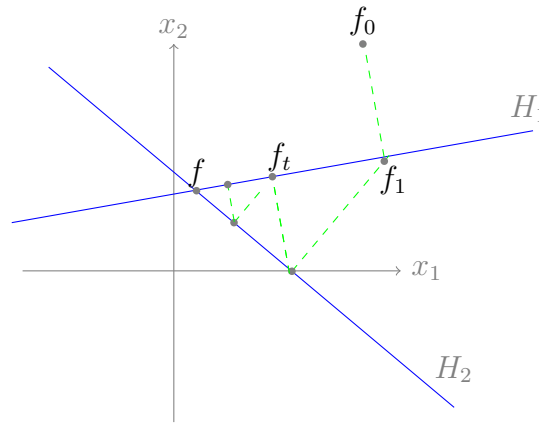
$$\begin{aligned} a_{11}f_1 + \dots + a_{1m}f_m &= p_1 \\ &\vdots \\ a_{j1}f_1 + \dots + a_{jm}f_m &= p_j \\ &\vdots \\ a_{n1}f_1 + \dots + a_{nm}f_m &= p_n. \end{aligned} \tag{3.61}$$

So stellt jede der obigen Gleichungen eine Hyperebene  $H_j = \{f \in \mathbb{R}^m \mid \langle a_j, f \rangle = p_j, j \in [1, n], n \in \mathbb{N}\}$  dar. Hier ist  $a_j$  die  $j$ -te Zeile von  $A$ . Damit liegt die Lösung von (3.61) genau in dem Schnittpunkt aller Ebenen, was in diesem Falle  $f \in \mathbb{R}^m$  bedeuten würde.

Zum Suchen von  $f$  bietet sich die Methode der Lotprojektion an. Dafür wählt man einen beliebigen Punkt  $f_0$  und projiziert einen Lot von  $f$  aus auf eine beliebige Hyperebene  $H_j$ . Der Schnittpunkt des Lots mit  $H_j$  ist dann  $f_1$ . Von  $f_1$  aus projiziert man auf die nächste Hyperebene, so initialisiert man  $f_2$  und so weiter, bis man an dem Schnittpunkt  $f$  angekommen ist (Abb. 3.17).

<sup>18</sup>Stefan Kaczmarz (1895-1939) polnischer Mathematiker.

Dieses Vorgehen beschreibt genau die Arbeitsweise des Algorithmus, also das Kaczmarz-Verfahren. Man unterscheidet zwischen dem vorwärts projektiven Verfahren und dem randomisierten. Der Unterschied besteht nur darin, dass bei dem ersten Verfahren die Projektionen auf die Ebenen der Reihe nach berechnet werden und in dem zweiten Verfahren werden die Ebenen zufällig ausgewählt.



**Abbildung 3.17.:** Eine Veranschaulichung der Arbeitsweise des Kaczmarz-Verfahrens in  $\mathbb{R}^2$ .

**Bemerkung 11.** Die Abbildung 3.17 zeigt ein Fall, wo die Lösung immer gefunden werden kann, denn der Schnittpunkt zweier Geraden ist immer exakt bestimmbar. Schon ab  $\mathbb{R}^n$  für  $n \geq 2$  ändert sich die Situation in den realen Fällen drastisch. Der Grund dafür ist, dass man ein mit Rauschen behaftetes System  $Af \approx p + \epsilon$  hat. Somit ist jede Hyperebene  $H_j$  um ein  $\epsilon_j$ ,  $j \in [1, n]$ ,  $n \in \mathbb{N}$  von dem Schnittpunkt verschoben. Das heißt, dass die Lösung  $\tilde{f} \approx f$  in einem kleinem Toleranzvolumen  $\prod_{j=1}^n \epsilon_j$  liegt und daher nur annähernd angebar ist.

In [0] wird gezeigt, dass das randomisierte Verfahren eine schnellere Konvergenz nachweist. Aus diesem Grund werden wir dieses Vorgehen in Betracht ziehen.

**Definition 9** (Randomisierter Kaczmarz-Algorithmus). Sei  $Af = p$  ein konsistentes, lineares Gleichungssystem. Sei  $f_0 \in \mathbb{R}^m$  ein beliebiger Startwert. Für  $t = 0, 1, 2, \dots$  ist der Randomisierte Kaczmarz-Algorithmus definiert als

$$f_{t+1} = f_t + \frac{p_j - \langle a_j, f_t \rangle}{\|a_j\|_2^2} a_j, \quad (3.62)$$

wobei  $a_j$  die  $j$ -te Zeile von  $A$  ist und  $p_j$  die dazugehörige Projektion.  $j$  wird zufällig ausgewählt.

Wir wollen nun (3.62) etwas besser verstehen. Dafür formen wir diese Gleichung etwas um und bekommen

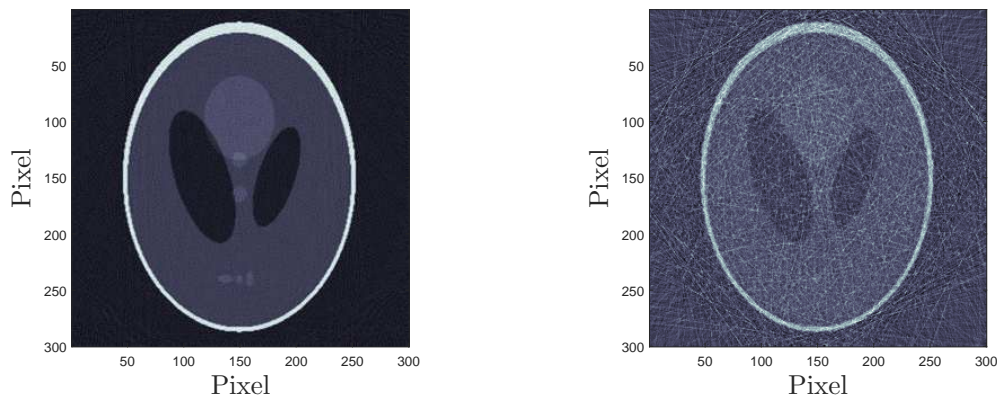
$$f_{t+1} = f_t + \left( \frac{p_j}{\|a_j\|_2} - \left\langle \frac{a_j}{\|a_j\|_2}, f_t \right\rangle \right) \frac{a_j}{\|a_j\|_2}.$$

Der eingeklammerte Teil der Gleichung stellt den Abstand von  $f_t$  zu der  $j$ -ten Hyperebene  $H_j$  in der Hesseschen Normalform Darstellung dar. Die Multiplikation

mit  $\frac{a_j}{\|a_j\|_2}$  liefert die Orthogonalität zu der  $H_j$ . Somit haben wir die oben beschriebene Lotprojektion, oder auch orthogonale Projektion auf  $H_j$  genannt.

Nun wollen wir uns von der Qualität des Verfahrens überzeugen, dessen genaue Implementierung in ?? (insbesondere die Abbruchbedingung ??) des Verfahrens) erläutert ist. Dazu sei die Abbildung 3.18 betrachtet. Die Ergebnisse in der Abbildung 3.18 wurden unter den gleichen Bedingungen, wie für die gefilterte Rückprojektion in den Abbildungen 3.13 für ungestörte und 3.15 (mittig) für gestörte Daten erzeugt.

Man erkennt, dass die Ergebnisse der iterativen Rekonstruktion optisch gleichviel (oder sogar mehr) Informationsgehalt bieten, wie die der gefilterten Rückprojektion.



(a) Iterative Rekonstruktion aus nicht verrauschten Daten von 3.9b.

(b) Iterative Rekonstruktion aus verrauschten Daten von 3.14b

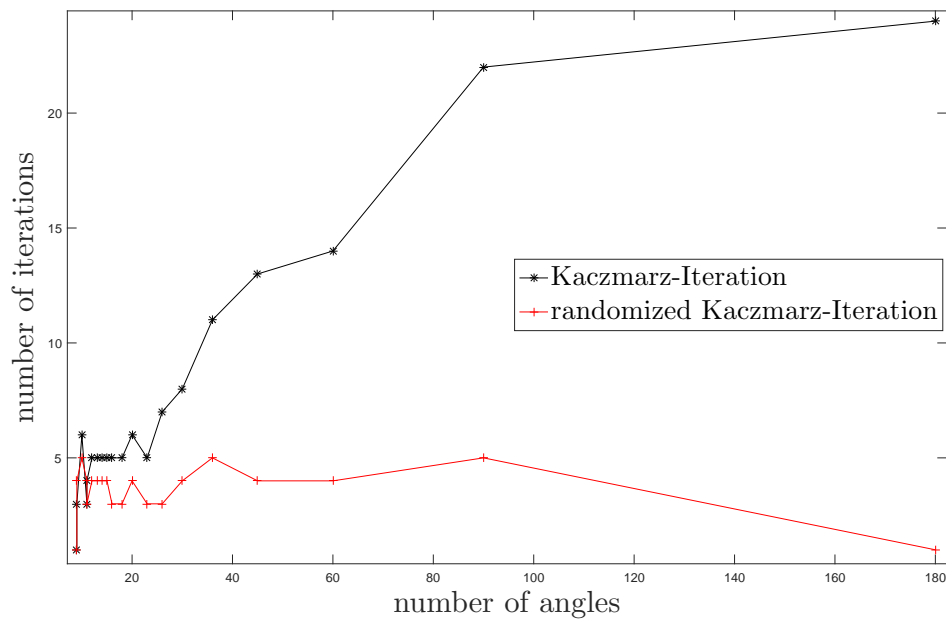
**Abbildung 3.18.:** Randomisierte iterative Rekonstruktionen (a) nicht verrauscht, (b) verrauscht. Die Projektionen in Beiden Fällen wurden für alle Winkel  $\theta$  von  $1^\circ$  bis  $180^\circ$  mit der Schrittweite von  $1^\circ$  berechnet. Detektorbreite ist  $k = 300$ .

Jetzt wollen wir experimentell an einem Beispiel nachweisen, dass das randomisierte Verfahren in der Tat schneller konvergiert als das vorwärts projektive. Für unsere Testzwecke wählen wir ein Phantombild mit der Größe  $50 \times 50$  Pixel (MATLAB, `phantom(50)`). Die kleine Größe des Phantombildes soll die Größe der Systemmatrix in annehmbaren Rahmen halten, da sonst zu große Matrizen entstehen, die sehr lange Rechenzeiten brauchen werden. Schließlich wollen wir nur die Konvergenz des Verfahrens untersuchen.

In der Abbildung 3.19 ist das Ergebnis des Tests aus verrauschten (Salzstreuer-Effekt) Daten zu sehen. Dabei wurde die Systemmatrix in jedem Testschritt vergrößert, indem sich die Winkelanzahl und somit die Projektionsanzahl erhöhten. Der erste Testschritt wurde mit einer Winkelschrittweite von  $20^\circ$ , also für  $q = 9$  Winkelpositionen durchgeführt. Die Detektorbreite wurde mit  $k = 50$  initialisiert, was insgesamt für  $A \in \mathbb{R}^{450 \times 2500}$  im ersten Testschritt ergeben hat. Der Test wurde bis  $q = 180^\circ$  mit der Winkelschrittweite von  $1^\circ$ , das heißt bis  $A \in \mathbb{R}^{9000 \times 2500}$  durchgeführt.

Die Abbruchbedingung musste für jeden Testschritt angepasst werden, da die





**Abbildung 3.19.:** Ein Vergleich des randomisierten und vorwärts projektiven Kaczmarz-Verfahrens. Der Test wurde für die Rekonstruktion eines  $50 \times 50$  Pixel großen Bildes durchgeführt. In jedem Testschritt wurde die Winkelanzahl erhöht, sodass die Systemmatrix immer größer wurde.

Größe von  $A$  schrittweise erhöht wurde und damit auch die Fehleranfälligkeit des Verfahrens. Aus (??) folgt  $\delta = \alpha\epsilon$ , wo  $\epsilon = \|p - \tilde{p}\|_2$  der Norm des Fehlers.  $\alpha$  wurde als Verhältnis der Zeilenanzahl von  $A$  im jetzt-Schritt zu davor-Schritt multipliziert mit dem Verhältnis der Winkelanzahl im jetzt-Schritt zu davor-Schritt aufgefasst. Da die Abbruchbedingung für beide Verfahren gleich war, kann man den Test als fair annehmen.

## Rekonstruktion durch SWZ

Das letzte Verfahren, das wir uns anschauen werden beruht auf der Singulärwertzerlegung einer Matrix. Bekanntlich besitzt jede Matrix  $A \in \mathbb{R}^{n \times m}$  (wir beschränken uns hier auf reelle Matrizen) mindestens eine SWZ, sodass  $A = U\Sigma V^*$  geschrieben werden kann.  $U \in \mathbb{R}^{m \times m}$  ist die *unitäre* Matrix des Bildes von  $A$ ,  $\Sigma \in \mathbb{R}^{m \times n}$  beinhaltet die Singulärwerte von  $A$  auf der Diagonale sonst überall Null und  $V^*$  ist die adjungierte der unitären Matrix des Kerns von  $A$ . Somit lässt sich die Gleichung (3.59) in der Form

$$Af = (U\Sigma V^*)f = p \quad (3.63)$$

schreiben. Aus der Definition der unitären Matrizen folgt  $U^{-1} = U^*$  und  $(V^*)^{-1} = (V^*)^* = V$ . Da  $\Sigma$  eine Diagonalmatrix mit Singulärwerten größer Null ist, kann  $\Sigma$



wie folgt invertiert werden

$$\Sigma^+ = \begin{cases} \frac{1}{\sigma_{ji}} & : \text{ falls } j = i \\ 0 & : \text{ sonst} \end{cases}. \quad (3.64)$$

Somit kann die Gleichung (3.63) invertiert werden, also

$$f = (V\Sigma^+U^*)p. \quad (3.65)$$

Schreibt man die Gleichung 3.65 in expandierter Form auf

$$f_i = \sum_{i=1}^m [\Sigma^+]_{ii} \langle p, u_i \rangle v_i, \quad (3.66)$$

was mit  $[\Sigma^+]_{ii} = \sigma_i^{-1}$  genau (3.36) entspricht.

Mit den Ergebnissen aus dem Kapitel 2 wissen wir, dass bereits kleine Fehler in dem Messprozess stark oszillierend auf die Lösung (3.66) wirken können. Somit bedarf es einer Regularisierung in der Rekonstruktion. Wir werden die *Tikhonov*<sup>19</sup>-Regularisierung benutzen, die wie folgt aufgeschrieben werden kann:

$$d_i = \frac{\sigma_i}{\lambda^2 + \sigma_i^2}. \quad (3.67)$$

Man erkennt, dass die Regularisierung für  $\lambda = 0$  keine Wirkung ergibt. Ist  $\lambda \geq 0$  so wirkt die Regularisierung dämpfend auf  $\sigma_i$  und verhindert bei großen Werten die Oszillation der Lösung. Wir ersetzen die Elemente auf der Diagonale von  $\Sigma^+$  mit (3.67) und benennen die Matrix mit  $D$

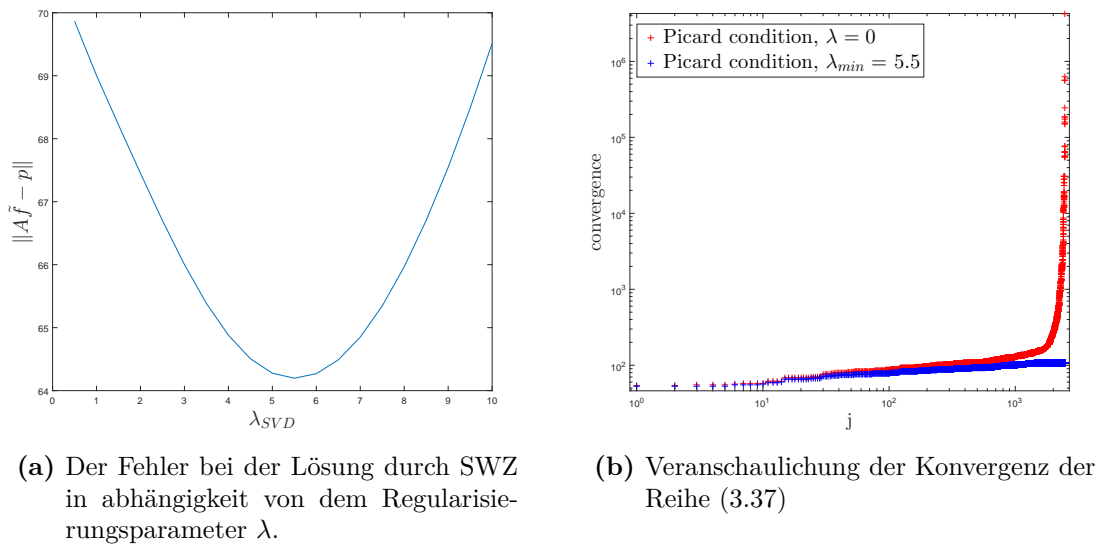
$$\tilde{f}_i = \sum_{i=1}^m [D]_{ii} \langle p, u_i \rangle v_i. \quad (3.68)$$

Wir wollen unsere Untersuchungen an dem Phantom-Beispiel (`phantom(50)`) weiterführen. Wie in dem vorherigen Abschnitt werden die Daten mit dem Salzstreuer-Effekt beaufschlagt, sowie anschließend auch die Messdaten. Dabei wurden die Projektionen für 50 Winkel mit der Schrittweite von  $3.6^\circ$  aufgenommen. Folglich bekommt man eine Systemmatrix  $A \in \mathbb{R}^{2500 \times 2500}$ .

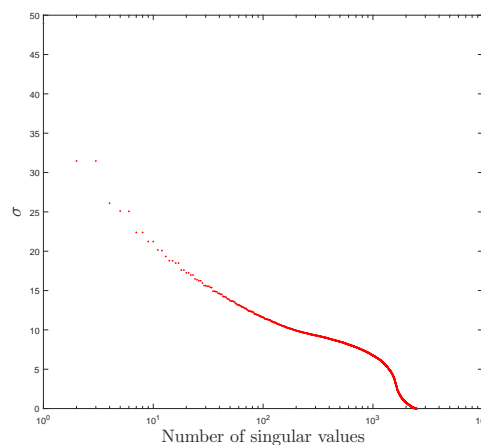
Nun wollen wir die Wirkung der Regularisierung (3.68) auf die Reihe (3.37) aus dem Satz 3 an unserem Beispiel verifizieren. Zunächst lassen wir das Verfahren (??) für verschiedene  $\lambda$  das Phantombild rekonstruieren und tragen den entstandenen Rekonstruktionsfehler  $\|A\tilde{f} - p\|_2$  gegen  $\lambda$  in einem Diagramm auf. Das Ergebnis ist in der Abbildung 3.20a zu sehen. Zu erkennen ist, dass der minimale Fehler etwa bei  $\lambda_{min} = 5.5$  liegt (Abb. 3.20a). Anschließend lassen wir die Reihe (3.37) für zwei Parameter laufen, also  $\lambda = 0$  und  $\lambda_{min} = 5.5$ . In der Abbildung 3.20b ist deutlich zu erkennen, dass die Reihe mit regularisierten Singulärwerten schneller konvergiert als die ohne Regularisierung.

Insgesamt zeigt die Abbildung 3.20b die Schlechtgestellttheit des hier betrachten Problems  $Af = p$  an einem konkreten Beispiel auf. Wir wissen, dass die Pcard-Bedingung (3) einen schnellen Abfall der Fourier-Koeffizienten  $\langle p, u_i \rangle$  fordert, was

<sup>19</sup>Andrei Nikolajewitsch Tichonow (1906-1993) russischer Mathematiker.

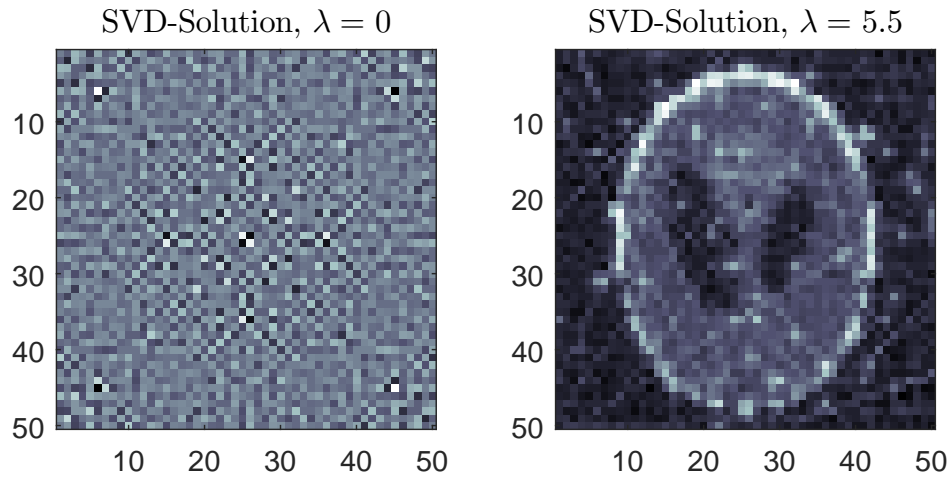


**Abbildung 3.20.:** (a) Rekonstruktion durch SWZ, für ein Phantombild ( $50 \times 50$ ), Mit der Detektorbreite  $k = 50$  und Winkelanzahl  $q = 50$ . (b) die Konvergenz der Reihe (3.37) blau mit Regularisierung  $\lambda_{min} = 5.5$  (das Minimum aus (a)), rot ohne Regularisierung.



**Abbildung 3.21.:** Die Folge der Singulärwerte für  $A$  aus 3.20a

eine stabile Lösung garantieren würde. Jedoch ist das bei Kompakten Operatoren, wie  $\mathcal{R}$  ist es nie der Fall. Der Grund ist die SWZ, bei der die Singulärwerte eine abfallende Folge bilden (Abb. 3.21). Was natürlich bei sehr kleinen Singulärwerten die Konvergenz der Reihe nicht fördert. Konvergiert die Reihe nicht, heißt das, dass die gemessenen Projektionen nicht in dem Bild der Operators  $\mathcal{R}$  liegen. Was im Falle der verrauschten Daten garantiert ist. Durch eine passende Regularisierung ( $\lambda = 5.5$ ) verschiebt man die verfälschten Fourier-Koeffizienten in die richtige Richtung und bekommt damit eine Minimum-Norm-Lösung. Was hier wieder zu der Abbildung 3.20a führt. Die Wirkung der Regularisierung ist in der Abbildung 3.22 offensichtlich.



**Abbildung 3.22.:** Die Rekonstruktion des Phantombildes mit der Größe  $50 \times 50$  Pixel. Links nicht regularisiert  $\lambda = 0$ , recht Regularisierung mit  $\lambda = 5.5$ .

Es ist natürlich sehr aufwendig, das Verfahren für verschiedene  $\lambda$  rechnen zu lassen, um eine Minimum-Norm-Lösung zu bekommen. In der Praxis gibt es ein heuristisches Verfahren mit dem man ein optimales  $\lambda_{opt}$  bestimmen kann. In der Abbildung (3.21) sehen wir, dass die Folge des Singulärwerte etwa bei  $10^3$  einen Knick nach unten macht, die Singulärwerte in diesem Bereich haben den Wert zwischen 4-6. Setzt man diese Werte für  $\lambda$  zu Regularisierung ein, so liefert das Verfahren die Minimum-Norm-Lösung.

# A. Implementierung

Hinweise zu konkreten Gegebenheiten, so etwas, wie FEniCS Package und dessen Nichtverfügbarkeit unter Windows, umsomehr die Benutzung des Dockers für diese Zwecke.

## B. Inhalte der vollen Masterarbeit

**Format:** GitHub Repository: Radmir Gesler Master Arbeit

### B.1. PDF-Dateien

- MA\_Gesler.pdf - gesamte Masterarbeit im PDF-Format

### B.2. Python Code

- Einzelne Implementierung findet man im Verzeichnis `/code/`

# Quellenverzeichnis

## Literatur

- [0] Klemens Burg, Herbert Haf und Friedrich Wille. *Funktionalanalysis und partielle Differentialgleichungen*. 1. Aufl. Teubner, 1991.
- [0] Thorsten M. Buzug. *Einführung in die Computertomographie*. 1. Aufl. Springer, 2004.
- [0] S. Handrock-Meyer. *Inverse Probleme für die Wärmeleitungsgleichung*. Zeitschrift für Analysis und ihre Anwendungen, S. 503-510, 1988.
- [0] Harro Heuser. *Funktionalanalysis*. 1. Aufl. B.G Teubner Stuttgart, 1975.
- [0] M. Zuhair Nashed und C. W. Groetsch. *A new approach to classification and regularization of ill-posed operator equations*. S. 53-75.
- [0] Lena Neufeld. „Randomisiertes Kaczmarz-Verfahren“. Bachelorarbeit. Münster: Westfälische Wilhelms-Universität, März 2015. URL: <https://www.uni-muenster.de/AMM/num/burger/teaching/bachelor.html>.
- [0] Andres Rieder. *Keine Probleme mit inversen Problemen*. 1. Aufl. Vieweg, 2003.
- [0] Ben Schweizer. *Partielle Differentialgleichungen*. 2. Aufl. Springer Spektrum, 2018.
- [0] Gengsheng Lawrence Zeng. *Medical Image Reconstruction*. 1. Aufl. Springer, 2009.

# Abbildungsverzeichnis

1.1.	Das Spektrum der elektromagnetischen Strahlung. Quelle: Wikipedia.	2
2.1.	Schematische Darstellung der Gleichung 2.1, mit $V$ das Volumen eines Mediums, $dV$ Rand (orange) von $V$ , $j$ Der Wärmestrom durch $V$ , $v \in \mathbb{R}^3$ Normalenvektor an $dV$ und $dS$ Tangente an $dV$ .	4
3.1.	Schematischer Aufbau einer Röntgenaufnahme im Draufsicht. Links ist das Schwächungsprofil nach Durchgang der Röntgenstrahlen durch ein Medium mit verschiedenen Dichten zu sehen (mit der Annahme, dass orangene Kreise höhere Dichte haben, als die graue Fläche).	19
3.2.	Das Prinzip der Bildrekonstruktion aus CT-Daten mit einer Rückprojektion (a) und zwei Rückprojektionen (b).	20
3.3.	Prinzipieller Aufbau eines CT-Scanners der dritten Generation.	21
3.4.	Das Schichtmodell eines Objekts als Grundlage zur dreidimensionalen Rekonstruktion.	21
3.5.	Intensitätsänderung der Röntgenstrahlung beim Passieren eines homogenen Objekts.	22
3.6.	Parametrisierung des Strahls $L$ auf dem Gebiet $\Omega$ .	24
3.7.	Ein Beispiel der Rückprojektion eines Projektionswerts $g_\theta(s_0)$ auf einen Punkt $x \in \Omega$ . Mit $g_\theta(s)$ ist hier eine Projektion bei einem festen Winkel $\theta$ bezeichnet, also $g_\theta(s) = g(s, \theta)$ .	29
3.8.	Anschauliche Darstellung der Diskretisierung der Dichtefunktion $f$ . Entlang des Röntgenstrahls (alle rot umrandeten Pixel) an der Stelle $s_i$ und zu einem festen Winkel $\theta_l = 0$ , werden die Werte von $f$ zu dem Wert $p(\tilde{s}_i, \theta_l)$ aufsummiert. $p(\tilde{s}, \theta_l)$ ist der Vektor zu allen Strahlen an den Stellen $\tilde{s}_i$ und somit die Diskretisierung von $g(s, \theta_l)$ .	36
3.9.	Links (a) ist ein Phantombild (Quelle: MATLAB, <code>phantom(300)</code> ) zu sehen, welches wir für alle Testzwecke in diesem Kapitel benutzen werden. Rechts (b) ist die Radontransformierte von (a) abgebildet. Die Transformation wurde über $\theta \in [0^\circ, 180^\circ]$ mit einer Schrittweite von $1^\circ$ durchgeführt. Das Detektorband hat eine Länge von $k = 300$ , wobei $k$ gleich der Anzahl der Tiefenpixel von (a) ist. Somit bestimmt $k$ auch die Anzahl der Strahlen, die durch das Phantombild geschickt worden sind.	38
3.10.	Links (a) ist die Radontransformierte wie in der Abb. 3.9b. Rechts (b) ist die ungefilterte Rückprojektion von (b).	39
3.11.	Eine Skizze (links) zur Verdeutlichung der Beziehung zweier Punkte $x, x_0 \in \Omega$ bezüglich ihrer kartesischen- und Polar-Koordinaten. Die dazugehörige Nebenrechnung (rechts) zeigt deren formelmäßigen Zusammenhang auf.	39

3.12. Eine Veranschaulichung des Fourier-Slice-Theorems. Links ist der Vorgang der Radon Transformation abgebildet, rechts seine Fouriertransformierte. . . . .	42
3.13. Links (a) das Rampenfilter, konstruiert passend zu dem Sinogramm aus 3.9b. Das Frequenzband des Filters ist durch die isometrische Eigenschaften der FT gleich der Detektorbreite, also ist auch gleich der Tiefe des Sinogramms. Rechts (b) ist die gefilterte Rückprojektion von 3.9b, durchgeführt unter der Einwirkung des Rampenfilters aus (a). . . . .	44
3.14. Links (a) ist ein Phantombild verrauscht mit einem Salzstreuer-Effekt (Quelle: MATLAB) dargestellt. Rechts (b) ist Die Radontransformierte von (a) und zusätzlich mit einem Salzstreuer-Effekt verrauscht. . . . .	45
3.15. Eine Rekonstruktion aus den Daten 3.14b unter der Einwirkung von drei verschiedenen Filtern. Unten links ist die Filterung mit dem <i>Ramachandran-Lakshminarayanan-Filter</i> , mittig mit dem <i>Shepp-Logan-Filter</i> und rechts mit dem <i>Hamming-Filter</i> dargestellt. . . . .	46
3.16. Eine Skizze zur Veranschaulichung des Aufbauprinzips der Abbildungsmatrix $A$ von $\mathcal{R}$ . . . . .	47
3.17. Eine Veranschaulichung der Arbeitsweise des Kaczmarz-Verfahrens in $\mathbb{R}^2$ . . . . .	49
3.18. Randomisierte iterative Rekonstruktionen (a) nicht verrauscht, (b) verrauscht. Die Projektionen in Beiden Fällen wurden für alle Winkel $\theta$ von $1^\circ$ bis $180^\circ$ mit der Schrittweite von $1^\circ$ berechnet. Detektorbreite ist $k = 300$ . . . . .	50
3.19. Ein Vergleich des randomisierten und vorwärts projektiven Kaczmarz-Verfahrens. Der Test wurde für die Rekonstruktion eines $50 \times 50$ Pixel großen Bildes durchgeführt. In jedem Testschritt wurde die Winkelanzahl erhöht, sodass die Systemmatrix immer größer wurde. . . . .	51
3.20. (a) Rekonstruktion durch SWZ, für ein Phantombild ( $50 \times 50$ ), Mit der Detektorbreite $k = 50$ und Winkelanzahl $q = 50$ . (b) die Konvergenz der Reihe (3.37) blau mit Regularisierung $\lambda_{min} = 5.5$ (das Minimum aus (a)), rot ohne Regularisierung. . . . .	53
3.21. Die Folge der Singulärwerte für $A$ aus 3.20a . . . . .	53
3.22. Die Rekonstruktion des Phantombildes mit der Größe $50 \times 50$ Pixel. Links nicht regularisiert $\lambda = 0$ , recht Regularisierung mit $\lambda = 5.5$ . . . . .	54



# Tabellenverzeichnis