



universität  
wien

# MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

**„Regularization by Orthogonalization and Frames“**

verfasst von / submitted by

Leon Frischauf BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of  
Master of Science (MSc)

Wien, August 2020 / Vienna, August 2020

Studienkennzahl lt. Studienblatt /  
degree programme code as it appears on  
the student record sheet:

A 066 821

Studienrichtung lt. Studienblatt /  
degree programme as it appears on  
the student record sheet:

Mathematik / Mathematics

Betreut von / Supervisor:

Univ.-Prof. Dr. Otmar Scherzer

## **Abstract**

In this thesis the method of regularization by projection is investigated, which allows to reconstruct the possible unbounded inverse of a linear operator using provided training data. The method is applied to the example of tomography. Different possible realizations of this approach using orthonormalization and frames are presented and investigated. Performance and errors of the variants are compared for different image data sets, which differ in size and number of images. The results show that the methods can be successfully applied for the reconstruction of images, but that there are limitations regarding the quality of the output, which result from the finiteness of the data (compared to convergence analysis), only approximately fulfilled model assumptions (linear independence of the image vectors) and numerical errors.

## Zusammenfassung

In dieser Arbeit wird die Methode der Regularisierung durch Projektion untersucht, die es erlaubt, die möglicherweise unbeschränkte Inverse eines linearen Operators über bereitgestellte Trainingsdaten zu rekonstruieren. Die Methode wird am Beispiel der Tomographie angewendet, wobei verschiedene mögliche Realisierungen dieses Ansatzes mittels Orthonormalisierung und Frames vorgestellt und untersucht werden. Performance und Fehler der Varianten werden für verschiedene Bilddatensätze, die sich in Größe und Anzahl der Bilder unterscheiden, verglichen. Die Ergebnisse zeigen, dass die Methoden erfolgreich für die Rekonstruktion von Bildern eingesetzt werden können, dass jedoch Grenzen hinsichtlich der Qualität der Ausgabe bestehen, die sich aus der Endlichkeit der Daten (im Vergleich zur Konvergenzanalyse), nur annähernd erfüllten Modellannahmen (lineare Unabhängigkeit der Bildvektoren) und numerischen Fehlern ergeben.

## **Acknowledgements**

My special thanks go to Dr. Andrea Aspri from RICAM, who partly supervised my thesis. His contributions are highly appreciated. Also without the continuous support of Univ.-Prof. Dr. Otmar Scherzer this work - especially during the COVID-19 crisis - would not have been possible.

Furthermore, I want to thank my family and my girlfriend Sophia for the ongoing support during my studies.

# Contents

<b>1</b>	<b>Introduction</b>	<b>5</b>
<b>2</b>	<b>Preliminaries</b>	<b>5</b>
2.1	Basic definitions . . . . .	5
2.1.1	The Moore-Penrose Generalized Inverse . . . . .	7
2.2	Introduction to spectral theory . . . . .	9
2.3	Introduction to frame theory . . . . .	18
<b>3</b>	<b>Regularization Operators</b>	<b>22</b>
3.1	Motivation . . . . .	22
3.2	Definition and properties . . . . .	23
3.3	Regularization by Projection . . . . .	25
<b>4</b>	<b>Regularization with training data</b>	<b>31</b>
<b>5</b>	<b>Simulations</b>	<b>32</b>
5.1	Algorithm . . . . .	32
5.1.1	Import of training data . . . . .	33
5.1.2	Gram Schmidt method . . . . .	33
5.1.3	Modified Gram Schmidt method . . . . .	34
5.1.4	Householder reflections . . . . .	34
5.1.5	QR MATLAB native . . . . .	35
5.1.6	Givens rotations . . . . .	36
5.1.7	Arnoldi iteration . . . . .	36
5.1.8	Backtransformation of test data . . . . .	36
5.1.9	Inverse Radon transformation . . . . .	36
5.2	Reconstruction via frames . . . . .	36
5.3	Reconstruction via frame iteration . . . . .	37
5.3.1	Numerical effort . . . . .	38
5.4	Numerical investigations . . . . .	39
5.4.1	Accuracy of the orthogonalization methods . . . . .	39
5.4.2	Visual observations . . . . .	42
<b>6</b>	<b>Summary and conclusion</b>	<b>52</b>

# 1 Introduction

The main idea of regularization operators is the solving of the linear equation  $Tx = y^\delta$  for  $x$  with bounded operator  $T$  and possibly noisy right-hand side  $y^\delta$ . Since the solution of the least-squared problem  $x^\dagger = T^\dagger y^\delta$ , which can in general be expressed by the (possibly unbounded) Moore-Penrose inverse  $T^\dagger$ , does not depend continuously on  $x$ , small perturbations of  $y^\delta$  can have a rather undesired impact on the solution  $x^\dagger$ . In even worse cases, the perturbed data is not even in the domain of  $T^\dagger$ . Various approaches can be used to overcome this issues. In the following thesis we present the *regularization by projection* approach with application to the Radon transform in image processing.

The next following Section 2 summarizes the main results of spectral and frame theory, which will be used in the applications afterwards. Section 3 appeals to the problem of regularization in general and present possible approaches and mathematical facts concerning this topic. In Section 4 we introduce the usage of training data in the regularization approach, which we finally use in Section 5 for our numerical experiments with image data sets. We summarize and conclude the results of the thesis in Section 6.

## 2 Preliminaries

### 2.1 Basic definitions

We start with the well-known Cauchy-Schwarz inequality (see [3, 18.1]):

**Lemma 2.1** (Cauchy-Schwarz inequality). *Let  $\mathcal{X}$  be a linear space with inner product and  $x, y \in \mathcal{X}$ , then*

$$|\langle x, y \rangle| \leq \sqrt{|\langle x, x \rangle|} \cdot \sqrt{|\langle y, y \rangle|} = \|x\| \cdot \|y\|.$$

*This inequality is an equality if and only if  $x, y$  are linearly dependent.*

Given a linear operator  $A$  between two Hilbert spaces, we recall the definition of its **adjoint operator**  $A^*$  as in [3, Chapter 58]:

**Definition 2.2** (Adjoint operator). *Let  $\mathcal{X}$  and  $\mathcal{Y}$  be Hilbert spaces and  $A : \mathcal{X} \rightarrow \mathcal{Y}$  a linear and bounded operator. Then there exists a linear and continuous operator  $A^* : \mathcal{Y} \rightarrow \mathcal{X}$ , which is uniquely defined by the equality*

$$\langle Ax, y \rangle = \langle x, A^*y \rangle \quad \forall x \in \mathcal{X}, y \in \mathcal{Y},$$

*which is called the **adjoint operator**.*

In [3, Theorem 58.1] some properties of the adjoint operator are given:

**Lemma 2.3** (Properties of adjoint operators). *Let  $\mathcal{X}, \mathcal{Y}$  and  $\mathcal{Z}$  be Hilbert spaces and  $A : \mathcal{X} \rightarrow \mathcal{Y}$  and  $B : \mathcal{Y} \rightarrow \mathcal{Z}$  be two linear and continuous operators. Then their adjoint operators  $A^* : \mathcal{Y} \rightarrow \mathcal{X}$  and  $B^* : \mathcal{Z} \rightarrow \mathcal{Y}$  fulfill:*

$$(i) \quad (BA)^* = A^*B^*$$

$$(ii) \quad (A^*)^* = A$$

*Proof.* (i). For  $x \in \mathcal{X}$  and  $z \in \mathcal{Z}$ , we apply the definition of the adjoint operator twice:

$$\langle BAx, z \rangle = \langle Ax, B^*z \rangle = \langle x, A^*B^*z \rangle \quad (1)$$

So the adjoint operator of  $BA$  is equal to  $A^*B^*$ .

(ii). For  $x \in \mathcal{X}$  and  $y \in \mathcal{Y}$  the defining property of the adjoint operator applied to  $A^*$  leads to:

$$\langle A^*y, x \rangle = \langle y, (A^*)^*x \rangle \quad (2)$$

Since the inner product is hermitian, we get on the other hand:

$$\langle A^*y, x \rangle = \overline{\langle x, A^*y \rangle} = \overline{\langle Ax, y \rangle} = \langle y, Ax \rangle \quad (3)$$

Comparison of the results together with the fact that the adjoint operator is uniquely defined by the equality in Definition 2.2, yields  $(A^*)^* = A$ .  $\square$

**Definition 2.4** (Self-adjoint). *If the target set of  $A$  is equal to its domain, i.e.  $A : \mathcal{X} \rightarrow \mathcal{X}$ , the operator is called **self-adjoint**, if  $A^* = A$ .*

We now define some concepts concerning the **eigenvalues** of an operator (see e.g. [3, Chapter 29]):

**Definition 2.5.** *For a linear operator  $T : \mathcal{X} \rightarrow \mathcal{Y}$ , we call the scalar  $\lambda \neq 0$  an **eigenvalue** of  $T$ , if there exists  $x \in \mathcal{X} \setminus \{0\}$  with  $Tx = \lambda x$ .  $x$  is called **eigenvector** of  $T$ . Since for all eigenvalues  $\lambda$ , the operator  $(\lambda I - T)$  is not injective (otherwise it would not be possible to find an  $x \neq 0$  with  $(\lambda I - T)x = 0 \Leftrightarrow Tx = \lambda x$ ) and the dimension of  $\mathcal{N}(\lambda I - T) > 0$  is called **multiplicity** of  $\lambda$ .*

Self-adjoint linear operators have special properties regarding their eigenvectors (see [3, Theorem 29.9]):

**Lemma 2.6.** *Eigenvalues of self-adjoint operators are always real and eigenvectors to different eigenvalues are orthogonal.*

This allows to define the following term:

**Definition 2.7** (Eigensystem). *For a self-adjoint operator  $A : \mathcal{X} \rightarrow \mathcal{X}$  we define the **eigensystem** as the set of all pairs  $(\lambda_i, x_i)$ , where  $\lambda_i \in \mathbb{R}$  are eigenvalues of  $A$  and  $x_i$  are the corresponding orthonormal eigenvectors of  $A$ .*

Note that due to the multiplicity of eigenvalues, there may exist more than one pair corresponding elements in the eigensystem with different (orthonormalized) eigenvectors. Furthermore, the family of pairs  $(\lambda_i, x_i)$  may be uncountable. This circumstance does not appear, if the operator  $A$  is additionally compact, which we will refer to in the next subsection. In the following subsection we present the concept of the **Moore-Penrose generalized inverse** of an operator and some of its properties. They correspond to the results of [1, Sec. 2.1] and [16, Chapter III].

### 2.1.1 The Moore-Penrose Generalized Inverse

We start with a definition about the context in which the Moore-Penrose generalized inverse is used: least-squares solution of linear equations.

**Definition 2.8.** *For a bounded linear operator  $T : \mathcal{X} \rightarrow \mathcal{Y}$  and  $y \in \mathcal{Y}$ , we call*

(i)  $x \in \mathcal{X}$  is called a **least-squares solution** of  $Tx = y$  if

$$\|Tx - y\| = \inf \{ \|Tz - y\| \mid z \in \mathcal{X} \}.$$

(ii)  $x \in \mathcal{X}$  is called a **best-approximate solution** of  $Tx = y$  if  $x$  is least-squares solution and additionally

$$\|x\| = \inf \{ \|z\| \mid z \text{ is a least-squares solution of } Tx = y \}$$

holds.

The following theorem defines the Moore-Penrose generalized inverse of the operator  $T$ :

**Theorem 2.9.** *Let  $T : \mathcal{X} \rightarrow \mathcal{Y}$  be linear. We define the restricted operator*

$$\tilde{T} := T|_{\mathcal{N}(T)^\perp} : \mathcal{N}(T)^\perp \rightarrow \mathcal{R}(T).$$

Since  $\mathcal{N}(\tilde{T}) = \{0\}$  and  $\mathcal{R}(\tilde{T}) = \mathcal{R}(T)$ ,  $\tilde{T}$  is injective and has an inverse  $\tilde{T}^{-1} : \mathcal{R}(T) \rightarrow \mathcal{N}(T)^\perp$ . The unique linear extension  $T^\dagger$  of this inverse to  $\mathcal{D}(T^\dagger) := \mathcal{R}(T) \dot{+} \mathcal{R}(T)^\perp$  is called **Moore-Penrose generalized inverse**. Furthermore for all  $y \in \mathcal{D}(T^\dagger)$  the result  $x^\dagger := T^\dagger y$  is the best-approximate solution of Definition 2.8 and all least-squares solutions are  $x^\dagger + \mathcal{N}(T)$ .

The following proposition characterises  $T^\dagger$  uniquely:

**Proposition 2.10.** *Let  $P$  and  $Q$  be the orthogonal projectors onto  $\mathcal{N}(T)$  and  $\overline{\mathcal{R}(T)}$ , respectively. Then  $T^\dagger : \mathcal{D}(T^\dagger) \rightarrow \mathcal{X}$  satisfies:*

- (i)  $\mathcal{R}(T^\dagger) = \mathcal{N}(T)^\perp$ ,
- (ii)  $TT^\dagger T = T$ ,
- (iii)  $T^\dagger TT^\dagger = T^\dagger$ ,
- (iv)  $T^\dagger T = I - P$ ,
- (v)  $TT^\dagger = Q|_{\mathcal{D}(T^\dagger)}$ .

The concept of the square root of a positive operator is given by the following lemma in [3, page 552]:

**Lemma 2.11** (Square root of operator). *Let  $A : \mathcal{X} \rightarrow \mathcal{X}$  be a positive operator, i.e.  $A$  is self-adjoint and it holds  $\langle Ax, x \rangle \geq 0$  for all  $x \in \mathcal{X}$ . Then there exists an operator  $B : \mathcal{X} \rightarrow \mathcal{X}$ , which fulfills  $B^2 = A$ . This operator is called the **square root** of  $A$  and is also written as  $B = \sqrt{A} = A^{\frac{1}{2}}$ .*

The next definition describes compact linear operators, a subset of all continuous operators:

**Definition 2.12** (Compact linear operator). *A linear operator  $K : \mathcal{X} \rightarrow \mathcal{Y}$  between two normed spaces  $\mathcal{X}$  and  $\mathcal{Y}$  is called a **compact** linear operator if for every bounded sequence  $(x_n)_{n \geq 1}$  in  $\mathcal{X}$ , the sequence  $(Kx_n)_{n \geq 1}$  has a convergent subsequence.*

We present an example of a compact linear operator (see [1, Sec. 2.2]):

*Example 2.13.* Let  $\Omega \subseteq \mathbb{R}^n$  and  $k \in \mathcal{L}^2(\Omega \times \Omega)$ . The integral operator  $K : \mathcal{L}^2(\Omega) \rightarrow \mathcal{L}^2(\Omega)$  with kernel  $k$  is defined by:

$$Kx(s) = \int_{\Omega} k(s, t)x(t) dt \quad (4)$$

Under the following two requirements, the integral operator  $K$  is compact:

- (i)  $\Omega$  is a compact set.
- (ii) There exist  $M > 0$  and  $\epsilon > 0$  so that

$$|k(s, t)| \leq \frac{M}{|s - t|^{n-\epsilon}} \quad (5)$$

holds for all  $s \neq t \in \Omega$ .

With the assumption that (i) and (ii) hold, integral operator  $K$  is compact iff  $k$  can be written as

$$k(s, t) = \sum_{i=1}^N \varphi_i(s)\psi_i(t), \quad s, t \in \Omega \quad (6)$$

with  $\varphi_i, \psi_i \in \mathcal{L}^2(\Omega)$  and  $N \in \mathbb{N}$ .

## 2.2 Introduction to spectral theory

Most of the definitions and results of this section can be found in [1, chapters 2.2 and 2.3] (unless otherwise indicated).

**Theorem 2.14.** *All self-adjoint compact linear operators  $K$  with eigensystem  $(\lambda_i, v_i)$  have the representation*

$$Kx = \sum_{i=1}^{\infty} \lambda_i \langle x, v_i \rangle v_i. \quad (7)$$

*In particular, compactness implies that the eigensystem defined in Def. 2.7 is countable.*

If the compact linear operator is not self-adjoint, the previous representation does not necessarily exist. Nevertheless, a decomposition with the following defined singular system is possible.

**Lemma 2.15.** *For any bounded linear operator  $K$  and its adjoint operator  $K^*$ , the compositions  $KK^* : \mathcal{Y} \rightarrow \mathcal{Y}$  and  $K^*K : \mathcal{X} \rightarrow \mathcal{X}$  are both self-adjoint operators.*

*Proof.* According to Lemma 2.3, it holds

$$\begin{aligned}(KK^*)^* &= (K^*)^*K^* = KK^*, \\ (K^*K)^* &= K^*(K^*)^* = K^*K.\end{aligned}$$

□

**Lemma 2.16.** *The self-adjoint operators  $KK^*$  and  $K^*K$  have the same non-negative eigenvalues  $\{\sigma_i^2\}_{i \in \mathbb{N}}$ . For the corresponding eigensystem  $\{(\sigma_i^2, v_i)\}_{i \in \mathbb{N}}$  of  $K^*K$ , the system  $\{(\sigma_i^2, u_i)\}_{i \in \mathbb{N}}$  with  $u_i := \frac{Kv_i}{\|Kv_i\|}$  is an eigensystem of  $KK^*$ . This means that  $\{\sigma_i^2\}_{i \in \mathbb{N}}$  are spectral values of those operators and belong to the point spectrum (see [3, Chapter 96]).*

**Definition 2.17** (Singular system). *For a compact linear operator  $K$ , let  $u_i, v_i$  the eigenvectors and  $\sigma_i^2$  eigenvalues of  $K^*K$  and  $KK^*$ , respectively according to Lemma 2.16. Further  $\sigma_i > 0$  denote the positive square roots of the values  $\sigma_i^2$  indicated in decreasing order with their multiplicity. The values  $\sigma_i$  are called **singular values** of the operator  $K$ . We denote furthermore  $(\sigma_i, u_i, v_i)_{i \in \mathbb{N}}$  as the **singular system** of  $K$ .*

In [5, Section 21.4], the **singular value decomposition** for the finite dimensional case of a  $n \times m$  matrix (which can be interpreted as a specific coordinate representation of a linear operator between two finite dimensional vector spaces) is stated:

**Lemma 2.18** (Singular value decomposition). *For each matrix  $A \in \mathbb{R}^{n \times m}$ , there exist two orthogonal matrices  $U \in \mathbb{R}^{n \times n}$  and  $V \in \mathbb{R}^{m \times m}$  such that*

$$A = U\Sigma V^T, \tag{8}$$

where  $\Sigma := \text{diag}(\sigma_1, \dots, \sigma_r) \in \mathbb{R}^{n \times m}$  is the generalized diagonal matrix of all  $r$  singular values of the matrix  $A$ .

The next lemma is some kind of a generalization of the previous lemma:

**Lemma 2.19.** *For a compact linear operator  $K$  and its singular system  $(\sigma_i, u_i, v_i)_{i \in \mathbb{N}}$ , the following equations hold:*

$$Kv_i = \sigma_i u_i \tag{9}$$

$$K^*u_i = \sigma_i v_i \tag{10}$$

$$Kx = \sum_{i=1}^{\infty} \sigma_i \langle x, v_i \rangle u_i, \quad x \in \mathcal{X} \tag{11}$$

$$K^*y = \sum_{i=1}^{\infty} \sigma_i \langle y, u_i \rangle v_i, \quad y \in \mathcal{Y} \tag{12}$$

**Remark 2.20.** Equations (11) and (12) are called *singular value expansions* and the infinite dimensional analogues of the singular value decomposition in Definition 2.18.

The basis for the following lemmata is the following well-known result about finite dimensional spaces, which can be found in [3, Theorem 11.4]:

**Theorem 2.21.** *Every finite dimensional subspace of a normed space is closed.*

This implies the lemmata (see [1, Sec. 2.2]):

**Lemma 2.22.** *If the compact linear operator  $K$  has finite dimensional range, only finitely many singular values exist. If this is not the case, then  $\lim_{i \in \mathbb{N}} \sigma_i$  is zero.*

**Lemma 2.23.** *For the compact linear operator  $K$ , the range  $\mathcal{R}(K)$  is closed iff it is finite dimensional.*

*Proof.* If  $\mathcal{R}(K)$  is finite dimensional, it is closed according to Theorem 2.21.

If  $\mathcal{R}(K)$  is closed, then it is also complete and according to Banach's open mapping theorem, the restricted operator on the orthogonal complement of the nullspace  $\mathcal{N}(K)^\perp$ ,  $K|_{\mathcal{N}(K)^\perp} : \mathcal{N}(K)^\perp \rightarrow \mathcal{R}(K)$ , is continuously invertible. Since the composition  $K \circ \left( K|_{\mathcal{N}(K)^\perp} \right)^{-1} = I_{\mathcal{R}(K)}$  is compact, the dimension of  $\mathcal{R}(K)$  is finite. □

**Proposition 2.24.** *If the compact linear operator  $K$  has infinite dimensional range  $\mathcal{R}(K)$ , the Moore-Penrose generalized inverse  $K^\dagger$  is a densely defined unbounded linear operator with closed graph.*

Given  $y \in \mathcal{Y}$ , consider the linear equation  $Kx = y$ , where  $x$  is the unknown. Then, for a compact linear operator  $K$  with non-closed range, the best approximation  $x^* = K^\dagger y$  does not depend continuously on the right hand side  $y$ .

The following theorem shows some basic properties of the Moore-Penrose inverse. The whole domain of  $K^\dagger$  is denoted by  $\mathcal{D}(K^\dagger)$ .

**Theorem 2.25.** *For the singular system  $(\sigma_i, u_i, v_i)$  of a compact linear operator  $K$  holds:*

1.  $y \in \mathcal{D}(K^\dagger) \iff \sum_{i=1}^{\infty} \frac{|\langle y, u_i \rangle|^2}{\sigma_i^2} < \infty$

2. For every  $y \in \mathcal{D}(K^\dagger)$ , the operator  $K^\dagger$  can be written as

$$K^\dagger y = \sum_{i=1}^{\infty} \frac{\langle y, u_i \rangle}{\sigma_i} v_i.$$

The following definitions are motivated by the aim of applying functions to self-adjoint operators. First we define the following orthogonal projector  $E_\lambda$ , which projects for every  $\lambda > 0$  onto the set  $\mathcal{E}_\lambda := \text{span} \{v_n | n \in \mathbb{N}, \sigma_n^2 < \lambda\} + \mathcal{N}(K^*K)$ , for every  $x \in \mathcal{X}$  as

$$E_\lambda x := \begin{cases} \sum_{\substack{i=1 \\ \sigma_i^2 < \lambda}}^{\infty} \langle x, v_i \rangle v_i + Px & \text{if } \lambda > 0, \\ 0 & \text{if } \lambda \leq 0. \end{cases} \quad (13)$$

Here,  $P$  denotes the orthogonal projection onto  $\mathcal{N}(K^*K)$ . The following two Lemmata contain some properties of this operator:

**Lemma 2.26.** *For  $\lambda > \sigma_1^2$  the image set of the projector  $E_\lambda$  fulfills*

$$\mathcal{E}_\lambda = \overline{\mathcal{R}(K^*K)} + \mathcal{N}(K^*K),$$

*since the set  $\{v_i\}_{i \in \mathbb{N}}$  spans  $\mathcal{R}(K^*K)$  because of the representation in (12).*

**Lemma 2.27** (monotonicity property). *For all  $\lambda \leq \mu$ , it holds*

$$\langle E_\lambda x, x \rangle \leq \langle E_\mu x, x \rangle \quad \forall x \in \mathcal{X}. \quad (14)$$

*Proof.* The only relevant case is  $0 < \lambda \leq \mu$ , otherwise  $E_\lambda = 0$  and the claim is fulfilled trivially. To see this let

$$\langle E_\lambda x, x \rangle = \sum_{\substack{i=1 \\ \sigma_i^2 < \lambda}}^{\infty} |\langle x, v_i \rangle|^2 + \|Px\|^2 \leq \sum_{\substack{i=1 \\ \sigma_i^2 < \mu}}^{\infty} |\langle x, v_i \rangle|^2 + \|Px\|^2 = \langle E_\mu x, x \rangle. \quad (15)$$

□

With these projections, we define the integration of a (piecewise) continuous function  $f$  over the so-called **spectral family**  $\{E_\lambda\}_{\lambda \in \mathbb{R}}$  for a singular system  $(\sigma_i, u_i, v_i)_{i \in \mathbb{N}}$  as (integral bounds are always  $-\infty$  and  $+\infty$ )

$$\int f(\lambda) dE_\lambda x := \sum_{i=1}^{\infty} f(\sigma_i^2) \langle x, v_i \rangle v_i, \quad (16)$$

$$\int f(\lambda) d\langle E_\lambda x, y \rangle := \sum_{i=1}^{\infty} f(\sigma_i^2) \langle x, v_i \rangle \langle y, v_i \rangle, \quad (17)$$

$$\int f(\lambda) d\|E_\lambda x\|^2 := \sum_{i=1}^{\infty} f(\sigma_i^2) |\langle x, v_i \rangle|^2. \quad (18)$$

for all  $x, y \in \mathcal{X}$ .

In the special case for the identity function  $f = \text{id}$ , we observe in (16) that

$$\int \lambda \, dE_\lambda x = \sum_{i=1}^{\infty} \sigma_i^2 \langle x, v_i \rangle v_i \quad (19)$$

and Theorem 2.14 together with Definition 2.17 show that this is furthermore equal to  $K^*Kx$ . With this summary we are finally able to define the application of functions to arguments  $K^*K$  in the form of the following definition:

$$f(K^*K) := \int f(\lambda) \, dE_\lambda := \sum_{i=1}^{\infty} f(\sigma_i^2) \langle \cdot, v_i \rangle v_i \quad (20)$$

Analogously, the previous steps can be done for the self-adjoint and compact operator  $KK^*$  instead of  $K^*K$ , what leads to the spectral family  $\{\bar{E}_\lambda\}_{\lambda \in \mathbb{R}}$  of the form

$$\bar{E}_\lambda y := \sum_{\substack{i=1 \\ \sigma_i < \lambda}}^{\infty} \langle y, u_i \rangle u_i + (I - Q)y \quad (21)$$

for  $\lambda > 0$  with  $(I - Q)$  denoting the orthogonal projector onto  $\mathcal{N}(KK^*)$ . Finally, (17) and (18) imply the properties

$$\langle f(K^*K)x, y \rangle = \int f(\lambda) \, d \langle E_\lambda x, y \rangle, \quad (22)$$

$$\|f(K^*K)x\|^2 = \int f^2(\lambda) \, d \|E_\lambda x\|^2. \quad (23)$$

In the following we skip the assumption of compactness and turn to the case of general linear self-adjoint operators  $A$  in a Hilbert space  $\mathcal{X}$ . These results can be found in [2, Chapter 3].

**Definition 2.28.** A family  $\{E_\lambda\}_{\lambda \in \mathbb{R}}$  orthogonal projections in  $\mathcal{X}$  is called a **spectral family** or else a **resolution of the identity** if it satisfies these conditions:

- (i)  $E_\lambda E_\mu = E_{\min\{\lambda, \mu\}}, \quad \lambda, \mu \in \mathbb{R},$
- (ii)  $E_{-\infty} = 0, E_{+\infty} = I,$  where  $E_{\pm\infty}x = \lim_{\lambda \rightarrow \pm\infty} E_\lambda x$  for all  $x \in \mathcal{X},$
- (iii)  $E_{\lambda+0} = E_\lambda,$  where  $E_{\lambda+0}x = \lim_{\varepsilon \rightarrow 0^+} E_{\lambda+\varepsilon}x$  for all  $x \in \mathcal{X}.$

In the following we state some properties of the **spectral family**. We start with a lemma.

**Lemma 2.29.** *The operator defined by  $E_{(\alpha,\beta]} := E_\beta - E_\alpha$  is also an orthogonal projector, i.e. the following two properties hold:*

$$(i) \quad E_{(\alpha,\beta]}^2 = E_{(\alpha,\beta]},$$

$$(ii) \quad E_{(\alpha,\beta]}^* = E_{(\alpha,\beta]}.$$

*Proof.* We show the two properties by using Definition 2.28 and the fact that  $E_\alpha$  and  $E_\beta$  are both orthogonal projections:

$$(i) \quad \underbrace{E_{(\alpha,\beta]} E_{(\alpha,\beta]}}_{\checkmark} = E_\beta - E_\alpha E_\beta - E_\beta E_\alpha + E_\alpha = E_\beta - E_\alpha - E_\alpha + E_\alpha = E_\beta - E_\alpha = E_{(\alpha,\beta]}$$

$$(ii) \quad E_{(\alpha,\beta]}^* = E_\beta^* - E_\alpha^* = E_\beta - E_\alpha = E_{(\alpha,\beta]} \checkmark$$

□

**Proposition 2.30.** *For a spectral family  $\{E_\lambda\}_{\lambda \in \mathbb{R}}$  the function*

$$\lambda \longmapsto \langle E_\lambda x, y \rangle \tag{24}$$

*is a function of bounded variation on every finite interval, with total variation  $V(\lambda; x, y)$  satisfying*

$$V(\lambda; x, y) \leq |x| \cdot |y|, \quad \forall x, y \in \mathcal{X}, \quad \lambda \in \mathbb{R}. \tag{25}$$

*Proof.* We numerate  $\lambda_1 < \lambda_2 < \dots < \lambda_n$  and use the orthogonal projection  $E_{(\alpha,\beta]}$  of Lemma 2.29. Applying the Cauchy-Schwarz inequality of Lemma 2.1 twice, we get

$$\sum_{j=2}^n |(E_{(\lambda_{j-1}, \lambda_j]} x, y)| = \sum_{j=2}^n |(E_{(\lambda_{j-1}, \lambda_j]} x, E_{(\lambda_{j-1}, \lambda_j]} y)| \tag{26}$$

$$\leq \sum_{j=2}^n |E_{(\lambda_{j-1}, \lambda_j]} x| \cdot |E_{(\lambda_{j-1}, \lambda_j]} y| \tag{27}$$

$$\leq \left( \sum_{j=2}^n |E_{(\lambda_{j-1}, \lambda_j]} x|^2 \right)^{1/2} \cdot \left( \sum_{j=2}^n |E_{(\lambda_{j-1}, \lambda_j]} y|^2 \right)^{1/2} \tag{28}$$

$$= \left( |E_{(\lambda_1, \lambda_n]} x|^2 \right)^{1/2} \cdot \left( |E_{(\lambda_1, \lambda_n]} y|^2 \right)^{1/2} \leq |x| \cdot |y|, \tag{29}$$

where the last line is implied by the property  $E_{(\lambda_{j-1}, \lambda_j]} \cdot E_{(\lambda_{i-1}, \lambda_i]} = 0, i \neq j$  and so (see [2, Eq. (3.12)]) for  $m > n$

$$|E_{(\lambda_n, \lambda_m]}x|^2 = \left| \sum_{i=n}^{m-1} E_{(\lambda_i, \lambda_{i+1}]}x \right|^2 = \sum_{i=n}^{m-1} |E_{(\lambda_i, \lambda_{i+1}]}x|^2. \quad (30)$$

□

**Corollary 2.31.** *Let  $\{E_\lambda\}_{\lambda \in \mathbb{R}}$  be a spectral family. Then, for all  $\lambda \in \mathbb{R}$ , there exist the operators*

$$E_{\lambda_+} = \lim_{\mu \rightarrow +\lambda} E_\mu, \quad E_{\lambda_-} = \lim_{\mu \rightarrow -\lambda} E_\mu.$$

*Proof.* Application of (30) implies that if  $\lambda \rightarrow_+ \lambda$ , then

$$\lim_{j,k \rightarrow \infty} |E_{(\lambda_j, \lambda_k]}x|^2 = 0. \quad (31)$$

For the operator  $E_{\lambda+0}$ , one can directly apply property (iii) of Definition 2.28. □

**Proposition 2.32.** *Let  $f$  be a continuous function on  $\mathbb{R}$  with complex function values and  $x \in \mathcal{X}$ . Then for  $\alpha < \beta, [\alpha, \beta] \subset \mathbb{R}$  it is possible to define the integral*

$$\int_{\alpha}^{\beta} f(\lambda) dE_{\lambda}x \quad (32)$$

*as the strong limit in  $\mathcal{X}$  of the Riemann sum:*

$$\sum_i f(\lambda'_i) E_{(\lambda_i, \lambda_{i+1}]}x \text{ where } \alpha = \lambda_1 < \lambda_2 < \dots < \lambda_n = \beta \text{ and } \lambda'_i \in (\lambda_i, \lambda_{i+1}), \quad (33)$$

*where the fineness of the Riemann sum  $\max_i |\lambda_{i+1} - \lambda_i|$  converges to zero.*

*Proof.* On every compact interval  $[\alpha, \beta]$ ,  $f$  is uniformly continuous. That means that for every  $\epsilon > 0$  exists  $\delta > 0$  so that for all  $\lambda, \lambda' \in [a, b]$  the implication

$$|\lambda - \lambda'| < \delta \Rightarrow |f(\lambda) - f(\lambda')| < \epsilon \quad (34)$$

holds. We consider two partitions of the interval  $[\alpha, \beta]$  with fineness  $< \delta$ , i.e.

$$\alpha = \lambda_1 < \lambda_2 < \dots < \lambda_n = \beta, \quad \max_i |\lambda_{i+1} - \lambda_i| < \delta, \quad (35)$$

$$\alpha = \mu_1 < \mu_2 < \dots < \mu_m = \beta, \quad \max_j |\mu_{j+1} - \mu_j| < \delta. \quad (36)$$

We denote the partition resulting from the union of  $\{\lambda_i\}_{i=1,\dots,n}$  and  $\{\mu_j\}_{j=1,\dots,m}$  by  $\{\nu_s\}_{s=1,\dots,p}$  with  $p \leq m+n$ . For  $\mu'_j \in (\mu_j, \mu_{j+1}]$ , we get with the uniform continuity of  $f$  the equality

$$\sum_i f(\lambda'_i) E_{(\lambda_i, \lambda_{i+1}]} x - \sum_j f(\mu'_j) E_{(\mu_j, \mu_{j+1}]} x = \sum_s \varepsilon_s E_{(v_s, v_{s+1}]} x, \text{ with } |\varepsilon_s| \leq 2\varepsilon, \quad (37)$$

where  $\varepsilon_s$  denotes the difference of the function values  $f(\lambda'_i)$  and  $f(\mu'_j)$  on the specific interval of the resulting partition  $\{\nu_s\}_{s=1,\dots,p}$ . This result can finally be estimated in the following

$$\left| \sum_s \varepsilon_s E_{(v_s, v_{s+1}]} x \right|^2 \leq \varepsilon^2 \left| \sum_s E_{(v_s, v_{s+1}]} x \right|^2 = \varepsilon^2 |E_{(\alpha, \beta]} x|^2 \leq \varepsilon^2 |x|^2.$$

This means that the difference in (37) converges to zero for  $\varepsilon \rightarrow 0$  and the sequence is Cauchy and therefore convergent. The existing limit can therefore be defined as the integral in (32).  $\square$

It is also possible to define improper integrals.

**Definition 2.33.** *For any continuous real function  $f$  and  $x \in \mathcal{X}$ , the improper integral*

$$\int_{-\infty}^{+\infty} f(\lambda) dE_\lambda x \quad (38)$$

*is defined as the (strong) limit*

$$\lim_{\substack{\alpha \rightarrow -\infty \\ \beta \rightarrow +\infty}} \int_\alpha^\beta f(\lambda) dE_\lambda x, \quad (39)$$

*in the case that this limit exists in  $\mathcal{X}$ .*

The existence of improper integrals is characterized by the following theorem. We skip the proof, which is given in [2, page 115 f.].

**Theorem 2.34.** *In the context of the previous definition, the following three conditions are equivalent:*

(i)  $\int_{-\infty}^{+\infty} f(\lambda) dE_\lambda x$  exists ,

$$(ii) \int_{-\infty}^{+\infty} |f(\lambda)|^2 d|E_\lambda x|^2 < +\infty,$$

$$(iii) y \mapsto F(y) := \int_{-\infty}^{+\infty} f(\lambda) d\langle E_\lambda y, x \rangle \quad \text{and } F \text{ is a linear functional.}$$

The next theorem allows to define a self-adjoint operator to any real-valued continuous function. Again the proof is given in [2, page 117].

**Theorem 2.35.** *Let  $f$  be an arbitrary real-valued continuous function,  $\mathcal{D}(A) \subseteq \mathcal{X}$  a domain defined by*

$$\mathcal{D}(A) = \left\{ x \in \mathcal{X} \mid \int_{-\infty}^{+\infty} f(\lambda)^2 d|E_\lambda x|^2 < +\infty \right\} \quad (40)$$

and  $A : \mathcal{D}(A) \longrightarrow \mathcal{X}$  the corresponding self-adjoint operator defined by

$$\langle Ax, y \rangle = \int_{-\infty}^{+\infty} f(\lambda) d\langle E_\lambda x, y \rangle, \quad \forall x \in \mathcal{D}(A), y \in \mathcal{X}. \quad (41)$$

Then  $\mathcal{D}(A)$  is a dense set in  $\mathcal{X}$  and this operator  $A$  is self-adjoint and fulfills  $AE_\lambda \supset E_\lambda A$  in the sense that  $AE_\lambda$  is an extension of  $E_\lambda A$  with domain  $\mathcal{X}$  instead of  $\mathcal{D}(A)$ .

A special case of the previous theorem leads to the **spectral representation** of the self-adjoint operator  $A$  and is given by the following:

**Corollary 2.36.** *We consider the special case of Theorem 2.35 where  $f(\lambda) = \lambda$ . Then we have*

$$\langle Ax, y \rangle = \int_{-\infty}^{+\infty} \lambda d\langle E_\lambda x, y \rangle, \quad x \in \mathcal{D}(A) \subset \mathcal{X}, \quad y \in \mathcal{X}, \quad (42)$$

$$\mathcal{D}(A) = \left\{ x \in \mathcal{X} \mid \int_{-\infty}^{+\infty} \lambda^2 d|E_\lambda x|^2 < +\infty \right\}. \quad (43)$$

In this case we formally write

$$A = \int_{-\infty}^{+\infty} \lambda dE_\lambda, \quad (44)$$

which is called the **spectral representation** of the self-adjoint operator  $A$  in  $\mathcal{X}$ .

**Corollary 2.37.** *For  $A = \int_{-\infty}^{+\infty} f(\lambda) dE_\lambda$ , it holds*

$$\|Ax\|^2 = \int_{-\infty}^{+\infty} f(\lambda)^2 d|E_\lambda x|^2, \quad x \in \mathcal{D}(A). \quad (45)$$

*Proof.* According to Theorem 2.35, it holds  $E_\lambda Ax = AE_\lambda x$  for all  $x \in \mathcal{X}$ . Using this, we have

$$\|Ax\|^2 = \langle Ax, Ax \rangle = \int_{\mathbb{R}} f(\lambda) \, d \langle E_\lambda x, Ax \rangle = \int_{\mathbb{R}} f(\lambda) \, d \langle AE_\lambda x, x \rangle \quad (46)$$

$$= \int_{\mathbb{R}} f(\lambda) \, d_\lambda \left( \int_{\mathbb{R}} f(\mu) \, d_\mu \langle E_\mu E_\lambda x, x \rangle \right) \quad (47)$$

$$\stackrel{\text{Def. 2.28}(i)}{=} \int_{\mathbb{R}} f(\lambda) \, d_\lambda \left\langle \int_{-\infty}^{\lambda} f(\mu) \, d_\mu (E_\mu x, x) \right\rangle \quad (48)$$

$$\stackrel{\text{Def. 2.28}(ii)}{=} \int_{\mathbb{R}} |f(\lambda)|^2 \, d |E_\lambda x|^2. \quad (49)$$

□

Finally we note the following proposition about the ranges of operators between two Hilbert spaces  $\mathcal{X}$  and  $\mathcal{Y}$ , which is proven in [1, Proposition 2.18]:

**Proposition 2.38.** *Let  $T : \mathcal{X} \rightarrow \mathcal{Y}$  be a linear bounded operator. Then*

$$\mathcal{R}(T^*) = \mathcal{R}\left((T^*T)^{\frac{1}{2}}\right), \quad (50)$$

where  $(T^*T)^{\frac{1}{2}}$  denotes the operator square root of Lemma 2.11.

## 2.3 Introduction to frame theory

In this chapter we introduce the basic properties of frames. Frames are, compared to a basis, not necessarily linear independent spanning sets of vectorspaces. The following definitions and results are based on [12, Chapter 1.3]. For the whole chapter we assume that  $\mathcal{X}_n$  is a finite dimensional real Hilbert space with  $\dim(\mathcal{X}_n) \leq n$ . First we give the formal definition of a frame.

**Definition 2.39.** *A family  $(\hat{x}^i)_{i=1,\dots,n} \subseteq \mathcal{X}_n$  is called a **frame** of  $\mathcal{X}_n$ , if there exist two constants  $A, B > 0$  with*

$$A\|x\|^2 \leq \sum_{i=1}^n |\langle x, \hat{x}^i \rangle|^2 \leq B\|x\|^2 \quad \text{for all } x \in \mathcal{X}_n. \quad (51)$$

*Those constants are called **frame bounds**.*

The following Lemma gives an alternative characterization for frames.

**Lemma 2.40.** *Let  $(\hat{x}^i)_{i=1,\dots,n}$  be a family of vectors in  $\mathcal{X}_n$ . Then the family is a frame of  $\mathcal{X}_n$ , iff it is a spanning set of  $\mathcal{X}_n$ .*

We now define three operators, which are quite important for handling with frames, even though they can be defined for an arbitrary family of vectors (not necessarily a frame) within a Hilbert space.

**Definition 2.41.** *With  $(\hat{x}^i)_{i=1,\dots,n} \subseteq \mathcal{X}_n$  we define the **analysis operator**  $T : \mathcal{X}_n \rightarrow \mathbb{R}^n$  by*

$$Tx := (\langle x, \hat{x}^i \rangle)_{i=1}^n, \quad x \in \mathcal{X}_n. \quad (52)$$

*The adjoint operator  $T^* : \mathbb{R}^n \rightarrow \mathcal{X}_n$  according to Definition 2.2 is called **synthesis operator**. Furthermore, the operator  $S : \mathcal{X}_n \rightarrow \mathcal{X}_n$  with  $S := T^*T$  is called **frame operator**. It is explicitly given by*

$$Sx = T^*Tx = \sum_{i=1}^n \langle x, \hat{x}^i \rangle \hat{x}^i, \quad x \in \mathcal{X}_n. \quad (53)$$

The following theorem states some properties of the frame operator:

**Theorem 2.42.** *The frame operator  $S$  of Definition 2.41 based on a frame  $(\hat{x}^i)_{i=1,\dots,n} \subseteq \mathcal{X}_n$  is a positive, self-adjoint and invertible operator. Given  $A, B$  as in Definition 2.39, the operator satisfies*

$$A \cdot \text{Id} \leq S \leq B \cdot \text{Id} \quad (54)$$

*in the sense that for all  $x \in \mathcal{X}_n$ , it holds*

$$\langle Ax, x \rangle \leq \langle Sx, x \rangle \leq \langle Bx, x \rangle. \quad (55)$$

The previous Lemma 2.40 guarantees that a frame is always a spanning set of the corresponding Hilbert space. The question arises, how an arbitrary element of the Hilbert space can be expressed as linear combination of the frame elements. The next theorem states, how this expression can be found with the help of the frame operator. It makes use of the invertibility of the frame operator  $S$  in the finite dimensional case (see [12, Sec. 1.4.2]).

**Theorem 2.43.** *Let  $(\hat{x}^i)_{i=1,\dots,n}$  be a frame in  $\mathcal{X}_n$  with frame operator  $S$ . Then, for every  $x \in \mathcal{X}_n$ , it holds*

$$x = \sum_{i=1}^n \langle x, \hat{x}^i \rangle S^{-1} \hat{x}^i = \sum_{i=1}^n \langle x, S^{-1} \hat{x}^i \rangle \hat{x}^i. \quad (56)$$

*In this case it is sufficient to invert  $S$  only on  $\text{span}(\hat{x}^i)_{i=1,\dots,n}$ .*

This last equality in the statement of this theorem motivates the following definition:

**Definition 2.44.** *In the setting of Theorem 2.43, the family  $(S^{-1}\hat{x}^i)_{i=1,\dots,n}$  is called the **canonical dual frame** of  $(\hat{x}^i)_{i=1,\dots,n}$ .*

Since  $\{\hat{x}^i\}$  is only a spanning set, there exist many different representations of  $x$ . In general, a representation similar to (56) can also be found without application of  $S^{-1}$ . So we define a general **dual frame**.

**Definition 2.45.** *For a frame  $(\hat{x}^i)_{i=1,\dots,n}$  in  $\mathcal{X}_n$ , the frame  $(\hat{x}^i)_{i=1,\dots,n}$  is called **dual frame**, if*

$$x = \sum_{i=1}^n \langle x, \hat{x}^i \rangle \hat{x}^i \quad \text{for all } x \in \mathcal{X}_n. \quad (57)$$

Even though the dual frame is not unique, the canonical dual frame has a special property, which is presented in the following proposition and the corollary below:

**Proposition 2.46.** *For a frame  $(\hat{x}^i)_{i=1,\dots,n}$  in  $\mathcal{X}_n$  we assume that  $x \in \mathcal{X}_n$  can be written as  $x = \sum_{i=1}^n a_i \hat{x}^i$  with appropriately chosen scalars  $a_i$ . Then the equality*

$$\sum_{i=1}^n |a_i|^2 = \sum_{i=1}^n |\langle x, S^{-1}\hat{x}^i \rangle|^2 + \sum_{i=1}^n |a_i - \langle x, S^{-1}\hat{x}^i \rangle|^2.$$

holds with the frame operator  $S$ .

*Proof.* With the analysis operator  $T$  of Definition 2.41, we obtain

$$(\langle x, S^{-1}\hat{x}^i \rangle)_{i=1}^n = (\langle S^{-1}x, \hat{x}^i \rangle)_{i=1}^n \in \text{ran } T$$

Since  $x = \sum_{i=1}^n a_i \hat{x}^i$ , it follows by calculating the scalar product of vectors in  $\mathbb{R}^n$

$$\begin{aligned} (a_i - \langle x, S^{-1}\hat{x}^i \rangle)_{i=1}^n \cdot Tx &= (a_i - \langle x, S^{-1}\hat{x}^i \rangle)_{i=1}^n \cdot (\langle x, \hat{x}^i \rangle)_{i=1}^n \\ &= \left\langle x, \underbrace{\sum_{i=1}^n a_i \hat{x}^i}_{=x} \right\rangle - \left\langle x, \underbrace{\sum_{i=1}^n \langle x, S^{-1}\hat{x}^i \rangle \hat{x}^i}_{=x} \right\rangle = 0 \end{aligned}$$

and therefore

$$(a_i - \langle x, S^{-1}\hat{x}^i \rangle)_{i=1}^n \in \ker T^* = (\text{ran } T)^\perp$$

Considering the trivial decomposition

$$(a_i)_{i=1}^n = (\langle x, S^{-1}\hat{x}^i \rangle)_{i=1}^n + (a_i - \langle x, S^{-1}\hat{x}^i \rangle)_{i=1}^n,$$

squaring both sides and using the orthogonality of the terms on the right hand side proves the claim. □

**Corollary 2.47.** *In the setting of the previous proposition, we assume that  $(\hat{x}^i)_{i=1,\dots,n}$  is an arbitrary dual frame. Then we have the inequality*

$$\|(\langle x, S^{-1}\hat{x}^i \rangle)_{i=1}^n\|_2 \leq \|(\langle x, \hat{x}^i \rangle)_{i=1}^n\|_2.$$

*This means that the canonical dual frame  $(S^{-1}\hat{x}^i)_{i=1,\dots,n}$  of Definition 2.44 is a dual frame that minimizes the  $\ell_2$ -norm.*

In the general case of frames, it is sometimes difficult, numerically instable and computationally expensive to calculate a dual frame analytically. Therefore it is interesting to observe algorithms that approximate a dual frame in iterative steps. One example for that purpose is called the **frame algorithm**:

**Proposition 2.48.** (Frame Algorithm) *Let  $A, B \in \mathbb{R}$  be the bounds of a frame  $(\hat{x}^i)_{i=1,\dots,n}$  and  $S$  the frame operator in the vector space  $\mathcal{X}_n$ . For  $x \in \mathcal{X}_n$  the sequence  $(y_j)_{j=0}^\infty$  given recursively by*

$$y_0 = 0, \quad y_j = y_{j-1} + \frac{2}{A+B} S(x - y_{j-1}) \quad \text{for all } j \geq 1$$

*converges to  $x$  with the rate*

$$\|x - y_j\| \leq \left( \frac{B-A}{B+A} \right)^j \|x\|, \quad j \geq 0.$$

Since the previous proposition requires the frame bounds  $A, B$  explicitly, which might not be known in some cases, we also present an alternative algorithm that is called the **Conjugate Gradient Method** in frame theory. The formulas in the method are similar to the conventional Conjugate Gradient Method for solving systems of linear equations given in [17, Chapter 2]. It constructs a sequence without the knowledge of  $A$  and  $B$ .

**Proposition 2.49.** (Conjugate Gradient Method) *Let  $(\hat{x}^i)_{i=1,\dots,n}$  be a frame in  $\mathcal{X}_n$  with frame operator  $S$ . For a given  $x \in \mathcal{X}_n$ , we define three sequences  $(y_j)_{j=0}^\infty$ ,  $(r_j)_{j=0}^\infty$  and  $(p_j)_{j=-1}^\infty \in \mathcal{X}_n$  and a sequence of scalars  $(\lambda_j)_{j=-1}^\infty$  by the initial values*

$$y_0 = 0, \quad r_0 = p_0 = Sx, \quad \text{and} \quad p_{-1} = 0$$

and recursively by

$$\lambda_j = \frac{\langle r_j, p_j \rangle}{\langle p_j, Sp_j \rangle}, \quad y_{j+1} = y_j + \lambda_j p_j, \quad r_{j+1} = r_j - \lambda_j Sp_j$$

and

$$p_{j+1} = Sp_j - \frac{\langle Sp_j, Sp_j \rangle}{\langle p_j, Sp_j \rangle} p_j - \frac{\langle Sp_j, Sp_{j-1} \rangle}{\langle p_{j-1}, Sp_{j-1} \rangle} p_{j-1}.$$

Then the sequence  $(y_j)_{j=0}^\infty$  converges to  $x$  with the rate

$$\| \|x - y_j\| \| \leq \frac{2\sigma^j}{1 + \sigma^{2j}} \| \|x\| \| \quad \text{with } \sigma = \frac{\sqrt{B} - \sqrt{A}}{\sqrt{B} + \sqrt{A}}$$

with the norm  $\| \cdot \| := \| S^{1/2}(\cdot) \|$ .

## 3 Regularization Operators

### 3.1 Motivation

The main idea of **regularization** is to find the solution  $x$  of the equation

$$Tx = y \tag{58}$$

for a given noisy data  $y^\delta$  of  $y$  fulfilling

$$\| y^\delta - y \| \leq \delta. \tag{59}$$

Even though the best approximation  $x^\dagger$  of the exact equation (58) can be written with the Moore-Penrose inverse (see Theorem 2.9) in the form  $x^\dagger = T^\dagger y$ , it is in general not the best idea to take  $T^\dagger y^\delta$  as solution of the noisy problem, if  $T^\dagger$  is unbounded. Furthermore it is possible that  $y^\delta$  does not even belong to the domain  $\mathcal{D}(T^\dagger)$  of  $T^\dagger$ . So the goal is to find a clever way to calculate an approximating solution  $x_\alpha^\delta$  of  $x^\dagger$ , which does depend continuously on the noisy data  $y^\delta$  and has furthermore

the property of converging to  $x^\dagger$  in the case that the noise level  $\delta$  decreases to zero. The trick is now to replace  $T^\dagger$  by the parameter dependent operator family  $\{R_\alpha\}$ , which depends on a regularization parameter  $\alpha$  and is continuous. Then the solution of the approximating problem is defined by  $x_\alpha^\delta := R_\alpha y^\delta$ . Since  $\delta \rightarrow 0$  should result in  $x_\alpha^\delta \rightarrow x^\dagger$  (which is called **pseudo convergence** of these approximations to the free noise solution),  $\alpha$  has to be somehow linked to  $\delta$ . The definition of  $R_\alpha$  and detailed properties are presented in the following subsection.

### 3.2 Definition and properties

**Definition 3.1.** Let  $T : \mathcal{X} \rightarrow \mathcal{Y}$  be a bounded linear operator between the Hilbert spaces  $\mathcal{X}$  and  $\mathcal{Y}$  and  $\alpha_0 \in (0, \infty]$ . For every  $\alpha \in (0, \alpha_0)$ , let  $R_\alpha : \mathcal{Y} \rightarrow \mathcal{X}$  be a continuous (not necessarily linear) operator. We call the operator family  $\{R_\alpha\}$  a **regularization** for  $T^\dagger$ , if for all  $y \in \mathcal{D}(T^\dagger)$  there exists a parameter choice rule  $(\delta, y^\delta) \mapsto \alpha(\delta, y^\delta)$  such that

$$\limsup_{\delta \rightarrow 0} \left\{ \left\| R_{\alpha(\delta, y^\delta)} y^\delta - T^\dagger y \right\| \mid y^\delta \in \mathcal{Y}, \|y^\delta - y\| \leq \delta \right\} = 0 \quad (60)$$

with  $\alpha : \mathbb{R}^+ \times \mathcal{Y} \rightarrow (0, \alpha_0)$  fulfilling

$$\limsup_{\delta \rightarrow 0} \left\{ \alpha(\delta, y^\delta) \mid y^\delta \in \mathcal{Y}, \|y^\delta - y\| \leq \delta \right\} = 0. \quad (61)$$

If a pair  $(R_\alpha, \alpha)$  fulfills (60) and (61) for a specific  $y \in \mathcal{D}(T^\dagger)$  (not necessarily for all  $y \in \mathcal{D}(T^\dagger)$ ), we call it a **regularization method** for  $Tx = y$ .

Dependent on their properties, we distinguish between two different types of parameter choice rules:

**Definition 3.2.** If the parameter choice rule  $\alpha(\delta, y^\delta)$  in Definition 3.1 does not depend explicitly on  $y^\delta$ , we call it an **a-priori parameter choice**. Otherwise it is called **a-posteriori parameter choice rule**.

The following theorem (with proof in [1, Theorem 3.3]) presents a connection between the properties of  $\alpha$  and  $T^\dagger$ .

**Theorem 3.3.** Let  $T : \mathcal{X} \rightarrow \mathcal{Y}$  be a bounded linear operator. If the parameter choice rule  $\alpha$  of the regularization  $\{R_\alpha\}$  only depends on  $y^\delta$  (not explicitly on  $\delta$ ) and the regularization method  $(R_\alpha, \alpha)$  is convergent for every  $y \in \mathcal{D}(T^\dagger)$ , then  $T^\dagger$  is bounded.

This means that the choice  $\alpha = \alpha(y^\delta)$  can never lead to a convergent realization method if  $T^\dagger$  is unbounded. Anyhow it is possible that such a parameter choice rule behaves suitable for finite noise levels  $\delta$ . The following proposition (with proof in [1, Proposition 3.4]) gives a hint, how regularization operators can actually be constructed.

**Proposition 3.4.** *If  $\{R_\alpha\}$  is a family of continuous (not necessarily linear) operators with  $R_\alpha \xrightarrow{\alpha \rightarrow 0} T^\dagger$  pointwise on  $\mathcal{D}(T^\dagger)$ . Then it is a regularization for  $T^\dagger$  and for every  $y \in \mathcal{D}(T^\dagger)$  and an a-priori parameter choice rule, there exists  $\alpha$  such that  $(R_\alpha, \alpha)$  is a convergent regularization method for the problem  $Tx = y$ .*

The converse of this proposition holds in the following sense:

For every convergent regularization method  $(R_\alpha, \alpha)$  and when  $y^\delta = y$ , Equation (60) implies immediately  $\lim_{\delta \rightarrow 0} R_{\alpha(\delta, y)}y = T^\dagger y$  for all  $y \in \mathcal{D}(T^\dagger)$ . Note that this limit holds in comparison to the previous proposition only for the  $\alpha$  values, which are in the range of the parameter choice strategy  $\alpha(\delta, y^\delta)$ . With the additional assumption of continuity of  $\delta \mapsto \alpha(\delta, y^\delta)$ , we also get the pointwise limit in the form  $R_\alpha y \xrightarrow{\alpha \rightarrow 0} T^\dagger y$  for all  $y \in \mathcal{D}(T^\dagger)$ .

The next proposition describes the behavior of the solution of the approximating problem  $x_\alpha$  in the case of regularizations  $R_\alpha$ , which are linear operators (i.e. **linear regularizations**). The proof is given in [1, Proposition 3.6].

**Proposition 3.5.** *If  $\{R_\alpha\}$  is a linear regularization and  $x_\alpha := R_\alpha y$ , which is defined for all  $y \in \mathcal{Y}$ . Then  $x_\alpha \xrightarrow{\alpha \rightarrow 0} T^\dagger y$  for all  $y \in \mathcal{D}(T^\dagger)$ . If additionally  $\sup\{\|TR_\alpha\| \mid \alpha > 0\} < \infty$  (i.e. the operator norms are uniformly bounded in  $\alpha$ ), then it holds additionally that  $\|x_\alpha\| \xrightarrow{\alpha \rightarrow 0} +\infty$  for all  $y \in \mathcal{Y} \setminus \mathcal{D}(T^\dagger)$ .*

In Proposition 3.4, we mentioned under which assumptions an a-priori parameter choice with a convergent regularization method  $(R_\alpha, \alpha)$  exists. Such a rule can be characterized in the following, at least for linear regularizations. The statement is proved in [1, Proposition 3.7].

**Proposition 3.6.** *For a linear regularization  $\{R_\alpha\}$  let  $\alpha = \alpha(\delta)$  be an a-priori parameter choice rule. Then  $(R_\alpha, \alpha)$  is a convergent regularization method iff*

$$\lim_{\delta \rightarrow 0} \alpha(\delta) = 0 \tag{62}$$

and

$$\lim_{\delta \rightarrow 0} \delta \|R_{\alpha(\delta)}\| = 0. \tag{63}$$

**Remark 3.7.** *If we replace (63) by the condition*

$$\limsup_{\delta \rightarrow 0} \delta \|R_{\alpha(\delta)}\| < +\infty, \quad (64)$$

*then  $(R_\alpha, \alpha)$  is only weakly convergent in the sense that for all sequences  $\delta_k \rightarrow 0$  and  $y_k \in \mathcal{Y}$  chosen so that  $\|y_k - y\| \leq \delta_k$  for all  $k$ ,  $\{R_{\alpha(\delta_k)}y_k\}$  converges weakly to  $T^\dagger y$ . If (64) does not hold, it is further possible to find sequences such that the norm of  $\{R_{\alpha(\delta)}y_k\}$  diverges. The proof of this statement can be found in [4].*

### 3.3 Regularization by Projection

In the previous section we presented some general properties of regularization. In this chapter, we start with a concrete possible method of implementing such a regularization, namely projection onto suitable subspaces. The idea is very simple and natural: Instead of finding the minimum-norm solution of  $Tx = y$  in  $\mathcal{X}$ , we choose some finite dimensional subspaces  $\mathcal{X}_1 \subset \mathcal{X}_2 \subset \mathcal{X}_3 \subset \dots$ , such that  $\bigcup_{n=1}^{\infty} \mathcal{X}_n$  is dense in  $\mathcal{X}$  and solve the problem

$$T_n x = y, \quad (65)$$

where  $T_n := TP_n$  and  $P_n$  denotes the orthogonal projection onto  $\mathcal{X}_n$ . The solution  $x_n := T_n^\dagger y$  is a stable approximation of  $x^\dagger$ , since  $T_n$  has finite dimensional and thus closed range (see Theorem 2.21) and the operator  $T_n^\dagger$  is bounded. Nevertheless, it is not guaranteed that  $x_n$  converges to  $x^\dagger$ . We now give an example, where this is not the case (see Example 3.8 [1, Example 3.19] and [6], respectively).

*Example 3.8.* For an infinite dimensional Hilbert space  $\mathcal{X}$ , let  $\{e_n\}$  be an orthonormal basis. Since each  $x \in \mathcal{X}$  can be written as  $x = \sum_{i=1}^{\infty} \xi_i e_i$ , we define the operator  $T : \mathcal{X} \rightarrow \mathcal{X}$  by:

$$T : \sum_{i=1}^{\infty} \xi_i e_i \mapsto \sum_{i=1}^{\infty} (a_i \xi_i + b_i \xi_1) e_i$$

with

$$a_i := \begin{cases} i^{-1}, & i \text{ odd,} \\ i^{-\frac{5}{2}}, & i \text{ even.} \end{cases} \quad b_i := \begin{cases} 0, & i = 1 \\ i^{-1}, & i > 1 \end{cases}$$

We show that this operator is linear, injective and compact: The linearity is clear, since for  $x, x' \in \mathcal{X}$  we have

$$\begin{aligned}
T(x + x') &= T\left(\sum_{i=1}^{\infty} \xi_i e_i + \sum_{i=1}^{\infty} \xi'_i e_i\right) \\
&= T\left(\sum_{i=1}^{\infty} (\xi_i + \xi'_i) e_i\right) \\
&= \sum_{i=1}^{\infty} (a_i(\xi_i + \xi'_i) + b_i(\xi_1 + \xi'_1)) e_i \\
&= \sum_{i=1}^{\infty} (a_i \xi_i + b_i \xi_1) e_i + \sum_{i=1}^{\infty} (a_i \xi'_i + b_i \xi'_1) e_i \\
&= T(x) + T(x')
\end{aligned}$$

and for a scalar  $\lambda$ , it holds

$$\begin{aligned}
T(\lambda x) &= T\left(\lambda \sum_{i=1}^{\infty} \xi_i e_i\right) \\
&= T\left(\sum_{i=1}^{\infty} \lambda \xi_i e_i\right) \\
&= \sum_{i=1}^{\infty} (a_i \lambda \xi_i + b_i \lambda \xi_1) e_i \\
&= \lambda \sum_{i=1}^{\infty} (a_i \xi_i + b_i \xi_1) e_i \\
&= \lambda T(x).
\end{aligned}$$

To show the injectivity, we assume  $T(x) - T(x') = T(x - x') = 0$  for  $x, x' \in \mathcal{X}$ . We observe

$$\begin{aligned}
0 &= T\left(\sum_{i=1}^{\infty} (\xi_i - \xi'_i) e_i\right) \\
&= \sum_{i=1}^{\infty} (a_i(\xi_i - \xi'_i) + b_i(\xi_1 - \xi'_1)) e_i
\end{aligned}$$

and since  $\{e_i\}$  is a linearly independent set, this implies  $a_i(\xi_i - \xi'_i) + b_i(\xi_1 - \xi'_1) = 0$  for all  $i \in \mathbb{N}$ . In the case  $i = 1$  we get  $1^{-1}(\xi_1 - \xi'_1) + 0(\xi_1 - \xi'_1) = \xi_1 - \xi'_1 = 0$  and therefore  $\xi_1 = \xi'_1$ . Using this for  $i \geq 2$ , we observe  $a_i(\xi_i - \xi'_i) + b_i(\xi_1 - \xi'_1) = a_i(\xi_i - \xi'_i) = 0$ . Since  $a_i \neq 0$  per definition we also get  $\xi_i = \xi'_i$  in this case and therefore  $x = x'$ , which finally implies the injectivity. For the compactness, we define finite dimensional subspaces of  $\mathcal{X}$  by  $\mathcal{X}_n := \text{span}\{e_1, \dots, e_n\}$  and the approximating operators  $T_n : \mathcal{X} \rightarrow \mathcal{X}_n$  by

$$T_n : \sum_{i=1}^{\infty} \xi_i e_i \mapsto \sum_{i=1}^n (a_i \xi_i + b_i \xi_1) e_i.$$

Since the range of  $T_n$  is finite dimensional, these operators are compact. We show that  $\|T - T_n\| \rightarrow 0$  for  $n \rightarrow \infty$  to prove also the compactness of  $T$ . For  $x \in \mathcal{X}$  we have

$$\begin{aligned} \|(T - T_n)x\|^2 &= \left\| \sum_{i=n+1}^{\infty} (a_i \xi_i + b_i \xi_1) e_i \right\|^2 \\ &= \sum_{i=n+1}^{\infty} |a_i \xi_i + b_i \xi_1|^2 \\ &\leq \sum_{i=n+1}^{\infty} a_i^2 |\xi_i|^2 + |\xi_1|^2 \sum_{i=n+1}^{\infty} b_i^2 \\ &\leq \sum_{i=n+1}^{\infty} a_i^2 \sum_{i=n+1}^{\infty} |\xi_i|^2 + \|x\|^2 \sum_{i=n+1}^{\infty} b_i^2 \\ &\leq \left( \sum_{i=n+1}^{\infty} a_i^2 + \sum_{i=n+1}^{\infty} b_i^2 \right) \|x\|^2. \end{aligned}$$

Since the infinite series  $\sum_{i=1}^{\infty} a_i^2$  and  $\sum_{i=1}^{\infty} b_i^2$  are both convergent since they are bounded by  $\sum_{i=1}^{\infty} i^{-2} < \infty$ , the upper expression in the bracket converges to zero for  $n \rightarrow \infty$ . This concludes the compactness.

For  $x^\dagger := \sum_{i=1}^{\infty} i^{-1} e_i$  (which is an element of  $\mathcal{X}$  because of  $\|x\|^2 = \sum_{i=1}^{\infty} i^{-2} < \infty$ ) we set

$$y := Tx^\dagger = T \left( \sum_{i=1}^{\infty} i^{-1} e_i \right) = \sum_{i=1}^{\infty} (a_i i^{-1} + b_i 1^{-1}) e_i = \sum_{i=1}^{\infty} (a_i i^{-1} + b_i) e_i,$$

then because of the injectivity the problem  $Tx = y$  has trivially the unique solution  $x^\dagger$ .

We seek for the best-approximate solution  $x_n := \sum_{i=1}^n \xi_{i,n} e_i$  in  $\mathcal{X}_n$ . The unknown coefficients in the representation of  $x_n$  can be calculated by the minimization problem

$$\|Tx_n - y\|^2 \rightarrow \min.$$

Inserting for  $x_n$  and  $y$  and applying the Parseval's identity leads to

$$\sum_{i=1}^n (a_i (\xi_i - i^{-1}) + b_i (\xi_1 - 1))^2 + \sum_{i=n+1}^{\infty} i^{-2} (1 + a_i - \xi_1)^2 \rightarrow \min, \quad (66)$$

so that  $(\xi_{1,n}, \dots, \xi_{n,n})$  is equal to  $(\xi_1, \dots, \xi_n)$ , for that (66) is minimized. The vanishing first partial derivatives lead to the solution

$$\begin{aligned} \xi_{1,n} &= 1 + \left( \sum_{i=n+1}^{\infty} a_i i^{-2} \right) \left( 1 + \sum_{i=n+1}^{\infty} i^{-2} \right)^{-1}, \\ \xi_{i,n} &= i^{-1} + (a_i i)^{-1} (\xi_{1,n} - 1), \quad 2 \leq i \leq n. \end{aligned}$$

We compute, that

$$\begin{aligned}
\|x_n - P_n x^\dagger\|^2 &= \left\| \sum_{i=1}^n \xi_{i,n} e_i - \sum_{i=1}^n i^{-1} e_i \right\|^2 \\
&= \left\| \sum_{i=1}^n (\xi_{i,n} - i^{-1}) e_i \right\|^2 \\
&= \sum_{i=1}^n |\xi_{i,n} - i^{-1}|^2 \\
&= |\xi_{1,n} - 1^{-1}|^2 + \sum_{i=2}^n |\xi_{i,n} - i^{-1}|^2 \\
&= |\xi_{1,n} - 1|^2 + \sum_{i=2}^n |(a_i i)^{-1} (\xi_{1,n} - 1)|^2 \\
&= \sum_{i=1}^n (a_i i)^{-2} (\xi_{1,n} - 1)^2 \\
&= \left( \sum_{i=1}^n (a_i i)^{-2} \right) \left( \sum_{i=n+1}^{\infty} a_i i^{-2} \right)^2 \left( 1 + \sum_{i=n+1}^{\infty} i^{-2} \right)^{-2}.
\end{aligned}$$

This shows that

$$\|x_n - P_n x^\dagger\| \sim n$$

and since  $\bigcup_{n=1}^{\infty} \mathcal{X}_n$  is dense in  $\mathcal{X}$  by definition and therefore  $P_n x^\dagger \rightarrow x^\dagger$ , this shows that  $x_n$  does not converge to  $x^\dagger$ .

Now we will present the **dual least-squares method** given in [1, 7]. This is a projection method with ensured convergence. It chooses a sequence  $\mathcal{Y}_1 \subset \mathcal{Y}_2 \subset \mathcal{Y}_3 \subset \dots$  of finite-dimensional subspaces of  $\overline{\mathcal{R}(T)} = \mathcal{N}(T^*)^\perp \subseteq \mathcal{Y}$ , whose union is dense in  $\mathcal{N}(T^*)^\perp$ . If we denote the orthogonal projectors onto  $\mathcal{Y}_n$  by  $Q_n$ , we define the approximating sequence  $(x_n)_{n \in \mathbb{N}}$  by the least-squares solution of  $T_n x = y_n$  with  $T_n := Q_n T$  and  $y_n := Q_n y$ . Then  $x_n$  is a stable approximation of  $x^\dagger$  for the same reasons above. This means that in contrast to the beginning of this section, the projections take place in the space  $\mathcal{Y}$  instead of  $\mathcal{X}$ . The connection to the initially space  $\mathcal{X}$  can be done via the adjoint operator  $T^*$ , as we show in the following theorem (see [1, Theorem 3.24], which gives an alternative characterization of  $x_n$ ).

**Theorem 3.9.** *Is  $y \in \mathcal{D}(T^\dagger)$  and  $x_n$  defined as above. If we define  $\mathcal{X}_n := T^*\mathcal{Y}_n$  and  $\Pi_n$  is the orthogonal projector onto  $\mathcal{X}_n$  in  $\mathcal{X}$ , then it holds*

$$x_n = \Pi_n x^\dagger \text{ and } x_n \xrightarrow{n \rightarrow \infty} x^\dagger.$$

*Proof.* Due to the definition of the spaces and projectors, we have  $\mathcal{N}(T_n) = \mathcal{R}(T^*Q_n)^\perp = (T^*\mathcal{Y}_n)^\perp = \mathcal{X}_n^\perp$ . Since  $\Pi_n$  is the projector onto  $\mathcal{X}_n$ ,  $I - \Pi_n$  projects onto  $\mathcal{X}_n^\perp$  and we get

$$Q_n T(I - \Pi_n) = T_n(I - \Pi_n) = 0.$$

If we denote the orthogonal projector onto  $\overline{\mathcal{R}(T)}$  by  $Q$ , we get

$$\|T_n x - Q_n y\| = \|Q_n T(x - x^\dagger)\| = \|Q_n T \Pi_n(x - \Pi_n x^\dagger)\|.$$

So  $\{\Pi_n x^\dagger\} + \mathcal{X}_n^\perp$  is the set of least-squares solutions of the equation  $T_n x = y$  and therefore  $x_n = \Pi_n x^\dagger$ . Since  $\mathcal{Y}_n \subset \mathcal{Y}_{n+1}$  implies  $\mathcal{X}_n \subset \mathcal{X}_{n+1}$  by definition,  $\bigcup_{n \in \mathbb{N}} \mathcal{X}_n$  is dense in  $\mathcal{N}(T)^\perp$  and  $x^\dagger \in \mathcal{N}(T)^\perp$ , we conclude  $x_n \xrightarrow{n \rightarrow \infty} x^\dagger$ .  $\square$

This theorem shows that with  $T_n$  defined as above, the set  $\{T_n^\dagger\}$  is a regularization operator according to Proposition 3.4. Even though this family of operators are not introduced with an explicit regularization parameter, the following theorem shows that the smallest singular value of the operator  $T_n$ , which we denote by  $\mu_n$  can be seen as a hidden regularization parameter. As in Sec. 3.2, we assume for the noisy data  $y^\delta$  that  $\|Q_n(y - y^\delta)\| \leq \delta$ . The noisy least-squares solution with  $y^\delta$  instead of  $y$  is denoted by  $x_n^\delta$ . Then we state:

**Theorem 3.10.** *Let  $y \in \mathcal{D}(T^\dagger)$  and  $\mu_n$  denote the smallest singular value of the above introduced operator  $T_n$ . We further assume that  $\delta/\mu_n \rightarrow 0$  as  $\delta \rightarrow 0$  and  $n \rightarrow \infty$ . Then it holds*

$$x_n^\delta \rightarrow x^\dagger \text{ as } \delta \rightarrow 0 \text{ and } n \rightarrow \infty.$$

*Proof.* We can estimate

$$\|x_n^\delta - x^\dagger\| \leq \|x_n - x^\dagger\| + \|x_n^\delta - x_n\| \tag{67}$$

$$= \|x_n - x^\dagger\| + \|T_n^\dagger Q_n(y - y^\delta)\| \tag{68}$$

$$\leq \|x_n - x^\dagger\| + \|T_n^\dagger\| \delta \tag{69}$$

$$\leq \|x_n - x^\dagger\| + \frac{\delta}{\mu_n}. \tag{70}$$

With Theorem 3.9 the assertion follows.  $\square$

## 4 Regularization with training data

In this section we present a very practical application of **Regularization by Projection** presented in Sec. 3.3. The context is the following (see also [8]): We aim to solve the equation  $Tx = y$  for unknown linear operator  $T$  between two Hilbert spaces  $\mathcal{X}, \mathcal{Y}$  and given right hand side  $y \in \mathcal{R}(T)$ . The only available information about  $T$  are so-called training pairs  $(\hat{x}^i, \hat{y}^i) \in \mathcal{X} \times \mathcal{Y}$  with  $T\hat{x}^i = \hat{y}^i$  for all  $i$ . We furthermore assume that  $\{\hat{x}^i\}_{i=1, \dots, n}$  are linearly independent for every  $n \in \mathbb{N}$ . With the notation  $\mathcal{X}_n := \text{span}\{\hat{x}^1, \dots, \hat{x}^n\}$  and  $\mathcal{Y}_n := \text{span}\{\hat{y}^1, \dots, \hat{y}^n\}$ , we get the finite dimensional subspaces of  $\mathcal{X}$  and  $\mathcal{Y}$ , respectively. With these spaces we can apply the theory at the beginning of Sec. 3.3. Assuming the injectivity of the operator  $T$ , we can give an explicit expression of the regularized solution  $x_n$  in terms of the training pairs. In order to do that, we need the following theorem.

**Theorem 4.1.** *Let  $T$  be injective. Then it holds*

$$T_n^\dagger = T^{-1}Q_n, \quad (71)$$

where  $Q_n$  is defined as in the previous section.

*Proof.* We proof the claim by showing (i)-(v) of Proposition 2.10 for the operator  $T^{-1}Q_n$ . Due to the injectivity of  $T$  and the fact that the training pictures  $\{\hat{x}^i\}_{i=1, \dots, n}$  are linearly independent, the set  $\{\hat{y}^i\}_{i=1, \dots, n}$  is also linearly independent and therefore a basis of  $\mathcal{Y}_n$ . Thus, for every  $y \in \mathcal{Y}$ ,  $Q_n y \in \mathcal{Y}_n$  can be written as

$$Q_n y = \sum_{i=1}^n \lambda_i(n) \hat{y}^i.$$

Using this, we get

$$x = T^{-1}Q_n y = T^{-1} \left( \sum_{i=1}^n \lambda_i(n) \hat{y}^i \right) = \sum_{i=1}^n \lambda_i(n) \hat{x}^i \in \mathcal{X}_n.$$

Since the last term in the equation is just a linear combination in  $\mathcal{X}_n$ , we can conclude that for any  $x \in \mathcal{X}_n$  we find some  $y \in \mathcal{Y}$ , so that  $x = T^{-1}Q_n y$ . This shows that

$$\mathcal{R}(T^{-1}Q_n) = \mathcal{X}_n. \quad (72)$$

For every  $x' \in \mathcal{N}(AP_n)$ , thus to the injectivity of  $T$ ,  $x' \in \mathcal{N}(P_n) = \mathcal{X}_n^\perp$ . Since those spaces are finite dimensional, this implies

$$\mathcal{N}(TP_n)^\perp = \mathcal{X}_n. \quad (73)$$

Equations (72) and (73) show (i) in Proposition 2.10. Equations in (ii)-(v) come from straightforward calculation, where we use the notation  $P_{\mathcal{U}_n^\perp}$  and  $P_{\mathcal{N}(TP_n)}$  for the projectors onto the space in the subscript:

- (ii)  $TP_n T^{-1} Q_n TP_n = TP_n T^{-1} TP_n = TP_n,$
- (iii)  $T^{-1} Q_n TP_n T^{-1} Q_n = T^{-1} Q_n T T^{-1} Q_n = T^{-1} Q_n,$
- (iv)  $T^{-1} Q_n TP_n = T^{-1} TP_n = P_n = I - P_{\mathcal{U}_n^\perp} = I - P_{\mathcal{N}(TP_n)},$
- (v)  $TP_n T^{-1} Q_n = T T^{-1} Q_n = Q_n = P_{\mathcal{R}(TP_n)}.$

□

Thanks to the previous theorem, we observe:

**Remark 4.2.** *The solution of the projected problem (65) can be written as*

$$x_n = T_n^\dagger y \stackrel{(71)}{=} T^{-1} Q_n y.$$

Let  $\{\underline{\hat{y}}^1, \dots, \underline{\hat{y}}^n\}$  be an orthonormal basis of  $\mathcal{Y}_n$  and  $\{\underline{\hat{x}}^1, \dots, \underline{\hat{x}}^n\}$  defined by  $\underline{\hat{x}}^i := T^{-1} \underline{\hat{y}}^i$  for  $i = 1, \dots, n$  the preimages of the orthonormal basis. Then the orthogonal projector  $Q_n$  can be written as

$$Q_n = \sum_{i=1}^n \langle \cdot, \underline{\hat{y}}^i \rangle \underline{\hat{y}}^i \tag{74}$$

and we get for the regularized solution  $x_n$  the representation

$$x_n = T^{-1} Q_n y = T^{-1} \sum_{i=1}^n \langle y, \underline{\hat{y}}^i \rangle \underline{\hat{y}}^i = \sum_{i=1}^n \langle y, \underline{\hat{y}}^i \rangle T^{-1} \underline{\hat{y}}^i = \sum_{i=1}^n \langle y, \underline{\hat{y}}^i \rangle \underline{\hat{x}}^i. \tag{75}$$

## 5 Simulations

In this section, we show a practical application of the results of the previous Chapter 4. As linear operator according to (58), we choose  $T : L^2(\Omega) \rightarrow L^2(\mathbb{R} \times [0, \pi])$  as the **Radon transform**  $R$  with a parallel beam geometry (see [18]).

### 5.1 Algorithm

The computation algorithm written in MATLAB consists of several parts:

### 5.1.1 Import of training data

We test the algorithm on three different data sets:

- Numbers
- Sunflowers

Each set includes  $n$  different pictures to the corresponding topic of equal image resolution consisting of  $N$  gray scale data points in  $[0, 1]$ , which can be represented as linearly independent elements of  $\mathbb{R}^N$ . They are used as training data and Radon transformed by the built in MATLAB-function for  $K$  different angles  $\theta_k \in [0, 2\pi)$ . The results are  $n$  different elements of  $\mathbb{R}^{M \times K} \simeq \mathbb{R}^{M \cdot K}$ , where  $M$  is the length of a Radon projection at a specific angle. In the case of quadratic pictures with size  $\sqrt{N} \times \sqrt{N}$ , this implies that  $M \approx \sqrt{2 \cdot N}$ , which is the length of the diagonal of the square. All pictures  $\hat{x}^i \in \mathbb{R}^N$  and Radon transformed  $\hat{y}^i \in \mathbb{R}^{M \cdot K}$  represent the **training data** pairs  $(\hat{x}^i, \hat{y}^i)_{i=1, \dots, n}$  described in Chapter 4. An overview to these introduced variables is given by Tab. 1.

$N$	Size of images
$n$	Number of training images
$M$	Length of Radon projection
$K$	Number of Radon angles

Table 1: Legend of the used variables in this chapter.

Note that in our following investigations, it holds  $(M \cdot K) \gg n$ , i.e. the picture size is large compared to the number of training pairs. We also use the variable  $n$  as regularization parameter according to Sec. 3.2. In order to perform the orthonormalization of the set  $\{\hat{y}^i\}_{i=1, \dots, n}$ , we use four different methods of the following subsections and compare their outcomes.

### 5.1.2 Gram Schmidt method

The vectors defined by

$$c^i := \hat{y}^i - \sum_{j=1}^{i-1} \langle \hat{y}^i, \hat{y}^j \rangle \hat{y}^j \quad i = 1, \dots, n \quad (76)$$

are an orthogonal system (see [5, page 754]). After normalizing, we get the orthonormal system by

$$\hat{y}^i := \frac{c^i}{\|c^i\|}. \quad (77)$$

The computational effort of this algorithm is  $\mathcal{O}(MKn^2)$  according to [11, Section 5.2.8].

### 5.1.3 Modified Gram Schmidt method

An alternative version of the Gram Schmidt method is given by [11, Section 5.2.8, Algorithm 5.2.5].

---

**Algorithm 1:** Modified Gram Schmidt

---

**input :** The matrix  $A \in \mathbb{R}^{M \cdot K \times n}$  with columns  $\{\hat{y}^i\}_{i=1, \dots, n}$

**output:** QR factorization  $Q \in \mathbb{R}^{M \cdot K \times n}$ ,  $R \in \mathbb{R}^{n \times n}$ , where  $Q$  contains the columns  $\{\hat{y}^i\}_{i=1, \dots, n}$

**for**  $k = 1 : n$  **do**

$R(k, k) = \|A(1 : M * K, k)\|_2;$

$Q(1 : M * K, k) = A(1 : M * K, k) / R(k, k);$

**for**  $j = k + 1 : n$  **do**

$R(k, j) = Q(1 : M * K, k)^T A(1 : M * K, j);$

$A(1 : M * K, j) = A(1 : M * K, j) - Q(1 : M * K, k)R(k, j);$

**end**

**end**

---

The orthonormal system is then given by the columns of the This version is a sounder computational procedure, since the calculation steps are rearranged in comparison to the straight forward calculation of (76) and (77) presented in the previous section. The computational effort of this algorithm is  $\mathcal{O}(MKn^2)$ .

### 5.1.4 Householder reflections

Applying Householder transformations in [9, Section 5.2.1, Algorithm 5.2.1] to the set, we receive again an orthonormal system of vectors.

---

**Algorithm 2:** Householder reflection vector

---

**input** : vector  $x \in \mathbb{R}^n$ **output:** Householder vector  $v \in \mathbb{R}^n$ ,  $\beta \in \mathbb{R}$ function:  $[v, \beta] = \text{house}(x)$ ; $n = \text{length}(x)$ ; $\sigma = x(2:n)^T x(2:n)$ ; $v = \begin{bmatrix} 1 \\ x(2:n) \end{bmatrix}$ ;**if**  $\sigma = 0$  **then**└  $\beta = 0$ **else**└  $\mu = \sqrt{x(1)^2 + \sigma}$ ;└ **if**  $x(1) \leq 0$  **then**└└  $v(1) = x(1) - \mu$ └ **else**└└  $v(1) = -\sigma / (x(1) + \mu)$ └  $\beta = 2v(1)^2 / (\sigma + v(1)^2)$ ;└  $v = v/v(1)$ ;

---

**Algorithm 3:** QR decomposition with Householder reflections

---

**input** : The matrix  $A \in \mathbb{R}^{M \cdot K \times n}$  with columns  $\{\hat{y}^i\}_{i=1,\dots,n}$ **output:** QR factorization  $Q \in \mathbb{R}^{M \cdot K \times n}$ ,  $R \in \mathbb{R}^{n \times n}$ , where  $Q$  contains the columns  $\{\hat{y}^i\}_{i=1,\dots,n}$ **for**  $j = 1 : n$  **do**└  $[v, \beta] = \text{house}(A(j : M \cdot K, j))$ ;└  $A(j : M \cdot K, j : n) = (I_{M \cdot K - j + 1} - \beta v v^T) A(j : M \cdot K, j : n)$  ;└ **if**  $j < M \cdot K$  **then**└└  $A(j + 1 : M \cdot K, j) = v(2 : M \cdot K - j + 1)$ └ **end****end**

---

The computational effort of this algorithm is  $\mathcal{O}(M^2 K^2 n - M K n^2 + n^3/3)$  (see [11, Section 5.2.1]).

### 5.1.5 QR MATLAB native

If we write all vectors  $\hat{y}^i$  as columns of a matrix and perform the MATLAB native QR decomposition, we also receive an orthonormal system with the same span as

the columns of the original matrix, since the columns of  $Q$  represent an orthonormal system.

### 5.1.6 Givens rotations

In [11, Section 5.2.3, Algorithm 5.2.2], the orthonormalization of the columns of a matrix via Givens rotations is presented. However, the application of this algorithm to a  $(M \cdot K) \times n$  matrix leads to a  $Q \in \mathbb{R}^{(M \cdot K) \times (M \cdot K)}$  orthogonal matrix, which is in our case  $(M \cdot K) \gg n$  very inefficient in terms of memory usage. In all other versions of the previous subsections, it is possible to calculate the  $QR$  decomposition with  $Q \in \mathbb{R}^{(M \cdot K) \times n}$  and  $R \in \mathbb{R}^{n \times n}$ . Therefore, we do not implement and skip this version in the in the following.

### 5.1.7 Arnoldi iteration

The Arnoldi algorithm in [12, Section 4.4] is also a possible orthonormalization method, but since it works on Krylov spaces of quadratic matrices, we do not use it in the context of this thesis.

### 5.1.8 Backtransformation of test data

Finally for each Radon transformed picture  $y$  of the test data, we are able to use the formula in (75) for the reverse transformation to the image space. Note that the orthonormal system  $\{\hat{y}^1, \dots, \hat{y}^n\}$  has to be backtransformed via the exact inverse Radon transform before, to receive the preimages  $\{\hat{x}^1, \dots, \hat{x}^n\}$  of Remark 4.2. The result can then be compared to the original image to investigate the quality of the procedure.

### 5.1.9 Inverse Radon transformation

Independent of the chosen orthonormalization method, we apply the MATLAB-function of the inverse Radon transform to the whole set  $\{\hat{y}^i\}_{i=1, \dots, n}$  afterwards to get the corresponding set  $\{\hat{x}^i\}_{i=1, \dots, n}$  according to Proposition 4.2.

## 5.2 Reconstruction via frames

In the previous section, we described the reconstruction of test images by orthonormalization of the training data and application of the formula in (75). In this section

we present a method, which avoids the necessity of this additional exact backtransformation of the orthogonalized system by avoiding the orthogonalization and directly the original training images and their radon transformations instead.

Since the set  $\{\hat{y}^1, \dots, \hat{y}^n\} \subseteq \mathbb{R}^{M \cdot K}$  is a basis set of  $\mathcal{Y}_n$  by definition, it is also a frame for  $\mathcal{Y}_n$  according to Section 2.3. A corresponding dual frame  $\{\hat{y}'^1, \dots, \hat{y}'^n\} \subseteq \mathbb{R}^{M \cdot K}$  is defined by the property

$$\langle \hat{y}'^i, \hat{y}^j \rangle = \delta_{ij},$$

which allows to represent each vector  $y \in \mathcal{Y}_n$  as

$$y = \sum_{i=1}^n \langle \hat{y}'^i, y \rangle \hat{y}^i.$$

A general  $y \in \mathbb{R}^{M \cdot K}$  can uniquely be written as  $y = y_{\mathcal{Y}_n} + y_{\mathcal{Y}_n^\perp}$  with  $y_{\mathcal{Y}_n} \in \mathcal{Y}_n$  and  $y_{\mathcal{Y}_n^\perp} \in \mathcal{Y}_n^\perp$ . We therefore receive the formula

$$\sum_{i=1}^n \langle \hat{y}'^i, y \rangle \hat{y}^i = \sum_{i=1}^n \langle \hat{y}'^i, y_{\mathcal{Y}_n} + y_{\mathcal{Y}_n^\perp} \rangle \hat{y}^i = \sum_{i=1}^n \langle \hat{y}'^i, y_{\mathcal{Y}_n} \rangle \hat{y}^i + \underbrace{\sum_{i=1}^n \langle \hat{y}'^i, y_{\mathcal{Y}_n^\perp} \rangle \hat{y}^i}_{=:\epsilon} = y_{\mathcal{Y}_n} + \epsilon. \quad (78)$$

The error  $\epsilon \in \mathbb{R}^{M \cdot K}$  can be reduced by choosing  $n$  sufficiently large. For our studies we combine the vectors  $\{\hat{y}^1, \dots, \hat{y}^n\}$  to a matrix  $\hat{Y} \in \mathbb{R}^{M \cdot K \times n}$  and use the *backslash* operator in MATLAB to get a matrix  $\hat{Y}' \in \mathbb{R}^{M \cdot K \times n}$  that fulfills

$$\hat{Y}'^T \cdot \hat{Y} = I.$$

The resulting columns of  $\hat{Y}'$  build then a dual basis  $\{\hat{y}'^1, \dots, \hat{y}'^n\} \subseteq \mathbb{R}^{M \cdot K}$ . With this, we can then calculate the backtransformed  $x_n$  similar to (75) by

$$x_n \approx \sum_{i=1}^n \langle \hat{y}'^i, y \rangle \hat{x}^i. \quad (79)$$

### 5.3 Reconstruction via frame iteration

In this section, we present another method which is based on frame theory in Section 2.3 and further the iterative construction methods of the dual frames in Propositions 2.48 and 2.49. Since we aim to solve the problem  $Tx = y$  for unknown  $x$  (see

Chapter 4) with training data as in Section 5.1.1, we need to find a way to calculate the frame operator (see Definition 2.41) of a vector  $Sx$  within the algorithms. Note at this point, that all frame results in Section 2.3 require  $x \in \mathcal{X}$ , i.e. that the test picture is in the span of the training pictures. This is in general not true, but a reasonable approach for a sufficient number of training pictures. The following derivation shows a possible realization for the Radon transform (see [14, Chapter 2.3]), which we will use in the numerical investigations later on: With the definition of the frame operator in (53) and the **filtered backprojection formula**

$$x = \frac{1}{4\pi} R^* (\mathcal{I}^{-1} \circ R) x, \quad (80)$$

for the Radon transform  $R$ , the Hilbert transform  $\mathcal{I}^{-1}$ , the backprojection  $R^*$  (adjoint with respect to  $\ell^2$  norm) and an arbitrary  $x$ , we get

$$\begin{aligned} Sx &= \sum_{i=1}^n \langle x, \hat{x}^i \rangle \hat{x}^i = \sum_{i=1}^n \langle x, R^*(R^{-1})^* \hat{x}^i \rangle \hat{x}^i = \sum_{i=1}^n \langle \underbrace{Rx}_{=y}, (R^{-1})^* \hat{x}^i \rangle \hat{x}^i \\ &= \frac{1}{4\pi} \sum_{i=1}^n \langle y, (\mathcal{I}^{-1} \circ R) \hat{x}^i \rangle \hat{x}^i = \frac{1}{4\pi} \sum_{i=1}^n \langle y, \mathcal{I}^{-1} \hat{y}^i \rangle \hat{x}^i. \end{aligned} \quad (81)$$

The result shows that it is sufficient to determine the filter  $\mathcal{I}^{-1}$  of the given training data  $\{\hat{y}^i\}_{i=1,\dots,n}$ . This is a subfunction in the Image Processing Toolbox of MATLAB. The remaining part of the iteration methods follows exactly the formulas in Propositions 2.48 and 2.49. Since the possible frame bounds  $A$  and  $B$  in 2.39 depend on the chosen test data and are therefore in general unknown in our case, we set  $A + B = 10^6$  for the algorithm according to Proposition 2.48. This value was determined experimentally and led to fairly good convergence. The number of iteration steps is fixed to  $10^4$  for both algorithms, since optical observations of the results lead to no significant change above  $10^3$  iteration steps, so we assume already very good convergence after  $10^4$  steps.

### 5.3.1 Numerical effort

In this section we analyse the numerical effort of the iterative frame algorithms and compare it to the effort of the orthonormalization methods. Since the computational effort of the Hilbert transform  $\mathcal{I}^{-1}$  is  $N \cdot K$  as given in [15] and we have to calculate additionally an inner product ( $\mathcal{O}(M \cdot K)$  flops) and a multiplication ( $\mathcal{O}(N)$  flops) in the sum of (81), the frame operator has a total effort of  $\mathcal{O}(n \cdot (M \cdot K + M \cdot K + N)) \simeq \mathcal{O}(n \cdot (2 \cdot M \cdot K + N))$ . This is calculated in  $i$  iteration steps, so we finally get a

computational effort of  $\mathcal{O}(i \cdot n \cdot (2 \cdot M \cdot K + N))$  for the whole iteration of the algorithms in Propositions 2.48 and 2.49. Note that this effort is just linear in  $n$ , which is the number of training pictures. If we compare that dependency to the convergence rates of Sec. 5.1.2–5.1.4, we see that the effort of the frame reconstruction is lower, since the orthonormalization step is avoided. We observe this reduced effort in Fig. 1. However, we will observe in the following sections, that this reduction of effort reduced also the quality of the reconstruction.

## 5.4 Numerical investigations

In this chapter we present the results of some numerical investigations.

### 5.4.1 Accuracy of the orthogonalization methods

In the orthogonalization step, we have four possible methods of sections 5.1.2–5.1.5. We now investigate the stability of the algorithms on the example of our imaging application. We assume the set  $\{\hat{\underline{y}}^i\}_{i=1,\dots,n}$  being available and analyse the error of the resulting orthonormal system. We first define an appropriate measure of this error, which quantifies the success of the orthogonalization procedure as a numerical value.

**Definition 5.1.** For a set  $\{\hat{\underline{y}}^i\}_{i=1,\dots,n} \subset \mathbb{R}^{M \cdot K}$ , which is assumed to be approximately orthonormal, we define the matrix  $\underline{Y} \in \mathbb{R}^{n \times n}$  by

$$(\underline{Y})_{ij} := \langle \hat{\underline{y}}^i, \hat{\underline{y}}^j \rangle. \quad (82)$$

Furthermore, we define the orthonormality error  $\epsilon_{ortho}$  by

$$\epsilon_{ortho} := \|\underline{Y} - I\|_1, \quad (83)$$

where  $I$  denotes the  $\mathbb{R}^{n \times n}$  identity matrix and  $\|\cdot\|_1$  maximum absolute column sum of the matrix.

**Remark 5.2.** For an orthonormal system  $\{\hat{\underline{y}}^i\}_{i=1,\dots,n}$ , we observe

$$(\underline{Y})_{ij} = \langle \hat{\underline{y}}^i, \hat{\underline{y}}^j \rangle = \delta_{ij} = I$$

and therefore  $\epsilon_{ortho} = 0$ . If the set is only an approximate orthonormal system, the matrix  $\underline{Y}$  is only close a identity matrix, so  $\epsilon_{ortho}$  becomes positive.

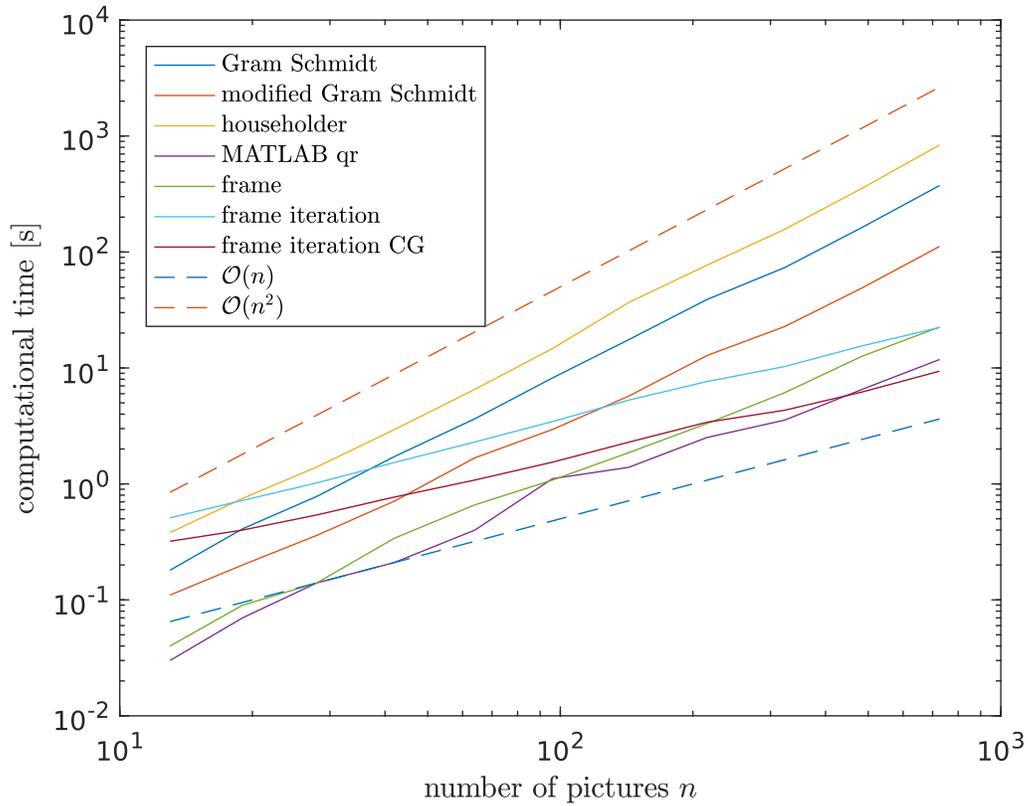
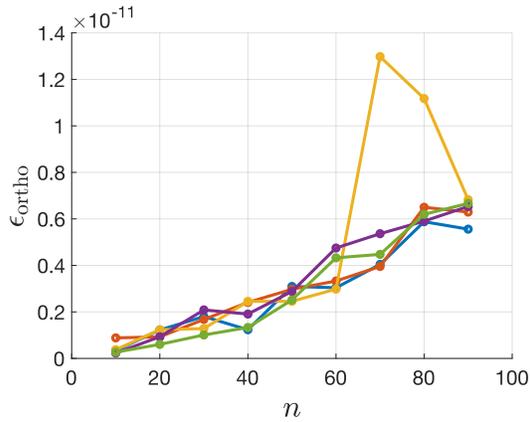
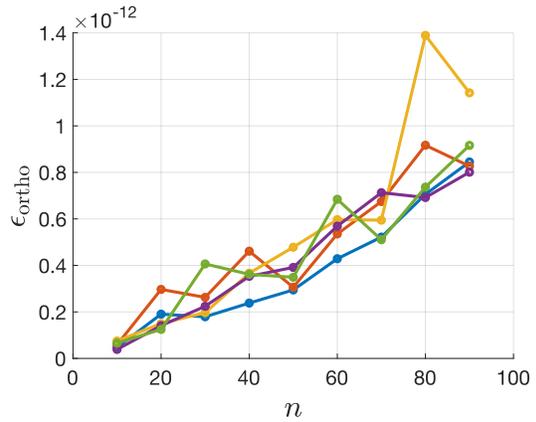


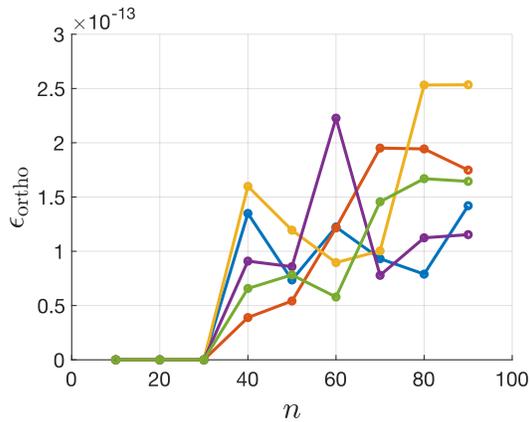
Figure 1: Computation time vs. number of training data for the different methods with orthonormalization and frames. Whereas the effort of all four methods based on orthonormalization increases at least quadratically, the effort of the frame iteration methods increases linearly with the number of pictures.



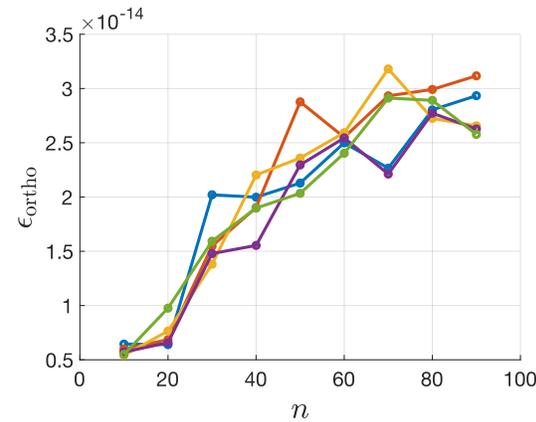
(a) Gram-Schmidt method



(b) Modified Gram-Schmidt method



(c) Householder reflections



(d) QR decomposition in MATLAB

Figure 2: Error  $\epsilon_{\text{ortho}}$  for different numbers of images  $n$ . Each color represents a different sequential order of the vectors, to which each method is applied.

We test the methods on the example of the Sunflower data set with different numbers of images and plot the error  $\epsilon_{\text{ortho}}$  over the number  $n$ . We furthermore investigate the impact of random permutation of the images, before the method is applied. The results can be observed in Fig. 2.

We see that the error increases with increasing number of images. The sequential order of the images generates slight deviations, but the increasing trend remains similar. We can furthermore observe that the numerical errors of the Gram Schmidt method ( $\sim 10^{-11}$ ) and modified Gram Schmidt method ( $\sim 10^{-12}$ ) are 2 to 4 magnitudes larger than the orthonormalization errors of the Householder reflections ( $\sim 10^{-14}$ ) and the QR decomposition in MATLAB ( $\sim 10^{-15}$ ).

### 5.4.2 Visual observations

In this section, we want to observe the reconstructed images of the methods visually. Applying the method on the Radon transforms of the test images, we can compare the original test image with the output of our algorithm. Again, this is possible with all orthonormalization and frame theory methods. Furthermore, we also compare the projected Radon transform  $Q_n y$  according to Equation (74) to the radon transform of the original image for the orthonormalization methods. In the frame methods, there is no similar equation available.

**Sunflowers** We use  $n = 726$  training images ( $150 \times 150$  pixels each) of the sunflower data set. Seven additional images, which are not part of the training images are used as test images. These test images contain 4 images with typical motives of sunflowers, where a good approximation on base of the training data is supposed and further 3 images with atypical content, which do not fit to the training data that well. On each test image, the reconstruction procedure is applied individually. The results can be seen in Figs. 3 and 5. The Figs. 4 and 6 show the same results for permuted images as input for the orthogonalization methods. We observe a better similarity of the pictures in Fig. 3, since due to the similarity of sunflowers, the radon transformations of sunflower motives can assumed to be "closer" to the finite dimensional subspace spanned by the training data, than other arbitrary motives. Therefore the projection error  $(I - Q_n)y$  of the map onto the subspace  $Q_n$  in (74) might be smaller and therefore the solution  $x_n$  in (75) lies closer to the original image. The outcome of the variant with permuted images look almost identical compared the to original results, so the sorting order of the sunflower pictures does not seem to affect the results significantly. Furthermore we could see, that the orthogonalization methods perform at almost the same level. In the results of the frame methods, we

clearly see the systematic error described in Sec. 5.2 and 5.3.

**Digits** Similar observations are done with a digits data set with  $n = 95,495,995$  and 4995 training images ( $28 \times 28$  pixels each) and 5 test images in Figs. 7–10. We see a clear improvement of the result with the frame method for increasing training images. We also observe that the reconstruction via frames of Sec. 5.2 improves a lot for increasing number of pictures. The reason for this behavior is the small value for  $\epsilon$  in (78), if the large number of pictures already leads to  $\mathcal{Y}_n \approx \mathcal{Y}$  and therefore  $y_{\mathcal{Y}_n^\perp}$  becomes very small.

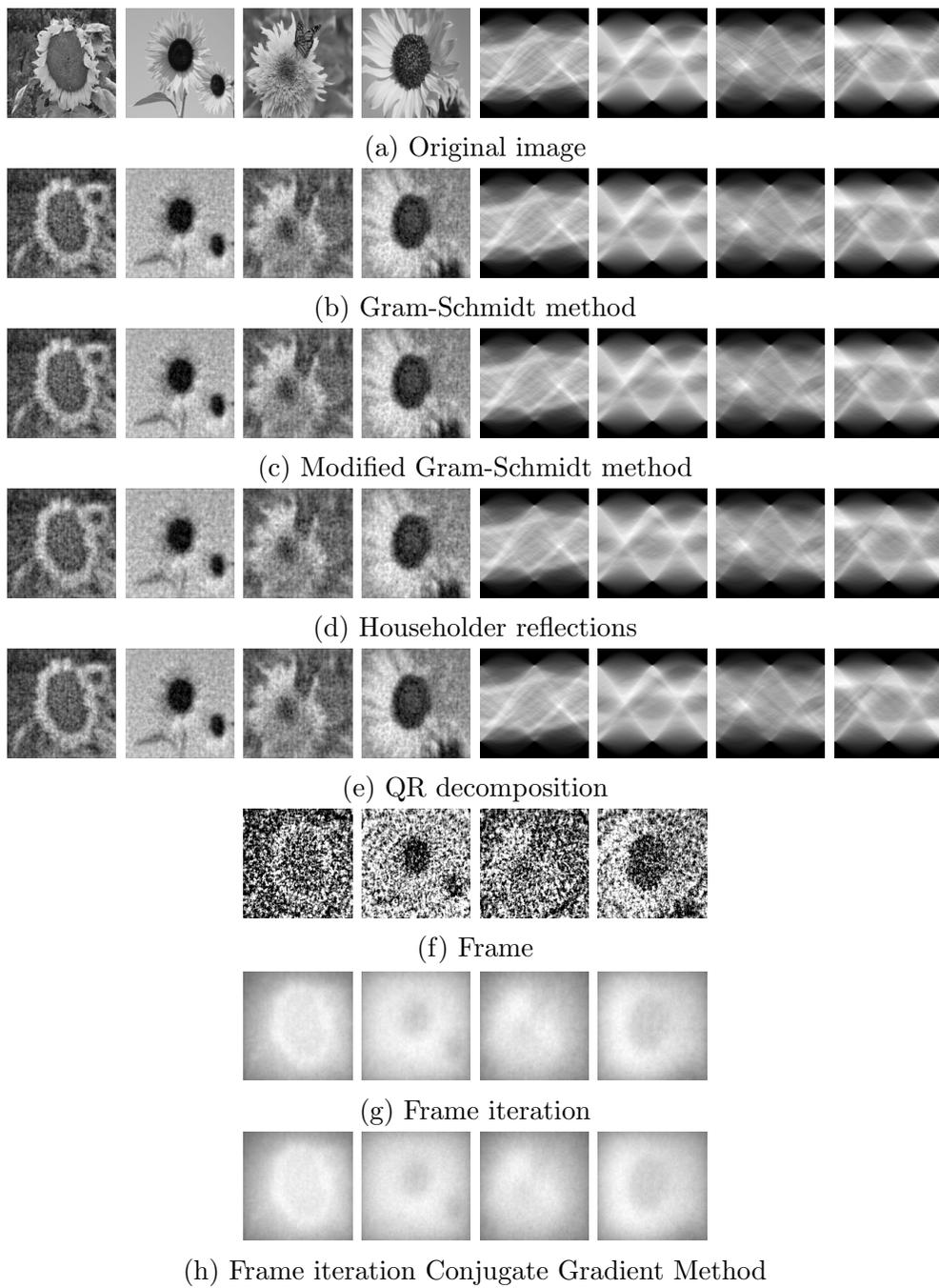


Figure 3: Reconstructed images via the different methods in comparison to the original image for sunflower pictures.

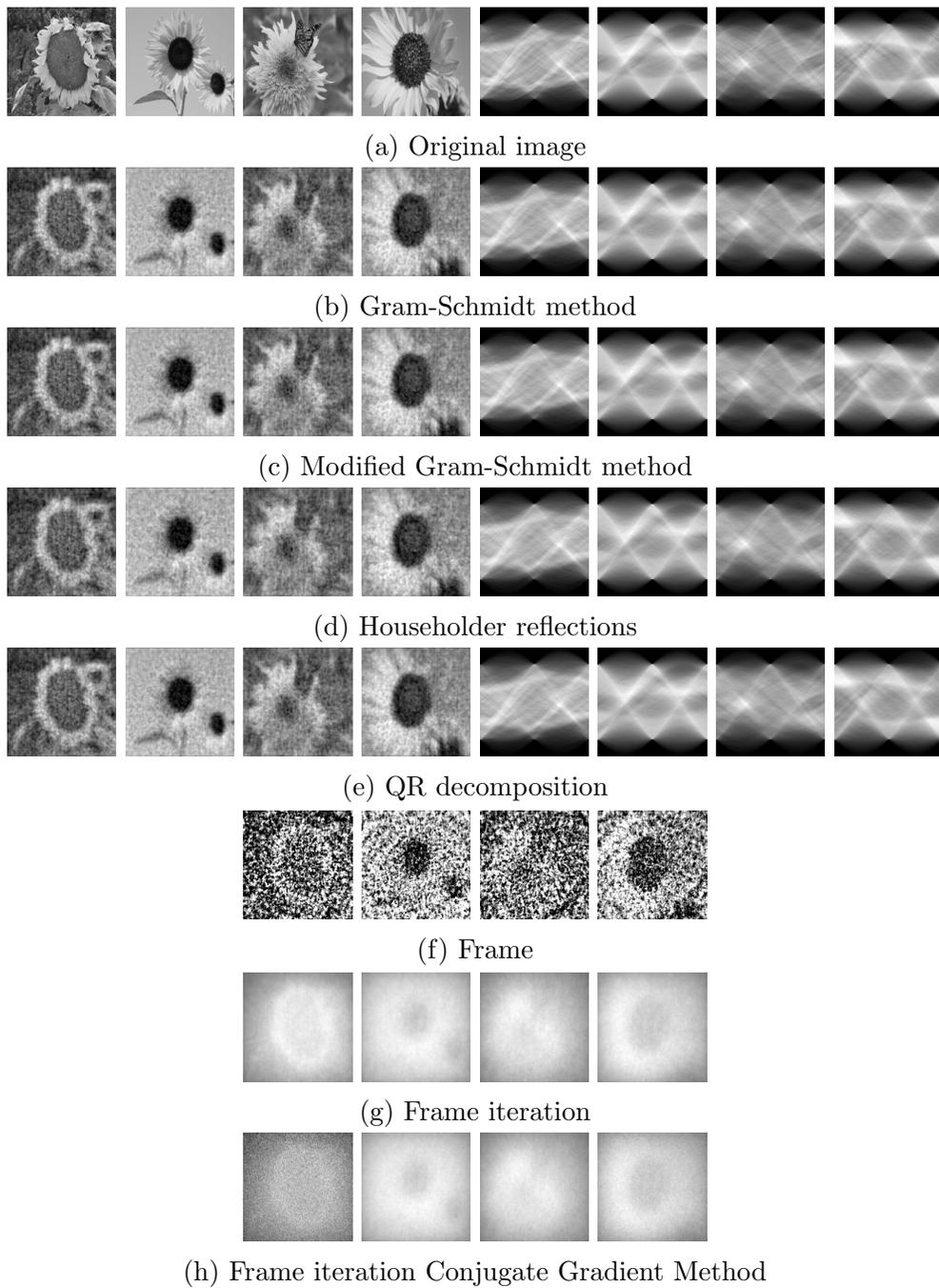


Figure 4: Reconstructed images via the different methods with permuted training image data set as input in comparison to the original image for sunflower pictures.

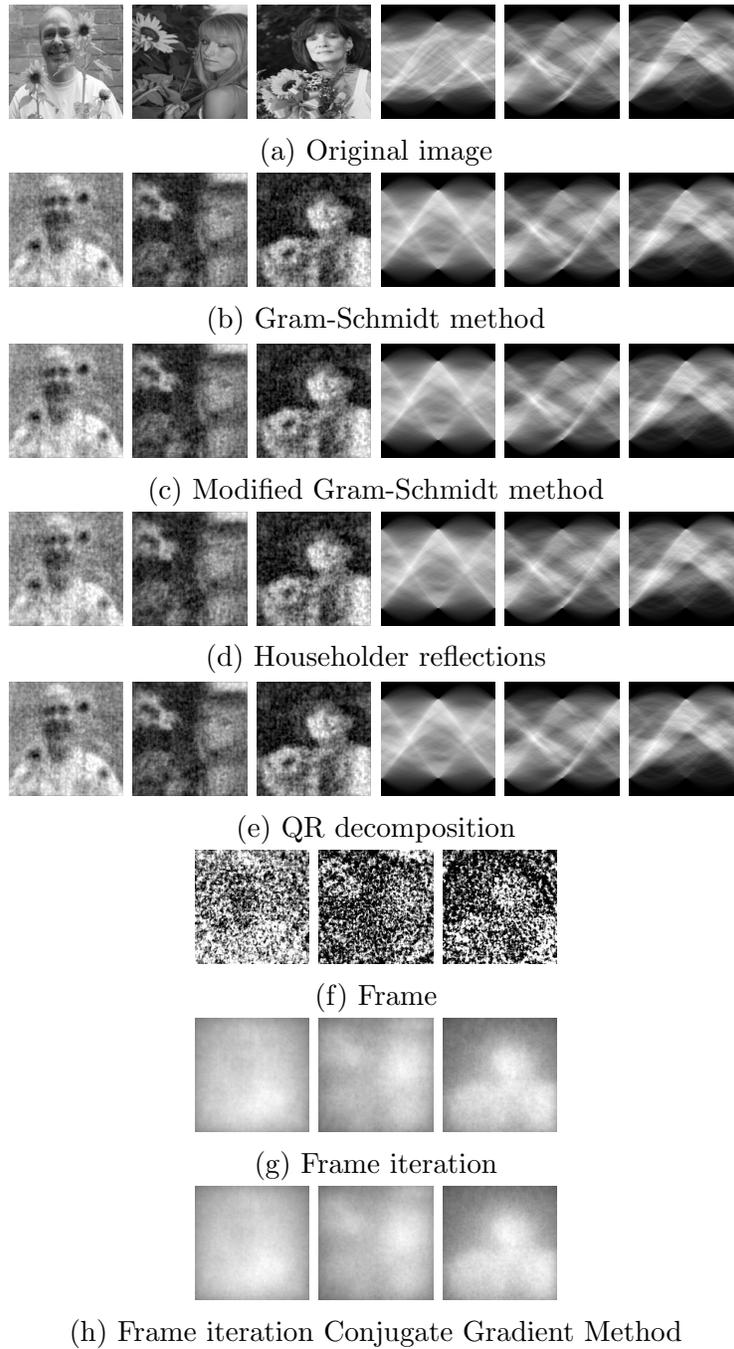


Figure 5: Reconstructed images via the different methods in comparison to the original image for untypical pictures of the data set, which contain a person and the sunflower motive is only incidental.

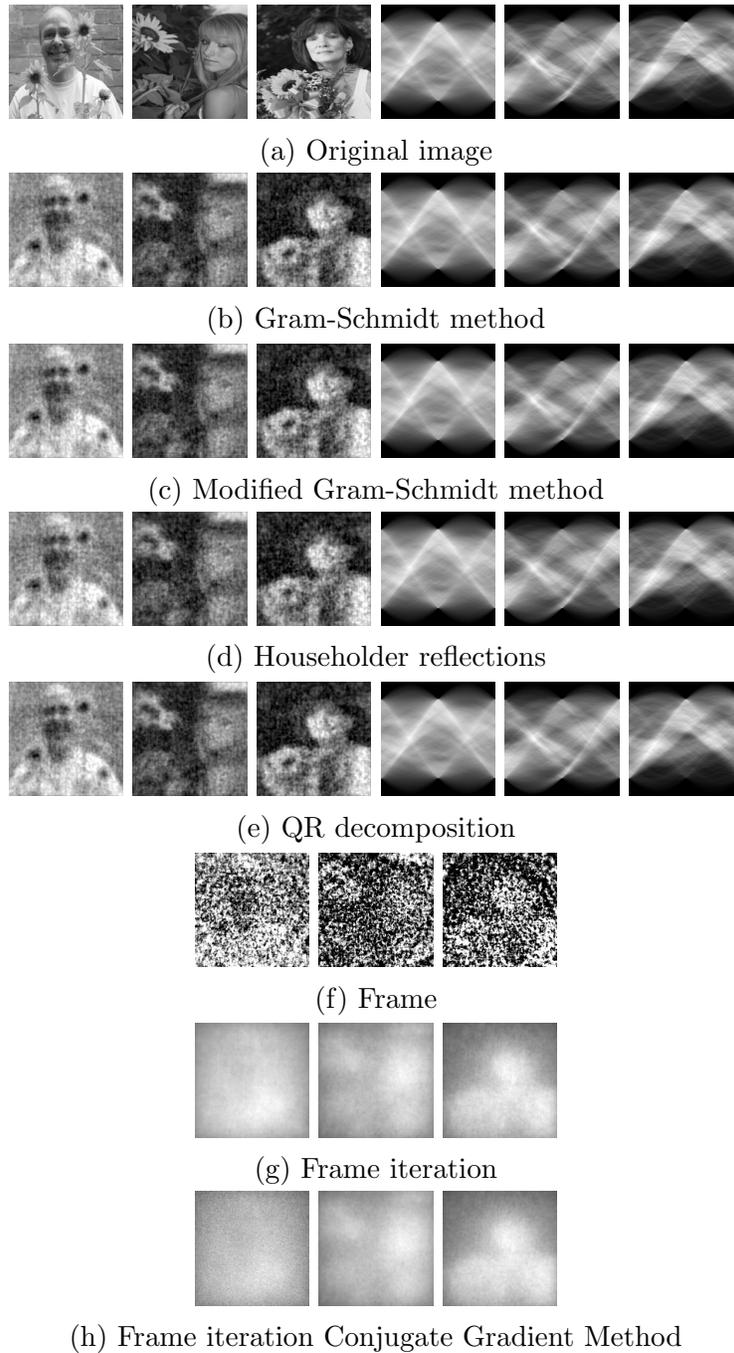
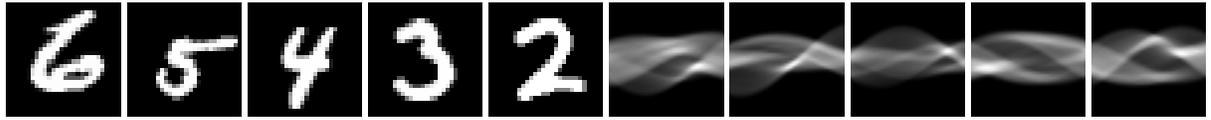
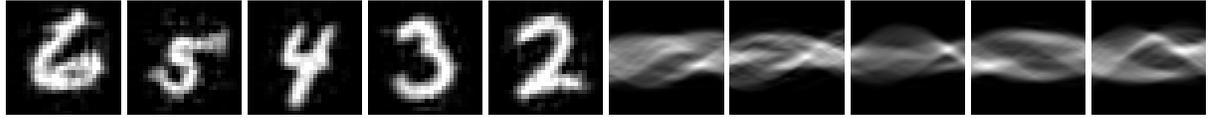


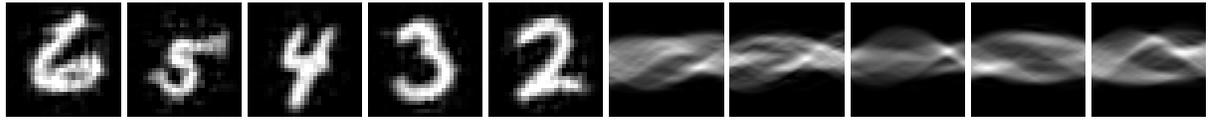
Figure 6: Reconstructed images via the different methods with permuted order of the training data as input in comparison to the original image for untypical pictures of the data set, which contain a person and the the sun flower motive is only incidental.



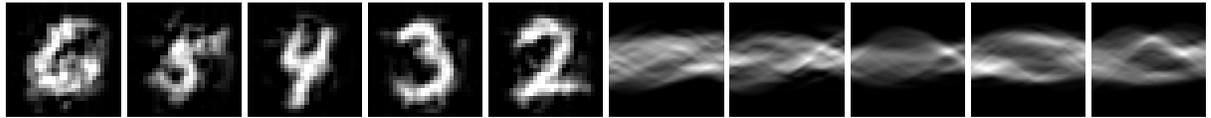
(a) Original image



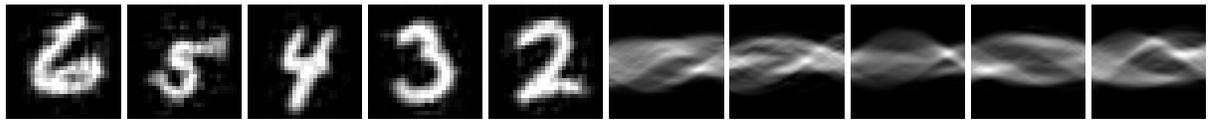
(b) Gram-Schmidt method



(c) Modified Gram-Schmidt method



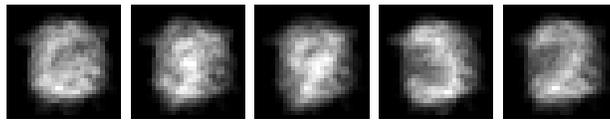
(d) Householder reflections



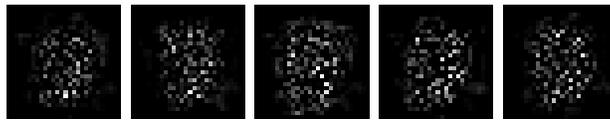
(e) QR decomposition



(f) Frame

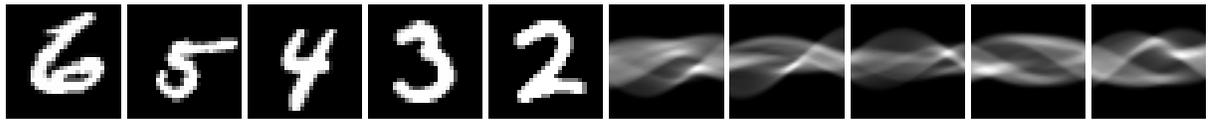


(g) Frame iteration

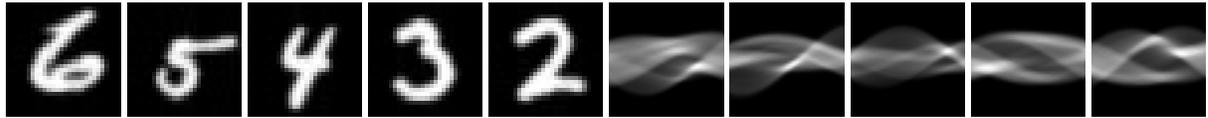


(h) Frame iteration Conjugate Gradient Method

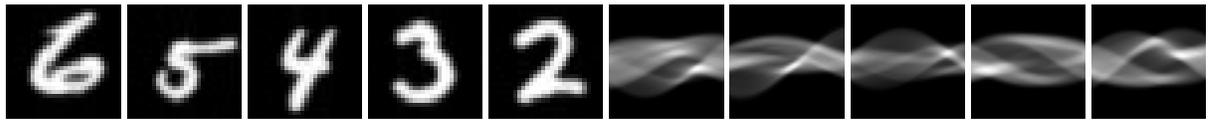
Figure 7: Number of training images: 100.



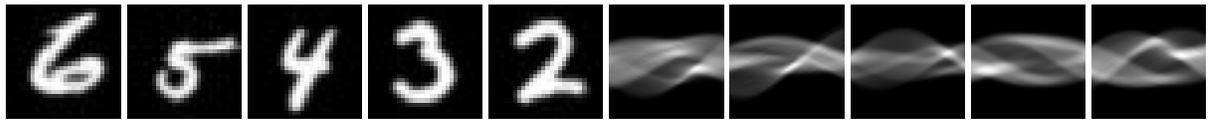
(a) Original image



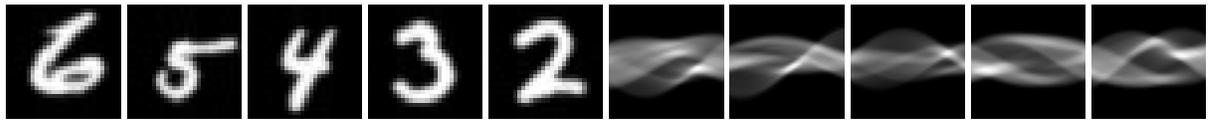
(b) Gram-Schmidt method



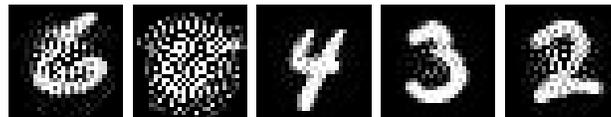
(c) Modified Gram-Schmidt method



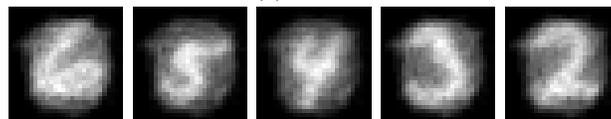
(d) Householder reflections



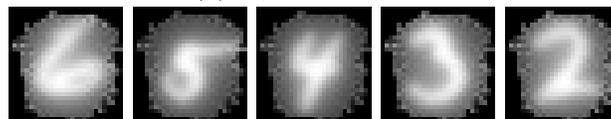
(e) QR decomposition



(f) Frame

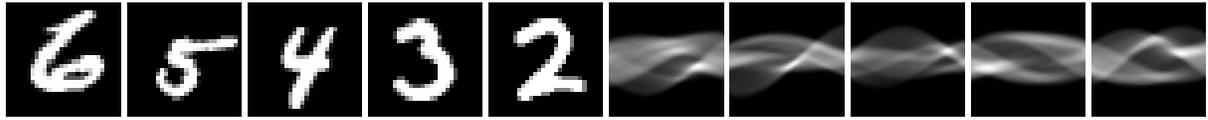


(g) Frame iteration

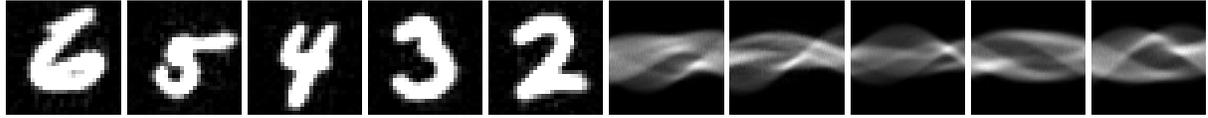


(h) Frame iteration Conjugate Gradient Method

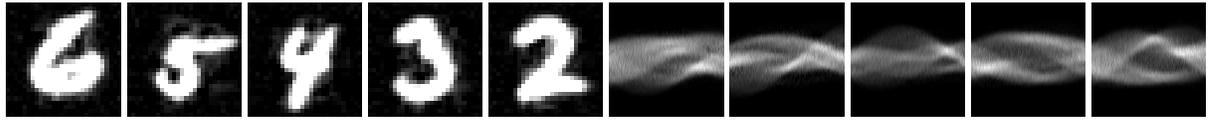
Figure 8: Number of training images: 500.



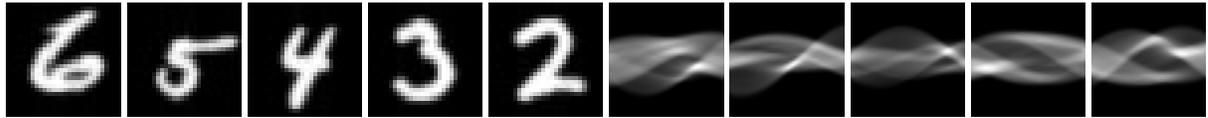
(a) Original image



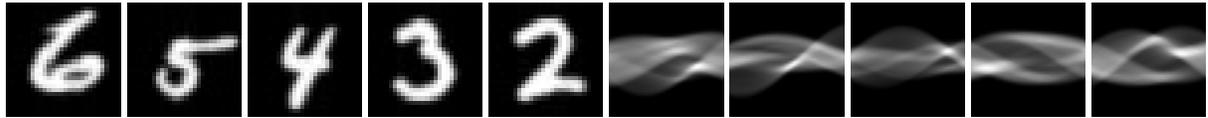
(b) Gram-Schmidt method



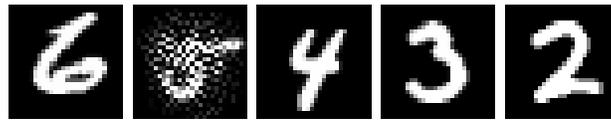
(c) Modified Gram-Schmidt method



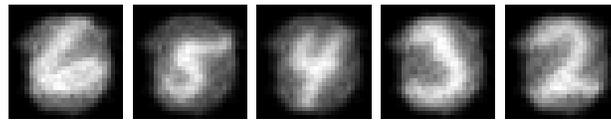
(d) Householder reflections



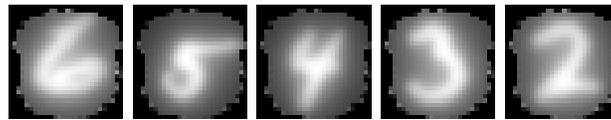
(e) QR decomposition



(f) Frame

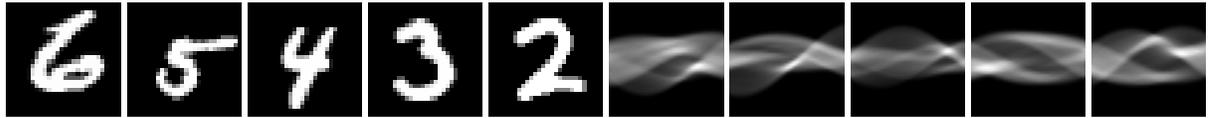


(g) Frame iteration

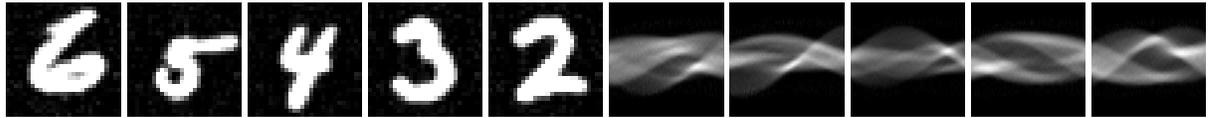


(h) Frame iteration Conjugate Gradient Method

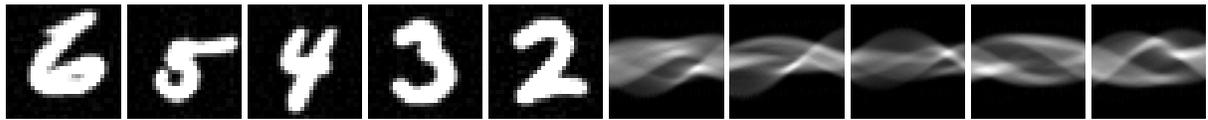
Figure 9: Number of training images: 1000.



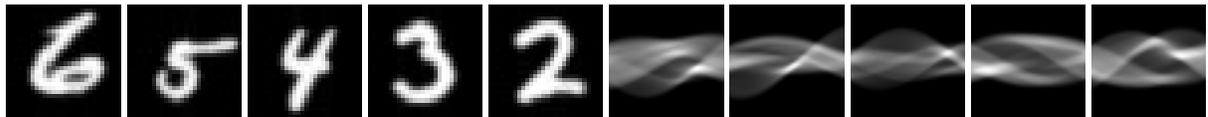
(a) Original image



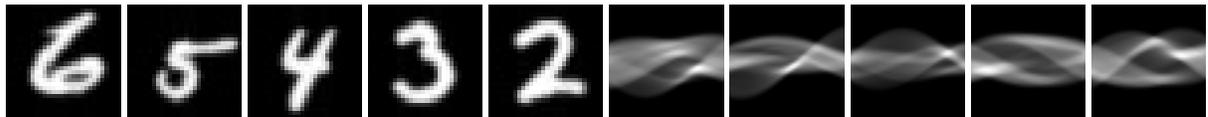
(b) Gram-Schmidt method



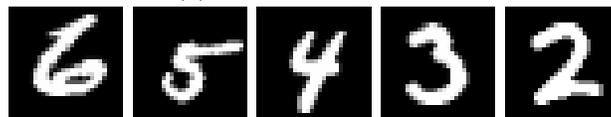
(c) Modified Gram-Schmidt method



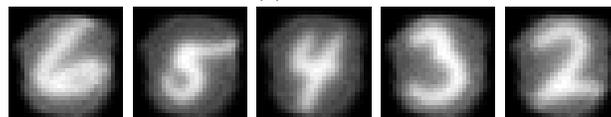
(d) Householder reflections



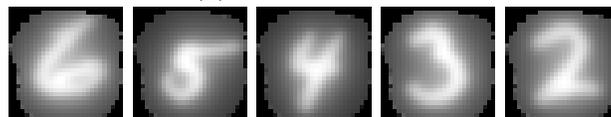
(e) QR decomposition



(f) Frame



(g) Frame iteration



(h) Frame iteration Conjugate Gradient Method

Figure 10: Number of training images: 5000.

## 6 Summary and conclusion

In this thesis, we presented and investigated the regularization by projection on the example of the Radon transform. In the first part of the work, we gave an introduction into spectral and frame theory to build a foundation of the following chapter about regularization methods. We have the goal to find an unknown operator, which maps between two sets. For given training data pairs, this operator can be approximated. As an application, we demonstrate the reconstruction of the inverse Radon transform with training data of various image data sets. The functionality of the methods was afterwards demonstrated on test data. An important component of the reconstruction was the orthonormalization of the data. Therefore different orthonormalization methods were used and both quantitatively and qualitatively compared. Furthermore, alternative approaches based on frame theory instead of orthonormalization were presented and discussed. We could show, that depending on the dimension of the vector spaces, which are given by the size and number of pictures, the methods have different properties in terms of performance and quality of the output. With a smaller amount of training data, we observe a better visual output with the orthogonalization methods, whereas a high amount of training data leads to a better output with frames. In the latter case, the frame iteration methods are even preferable in terms of computational effort, since they increase only linearly with the number of pictures, whereas the orthogonalization does at least quadratic. We furthermore observed the numerical precision of the different orthogonalization methods in MATLAB arithmetic and compared the residual of the orthogonal (up to rounding errors) outcomes and gave several visual examples to illustrate the functional of our methods.

## References

- [1] Engl, Heinz W, et al. Regularization of Inverse Problems. Dordrecht, Kluwer Academic Publisher, 2000.
- [2] Dautray, Robert, and Jacques Louis Lions. Mathematical Analysis and Numerical Methods for Science and Technology. 3, Spectral Theory and Applications. Berlin; Heidelberg; New York; Barcelona; Hong Kong; London, Springer, 2000.
- [3] Harro Heuser. Funktionalanalysis: Theorie und Anwendung. Stuttgart, Teubner, 1986.
- [4] Engl, Heinz W. Necessary and Sufficient Conditions for Convergence of Regularization Methods for Solving Linear Operator Equations of the First Kind. Numerical Functional Analysis and Optimization, vol. 2, 201-222, 1981.
- [5] Arens, Tilo. Mathematik. Springer Berlin Heidelberg, 2018.
- [6] Seidman, Thomas I. "Nonconvergence Results for the Application of Least-Squares Estimation to Ill-Posed Problems." Journal of Optimization Theory and Applications, vol. 30, no. 4, Apr. 1980, pp. 535–547, 10.1007/bf01686719.
- [7] Natterer, Frank. "Regularisierung Schlecht Gestellter Probleme Durch Projektionsverfahren." Numerische Mathematik, vol. 28, no. 3, Sept. 1977, pp. 329-341, 10.1007/bf01389972.
- [8] Andrea Aspri and Yury Korolev and Otmar Scherzer. "Data driven regularization by projection", arXiv:1909.11570v1, 2019.
- [9] Karpfinger, Christian. Höhere Mathematik in Rezepten: Begriffe, Sätze Und Zahlreiche Beispiele in Kurzen Lerneinheiten. Berlin, Springer Spektrum, 2017.
- [10] Mo Chen (2020). Pattern Recognition and Machine Learning Toolbox (<https://www.github.com/PRML/PRMLT>), GitHub. Retrieved April 28, 2020.
- [11] Golub, Gene H. and Van Loan, Charles F.. Matrix Computations, Third Edition. Baltimore, The John Hopkins University Press, 1996.
- [12] Casazza, Peter G. and Kutyniok, Gitta. Finite Frames: Theory and Applications. New York, Springer Science, Business Media, 2013.
- [13] Quarteroni, Alfio and Sacco, Riccardo and Saleri, Fausto. Numerical Mathematics. New York, Springer-Verlag, 2000.

- [14] Scherzer, Otmar. Lecture notes to Inverse Problems, 2017.
- [15] Basu, Samit and Bresler, Yoram.  $O(N^2 \log_2 N)$  Filtered Backprojection Reconstruction Algorithm for Tomography. IEEE Transactions on Image Processing 9(10):1760-1773, 2000.
- [16] Nashed, M. Zuhair. Generalized Inverses and Applications, Proceeding of an Advanced Seminar. New York, Academic Press, 1976.
- [17] Kelley, C. T. Iterative Methods for Linear and Nonlinear Equations. Society for Industrial and Applied Mathematics. Philadelphia, SIAM, 1995.
- [18] Kuchment, Peter. The Radon Transform and Medical Imaging. Society for Industrial and Applied Mathematics. Philadelphia, SIAM, 2014.