Radar Target Classification via Sparse Decomposition

DISSERTATION
zur Erlangung des Grades eines Doktors
der Ingenieurwissenschaften

vorgelegt von
Simon Wagner M.Sc.

eingereicht bei der Naturwissenschaftlich-Technischen Fakultät
der Universität Siegen
Siegen 2022

Betreuer und erster Gutachter
Prof. Dr.-Ing. Joachim Ender
Universität Siegen

Zweiter Gutachter
Prof. Dr.-Ing. Otmar Loffeld
Universität Siegen

Tag der mündlichen Prüfung
20. April 2022

Für Katharina, Levi und Amalia

Zusammenfassung

Die Erkennung und Identifizierung von Zielen sind entscheidende Schritte in der Kette der Radarsignalverarbeitung. Aufgrund der hohen Auflösung sind abbildende Radare für diese Aufgaben gut geeignet. In dieser Arbeit wird ein Verfahren vorgestellt, das auf der Zerlegung von Radarbildern in verschiedene Streumechanismen basiert. Das Ziel dieses Verfahrens ist die Identifizierung von spezifischen Streuzentren bestimmter Ziele. Diese spezifischen Streuzentren können in weiteren Schritten der Identifizierung genutzt werden um z.B. die Menge der möglichen Ziele einzuschränken.

Als Anwendungsbeispiel wird in dieser Arbeit die Trennung der Echos von Strahltriebwerken sowie von isotropen Streuzentren genutzt. Zur Modellierung der Echos von Strahltriebwerken wird ein Wellenleitermodell genutzt, welches sehr verbreitet in der Modellierung von Radarechos ist. Die zum Verständnis des Modells notwendigen Grundlagen der Ausbreitung elektromagnetischer Wellen in Wellenleitern werden im Rahmen dieser Arbeit erläutert. Abschließend steht ein universelles Modell zur Verfügung, das mit beliebigen Wellenformen kombiniert werden kann. Beispielhaft wird die Veränderung des Zeit-Frequenz-Verhaltens einer Chirp-Wellenform dargestellt.

Die Trennung der Echos selbst erfolgt mit einem Algorithmus aus dem Bereich der dünnbesetzten Gleichungssysteme. Aus diesem Bereich der Mathematik werden einige grundlegende Zusammenhänge sowie elementare Lösungsverfahren vorgestellt. Das Verfahren das hier zum Einsatz kommt, wird als morphologische Komponentenanalyse bezeichnet, und wurde entwickelt um Signale in Komponenten mit unterschiedlicher Struktur zu zerlegen. Aufgrund dieser Historie erscheint dieses Verfahren als besonders geeignet, das zugrundeliegende Problem zu lösen, und wird daher im Detail vorgestellt.

Das grundsätzliche Vorgehen des Verfahrens wird anschließend durch eine Simulation mit verschiedenen kanonischen Formen, d.h. Punkten, Quadraten, Kreisen, Kreuzen und Linien, dargestellt. Um die Fähigkeit des Algorithmus zur Trennung von Triebwerksechos und isotropen Punktstreuern zu bewerten wird eine Simulation mit diesen beiden Arten von Echos durchgeführt. Abschließend wird ein realer Datensatz des Tracking and Imaging Radars (TIRA) des Fraunhofer FHR verwendet, um Ergebnisse mit einem realen Radarsystem zu zeigen.

Abstract

Recognition and Identification of targets are crucial steps in the radar signal processing chain. Due to their high resolution, imaging radars are a well suitable choice for these tasks. This thesis presents a framework, which is based on sparse decomposition of radar images, to identify different kinds of scattering mechanisms. The aim of this framework is the identification of specific parts of certain targets. These specific scattering centers can be used in further steps of identification, e.g. to limit the number of possible targets.

As application example, the separation of echoes from jet engines and isotropic scattering centers is used in this thesis. To model the echoes from jet engines a waveguide model is used, which is common in the modeling of radar echoes. The basic principles of electromagnetic wave propagation in waveguides, which are necessary to understand the model, are explained in this thesis. Finally, a universal model is available, which can be combined with arbitrary waveforms. As an example the change of the time-frequency behavior of a chirp waveform is shown.

The separation of the echoes themselves is done with an algorithm from the domain of solving sparse systems of equations. From this area of mathematics, some basic relationships and elementary solution methods are presented. The method used in this thesis is called morphological component analysis and was developed to decompose signals into components with different structures. Based on this history, this method seems to be particularly suitable to solve the underlying problem and is therefore presented in detail.

The basic procedure of the method is then illustrated by a simulation with different canonical forms, i.e. points, squares, circles, crosses and lines. To evaluate the ability of the algorithm to separate engine echoes and isotropic point scatterers, a simulation with these two types of echoes is performed. Finally, a real data set of the Tracking and Imaging Radar (TIRA) of Fraunhofer FHR is used to show results with a real radar system.

Parts of the results of this dissertation have been published in:

Simon Wagner, Joachim Ender: *Scattering Identification in ISAR Images via Sparse Decomposition*, 2022 IEEE Radar Conference (RadarConf22), 2022

Simon Wagner, Joachim Ender: *Jet engine recognition via sparse decomposition of ISAR images*, Chapter 6 in *New Methodologies for Understanding Radar Data*, edited by A.K. Mishra and S. Brüggenwirth, Scitech Publishing, The Institution of Engineering and Technology, p. 253 – 284, Stevenage, UK, 2021

Simon Wagner, Joachim Ender: *Recognition of jet engines via sparse decomposition of ISAR images using a waveguide scattering model*, International Journal of Microwave and Wireless Technologies, Vol. 9, No. 6, Cambridge University Press, p. 1339 – 1343, 2017

Simon Wagner, Franz Dommermuth, Joachim Ender: *Detection of Jet Engines via Sparse Decomposition of ISAR Images for Target Classification Purposes*, 13th European Radar Conference (Eurad 2016), p. 77 – 80, 2016

Simon Wagner: *Morphological Component Analysis in SAR Images to Improve the Generalization of ATR Systems*, Compressed Sensing Theory and its Applications to Radar, Sonar and Remote Sensing, 3rd International Workshop on (CoSeRa), p. 46 – 50, 2015

Further publications of the author not related to this thesis:

Simon Wagner, Winfried Johannes, Stefan Brüggenwirth: *Background Subtraction by a Contractive Autoencoder to Detect Small Targets in Clutter*, NATO STO Specialists' Meeting SET-273 on Multidimensional Radar Imaging for ATR, 2021

Simon Wagner, Chandana Panati, Stefan Brüggenwirth: *Fool the COOL - On the Robustness of Deep Learning SAR ATR Systems*, 2021 IEEE Radar Conference (RadarConf21), 2021

Stefan Brüggenwirth, Simon Wagner, Tanja Bieker, Nicola Battisti, Vincenzo Rispoli, Mario Greco, Gianpaolo Pinelli, Davide Cataldo, Marco Martorella: *Results on Super-Resolution and Target Identification techniques from the SPERI Project*, IEEE Aerospace and Electronic Systems Magazine, Vol. 36, No. 3, p. 24 – 35, 2021

Stefan Brüggenwirth, Simon Wagner: *Theoretical foundations of deep learning*, Chapter 3 in *Deep Neural Network Design for Radar Applications*, edited by S.Z. Gurbuz, Scitech Publishing, The Institution of Engineering and Technology, p. 69 – 96, Stevenage, UK, 2020

Simon Wagner, Stefan Brüggenwirth: *Deep neural network design for SAR/ISAR-based automatic target recognition*, Chapter 8 in *Deep Neural Network Design for Radar Applications*, edited by S.Z. Gurbuz, Scitech Publishing, The Institution of Engineering and Technology, p. 243 – 277, Stevenage, UK, 2020

Stefan Brüggenwirth, Marcel Warnke, Simon Wagner, Kilian Barth: *Cognitive Radar for Classification*, IEEE Aerospace and Electronic Systems Magazine, Vol. 34, No. 12, p. 30 – 38, 2019

Simon Wagner, Winfried Johannes: *Target Detection using Autoencoders in a Radar Surveillance System*, International Radar Conference (RADAR), 2019

Stefan Brüggenwirth, Marcel Warnke, Christian Bräu, Simon Wagner, Tobias Müller, Pascal Marquardt, Fernando Rial: *Sense Smart, Not Hard: A Layered Cognitive Radar Architecture*, Chapter 7 in *Topics in Radar Signal Processing*, edited by G. Weinberg, IntechOpen, p. 167 – 196, London, UK, 2018

Dario Ligori, Simon Wagner, Luca Fabbrini, Mario Greco, Tanja Bieker, Gianpaolo Pinelli, Stefan Brüggenwirth: *Nonparametric ISAR autofocusing via entropy-based Doppler centroid search*, IEEE Geoscience and Remote Sensing Letters, Vol. 15, No. 11, p. 1725 – 1729, 2018

Simon Wagner, Kilian Barth, Stefan Brüggenwirth: *A deep learning SAR ATR System using regularization and prioritized classes*, 2017 IEEE Radar Conference (RadarConf17), p. 772 – 777, 2017

Simon Wagner: *SAR ATR by a combination of convolutional neural network and support vector machines*, IEEE Transactions on Aerospace and Electronic Systems, Vol. 52, No. 6, p. 2861 – 2872, 2016

Simon Wagner: *Increasing the Performance of ATR Systems using Artificial Training Data*, NATO STO Specialists' Meeting SET-228 on Radar Imaging for Target Identification, 2015

Simon Wagner: *Combination of Convolutional Feature Extraction and Support Vector Machines for Radar ATR*, 17th International Conference on Information Fusion (FUSION), p. 1 – 6, 2014

Contents

Chapter 1

Introduction

Within the big field of radar there exist many research fields, one of the most basic is the detection of targets, but after the detection of a target an open question remains: "What has been detected?". The task of classification and identification of radar targets is the research field in which this work is placed.

1.1 The need for Classification Systems

The use of surveillance systems, like air traffic control or driver assistance to avoid collisions on the street, has become part of everyday life. In all of these systems it is necessary to know the number and types of other traffic participants to decide how to act next. In tracking systems, classifiers can be used to support the tracker in ambiguous situations [NKOW02] and by that enhance the safety of transport and navigation systems [Neb95].

In modern air traffic the use of transponders has become standard, and in many airspaces mandatory, as identification method. These transponders are used to communicate with so-called secondary radar systems and are therefore communication rather than radar systems. First the airplane is detected by the primary radar, the actual radar system, and after that the secondary radar identifies the target. These transponders are a sort of *Identification Friend or Foe* (IFF) system, which are used in civil and military airplanes [Bow85]. Unfortunately, there are some examples in history where the IFF system was not working properly and caused disastrous accidents. One example is the shoot-down of Iran Air 655 in 1988 [Fis13], where according to the official statement, the Airbus was mistaken as an F-14 fighter aircraft due to a wrong identification [Neb95]. More recent examples for accidental downings are the Malaysia Airlines flight MH17 over the Ukraine in 2014 [Dut15] and the Ukraine International Airlines flight PS752 from Tehran, Iran to Ukraine in January 2020, which has been mistaken as a cruise missile [BBC20].

These examples show that the IFF system is not completely secure and that further methods, which do not rely on a communication between the radar and the target, are necessary. These methods are called *noncooperative methods*, and a recent example, where this kind of system would have been helpful is the Malaysia Airlines flight MH370. In that case the IFF system was switched off on purpose and the aircraft could not be identified anymore by the secondary radar. It is assumed, but not confirmed, that the aircraft was detected heading towards the Indian Ocean by a military radar [MBK16]. If this radar would have been capable of noncooperative identification, the search area could have been specified earlier.

1.2 State-of-the-Art in Target Classification and Scattering Identification

Radar target recognition has been studied for a long time already. In this section, a review of a part of the corresponding literature is given and no claims to completeness are made. The basic application of the algorithms presented in Section 1.2.1 is the comparison of a measured radar signature with a database of known targets. There are mainly two groups of systems, the feature-based systems and the signature comparison systems. Both need a database of known targets to perform a comparison of the entries of the database with the measured data. The difference between both systems is the representation of the data. The feature-based systems transform the measured data into a feature space and perform the comparison with the known data in this feature space, which is normally much faster than the direct comparison of the measured radar signatures. The drawback of the feature-based methods is that the design of appropriate features either requires a training using a machine learning algorithm or a human expert that designs the feature extraction and classification manually.

The algorithms presented in Section 1.2.2 are aiming at the identification of different scattering mechanisms and not at the identification of specific targets. The results of these algorithms can be used as part of a feature-based identification, but the presented literature is not focused on that topic. Therefore, the overview is split into these two parts. These scattering identification methods are closer related to the topic of this thesis than the methods of Section 1.2.1, but the target classification remains the goal of the presented algorithm and therefore a short overview of available methods is given below.

1.2.1 Target Classification

The most obvious way of classification with radar signatures is a nearest neighbor approach, which is for example used in [HP93, RS95, Dom96] for air targets. While [HP93] and [Dom96] use High Range Resolution (HRR) profiles for classification, [RS95] uses Inverse Synthetic Aperture Radar (ISAR) images as input. The implementation of the classifiers is comparable, except that [HP93] uses correlation filters, which are designed as an average HRR profile over a certain aspect angle interval and [RS95, Dom96] compare unknown test data with all stored reference data around the estimated aspect angle. Additionally to the correlation filter, a majority voting over a sequence of HRR profiles is used in [HP93] to achieve a more robust result.

The benefit of using sequences of HRR profiles is also shown in [JO00] and [Bie08]. The latter uses a classifier similar to the one in [Dom96], but the results of consecutive HRR profiles are fused to improve the results. As fusing methods, a Bayes approach, majority voting and a Dempster-Shafer decision support method are used. In [JO00] a combination of a tracking and a classification system is presented, which uses a Bayes approach for classification. The initial estimate of target type and orientation is either established or rejected using the classification results of the following HRR profiles. The combination of tracking and classification is also exploited in [NKOW02], which uses a HRR-based template matching classifier comparable to [HP93] in ambiguous ground target tracking situations. For example if two or more vehicles merge to a single track on the road, a classifier can be used when these targets split again to reconstruct the track of each target.

Examples for feature-based methods can also be found in the literature. A prominent example is the extraction of strong peaks and their relative position along the HRR profile. This is done for example in [MW99] and [HWS10]. Both use a statistical classifier based on the *peak location probability function* and the *peak amplitude probability density function*, which are estimated from a training set of known targets.

A feature-based approach that is not aiming at a reduced dimension can be found in [KSK05], which is focusing on a classifier with rotational invariance instead of a reduced complexity. In this work, radar images are transformed into a polar coordinate system with the center of the target as the origin and radius and angle as axes of the system. As features, the projection of the target on each axis, i.e radius and angle, and the twenty largest elements of a principal component analysis (PCA) of the transformed images are used.

A completely different approach that is not based on high range resolution, but on high Doppler resolution is the classification with micro-Doppler features. For aircrafts, the most prominent feature of this kind is the *Jet Engine Modulation* (JEM) [Tai05], which uses the modulation of the radar echo caused by the blades of jet engines. With these modulation features, the number of blades and the rotational speed can be estimated, which can be used to limit the number of possible engines and with that the number of possible targets. This means that identification of the target is generally not possible, but a list of candidates can be produced.

These modulation features can also be used for rotor engines of helicopters or drones. In these cases, also the size of the blades can be estimated from the covered Doppler spectrum. An example for this approach can be found in [dWHM14], which uses a singular value decomposition of measured spectrograms to extract rotational speed and size of the blades. A combination of these modulation features with HRR profiles can be found in [MEA$^+$14], which uses these features to identify three different drones from simulated data.

These classical methods are studied already for many years and many applications can be found in the literature. In [CW07] the target classification using JEM was described as the most mature, since it is already applied for many years. One of the first reports is [Gar61], which was published in 1961. Nevertheless, over the last years, since 2014, an increased interest in image-based classification can be seen in the literature. This increased interest is related to the popularity of *Deep Learning* methods, which give today's state-of-the-art results for almost all available datasets. To the knowledge of the author, the first publication on deep learning related to radar target recognition is [Wag14], which uses a convolutional neural network to identify ten different targets in SAR images. Over the last years a vast amount of papers related to target recognition using deep learning were published, some examples are [CWXJ16, ØKCL16, Wag16, KS18, BAB$^+$20]. The success of these methods is based on the intrinsic structure of the used networks, which use convolutional filters to extract an internal representation of the input data. This internal representation should contain all relevant information about the input data and should be sufficient for the classification. Therefore, these methods can be seen as feature-based methods, but the extraction of the features is done automatically using the input data, which is normally two-dimensional, i.e. radar images. One reason for the efficiency of this internal representation are the nonlinearities in the convolutional part of the network. These nonlinearities allow the network to represent far more complicated functions than for example a PCA, which can be modeled as a single layer linear neural network.

1.2.2 Scattering Identification

In this thesis, the focus is on scattering identification in high resolution radar images rater than the actual target identification. Therefore, a part of literature focusing on this topic is given here separately, as it was already mentioned above. Yet the use of scattering properties is not limited to high resolution imaging, resonance effects that appear in the lower frequency bands can also be used for identification. An example is given in [CdBS10], which itself includes a short review on this field. The general idea is to use a radar with a wavelength in the region of the size of the air-intake of an aircraft, since this will cause resonances that can be detected in the echo signal. These resonances can be sufficient to identify targets from a list of potential targets using a matching process between the measured data and a database.

In the field of imaging radar, the oldest and thus most mature technique to identify scattering mechanisms is the use of polarimetric data with an appropriate decomposition [CP96]. To use these decomposition theorems, fully polarimetric measurements are necessary, i.e. two coherent transmitters and two receivers with vertical and horizontal polarization of the antenna. With this constellation, a 2×2 scattering matrix consisting of the echos in co-polarized and cross-polarized combinations can be defined for each pixel in the radar image. For this scattering matrix, several coherent decomposition theorems exist that allow the identification of single, double and multiple bounce scattering centers. An example with two different decompositions is given in [KLBA06], which presents the decomposition of the measurements of a bridge during construction and after completion. In the given examples, the difference between the cables, the deck and the supporting structures is clearly visible and shows the potential of fully polarimetric measurements.

Nevertheless, the use of fully polarimetric systems is still not very widespread due to the high cost and the demanding calibration needed for these systems. Therefore, most research related to scattering identification over the last years was focused on single channel radar systems.

One of the first works pointing in this direction is [PM97] that introduced a parametric signal model for *attributed scattering centers*. These scattering centers are namely corners, edges, flat plates and curved surfaces and the response of them is derived by the geometric theory of diffraction. Additionally to these models also a feature extraction method was introduced and some examples with SAR images are shown to validate the derived models. The application of these parametric scattering center models for target recognition can be found for example in [CMP00], which uses a SAR dataset consisting of ten targets to perform a classification. For the presented algorithm, a target is modeled as a collection of attributed scattering centers, whose position and type are extracted from the measurements and afterwards matched with the predicted features based on a target hypothesis. The application of attributed scattering centers can also be found in more recent publications like [DWH$^+$17] or [DWMY17]. A three-dimensional extension of these models, also for bistatic observations, was published in [Jac09], which also gives a maximum likelihood estimator to identify these scattering centers in 3-D radar images.

An application that is using sparse decomposition of radar measurements to identify different scattering targets is presented in [dWvR16]. In this publication, a through-the-wall radar is used to perform stand-off measurements of a house, and a sparse decomposition algorithm is used to identify corners and planar areas of the walls inside the house. The model of these scattering centers does not only rely on the magnitude, rather the phase is exploited as well. A planar wall should produce a linear phase, while a corner is modeled with a quadratic phase term. A comparable approach is chosen in [ES17], which uses the method of stationary phase to create models for different geometric structures and a sparse reconstruction algorithm to locate these models in measured ISAR images. The data used is from a car on a turntable and the models are curvature structures on a car. Additionally to the decomposition of a single image, the recorded data from 360° is split into several parts and the aspect angle dependent behavior of scattering centers is exploited for recognition. This separation of the measured data is one possibility to process data from large observation angles for scattering identification. Other examples to model and identify aspect dependent scattering centers is given in [ÇSÖ$^+$14] and the references therein.

Another example that uses sparse representation of radar images is [SÇMs11], which presents a framework that is able to handle several different scattering models in the complex measurement domain. This work is, although it focuses on the improvement of the imaging result and is limited to wavelets and point scattering centers together with [dWvR16], one of the closest to the work presented in this thesis.

1.3 Motivation and Contribution

In Section 1.1 was already shown that reliable classification methods are needed to avoid tragic accidents and support decision making in surveillance systems. Therefore, the main motivation of this thesis is to develop an intermediate step between the detection of targets and the classification methods presented in the overview of Section 1.2.1. This overview showed that most classifiers that are able to identify targets rely on a list of potential targets that can be discriminated. This list of potential targets is either created by operational constraints or by the data that is available.

One application of the framework presented in this thesis is a feature extraction algorithm that can be used to create a list of potential targets if only a small amount of a priori information is available. The presented method is able to gather information about the number and position of jet engines and can thus distinguish between a two-engine and a four-engine airplane, or between an aircraft with engines mounted under the wing and one with engines at the fuselage. The number and position of engines is thus a feature that can be used for the classification of the aircraft itself. The presented algorithm relies on the dispersion effect that occurs inside jet engines while the electromagnetic wave of the radar is propagating inside the air intake of engines. The relevance of this kind of scattering was already mentioned for example in [RH96], but so far this effect is considered mostly as imaging artifact [Bor99].

In this thesis, the effect is modeled in the frequency domain and is used additionally to the standard point scatterer model in a sparse decomposition framework, which is capable of using several models at the same time and decompose the ISAR image into the corresponding parts. To the knowledge of the author, this is the first time that an approach like this is presented with a physically motivated model and used in a target recognition context. In the literature, only frameworks with dictionaries[1], which are known to give an efficient representation in the image domain, e.g. wavelets, can be found [SÇMs11].

The validation of the presented algorithm will be done with ISAR images, which are measured with the Tracking and Imaging Radar (TIRA) of the Fraunhofer Institute for High Frequency Physics and Radar Techniques (FHR).

The main contributions of this thesis are summarized as follows.

1. Design of an algorithm that can be used as intermediate step in an air target recognition scenario to create a list of potential targets.

2. Design of a complex valued dictionary for air intakes and jet engines based on a waveguide scattering model to localize these in ISAR images.

3. Extension of a sparse decomposition algorithm called morphological component analysis to the complex valued radar raw data domain.

---

[1] The term *dictionary* will be explained in Section 4.1.

4. Validation of the presented algorithm with simulations and real ISAR images measured with the TIRA system of the Fraunhofer FHR.

Parts of this thesis have been published before. In [Wag15] a variation of the morphological component analysis, which will be presented in Section 4.3, was used with real valued SAR data to separate targets and clutter. The algorithm was applied in this case with a wavelet and a cosine transform to represent targets and clutter, respectively. The detection of jet engines using a waveguide model was presented for the first time in [WDE16] by the author. In this work, the decomposition was also applied in the image domain with an image domain approximation of the model presented in Section 3.2 and the decomposition was performed by a greedy approach called matching pursuit, which will be shown in Section 4.2.1. The work in [WE17] was an extension of the work in [WDE16], giving more details about the model and the decomposition, but without new results compared to [WDE16].

Actual results of this thesis were pubished for the first time in [WE21] with a detailed derivation of the model, which is comparable to Chapter 3. As results, parts of the content from Sections 5.1 and 5.3 are presented in [WE21]. The latest results, including the results from Section 5.2, were presented in [WE22] together with the latest development of the sparse decomposition framework.

1.4 Outline

The remainder of this dissertation is organized as follows. In Chapter 2 to 4, the background material for this dissertation is covered. From these chapters, Chapter 2 gives the basics on radar and Inverse Synthetic Aperture Radar (ISAR) imaging. In Chapter 3, first a review of the propagation of electromagnetic waves in waveguides is given and afterwards the model that is used in this dissertation for the jet engine scattering is formulated. The foundations and methods of sparse representation are described in Chapter 4 together with the framework that is used for the discrimination of the different scattering mechanisms. The results of this framework and the main contribution of this thesis are given in Chapter 5. Finally Chapter 6 concludes this thesis and shows some possible directions for future work using the obtained results.

Chapter 2

Radar and ISAR Imaging Principles

The purpose of this chapter is to present the fundamentals of radar in general and of ISAR imaging as detailed as necessary to understand the context and application of the later presented algorithm. For this reason, the basic elements of radar with an emphasis on high resolution radar will be presented in Section 2.1. The core concepts of ISAR imaging are shown in Section 2.2.

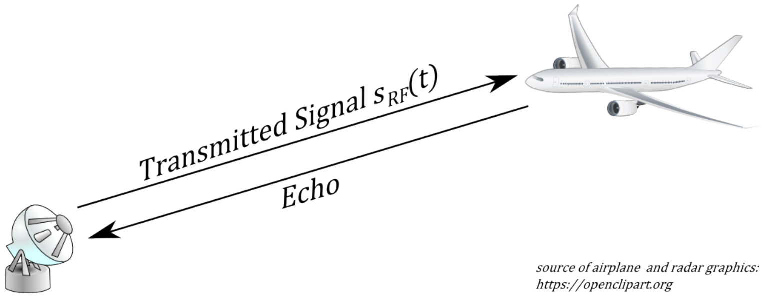
2.1 High Range Resolution Radar

The most important part of high resolution radar are the range compression waveforms, but to understand the motivation for these waveforms some knowledge of the fundamental radar elements is required. These fundamentals are limited here to the basic elements of radar transmitter and receiver and are described in Section 2.1.1 from a signal processing perspective without considering hardware constraints. This way is chosen, since this thesis introduces a new signal processing method that works independently from the chosen hardware implementation, and a detailed understanding of the radar hardware is not necessary to understand radar imaging theory.

Section 2.1.2 describes the range determination of a target with a basic pulsed rectangular waveform. The already mentioned range compression waveforms, especially the method that is used in the TIRA data, are described in Section 2.1.3. These methods are necessary to achieve range resolutions below the typical sizes of aircrafts, which means that these methods are mandatory to produce images of air targets.

2.1.1 Fundamentals of the Radar Transmitter and Receiver

The principal function of a radar is to determine the location of targets by transmitting a signal and receiving echos of this signal. The sources of these echos can have various origins, it might be man-made targets, like airplanes, but also birds can cause echos strong enough that a radar can detect it. Also natural objects, like trees or the ground cause sometimes strong reflections that are visible in the processed radar output. These unwanted echos are normally referred to as 'clutter' and several methods have been presented in the past to suppress this stationary clutter [Sko01]. For this thesis, clutter is not relevant, since the produced images are from in flight targets, which means that there is no clutter near the targets. The principal operation of a radar, transmitting a signal and receiving an echo, is depicted in Figure 2.1.1



source of airplane and radar graphics: https://openclipart.org

Figure 2.1.1: Basic radar principle

The graphic in Figure 2.1.1 shows that the radar transmits a signal $s_{RF}(t)$ through the air, thus this signal has to follow the boundary conditions given by Maxwell's equations (see Section 3.1 or for example [Geo97]). These boundary conditions require a real valued signal and the common form that is used in radar is [CB09]

$$\begin{aligned} s_{RF}(t) &= \Re\left\{s(t)e^{j2\pi f_0 t}\right\} \\ &= \Re\{s(t)\}\cos(2\pi f_0 t) - \Im\{s(t)\}\sin(2\pi f_0 t), \end{aligned} \tag{2.1.1}$$

where $s(t)$ is the complex envelope, i.e. the *waveform* of the transmitted signal. The frequency f_0 is the *carrier frequency* of the radar signal and the typical range for radar operations is the range of 1 GHz to 100 GHz. The generation of $s_{RF}(t)$ is done by the so called *quadrature modulator* (QM), which basic structure is shown in Figure 2.1.2. With this structure any arbitrary complex waveform can be converted to a real valued bandpass signal that can be transmitted via an antenna. The choice for this modulation scheme becomes clear in combination with the receiver structure shown below.

The main function of the receiver is to remove the carrier frequency of the echo signals and to extract the complex envelope. These steps are done since the processing of the data is performed in the *baseband*, which is located in the spectrum around the frequency of zero and defined by the bandwidth of the waveform. This transformation is done to reduce the requirements on the processing hardware, since processing of the transmitted signal directly would be very demanding and there would be no benefit from a signal processing perspective, since the information is completely contained in the modulation of the signal.

Graphically speaking, this transformation from the carrier frequency to the baseband represents a shifting in the spectrum of the frequencies around f_0 to the area around the frequency $f = 0$ in the center of the spectrum. But if the carrier frequency is simply removed from the spectrum, the frequencies around f_0 will overlap with the

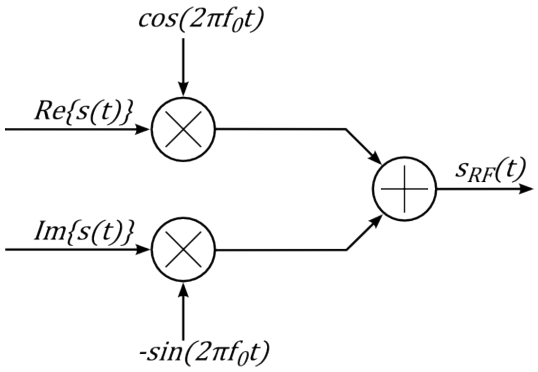


Figure 2.1.2: Basic structure of a quadrature modulator

frequencies around $-f_0$, therefore this operation cannot be realized by simply removing the f_0 term from (2.1.1). The solution is the so called *equivalent low-pass signal*, which is the result of a two-step procedure. First the spectrum of the signal is changed into a single sided spectrum that only contains positive frequencies and afterwards the single sided spectrum is shifted towards the negative frequencies by the amount of f_0.

The first step of creating a single sided spectrum requires an analytic time signal, which means that the real and imaginary parts must form a Hilbert transform pair

$$s_{Analytic}(t) = s_{RF}(t) + j\mathcal{H}\{s_{RF}(t)\}. \tag{2.1.2}$$

The imaginary part of (2.1.2) is called the *quadrature (Q-) component* and the model is called *quadrature model* [Bor99]. The real part of the analytic signal is called *in-phase (I-) component*. A general problem of the Hilbert transform for practical applications is that it cannot be implemented in real time and therefore an approximation must be used.

For narrow-band signals, which are used in the radar applications considered here, the standard method to approximate the Hilbert transform and the shifting to baseband is the *quadrature demodulator* (QDM) [OL07]. The QDM realizes the two steps mentioned above, creating a single sided spectrum and removing the carrier frequency.

The basic structure of a QDM is shown in Figure 2.1.3, which shows a multiplication of the received signal with $\cos(\cdot)$ and $\sin(\cdot)$ functions. To understand why this structure extracts the transmitted waveform of the QM, some trigonometric identities are needed, precisely

$$\begin{aligned}\cos^2(2\pi f_0 t) &= \frac{1}{2}(1+\cos(4\pi f_0 t)) \quad (2.1.3)\\ \sin^2(2\pi f_0 t) &= \frac{1}{2}(1-\cos(4\pi f_0 t)) \quad (2.1.4)\\ \sin(2\pi f_0 t)\cdot\cos(2\pi f_0 t) &= \frac{1}{2}\sin(4\pi f_0 t). \quad (2.1.5)\end{aligned}$$

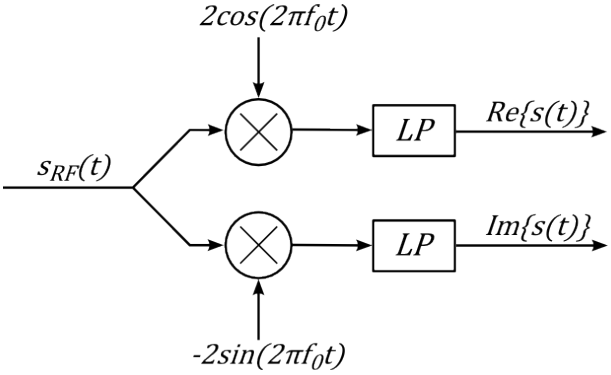


Figure 2.1.3: Basic structure of a quadrature demodulator

The output of the multiplications in the QDM give again the real and imaginary part of the complex envelope

$$\begin{aligned}s_{RF}(t)2\cos(2\pi f_0 t) &= 2\Re\{s(t)\}\cos^2(2\pi f_0 t) - 2\Im\{s(t)\}\sin(2\pi f_0 t)\cos(2\pi f_0 t)\\ &= \Re\{s(t)\}(1+\cos(4\pi f_0 t)) - \Im\{s(t)\}\sin(4\pi f_0 t)\\ &= \Re\{s(t)\} + \Re\{s(t)\}\cos(4\pi f_0 t) - \Im\{s(t)\}\sin(4\pi f_0 t)\end{aligned} \quad (2.1.6)$$

and

$$\begin{aligned}s_{RF}(t)(-2\sin(2\pi f_0 t)) &= -2\Re\{s(t)\}\cos(2\pi f_0 t)\sin(2\pi f_0 t) + 2\Im\{s(t)\}\sin^2(2\pi f_0 t)\\ &= -\Re\{s(t)\}\sin(4\pi f_0 t) + \Im\{s(t)\}(1-\cos(4\pi f_0 t))\\ &= \Im\{s(t)\} - \Im\{s(t)\}\cos(4\pi f_0 t) - \Re\{s(t)\}\sin(4\pi f_0 t).\end{aligned} \quad (2.1.7)$$

The unwanted terms at $\pm 2f_0$ are removed with the low-pass filters shown in Figure 2.1.3. The output of the QDM corresponds thus to the input of the QM if the modulated signal is fed directly into the QDM.

Since the first step of the QM and QDM is the creation of an analytic signal with the I- and Q-components, common alternative names for the QM and QDM are IQ modulator and demodulator, respectively.

2.1.2 Range Determination with a Pulsed Waveform

2.1.2.1 Waveform Description and Properties

As shown above, an IQ modulator can be used to submit any arbitrary waveform over the air. In this section the most basic waveform, a rectangular pulse of duration τ_P and amplitude A

$$s(t) = A \cdot rect\left(\frac{t}{\tau_P}\right), \tag{2.1.8}$$

will be used as an example to illustrate the basic relations in range determination of targets[1]. An illustration of a pulse train with this waveform is shown in Figure 2.1.4.

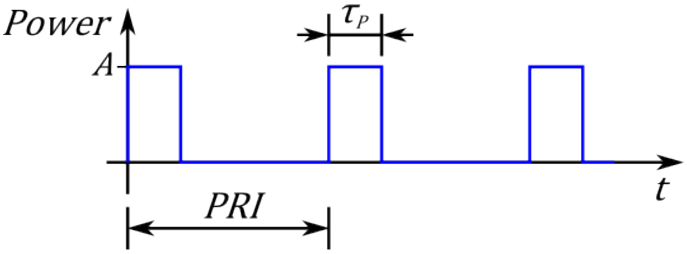


Figure 2.1.4: Train of rectangular pulses

The pulses in Figure 2.1.4 are submitted with a fixed period, which is called *pulse repetition interval* (PRI) and the inverse of the PRI is the *pulse repetition frequency* (PRF). According to the basic radar principle described in Figure 2.1.1, the radar is waiting for an echo of the target and the distance to the target is determined by the time delay Δt between the transmitted and the received pulse. The distance R to the target is calculated by

$$R = \frac{c\Delta t}{2}, \tag{2.1.9}$$

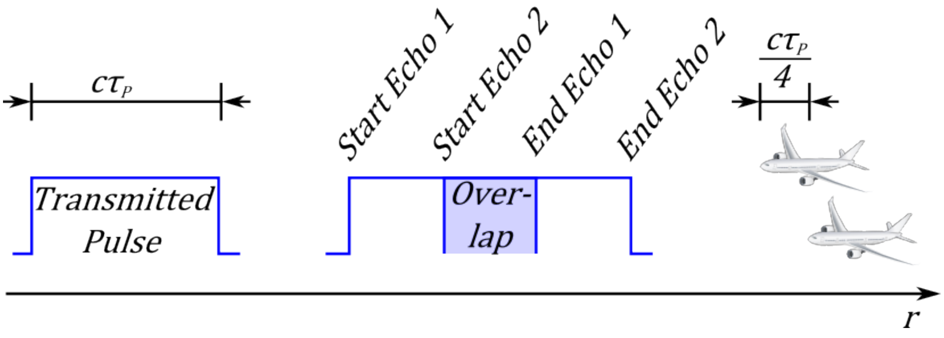
where c is the propagation speed of electromagnetic waves, i.e. the speed of light. The factor of two is needed to compensate the round trip delay. From this formula, the range resolution ΔR of this basic waveform follows as the minimum distance between two targets that produces two separable echos in the received signal. The value of the range resolution is given by

---

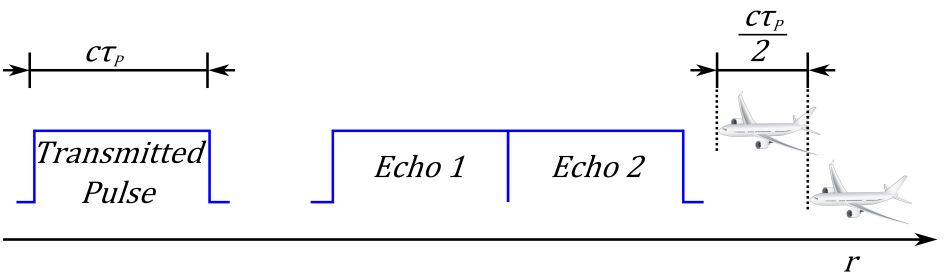
[1] The rectangular function $rect(t/\tau_P)$ is defined as 1 from $-\tau_P/2$ to $\tau_P/2$ and zero elsewhere.

$$\Delta R = \frac{c\tau_P}{2} \tag{2.1.10}$$

and the reason for this value is shown in Figure 2.1.5. In plot (a) two closely spaced targets produce an overlap in the echo signal, and in plot (b) the separation of the targets produces two exactly successive echo pulses and is thus the minimum distance that should be present between two targets. It is assumed in both cases that the reflection appears only at the nose of the depicted aircraft.



(a) Non-separable targets



source of airplane graphic: https://openclipart.org

(b) Separable targets

Figure 2.1.5: Range resolution of a rectangular waveform

Since the spectrum of a rectangular function is given by a sinc-function and if the bandwidth B of the waveform is interpreted as the frequencies up to the first zero of the sinc, which is given by $1/\tau_P$, the range resolution can also be written in terms of the bandwidth of the waveform by

$$\Delta R = \frac{c\tau_P}{2} = \frac{c}{2B}. \tag{2.1.11}$$

It will shown in Section 2.1.3 that the dependency of the range resolution on the bandwidth holds for any waveform and thus the range resolution is improved by increasing the bandwidth of the waveform.

Another point that should be mentioned is the unambiguous range of a radar. Figure 2.1.4 shows that a radar is transmitting a series of pulses with a fixed interval and normally an echo is assigned to the pulse that was transmitted before the echo arrives. Therefore, the maximum range R_{max} that can be determined unambiguously by a radar is limited by the PRF and is given by

$$R_{max} = \frac{c}{2 \cdot \mathrm{PRF}}. \tag{2.1.12}$$

To detect targets up to the given maximum range, the signal must be sampled continuously. However, in Section 2.1.3.2, a method will be described that reduces the unambiguous range to a small window, but keeps the range resolution with a sampling frequency much smaller than the bandwidth of the radar waveform. This processing scheme is also applied to create the data that is used later in this thesis. However, before this method will be explained, a general filtering algorithm to improve the target detection in noisy signals will be shown in the next part.

2.1.2.2 Matched Filter

The visualization of the range resolution in Figure 2.1.5 shows the ideal case that the echo is the exact reversed copy of the transmitted signal without any noise in the signal. The assumption that the echo has the same shape as the transmitted signal is a common assumption in radar and assumes non-dispersive scattering, which is true for most types of scattering. An exception of this model and the consequences for the radar signal will be shown in Chapter 3.

In this section the optimal and commonly used filtering method to improve the signal to noise ratio (SNR) of the radar signal is presented. This filtering method is very widespread since it gives the optimal SNR for known signals in white noise, which is the task in radar signal processing using the common assumptions.

The SNR of a radar system can be calculated with the so called *radar equation*, which relates the parameters of the system to the range R and the strength of the target. The strength of the target is given by the *radar cross section* (RCS) σ_{RCS}, which is defined as the projected area of a sphere that would lead to the same echo. One of the results of the radar equation is that the received energy is decreasing with the factor R^4, i.e. a doubling of the target range would lead to a reduction of 12dB in the signal power. With the very high ranges that appear in air surveillance systems and the unavoidable system noise, a filtering method to optimize the SNR is necessary.

The presented method is called *matched filter* and improves the detectability of a target by maximizing the output peak-SNR of the radar receiver [Sko01]. The use of the peak-SNR indicates already that the energy of the received waveform is compressed into a single value at a certain time t_0; a point that will be used again in Section 2.1.3.1.

For the derivation of the matched filter, it is assumed that the received signal $x(t)$ is the sum of the transmitted signal $s(t)$, which is scaled in amplitude by the factor α due to target strength and range, and independent white Gaussian noise $N(t)$ to represent the system noise. Furthermore, it is assumed that the received signal has been digitized and is organized in an N-dimensional vector $\mathbf{x}$. This vector is the sum of the digitized transmitted signal $\mathbf{s}$ and the noise vector $\mathbf{N}$, which remains to be Gaussian white noise bandlimited according to the sampling frequency, i.e.

$$\mathbf{x} = \alpha\mathbf{s} + \mathbf{N}. \tag{2.1.13}$$

The noise is assumed to have a mean value of zero, which leads to a covariance matrix R_N of the form $\sigma^2\mathbf{I}_N$. The aim is to find an optimal weight vector $\mathbf{w}$ that maximizes the output SNR, which is given by

$$SNR(\mathbf{w}) = \frac{|\alpha\mathbf{w}^H\mathbf{s}|^2}{\mathbb{E}\{|\mathbf{w}^H\mathbf{N}|^2\}} = \frac{|\alpha|^2|\mathbf{w}^H\mathbf{s}|^2}{\mathbb{E}\{\mathbf{w}^H\mathbf{N}\mathbf{N}^H\mathbf{w}\}} = \frac{|\alpha|^2|\mathbf{w}^H\mathbf{s}|^2}{\mathbf{w}^H\mathbb{E}\{\mathbf{N}\mathbf{N}^H\}\mathbf{w}}. \tag{2.1.14}$$

The expectation in the denominator of (2.1.14) is the above mentioned covariance matrix of the additive noise, thus the expression can be simplified as

$$SNR(\mathbf{w}) = \frac{|\alpha|^2|\mathbf{w}^H\mathbf{s}|^2}{\mathbf{w}^H\sigma^2\mathbf{I}_N\mathbf{w}} = \frac{|\alpha|^2|\mathbf{w}^H\mathbf{s}|^2}{\sigma^2||\mathbf{w}||^2}. \tag{2.1.15}$$

The output of (2.1.15) can be expanded with the expression $||\mathbf{s}||^2/||\mathbf{s}||^2$, which results in

$$SNR(\mathbf{w}) = \frac{|\alpha|^2|\mathbf{w}^H\mathbf{s}|^2}{\sigma^2||\mathbf{w}||^2} \cdot \frac{||\mathbf{s}||^2}{||\mathbf{s}||^2} = \frac{|\alpha|^2||\mathbf{s}||^2}{\sigma^2} \cdot |\rho(\mathbf{w},\mathbf{s})|^2, \tag{2.1.16}$$

where $\rho(\mathbf{w},\mathbf{s})$ is the correlation coefficient of the weight vector and the transmitted signal

$$\rho(\mathbf{w},\mathbf{s}) = \frac{\mathbf{w}^H\mathbf{s}}{||\mathbf{w}||\,||\mathbf{s}||}. \tag{2.1.17}$$

The correlation coefficient ranges between -1 and 1 and the magnitude squared correlation coefficient in (2.1.16) reaches its maximum if $\mathbf{s}$ and $\mathbf{w}$ are collinear [BV18]. Therefore, the solution for the optimal weight vector is given as

$$\mathbf{w}_{opt} = argmax\{SNR(\mathbf{w}) : \mathbf{w} \in \mathbb{C}\} = w\mathbf{s}. \tag{2.1.18}$$

This means that the filter coefficients of the matched filter are a scaled version of the transmitted signal, since this will give the maximum correlation output. The filter operation is formally given for a transmitted signal with a finite length K as

$$\mathbf{y}[n] = \sum_{k=0}^{K} \mathbf{x}^*[n-k]\mathbf{s}[k], \tag{2.1.19}$$

where $\mathbf{x}$ is the received signal, $\mathbf{s}$ is the known transmitted waveform and $\mathbf{y}$ is the filter output. The operation principle is qualitatively shown in Figure 2.1.6.

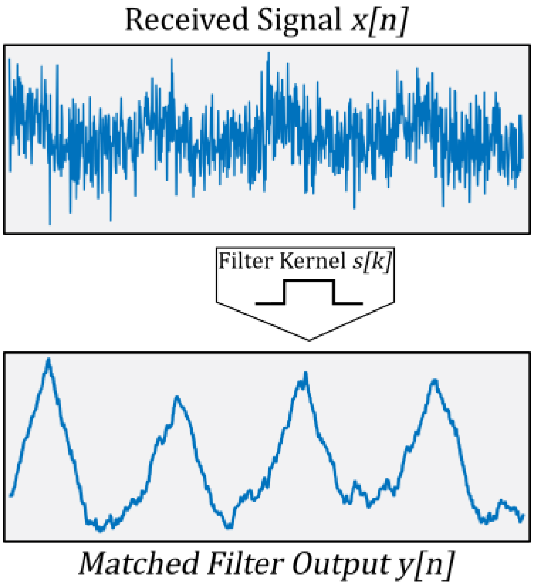


Figure 2.1.6: Matched filter principle

The visualization shows that the matched filter has a strong denoising effect and compresses the lengthy rectangular time signal into a single point with a maximum SNR. This point is reached at the end of the waveform in the received signal when the maximum overlay between filter and received pulse is present. This compression effect will be the basis for the techniques presented in the next section.

2.1.3 Range Compression Techniques

Up to this point, the signal processing to determine the range of a target with a pulsed waveform has been described. The major problem of this kind of waveform is the range resolution, since it was shown in (2.1.11) that the range resolution of the system is proportional to the length of the pulse. This leads to the conclusion that a very short pulse should be used to achieve a very high resolution. This is done for example in near range systems like ground penetrating radar with a typical pulse duration between 200 ps and 20 ns [RLNP11]. In high resolution air surveillance systems, this rectangular waveform is not very common, since the R^4 range dependency of the received energy would require a very high peak power of the pulse to operate at appropriate distances. Therefore, the use of wideband waveforms became popular, which transmit a longer pulse with a modulation to achieve a high bandwidth. This high bandwidth is used in the receiver to achieve the range resolution of an equivalent rectangular waveform according to (2.1.11).

Two kind of waveforms are commonly used in *high range resolution* (HRR) radars, the *linear frequency modulated* (LFM) and the *stepped frequency* waveform. The latter synthesizes a high bandwidth by submitting a train of narrowband pulses with shifted carrier frequencies, which puts some requirements on the stationarity of the observed target. The LFM waveform transmits the used bandwidth in a single pulse using frequency modulation and since this is the waveform that is used in the TIRA system, the description in this thesis is limited to this waveform. A detailed description of the signal processing for stepped frequency waveforms can be found for example in [Weh95] or [Mah05].

The LFM waveform is also called *chirp* and in the notation of (2.1.1), this waveform is written as

$$s(t) = A \cdot rect\left(\frac{t}{\tau_P}\right) \cdot e^{j2\pi(f_0 t + \frac{\mu}{2}t^2)}, \tag{2.1.20}$$

where μ is the so called *chirp rate* and f_0 is the carrier frequency. The chirp rate defines the slope of the frequency modulation, which becomes visible by calculating the instantaneous frequency of the transmitted chirp from its phase term by

$$f(t) = \frac{1}{2\pi}\frac{\partial}{\partial t}2\pi(f_0 t + \frac{\mu}{2}t^2) = f_0 + \mu t. \tag{2.1.21}$$

In (2.1.21) the waveform and the carrier frequency have been combined to a single expression. The chirp rate is related to the time bandwidth product of the waveform and is calculated by the ratio of the covered frequencies[2] B and the pulse length τ_P

$$\mu = \frac{B}{\tau_P}. \tag{2.1.22}$$

In the following sections, the two common methods used in the receiver to produce HRR profiles with the chirp waveform are presented.

2.1.3.1 Matched Filtering

As it was already shown in Section 2.1.2.2, the matched filter is the optimal filter with regard to the SNR and compresses the presence of a target to a single point in the output signal. Furthermore, the derivation of the matched filter was independent of the waveform and therefore gives the optimum SNR for any waveform, also the chirp waveform. A point not considered specifically before is the *point spread function* (PSF) of a certain waveform. In the example of Figure 2.1.6, it is visible that the matched filter output of a rectangular pulse is a triangular function. Generally, the output of the matched filter is determined by the autocorrelation of the transmitted signal [CM14], which follows directly from the derivation of the matched filter.

The derivation of the matched filter in Section 2.1.2.2 was shown with discrete samples of the signal to simplify the result with the correlation coefficient. The notation

---

[2] In general, the range of covered frequencies does not coincide with the bandwidth of a signal. However, for a chirp waveform it can be shown that this is the case, for example with the use of Fresnel integrals [Bor99].

of the chirp in (2.1.20) is in continuous time domain and also the result of the matched filter here will be shown in a continuous time notation. The transfer from the discrete to the continuous domain is done via the Cauchy-Schwarz inequality, which is given here without proof for two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ of the same dimension[3]. In the discrete domain the inequality is written as

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq ||\mathbf{x}||\,||\mathbf{y}||, \tag{2.1.23}$$

where $\langle \cdot, \cdot \rangle$ represents the inner product of two vectors. The result of this equation was already used indirectly in the correlation coefficient to obtain the general result of the matched filter. The equality of (2.1.23) holds, if both vectors are collinear. In the continuous time domain, the Cauchy-Schwarz inequality is written for square-integrable functions[4] $x(t)$ and $y(t)$ as [Unb02]

$$\left| \int_a^b x(t) y^*(t) dt \right|^2 \leq \int_a^b |x(t)|^2 dt \int_a^b |y(t)|^2 dt. \tag{2.1.24}$$

The equality in the continuous case also holds for proportional functions $x(t)$ and $y(t)$ and the implication for the continuous matched filter is that the complex conjugated of the transmitted waveform must be used, which is in the discrete case implicit due to the definition of the inner product. Under the assumption that the target does not change the waveform, it follows that the output y_{MF} of the matched filter gives the PSF of the chirp by the autocorrelation

$$y_{MF}(t) = \int_{-\infty}^{\infty} s^*(\tau - t) s(\tau) d\tau. \tag{2.1.25}$$

With the definition of the chirp in (2.1.20) follows[5]

$$y_{MF}(t) = \int_{-\infty}^{\infty} rect\left(\frac{\tau - t}{\tau_P}\right) rect\left(\frac{\tau}{\tau_P}\right) e^{j2\pi(f_0 t + \frac{\mu}{2}(\tau^2 - (\tau - t)^2))} d\tau. \tag{2.1.26}$$

Since the autocorrelation is symmetric, (2.1.26) can be solved for $t \geq 0$ and extended to the negative time domain. The two rectangular functions limit the signal to $|t| \leq \tau_p$ and the output of the matched filter vanishes for $|t| > \tau_p$. Therefore, the limits of the integral can be changed to the remaining interval and the simplified equation is given by

$$y_{MF}(t) = e^{j2\pi(f_0 t - \frac{\mu}{2} t^2)} \int_{t - \tau_P/2}^{\tau_P/2} e^{j2\pi\mu\tau t} d\tau. \tag{2.1.27}$$

The solution of the integration in (2.1.27) is given as [KPDA60]

---

[3] A formal proof can be found in textbooks on linear algebra, e.g. [BV18].

[4] A function is called square-integrable if the integral of the square of the absolute value is finite.

[5] The amplitude A of the waveform is ignored here, since it will not change the shape of the PSF.

$$y_{MF}(t) = e^{j2\pi f_0 t}\frac{\sin\left(\pi\mu(\tau_P t - t^2)\right)}{\pi\mu t}, \tag{2.1.28}$$

which possesses in baseband the envelope

$$\tau_p\frac{\sin\left(\pi(B|t| - \mu t^2)\right)}{\pi B|t|}, \tag{2.1.29}$$

which is simplified with the definition of μ in (2.1.22) and extended to negative times.

Normally the output of the matched filter is given in an approximated version as a sinc function, where the $\pi\mu t^2$ term in the argument of the sine function is ignored and therefore, the output can be written as

$$y_{MF}(t) = \tau_p\frac{\sin\left(\pi(B|t| - \mu t^2)\right)}{\pi B|t|} \approx \tau_p\frac{\sin(\pi B|t|)}{\pi B|t|} = \tau_P sinc(Bt). \tag{2.1.30}$$

Beside the scaling with the pulse length, the main result is that the matched filter will produce a sinc-like output and the width of the main lobe depends only on the bandwidth of the chirp. To achieve a definition of resolution that fits the range resolution of the rectangular pulse in (2.1.11), the 3dB point of the sinc main lobe is used as the reference point that two targets can be distinguished. This means that two targets must be separated by at least the distance that is produced by a matched filter with neighboring sinc functions that intersect at the corresponding 3dB points. This distance ΔR is given by

$$\Delta R = \frac{c}{2B}, \tag{2.1.31}$$

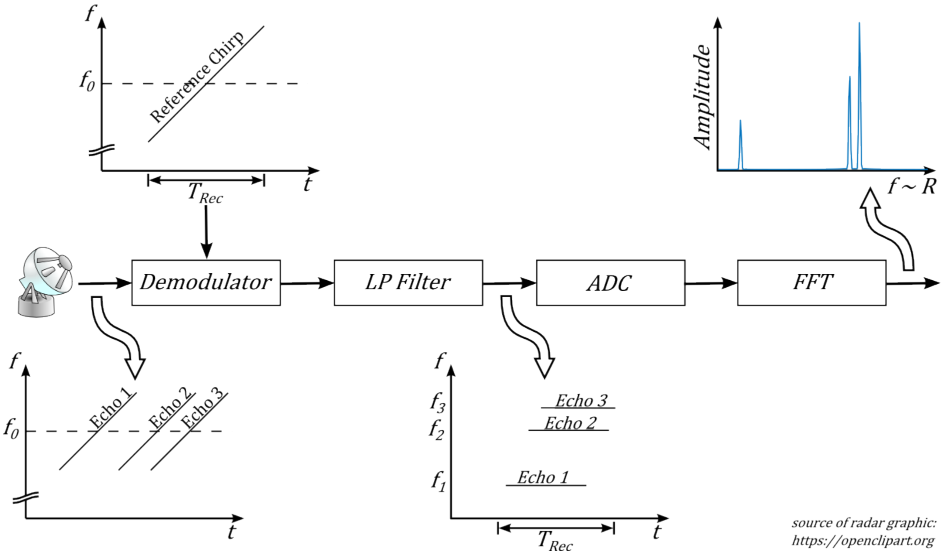
which corresponds to the range resolution ΔR of the rectangular function. The application of the matched filter is equivalent to a submission of a rectangular pulse with a length defined by the bandwidth of the chirp. Therefore, the input of the matched filter is sometimes called *uncompressed range profile* and the output is called *compressed range profile*, since the matched filter acts like a filter that compresses all frequencies of the chirp into a pulse of equivalent length.

The matched filter gives the theoretical optimum SNR for an isotropic point target and white noise, but a drawback is that the signal must be sampled according to the Nyquist rate with at least the bandwidth of the signal, which can be very high in modern radar systems. A technique to avoid these high sampling rates but still achieves the high range resolution is presented in the next section.

2.1.3.2 Deramping

The matched filter processing described in the previous section requires a continuous recording of the incoming data with at least the bandwidth of the radar signal. In this way, the radar data can be recorded over the unambiguous range interval given in

(2.1.12), but has the drawback of the high sampling rate and additionally produces a lot of data that eventually must be stored. An alternative method to achieve the high resolution of a chirp waveform with reduced sampling requirements is the *deramping* method. This processing scheme reduces the required sampling frequency enormously with the cost of a reduced observable range interval. Its basic steps are shown in Figure 2.1.7.



source of radar graphic:
https://openclipart.org

Figure 2.1.7: Deramping procedure

The core idea is that a complex conjugated replica $r(t)$ of the transmitted chirp is used in the analog demodulation of the received echo and therefore, the reception of a chirp will result in a constant frequency. For a chirp waveform with carrier frequency f_0, see (2.1.20), the replica, which is started with a delay t_0, is given by

$$r(t) = e^{-j2\pi\left(f_0(t-t_0)+\frac{\mu}{2}(t-t_0)^2\right)}. \tag{2.1.32}$$

In principle, the replica is not restricted in time and can have a higher bandwidth than the transmitted chirp, but always the same chirp rate. The mixing, i.e. multiplication, of this replica with the time delayed echo $s(t-\Delta t)$ of the transmitted LFM pulse[6] caused by a point target at range R

$$s(t-\Delta t) = rect\left(\frac{t-\Delta t}{\tau_P}\right) e^{j2\pi\left(f_0(t-\Delta t)+\frac{\mu}{2}(t-\Delta t)^2\right)} \tag{2.1.33}$$

gives the result

---

[6] The amplitude of the echo is ignored, since it is not important for the understanding of the deramping.

$$r(t)s(t-\Delta t) = rect\left(\frac{t-\Delta t}{\tau_P}\right) e^{j2\pi\left(f_0(t_0-\Delta t)+\frac{\mu}{2}(\Delta t^2-t_0^2)\right)} e^{j2\pi\mu(t_0-\Delta t)t}. \tag{2.1.34}$$

The product of (2.1.34) can be split into three parts, the first one is the rect function, which limits the output in time. The second factor is an exponential term that is independent of time and thus represents a phase factor. The important part is the second exponential term, which depends on t and is thus a constant frequency proportional to the difference $(t_0 - \Delta t)$. If the time delay t_0 is considered as the delay to a reference point R_0, given by $t_0 = 2R_0/c_0$, the frequency of the output signal can be written as

$$f(t) = \frac{1}{2\pi}\frac{\partial}{\partial t} 2\pi\mu(t_0-\Delta t)t = \mu(t_0-\Delta t) = \frac{2\mu}{c_0}(R_0 - R). \tag{2.1.35}$$

Equation 2.1.35 shows that the frequency of the output signal is proportional to the relative distance of the actual scatterer, which has caused the echo, and the reference point defined by the starting time of the reference chirp. Therefore, the bandwidth of the deramping output depends on the *range window* that should be covered and is normally much smaller than the bandwidth of the waveform itself. This reduces the demands on the analog to digital conversion hardware compared to a matched filter. A further benefit of this technique is that the range resolution calculated above as $c_0/2B$ is still valid; the drawback of this method is that the areas outside the range window are lost. To convert the output signal into an HRR profile only a Fourier transform is necessary. A HRR profile measured with the TIRA system, which uses the deramping method, is shown in Figure 2.1.8. The profile belongs to the data that is used in this thesis. It was measured while the target was flying more than 100 km away, but the range axis shows only around 40 m. It should be mentioned that the original profile measured by the TIRA system is around 200 m [Meh96] and the example is cropped to the significant part containing the target.

To achieve a more handy signal model for the later application, some steps are necessary. The first one is the introduction of the k-space notation that uses the wavenumber k, which is defined by $2\pi f/c_0$ and known from the propagation of electromagnetic waves [Orf16]. For the chirp, or any other waveform, the wavenumber k_r of the propagating wave can be split into two parts, the first one k_0 represents the contribution of the carrier frequency f_0 and the second one k_b the baseband contribution μt. Therefore, the wavenumber can be written as

$$k_r = k_0 + k_b = \frac{2\pi}{c_0}(f_0 + \mu t), \tag{2.1.36}$$

which corresponds to the multiplication of the instantaneous chirp frequency of (2.1.21) with the constant factor $2\pi/c_0$. Using this definition, the output (2.1.34) of the deramping process can be idealized in k-space as

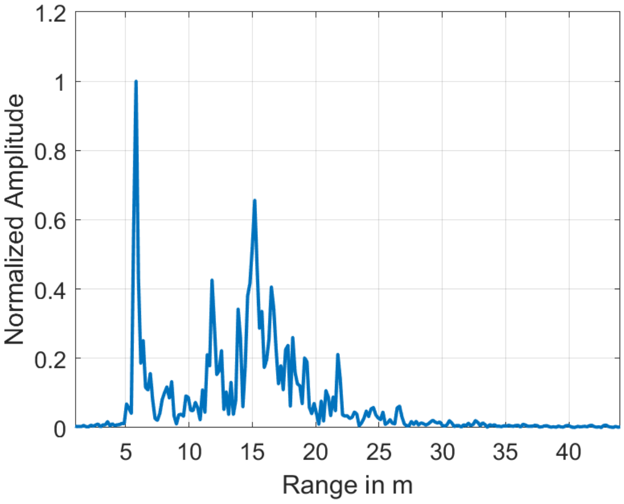


Figure 2.1.8: Example of a TIRA HRR profile

$$s(k_r, R) = e^{-j2k_r R}, \tag{2.1.37}$$

which is a normalized signal model for a point scatterer at range R [End10] and is called *normal form* from here on. The quadratic phase term of (2.1.34) is ignored in (2.1.37) and since the range variable R is used directly, the reference range R_0 is set to zero. In this case a chirp with an infinite duration would be needed, but this expression is the most general form and can be used for any waveform. For the deramping process, a normal form with relative range $\Delta R = R - R_0$ can be introduced as

$$s(k_r, R) = e^{-j2k_r \Delta R}. \tag{2.1.38}$$

Nevertheless, in the remainder of this thesis, the general expression of (2.1.37) is used and for the later application in Chapter 5, the range is converted to the relative range.

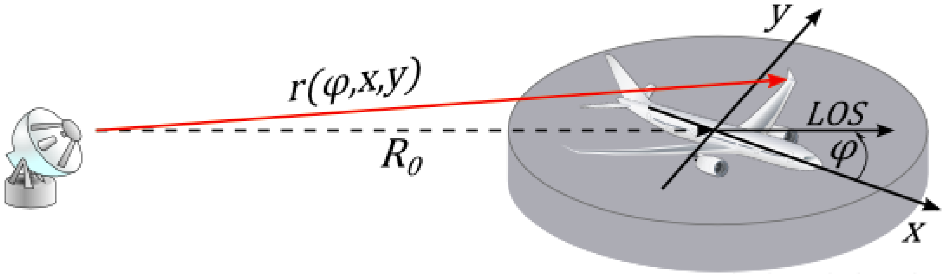
The benefit of the normal form is that the temporal signal of (2.1.34) is converted to a spatial signal and thus a spatial Fourier transform can be used to create the HRR profile. This step introduces the range directly in the signal model and therefore, this model will also be used in the next section on ISAR imaging. The extension of this model to two dimensions gives as result the model that will be used later in the decomposition framework.

2.2 Basics of ISAR Imaging

In this section the fundamentals of ISAR imaging are introduced as far as necessary to understand the basic signal processing. As a starting point, the basic turntable imaging will be introduced in Section 2.2.1, since the aim of ISAR imaging is to reproduce this situation in in-flight measurements of an aircraft. The basic elements of aircraft imaging will be presented in Section 2.2.2 without giving details of some necessary steps to reproduce the turntable measurement.

2.2.1 Turntable Imaging

Turntable imaging is used as a reference for ISAR imaging, since the processing is the same, although the situation in ISAR is far more complicated as will be shown later. The turntable is used to induce a motion of the target, since for any synthetic aperture radar imaging system some sort of motion is necessary. In ISAR imaging the target is moving, while the radar is stationary, which is shown in Figure 2.2.1 for a turntable measurement. This situation is contrary to the situation in the more familiar SAR imaging, where a moving platform is used to image a certain static region of interest. This reversed situation was also the motivation for the name *Inverse* SAR.



source of airplane and radar graphics:
https://openclipart.org

Figure 2.2.1: Turntable measurement

The situation on a turntable is simplified in many ways compared to in-flight measurements. The most obvious one is the known movement of the target with a fixed rotational speed ω and the rotation is limited to the x-y plane and no z-component of the motion is present. To derive the resulting radar signal of the turntable situation, it is assumed that the radar works in far-field condition and the distance R_0 to the radar is much larger than the extent of the target. The center of the x-y coordinate system lies in the geometric center of the target and the x-axis corresponds to the line of sight (LOS) of the radar for an azimuth angle $\varphi = 0$.

For the derivation of a signal model, a scatterer at position (x, y) in the local coordinate system is assumed. For $R_0 \gg x, y$, the distance from the radar to the scatterer r can be approximated as

$$
\begin{aligned}
r(\varphi, x, y) &= \sqrt{(R_0 \cos\varphi + x)^2 + (R_0 \sin\varphi + y)^2} \\
&\approx R_0 + x\cos\varphi + y\sin\varphi, \qquad (2.2.1)
\end{aligned}
$$

which can be included in the normal form of (2.1.37). To get a signal model in the local coordinates x and y, the distance R_0 is assumed to be known and subtracted from the scatterer range. The normal form for a scatterer in x-y-coordinates is thus given as

$$s(k_r, \varphi, x, y) = e^{-j2k_r(r(\varphi,x,y)-R_0)} \approx e^{-j2k_r(x\cos\varphi + y\sin\varphi)}. \qquad (2.2.2)$$

To calculate the received signal from the model, the target is assumed to be represented by the reflectivity distribution $a(x, y)$. With this distribution, the received noise free signal can be written as

$$y(k_r, \varphi) \approx \iint e^{-j2k_r(x\cos\varphi + y\sin\varphi)} a(x, y) dxdy. \qquad (2.2.3)$$

The next step is a variable substitution introducing the variables k_x and k_y that is known as *polar reformatting* and is realized by

$$
\begin{aligned}
k_x &= 2k_r \cos\varphi \qquad (2.2.4) \\
k_y &= 2k_r \sin\varphi. \qquad (2.2.5)
\end{aligned}
$$

With the definitions of (2.2.4) and (2.2.5), the variables k_r and φ can be expressed as

$$
\begin{aligned}
k_r &= \frac{1}{2}\sqrt{k_x^2 + k_y^2} \qquad (2.2.6) \\
\varphi &= \cot^{-1}\left(\frac{k_x}{k_y}\right). \qquad (2.2.7)
\end{aligned}
$$

Using this variable substitution, the received signal of (2.2.3) can be written as

$$y(k_x, k_y) = \iint e^{-j(k_x x + k_y y)} a(x, y) dxdy, \qquad (2.2.8)$$

which corresponds to the spatial Fourier transform of $a(x, y)$ in the variables k_x and k_y. Therefore, an estimate $\hat{a}(x, y)$ of the target reflectivity, i.e. an image of the target, is calculated by an inverse 2-d Fourier transform of the received data collected over a certain range of aspect angles φ.

The resolution of the estimate $\hat{a}(x, y)$ is defined by the spatial sampling theorem and is given by (2.1.31) for the down-range direction. The resolution of the cross range dimension ΔCR, i.e. the dimension perpendicular to the line of sight is given by

$$\Delta CR = \frac{\lambda_0}{2\varphi_{int}}, \tag{2.2.9}$$

where λ_0 is the wavelength of the carrier frequency and φ_{int} is the *integration angle*, i.e. the covered range of aspect angles.

For the practical implementation normally a Fast Fourier Transform (FFT) is applied. Since the measurements described above are sampled on a circular ring segment, a direct application of the FFT, which works in rectangular coordinates, will give blurring effects in the image, especially for large spans of φ. To achieve an image without these blurring effects, the standard procedure is to perform an interpolation from polar coordinates to a rectangular grid. Nevertheless, the direct application of the FFT is also very widespread, since the errors that occur due to the wrong grid are negligible compared to errors due to imperfect knowledge of the target motion. The direct application of the FFT without interpolation is known as *range-Doppler-imaging* and the ISAR images measured by the TIRA system shown in this thesis are generated with this technique without interpolation. The reasons why this simplified processing is used are discussed in the next section, which shows some problems that arise for in-flight ISAR imaging of aircrafts.

2.2.2 Imaging of an Aircraft

Compared to the situation of turntable imaging, the situation of in-flight measurements is far more complicated. The first difference is of course the unknown movement of the target. The rotation of the turntable experiment can also be described by a rotational vector $\mathbf{\Omega}$,which points in the direction of the z-coordinate and is of constant length, since the rotation is around the z-axis with a constant rotational velocity. An aircraft can principally rotate around each axis during the flight, which gives three rotation vectors. The movements around the different axes of an airplane are called *roll*, *pitch* and *yaw* and are depicted in Figure 2.2.2 together with the overall rotation vector $\mathbf{\Omega}_{Ang}$, which is the vector sum of the three rotation vectors[7].

Another important difference to turntable imaging is the varying range R_0 between the radar and the center of the target. This range history must be estimated during the measurement and compensated as shown in (2.2.2). This range estimation can be done for example from the tracking information of the target, but is not considered further here. Details about range tracking can be found in literature focusing on ISAR, e.g. [CM14]. Another effect of the varying range is that the aspect angle of the target is changing, as long as the target is not flying along the LOS of the radar or an exact

---

[7] Roll, pitch and yaw have an intrinsic dependency on the maneuver the target is flying, which is ignored here. For the derivation here, only the fact that all three have a contribution to the resulting rotational vector is important. Furthermore, the depicted movement is not supposed to be realistic for an airliner.

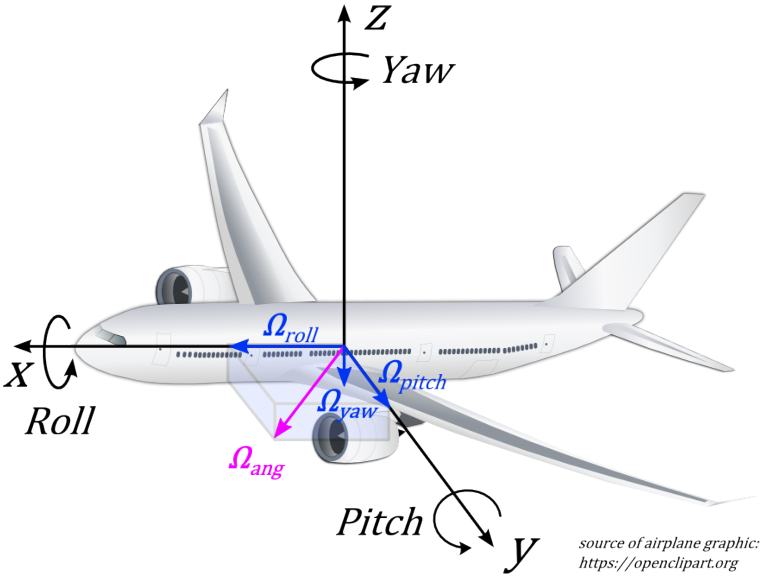


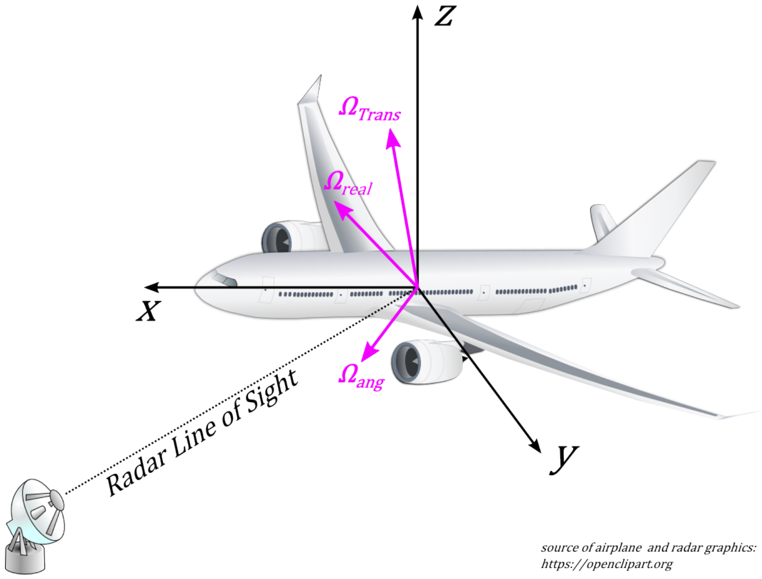
Figure 2.2.2: Aircraft with rotation vector due to its own motion

circle around it. In these two cases the aspect angle will remain constant during the measurement. The changing aspect angle due to the translational motion of the target can also be modeled as a rotation, which is represented by the rotation vector $\mathbf{\Omega}_{Trans}$. This rotation vector is shown in Figure 2.2.3 together with the rotation vector from Figure 2.2.2. The vector sum of both rotation vectors gives the actual rotation vector $\mathbf{\Omega}_{real}$ of the target. The changing aspect angle due to the varying range allows ISAR imaging even if no roll, pitch or yaw movement is present. In fact, the estimation of the rotation vector becomes easier if the target is flying a straight line and is not performing any maneuvers.

In (2.2.3) the target was modeled as the two-dimensional reflectivity distribution $a(x, y)$, but a real target is a three-dimensional object with a corresponding 3-d reflectivity distribution. In case of ISAR imaging a projection of the true reflectivity distribution on two dimensions is measured. The *image projection plane* (IPP) is given here without proof, further details can be found for example in [CM14] or [Weh95].

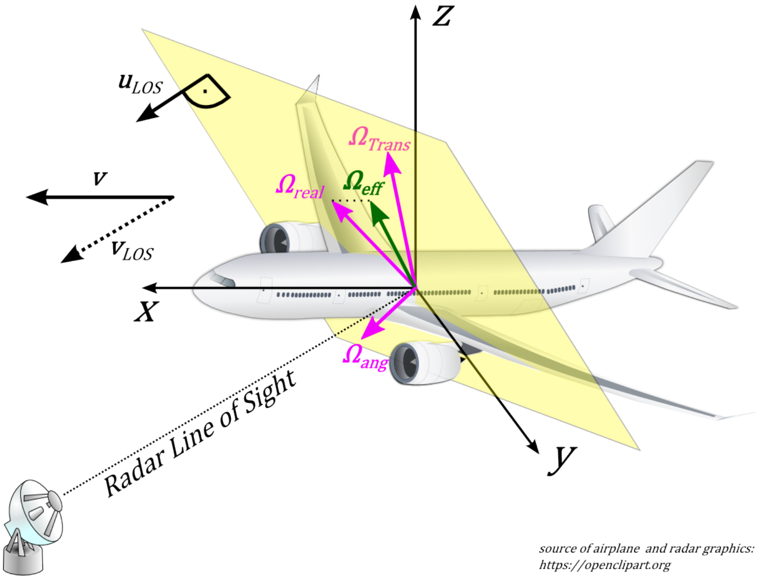
The IPP is defined by the so-called *effective rotation vector*, which is itself the result of a projection of the true rotation vector of the target on the plane that is perpendicular to the LOS of the radar. This projection is shown as an intermediate step in Figure 2.2.4, which uses the true rotation vector from Figure 2.2.3.

This effective rotation vector is the decisive vector that defines the scaling and projection of the resulting ISAR image. It was shown in (2.2.8) that the coverage of the aspect angle during the measurement defines the resolution of the cross-range dimension. It is known from the theory of Fourier transform that the span of the covered frequencies defines the resolution in the spatial domain [OL07] and it was shown in



source of airplane and radar graphics:
https://openclipart.org

Figure 2.2.3: Aircraft with real rotation vector



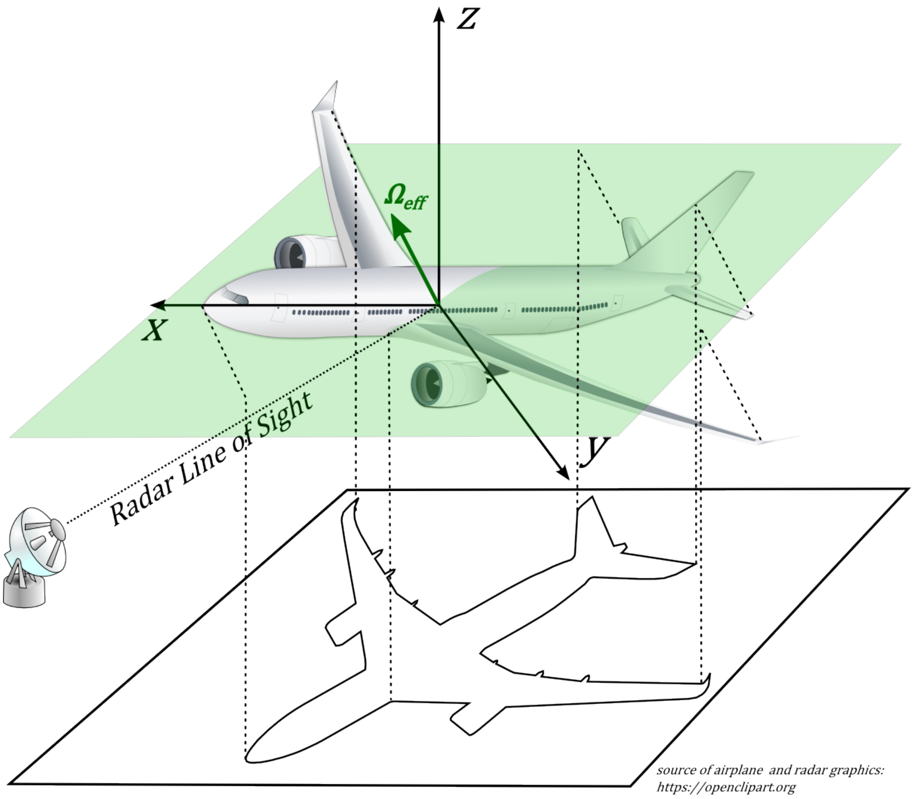
source of airplane and radar graphics:
https://openclipart.org

Figure 2.2.4: Projection of the real rotation vector on the plane perpendicular to the LOS

(2.2.4) and (2.2.4) that the aspect angle is directly related to k_x and k_y, which define the domain of the spatial Fourier transform in (2.2.8). Therefore, the resolution and the scaling of the cross-range axis in the resulting image is given by the covered span of aspect angles, since the image is processed by an inverse Fourier transform. This means that the knowledge of the effective rotation vector is necessary to scale the cross-range axis and to estimate the width of the target. With a monostatic radar like the TIRA system only the magnitude of the effective rotation vector can be estimated, the direction of the effective rotation vector is ambiguous in that case, since it is always possible to find more than one effective rotation vector that can cause the changes of the data during the measurement [Bri16]. One way to support the estimation of the effective rotation vector is the use of tracking data, which can be used with assumptions on the flight mechanics to give a reasonable estimate. Nevertheless, the problem of motion estimation of targets is an ongoing research topic, e.g. [Bri16], which is not covered here in detail.

The main reason why the direction of the effective rotation vector is so important is the IPP of the ISAR image. The IPP is given as the plane that is perpendicular to the effective rotation vector. This principle is visualized in Figure 2.2.5, which shows that knowledge about the orientation of the IPP is necessary to estimate the dimensions of the target, since the projected length of the target is inverse proportional to the cosine of the angle between the target axis and the IPP.

A point that has not been questioned so far is the validity of the model that the echo produced by the target looks exactly like the transmitted waveform. Although this common assumption of isotropic scattering is the basis of most ISAR processing algorithms, there are types of scattering that change the waveform as will be shown in the next chapter.



source of airplane and radar graphics:
https://openclipart.org

Figure 2.2.5: Aircraft with projection on the image projection plane

Chapter 3

Waveguide Scattering Model

In this chapter the signal model for jet engines is derived. Since this model is physically motivated, the fundamentals of the propagation of electromagnetic waves in waveguides are described in Section 3.1. The actual model that will be used later is shown in Section 3.2 and an example for a scattered chirp waveform is given in Section 3.3.

3.1 Propagation of Electromagnetic Waves in Waveguides

Before the special case of an electromagnetic wave inside a waveguide, which is confined by conducting walls and thus creates certain boundary conditions is described in Section 3.1.2, some basic electromagnetic theory is given in Section 3.1.1. The content of these two sections can also be found in textbooks like [Orf16], [Str41] or [Geo97], therefore citations for certain formulas or statements are only given if they are outside of the basic content. The description is given for completeness and is reduced and tailored to the elements that are necessary for the signal propagation model of Section 3.1.3.

3.1.1 The Electromagnetic Field in Free Space

Since the waveguide model is derived for jet engines, which are filled with air, the simplified *Maxwell equations* for the source-free case[1]

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{3.1.1}$$

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} \tag{3.1.2}$$

$$\nabla \cdot \mathbf{D} = 0 \tag{3.1.3}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{3.1.4}$$

---

[1] The nabla operator ∇ is a vector that contains the partial derivative operators of the corresponding components, i.e. $\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)^T$.

are used for further derivations.

Equation (3.1.1) is called *Faraday's law of induction*, Equation (3.1.2) is known as *Ampere's law* and Equations (3.1.3) and (3.1.4) are called *Gauss' laws* of the electric and magnetic field, respectively. The quantities **E** and **H** denote the electric and magnetic *fields*, while **D** and **B** describe the electric and magnetic *flux densities*.

The electric and magnetic flux densities **D** and **B** are related to the electric and magnetic field by the *permittivity* ϵ and *permeability* μ of the medium via

$$\mathbf{D} = \epsilon\mathbf{E} \tag{3.1.5}$$

$$\mathbf{B} = \mu\mathbf{H}. \tag{3.1.6}$$

Since only free space propagation is considered here, the values ϵ_0 and μ_0 of vacuum are used throughout the thesis. This approximation is done since dry air has similar electromagnetic properties as vacuum [CB09]. Performing a curl operation in (3.1.1) and using the result of (3.1.2) together with the relations in (3.1.5) and (3.1.6) shows that the electric field satisfies

$$\nabla \times \nabla \times \mathbf{E} = -\mu_0\epsilon_0\frac{\partial^2\mathbf{E}}{\partial t^2}, \tag{3.1.7}$$

which can be written as

$$\nabla^2\mathbf{E} = \mu_0\epsilon_0\frac{\partial^2\mathbf{E}}{\partial t^2} \tag{3.1.8}$$

using the vector identity $\nabla \times (\nabla \times \mathbf{E}) = \nabla(\nabla \cdot \mathbf{E}) - \nabla^2\mathbf{E}$ and $\nabla \cdot \mathbf{E} = 0$, which follows from (3.1.3) and (3.1.5). Equation (3.1.8) is a wave equation [Far93] with the propagation speed c given by

$$c = \frac{1}{\sqrt{\epsilon_0\mu_0}}. \tag{3.1.9}$$

With (3.1.9), the wave equation (3.1.8) can be written in the known form

$$\nabla^2\mathbf{E} - \frac{1}{c^2}\frac{\partial^2\mathbf{E}}{\partial t^2} = 0. \tag{3.1.10}$$

In the same way, Equation (3.1.2) can be changed with the result of (3.1.1) to a wave equation for the magnetic field

$$\nabla^2\mathbf{H} - \frac{1}{c^2}\frac{\partial^2\mathbf{H}}{\partial t^2} = 0. \tag{3.1.11}$$

Therefore, each component of the electric and magnetic field satisfies a scalar wave equation in cartesian coordinates. Furthermore, the electric and the magnetic field is perpendicular to each other, which follows directly from the first two Maxwell equations. This perpendicularity together witch the wave equations give the basic form of an electromagnetic wave, which is depicted in Figure 3.1.1.

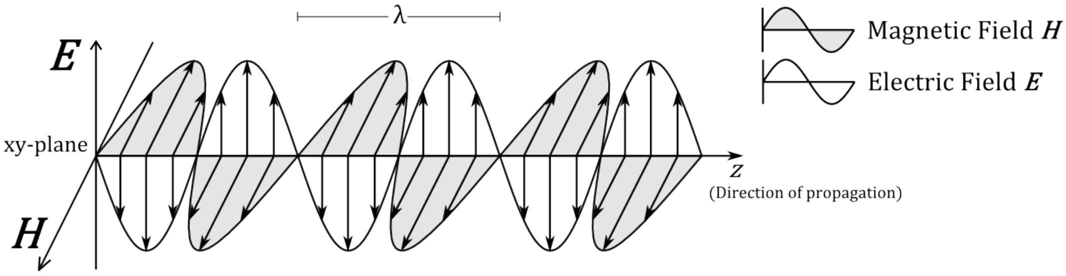


Figure 3.1.1: Electromagnetic wave in free space

Figure 3.1.1 also shows the principle of a *plane wave.* In free space it is assumed that the value of **E** and **H** depend only on the direction of propagation, which is z throughout this section on waveguides. This means that all over the xy-plane **E** and **H** have constant values and $E_z = H_z = 0$. The full space-time forms of **E** and **H** of the free space propagating wave are given by

$$\mathbf{E}(z,t) = \mathbf{E}(0,0)e^{j(\omega t - k_z z)} \qquad \mathbf{H}(z,t) = \mathbf{H}(0,0)e^{j(\omega t - k_z z)}, \tag{3.1.12}$$

where ω denotes the angular frequency and k_z is the z-component of the wavenumber, which is defined by

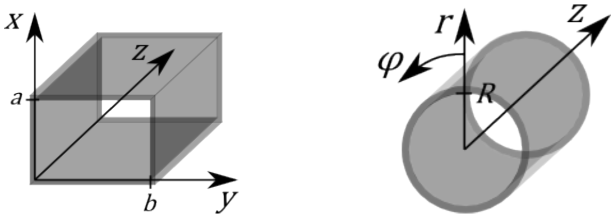
$$k_z = \frac{2\pi}{\lambda} = \frac{\omega}{c}. \tag{3.1.13}$$

In Equation 3.1.13, λ is the *wavelength* and every cut along the z-axis in Figure 3.1.1 parallel to the xy-plane in the distance of λ will result in the same values for **E** and **H**. These cuts are called *phase fronts* and are the reason for the name plane waves, since at every constant z value, a plane with constant **E** and **H** can be found.

Free space waves are called *transversal electromagnetic waves* (TEM-waves), which refers to the components that are transversal, i.e. perpendicular, to the direction of propagation. **E** and **H** are perpendicular to the direction of propagation, since there are no components in z-direction, i.e. $H_z = E_z = 0$. In the next section, it will be shown that TEM-waves are not possible inside waveguides due to the boundary conditions.

3.1.2 The Electromagnetic Field inside a Waveguide

The derivation of the electromagnetic field is shown for rectangular waveguides in cartesian coordinates, since this notation is more intuitive than the notation in cylindrical coordinates. All results from the cartesian coordinates can be transferred to cylindrical waveguides, which is shown at the end of this part. The coordinate systems for both kinds of waveguides are shown in Figure 3.1.2.



(a) Rectangular (b) Cylindrical

Figure 3.1.2: Waveguide shapes and the corresponding coordinate systems

The confining walls in the cartesian system are assumed to coincide with the x- and the y-axis of the coordinate system. These confining walls posses an inner diameter of a on the x-axis and b on the y-axis, whereas in the cylindrical system, the waveguide has a certain circular opening diameter d. As mentioned already in Section 3.1.1, the direction of propagation is the z-axis and the waveguide is filled with air, i.e. $\mu = \mu_0$ and $\epsilon = \epsilon_0$.

Furthermore, the material of the waveguides is assumed to be a *perfect electric conductor* (PEC), which means a lossless material. From this assumption follows that the electric field is perpendicular and the magnetic field is parallel to these walls at their surface [Orf16]. The existence of these walls implies also that the assumption of plane waves, as in free space, is not possible here. Since the PEC walls must have the same electric potential everywhere in a xy-plane, at least one component, either $\mathbf{E}$ or $\mathbf{H}$, must vary somehow between these walls, otherwise there will be only a static field and no wave. Due to this reason, a cut in the xy-plane must have different values in at least one component and compared to the TEM-waves of Section 3.1.1 none of the components of $\mathbf{E}$ and $\mathbf{H}$ can assumed to be zero.

Since the confining waveguide walls are only in the x- and y-direction of the coordinate system, there is no restriction on the wave propagation along the z-direction, i.e. it is basically possible that a wave propagates with a wavenumber k_z along the z-direction. This last assumption is listed and numbered here for an overview and reference to them in the later derivations together with the assumptions and boundary conditions mentioned before:

1. The Waveguide is filled with air, $\mu = \mu_0$ and $\epsilon = \epsilon_0$.
2. Confining walls are PEC materials, i.e. electric field lines are perpendicular and magnetic field lines are parallel to the conducting surface.
3. $\mathbf{E}$ and $\mathbf{H}$ do not only depend on the direction of propagation, $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\mathbf{E} \neq 0$ and $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\mathbf{H} \neq 0$.
4. None of the components is assumed to be zero, $E_x, E_y, E_z \neq 0$ and $H_x, H_y, H_z \neq 0$.

5. A lossless wave propagation along the z-direction is possible and represented by the factor $e^{-jk_z z}$.

3.1.2.1 Solution inside a Rectangular Waveguide

The starting point for the derivation of the field inside a waveguide is the wave equation in (3.1.10). The only time dependency of the waves considered here is a $e^{j\omega t}$ term, which means that the derivative with respect to time can be replaced by $j\omega$ and thus can (3.1.10) also be written as

$$\nabla^2 \mathbf{E} + \frac{\omega^2}{c^2}\mathbf{E} = 0. \tag{3.1.14}$$

The expression in (3.1.14) is a wave equation without time derivatives and is called *Helmholtz equation*[2]. In the free space case with TEM waves, i.e. $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\mathbf{E} = 0$, the only spatial dependency is given by the $e^{-jk_z z}$ term and the Helmholtz equation is reduced to

$$\left(\frac{\omega^2}{c^2} - k_z^2\right)\mathbf{E} = 0. \tag{3.1.15}$$

With the definition (3.1.13) and (3.1.15) can be seen that (3.1.12) fulfills the wave equation under free space conditions . To formulate the field inside waveguides, the Laplacian operator in (3.1.14) can be separated into a *longitudinal* term, i.e. in the direction of propagation, and a *transversal* term ∇_T^2. According to the above mentioned conditions 3 and 5, the spatial dependency in z-direction is given by $e^{-jk_z z}$ and thus the Laplacian in this direction is reduced to $-k_z^2$. With this condition, the Laplace operator is written as

$$\nabla^2 = \underbrace{\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}}_{transversal} + \underbrace{\frac{\partial^2}{\partial z^2}}_{longitudinal} = \nabla_T^2 + \frac{\partial^2}{\partial z^2} = \nabla_T^2 - k_z^2. \tag{3.1.16}$$

With this separation, the Helmholtz equation can be written as

$$\nabla_T^2\mathbf{E} - k_z^2\mathbf{E} + \frac{\omega^2}{c^2}\mathbf{E} = \nabla_T^2\mathbf{E} + \left(\frac{\omega^2}{c^2} - k_z^2\right)\mathbf{E} = 0. \tag{3.1.17}$$

Since the expression $\left(\frac{\omega^2}{c^2} - k_z^2\right)$ will be used very often in the following text, the symbol κ_c^2 is used for this expression

$$\kappa_c^2 = \left(\frac{\omega^2}{c^2} - k_z^2\right), \tag{3.1.18}$$

---

[2] Generally, any differential equation for an arbitrary vector field $\mathbf{A}$ with the scalar value λ of the form $\nabla^2\mathbf{A} + \lambda\mathbf{A} = 0$ is called Helmholtz equation [Far93]

and thus the Helmholtz equation for waveguides can be shortly noted as

$$\nabla_T^2 \mathbf{E} + \kappa_c^2 \mathbf{E} = 0. \tag{3.1.19}$$

Equation (3.1.19) specifies an eigenvalue problem for the transversal Laplace operator. The solution of this equation gives the eigenfunctions **E** with the eigenvalues κ_c, which determine a field form that can propagate along the z-direction inside the waveguide. To find a solution for this problem, the partial differential equation (PDE) for the z-component of the electric field

$$\frac{\partial^2 E_z}{\partial x^2} + \frac{\partial^2 E_z}{\partial y^2} + \kappa_c^2 E_z = 0 \tag{3.1.20}$$

must be solved and the approach of variable separation that is used here can only justified by the later success. The component E_z is decomposed into a part $X(x)$ that only depends on x and $Y(y)$, which only depends on y. Using these terms, the amplitude $\hat{E}$ and the propagation factor $e^{-jk_z z}$, the field can be written as

$$E_z(x,y,z) = \hat{E} X(x) Y(y) e^{-jk_z z}. \tag{3.1.21}$$

Since the amplitude and the propagation factor are the same for all terms in (3.1.20), they are not relevant for the determination of the possible field forms and can be removed for simplicity. After this reduction, only $X(x)$ and $Y(y)$ remain and the PDE (3.1.20) becomes

$$\frac{\partial^2 X(x)}{\partial x^2} Y(y) + X(x) \frac{\partial^2 Y(y)}{\partial y^2} + \kappa_c^2 X(x) Y(y) = 0. \tag{3.1.22}$$

By the multiplication of (3.1.22) with $\frac{1}{X(x)Y(y)}$, a PDE with terms that depend only on a single variable is created

$$\frac{\partial^2 X(x)}{\partial x^2} \frac{1}{X(x)} + \frac{\partial^2 Y(y)}{\partial y^2} \frac{1}{Y(y)} + \kappa_c^2 = 0. \tag{3.1.23}$$

Collecting terms, and the use of the variable substitution

$$k_x^2 = \frac{\partial^2 Y(y)}{\partial y^2} \frac{1}{Y(y)} + \kappa_c^2 \tag{3.1.24}$$

$$k_y^2 = \frac{\partial^2 X(x)}{\partial x^2} \frac{1}{X(x)} + \kappa_c^2 \tag{3.1.25}$$

yield the two ordinary differential equations (ODEs)

$$\frac{\partial^2 X(x)}{\partial x^2} + k_x^2 X(x) = 0 \tag{3.1.26}$$

$$\frac{\partial^2 Y(y)}{\partial y^2} + k_y^2 Y(y) = 0. \tag{3.1.27}$$

These ODEs are oscillating ODEs and the solution must fulfill the boundary condition 2, which means that the tangential components of the electric field must be zero along the surface of the waveguide

$$E_y = E_z = 0 \quad \text{at} \quad x = 0 \text{ and } x = a \tag{3.1.28}$$

$$E_x = E_z = 0 \quad \text{at} \quad y = 0 \text{ and } y = b. \tag{3.1.29}$$

The solution of (3.1.26) and (3.1.27) can be given directly as a $\sin(k_x x)$ and $\sin(k_y y)$, where k_x and k_y can be only discrete values that guarantee that the boundary conditions (3.1.28) and (3.1.29) are fulfilled

$$X(x) = \sin(k_x^{TM} x) \quad \text{with} \quad k_x^{TM} = m\frac{\pi}{a} \quad m = 1, 2, 3, \dots \tag{3.1.30}$$

$$Y(y) = \sin(k_y^{TM} y) \quad \text{with} \quad k_y^{TM} = n\frac{\pi}{b} \quad n = 1, 2, 3, \dots . \tag{3.1.31}$$

In case of $m = 0$ or $n = 0$ the expression in (3.1.21) will be zero and there will be no propagating wave. The superscript TM with k_x and k_y is used to distinguish this solution for the E-modes from the later shown solution of the H-modes. The derivation until here shows that a component of the electric field in direction of propagation is possible, which is contrary to the TEM-wave in Section 3.1.1.

To solve the Helmholtz equation for the magnetic field **H** an expression for the transversal components x and y that depends only on the z-component is necessary to formulate the boundary conditions for the PDE. To get this formulation, (3.1.1) and (3.1.2) are used together with $\frac{\partial}{\partial t} = j\omega$ and $\frac{\partial}{\partial z} = -jk_z$, which were already used to create (3.1.19). The components of the first Maxwell equation (3.1.1) give

$$\frac{\partial E_z}{\partial y} + jk_z E_y = -j\omega\mu H_x \tag{3.1.32}$$

$$-jk_z E_x - \frac{\partial E_z}{\partial x} = -j\omega\mu H_y \tag{3.1.33}$$

$$\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} = -j\omega\mu H_z. \tag{3.1.34}$$

The second Maxwell equation can be written as

$$\begin{aligned} \frac{\partial H_z}{\partial y} + jk_z H_y &= j\omega\epsilon E_x && (3.1.35) \\ -jk_z H_x - \frac{\partial H_z}{\partial x} &= j\omega\epsilon E_y && (3.1.36) \\ \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} &= j\omega\epsilon E_z. && (3.1.37) \end{aligned}$$

With (3.1.9) and (3.1.18), expressions for E_x, E_y, H_x and H_y can be found that only depend on E_z and H_z

$$\begin{aligned} E_x &= \frac{-j}{\kappa_c^2}\left(k_z\frac{\partial E_z}{\partial x} + \omega\mu\frac{\partial H_z}{\partial y}\right) && (3.1.38) \\ E_y &= \frac{j}{\kappa_c^2}\left(-k_z\frac{\partial E_z}{\partial y} + \omega\mu\frac{\partial H_z}{\partial x}\right) && (3.1.39) \\ H_x &= \frac{j}{\kappa_c^2}\left(\omega\mu\frac{\partial E_z}{\partial y} - k_z\frac{\partial H_z}{\partial x}\right) && (3.1.40) \\ H_y &= \frac{-j}{\kappa_c^2}\left(\omega\mu\frac{\partial E_z}{\partial x} + k_z\frac{\partial H_z}{\partial y}\right). && (3.1.41) \end{aligned}$$

With these equations, the boundary conditions for the magnetic field can be formulated and since the tangential components of the electric field must still vanish, the boundary conditions (3.1.28) and (3.1.29) must still be fulfilled. From (3.1.28) follows with (3.1.39) that

$$\frac{\partial H_z}{\partial x} = 0 \quad \text{at} \quad x = 0 \text{ and } x = a. \qquad (3.1.42)$$

In the same way follows from (3.1.29) and (3.1.38) that

$$\frac{\partial H_z}{\partial y} = 0 \quad \text{at} \quad y = 0 \text{ and } y = b. \qquad (3.1.43)$$

To find the solution for the magnetic field, the steps from (3.1.14) to (3.1.27) are repeated with **H** instead of **E**. The physical meaning of the different boundary conditions is that not the longitudinal component itself must fulfill the condition, but the derivative with respect to the transversal component. This means that the derivatives must be sinus functions, i.e. the dependency of the z-component of **H** on x and y is given by cosine functions

$$\begin{aligned} X(x) &= \cos(k_x^{TE}x) \qquad \text{with} \qquad k_x^{TE} = m\frac{\pi}{a} \quad m = 0, 1, 2, \ldots && (3.1.44) \\ Y(y) &= \cos(k_y^{TE}y) \qquad \text{with} \qquad k_y^{TE} = n\frac{\pi}{b} \quad n = 0, 1, 2, \ldots. && (3.1.45) \end{aligned}$$

In this case it is possible that either n or m is zero, but $m = n = 0$ is not possible, since this would result in a static field and there would be no propagating wave. As already mentioned above for the electric field, this derivation shows that a component of the magnetic field in direction of propagation is possible in waveguides. The different values of k_x and k_y define *modes*, which determine the field in x and y direction of the z component of the propagating field as

$$\begin{aligned} E_z^{(m,n)}(x,y,z) &= \hat{E}\sin(k_x^{TM}x)\sin(k_y^{TM}y)e^{-jk_z z} \\ &= \hat{E}\sin(m\frac{\pi}{a}x)\sin(n\frac{\pi}{b}y)e^{-jk_z z} \qquad (3.1.46) \\ H_z^{(m,n)}(x,y,z) &= \hat{H}\cos(k_x^{TE}x)\cos(k_y^{TE}y)e^{-jk_z z} \\ &= \hat{H}\cos(m\frac{\pi}{a}x)\cos(n\frac{\pi}{b}y)e^{-jk_z z}. \qquad (3.1.47) \end{aligned}$$

Dependent on the components that are present in the direction of propagation, four cases can be distinguished

$$E_z = 0,\ H_z = 0 \qquad \text{TEM-mode} \qquad (3.1.48)$$
$$E_z \neq 0,\ H_z = 0 \qquad \text{TM- or E-mode} \qquad (3.1.49)$$
$$E_z = 0,\ H_z \neq 0 \qquad \text{TE- or H-mode} \qquad (3.1.50)$$
$$E_z \neq 0,\ H_z \neq 0 \qquad \text{Hybrid mode.} \qquad (3.1.51)$$

As already mentioned above, the existence of TEM-modes is not possible inside of waveguides, since the solutions of the Helmholtz equations require a z-component of at least one filed component, either $\mathbf{E}$ or $\mathbf{H}$. The coexistence of E_z and H_z is not considered in this thesis, since it require a model with finite conductivity of the confining walls, which violates assumption 2 that was already used in the derivation above. Due to this reasons, the only remaining modes for the model used here are TM- and TE-modes. An example for a vectorial plot of $\mathbf{E}$ and $\mathbf{H}$ is shown in Figure 3.1.3; the electric field has a component in z-direction and the vector of the magnetic field lies within the xy-plane, thus it is a TM- or E-mode.

With the expressions (3.1.46) and (3.1.47), the x and y components of the field can be calculated using (3.1.38) to (3.1.41). Since it is not possible to draw the complete electrical field within a waveguide, a visualization of the E_y component of an $\mathbf{E}^{(3,n)}$ mode for a certain y-value is depicted in Figure 3.1.4.

The order $m = 3$ in Figure 3.1.4 can be seen from the 1.5 full periods in x-direction. The order in y-direction can not be seen from the picture and the shape of the E_y will be the same for all values of y, but the magnitude of this component will oscillate with the spatial frequency defined by k_y^{TM}.

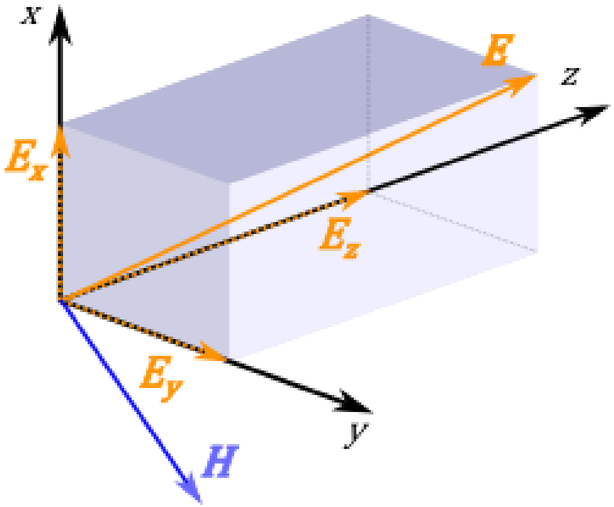


Figure 3.1.3: Components of the electric field inside a waveguide

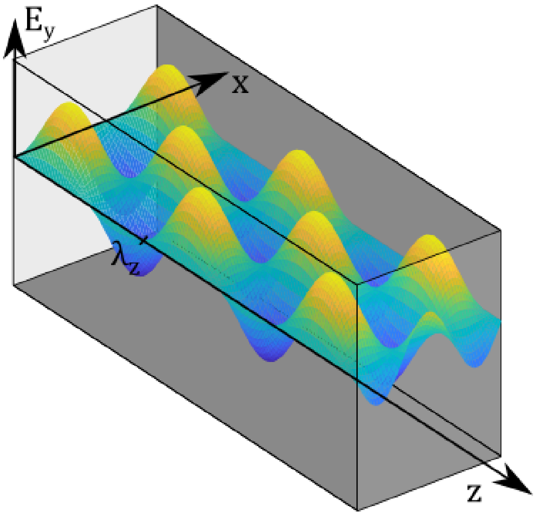


Figure 3.1.4: y component of the electric field inside a waveguide at a fixed value of y

3.1.2.2 Physical Interpretation of the Solution

To give a physical interpretation of this mathematical solution, some more mathematical conclusions that follow directly from the derivation before are given. Since the results are the same for E- and H-modes, the superscripts of k_x and k_y are removed from here on and first it is shown that the term $k_x^2+k_y^2$ will result in κ_c for the approach chosen above. This can be seen for example from the definition of k_x^2 in (3.1.24) and the solution of $Y(y)$ in (3.1.31). The second derivative of $Y(y)$ is given as

$$\frac{\partial^2 Y(y)}{\partial y^2} = -k_y^2 \sin(k_y y), \tag{3.1.52}$$

which can be used in the definition of k_x^2 and thus gives directly the already mentioned result

$$\kappa_c^2 = k_x^2 + k_y^2. \tag{3.1.53}$$

This result can also be obtained with all other combinations of the k_x^2 and k_y^2 definitions and solutions for $X(x)$ and $Y(y)$ above. This result can be used to better understand the wave propagation inside waveguides by inserting (3.1.53) in (3.1.18), which results in

$$k_z = \sqrt{\frac{\omega^2}{c^2} - \kappa_c^2} = \sqrt{\frac{\omega^2}{c^2} - k_x^2 - k_y^2} \tag{3.1.54}$$

as an expression for the propagation factor k_z. The term ω/c was defined in Section 3.1.1 as the wavenumber k_z for TEM-waves in free space. Since the symbol k_z will also be used here for the propagation constant in z-direction, the expression ω/c is redefined here as the *free space wavenumber* k_0

$$k_0 = \frac{\omega}{c}. \tag{3.1.55}$$

With these terms and the discrete values of k_x and k_y, k_z can be written as

$$k_z = \sqrt{k_0^2 - \left(\frac{m\pi}{a}\right)^2 - \left(\frac{n\pi}{b}\right)^2} \tag{3.1.56}$$

for a certain k_0. The expression in (3.1.56) gives the propagation factor k_z for a certain frequency ω and has different values for the possible values of k_x and k_y. The term k_z is also called *waveguide-wavenumber*. It is also visible that k_z can become imaginary if k_0 is smaller than κ_c, which means that the propagation term $e^{-jk_z z}$ becomes real and will result in an exponential damping of the wave along the z-direction[3]. Due to this connection, the value of κ_c gives a lower limit for the frequencies that can propagate without losses inside a waveguide and is thus called *cutoff-wavenumber*. Modes with a

---

[3] Mathematically, also an exponential increase over the z-direction would be possible, but this will result in infinite energy and is physically not relevant [Geo97]

cutoff-wavenumber higher than k_0 are called *evanescent modes*, which means that for every k_0 a limited number of propagating modes exist, but always an infinite number of evanescent modes. Analogous to the cutoff-wavenumber, a cutoff-wavelength λ_c can be defined as

$$\lambda_c = \frac{2\pi}{\kappa_c}. \tag{3.1.57}$$

To simplify notation in the upcoming part, a dimensionless factor Λ is defined here as

$$\Lambda := \sqrt{1 - \left(\frac{\kappa_c}{k_0}\right)^2} = \sqrt{1 - \left(\frac{\lambda_0}{\lambda_c}\right)^2}, \tag{3.1.58}$$

where λ_0 is the free space wavelength that belongs to k_0. Above the cutoff-wavenumber, the factor Λ is real-valued with a magnitude smaller than one. Below the cutoff, it becomes imaginary and thus k_z can also be written as

$$k_z = \sqrt{\frac{\omega^2}{c^2} - \kappa_c^2} = k_0 \Lambda. \tag{3.1.59}$$

With this expression, a formula for the wavelength λ_z in the direction of propagation can be given as

$$\lambda_z = \frac{2\pi}{k_z} = \frac{\lambda_0}{\Lambda}, \tag{3.1.60}$$

which means that inside the waveguide the wavelength in the direction of propagation differs from the one in free space for a given frequency. This indicates that the propagation speed in z-direction differs from the speed of light c_0 and also depends on the frequency since the factor Λ shows strong variations over frequency[4]. This difference in the propagation speed is the main reason for the *waveguide dispersion*, which means that different frequencies travel at different velocities and thus a bandpass signal gets spread in time and the shape of the envelope gets distorted. In general these effects are analyzed by the *group velocity*

$$v_G = \frac{\partial \omega}{\partial k_z} = \left(\frac{\partial k_z}{\partial \omega}\right)^{-1} = \left(\frac{\partial}{\partial \omega}\sqrt{\frac{\omega^2}{c^2} - \kappa_c^2}\right)^{-1} = c_0 \Lambda, \tag{3.1.61}$$

which gives the propagation speed for the signal contained in the wave [Str41][5]. From

---

[4] It is important to notice the dependency on the coordinate z here, because waves will also travel inside waveguides with the speed of light and only seemingly slower in the direction of propagation. This becomes clearer when it is shown that the propagation in z-direction is a sum of two TEM-waves that bounce between the walls of the waveguide. For more details on this topic the reader is referred to the literature, e.g. [Geo97].

[5] The group velocity does not give the "signal velocity" in general, only in the cases of no or normal dispersion, which means that the phase velocity is higher than the group velocity [Str41]. Since this is the case for waveguides, this equivalence can be assumed. Inside waveguides, the phase

(3.1.61) follows that information inside of waveguides propagates slower than the speed of light and the actual velocity depends on the frequency. An exemplary plot of the group velocity of a toy example, which will be used later again, is shown in Figure 3.1.5.

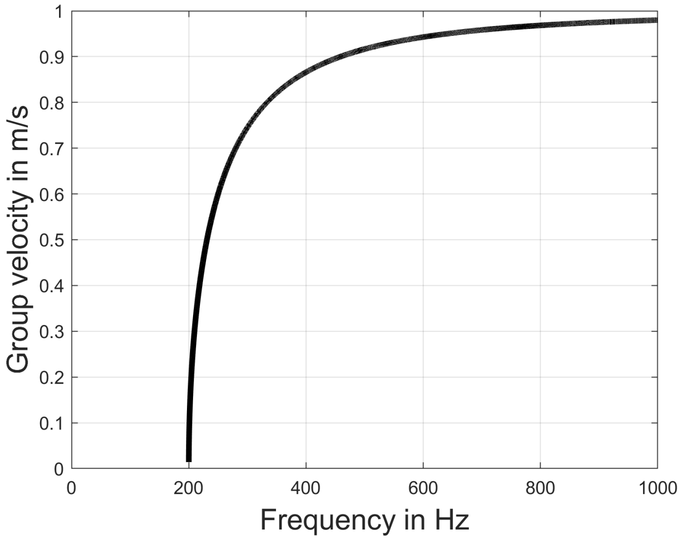


Figure 3.1.5: Group velocity for a mode with a cutoff-frequency of 200 Hz

In this example, the group velocity for a mode with a cutoff-frequency of 200 Hz is shown and it is visible that below the cutoff no graph is shown since a propagation is not possible in this region. Exactly at the cutoff-frequency the velocity is zero, which corresponds to a standing wave, and with increasing frequency, v_G approaches the free space velocity c_0 that is 1m/s in this example.

So far only one mode was considered during propagation, but from the solution of the Helmholtz equation can be seen that with a growing dimension of the waveguide more propagating modes are possible. If, for example, the frequency of the wave is higher than the cutoff-frequency of the first two modes, the energy will be split between these and the resulting wave is a sum of these two modes. For the electric field with two TM-modes this is written as

$$\mathbf{E}(z,t) = \mathbf{E}(0,0)\left(S_{TM}^{(1)} e^{j\left(\omega t - k_z^{(1)} z\right)} + S_{TM}^{(2)} e^{j\left(\omega t - k_z^{(2)} z\right)}\right). \tag{3.1.62}$$

The terms $S_{TM}^{(1)}$ and $S_{TM}^{(2)}$ determine the energy spread between the two modes and a formula to calculate these values will be given in Chapter 3.2. The waveguide-wavenumbers $k_z^{(1)}$ and $k_z^{(2)}$ characterize the propagation of the corresponding mode and a wave is in general the sum of all modes with wavenumbers $k_z^{(m,n)}$, which have

---

velocity v_{ph} and the group velocity v_G are related by $v_{ph} v_G = c_0^2$.

a cutoff frequency lower than the frequency of the propagating wave. In case of the electric field for example, this dependency is given by

$$\mathbf{E}(z,t) = \mathbf{E}(0,0) \sum_{m=1}^{M} \sum_{n=1}^{N(m)} S_{TM}^{(m,n)} e^{j\left(\omega t - k_z^{(m,n)} z\right)}, \tag{3.1.63}$$

where M and $N(m)$ are given by the maximum cutoff-wavenumber below the propagating free space wavenumber

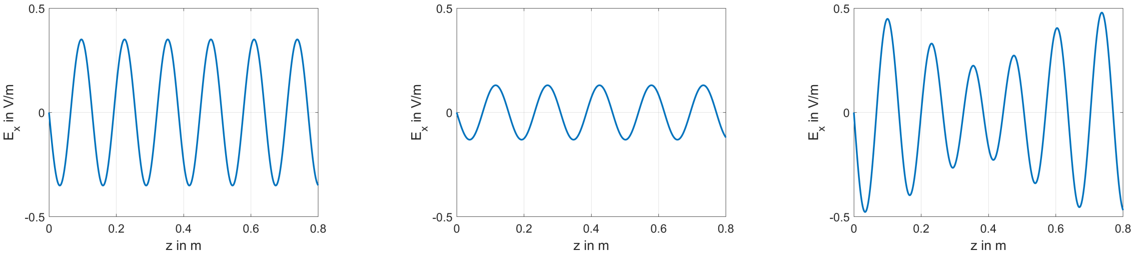
$$(M, N(m)) = \max\{(m,n) | \kappa_c(m,n) < \frac{\omega_{max}}{c}\}. \tag{3.1.64}$$

The selection of N depends on m due to the quadratic structure of κ_c in (3.1.53) and the definitions of k_x and k_y, i.e. a small value of m and thus k_x will allow a high value of n, which defines k_y. With high values of m, the situation is reversed, only small values of n, i.e k_y, are possible. This dependency of N on m is also valid for other geometries, as will be shown in Section 3.1.2.3 for cylindrical waveguides, although the definition of κ_c is different and the dependency is not quadratic anymore. For rectangular waveguides, the maximum order M and $N(m)$ are the same for E- and H-modes and do not have to be discriminated.

Since all modes in (3.1.63) possess a different group velocity, a bandpass signal with a high carrier frequency will show a strong dispersion while traveling through a waveguide with a rather large geometry. To illustrate this difference in the propagation speed of different modes, an example of a traveling wave is shown in Figure 3.1.6. In this example, a wave with a certain frequency ω_0 propagates as the sum of two distinct modes. The example shows the x-component E_x of the electric field $\mathbf{E}$ for fixed x_0 and y_0 values along the z-direction, Figures 3.1.6 (a) and (b) show E_x for the two modes individually and Figure 3.1.6 (c) shows the modal sum. The individual plots show the different wavelengths of the two modes along the z-direction. The first mode, the lower mode, ends at $z = 2$m in the seventh "valley" of the wave and the second, the higher mode, ends in the sixth valley at $z = 2$m. It is important to notice that these two plots show the same frequency in the time domain, since the different wavelength are compensated by the different propagation velocities.

From the modal sum of these two modes it becomes clear why the energy of the wave is spread out in the frequency domain when the wave is going from the inside of a waveguide back into free air, as is the case for the jet engine model in Section 3.2. The plot in Figure 3.1.6 (c) is the sum of the field values of the individual plots and thus the spatial wavenumber is also the sum of the two individual modes. Therefore the single frequency at the entrance of the waveguide is shifted to two new frequencies, which correspond to the wavelength of the propagating modes in z-direction, when this field form begins to propagate in free space.

In the next part, the solution for circular waveguides is shown shortly, since common jet engines mounted at airliners and small aircrafts have a circular dimension and therefore these waveguides will be used for the model in Section 3.2.



(a) Mode 1 (b) Mode 2 (c) Modal sum

Figure 3.1.6: Example of a traveling wave as the sum of two modes

3.1.2.3 Transfer to Cylindrical Waveguides

The deviation of the solution for circular waveguides is similar to the solution of rectangular waveguides, the difference is the coordinate system. Instead of cartesian coordinates, a cylindrical coordinate system is used and the starting point of the deviation is again the Helmholtz equation (3.1.14)

$$\nabla^2 \mathbf{E} + \frac{\omega^2}{c^2}\mathbf{E} = 0$$

together with the boundary condition 2 for the electric and magnetic field above. These boundary conditions must be fulfilled over the inner surface at $r = R = D/2$ and can be written as

$$E_\varphi = E_z = 0 \qquad \text{at} \qquad r = R = \frac{D}{2} \tag{3.1.65}$$

$$H_r = \frac{\partial H_\varphi}{\partial r} = \frac{\partial H_z}{\partial r} = 0 \qquad \text{at} \qquad r = R = \frac{D}{2}. \tag{3.1.66}$$

In the cylindrical coordinate system, the Laplace operator ∇^2 is changed to [BSMM08]

$$\nabla^2 = \underbrace{\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \varphi^2}}_{transversal} + \underbrace{\frac{\partial^2}{\partial z^2}}_{longitudinal}, \tag{3.1.67}$$

which allows to use again the approach of variable separation with the polar coordinates φ and r. The according functions are $\Phi(\varphi)$ and $R(r)$, which are used together with the notation $\frac{\partial}{\partial z} = -jk_z$ and the definition of κ_c from (3.1.18) to achieve the PDE

$$\left(\frac{\partial^2 R(r)}{\partial r^2} + \frac{1}{r}\frac{\partial R(r)}{\partial r}\right)\Phi(\varphi) + \left(\frac{1}{r^2}\frac{\partial^2 \Phi(\varphi)}{\partial \varphi^2}\right)R(r) + \kappa_c^2 \Phi(\varphi) R(r) = 0, \tag{3.1.68}$$

where E_z is replaced by $\Phi(\varphi) \cdot R(r)$. This expression is now multiplied with $\frac{r^2}{\Phi(\varphi)R(r)}$ and terms are collected according to the independent variable

$$\frac{r^2}{R(r)}\left(\frac{\partial^2 R(r)}{\partial r^2} + \frac{1}{r}\frac{\partial R(r)}{\partial r} + \kappa_c^2 R(r)\right) = -\frac{1}{\Phi(\varphi)}\frac{\partial^2 \Phi(\varphi)}{\partial \varphi^2}. \tag{3.1.69}$$

Since the left side of (3.1.69) is independent of φ it could be considered as constant for the solution of $\Phi(\varphi)$, therefore the term m^2 is used for the left side, what yields the oscillating ODE

$$\frac{\partial^2 \Phi(\varphi)}{\partial \varphi^2} + m^2 \Phi(\varphi) = 0 \tag{3.1.70}$$

similar to (3.1.26) and (3.1.27) for rectangular waveguides. The difference to the rectangular system is that both $\sin(m\varphi)$ and $\cos(m\varphi)$ are solutions to this ODE for $\Phi(\varphi)$ since there is no boundary condition, which excludes one of the two. Although there is the physical boundary condition that the values at $\varphi = 0$ and at all multiples of 2π must be the same, because they refer to the same point in the coordinate system and the value of E_z must be unambiguous. Therefore only integer values are allowed for m and in the following the solution $\sin(m\varphi)$ is ignored since it is the same as the $\cos(m\varphi)$ solution with the reference angle shifted by $\pi/2$.

To get the solution for $R(r)$, the right side of (3.1.69) is abbreviated with m^2 and in this way an ODE

$$\frac{\partial^2 R(r)}{\partial r^2} + \frac{1}{r}\frac{\partial R(r)}{\partial r} + \left(\kappa_c^2 - \frac{m^2}{r^2}\right) R(r) = 0, \tag{3.1.71}$$

which is independent of φ is created. The general solution of (3.1.71) is given by [Far93] as

$$R(r) = A \cdot J_m(\kappa_c r) + B \cdot Y_m(\kappa_c r), \tag{3.1.72}$$

with A and B as arbitrary magnitudes of the *m-th order Bessel function of the first kind* $J_m(\cdot)$ and the *m-th order Bessel function of the second kind* $Y_m(\cdot)$.

The Bessel function of the second kind can be ignored here, since they diverge for $r \to 0$ and this would lead again to a wave with infinite energy as before for the negative solution of k_z below the cutoff-frequency. Because of that, the Bessel function of the first kind is the only relevant term in (3.1.72). Since m must be an integer value according to the solution of (3.1.70), the zeros of the m-th order Bessel function define the possible modes inside circular waveguides. By setting $A = E_0$ and combining the solutions of $\Phi(\varphi)$ and $R(r)$, the final form of E_z is achieved

$$E_z(r, \varphi) = \hat{E} J_m(\kappa_c r)\cos(m\varphi)e^{-jk_z z}. \tag{3.1.73}$$

The other field components can be calculated from the z-component, similar to Equations (3.1.38) to (3.1.41), which also requires the z-component of the magnetic field that is determined below. First, the cutoff-frequencies for the E-modes are derived from the boundary condition that E_z must be zero at $r = R$, which implies that $J_m(\kappa_c R)$ must be zero and κ_c^{TM} is given as the quotient of P_{mn}, the n-th zero of the m-th Bessel function, and the radius R of the waveguide

$$\kappa_c^{TM} = \frac{P_{mn}}{R}. \tag{3.1.74}$$

The first three Bessel functions of the first kind from order zero to two are shown in Figure 3.1.7 and the zeros of these functions can be calculated numerically and in this way the cutoff-frequencies of the E-modes can be determined. The first zero of $J_m(x)$ for $m \geq 1$ at $x = 0$ is not considered for this calculation.

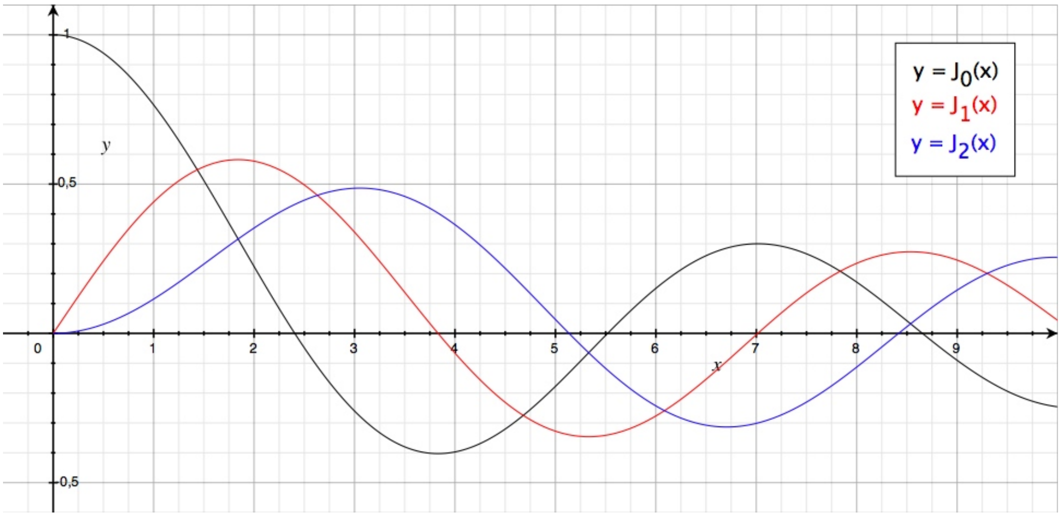


Figure 3.1.7: First three Bessel functions

For H-modes the solution of the Helmholtz equation is principally the same as before and the solution can be given as

$$H_z(r, \varphi) = \hat{H} J_m(\kappa_c r) \cos(m\varphi) e^{-jk_z z}. \tag{3.1.75}$$

The difference to the E-modes can be found in the boundary conditions (3.1.66) that require the derivative of H_z with respect to r to be zero. Since the only dependency on r in (3.1.75) can be found in the term $J_m(\kappa_c r)$, the derivation of this term determines the cutoff-frequencies for H-modes in cylindrical waveguides. The actual calculation is done in the same way as above for the E-modes by dividing the zeros P'_{mn} of the Bessel function derivative by the radius R of the waveguide

$$\kappa_c^{TE} = \frac{P'_{mn}}{R}. \tag{3.1.76}$$

The derivatives of the first three Bessel functions are shown in Figure 3.1.8.

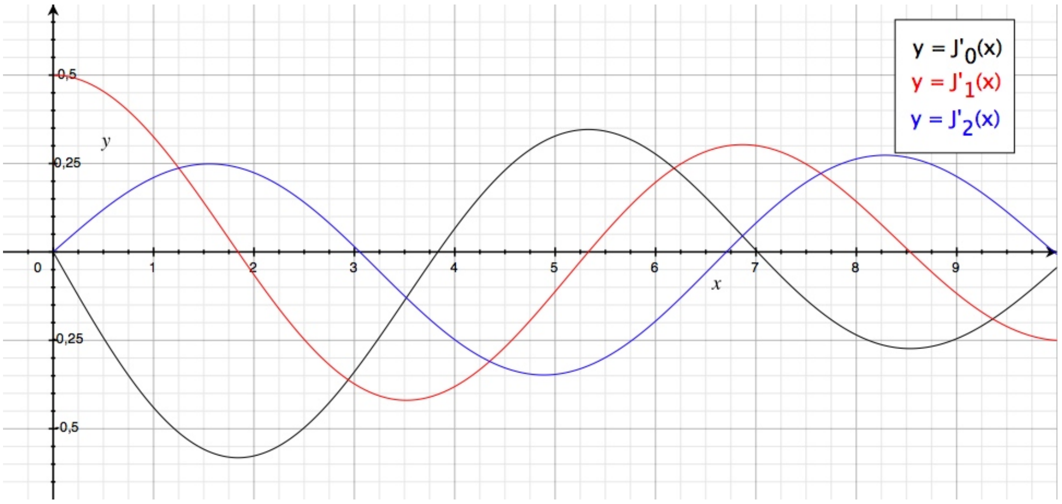


Figure 3.1.8: Derivatives of the first three Bessel functions

Before the final jet engine model is derived in Section 3.2, a toy example summarizing the results of this section and introducing the signal-theoretic model that will be used later is given in the following part.

3.1.3 Example of Chirp Propagation inside a Waveguide

In Sections 3.1.1 and 3.1.2 all fields are given in the time domain, but since it is known from signal processing that processing in the frequency domain is often faster and easier, the Fourier transform $\tilde{\mathbf{E}}$ of the initial waveform $\mathbf{E}(z=0,t)$ is given as

$$\tilde{\mathbf{E}}(0,\omega) = \int_{-\infty}^{\infty} \mathbf{E}(0,t)e^{-j\omega t}dt. \tag{3.1.77}$$

The initial waveform $\mathbf{E}(z=0,t)$ can be of arbitrary form, e.g. a chirp or a continuous wave with a single frequency. If now a propagation factor $e^{-jk_z z}$ is added to the waveform, the Fourier transform at position z is calculated as

$$\tilde{\mathbf{E}}(z,\omega) = \int_{-\infty}^{\infty} \mathbf{E}(0,t)e^{-j(\omega t+k_z z)}dt. \tag{3.1.78}$$

Since the term $e^{-jk_z z}$ in (3.1.78) does not depend on t it can be written in front of the integral, what results in

$$\tilde{\mathbf{E}}(z,\omega) = e^{-jk_z z}\int_{-\infty}^{\infty} \mathbf{E}(0,t)e^{-j\omega t}dt = e^{-jk_z z}\tilde{\mathbf{E}}(0,\omega). \tag{3.1.79}$$

The expression in (3.1.79) can be considered as the output of a linear system with the transfer function $e^{-jk_z z}$ [Orf16]

$$H(z,\omega) = e^{-jk_z(\omega)z}, \tag{3.1.80}$$

which allows to write the waveform after propagation to point z similar to linear time invariant systems known from signal processing in the frequency domain as

$$\tilde{\mathbf{E}}(z,\omega) = H(z,\omega)\tilde{\mathbf{E}}(0,\omega). \tag{3.1.81}$$

The expression for $k_z(\omega)$ is given in Section 3.1.2.2 for waveguides as $\sqrt{\frac{\omega^2}{c^2} - \kappa_c(m,n)^2}$ for mode (m,n). In the example at the end of Section 3.1.2.2 was already shown that a the electric field is traveling inside a waveguide as a sum of the propagating modes and, therefore, the general transfer $H(z,\omega)$ function is a sum of E- and H-modes given by

$$H(z,\omega) = \underbrace{\sum_{m=0}^{M_{TE}} \sum_{n=0}^{N_{TE}(m)} S_{TE}^{(m,n)} e^{-jk_z^{(m,n)}z}}_{TE-modes} + \underbrace{\sum_{m=0}^{M_{TM}} \sum_{n=1}^{N_{TM}(m)} S_{TM}^{(m,n)} e^{-jk_z^{(m,n)}z}}_{TM-modes}. \tag{3.1.82}$$

The upper bounds of the order of the modes $M_{TE}, N_{TE}(m), M_{TM}$ and $N_{TM}(m)$ are the same for E- and H-modes in case of a rectangular waveguide, but differ in case of a circular one. They are given in both cases by the mode with the highest cutoff-frequency below the maximum frequency ω_{max} by (3.1.64).

With this transfer function, the output of a waveguide can be calculated for an arbitrary waveform at the input. From a signal processing perspective, it is not useful to have the expression separated for the electric and magnetic field components of the traveling wave. Thus from now on, a signal $s(t,z)$ is used instead of the field components. This substitution is possible, since the components of the electric and magnetic field follow the signal and thus the signal propagates as the sum of all possible TE- and TM-modes. This sum of TE- and TM-modes is also the reason that the summing terms in (3.1.82) start at zero and not at one, since for TM modes in rectangular waveguides either m or n can be zero. For $m = n = 0$, the amplitude $S^{(0,0)}$ is set to zero because it defines a mode that is physically not possible. The dependency of N on m was already explained in the context of (3.1.64).

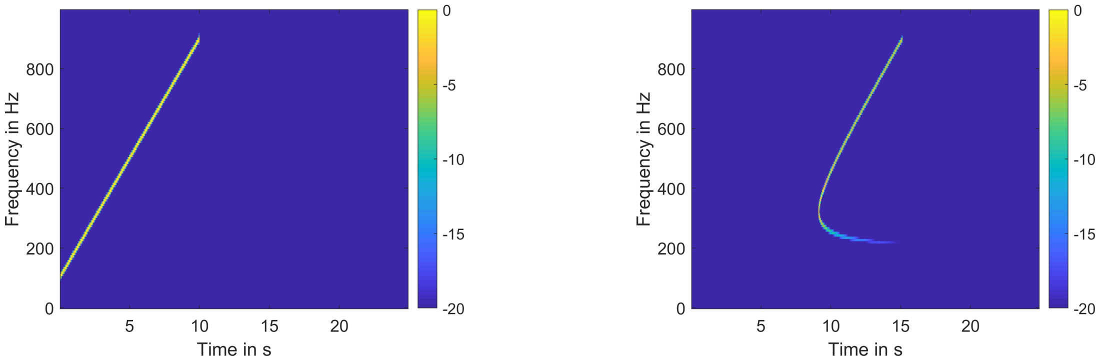
To simplify the notation, a single index m is used from here on as index for the different modes, which is running over all E- and H-modes. Using this notation, the transfer function is written as

$$H(z,\omega) = \sum_{m=1}^{M} S^{(m)} e^{-jk_z^{(m)}z}. \tag{3.1.83}$$

The calculation of the amplitudes of the different modes $S^{(m)}$ is not considered here, a way to calculate these values is given in Section 3.2.

To show the effects of waveguide propagation, a toy example with a chirp waveform is shown below. The chirp has a bandwidth of 800 Hz and runs from 100 Hz to 900 Hz, the spectrogram of this chirp is shown in Figure 3.1.9 (a). It can be seen that the duration of the chirp is 10 seconds. The free space velocity is 1 m/s and the waveguide has a length of 5 m. These values are picked to create a vivid, rather than a realistic example. A realistic example will be given in Section 3.3.

In the first part of this example a single propagating mode is assumed with a cutoff-frequency of 200 Hz. The group velocity for this mode was already depicted in Figure 3.1.5. The resulting spectrogram at the output of the waveguide in Figure 3.1.9 (b), which shows that frequencies below the cutoff are removed from the waveform. The frequencies close, but above the cutoff, are extremely slowed down that they get passed by the higher frequencies inside the waveguide.

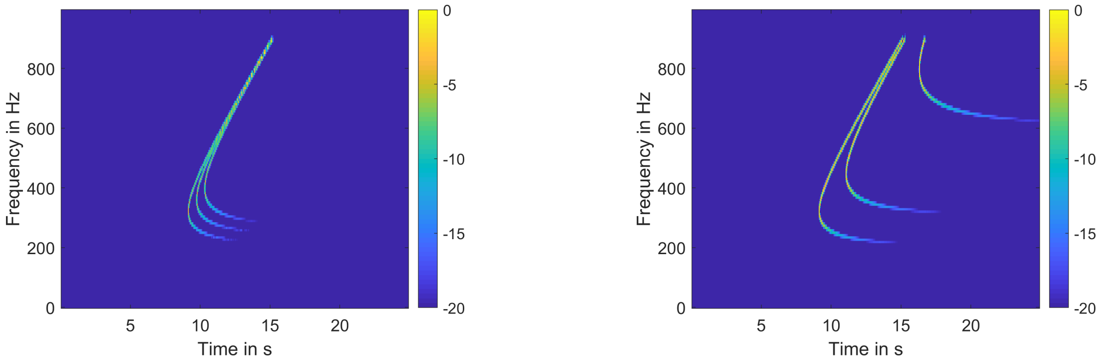


(a) Spectrogram input signal (b) Spectrogram output signal

Figure 3.1.9: Time-frequency representation of a chirp before and after propagation through a hypothetic waveguide with a single propagating mode

In the second part of this example, the waveform propagates as the sum of three modes with different cutoff-frequencies, but with the same amplitudes $S^{(m)}$. In Figure 3.1.10 (a) a hypothetic waveguide with three possible modes close to each other is assumed. In this plot, the low frequencies close the cutoff-frequencies show a strong dispersion, while the straight line of the chirp is still visible at the higher frequencies. In the spectrogram in Figure 3.1.10 (b), the cutoff-frequencies are more apart from each other. This results in the appearance of three chirps at the output of the waveguide, since the group velocity at the high frequencies differs also quite a lot. In fact, the example in Figure 3.1.10 (a) also shows three chirps, but the difference at the high frequencies is that small that they cannot be distinguished by the eye.

In this example and also in the later applications a simplification regarding the velocity of propagation is done. In reality, there will always be a small part of the energy of the wave travel with the speed of light, which is called the *wavefront*. The velocity



(a) Spectrogram output signal with $f_{c1} = 200\,\mathrm{Hz}$, $f_{c2} = 230\,\mathrm{Hz}$ and $f_{c3} = 260\,\mathrm{Hz}$

(b) Spectrogram output signal with $f_{c1} = 200\,\mathrm{Hz}$, $f_{c2} = 300\,\mathrm{Hz}$ and $f_{c3} = 600\,\mathrm{Hz}$

Figure 3.1.10: Time-frequency representation of a chirp before and after propagation through a hypothetic waveguide with three propagating modes

of this wave front is limited by the speed of light and is called *wave front velocity*. This means that after $t = \frac{z}{c_0}$ already a part of the waveform appears at the output of the waveguide, followed actually by other so called *precursors* or *forerunners* [Str41]. These forerunners are ignored in this work, since the amplitude of these forerunners is relatively low compared the remaining signal.

For a waveform traveling with light speed, as in the radar case, already very small differences in the group velocity will cause visible differences at the output of the radar. To show this effect, an example with a realistic waveform and a realistic target will be shown in Section 3.3. But before this example is given, the final jet engine scattering model will be shown in the next section.

3.2 Jet Engine Scattering Model

Modeling jet engines as waveguides and using modal decompositions of the incident wave is a common way to determine the RCS of engines and aircrafts [Ana03]. Some applications that use this kind of model in the imaging domain can be found in the literature, for example in [Bor97], where it was used to remove the range extended returns from an ISAR image and increase the matching score of a template matching target recognition system. Another example can be found in [TL96], which uses a time-frequency transformation based ISAR imaging algorithm to produce several ISAR images that differ in the used frequency. Furthermore, parts that are unstable across the chosen frequency range are removed from the image. In this way, an image that only consists of stable point scatterers is created and it was also shown that the parts that were removed from the image show the highest variation in frequencies that are very near to the theoretical modes of the inlet structure. The approach in this work is different, since the range extended return is not considered as an artifact that should be removed, but as a feature that should be used for classification.

The basic idea is to model the engine structure as a terminated waveguide and thus allow a modal decomposition of the waveform inside the engine. Using this approach, the modal sum is calculated only once for a given jet engine and can be stored for the classification process. Further, an analytical model is available and therefore, the computational demands are reduced. The termination of the waveguide is realized by the fan or the first stage of the compressor, dependent on the engine. The modulation due to the rotation of the blades, which was introduced in Section 1.2.1 as JEM, is not considered here. Although JEM has an effect on ISAR images, the so called JEM lines, it is not used in the model, since these JEM lines appear perpendicular to the range dimension. The model used here is limited to the range dimension and the use of the modulation effect in the model might be considered in the future as extension of this work.

To keep the notation that is used for point scatterer, the normal form

$$s(k_r, R) = e^{-jk_r R} \tag{3.2.1}$$

must be extended by the transfer function (3.1.83) of a waveguide with length L

$$H(L, k_{wg}) = \sum_{m=1}^{M} S^{(m)} e^{-j2k_{wg}^{(m)} L}, \tag{3.2.2}$$

where the term k_z is replaced by k_{wg}, since the direction of propagation is not z anymore. Instead of z, the length L of the waveguide is used as the propagation distance of the waveguide, the factor 2 stands for the two way propagation inside the engine. Combining the terms above gives the normal form of waveguide scatterers

$$\begin{aligned} s(k_r, k_{wg}, R, L) &= \sum_{m=1}^{M} S^{(m)} e^{-j(k_r R + 2k_{wg}^{(m)} L)} \\ &= e^{-jk_r R} \sum_{m=1}^{M} S^{(m)} e^{-j2k_{wg}^{(m)} L}. \end{aligned} \tag{3.2.3}$$

The first term in the exponent of (3.2.3) represents the distance of the radar to the entrance of the engine, where the opening structure also gives a strong point-like scattering effect. The different scattering effects of an engine are depicted in Figure 3.2.1, which shows three different types of scattering for an incident electric field $\mathbf{E}^{in}$ [MW93]. The first one is the point-like scattering at the opening structure $\mathbf{E}^{s}_{rim}$, the second one is the waveguide scattering from the fan $\mathbf{E}^{s}_{wg}$ and the third one is the point-like scattering from the exterior of the engine $\mathbf{E}^{s}_{ext}$.

The first two of the scattering mechanisms above are relevant for the model used here,

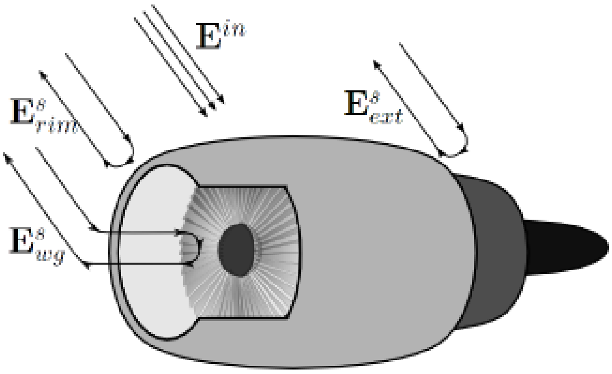
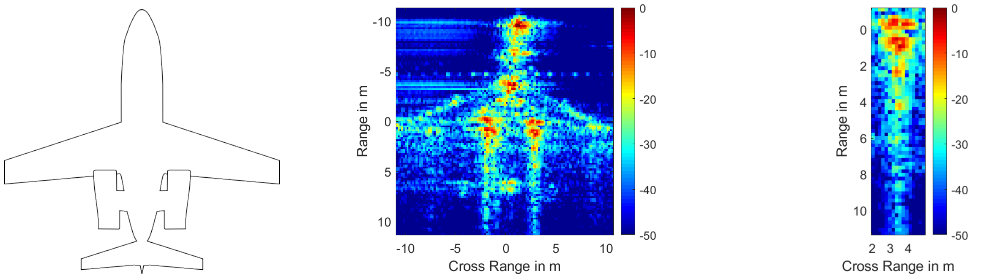


Figure 3.2.1: Scattering mechanisms of a jet engine

the scattering from the exterior is ignored, since this scattering can be very diffuse and might not contribute in the backscattering direction. The scattering of the outer rim and the cavity should always be present for this kind of engine. An example of a real ISAR image with a zoomed plot of the jet engine is shown in Figure 3.2.2. The data was measured with the TIRA system of Fraunhofer FHR, the bandwidth is 800 MHz and the processing interval was selected to achieve a corresponding cross-range resolution.



(a) Shape of target (b) ISAR image (c) Magnification of right engine

Figure 3.2.2: Example of a real ISAR image with range extended returns

In the zoomed image in Figure 3.2.2 two strong point-like scattering centers can be seen close to each other in the down-range direction. The first one corresponds to the opening of the jet engine, the second one is the result of the fast propagating modes that generate an echo from the fan. The higher modes with a slower group velocity produce the ponytail behind the physical position of the engine. Since the matched filter in the processing chain is fitted to a wave that is propagating with the speed of light, the slow group velocity of the higher modes inside a waveguide has the same effect as a longer distance of a wave traveling with the speed of light.

A point that was not yet considered is the calculation of the weighting parameters

$S^{(m,n)}$ in the model and one method to calculate these is given in the next part.

3.2.1 Calculation of the Mode Distribution

To calculate the amplitudes of the different modes, the modes that are possible for a given frequency and a certain waveguide opening must be calculated. This is done by using (3.1.74) and (3.1.76) and for the examples we use here openings of 60 and 90 cm are assumed for the engine and as frequency 10 GHz (X-Band) and 16.7 GHz (Ku-Band) are chosen. The possible modes are given graphically in Figure 3.2.3. The yellow areas in the plots represent the possible modes, while the blue areas possess a cutoff-frequency higher than the propagating frequency and thus represent evanescent modes, which are ignored in this model. The axes of the plots correspond to m and n in (3.1.74) and (3.1.76) and it can be seen that the number of possible modes grows with frequency and dimension of the waveguide.

To calculate an approximation of the amplitudes of the different modes for a cylindrical waveguide, the result of [Hua83] is presented here. For a detailed derivation of the following formulas, the reader is referred to the original paper and the references therein. The main parameters for the calculation are the free space wavenumber k_0 of the transmitted signal, the radius R of the engine and the angle Θ between the main axis of the engine and the incident wave. This means that an illumination from exactly nose on results in an angle of 0° and from broadside in 90°. The given results should hold within the region $0° < \Theta < 80°$.

The amplitudes are calculated differently for TE- and TM-modes, for TE-modes by

$$S^{(m,n)} = j \cdot \left(j^n N_{(m,n)} n \left(1 + \frac{k_{wg}^{(m,n)}}{k_0} \cos(\Theta) \right) \frac{J_n \left(k_0 R \sin(\Theta) \right)}{\sin(\Theta)} \right)^2 \tag{3.2.4}$$

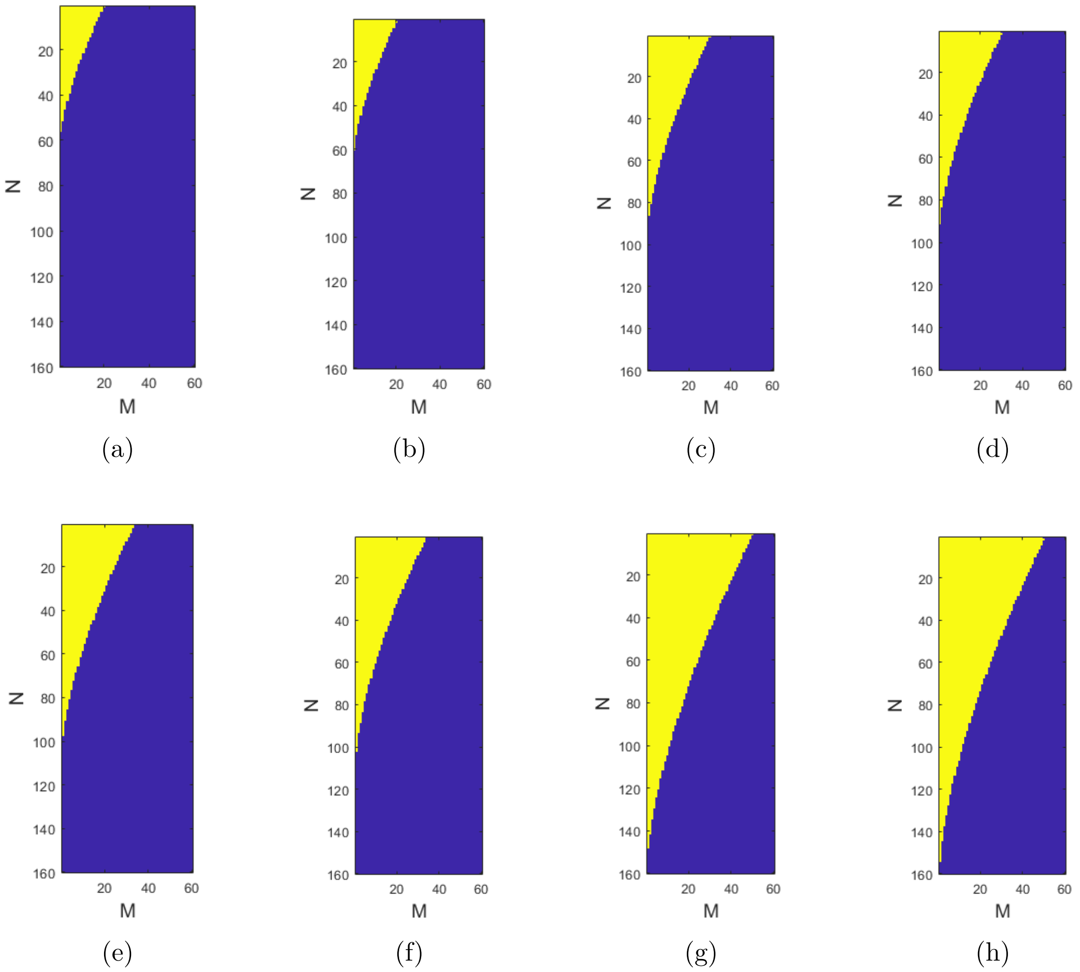
and for TM-modes by

$$S^{(m,n)} = j \cdot \left(j^n N'_{(m,n)} \frac{1}{\sin(\Theta)} \left(\frac{k_{wg}^{(m,n)}}{k_0} + \cos(\Theta) \right) \frac{J_n(k_0 R \sin(\Theta))}{1 - \left(\frac{P_{mn}}{k_0 R \sin(\Theta)} \right)^2} \right)^2 . \tag{3.2.5}$$

J_n and J'_n denote the the n-th order Bessel function of the first kind and its derivative with the zeros P_{mn} and P'_{mn} as given at the end of Section 3.1.2.3. The normalization terms $N_{(m,n)}$ and $N'_{(m,n)}$ are given by

$$N_{(m,n)} = \frac{1}{\sqrt{k_{wg}^{(m,n)} \epsilon_{0n} (P'^2_{mn} - n^2)}} \tag{3.2.6}$$

and



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 3.2.3: Possible E- and H-modes for different examples; (a) E-modes for 10 GHz and 60 cm opening; (b) H-modes for 10 GHz and 60 cm opening; (c) E-modes for 10 GHz and 90 cm opening; (d) H-modes for 10 GHz and 90 cm opening; (e) E-modes for 16.7 GHz and 60 cm opening; (f) H-modes for 16.7 GHz and 60 cm opening; (g) E-modes for 16.7 GHz and 90 cm opening; (h) H-modes for 16.7 GHz and 90 cm opening

$$N'_{(m,n)} = \frac{1}{\sqrt{k_{wg}^{(m,n)} \epsilon_{0n}}}, \tag{3.2.7}$$

with

$$\epsilon_{0n} = \begin{cases} 2 & : \quad n = 0 \\ 1 & : \quad n \neq 0 \end{cases} . \tag{3.2.8}$$

Using these equations, the amplitudes of the different modes can be calculated. An example for this mode distribution is depicted in Figure 3.2.4, where the values for the frequency and opening geometry are picked from the example before as $f = 16.7\,\text{GHz}$ and an opening diameter of 90 cm.

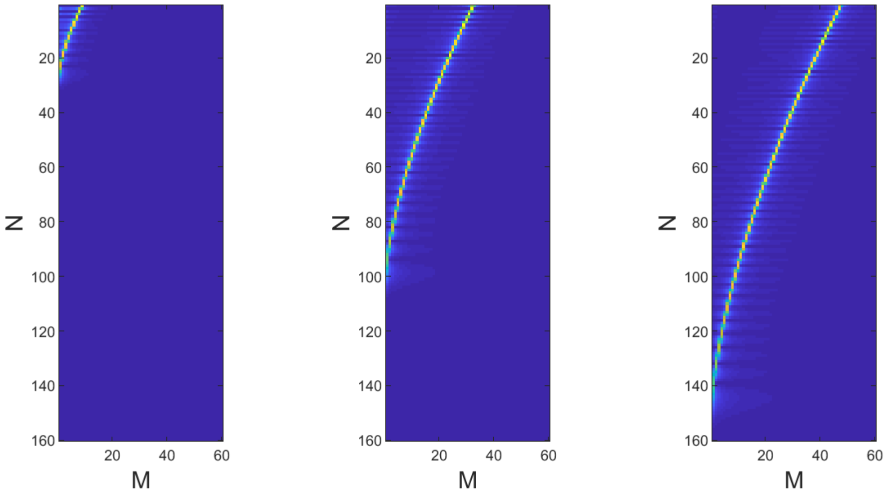
From the plots in Figure 3.2.4 can be seen that a higher aspect angle will lead to higher modes, which can be explained by the method of *mode matching* [Wex67]. This principle can be explained graphically by a wave, which travels along the opening of a waveguide. If the aspect angle is high, which means that the wave travels almost perpendicular to the opening of the waveguide, the maximum number of minima and maxima of the electric field for example can be found along the opening geometry. This number of minima and maxima matches the field structure of a certain mode that will therefore start to propagate.

The results presented above can principally be used to determine the coefficients of the jet engine scattering model, but this method has a major drawback, since it is based on the assumption of a perfect waveguide structure, which is not true for jet engines. It is known from waveguide theory that any discontinuity along the inner surface of a waveguide will excite all modes, also the evanescent ones [Wex67, Coo71]. From this point of view, it is very likely that inside a jet engine the propagating modes are equally distributed, since a large number of discontinuities can be found inside the air intake of engines, e.g. small things like screws or large ones like the air bypass. Therefore, the method to calculate the coefficients $S^{(m,n)}$, which was presented for completeness, will not be used in the further presented work. Instead an equal distribution of the propagating modes is assumed and evanescent modes are ignored. To validate the model and the given assumptions, a comparison of the example in Figure 3.2.2 and the model produced with the TIRA parameters is given in the next section.

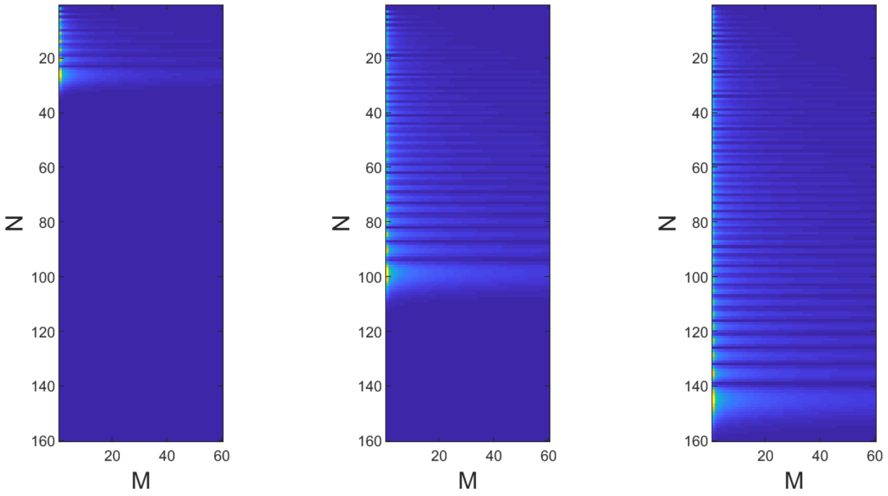
3.3 Scattering of a Chirp Waveform

An example of an ISAR image that was measured with the TIRA system can be seen in Figure 3.2.2 in the previous section. In this section, a comparison between this example and the output of the model with TIRA parameters is given. The parameter of the TIRA system are repeated here, as they are given in [Meh96]:

- Bandwidth $B = 800\,\text{MHz}$



(a) $\Theta = 10°$ (b) $\Theta = 40°$ (c) $\Theta = 70°$



(d) $\Theta = 10°$ (e) $\Theta = 40°$ (f) $\Theta = 70°$

Figure 3.2.4: E- (upper row) and H-mode (lower row) distribution for 16.7 GHz and 90 cm opening with different aspect angles

- Pulse duration $\tau_p = 256\,\mu\mathrm{s}$
- Carrier frequency $f_0 = 16.7\,\mathrm{GHz}$
- Sampling frequency $f_s = 4\,\mathrm{MHz}$
- 1024 samples per range profile

These numbers give an unambiguous range of 192 m and a range resolution of 18.75 cm. The parameter of the opening geometry, i.e. the diameter of the jet engine, is the same as in the example in Figure 3.2.4, i.e. 90 cm opening. Since no ground truth for this target is available, a waveguide length of 40 cm is used here, which gives a total propagation distance of 80 cm.

To simulate the scattering from a jet engine, an additional point scatterer is added, which represents the scattering from the opening rim. The signal model for this point scatterer is the normal form (2.1.37), which is combined with the waveguide model (3.2.3). To balance the energy between these two parts, a normalization constant ζ that guarantees the same peak value in both components is used. The final one-dimensional model is formulated as

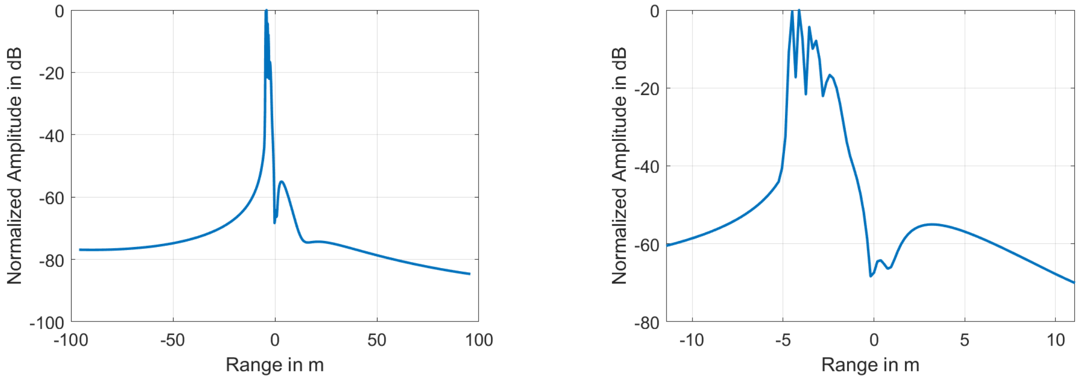
$$s(k_r, R, R_{wg}, L) = e^{-jk_rR}\left(\frac{1}{\zeta} + \sum_{m=1}^{M} e^{-jk_r\Lambda^{(m)}L}\right), \tag{3.3.1}$$

where $\Lambda^{(m)}$ is the dimensionless factor from (3.1.58), which also contains the implicit dependency on the waveguide radius R_{wg}. L is the length of the waveguide and R is the distance from the radar to the opening of the engine. The contribution of the M propagating modes in this example are, as it was already mentioned at the end of Section 3.2, equally distributed. The parameter ζ is chosen to give the same peak value at the opening of the engine and at the first reflection, as it was mentioned above.

A visualization of this model and thus the PSF for this kind of scattering is shown in Figure 3.3.1. The plot on the left side shows the complete unambiguous range and on the right a zoom in on the engine position shows the decay of the amplitude over range. The maximum amplitude is normalized to 0 dB in both plots and is reached at two points, which mark the position of the point scatterer and the first return of the waveguide scattering.

The values of the plot in Figure 3.3.1 (b) show that roughly 2.5 m behind the physical entrance, the amplitude is 20 dB below the maximum. The entrance of the engine is placed at −4.5 m in range. Since most of the propagating modes are close to the velocity of light, the amplitudes close to the entrance of the engine are the highest ones along the PSF. The slower propagating modes result in the smaller amplitudes behind the physical position of the engine. Approximately 5 m behind the entrance, a cancellation effect appears for this constellation of parameters.

For the two-dimensional PSF, the model in (2.2.8) is extended by the waveguide propagation transfer function (3.2.3). This results in the duct delayed returns in range



(a) Full range axis (b) Zoomed to relevant area

Figure 3.3.1: Calculated PSF for jet engine scattering

and a sinc-function in the cross-range domain, which is defined by the movement of the target (see Section 2.2). Additionally to the waveguide scattering, a point scatterer at the position of the entrance of the waveguide is used in the model. This corresponds to the rim reflection shown in Figure 3.2.1 and was also done in (3.3.1). To formulate the model, the point scatterer model

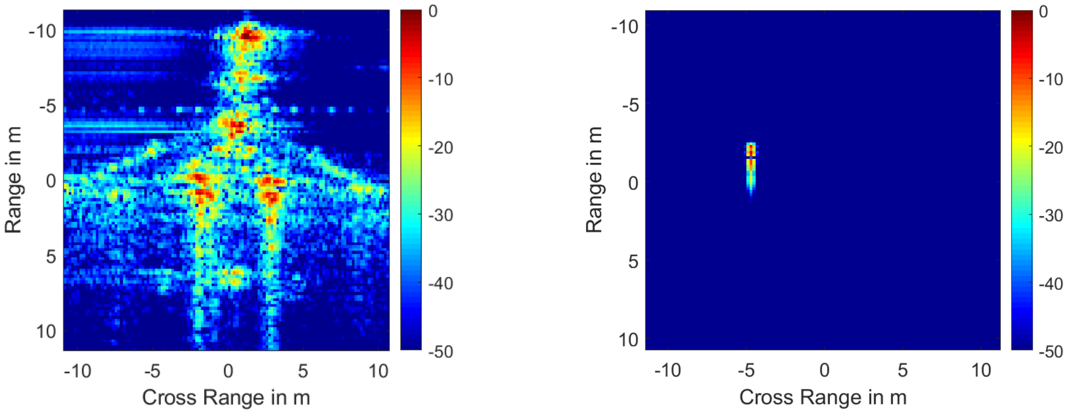
$$s(k_x, k_y, x, y) = e^{-j(k_x x + k_y y)} \tag{3.3.2}$$

from Section 2.2.1 is used as a starting point. This model is then extended with the waveguide model (3.2.3) in x-direction, which corresponds to the LOS. The waveguide model at position (x, y) is thus given as

$$\begin{aligned} s(k_x, k_y, x, y, R_{wg}, L) &= \sum_{m=1}^{M} e^{-j(k_x x + k_y y + 2k_{wg}^{(m)} cos(\varphi) L)} \\ &= \sum_{m=1}^{M} e^{-j(k_x x + k_y y + 2k_r cos(\varphi) \Lambda^{(m)} L)} \\ &= \sum_{m=1}^{M} e^{-j(k_x (x + \Lambda^{(m)} L) + k_y y)} \\ &= e^{-j(k_x x + k_y y)} \sum_{m=1}^{M} e^{-j k_x \Lambda^{(m)} L}. \end{aligned} \tag{3.3.3}$$

This shows that the waveguide echo is similar to the standard point scatterer model in (3.3.2) corrected by the sum that contains all propagating modes along the waveguides. For the final model that will be used later, a point scatterer at the position of the inlet is added to the waveguide propagation, which is modified in magnitude by the normalization constant ζ similar to the one-dimensional case above

$$s(k_x, k_y, x, y, R_{wg}, L) = e^{-j(k_x x + k_y y)} \left(\frac{1}{\zeta} + \sum_{m=1}^{M} e^{-jk_x \Lambda^{(m)} L} \right). \tag{3.3.4}$$



(a) Real ISAR image (b) Jet engine PSF

Figure 3.3.2: Comparison of an ISAR image and a 2D-PSF of a jet engine

Although the ponytail effect seems to be shorter in the calculated jet engine PSF, the similarity between it and the real ISAR image is clearly visible in Figure 3.3.2. With this model, the next step for an automatic recognition of this kind of scattering is an algorithm that is able to distinguish these scatterers from ordinary point scatterers. The theoretical background for this algorithm will be presented in the next chapter and the application to radar data will be shown in Chapter 5.

Chapter 4

Sparse Representation

Over the last years, the topic of *sparse representation* (SR) became very popular since many applications showed that this technique can be used in wide variety of areas. Together with these applications, a lot of attention was given to the theoretical foundations and therefore, a well understood framework is now available. In the first part of this chapter the basics of SR will be described as far as they are necessary for the decomposition framework used later in this chapter. As a first step, the basic terms and notations will be given in Section 4.1. In Section 4.2 an overview of the available algorithms is given and in Section 4.3 the algorithm that is chosen for the applications in the upcoming chapter is described in detail.

4.1 Terminology and Problem Description

The aim of the SR framework is to represent an input signal by a linear combination of elementary signals. This input signal is limited here to a finite dimensional vector $\boldsymbol{x} \in \mathbb{C}^N$, representing a complex valued sampled signal. The elementary signals that are used to approximate the signal are collected as columns of a Matrix $\boldsymbol{\Phi} \in \mathbb{C}^{N \times M}$, which is called *dictionary*. The weighting of these elementary signals are represented by the coefficients of the vector $\boldsymbol{\alpha} \in \mathbb{C}^M$

$$\boldsymbol{x} = \boldsymbol{\Phi}\boldsymbol{\alpha}. \tag{4.1.1}$$

The single columns $\boldsymbol{\varphi}_i$ of the dictionary are called *atoms* and to represent any arbitrary signal, these atoms must span the $\mathbb{C}^N$, i.e. the dictionary $\boldsymbol{\Phi}$ must be a full rank matrix. If the number of atoms M is the same as the dimension N, the dictionary is a basis of the signal space. If $M > N$, the matrix $\boldsymbol{\Phi}$ is called an *overcomplete dictionary*. These overcomplete dictionaries introduce the problem of linear dependency in (4.1.1), but hopefully allow a representation of the input signal with less elements.

If the input signal $\boldsymbol{x}$ can be exactly reproduced with $s \ll N$ atoms of the dictionary, the coefficient vector $\boldsymbol{\alpha}$ of this representation is called *s-sparse*. In most practical applications, this ideal case of an exactly s-sparse coefficient vector is not possible due to noise and model imperfections and therefore, an error η between the input signal and the sparse approximation is allowed, i.e.

$$||\boldsymbol{x} - \boldsymbol{\Phi\alpha}||_2 \leq \eta. \tag{4.1.2}$$

With this error term, the expression $\boldsymbol{\Phi\alpha}$ is strictly speaking not a sparse representation anymore, since the signal is not exactly represented, it is only approximated. The approximation with the minimum η for a given s, i.e. the number of used atoms, is called *best s-term approximation*. For this approximation problem, the expression compressible representation is introduced and the vector $\boldsymbol{x} \in \mathbb{C}^N$ is called compressible if its best s-term approximation decays quickly in s [FR13]. Although the expression sparse representation is strictly speaking not correct for the mentioned best s-term approximation that is calculated in most cases, it is mostly used synonymously in the literature and also in this thesis the terms SR and best s-term approximation are used equivalently.

To formulate the mathematical optimization problem that must be solved to find the best s-term approximation, a common mathematical tool, namely the ℓ_p norm $||\cdot||_p$, is introduced here. In (4.1.2) the ℓ_2 norm $||\cdot||_2$, which is also known as the Euclidean norm, was already used. This norm is a special case of the more general ℓ_p *norm*, which is defined for $p \in [1, \infty]$ as [DDEK12]

$$||\boldsymbol{\alpha}||_p := \begin{cases} \left(\sum_{n=1}^{N} |\alpha_n|^p\right)^{\frac{1}{p}} & \text{for } p \in [1, \infty) \\ \max\limits_{n=1,2,\ldots,N} |\alpha_n| & \text{for } p = \infty \end{cases} \tag{4.1.3}$$

with the Euclidean norm as a special case for $p = 2$. The ℓ_p norm can also be extended to the case $0 < p < 1$, which leads to a violation of the triangle inequality and thus to a quasinorm. For the SR framework, the p-th power of the ℓ_p norm for the case $p \to 0$ is very important, since it can be shown that this case gives the number of nonzero elements of the vector $\boldsymbol{\alpha}$.

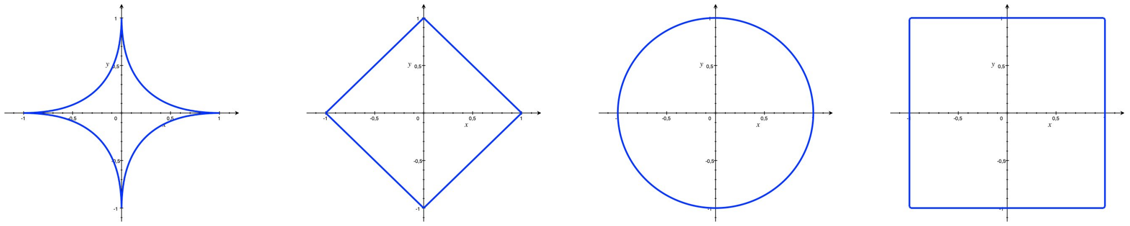
The index of the nonzero elements in $\boldsymbol{\alpha}$ is called *support* of $\boldsymbol{\alpha}$ and is written as supp($\boldsymbol{\alpha}$). As already mentioned, it can be shown that [FR13]

$$||\boldsymbol{\alpha}||_p^p := \sum_{n=1}^{N} |\alpha_n|^p \xrightarrow[p \to 0]{} \sum_{n=1}^{N} \mathbf{1}_{\{\alpha_n \neq 0\}} = |\text{supp}(\boldsymbol{\alpha})|, \tag{4.1.4}$$

where $\mathbf{1}_{\{\alpha_n \neq 0\}}$ is a so called *indicator function* that gives a value of one for $\alpha_n \neq 0$ and zero for $\alpha_n = 0$, i.e.

$$\mathbf{1}_{\{\alpha_n \neq 0\}} := \begin{cases} 1 & \text{for } \alpha_n \neq 0 \\ 0 & \text{for } \alpha_n = 0 \end{cases}. \tag{4.1.5}$$

Therefore, the notation $||\boldsymbol{\alpha}||_0 := |\text{supp}(\boldsymbol{\alpha})|$ is used to calculate the cardinality of the support of $\boldsymbol{\alpha}$. Although this is not even a quasinorm and the notation $||\boldsymbol{\alpha}||_0^0$ would be more appropriate [FR13], it is common to use the expression ℓ_0 *norm* for the number of nonzero elements of $\boldsymbol{\alpha}$. In Figure 4.1.1 a visualization for different values of p in (4.1.3) is shown. These images show unit circles in different norms and it can be seen that for the case $p \to \infty$ the maximum norm is achieved and that for $p \to 0$ the unit circle changes to a cross that indicates the presence of a single value.



(a) $\ell_{0.5} - norm$ (b) $\ell_1 - norm$ (c) $\ell_2 - norm$ (d) $\ell_{10000} - norm$

Figure 4.1.1: Unit circles in different ℓ_p-norms

Therefore, the mathematical problem that must be solved to determine the sparsest solution of (4.1.1) can be written as

$$\underset{\boldsymbol{\alpha}}{\text{minimize}} \;\; ||\boldsymbol{\alpha}||_0 \quad \text{s.t.} \quad \boldsymbol{x} = \boldsymbol{\Phi\alpha} \tag{4.1.6}$$

and for (4.1.2), which is the more relevant case from a practical point of view, the corresponding problem is

$$\underset{\boldsymbol{\alpha}}{\text{minimize}} \;\; ||\boldsymbol{\alpha}||_0 \quad \text{s.t.} \quad ||\boldsymbol{x} - \boldsymbol{\Phi\alpha}||_2 < \eta. \tag{4.1.7}$$

Both problems, i.e. (4.1.6) and (4.1.7), are generally intractable, since both problems are NP-hard [FR13]. An alternative formulation for the problem in (4.1.7) can be found by combining the ℓ_0 term and the constraint using a multiplicative Lagrange factor λ to

$$\boldsymbol{\alpha}^* = \underset{\boldsymbol{\alpha}}{\text{argmin}} \;\; \lambda||\boldsymbol{\alpha}||_0 + \frac{1}{2}||\boldsymbol{x} - \boldsymbol{\Phi\alpha}||_2^2, \tag{4.1.8}$$

where $\boldsymbol{\alpha}^*$ is the optimal solution. This notation will, for some choice of the parameter λ, yield the same result as (4.1.7), but since these values of λ are generally unknown, the notation of (4.1.7) is a more natural parameterization and in most applications the more appropriate using for example an estimated noise level for the parameter η. The second term in (4.1.8), i.e. the squared ℓ_2 norm of the difference between $\boldsymbol{x}$ and $\boldsymbol{\Phi\alpha}$, is called *data fidelity* term, while the first term, i.e the ℓ_0 norm with the weight λ, is called *penalty term*.

Several algorithms have been presented over the last years to solve the problems above approximately and in some cases exactly. The success of these algorithms depends heavily on the structure of the dictionary $\boldsymbol{\Phi}$. Therefore, several different characterizations of the dictionary that should assess the success of the SR algorithm in advance were introduced over the last years, for example the *restricted isometry property* and the *null space property.* These different characteristics are used to derive sparsity levels that are guaranteed to be recovered from an arbitrary vector $\boldsymbol{x}$ with a given dictionary $\boldsymbol{\Phi}$ and most results are actually related to the convex relaxation that will be presented in Section 4.2.2 together with one example of these theoretical bounds. The drawback is that these theoretical bounds cannot be calculated for an arbitrary dictionary, they can only be shown for dictionaries that are designed according to certain rules [DDEK12]. Furthermore, these bounds are very pessimistic compared to empirical results [Ela10] and therefore, the characteristics mentioned above are not used for the dictionary design in this thesis. Other than these theoretical properties, there exist characteristics like the *spark* or the *mutual coherence* that can be directly calculated from a given dictionary. However, the only dictionary characteristic that is used within this chapter is the mutual coherence $\mu(\boldsymbol{\Phi})$, which is more intuitive than the other mentioned so far. It is defined by

$$\mu(\boldsymbol{\Phi}) = \max_{i \neq j} \frac{|\boldsymbol{\varphi}_i^H \boldsymbol{\varphi}_j|}{||\boldsymbol{\varphi}_i||_2 ||\boldsymbol{\varphi}_j||_2}, \tag{4.1.9}$$

which gives a maximum value of 1 if the dictionary contains at least two linear dependent atoms φ_i and φ_j and a minimum value of

$$\mu(\boldsymbol{\Phi}) \geq \sqrt{\frac{M - N}{N(M-1)}}, \tag{4.1.10}$$

where M and N are the dimensions of the dictionary.

The mutual coherence characterizes the dependence between the columns of the dictionary $\boldsymbol{\Phi}$ and in the design process of a dictionary a small value is desired, since a lower value guarantees a better result of the SR [Ela10]. Precisely, if the sparsity s of $\hat{\boldsymbol{\alpha}}$, which represents the output of an arbitrary algorithm, is below an upper bound given by

$$s < \frac{1}{2} \left(1 + \frac{1}{\mu(\boldsymbol{\Phi})} \right), \tag{4.1.11}$$

it is necessarily the unique solution of (4.1.6). This means that a dictionary with a low coherence allows a less sparse signal to be recovered exactly [SMF10]. This is one example of a theoretical convergence guarantee and although many of these results exist, most of them are still too pessimistic compared to empirical results [Ela10] and are therefore not considered in this thesis.

In the next section, some prototype algorithms are presented, which are used regularly to solve (4.1.6) and (4.1.7).

4.2 Algorithmic Principles

Over the last years, many algorithms were introduced to solve the problem of finding the sparsest representation using an overcomplete dictionary, as it was presented in Section 4.1. The purpose of this chapter is not to give a full overview of all available algorithms, but to show the principles of three different groups of algorithms. These groups are namely the *greedy algorithms*, the *optimization-based algorithms* and the *iterative shrinkage algorithms*. A more extensive review of SR algorithms can be found for example in [ZXY$^+$15]. Only the method that is chosen in the later presented applications is described in detail in Section 4.3.

For the description of the algorithms, the following notation is used. The common output of all algorithms is an estimate $\hat{\boldsymbol{\alpha}}$ of the true sparse vector $\boldsymbol{\alpha}$. If this estimate is build up in an iterative way, a superscript with square brackets is used to determine the estimate of the current iteration, i.e. $\hat{\boldsymbol{\alpha}}^{[\nu]}$ is the estimate of the sparse vector $\boldsymbol{\alpha}$ at iteration ν. The difference between the signal $\boldsymbol{x}$ and the reconstruction using the dictionary $\boldsymbol{\Phi}$ at iteration ν is given by the current *residual*

$$\boldsymbol{r}^{[\nu]} = \boldsymbol{x} - \boldsymbol{\Phi}\hat{\boldsymbol{\alpha}}^{[\nu]}. \tag{4.2.1}$$

Finally, the support of the estimated sparse coefficient vector at iteration ν is written as $\mathrm{supp}(\hat{\boldsymbol{\alpha}}^{[\nu]})$.

4.2.1 Greedy Methods

The methods of this section are iterative methods that build up an estimate $\hat{\boldsymbol{\alpha}}$ of the true sparse vector $\boldsymbol{\alpha}$ using locally optimal decisions at each step. These algorithms, which are also called *greedy pursuits*, give the weakest theoretical performance bounds compared to the other methods of Section 4.2.2 and 4.2.3, but are very fast and applicable to very large datasets and are therefore widely used [BDR12].

The general framework of these greedy methods consists of two steps at each iteration, first one or more atoms of the dictionary are selected and secondly, the current estimated coefficient vector is updated. The selection of the atoms is solely based on the magnitude of the projection $\boldsymbol{g}^{[\nu]}$ of the current residual on all atoms of the dictionary

$$\boldsymbol{g}^{[\nu]} = \boldsymbol{\Phi}^H \boldsymbol{r}^{[\nu]}. \tag{4.2.2}$$

From this projection, the maximum magnitudes are used and the corresponding atoms are added to the support of the estimate $\hat{\boldsymbol{\alpha}}^{[\nu]}$. The number of elements that are added to the support at each iteration depends on the specific algorithm. The update of the coefficient vector is done in a way that decreases the cost function

$$F(\hat{\boldsymbol{\alpha}}^{[\nu]}) = ||\boldsymbol{r}^{[\nu]}||_2^2 = ||\boldsymbol{x} - \boldsymbol{\Phi}\hat{\boldsymbol{\alpha}}^{[\nu]}||_2^2, \tag{4.2.3}$$

which consists only of the data fidelity term, i.e. the squared ℓ_2-norm of the residual.

This basic structure is summarized in Algorithm 4.1, which also uses a parameter S that determines the maximum number of iterations.

Algorithm 4.1: Greedy pursuit framework

```
Input: $\boldsymbol{x}$, $\boldsymbol{\Phi}$, $\eta$, S
Init: $\hat{\boldsymbol{\alpha}}^{[0]} = \mathbf{0}$, $\boldsymbol{r}^{[0]} = \boldsymbol{x}$, $\nu = 1$;
while stopping criteria is not met do
    $\boldsymbol{g}^{[\nu]} = \boldsymbol{\Phi}^H \boldsymbol{r}^{[\nu-1]}$;
    Update $\hat{\boldsymbol{\alpha}}^{[\nu]}$ based on $\boldsymbol{g}^{[\nu]}$;
    $\boldsymbol{r}^{[\nu]} = \boldsymbol{x} - \boldsymbol{\Phi}\hat{\boldsymbol{\alpha}}^{[\nu]}$;
    if $||\boldsymbol{r}^{[\nu]}||_2^2 \leq \eta$ or $\nu = S$ then
        stop pursuit;
    else
        $\nu = \nu + 1$;
    end
end
Output: $\hat{\boldsymbol{\alpha}}^{[\nu]}$, $\boldsymbol{r}^{[\nu]}$
```

In the following, two concrete examples of greedy pursuit algorithm are presented. The first one is *matching pursuit* (MP), which was the first pursuit algorithm presented for signal approximation [MZ93]. In this algorithm only one atom $\boldsymbol{\varphi}_i$ of the dictionary is selected at each iteration and added to the current support. In the coefficient vector $\hat{\boldsymbol{\alpha}}^{[\nu]}$, only the element α_i that belongs to the selected atom is updated. The selection of $\boldsymbol{\varphi}_i$ is based on the magnitude of (4.2.2) normalized by the ℓ_2-norm of the atoms. This normalization is done to select the dictionary element that removes the highest amount of energy from the current residual term. In most practical applications this normalization step is included in the design of the dictionary, i.e. all atoms posses an ℓ_2-norm of one. An algorithmic description of MP is given in Algorithm 4.2.

The second algorithm that is presented here is an extension of the MP, known as *orthogonal matching pursuit* (OMP). The structure is very similar to the one of the standard MP, but with the difference that at the end of each iteration step the signal $\boldsymbol{x}$ is projected orthogonally on the columns of $\boldsymbol{\Phi}$ associated with the current support $\mathrm{supp}(\hat{\boldsymbol{\alpha}}^{[\nu]})$, which is denoted here by the set $\Omega^{[\nu]}$. In this way the approximation is optimized with respect to the support selected up to the current iteration [Tro04]. This optimization is realized by a multiplication of the signal $\boldsymbol{x}$ with the pseudo-inverse of the support restricted dictionary $\boldsymbol{\Phi}_{\Omega^{[\nu]}}$, which consists only of the columns $\boldsymbol{\varphi}_i$ with $i \in \Omega^{[\nu]} = \mathrm{supp}(\hat{\boldsymbol{\alpha}}^{[\nu]})$. In this way all nonzero elements of $\hat{\boldsymbol{\alpha}}^{[\nu]}$ are allowed to change to

Algorithm 4.2: Matching pursuit (MP)

```
Input: $\boldsymbol{x}$, $\boldsymbol{\Phi}$, $\eta$, S
Init: $\hat{\boldsymbol{\alpha}}^{[0]} = \mathbf{0}$, $\boldsymbol{r}^{[0]} = \boldsymbol{x}$, $\nu = 1$;
while stopping criteria is not met do
    $\boldsymbol{g}^{[\nu]} = \boldsymbol{\Phi}^H \boldsymbol{r}^{[\nu-1]}$;
    $i^{[\nu]} = \text{argmax}_i |g_i^{[\nu]}| / ||\boldsymbol{\varphi}_i||_2$;
    $\hat{\alpha}_{i^{[\nu]}}^{[\nu]} = \hat{\alpha}_{i^{[\nu]}}^{[\nu-1]} + g_{i^{[\nu]}}^{[\nu]} | / ||\boldsymbol{\varphi}_{i^{[\nu]}}||_2^2$;
    $\boldsymbol{r}^{[\nu]} = \boldsymbol{r}^{[\nu-1]} - \boldsymbol{\varphi}_{i^{[\nu]}} g_{i^{[\nu]}}^{[\nu]} | / ||\boldsymbol{\varphi}_{i^{[\nu]}}||_2^2$;
    if $||\boldsymbol{r}^{[\nu]}||_2^2 \leq \eta$ or $\nu = S$ then
        stop pursuit;
    else
        $\nu = \nu + 1$;
    end
end
Output: $\hat{\boldsymbol{\alpha}}^{[\nu]}$, $\boldsymbol{r}^{[\nu]}$
```

solve the least squares minimization problem

$$\hat{\boldsymbol{\alpha}}_{\Omega^{[\nu]}}^{[\nu]} = \underset{\tilde{\boldsymbol{\alpha}}_{\Omega^{[\nu]}}}{\text{argmin}} \ ||\boldsymbol{x} - \boldsymbol{\Phi}_{\Omega^{[\nu]}} \tilde{\boldsymbol{\alpha}}_{\Omega^{[\nu]}}||_2^2. \tag{4.2.4}$$

This additional minimization problem guarantees that the residual is always orthogonal to all currently selected atoms and therefore, no atom is selected twice by the OMP algorithm. The OMP is summarized in Algorithm 4.3.

These two basic structures form the basis for many other greedy algorithms that are not shown here. For further details the reader is referred to the literature, e.g. [BDR12] or [ZXY$^+$15]. In the following part of this thesis, the algorithms that are based on mathematical optimization are shortly introduced.

4.2.2 Convex Optimization Methods

It was already mentioned at the end of Section 4.1 that the ℓ_0-norm minimization problems are NP-hard and thus intractable in practice. The first approaches in Section 4.2.1 used locally optimal heuristic measures to build the support of the underlying problem. In this chapter, no specific algorithm will be shown, but the mathematical principles to find the optimal solution will be introduced. For this principle the initial formulations of (4.1.6) and (4.1.7) are replaced by

Algorithm 4.3: Orthogonal matching pursuit (OMP)

```
Input: $\boldsymbol{x}$, $\boldsymbol{\Phi}$, $\eta$, S
Init: $\hat{\boldsymbol{\alpha}}^{[0]} = \mathbf{0}$, $\boldsymbol{r}^{[0]} = \boldsymbol{x}$, $\Omega^{[0]} = \mathrm{supp}(\hat{\boldsymbol{\alpha}}^{[0]}) = \emptyset$, $\nu = 1$;
while stopping criteria is not met do
    $\boldsymbol{g}^{[\nu]} = \boldsymbol{\Phi}^H \boldsymbol{r}^{[\nu-1]}$;
    $i^{[\nu]} = \mathrm{argmax}_i |g_i^{[\nu]}| / ||\boldsymbol{\varphi}_i||_2$;
    $\Omega^{[\nu]} = \Omega^{[\nu-1]} \cup i^{[\nu]}$;
    $\hat{\boldsymbol{\alpha}}^{[\nu]}_{\Omega^{[\nu]}} = \boldsymbol{\Phi}^{\dagger}_{\Omega^{[\nu]}} \boldsymbol{x}$;
    $\boldsymbol{r}^{[\nu]} = \boldsymbol{x} - \boldsymbol{\Phi}\hat{\boldsymbol{\alpha}}^{[\nu]}$;
    if $||\boldsymbol{r}^{[\nu]}||_2^2 \leq \eta$ or $\nu = S$ then
        stop pursuit;
    else
        $\nu = \nu + 1$;
    end
end
Output: $\hat{\boldsymbol{\alpha}}^{[\nu]}$, $\boldsymbol{r}^{[\nu]}$
```

$$\hat{\boldsymbol{\alpha}} = \underset{\boldsymbol{\alpha}}{\mathrm{argmin}} \;\; ||\boldsymbol{\alpha}||_1 \quad \text{s.t.} \quad \boldsymbol{x} = \boldsymbol{\Phi}\boldsymbol{\alpha} \tag{4.2.5}$$

and

$$\hat{\boldsymbol{\alpha}} = \underset{\boldsymbol{\alpha}}{\mathrm{argmin}} \;\; ||\boldsymbol{\alpha}||_1 \quad \text{s.t.} \quad ||\boldsymbol{x} - \boldsymbol{\Phi}\boldsymbol{\alpha}||_2 < \eta. \tag{4.2.6}$$

The particular choice of $p = 1$ in the minimization problem has been done on purpose, since this is the so called *convex relaxation* of (4.1.6) and (4.1.7). It can be shown that this choice will lead to the sparsest solutions under certain conditions, of which only one is given here without proof to show the basic relationships.

This condition is related to the null space property mentioned in Section 4.1. The null space $\mathcal{N}(\boldsymbol{\Phi})$ of a matrix $\boldsymbol{\Phi}$ is given by the vectors $\boldsymbol{u}$ that create a vector of zeros $\mathbf{0}$ when multiplied with the matrix, i.e.

$$\mathcal{N}(\boldsymbol{\Phi}) = \left\{ \boldsymbol{u} \in \mathbb{C}^N : \boldsymbol{\Phi}\boldsymbol{u} = \mathbf{0} \right\}. \tag{4.2.7}$$

With the vectors creating the null space, a sufficient condition can be given that guarantees that the solution calculated with (4.2.5) or (4.2.6) is the same as the solution of (4.1.6) or (4.1.7). This condition gives an upper bound of the sparsity of the solution $\hat{\boldsymbol{\alpha}}$ that guarantees the equivalence of the ℓ_0- and ℓ_1-minimization

$$\sqrt{||\hat{\boldsymbol{\alpha}}||_0} \leq \min \left\{ \frac{1}{2} \frac{||\boldsymbol{u}||_1}{||\boldsymbol{u}||_2} : \boldsymbol{u} \in \mathcal{N}(\boldsymbol{\Phi}) \setminus \{\boldsymbol{0}\} \right\}. \tag{4.2.8}$$

This and other conditions can be found in the literature, e.g. in [Hag20] and the references therein.

For values of $p > 1$ even 1-sparse vectors will not be found by a specific algorithm and for $p < 1$, the problem is not convex and also NP-hard in general. The convexity of the problem is a necessary condition to guarantee a unique solution of (4.2.5) and (4.2.6), which allows the application of convex optimization tools [BV04]. The special case of the ℓ_1 norm allows for example the use of linear programming techniques [Ela10]. The principle of ℓ_1-minimization for signal decomposition is also known as *basis pursuit* for (4.2.5) and as *basis pursuit denoising* for (4.2.6) [Che95]. For both principles several algorithms are available to solve this problem and an overview can be found for example in [ZXY$^+$15].

An alternate version of (4.2.6) that should be mentioned is the combination of the ℓ_1-term and the data fidelity term in a single equation using a multiplicative parameter

$$\hat{\boldsymbol{\alpha}} = \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \ \lambda ||\boldsymbol{\alpha}||_1 + \frac{1}{2} ||\boldsymbol{x} - \boldsymbol{\Phi}\boldsymbol{\alpha}||_2^2, \tag{4.2.9}$$

which is comparable to (4.1.8). Both formulations of (4.1.8) and (4.2.9) have in common that the parameter λ is generally unknown, but both formulations posses a closed form solution if $\boldsymbol{\Phi}$ is a unitary matrix, i.e. a orthonormal basis, which will be utilized in the next section.

In this thesis, methods based on convex optimization, e.g. interior point methods or the simplex algorithm, will not be used in the application and therefore, they are not further investigated. The main reason for this are the high computational demands and the fact that the application aims at a decomposition of a signal into parts that originate from different sources. For this decomposition task, a tailored algorithm is available that uses an iterative thresholding scheme, which will be described generally in the following section.

4.2.3 Iterative Thresholding Methods

In this section, iterative thresholding methods will be presented, which are between the two previous algorithm types. They can be seen as greedy methods, since they principally follow the basic scheme of Algorithm 4.1. However, the algorithms can reach global convergence if the step size is chosen properly, but not necessarily to the global optimum, since one of the algorithms uses the non-convex ℓ_0 norm in its cost function. The algorithms mentioned in Section 4.2.2 use always the convex relaxation with the ℓ_1 norm and if the given problems fulfills the conditions on the sparsity level for certain dictionaries, the convex optimization algorithms will give the optimal solution. Therefore, the iterative thresholding algorithms are presented separately because of their improved recovery guarantees and stronger theoretical foundations compared to the greedy algorithms in Section 4.2.1, although one of the algorithms uses a non-convex cost function. Furthermore,the chosen algorithm that will be presented in Section 4.3 and used in the applications of Chapter 5 is of this type, which justifies a closer look at these methods.

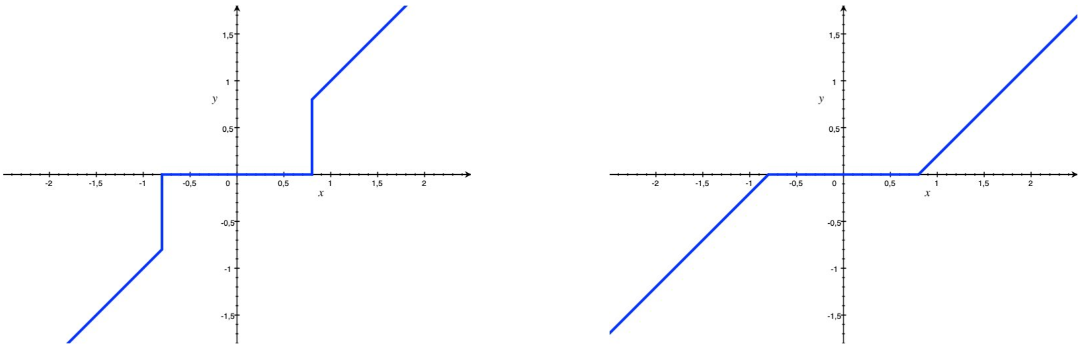
Before the algorithms are shown, two basic thresholding operations will be introduced. The first one $H_T(x)$ is called *hard thresholding* and keeps the argument x unchanged if the magnitude is above a certain threshold T

$$H_T(x) = \begin{cases} x & \text{for } |x| \geq T \\ 0 & \text{for } |x| < T. \end{cases} \tag{4.2.10}$$

The second thresholding operation $S_T(x)$ is called *soft thresholding* or *shrinkage* and is defined by

$$S_T(x) = \begin{cases} \frac{x}{|x|}(|x| - T) & \text{for } |x| \geq T \\ 0 & \text{for } |x| < T. \end{cases} \tag{4.2.11}$$

Both thresholding operations can be applied to real-valued and complex-valued numbers. In case of complex numbers, the thresholding operation is applied on the magnitude while the phase is kept unchanged. Further can both operations also be applied to vectors, in this case each element of the vector is treated independently. A visualization of both thresholding methods is shown in Figure 4.2.1 for real numbers. The output on the y-axis of Figure 4.2.1 (b) clearly shows the shrinkage effect of the soft thresholding operation.



(a) Hard thresholding

(b) Soft thresholding

Figure 4.2.1: Hard and soft thresholding

These thresholding operations possess a relation to SR that will be shown for orthonormal dictionaries before the algorithms that actually solve SR problems with overcomplete dictionaries are formulated. It was already mentioned in Section 4.2.2 that closed form solutions for (4.1.8) and (4.2.9) exist if the dictionary is a unitary matrix. In this case follows with the definition $\tilde{\boldsymbol{\alpha}} = \boldsymbol{\Phi}^{-1}\boldsymbol{x}$ a reduced form of both problems given as

$$\boldsymbol{\alpha}^* = \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \; h(\boldsymbol{\alpha}) = \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \; \lambda||\boldsymbol{\alpha}||_p + \frac{1}{2}||\tilde{\boldsymbol{\alpha}} - \boldsymbol{\alpha}||_2^2 \tag{4.2.12}$$

with $p = 0$ or $p = 1$[1]. This problem can also be seen as a data approximation with ℓ_0 or ℓ_1 penalty. In case of the ℓ_0 penalty, the problem can also be written as

$$\begin{aligned} \boldsymbol{\alpha}^* = \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \; h(\boldsymbol{\alpha}) &= \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \; \lambda||\boldsymbol{\alpha}||_0 + \frac{1}{2}||\tilde{\boldsymbol{\alpha}} - \boldsymbol{\alpha}||_2^2 \\ &= \underset{\alpha_1,\alpha_2,\ldots,\alpha_N}{\operatorname{argmin}} \; \lim_{p\to 0} \sum_{n=1}^{N} \lambda|\alpha_n|^p + \frac{1}{2}|\tilde{\alpha}_n - \alpha_n|^2, \end{aligned} \tag{4.2.13}$$

which uses the definition of the ℓ_0 norm in (4.1.4) and gives N independent problems that can be optimized separately [SED04]. To find a solution to this problem, the formulation of (4.2.13) is split into two parts $h_1(\alpha_n) = \lim_{p\to 0} \lambda|\alpha_n|^p$ and $h_2(\alpha_n) = \frac{1}{2}|\tilde{\alpha}_n - \alpha_n|^2$ and thus (4.2.13) can also be written as

$$\boldsymbol{\alpha}^* = \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \; h(\boldsymbol{\alpha}) = \underset{\alpha_1,\alpha_2,\ldots,\alpha_N}{\operatorname{argmin}} \; \sum_{n=1}^{N} h_1(\alpha_n) + h_2(\alpha_n). \tag{4.2.14}$$

To find the optimal solution $\boldsymbol{\alpha}^*$ for (4.2.14), the minimal values for $h_1(\alpha_n)$ and $h_2(\alpha_n)$ are used. $h_1(\alpha_n)$ can also be substituted by the indicator function given in (4.1.5) multiplied with λ and this will give a minimum of zero for $\alpha_n = 0$, otherwise a constant cost of λ is created. $h_2(\alpha_n)$ is a quadratic function with a minimum of zero for $\alpha_n = \tilde{\alpha}_n$, otherwise costs of $\frac{1}{2}|\tilde{\alpha}_n - \alpha_n|^2$ are created. In the optimal case of $h_1(\alpha_n)$, i.e. $\alpha_n = 0$, the second term produces costs of $\frac{1}{2}|\tilde{\alpha}_n|^2$. The choice of $\boldsymbol{\alpha}^*$ is determined by the point where $\frac{1}{2}|\tilde{\alpha}_n|^2$ reaches a value of λ, since from this point on, the choice $\alpha_n^* = \tilde{\alpha}_n$ would produce lower costs. This point is given by

$$\begin{aligned} \frac{1}{2}|\tilde{\alpha}_n|^2 &= \lambda \\ |\tilde{\alpha}_n| &= \sqrt{2\lambda} \end{aligned} \tag{4.2.15}$$

and thus the optimal solution $\boldsymbol{\alpha}^*$ for (4.2.14) is given by

$$\alpha_n^* = \begin{cases} \tilde{\alpha}_n & \text{for } |\tilde{\alpha}_n| \geq \sqrt{2\lambda} \\ 0 & \text{for } |\tilde{\alpha}_n| < \sqrt{2\lambda}, \end{cases} \tag{4.2.16}$$

---

[1] A solution for the interval $0 < p < 1$ can be found in [SMF10] and the references therein.

which is the hard thresholding operation with the threshold $\sqrt{2\lambda}$. This equivalence is used for example in image denoising with wavelets [SED04], which assumes that the signal energy can be represented by a few coefficients and the remainder contains only noise.

A comparable approach is chosen to find the optimal solution $\boldsymbol{\alpha}^*$ of (4.2.9). This problem can also be separated into N independent optimization problems and for the special case of a unitary dictionary with $\tilde{\boldsymbol{\alpha}} = \boldsymbol{\Phi}^{-1}\boldsymbol{x}$, the reduced form is

$$\begin{aligned} \boldsymbol{\alpha}^* = \underset{\boldsymbol{\alpha}}{\operatorname{argmin}}\ h(\boldsymbol{\alpha}) &= \underset{\boldsymbol{\alpha}}{\operatorname{argmin}}\ \lambda||\boldsymbol{\alpha}||_1 + \frac{1}{2}||\tilde{\boldsymbol{\alpha}} - \boldsymbol{\alpha}||_2^2 \\ &= \underset{\alpha_1,\alpha_2,\dots,\alpha_N}{\operatorname{argmin}} \sum_{n=1}^{N} \lambda|\alpha_n| + \sum_{n=1}^{N} \frac{1}{2}|\tilde{\alpha}_n - \alpha_n|^2. \end{aligned} \tag{4.2.17}$$

To find the optimal solution, the problem is further simplified to a single component written as the sum of a real and imaginary part $\alpha_R + j\alpha_I$

$$\alpha_R^*, \alpha_I^* = \underset{\alpha_R,\alpha_I}{\operatorname{argmin}}\ h(\alpha_R,\alpha_I) = \underset{\alpha_R,\alpha_I}{\operatorname{argmin}}\ \lambda\sqrt{\alpha_R^2+\alpha_I^2} + \frac{1}{2}\left((\tilde{\alpha}_R-\alpha_R)^2 + (\tilde{\alpha}_I-\alpha_I)^2\right). \tag{4.2.18}$$

If α_R^* and α_I^* are the optimal solution of (4.2.17) and if $(\alpha_R^*)^2 + (\alpha_I^*)^2 > 0$, the derivative of real and imaginary part must be zero, i.e.

$$\begin{aligned} \frac{\partial h(\alpha_R,\alpha_I)}{\partial \alpha_R} &= 0 = \lambda\frac{\alpha_R}{\sqrt{\alpha_R^2+\alpha_I^2}} + \alpha_R - \tilde{\alpha}_R \\ \frac{\partial h(\alpha_R,\alpha_I)}{\partial \alpha_I} &= 0 = \lambda\frac{\alpha_I}{\sqrt{\alpha_R^2+\alpha_I^2}} + \alpha_I - \tilde{\alpha}_I. \end{aligned} \tag{4.2.19}$$

With $c := \sqrt{\alpha_R^2 + \alpha_I^2}$, it follows that

$$\begin{aligned} \alpha_R &= 0 = \frac{c\tilde{\alpha}_R}{\lambda + c} \\ \alpha_I &= 0 = \frac{c\tilde{\alpha}_I}{\lambda + c}. \end{aligned} \tag{4.2.20}$$

If the last two equations are added, it follows

$$c = \sqrt{\alpha_R^2 + \alpha_I^2} = \frac{c}{\lambda + c}\sqrt{\tilde{\alpha}_R^2 + \tilde{\alpha}_I^2}, \tag{4.2.21}$$

and therefore

$$\sqrt{\alpha_R^2 + \alpha_I^2} = |\alpha| = \sqrt{\tilde{\alpha}_R^2 + \tilde{\alpha}_I^2} - \lambda = |\tilde{\alpha}| - \lambda. \tag{4.2.22}$$

The solution is thus a reduction of the magnitude by λ and the argument is kept unchanged for $|\tilde{\alpha}| \geq \lambda$. For $|\tilde{\alpha}| < \lambda$ the output must be zero, since the magnitude cannot be negative. This solution corresponds to the definition of the shrinkage operator in (4.2.11) with the threshold λ, i.e.

$$\alpha_n^* = \begin{cases} \frac{\tilde{\alpha}}{|\tilde{\alpha}|}(|\tilde{\alpha}| - \lambda) & \text{for } |\tilde{\alpha}| \geq \lambda \\ 0 & \text{for } |\tilde{\alpha}| < \lambda. \end{cases} \tag{4.2.23}$$

To formulate iterative thresholding algorithms, the problems given in (4.1.8) and (4.2.9) must be changed to the form of (4.2.13) and (4.2.17). Since the ℓ_p-constraint is kept in both problems, the data fidelity term must be changed to a formulation in the image domain, which is shown below.

For the following derivation, the dictionary $\mathbf{\Phi}$, which was defined as a complex-valued matrix of size $N \times M$, is changed to real-valued matrix of size $2N \times 2M$ with independent real and imaginary parts. Further are the real and imaginary parts of the vectors $\boldsymbol{x}$ and $\boldsymbol{\alpha}$ decoupled and combined to vectors of the size $2N$ and $2M$ respectively. The real-valued matrix and vectors are given by [End10]

$$\mathbf{\Phi}_{RV} = \begin{pmatrix} \mathrm{Re}\{\mathbf{\Phi}\} & -\mathrm{Im}\{\mathbf{\Phi}\} \\ \mathrm{Im}\{\mathbf{\Phi}\} & \mathrm{Re}\{\mathbf{\Phi}\} \end{pmatrix}, \quad \boldsymbol{x}_{RV} = \begin{pmatrix} \mathrm{Re}\{\boldsymbol{x}\} \\ \mathrm{Im}\{\boldsymbol{x}\} \end{pmatrix}, \quad \boldsymbol{\alpha}_{RV} = \begin{pmatrix} \mathrm{Re}\{\boldsymbol{\alpha}\} \\ \mathrm{Im}\{\boldsymbol{\alpha}\} \end{pmatrix}. \tag{4.2.24}$$

These definitions will give the same results for real and imaginary parts as the complex multiplication and are necessary to allow a Taylor expansion of $f(\boldsymbol{\alpha}) = \frac{1}{2}||\boldsymbol{x} - \mathbf{\Phi}\boldsymbol{\alpha}||_2^2$ at a point $\boldsymbol{\alpha}^t$. The index RV is omitted since the result can be applied directly with complex-valued vectors and dictionary [HGT06]. The first three coefficients of the Taylor series are

$$f(\boldsymbol{\alpha}) \approx f(\boldsymbol{\alpha}^t) + (\boldsymbol{\alpha} - \boldsymbol{\alpha}^t)\nabla f(\boldsymbol{\alpha}^t) + \frac{1}{2}(\boldsymbol{\alpha} - \boldsymbol{\alpha}^t)^T \boldsymbol{H}_f(\boldsymbol{\alpha}^t)(\boldsymbol{\alpha} - \boldsymbol{\alpha}^t), \tag{4.2.25}$$

where $\nabla f(\boldsymbol{\alpha}^t)$ is the gradient and $\boldsymbol{H}_f(\boldsymbol{\alpha}^t)$ is the Hessian matrix. Both terms can be calculated as

$$\nabla f(\boldsymbol{\alpha}) = -\mathbf{\Phi}^T(\boldsymbol{x} - \mathbf{\Phi}\boldsymbol{\alpha}) \tag{4.2.26}$$

and

$$\boldsymbol{H}_f(\boldsymbol{\alpha}) = \mathbf{\Phi}^T\mathbf{\Phi}. \tag{4.2.27}$$

With these terms and the definition of $f(\boldsymbol{\alpha})$, the Taylor series becomes

$$f(\boldsymbol{\alpha}) \approx \frac{1}{2}||\boldsymbol{x} - \boldsymbol{\Phi}\boldsymbol{\alpha}^t||_2^2 - (\boldsymbol{\alpha} - \boldsymbol{\alpha}^t)\boldsymbol{\Phi}^T(\boldsymbol{x} - \boldsymbol{\Phi}\boldsymbol{\alpha}) + \frac{1}{2}(\boldsymbol{\alpha} - \boldsymbol{\alpha}^t)^T\boldsymbol{\Phi}^T\boldsymbol{\Phi}(\boldsymbol{\alpha} - \boldsymbol{\alpha}^t), \quad (4.2.28)$$

The next step is an approximation of the Hessian matrix, which guarantees that $f(\boldsymbol{\alpha})$ is strictly convex. A sufficient condition is that the Hessian matrix must be positive-definite, which is fulfilled for a positive multiple of the identity matrix $\boldsymbol{I}$[2]. Therefore, the expression $\boldsymbol{\Phi}^T\boldsymbol{\Phi}$ is replaced by $\frac{1}{\tau}\boldsymbol{I}$ and the Taylor approximation becomes

$$\begin{aligned} f(\boldsymbol{\alpha}) &\approx \frac{1}{2}||\boldsymbol{x} - \boldsymbol{\Phi}\boldsymbol{\alpha}^t||_2^2 - (\boldsymbol{\alpha} - \boldsymbol{\alpha}^t)\boldsymbol{\Phi}^T(\boldsymbol{x} - \boldsymbol{\Phi}\boldsymbol{\alpha}) + \frac{1}{2\tau}(\boldsymbol{\alpha} - \boldsymbol{\alpha}^t)^T(\boldsymbol{\alpha} - \boldsymbol{\alpha}^t) \\ &= \frac{1}{2}||\boldsymbol{x} - \boldsymbol{\Phi}\boldsymbol{\alpha}^t||_2^2 - (\boldsymbol{\alpha} - \boldsymbol{\alpha}^t)\boldsymbol{\Phi}^T(\boldsymbol{x} - \boldsymbol{\Phi}\boldsymbol{\alpha}) + \frac{1}{2\tau}||\boldsymbol{\alpha} - \boldsymbol{\alpha}^t||_2^2 \\ &= Q_t(\boldsymbol{\alpha}, \boldsymbol{\alpha}^t). \end{aligned} \quad (4.2.29)$$

Since the dictionary $\boldsymbol{\Phi}$ is overcomplete, the identity matrix is actually not possible for the term $\boldsymbol{\Phi}^T\boldsymbol{\Phi}$. For this reason, the parameter τ is introduced to control the iterative algorithm below. An adequate choice for τ with the given dictionary $\boldsymbol{\Phi}$ is an upper bound of

$$\tau \leq \frac{1}{\lambda_{max}\left(\boldsymbol{\Phi}^H\boldsymbol{\Phi}\right)}, \quad (4.2.30)$$

where $\lambda_{max}(\boldsymbol{A})$ is the largest eigenvalue of the matrix $\boldsymbol{A}$ [Ela10]. With the approximation of $f(\boldsymbol{\alpha})$ in (4.2.29), the cost function (4.2.12) that should be minimized can be written as

$$\hat{\boldsymbol{\alpha}}^{t+1} = \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \; \lambda||\boldsymbol{\alpha}||_p + Q_t(\boldsymbol{\alpha}, \hat{\boldsymbol{\alpha}}^t), \quad (4.2.31)$$

where the iterative behavior is already indicated by the superscripts $t+1$ and t. But to achieve a structure comparable to the one in (4.2.17), a reformulation of $Q_t(\boldsymbol{\alpha}, \hat{\boldsymbol{\alpha}}^t)$ is necessary.

$$\begin{aligned} Q_t(\boldsymbol{\alpha}, \hat{\boldsymbol{\alpha}}^t) &= \frac{1}{2}||\boldsymbol{x} - \boldsymbol{\Phi}\hat{\boldsymbol{\alpha}}^t||_2^2 - (\boldsymbol{\alpha} - \hat{\boldsymbol{\alpha}}^t)\boldsymbol{\Phi}^T(\boldsymbol{x} - \boldsymbol{\Phi}\boldsymbol{\alpha}) + \frac{1}{2\tau}||\boldsymbol{\alpha} - \hat{\boldsymbol{\alpha}}^t||_2^2 \\ &= \frac{1}{2}||\boldsymbol{x} - \boldsymbol{\Phi}\hat{\boldsymbol{\alpha}}^t||_2^2 + \frac{1}{2\tau}||\boldsymbol{\alpha} - \hat{\boldsymbol{\alpha}}^t + \tau\boldsymbol{\Phi}^T(\boldsymbol{x} - \boldsymbol{\Phi}\hat{\boldsymbol{\alpha}}^t)||_2^2 - \frac{\tau}{2}||\boldsymbol{\Phi}^T(\boldsymbol{x} - \boldsymbol{\Phi}\hat{\boldsymbol{\alpha}}^t)||_2^2 \\ &= \frac{1}{2\tau}||\left(\hat{\boldsymbol{\alpha}}^t + \tau\boldsymbol{\Phi}^T(\boldsymbol{x} - \boldsymbol{\Phi}\hat{\boldsymbol{\alpha}}^t)\right) - \boldsymbol{\alpha}||_2^2 + B(\hat{\boldsymbol{\alpha}}^t) \end{aligned} \quad (4.2.32)$$

---

[2] A matrix $\boldsymbol{A}$ is called *positive-definite* if the result of $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}$ is always positive, except for $x_1 = x_2 = \cdots = x_n = 0$. A necessary condition that must be fulfilled for a positive-definite matrix $\boldsymbol{A}$ is that all elements on the main diagonal must be positive, i.e. $a_{ii} > 0$, $i = \{1, 2, \ldots, n\}$. A sufficient condition is that all eigenvalues of $\boldsymbol{A}$ are positive [BSMM08]

with

$$B(\hat{\boldsymbol{\alpha}}^t) = \frac{1}{2}||\boldsymbol{x} - \boldsymbol{\Phi}\hat{\boldsymbol{\alpha}}^t||_2^2 - \frac{\tau}{2}||\boldsymbol{\Phi}^T(\boldsymbol{x} - \boldsymbol{\Phi}\hat{\boldsymbol{\alpha}}^t)||_2^2. \tag{4.2.33}$$

The term $B(\hat{\boldsymbol{\alpha}}^t)$ in (4.2.32) is constant with respect to $\boldsymbol{\alpha}$ and can be omitted in the minimization problem. Therefore the final minimization problem is given by

$$\hat{\boldsymbol{\alpha}}^{t+1} = \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \ \lambda||\boldsymbol{\alpha}||_p + \frac{1}{2\tau}|| \left(\hat{\boldsymbol{\alpha}}^t + \tau\boldsymbol{\Phi}^T(\boldsymbol{x} - \boldsymbol{\Phi}\hat{\boldsymbol{\alpha}}^t)\right) - \boldsymbol{\alpha}||_2^2, \tag{4.2.34}$$

which is of the form of (4.2.17) with the variable substitution

$$\boldsymbol{\theta}(\hat{\boldsymbol{\alpha}}^t) = \hat{\boldsymbol{\alpha}}^t + \tau\boldsymbol{\Phi}^T(\boldsymbol{x} - \boldsymbol{\Phi}\hat{\boldsymbol{\alpha}}^t). \tag{4.2.35}$$

To achieve a solution for (4.2.34), either the hard thresholding or shrinkage operation is applied iteratively to the term $\boldsymbol{\theta}(\hat{\boldsymbol{\alpha}}^t)$ with a threshold according to the choice of p. This can be seen as a shift of the Taylor approximation until convergence and the minimum are reached. This basic step of the algorithm is actually an optimization in itself with the closed form solution given in (4.2.16) and (4.2.23). The basic iterative thresholding is summarized in Algorithm 4.4, where the index t of the incremental Taylor approximation is replaced by the variable ν to have the same notation as above.

Algorithm 4.4: Iterative thresholding

```
Input: $\boldsymbol{x}$, $\boldsymbol{\Phi}$, $\eta$, S, T
Init: $\hat{\boldsymbol{\alpha}}^{[0]} = \boldsymbol{0}$, $\boldsymbol{r}^{[0]} = \boldsymbol{x}$, $\nu = 1$, $\tau = \frac{1}{\lambda_{max}\left(\boldsymbol{\Phi}^H\boldsymbol{\Phi}\right)}$;
while stopping criteria is not met do
    $\hat{\boldsymbol{\alpha}}^{[\nu]} = \text{Threshold}_T\left(\boldsymbol{\alpha}^{[\nu-1]} + \tau\boldsymbol{\Phi}^H\boldsymbol{r}^{[\nu-1]}\right)$;
    $\boldsymbol{r}^{[\nu]} = \boldsymbol{x} - \boldsymbol{\Phi}\hat{\boldsymbol{\alpha}}^{[\nu]}$;
    if $||\boldsymbol{r}^{[\nu]}||_2^2 \leq \eta$ or $\nu = S$ then
        stop thresholding;
    else
        $\nu = \nu + 1$;
    end
end
Output: $\hat{\boldsymbol{\alpha}}^{[\nu]}$, $\boldsymbol{r}^{[\nu]}$
```

The operation Threshold_T in Algorithm 4.4 represents either the hard thresholding $H_T(\cdot)$ or the shrinkage operation $S_T(\cdot)$ dependent on the choice of p. In case of hard thresholding ($p = 0$), the algorithm is called *iterative hard thresholding* (IHT) and with soft thresholding ($p = 1$) it is called *iterative shrinkage thresholding* (IST).

A variation of the IST that improves the convergence properties of the algorithm is the *fast iterative shrinkage thresholding algorithm* (FISTA), which is given without proof here. A detailed description of this algorithm can be found in [BT09], here only the core idea is sketched. The main element of FISTA is a refinement of the coefficient vector $\hat{\boldsymbol{\alpha}}^{[\nu]}$ by linear backtracking research via a linear combination of the previous two iteration points $\hat{\boldsymbol{\alpha}}^{[\nu]}$ and $\hat{\boldsymbol{\alpha}}^{[\nu-1]}$ given by

$$\hat{\boldsymbol{\alpha}}^{[\nu]} = \hat{\boldsymbol{\alpha}}^{[\nu]} + \frac{\xi^{[\nu]} - 1}{\xi^{[\nu+1]}} \left(\hat{\boldsymbol{\alpha}}^{[\nu]} - \hat{\boldsymbol{\alpha}}^{[\nu-1]} \right). \tag{4.2.36}$$

The parameter ξ must be a positive sequence that satisfies $\xi^{[\nu]} \geq (\nu + 1)/2$, which is guaranteed by the update rule

$$\xi^{[\nu+1]} = \frac{1 + \sqrt{1 + 4(\xi^{[\nu]})^2}}{2} \tag{4.2.37}$$

and the initialization $\xi^{[1]} = 1$.

To change the IST algorithm to the FISTA, the update (4.2.36) is included between the calculation of $\hat{\boldsymbol{\alpha}}^{[\nu]}$ and the residual in Algorithm 4.4, which requires the calculation of (4.2.37) beforehand. Although this scheme is only proven to work with the IST, it will be also used with the IHT in the algorithm presented in the next section. This step can only be justified here with the success in the results that will be presented below and in the next chapter.

As it was already mentioned above, both the IHT and IST with the FISTA variation reach global convergence if the step size size τ is chosen in the given way. However, for the IHT it is only proven that the algorithm can reach a critical point, i.e. a point with a gradient of zero, of the cost function (4.1.8) [ABS13], which is non-convex. Therefore, it is not guaranteed that the algorithm will reach the global optimum. In case of FISTA, the global optimum can be reached [ZXY$^+$15], since the cost function is convex and if the conditions on the sparsity mentioned in Section 4.2.2 are fulfilled, the global optimum of the convex relaxation coincides with the optimum of (4.1.8). A point that distinguishes FISTA from the convex optimization methods mentioned in Section 4.2.2 is the non-monotonicity of its convergence. It is possible that FISTA exceeds the optimum value and a rippling behavior occurs towards the end of the algorithm around the true optimum [CAD18].

4.3 Morphological Component Analysis

In this section, an algorithm that is specifically designed to decompose signals into different components is introduced. The algorithm is called *morphological component analysis* (MCA) and will be introduced in Section 4.3.1 together with the principle of *morphological diversity*. These theoretical foundations will be visualized by a basic image processing example in Section 4.3.2.

4.3.1 Model Assumptions and Algorithm Description

The starting point of this algorithm is also (4.1.1), which can be extended by a noise term $\boldsymbol{n}$ to

$$\boldsymbol{x} = \boldsymbol{\Phi}\boldsymbol{\alpha} + \boldsymbol{n}. \tag{4.3.1}$$

This noise term is the reason that normally (4.1.2) is used to find the sparse representation of the input data $\boldsymbol{x}$. The model of morphological diversity assumes that the input data can be modeled as the sum of K different components $\boldsymbol{x}_{(k)}$ [SMF10], i.e.

$$\boldsymbol{x} = \sum_{k=1}^{K} \boldsymbol{x}_{(k)}. \tag{4.3.2}$$

Each $\boldsymbol{x}_{(k)}$ is called a *morphological component* and should be sparsely represented by a given dictionary $\boldsymbol{\Phi}_{(k)}$ with the coefficient vector $\boldsymbol{\alpha}_{(k)}$. Up to this point, the dictionary was assumed to be one large matrix $\boldsymbol{\Phi}$, which is still valid here, but this matrix can now be divided into several subdictionaries $\boldsymbol{\Phi}_{(k)}$ and each component is represented as

$$\boldsymbol{x}_{(k)} = \boldsymbol{\Phi}_{(k)}\boldsymbol{\alpha}_{(k)} \quad \forall\, k \in \{1, \dots, K\}. \tag{4.3.3}$$

Each *subdictionary* $\boldsymbol{\Phi}_{(k)}$ must provide a sparse representation of the component it is serving, while being inefficient representing the other contents. One possibility to ensure this behavior is to use dictionaries, which are *mutually incoherent*, i.e. the combined dictionary possesses a low mutual coherence. The choice of the components depends on the given problem. In a radar context, the vector $\boldsymbol{x}$ is given by the recorded radar data and in Chapter 5 it will be pointed out that these components can be different scattering mechanisms. The vectors $\hat{\boldsymbol{\alpha}}_{(k)}$ estimate the scene, i.e. the coordinates, of the specific scattering mechanisms and the sum over all $\boldsymbol{\Phi}_{(k)}\hat{\boldsymbol{\alpha}}_{(k)}$ gives the noise free radar data as sum of the different scattering mechanisms in the ideal case. The different subdictionaries $\boldsymbol{\Phi}_{(k)}$ do not have to be overcomplete, in fact undercomplete subdictionaries are also possible. The concatenation of all subdictionaries creates an overcomplete dictionary and thus an underdetermined problem.

The idea of the presented algorithm, which is called *morphological component analysis* (MCA) is the estimation of the components $\boldsymbol{\alpha}_{(k)}$ by solving the constraint optimization

$$\hat{\boldsymbol{\alpha}}_{(1)}, \dots, \hat{\boldsymbol{\alpha}}_{(K)} = \underset{\boldsymbol{\alpha}_{(1)},\dots,\boldsymbol{\alpha}_{(K)}}{\operatorname{argmin}} \sum_{k=1}^{K} ||\boldsymbol{\alpha}_{(k)}||_p \quad \text{s.t.} \quad ||\boldsymbol{x} - \sum_{k=1}^{K} \boldsymbol{\Phi}_{(k)}\boldsymbol{\alpha}_{(k)}||_2^2 < \eta. \tag{4.3.4}$$

The parameter η in the constraint function is used to take noise into account.

Assumed that each dictionary gives a sparse representation of its specific content, Equation 4.3.4 gives the coefficient vectors $\boldsymbol{\alpha}_{(k)}$ of the different components, which are concatenated to the overall coefficient vector $\boldsymbol{\alpha}$ in (4.3.1). It should be mentioned that this algorithm should not be mistaken with the morphological operations known from image processing [Pra07]. Although MCA was also introduced in the context of image processing [SED04], the term *morphology* refers to the general study of structures underlying measured data in this case.

A solution of 4.3.4 can be achieved by iterative hard thresholding ($p = 0$) or iterative soft thresholding ($p = 1$) [SMF10], which were presented in Section 4.2.3. The difference compared to the general scheme of the algorithms presented above is the now given structure of the dictionary. In Section 4.2.3, the dictionary was assumed to be one large matrix, which is still valid here, but this matrix can now be divided into several subdictionaries $\boldsymbol{\Phi}_{(k)}$. To solve this problem, the so called *block coordinate relaxation* (BCR) method is exploited. This method was introduced for unitary subdictionaries, i.e. orthonormal transformations as in the derivations of the optimal solutions of (4.2.12) in Section 4.2.3, and the method is applied iteratively with all but one $\boldsymbol{\alpha}_{(k)}$ fixed [BST98]. In this case, (4.3.4) can be written for $p = 1$ as

$$\hat{\boldsymbol{\alpha}}_{(l)} = \underset{\boldsymbol{\alpha}_{(l)}}{\operatorname{argmin}} \, ||\boldsymbol{\alpha}_{(l)}||_1 \quad \text{s.t.} \quad \left|\left| \left(\boldsymbol{x} - \sum_{k=1, k \neq l}^{K} \boldsymbol{\Phi}_{(k)} \boldsymbol{\alpha}_{(k)} \right) - \boldsymbol{\Phi}_{(l)} \boldsymbol{\alpha}_{(l)} \right|\right|_2^2 < \eta, \qquad (4.3.5)$$

which is of the same type as (4.2.17) that is known to possess the closed form solution in (4.2.23). The iterative application of (4.3.5) will converge to the optimum of (4.3.4) for $p = 1$ if the dictionary $\boldsymbol{\Phi}$ is built of unitary subdictionaries $\boldsymbol{\Phi}_{(k)}$, which guarantees a unique reconstruction for each subdictionary [SED04]. Although a convergence guarantee cannot be given for overcomplete subdictionaries, the careful design of them considering a low mutual coherence between the subdictionaries gives an intuitive justification that the algorithm should pick only salient features for the current component in each iteration [SMF10].

To realize that the algorithm picks only salient features in each iteration, it is applied with a decreasing threshold from one iteration to the next. This scheme using a decreasing threshold results in the desired coarse to fine behavior of the decomposition. Therefore, an initial maximum and a minimal threshold must be determined and the actual threshold is varied between these two points. The two standard schemes are either a linear or an exponential decrease determined by the predefined maximum number of iterations. To determine the initial threshold, a common strategy is to set it to a certain value that only coefficients from a single subdictionary are chosen at the first iteration. The algorithm can also be stopped early if the targeted error bound is achieved before the minimal threshold is reached. Additionally to these standard schemes, adaptive methods can be used to determine the threshold in each iteration [BSF^{+}07]. The structure of MCA is shown in Algorithm 4.5 using the shrinkage operation.

Algorithm 4.5: Morphological component analysis

```
Input: $\boldsymbol{x}$, $\boldsymbol{\Phi} = \left[\boldsymbol{\Phi}_{(1)}, \boldsymbol{\Phi}_{(2)}, \ldots, \boldsymbol{\Phi}_{(K)}\right]$, $\eta$, S
Init: $\hat{\boldsymbol{\alpha}}_{(k)}^{[0]} = \mathbf{0}\ \forall k$, $\boldsymbol{r}^{[0]} = \boldsymbol{x}$, $\xi^{[1]} = 1$, $\nu = 1$, $\tau = \frac{1}{\lambda_{max}\left(\boldsymbol{\Phi}^H\boldsymbol{\Phi}\right)}$;
      let $k^* = \max_{k} ||\boldsymbol{\Phi}_{(k)}^H \boldsymbol{x}||_\infty$, then $T^{[1]} = \max_{k \neq k^*} ||\boldsymbol{\Phi}_{(k)}^H \boldsymbol{x}||_\infty$;
while stopping criteria is not met do
    $\xi^{[\nu+1]} = \frac{1+\sqrt{1+4(\xi^{[\nu]})^2}}{2}$;
    for $k = 1$ to $K$ do
        $\hat{\boldsymbol{\alpha}}_{(k)}^{[\nu]} = S_T\left(\boldsymbol{\alpha}_{(k)}^{[\nu-1]} + \tau\boldsymbol{\Phi}_{(k)}^H \boldsymbol{r}^{[\nu-1]}\right)$;
        $\hat{\boldsymbol{\alpha}}_{(k)}^{[\nu]} = \hat{\boldsymbol{\alpha}}_{(k)}^{[\nu]} + \frac{\xi^{[\nu]}-1}{\xi^{[\nu+1]}}\left(\hat{\boldsymbol{\alpha}}_{(k)}^{[\nu]} - \hat{\boldsymbol{\alpha}}_{(k)}^{[\nu-1]}\right)$;
    end
    $\boldsymbol{r}^{[\nu]} = \boldsymbol{x} - \sum_{k=1}^{K} \boldsymbol{\Phi}_{(k)} \hat{\boldsymbol{\alpha}}_{(k)}^{[\nu]}$;
    if $||\boldsymbol{r}^{[\nu]}||_2^2 \leq \eta$ or $\nu = S$ then
        stop MCA;
    else
        $\nu = \nu + 1$;
        Set $T^{[\nu]}$ according to chosen updating scheme;
    end
end
Output: $\hat{\boldsymbol{\alpha}}_{(1)}^{[\nu]}, \ldots, \hat{\boldsymbol{\alpha}}_{(K)}^{[\nu]}$, $\boldsymbol{r}^{[\nu]}$
```

This iterative behavior with a decreasing threshold also allows to correct separations from early iterations, i.e. content of the data $\boldsymbol{x}$ is allowed to change from one morphological component to another during the iterations. This is a result of the block-wise application of the thresholding for each subdictionary with the residual of the previous iteration. Further, the algorithm can also be used with the hard instead of the soft thresholding, which would lead to the minimization of the ℓ_0 norm in (4.3.4). Also a combination of both thresholding methods is possible, in [SED04] it is mentioned that the best results are often reached if the soft thresholding is replaced by hard thresholding towards the end of the decomposition.

It should be mentioned that the MCA in Algorithm 4.5 differs from many MCA descriptions in the literature, for example in [FSED10], since these notations are often limited to image processing applications and assume dictionaries, which can be replaced by a fast transform, e.g. undecimated wavelet transform. In this case a faster implementation is possible that is applied in the image domain and is not minimizing the ℓ_p norm of the coefficient vector. Furthermore, the references above use an MCA implementation without the FISTA iteration scheme. To visualize the MCA decomposition a basic image processing example is shown below.

Another point that should be mentioned is the difference of this algorithm to so called *group sparsity algorithms*, like the group Least Absolute Shrinkage and Selection Operator (LASSO) [YL06]

$$\hat{\boldsymbol{\alpha}}_{(1)}, \ldots, \hat{\boldsymbol{\alpha}}_{(K)} = \underset{\boldsymbol{\alpha}_{(1)}, \ldots, \boldsymbol{\alpha}_{(K)}}{\operatorname{argmin}} \frac{1}{2} \left\| \boldsymbol{x} - \sum_{k=1}^{K} \boldsymbol{\Phi}_{(k)} \boldsymbol{\alpha}_{(k)} \right\|_2^2 + \lambda \sum_{k=1}^{K} \sqrt{M_{(k)}} \left\| \boldsymbol{\alpha}_{(k)} \right\|_2 , \tag{4.3.6}$$

where $M_{(k)}$ is the number of atoms in the k-th subdictionary. The penalty term in this equation is the sum of the ℓ_2 norm of the groups in the decomposition. If the group size is reduced to one element, this penalty will be the same as the ℓ_1 penalty before. Therefore it can be seen as a sparse approximation for coefficient groups, i.e. the algorithm handles a whole group as one element in the sparsity penalty. As a result of this formulation, the number of groups is minimized, but all coefficients in the selected groups are allowed to be nonzero [SFHT13].

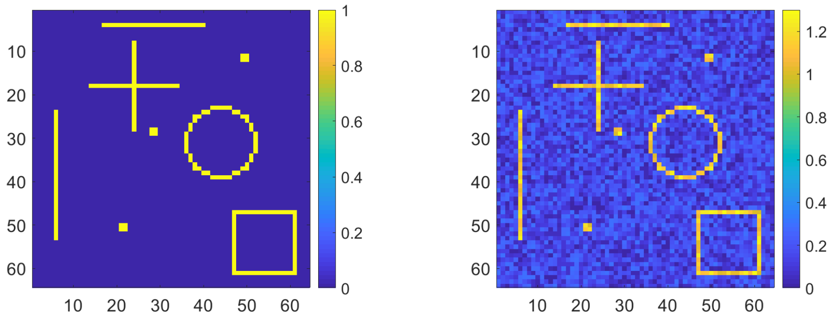
Recalling that the aim of this thesis is the detection of jet engines in ISAR using the waveguide model from Chapter 3 and their separation from point scattering centers, a coefficient vector with only nonzero elements would detect jet engines in the whole ISAR image. Therefore, this algorithm seems to be not appropriate for the given task. A possible solution to this problem is the sparse group LASSO [SFHT13]

$$\hat{\boldsymbol{\alpha}}_{(1)}, \ldots, \hat{\boldsymbol{\alpha}}_{(K)} = \underset{\boldsymbol{\alpha}_{(1)}, \ldots, \boldsymbol{\alpha}_{(K)}}{\operatorname{argmin}} \frac{1}{2N} \left\| \boldsymbol{x} - \sum_{k=1}^{K} \boldsymbol{\Phi}_{(k)} \boldsymbol{\alpha}_{(k)} \right\|_2^2 + (1-\beta)\lambda \sum_{k=1}^{K} \sqrt{M_{(k)}} \left\| \boldsymbol{\alpha}_{(k)} \right\|_2 + \beta\lambda \left\| \boldsymbol{\alpha} \right\|_1 , \tag{4.3.7}$$

where N is the dimension of $\boldsymbol{x}$, $\boldsymbol{\alpha}$ is the concatenation of the $\boldsymbol{\alpha}_{(k)}$ and $\beta \in [0, 1]$ is an additional optimization parameter to create a convex combination of the two penalties. For $\beta = 0$, the expression is reduced to the group LASSO and for $\beta = 1$, the Lagrangian form of (4.3.4) for $p = 1$ is created. The additional ℓ_1 penalty is used to create sparse coefficients in the sparse groups. This would solve the problem of equally distributed detection over the whole ISAR image, but the penalty of sparse groups might removes certain scattering mechanisms completely. Since the aim is to design a dictionary only with scattering mechanisms that are expected in the data, this variation of the group LASSO, and group sparsity algorithms in general, are also not appropriate for the task of this thesis.

4.3.2 An Image Processing Example

To visualize the decomposition task of the MCA, an image with different shapes is created and decomposed according to these shapes. The test image is depicted in Figure 4.3.1 (a), which is a binary image with a value of zero for the background and one for the shapes. It is visible that the image is a composition of dots, sqares, crosses, circles, horizontal lines, and vertical lines. For the actual decomposition example a noisy version of this image is used, which is shown in Figure 4.3.1 (b). The noise is equally distributed noise in the interval [0, 0.3] and was added to the original image.



(a) Original

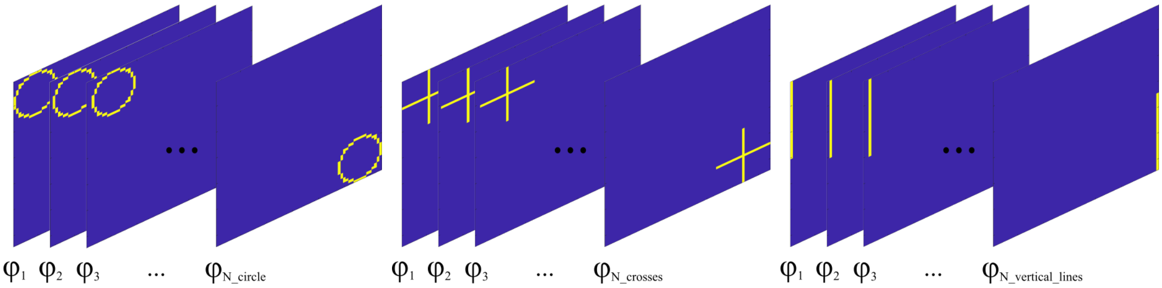
(b) With noise

Figure 4.3.1: Test image

In this example, six different subdictionaries are designed, one for each shape. The columns of these subdictionaries, i.e. the atoms, are scaled to give a value of one if the atom matches the object in the image. To perform the multiplication $\boldsymbol{\Phi}_{(k)}^H \boldsymbol{x}$ with the image as vector $\boldsymbol{x}$, the image is vectorized by column stacking. Since the data in this example is real-valued, the Hermitian of the dictionary can be replaced by the transposed. Three of these six subdictionaries are visualized in Figure 4.3.2, which shows the atoms without column stacking.

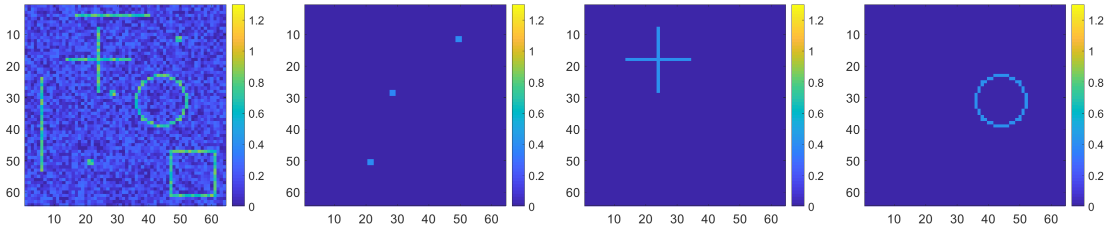
For the implementation of this example, the minimum threshold was set to 0.1, which is inside the noise level. The maximum threshold was determined in the way given in Algorithm 4.5, and 100 iterations with an exponential decrease are used. The stopping criteria using η was not implemented to show the influence of a threshold value within the noise level. The thresholding operation itself was realized by the shrinkage operator.

For the visualization of the decomposition, the reconstructions in the image domain $\boldsymbol{\Phi}_{(k)}\hat{\boldsymbol{\alpha}}_{\boldsymbol{(k)}}$ are shown in Figures 4.3.3 to 4.3.5 (b) to (d) and (f) to (h) together with the remaining residual in plot (a) and the sum of the components $\sum_{k=1}^{6} \boldsymbol{\Phi}_{(k)}\hat{\boldsymbol{\alpha}}_{\boldsymbol{(k)}}$ in plot (e). The decomposition is shown after 20, 60 and 100 iterations, which is the final result. All plots are in the same dynamic range,which is matched to the maximum value in the noisy input image, to allow a comparison between the different Figures.

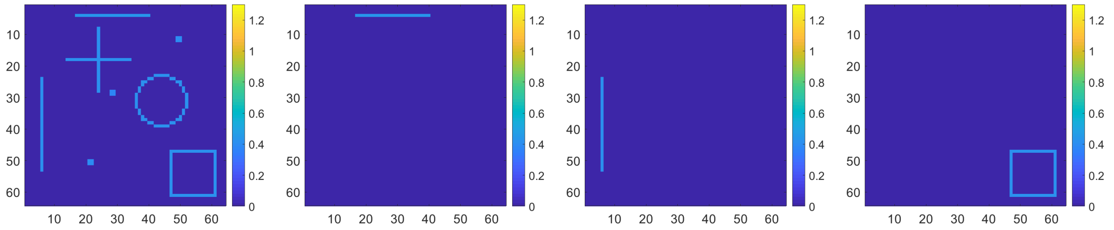


(a) Circles (b) Crosses (c) Vertical lines

Figure 4.3.2: Visualization of three subdictionaries

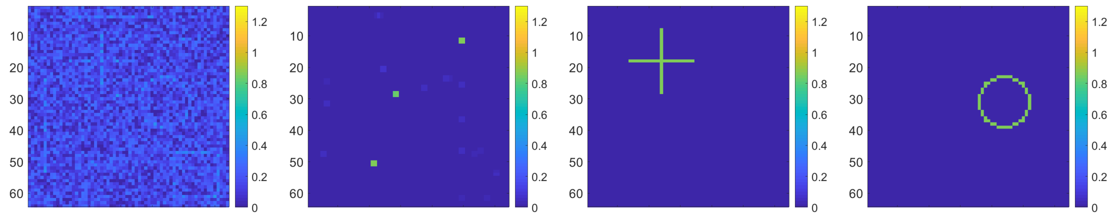


(a) Residual (b) Dots (c) Crosses (d) Circles

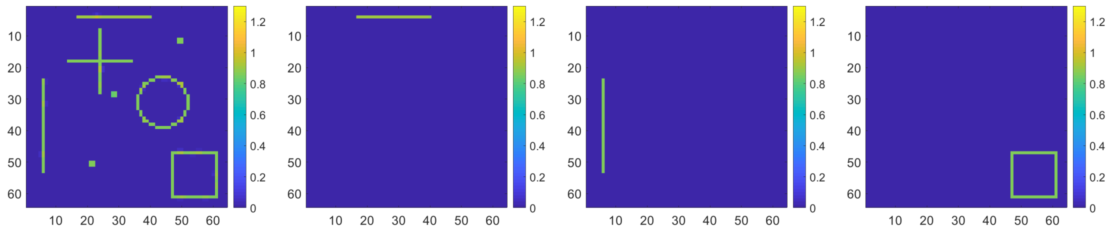


(e) Components sum (f) Horizontal lines (g) Vertical lines (h) Squares

Figure 4.3.3: MCA Example after 20 Iterations

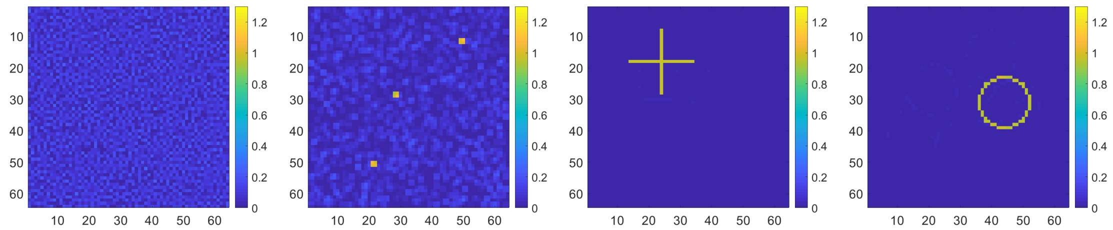


(a) Residual (b) Dots (c) Crosses (d) Circles

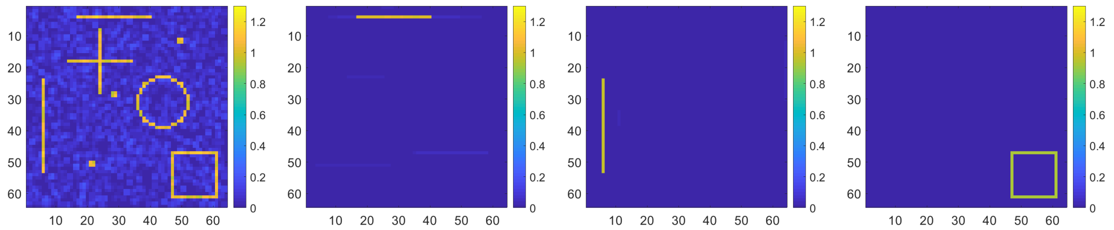


(e) Components sum (f) Horizontal lines (g) Vertical lines (h) Squares

Figure 4.3.4: MCA Example after 60 Iterations



(a) Residual (b) Dots (c) Crosses (d) Circles



(e) Components sum (f) Horizontal lines (g) Vertical lines (h) Squares

Figure 4.3.5: MCA Example after 100 Iterations

It can be seen that the contribution of the single shapes is increasing constantly during the decomposition algorithm. The color intensity in the plots of the single components and the components sum of the current iteration is increasing, while the components are removed from the residual. The influence of the noise can be seen in the images of Figure 4.3.5, which is the final result. Since the threshold is reduced towards the end of the algorithm to values within the noise level, noisy elements become visible in the reconstruction of the different shapes. Especially the dictionary of the dots, which are produced of 2×2 pixel areas, is prone to noise due to the high similarity with the canonical single pixel basis. The other elements posses a higher averaging effect, although some artifacts are visible in all components.

The next step is to apply the MCA algorithm to radar data, which is done in the experiments of the following chapter.

Chapter 5

Decomposition of Radar Images

This chapter presents the results obtained with the methods and models of the previous chapters and is split into three parts. The first one is a simulation with scatterers of different shapes with the aim to decompose the complex-valued raw data according to the scattering centers. This simulation is comparable to the example in Section 4.3.2, which was applied in the real-valued image domain. Results for this simulation are presented in Section 5.1.

The second part in Section 5.2 presents a simulation using the model of Chapter 3 together with the standard isotropic point scattering model. This simulation aims at validating the model and showing the capabilities of the algorithm to separate both kinds of scattering. Further, the sensitivity of different models and algorithmic parameters is investigated.

In the last experiment, in Section 5.3, the MCA is applied to in-flight ISAR images with the purpose to recognize and localize jet engines in these images. This experiment is done with data measured by the Fraunhofer TIRA system.

5.1 Simulation of Objects with Different Shapes

The first experiment of this chapter is comparable to the image processing example in Section 4.3.2. The main difference is the data domain in which the decomposition is applied. The example of Section 4.3.2 has used real-valued image data, while the decomposition here is applied to complex-valued data in the k_x-k_y-domain that was introduced in Section 2.2.1. Therefore, the example given here is closer to a radar signal processing application. The simulated data is created with a discretized version of (2.2.8), which is defined by the discrete sampling times t_i of the radar receiver and the observation angles φ_i of the measurement.

The radar parameters used in this section are the same as in Section 3.3 and are chosen accordingly to the TIRA parameters given in [Meh96]. These parameters are a center frequency of $16.7\,\mathrm{GHz}$, a bandwidth of $800\,\mathrm{MHz}$ and a chirp duration of $256\,\mu\mathrm{s}$. The integration angle was chosen with (2.2.9) to achieve the same resolution in cross-range as in range. The center integration angle is set to zero degrees and in both dimensions 64 bins are used, i.e. the simulated data is given as a complex-valued

64 × 64 matrix. A plot of the scene that is used for the simulations is shown in Figure 5.1.1.

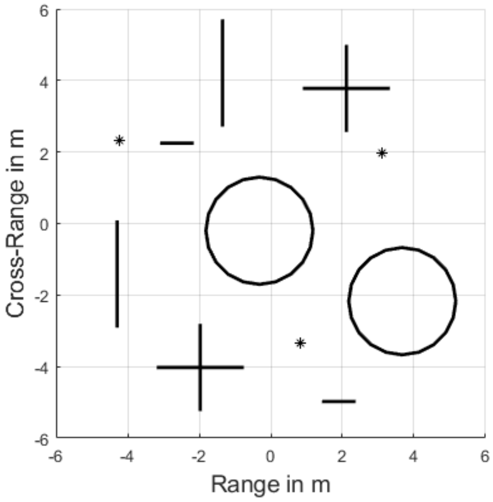


Figure 5.1.1: Plot of the scene used for the simulation

This plot shows two circles, two crosses, two vertical lines, two horizontal lines and three point scatterers, which are marked by the $*$ sign. For this simulation these targets are modeled as a collection of point scatterers with a difference in range and cross-range that fits the size of the corresponding resolution cells. For a more challenging decomposition a small random shift value is added to the position to create an off-grid problem. It should be mentioned that the on-grid simulation is of limited use to verify the robustness of the decomposition, since the algorithm performs the inverse operation with the same model as the forward operation that is used to create the data. In this case the added noise is the only factor that limits the performance of the algorithm. In general, the decomposition is a so called *inverse problem*, i.e. the aim is to find the sources that cause the noisy observations. The effect that the same model is used to synthesize the data as for the inversion is also called the *inverse crime* [Wir04]. Therefore, the off-grid simulation is added, since in this case, it is almost impossible that an element of the dictionary matches the forward model exactly and the inverse crime is avoided. Results will be presented for both cases, i.e. scattering centers on the grid and off-grid.

The basic formula for this simulation is (2.2.8), which is repeated here for convenience

$$y(k_x, k_y) = \iint e^{-j(k_x x + k_y y)} a(x, y) dx dy. \tag{5.1.1}$$

For the implementation of the algorithm, a discrete approximation of this formula is necessary to create the entries of the different dictionaries and vectors. Therefore, the discrete version [ÇSÖ$^+$14]

$$y(k_x, k_y) = \sum_{m=1}^{M} \sum_{n=1}^{N} e^{-j(k_x x_m + k_y y_n)} a(x_m, y_n). \tag{5.1.2}$$

is used here with real valued amplitudes, i.e. $a(x_m, y_n)$, of the scattering centers. To calculate the discrete k_x and k_y values, the integration angle φ_{int} must be determined. For this calculation, equation 2.2.9 is used with the cross-range resolution ΔCR set to the range resolution defined by the bandwidth. With this choice, the integration angle is calculated using the wavelength λ_0 of the center frequency as

$$\varphi_{int} = \frac{\lambda_0}{2\Delta CR}. \tag{5.1.3}$$

The integration angle is represented by a vector with 64 entries bounded by $\pm\varphi_{int}/2$ and centered around zero that represent the measurement angles of the simulation. The vector k_r is given by the center frequency f_0 and the bandwidth B by

$$k_r = \frac{2\pi}{c_0}\left(f_0 + B\frac{[-32, -31, \ldots 31]}{64}\right), \tag{5.1.4}$$

where $[-32, -31, \ldots 31]$ is a vector with integer values from -32 to 31 representing the sampling in time and c_0 is the speed of light. The length of this vector is predefined by the dimension of the problem. This formulation is possible since the simulation time t of (2.1.36) corresponds to the pulse length τ_P here and therefore, the pulse length is canceled in the equation.

With the values of k_r and φ_{int}, the values for k_x and k_y can be calculated as two-dimensional fields with one specific φ-value for each row. A visualization of the covered k-space is shown in Figure 5.1.2, which shows a slight curvature of a ring segment due to the polar format of the simulated data. The actual integration angle that is needed to achieve the same resolution in cross-range as in range is calculated as 2.78°. Since this angle is rather small, the depicted ring segment looks almost square.

The k-domain data is calculated for each of the shapes in Figure 5.1.1 and stored for later comparison with the decomposition results. The overall received signal is calculated as the sum of all shapes and a noise term is added to achieve a predefined SNR[1], which is varied here between -10 and 20 dB. A detailed model of the simulated data for the different shapes will be given below in Section 5.1.1, where the dictionary design for the different shapes is described.

Since the modeled signal represents the complex-valued data of a radar receiver, it cannot be plotted directly. For this reason the real part of the signal is shown in Figure 5.1.3 for three different SNR levels together with the magnitude plots of the scene. These scene plots are calculated by a matrix-vector multiplication of the stacked data in the upper row with the standard signal model, i.e. a matched filter designed

---

[1] For the SNR, the ratio between the average magnitude A in the raw data domain and the noise standard deviation σ is used here. To reach a specific SNR in db (SNR_{dB}), the noise standard deviation is calculated by $\sigma = A/10^{SNR_{dB}/10}$.

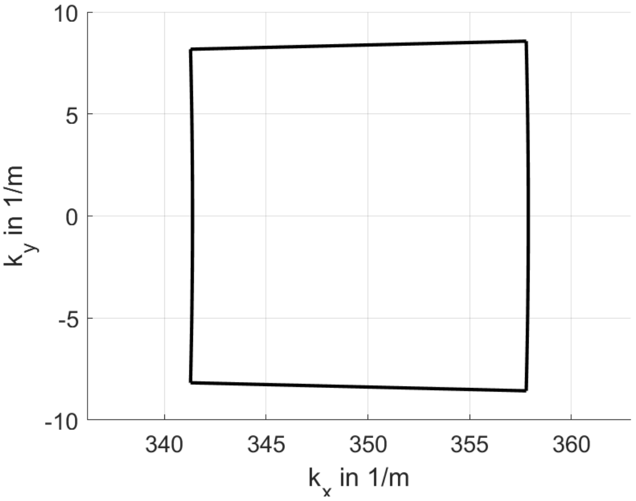


Figure 5.1.2: k-domain covered within this simulation

with (5.1.1) and the parameters that have been used for the simulation. This matched filter will be also used as the point scatterer dictionary, and more details on the design of it can be found in Section 5.1.1. To create a vector from the data in the upper row, the columns of the matrices are stacked over each other. There are also other options to create vectors from the data, the only point is that the dictionary elements must be vectorized in the same way. Details on the dictionary design will be also given below.

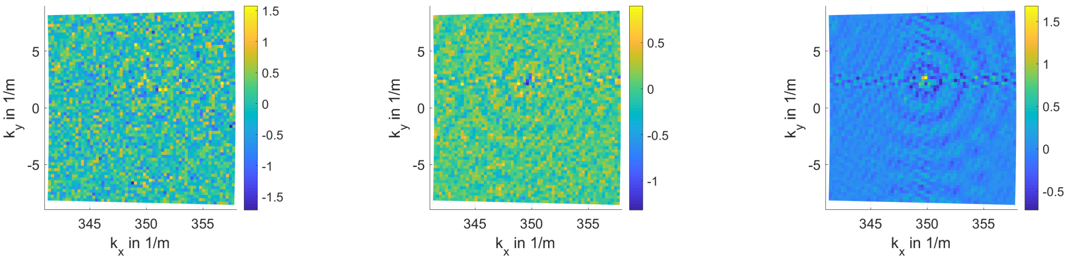
Compared to the positions shown in Figure 5.1.1, the positions of the targets in the examples of Figure 5.1.3 are slightly shifted to move the targets to arbitrary positions between the sampling grid. This shift is the reason for the several visible sidelobes in the reconstructed images.

5.1.1 Dictionary Design

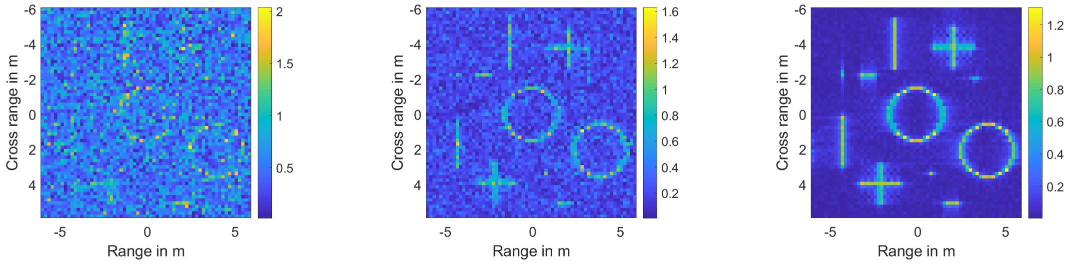
The MCA decomposition of Section 4.3 is applied here with six different subdictionaries, i.e. one for each scatterer type that is present in the scene plus an additional dictionary of squares, which is used to verify the robustness of the decomposition. The coefficient vector of the latter dictionary should still be zero at the end of the decomposition, since no square shapes are present in the scene.

The basic signal model is the exponential term in (5.1.1) with the fields of k_x and k_y multiplied with specific values of x and y given as multiples of the range and cross-range resolution. The overall received signal of a specific scatterer shape is given by the coherent sum of several independent point scatterers, e.g. one scattering center for a point scatterer and five scattering centers for the horizontal lines in Figure 5.1.1.

In this way, the dictionary atoms are modeled with x and y values set according



(a) k-domain (-4dB SNR) (b) k-domain (0dB SNR) (c) k-domain (10dB SNR)



(d) Image (-4dB SNR) (e) Image (0dB SNR) (f) Image (10dB SNR)

Figure 5.1.3: Simulated k-domain data and reconstruction with different SNR values

to the relative point scatterer position of the different target shapes starting at every possible grid position. These positions are given in multiples of the resolution cells, i.e. they are placed on the measurement grid, and an atom is only constructed if the shape fits completely into the image. To convert the problem to the structure of a matrix-vector multiplication, the received signal is changed from a matrix to a vector by column-stacking and the same method is applied to the dictionaries by stacking the columns to create an atom of the dictionary.

As it was already mentioned, the simulation is performed with the scatterers on the sampling grid as well as off the grid. To consider scatterers that are between sampling points of the grid, the dictionaries are extended by a factor of two or four in each dimension by including additional atoms to the dictionary that represent scatterers at points between the resolution cells. This grid oversampling is depicted in Figure 5.1.4 for an oversampling factor of two in each dimension. The black circles indicate the original sampling points on the grid with values for x and y according to the resolution of the radar. The blue circles indicate atoms with x and y values finer than the radar resolution, i.e with these atoms, the dictionary possesses super-resolution capabilities and the relative position between targets can be determined more accurately. It should be mentioned that the oversampling refers only to the starting point of an atom, the relative distances between the scattering centers of an atom are the same for all elements.

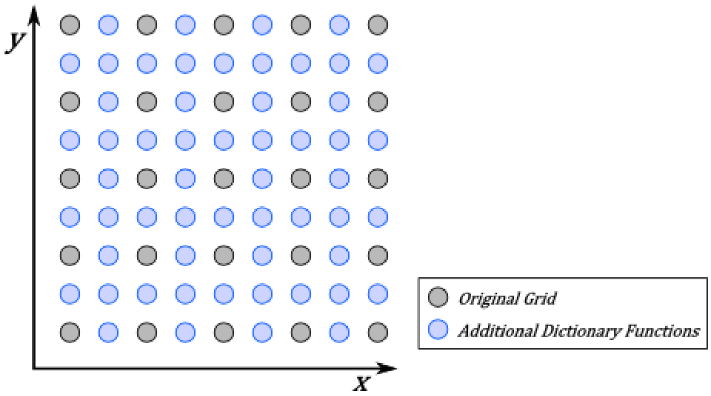


Figure 5.1.4: Schematic representation of the sampling grid with oversampling

Another option to take the off-grid problen into account is the multiplication with arbitrary range and cross range values in the frequency domain. This would require another search step during the decomposition to find the best matching x_m and y_n values in (5.1.2). This search would increase the complexity of the algorithm and is not considered in this thesis, but should be mentioned as an option to improve the results in a future work.

All atoms are calculated with (5.1.2), i.e. the discretised version of (5.1.1). In case of the oversampled dictionary, the values of x and y are positions between the theoretical sampling points given by the bandwidth of k_x and k_y. The distribution $a(x_m, y_n)$ of the scattering centers corresponds to the already mentioned collection of point scattering centers for each shape. With this model it is not possible to scale the different shapes during the decomposition. If shapes appear in different scales, a new dictionary must be designed for each scale. The design of the different dictionaries is described next, starting with the elementary point scatterer dictionary.

5.1.1.1 Point Scatterer Dictionary

The basic function for an atom of a point scatterer $\varphi_i^{(PS)}$ at position (x_i, y_i) is given as

$$\varphi_i^{(PS)} = e^{-j(k_x x_i + k_y y_i)}. \tag{5.1.5}$$

To fit this function into the point scatterer dictionary $\Phi_{(PS)}$ it is vectorized by column stacking in the same way as the simulated received signal of the scene[2]. With the grid values for (x_i, y_i), the basic dictionary is built, while for the oversampled dictionaries $\Phi_{(PS\ SR2)}$ and $\Phi_{(PS\ SR4)}$ with two- and fourfold oversampling in each dimension, respectively, additional points are added accordingly to the principle shown in Figure 5.1.4.

---

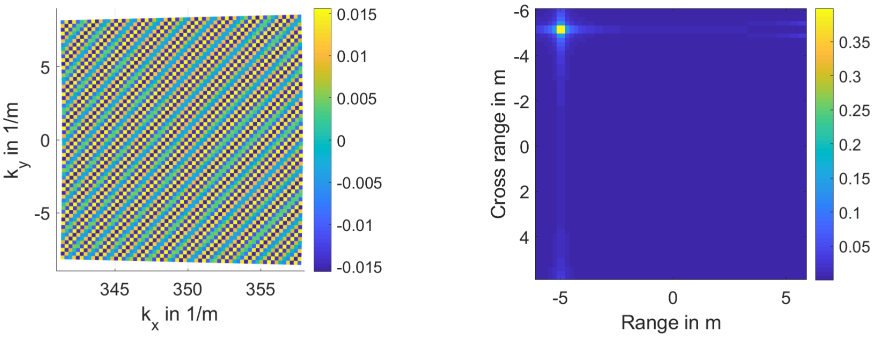
[2] In principle any vectorization is possible as long as it is applied consistently for all matrices that occur in the decomposition problem.

The size of the dictionary scales with the squared value of the oversampling factor. This means that the dictionary $\Phi_{(PS\;SR2)}$ is roughly four times the size of the standard one and $\Phi_{(PS\;SR4)}$ is around 16 times larger.

The matched filter results $\hat{a}(x, y)$ above are created by a matrix-vector multiplication of the simulated raw data $y(k_x, k_y)$ with the conjugate transpose of the dictionary $\Phi^H_{(PS)}$, i.e.

$$\hat{a}(x, y) = \Phi^H_{(PS)} y(k_x, k_y). \tag{5.1.6}$$

In the same way, the data is projected on the dictionary during the decomposition in Section 5.1.2. An example of the real part of a dictionary function and the corresponding image domain is shown in Figure 5.1.5. In this case the matched filter of (5.1.6) was applied to one of its own functions.



(a) k-domain (real part)

(b) Image domain

Figure 5.1.5: Example of a point scattering dictionary function

5.1.1.2 Vertical and Horizontal Lines Dictionaries

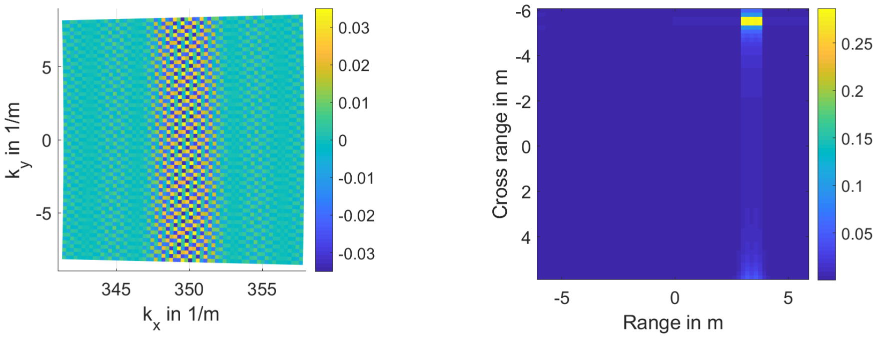
Since the different shapes are simulated as a collection of independent point scatterers, the dictionaries $\Phi_{(VL)}$ and $\Phi_{(HL)}$ of vertical and horizontal lines use a sum of several point atoms calculated by (5.1.5) as their basic atom structure. In case of the vertical lines sixteen point scatterers are used, while the horizontal lines only use five scattering centers. Therefore, the atoms $\varphi_i^{(HL)}$ and $\varphi_i^{(VL)}$ are calculated as

$$\varphi_i^{(HL)} = \sum_{\nu=0}^{4} e^{-j(k_x(x_i+\nu\Delta R)+k_y y_i)} \tag{5.1.7}$$

and

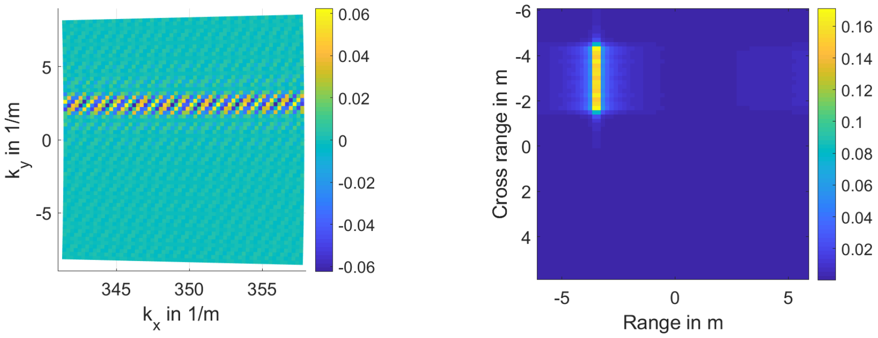
$$\varphi_i^{(VL)} = \sum_{\nu=0}^{15} e^{-j(k_x x_i+k_y(y_i+\nu\Delta CR))}. \tag{5.1.8}$$

To create the oversampled line dictionaries, the procedure is the same as for the point scatterer dictionary, i.e. the staring point of the shape is shifted by a fraction of a resolution cell according to the scheme of Figure 5.1.4. The number of atoms in the line dictionaries is smaller than in $\Phi_{(PS)}$, since the atom is only calculated if the whole shape fits in the image. This means that there are no starting points of the vertical lines in the last fifteen rows and no starting points of horizontal lines in the last four columns of the discretized scene. Examples of both dictionaries are depicted in Figures 5.1.6 and 5.1.7. The differences in the numerical values between the different dictionaries is a result of the normalization of the ℓ_2 norm of the atoms.



(a) k-domain (real part) (b) Image domain

Figure 5.1.6: Example of a horizontal line dictionary function

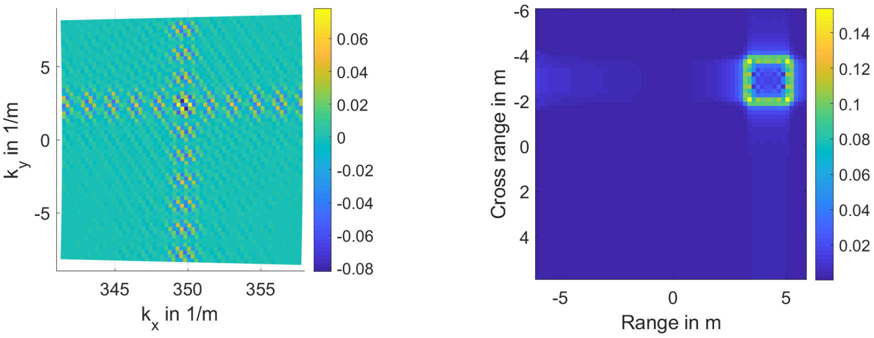


(a) k-domain (real part) (b) Image domain

Figure 5.1.7: Example of a vertical line dictionary function

5.1.1.3 Squares and Crosses Dictionaries

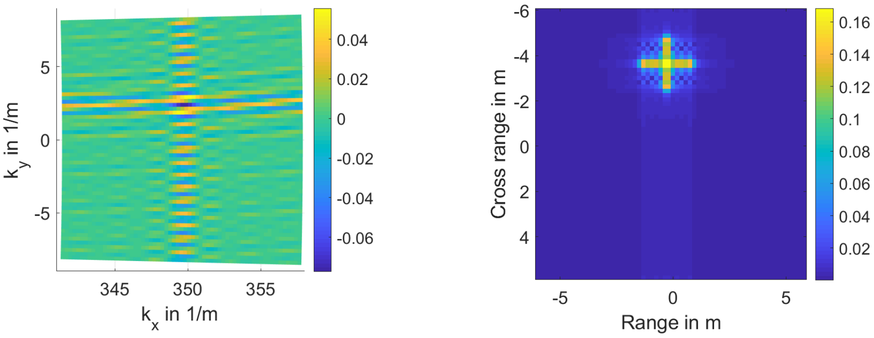
The dictionaries for crosses and squares are calculated by adding the models of horizontal and vertical lines. For squares the echo of two vertical and two horizontal lines are added and for crosses the echoes of one horizontal and two vertical lines are added. The split of one of both lines of the cross into two parts is necessary to avoid a scattering center with the doubled amplitude in the center of the cross. In Figures 5.1.8 and 5.1.9 example atoms of both dictionaries are shown.



(a) k-domain (real part)

(b) Image domain

Figure 5.1.8: Example of a squares dictionary function



(a) k-domain (real part)

(b) Image domain

Figure 5.1.9: Example of a crosses dictionary function

5.1.1.4 Circle Dictionary

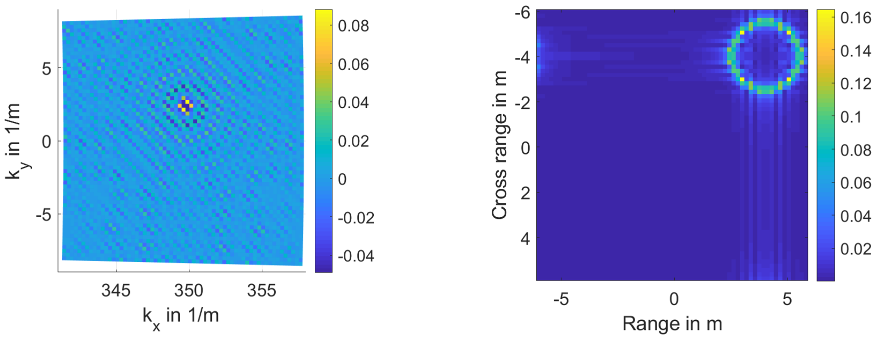
In case of the circle dictionary, the collection of point scatterers depends on the distance of (x_i, y_i) to the center of the circle. The k-domain data for this scattering shape is a conditional sum over the set Ω_i based on the distance of the grid point $(\tilde{x}, \tilde{y})$ to the center of the circle. The radius r_{circle} is given as a multiple of the resolution cell size ΔR and it is assumed that range and cross-range posses the same resolution, which is the case in this simulation. Therefore, the set of points belonging to the circle around (x_i, y_i) is given as

$$\Omega_i = \left\{ (\tilde{x}, \tilde{y}) \;\middle|\; \left[\frac{\sqrt{(\tilde{x} - x_i)^2 + (\tilde{y} - y_i)^2}}{\Delta R} \right] = \frac{r_{circle}}{\Delta R} \right\}, \tag{5.1.9}$$

where $[\cdot]$ is the *rounding* operator. The k-domain data of the circle atom $\varphi_i^{(Circle)}$ is calculated as

$$\varphi_i^{(Circle)} = \sum_{\Omega_i} e^{-j(k_x \tilde{x} + k_y \tilde{y})}. \tag{5.1.10}$$

To create the oversampled dictionaries, the complete set Ω_i is shifted by the appropriate size in each dimension, e.g. adding half of a resolution cell to each x-value to create a circle centered between two range cells. An example of this dictionary is depicted in Figure 5.1.10.



(a) k-domain (real part)

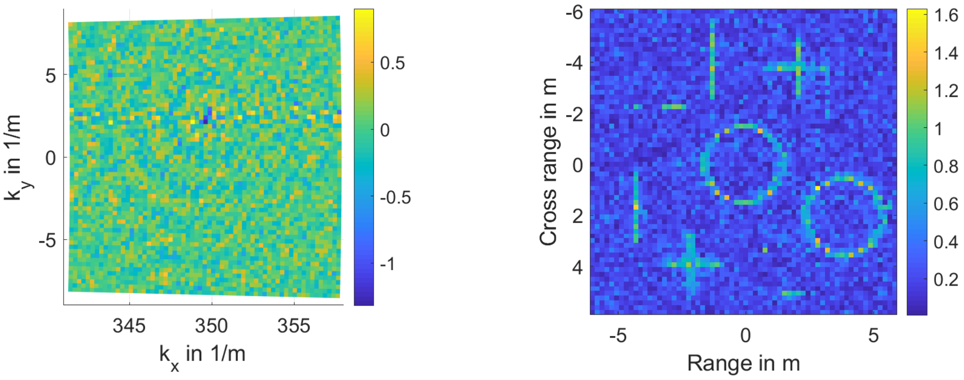
(b) Image domain

Figure 5.1.10: Example of a circle dictionary function

5.1.2 Results

The aim of this simulation is the decomposition of the data shown in Figure 5.1.3 using the MCA algorithm of Section 4.3.1 with the dictionaries presented in Section 5.1.1. In this simulation, the main parameters of the algorithm are an exponential decrease of the threshold T within 100 iterations, which is the same as in the example of Section 4.3.2. The minimal threshold is calculated as three times the noise standard deviation, which is known in this simulation. The chosen lower thresholding bound was chosen to avoid noise in the output components of the algorithm. However, for high SNR values, the numerical value will be very low and the lower bound must be adapted in these cases. This phenomenon of a erroneous reconstruction for signals with low noise will be discussed below. Further, the scene is simulated first on the grid and second with a random shift in the range of ±0.5 range cells for each object, i.e. each object is shifted by a certain amount within one range cell. Three different dictionary sizes are utilized, i.e. without oversampling and with oversampling factors of two and four. This gives 6 different combinations for the simulation, which is performed for each combination with fifteen different SNR values from -10 to 20 dB in steps of 2 dB. Therefore, only some exemplary results can be shown here.

The first example shows the intermediate steps of the decomposition of the off-grid scene using dictionaries with an oversampling factor of two and an SNR of 0 dB. The calculated raw data and the matched filter imaging result are shown in Figure 5.1.11.

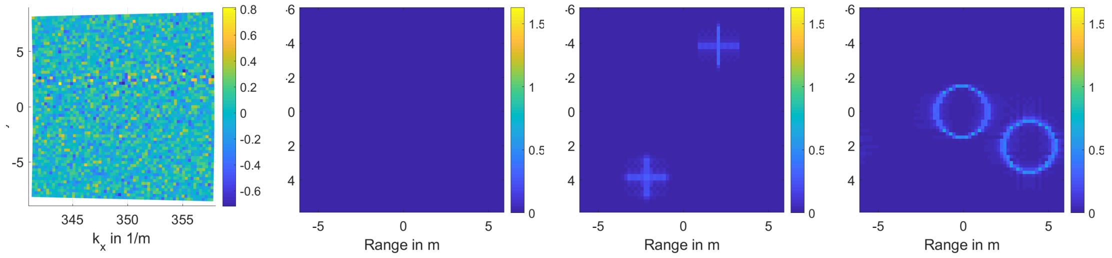


(a) Simulated k-domain (b) Matched filter image

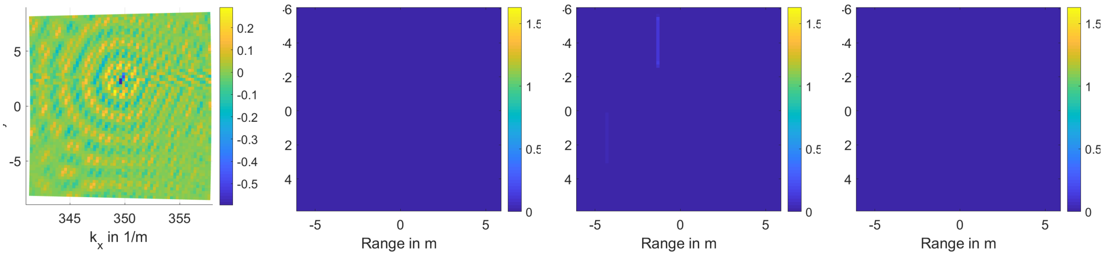
Figure 5.1.11: Input of the MCA decomposition and corresponding image

The decomposition is shown from Figure 5.1.12 to 5.1.14 for 20, 60 and 100 iterations,respectively. This visualization is comparable to the presentation in Section 4.3.2 with the difference that the residual in plot(a) and the component sum in plot (e) are now in the complex-valued k-domain. Therefore, only the real part is depicted for these two plots.

This example shows already that this decomposition has a strong denoising effect and that the images of the individual shapes contain mainly the corresponding shapes from the simulation. The image related to the squares dictionary does not show any significant content. Although the amplitude of the scattering centers is not reconstructed perfectly, the result clearly indicates the recognition of the different shapes. The difference in the amplitude of the, for example, cross shape scatterers is probably due to the off-grid simulation. The cross in the upper part of the image is more clearly visible than the cross in the lower part, this indicates that the physical position of the upper

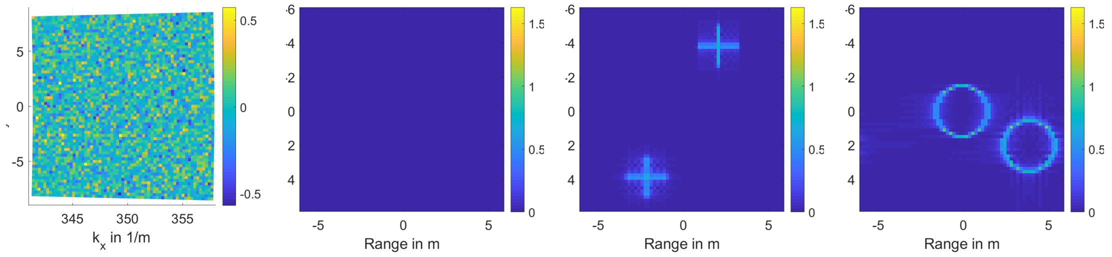


(a) Residual (b) Dots (c) Crosses (d) Circles

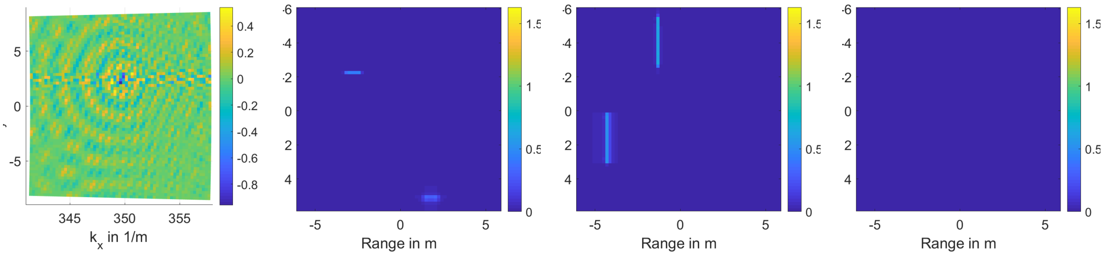


(e) Components sum (f) Horizontal lines (g) Vertical lines (h) Squares

Figure 5.1.12: MCA of radar simulation after 20 Iterations. Residual and component sum in k-domain, the components itself are plotted in the image domain.

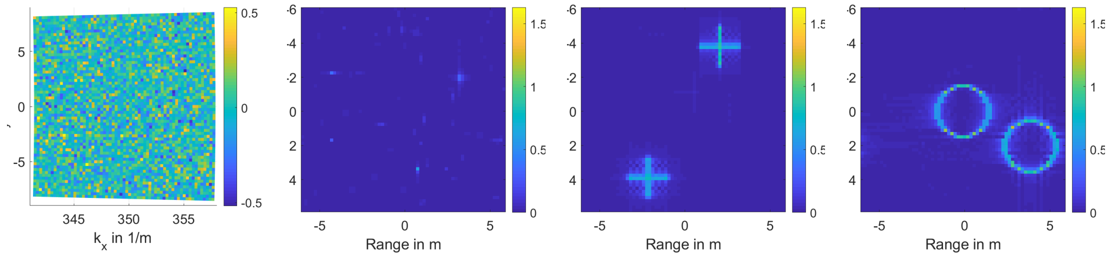


(a) Residual (b) Dots (c) Crosses (d) Circles

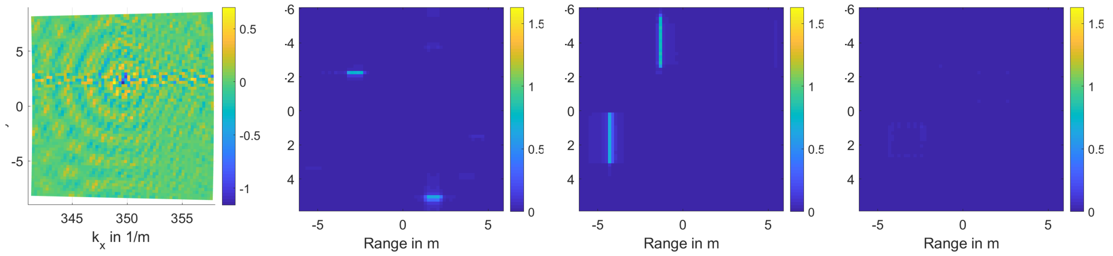


(e) Components sum (f) Horizontal lines (g) Vertical lines (h) Squares

Figure 5.1.13: MCA of radar simulation after 60 Iterations. Residual and component sum in k-domain, the components itself are plotted in the image domain.



(a) Residual (b) Dots (c) Crosses (d) Circles



(e) Components sum (f) Horizontal lines (g) Vertical lines (h) Squares

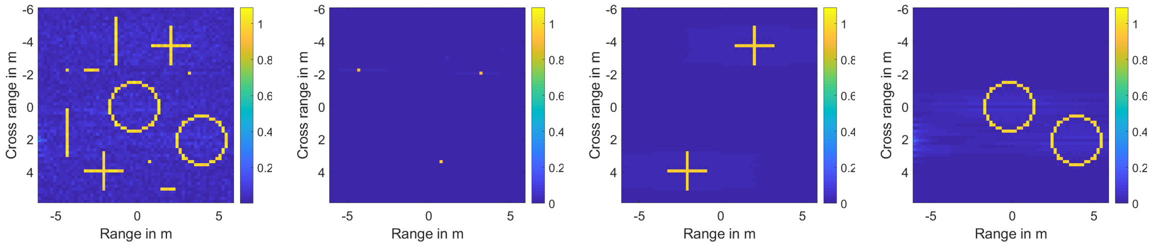
Figure 5.1.14: MCA of radar simulation after 100 Iterations. Residual and component sum in k-domain, the components itself are plotted in the image domain.

cross was closer to a dictionary function, i.e. the x_m and y_n in (5.1.2), than the lower one. The denoising effect is also visible in the k-domain data of the decomposition. A comparison of the simulated raw data in Figure 5.1.11 (a) and the component sum in Figure 5.1.14 (e) clearly shows less noise in the k-domain data after the decomposition. Furthermore, the underlying structure is not visible anymore in the final residual in Figure 5.1.14 (a), which seems to contain only noise. The detected objects are of the same size as the objects in the input data because the atoms are matched to the data.

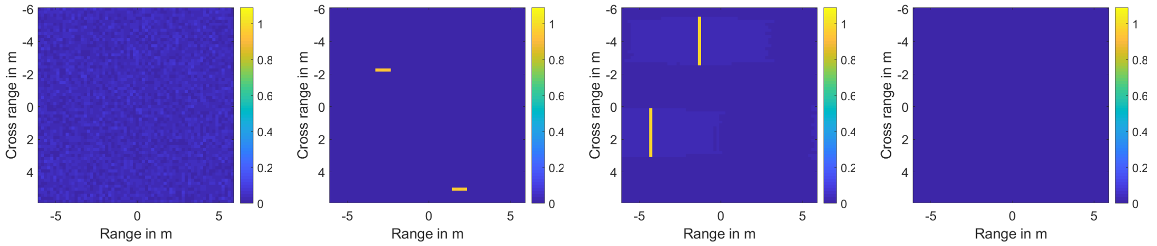
Since it is impractical to show the decomposition in this way for all mentioned combinations, the final decomposition result is shown for two more examples in Figures 5.1.15 and 5.1.16. The example in Figure 5.1.15 is simulated with the targets on the grid, without dictionary oversampling and with an SNR of 8 db. The results in Figure 5.1.16 are obtained for an off-grid simulation with an dictionary oversampling factor of four and an SNR of 14 dB.

The decomposition results of Figure 5.1.15 show the almost ideal decomposition and denoising if the model is known perfectly. The input and the residual data are shown in both figures in the image domain to get a better impression of the decomposition and the residual is shown in a different dynamic, which can be seen from the colorbar. This was done to visualize that only noise is contained in the residual.

The results of Figure 5.1.16 are comparable to the results of Figure 5.1.14, since the data is also off-grid and oversampling is used. Compared to the visualization above,

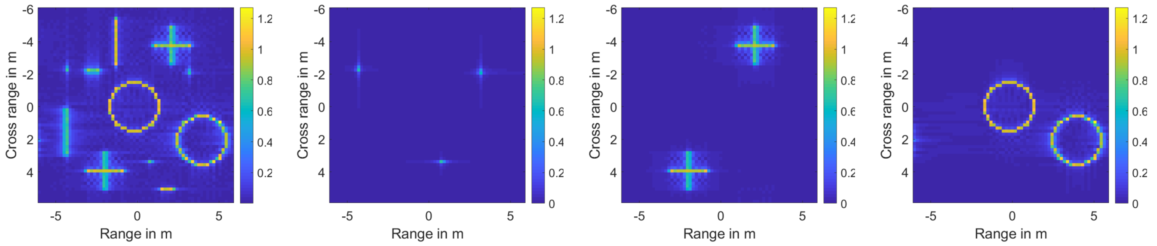


(a) Input (b) Dots (c) Crosses (d) Circles

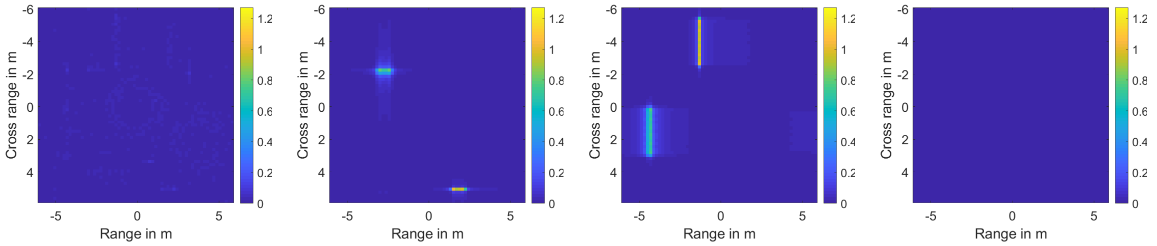


(e) Residual (f) Horizontal lines (g) Vertical lines (h) Squares

Figure 5.1.15: MCA of on-grid simulation with 8 dB SNR and without oversampling. All plots are shown in the image domain.



(a) Input (b) Dots (c) Crosses (d) Circles

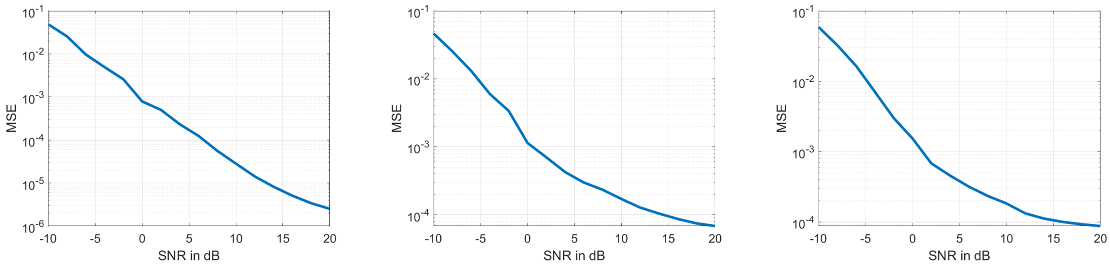


(e) Residual (f) Horizontal lines (g) Vertical lines (h) Squares

Figure 5.1.16: MCA of off-grid simulation with 14 dB SNR and fourfold dictionary oversampling. All plots are shown in the image domain.

the residual in the image domain shows that it still contains some image content. Nevertheless, the different shapes have all been recognized by the specific dictionaries and the dictionary of squares did not create confusion between the different components.

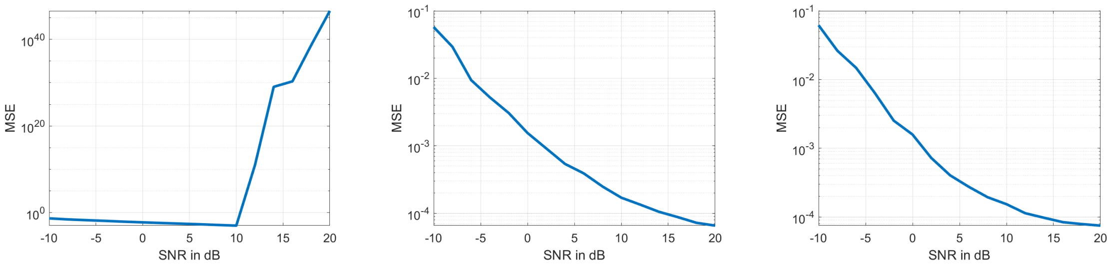
For a more compact presentation of the remaining results, the mean square error (MSE) of the reconstructed k-domain data compared to the stored reference values is depicted in Figure 5.1.17 for the different oversampling factors with on-grid and off-grid targets. For this figure, the sum of all reconstructed components is compared to the sum of the reference data in k-domain. The upper row shows the results for the on-grid simulation of the targets and the lower row shows the results of the off-grid simulations. Each plot shows the results over the SNR range from $-10\,\mathrm{dB}$ to $20\,\mathrm{dB}$.



(a) No oversampling, on-grid

(b) Twofold oversampling, on-grid

(c) Fourfold oversampling, on-grid



(d) No oversampling, off-grid

(e) Twofold oversampling, off-grid

(f) Fourfold oversampling, off-grid

Figure 5.1.17: MSE in the raw data domain over SNR

The results in the upper row of Figure 5.1.17 show that the use of oversampling leads to worse results if the model is perfectly matched to the targets on the grid. This effect is probably due to the increased coherence in the dictionaries if oversampling is used. In this case, it happens that several neighboring atoms are selected, although one of them perfectly matches the target in the scene.

The results in the lower row of Figure 5.1.17 on the other hand show the benefit of oversampling in the more realistic case of off-grid targets. From Figure 5.1.17 (d) can be seen that without oversampling, for this choice of the thresholding bound, the algorithm diverges for an SNR of more than 10 dB. This divergence does not happen with the oversampled dictionaries and the overall best result is achieved with an oversampling factor of two. Although the increase of the MSE with the SR4 dictionary is small, which is probably due to the effect that more neighboring atoms are selected in this dictionary.

The divergence without oversampling can be explained by the selection of neighboring on-grid atoms for a target between these grid points. These on-grid atoms are selected at the same iteration with a comparable magnitude, which leads to a feedback in the residual and thus to an exponential increase of the magnitude of these atoms in the decomposition. Alternatively to the use of oversampling would be an increase of the minimal threshold in the algorithm to avoid the divergence effect. This change would avoid the selection of the neighboring atoms at a certain iteration, which happened to be one of the last in these experiments. The experimental determination of a minimal threshold is part of the simulations in the next section.

5.2 Jet Engine Recognition in Simulated ISAR Images

The application in this and the following section is related to the preparation of a list of potential targets in an air surveillance scenario. Therefore, the aim is not the identification of a specific target, but to find features that can be used to limit the number of possible targets. The feature that is chosen here is the presence of jet engines in an ISAR image of a potential target. To find these engines, the model of Chapter 3 is used together with a standard point scatterer model to design two distinct subdictionaries and apply the MCA for the decomposition of the target. The decomposition is performed in the k-domain, but results are shown only in the image domain for better interpretability.

5.2.1 Dictionary Design

For the application here, two subdictionaries must be designed. The first one is the same point scattering dictionary as it was described in Section 5.1.1.1 above. This dictionary is used to describe the standard isotropic nondispersive scattering mechanisms in the image.

The second dictionary is the waveguide dictionary that uses model (3.3.4), which was presented in Chapter 3. With this model, a dictionary atom at position (x_i, y_i) is given by

$$\varphi_i^{(WG)} = e^{-j(k_x x_i + k_y y_i)} \left(\frac{1}{\zeta} + \sum_{m=1}^{M} e^{-jk_x \Lambda^{(m)} L} \right), \tag{5.2.1}$$

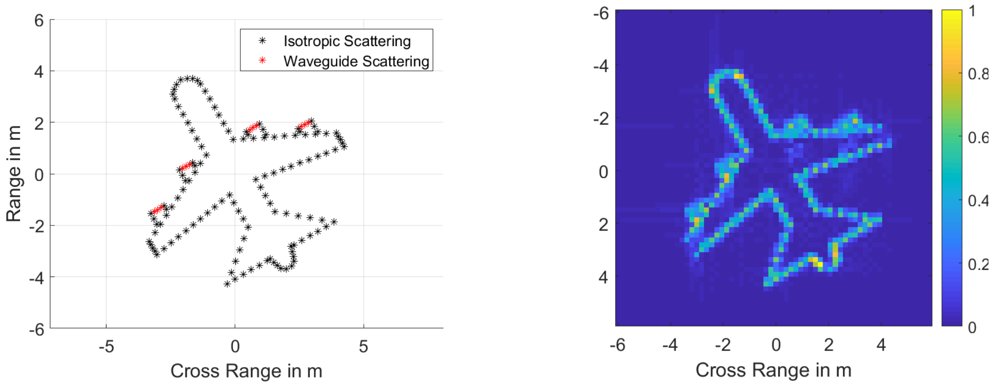
which can be interpreted as a point scattering atom that is corrected by the term in brackets.

The variable $\Lambda^{(m)}$ is a dimensionless factor describing the dispersion for each mode and L is the length of the waveguide. In the correcting term of the waveguide model only the wavenumber in range direction k_x appears, since the dispersion shows only effects in range dimension. The variable ζ balances the magnitude of the point scattering effect at the entrance of the engine and the waveguide scattering. The value of ζ was set to achieve the same peak value in the point scattering and the waveguide scattering of the atom. M is the number of propagating modes inside the engine. Both dictionaries are implemented with an oversampling factor of two.

Since the PSF of this waveguide model produces a ponytail-like shape in the ISAR image, atoms are only calculated for positions that allow that this ponytail is visible in the resulting image. Therefore an offset of 15 range cells is used within the simulations in this section, i.e. within the last 15 range cells of the image no waveguide scattering can be detected. This number was chosen arbitrarily by a visual comparison in an appropriate dynamic range of the jet engine PSF. In principle, the PSF contains scattering centers to the end of the image independent of its position.

5.2.2 Simulation Results

To evaluate whether the MCA algorithm is able to separate these two dictionaries, a simulation with the presented models is done. For this simulation, an aircraft is modeled as a mixture of point targets and waveguide scattering at the opening of the jet engines. The arrangement of the scattering centers of the model together with the ISAR image is shown in Figure 5.2.1. The given physical dimension of the aircraft follow from the bandwidth and number of range cells used in this simulation and are not supposed to be realistic for this kind of target.

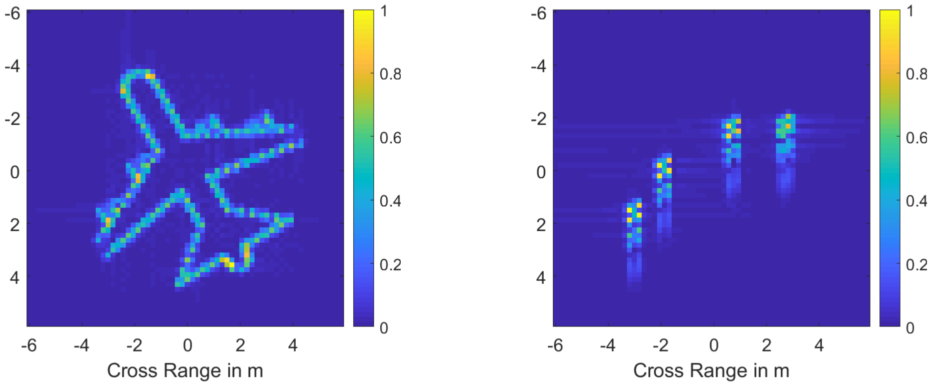


(a) 2D target model (b) Image domain

Figure 5.2.1: Aircraft target model and ISAR image

The red markers in Figure 5.2.1 (a) show the position of the waveguide scattering centers, but it should be mentioned that at these positions, a further point scattering center is added to the model. Therefore, the magnitude at these positions is comparably high and the scattering centers are much closer together along the entrance of the jet engine. These actions are done to get closer to the common assumption that the engines have a very high impact on the full RCS of the target [FDST+15].

To measure the quality of the decomposition result of this simulation, reference images for point and waveguide scatterers are computed for this aircraft model. This method was also used in the simulations of Section 5.1 and two exemplary reference images are shown in Figure 5.2.2 for an aspect angle of 30°.



(a) Point scatterer (b) Waveguide scatterer

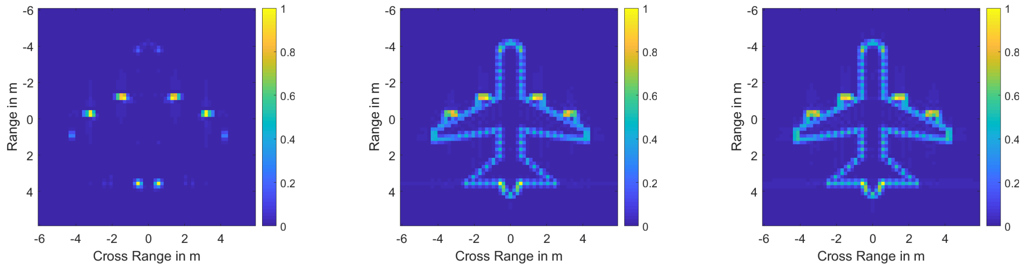
Figure 5.2.2: Reference images to evaluate simulation results

The images in Figure 5.2.1 and 5.2.2 are processed with the standard tomographic model of (2.2.8) and a Taylor weighting in the raw data domain. This means that the image is produced via a matrix-vector multiplication using the point scatterer dictionary without oversampling and the vectorized form of the raw data. The Taylor weighting is applied before the raw data is stacked into a vector via a multiplication with a two-dimensional Taylor window. Since the simulation below is performed with four different aspect angles, the aspect angle of the reference images are changed accordingly.

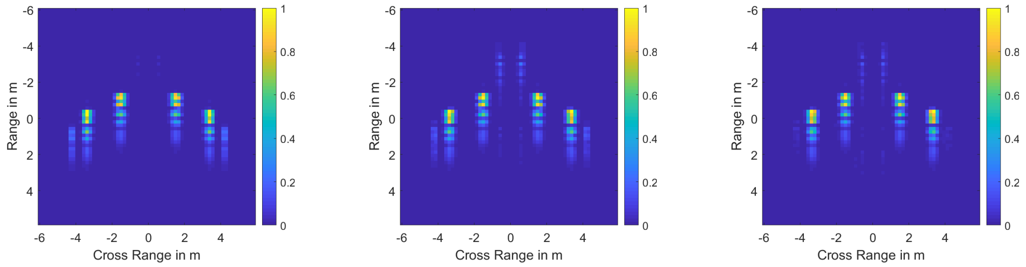
5.2.2.1 Simulation with Matching Models

The first simulation that is shown here uses the same model parameters for both dictionary and aircraft model, i.e the model is perfectly matched to the target. This experiment is done with four different aspect angles, varying from 0° to 60° in steps of 15°. Each of these four scenarios is simulated with sixteen different SNR values from $-10\,\text{dB}$ to $20\,\text{dB}$ in steps of $2\,\text{dB}$, which are the same SNRs as for the simulation in Section 5.1.2. The presentation of the decomposition results is also comparable to the presentation in Section 5.1.2, i.e. two examples with results from intermediate iterations are shown and the remainder is summarized in plots of the MSE of the recovered raw date over the SNR. It should be mentioned that independent of the matching between model and target, the problem is still an off-grid problem due to the off-grid positions of the scattering centers. Therefore, the inverse crime that was mentioned in Section 5.1 is already avoided here and also below in the results using the mismatched models.

The first results, given as scattering components in the image domain, are shown in Figure 5.2.3 with an SNR of 12 dB and for an aspect angle of 0°. In this figure, the final result after 100 iterations is depicted with two intermediate results after 20 and 60 iterations.



(a) Point targets, 20 iterations (b) Point targets, 60 iterations (c) Point targets, 100 iterations



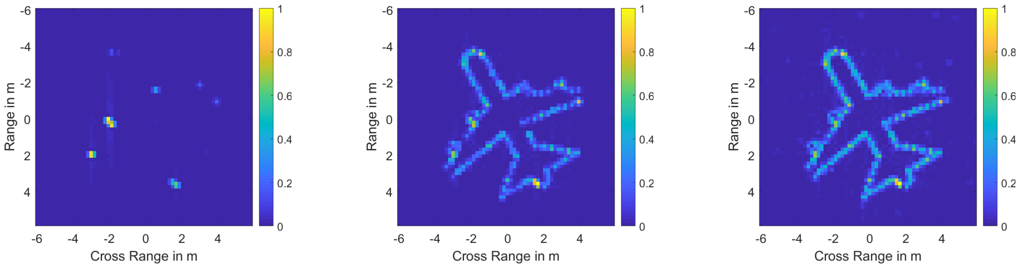
(d) Waveguides, 20 iterations (e) Waveguides, 60 iterations (f) Waveguides, 100 iterations

Figure 5.2.3: Simulation results at 0° aspect and 12 dB SNR

The results in Figure 5.2.3 show that the engine scattering mechanism can be separated, but some confusion with isotropic scattering along straight lines in range direction appears. However, it is also visible that the false alarms at the left and right wingtips are reduced from the center image at 60 iterations to the final result. This shows the correcting behavior of the algorithm and for this SNR level, the target is clearly visible in the point scattering image.

The MCA parameters of this and the following simulation are the same as in Section 5.1.2, i.e. an exponential decay of the threshold and the shrinkage operation. The depicted examples are created with a minimal threshold of three times the noise standard deviation. This definition is possible, since the exact noise standard deviation is known in this simulation.

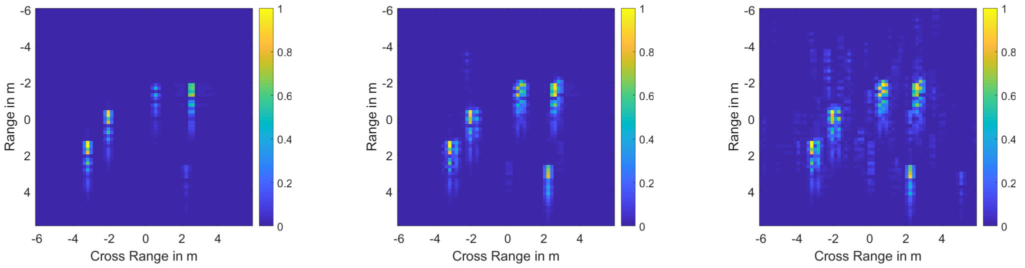
The second example with matching dictionary and target models is shown in Figure 5.2.4. It is simulated with an aspect angle of 30° and an SNR of 2 dB, which is a more challenging setting than the first example.



(a) Point targets, 20 iterations

(b) Point targets, 60 iterations

(c) Point targets, 100 iterations



(d) Waveguides, 20 iterations

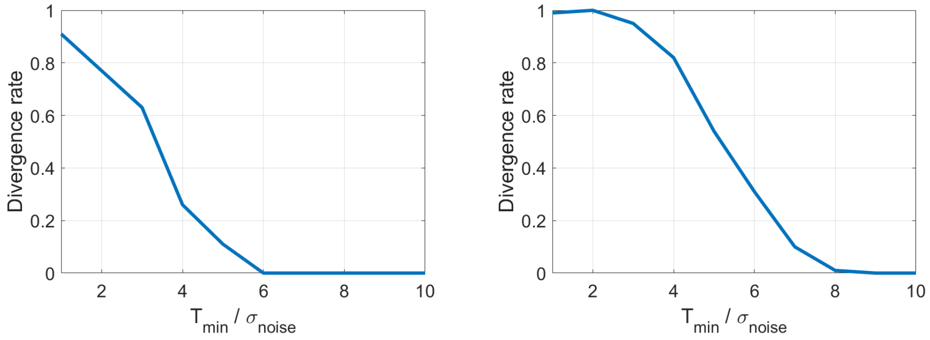
(e) Waveguides, 60 iterations

(f) Waveguides, 100 iterations

Figure 5.2.4: Simulation results at 30° aspect and 2 dB SNR

It is visible that the higher amount of noise in Figure 5.2.4 leads to a larger amount of false alarms compared to the previous example. Further, the magnitude of some of the false alarms is comparable to the magnitude of the correct positions. One option to reduce the number of false alarms, especially in the waveguide image, is the use of a higher value for the minimal threshold. This can be seen by comparing Figures 5.2.4 (e) and 5.2.4 (f), which show the waveguide images at 60 and 100 iterations. In the intermediate result, which belongs to a larger threshold, the number of false alarms is smaller than in the final result. Nevertheless, the high magnitude false alarm will not be effected from this change and already shows a limitation of this application in low SNR scenarios.

Another reason to increase the minimal threshold is observed in high SNR simulations with aspect angles of 15° and 30°. In these cases, the algorithm becomes numerically unstable towards the end of the decomposition and diverges. The same effect appeared already in the simulations of Section 5.1.2 for off-grid targets with high SNR and without oversampling in the dictionaries. To assess this divergence problem, the 20 dB SNR simulations with 15° and 30° aspect angle are repeated with different minimal thresholding values. This bound is varied from one to ten times the noise standard deviation and a divergence rate is calculated as the ratio of the trials that failed and the total number of trials, which is 100 for each minimal thresholding bound. The results of this simulation are summarized in Figure 5.2.5. A trial is counted as divergence trial if the MSE of the reconstruction is at the end at least ten times higher than the minimum MSE value during the decomposition.

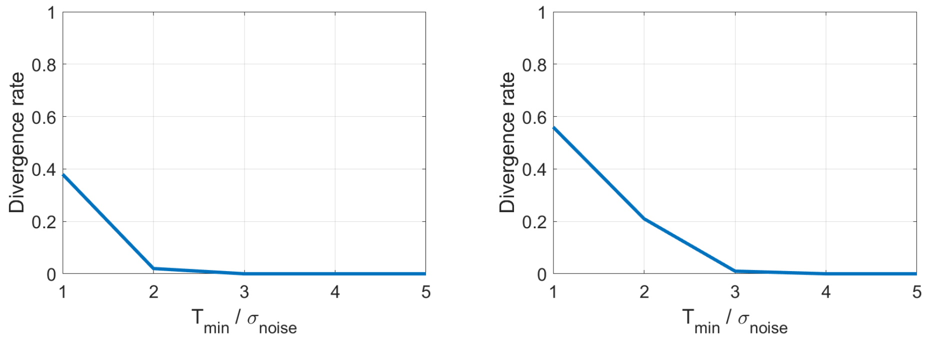


(a) 15° aspect angle (b) 30° aspect angle

Figure 5.2.5: Divergence rate of the 15° and 30° aspect angle simulation over the lower thresholding bound

The plots in Figure 5.2.5 show that the problem is present for both aspect angles, but in case of the 30° aspect angle, a minimal threshold value of nine times the noise standard deviation is necessary to achieve a convergence rate of 100%, while in the other case a factor of six is enough. For comparison with the 0° and 45° cases, the analysis is repeated with the lower bound varied from one to five times the noise standard deviation. The results of this simulation are shown in Figure 5.2.6 and show that the divergence rate is significantly lower in these cases. In case of 0° aspect angle, only two out of 100 trials diverged for a minimal threshold of two times the noise standard deviation and for a factor of three or larger, all trials converged. In case of 45° aspect angle, one out of 100 trials failed with a lower thresholding bound of three times the noise standard deviation and with a higher minimal threshold all trials converged.

Because of the convergence problems, the experiment using the varying SNR values is repeated for the aspect angles of 15° and 30° with a minimal threshold of six and nine times the noise standard deviation, respectively. Results of this simulation are shown together with the results of the 0° and 45° aspect angle simulation in Figure 5.2.7. The difference in the minimal threshold has the highest influence in the simulations using a low SNR. The flat segment in this part of the curve indicates that no reconstruction is



(a) 0° aspect angle (b) 45° aspect angle

Figure 5.2.6: Divergence rate of the 0° and 45° aspect angle simulation over the lower thresholding bound

done and therefore, no scatterers are detected at all. In the regions of relatively high SNR, the MSE is comparable between the different aspect angles, although the MSE reached for the aspect angles 15° and 30° is of course higher than the ones using the lower minimal threshold.

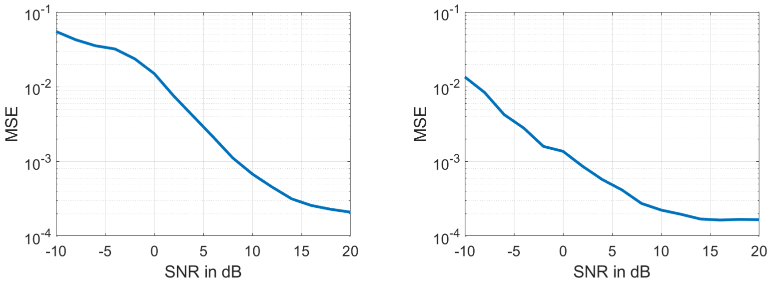
The use of a higher value for the minimal threshold can thus be summarized as follows: The number of false alarms is decreased with the cost that the waveguide scattering centers are less accurately reconstructed in case of a sufficient SNR.

The simulation in this section was limited in the sense that data and model matched perfectly, even if the scattering centers are placed at off-grid positions. This situation is not realistic in an operational scenario with unknown targets and therefore, the next step is an analysis of the sensitivity to the model parameters, i.e. the radius and length of the modeled waveguide. To quantify the dependency on these parameters, a simulation with a mismatch between them is presented next.

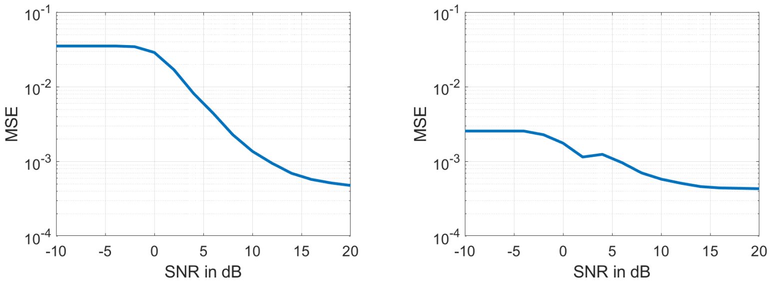
5.2.2.2 Simulation with Mismatched Models

For the simulation in this part, the same target model as above is used with an aspect angle of 15°. The SNR is set to 10 dB and the soft thresholding operator is used with an exponential decrease of the threshold. The minimal threshold is set to three times the noise standard deviation, which is possible since the convergence problems described above appeared only for higher SNR values. The parameters of the dictionary are set to a radius of 0.45 m and 0.4 m for the length. The actual parameters that are simulated in the scene are varied in steps of 0.01 m from 0.4 m to 0.5 m in case of the radius and from 0.35 m to 0.45 m for the length. Both dictionaries, the isotropic and the waveguide dictionary, are implemented with an oversampling factor of two.

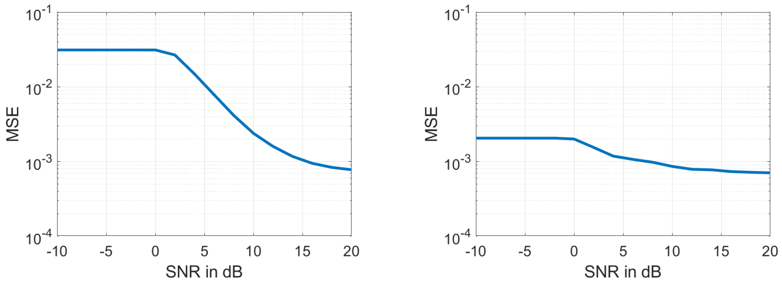
The main results are shown in Figure 5.2.8 as a plot of the MSE of the waveguide component. The plot of the isotropic scattering reconstruction is comparable and is



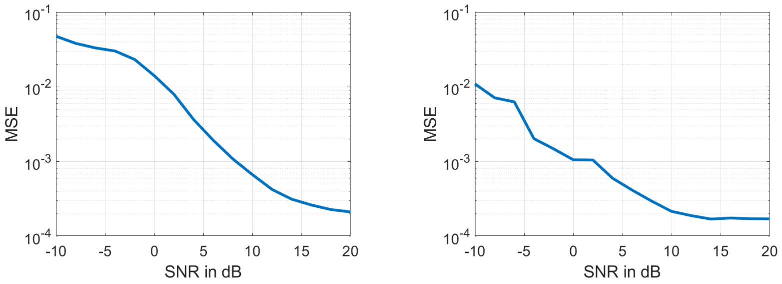
(a) 0° aspect, point targets (b) 0° aspect, waveguides



(c) 15° aspect, point targets (d) 15° aspect, waveguides



(e) 30° aspect, point targets (f) 30° aspect, waveguides



(g) 45° aspect, point targets (h) 45° aspect, waveguides

Figure 5.2.7: MSE in the raw data domain over SNR

not shown here.

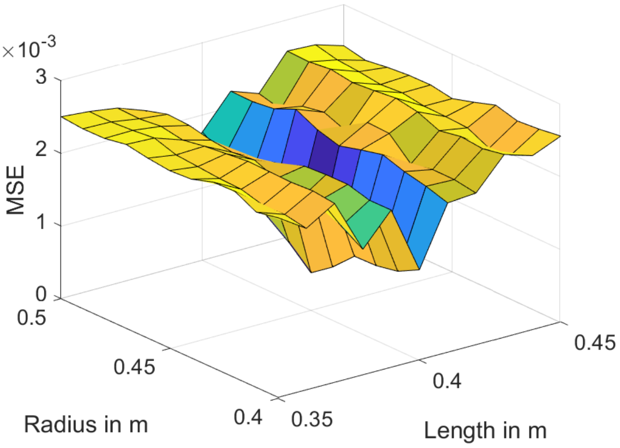


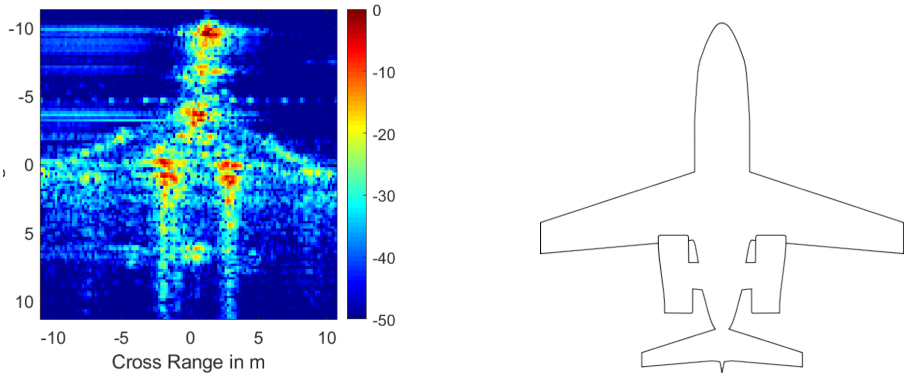
Figure 5.2.8: MSE of the waveguide component with a mismatch between model and simulation

The surface of the plot in Figure 5.2.8 shows that the length has a much stronger effect than the radius of the model. If the dictionary parameters match the model parameters in the scene, the minimum value is reached and a small variation of 0.1 m in the length of the model gives a significant increase in the MSE. If the length of the model is matched to the scene, the mismatch of the radius also leads to an increase of the MSE, but not as steep as in case of a mismatch in length. In fact, the increase of the MSE in direction of the length is 5.98 and 7.57 times larger than in the direction of the radius, depending on the direction of variation.

This simulation shows that is very important that the model is matched to the real engine, otherwise a proper reconstruction is not possible. This is highly relevant for the application to real data in the next section, since the true parameters of the target in the real ISAR image are not known. It cannot be said at this point whether an identification of different engines with this decomposition is possible. Even if different parameters lead to different results in the reconstruction, the absolute amount of the engine scattering is not known in practice and thus a criteria like the MSE cannot be used for classification. For the application to real data, a heuristic will be used to find the best matching parameters. This heuristic and the corresponding results are given in the following section.

5.3 Recognition of Jet Engines in Real ISAR Images

To verify the applicability of the algorithm to real data, ISAR images that were measured with the TIRA system of Fraunhofer FHR are used as input data for the decomposition. One example of these images is shown in Figure 5.3.1 (a) with a schematic drawing of the aircraft in Figure 5.3.1 (b).



(a) ISAR image (b) Shape of target

Figure 5.3.1: Real ISAR image used for MCA decomposition

The schematic drawing of the target shows two engines mounted at the fuselage of the aircraft with the openings above the wings. This constellation is also visible in the image, which shows two strong reflections at these positions and range extended returns from the engine to the end of the image. Therefore, the ideal output of the decomposition framework would show components of the waveguide dictionary only at these positions. The aircraft itself is a small business jet, but precise parameters of the engine are not known to the author.

Since the TIRA system uses the deramping process described in Section 2.1.3.2, scattering centers at different ranges are converted to constant frequencies. Therefore, the output of the deramping process is interpreted as data in k-space as it was also done in Section 2.1.3.2. It was also mentioned in Section 2.2.2 that the integration angle is not known exactly in an in-flight ISAR measurement and the given cross range scaling in the image is only an estimate of the true size of the target by the estimated aspect angle. This angle is also used to create the dictionary elements in k-space. The elevation angle is assumed to be constant during the measurement, since the distance to the target of approximately 160 km is relatively large compared to the target altitude of 7.5 km. Therefore, only the azimuth aspect angle is used in the model, which is $\sim 4.9°$ for the center angle in the image of Figure 5.3.1.

To calculate the atoms of the two subdictionaries, the parameters of the system are repeated here for convenience:

- Bandwidth $B = 800\,\text{MHz}$
- Pulse duration $\tau_p = 256\,\mu\text{s}$
- Carrier frequency $f_0 = 16.7\,\text{GHz}$
- Sampling frequency $f_s = 4\,\text{MHz}$
- 1024 samples per range profile

The last two parameters, i.e. the sampling frequency and the number of samples belong to the full covered range of the system. The image in Figure 5.3.1 (a) shows only a reduced amount of 121 range cells, which is enough to include the whole target and reduces the computational cost. To create the data to process the image, the full range profile is cut to the range cells that actually contain the target. An inverse Fourier transform of this reduced range profile creates the same output as a deramping process tailored to this reduced range with a reduced sampling rate $\tilde{f}_s$, which is given in this case as

$$\tilde{f}_s = f_s \cdot \frac{121}{1024}. \tag{5.3.1}$$

With these parameters, the dictionaries, which are the same as in the simulations of the previous section, can be calculated. Similar to the simulations before, both dictionaries are implemented with an oversampling factor of two. The main difference between the dictionaries used for simulations and the real data is the size of the dictionary. Using the real data, the dimension of both dictionaries is larger due to the larger image dimension. The decomposition itself is also done in k-space, although results are only shown in the image domain for better interpretability.

As Section 5.2.2.2 clearly demonstrated, the parameters of the model should be known as exactly as possible to achieve the best results of the decomposition. Since these parameters are not known, a grid search with several radius-length combinations was performed to find the maximum correlation between these model parameters and the image in Figure 5.3.1 (a). The search interval for the radius was set from 30 cm to 45 cm in steps of 0.5 cm and the length interval was from 40 cm to 70 cm, also in steps of 0.5 cm. For all combinations of this grid, a dictionary was designed and the maximum correlation coefficient for each of these dictionaries was stored and the maximum of all dictionaries determined the radius and length that were used in the decomposition. This leads to a radius of 0.39 m and a length of 0.675 m. These parameters seem to be realistic if they are compared to the spread of the engine scattering centers in the image considering the dimension of one range cell[3].

Beside the parameters of the model, the parameters of the algorithm must be determined. To determine the minimal threshold, the noise standard deviation is estimated

---

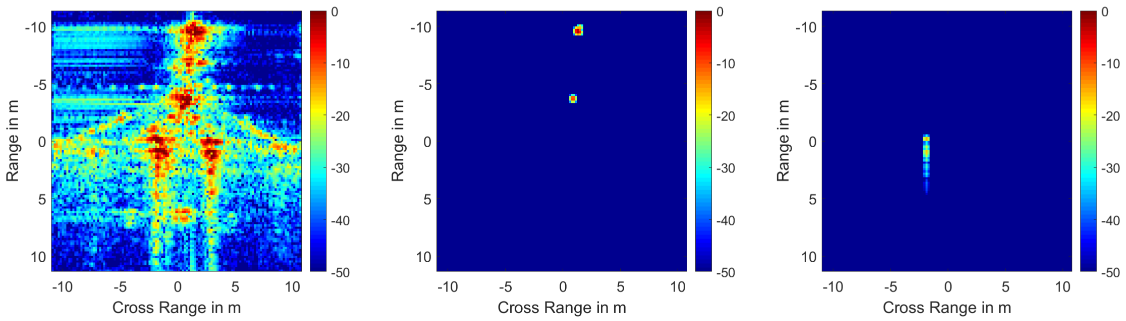
[3] The cross-range size of the range cells cannot be determined exactly due to the uncertainty in the estimated integration angle during the ISAR imaging process, see Section 2.2.2.

as the median absolute deviation (MAD) of the coefficients in the finest resolution band of an arbitrary wavelet. Here a Daubechies wavelet with a filter length of four is used. This estimation is based on the assumption that most of the image content is represented by the low frequency bands and that the finest resolution band contains only noise. The details of this method can be found for example in [SMF10]. However, since most of the undesired echos in the image in Figure 5.3.1 (a) are not related to noise, but to artifacts of the imaging process that also create coefficients in the lower frequency bands, the estimated noise standard deviation must be arbitrarily multiplied by 20 to achieve a satisfactory result. This factor was determined empirically by optical comparison of decomposition results with different lower bounds. With lower values for this factor, the number of false alarms starts to increase towards the end of the decomposition. However, even with this high value compared to the simulations before, more than 80% of the signal energy is covered by the decomposition, as will be shown below.

Similar to the minimal threshold value, the remaining parameters of the algorithm were also determined experimentally. The number of iterations is set to 400 and the hard thresholding was chosen in the iterative algorithm. The hard thresholding was beneficial, since the main goal was to find the position of the air intakes, which produce a rather large echo in the data. Therefore, the algorithm can detect these scattering centers in an early stage of the algorithm and the hard thresholding removes these strong scattering centers as soon as they are detected. With this choice, it is harder for the algorithm to correct its choice, but it was the superior method in this case. A condition for the hard thresholding is the isolated strong scattering center, which is fulfilled for this kind of aircraft with jet engines. If the algorithm is able to detect the engines if these are covered and the air intake is part of the fuselage, e.g. a fighter aircraft, cannot be said at this point. This situation might also require further improvements of the model, since the air intake will be several meters long and change its geometry from the intake to the engine.

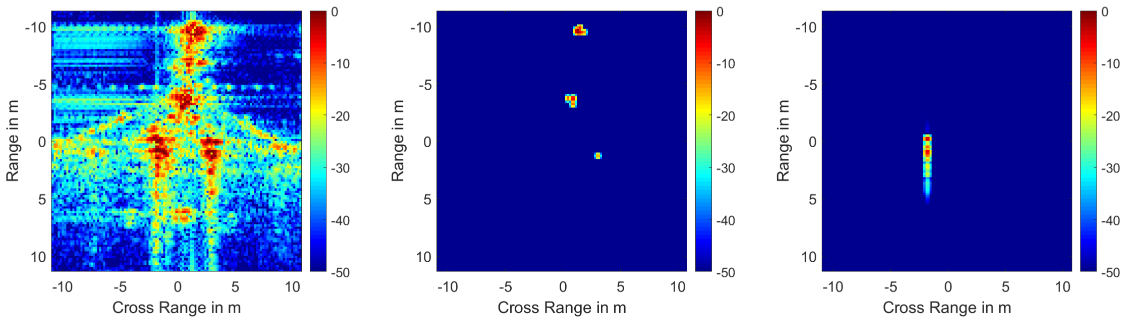
The decomposition of the example in Figure 5.3.1 (a) is shown from Figure 5.3.2 to Figure 5.3.6 from an early stage after 10 iterations to the final result with three intermediate steps after 50, 100 and 200 iterations. All of these figures are in the same dynamic range for a better comparison of the components.

The components in Figures 5.3.2 and 5.3.3 at the beginning of the algorithm contain only a few dominant scatterers and after 100 iterations, the added echoes in both components are in the same area of the image, which contains most of the energy. Nevertheless, the separation of the strong scattering centers into waveguides and point scattering is already visible in these early iterations of the algorithm. In the middle of the decomposition, at 200 iterations, some more small point scattering centers are added to this component, while the waveguide scattering is still concentrated in the area of the engine inlets. The errors in the decomposition towards the end of the algorithm indicate that the real data shows the same confusion between waveguides and straight lines in the image as it was observed in the simulations above. However,



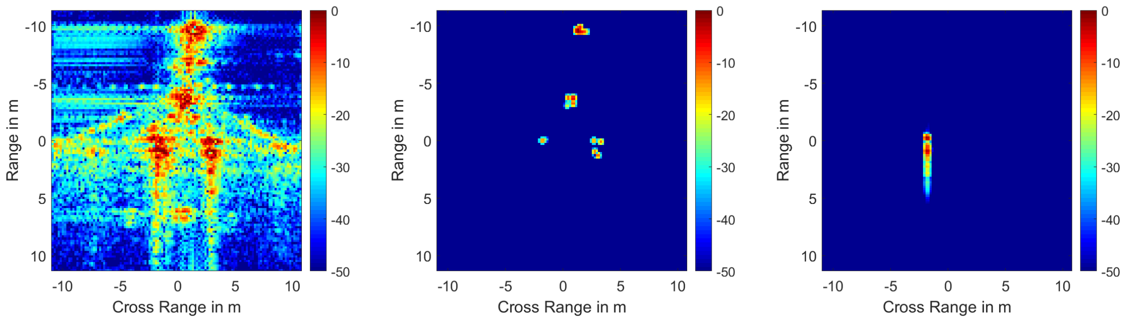
(a) Residual (b) Point scatterer (c) Waveguide scatterer

Figure 5.3.2: MCA of a real ISAR image after 10 Iterations



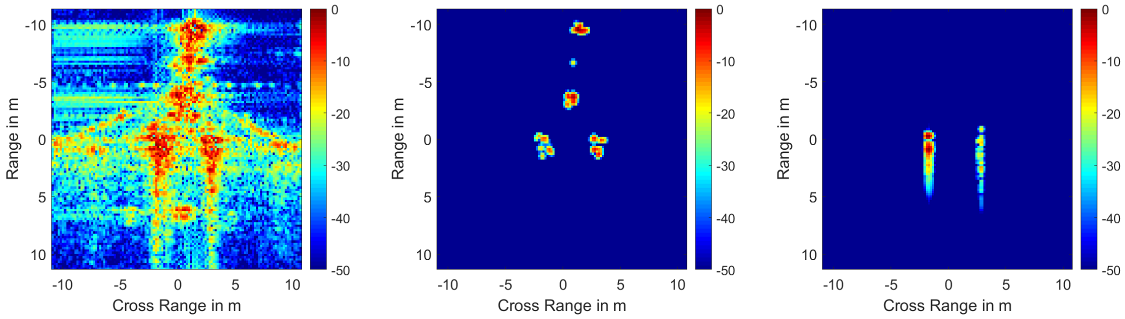
(a) Residual (b) Point scatterer (c) Waveguide scatterer

Figure 5.3.3: MCA of a real ISAR image after 50 Iterations



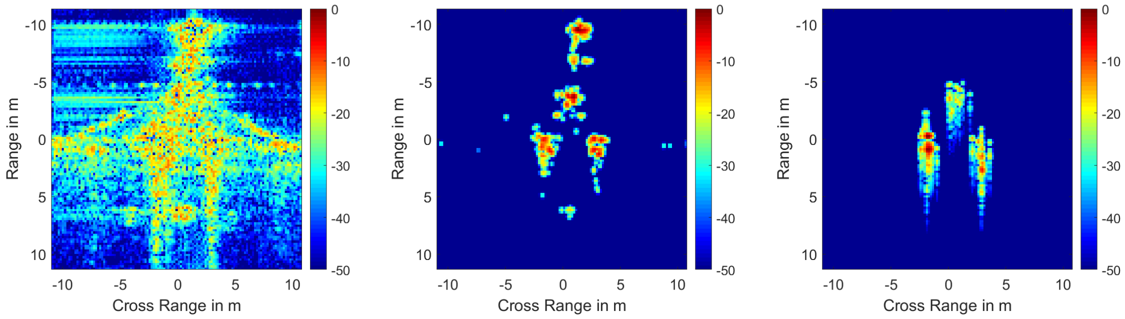
(a) Residual (b) Point scatterer (c) Waveguide scatterer

Figure 5.3.4: MCA of a real ISAR image after 100 Iterations



(a) Residual (b) Point scatterer (c) Waveguide scatterer

Figure 5.3.5: MCA of a real ISAR image after 200 Iterations



(a) Residual (b) Point scatterer (c) Waveguide scatterer

Figure 5.3.6: MCA of a real ISAR image after 400 Iterations

the results of this example show that it is possible to separate jet engines and isotropic scattering in ISAR images using this approach, although the results are not perfect.

To evaluate how much of the energy is covered by the selected atoms, the ℓ_2 norm of the vectorized residual during the decomposition is shown in Figure 5.3.7. The ℓ_2 norm corresponds to the square root of the signal energy and the dashed vertical lines mark the iterations that belong to the intermediate images shown from Figure 5.3.2 to 5.3.5.

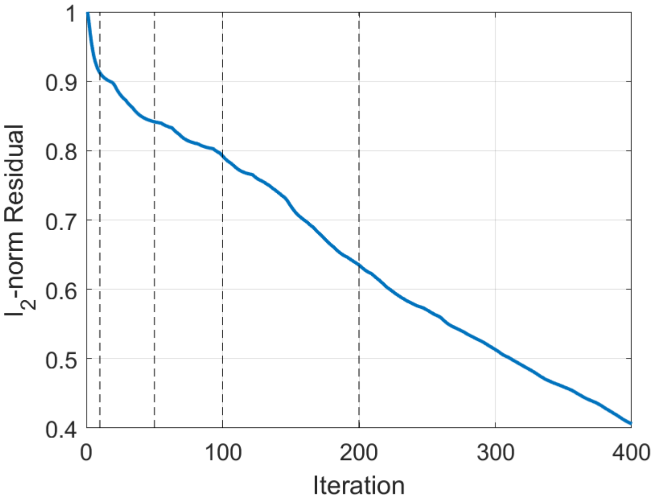
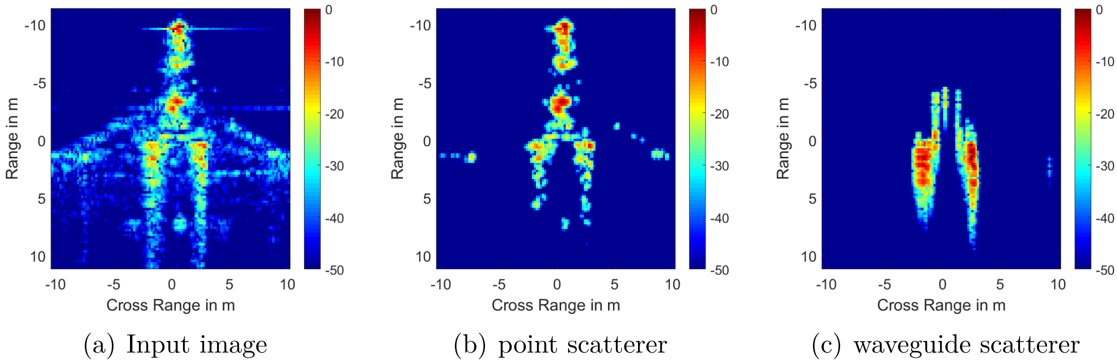


Figure 5.3.7: Norm of the residual during the ISAR decomposition

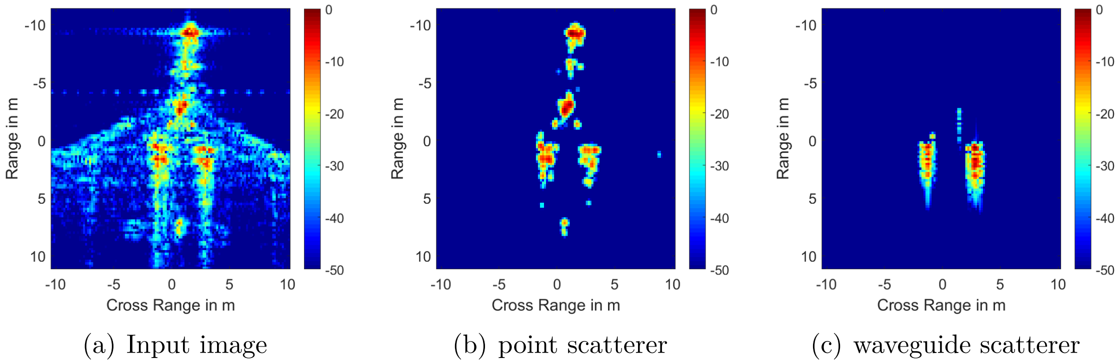
The curve in Figure 5.3.7 shows that the ℓ_2 norm of the residual is reduced by almost 60% after 400 iterations, which corresponds to almost 84% of the energy, although the amount of scattering centers is rather low in the component images. This is an indicator for the common assumption that the main energy is concentrated in a few scattering centers, which can be also seen from the marked iterations. After the first steep decrease within the first ten iterations 9.05% of the ℓ_2 norm is already covered by the decomposition. The next selection of scattering centers between ten and fifty iterations reduces the norm of the residual by 6.84% and from there on, the decrease of the ℓ_2 norm is approximately linear. After 100 iterations, the norm of the residual is reduced to 79.01% of the initial value and after 200 iterations to 63.35%.

Finally, the decomposition is applied to three more ISAR images of this target. These images are shown together with the final component images in Figure 5.3.8 to 5.3.10. The center azimuth aspect angle of the different images are $\sim 3°$, $\sim 6°$ and $\sim 10°$ respectively. The results are comparable to the example above, since the engines are clearly visible in the waveguide images, but in all images a small number of false alarms is visible. These results also indicate that the algorithm should be robust against cases where no range extended returns are visible in the image, since the number of false alarms is small. This situation would be comparable to the simulation in Section 5.1.2, where no target of the square shape was visible.



(a) Input image (b) point scatterer (c) waveguide scatterer

Figure 5.3.8: MCA of a real ISAR image (Example 2)

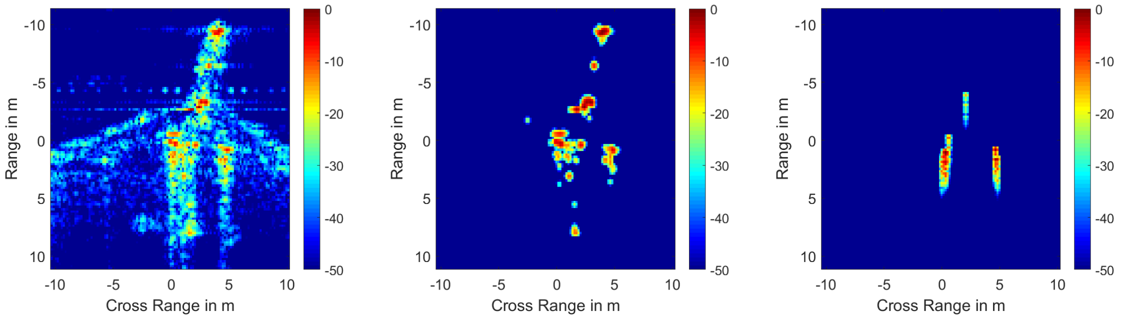


(a) Input image (b) point scatterer (c) waveguide scatterer

Figure 5.3.9: MCA of a real ISAR image (Example 3)

These additional examples differ in the amount of the signal energy that is covered by the algorithm. The component images in Figure 5.3.8 reduce the ℓ_2-norm of the residual to 26.06% of the initial value, while the residual of Figure 5.3.9 is reduced to 39.6%, which is comparable to the detailed example above. Finally, the ℓ_2 norm of the residual of the example in Figure 5.3.10 is reduced to 45.11% of its initial value.

It should be mentioned that the parameters of the algorithm are optimized for each of these examples. Although all of these results are achieved with hard thresholding and 400 iterations of the algorithm, the minimal threshold value is changed in the last example in Figure 5.3.10 to ten times the estimated noise standard deviation. This value is closer to the values of the simulations in Section 5.2 and is most likely due to the reduced imaging artifacts in this example.



(a) Input image (b) point scatterer (c) waveguide scatterer

Figure 5.3.10: MCA of a real ISAR image (Example 4)

Further to the algorithm parameters, the parameters of the model, i.e. radius and length, are optimized for each image with the same method as for the examples above. The parameters are chosen to maximize the ratio of the response of the waveguide model in the area of the engines to the central part of the aircraft. This optimization gives for the radius three different values of 0.42 m, 0.39 m and 0.3 m and for the length 0.53 m, 0.56 m and 0.47 m. These numbers show that radius and length of the model in the examples are not the same, but in the same order of size. The reduced radius and length in the example of Figure 5.3.10 might be also an effect due to the change of the aspect angle compared to the other images. This indicates that the model should be further improved in the future, but the results show that a decomposition of ISAR images is possible.

Chapter 6

Conclusion and Future Work

This thesis has leveraged previous work on jet engine models and sparse decomposition to study the ability to detect this kind of scattering in ISAR images. The main contribution of this work is the proof of principle that it is possible to detect jet engines in ISAR images via the presence of range extended returns that are produced by duct delays and a waveguide scattering model.

The fundamentals of the waveguide model are developed in Chapter 3, which presents the propagation of electromagnetic waves inside waveguides. In Section 3.3, the developed model is applied to show the scattering of a chirp waveform, which is used in the application of Chapter 5.

In Chapter 4, the sparse decomposition framework and the MCA algorithm are presented. This algorithm needs a signal model for each component that should be part of the decomposition. These models should be very efficient presenting the content in the image related to them and very inefficient in presenting content related to the other models. Efficiency refers in this case to the number of coefficients that are necessary to represent the content. If this condition is met, a sparse decomposition of the image is possible and the allocation of coefficient and dictionary allows the decomposition of the image into different parts related to the different models.

The application of this framework to radar data is shown in Chapter 5, which is separated into three parts. All of these parts apply the decomposition framework in the complex-valued raw data domain. The first part in Section 5.1 presents a simulation with scatterers of different shape to show the potential of the algorithm to separate several kinds of scattering. The remaining two parts of that chapter are related to the separation of point and waveguide scattering centers in ISAR images. First, a simulation with the waveguide model of Chapter 3 and the standard isotropic point scattering model is done to show that the separation of these two models is feasible. After that, the decomposition is applied to real ISAR data that was measured with the TIRA system of Fraunhofer FHR.

Both, the simulation and the application to real data, show comparable results, i.e. the waveguide scattering can be separated from the point scattering if the SNR is high enough and the model parameter match the data. The simulations show that a low SNR leads to confusion between these two kind of scattering and that knowledge of the model parameters, i.e. radius and length of the engine, are crucial for the decomposition. Especially, a mismatch of the length in the waveguide model and data leads to errors in the reconstruction. Therefore, a heuristic is used with the real data that maximizes the response of the waveguide scattering to find the optimal parameter of the model. The waveguide components of the real data clearly show peaks at the engine positions with some minor components along the fuselage of the aircraft.

The results show that a decomposition into different kind of scattering is possible, if certain requirements are fulfilled. At the same time, it becomes clear that this thesis can only be seen as a step in the direction of physically motivated model-based signal processing. The sensitivity of the results to the model parameters reveals a lack of robustness that should be investigated in the future. At the same time, the jet engine model itself must be improved, since it is following the physics of a perfect waveguide. Although the model matches the real data already in a good way, it is also clear that a real jet engine is not an optimal waveguide. Furthermore, as it was already mentioned in Section 3.2, the rotation of the fan is not considered here. A model including the JEM effect is another option to improve the results of the presented algorithm. If the JEM and other effects are considered in the model, an identification of a certain engine type might be possible.

Another option to improve the results is the sparse decomposition algorithm. One reason for the confusion between point and waveguide scattering might be the similarity in the model, i.e. the coherence of the dictionaries. An algorithm with a high degree of invariance against coherent dictionaries might lead to more robust results with these dictionaries.

The engine dictionary, i.e. the variability of the model, is also a possibility to improve the decomposition. With an improved model, the free parameters will probably increase. In this thesis, the only free parameters during the decomposition were the x- and y-coordinates of the scattering center. If, for example, the aspect angle is taken into account, the mode distribution shown in Section 3.2.1 becomes more important and the dictionary size must be increased since the aspect angle can only be estimated during the imaging process. The use of further parameters, e.g. radius and length, would further increase the size of the dictionary to a size that can probably not be handled anymore in an efficient manner. In that case a dynamic dictionary design during the decomposition is an option to compute a solution. In this thesis, the radius and length used in the dictionary are determined in a heuristic way before the actual decomposition algorithm is applied. With a dynamic dictionary approach, the determination of the model parameters become a part of the problem itself and would be calculated during the decomposition.

List of Figures

List of Acronyms

ATR	Automatic Target Recognition
BCR	Block Coordinate Relaxation
BPDN	Basis Pursuit De-Noising
CNN	Convolutional Neural Network
FFT	Fast Fourier Transform
HRR	High Range Resolution
IFF	Identification Friend or Foe
IHT	Iterative Hard Thresholding
IPP	Image Projection Plane
ISAR	Inverse Synthetic Aperture Radar
IST	Iterative Shrinkage Thresholding
JEM	Jet Engine Modulation
LASSO	Least Absolute Shrinkage and Selection Operator
LOS	Line Of Sight
MAD	Median Absolute Deviation
MCA	Morphological Component Analysis
MP	Matching Pursuit
MSE	Mean Square Error
NCTI	Non Cooperative Target Identification
NCTR	Non Cooperative Target Recognition
ODE	Ordinary Differential Equation

OMP	Orthogonal Matching Pursuit
PCA	Principal Component Analysis
PDE	Partial Differential Equation
PEC	Perfect Electric Conductor
PRF	Pulse Repetition Frequency
PRI	Pulse Repetition Interval
PSF	Point Spread Function
QDM	Quadrature Demodulator
QM	Quadrature Modulator
RCS	Radar Cross Section
SAR	Synthetic Aperture Radar
SNR	Signal to Noise Ratio
SR	Sparse Representation
TIRA	Tracking and Imaging Radar

Acknowledgment

First, I would like to thank Prof. Dr. Joachim Ender for the supervision and patience over the last years and also Prof. Dr. Otmar Loffeld for his support at the University of Siegen. Special thanks go to Dr. Holger Nies from the University of Siegen for his support with all administrative work. At FHR, I would like all my colleagues for their time and also for the feedback I received over the last years regarding this work. Especially the notation and presentation of the sparse representation algorithms have improved a lot due to the help of Dr. Kilian Barth and Dr. Sebastian Durst. Many thanks also to Dr. Christos Liontas and Dr. Stefano Turso for their time explaining waveguide propagation to me. For the time and freedom I had to work on this thesis, I would like to thank Dr. Tanja Bieker and Dr. Stefan Brüggenwirth. I also thank all other colleagues I did not mention here and also all my friends and my family for the support during the last years.

Bibliography

[ABS13] Attouch, Hedy, Jérôme Bolte, and Benar Fux Svaiter: *Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward–backward splitting, and regularized gauss–seidel methods.* Mathematical Programming, 137:91 – 129, 2013.

[Ana03] Anastassiu, H.T.: *A review of electromagnetic scattering analysis for inlets, cavities, and open ducts.* Antennas and Propagation Magazine, IEEE, 45(6):27–40, Dec 2003.

[BAB$^+$20] Belloni, Carole, Nabil Aouf, Alessio Balleri, Jean Marc Le Caillec, and Thomas Merlet: *Pose-informed deep learning method for SAR ATR.* IET Radar, Sonar & Navigation, March 2020, ISSN 1751-8784. `https://digital-library.theiet.org/content/journals/10.1049/iet-rsn.2019.0615`.

[BBC20] BBC: *Iran plane crash: What we know about flight PS752.* https://www.bbc.com/news/world-middle-east-51047006, 2020.

[BDR12] Blumensath, Thomas, Michael E. Davies, and Gabriel Rilling: *Greedy algorithms for compressed sensing.* In Eldar, Yonina C. and Gitta Kutyniok (editors): *Compressed Sensing - Theory and Applications*, chapter 8, pages 348 – 393. Cambridge University Press, New York, 2012.

[Bie08] Bieker, T.: *Statistical evaluation of decision-level fusion methods for non-cooperative target identification by radar signatures.* In *Information Fusion, 2008 11th International Conference on*, pages 1–7, June 2008.

[Bor97] Borden, Brett: *An observation about radar imaging of re-entrant structures with implications for automatic target recognition.* Inverse Problems, 13(6):1441–1452, 1997.

[Bor99] Borden, Brett: *Radar Imaging of Airborne Targets.* Institute of Physics Publishing, Bristol, UK and Philadelphia, PA, USA, 1999.

[Bow85] Bowden, Lord: *The story of IFF (identification friend or foe).* Physical Science, Measurement and Instrumentation, Management and Education - Reviews, IEE Proceedings A, 132(6):435–437, October 1985.

[Bri16] Brisken, Stefan: *Motion Estimation and Imaging with Multistatic Inverse Synthetic Aperture Radar.* PhD thesis, University of Siegen, 2016.

[BSF⁺07] Bobin, J., J L Starck, J.M. Fadili, Y. Moudden, and D.L. Donoho: *Morphological component analysis: An adaptive thresholding strategy.* Image Processing, IEEE Transactions on, 16(11):2675–2681, Nov 2007.

[BSMM08] Bronstein, I.N., K.A. Semendjajew, G. Musiol, and H. Mühlig: *Taschenbuch der Mathematik.* Harri Deutsch Verlag, 7. auflage edition, 2008.

[BST98] Bruce, Andrew G., Sylvain Sardy, and Paul Tseng: *Block coordinate relaxation methods for nonparamatric signal denoising.* In *Proc. of SPIE*, volume 3391, 1998.

[BT09] Beck, A. and M. Teboulle: *A fast iterative shrinkage-thresholding algorithm for linear inverse problems.* SIAM Journal on Imaging Sciences, 2(1):183–202, 2009.

[BV04] Boyd, Stephen and Lieven Vandenberghe: *Convex Optimization.* Cambridge University Press, 2004.

[BV18] Boyd, Stephen and Lieven Vandenberghe: *Introduction to Applied Linear Algebra.* Cambridge University Press, 2018.

[CAD18] Catalina, Alejandro, Carlos M. Alaíz, and José R. Dorronsoro: *Revisiting fista for lasso: Acceleration strategies over the regularization path.* In *European Symposium on Artificial Neural Networks, Computational Intelligenceand Machine Learning ESANN*, 2018.

[CB09] Cheney, Margaret and Brett Borden: *Fundamentals of radar imaging.* Siam, 2009.

[CdBS10] Chauveau, Janic, Nicole de Beaucoudrey, and Joseph Saillard: *Use of resonance parameters of air-intakes for the identification of aircrafts.* In Kouemou, Guy (editor): *Radar Technology*, chapter 13. IntechOpen, Rijeka, 2010.

[Che95] Chen, Shaobing: *Basis Pursuit.* Dissertation Doctor of Philosophy, Stanford University, November 1995.

[CM14] Chen, Victor C. and Marco Martorella: *Inverse Synthetic Aperture Radar Imaging.* SciTech Publishing, Edison, NJ, 2014.

[CMP00] Chiang, Hung Chih, R.L. Moses, and L.C. Potter: *Model-based classification of radar images.* Information Theory, IEEE Transactions on, 46(5):1842–1854, Aug 2000.

[Coo71] Cooper, D. N.: *Complex propagation coefficients and the step discontinuity in corrugated cylindrical waveguide.* Electronics Letters, 7(5):135–136, March 1971.

[CP96] Cloude, S. R. and E. Pottier: *A review of target decomposition theorems in radar polarimetry.* IEEE Transactions on Geoscience and Remote Sensing, 34(2):498–518, March 1996.

[ÇSÖ+14] Çetin, M., I Stojanović, N.O. Önhon, K.R. Varshney, S. Samadi, W.C. Karl, and AS. Willsky: *Sparsity-driven synthetic aperture radar imaging: Reconstruction, autofocusing, moving targets, and compressed sensing.* Signal Processing Magazine, IEEE, 31(4):27–40, July 2014.

[CW07] Chadwick, J. and G.L. Williams: *Air target identification - concept to reality.* In *Radar Systems, 2007 IET International Conference on*, pages 1–5, Oct 2007.

[CWXJ16] Chen, S., H. Wang, F. Xu, and Y. Q. Jin: *Target classification using the deep convolutional networks for SAR images.* IEEE Transactions on Geoscience and Remote Sensing, 54(8):4806–4817, Aug 2016.

[DDEK12] Davenport, Mark A., Marco F. Duarte, Yonina C. Eldar, and Gitta Kutyniok: *Introduction to compressed sensing.* In Eldar, Yonina C. and Gitta Kutyniok (editors): *Compressed Sensing - Theory and Applications*, chapter 1, pages 1 – 64. Cambridge University Press, New York, 2012.

[Dom96] Dommermuth, F.: *Target Classification Using Radar Range Profiles (in German).* Frequenz, 50:157–164, July, August 1996.

[Dut15] Dutch Safety Board: *Mh17 crash.* Technical report, Dutch Safety Board, 2015.

[DWH+17] Ding, B., G. Wen, X. Huang, C. Ma, and X. Yang: *Target recognition in synthetic aperture radar images via matching of attributed scattering centers.* IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 10(7):3334–3347, July 2017.

[dWHM14] Wit, J.J.M. de, R.I.A. Harmanny, and P. Molchanov: *Radar micro-Doppler feature extraction using the singular value decomposition.* In *Radar Conference (Radar), 2014 International*, pages 1–6, Oct 2014.

[DWMY17] Ding, B., G. Wen, C. Ma, and X. Yang: *Decision fusion based on physically relevant features for SAR ATR.* IET Radar, Sonar Navigation, 11(4):682–690, 2017.

[dWvR16] Wit, J.J.M. de and W.L. van Rossum: *Extraction of building features from stand-off measured through-wall radar data.* Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of, 9(1):149–158, Jan 2016.

[Ela10] Elad, Michael: *Sparse and Redundant Representations.* Springer Science+Business Media, New York, NY, 2010.

[End10] Ender, Joachim H. G.: *On compressive sensing applied to radar.* Signal Processing, 90(5):1402 – 1414, 2010.

[ES17] Ender, J. and R. Sommer: *Compressive sensing techniques applied to multilook isar images.* In *2017 18th International Radar Symposium (IRS)*, pages 1–10, June 2017.

[Far93] Farlow, Stanley J.: *Partial Differential Equations for Scientists and Engineers.* Dover Publications, 1993.

[FDST+15] Fromentin-Denozière, B., J. Simon, A. Tzoulis, F. Weinmann, H.T. Anastassiu, D. Escot Bocanegra, D. Poyatos Martínez, R. Fernández Recio, and A. Zdunek: *Comparative study of miscellaneous methods applied to a benchmark, inlet scattering problem.* Radar, Sonar Navigation, IET, 9(3):342–354, 2015.

[Fis13] Fisher, Max: *The forgotten story of iran air flight 655.* https://www.washingtonpost.com/news/worldviews/wp/2013/10/16/the-forgotten-story-of-iran-air-flight-655/, October 2013.

[FR13] Foucart, Simon and Holger Rauhut: *A Mathematical Introduction to Compressive Sensing.* Applied and Numerical Harmonic Analysis. Springer Science+Business Media, New York, 2013.

[FSED10] Fadili, J.M., J L Starck, M. Elad, and D.L. Donoho: *MCALab: Reproducible research in signal and image decomposition and inpainting.* Computing in Science Engineering, 12(1):44–63, Jan 2010.

[Gar61] Gardner, R.E.: *Doppler spectral characteristics of aircraft radar targets at S-band.* Technical Report 5656, Naval Research Laboratory, Washington DC, Aug 1961.

[Geo97] Georg, Otfried: *Elektromagnetische Wellen (in German).* Springer-Verlag, Berlin, 1997.

[Hag20] Hage, Dunja: *Neue Lösungsstrategien für l1-Minimierungsprobleme mit Kalman-Filtern (in German).* Dissertation Dr. Ing., University of Siegen, 2020.

[HGT06] Herrity, K.K., A.C. Gilbert, and J.A. Tropp: *Sparse approximation via iterative thresholding.* In *Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on*, volume 3, pages III624–III627, May 2006.

[HP93] Hudson, S. and D. Psaltis: *Correlation filters for aircraft identification from radar range profiles.* Aerospace and Electronic Systems, IEEE Transactions on, 29(3):741–748, Jul 1993.

[Hua83] Huang, C. C.: *Simple formula for the RCS of a finite hollow circular cylinder.* Electronics Letters, 19(20):854–856, September 1983.

[HWS10] Haumtratz, T., J. Worms, and J. Schiller: *Classification of air targets including a rejection stage for unknown targets.* In *Radar Symposium (IRS), 2010 11th International*, pages 1–4, June 2010.

[Jac09] Jackson, Julie Ann: *Three-Dimensional Feature Models for Synthetic Aperture Radar and Experiments in Feature Extraction.* Dissertation Doctor of Philosophy, The Ohio State University, 2009.

[JO00] Jacobs, S.P. and Joseph A. O'Sullivan: *Automatic target recognition using sequences of high resolution radar range-profiles.* Aerospace and Electronic Systems, IEEE Transactions on, 36(2):364–381, Apr 2000.

[KLBA06] Krogager, E., J. Lee, W. Boerner, and T. L. Ainsworth: *Polarimetric analysis of radar signature of a manmade structure.* In *2006 International Radar Symposium*, pages 1–4, May 2006.

[KPDA60] Klauder, J.R., A.C. Price, S. Darlington, and W.J. Albersheim: *The theory and design of chirp radars.* The Bell System Technical Journal, 39(4):745 – 808, July 1960.

[KS18] Kechagias-Stamatis, Odysseas: *Target recognition for synthetic aperture radar imagery based on convolutional neural network feature fusion.* Journal of Applied Remote Sensing, 12:046025–1–17, 2018.

[KSK05] Kim, Kyung Tae, Dong Kyu Seo, and Hyo Tae Kim: *Efficient classification of ISAR images.* Antennas and Propagation, IEEE Transactions on, 53(5):1611–1621, May 2005.

[Mah05] Mahafza, Bassem R.: *Radar Systems Analysis and Design Using MATLAB.* Chapman & Hall/CRC, 2nd edition edition, 2005.

[MBK16] Maier-Borst, Haluka and Sybille Klormann: *Steuerte der Pilot MH370 bewusst ins Meer? (in German).* https://www.zeit.de/wissen/2014-03/mh370-malaysia-airlines-zwischenstand/komplettansicht, August 2016.

[MEA$^+$14] Molchanov, P., K. Egiazarian, J. Astola, A. Totsky, S. Leshchenko, and M.P. Jarabo-Amores: *Classification of aircraft using micro-Doppler bicoherence-based features.* Aerospace and Electronic Systems, IEEE Transactions on, 50(2):1455–1467, April 2014.

[Meh96] Mehrholz, Dieter: *Ein Verfolgungs- und Abbildungsradarsystem zur Beobachtung von Weltraumobjekten (in German).* Frequenz, 50(7-8):138–146, Juli 1996.

[MW93] Moghaddar, A. and E. K. Walton: *Time-frequency distribution analysis of scattering from waveguide cavities.* IEEE Transactions on Antennas and Propagation, 41(5):677–680, May 1993.

[MW99] Mitchell, R.A. and John J. Westerkamp: *Robust statistical feature based aircraft identification.* Aerospace and Electronic Systems, IEEE Transactions on, 35(3):1077–1094, Jul 1999.

[MZ93] Mallat, Stephane G. and Zhifeng Zhang: *Matching pursuits with time-frequency-dictionaries.* IEEE Transactions on Signal Processing, 41(12):3397 – 3415, December 1993.

[Neb95] Nebabin, Viktor Georgieviech: *Methods and Techniques of Radar Recognition.* Artech House, Norwood, MA, 1995.

[NKOW02] Nguyen, Duy H., John H. Kay, Bradley J. Orchad, and Robert H. Whiting: *Classification and tracking of moving ground vehicles.* Lincoln Laboratory Journal, 13(2):275 – 308, 2002.

[ØKCL16] Ødegaard, Nina, Atle Onar Knapskog, Christian Cochin, and Jean Christophe Louvigne: *Classification of ships using real and simulated data in a convolutional neural network.* In *Radar Conference (RadarConf), 2016 IEEE*, May 2016.

[OL07] Ohm, Jens Rainer and Hans Dieter Lüke: *Signalübertragung - Grundlagen der digitalen und analogen Nachrichtenübertragungssysteme (in German).* Springer-Verlag, 10th edition, 2007.

[Orf16] Orfanidis, Sophocles J.: *Electromagnetic Waves and Antennas.* http://www.ece.rutgers.edu/ orfanidi/ewa/, 2016.

[PM97] Potter, L.C. and R.L. Moses: *Attributed scattering centers for SAR ATR.* Image Processing, IEEE Transactions on, 6(1):79–91, Jan 1997.

[Pra07] Pratt, William K.: *Digital Image Processing.* John Wiley & Sons, 4th edition, 2007.

[RH96] Rihaczek, A. W. and S. J. Hershkowitz: *Man-made target backscattering behavior: applicability of conventional radar resolution theory.* IEEE Transactions on Aerospace and Electronic Systems, 32(2):809–824, April 1996.

[RLNP11] Rial, F. I., H. Lorenzo, A. Novo, and M. Pereira: *Checking the signal stability in GPR systems and antennas.* IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 4(4):785–790, Dec 2011.

[RS95] Rosenbach, K. and J. Schiller: *Identification of aircraft on the basis of 2-D radar images.* In *Radar Conference, 1995., Record of the IEEE 1995 International*, pages 405–409, May 1995.

[SÇMs11] Samadi, S., M. Çetin, and M.A. Masnadi-shirazi: *Sparse representation-based synthetic aperture radar imaging.* Radar, Sonar Navigation, IET, 5(2):182–193, Feb 2011.

[SED04] Starck, Jean Luc, Michael Elad, and David Donoho: *Redundant multiscale transforms and their application for morphological component separation.* In *Advances in Imaging and Electron Physics*, volume 132, pages 287 – 348. Elsevier, 2004.

[SFHT13] Simon, Noah, Jerome Friedman, Trevor Hastie, and Robert Tibshirani: *A sparse-group lasso.* Journal of Computational and Graphical Statistics, 22(2):231–245, 2013. `https://doi.org/10.1080/10618600.2012.681250`.

[Sko01] Skolnik, Merrill I.: *Introduction to Radar Systems.* McGraw Hill, 3rd edition edition, 2001.

[SMF10] Starck, Jean Luc, Fionn Murtagh, and Jalal M. Fadili: *Sparse Image and Signal Processing.* Cambridge University Press, New York, NY, 2010.

[Str41] Stratton, Julius Adams: *Electromagnetic Theory.* International Series in pure and applied physics. McGraw Hill, New York, NY, USA, 1941.

[Tai05] Tait, Peter: *Introduction to Radar Target Recognition*, volume 18 of *Radar, Sonar, Navigation and Avionics Series.* The Institution of Electrical Engineers (IEE), Stevenage, UK, 2005.

[TL96] Trintinalia, L. C. and Hao Ling: *Extraction of waveguide scattering features using joint time-frequency ISAR.* IEEE Microwave and Guided Wave Letters, 6(1):10–12, Jan 1996.

[Tro04] Tropp, Joel Aaron: *Topics in Sparse Approximation.* Dissertation Doctor of Philosophy, University of Texas, August 2004.

[Unb02] Unbehauen, Rolf: *Systemtheorie 1 - Allgemeine Grundlagen, Signale und lineare Systeme im Zeit- und Frequenzbereich (in German)*. Oldenbourg Wissenschaftsverlag GmbH, 8. auflage edition, 2002.

[Wag14] Wagner, Simon: *Combination of convolutional feature extraction and support vector machines for radar ATR*. In *Information Fusion (FUSION), 2014 17th International Conference on*, July 2014.

[Wag15] Wagner, S.: *Morphological component analysis in SAR images to improve the generalization of ATR systems*. In *Compressed Sensing Theory and its Applications to Radar, Sonar and Remote Sensing (CoSeRa), 2015 3rd International Workshop on*, pages 46–50, June 2015.

[Wag16] Wagner, Simon A.: *SAR ATR by a combination of convolutional neural network and support vector machines*. IEEE Transactions on Aerospace and Electronic Systems, 52(6):2861–2872, December 2016.

[WDE16] Wagner, Simon, Franz Dommermuth, and Joachim Ender: *Detection of jet engines via sparse decomposition of ISAR images for target classification purposes*. In *Proceedings of the 13th European Radar Conference*, pages 77 – 80, London, UK, October 2016.

[WE17] Wagner, Simon and Joachim Ender: *Recognition of jet engines via sparse decomposition of ISAR images using a waveguide scattering model*. International Journal of Microwave and Wireless Technologies, 9(6):1339–1343, 2017.

[WE21] Wagner, Simon and Joachim Ender: *Jet engine recognition via sparse decomposition of ISAR images*. In Mishra, Amit Kumar and Stefan Brüggenwirth (editors): *New Methodologies for Understanding Radar Data*, pages 253–284. Scitech Publishing, 2021.

[WE22] Wagner, Simon and Joachim Ender: *Scattering identification in isar images via sparse decomposition*. In *2022 IEEE Radar Conference (RadarConf22)*, pages 1–6, 2022.

[Weh95] Wehner, Donald R.: *High-Resolution Radar*. Artech House, 2nd edition edition, 1995.

[Wex67] Wexler, A.: *Solution of waveguide discontinuities by modal analysis*. IEEE Transactions on Microwave Theory and Techniques, 15(9):508–517, September 1967.

[Wir04] Wirgin, Armand: *The inverse crime*, 2004. `https://arxiv.org/abs/math-ph/0401050`.

[YL06] Yuan, Ming and Yi Lin: *Model selection and estimation in regression with grouped variables*. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 68(1):49–67, 2006. `https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9868.2005.00532.x`.

[ZXY$^+$15] Zhang, Zheng, Yong Xu, Jian Yang, Xuelong Li, and D. Zhang: *A survey of sparse representation: Algorithms and applications*. Access, IEEE, 3:490–530, 2015.

Index