Uppsala University
Signals and Systems

Time-domain Reconstruction Methods for Ultrasonic Array Imaging

A Statistical Approach

Fredrik Lingvall

UPPSALA UNIVERSITY 2004

Dissertation for the degree of Doctor of Philosophy
in Electrical Engineering with specialization in Signal Processing, 2004.

ABSTRACT
Lingvall, F., 2004. Time-domain Reconstruction Methods for Ultrasonic Array Imaging: A Statistical Approach, 178 pp. Uppsala. ISBN 91-506-1772-9.

This thesis is concerned with reconstruction techniques for ultrasonic array imaging based on a statistical approach. The reconstruction problem is posed as the estimation of an image consisting of scattering strengths. To solve the estimation problem, a time-domain spatially-variant model of the imaging system is developed based on the spatial impulse response method. The image reconstruction is formulated as a linearized inverse-scattering problem in which the time and space discrete natures of the imaging setup as well as measurement uncertainties are taken into account. The uncertainties are modeled in terms of a Gaussian distribution. The scattering strengths are estimated using two prior probability density functions (PDF's), Gaussian and exponential PDF's. For the Gaussian PDF, the maximum a posteriori (MAP) approach results in an analytical solution in the form of a linear spatio-temporal filter which deconvolves the diffraction distortion due to the finite-sized transducer. The exponential distribution leads to a positivity constrained quadratic programming (PCQP) problem that is solved using efficient optimization algorithms. In contrast to traditional beamforming methods (based on delay-and-summation), the reconstruction approach proposed here accounts both for diffraction effects and for the transducer's electro-mechanical characteristics. The simulation and experimental results presented show that the performances of the linear MAP and nonlinear PCQP estimators are superior to classical beamforming in terms of resolution and sidelobe level, and that the proposed methods can effectively reduce spatial aliasing errors present in the conventional beamforming methods.

Keywords: Ultrasonic array imaging, synthetic aperture imaging, maximum a posteriori estimation, linear minimum mean squared error estimation, SAFT, Bayesian image reconstruction.

Fredrik Lingvall, Signals and Systems, Uppsala University, P O Box 528, SE-751 20 Uppsala, Sweden. E-mail: fredrik.lingvall@signal.uu.se.

This thesis has been prepared using LaTeX.

ISBN 91-506-1772-9
Printed in Sweden by Eklundshofs Grafiska AB,Uppsala, 2004
Distributed by Signals and Systems, Uppsala University, Uppsala, Sweden.

To my parents

Acknowledgments

Many persons have supported and encouraged me in my work as a Ph.D student. I would especially like to thank my supervisor Prof. Tadeusz Stepinski for all his help, encouragement, and for giving me the opportunity to work in this very inspiring group. I also want to express my sincere gratitude to Dr. Tomas Olofsson for his valuable guidance and support. Much of what is presented in this thesis is the result of stimulating discussions with Tomas Olofsson.

I want to thank Dr. Ping Wu and Prof. Bogdan Piwakowski who have helped me the most in understanding acoustic wave propagation and acoustic modeling. It has been very interesting and instructive to develop the acoustic modeling software used in this thesis in cooperation with Bogdan Piwakowski.

I am very grateful to Prof. Anders Ahlén for proof-reading (and improving) the manuscript, as well as encouraging me when everything seemed hopeless.

I also acknowledge all the people that work, and have worked, at the Signals and Systems group who have made Magistern such a fun place to work at. In particular, Dr. Mats Gustafsson and Dr. Daniel "kungligt" Asraf for, among other things, our many valuable discussions, and Ove Ewerlid, for improving my Linux skills. I'm also grateful for the support of my "teammates" Mattias Johansson and Nilo Casimiro Ericsson, and to Dr. Sorour Falahati and Prof. Mikeal Sternard for keeping my spirit up while writing this thesis!

I'm also very grateful to my biking and climbing friends, where our many mountainbike rides, trips to the French Alps and the Norwegian icefalls have helped me "re-charging my batteries".

Finally, I would like to express my deepest gratitude to my parents and to my two brothers for your never-ending support. Thank you!

Fredrik Lingvall
Uppsala, September 2004.

CONTENTS

Glossary

Notational conventions

Matrices and vectors are denoted using boldface Roman letters and scalars are denoted using lowercase Roman or Greek letters throughout the thesis. All matrices are expressed using boldface upper-case letters and all vectors are, by convention, column vectors.

Two types of vectors occur in the thesis. One type denote the positional vectors that are used when presenting the sound model. These three dimensional vectors are expressed using $\mathbf{r}$ and sometimes have a sub and/or super script. For instance, $\mathbf{r}_{\mathrm{T}}$ denotes the position of a point-like target.

The vectors of the second type are of higher dimension. These vectors typically consist of vectorized matrices or of samples from a signal or noise sequence. A vectorized matrix $\mathbf{a}$ is defined by

$$\mathbf{a} \triangleq \mathrm{vec}(\mathbf{A}) = [\mathbf{a}_0^T \mathbf{a}_1^T \cdots \mathbf{a}_J^T],$$

where $\mathbf{a}_j$ denotes the jth column in $\mathbf{A}$ and J is the total number of columns in $\mathbf{A}$.

Symbols

Below the most frequently used symbols in the thesis are summarized:

$*$	Convolution in time
$\doteq$	Rounding towards the nearest integer
$\triangleq$	Equality by definition
$\mathbf{A}^T$	The transpose of matrix $\mathbf{A}$
$\mathbf{a}_j$	The jth column of matrix $\mathbf{A}$
$(\mathbf{A})_{m,n}$	Element (m, n) in matrix $\mathbf{A}$
$\mathrm{vec}(\cdot)$	Operator that transforms a matrix lexicographically to a vector
E	The expectation operator with respect to signals
$\|\|\mathbf{x}\|\|^2$	The L_2 norm of the vector $\mathbf{x}$ ($\|\|\mathbf{x}\|\|^2 = \mathbf{x}^T\mathbf{x}$)
$\mathrm{tr}\{\cdot\}$	The trace of a matrix ($\mathrm{tr}\{\mathbf{A}\} = \sum_i (\mathbf{A})_{i,i}$)
$\|\mathbf{A}\|$	The determinant of the matrix $\mathbf{A}$
t	Time
k	Discrete-time index
f	Frequency in Hertz

ω	Angular frequency ($\omega = 2\pi f$)
F_s	Sampling frequency
T_s	Sampling period ($T_\mathrm{s} = 1/F_\mathrm{s}$)
$\mathbf{C}_\mathrm{e}$	Covariance matrix of $\mathbf{e}$
$\mathbf{C}_\mathrm{o}$	Covariance matrix of $\mathbf{o}$
$\mathbf{I}$	The identity matrix
$\mathbf{0}$	The null matrix of suitable dimension
$\delta(t)$	Delta function (in continuous-time)
w_n	Apodization weight
$v_\mathrm{n}(t)$	Normal velocity
S_R^b	Surface of a transducer in receiving (backward) mode
S_R^f	Surface of a transducer in transmit (forward) mode
c_p	Longitudinal (primary) sound speed
ρ_0	Medium density
$\mathbf{r}$	A general field point
$\mathbf{r}_\mathrm{T}$	A field point of a point-like target
$\mathbf{r}_\mathrm{R}^\mathrm{r}$	A point on the receive (backward) transducer's surface
$\mathbf{r}_\mathrm{R}^\mathrm{f}$	A point on the transmit (forward) transducer's surface

Abbreviations

ADC	Analog-to-Digital Converter
DREAM	Discrete REpresentation Array Modeling
DAS	Delay-and-Sum
CW	Continuous Wave
FIR	Finite Impulse Response
HPBW	Half-power Bandwidth
IID	Independent and Identically Distributed
IIR	Infinite Impulse Response
LS	Least Squares
LTI	Linear Time-Invariant
MAP	Maximum *A Posteriori*
MF	Matched Filter
MMSE	Minimum Mean Squared Error
ML	Maximum Likelihood
MSE	Mean Squared Error
MCMC	Markov Chain Monte-Carlo
NDE	Nondestructive Evaluation

NDT	Nondestructive Testing
PA	Phased Array
PDE	Partial Differential Equation
PDF	Probability Density Function
PSF	Point Spread Function
QP	Quadratic Programming
RF	Radio Frequency
ROI	Region-of-Interest
SAI	Synthetic Aperture Imaging
SAFT	Synthetic Aperture Focusing Technique
SDH	Side-drilled Hole
SNR	Signal-to-Noise Ratio
SAR	Synthetic Aperture Radar
SAS	Synthetic Aperture Sonar
TOFD	Time-of-Flight Diffraction

CHAPTER 1

Introduction

ACOUSTIC array imaging is a technique used in many fields such as medical diagnostics, sonar, seismic exploration, and non-destructive testing. The objective with acoustic imaging, in general, is to localize and characterize objects by means of the acoustic scattered field that they generate. An array of sensors facilitates this task since the diversity introduced by transmitting and receiving acoustic energy from many positions enables localization of the scatterers.[1] In other words, by using an array it is possible to both localize remote objects and estimate their scattering strengths.

In contrast to electromagnetic array imaging, where narrow band signals with wavelengths exceeding the array antenna elements are used, broadband pulsed waveforms are commonly used in acoustic array imaging. The wavelength corresponding to the center frequency of these pulses is often of the same size as the array elements. This make diffraction effects apparent resulting in a frequency dependent beam directivity of the array elements.

Array imaging is, in this thesis, understood as a process which aims at extracting information about the scattering strength of an object under study. A realistic model of the imaging system is required to perform a successful imaging.

The aim of this thesis is to develop and explore new model based imaging

[1] Acoustic arrays have also become popular since they are less expensive than, for example, X-ray computerized tomography (CT) or magnetic resonance imaging (MRI) systems [1].

methods by utilizing the spatial diversity inherent in array systems by taking diffraction effects as well as the broadband pulsed waveforms into account. This is performed using a model-based approach where the model is obtained from the known geometry and electro-mechanical characteristics of the array.

A suitable model can be obtained by solving the *forward problem* of wave propagation. The main objective in acoustic array imaging is, however, finding a solution to the backward, or the *inverse problem*, which is defined here as image reconstruction from the ultrasound data. More specifically, the image under consideration, which consists of the scattering strength at every observation point in the region-of-interest (ROI), has to be estimated from the received acoustic field. That is, the image cannot be measured directly, it can only be observed indirectly through the scattered field measured at the (limited) array aperture which, furthermore, is distorted due to the electro-mechanical characteristics of the measurement system. Therefore, the image has to be estimated, or reconstructed, from the observed data. This can be seen as a process of removing the effects of the observation mechanism, which in our case is the processing needed to compensate for the wave propagation effects and the electro-mechanical properties of the array.

Traditionally, acoustic array data has been processed by means of spatial filtering, or *beamforming*, to obtain an image of the scattering objects [2–4]. Beamforming using the classical time-domain delay-and-sum method is analogous to the operation of an acoustical lens and it can be performed efficiently using delay-line operations in real time or using post-processing as in synthetic aperture systems [5–8]. Conventional beamforming, which is essentially based on a geometrical optics approach [9], is computationally attractive due to its simplicity but is has several inherent drawbacks. In particular, conventional beamforming does not perform well in situations where the ultrasound data is obtained using an array that is sparsely sampled which, unfortunately, is quite common in many applications.

The main motivation for developing the reconstruction algorithms discussed in this thesis is the need to improve the performance of traditional ultrasonic array imaging methods. To overcome, or at least alleviate, the problems with conventional beamformers, a model based time-domain approach is proposed here. It takes into account the diffraction pattern of an arbitrary array setup. This model accounts for all linear effects, such as, transducer size effects, side- and grating lobes, focusing, and beam steering. By a proper design of the reconstruction algorithm it should be possible to compensate such effects and thereby achieve reconstruction performance

superior to conventional array imaging.

The rest of the chapter is organized as follows: In Section 1.1, some basic concepts in ultrasonic array imaging are presented. In particular an introduction to elastic wave propagation, scattering and array image formation is given. This is followed by Section 1.2, where conventional array imaging, or beamforming is discussed. Section 1.3 presents the general formulation of the array imaging problem, and in Section 1.4 the contributions of this thesis are summarized. A short review of earlier related work is also presented. Finally, Section 1.5 gives the outline of this thesis.

1.1 Basic Ultrasonic Imaging Concepts

The basic physical property behind ultrasonic imaging is that an impinging acoustical wave is scattered by a discontinuity, that is, a change in the acoustic impedance of an otherwise homogeneous object. The received scattered wave is then converted to an electric signal which normally is sampled and stored in a computer, or a dedicated ultrasonic instrument.

The scattering is governed by a change in the density, ρ, and sound speed, c, of the inspected specimen. The reflected and transmitted pressure amplitudes are described by the transmission and reflection coefficients, T_{t} and T_{r} respectively. For a planar discontinuity, these coefficients are given by the acoustic impedance, $z_1 = \rho_1 c_1$, before the discontinuity, and the acoustic impedance, $z_2 = \rho_2 c_2$, after the discontinuity, where ρ_1 and ρ_2 are the respective densities and c_1 and c_2 the corresponding sound speeds. Let p_{i} denote the incident pressure. Then the pressure of the reflected wave, p_{r}, is given by [10]

$$p_{\mathrm{r}} = T_{\mathrm{r}} p_{\mathrm{i}} = \frac{z_2 - z_1}{z_2 + z_1} p_{\mathrm{i}} \tag{1.1}$$

and the pressure of the transmitted wave, p_{t}, is

$$p_{\mathrm{t}} = T_{\mathrm{t}} p_{\mathrm{i}} = \frac{2z_2}{z_2 + z_1} p_{\mathrm{i}}. \tag{1.2}$$

For a *hard* discontinuity we have that $z_2 > z_1$ and and for a *soft* $z_2 < z_1$. For soft scatterers, the reflection coefficient is negative, which physically results in a phase shift at the boundary of the scatterer.[2]

[2]Later when the reconstruction algorithms are discussed it will be shown that the reconstruction performance can be improved if it is known *a priori* whether the scatterer is hard or soft.

The acoustic waves are generated using a *transducer* which converts an electrical waveform to an acoustical (or elastic) wave and vice versa.[3] The driving input signal to the transducer is normally a sinusoidal signal, resulting in a narrowband continuous wave (CW), or a pulsed broadband signal.

In CW imaging it is difficult to separate the scattered and the incident field. This is normally solved by using tone-burst signals that still are narrowband but allow for a sufficient time separation between the transmitted and received signals.[4] Henceforth in this thesis it will be assumed that there always is a sufficient time separation between the transmitted and received waveforms.

1.1.1 Ultrasonic Transducers

Good acoustic coupling between the transducer and the specimen/medium is required for an acoustic wave to propagate efficiently into the specimen. Therefore many types of transducers are available to suit different applications. The two perhaps most common types are contact and immersion transducers. Contact transducers are common in non destructive evaluation (NDE) applications where the transducer is placed on the surface of the specimen, normally with a couplant gel or water between the transducer and the specimen's surface to provide sufficient wave transfer into the specimen.

The most common type of transducer is the piezoelectric transducer which is used for both contact and immersion measurements. A typical piezoelectric contact transducer is shown in Figure 1.1. The piezoelectric crystal is plated on both sides to create electrodes where the driving input signal can be applied. If an electrical signal is applied to the electrodes, then the crystal will expand and thereby generate an acoustic wave. The crystal of a contact transducer is also normally protected with a ceramic wear plate. Moreover, the backface of the crystal is loaded with a highly attenuating medium, the so-called *backing*, which controls the shape and duration of the output waveform. Without the backing the transducer would have a very distinct resonant frequency resulting in a narrowband system, that is, the transducer would have a long impulse response which deteriorates the

[3] If one makes the comparison with sound waves then the transducer acts both as a microphone and a as loudspeaker.

[4] In some applications, for example sonar, frequency sweeped chirp signals are also common.

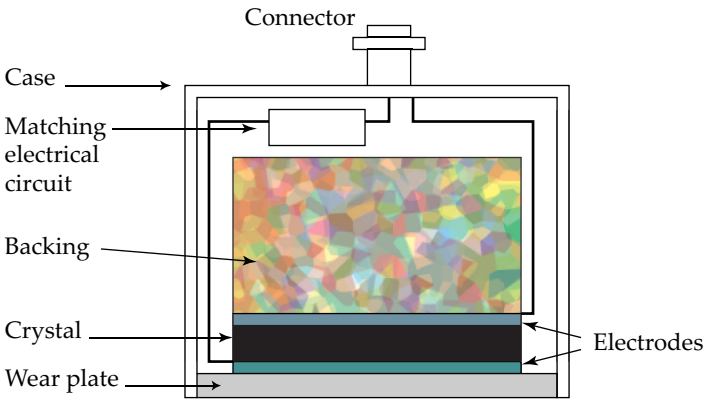


Figure 1.1: Schematic view of a piezoelectric contact transducer. The piezoelectric crystal is protected on the front side with a ceramic wear plate and attached to a highly attenuating backing medium on its backface to reduce resonant oscillations.

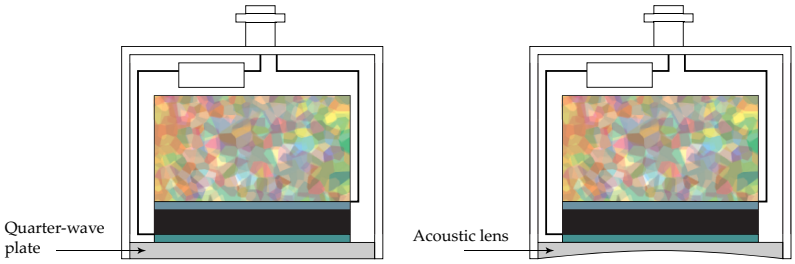
temporal resolution in many systems.

Immersion transducers are similar to contact transducers with the difference that an immersion transducer has often a quarter-wave matching layer on its surface instead of the ceramic wear plate, see Figure 1.2(a). The quarter-wave matching is specifically designed to allow an efficient transfer of the acoustic energy into the medium (i.e., water). Immersion transducers may also have an acoustic lens attached to the front crystal surface, as illustrated in Figure 1.2(b), for concentrating the acoustic field to a narrow region.

1.1.2 Ultrasonic Image Presentation Methods

Acoustic imaging is in many respects similar to electromagnetic imaging. In particular *pulse-echo* imaging is similar to radar in the sense that in both fields the round-trip time of the pulse is utilized to estimate the distance to the scatterer. More specifically, the arrival time of the return echo provides range information of the scatterer and by acquiring the signals from different antenna/transducer positions, the scatterer's bearing can also be estimated.

A typical immersion pulse-echo experimental setup is depicted in Figure 1.3(a). The received digitized radio frequency (RF) waveforms are known as *A-scans* (Amplitude-scans) and are the basic data unit for broadband acoustic imaging; an A-scan can be visualized as an oscilloscope presenta-



(a) A piezoelectric immersion transducer. The quarter-wave plate is designed to allow an efficient energy transfer into the water.

(b) A piezoelectric focused immersion transducer. The acoustic lens concentrates the acoustic energy to a narrow region.

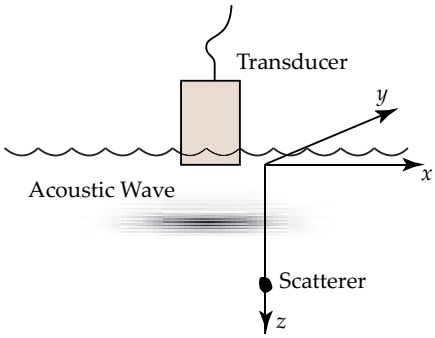
Figure 1.2: Illustration of immersion transducers.

tion of the received waveform as shown in shown in Figure 1.3(b).

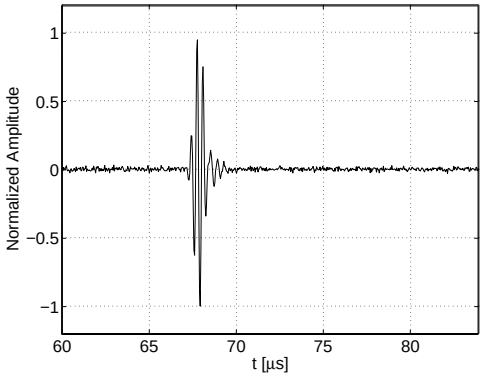
As mentioned above, the round-trip time of the echo yields information about the range to the scatterer. This information is, however, ambiguous since many scatterer positions can result in identical round-trip times. Therefore, to enable localizing scatterers, it is necessary to sample the wave field spatially by moving the transducer or by using several transducers. An image obtained by a linear movement of the transducer and coding the received waveform amplitudes, from many transducer positions, in colors or as gray levels is usually denoted a *B-scan*. An example B-scan is shown in Figure 1.3(c) and if the scatterer is located along the scanning direction, the distance, z_{sc}, between the transducer and the scatterer can be computed from the shortest roundtrip-time, t_{rt}, found in the B-scan, that is, $t_{\mathrm{rt}} = \frac{2z_{\mathrm{sc}}}{c}$.

If the scatterers are located in a certain volume then a surface scan can be performed resulting in a 3D data volume. A common technique to visualize 3D data in acoustical imaging is the so-called *C-scan*. A C-scan can be thought of as a photograph taken from above the specimen, obtained by taking the maximum value of the received A-scans in a particular time-frame. Figure 1.3(d) shows an example C-scan and it can now be seen that the scattering energy is concentrated, or "has a peak", at $x = y = 0$.

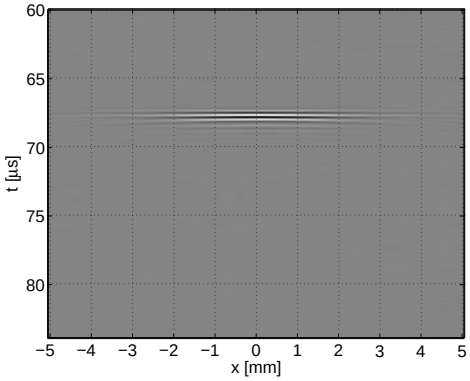
The resolution in B- and C-scan imaging is determined primarily by the beam shape of the used transducer. Using a highly diffracting (small) transducer results in poor lateral resolution and using a large or focused



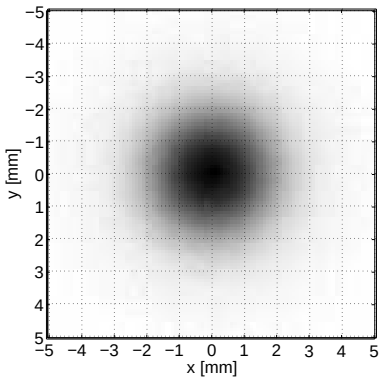
(a) A typical immersion pulse-echo measurement setup.



(b) A-scan (Amplitude-scan) plot.



(c) B-scan image obtained by a horizontal movement of the transducer. The amplitudes of the waveforms corresponding to each transducer position is coded as gray levels in the B-scan.



(d) C-scan image obtained by scanning the transducer in the xy-plane. A C-scan is formed by taking the maximum value of the received A-scans in a particular time-frame.

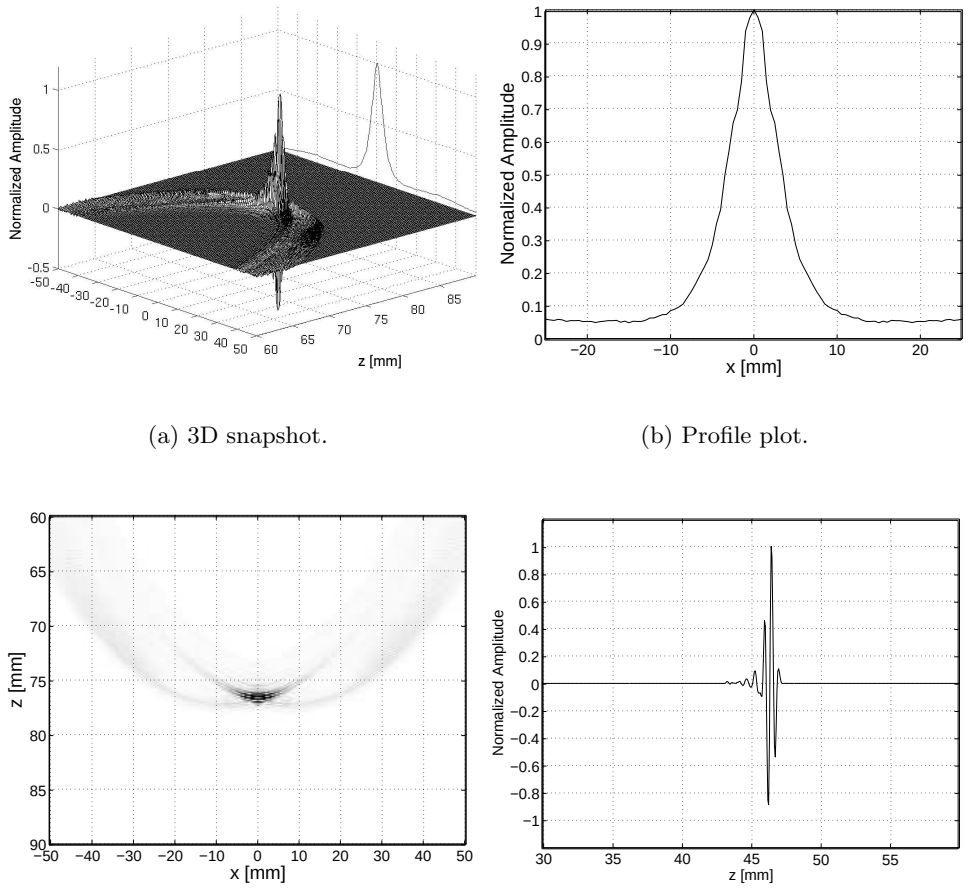
Figure 1.3: Common data presentations types in pulse-echo ultrasonic imaging.

transducer improves the resolution. This is analogous with optical imaging where the sharpness of the image is improved by using a lens. In acoustical imaging, the acoustic lens compensates for the arrival times of the sound waves originating from the focal zone of the lens, yielding a coherent summation of those waves. The width and depth of the focal zone depends of the aperture and focusing depth of the lens. A larger lens has a more narrow focal zone than a small lens focused at the same depth [9]. Acoustic imaging using physical lenses may therefore require repeated scanning with different large aperture lenses, focused at different depths, to obtain volume data with sufficient resolution.

A-, B- and C-scan imaging are the most rudimentary methods used in acoustical imaging where almost no processing of the data is required. They enable, by simple means, visualization of the scattering strengths but their usefulness is also, somewhat limited. In particular, the user that observes the images must be skilled enough to be able to interpret the results, which can become rather complicated since responses from many scatterers may be superimposed. The mechanical scanning required to form an image when using physical lenses is also time consuming which limits the image frame rate. For these and other reasons, imaging using physical lenses is often impractical or even unfeasible. Therefore, a considerable interest has been observed in lensless acoustic imaging and array imaging where focusing can be performed after acquiring data, as in synthetic aperture imaging or in real time by using physical array systems. It will be shown later in this thesis that the use of computer analysis can yield further improvements since signal processing can contribute to a substantial improvement of the resulting images.

Image Display Methods for Broadband Sound Fields and Signals

Properties, such as, lateral- and temporal resolution, contrast etc are important measures in ultrasonic imaging. For CW signals and far-field analysis there exist well established presentation methods, as we will describe in Section 1.2 below, but for broadband and near-field data special presentation methods are required. To facilitate evaluation of broadband data, four different display methods are used in this thesis, each one of which being suitable for a different purpose. The display methods are shown in Figure 1.4 using example waveform data from a 1D array simulation. Figure 1.4(a) shows 3D snapshot graph of a (simulated) broadband waveform at one time in-



(a) 3D snapshot.

(b) Profile plot.

(c) 2D snapshot.

(d) On-axis waveform.

Figure 1.4: Image display methods used in throughout the thesis.

stant. Also shown in the 3D graph is the profile of the acoustic field, that is, a projection of the maximum amplitudes of the field on the x-axis. This profile plot is also shown in Figure 1.4 (b). In Figure 1.4(c) a 2D image of the waveform is depicted, whereas in Figure 1.4(d) the on-axis waveform is displayed. These four display methods will be used to present both pressure waveforms, of the type shown in Figure 1.4, and reconstructions, that is, an image or estimate of the insonified object's scattering strength obtained by processing ultrasound data.

The profile plot can be seen as the broadband counterpart to the CW beampattern plot which is suitable for studying the lateral resolution or the contrast. The temporal or on-axis resolution is most easily studied using plots of the type shown in Figure 1.4(d). The 2D and 3D plots are suitable to obtain an overall view of a wavefield or a reconstruction.

1.1.3 Wave Modes of Elastic Waves

In addition to the comparison with radar imaging above, it should be noted that, even though radar and acoustic pulse-echo imaging have many similarities there are also some fundamental differences. The most apparent difference is that elastic waves, as opposed to electromagnetic waves have many propagation modes.

The two dominating modes in acoustic imaging are *longitudinal* or *pressure* waves and *shear* or *transversal* waves [10, 11]. In longitudinal wave propagation, the particle displacement is in the same direction as the propagation whilst the particle displacement is orthogonal to the propagation direction for shear waves. In liquids and gases, shear wave propagation is strongly attenuated and only the longitudinal mode will therefore propagate. In solids, both longitudinal and shear waves can propagate. The wave modes have different sound speeds and the longitudinal sound speed, c_{p}, in metals is roughly twice the shear wave speed.

Furthermore, elastic wave propagation is complicated by the fact that *mode conversion* occurs at a non-normal reflection and refraction at a discontinuity inside a test specimen, e.g., a metal test object. A typical example, from NDE imaging, is the multiple mode conversions that can occur at cracks [12]. There, mode conversion generates both longitudinal and shear diffracted waves at the tips of the crack. A surface wave may also be generated along the crack, which has different sound speed compared to longitudinal and shear waves, resulting in additional diffracted waves at the crack ends [13]. The resulting received signals may therefore consist of a mix of different wave modes which may be very difficult to discriminate. Fortunately, in many applications the echoes, corresponding to each wave mode can be resolved since they will arrive with a time separation caused by the different sound speeds of the wave modes.

In the work presented in this thesis all measurements have been performed in water in order to avoid mode converted waves and thereby simplifying the interpretation of the obtained results. The only exception where longitudinal-shear-longitudinal mode conversion could possibly occur was

in an experiment with scatterers inside an immersed copper block. That experiment was, however, performed in such way that only the non-mode converted longitudinal waves could arrive in the given time frame.

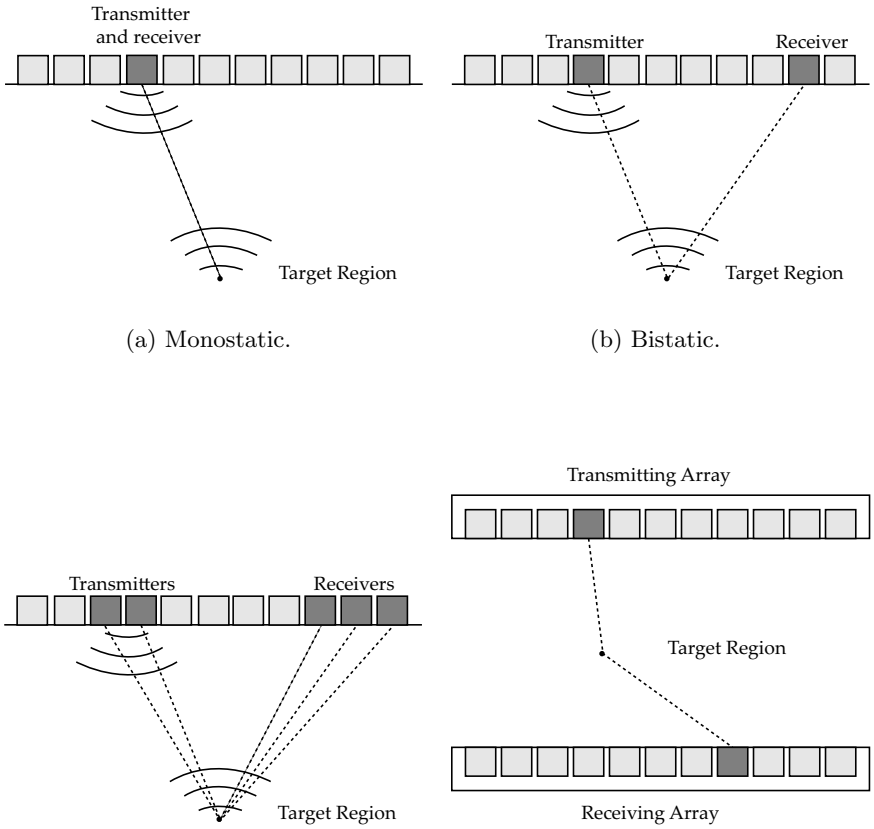
1.1.4 Ultrasonic Measurement Configurations

So far, we have mostly discussed acoustic imaging using a single, possible scanned, transducer. Such imaging systems are cost effective but the acquisition frame rate is insufficient for many applications. Also, the resolution in B-scan images, obtained from mechanically scanned systems, is defined by the beampattern of the transducer. As noted above, high resolution B-scan imaging therefore normally requires focused transducers and, consequently, the focusing depth cannot be changed during acquisition in a flexible manner. On the other hand, array imaging does not have these limitations and therefore, a considerable interest in array imaging methods has been observed.

Array imaging methods can be classified in two groups: *synthetic aperture imaging* and *physical array imaging*. In synthetic aperture imaging the array aperture is obtained by processing sequentially recorded data from different transducer positions, typically using a small transducer with a wide beampattern. Thus, transmit focusing cannot be achieved in synthetic array imaging and focusing can only be performed in the receive mode on the recorded data. On the other hand, a physical array comprising several transducer elements allows for both steering and focusing of the transmitted beam. A physical array can also be rapidly re-focused enabling fast beam sweeping systems. Physical array systems have therefore become popular, in particular, in medical imaging where real-time imaging is of great interest.

Historically, a distinction has been made between *linear* and *phased* (physical) arrays. A linear array system does not have any delay-line circuitry that enables focusing of the transmit beam. Instead, linear arrays are used for fast electronic scanning by switching among the active elements of the array. Phased array systems, on the other hand, are able to both steer and focus the beam by applying suitable time delays on the driving input signals to the array elements. The array that has been used in the experiments presented in this thesis is of the latter type and is consequently capable of focusing the beam.

An array system, whether synthetic or physical, can be classified according to its configuration as shown in Figure 1.5. Monostatic configurations, shown in Figure 1.5(a), are common in synthetic aperture imaging where the



(a) Monostatic.

(b) Bistatic.

(c) Multistatic.

(d) Transmission.

Figure 1.5: Common array measurement methods.

same transducer is used both in transmit and receive mode. Bistatic systems, also known as pitch-catch systems, are used in, for example, NDE applications to detect diffraction echos from crack tips, i.e., *time-of-flight diffraction* (TOFD) imaging [14]. A bistatic system is shown in Figure 1.5(b). Physical arrays can be classified as multistatic, see Figure 1.5(c), since many elements are used simultaneously both in transmit and receive mode. Finally, in transmission imaging, shown in Figure 1.5(d), the specimen is located between two transducers or physical arrays.

The above presented nomenclature will be used throughout this thesis. Note however that we will primarily consider monostatic and multistatic

configurations.

1.2 Conventional Array Imaging

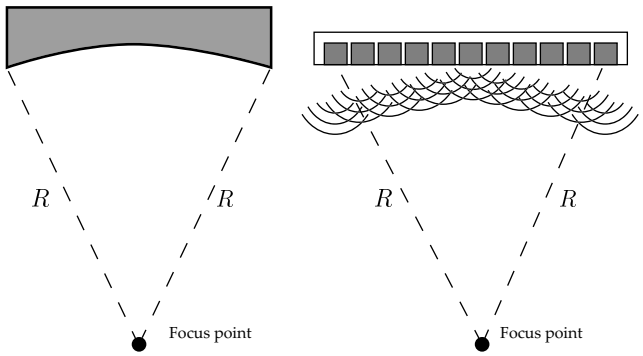
In order to obtain a high resolution image from array data, the data must be compensated for the effect of the wave propagation and the array imaging system. The classical method to solve this problem is *array beamforming*. In this section we will present the basic properties of classical beamforming. This is done in order to introduce the reader not familiar with array imaging to the subject, but also to introduce concepts used later in the thesis.

As mentioned above, the operation principle of a beamformer in conventional array imaging is the same as that of an acoustic lens. The basic idea is to impose delays on the transmitted and/or received signals so that the transmitted acoustic energy or the received signals are combined coherently at a focusing point while other signals are combined incoherently. A beamformer can be implemented using both analog delay-line circuits or using digital hardware [15, 16].

The transmit focusing operation of an acoustical lens and a phased array is illustrated in Figure 1.6. Focusing of the acoustical lens, shown in Figure 1.6(a), is achieved by the curvature of the aperture whereas the phased array is focused at the same point by imposing time-delays that corresponds to the curvature of the lens.

Reception of the beamformer is analogous to the transmission process and the delays of the received signals can be implemented using delay-line circuitry and a summing amplifier or by digital shift operations. The operation of a receive mode beamformer is illustrated by the block-diagram shown in Figure 1.7. The time-delays, for the corresponding array elements, are denoted τ_n for $n = 0, 1, \ldots, L-1$, where L is the number of array elements. The output signal for the nth element is denoted $u_{\mathrm{o}}^{(n)}(t)$ and w_n is an aperture tapering (or apodization) weight. Apodization, which will be discussed in more detail below, is a somewhat crude method, originating from spectral analysis, used to control the sidelobe levels in linear beamforming.

The simple delay-and-sum (DAS) operation of the beamformer does not take any directivity of the array elements into account, hence the elements are treated as point sources/receivers. Assuming that the point source element model is viable for the array, the performance of the DAS beamformer can be analyzed in the frequency domain by considering a sum of phase-shifted spherical waves. The response for a sinusoidal source, at the point



(a) Physical transducer with curvature radius R

(b) Phased array focused at depth R by applying suitable delays to the transmitted signals for each array element.

Figure 1.6: Illustration of transmit focusing using an acoustical lens and a phased array.

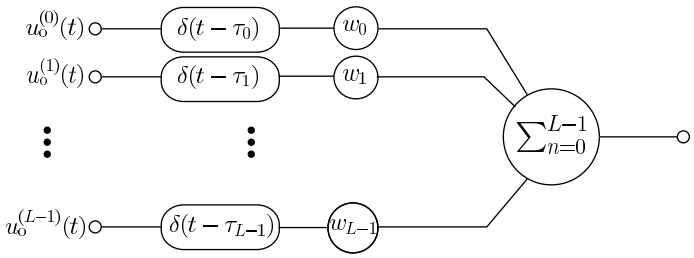


Figure 1.7: Block diagram of a conventional delay-and-sum receive mode beamformer.

(x, z), can then be expressed as [9]

$$H(x,z) = \sum_{n=0}^{L-1} \frac{e^{j\omega(t-R_n/c_\mathrm{p})}}{R_n} w_n e^{j\omega\tau_n}, \tag{1.3}$$

where c_p is the sound speed, w_n an apodization weight, and $R_n = \sqrt{(x_n - x)^2 + z^2}$ is the distance from the observation point to the nth array element positioned at $(x_n, 0)$, see Figure 1.8. The phase-shift due to

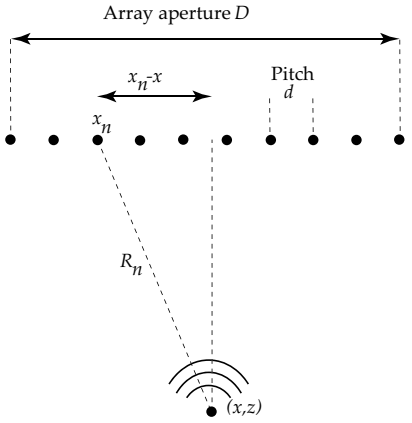


Figure 1.8: A 1D array with an array pitch d and an aperture D.

the propagation distance from the observation point to the nth element is given by $\omega R_n/c_\mathrm{p}$ and the phase-shift $\omega\tau_n$ is due to the focusing operation.

Conventional Beampattern Analysis

In this thesis we are primarily concerned with broadband array imaging and modeling, since broadband modeling offers an accurate description of the imaging system. However, array imaging performance is often analyzed using narrowband and far-field approximations. We will discuss the differences between narrowband and broadband analysis in later chapters of this thesis, and a short introduction to narrowband analysis is, therefore, given here.

To simplify the analysis the *paraxial approximation* is normally applied which means keeping only up to second order terms in the Taylor expansion of R_n in the phase term in (1.3) and only the zero-order term for the amplitude term, $\frac{1}{R_n}$, and also assuming that $(x_n - x)^2 \ll z^2$. In such case a R_n reduces to $R_n \approx z + \frac{(x-x_n)^2}{2z}$ in the phase term.

Consider now focusing at the point $(0, z)$. Using the paraxial approximation, the focusing delays are given by

$$\tau_n \cong \frac{1}{c_\mathrm{p}}(z + \frac{x_n^2}{2z}), \tag{1.4}$$

and the response of a source at (x, z), after focusing will be

$$H(x, z) \propto e^{j\omega(t-\frac{x^2}{2zc_\mathrm{p}})} \sum_{n=0}^{L-1} e^{\frac{j\omega x x_n}{zc_\mathrm{p}}}. \tag{1.5}$$

After performing the summation in (1.5) the following expression defining the magnitude of $H(x, z)$, or the *beampattern* is obtained

$$|H(x, z)| \propto \left| \frac{\sin(\frac{L\pi xd}{\lambda z})}{\sin(\frac{\pi xd}{\lambda z})} \right|, \tag{1.6}$$

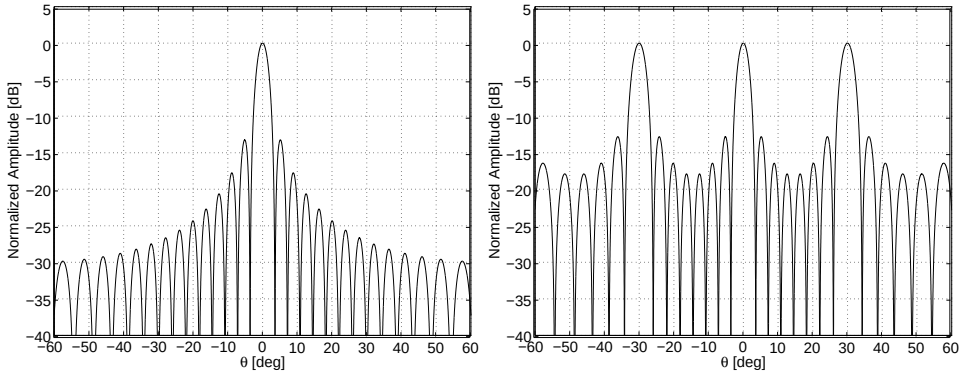
where λ is the acoustic wavelength. The beampattern can also be expressed as a function of the angle, θ, from the center axis of the array to the source point, (x, z), which in the far-field is given by

$$|H(\theta)| \propto \left| \frac{\sin(\frac{L\pi \sin(\theta)d}{\lambda})}{\sin(\frac{\pi \sin(\theta)d}{\lambda})} \right|. \tag{1.7}$$

To illustrate the behavior of the DAS beamformer, two normalized beampatterns, computed using (1.7) for a phased array focused at $(0, z)$, are shown in Figure 1.9. The beampatterns obtained for the same aperture $D = 8\lambda$ are plotted as a function of the angle θ for two different array pitches, $d = \lambda/2$ and $d = 2\lambda$. Noticeable is that the array pitch does not influence the main beam width or the amplitude of the sidelobes. The width of the main beam or the lateral resolution can only be improved by increasing the aperture, D. To see this, consider the standard criterion for resolving two point sources known as the *Rayleigh criterion* [9]. The Rayleigh criterion is based on the idea that two point sources of equal amplitude can be separated if the first source is placed at the maximum point of the beampattern, $x = 0$, and the second source at a point where the beampattern is zero. The first zero of the beampattern occur at the distance

$$d_x = \frac{\lambda z}{Ld} = \frac{\lambda z}{D} = \lambda F, \tag{1.8}$$

where $F = z/D$ is the so-called F-number. The Rayleigh criterion, which is a rough measure of the lateral resolution attainable by a physical lens or



(a) $d = \lambda/2$. (b) $d = 2\lambda$.

Figure 1.9: Beampatterns for a phased array focused at $x = 0$, with an aperture of 8λ, and $w_n \equiv 1$. The beampattern in (a) is for an array with a pitch $d = \lambda/2$ and in (b) for an undersampled array with a pitch $d > \lambda/2$.

a classical array imaging, states that the resolution is determined by the product of the wavelength and the F-number.

Another characteristic feature of classical array imaging can be seen in Figure 1.9(b). Here, the array pitch d is 2λ, and aliased lobes are apparent at $\theta = 30$ deg. These so-called *grating lobes* occurs due to the fact that the sinusoid signal arrives in phase, at all array elements, also for other angles than the focusing direction if d is too large. If strong scatterers are present in the grating lobes, they result in ghost responses in the beamformed image, which is highly undesired. To avoid grating lobes the *array sampling theorem* criterion has to be fulfilled, that is, the array pitch, d, must be less than $\lambda/2$ [9]. Arrays that have a pitch that is larger than $\lambda/2$ are referred to as *undersampled* or *sparse arrays*.

The sidelobes seen in Figure 1.9 is a characteristic feature of classical beamforming, which severely limits contrast in the images. This effect is most apparent when imaging weakly scattering objects in the presence of strong ones. As can be seen in Figure 1.9, the first sidelobe is only about 13 dB below the main lobe, and a strong scatterer in the side-lobe may, therefore, completely obscure a week scatterer in the main lobe.

As mentioned above, the sidelobe levels can be controlled by performing a smooth apodization of the aperture. This is a technique known as windowing

in spectral analysis where the commonly used window functions are, for example, Hamming, Hanning, and Gaussian functions. Figure 1.10 shows an example of apodization where a Hanning-apodized aperture has been compared to an un-apodized, or rectangular-apodized, aperture. Clearly,

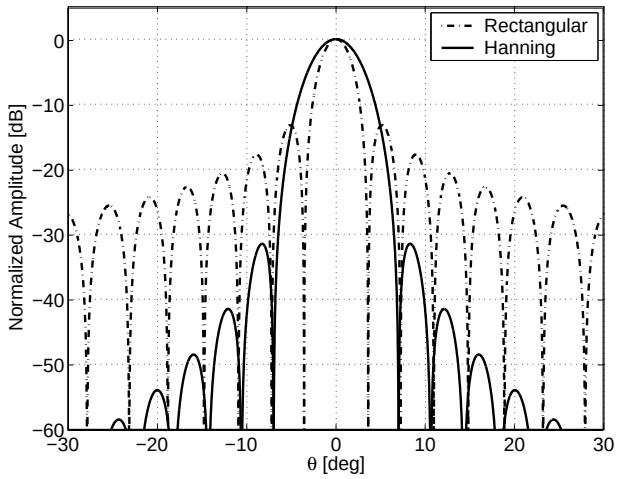


Figure 1.10: Beampatterns for an un-apodized phased array (dash-dotted line) and a Hanning-apodized phased array. The apodization procedure has reduced the sidelobe levels at the expense of a wider main lobe.

the sidelobe levels have been decreased substantially by the apodization procedure. This has, however, occurred at the expense of a significantly increased width of the main lobe. This result can be explained by the fact that apodization effectively reduces the aperture size since apodization functions normally have low weight coefficients near the ends of the aperture. In some applications apodization may be a viable technique but if the data is acquired using a physical array it is highly undesirable to discard a significant portion of the acoustic energy by tapering the aperture.

Above, we have presented the most important properties of classical CW beamforming in order to introduce concepts common in acoustic array imaging. This thesis is, however, concerned with broadband acoustic imaging where the above presented simplified analysis is insufficient. In particular, for broadband signals the grating- and sidelobes will not appear as distinctly as for the CW case, shown in Figure 1.9, since the amplitude of many frequencies will be super-imposed. From (1.6) it can be seen that the beampattern is a function of the wavelength, λ, and the location of side- and

grating lobes will, therefore, change with frequency. Thus, super-imposing the amplitude of many frequencies will smoothen the beampattern [9].

Furthermore, classical array imaging has many deficiencies that limit its usefulness, especially for sparse array data. Particular problems associated with conventional array imaging are:

- The array elements are treated as point sources. That is, the DAS operation does not compensate for the specific diffraction effects due to finite sized array elements.

- The distance between the array elements, array pitch, must be less than half the acoustic wavelength to avoid artifacts from aliased lobes.

- From (1.8) it is apperent that the resolution is determined mainly by the array aperture. In particular, the lateral resolution is not dependent on the quality of the signal, measured for instance by the signal-to-noise ratio (SNR). Nor does the time available for acquiring data influence the resolution (see also [4, 17]). This implies that gathering data over a longer period of time would not improve lateral resolution.

- The sidelobe levels can only be reduced by tapering or apodizing the aperture which has the cost of a wider main-lobe.

- The range- or temporal resolution is determined by the interrogating wavelet. This fact was not mentioned in the discussion above, but the interrogating wavelet is determined by the electro-mechanical properties of the array elements and the driving input signal which results in a band-limited waveform. The temporal resolution is then roughly given by the length of that waveform [18].

The first two issues above are perhaps the most severe since they impose restrictions on the array design, that is: 1) the element size must be small in comparison to the acoustical wavelength and, 2) the array sampling theorem limits of the array pitch. It is of great interest to relax these two restrictions since this will allow for larger sized array elements, capable of transmitting more acoustic power, as well as sparse array designs.

In many medical and NDE applications, where the available apertures are only a few tens of wavelengths and the F-numbers are in the order of unity or more, the resolution achieved using classical beamforming is rather limited. It is, therefore, desirable to develop reconstruction algorithms that

are not restricted by the Rayleigh limit (1.8).[5]

Classical beamforming using apodization is today, essentially, a reminiscence from the past when high-resolution imaging was performed by using lenses. When fast computers are available beamforming in its classical form is not any longer optimal in any sense of *optimal information processing* and as E.T. Jaynes stated [29]: "once one is committed to using a computer to analyze the data, the high-resolution imaging problem is completely changed". The E.T. Jaynes' paper is concerned with optical imaging but the results are equally applicable to acoustic imaging. The conclusion from this paper is nicely summarized by:

1. Keep your optical system clear and open, gathering the maximum possible amount of light (i.e., information).
2. Don't worry about wiggles in the point-spread function; the computer will straighten them out far better than apodization could ever have done, at a small fraction of the cost.
3. For the computer to do this job well, it needs only to know the actual point-spread function $G(x)$, whatever it is. So get the best measurement of $G(x)$ that you can, and let the computer worry about it from then on.
4. What is important to the computer is not the spatial extent of the point-spread function, but its extent in Fourier transform space; over how large "window" in k-space does the PSF give signals above the noise level, thus delivering relevant information to the computer? Apodizing contracts this window by denying us information in high spatial frequencies, associated with the sharp edge of the pupil function. But this is just the information most crucial for resolving fine detail! In throwing away information, it is throwing away resolution. Apodization does indeed "remove the foot;" but it does it by shooting yourself in the foot.

Translated to acoustical imaging this means that one should transmit as

[5]Resolution improvement beyond the Rayleigh limit, is often referred to as *super-resolution* or *wavefield extrapolation* [17]. Super-resolution has been the topic for intensive research in both electromagnetic and acoustic imaging [19–28].

much acoustic energy as possible in all directions[6] in order to maximize the signal-to-noise ratio and then let the computer compensate for interrogated waveforms and diffraction effects.

Our approach which is discussed in more detail in the next section will follow to E.T. Jaynes' suggestions.

1.3 Problem Formulation

The general objective of array imaging, that is, extracting information on the scattering strength of an object under study is, as discussed above, traditionally performed using beamforming. The need for new and more powerful imaging methods is due to the unsatisfactory performance of the traditional DAS beamforming observed especially in the near field.

In the approach presented here the imaging problem is formulated as an *image reconstruction* problem. The basic idea is to create a model of the imaging system that describes the wave propagation and all other properties of the array system with sufficient accuracy compensate for the image degradation introduced by the array system. Problems of this type, known as inverse problems, often require a sophisticated mathematical tools to obtain a satisfactory solution. This is particularly evident for noisy and incomplete data. In particular, when pursuing image reconstruction one must bear in mind that [31]:

- Obtaining true solutions from the imperfect data is impossible.
- Computational and methodological complexity has to be balanced against the quality of the results.

The first issue is related to the fact that the reconstruction performance depends on the amount of information or evidence that is available. That is, the reconstruction performance is directly related to the available background information regarding the scatterers and the imaging system as well as the evidence contained in the data. In particular, if the model does not fit the data well, then the reconstruction performance will be poor, even if the data is almost noise free. On the other hand, an excellent reconstruction performance can be obtained, even for very noisy data, if a strong *a priori* information of the scattering amplitudes is available. An optimal

[6] In medical applications upper limits on average pressure amplitude is often imposed to avoid damage and pain to the patient [30].

strategy should, therefore, properly process both the prior information and the information contained in the data.

The second issue is due to the practical reality: the computational capacity and the available memory resources are limited, which requires compromises in the design of the reconstruction algorithm. The reconstruction algorithms discussed in this thesis are more computationally demanding than classical beamforming but all the algorithms have been implemented on as-of-today standard PC hardware.

As pointed out above, it is vital for the reconstruction performance that both the model is accurate and the reconstruction algorithm makes use of all the available information. These two topics are discussed further below.

1.3.1 The Image Formation Model

The array imaging system has been modeled as a linear time-invariant (LTI) system. Such a system can be modeled using convolutions of the array input signals and the impulse responses corresponding to each array element and observation point. However, it should be emphasized that, even if the system is time-invariant it is in general not position invariant. That is, the impulse responses, corresponding to each array element are functions of the position of the observation point.

To simplify the discussion we consider a pulse-echo measurement using a single array element. Furthermore, we consider scattering from a single scatterer at observation point $\mathbf{r}$. The scatterer is assumed to have a scattering strength $o(\mathbf{r})$. The received signal, $u_{\mathrm{o}}(k)$, can then be modeled as a noise corrupted convolution

$$u_{\mathrm{o}}(t) = o(\mathbf{r})h(\mathbf{r}, t) * u_{\mathrm{i}}(t) + e(t), \tag{1.9}$$

where $h(\mathbf{r}, t)$ is the double-path impulse response of the overall ultrasonic system, $u_{\mathrm{i}}(t)$ is the input signal driving the transducer element and, $e(t)$ is the noise.[7] In order to obtain an accurate model for broadband array imaging the impulse response, $h(\mathbf{r}, t)$, must account for both the diffraction effects associated with the transducer element and the electro-acoustic properties of the transducer. The impulse responses are therefore generally spatially variant.

[7]The "noise" term describes everything that is observed in the data that can not be predicted by the linear model. There is no judgment made whether the error is "random" or "systematic". It can, however, be understood that the noise term is a combination of modeling errors and thermal measurement noise.

To obtain a model useful for array imaging it is not sufficient to consider only a single observation point. If we confine our model to be linear, then (1.9) can easily be extended by summing contributions from many observation points in a suitable region. Throughout this thesis it is assumed that the scatterer's locations are known approximately so that a ROI can be defined within the medium, or specimen, before any experiments are conducted.

Note that if the imaging system is properly modeled all scatterers illuminated by the incident field are confined within the ROI. Otherwise, the fields originating from the scatterers outside the ROI may cause large modeling errors. In particular, it is impossible for the reconstruction algorithm to map the scatterers outside the ROI to the correct location.

The single A-scan model can now be extended to include several elements or transducer positions. The extended discretized model can be expressed as

$$\begin{aligned}\mathbf{y} &= \begin{bmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_{L-1} \end{bmatrix} = \mathbf{P}\mathbf{o} + \mathbf{e} \\ &= \begin{bmatrix} \mathbf{P}_{d(0,0)} & \mathbf{P}_{d(1,0)} & \cdots & \mathbf{P}_{d(N-1,0)} \\ \mathbf{P}_{d(1,0)} & \mathbf{P}_{d(1,1)} & \cdots & \mathbf{P}_{d(N-1,1)} \\ \vdots & \vdots & & \vdots \\ \mathbf{P}_{d(0,L-1)} & \mathbf{P}_{d(1,L-1)} & \cdots & \mathbf{P}_{d(N-1,L-1)} \end{bmatrix} \begin{bmatrix} \mathbf{o}_0 \\ \mathbf{o}_1 \\ \vdots \\ \mathbf{o}_{N-1} \end{bmatrix} + \begin{bmatrix} \mathbf{e}_0 \\ \mathbf{e}_1 \\ \vdots \\ \mathbf{e}_{L-1} \end{bmatrix}\end{aligned} \tag{1.10}$$

where

$$\mathbf{y}_n = \begin{bmatrix} \mathbf{h}_0^{d(\tilde{n},n)} & \mathbf{h}_1^{d(\tilde{n},n)} & \cdots & \mathbf{h}_{M-1}^{d(\tilde{n},n)} \end{bmatrix} \begin{bmatrix} (\mathbf{O})_{0,\tilde{n}} \\ (\mathbf{O})_{1,\tilde{n}} \\ \vdots \\ (\mathbf{O})_{M-1,\tilde{n}} \end{bmatrix} = \mathbf{P}_{d(\tilde{n},n)}\mathbf{o}_{\tilde{n}}. \tag{1.11}$$

The vector $\mathbf{y}_n$ is the A-scan acquired using a transducer at $(x_n, 0)$, $\mathbf{h}_m^{d(\tilde{n},n)}$ is the discrete-time system impulse response corresponding to the observation point $(x_{\tilde{n}}, z_m)$ for the nth transducer position, $\mathbf{o}$ is the vectorized scattering image, and $\mathbf{e}$ is the noise vector. The model (1.10) is described in detail in Chapter 4. The *propagation matrix*, $\mathbf{P}$, contains the impulse responses for *all* observation points to the corresponding array elements. Thus, the model (1.10) is a fully *spatial variant* model of the imaging system.

The benefit of using a spatial variant model, compared to a spatial *invariant* approximations, is that the model is equally valid in the near-field as in the far-field of the aperture.[8] A reconstruction method based on the model, (1.10), should therefore be applicable both in the near- and the far-field.

1.4 Thesis Contributions and Relation to Earlier Work

Ultrasonic imaging has been an active research area for at least 40 years [32]. A substantial part of that research has been on ultrasonic array imaging and beamforming. A considerable effort has been spent to improve the beamforming performance. Here, we present a selection of previous research related to the array imaging problem treated in this thesis.

The previous efforts can roughly be categorized as 1D and 2D processing methods. The 1D methods aim at improving the temporal or range resolution in conventional beamforming. Two common approaches are:

Pulse compression: In synthetic aperture radar (SAR) and synthetic aperture sonar (SAS) applications, the incoming echos are usually time- or range-compressed by correlating them with the transmitted waveform [33–35]. If the correlation pattern has a sharp peak, then the temporal resolution is increased. In SAS frequency sweeped so-called chirp signals are often used, and it usually results in a sharp correlation pattern [36].

1D Deconvolution: In medical and NDT applications it is most common to use broadband pulse excitation and the pulse shape is therefore mainly determined by the electro-mechanical properties of the used transducer. In this case pulse compression will in general not improve temporal resolution much. For these applications 1D deconvolution techniques can be utilized to improve temporal resolution [37, 38].

Various 2D methods have been utilized to improve both temporal and lateral resolution. Two examples are:

[8] The drawback of using a spatial variant model is the rather large amount of computer memory that is required to store all impulse responses.

2D Deconvolution with spatial invariant kernels: 2D deconvolution methods have been applied to ultrasonic array imaging where the forward problem has been approximated using a spatial invariant convolution kernel to allow standard Wiener filtering techniques [39–42].

Inverse filtering methods: Inverse filtering methods have recently been applied to acoustic imaging. In these methods a deterministic approach is taken and the inverse is found by means of singular value decomposition techniques [43–45].

In addition to the approaches presented above, Bayesian linear minimum mean squared error (MMSE) methods have been applied to ultrasonic array imaging, but only for narrow band signals and for simplified models that do not consider the array element size [21–23].

However, no reports on methods using a fully spatial variant model that use a realistic model for finite-sized transducers has been presented so far.

Thesis Contribution

An important contribution of this thesis is a new approach to ultrasonic imaging which incorporates the time-domain ultrasonic model of the imaging system into the Bayesian estimation framework. The most commonly used approach to ultrasonic imaging involves the design of focusing and steering ultrasonic beams using discrete-time focusing laws applied to finite-sized transducers. This approach, which is inspired by geometrical optics, aims at creating analogues of lens systems using ultrasonic array systems. Although formal solutions treating imaging as a deterministic inverse problem have been proposed before, see for example, [43, 45], presenting the ultrasonic imaging as an estimation problem and solving it using a realistic discrete-time model is new.

The contributions of this thesis is described in more detail next.

Model: A linear discrete-time matrix model of the imaging system taking into account model uncertainties has been developed.

Problem Statement: A new approach to ultrasonic imaging has been proposed. It globally optimizes the imaging performance taking into account parameters of the imaging system and the information about the

ROI and errors known a priori.[9] The proposed approach is more general than the traditional imaging methods that maximize the signal-to-noise ratio for a certain point in the ROI using a simplified geometrical optics model of the imaging system.

Solution: Estimation of the scattering strengths has been performed using tools from Bayesian estimation theory. The imaging can be seen as extracting information about the targets in ROI contained in the measurements performed using the imaging setup in presence of measurement errors.

Special Cases: A linear MMSE filter is derived, as a special case of the proposed approach, using Gaussian measurement noise and a Gaussian probability density function (PDF) for the scatterers in the ROI. The linear MMSE solution takes the form of a spatio-temporal filter which deconvolves the distortion caused by the transducer diffraction effects in the ROI using the information contained in the transducer's spatial impulse responses (SIRs). Furthermore, a non-linear (maximum a posteriori) MAP estimator is proposed for imaging targets with positive scattering strengths.

Experiments: The new imaging algorithms have been used on simulated and real ultrasonic data. The algorithms have been applied to monostatic synthetic aperture imaging as well as to parallel array imaging.

Parts of the results presented in this thesis have been published in the following papers:

- F. Lingvall
  "A Method of Improving Overall Resolution in Ultrasonic Array Imaging using Spatio-Temporal Deconvolution" Ultrasonics, Volume 42, pp. 961–968, April 2004.

- F. Lingvall, T. Olofsson and T. Stepinski
  "Synthetic Aperture Imaging using Sources with Finite Aperture—Deconvolution of the Spatial Impulse Response", Journal of the Acoustic Society of America, JASA, vol. 114 (1), July 2003, pp. 225-234.

- F. Lingvall and T. Stepinski
  "Compensating Transducer Diffraction Effects in Synthetic Aperture

[9] Optimization of the input signals driving the array elements are not considered in this work.

Imaging for Immersed Solids", IEEE International Ultrasonic Symposium, October 8-11, Munich, Germany, 2002.

- F. Lingvall, T. Olofsson, E. Wennerström and T. Stepinski "Optimal Linear Receive Beamformer for Ultrasonic Imaging in NDE", Presented at the 6th World Congress of NDT, Montreal, August, 2004.

- T. Stepinski and F. Lingvall "Optimized Algorithm for Synthetic Aperture Imaging", Presented at the IEEE International Ultrasonic Symposium, 24–27 August, Montréal, Canada, 2004.

1.5 Outline of the Thesis

The thesis is divided in three major parts: *Ultrasound Theory and Modeling*, *Methods based on Delay-and-Sum Focusing*, and *Bayesian Ultrasonic Image Reconstruction*. The first part consists of Chapters 2–4, where Chapter 2 and 3 include a comprehensive introduction to acoustic wave propagation theory and the impulse response method, whereupon the discrete-time model, presented in Chapter 4, is based on. The second part consists of Chapter 5 and 6 where classical array imaging is treated and, in particular, the broadband properties are discussed. These methods are then used as benchmarks for the Bayesian reconstruction methods presented in the last part, consisting of Chapters 7 and 8.

Chapter 2: Ultrasound Wave Propagation

In this chapter we present the time domain model used for modeling the transducer radiation process in this thesis. This chapter gives a comprehensive introduction to time domain acoustic modeling for anyone working in the interdisciplinary field of ultrasonic signal processing. Here, the wave propagation phenomena are expressed using a convolutional formulation which should be familiar to people working with signal processing. This chapter also includes the definition of the spatial impulse response which plays a central role in the impulse response method.

Chapter 3: The Spatial Impulse Response

Here the spatial impulse response (SIR) concept is studied in more detail. In particular, the role of the SIRs as a spatial variant filter is discussed. Situations where the spatial impulse responses can be computed analytically is reviewed. The approach that is used to sample analytical SIRs is also presented. Moreover it also includes a description of the numerical method used for situations when no analytical solutions can be found. For such situations the author has developed software tool which is described at the end of the chapter.

Chapter 4: A Discrete Linear Model for Ultrasonic Imaging

This chapter presents the spatial variant discrete model which constitutes the basis for the Bayesian reconstruction methods utilized in this thesis. We describe the modeling of the combined transmission, reception, and the scattering processes. The model is expressed using a matrix formalism which facilitates the derivation of the reconstruction algorithms presented in the last part of the thesis.

Chapter 5: The Synthetic Aperture Focusing Technique

The synthetic aperture focusing technique is a widespread method used in ultrasonic array imaging. This chapter reviews the basic properties of the method with a focus on broadband characteristics and the specific diffraction effects of finite sized transducers.

Chapter 6: Real Aperture Array Beamforming

This chapter presents some aspects of real aperture array imaging, that is, array imaging with physical arrays. In particular the effects that finite sized array elements have on the transmitted acoustic field is discussed. This chapter also introduces the concept of parallel array imaging.

Chapter 7: Ultrasonic Image Reconstruction Methods

The proposed novel Bayesian reconstruction methods for ultrasonic array imaging is introduced here. The chapter also contains a comparison of the

Bayesian methods with common methods for solving so-called ill-posed inverse problems.

Chapter 8: Applications

The aim of this chapter is to report results of experiments and simulations for both synthetic and real array apertures, comparing the Bayesian estimators presented in Chapter 7 with conventional beamforming methods. Aspects treated include the size effects of finite-sized transducers, sparse sampled or under-sampled arrays, as well as parallel array imaging.

Chapter 9: Performance of the Optimal Linear Estimator

The aim of this chapter is to show that Bayesian analysis constitute a useful tool for designing and evaluating ultrasonic array systems. This is due to the fact that Bayesian methods provide a measure of the accuracy of the estimates.

Chapter 10: Summary and Future Work

Here conclusions are drawn and directions for future research are indicated.

1.6 Financial Support

We gratefully acknowledge the financial support obtained from the Swedish Nuclear Fuel and Waste Management Co. (SKB).

Part I

Ultrasound Theory and Modeling

CHAPTER 2

Ultrasound Wave Propagation

THIS thesis is concerned with acoustic array signal processing and for that purpose a model of the acoustic array imaging sytem is needed. In particular, the transmission, propagation, scattering as well as the reception processes must be modeled. The purpose of acoustic wave modeling is to calculate the acoustic field in a medium given the field sources, geometry, and the initial and boundary conditions of the problem. This typically involves finding a solution to the wave equation. As mentioned in Chapter 1, only longitudinal waves are considered in this thesis, and longitudinal waves, both in fluids and solids, are well described by the linear *acoustic* wave equation.[1]

This chapter is concerned with modeling the transmission and propagation processes or, in other words, the acoustic radiation process of an ultrasonic transducer. The scattering and reception process will be discussed later in Chapter 4. It will be shown here that the pressure at an observation point can be expressed as a convolution between the normal velocity on the transducer surface and an *impulse response* that is determined by the geometry of the transducer and the boundary conditions. This impulse response, known as the *spatial impulse resonse* (SIR), is the foundation of

[1]Linear wave propagation describes many acoustic phenomena surprisingly well. Wave propagation in a fluid will be nearly linear if the fluid is uniform and in equilibrium, and the viscosity and thermal conduction can be neglected, as well as if the acoustic pressure generated by a transducer is small compared to the equilibrium pressure [46]. In ultrasonic array imaging linear wave propagation is normally assumed which also is the case in this thesis.

the ultrasonic wave modeling performed here.

The chapter is organized as follows: The acoustic wave equation is introduced in Section 2.1, and Section 2.2 reviews the method of Green's functions for finding solutions to the wave equation. Particlar solutions, given by the initial and boundary conditions, are discussed in Section 2.3. The presented solutions take the form of three integral expressions that correspond to, the rigid baffel, the soft baffle, and the free-space boundary conditions, respectively. All these integral formulas can be expressed in a convolutional form which is attractive since wave propagation can now be modeled using linear systems theory. In Section 2.4 it is finally shown how the sound field can be explicitly expressed as convolutions between the transducer's normal velocity and the SIRs.

2.1 The Acoustic Wave Equation

Wave propagation in an isotropic compressible medium containing an acoustic source can be described by the acoustic wave equation [46],

$$\nabla^2 p(\mathbf{r},t) - \frac{1}{c_{\mathrm{p}}^2}\frac{\partial^2}{\partial t^2} p(\mathbf{r},t) = f_p(\mathbf{r},t), \tag{2.1}$$

where the scalar function, $p(\mathbf{r},t)$ is the pressure at $\mathbf{r}$ and time t, c_{p} is the sound speed of the medium, and $f_p(\mathbf{r},t)$ is the source density, or the driving function of the source. The wave equation can also be expressed in terms of a velocity potential $\phi(\mathbf{r},t)$,

$$\nabla^2 \phi(\mathbf{r},t) - \frac{1}{c_{\mathrm{p}}^2}\frac{\partial^2}{\partial t^2} \phi(\mathbf{r},t) = f_\phi(\mathbf{r},t). \tag{2.2}$$

The pressure and the potential is related through,

$$p(\mathbf{r},t) = \rho_0 \frac{\partial \phi(\mathbf{r},t)}{\partial t}, \tag{2.3}$$

and ρ_0 is the equilibrium density of the medium which is assumed to be constant.

The wave equation (2.1) describes the wave propagation in space and time for a driving function $f_p(\mathbf{r},t)$. The driving function can be seen as the input signal to the acoustical system, and a solution to the wave equation is the response in space and time for that input signal.

Here we are interested in obtaining a solution where the input signal is the normal velocity on the active area of the transducer. An analythical solution to the wave equation is in general difficult to obtain except for some special cases. Fortunately, a piezoelectric transducer can be well described as a rigid piston due to the large difference in acoustic impedance between the piezoelectric material and the medium (water). Under such conditions, a solution can be found by the method of Green's functions which is discussed in the next section.

2.2 Green's Functions

In the theory of linear time-invariant systems, the impulse response of a system fully describes its properties. Green's function has an analogous property for the wave equation. To see this, let us first introduce the differential operator,

$$\mathcal{L} = \nabla^2 - \frac{1}{c_{\mathrm{p}}^2}\frac{\partial^2}{\partial t^2}. \tag{2.4}$$

The wave equation (2.1) can then be expressed as

$$\mathcal{L}p(\mathbf{r},t) = f_p(\mathbf{r},t). \tag{2.5}$$

Since $\mathcal{L}$ is a differential operator, the inverse operator must be an integral operator. Green's function is the kernel for this integral operator. Physically Green's function is the solution to a linear partial differential equation (PDE) for a unit impulse disturbance [47]. Here the PDE is the wave equation, (2.1), which becomes,

$$\nabla^2 g(\mathbf{r}-\mathbf{r}_0, t-t_0) - \frac{1}{c_{\mathrm{p}}^2}\frac{\partial^2}{\partial t^2} g(\mathbf{r}-\mathbf{r}_0, t-t_0) = \delta(\mathbf{r}-\mathbf{r}_0, t-t_0) \tag{2.6}$$

for a unit impulse source, $\delta(\mathbf{r},t)$, at $\mathbf{r}_0$ applied at time t_0. If $g(\mathbf{r}-\mathbf{r}_0, t-t_0)$ is known for the problem given by (2.1) and the boundary- and initial conditions, the solution for an arbitrary source $f_p(\mathbf{r},t)$ can be obtained by the following procedure:

First, multiply (2.1) with $g(\mathbf{r},t)$, and (2.6) with $p(\mathbf{r},t)$, and then the substract the results, to obtain,

$$\begin{aligned} &g(\mathbf{r}-\mathbf{r}_0, t-t_0)\nabla^2 p(\mathbf{r},t) - p(\mathbf{r},t)\nabla^2 g(\mathbf{r}-\mathbf{r}_0, t-t_0) \\ &\quad = p(\mathbf{r},t)\delta(\mathbf{r}-\mathbf{r}_0, t-t_0) - g(\mathbf{r}-\mathbf{r}_0, t-t_0) f(\mathbf{r},t). \end{aligned} \tag{2.7}$$

Interchange $\mathbf{r}$ with $\mathbf{r}_0$ and t with t_0, and integrate with respect to t_0 and $\mathbf{r}_0$ over time and the volume V defining our region-of-interest and use the sifting property of the delta function. We thus obtain, [48]

$$\begin{aligned}\int_{-\infty}^{t^+}\int_V \Big(g(\mathbf{r}-\mathbf{r}_0,t-t_0)\nabla^2\phi(\mathbf{r}_0,t_0) - p(\mathbf{r}_0,t_0)\nabla^2 g(\mathbf{r}-\mathbf{r}_0,t-t_0)\Big)d\mathbf{r}_0\,dt_0 \\ = p(\mathbf{r},t) - \int_{-\infty}^{t+}\int_V g(\mathbf{r}-\mathbf{r}_0,t-t_0)f(\mathbf{r}_0,t_0)d\mathbf{r}_0\,dt_0,\end{aligned} \tag{2.8}$$

where the notation t^+ means $t+\epsilon$, for an arbitrary small ϵ, thereby avoiding an integration ending at the peak of a delta function.

Now, by using Green's theorem [17], stating that two functions, $u(\mathbf{r},t)$ and $v(\mathbf{r},t)$, with continous first and second order partial derivatives (within, and on a surface S inclosing a volume V) are related through

$$\begin{aligned}&\int_V \big(u(\mathbf{r},t)\nabla^2 v_{\mathrm{n}}(\mathbf{r},t) - v_{\mathrm{n}}(\mathbf{r},t)\nabla^2 u(\mathbf{r},t)\big)\,d\mathbf{r} \\ &= \int_S \left(u(\mathbf{r},t)\frac{\partial v_{\mathrm{n}}(\mathbf{r},t)}{\partial n} - v_{\mathrm{n}}(\mathbf{r},t)\frac{\partial u(\mathbf{r},t)}{\partial n}\right) dS,\end{aligned} \tag{2.9}$$

the volume integral on the left hand side of (2.8) can be turned into a surface integral so that the pressure, $p(\mathbf{r},t)$, can be found to be

$$\begin{aligned}p(\mathbf{r},t) = &\int_{-\infty}^{t^+}\int_V g(\mathbf{r}-\mathbf{r}_0,t-t_0)f_p(\mathbf{r}_0,t_0)d\mathbf{r}_0\,dt_0 + \\ &\int_{-\infty}^{t^+}\int_S \Big(g(\mathbf{r}-\mathbf{r}_0,t-t_0)\frac{\partial}{\partial n}p(\mathbf{r}_0,t_0) \\ &- p(\mathbf{r}_0,t_0)\frac{\partial}{\partial n}g(\mathbf{r}-\mathbf{r}_0,t-t_0)\Big)dS\,dt_0.\end{aligned} \tag{2.10}$$

Here n is the outward normal on S. Eq. (2.10) is an integral operator that expresses the pressure at $\mathbf{r}$ both for a source within the volume V and on the surface of V.

When modeling the acoustic radiation from an ultrasonic transducer the source, i.e., the transducer, is located on S and and the source term inside V is then not present, $f_p(\mathbf{r},t)=0$. Eq. (2.10) then reduces to the time-domain

Helmholtz-Kirchhoff, or the *Helmholtz* equation

$$p(\mathbf{r},t) = \int_{-\infty}^{t^+}\int_S \Big(p(\mathbf{r}_0,t_0)\frac{\partial g(\mathbf{r}-\mathbf{r}_0,t-t_0)}{\partial n} - g(\mathbf{r}-\mathbf{r}_0,t-t_0)\frac{\partial p(\mathbf{r}_0,t_0)}{\partial n}\Big) dS\, dt_0. \tag{2.11}$$

To obtain an expression for the pressure waveform in the transmit process we need to consider the initial- and boundary conditions which are discussed next.

2.3 Boundary Conditions and Related Integrals

The Helmholtz equation (2.11) may have different forms for different boundary conditions. The Rayleigh integral, the Rayleigh-Sommerfeld integral, and the Kirchhoff integral are three forms for three different boundary conditions that we will considered here.

2.3.1 The Rayleigh Integral

The boundary conditions determine the impulse response, i.e., Green's function, for the problem at hand. If a point source is in free space, then the radiated waves will spread spherically and $g(\mathbf{r}-\mathbf{r}_0,t-t_0)$ is given by the *free-space* Green's function [46]

$$g^{\mathrm{F}}(\mathbf{r}-\mathbf{r}_0,t-t_0) = \frac{\delta\left(t-t_0-|\mathbf{r}-\mathbf{r}_0|/c_p\right)}{4\pi|\mathbf{r}-\mathbf{r}_0|}, \tag{2.12}$$

where the super-script, F, denotes free-space conditions.

If some part of S is rigid, then the normal derivative of the pressure on that part of the boundary is zero. A wave can now be reflected by the rigid, or hard, boundary so the pressure at $\mathbf{r}$, from a source at $\mathbf{r}_0$, is for a planar boundary, the sum of two spherical waves: the direct and the reflected wave, see Figure 2.1. Green's function then becomes

$$g^{\mathrm{H}}(\mathbf{r}-\mathbf{r}_0,t-t_0) = \frac{\delta\left(t-t_0-|\mathbf{r}-\mathbf{r}_0|/c_p\right)}{4\pi|\mathbf{r}-\mathbf{r}_{\mathrm{O}}|} + \frac{\delta\left(t-t_0-|\mathbf{r}-\mathbf{r}_0'|/c_p\right)}{4\pi|\mathbf{r}-\mathbf{r}_0'|}, \tag{2.13}$$

where the super-script, H, denotes hard or rigid boundary conditions.

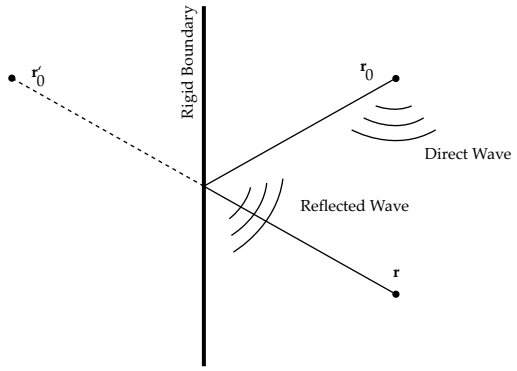


Figure 2.1: Illustration of a wave reflected by a rigid boundary for a source at $\mathbf{r}_0$.

If the boundary is located in the xy-plane the point $\mathbf{r}_0' = (x_0, y_0, -z_0)$ is then the mirror point of $\mathbf{r}_0$. On the boundary, $z_0 = 0$, the normal derivative of Green's function (2.13) vanishes, hence the contribution from the rigid plane to the surface integral in (2.11) will be zero and (2.13) becomes

$$g^{\mathrm{H}}(\mathbf{r}-\mathbf{r}_0, t-t_0) = \frac{\delta\left(t-t_0-|\mathbf{r}-\mathbf{r}_0|/c_p\right)}{2\pi|\mathbf{r}-\mathbf{r}_{\mathrm{O}}|} = 2g^{\mathrm{F}}(\mathbf{r}-\mathbf{r}_0, t-t_0), \quad (2.14)$$

with

$$\frac{\partial}{\partial n} g^{\mathrm{H}}(\mathbf{r}-\mathbf{r}_0, t-t_0) = 0. \quad (2.15)$$

Now, if the source, i.e., the transducer, is located in the rigid plane, often referred to as the *rigid baffle* as illustrated in Figure 2.2, then the baffle (S_{B}) will not contribute to the field.

By using the *Sommerfeld radiation condition*, where one lets the volume V expand to infinity, the contributions from the remaining part of the surface (depicted $\sum$ in Figure 2.2) will cancel [17]. Thus, the only area that will contribute to the surface integral is the active area of the transducer. If the transducer is planar, then the wave in the vicinity will also be planar, resulting in that the normal derivative of the pressure, its time derivative, and the time derivative of the source velocity will be proportional [9, 49]

$$\frac{\partial p(\mathbf{r}_0, t)}{\partial n} = -\frac{1}{c_{\mathrm{p}}}\frac{\partial p(\mathbf{r}_0, t)}{\partial t} = -\rho_0 \frac{\partial}{\partial t} v_{\mathrm{n}}(\mathbf{r}_0, t). \quad (2.16)$$

Substituting (2.16) and (2.14) into (2.11) results in the well known *Rayleigh*

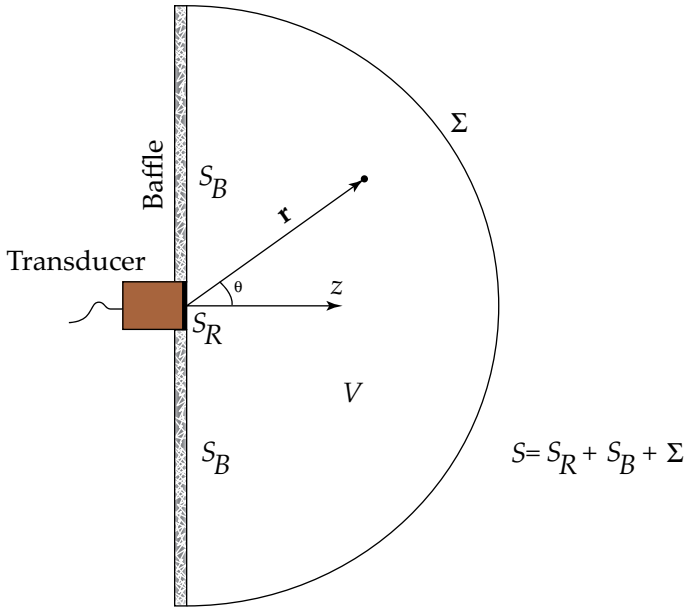


Figure 2.2: Illustration of the decomposition of the boundary S enclosing the volume V.

integral [50]

$$p(\mathbf{r},t) = \rho_0 \frac{\partial}{\partial t} \int_{-\infty}^{t^+} \left(\int_{S_\mathrm{R}} v_\mathrm{n}(\mathbf{r}_0,t_0) \frac{\delta\left(t - t_0 - |\mathbf{r}-\mathbf{r}_0|/c_p\right)}{2\pi|\mathbf{r}-\mathbf{r}_0|} dS_\mathrm{R} \right) dt_0 \quad (2.17)$$

which is commonly used to model acoustic radiation from piezoelectric ultrasonic transducers.

2.3.2 The Rayleigh-Sommerfeld Integral

If we consider a baffle that is acoustically soft, the so-called *pressure-release* baffle conditions, then the pressure on the baffle will be zero. Green's func-

tion and its normal derivative on the baffle ($\mathbf{r}_0' = \mathbf{r}_0$) will then be [2]

$$\begin{aligned} g^{\mathrm{S}}(\mathbf{r}-\mathbf{r}_0, t-t_0) &= \frac{\delta\left(t-t_0-|\mathbf{r}-\mathbf{r}_0|/c_p\right)}{4\pi|\mathbf{r}-\mathbf{r}_{\mathrm{O}}|} - \frac{\delta\left(t-t_0-|\mathbf{r}-\mathbf{r}_0'|/c_p\right)}{4\pi|\mathbf{r}-\mathbf{r}_0'|} \\ &= 0 \\ \frac{\partial}{\partial n} g^{\mathrm{S}}(\mathbf{r}-\mathbf{r}_0, t-t_0) &\approx \frac{\delta'\left(t-t_0-|\mathbf{r}-\mathbf{r}_0|/c_p\right)}{4\pi c_{\mathrm{p}}|\mathbf{r}-\mathbf{r}_0|} 2\cos(\theta) \end{aligned} \tag{2.18}$$

where θ is the angle between the normal vector $\mathbf{n}$ and $\mathbf{r}-\mathbf{r}_0$ and $\delta'(\cdot)$ is the first derivative of the delta function.[3] By substituting (2.18) and $p(\mathbf{r}_o, t) = 0$ into (2.11), the Helmholtz equation reduces to the *Rayleigh-Sommerfeld* integral formula

$$p(\mathbf{r},t) = \rho_0 \frac{\partial}{\partial t} \int_{-\infty}^{t^+} \left(\int_{S_{\mathrm{R}}} v_{\mathrm{n}}(\mathbf{r}_0, t_0) \frac{\delta\left(t-t_0-|\mathbf{r}-\mathbf{r}_0|/c_p\right)}{4\pi|\mathbf{r}-\mathbf{r}_0|} 2\cos(\theta) dS_{\mathrm{R}} \right) dt_0. \tag{2.19}$$

2.3.3 The Kirchhhoff Integral

For the free-space case, none of the terms in the Helmholtz equation disappears and the normal derivative at the plane source of the free-space Green's function becomes

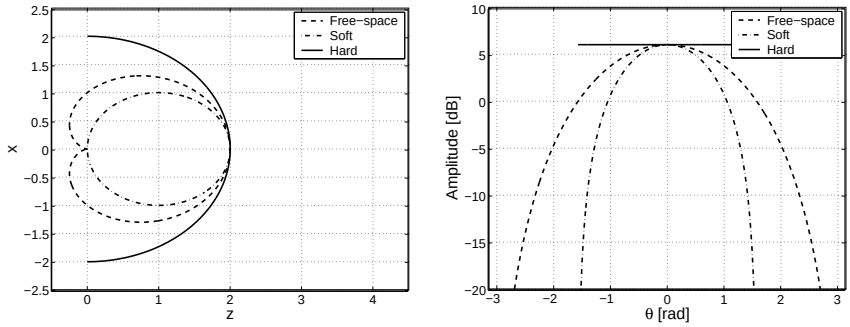
$$\frac{\partial}{\partial n} g^{\mathrm{S}}(\mathbf{r}-\mathbf{r}_0, t-t_0) \approx \frac{\delta'\left(t-t_0-|\mathbf{r}-\mathbf{r}_0|/c_p\right)}{4\pi c_{\mathrm{p}}|\mathbf{r}-\mathbf{r}_0|} \cos(\theta), \tag{2.20}$$

where the super-script, S, denotes soft boundary conditions. Substituting the free-space Green's function and its derivative (2.20) into Helmholtz equation results in the *Kirchhoff integral* formula

$$p(\mathbf{r},t) = \rho_0 \frac{\partial}{\partial t} \int_{-\infty}^{t^+} \left(\int_{S_{\mathrm{R}}} v_{\mathrm{n}}(\mathbf{r}_0, t_0) \frac{\delta\left(t-t_0-|\mathbf{r}-\mathbf{r}_0|/c_p\right)}{4\pi|\mathbf{r}-\mathbf{r}_0|} (1-\cos(\theta)) dS_{\mathrm{R}} \right) dt_0. \tag{2.21}$$

In summary, by comparing (2.17), (2.19), and (2.21) tha difference between the three different boundary conditions is a direction dependent coefficient, $\alpha(\theta)$, which is $(1-\cos(\theta))$ for the free-space conditions, $2\cos(\theta)$ for the soft baffle, and a constant 2 for the hard baffle. The angular dependence of the direction coefficient is shown in Figure 2.3. Normally the

[2] The approximation for the normal derivative of Green's function is valid for $|\mathbf{r}-\mathbf{r}_0| \gg \frac{\lambda}{2\pi}$ where λ is the wavelength corresponding to the highest frequency in the frequency band



(a) Polar diagram.

(b) Directional coefficient as a function of θ.

Figure 2.3: Angular dependence of the directional coefficient, $\alpha(\theta)$, for: a hard baffle $\alpha(\theta) = 2$, a soft baffle $\alpha(\theta) = 2\cos(\theta)$, and for free-space boundary conditions $\alpha(\theta) = 1 - cos(\theta)$.

hard baffle conditions (i.e., the Rayleigh integral) is used for modeling wave propagation using an piezoelectric ultrasonic transducer immersed in water.

2.4 The Convolution Integral

In Section 2.3 we found that the Kirchhhoff, Rayleigh, and Rayleigh-Sommerfeld integrals all shared the form

$$p(\mathbf{r},t) = \rho_0 \frac{\partial}{\partial t} \int_{-\infty}^{t+} \left(\int_{S_\mathrm{R}} v_\mathrm{n}(\mathbf{r}_0,t_0) \frac{\delta\left(t - t_0 - |\mathbf{r}-\mathbf{r}_0|/c_p\right)}{4\pi|\mathbf{r}-\mathbf{r}_0|} \alpha(\theta) dS_\mathrm{R} \right) dt_0 \tag{2.22}$$

where $\alpha(\theta)$ is the direction coefficient that depends on the boundary condition. If the source velocity is constant over the active area of the transducer (i.e., $v_\mathrm{n}(\mathbf{r}_0,t_0) = v_\mathrm{n}(t_0)$), then $v_\mathrm{n}(\mathbf{r}_0,t_0)$ can be moved outside the surface

of interest [51].

[3]The derivative of a delta function has the property $\int_{-\infty}^{\infty} \delta'(x) f(x+a) dx = -f'(a)$ [48].

integral, yielding

$$\begin{aligned} p(\mathbf{r},t) &= \rho_0 \frac{\partial}{\partial t} \int_{-\infty}^{t+} v_{\mathrm{n}}(t_0) \left(\int_{S_{\mathrm{R}}} \frac{\delta\left(t - t_0 - |\mathbf{r} - \mathbf{r}_0|/c_p\right)}{2\pi|\mathbf{r} - \mathbf{r}_{\mathrm{R}}|} \alpha(\theta) dS_{\mathrm{R}} \right) dt_0 \\ &= \rho_0 \frac{\partial}{\partial t} v_{\mathrm{n}}(t) * \left(\int_{S_{\mathrm{R}}} \frac{\delta\left(t - |\mathbf{r} - \mathbf{r}_0|/c_p\right)}{2\pi|\mathbf{r} - \mathbf{r}_0|} \alpha(\theta) dS_{\mathrm{R}} \right). \end{aligned} \tag{2.23}$$

Thus, the pressure at $\mathbf{r}$ can be expressed as a temporal convolution between the two factors,

$$\rho_0 \frac{\partial}{\partial t} v_{\mathrm{n}}(t), \tag{2.24}$$

and

$$\int_{S_{\mathrm{R}}} \frac{\delta\left(t - |\mathbf{r} - \mathbf{r}_0|/c_p\right)}{4\pi|\mathbf{r} - \mathbf{r}_0|} \alpha(\theta) dS_{\mathrm{R}}. \tag{2.25}$$

The normal velocity, $v_{\mathrm{n}}(t_0)$, depends on the electro-mechanical transfer function of the transducer combined, i.e., convolved, with the excitation signal of the transducer. If the excitation signal is a delta function, then the factor (2.24) can be seen as the forward *electrical impulse response* of the transducer.

The second factor, (2.25), describes the acoustic wave propagation from the tranducer surface to the observation point, $\mathbf{r}$. This factor is then the response at $\mathbf{r}$ that one would obtain for an impulsive transducer surface velocity. The factor (2.25) is commonly called the *spatial impulse response* (SIR). The spatial impulse response depends of the shape and size of the active area (S_{R}) of the transducer, and it is this factor that determines the directivity patterns of the transducer. The spatial impulse response has a central role in this thesis and Chapter 3 is devoted entirely to this issue.

CHAPTER 3

The Spatial Impulse Response

THE spatial impulse response (SIR) concept plays a central role in this thesis. It is the basis of the discrete wave propagation model on which the reconstruction algorithms introduced in Chapter 7 is based on. The SIR method, also known as the *impulse response method*, is established from linear acoustics. Thus, wave propagation and scattering are treated as a linear time-invariant (LTI) system. It is a time domain approach to model acoustic radiation and reception.

The impulse response method is based on the Rayleigh integral formula derived in Chapter 2 because the transducer is normally treated as a planar rigid piston. For a planar piston the normal velocity all over the transducer's active area is uniform. The pressure at an observation point can, therefore, be described by the convolution of normal velocity and SIR of the transducer, where the normal velocity is determined by the driving input signal and the electro-mechanical characteristics of the transducer.

The commonly used transducers all have a band-pass characteristics. Thus, the frequencies outside the active frequency band will therefore not be sent into the insonified medium. Also this effect can be modeled as a convolution of the transducer's electro-mechanical impulse response and the input signal, and the combined imaging system can thus be modeled as a LTI system. Consequently, if the electro-mechanical impulse response of the transducer, the transducer shape, and the input signal are known, the total system response of the transducer can be modeled by linear filtering.

Evidently, to model the ultrasonic imaging system, the SIRs associated with the transducer are needed. Analytical solutions to SIRs exist for a few geometries, but one must in general resort to numerical methods. Section 3.1 discusses some important cases where analytical solutions do exist. These *time continuous* solutions are however not practical since all signals acquired are normally sampled and time discrete models are therefore needed. Section 3.2 describes the sampling approach of the analytical SIRs used in this thesis. The total system response is determined by both the electrical impulse response and the SIRs. The properties and influence of the electrical impulse response are discussed in Section 3.3.

As mentioned above, there exist no analytical analytical solutions for many transducer geometries, and in such situation a method based on the *discrete representation* (DR) computational concept is used here [52, 53]. The DR method is very flexible in the sense that complex transducer shapes as well as arbitrary focusing methods easily can be modeled. Another benefit of the DR method is that the SIRs are directly computed in a discrete form. A short description of the DR method, and a numerical package based on the DR method that has been developed by the author, is presented in the final Section 3.4 of this chapter.

3.1 Analytical Solutions for some Canonical Transducer Types

In this section a few analytical SIR solutions are presented. These solutions will be used later in scattering and transducer radiation modeling problems. The first solution considered is the point source that is commonly used in array imaging. In this thesis, the point source model is mainly used to modeling scattering which is discussed in more detail in Chapter 4.[1] The other analytical solutions considered are the line and the circular disc solutions that will be used to model synthetic aperture experiments in Chapters 5 and 8.

3.1.1 The Point Source

A point source is a common model used for a source that has an aperture considerably smaller than the acoustical wavelength. A point source emits

[1] In pulse-echo imaging a scatterer can be seen a source of the scattered wave field.

spherical waves, and thus has an angle-independent SIR which is given by the free-space Green's function, cf. (2.12),

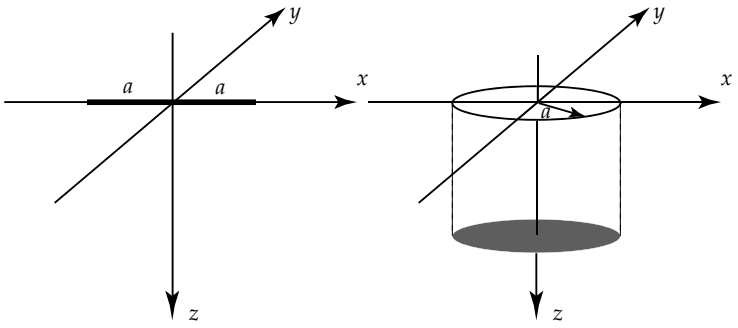
$$h(\mathbf{r}, t) = s_{\mathrm{e}} \frac{\delta\left(t - |\mathbf{r} - \mathbf{r}_{\mathrm{T}}|/c_{\mathrm{p}}\right)}{4\pi|\mathbf{r} - \mathbf{r}_{\mathrm{T}}|}, \tag{3.1}$$

where $\mathbf{r}_{\mathrm{T}}$ is the position of the point source and $\mathbf{r}$ the position of the observation point. Since the SIR is a delta function the pulse shape will not depend on the observation position, only the pulse's amplitude will be scaled with the factor $1/(4\pi|\mathbf{r} - \mathbf{r}_{\mathrm{T}}|)$. The scale factor, s_{e} in (3.1), is the cross-sectional area or the *elementary surface* of the point source [54]. Note that the elementary surface must be small enough so that the source behaves as a unit point source in all aspects except for the scale factor s_{e}.

In traditional ultrasonic array imaging, the array elements are treated as point sources and the total array response is modeled as a superposition of delayed point source responses. The point source is also commonly used to model scattering. In this case s_{e} is thought of as the cross-section area of a small scatterer.

3.1.2 The Line Source

A line source, or a uniformly excited line strip, can be used as an approximation of a narrow strip transducer or for modeling scattering from a wire target that is commonly used in ultrasonic imaging. Here the line strip,



(a) A line source. (b) A circular disc source.

Figure 3.1: Geometries of a line and a circular disc source.

with length $2a$, is assumed to be located on the x-axis, centered at $x = 0$, see Figure 3.1(a). The SIR of a uniformly excited line strip is given by [55]

$$\text{for } |x| \leq a,$$
$$h(x,z,t) = s_{\mathrm{e}} \cdot \begin{cases} 0, & t < t_z \\ 2\frac{1}{\sqrt{t^2 - t_z^2}}, & t_z < t \leq t_1 \\ \frac{1}{\sqrt{t^2 - t_z^2}}, & t_1 < t \leq t_2 \\ 0, & t > t_2 \end{cases} \tag{3.2}$$
$$\text{and for } |x| > a,$$
$$h(x,z,t) = s_{\mathrm{e}} \cdot \begin{cases} 0, & t \leq t_1 \\ \frac{1}{\sqrt{t^2 - t_z^2}}, & t_1 < t \leq t_2 \\ 0, & t > t_2 \end{cases}$$

where $t_z = z/c_{\mathrm{p}}$ is the start time of the SIR when $|x| \leq a$ and $t_{1,2} = t_z\sqrt{1 + \left(\frac{a \mp |x|}{z}\right)^2}$. Due to the rotational symmetry of the SIRs for a line source, only the two-dimensional case is considered here. That is, if the SIRs in the xz-plane are known than all other SIRs can be found by a rotation of the plane around the x-axis.[2]

Note that similar to the point source it is necessary to use a weight, s_{e}, for, in this case, the yz-direction cross-sectional area of the line source.

3.1.3 The Circular Disc

Circular transducers are commonly used in ultrasonic imaging applications. Typical examples are NDT/NDE, sonar, and scanned synthetic aperture imaging applications.

The SIR of a circular disc has an analytical solution when the observation point is inside the aperture of the disc, $\sqrt{x^2 + y^2} \leq a$ where a is the transducer radius, and when the observation point is outside the aperture, see Figure 3.1(b) [56]. The disc is assumed to be located in the xy-plane centered at $x = y = 0$. Let first r denote the distance from the center axis

[2]The 3D solution for a line source can be found in [49].

of the disc, $r = \sqrt{x^2 + y^2}$, then the disc SIR is given by

$$\text{for } r \leq a$$

$$h(\mathbf{r}, t) = \begin{cases} 0, & t \leq t_z \\ c_\mathrm{p}, & t_z \leq t \leq t_1 \\ \frac{c_\mathrm{p}}{\pi} \cos^{-1}\left(c_\mathrm{p}^2 \frac{t^2 - t_z^2 + t_r^2 - a/c_\mathrm{p}^2}{2t_r\sqrt{t^2 - t_z^2}} \right), & t_1 < t \leq t_2 \\ 0, & t > t_2 \end{cases} \tag{3.3}$$

$$\text{for } r > a$$

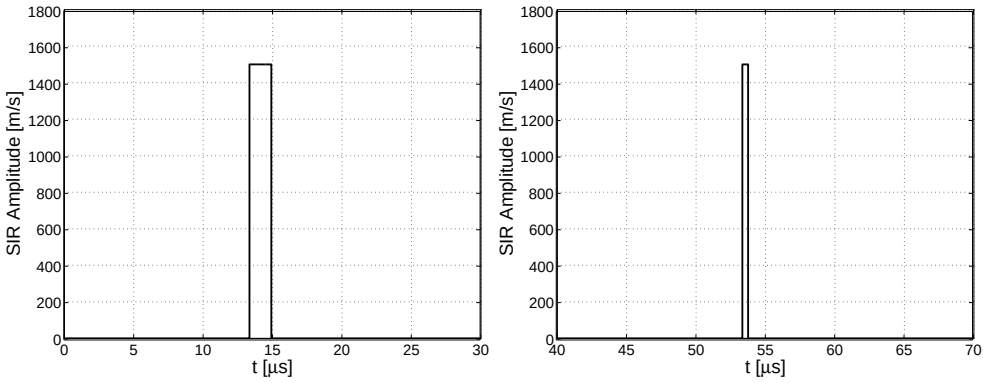
$$h(\mathbf{r}, t) = \begin{cases} 0, & t \leq t_1 \\ \frac{c_\mathrm{p}}{\pi} \cos^{-1}\left(c_\mathrm{p}^2 \frac{t^2 - t_z^2 + t_r^2 - a/c_\mathrm{p}^2}{2t_r\sqrt{t^2 - t_z^2}} \right), & t_1 < t \leq t_2 \\ 0, & t > t_2 \end{cases}$$

where $t_z = z/c_\mathrm{p}$ is the earliest time that the wave reaches the observation point $\mathbf{r}$ when $r \leq a$, $t_r = r/c_\mathrm{p}$, and $t_{1,2} = t_z\sqrt{1 + \left(\frac{a \mp r}{z}\right)^2}$ are the propagation times corresponding to the edges of the disc that are closest and furtherest away from $\mathbf{r}$, respectively.

Noticeable is that the pulse amplitude of the on-axis SIR is constant regardless of the distance to the observation point.[3] The duration of the on-axis SIR is given by $t_1 = a/c_\mathrm{p}$ at $z = 0$. As the distance increases the duration, $t_1 - t_z$, of the SIR becomes shorter, and for large z it approaches to the delta function. The transducer size effects are therefore most pronounced in the near-field. This is illustrated in Figure 3.2 where the on-axis SIRs at $z = 20$ and $z = 80$ mm, respectively are shown. The duration at $z = 20$ is longer than that of $z = 80$ and if the distance, z, increases then the on-axis SIR will approach to a delta function, cf. Figures 3.2(a) and (b).

The analytical SIR solutions presented above are the ones that will be utilized in later chapters of this thesis, and additionally, a solution for a rectangular transducer is presented in Appendix 3.A.

[3]The on-axis SIR has duration $t_1 - t_z$ with the constant amplitude c_p in the time interval $t_z \leq t \leq t_1$.



(a) Spatial impulse response at $z = 20$ mm.

(b) Spatial impulse response at $z = 80$ mm.

Figure 3.2: On-axis spatial impulse responses for a 10 mm disc where the sound speed, c_p, was 1500 m/s.

3.2 Sampling Analytical Spatial Impulse Responses

The analytical expressions for the SIRs discussed above must be converted to a discrete form in order to be useful for digital signal processing. A proper discrete representation of the SIRs is necessary so that when the sampled SIR is convolved with the normal velocity the resulting waveform can faithfully represent the sampled measured waveform.

The analytical SIRs have an infinite bandwidth due to the abrupt amplitude changes that, for example, could be seen in the line and disc solutions above. In some situations the duration of a SIR may even be shorter than the sampling interval, T_s, and it is therefore not sufficient to sample the analytical SIRs by simply taking the amplitude at the sampling instants, t_k, since the SIR may actually be zero those time instants. The SIRs are however convolved with a band-limited normal velocity, hence the resulting pressure waveform must also be band-limited, cf. (2.23). Consequently, we only need to sample the SIR in such way that the band-limited recieved A-scans are properly modeled.

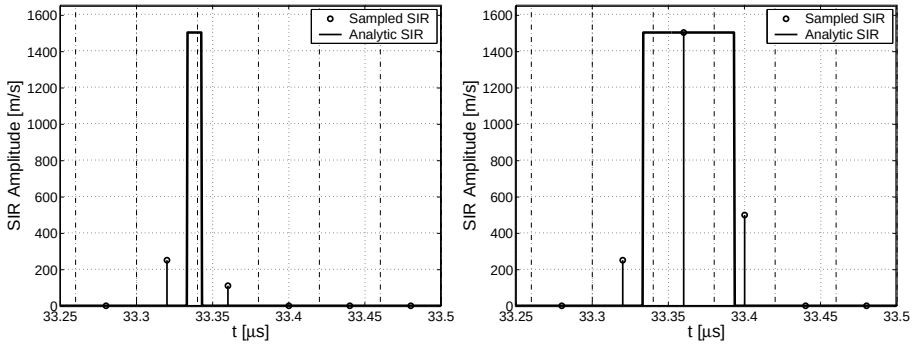
Here, the SIR sampling process is performed by collecting all contributions from the continuous time SIR in the corresponding sampling interval $[t_k - T_s/2, t_k + T_s/2]$. A discrete version of the a time continuous SIR is then

obtained by summing all contributions from the SIR in the actual sampling interval. That is, the sampled SIR is defined as

$$h(\mathbf{r}, t_k) \triangleq \frac{1}{T_\mathrm{s}} \int_{t_k - T_\mathrm{s}/2}^{t_k + T_\mathrm{s}/2} h(\mathbf{r}, t')dt'. \tag{3.4}$$

The division by T_s retains the same unit (m/s) of the sampled SIR as the continuous one. The amplitude of the sampled SIR, at time t_k, is then the mean value of the continuous SIR in the corresponding sampling interval, $[t_k - T_\mathrm{s}/2, t_k + T_\mathrm{s}/2]$. Also, as seen from the analytical solutions above, the SIRs always have a finite length as the transducer has a finite size. The sampled SIRs are therefore naturally represented by finite impulse response filters (FIR).

The effect of the sampling scheme (3.4) is illustrated in Figure 3.3 for two discs with radii 1.2 and 3 mm, respectively, where the sampling interval, T_s, was 0.04μs. In Figure 3.3(a) the analytic SIR is shorter than the sampling



(a) Continuous and sampled spatial impulse responses of a circular disc with radius $r = 1.2$ mm.

(b) Continuous and sampled spatial impulse responses of a circular disc with radius $r = 3$ mm.

Figure 3.3: Illustration of the spatial impulse response sampling method used in this thesis. The continuous and sampled on-axis SIRs for two discs with radii 1.2 and 3 mm, respectively are shown where the sampling interval, T_s, was 0.04μs.

interval, T_s. The max amplitude of the discrete SIR is therefore lower than then the max amplitude of the continuous SIR. If the duration of the analytic SIR is longer than the sampling interval, as for the 3 mm disc shown in

Figure 3.3(b), then the max amplitudes of the on-axis sampled and analytic disc SIRs will be the same.

3.3 The Pressure Field

In Chapter 2 it was shown that the pressure waveform at an observation point $\mathbf{r}$ can be modeled as a convolution of the SIR, $h(\mathbf{r}, t)$, corresponding to $\mathbf{r}$ with the normal velocity waveform at the transducer surface,

$$p(\mathbf{r}, t) = h(\mathbf{r}, t) * v(t). \tag{3.5}$$

Since all SIRs of practical interest have a low-pass character, there will only be significant acoustic energy at those observation points where the frequency-bands of the SIR and the normal velocity overlap. The bandwidth of the SIR becomes narrower when the observation point is further away the from the center-axis of the aperture because the duration of the SIR gets longer. This makes the beampattern frequency dependent and a transducer with low center frequency has therefore a wider beampattern than a transducer with high center frequency for a fixed size of the transducer.

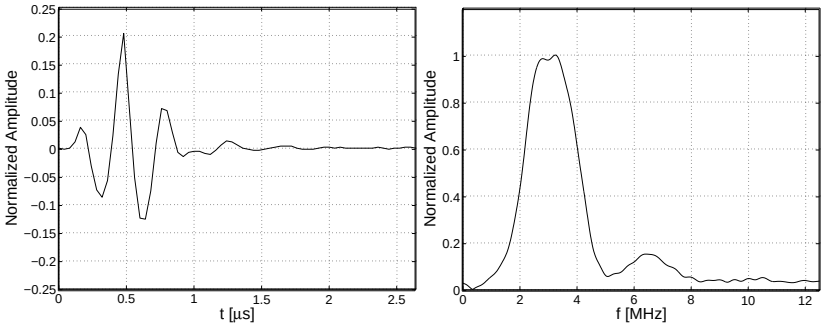
As an example of this effect, simulation results for two circular disc transducers, with radii $r = 1$ mm and $r = 5$ mm respectively, are shown in Figures 3.5 and 3.6. An impulse excitation is used here and the normal velocity is therefore given by the forward electro-mechanical, or electrical, impulse response of the transducer, that is,

$$v(t) = h_{\text{ef}}(t) * \delta(t), \tag{3.6}$$

where $h_{\text{ef}}(t)$ denotes the transducer's forward electrical impulse response. The electrical impulse response used here is shown in Figure 3.4, and the center frequency is about 3 MHz which corresponds to a wavelength of 0.5 mm.[4]

Figure 3.5 shows the on-axis responses for the two transducer sizes and the corresponding normalized spectra for the SIRs, the electrical impulse response, and the SIRs convolved with the electrical impulse response.

[4]The electrical impulse response was measured with a Precision Acoustics PVDF hydrophone. The hydrophone has a bandwidth that is considerably larger than the bandwidth of the transducer and will, therefore, not affect the measurement of the transducer's electrical impulse response.



(a) The electrical impulse response.

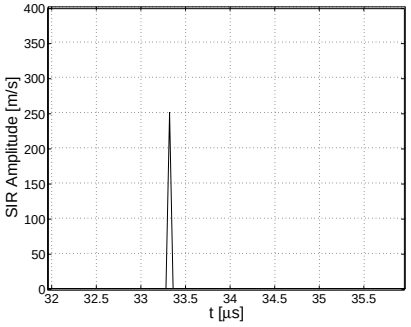
(b) Amplitude spectrum.

Figure 3.4: The (forward) electrical impulse response of the array transducer that has been used in the experiments in this thesis. The electrical impulse response were measured with a broadband PVDF hydrophone manufactured by Precision Acoustics.

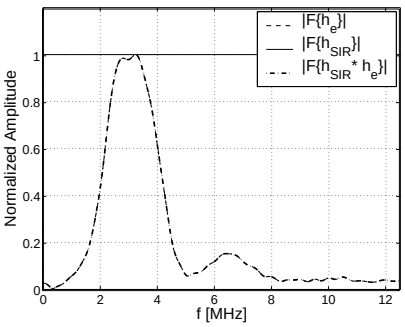
The plots in Figure 3.5, repeated for the off-axis case $x = 10$ mm, are shown in Figure 3.6. We can make the following observations from Figures 3.5 and 3.6: firstly, as noted above, a small transducer has a shorter SIR than a large one and is therefore more broadband, and secondly, the off-axis SIRs are longer than the on-axis ones. The acoustic energy is therefore more concentrated below the center of the transducer. The low pass-filtering effect of the SIRs can clearly be seen in Figure 3.6(d) where most of the transducer's electro-mechanical amplitude spectrum, $|\mathcal{F}\{h_{\mathrm{e}}(t)\}|$, is outside the pass-band of the SIR. A consequence of the SIRs' filtering effect is thus that a transducer with a low center frequency will have a wider beam than a transducer with a high center frequency. This is well known in classical sound field modeling.

3.4 A Numerical Method for Computing Spatial Impulse Responses

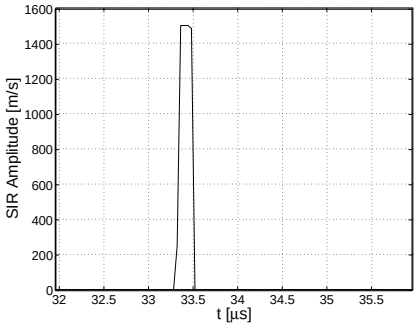
As mentioned in the introduction of this chapter, the analytical spatial impulse responses are only available for a few simple transducer geometries. Therefore, for a transducer with an arbitrary geometry a numerical method



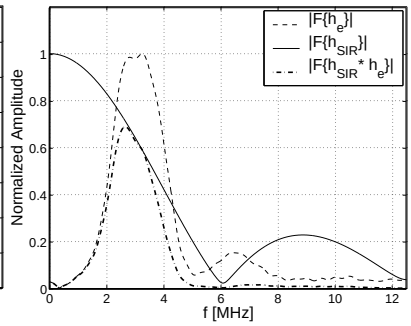
(a) Sampled spatial impulse response for a disc radius of 1 mm.



(b) Amplitude spectrum of the on-axis SIR for a disc of radius 1 mm.

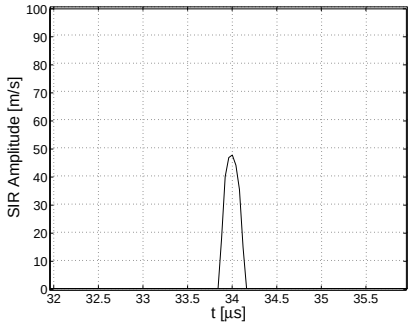


(c) Sampled spatial impulse response for a disc radius of 5 mm.

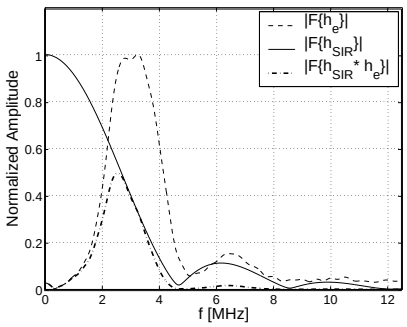


(d) Amplitude spectrum of the on-axis SIR for a disc of radius 5 mm.

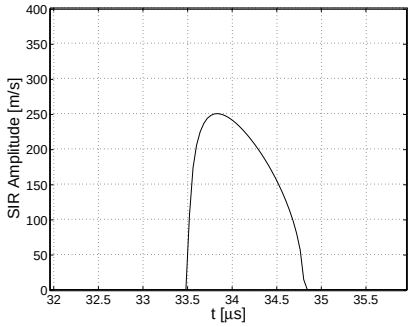
Figure 3.5: On-axis sampled spatial impulse responses for a circular disc with radii 1 and 5 mm at (x=0,z=50mm) and the corresponding amplitude spectra. The solid lines are the normalized spectrum of the SIRs, the dashed lines are the normalized spectra of the electrical impulse response, and the dashed-dotted lines are the spectra of the convolved SIRs and the electrical impulse response.



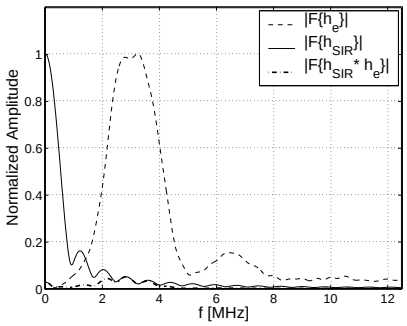
(a) Sampled spatial impulse response for a disc radius of 1 mm.



(b) Amplitude spectrum of the off-axis SIR for a disc of radius 1 mm.



(c) Sampled spatial impulse response for a disc radius of 5 mm.



(d) Amplitude spectrum of the off-axis SIR for a disc of radius 5 mm.

Figure 3.6: Off-axis sampled spatial impulse responses for a circular disc with radii 1 and 5 mm at (x=10 mm,z=50 mm) and the corresponding amplitude spectra. The solid lines are the normalized spectra of the SIRs, the dashed lines are the normalized spectra of the electrical impulse response, and the dashed-dotted lines are the spectra of the convolved SIRs and the electrical impulse response.

must be used. The method is used throughout this thesis, the *discrete representation* (DR) method, is based on a discretization of the Rayleigh integral formula, (2.17). In the DR method, the radiating surface is divided into a set of small surface elements, and the surface integral in the Rayleigh

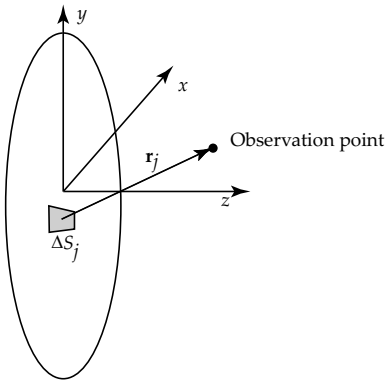


Figure 3.7: Geometry and notations for the discrete representation method.

formula is replaced by a summation, as illustrated in Figure 3.7. The DR method facilitates computation of SIRs for non-uniform excitation, apodization of the aperture, and arbitrary focusing laws since each surface element can be assigned a different normal velocity, apdodization or time-delay. The DR computational concept can therefore be used for computing SIRs for an arbitrary transducer shape or array layout [52, 53].

First, let the total transducer surface be divided into a set of J surface elements $\{\Delta S_0, \Delta S_1, \ldots, \Delta S_{J-1}\}$. Furthermore, let w_j denote an aperture weight, and $R_j = |\mathbf{r} - \mathbf{r}_j|$ the distance from the jth surface element to the observation point. The discrete SIR can now be approximated by

$$\begin{aligned} h(\mathbf{r}, t_k) &= \frac{1}{2\pi} \sum_{j=0}^{J-1} \frac{w_j \delta(t_k - R_j/c_\mathrm{p} - d_j)}{R_j} \Delta S_j \\ &= \sum_{j=0}^{J-1} a_j \delta(t_k - R_j/c_\mathrm{p} - d_j), \end{aligned} \tag{3.7}$$

where d_j is a user defined focusing delay and $t_k = kT_\mathrm{s}$, for $k = 0, 1, \ldots K-1$. The amplitude scaling factor

$$a_j = \frac{w_j \Delta S_j}{2\pi R_j} \tag{3.8}$$

in (3.7) represents the amplitude of the impulse response for an elementary surface at $\mathbf{r}_j$ excited by a Dirac pulse, cf. (3.1). Hence, the total response at time t_k is a sum of contributions from those elementary surface elements, ΔS_j, whose response arrive in the time interval $[t_k - T_s/2, t_k + T_s/2]$.

The accuracy of the method depends on the size of the discretization surfaces ΔS_j. It should however be noted that high frequency numerical noise due to the surface discretization is in practice not critical since the transducer's electrical impulse response has a bandwidth in the low frequency range (for a further discussion see [53]). Also, these errors are small if the elementary surfaces, ΔS_j, are small. The DR-method is very flexible in the sense that beam steering, focusing, apodization, and non-uniform surface velocity can easily be included in the simulation.

The author has developed a software tool, e.g. the DREAM (Discrete REpresentation Array Modeling) Toolbox, based on the DR method in cooperation with Prof. Bogdan Piwakowski at Lille Groupe ≪ Electronique - Acoustique ≫, Institut d'Electronique et de Micro-électronique du Nord, France.[5] The DREAM toolbox has support for many different transducer types and array layouts, including arrays with circular, rectangular, and cylindrical focused elements to name a few. Numerical routines from the DREAM toolbox have been used for all the numerical SIR computations in this thesis when analytical solutions are not available.[6]

[5] The DREAM toolbox is available on-line at `http://www.signal.uu.se/Toolbox/dream/`.

[6] Currently, the DREAM toolbox only has support for computing SIRs in homogeneous media. In this thesis experiments are performed also using an immersed copper block. To compute SIRs for this case an algorithm inspired by the DR method is used which is described in Appendix 3.B.

3.A The Rectangular Source

Rectangular sources are also very common in ultrasonic imaging. Array transducers is a typical example where the array elements often have a rectangular active area.

The analytic SIR for a rectangular transducer can be computed by using the solution for an observation point at a corner point under the rectangular aperture, see illustration in Figure 3.8. The solution for the corner point is

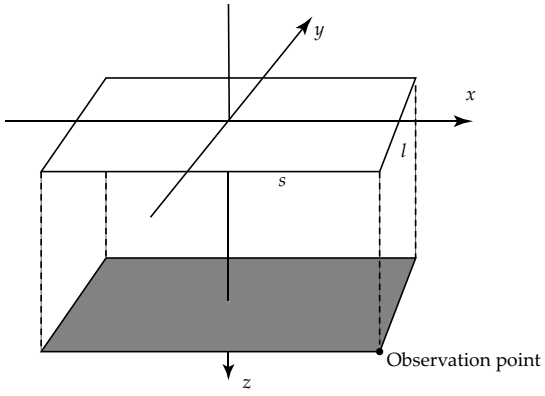


Figure 3.8: Geometry of a rectangular aperture located in the xy-plane. The y-direction width is given by l and x-direction width by s. Also illustrated is the location of an observation point located under a corner of the rectangular aperture.

given by [57]

$$
\begin{aligned}
h(\mathbf{r},t) = \frac{c_\mathrm{p}}{2\pi}\Bigg(&\frac{\pi}{2}\Pi(\tau^\mathrm{I},\tau^\mathrm{IV}) \\
&- \cos^{-1}\left(\frac{s}{c_\mathrm{p}\sqrt{t^2-t_z^2}}\right)\Pi(\tau^\mathrm{II},\tau^\mathrm{IV}) \\
&- \cos^{-1}\left(\frac{l}{c_\mathrm{p}\sqrt{t^2-t_z^2}}\right)\Pi(\tau^\mathrm{III},\tau^\mathrm{IV})\Bigg).
\end{aligned}
\tag{3.9}
$$

where l is the y-direction width and s the x-direction width of the rectangle. The function $\Pi(\cdot,\cdot)$ is the "pulse function" defined as $\Pi(t_1,t_2)=1$ when $t\in[t_1,t_2]$ and $\Pi(t_1,t_2)=0$ otherwise. The time delays, $\tau^\mathrm{I},\ldots,\tau^\mathrm{IV}$, defined

by

$$\begin{aligned}\tau^{\mathrm{I}} &= t_z \\ \tau^{\mathrm{II}} &= \sqrt{z^2+s^2}/c_{\mathrm{p}} \\ \tau^{\mathrm{III}} &= \sqrt{z^2+l^2}/c_{\mathrm{p}} \\ \tau^{\mathrm{IV}} &= \sqrt{z^2+s^2+l^2}/c_{\mathrm{p}},\end{aligned} \tag{3.10}$$

are the travel times of the pulse from the four corner points of the rectangle to the observation point, respectively.

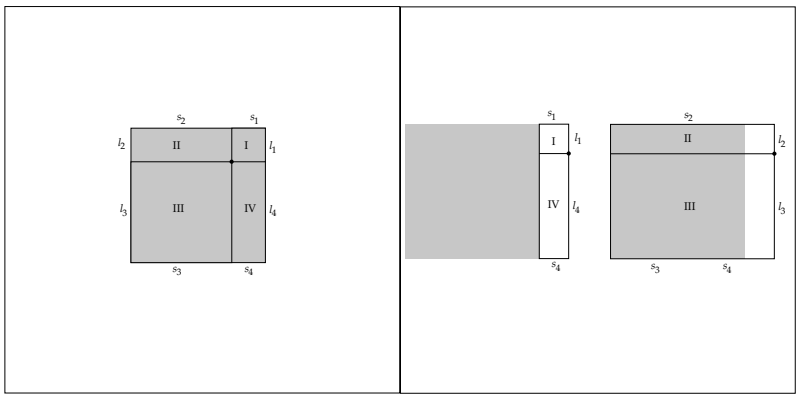
The response at an arbitrary observation point can now be obtained by adding and/or subtracting responses from rectangular apertures. There are four possible cases which are illustrated in Figure 3.9: In Figure 3.9(a) the observation point is strictly inside the aperture, in Figure 3.9(b) x is outside the aperture, in Figure 3.9(c) y is outside the aperture, and in Figure 3.9(d) both x and y is outside the aperture.

For the first case, when the observation point is inside the aperture, the SIR is obtained by summing the contributions from the rectangular sub-apertures I,. . .,IV. When x is outside the aperture the SIR is obtained by first summing contributions from the apertures II and III and then subtracting the two contributions from rectangle I and IV that is not a part of the true aperture. The SIRs for the two remaining cases are obtained in a similar fashion.

The SIR for the a rectangular aperture can finally be expressed as

$$\begin{aligned}h(\mathbf{r},t) = \frac{c_{\mathrm{p}}}{2\pi}\sum_{i=1}^{4} g_i \cdot \Bigg(&\frac{\pi}{2}\Pi(\tau_i^{\mathrm{I}},\tau_i^{\mathrm{IV}}) \\ &- \cos^{-1}\left(\frac{s_i}{c_{\mathrm{p}}\sqrt{t^2-t_z^2}}\right)\Pi(\tau_i^{\mathrm{II}},\tau_i^{\mathrm{IV}}) \\ &- \cos^{-1}\left(\frac{l_i}{c_{\mathrm{p}}\sqrt{t^2-t_z^2}}\right)\Pi(\tau_i^{\mathrm{III}},\tau_i^{\mathrm{IV}})\Bigg)\end{aligned} \tag{3.11}$$

where $g_i = 1$ or -1 depends on the location of the observation point.



(a) Both x and y inside the aperture.

(b) x outside the aperture.

(c) y outside the aperture.

(d) Both x and y outside the aperture.

Figure 3.9: The four possible observation point positions under a rectangular aperture. The gray-shaded area denotes the transducer's active area.

3.B An Algorithm for Computing SIRs in Immersed Solids

One of the experiments presented in Chapter 8 is performed using an immersed copper block. To be able to use the model based methods that is introduced in Chapter 7 we need to compute the SIRs for this setup. The numerical method used in this thesis for this case is based on a line source model with adjustments for the different sound speeds of the two media and the refraction effects at the water-solid interface, as described below.

Wave propagation in elastic solids are, as mentioned in Chapter 1, more complicated than propagation in fluids since many types of waves can propagate. In the bulk of the solid both longitudinal and shear waves can propagate. The longitudinal and shear waves have different sound speeds where the longitudinal sound speed, $c_{\mathrm{p}}^{\mathrm{Cu}}$, in copper is around 4600 m/s and the shear wave sound speed, $c_{\mathrm{s}}^{\mathrm{Cu}}$, is 2250 m/s. Thus, longitudinal waves in copper travels roughly two times faster than shear waves.

As also mentioned in Chapter 1, mode conversion occurs at interfaces were the acoustical impedance changes abruptly, resulting in both longitudinal and shear wave components. In order to successfully model wave propagation in an immersed copper block these mode conversion must be taken into account.

To model the wave propagation, at the water-copper interface, consider a harmonic plane wave impinging on the interface. The incident longitudinal wave will both be reflected by the copper surface and transmitted into the block. The transmitted wave will, due to mode conversion, consist of both a longitudinal and a shear wave, see Figure 3.10. The angle of the transmitted longitudinal wave, ϕ_{P}, and the angle of the shear wave, ϕ_{S}, are given by the *generalized Snell's* law [10].[7]

$$\frac{\sin\phi^{\mathrm{i}}}{c_{\mathrm{p}}} = \frac{\sin\phi_{\mathrm{P}}}{c_{\mathrm{p}}^{\mathrm{Cu}}} = \frac{\sin\phi_{\mathrm{S}}}{c_{\mathrm{s}}^{\mathrm{Cu}}}, \tag{3.12}$$

where ϕ^{i} is the angle of the incident wave. The energy from the incident wave will consequently be distributed between the reflected wave and the two transmitted waves.

Evidently it is more difficult to model wave propagation in immersed elastic solids compared to the acoustic case. To simplify the problem, mea-

[7] Longitudinal waves are also known as primary waves, or P-waves, and shear waves are known as secondary waves, or S-waves.

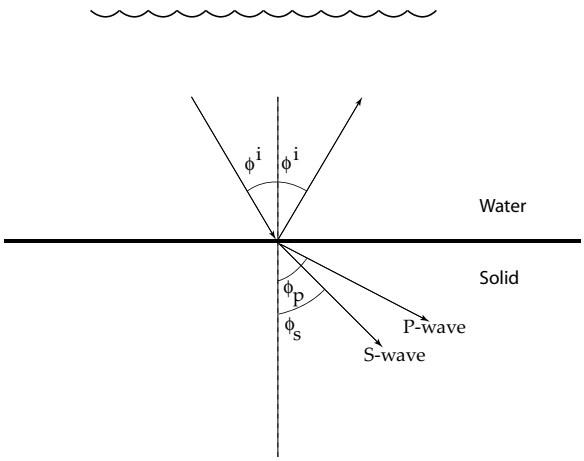


Figure 3.10: Reflection and refraction of a plane pressure wave at a fluid-solid interface.

surements are usually performed in such way that the three waves received by the transducer can be separated in time so that only one wave type needs to be considered. In the experiments presented in this thesis only the transmitted longitudinal wave mode is used and the wave propagation can, therefore, be modeled using the corresponding SIRs. To simplify the modeling further a line source was used to model the transducer

The line source SIRs were obtained by performing a numerical integration over the length of the line source with a method similar DR concept used in the DREAM software tool described in Section 3.4. The amplitude of the SIR at time t_k is proportional to l_k/R_k, that is,

$$h(\mathbf{r}, t_k) \propto \frac{l_k}{R_k}, \tag{3.13}$$

where l_k is the length (e.g. the width of the area) of the transducer that is active in the sampling interval $[t_k - T_s/2, t_k + T_s/2]$ and R_k is the propagation distance from the center of l_k to the observation point, see illustration in Figure 3.11. The length l_k and the propagation distance R_k are computed using the generalized Snell's law (3.12) and a numerical search algorithm.

Note that here, the line source it positioned along the x-axis and, therefore, only the x-direction size effects are considered. In the y-direction the transducer is treated as a point source and the y-direction size effects are thus not modeled. However, since the transducer used in these experiments

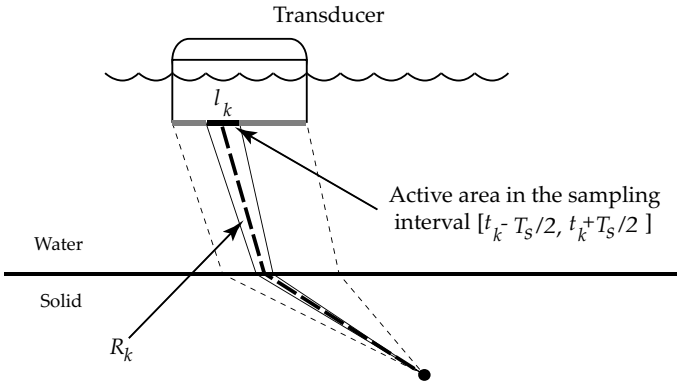


Figure 3.11: Illustration of the numerical method for computing spatial impulse responses for measurements performed using immersed solids. The amplitude of the SIR at time t_k is proportional to l_k/R_k where l_k is the length (width) of the area of the transducer that is active in the sampling interval $[t_k - T_s/2, t_k + T_s/2]$ and R_k is the propagation distance from the center of l_k to the observation point.

is concave,[8] and the targets were located at the transducer's geometrical focal point, the y-direction characteristics will be similar to a point source. That is, at the focal point waves from all points on the transducer surface arrive in phase and the shape, but not the amplitude, of the waveform will, therefore, resemble the pulse shape from a point source.

[8] A picture of the transducer is shown in Figure 8.1.

CHAPTER 4

A Discrete Linear Model for Ultrasonic Imaging

THE objective of this chapter is to derive a discrete-time linear model for pulse-echo ultrasonic imaging based on the SIR method discussed in Chapter 3. To obtain a model for pulse-echo imaging we need to consider, the transducer radiation, the scattering and the reception processes. The reception process is similar to the transmission process and can therefore also be modeled using the SIR method. The only task that remains before the model is complete is thus to model the scattering process, and here, the scattering has been modeled a superposition of responses from point-scatterers.

The present model takes both the diffraction and focusing effects into account. As was discussed in Chapter 3 the diffraction due to finite-sized transducers results in a low-pass filtering effect, which often is neglected in traditional array imaging. Modeling of this effect is therefore important, in particular for the performance of the reconstruction methods discussed in Chapter 7 and 8.

This chapter is organized as follows: In Section 4.1, the model assumptions are discussed. In Section 4.2, a continuous-time convolution model for point-like target is derived that is based on the SIR method discussed in Chapter 3, and finally, the sampling of the model is described in Section 4.3.

4.1 Model Assumptions

The SIR method is based on several assumptions in the modeling of transducer radiation and reception. Since the imaging model also includes the scattering process, more assumptions are needed to obtain a realizable model for our purposes.

As we mentioned in Chapter 1, scattering can become arbitrarily complex since many wave modes can be generated at, for example, a crack in a specimen. This makes it inherently difficult to model scattering and, to obtain a simple model, one therefore has to resort to approximations.

To clarify the approximations, and also to make it easier to understand the possible discrepancies between the model and physical imaging systems, the most important model assumptions are listed below:

Assumption 4.1 The medium is assumed to be homogeneous and isotropic except where a scatterer is located.

Assumption 4.2 The wave propagation is assumed to be linear.

Assumption 4.3 The transducer is to be baffled with an infinite rigid plane. The transducer radiation can, therefore, be modeled using the SIR concept, discussed in Chapters 2 and 3.

Assumption 4.4 The transducer can be treated as a uniform rigid piston.

Assumption 4.5 The scattered field can be modeled as a superposition of contributions from point scatterers.

Assumption 4.6 Mode conversion effects can be neglected.

Assumption 4.7 Multiple scattering is assumed to be very weak and can thus be neglected.[1]

[1] If the superposition principle holds then the combined response from two objects should be the sum of the individual responses. This is however a simplification since the scattered field from the first object may be re-scattered by other objects. This will, however, result in a non-linear forward model.

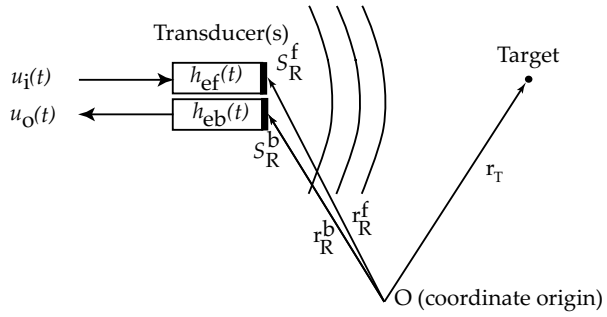


Figure 4.1: Schematic view of the ultrasonic pulse-echo measurement system. The system has a forward electro-mechanical impulse response $h_{ef}(t)$, a backward electro-mechanical impulse response $h_{eb}(t)$, and the corrsponding forward and backward active areas are denoted S_R^f and S_R^b, respectively.

Assumption 4.8 Shadowing effects can be neglected. The acoustic field behind a strong scatterer will be weakened since a substantial portion of the acoustic energy is back-scattered, and this effect is not considered here.

The above assumptions reveal that the response from a scatterer at a certain observation point is independent of those from other scatterers. This enables the array imaging system to be treated as a linear and time-invariant (LTI) system where at each observation point the impulse response is determined by the transducer shape, apodization, and focusing delay.

4.2 A Linear Imaging Model

The linear model of an imaging system is derived based on a point target at position $\mathbf{r}$, with an elementary surface s_e, as depicted in Figure 4.1. We consider a typical piezoelectric transducer that has a *forward* electro-mechanical impulse response denoted $h_{ef}(t)$ and a *backward* electro-mechanical impulse response $h_{eb}(t)$, respectively. We also denote the electrical input signal to the transducer $u_i(t)$, and its output $u_o(t)$.

Since the transducer is treated as uniform rigid piston the normal velocity, $v_n(t)$, on its active area S_R is space independent, and can therefore be

expressed as

$$v_{\mathrm{n}}(t) = h_{\mathrm{ef}}(t) * u_{\mathrm{i}}(t). \tag{4.1}$$

In this case the pressure field at $\mathbf{r}$ is given by the Rayleigh integral (2.17),

$$p(\mathbf{r},t) = \rho_0 v_{\mathrm{n}}(t) * \frac{\partial}{\partial t}\left(\int_{S_{\mathrm{R}}^{\mathrm{f}}} \frac{\delta\left(t - (|\mathbf{r}-\mathbf{r}_{\mathrm{R}}^{\mathrm{f}}|/c_{\mathrm{p}})\right)}{2\pi|\mathbf{r}-\mathbf{r}_{\mathrm{R}}^{\mathrm{f}}|} dS_{\mathrm{R}}^{\mathrm{f}}\right), \tag{4.2}$$

where $S_{\mathrm{R}}^{\mathrm{f}}$ denotes the transmit, or forward, aperture and $\mathbf{r}_{\mathrm{R}}^{\mathrm{f}} \in S_{\mathrm{R}}^{\mathrm{f}}$.[2] Letting $h^{\mathrm{f}}(\mathbf{r},t)$ denote the SIR associated with the transmit aperture, called the *forward* SIR, defined as,

$$h^{\mathrm{f}}(\mathbf{r},t) \triangleq \int_{S_{\mathrm{R}}^{\mathrm{f}}} \frac{\delta\left(t - (|\mathbf{r}-\mathbf{r}_{\mathrm{R}}^{\mathrm{f}}|/c_{\mathrm{p}})\right)}{2\pi|\mathbf{r}-\mathbf{r}_{\mathrm{R}}^{\mathrm{f}}|} dS_{\mathrm{R}}^{\mathrm{f}}, \tag{4.3}$$

then the pressure at $\mathbf{r}$ can be expressed

$$p(\mathbf{r},t) = \rho_0 v_{\mathrm{n}}(t) * \frac{\partial}{\partial t} h^{\mathrm{f}}(\mathbf{r},t). \tag{4.4}$$

Thus, the forward SIR essentially specifies a *spatial variant* LTI filter that relates the normal velocity to the pressure at an observation point. Hence, the SIR concept enables the transducer radiation to be treated as linear filtering.

To model the reception process the scattering from a point target, received by a transducer with a surface $S_{\mathrm{R}}^{\mathrm{b}}$, is considered. In this case we can see a scatterer as source.[3] Since the SIR for a point source takes the form of a delta function, cf. (3.1), the scattered field is given by the convolutional form [54],

$$p(\mathbf{r},t) = \frac{s_{\mathrm{e}}}{c_{\mathrm{p}}} \frac{\partial}{\partial t}\left(p_{\mathrm{inc}}(\mathbf{r}_{\mathrm{T}},t) * \frac{\delta(t - |\mathbf{r}-\mathbf{r}_{\mathrm{T}}|/c_{\mathrm{p}})}{4\pi|\mathbf{r}-\mathbf{r}_{\mathrm{T}}|}\right), \tag{4.5}$$

where $p_{\mathrm{inc}}(\mathbf{r}_{\mathrm{T}},t)$ is the incident pressure given by (4.4) and $\mathbf{r}_{\mathrm{T}}$ is the position of the point-like target.

In reception the piezoelectric transducer is sensitive to the total instantaneous pressure, denoted $\langle p \rangle(\mathbf{r},t)$, on its surface. That is, the total pressure

[2] We introduce a separate notation for the transmit and receive processes to allow the possibility for a different transmit and receive aperture, even though the same aperture is used in transmit and receive in many cases.

[3] It is assumed that there is a sufficient time separation between the transmitted and scattered fields.

is the incident pressure, at time t, integrated over the transducer's active *backward* area, $S_{\mathrm{R}}^{\mathrm{b}}$ [54]. By using the point scattering model, (4.5), the total pressure received by the transducer can then be expressed as

$$\langle p\rangle(\mathbf{r}_{\mathrm{T}},t)=\frac{s_{\mathrm{e}}}{2c_{\mathrm{p}}}p_{\mathrm{inc}}(\mathbf{r}_{\mathrm{T}},t)*\frac{\partial}{\partial t}\left(\int_{S_{\mathrm{R}}^{\mathrm{b}}}\frac{\delta\left(t-(|\mathbf{r}_{\mathrm{T}}-\mathbf{r}_{\mathrm{R}}^{\mathrm{b}}|/c_{\mathrm{p}})\right)}{2\pi|\mathbf{r}_{\mathrm{T}}-\mathbf{r}_{\mathrm{R}}^{\mathrm{b}}|}dS_{\mathrm{R}}^{\mathrm{b}}\right). \tag{4.6}$$

If we also define the *backward* SIR,

$$h^{\mathrm{b}}(\mathbf{r},t)\triangleq\int_{S_{\mathrm{R}}^{\mathrm{b}}}\frac{\delta\left(t-(|\mathbf{r}-\mathbf{r}_{\mathrm{R}}^{\mathrm{b}}|/c_{\mathrm{p}})\right)}{2\pi|\mathbf{r}-\mathbf{r}_{\mathrm{R}}^{\mathrm{b}}|}dS_{\mathrm{R}}^{\mathrm{b}}, \tag{4.7}$$

and substitute both the forward and backward SIRs into (4.6), then (4.6) is reduced to

$$\begin{aligned}\langle p\rangle(\mathbf{r}_{\mathrm{T}},t)&=\frac{s_{\mathrm{e}}\rho_0}{2c_{\mathrm{p}}}v_{\mathrm{n}}(t)*\frac{\partial}{\partial t}h^{\mathrm{f}}(\mathbf{r},t)*\frac{\partial}{\partial t}h^{\mathrm{b}}(\mathbf{r},t)\\&=\frac{s_{\mathrm{e}}\rho_0}{2c_{\mathrm{p}}}\frac{\partial^2}{\partial t^2}h^{\mathrm{f}}(\mathbf{r},t)*h^{\mathrm{b}}(\mathbf{r},t)*v_{\mathrm{n}}(t).\end{aligned} \tag{4.8}$$

The output $u_{\mathrm{o}}(t)$ is related to total pressure by a convolution between the backward electro-mechanical impulse response $h_{\mathrm{eb}}(t)$ and the total pressure, $\langle p\rangle(\mathbf{r},t)$, that is

$$u_{\mathrm{o}}(t)=h_{\mathrm{eb}}(t)*\langle p\rangle(\mathbf{r},t). \tag{4.9}$$

If we now define the double-path (DP) SIR, $h^{\mathrm{DP}}(\mathbf{r},t)$, as

$$h^{\mathrm{DP}}(\mathbf{r},t)\triangleq h^{\mathrm{f}}(\mathbf{r},t)*h^{\mathrm{b}}(\mathbf{r},t), \tag{4.10}$$

then, (4.9) reduces to

$$u_{\mathrm{o}}(t)=s_{\mathrm{e}}\frac{\rho_0}{2c_{\mathrm{p}}}h_{\mathrm{eb}}(t)*\frac{\partial^2}{\partial t^2}h^{\mathrm{DP}}(\mathbf{r},t)*v_{\mathrm{n}}(t). \tag{4.11}$$

Furthermore, if we define the double-path electrical impulse response, $h_{\mathrm{e}}(t)$, as

$$h_{\mathrm{e}}(t)\triangleq\frac{\rho_0}{2c_{\mathrm{p}}}\frac{\partial^2}{\partial t^2}h_{\mathrm{eb}}(t)*h_{\mathrm{ef}}(t), \tag{4.12}$$

and insert (4.1) and (4.12) into (4.11), then the output, $u_{\mathrm{o}}(t)$, can finally be expressed as

$$\begin{aligned}u_{\mathrm{o}}(t)&=s_{\mathrm{e}}h^{\mathrm{DP}}(\mathbf{r},t)*h_{\mathrm{e}}(t)*u_{\mathrm{i}}(t)\\&=s_{\mathrm{e}}h^{\mathrm{sys}}(\mathbf{r},t)*\delta(t),\end{aligned} \tag{4.13}$$

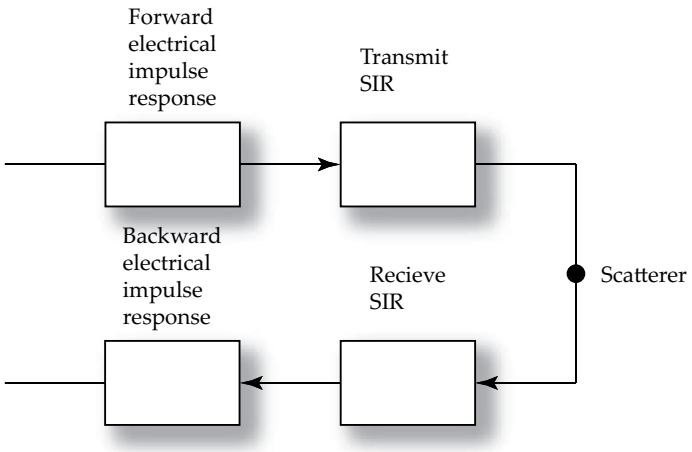


Figure 4.2: Block scheme view of the ultrasonic pulse-echo measurement system shown in Figure 4.1.

where $h^{\text{sys}}(\mathbf{r}, t) \triangleq h^{\text{DP}}(\mathbf{r}, t) * h_{\text{e}}(t) * u_{\text{i}}(t)$ is the *system* impulse response.

Eq. (4.13) states that the output, for a point scatterer at $\mathbf{r}$, is a convolution between the double-path spatial impulse response associated with the transmit and receive aperture, the electrical impulse response, and the input signal, scaled with the scattering amplitude given by the cross-section area, s_{e}, of the target. The total imaging system can therefore be seen as a LTI system where the SIRs and the electro-mechanical impulses responses acts as linear filters, as illustrated in Figure 4.2.

In conclusion, if the shape of the transducer surface, the electrical impulse response, and the positions of the targets are known, the (noise free) output $u_{\text{o}}(t)$ can be modeled using (4.13).

4.3 A Discrete Two-dimensional Model

To obtain a realistic model for an array imaging system we need to consider many transducer positions and the possibility for more than one scatterer. As stated in Assumption 4.5 above, we will model the scattering response as a superposition of contributions from point-like targets. To simplify the analysis we consider the two-dimensional case where we assume that scatterers are found only in the xz-plane, which is our ROI. That is, the responses from targets outside the xz-plane are neglected in the model.

If we consider a single transducer, located on the x-axis ($y = z = 0$), then the signal received by the transducer at position x can, by using superposition and (4.13), be expressed as

$$\begin{aligned} u_{\mathrm{o}}(x,t) &= \sum_{\mathbf{r}\in\mathcal{T}} s_{\mathrm{e}}(\mathbf{r}) h_x^{\mathrm{DP}}(\mathbf{r},t) * h_{\mathrm{e}}(t) * u_{\mathrm{i}}(t) \\ &= \sum_{\mathbf{r}\in\mathcal{T}} s_{\mathrm{e}}(\mathbf{r}) h_x^{\mathrm{sys}}(\mathbf{r},t), \end{aligned} \tag{4.14}$$

where $\mathcal{T}$ is the set of positions of the targets insonified by the ultrasound from the transducer, $h_x^{\mathrm{DP}}(\mathbf{r},t)$ denotes the double-path SIR associated with the transducer position, x, and $h_x^{\mathrm{sys}}(\mathbf{r},t)$ is the corresponding system impulse response. We have also introduced the notation $s_{\mathrm{e}}(\mathbf{r})$ in (4.14) to allow point scatterers with different cross-sectional areas.

A discrete-time version of (4.14) can be obtained by sampling the SIRs and the electrical impulse response, and replacing the continuous time convolutions in (4.14) by discrete-time convolutions. Furthermore, a discrete representation of the scattering strengths is obtained by defining the so-called *object function*, $o(\mathbf{r})$, as

$$o(\mathbf{r}) \triangleq \begin{cases} s_{\mathrm{e}}(\mathbf{r}) & \text{for } \mathbf{r} \in \mathcal{T} \\ 0 & \text{otherwise.} \end{cases} \tag{4.15}$$

The discretized $o(\mathbf{r})$ only takes values at the sampling points, hence, $o(\mathbf{r})$ can be represented by a matrix $\mathbf{O}$, see illustration in Figure 4.3. The element $(\mathbf{O})_{m,\tilde{n}}$ then represents the scattering strength at the observation point $(x_{\tilde{n}}, z_m)$ which is the sum of the contributions from all scatterers inside the sampling cell,

$$(\mathbf{O})_{m,\tilde{n}} = \sum_{\mathbf{r}\in\mathcal{T}_{m,\tilde{n}}} o(\mathbf{r}), \tag{4.16}$$

where $\mathcal{T}_{m,\tilde{n}}$ denotes the set of positions of the targets inside the sampling cell at $(x_{\tilde{n}}, z_m)$.

If we consider the response from one point scatterer at $(x_{\tilde{n}}, z_m)$ the sampled output signal, $\mathbf{y}_n = [u_{\mathrm{o}}(x_n,t_0)\ u_{\mathrm{o}}(x_n,t_1)\cdots u_{\mathrm{o}}(x_n,t_{K-1})]^T$, is then given by the multiplication of $(\mathbf{O})_{m,\tilde{n}}$ with the discrete system impulse response vector, $\mathbf{h}_m^{d(\tilde{n},n)}$, obtained by sampling the system impulse response $h^{\mathrm{sys}}(\mathbf{r},t)$, cf. (4.13). That is, the discrete A-scan vector, $\mathbf{y}_n$, can be expressed

$$\mathbf{y}_n = \mathbf{h}_m^{d(\tilde{n},n)} (\mathbf{O})_{m,\tilde{n}}, \tag{4.17}$$

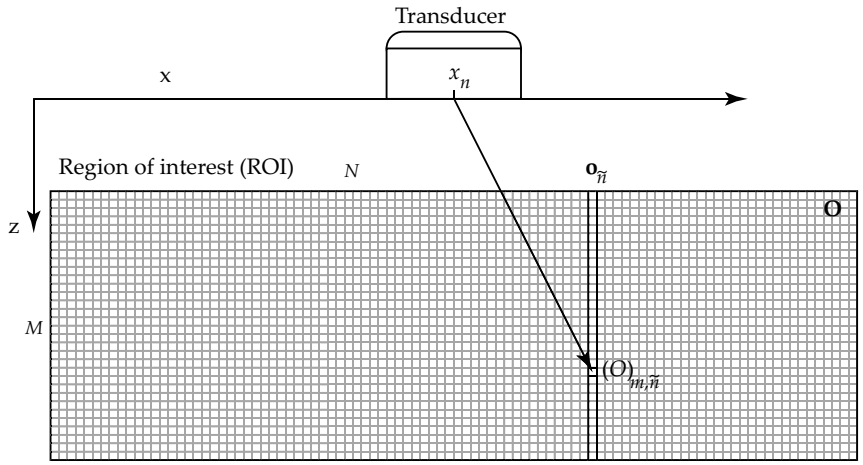


Figure 4.3: Illustration of the discretization of the region of interest.

where

$$d(\tilde{n}, n) \triangleq x_{\tilde{n}} - x_n, \tag{4.18}$$

indicates that the system impulse response, $\mathbf{h}_m^{d(\tilde{n},n)}$, depends on the horizontal distance between the transducer x_n and the observation point $x_{\tilde{n}}$.

The output signal, when all observation points along the line defined by $x = x_{\tilde{n}}$ is considered,[4] can now be found by superposition,

$$\mathbf{y}_n = \sum_{m=0}^{M-1} \mathbf{h}_m^{d(\tilde{n},n)} (\mathbf{O})_{m,\tilde{n}}, \tag{4.19}$$

which can be expressed in a compact matrix-vector form

$$\begin{aligned} \mathbf{y}_n &= \begin{bmatrix} \mathbf{h}_0^{d(\tilde{n},n)} & \mathbf{h}_1^{d(\tilde{n},n)} & \cdots & \mathbf{h}_{M-1}^{d(\tilde{n},n)} \end{bmatrix} \begin{bmatrix} (\mathbf{O})_{0,\tilde{n}} \\ (\mathbf{O})_{1,\tilde{n}} \\ \vdots \\ (\mathbf{O})_{M-1,\tilde{n}} \end{bmatrix} \\ &\triangleq \mathbf{P}_{d(\tilde{n},n)} \mathbf{o}_{\tilde{n}}. \end{aligned} \tag{4.20}$$

Since the transducer will receive echos from targets at different x-positions

[4] cf. Figure 4.3.

a summation must be performed over $\tilde{n}$,[5]

$$\mathbf{y}_n = \sum_{\tilde{n}=0}^{N-1} \mathbf{P}_{d(\tilde{n},n)} \mathbf{o}_{\tilde{n}}. \tag{4.21}$$

In reality the received signal will be contaminated by measurement noise and quantization errors, and the SIR method is as discussed above based several approximations. Hence, the model will will never describe the true imaging system perfectly. Here we model this uncertainty as a noise. If the noise is assumed to be additive then an A-scan measurement vector can be expressed as

$$\mathbf{y}_n = \sum_{\tilde{n}=0}^{N-1} \mathbf{P}_{d(\tilde{n},n)} \mathbf{o}_{\tilde{n}} + \mathbf{e}_n \tag{4.22}$$

where $\mathbf{e}_n$ is the noise vector corresponding to the nth transducer position. Eq. (4.22) is then a discrete linear model for a single A-scan measurement for the imaging system.

The model (4.22) can be extended to a 2D B-scan by first defining L as the number of A-scan measurements in the B-scan, resulting in a $K \times L$ B-scan matrix $\mathbf{Y}$. Then by lexicographically vectorizing $\mathbf{O}$ and $\mathbf{Y}$, $\mathbf{o} = \text{vec}(\mathbf{O})$ and $\mathbf{y} = \text{vec}(\mathbf{Y})$, the B-scan model is obtained by appending the $\mathbf{y}_n$s, for $n = 0, 1, \ldots, L-1$, according to

$$\begin{aligned} \mathbf{y} &= \begin{bmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_{L-1} \end{bmatrix} = \begin{bmatrix} \mathbf{P}_{d(0,0)} & \mathbf{P}_{d(1,0)} & \cdots & \mathbf{P}_{d(N-1,0)} \\ \mathbf{P}_{d(1,0)} & \mathbf{P}_{d(1,1)} & \cdots & \mathbf{P}_{d(N-1,1)} \\ \vdots & \vdots & & \vdots \\ \mathbf{P}_{d(0,L-1)} & \mathbf{P}_{d(1,L-1)} & \cdots & \mathbf{P}_{d(N-1,L-1)} \end{bmatrix} \begin{bmatrix} \mathbf{o}_0 \\ \mathbf{o}_1 \\ \vdots \\ \mathbf{o}_{N-1} \end{bmatrix} + \begin{bmatrix} \mathbf{e}_0 \\ \mathbf{e}_1 \\ \vdots \\ \mathbf{e}_{L-1} \end{bmatrix} \\ &= \mathbf{P}\mathbf{o} + \mathbf{e} \end{aligned} \tag{4.23}$$

where $\mathbf{P}$ has dimension $KL \times MN$.

Eq. (4.23) is a discrete linear model for 2D ultrasonic imaging when the backscattering is considered as a sum of responses from point targets. The ith column, $\mathbf{p}_i$, in $\mathbf{P}$ is the response from a unit area point target at the image point corresponding to the ith element, o_i, in $\mathbf{o}$.

The model takes the diffraction effects from a finite sized aperture into account provided that the surface is planar. The model (4.23) can now be

[5] cf. Eq. (4.14).

used to obtain simulated B-scan data by simply inserting an object function into (4.23) and rearranging $\mathbf{y}$ into an image. The model (4.23) is the foundation for the reconstruction methods discussed in Part III of this thesis.

Part II

Methods based on Delay-and-sum Focusing

CHAPTER 5

The Synthetic Aperture Focusing Technique

SYNTHETIC aperture imaging (SAI) was developed to improve resolution in the along track direction for side-looking radar. The idea was to record data from a sequence of pulses from a single moving real aperture and then, with suitable computation, combine the signals so the output can be treated as a much larger aperture. The first synthetic aperture radar (SAR) systems appeared in the beginning of the 1950's [58, 59]. Later on the method has carried over to ultrasound imaging in areas such as synthetic aperture sonar (SAS) [36], medical imaging, and nondestructive testing [5, 6].

In this chapter we will study traditional delay-and-sum (DAS) based ultrasonic SAI for the purpose of familiarizing the reader with the characteristics of the technique. In particular, we will focus on broadband SAI and the reconstruction performance with respect to spatial sampling and the size of the physical transducer. In Chapter 8, we will then compare the DAS based SAI method to the reconstruction methods that is presented in Chapter 7.

SAI refers to a process in which the focal properties of a large-aperture focused transducer is synthesized by processing sequentially recorded data, often using a small-aperture transducer that has been scanned over a large area. Most of the algorithms used in ultrasonic SAI have been developed for SAR applications which are characterized by narrowband transmissions and very high propagation speeds, i.e., the speed of light. Due to the high propagation speed in SAR, it is rarely a problem to acquire data at a suffi-

cient rate to avoid spatial aliasing. However, in ultrasound where the sound speed ranges from 340 m/s (air) to about 6000 m/s (metals), the low propagation speeds results in a long round-trip time of the echos that consequently results in a low scanning speed. This is in conflict with the need to scan as fast as possible to maximize the frame rate. This leads to a aperture undersampling problem that limits widespread use of ultrasonic SAI [36].[1]

Other differences between SAR and ultrasound SAI are that, in ultrasound, the pulses are often broadband, and the finite size of the scanning transducer can cause a significant beam directivity. The latter issue is in particular pronounced in NDT and medical applications where measurements often are performed in the near-field where the transducer size effects may be significant.

Different measures have been taken to improve the performance of the SAR based methods for conditions found in ultrasonics. Typically different correlation or pulse compression methods are used to improve temporal resolution. In medical applications, requirements concerning frame rate have led to solutions where a combination of a conventional electronic array focusing and SAI imaging is used [61]. Recently, there has also been a considerable interest in multi-element SAI applied to medical array systems [8, 62–64]. These systems are defocused to emulate a single-element spatial response with high acoustic power. This technique involves treating the focal point of the transducer as a virtual source [65]. The virtual source, which is aimed to produce approximately spherical waves over a certain aperture angle, can be treated as a point source [62, 63].

In this chapter we will concentrate on the overly dominating method for combining the signals in ultrasonic synthetic aperture imaging—the synthetic aperture focusing technique (SAFT). We will confine the discussion to the monostatic case where a single unfocused transducer is linearly scanned along the horizontal axis. The SAFT method has both time-domain [5, 66] and frequency-domain [67, 68] implementations but here we shall focus our interest on the time-domain algorithm.

This chapter is organized as follows: In Section 5.1 the classical time-domain SAFT algorithm is introduced. The performance with respect to transducer size and spatial sampling is then illustrated in Section 5.2. In Section 5.3, a matrix form of the SAFT algorithm is derived and in, Section 5.4, a matched filter interpretation of SAFT is given. Finally, in Sec-

[1] The problem can be relaxed somewhat due to the directivity of the transducer which will attenuate the spatial aliasing effects [60].

tion 5.5 some remarks are presented.

5.1 The SAFT Algorithm

The conventional time-domain SAFT algorithm performs synthetic focusing by means of coherent summations, of responses from point scatterers, along hyperbolas.[2] These hyperbolas simply express the distances, or time-delays, from transducer positions in the synthetic aperture to the observation points, see illustration in Figure 5.1.

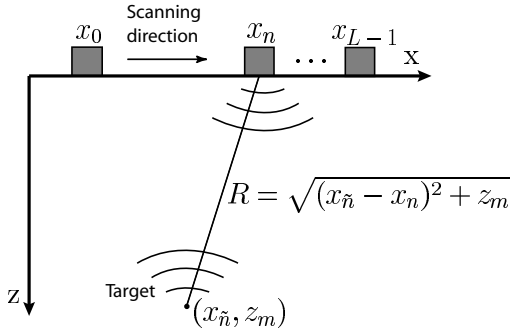


Figure 5.1: Typical measurement setup for a SAFT experiment. The transducer is mechanically scanned along the x-axis and at each sampling position, $x_n = nd$, $n = 0, 1, \ldots, L-1$, a data vector (A-scan) of length K is recorded. The distance between the transducer, at $(x_n, z = 0)$, and the observation point, $(x_{\tilde{n}}, z_m)$, is given by R.

More specifically, to achieve focus at an observation point $(x_{\tilde{n}}, z_m)$, the SAFT algorithm time shifts and performs a summation of the received signals $u_{\mathrm{o}}(x_n, t)$ measured at transducer positions x_n for all n in the synthetic aperture. The time shifts which aim to compensate for differences in pulse traveling time, are simply calculated using the Pythagorean theorem and the operation is commonly expressed in the continuous time form [62, 69–71]

$$o(x_{\tilde{n}}, z_m) = \sum_n w_n u_{\mathrm{o}}(x_n, \frac{2}{c_{\mathrm{p}}}\sqrt{(x_{\tilde{n}} - x_n)^2 + z_m^2}). \tag{5.1}$$

where $o(x_{\tilde{n}}, z_m)$ is the beamformed image.[3]

[2] Linear scanning of the transducer is assumed here.

[3] The beamformed image $o(x_{\tilde{n}}, z_m)$ and the object function introduced in Chapter 4 is

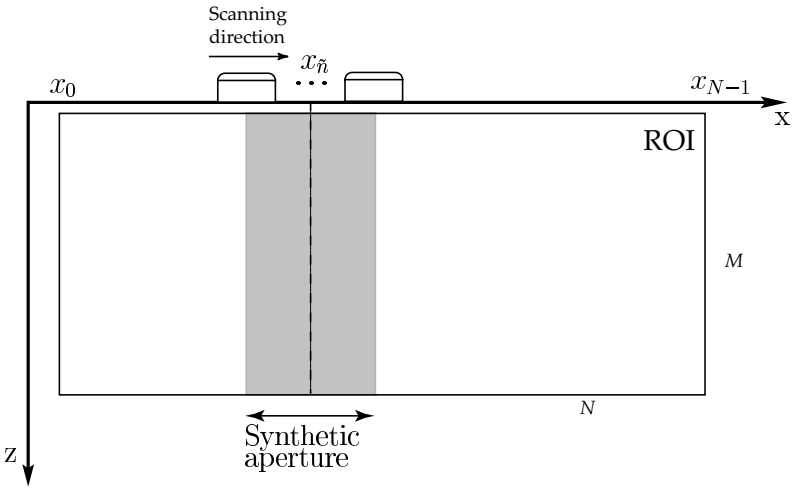


Figure 5.2: Illustration of spatial sampling of the region-of-interest. The ultrasonic B-scan data, $\mathbf{Y}$, are recorded for $x \in [x_0, x_{N-1}]$, shown as the solid-line rectangle. The reconstructed image, $\hat{\mathbf{O}}$, is calculated for every point in the figure. The shaded part of the ROI is the area that is used for reconstruction of the image points along the vertical line defined by $x = x_{\tilde{n}}$.

As mentioned in Chapter 1, apodization weights, w_n, can be introduced to reduce the sidelobe levels but at the cost of a broader main lobe. If the apodization weights, w_n, is equal to 1 then the far-field beampattern, see (1.6), becomes sinc-shaped for harmonic excitation [9, 62].

Normally the received signals are digitized and a discrete version of (5.1) can be obtained as follows: First, let d denote the distance between the consecutive transducer positions along the scanning direction, that is, the synthetic array pitch. Then, an ultrasonic data vector (A-scan) of length K is acquired at each transducer position, $x_n = nd$, $n = 0, 1, \ldots, L-1$. This results in a $K \times L$ B-scan measurement matrix $\mathbf{Y}$. Also, let the synthetic aperture consist of $\tilde{L}$ terms, yielding a width of the synthetic aperture of $(\tilde{L}-1)d$. Secondly, let the discretized region-of-interest (ROI) be represented by an $M \times N$ observation matrix $\mathbf{O}$ which is illustrated in Figure 5.2.[4]

not defined in the same way. However, we use the same notation for the object function and the beamformed image since they have a similar meaning.

[4]The image is discretized in the same manner as for the linear model discussed in Chapter 4.

The discrete SAFT algorithm (5.1) can now be expressed as

$$(\hat{\mathbf{O}})_{m,\tilde{n}} = \sum_{n=\tilde{n}-\tilde{L}/2}^{\tilde{n}+\tilde{L}/2} w_n(\mathbf{Y})_{k,n}. \tag{5.2}$$

The discrete time vector index, k, is a function of the distance between the transducer position, x_n, and the observation point, $(x_{\tilde{n}}, z_m)$, that is given by

$$k \doteq \left(\frac{2}{c_\mathrm{p}} \sqrt{(x_{\tilde{n}} - x_n)^2 + z_m^2} \right) \frac{1}{T_\mathrm{s}}. \tag{5.3}$$

Note that since the signal is discrete, the time-shifts in (5.1) must be rounded towards the nearest sampling instance which introduces some errors. The operator $\doteq$ denotes that k is rounded toward the nearest integer.[5]

5.2 Performance of The Synthetic Aperture Focusing Technique

Here a brief analysis of the SAFT performance is given. Since one of the main topics in this thesis is compensation of beampattern and diffraction effects, we will treat these aspects for the SAFT algorithm below.

In SAFT experiments, the scanning transducer is normally small for the purpose of creating an almost spherical beampattern. However, in many situations a finite sized transducer must be used to obtain sufficient acoustic power. The diffraction effects due to a finite aperture will then reduce performance, especially in the near-field. Undersampling may also cause problems with grating lobes, in particular for narrowband systems. Some important parameters for the SAFT performance are:

- The bandwidth of the transducer.
- The size of the transducer's active area.
- The synthetic array pitch.
- The F-number.
- The signal-to-noise ratio (SNR).

[5]It is common to perform interpolation or to oversample in order to reduce the rounding errors.

Below, the first four of these topics are discussed. The last one is postponed to Chapter 8.

5.2.1 Spatial Undersampled Broadband Data

As mentioned in Chapter 1, the results from conventional far-field beampattern analysis of an undersampled array predict aliased lobes with the same amplitude as the main lobe. However, for broadband imaging, spatial aliasing will in general be less severe, as shall be shown below.

Recall from Chapter 1 that, for a harmonic excitation signal and an array pitch, d, that is larger than half the wavelength, $\lambda/2$, there will be strong grating lobes in the angle interval -90 to 90 degrees outside the main lobe. The grating lobes may severely deteriorate the processed image, since energy from a target within the grating lobe will leak into the main lobe resulting in spurious responses. The far-field beampattern,

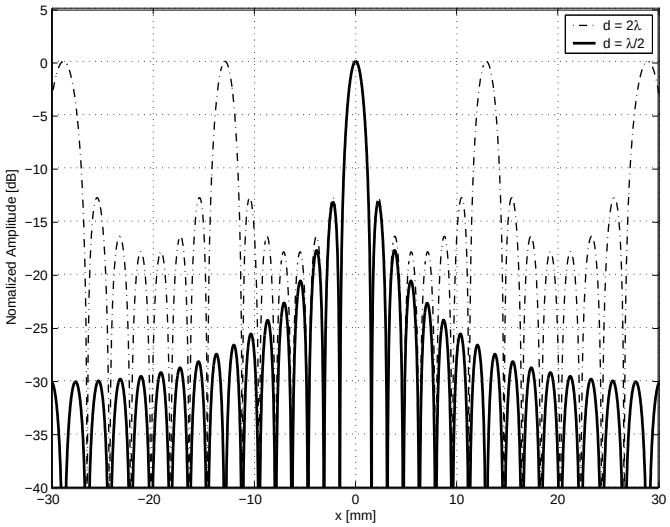
$$|H(x,z)| = \left| \frac{\sin\left(\tilde{L} d \frac{\pi x}{\lambda z}\right)}{\sin\left(d \frac{\pi x}{\lambda z}\right)} \right|, \tag{5.4}$$

has a maximum at $x = 0$ but also for those points where the phase shifts of the received harmonic signals is a multiple of the wavelength at all transducer positions. That is, all received signals arrive in phase for those points. If $d < \lambda/2$ then this will not occur.

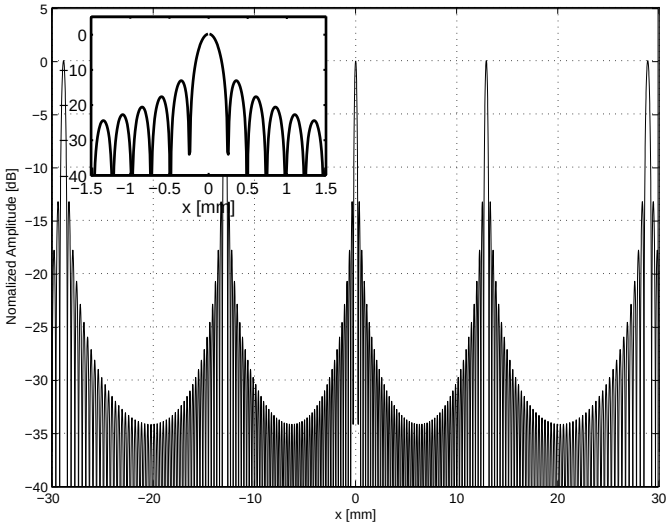
Two example beampatterns computed using (5.4), for a point target at depth $z = 50$ mm, is shown in Figure 5.3. The synthetic array pitch, d, was 1 mm, and $\lambda = 0.5$ mm. Due to the undersampling grating lobes can be found around ± 14 and ± 28 mm for both apertures.

The beampattern in Figure 5.3 is valid also for tone-burst excitation provided that the pulse length is sufficiently long. That is, if the sinusoidal signal is too short, then all transducer positions will not contribute to the total output. If this is the case, then the amplitude in the grating lobe will drop and the effects of undersampling in SAFT is therefore less severe compared to harmonic excitation [9, p. 265].

Normally, in NDE and many other applications, broadband pulses are used and effects of undersampling will then depend of the pulse shape and length [35, 72]. This effect is shown in Figure 5.5 where three different synthetic pulses, shown in Figure 5.4, have been used. The three pulses had a half-power bandwidth (HPBW) of 89%, 28%, and 12%, respectively.



(a) Beampattern for a synthetic aperture of 8 mm. $d = 2\lambda$ (dash-dotted) and $d = \lambda/2$ (solid).



(b) Beampattern for a synthetic aperture 50 mm ($d = 2\lambda$).

Figure 5.3: Beampattern for two synthetic apertures (8 and 50 mm) with an array pitch, $d = 1$ mm, for a target at $z = 50$ mm, and the wavelength, λ, was 0.5 mm. The inserted figure in (b) is an amplification of the beampattern around $x = 0$.

The bandwidth of 89% can be considered as high and 12% as relatively low. The sound field was simulated using the DREAM method for a circular transducer and the simulation parameters are presented in Table 5.1.

Aperture [mm]	50
Depth z [mm]	50
Transducer diameter [mm]	0.4
Sound speed c_{p} [m/s]	1500

Table 5.1: Parameters for the simulation results shown in Figure 5.5.

It is clear that the amplitudes of the grating lobes have been decreased compared to the narrowband analysis for all three pulses. Using a pulse with a HPBW of 89% resulted, for example, in a grating lobe level below -30 dB. The explanation for this behavior is simply that if the pulse is shorter than $Ld\sin(\theta)$, where θ is the grating lobe angle, the pulse will not be present "simultaneously" at all transducer positions in the synthetic aperture and the grating lobe amplitude will therefore drop. However, the pulse length cannot become arbitrarily short since the bandwidth of the array system is limited. Grating lobes will therefore always be a problem in undersampled DAS based array imaging.[6]

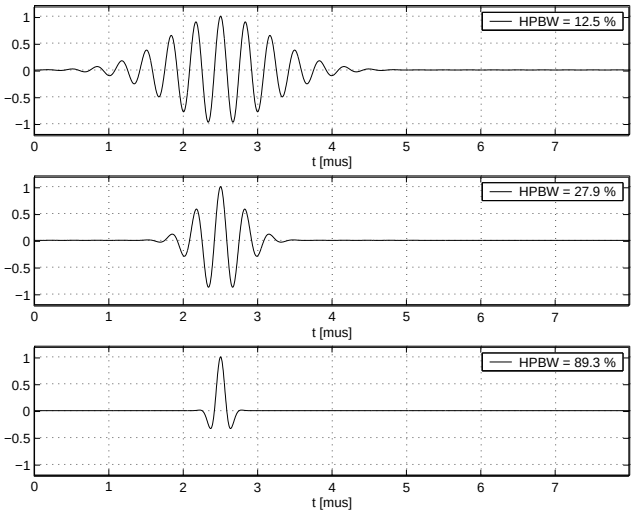
Note also that even though a shorter pulse alleviates the problem with grating lobes, the lateral resolution does not improve significantly by using a short pulse, cf. Figures 5.3(b) and 5.5. The lateral resolution depends mostly on the synthetic aperture and the transducer size as will be discussed in the next subsection.

5.2.2 Finite Sized Transducers and Lateral Resolution

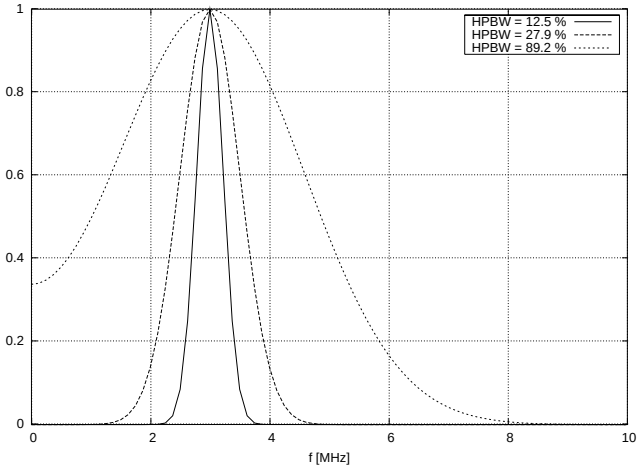
As mentioned earlier in this chapter the transducer is assumed to be a point source and receiver in the SAFT algorithm. The validity of this assumption depends on the size of the transducer in relation to the distance to the focal point as well as the transducer's center frequency and bandwidth.

As an example of typical SAFT performance for finite-sized transducers,

[6]It can also be seen in Figure 5.4 that the grating lobe pattern is smoother than the narrowband pattern, shown in Figure 5.3, since the amplitudes of many frequencies superimpose and since the grating lobes occur at different angles for different frequencies.



(a) Synthetic impulse responses.



(b) Amplitude spectra.

Figure 5.4: Synthetic impulse responses and their corresponding amplitude spectra. The half-power bandwidths of the pulses are 12%, 28%, and 89%, respectively.

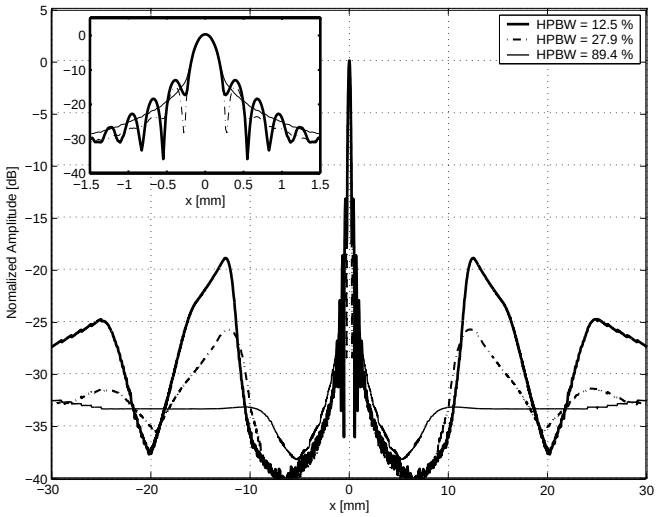


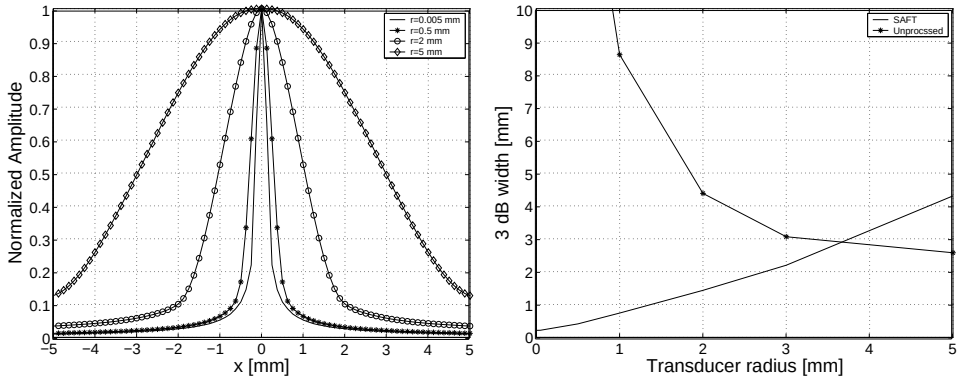
Figure 5.5: Beampattern profiles of broadband SAFT processed data with an aperture of 50 mm for the three pulses shown in Figure 5.4. The thin line corresponds to a half-power bandwidth (HPBW) of 89%, the dash-dotted line to a HPBW of 28%, and the thick lines to HPBW of 12%. The inserted figure is an amplification of the beam profiles around $x = 0$.

a simulation using the DREAM method for a short pulse[7] and synthetic aperture of 50 mm, for a target at $z = 50$ mm, was performed. The beam profile and 3 dB lobe width of the SAFT processed data are shown in Figure 5.6 for a transducer radius ranging from 0.005 to 5 mm. The 0.005 mm transducer can be considered as very small and the 5 mm transducer as large. That is, the 5 mm transducer is large enough to manifest strong beam directivity effects.

As can be seen in Figure 5.6, the resolution clearly deteriorates as the size of the transducer increases. There is nearly a linear relationship between beam width and transducer size, cf. Figure 5.6(b). Note that in this example, if the transducer is larger than approximately 3.5 mm the lateral resolution becomes even worse than the unprocessed data.

To further illustrate the behavior of the SAFT algorithm, B-scan images of unprocessed and SAFT processed simulated data is shown in Figure 5.7. Results from two transducer sizes are shown where the first one is

[7]The pulse with the half power bandwidth (HPBW) of 89% shown in Figure 5.4 was used.



(a) Broadband beam profile for 4 different transducer sizes.

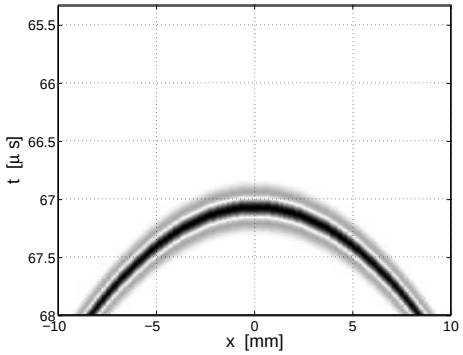
(b) 3 dB beam width as function of transducer size for unprocessed- and SAFT-processed data.

Figure 5.6: Lateral resolution for SAFT processed data for transducer sizes from 0.005 to 5 mm. The simulations were performed with an aperture of 50 mm for a target at $z = 50$ mm.

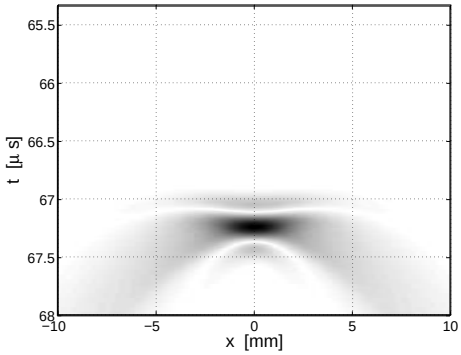
small (0.005 mm) and the second large enough to manifest strong transducer size effects. The unprocessed data for the small transducer, shown in Figure 5.7(a), shows the typical hyperbolic form resulting from linear scanning when the transducer emits nearly spherical waves. It is this pattern that the hyperbola in the SAFT algorithm (5.1) is matched against. The data from the large transducer (5 mm) shown in Figure 5.7(b) deviates, however, significantly from the hyperbolic form. The SAFT performance is therefore poor for the large transducer whilst the result from the small transducer shows a significant improvement in lateral resolution compared to the unprocessed data, cf. Figures 5.7(d) and 5.7(c). Consequently, the diffraction effects of the transducer can severely deteriorate the lateral resolution if care is not taken.

5.3 Matrix Formulation of the SAFT Algorithm

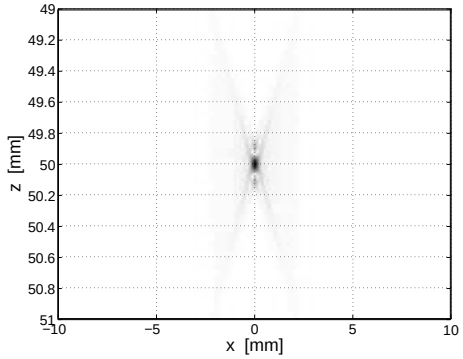
The time-domain SAFT algorithm can also be expressed using a matrix formalism. Here this is done in the interest of clarity. Furthermore, the matrix formulation of SAFT enables a direct comparison with the later



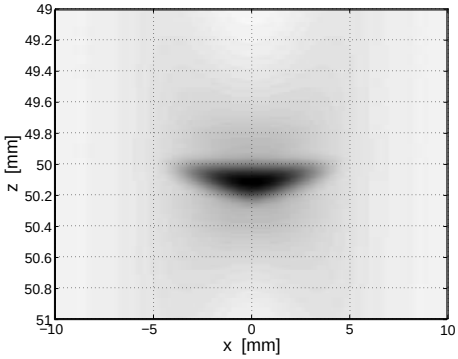
(a) Unprocessed B-scan for the 0.005 mm transducer.



(b) Unprocessed B-scan for the 5 mm transducer.



(c) SAFT processed image for the 0.005 mm transducer.



(d) SAFT processed image for the 5 mm transducer.

Figure 5.7: B-mode images of unprocessed and SAFT processed simulated data for two transducer sizes. The simulations were performed with an aperture of 50 mm for a target at $z = 50$ mm using a wide band pulse excitation with a HPBW of 89%.

proposed algorithms for better illustration of the relations between them.

Let $\hat{\mathbf{O}}$ denote the SAFT reconstructed image. The reconstructed image takes the values at the spatial sampling points $x_{\tilde{n}}, z_m$) for $\tilde{n} = 0, \ldots, L-1$ and $m = 0, \ldots, M-1$, respectively.

Again, let $\mathbf{y}_n = [u_o(x_n, t_0)\ u_o(x_n, t_1) \cdots u_o(x_n, t_{K-1})]^T$ denote the nth A-scan column vector in $\mathbf{Y}$. The DAS operation can now be expressed as a summation of matrix-vector multiplications, that is, the $\tilde{n}$th column, $\hat{\mathbf{o}}_{\tilde{n}}$, in $\hat{\mathbf{O}}$ can be expressed as

$$\hat{\mathbf{o}}_{\tilde{n}} = \sum_{n=\tilde{n}-\tilde{L}/2}^{\tilde{n}+\tilde{L}/2} \mathbf{D}_{d(\tilde{n},n)} \mathbf{y}_n \tag{5.5}$$

where $d(\tilde{n}, n)$ is defined in (4.18) and $\mathbf{D}_{d(\tilde{n},n)}$ is given by

$$(\mathbf{D}_{d(\tilde{n},n)})_{m,j} = \begin{cases} w_{d(\tilde{n},n)} & \text{for } j = k \ \ (\text{with } k \text{ defined in (5.3)}), \\ 0 & \text{otherwise.} \end{cases} \tag{5.6}$$

The apodization weight $w_{d(\tilde{n},n)}$ can be a rectangular window function, i.e., $w_{d(\tilde{n},n)}$ is identical to 1, but often it is a decreasing function of the distance $d(\tilde{n}, n)$.

By following the notation introduced in Chapter 4, where the data vector was defined as $\mathbf{y} = \text{vec}(\mathbf{Y}) \triangleq [\mathbf{y}_0^T \mathbf{y}_1^T \cdots \mathbf{y}_{L-1}^T]^T$, and the vectorized image as $\hat{\mathbf{o}} = \text{vec}(\hat{\mathbf{O}})$, and by also noting that $d(\tilde{n}, n) = d(\tilde{n}+i, n+i)$ for any integer i, the SAFT algorithm can finally be expressed as

$$\begin{aligned} \hat{\mathbf{o}} = & \begin{bmatrix} \mathbf{D}_{d(i-\tilde{L}/2,i)} & & & & \\ \vdots & \mathbf{D}_{d(i-\tilde{L}/2,i)} & & \mathbf{0} & \\ \mathbf{D}_{d(i+\tilde{L}/2,i)} & \vdots & & & \\ & \mathbf{D}_{d(i+\tilde{L}/2,i)} & \ddots & & \\ & & & \ddots & \mathbf{D}_{d(i-\tilde{L}/2,i)} \\ & \mathbf{0} & & & \vdots \\ & & & & \mathbf{D}_{d(i+\tilde{L}/2,i)} \end{bmatrix} \begin{bmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \\ \\ \vdots \\ \\ \mathbf{y}_{L-1} \end{bmatrix} \\ = & \mathbf{D}_{\text{SAFT}}^T \mathbf{y}, \end{aligned} \tag{5.7}$$

for any i such that $i-\tilde{L}/2$ and $i+\tilde{L}/2$ is in the ROI. Note also that k in (5.6) is a function of $(x_{\tilde{n}} - x_n)^2$, implying the symmetry relation $\mathbf{D}_{d(n-\tilde{n},n)} = \mathbf{D}_{d(n+\tilde{n},n)}$. The matrix $\mathbf{D}_{\text{SAFT}}$ is a $KL \times MN$ sparse block matrix.

5.4 Matched Filter Interpretation

The SAFT algorithm can be viewed as an implementation of a matched filter (MF) [6, 9], where the filter associated with image point $(x_{\tilde{n}}, z_m)$ is given by the hyperbolic function (5.3). The matched filter is the linear filter that maximizes the SNR for a *known signal* with *known* arrival time [73, 74]. Consider the one-dimensional discrete time case with a signal $\mathbf{s}$ in Gaussian noise,

$$\mathbf{y} = \mathbf{s} + \mathbf{e}. \tag{5.8}$$

The filter, $\mathbf{h}$, that maximizes the SNR is given by [73]

$$\mathbf{h} \propto \mathbf{C}_{\text{e}}^{-1}\mathbf{s}, \tag{5.9}$$

where $\mathbf{C}_{\text{e}}$ is the noise covariance matrix. If the noise is white ($\mathbf{C}_{\text{e}} = \sigma_{\text{e}}^2\mathbf{I}$) the MF output, $\mathbf{h}^T\mathbf{y} = \mathbf{s}^T\mathbf{y}$, is simply a correlation with the signal prototype $\mathbf{s}$. In the SAFT case, the hyperbola, defined in (5.3), is the 2D MF for the response from a *single* point target at image point $(x_{\tilde{n}}, z_m)$, given that the transducer can be well approximated as a point source and that the its forward and backward electro-mechanical impulse responses, $h_{\text{ef}}(t)$ and $h_{\text{eb}}(t)$, are delta functions, cf. Figure 4.1. Under these circumstances the SNR is expected to increase after SAFT processing

Each column in the SAFT matrix, $\mathbf{D}_{\text{SAFT}}$, defined in (5.7) contains the matched filter for the corresponding image point. That is, by using (5.7) the SAFT matched filter outputs are computed simultaneously for all image points.

Using the linear model derived in Chapter 4 the image obtained after SAFT processing can be modeled as

$$\begin{aligned} \hat{\mathbf{o}} =& \mathbf{D}_{\text{SAFT}}^T\mathbf{y} \\ =& \mathbf{D}_{\text{SAFT}}^T(\mathbf{P}\mathbf{o} + \mathbf{e}), \end{aligned} \tag{5.10}$$

where, $\mathbf{y} = \mathbf{P}\mathbf{o} + \mathbf{e}$, was introduced in (4.23). The columns in the matrix $\mathbf{D}_{\text{SAFT}}^T$ acts as a masks that "picks out" the samples in $\mathbf{y}$ corresponding to the hyperbolas given by (5.3).

5.5 Remarks

Note that for the SAFT algorithm to be a true matched filter, $\mathbf{D}_{\text{SAFT}}^T$ must equal $\mathbf{P}^T$ otherwise will the SAFT filters not match the true response exactly. Recall from Chapter 4 that the ith column, $\mathbf{p}_i$, in the propagation matrix $\mathbf{P}$ corresponds to the response from a unit area point target at the image point represented by the ith element in the vectorized image $\mathbf{o}$. The MF for that image point is then $\mathbf{p}_i^T\mathbf{y}$ This transpose operation can be seen as a *time-reversal* of the propagation operator that maximizes the amplitude of the acoustic wave at a given time and location [75].

In the far-field, and for very broadband systems, the approximation $\mathbf{D}_{\text{SAFT}}^T \approx \mathbf{P}^T$ may be reasonable but in the near-field there are a number of issues that only make the SAFT algorithm an approximation of a MF:

- The electrical impulse response of the transducer filters the signal so that even if the input signal is a delta function the temporal waveform will not resemble a delta function. Since SAFT uses delta functions (time-delays) temporally the SAFT MF will not match the true waveform.
- SAFT treats the transducer as a point source and receiver. If the transducer has a finite size the SIRs associated with the transducer will smoothen waveform which also leads to a mismatch between the SAFT filters and the true response.

It should also be emphasized that MFs is only optimal in the sense that they maximize the SNR at *one* point when *one* signal is present. A MF processing will therefore not necessarily result in high resolution images since energy from scatterers nearby to the observation point may leak into the focal point, through the side- or grating lobes associated with the MF [44].

CHAPTER 6

Real Aperture Array Imaging

REAL aperture array imaging plays an important role in ultrasonic imaging.[1] The reason for the popularity is the very high flexibility of this technique. Real arrays makes it possible to design very fast electronic beam scanning systems as well as fast switching of the focusing properties of the beam.

In this chapter we will review some important properties of real aperture array imaging. We focus on two topics, the transmit process of a phased array,[2] and parallel receive beamforming. The purpose of this chapter is mainly to familiarize the reader with some characteristics of broadband real aperture array imaging that we later will come back to in the following chapters in this thesis.

Note that, we will only consider classical delay-line type of systems that focus the beam by means of imposing time delays on the signals driving the array elements. That is, we will not consider the type of systems that have started to appear in the last decade that are capable of exciting the array elements with arbitrary waveforms [76, 77].

This chapter is organized as follows: In Section 6.1, classical CW far-field array modeling is discussed. In particular, beam steering and its influence

[1] By real aperture array imaging we mean imaging with physical arrays.

[2] The receive process is not discussed here since it is very similar to the receive process in synthetic aperture imaging discussed in Chapter 5. The main difference between a real and a synthetic aperture is that a real aperture allows for simultaneous acquisition from all array elements while using a synthetic aperture one acquire the data sequentially.

on grating lobes are mentioned. In Section 6.2, some important properties of broadband array imaging are treated, in particular, the influence of the array element size and the interrogated wave form is discussed. Finally in Section 6.2, parallel array imaging is briefly introduced and a matrix expression for a parallel DAS beamformer is presented.

6.1 Conventional Narrowband Transmit Beamforming Analysis

In Chapter 1, conventional delay-and-sum (DAS) beamforming was introduced. In particular, the effects of apodization as well as spatial sampling consequences, such as grating- and sidelobes are discussed. We will, therefore, assume that the reader is already familiar with these concepts. In this section we will extend the analysis to transmit beamforming, still treating the array elements as point sources and assuming CW excitation. These assumptions will then be relaxed in the next section where broadband array imaging, with finite sized elements, are discussed. This is done here in order to show the differences between the common CW far-field beam pattern analysis, where idealized arrays with point-like elements are assumed, and a broadband excitation with more realistic array models.

Consider a 1D phased array consisting of L elements equally spaced along the x-axis with an array pitch d, see Figure 1.8 in Chapter 1. The position vector, $\mathbf{r}_n$, for the nth element is then $\mathbf{r}_n = (x_n, 0, 0)$. Since the waves will spread spherically from the point-like elements, the wave field at an observation point, $\mathbf{r}$, will become a sum of the spherical wave contributions from each element. If we also apply the paraxial approximation introduced in Section 1.2 for the amplitude terms of the waves, then the CW field at $\mathbf{r}$ can be approximated by

$$H(\mathbf{r}, t) \approx \sum_{n=0}^{L-1} w_n e^{j\omega(t - \frac{|\mathbf{r}-\mathbf{r}_n|}{c_\mathrm{p}} + \tau_n)}, \tag{6.1}$$

where w_n is the apodization weight for the nth element and τ_n is the corresponding focusing delay. Focusing at a point $\mathbf{r}_\mathrm{F}$ can now achieved by assigning the time delay, $\tau_n = \frac{|\mathbf{r}_\mathrm{F}-\mathbf{r}_n|}{c_\mathrm{p}}$, for $n = 0, 1, \ldots, L-1$. If $\mathbf{r}_\mathrm{F}$ is located on the z-axis then the array is focused but not steered, otherwise the array is also steered.

To show the effect of focusing and beam steering, consider a 16 element phased array with a pitch $d = 2\lambda$ and focused at z=50 mm (c_p = 1500

m/s, $\lambda = 0.5$ mm). This is an undersampled array and grating lobes will therefore appear. Two example CW fields for the array are in shown Figure 6.1. Figure 6.1(a) shows the wave field when the array is focused but not steered and the corresponding beam pattern, at $z = 50$ mm, is shown in Figure 6.1(c). Similarily, the wave field for the array steered -20 degrees, is shown in Figure 6.1(b) and the corresponding beam pattern is shown in Figure 6.1(d).

Clearly, the undersampling manifests itself in several grating lobes both for the non-steered and the steered array. Also, as can be seen in Figure 6.1(b) and (d), beam steering not only steers the main lobe but also the grating lobes. Measurements often are performed in such way that it is more likely for a scatterer to be located close to the center of the array and the problem with ghost responses due to grating lobes are, therefore, more severe for steered systems than for non-steered systems.

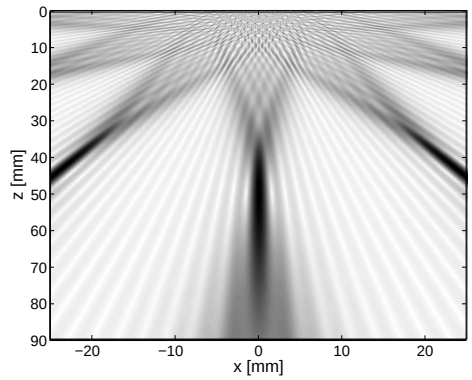
Note, however, that the model (6.1) neither considers the beam directivity of the elements nor the fact that the length of a the transmit pulse—if a broadband pulse is used—also influences the grating- and sidelobe amplitudes. These topics are further discussed in the next section.

6.2 Broadband Transmit Beamforming with Finite Sized Array Elements

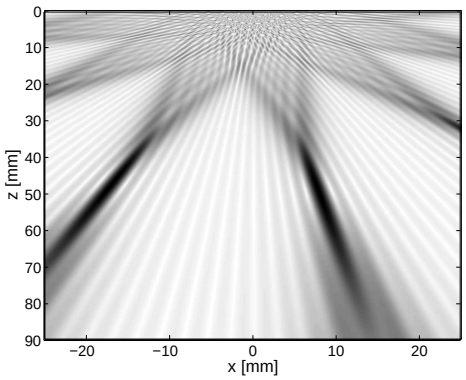
In Chapter 5 it was shown that the transducer size had a large impact on the lateral resolution for SAFT processed data. In this section we will discuss some of the effects that the transducer size, or more specifically, the array element size has on the transmitted wave field for broadband excitation of a real aperture array.

In the narrow band analysis in the previous section we saw that by steering the beam also the grating lobes was steered. Due to the paraxial approximation, the max amplitude in the main beam and in the grating lobes were the same, cf. Figure 6.1(b). For a more realistic array, with finite sized array elements, this will in general not be the case [78]. To see this, let us study an example.

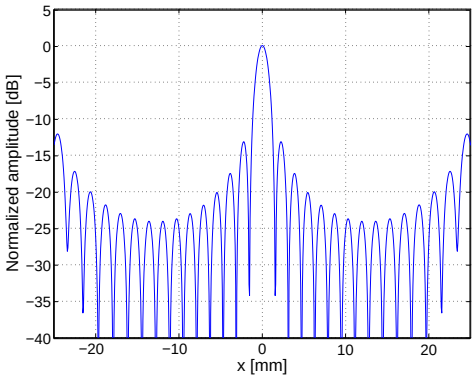
Consider again a phased array with the same pitch and aperture as was used in the CW example in the previous section, that is, a 16 element array with a pitch of 1 mm focused at z =50 mm. But now the elements have a finite size; they are circular with a diameter of 0.9 mm. Also, a



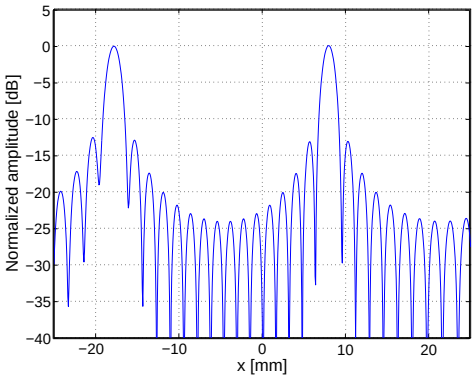
(a) CW field for a non-steered phased array focused at z=50 mm.



(b) CW field for a steered phased array focused at ($x = -18$ mm,$z = 50$ mm).



(c) Beam pattern at $z = 50$ mm for a non-steered phased array.



(d) Beam pattern at $z = 50$ mm for steered phased array.

Figure 6.1: Continuous wave fields and beam patterns for an undersampled 1D phased array. The wavelength, λ, was 0.5 mm and a 16 element array with a pitch of 2λ was simulated.

broadband pulse is used, which is shown in Figure 3.4(a) in Chapter 3. The pulse has a center frequency of approximately 3 MHz which corresponds to a wavelength, λ, of 0.5 mm. Thus, the array is undersampled and grating lobes should appear.

Simulated wave field snapshots, at three time instants, for this array setup are presented in Figure 6.2. The snapshots where computed using the DREAM method, for the array focused at $z = 50$ mm, both for an un-steered and a steered configuration.[3]

By comparing the CW interference patterns, shown in Figure 6.1(a) and (b), with the broadband fields, shown in Figure 6.2(a) and (b), one can observe that the somewhat chaotic interference patterns close to the array seen in the CW images cannot be observed in the broadband snapshots. Instead, the waves are almost planar shortly after they have left the array for the broadband case.

As time evolves, and the waves approaches the focal points, the wave fields become more concentrated. When they have reached $z = 50$, shown in Figure 6.2(e) and (f), the waves fields are rather narrow.

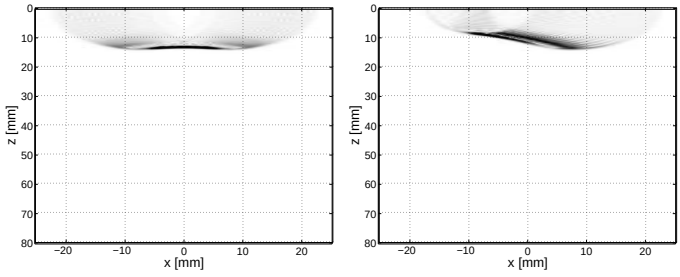
However, due to the undersampling of the array aperture, one can also observe significant acoustic energy outside the main beams, see Figure 6.2(c)–(f). This is more easily seen in the profile plots of the wave fields shown in Figure 6.3. These profile plots were taken when the waves had reached the focal points. By comparing the profile plots, shown in Figure 6.3, with the CW beampatterns, shown in Figure 6.1(c) and (d), one can notice that the amplitudes of the grating lobes, with respect to the main lobes, are significantly lower in the broadband profile plots. The reason for this behavior is twofold:

- The length of the pulse is only about two cycles (of the center frequency), cf. Figure 3.4(a). Hence, only the signal from two elements will add constructively at the grating lobe, in contrary to CW excitation, where the signals from all elements adds in phase at the grating lobes [9].

- The beam directivity of the elements results in significantly reduced amplitudes at observation points far from corresponding array element.

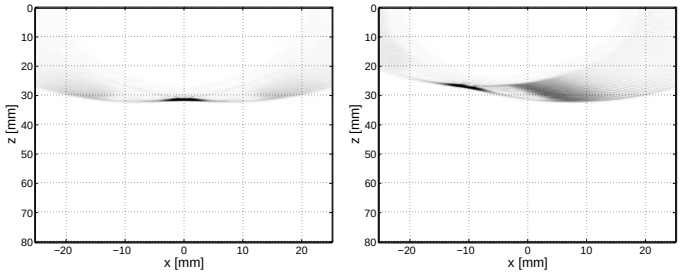
The beam directivity of the elements will, therefore, influence the beam steering properties of the array; the element's beam directivity essentially

[3]The array was steered -20 degrees.



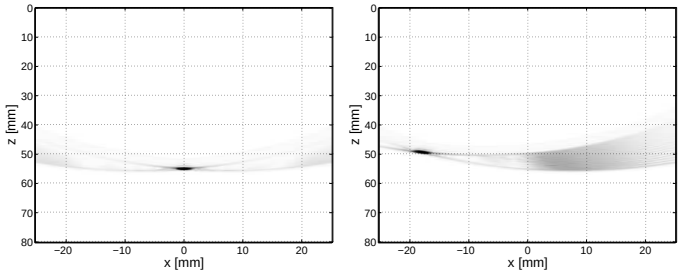
(a) Snapshot at $t = 10\mu$s a for a non-steered phased array.

(b) Snapshot at $t = 10\mu$s for a steered phased array.



(c) Snapshot at $t = 22\mu$s for a non-steered phased array.

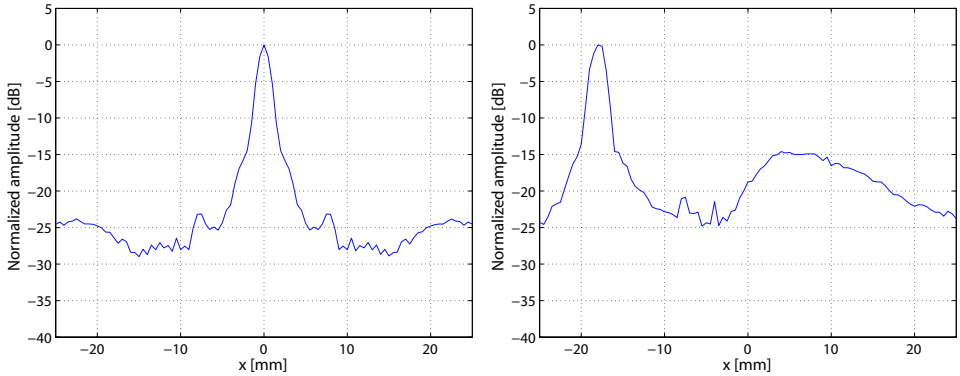
(d) Steered Array. Snapshot at $t = 22\mu$s.



(e) Snapshot at $t = 37.6\mu$s for a non-steered phased array (where the wave is close to the focal point).

(f) Snapshot at $t = 37.6\mu$s for a steered array (where the wave is close to the focal point).

Figure 6.2: Broadband field snapshots for an undersampled phased array with circular array elements. The center frequency of the broadband pulse corresponds approximately to a wavelength, λ, of 0.5 mm and a 16 element array with a pitch of 2λ was simulated.



(a) Non-steered Array. Profile at $t = 37.6\mu$s

(b) Steered Array. Profile at $t = 37.6\mu$s

Figure 6.3: Broadband profiles for the wave fields shown in Figure 6.2(e) and (f) respectively.

reduces the effective aperture of the array. To illustrate this, consider the profile plot of the wave field for one element, at $x = 7.5$ mm, shown in Figure 6.4. The profile plot was taken at $t = 34.1\mu$s and the pulse has then traveled approximately 50 mm in the medium. As can be observed, the amplitude at $x = -10$ mm is about 20 dB lower than directly under the array element, at $x = 7.5$. Thus, if the steering angle is high then the contributions from elements at a horizontal distance far from the focal point will become very low.

From the discussion above it is clear that many factors must be taken into account when choosing a proper array layout for broadband ultrasonic array imaging. The simplified far-field CW model (6.1) is often not sufficient since many important properties are not considered by that model. To successfully analyze broadband array imaging one must in particular consider the size effects of the array elements as well as the waveform of interrogated pulse.

We will come back to these topics in Chapters 7 and 8 when we discuss reconstruction methods for ultrasonic array imaging.

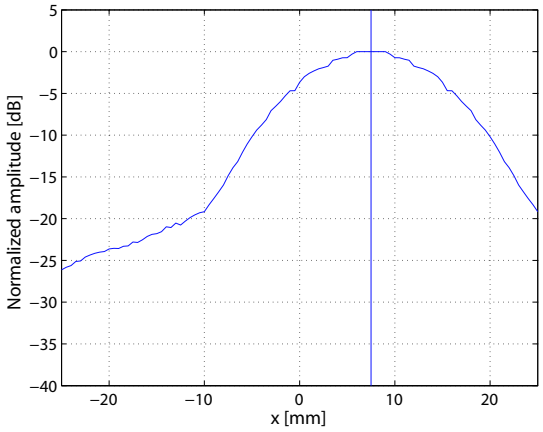


Figure 6.4: Broadband beam profile for a single circular array element with a diameter of 0.9 mm at $t = 34.1\mu s$. The vertical line indicates the position of the array element.

6.3 Parallel Receive Beamforming

Parallel beamforming is a method developed to increase the frame rate of array imaging systems. The basic idea is the transmit a broader-than-normal beam and then beamform several image lines from that transmission. This will decrease the number of re-transmissions needed to cover the whole ROI, thus increasing the image frame rate [79, 80].

In this section we will consider parallel DAS beamforming whilst more advanced approaches will be discussed in Chapter 8.

A parallel receive mode DAS beamformer consists of several delay-line units, thus enabling focusing at many points. A schematic block diagram of a parallel DAS beamformer is shown in Figure 6.5. Similar to the SAFT algorithm, the parallel DAS beamformer can be expressed using matrix formalism as shown below.

Let us for simplicity consider 2D parallel beamforming where the ROI has been discretized and represented by an $M \times N$ matrix $\mathbf{O}$ as described in Chapter 4.

Consider again a 1D array with L elements positioned along the x-axis. Let $\mathbf{y}_n$ denote the A-scan received by the nth array element, $\mathbf{y}_n = [u_o(x_n, t_0)$ $u_o(x_n, t_1) \cdots u_o(x_n, t_{K-1})]^T$, and $\mathbf{y}$, the (vectorized) B-scan containing the A-scan data for all elements, $\mathbf{y} \triangleq [\mathbf{y}_0^T \mathbf{y}_1^T \cdots \mathbf{y}_{L-1}^T]^T$. Then, the operation of

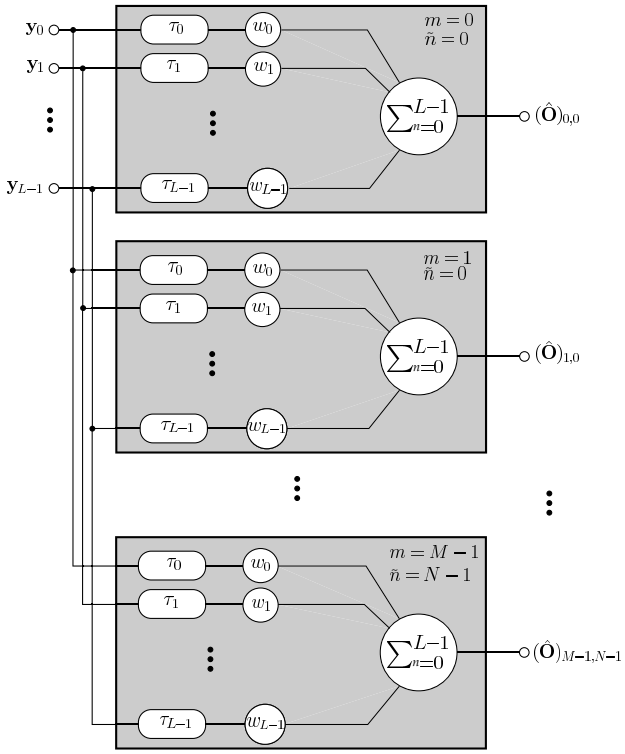


Figure 6.5: Block scheme of a parallel delay-and-sum beamfomer. Each delay-line unit is focused at a different observation point.

one of the delay-line units, focused at the observation point $(x_{\tilde{n}}, z_m)$, can be expressed as

$$\begin{aligned}(\hat{\mathbf{O}})_{m,\tilde{n}} &= \sum_{n=0}^{L-1} w_n u_{\mathrm{o}}(x_n, \tau_n(x_{\tilde{n}}, z_m)) \\ &= \sum_{n=0}^{L-1} w_n \mathbf{y}_n^T \mathbf{d}_n^{(m,\tilde{n})} \\ &= \begin{bmatrix} w_0\mathbf{d}_0^{(m,\tilde{n})} \\ w_1\mathbf{d}_1^{(m,\tilde{n})} \\ \vdots \\ w_{L-1}\mathbf{d}_{L-1}^{(m,\tilde{n})} \end{bmatrix}^T \mathbf{y},\end{aligned} \tag{6.2}$$

where w_n is an apodization weight, $\tau_n(x_{\tilde{n}}, z_m)$ is the focusing delay for fo-

cusing at the observation point $(x_{\tilde{n}}, z_m)$, and $\mathbf{d}_n^{(m,n)}$ is a vector of length K which is zero everywhere except at the kth index, $k \doteq \frac{\tau_l(x_n,z_m)}{T_\mathrm{s}}$, where it is 1.

Now, the total vectorized image, $\hat{\mathbf{o}} = \mathrm{vec}(\hat{\mathbf{O}})$, for all observation points, can be expressed by appending the *delay vectors*, $\mathbf{d}_n^{(m,\tilde{n})}$, for the remaining observation points, according to

$$\begin{aligned}\hat{\mathbf{o}} &= \begin{bmatrix} w_0\mathbf{d}_0^{(0,0)} & w_0\mathbf{d}_0^{(1,0)} & \cdots & w_0\mathbf{d}_0^{(M-1,N-1)} \\ w_1\mathbf{d}_1^{(0,0)} & w_1\mathbf{d}_1^{(1,0)} & \cdots & w_1\mathbf{d}_1^{(M-1,N-1)} \\ \vdots & \vdots & & \vdots \\ w_{L-1}\mathbf{d}_{L-1}^{(0,0)} & w_{L-1}\mathbf{d}_{L-1}^{(1,0)} & \cdots & w_1\mathbf{d}_{L-1}^{(M-1,N-1)} \end{bmatrix}^T \mathbf{y} \\ &= \mathbf{K}_{\mathrm{DAS}}\mathbf{y}.\end{aligned} \tag{6.3}$$

The operation described by (6.3) is similar to that of a *dynamic focus* beamformer. A dynamic focusing beamformer has time-dependent focus delays which, in receive mode, adapts focusing at the same rate as the wave propagates. This results in an extended region where the beam is focused compared to a fixed focus beamformer. The matrix $\mathbf{K}_{\mathrm{DAS}}$ in (6.3) acts as a dynamic focus beamformer that focuses, in parallel, at *all* observation points.

Part III

Bayesian Ultrasonic Image Reconstruction

CHAPTER 7

Ultrasonic Image Reconstruction Methods

RECALL from Chapter 1, that the main objective in all ultrasonic imaging applications is that we want to gain knowledge of an object under study by means of studying the ultrasonic scattering from the object. We perform experiments and by observing the recorded data we draw conclusions whether, for example, a defect is critical or not or, in a medical application, whether surgery must be performed or not. The data can however be difficult to interpret for a human operator since the wave propagation in the medium can be complex and the measurement apparatus may distort the received scattering information. That is, the user must have deep knowledge about the measurement system in order to draw accurate conclusions from the experiment.

The focusing methods discussed in Chapters 5 and 6 can alleviate this process to some degree. Focusing, or beamforming, is performed mainly for two reasons:

- The obtained ultrasonic images are easier to interpret when the acoustic energy is concentrated to a small region.
- The acoustic energy, in the region that we want to study, must be high enough to obtain a sufficient SNR since, otherwise, the data will not contain enough information regarding the scattering objects.

It should however be noted that receive mode beamforming does not give us any additional information than what is already in the unprocessed data

received from the individual array elements, rather the opposite. This fact is known as the *data processing inequality* which states that: no processing of the data, deterministic or random, can increase the information that the data contains of the parameter we want to estimate [81]. In practice, all processing that we perform on the data will in fact destroy some amount of information due limited numerical precision, etc.

A relevant question that we can ask us is: *How should we process our data to utilize its information about the scattering strengths in the best possible way to obtain a result that can be easily understood by a human?* In many applications DAS based methods may be sufficient for the particular purpose but they are not optimal in the sense of utilizing all available information. A reasonable answer to the question formulated above is that we should use all information that is available, *both from data and from prior knowledge*. We know, for example, many properties of the measurement system and the linear model described in Chapter 4 gives us a tool for modeling the system's "distortion". This knowledge can then be used to compensate for the distortion and thereby obtaining images that should be easier for a user to understand. In other words, if we formulate the problem as an *image reconstruction problem* we may be able to improve the image quality compared to traditional DAS beamforming.

The system model is however not the only prior information that we have. We also know that the scattering amplitudes cannot become arbitrarily large since the ultrasonic transducer emits a finite amount of acoustic energy. In some cases we may even know the sign of the scattering amplitude. An optimal reconstruction method, that accounts for all information, should therefore use both the prior information and the information from the experiments.

We must bear in mind that we cannot expect to fully recover the true scattering image (with arbitrary precision) since we never have complete information. The data are acquired using a limited aperture, often from a a 1D array, and we cannot expect to reconstruct a 3D object based on 2D data. Furthermore, the data are always corrupted by thermal noise and the models we use do always have limitations. It is evident that there will always be an amount of uncertainty in our estimates. This lack of complete information usually manifest itself in making the problem *ill-posed*. That is, if we deductively try to find an inverse to the propagation operator, then this inverse will not be stable in the sense that a small disturbance can cause a large change in the estimates. When solving inverse problems deductively one must therefore take special measures to stabilize the solution. This

process is usually called regularization and a popular method is the so-called Tikhonov regularization [31]. Regularization typically involves tuning parameters to compromise between fidelity to the data and the stability of the solution.

Another approach is to view the problem as a Bayesian inference problem. In the Bayesian philosophy the prior information has a central role and, for inference problems, the regularization parameters can be directly connected to the prior knowledge. In fact under certain conditions, the Tikhonov regularization schemes can be seen as a special case of the linear minimum mean squared estimator derived using Bayesian analysis, although Tikhonov regularization is based on totally different assumptions.

As mentioned in Chapter 1, ultrasonic imaging experiments easily produce large data volumes and the dimensionality of the problems are often very high. Practically, there is a trade off between computational complexity and the quality of the results. The reconstruction methods proposed in this chapter, have been developed with that in mind. The problem has been formulated from a Bayesian perspective where regularization parameters can be deduced from the prior information of the scattering strength, typically the scattering mean value or the variance, and the variance of the noise.

In this chapter two Bayesian estimators are derived, one linear and one non-linear. The linear estimator is based on the assumption that the mean and variance of the scattering strengths are known. This results in the linear minimum mean-squared error (MMSE) estimator. The non-linear estimator is based on the assumptions that we are imaging objects with positive scattering strengths that we know the mean of. It is shown that the resulting maximum *a posteriori* (MAP) solution is found by solving a quadratic programming (QP) problem. Also found in this chapter is a comparison of the two Bayesian methods with common methods for solving ill-posed inverse problems.

The chapter is organized as follows: In Section 7.1, a short introduction to Bayesian inference is presented and the maximum entropy principle, used for assigning prior probability density functions, is introduced. This section also includes derivations of the optimal linear MMSE estimator, as well as the optimal non-linear exponential MAP estimator. In section, 7.2, some common methods for solving ill-conditioned image reconstruction problems are reviewed and the relation to the Bayesian estimators are discussed. In particular it is shown that, under certain conditions, these methods can be seen as a special case of the linear Bayesian estimator discussed in Sec-

tion 7.1. This section also contains a discussion of the maximum likelihood estimator and its properties. Finally, in Section 7.3, some concluding remarks are given.

7.1 Bayesian Image Reconstruction

In Bayesian estimation theory a probability density function (PDF) describes the uncertainty, or the degree of belief, of the true value of the variable under study [82]. In other words, a PDF is thought of as a *carrier of information*. In our case we want to study the scattering strength at each image point described by the vector $\mathbf{o}$. To accomplish this we perform an experiment that gives us some data $\mathbf{y}$. Note, that our knowledge regarding $\mathbf{o}$ is usually not only coming from the data. Typically we know, without seeing any data, that $\mathbf{o}$ cannot have an arbitrary scattering amplitude; the scattering amplitude must at least be limited to some interval. The PDF, $p(\mathbf{o}|I)$, that describes our knowledge before seeing the data is usually called "the prior" and I denotes any background information that we have regarding $\mathbf{o}$. When we have performed an experiment we have received more information about $\mathbf{o}$ and the updated PDF *after* seeing the data is given by the *posterior* PDF, $p(\mathbf{o}|\mathbf{y}, I)$. If the measurement contained relevant (new) information of $\mathbf{o}$ then the posterior PDF should be more concentrated around the true value of $\mathbf{o}$.[1]

Bayes' theorem,

$$p(\mathbf{o}|\mathbf{y}, I) = p(\mathbf{o}|I)\frac{p(\mathbf{y}|\mathbf{o}, I)}{p(\mathbf{y}|I)}, \tag{7.1}$$

describes how to update our knowledge of the variable, $\mathbf{o}$, when new data, $\mathbf{y}$, is observed. The conditional PDF $p(\mathbf{y}|I)$ is a normalization factor not dependent on $\mathbf{o}$ and $p(\mathbf{y}|\mathbf{o}, I)$ describes the *likelihood* of seeing the data $\mathbf{y}$ given $\mathbf{o}$.

In this thesis the measurement errors are assumed to be Gaussian and $p(\mathbf{y}|\mathbf{o}, I)$ will therefore be given by the multidimensional Gaussian distribution,

$$p(\mathbf{y}|\mathbf{o}, I) = \frac{1}{(2\pi)^{\frac{MN}{2}}|\mathbf{C}_\mathrm{e}|^{1/2}} e^{-\frac{1}{2}(\mathbf{y}-\mathbf{P}\mathbf{o})^T \mathbf{C}_\mathrm{e}{}^{-1}(\mathbf{y}-\mathbf{P}\mathbf{o})}, \tag{7.2}$$

[1] If the prior PDF and the posterior PDF are identical we know that the data contains no new information of the variable.

where $\mathbf{P}$ is the propagation matrix from (4.23) and the size of the image to reconstruct is $M \times N$.

The two most common estimators in Bayesian estimation applications are the *maximum a posteriori* (MAP) estimator and the minimum mean square error (MMSE) estimator. The MAP estimate is found by maximizing $p(\mathbf{o}|\mathbf{y}, I)$ and take the maximizing argument as the estimate,

$$\hat{\mathbf{o}}_{\text{MAP}} = \arg\max_{\mathbf{o}} p(\mathbf{o}|\mathbf{y}, I) = \arg\max_{\mathbf{o}} p(\mathbf{o}|I)p(\mathbf{y}|\mathbf{o}, I), \tag{7.3}$$

where (7.1) was used in the last equality. The MAP estimate is thus the value $\mathbf{o}$ that is most probable given both data and prior information.

The MMSE estimate is given by

$$\hat{\mathbf{o}}_{\text{MMSE}} = \arg\min_{\mathbf{o}} \ \mathrm{E}\{\|\mathbf{o} - \hat{\mathbf{o}}\|^2\} \tag{7.4}$$

where $\mathrm{E}\{\cdot\}$ is the expectation operator. It should be emphasized that it is only when both $\mathbf{o}$ and $\mathbf{e}$ are Gaussian that the MMSE estimate is practically tractable. The reason is that when both $\mathbf{o}$ and $\mathbf{e}$ are Gaussian the MMSE estimate will be a linear combination of the data. If $\mathbf{o}$ and $\mathbf{e}$ are not Gaussian, then Markov chain Monte-Carlo (MCMC) methods can be used to draw samples from $p(\mathbf{o}|\mathbf{y}, I)$ and then average over these samples [83]. This is however very time consuming for high dimensional problems and is therefore not practical for the applications described in this thesis.

Furthermore, if both $\mathbf{o}$ and $\mathbf{e}$ are Gaussian then the MAP and the linear MMSE estimate will be identical. Although this is well known we present the respective solutions here for practical reasons. Consider first the linear MMSE estimate, defined as the estimate $\hat{\mathbf{o}} = \mathbf{K}\mathbf{y}$ that minimizes the criterion,

$$J_{\text{MSE}}(\mathbf{K}) = \mathrm{E}\{\|\mathbf{o} - \mathbf{K}\mathbf{y}\|^2\}. \tag{7.5}$$

with $\mathbf{y} = \mathbf{P}\mathbf{o} + \mathbf{e}$. The linear MMSE estimator, $\mathbf{K}_{\text{MMSE}}$, is then given by

$$\begin{aligned}
\mathbf{K}_{\text{MMSE}} &= \arg\min_{\mathbf{K}} \ \mathrm{E}\{\|\mathbf{o} - \mathbf{K}\mathbf{y}\|^2\} \\
&= \arg\min_{\mathbf{K}} \ \mathrm{E}\{(\mathbf{o} - \mathbf{K}\mathbf{y})^T(\mathbf{o} - \mathbf{K}\mathbf{y})\} \\
&= \arg\min_{\mathbf{K}} \ \mathrm{tr}\{\mathbf{C}_{\text{o}}\} - 2\,\mathrm{tr}\{\mathbf{K}^T\mathbf{C}_{\text{o}}\mathbf{P}^T\} + \mathrm{tr}\{\mathbf{K}\mathbf{P}\mathbf{C}_{\text{o}}\mathbf{P}^T\mathbf{K}^T\} \\
&\quad + \mathrm{tr}\{\mathbf{K}\mathbf{C}_{\text{e}}\mathbf{K}^T\}
\end{aligned} \tag{7.6}$$

where we assumed that $\mathbf{o}$ and $\mathbf{e}$ are mutually independent with covariance matrices $\mathbf{C}_{\text{o}} = \mathrm{E}\{\mathbf{o}\mathbf{o}^T\}$ and $\mathbf{C}_{\text{e}} = \mathrm{E}\{\mathbf{e}\mathbf{e}^T\}$, respectively. Taking the derivative with respect to $\mathbf{K}$ and equating to zero in (7.6) results the linear MMSE

estimate

$$\hat{\mathbf{o}}_{\text{MMSE}} = \mathbf{K}_{\text{MMSE}}\mathbf{y} = \mathbf{C}_{\text{o}}\mathbf{P}^T(\mathbf{P}\mathbf{C}_{\text{o}}\mathbf{P}^T + \mathbf{C}_{\text{e}})^{-1}\mathbf{y} \tag{7.7}$$

The MAP estimate can be found by again assuming that $\mathbf{o}$ and $\mathbf{e}$ are zero mean Gaussian. Now take the negative logarithm of (7.1), excluding all terms not dependent on $\mathbf{o}$, which results in the criterion

$$J_{\text{MAP}}(\mathbf{o}) = \frac{1}{2}\mathbf{o}^T{\mathbf{C}_{\text{o}}}^{-1}\mathbf{o} + \frac{1}{2}(\mathbf{y} - \mathbf{P}\mathbf{o})^T{\mathbf{C}_{\text{e}}}^{-1}(\mathbf{y} - \mathbf{P}\mathbf{o}) \tag{7.8}$$

Differentiating (7.8) with respect to $\mathbf{o}$ and equating to zero gives the MAP solution

$$\hat{\mathbf{o}}_{\text{MAP}} = (\mathbf{C}_{\text{o}}^{-\mathbf{1}} + \mathbf{P}^T\mathbf{C}_{\text{e}}^{-\mathbf{1}}\mathbf{P})^{-1}\mathbf{P}^T\mathbf{C}_{\text{e}}^{-\mathbf{1}}\mathbf{y} = \mathbf{K}_{\text{MAP}}\mathbf{y}. \tag{7.9}$$

The equivalence of (7.9) and (7.7) is discussed in Appendix 7.A.1. This equivalence is of practical interest since the matrix inverse in (7.7) and the inverse $(\mathbf{C}_{\text{o}}^{-\mathbf{1}} + \mathbf{P}^T\mathbf{C}_{\text{e}}^{-\mathbf{1}}\mathbf{P})^{-1}$ in (7.9) may be of different size.[2] Recall from Chapter 4 that the A-scan dimension is K and the number of measurements is L, resulting in a $KL \times KL$ noise covariance matrix $\mathbf{C}_{\text{e}}$, and that the number of observation points is MN, resulting in an $MN \times MN$ covariance matrix $\mathbf{C}_{\text{o}}$. Hence, if KL is larger than MN, then it is less computationally demanding to use the MMSE form compared to the MAP form and vice versa.[3]

Another important feature of Bayesian methods is that we obtain a measure of the precision of the estimate. For Gaussian $\mathbf{o}$ and $\mathbf{e}$, the PDF for the error, $\epsilon = \hat{\mathbf{o}}_{\text{MMSE}} - \mathbf{o}$, for the linear MMSE estimate is also Gaussian with zero mean and the covariance matrix,

$$\mathbf{C}_\epsilon = \text{E}\{(\hat{\mathbf{o}}_{\text{MMSE}} - \mathbf{o})(\hat{\mathbf{o}}_{\text{MMSE}} - \mathbf{o})^T\} = \text{E}\{\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T\}, \tag{7.10}$$

given by (see also Appendix 7.A.2) [84]

$$\begin{aligned}\mathbf{C}_\epsilon &= \mathbf{C}_{\text{o}} - \mathbf{C}_{\text{o}}\mathbf{P}^T(\mathbf{C}_{\text{e}} + \mathbf{P}\mathbf{C}_{\text{o}}\mathbf{P}^T)^{-1}\mathbf{P}\mathbf{C}_{\text{o}} \\ &= (\mathbf{C}_{\text{o}}^{-\mathbf{1}} + \mathbf{P}^T\mathbf{C}_{\text{e}}^{-\mathbf{1}}\mathbf{P})^{-1}.\end{aligned} \tag{7.11}$$

The mean squared error (MSE) is the sum of the diagonal elements of $\mathbf{C}_\epsilon$

$$J_{\text{MSE}} \triangleq \text{tr}\{\mathbf{C}_\epsilon\}. \tag{7.12}$$

[2] A matrix inverse requires a number of operations in the order of n^3 where n is the dimension of the matrix.

[3] A typical A-scan dimension, for the experiments presented later in this thesis, is $K = 1000$, and $L = 16$ for the real aperture array experiments. If $\mathbf{O}$, for example, is an 100×50 image, then $\mathbf{P}$ has the dimension 16000×5000, which is a normal size of the problems treated in this thesis.

If the noise energy is large, that is $\mathbf{C}_\mathrm{e} \gg \mathbf{P}\mathbf{C}_\mathrm{o}\mathbf{P}^T$, then $\mathbf{C}_\epsilon$ is approximately equal to $\mathbf{C}_\mathrm{o}$. We know then that our reconstruction performance is poor and we should choose to design the experiment differently. Moreover, since each diagonal element in $\mathbf{C}_\epsilon$ corresponds to the variance at an observation point, the expected error due to a chosen focusing method can be evaluated by organizing the diagonal elements into an $M \times N$ image. By observing the expected mean-squared error image it can then be directly seen in what regions the reconstruction performance is acceptable and where it is not. This topic will be further discussed in Chapter 9.

7.1.1 Assigning Prior Probability Density Functions — The Maximum Entropy Principle

So far we have not discussed how to obtain the prior $p(\mathbf{o}|I)$. In the linear MAP estimate (7.9) it was assumed that the prior was Gaussian. The Gaussian PDF is uniquely determined by the mean vector and the covariance matrix; hence it seems reasonable to assign a Gaussian prior if the only information about $\mathbf{o}$ we have is the mean and the variance. The question is if this assumption can be justified?

If we have no prior information of the scattering amplitudes, then all amplitudes are equally likely. This corresponds to a uniform distribution with infinite support. Hence the uniform PDF is the most "uninformative" of all distributions; the uniform distribution represents the most noncommittal distribution since no amplitude is favored over any other. However, by using a uniform prior PDF with infinite support we have not used that we know the mean and variance which is undesirable. A reasonable requirement is thus that the prior PDF should be maximum non-committal with respect to the, in this case, known mean and variance. Hence, we want the prior be as uninformative as possible but take into account the information that we actually have.

To judge how uninformative a PDF is we need a consistent measure of its "uniformity". This measure can then me maximized with respect to the constraints given by our prior information; thus yielding a prior that is honest about what we know. The measure should be continuous, consistent, and a monotonic increasing function in the sense that if n events are equally likely, then the measure should increase if n increases. That is, we are more uncertain which event that will occur if there are many possibilities than when there are few. The only known function with these properties is the

information entropy [82]

$$H(p(x)) = -\int_{-\infty}^{\infty} p(x) \log p(x) dx. \tag{7.13}$$

Thus by maximizing (7.13) with respect to constraints, given by our *a priori* knowledge, a PDF should be obtained that is the most honest with regard to what we know and what we do not know.

Note that the optimization is with respect to a *function* $p(x)$, required to have the property

$$\int_{-\infty}^{\infty} p(x) dx = 1, \tag{7.14}$$

since it is a probability density function. If, as assumed above, we know the first and second order moments of the PDF, then these two additional constraints enter in a similar fashion as the constraint (7.14).

The problem of finding a PDF that maximizes (7.13) with respect to constraints can be solved with the method of Lagrange multipliers [85]. Note that we are searching for a *stationary function* here; thus, finding the distribution that maximizes the entropy (7.13) given the constraints is a *variational* problem. The problem is to extremizing, or finding an *extremal*, to a functional, $J(\cdot)$, of the form

$$J(p(x)) = \int_a^b L(x, p(x), x', p'(x)) dx \tag{7.15}$$

where the function $L(\cdot)$ is the, so-called, *Lagrangian*. An extremal, $p(x)$, to (7.15) must then satisfy the *Euler* or *Euler-Lagrange* equation [47]

$$L_{p(x)} - \frac{d}{dx} L_{p'(x)} = 0 \tag{7.16}$$

where $L_{p(x)}$ is the derivative of the Lagrangian with respect to $p(x)$ and $L_{p'(x)}$ is the derivative with respect to $p'(x)$.

Below we present three common maximum entropy distributions, obtained from different moment constraints where, in particular, the last two examples is of great interest for ultrasonic image reconstruction purposes.

Three Maximum Entropy Distributions

Let us study three examples where we in the first only know that our variable is bounded to an interval, whereas in the second example we know the mean

and that the variable is positive. Finally in the last example we know both the mean and the variance.

Example 7.1 Known interval, $x \in [a, b]$.

If the only prior knowledge is that the variable x is known to be in the interval $[a, b]$ the only constraint is that $p(x)$ must fulfill (7.14) in that interval. Thus, the functional to maximize is

$$\begin{aligned} J(p(x)) &= H(p(x)) - \lambda \left(\int_a^b p(x)dx - 1 \right) \\ &= \int_a^b \left(-p(x) \log p(x) - \lambda p(x) \right) dx + \lambda. \end{aligned} \tag{7.17}$$

The derivatives of the Lagrangian with respect to $p(x)$ and $p'(x)$, respectively, then become

$$\begin{aligned} L_{p(x)} &= -\log p(x) - 1 - \lambda \\ L_{p'(x)} &= 0. \end{aligned} \tag{7.18}$$

Inserting (7.18) into the Euler equation (7.16) results in a constant PDF

$$p(x) = e^{(1+\lambda)} = C. \tag{7.19}$$

Using the single constraint,

$$\begin{aligned} \int_a^b p(x)dx &= [xC]_a^b \\ &= C(b-a) = 1, \end{aligned} \tag{7.20}$$

results in

$$p(x) = \frac{1}{b-a} \tag{7.21}$$

which is the uniform distribution.

Thus, if it is only known that x is within an interval, then the PDF with maximum entropy is the uniform distribution, which of course coincides with the intuition.

Example 7.2 Known mean, $\bar{m}$, and x positive.

If the variable is positive with know mean, then the entropy functional becomes

$$\begin{aligned} J(p(x)) &= H(p(x)) - \lambda \left(\int_0^\infty p(x)dx - 1 \right) - \mu \left(\int_0^\infty x\, p(x)dx - \bar{m} \right) \\ &= \int_0^\infty \left(-p(x) \log p(x) - \lambda p(x) - \mu\, x\, p(x) \right) dx + \lambda + \mu \bar{m} \end{aligned} \tag{7.22}$$

where $\bar{m}$ denotes the mean value and λ and μ are Lagrange multipliers.

The derivatives of the Lagrangian now become

$$\begin{aligned} L_{p(x)} &= -\log p(x) - 1 - \lambda - \mu x \\ L_{p'(x)} &= 0 \end{aligned} \tag{7.23}$$

which, after insertion into (7.16), results in

$$p(x) = \exp(-1 - \lambda - \mu x) \tag{7.24}$$

Using the first constraint gives

$$\begin{aligned} \int_0^\infty p(x)dx &= \left[\exp(-1-\lambda) \frac{1}{\mu} e^{-\mu x} \right]_0^\infty \\ &= \exp(-1-\lambda) \frac{1}{\mu} = 1, \end{aligned} \tag{7.25}$$

and the second constraint,

$$\begin{aligned} \int_0^\infty x p(x)dx &= \left[\exp(-1-\lambda) \frac{(-\mu x - 1) e^{-\mu x}}{\lambda^2} \right]_0^\infty \\ &= \exp(-1-\lambda) \frac{1}{\mu^2} = \bar{m}. \end{aligned} \tag{7.26}$$

Combining (7.25) and (7.26) determines the value of the first Lagrange multiplier, $\lambda = \frac{1}{\bar{m}}$, and the PDF becomes

$$p(x) = \frac{1}{\bar{m}} \exp(-\frac{1}{\bar{m}} x). \tag{7.27}$$

which is the exponential distribution.

Now we return to the question whether we could assign a Gaussian prior if the mean and variance are known. This is discussed in the following example.

Example 7.3 Known mean, $\bar{m}$, and variance, σ^2.

The functional to maximize now has three Lagrange multipliers corresponding respectively to, the condition that $p(x)$ should fulfill (7.14), and the mean and variance constraints, which results in the entropy functional,

$$\begin{aligned} J(p(x)) = H(p(x)) - \lambda \left(\int_{-\infty}^{\infty} p(x)dx - 1 \right) - \mu \left(\int_{-\infty}^{\infty} x\, p(x)dx - \bar{m} \right) - \\ \eta \left(\int_{-\infty}^{\infty} (x-\bar{m})^2 p(x) - \sigma^2 \right) \\ = \int_{-\infty}^{\infty} \left(-p(x) \log p(x) - \lambda p(x) - \mu\, x\, p(x) - (x-\bar{m})^2 p(x) \right) dx + \\ \lambda + \mu\bar{m} + \eta\sigma^2 \end{aligned} \tag{7.28}$$

The derivatives of the Lagrangian then become

$$\begin{aligned} L_{p(x)} &= -\log p(x) - 1 - \lambda - \mu x - \eta(x-\bar{m})^2 \\ L_{p'(x)} &= 0. \end{aligned} \tag{7.29}$$

Now, let $\tilde{\lambda} = 1 + \lambda$ in (7.29), and then insert (7.29) into the Euler equation (7.16). We thus obtain

$$p(x) = \exp(\tilde{\lambda} + \mu x + \eta(x-\bar{m})^2). \tag{7.30}$$

By using the first constraint,

$$\int_{-\infty}^{\infty} p(x)dx = \exp\left(\frac{1}{4} \frac{4\tilde{\lambda}\eta - \mu^2 - 4\mu\eta}{\eta} \right) \frac{\sqrt{\pi}}{\sqrt{-\eta}} = 1, \tag{7.31}$$

and the second constraint,

$$\begin{aligned} \int_{-\infty}^{\infty} p(x)dx = \exp\left(\frac{1}{4} \frac{4\tilde{\lambda}\eta - \mu^2 - 4\mu\eta}{\eta} \right) \frac{\sqrt{\pi}}{\sqrt{-\eta}} \\ \frac{(-\mu + 2\eta\bar{m})}{2\eta} = \bar{m}, \end{aligned} \tag{7.32}$$

gives that $-\mu + 2\eta\bar{m} = 2\eta\bar{m}$, hence $\mu = 0$. By inserting $\mu = 0$ into (7.30), the third constraint becomes

$$\int_{-\infty}^{\infty} (x - \bar{m})^2 p(x) dx = -\frac{e^{\tilde{\lambda}}\sqrt{\pi}}{2\eta\sqrt{-\eta}} = \sigma^2. \tag{7.33}$$

Now inserting $\mu = 0$ also in (7.31) gives that

$$\frac{e^{\tilde{\lambda}}\sqrt{\pi}}{\sqrt{-\eta}} = 1 \tag{7.34}$$

and combining (7.34) and (7.33) yields

$$\begin{aligned} \eta &= -\frac{1}{2\sigma^2} \\ e^{\tilde{\lambda}} &= \frac{\sqrt{-\eta}}{\sqrt{\pi}} = \frac{1}{\sqrt{2\pi}\sigma}. \end{aligned} \tag{7.35}$$

Finally, inserting (7.35) and $\mu = 0$ into (7.30) yields the PDF

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{1}{2\sigma^2}(x - \bar{m})^2) \tag{7.36}$$

which is the Gaussian distribution.

Evidently, a Gaussian prior is the maximum entropy PDF if the first two moments are known.

In the next two sub-sections we will present two Bayesian estimators that are based on Gaussian and exponential prior PDFs, respectively.

7.1.2 The Optimal Linear Estimator

The Gaussian prior can, as described above, be motivated by being a maximum entropy distribution for known mean and variance. The Gaussian prior also leads to the (closed form) MAP estimator (7.9). The form of the MAP estimator,

$$\hat{\mathbf{o}} = \mathbf{K}_{\text{MAP}}\mathbf{y}, \tag{7.37}$$

has an attractive characteristic for practical computation. Each row, $\mathbf{k}_i$, in $\mathbf{K}_{\text{MAP}}$ constitutes of the optimal linear estimator for the corresponding observation point. The optimal estimator for the ith element in $\mathbf{o}$ can be

computed as the inner product $\mathbf{k}_i^T\mathbf{y}$ independently of the estimates for all other observation points. Thus, the optimal linear filtering can be performed in parallel using many CPUs thereby improving real-time performance.

Also, as will be further discussed in Section 7.2, the optimal linear estimator has similarities to standard Tikhonov type of regularization. To understand this, first recall the equivalence between the linear MMSE and the MAP estimate

$$\begin{aligned}\mathbf{K}_{\text{MAP}} &= (\mathbf{C}_{\text{o}}^{-\mathbf{1}} + \mathbf{P}^T\mathbf{C}_{\text{e}}^{-\mathbf{1}}\mathbf{P})^{-1}\mathbf{P}^T\mathbf{C}_{\text{e}}^{-\mathbf{1}} \\ &= \mathbf{C}_{\text{o}}\mathbf{P}^T(\mathbf{P}\mathbf{C}_{\text{o}}\mathbf{P}^T + \mathbf{C}_{\text{e}})^{-1} = \mathbf{K}_{\text{MMSE}}.\end{aligned} \tag{7.38}$$

The noise covariance matrix, $\mathbf{C}_{\text{e}}$ in (7.38), can be seen as a regulizer that stabilizes the solution. To see this more clearly, consider a singular value decomposition (SVD) of the propagation operator $\mathbf{P}$. The SVD of $\mathbf{P}$ is given by

$$\mathbf{P} = \mathbf{U}\mathbf{D}\mathbf{V}^T, \tag{7.39}$$

where the columns of $\mathbf{U}$ are the eigenvectors of $\mathbf{P}\mathbf{P}^T$, the columns of $\mathbf{V}$ are the eigenvectors of $\mathbf{P}^T\mathbf{P}$, and the *singular values* on the diagonal of $\mathbf{D}$ are the square roots of the eigenvalues of both $\mathbf{P}\mathbf{P}^T$ and $\mathbf{P}^T\mathbf{P}$. If we let s_i denote the ith singular value and we also assume that the elements in both $\mathbf{o}$ and $\mathbf{e}$ are uncorrelated so that $\mathbf{C}_{\text{e}} = \sigma_{\text{e}}^2\mathbf{I}$ and $\mathbf{C}_{\text{o}} = \sigma_{\text{o}}^2\mathbf{I}$, then by using (7.39), (7.38) becomes

$$\begin{aligned}\mathbf{K}_{\text{MAP}} &= \sigma_{\text{o}}^2\mathbf{P}^T(\sigma_{\text{o}}^2\mathbf{P}\mathbf{P}^T + \sigma_{\text{e}}^2\mathbf{I})^{-1} \\ &= \mathbf{P}^T(\mathbf{P}\mathbf{P}^T + \sigma_{\text{e}}^2/\sigma_{\text{o}}^2\mathbf{I})^{-1} \\ &= \mathbf{V}\begin{bmatrix} \frac{s_0}{s_0^2+\sigma_{\text{e}}^2/\sigma_{\text{o}}^2} & & & \\ & \frac{s_1}{s_1^2+\sigma_{\text{e}}^2/\sigma_{\text{o}}^2} & \mathbf{0} & \\ & & \ddots & \\ & \mathbf{0} & & \frac{s_{MN-1}}{s_{MN-1}^2+\sigma_{\text{e}}^2/\sigma_{\text{o}}^2} \end{bmatrix}\mathbf{U}^T.\end{aligned} \tag{7.40}$$

The factors $\frac{s_i}{s_i^2+\sigma_{\text{e}}^2/\sigma_{\text{o}}^2}$ in (7.40) are of the form

$$\frac{s_i}{s_i^2 + \alpha}, \tag{7.41}$$

where the term α stabilizes the solution. That is, even if the some of the singular values are close to zero, which potentially could lead to a division by zero, the factor α ensures that (7.41) is always bounded. The particular choice $\alpha = \sigma_{\text{e}}^2/\sigma_{\text{o}}^2$ is the optimal choice of the stabilizer; that is, the choice that results in the minimum mean-squared error of the reconstructions.

7.1.3 The Optimal Beamformer with Exponential Prior

As discussed in Example 7.2 above the maximum entropy PDF for a positive variable with known mean is given by an exponential distribution. It may therefore be motivated to assign an exponential distribution for the ultrasonic scattering when it is known that the scattering strength must be positive.[4]

If we assume that the prior PDF for $\mathbf{o}$ is exponential, and also that all $\mathbf{o}_i$s are independent identically distributed (IID), then the prior is given by

$$p(\mathbf{o}|I) = \Pi_{i=0}^{MN-1} \lambda_{\mathrm{o}} \exp(-\lambda_{\mathrm{o}} o_i). \tag{7.42}$$

If $\mathbf{e}$ is also Gaussian then the resulting posterior PDF becomes

$$\begin{aligned} p(\mathbf{o}|\mathbf{y}, I) \propto & \frac{1}{(2\pi)^{\frac{MN}{2}} |\mathbf{C}_{\mathrm{e}}|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{y}-\mathbf{P}\mathbf{o})^T {\mathbf{C}_{\mathrm{e}}}^{-1}(\mathbf{y}-\mathbf{P}\mathbf{o})\right) \\ & \Pi_{i=0}^{MN-1} \lambda_{\mathrm{o}} \exp(-\lambda_{\mathrm{o}} o_i). \end{aligned} \tag{7.43}$$

The MAP estimate can then be found by maximizing (7.43) subject to the constraint that all o_i must be positive, or equivalently by taking the negative logarithm of (7.43), and minimizing

$$\begin{aligned} -\log p(\mathbf{o}|\mathbf{y}, I) =& C + \frac{1}{2}(\mathbf{y}-\mathbf{P}\mathbf{o})^T {\mathbf{C}_{\mathrm{e}}}^{-1}(\mathbf{y}-\mathbf{P}\mathbf{o}) + \sum_{i=0}^{MN-1} \sigma_{\mathrm{o}}^2 o_i \\ =& C + \frac{1}{2}(\mathbf{y}-\mathbf{P}\mathbf{o})^T {\mathbf{C}_{\mathrm{e}}}^{-1}(\mathbf{y}-\mathbf{P}\mathbf{o}) + \lambda_{\mathrm{o}} \mathbf{1}^T \mathbf{o}, \end{aligned} \tag{7.44}$$

subject to $o_i > 0$, where C is a factor independent of $\mathbf{o}$ and $\mathbf{1} = [1\ 1 \cdots 1]^T$.

Note that due to the constraints the estimate can no longer be expressed in a closed form as was the case for the linear MMSE estimate (7.7). The MAP estimate, given by minimizing (7.44) must therefore be found by means of an iterative optimization procedure. However, for optimization problems of this type there exist efficient numerical algorithms as will be discussed in more detail in the next chapter.

7.2 Focusing as an Inverse Problem

Recently there has been an interest in improving the reconstruction performance using inverse filtering techniques in ultrasonic imaging, [43–45].

[4] An example where the scattering amplitude must be positive, which is considered in the next chapter, is array imaging of hard steel wire targets immersed in water.

In order to relate such inverse filtering methods to the Bayesian methods discussed above, a brief introduction the inverse filtering is presented below.

Let us regard the image reconstruction problem as an inverse problem where we seek the inverse operator to remove the effects of the observation mechanism. We assume that the true image formation system is given by the linear model

$$\mathbf{y} = \mathbf{P}\mathbf{o} + \mathbf{e}. \tag{7.45}$$

The propagation matrix, $\mathbf{P}$, introduces a distortion or a degradation of the image $\mathbf{o}$ which we want to compensate for. Now a function $F(\cdot)$ is sought, which is the inverse function of $\mathbf{P}$. If $\mathbf{P}$ is square and invertible (all eigenvalues larger than 0) then the function $F(\cdot)$ is simply $\mathbf{P}^{-1}$. If $\mathbf{P}$ is not square a pseudo or generalized inverse can be used,

$$\hat{\mathbf{o}} = \mathbf{P}^{+}\mathbf{y}. \tag{7.46}$$

The pseudo inverse solution can be found by solving the equation system

$$\mathbf{P}^T\mathbf{y} = \mathbf{P}^T\mathbf{P}\mathbf{o}, \tag{7.47}$$

or equivalently by solving the least-squares (LS) optimization problem

$$\begin{aligned} J_{\text{LS}} &= \|\mathbf{y} - \mathbf{P}\mathbf{o}\|^2 \\ &= \mathbf{y}^T\mathbf{y} - 2\mathbf{P}^T\mathbf{y} + \mathbf{o}^T\mathbf{P}^T\mathbf{P}\mathbf{o}. \end{aligned} \tag{7.48}$$

Eq. (7.48) is minimized by

$$\hat{\mathbf{o}}_{\text{LS}} = (\mathbf{P}^T\mathbf{P})^{-1}\mathbf{P}^T\mathbf{y}. \tag{7.49}$$

which is the familiar least squares (LS) estimate.

A requirement for the existence of (7.49) is that the *normal* matrix, $\mathbf{P}^T\mathbf{P}$, is regular. Recall from Chapter 4 that $\mathbf{P}$ is a $KL \times MN$ matrix where MN is the dimension of $\mathbf{o}$ and KL is the dimension of $\mathbf{y}$. If $KL < MN$, then $\mathbf{P}^T\mathbf{P}$ will have rank KL and is therefore not invertible. Even though the normal matrix has full rank, some eigenvalues of $\mathbf{P}^T\mathbf{P}$ will usually be very close to zero. Inverting $\mathbf{P}^T\mathbf{P}$ will therefore result in a variance of the LS estimates that is unacceptable.

Inverse problems of this type in know as *ill-conditioned* or *ill-posed* problems and special precautions must be taken when solving such problems. For a solution to be acceptable, it must satisfy three conditions, known as the *Hadamard conditions*, which are: existence, uniqueness, and the physical condition of stability or continuity with respect to the data [31]. There exist several approaches to ensure stability of the solution and two common methods are discussed in the following two sub-sections.

7.2.1 Singular Value Decomposition Inversion

A common technique to solve ill-posed inverse problems is by means of a performing a SVD of the propagation operator. Recall from Section 7.1.2 that the SVD of the propagation operator, $\mathbf{P}$, is given by

$$\mathbf{P} = \mathbf{U}\mathbf{D}\mathbf{V}^T. \tag{7.50}$$

To avoid strong noise amplification, as a result of small singular values, it is thus unwanted to use a direct inverse of the normal matrix, $\mathbf{P}^T\mathbf{P}$, using all eigenvectors and singular values.

A common technique is then to truncate the singular value decomposition and only use the eigenvectors and corresponding singular values over a certain threshold, $R = i_{\max}$, in the inverse operator. The remaining singular values, for $i > R$, are then set identically to zero. A pseudo-inverse of $\mathbf{P}$ can then be found by first noting that $\mathbf{U}$ and $\mathbf{V}$ are orthonormal matrices so their corresponding inverse is given by $\mathbf{U}^T$ and $\mathbf{V}^T$, and since $\mathbf{D}$ is diagonal the inverse is simply obtained by taking the reciprocal of each diagonal element s_i. The truncated pseudo-inverse is thus given by

$$\begin{aligned}\hat{\mathbf{o}}_{\text{SVD}} &= \mathbf{V}\mathbf{D}^{+}\mathbf{U}^T\mathbf{y} \\ &= \mathbf{V}\begin{bmatrix} s_0^{-1} & & & & & & \\ & s_1^{-1} & & & \mathbf{0} & & \\ & & \ddots & & & & \\ & & & s_R^{-1} & & & \\ & & & & 0 & & \\ & & \mathbf{0} & & & \ddots & \\ & & & & & & 0 \end{bmatrix}\mathbf{U}^T\mathbf{y}.\end{aligned} \tag{7.51}$$

The strong noise amplification, due to a devision with a singular value close to zero, is thereby avoided resulting in a more stable solution [44, 86]. There is however no rationale on how to select the threshold, R, when the singular values are considered too small. In particular, the method does not consider the noise in (7.45) nor any prior information regarding the scattering strengths.

7.2.2 Tikhonov Regularization

A classical regularization method introduced by Tikhonov and Arsenin [87] is to add constraints on the solution to avoid excessive noise amplification.

The estimate is obtained by solving the optimization problem

$$\hat{\mathbf{o}} = \arg\min_{\mathbf{o}} \|\mathbf{y} - \mathbf{P}\mathbf{o}\|^2 + \alpha Q(\mathbf{o}), \tag{7.52}$$

where $Q(\cdot)$ is a discrete functional of the form $Q(\mathbf{o}) = \|\mathbf{R}\mathbf{o}\|^2$. Taking the derivative of (7.52) with respect to $\mathbf{o}$ and equating to zero results in

$$\hat{\mathbf{o}}_{\text{MNLS}} = (\mathbf{P}^T\mathbf{P} + \alpha\mathbf{R}^T\mathbf{R})^{-1}\mathbf{P}^T\mathbf{y}. \tag{7.53}$$

A common choice is to let $\mathbf{R}$ be a differential operator which, then leads to smoother estimates since rapid changes in $\mathbf{o}$ result in a large value of $\|\mathbf{R}\mathbf{o}\|^2$.

A special case of (7.53) is the minimum norm least-squares (MNLS) solution where $\mathbf{R} = \mathbf{I}$. Eq. (7.52) then simplifies to

$$\hat{\mathbf{o}} = \arg\min_{\mathbf{o}} \|\mathbf{y} - \mathbf{P}\mathbf{o}\|^2 + \alpha\|\mathbf{o}\|^2. \tag{7.54}$$

For this choice of $Q(\mathbf{o})$ there will be a penalty for large values of $\mathbf{o}$.

The MNLS solution can be compared to the SVD method, (7.51), by using the singular value decomposition (7.39) in (7.53) with $\mathbf{R} = \alpha\mathbf{I}$,

$$\begin{aligned}
\hat{\mathbf{o}}_{\text{MNLS}} &= (\mathbf{P}^T\mathbf{P} + \alpha\mathbf{I})^{-1}\mathbf{P}^T\mathbf{y} \\
&= \mathbf{V}(\mathbf{D}^2 + \alpha\mathbf{I})^{-1}\mathbf{D}\mathbf{U}^T\mathbf{y} \\
&= \mathbf{V}\begin{bmatrix} \frac{s_0}{s_0^2+\alpha} & & & \\ & \frac{s_1}{s_1^2+\alpha} & & \mathbf{0} \\ & & \ddots & \\ \mathbf{0} & & & \frac{s_{MN-1}}{s_{MN-1}^2+\alpha} \end{bmatrix}\mathbf{U}^T\mathbf{y}.
\end{aligned} \tag{7.55}$$

The factor α protects from a division by a factor close to zero thereby stabilizing the solution.

As already mentioned in the introduction to this chapter, Tikhonov regularization can be seen as a special case of the linear MMSE, or linear MAP, solution. To see this compare the MNLS solution (7.55) and the MAP solution (7.40). If we choose α, in (7.55), as $\sigma_{\text{e}}^2/\sigma_{\text{o}}^2$ then (7.55) and (7.40) becomes identical. Thus, for the particular choice $\alpha = \sigma_{\text{e}}^2/\sigma_{\text{o}}^2$ the MNLS solution (7.55) will also be the minimum mean-squared error solution.

7.2.3 The Maximum Likelihood Estimator

As a final remark we also consider the maximum likelihood (ML) estimator due to its popularity in the statistical literature. A reason for the popularity is that the ML estimate is unbiased and have the asymptotically smallest covariance matrix of all the unbiased estimators. The ML estimate is defined as the maximum value of the likelihood function, given by

$$\hat{\mathbf{o}} = \arg\max_{\mathbf{o}} \; p(\mathbf{y}|\mathbf{o}, I). \tag{7.56}$$

Here I denotes our background information, i.e., the model (7.45).

For a Gaussian measurement noise the likelihood function is given by

$$p(\mathbf{y}|\mathbf{o}, I) = \frac{1}{(2\pi)^{\frac{MN}{2}}|\mathbf{C}_\mathrm{e}|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{y}-\mathbf{P}\mathbf{o})^T {\mathbf{C}_\mathrm{e}}^{-1}(\mathbf{y}-\mathbf{P}\mathbf{o})\right). \tag{7.57}$$

Eq. (7.57) is then maximized by first taking the logarithm,

$$\log\left(p(\mathbf{y}|\mathbf{o}, \mathbf{I})\right) = -\log\left((2\pi)^{\frac{MN}{2}}|\mathbf{C}_\mathrm{e}|^{1/2}\right) - \frac{1}{2}(\mathbf{y}-\mathbf{P}\mathbf{o})^T {\mathbf{C}_\mathrm{e}}^{-1}(\mathbf{y}-\mathbf{P}\mathbf{o}), \tag{7.58}$$

and then equating the derivative with respect to $\mathbf{o}$, given by

$$\frac{\partial\left(-\log p(\mathbf{y}|\mathbf{o})\right)}{\partial \mathbf{o}} = -\mathbf{P}^T {\mathbf{C}_\mathrm{e}}^{-1}\mathbf{y} + \mathbf{P}^T {\mathbf{C}_\mathrm{e}}^{-1}\mathbf{P}\mathbf{o} \tag{7.59}$$

to zero. This results in the solution

$$\hat{\mathbf{o}}_{\mathrm{ML}} = (\mathbf{P}^T {\mathbf{C}_\mathrm{e}}^{-1}\mathbf{P})^{-1}\mathbf{P}^T {\mathbf{C}_\mathrm{e}}^{-1}\mathbf{y}. \tag{7.60}$$

If the noise is white, $\mathbf{C}_\mathrm{e} = \sigma_\mathrm{e}^2\mathbf{I}$, then (7.60) reduces to

$$\hat{\mathbf{o}} = (\mathbf{P}^T\mathbf{P})^{-1}\mathbf{P}^T\mathbf{y}, \tag{7.61}$$

which is identical to the LS estimate (7.49). Thus, the ML estimate has the same stability problems as the LS estimate: the low bias of the ML solution is traded against an (unacceptable) high variance of the estimates.

The problem with the ML estimator, and the estimators discussed above this section, is that they do not utilize the essential prior information. Therefore, more or less *ad hoc* methods were needed to regularize the solution.

7.3 Concluding Remarks

In this chapter, two Bayesian reconstruction methods were introduced, the linear MMSE and the non-linear MAP estimators. These estimators were considered for the purpose of improving the imaging performance of classical beamforming methods. Since the two Bayesian methods are based on a more realistic model of the imaging system than traditional DAS beamforming, as well as prior information of the scattering strengths is taken into account, the performance is expected to improve compared to traditional beamforming. This will be verified in Chapter 8, where the reconstruction performance of the two Bayesian methods will be further studied.

7.A Some Matrix Derivations

7.A.1 Equivalence of MAP and Linear MMSE Estimators

To prove the equivalence of MAP and linear MMSE estimators, when $\mathbf{o}$ and $\mathbf{e}$ are Gaussian, we use the matrix inversion lemma [88]

$$(\mathbf{A}+\mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{C}^{-1}+\mathbf{DA}^{-1}\mathbf{B})^{-1}\mathbf{DA}^{-1}. \tag{7.62}$$

First, let $\mathbf{A} = \mathbf{C}_\mathrm{e}$, $\mathbf{B} = \mathbf{P}$, $\mathbf{C} = \mathbf{C}_\mathrm{o}$, and $\mathbf{D} = \mathbf{P}^T$, then using (7.62), the factor $(\mathbf{C}_\mathrm{e}+\mathbf{PC}_\mathrm{o}\mathbf{P}^T)^{-1}$ in the MAP estimate (7.3) is

$$(\mathbf{C}_\mathrm{e}+\mathbf{PC}_\mathrm{o}\mathbf{P}^T)^{-1} = \mathbf{C}_\mathrm{e}^{-\mathbf{1}} - \mathbf{C}_\mathrm{e}^{-\mathbf{1}}\mathbf{P}(\mathbf{C}_\mathrm{o}^{-\mathbf{1}}+\mathbf{P}^T\mathbf{C}_\mathrm{e}^{-\mathbf{1}}\mathbf{P})^{-1}\mathbf{P}^T\mathbf{C}_\mathrm{e}^{-\mathbf{1}}. \tag{7.63}$$

Now let $\mathbf{\Gamma} = (\mathbf{C}_\mathrm{o}^{-\mathbf{1}}+\mathbf{P}^T\mathbf{C}_\mathrm{e}^{-\mathbf{1}}\mathbf{P})^{-1}$ then the MAP estimator can be expressed as

$$\mathbf{K}_\mathrm{MAP} = \mathbf{C}_\mathrm{o}\mathbf{P}^T(\mathbf{C}_\mathrm{e}^{-\mathbf{1}} - \mathbf{C}_\mathrm{e}^{-\mathbf{1}}\mathbf{P\Gamma P}^T\mathbf{C}_\mathrm{e}^{-\mathbf{1}}), \tag{7.64}$$

and the linear MMSE estimator (7.7), as

$$\mathbf{K}_\mathrm{LMMSE} = \mathbf{\Gamma P}^T\mathbf{C}_\mathrm{e}^{-\mathbf{1}}. \tag{7.65}$$

Taking the difference between (7.64) and (7.65),

$$\begin{aligned}
\mathbf{K}_\mathrm{MAP} - \mathbf{K}_\mathrm{LMMSE} &= \mathbf{C}_\mathrm{o}\mathbf{P}^T(\mathbf{C}_\mathrm{e}^{-\mathbf{1}} - \mathbf{C}_\mathrm{e}^{-\mathbf{1}}\mathbf{P\Gamma P}^T\mathbf{C}_\mathrm{e}^{-\mathbf{1}}) - \mathbf{\Gamma P}^T\mathbf{C}_\mathrm{e}^{-\mathbf{1}} \\
&= (\mathbf{C}_\mathrm{o} - \mathbf{C}_\mathrm{o}\mathbf{P}^T\mathbf{C}_\mathrm{e}^{-\mathbf{1}}\mathbf{P\Gamma} - \mathbf{\Gamma})\mathbf{P}^T\mathbf{C}_\mathrm{e}^{-\mathbf{1}} \\
&[\mathbf{C}_\mathrm{o} = \mathbf{C}_\mathrm{o}\mathbf{\Gamma}^{-1}\mathbf{\Gamma}] \\
&= (\mathbf{C}_\mathrm{o}\mathbf{\Gamma}^{-1} - \mathbf{C}_\mathrm{o}\mathbf{P}^T\mathbf{C}_\mathrm{e}^{-\mathbf{1}}\mathbf{P} - \mathbf{I})\mathbf{\Gamma P}^T\mathbf{C}_\mathrm{e}^{-\mathbf{1}} \\
&= (\mathbf{C}_\mathrm{o}(\mathbf{C}_\mathrm{o}^{-\mathbf{1}}+\mathbf{P}^T\mathbf{C}_\mathrm{e}^{-\mathbf{1}}\mathbf{P}) - \mathbf{C}_\mathrm{o}\mathbf{P}^T\mathbf{C}_\mathrm{e}^{-\mathbf{1}}\mathbf{P} - \mathbf{I})\mathbf{\Gamma P}^T\mathbf{C}_\mathrm{e}^{-\mathbf{1}} \\
&= (\mathbf{I} + \mathbf{C}_\mathrm{o}\mathbf{P}^T\mathbf{C}_\mathrm{e}^{-\mathbf{1}}\mathbf{P} - \mathbf{C}_\mathrm{o}\mathbf{P}^T\mathbf{C}_\mathrm{e}^{-\mathbf{1}}\mathbf{P} - \mathbf{I})\mathbf{\Gamma P}^T\mathbf{C}_\mathrm{e}^{-\mathbf{1}} \\
&= \mathbf{0},
\end{aligned} \tag{7.66}$$

and the equivalence is proved.

7.A.2 Performance of the Linear MMSE and MAP Estimators

The performance of the estimator is measured with the error $\boldsymbol{\epsilon} = \hat{\mathbf{o}} - \mathbf{o}$. Since $\mathbf{o}$ is Gaussian and $\hat{\mathbf{o}}$ is a linear transformation of a Gaussian variable

$\boldsymbol{\epsilon}$ must also be Gaussian. Let $\mathbf{K}$ denote the MAP estimator (7.9). The error covariance matrix is then obtained as

$$\begin{aligned}
\mathbf{C}_\epsilon &= \mathrm{E}\{\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T\} = \mathrm{E}\{(\mathbf{K}\mathbf{y}-\mathbf{o})(\mathbf{K}\mathbf{y}-\mathbf{o})^T\} \\
&= \mathrm{E}\{\mathbf{o}\mathbf{o}^T\} - \mathrm{E}\{\mathbf{K}(\mathbf{P}\mathbf{o}+\mathbf{e})\mathbf{o}^T\} - \mathrm{E}\{\mathbf{o}(\mathbf{e}^T+\mathbf{o}^T\mathbf{P}^T)\mathbf{K}^T\} + \\
&\quad \mathrm{E}\{\mathbf{K}(\mathbf{P}\mathbf{o}+\mathbf{e})(\mathbf{e}^T+\mathbf{o}^T\mathbf{P}^T)\mathbf{K}^T\} \\
&= \mathbf{C}_\mathrm{o} - \mathbf{K}\mathbf{P}\mathbf{C}_\mathrm{o} - \mathbf{C}_\mathrm{o}\mathbf{P}^T\mathbf{K}^T + \mathbf{K}(\mathbf{P}\mathbf{C}_\mathrm{o}\mathbf{P}^T + \mathbf{C}_\mathrm{e})\mathbf{K}^T \\
&= \mathbf{C}_\mathrm{o} - \mathbf{C}_\mathrm{o}\mathbf{P}^T(\mathbf{P}\mathbf{C}_\mathrm{o}\mathbf{P}^T + \mathbf{C}_\mathrm{e})^{-1}\mathbf{P}\mathbf{C}_\mathrm{o} \\
&= (\mathbf{C}_\mathrm{o}^{-\mathbf{1}} + \mathbf{P}^T\mathbf{C}_\mathrm{e}^{-\mathbf{1}}\mathbf{P})^{-1}
\end{aligned} \tag{7.67}$$

where we have used the properties that $\mathbf{o}$ and $\mathbf{e}$ are uncorrelated, and that the inverse of a symmetric matrix also is symmetric. The last identity in (7.67) follows directly from the matrix inversion lemma (7.62).

CHAPTER 8

Applications

IN the previous chapters, ultrasonic image reconstruction was formulated as a linearized inverse-scattering problem, to which both linear and non-linear approaches to were suggested. The simulations in Chapters 5 and 6 showed that traditional beamforming methods produced images with inadequate quality in cases involving spatially sparse sampled data and where the specific diffraction effects due to finite-sized array elements were apparent.

The aim of this chapter is to report results of experiments and simulations for both synthetic and real array apertures, comparing the Bayesian estimators presented in the previous chapter with conventional beamforming methods. Aspects treated in this chapter includes the size effects of finite-sized transducers, sparsely sampled, or under-sampled arrays, as well as parallel array imaging.

The experiments have been conducted using a phased array system and a standard immersion tank. The broadband data were acquired using a 1D array, employing both electronic and mechanical scanning. To make the reconstruction results easy to interpret, wire targets and side-drilled holes were used. The simulations in this chapter were designed to emulate the physical array system. This was done in order to facilitate the comparison with experiments.

For practical reasons, all investigations, both experimental and simulated, have been conducted on two-dimensional data only. In this way the memory and CPU requirements are reduced. Note however that there is

nothing that theoretically limits the Bayesian methods discussed here to be applied to full 3D data.

This chapter is organized as follows: In section 8.1, the array system used for the measurements is described. This section also contains a short description of how the SAI and the parallel imaging experiments have been conducted. Section 8.2 describes the SAI experiments, for investigation of the lateral resolution, when using finite-sized transducers. In Section 8.3, the parallel array experiments are discussed. Topics treated are grating lobe suppression and estimation performance, both with respect to different focusing methods and to signal-to-noise ratios. This section includes a comparison of the DAS, matched filter, and the linear MMSE methods as well as the non-linear MAP estimator. Finally, in Section 8.5, some conclusions from the experiments are given.

8.1 Experimental Equipment

The experimental data for this thesis were acquired using an ALLIN array system manufactured by NDT Systems. This system is equipped with an electronic beamformer that allows for focusing in both the transmit and the receive operation. The driving unit can use a maximum of 32 elements simultaneously, but only with 16 unique focusing delays, and beam steering can therefore only be performed using 16 elements. The system has an analog multiplexer that controls which array inputs are connected to the 8-bit analog-to-digital converter (ADC).[1]

Since the data, due to the analog multiplexer, can be acquired by electronically switching between different array elements, a very fast scanning of the beam is possible. The system can also perform mechanical scanning by moving the transducer using a step-motor controlled robot.

In the experiments, the ALLIN system was connected to a 64 element concave 1D array for immersion testing, manufactured by Imasonic. This array has a dimension of 64×33.5 mm, where each individual element is 0.9×33.5 mm. Each array element is geometrically focused in the y-direction at the depth $z = 190$ mm. The array pitch, d, is 1 mm and the elements are thus separated by 0.1 mm. The array has a center frequency at approximately 3 MHz (see Figure 3.4). This corresponds to $\lambda = 0.5$ mm in water. The array pitch is then approximately 2λ and the array is under-sampled according

[1] The 8-bit ADC limits the dynamic range of the system to 48 dB.

to the array sampling theorem, cf. Chapter 1. The Imasonic array and the ALLIN driving unit are shown in Figure 8.1. Figure 8.1(a) shows a schematic view of the Imasonic array and a photograph is shown in Figure 8.1(b). A photograph of the ALLIN system cabinet is shown in Figure 8.1(c) where also a part of the robot can be seen in the background.

This system was used in experiments for both synthetic aperture imaging (SAI) and parallel array beamforming. The SAI experiments were performed by means of mechanical scanning with a different number of active elements thus mimicking single scanned transducers with different sizes but with almost identical electrical characteristics. This measurement method facilitates the comparison of SAI experiments with different transducer sizes since the electrical impulse responses are the same regardless of the transducer size.

As discussed in Chapter 6, a parallel beamformer uses multiple focusing laws to beamform many points simultaneously. To acquire data for parallel imaging, the received signals from all elements must be available. Since the ALLIN system only has one ADC, parallel beamforming experiments cannot be performed from data acquired using a single emission. A parallel data acquisition can, however, be emulated by using the analog multiplexer by sequentially acquire data from each element with the array at the same position. All parallel array imaging experiments have been conducted using this technique.

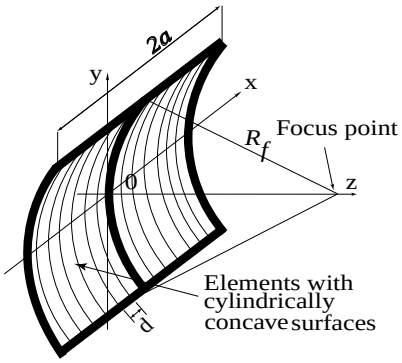
8.2 Ultrasonic Synthetic Aperture Imaging

In this section the specific effects due to transducer element size has been studied. SAI experiments has been chosen since the transducer size does not impose any restrictions on the array pitch, which is the case for physical arrays.

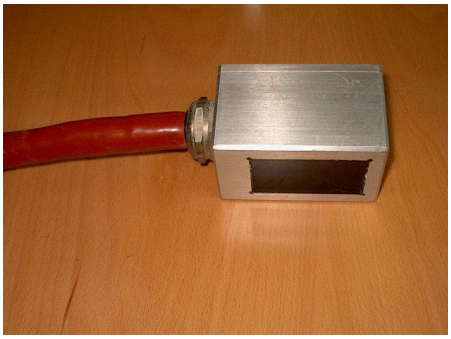
Below, the propagation matrix for monostatic SAI is presented which is then used in the experiments.

8.2.1 The Monostatic Propagation Matrix

The propagation matrix described in Chapter 4 was derived for a general measurement setup, allowing for transmission and reception using arbitrary combinations of transducers. In the SAI experiments presented in this thesis a single scanned axisymmetric transducer was used. In this case the acoustic



(a) Geometry of the 64 element concave 1D array.



(b) The 64 element concave Imasonic array.

(c) The ALLIN array system unit.

Figure 8.1: The ALLIN array system and the 64 element Imasonic array. The array elements are 0.9 mm wide in the x-direction and 33.5 mm in the y-direction, the array pitch d is 1 mm, and the array is geometrically focused at 190 mm. The driving unit can use 32 elements simultaneously but only with 16 unique focusing delays and the data are acquired using an 8-bit analog-to-digital converter.

transmit and reception processes will be identical, that is, the transmit and receive SIRs will be the same. Recall that the A-scan signal, $\mathbf{y}_n$, received by a transducer at x_n, is given by (4.21), that is,

$$\mathbf{y}_n = \sum_{\tilde{n}=0}^{N-1} \mathbf{P}_{d(\tilde{n},n)} \mathbf{o}_{\tilde{n}} + \mathbf{e}_n. \tag{8.1}$$

Due to the symmetry of the transducer, the $\mathbf{P}_{d(\tilde{n},n)}$ matrices for vertical lines at the same horizontal distance from the transducer will be identical, that is,

$$\mathbf{P}_{d(n-i,n)} = \mathbf{P}_{d(n+i,n)} \tag{8.2}$$

for an integer i in the ROI.[2] Moreover, the matrix $\mathbf{P}_{d(\tilde{n},n)}$ is also invariant with respect to a parallel shift of the position of the transducer, in the sense that

$$\mathbf{P}_{d(\tilde{n}+i,n+i)} = \mathbf{P}_{d(\tilde{n},n)}. \tag{8.3}$$

Furthermore, the number of elements in the propagation matrix can be decreased by noting that the contributions for observation points at a large distances from the transducer will be small due to the beam directivity of the transducer (cf. the discussion in Section 3.3). Therefore, the summation in (8.1) can be truncated without introducing any significant errors. That is, (8.1) can be simplified to

$$\mathbf{y}_n \approx \sum_{\tilde{n}=n-\tilde{L}/2}^{n+\tilde{L}/2} \mathbf{P}_{d(\tilde{n},n)} \mathbf{o}_{\tilde{n}} + \mathbf{e}_n, \tag{8.4}$$

where $\tilde{L}$ is the number transducer positions that significantly contribute to the response from targets located along the vertical line at $x_{\tilde{n}}$.

Using (8.4), and the symmetries (8.2) and (8.3), in (4.23), the total propagation matrix for monostatic SAI imaging can finally be expressed,

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_{d(i-\tilde{L}/2,i)} & \cdots & \mathbf{P}_{d(i+\tilde{L}/2,i)} & & \\ & \mathbf{P}_{d(i-\tilde{L}/2,i)} & \cdots & \mathbf{P}_{d(i+\tilde{L}/2,i)} & \mathbf{0} \\ & & \ddots & \ddots & \\ \mathbf{0} & & & & \\ & & & \mathbf{P}_{d(i-\tilde{L}/2,i)} & \cdots \quad \mathbf{P}_{d(i+\tilde{L}/2,i)} \end{bmatrix}, \tag{8.5}$$

for any i such that $i - \tilde{L}/2$ and $i + \tilde{L}/2$ is in the ROI.

[2] It is assumed that the ROI is spatially sampled at regular intervals.

8.2.2 Lateral Resolution and Finite Sized Transducers

A deficiency of the SAFT method comes from treating the scanning transducer as a point source. This results in a mismatch of the SAFT filter given in (5.7) if the transducer has a finite size, as was demonstrated in Chapter 5. To study only the influence of the transducer size on the lateral resolution, a simplified propagation matrix, $\mathbf{S}_{\text{SIR}}$, can be used which only contains the double-path SIRs and not the electrical impulse responses. This essentially assumes the propagation model

$$\mathbf{y} = \mathbf{S}_{\text{SIR}}\mathbf{o} + \mathbf{e} \tag{8.6}$$

where $\mathbf{S}_{\text{SIR}}$ has the same structure as the propagation matrix, $\mathbf{P}$ in (8.5).

The filter matrix in the linear MMSE estimator, for the simplified model (8.6), is then given by

$$\mathbf{K}_{\text{MMSE}} = \mathbf{C}_{\text{o}}\mathbf{S}_{\text{SIR}}^T(\mathbf{S}_{\text{SIR}}\mathbf{C}_{\text{o}}\mathbf{S}_{\text{SIR}}^T + \mathbf{C}_{\text{e}})^{-1}. \tag{8.7}$$

Since the model (8.6) does not include the electrical impulse response, the filter (8.7) should now only compensate for the specific diffraction effects due to the finite size of the active area of the transducer.

Below, two experiments are presented, one conducted using wire targets and one using an immersed copper block with side-drilled holes. The experiments were designed such that transducer size effects was apparent.

Experiment 1: Wire Targets

In this experiment, the lateral resolution for the simplified linear MMSE estimator (8.7) was evaluated on measured data, and compared to SAFT processing using the same synthetic aperture.

The data were acquired using the Imasonic array imaging a wire target[3] at $z = 50$ mm. The array was mechanically scanned along a line, and the data were acquired with a pitch $d = 1$ mm and a temporal sampling period, T_{s}, of 0.04 μs (=25 MHz). The synthetic aperture consisted of 31 A-scans for both the SAFT and the MMSE algorithm.

Two transducer sizes were considered: 1 mm and 16 mm, respectively. The latter was obtained by firing 16 elements of the Imasonic array simultaneously. The SIRs associated with the concave array, $\mathbf{S}_{\text{SIR}}$, used in (8.7),

[3] A steel pin with a diameter of 0.3 mm was used.

were computed using the DREAM method described in Chapter 3. Furthermore, in all experiments and simulations presented in this thesis, the following are assumed:

Assumption 8.1 The matrices $\mathbf{C}_{\mathrm{o}}$ and $\mathbf{C}_{\mathrm{e}}$ are assumed to have the form $\mathbf{C}_{\mathrm{o}} = \sigma_{\mathrm{o}}^2\mathbf{I}$ and $\mathbf{C}_{\mathrm{e}} = \sigma_{\mathrm{e}}^2\mathbf{I}$, respectively.

Assumption 8.1 states that the elements in both $\mathbf{o}$ and $\mathbf{e}$ are uncorrelated, which is reasonable if we do not have any information that indicates any dependence between the elements in $\mathbf{o}$ and $\mathbf{e}$. Using this assumption (8.7) is simplified to

$$\mathbf{K}_{\mathrm{MMSE}} = \mathbf{S}_{\mathrm{SIR}}^T(\mathbf{S}_{\mathrm{SIR}}\mathbf{S}_{\mathrm{SIR}}^T + \sigma_{\mathrm{e}}^2/\sigma_{\mathrm{o}}^2\mathbf{I})^{-1}. \tag{8.8}$$

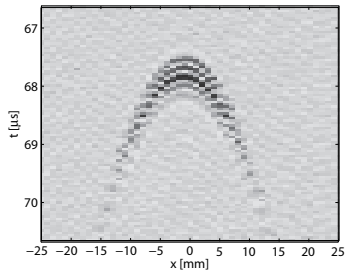
In this experiment, the simplified model (8.6) is used (that do not take the electrical impulse response into account) which introduces model errors. These errors are difficult to quantify and it is, consequently, also be difficult to determine the noise variance σ_{e}^2 since the noise is a combination the model and measurement errors. Therefore, the ratio $\mu = \sigma_{\mathrm{e}}^2/\sigma_{\mathrm{o}}^2$ was simply used as a tuning parameter, and it was adjusted manually to obtain the best visible performance.

The unprocessed data and the reconstruction results are shown in Figure 8.2. Figures 8.2(a) and (b) show the unprocessed B-scan data, corresponding to the 1 mm and the 16 mm transducer, respectively. The B-scan for the 1 mm transducer has the typical hyperbolic shape associated with small transducers that emits nearly spherical waves. The B-scan from 16 mm transducer, shown in Figure 8.2(b), shows however strong effects of the beam directivety.

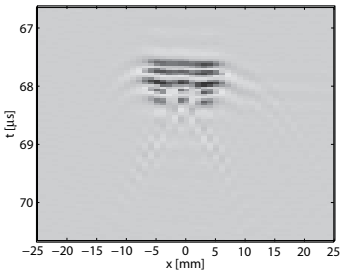
The reconstruction results for the SAFT processed 1 mm data, shown in Figure 8.2(c), shows a significant improvement in lateral resolution compared to the unprocessed data. However, the SAFT reconstructed image for the 16 mm transducer, shown in Figure 8.2(d) shows no improvement in lateral resolution compared to the unprocessed data. That is, SAFT performance deteriorates when the transducer size inceases.

The MMSE reconstructions, shown in Figures 8.2(e) and (f), shows, on the other hand, high lateral resolution for both transducer sizes. Thus, the MMSE estimator has sucessfully compensated for the diffraction effects of the transducer.

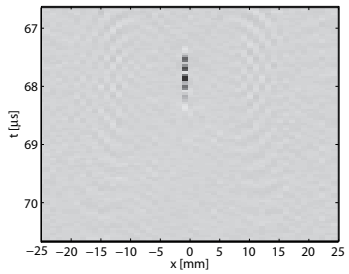
One can also finally note that the temporal, or range resolution, is roughly the same in the unprocessed B-scan data and the processed images



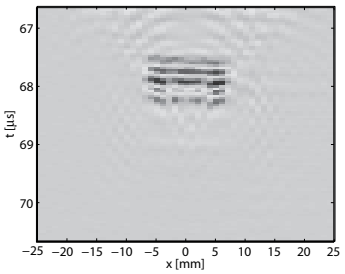
(a) Unprocessed B-scan data: 1mm transducer.



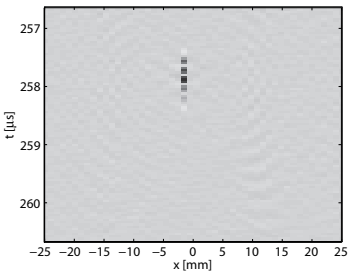
(b) Unprocessed B-scan data: 16 mm transducer.



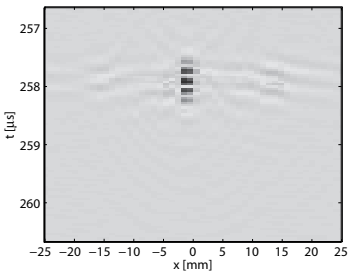
(c) SAFT processed data: 1mm transducer.



(d) SAFT processed data : 16 mm transducer.



(e) MMSE processed data: 1 mm transducer.



(f) MMSE processed data: 16 mm transducer.

Figure 8.2: Comparison between the SAFT and the simplified MMSE estimator, (8.7), on measured data from a wire target (steel pin with a diameter of 0.3 mm) at $z = 50$ mm using a phased array.

for both methods. This was expected since the conventional SAFT algorithm do compensate for the temporal waveform and, as mention above, we intentionally did not include the electrical impulse response in the model (8.7) that the MMSE estimator is based on.

Experiment 2: Side Drilled Holes in an Immersed Copper Block

The objective with this experiment was to evaluate the performance of the MMSE method in comparision to classical focusing methods in a typical NDT application. In NDT applications it is often important to be able to resolve closely spaced objects in order to, for example, identify critical defects. For that purpose, measurements were performed using a copper block with four pairs of side-drilled holes (SDHs). The copper block was immersed in water and the distance between the array and the block was adjusted so that the SDHs in the copper block were located at the geometric focus of the concave 1D array.[4] The spacing between the SDHs in the block ranged from 1 to 7 mm, see Figure 8.3. Below, the SDH spacing is denoted

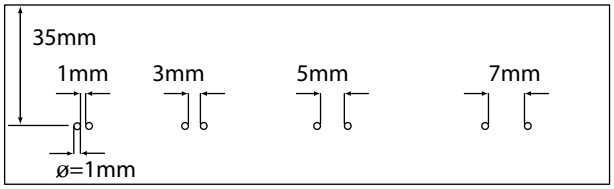


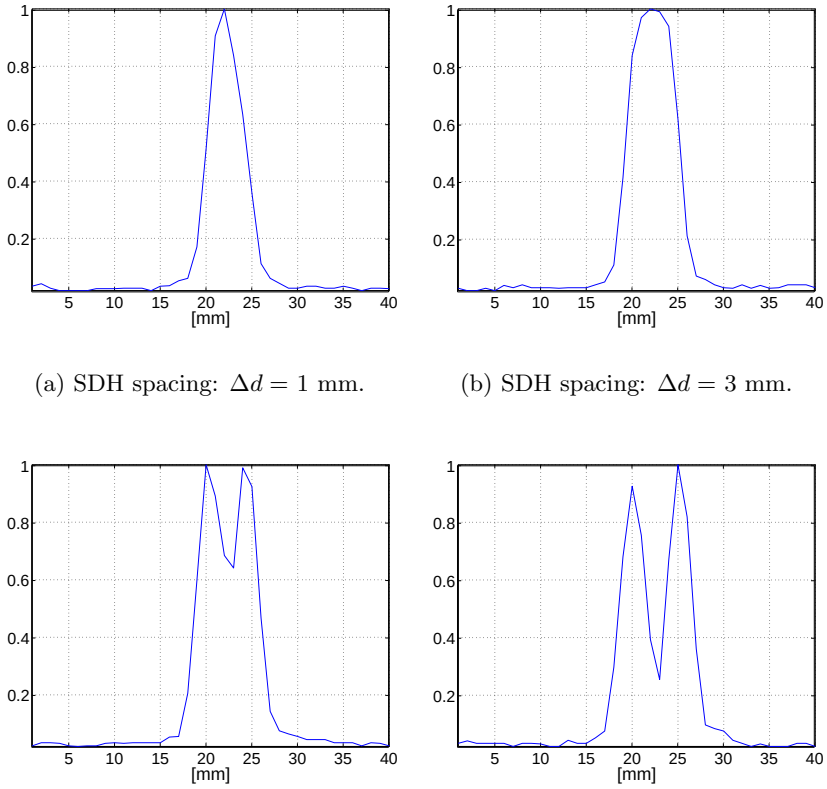
Figure 8.3: Copper test block with four pairs of side drilled holes.

Δd (which should not be confused with the array pitch d).

Three approaches for imaging were examined in this experiment. First, conventional phased array (PA) beamforming, then SAFT, and finally the MMSE method.

The PA method was used as a bench-mark, and the array was focused at the depth of the SDHs with 32 elements using the hardware beamformer in the ALLIN system. Results from the PA measurements, in the form of profile plots are shown in Figure 8.4. As can be seen in Figures 8.4(a) and (b), the PA method did not resolve the two SDHs with $\Delta d = 1$ and 3 mm,

[4]The Imasonic array was manufactured specificially for imaging immersed specimens.



(a) SDH spacing: $\Delta d = 1$ mm.

(b) SDH spacing: $\Delta d = 3$ mm.

(c) SDH spacing: $\Delta d = 5$ mm.

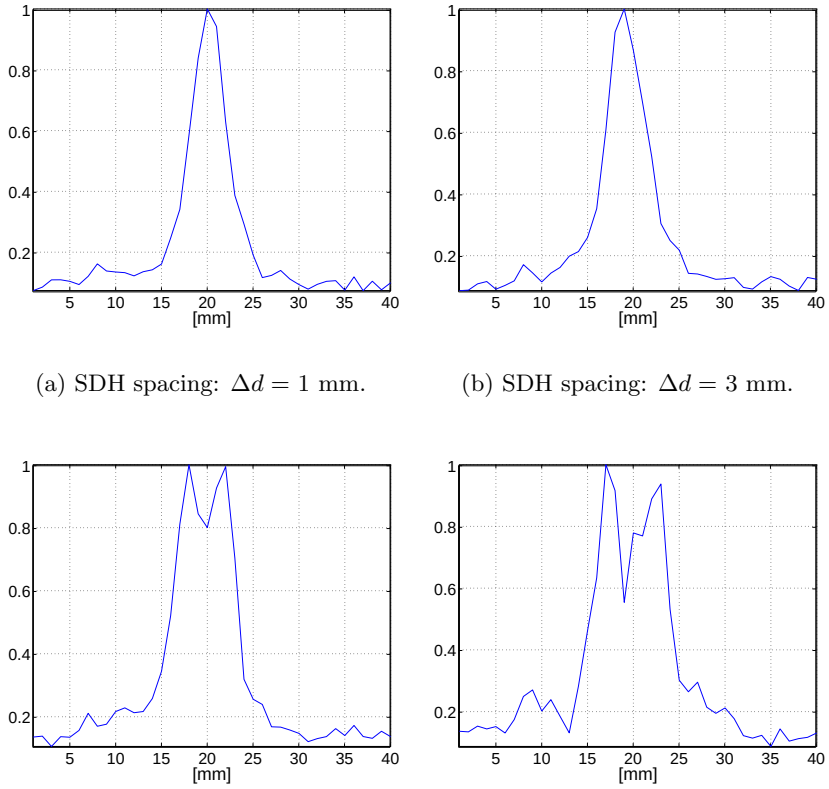
(d) SDH spacing: $\Delta d = 7$ mm.

Figure 8.4: Profile plots from side-drilled hole data beamformed using the Imasonic phased array focused at 35 mm in a copper block using 32 elements.

but when the SDH spacing was increased to 5 and 7 mm the SDHs could be distinguished, see Figure 8.4(c) and (d).

The SAI data were acquired using a single element of the array, and the width of the transducer were therefore approximately 1 mm. The pitch and sampling frequency were the same as for the PA data above, that is, $d = 1$ mm and $T_s = 0.04\mu$s, and the data were acquired by means of mechanical scanning in the same manner as in Experiment 1.

The SAI results for the SAFT method are shown in Figure 8.5 and the



(a) SDH spacing: $\Delta d = 1$ mm.

(b) SDH spacing: $\Delta d = 3$ mm.

(c) SDH spacing: $\Delta d = 5$ mm.

(d) SDH spacing: $\Delta d = 7$ mm.

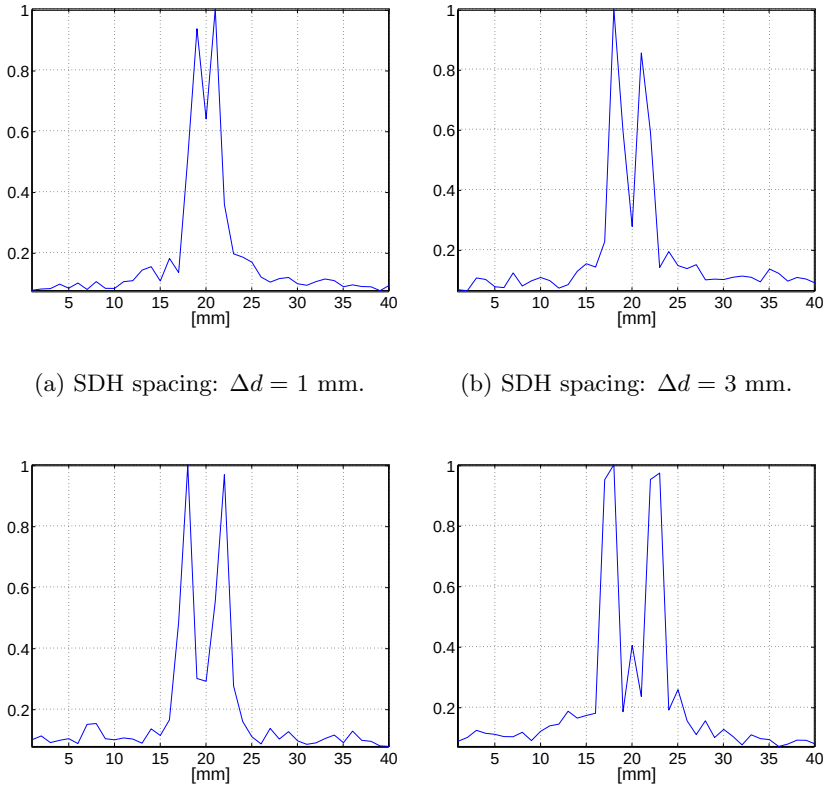
Figure 8.5: Profile plots from side-drilled hole data measured with a 1 mm transducer and processed using the SAFT algorithm.

results for MMSE algorithm are shown in and 8.6.[5]

If we compare the DAS based PA and SAFT results, cf. Figures 8.4 and 8.5, their performance are very similar.[6] That is, both methods did only resolve the SDHs with a spacing of 5 and 7 mm. This is not unexpected since both methods are based on DAS processing and the aperture was almost the same in both experiments. However, from Figure 8.6, where the results from

[5]The SIRs corresponding to the observation points in the immersed solid, needed for the MMSE method, were computed using the method described in Appendix 3.B.

[6]The noise level is higher in the SAFT processed plots since only one element is used whereas the PA uses 32 elements and therefore transmits more energy.



(a) SDH spacing: $\Delta d = 1$ mm.

(b) SDH spacing: $\Delta d = 3$ mm.

(c) SDH spacing: $\Delta d = 5$ mm.

(d) SDH spacing: $\Delta d = 7$ mm.

Figure 8.6: Profile plots from side-drilled hole data measured with a 1 mm transducer and processed using the simplified MMSE estimator, (8.7).

the MMSE estimator are presented, it is clear that all four pairs of SDHs were resolved.

It is evident from the Experiments 1 and 2 above, that the MMSE method successfully can compensate for the size effects of a finite-sized transducer resulting in a significantly improved lateral resolution compared to DAS based processing. Thus, the MMSE method enables more flexible measurement setups since transducers with larger active areas can be used without loosing too much resoltion.

8.3 Parallel Array Imaging

Recall from Chapter 6 that in parallel beamforming several receive focusing filters are used simultaneously to focus the array at many points on data from a single transmission [79, 80, 89]. The benefit of this approach is an increased frame rate compared to, for example, sweeped beam imaging where many transmissions are needed. Typically a transmit beam that is wider then normal is used and a few lines of the total ROI is then reconstructed using parallel receive beamforming [79]. The complete ROI could, in principle, be reconstructed from one transmitted pulse provided that the transmit beam is wide enough to illuminate the ROI. There are, however, practical considerations that make very wide beam illumination difficult. Firstly, it may be difficult to generate a sufficiently wide beam, and secondly ghost responses due to the grating lobes are stronger for DAS beamformed images when a wide beam is used compared to using a sweeped narrow beam [51].

As we will show below, model based reconstruction methods can however alleviate some of these problems since they can take both the transmission and reception processes into account and compensate for any beampattern. The reconstruction methods discussed in this thesis, that are based on the impulse response method, can in principle compensate for grating lobes or any other beam pattern, provided that the corresponding spatial impulse responses can be computed. Moreover the reconstruction methods discussed in Chapter 7, that are based on the linear model (4.23), reconstructs the full ROI from a single set of signals, received by the respective array elements, and are therefore also inherently parallel in nature.

Below we will study the linear MMSE estimator and the non-linear MAP estimator with exponential prior in a situation where the array is undersampled in order to center the attention on the lobe aliasing problems that may be apparent in such situations.

8.3.1 The Propagation Matrix for Parallel Receive Imaging

As discussed in Chapters 1 and 6, in phased array imaging the beam can be focused or steered by applying suitable delays of the signals driving the array elements. These delays will then also delay the SIRs for each array element. The receive and transmit SIRs will therefore be different in contrast to monostatic SAI where the transmit and receive SIRs were identical.

The system impulse response $h^{\text{sys}}(\mathbf{r}, t)$ for the nth receive element can

be computed as a superposition of all delayed transmit SIRs convolved with the SIR for the nth receive element and the electrical impulse response (cf. (4.13)),

$$h^{\text{sys}}(\mathbf{r},t) = \underbrace{\left(\sum_{p=0}^{P-1} h_p^{\text{f}}(\mathbf{r},t) * \delta(t-\tau_p)\right)}_{\text{total forward SIR}} * \underbrace{h_n^{\text{b}}(\mathbf{r},t)}_{n\text{th backward SIR}} * h_{\text{e}}(t), \tag{8.9}$$

where P in the number of transmit elements and τ_p is the focusing delay corresponding to the pth transmit element. The received discrete waveform, from a target at an observation $(x_m, z_{\tilde{n}})$, can then be expressed as

$$\mathbf{y}_n = \mathbf{h}_m^{(\tilde{n},n)}(\mathbf{O})_{m,\tilde{n}} \tag{8.10}$$

where the vector $\mathbf{h}_m^{(\tilde{n},n)}$ is the discrete system impulse response for the nth receive element.

Similar to (4.22) can the noise corrupted received discrete waveform be expressed as

$$\begin{aligned}
\mathbf{y}_n &= \sum_{\tilde{n}=0}^{N-1} \mathbf{P}_{d(\tilde{n},n)}\mathbf{o}_{\tilde{n}} + \mathbf{e}_n \\
&= \begin{bmatrix} \mathbf{h}_0^{(0,n)} & \mathbf{h}_1^{(0,n)} & \cdots & \mathbf{h}_{M-1}^{(0,n)} & \mathbf{h}_0^{(1,n)} & \cdots & \mathbf{h}_{M-1}^{(N-1,n)} \end{bmatrix} \mathbf{o} + \mathbf{e}_n \\
&= \mathbf{P}_n\mathbf{o} + \mathbf{e}_n,
\end{aligned} \tag{8.11}$$

and, by appending the received signals from all L receive elements, the total array imaging system can finally be modeled as

$$\begin{aligned}
\mathbf{y} = &\begin{bmatrix} \mathbf{P}_0 \\ \mathbf{P}_1 \\ \\ \vdots \\ \mathbf{P}_{L-1} \end{bmatrix} \begin{bmatrix} \mathbf{o}_0 \\ \mathbf{o}_1 \\ \\ \vdots \\ \\ \mathbf{o}_{N-1} \end{bmatrix} + \begin{bmatrix} \mathbf{e}_0 \\ \mathbf{e}_1 \\ \\ \vdots \\ \mathbf{e}_{L-1} \end{bmatrix} \\
= &\mathbf{P}\mathbf{o} + \mathbf{e}.
\end{aligned} \tag{8.12}$$

The $KL \times MN$ propagation matrix, $\mathbf{P}$, in (8.12) now describes both the focused transmission process and the reception for an arbitrary focused array.

Note that the model (8.12) is valid for *any* transmit and receive aperture, provided that the corresponding receive and transmit SIRs can be computed.

The model (8.12) is therefore not restricted to using the same elements both in receive and transmit, and it can be used to model both dense- and sparse, or under-sampled, arrays.

8.3.2 The Optimal Linear Estimator

The optimal linear estimator can for an array layout can now be found by inserting the propagation matrix from (8.12) into (7.6). In the experiments performed here a simplified form of (7.6) has been used were the MMSE filter matrix, $\mathbf{K}_{\text{MMSE}}$, for the estimator was computed using Assumption 8.1. Using this assumption (7.6) reduces to

$$\mathbf{K}_{\text{MMSE}} = \mathbf{P}^T(\mathbf{P}\mathbf{P}^T + \mu\mathbf{I})^{-1}, \tag{8.13}$$

where $\mu = \sigma_{\text{e}}^2/\sigma_{\text{o}}^2$. The μ parameter in (8.13) is related to the "confidence" that the MMSE estimator has to the data. If the noise is very large then μ will also be large and the factor $\mu\mathbf{I}$ will dominate the inverse $(\mathbf{P}\mathbf{P}^T + \mu\mathbf{I})^{-1}$ in (8.13). The MMSE estimator will in the limit $\mu \to \infty$ then reduce to $\mathbf{P}^T/\mu$. The gain of the MMSE estimator decreases when σ_{e}^2, and μ, increases. Furthermore, the estimate will tend toward zero, which is the *a priori* mean value of $\mathbf{o}$. In other words for this case, the measurement vector $\mathbf{y}$ contains no new information about $\mathbf{o}$ and the estimate is only based on our prior information.

For the opposite case, when μ approaches zero, the MMSE estimator (8.13) reduces to $\mathbf{P}^T(\mathbf{P}\mathbf{P}^T)^{-1}$. This will however rarely be the case, since we always have uncertainties in our model, thus, we need to set μ to a positive value.[7] Below, we will study the performance of the MMSE estimator and the influence of the μ parameter.

8.3.3 Grating Lobe Suppression of the Optimal Linear Estimator

A very important factor that can severely deteriorate image quality in traditional DAS beamforming is the presence of grating lobes, as discussed in the Chapters 5 and 6. To evaluate the MMSE estimator (8.13) with respect to grating lobe suppression, measurements and simulations were performed using the setup shown in Figure 8.7. A steered array configuration was cho-

[7] Remember that the noise, $\mathbf{e}$, is a combination of all errors of the model (8.12), including measurement noise and model errors.

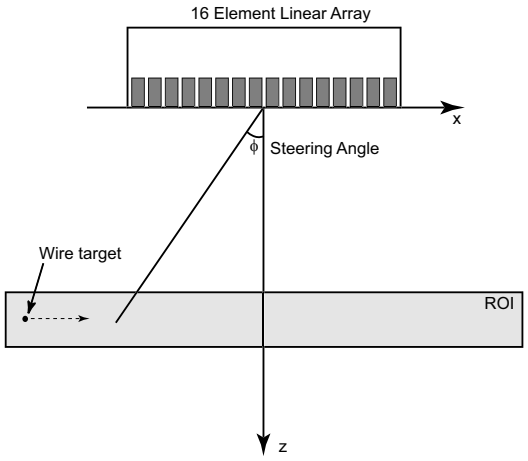


Figure 8.7: Illustration of the measurement setup for evaluation of the optimal linear parallel estimator. A 16 element phased array with an array pitch of 1 mm, focused at $z = 50$ mm and steered -20 degrees, was used.

sen since by steering the array, the grating lobes will also be steered and therefore appear more clearly as discussed in Chapter 6.[8]

A snapshot of the acoustic field simulated with the DREAM method for a 16 element steered array setup is shown in Figure 8.8.[9] The snapshot was taken slightly before the pulse reached the focal point at approximately ($x =$ -18 mm, z=50 mm). The acoustic energy in the region around $x = 8$ mm is due to the grating lobe of the under-sampled array. This can also be seen as the peak at $x = 8$ mm in the profile plot shown in Figure 8.8(b).

To study the behavior of the MMSE method, simulated data were obtained by using the linear model (8.12) were the noise vector $\mathbf{e}$ was simulated as white with Gaussian amplitude distribution. The SNR used in the simulations below was defined as

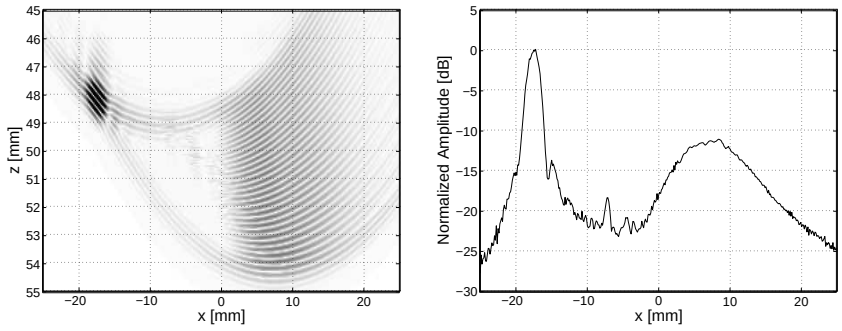
$$\eta \triangleq \frac{\mathbf{y}^T\mathbf{y}}{\mathbf{e}^T\mathbf{e}}, \tag{8.14}$$

where the signal energy, $\mathbf{y}^T\mathbf{y}$, computed using (8.12) with $\mathbf{e} = 0$, is given by

$$\mathbf{y}^T\mathbf{y} = \mathbf{o}^T\mathbf{P}^T\mathbf{P}\mathbf{o}. \tag{8.15}$$

[8] As also noted in Chapter 6 the array elements have a finite size, and their beam directivity will attenuate energy coming from positions at a high angle with respect to the vertical center axis. By steering the array, the angle between grating lobes and the center axis will be reduced and the response from targets in the grating lobe will be stronger.

[9] The under-sampled Allin array setup was simulated.



(a) Wave field snapshot at $t = 37.3$ μs.

(b) Profile plot of the wave field.

Figure 8.8: Snapshot of the acoustic field at -25 mm $< x <$ 25 mm, 45 mm $< z <$ 55 mm for a 16 element array with 0.9 mm wide elements and an array pitch of 1 mm, focused at $z = 50$ mm and steered -20 deg.

Thus, the signal energy, $\mathbf{y}^T\mathbf{y}$, is the total energy received by all array elements. Note that the received signal energy, when imaging a single point target, will depend on the position of the target. The SNR will therefore also be a function of the target position and, consequently, the SNR will be low when the target is in regions with low acoustic energy and high when the target is in the main or the grating lobe.

Experiment 1: Illustration of Grating Lobe Suppression of the MMSE Estimator

In the first experiment we consider the beamforming performance at the vertical image line where the focal point resides. If a target is horizontally moved from left to right under the array, cf. Figure 8.7, the reconstructed scattering amplitudes at that line should ideally be zero for all target positions except for the one where the target position and the position of the image line coincides. Normally, this will not be the case since, due to side- and grating lobes, energy will leak into to the main lobe resulting in ghost responses and a loss in resolution and image contrast.

Figure 8.9(a) shows the results for a simulated 16 element array, and Figure 8.9(b) shows the corresponding results for data measured with the Imasonic array. The max amplitudes of the beamformed images at the line

corresponding to the main lobe position, x = -18 mm, were plotted as a function of the target position. The DAS beamformer used a fixed receive focus at (z=50 mm, $x = -18$ mm), and the optimal filters for the corresponding image line were used in the MMSE estimator. In the simulations, the noise power was constant for all target positions where the SNR, defined by (8.14), was 15.8 dB when the target was at the focal point.

To show the influence of μ in (8.13), when μ is treated as a tuning parameter, the max amplitude was plotted as a function of the position of a unit area point target for five values of μ: $\mu = 10^2$, 10^3, 10^4, 10^6, and 10^9, respectively, see Figure 8.9(a). For low values of μ the MMSE filter relies strongly on the measured data, resulting in a significant noise amplification. This is seen in Figure 8.9(a), for $\mu = 10^2$, where the noise floor only is about 7 dB lower than the maximum amplitude at the focal point. Increasing μ lowers the noise sensitivity but it also results in a worse suppression of the grating lobe, which can be seen for $\mu = 10^6$ and $\mu = 10^9$. A good trade-off between the noise sensitivity and grating lobe suppression is in this case $\mu = 10^4$ where the noise and grating lobe levels are roughly 25 dB lower than the target signal. Note that the MMSE beamformer suppressed the grating lobe better for all the examined μ values. It is also noticeable that the MMSE reconstructed image has higher lateral resolution than the DAS beamformed data.

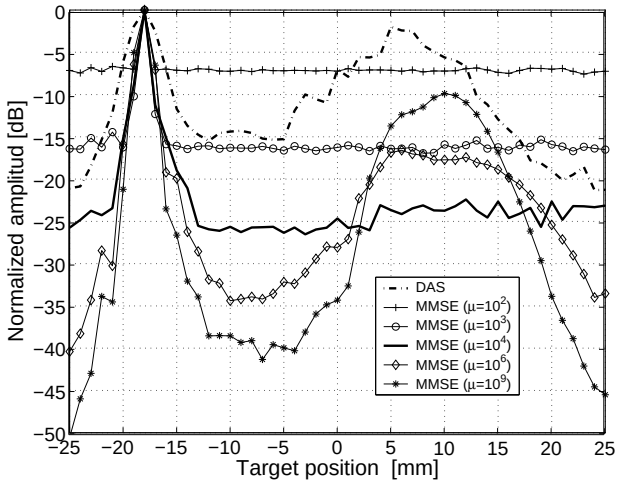
A similar behavior can be seen in the experimental results, shown in Figure 8.9(b), where $\mu = 10^2$ was used for the MMSE estimates.[10] Clearly, the results from MMSE method show both higher resolution and a lower amplitude at the grating lobe, even though model errors degraded the MMSE reconstructions somewhat.[11]

The fact the MMSE estimates becomes more similar to DAS beamformed data for large values of μ, can be explained by the fact that the MMSE estimator reduces to a matched filter, $\mathbf{P}^T/\mu$, when μ is large (cf. Section 5.4). Since the DAS beamformer approximately is a matched filter as well, their performance will be similar. Thus, if the variance of the measurement noise is very high, then the MMSE estimator will reduce to a matched filter, which is known to maximize the SNR.

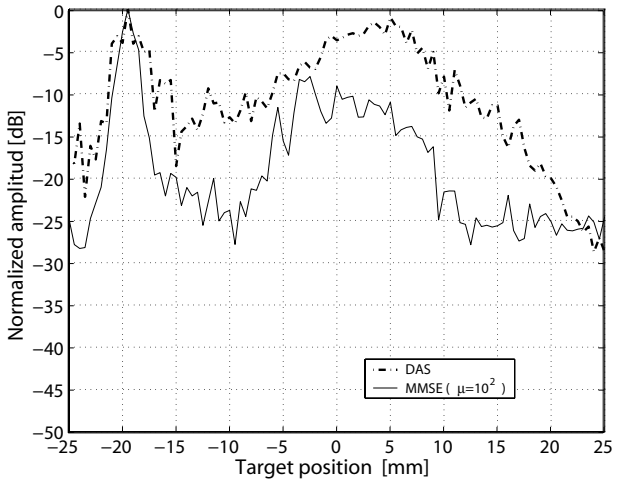
From the results presented above it can be concluded that the MMSE estimator is capable of eliminating most of the grating lobe effects, if the

[10] $\mu = 10^2$ resulted in the best visible performance.

[11] The synchronization in the ALLIN system is sometimes inexact which introduces modeling errors.



(a) Results for simulated data (SNR= 15.8 [dB]).



(b) Results for measured data using the Imasonic array and a 0.3 mm wire target.

Figure 8.9: Normalized maximum amplitude as a function of horizontal target position for a 16 element 1D array focused at z =50 mm and steered at an angle of -20 degrees. The point target was located at depth $z = 50$ mm and the SNR in (a), for the target at $x = -18$ mm, was 15.8 dB.

SNR is sufficient. We will continue to investigate these properties of the MMSE estimator in the next experiment presented below.

Experiment 2: Illustration of Grating Lobe Suppression for Parallel Array Imaging

In the second experiment the parallel DAS beamformer (5.7) and the MMSE estimator (8.13) were used to reconstruct a full image from a single emission from a wire target at a fixed position. More precisely, the entire 50 mm wide ROI was reconstructed from one emission.[12] Again, a 16 element 1D array, focused at z=50 mm and steered -20 degrees in transmission, was used.

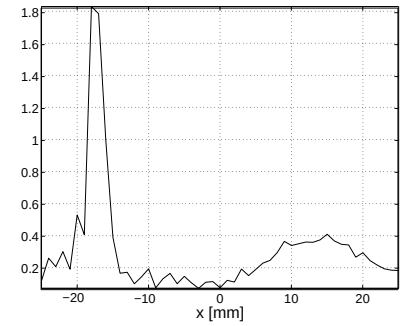
The profile plots of the reconstruction performance for two target positions was examined; the first one with the target in the main lobe at $x = -18$ mm, shown in Figure 8.10, and the second with the target located in the grating lobe at $x = 10$ mm, shown in Figure 8.11. The noise variance was constant in the simulations, which resulted in an SNR of 12.5 dB when the target was in the main lobe and a SNR of 10.7 dB for target in the grating lobe, where the SNR is defined in (8.14).

As can be seen in Figure 8.10, the DAS beamformed data has a clear peak at the grating lobe position, around $x = 10$ mm, for both the simulated and the measured data. This can not be seen in the MMSE processed images, presented in Figure 8.10(b) and (d). The same behavior is even more pronounced when the target is in the grating lobe. Then, the parallel DAS processed data shows two strong peaks at both the main lobe and the grating lobe positions, even though there is only one target located in the grating lobe.

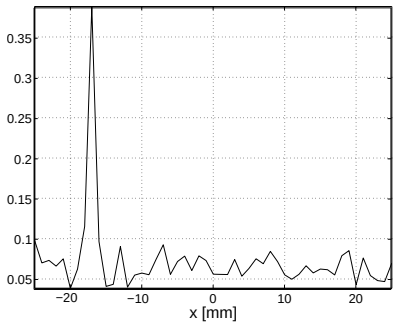
Also here the lateral resolution is superior for the MMSE method compared to the parallel DAS beamformer, and the width of the target peak in the MMSE reconstructed image is roughly half the width of the corresponding DAS image.

Thus, it can be concluded that the MMSE method can successfully compensate for grating lobes resulting from the use of under-sampled arrays. Essentially, the parallel MMSE beamformer can reconstruct a target regardless if the target is in the main or the grating lobe. The performance depends mainly on the acoustic energy at the particular target position. This topic will be discussed further in Chapter 9.

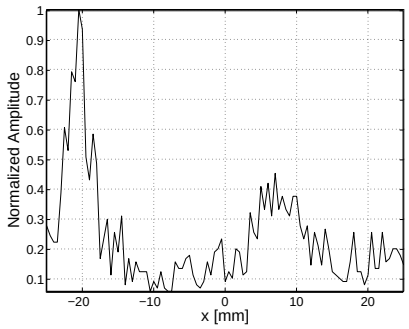
[12] As in Experiment 1 above, the "single" emission was emulated due to the limitations of the ALLIN system.



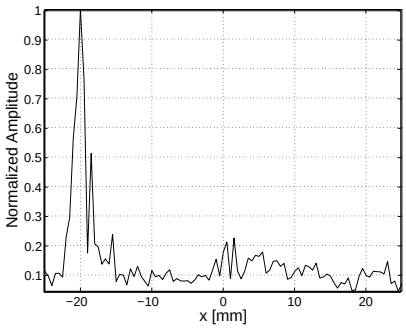
(a) Parallel DAS: Simulated data (SNR = 12.5 dB).



(b) MMSE: Simulated data (SNR = 12.5 dB).



(c) Parallel DAS: Measured data.

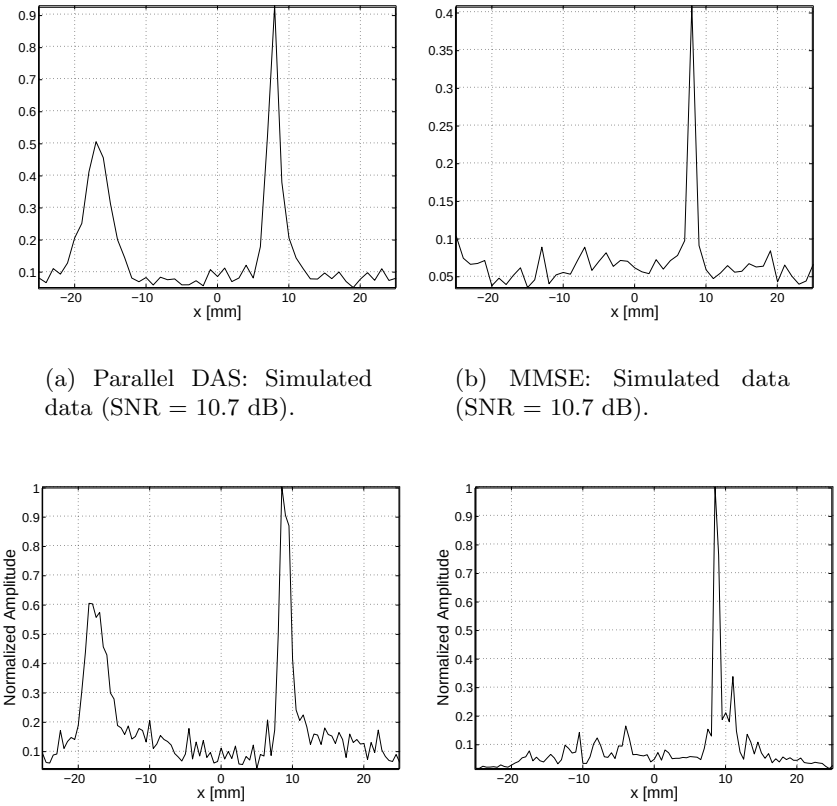


(d) MMSE: Measured data.

Figure 8.10: Profile plots for parallel beamformed data with a point target in the main lobe using a 16 element 1D array focused at $z = 50$ mm and steered -20 degrees.

8.4 Parallel Imaging with Exponential Priors

So far it has been assumed that $\mathbf{o}$ could take both negative and positive values. Such an assumption is reasonable if the inspected object may have an acoustic impedance that can be both higher and lower from that of the surrounding medium. In some applications it may however be appropriate to assume that the scattering strength only can take positive values. Consider, for example, a similar experiment to the experiments discussed in Section 8.3.2 above. There it is known *a priori* that the wire targets have a



(a) Parallel DAS: Simulated data (SNR = 10.7 dB).

(b) MMSE: Simulated data (SNR = 10.7 dB).

(c) Parallel DAS: Measured data.

(d) MMSE: Measured data.

Figure 8.11: Profile plots for parallel beamformed images with a point target in the grating lobe using a 16 element 1D array focused at $z = 50$ mm and steered -20 degrees.

much higher acoustic impedance than the surrounding medium. Thus, the resulting scattering amplitude must be positive. By using this information, the reconstruction performance should be improved, especially for band limited systems where linear deconvolution typically leads to ringings [38].

A drawback of imposing the positivity restriction on the scattering amplitude is that a closed form expression for a MAP estimator can no longer be found. This can be a serious drawback in ultrasonic array imaging applications due to the usually high dimensionality in these problems. However, in some cases the MAP solution has a structure that enables the use of very

efficient numerical methods. One such example is when the measurement noise is Gaussian and the prior distribution for $\mathbf{o}$ is the exponential PDF.[13]

Recall from Section 7.1.3 that the MAP estimate for an exponential prior and Gaussian measurement noise can be found by solving the optimization problem

$$\begin{aligned} \hat{\mathbf{o}} = \arg\min_{\mathbf{o}} \; & \frac{1}{2}(\mathbf{y}-\mathbf{P}\mathbf{o})^T {\mathbf{C}_{\mathrm{e}}}^{-1}(\mathbf{y}-\mathbf{P}\mathbf{o}) + \lambda_{\mathrm{o}}\mathbf{1}^T\mathbf{o} \\ \text{subject to} \quad & o_i \geq 0 \; \forall i. \end{aligned} \tag{8.16}$$

Eq. (8.16) is a quadratic programming (QP) problem with inequality constraints. QP problems have the important property that they always can be solved in a finite number of iterations, if a solution exists [85]. Here, the QP problem is convex since the Hessian matrix, $\mathbf{P}^T\mathbf{C}_{\mathrm{e}}^{-1}\mathbf{P}$ is positive semidefinite which further facilitates the solution of the problem.

The estimate obtained by solving (8.16) will hereafter be denoted the *positivity constrained quadratic programming*, or PCQP, estimate. In the two experiments described below the PCQP estimate has been compared with the DAS beamformer, the matched filter (MF), and the linear MMSE estimator. The image to estimate consisted of about 5000 unknowns and all computations were performed, on as-of-today, standard PC hardware in less than an hour.

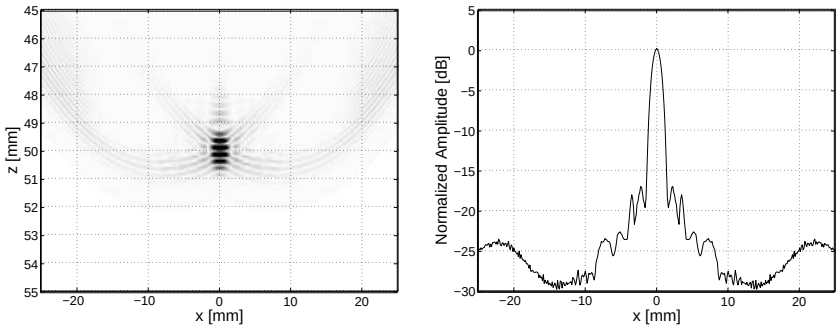
Below, the performance of the PCQP estimator is illustrated, using both simulated data and experimental data acquired using the Imasonic array.

Experiment 1: Focused Array using Simulated Data

In this experiment the PCQP estimate was compared with the parallel DAS beamformer, (5.7), the MMSE estimator, (8.13), and the matched filter, $\mathbf{P}^T$. The Imasonic array was simulated using 16 elements, focused at $z = 50$ mm. A snapshot of the acoustic field around the focal point for this setup is shown in Figure 8.12. Since the array is under-sampled one can notice two grating lobes, at approximately $x = \pm 23$ mm, but most of the acoustic energy is focused around the center axis $x = 0$.

Three simulations were performed, with SNR = 30, 10, and -20 dB, were the SNR is defined by (8.14). The true scattering image, $\mathbf{o}$, had five pairs of unit area point targets located at the horizontal positions x=-25, -13, 0, 13,

[13]The exponential PDF is also reasonable in the sense that it is the maximum entropy distribution if the mean is known as discussed in the previous chapter.



(a) Wave field snapshot at $t = 37.3$ μs.

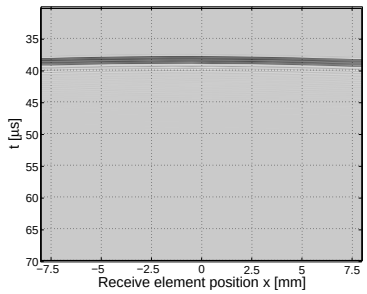
(b) Profile plot of the wave field.

Figure 8.12: Snapshot of the acoustic field at -25 mm $< x <$ 25 mm, 45 mm $< z <$ 55 mm for a 16 element array with 0.9 mm wide elements and an array pitch of 1mm focused at $z = 50$ mm.

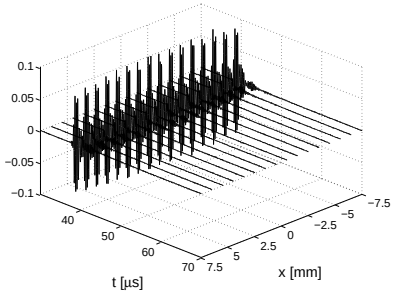
and 25 mm, and vertically at $z = 50$ and $z = 50.5$ mm, respectively. The vertical distance between the point targets is then roughly one wavelength of the corresponding center frequency of the array.

To obtain a fair comparison we need to set the parameters in the MMSE and the PCQP so that the expected energy using both estimators is the same. The prior that is determined by the mean value, λ_o, used in the PCQP estimator was based on the assumption that we can expect to find one unit area point target in the ROI. The mean scattered energy is the sum of the squared mean value and the variance, which for a single point target becomes $1^2/MN = \lambda_o^2 + \lambda_o^2$ since the variance is given by λ_o^2 for the exponential PDF. Thus, λ_o is given by $1/\sqrt{2MN}$. By a similar reasoning the variance, σ_o^2, used in the linear MMSE estimator is given by $\sigma_o^2 = 1/MN$.

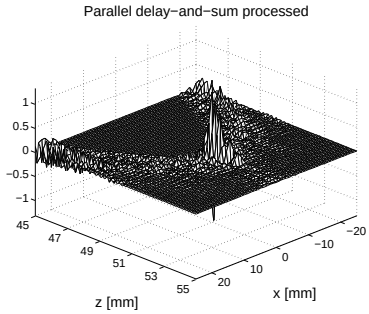
The reconstruction results, for the respective SNRs, are displayed in Figures 8.13, 8.14, and 8.15. If we study the DAS and matched filter (MF) beamformed images it is evident that even for the data with the relatively high SNR of 30 dB, shown in Figure 8.13, these two methods can only resolve the targets located close to the focal point. On the contrarym, the MMSE and PCQP estimators were able to resolve all targets. The PCQP estimator had, as it can be seen in 8.13(f), very low error in the scattering amplitudes at all target positions. The MMSE estimator underestimated



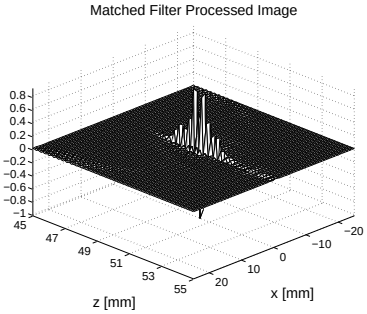
(a) Simulated data (B-mode image).



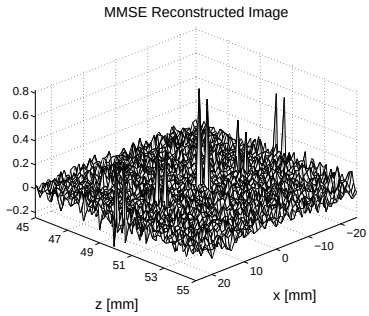
(b) Simulated data.



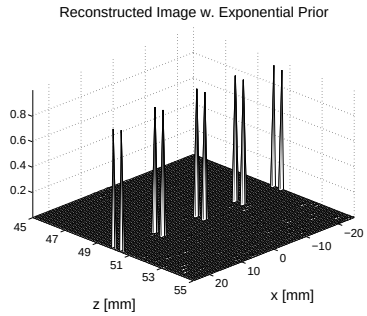
(c) DAS.



(d) Matched filter processed image.



(e) MMSE.



(f) PCQP.

Figure 8.13: Simulated data, with an SNR of 30 dB for five pairs of point targets at x=-25, -13,0 13, and 25 mm, at $z = 50$ and $z = 50.5$ mm, respectively, and the reconstructions using the DAS, MF, MMSE, and PCQP algorithms.

the amplitudes somewhat in the regions where the acoustic power was low[14], which is a reasonable behavior since the SNR in those regions is lower than at the focusing region and the estimates are therefore more uncertain.

When decreasing the SNR to 10 dB the PCQP estimate, shown in Figure 8.14, starts to show a similar behavior as the MMSE estimator, that is, it becomes more "cautious" at regions with low acoustic power. This is even more evident in the reconstructions from the rather low SNR of -20 dB shown in Figure 8.15. Now, both the MMSE and PCQP estimators only "trust" responses from the focusing region and all other responses are close to zero. It is also noticeable that the MMSE reconstruction becomes more similar to the MF output as the SNR decreases, see Figures 8.15(d) and (e). This is due to the fact, mentioned above, that the noise covariance matrix, $\mathbf{C}_{\mathrm{e}}$, dominates the factor $(\mathbf{P}\mathbf{C}_{\mathrm{o}}\mathbf{P}^T + \mathbf{C}_{\mathrm{e}})^{-1}$ in the MMSE estimator when the SNR is low and the MMSE filter is then approximately $\mathbf{K}_{\mathrm{MMSE}} \approx \sigma_{\mathrm{o}}^2\mathbf{P}^T(\sigma_{\mathrm{e}}^2\mathbf{I})^{-1} = \mathbf{P}^T\sigma_{\mathrm{o}}^2/\sigma_{\mathrm{e}}^2$.[15]

Noteworthy in Figure 8.15(f), is also the remarkable high ability of the PCQP estimator to suppress the measurement noise. Even though the received signals were completely hidden in the noise, see Figures 8.15(a) and (b), the PCQP estimator managed to predict the two targets at the focal point with a noticeable high contrast in the reconstructed image.

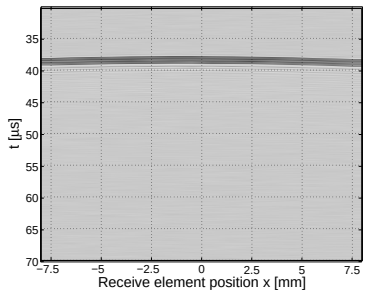
It is evident from these simulations that the two model based MMSE and PCQP methods are superior to traditional DAS based beamforming, in the sense that DAS processing does not perform well for targets outside the focusing region, whereas the two model based MMSE and PCQP methods can reconstruct the targets with high temporal and lateral resolution provided that the SNR is sufficiently high.

Noticeable is also that, in contrast to the DAS and the MF beamformers, the overall amplitudes of the estimates become lower for both the MMSE and the PCQP estimators as the noise variance increases. In essence, the data become less informative regarding the scattering amplitudes, $\mathbf{o}$, when the noise variance increases and the MMSE and PCQP estimators therefore becomes more cautious and underestimates the amplitudes.[16] This is a reasonable behavior which gives us a warning that the quality of the data may not be sufficient for our particular purpose.

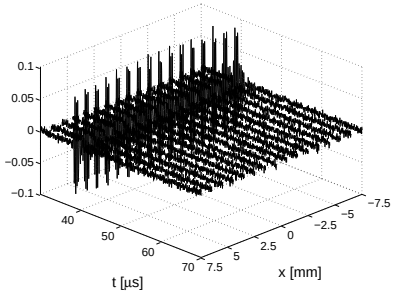
[14] The true target amplitude was 1 in this case.

[15] The gain of the MMSE estimator is lower than that of the MF due to the factor $\sigma_{\mathrm{o}}^2/\sigma_{\mathrm{e}}^2$, which will be low if the noise variance is high.

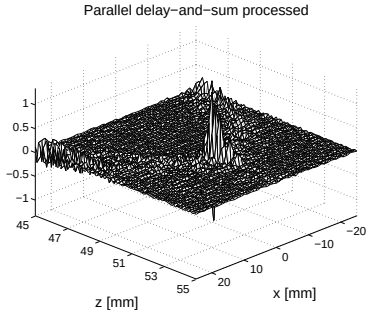
[16] The amplitudes goes toward the *a priori* mean value, which was 0 for the MMSE estimates and $1/\sqrt{2MN}$ for the PCQP estimates in this case.



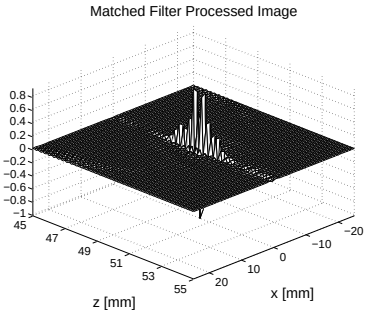
(a) Simulated data (B-mode image).



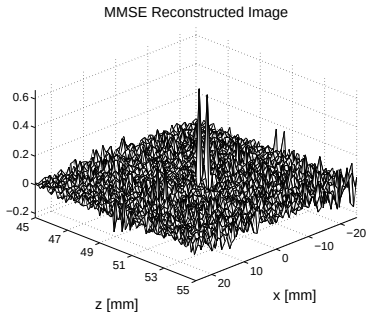
(b) Simulated data.



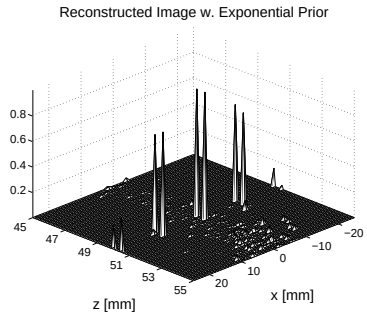
(c) DAS.



(d) Matched filter processed image.

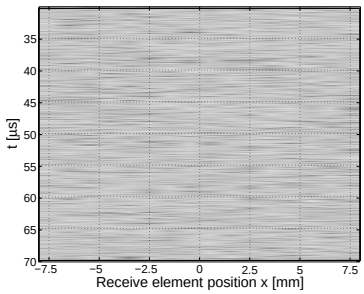


(e) MMSE.

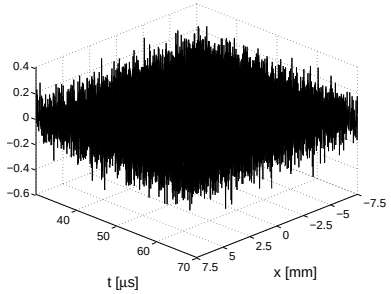


(f) PCQP.

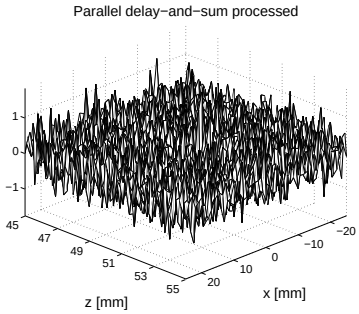
Figure 8.14: Simulated data, with an SNR of 10 dB for five pairs of point targets at x=-25, -13,0 13, and 25 mm, at $z = 50$ and $z = 50.5$ mm, respectively, and the reconstructions using the DAS, MF, MMSE, and PCQP algorithms.



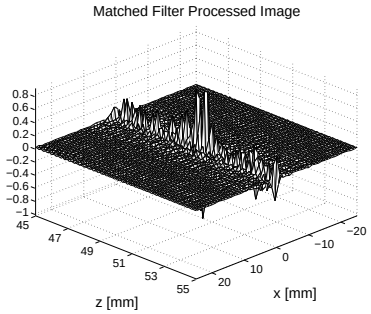
(a) Simulated data (B-mode image).



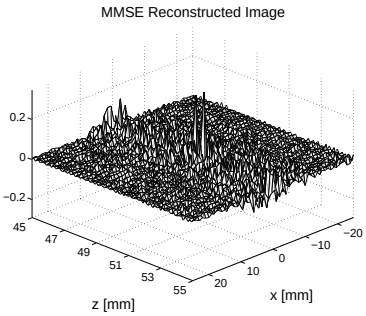
(b) Simulated data.



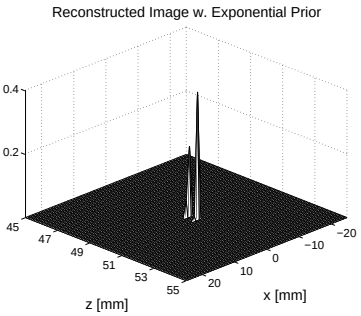
(c) DAS.



(d) Matched filter processed image.



(e) MMSE.



(f) PCQP.

Figure 8.15: Simulated data, with an SNR of -20 dB for five pairs of point targets at x=-25, -13,0 13, and 25 mm, at $z = 50$ and $z = 50.5$ mm, respectively, and the reconstructions using the DAS, MF, MMSE, and PCQP algorithms.

The general observation from these simulations is that it is the spatial distribution of acoustic energy that is important for the reconstruction performance, not the actual pulse shape. More specifically, the reconstruction performance of the model based methods employed here mostly depends on the SNR in the particular region[17] and not the length or shape of the waveform. Moreover, as demonstrated above, the MMSE and PCQP estimators could estimate targets outside the focusing region, where the transmitted waveforms from each element are not in phase, provided that the SNR was sufficient.

Experiment 2: Steered Array using Measured Data from Wire Targets

To verify the above mentioned conclusions regarding the ability of the Bayesian estimators to reconstruct targets outside the main beam, an experiment with an under-sampled steered array is presented here. As we have mentioned earlier, such an array will transmit energy both in the focusing direction and in the directions of the grating lobes. A relevant question is therefore if the Bayesian methods also are able to accurately estimate the scattering strength from a target located in a grating lobe?

For this purpose, the parallel DAS, MMSE and the PCQP methods were used to reconstruct data from a 16 element array, focused at $z = 50$ and steered at -20 degrees, acquired using the ALLIN system and the Imasonic array. A snapshot of the resulting wave field from this setup is presented in Figure 8.8 above. The measurements were performed using wire targets, positioned both in the main lobe at $x = -18$ mm and in the grating lobe at $x = 10$ mm, at depth $z = 50$ mm, see the illustration in Figure 8.16. The reconstruction results are shown in Figure 8.17.

The results for the DAS method, shown in Figures 8.17(a) and (b), again demonstrate the typical leakage from the grating lobe associated with conventional beamforming. The temporal resolution for the DAS method is roughly the same as the length of the acoustic pulse for the target in the main lobe.[18] Also, the temporal resolution for the target in the grating lobe is very poor which can be explained by the fact that the acoustic waveform is much longer in the grating lobe than at the focal point, see Figure 8.8(a).

As can be seen in Figures 8.17 (c)–(f), both the MMSE and the expo-

[17]The bandwidth of the system is also of importance [38].

[18]This is expected since in ordinary DAS processing there is no compensation for the transmitted pulse waveform.

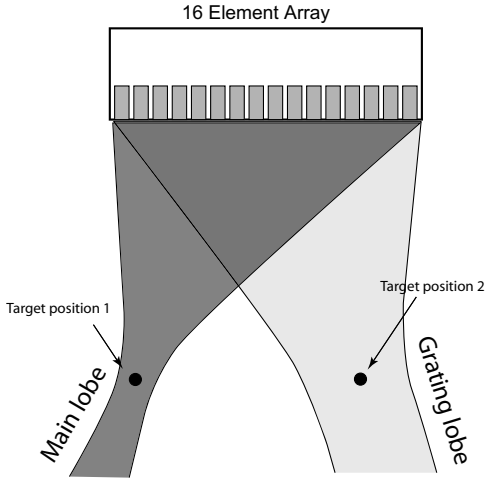
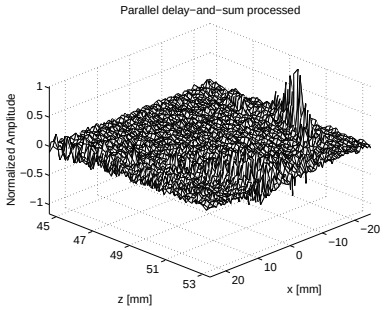


Figure 8.16: Illustration of the measurement setup for Experiment 2. The under-sampled 16 element array was focused at $z = 50$ mm and steered -20 degrees. A snapshot of the wave field for this configuration is presented in Figure 8.8.

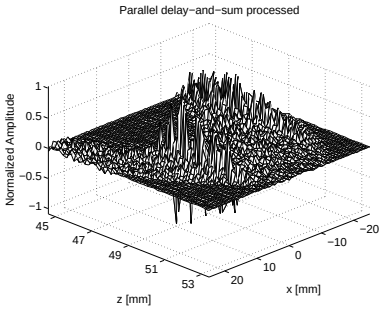
nential MAP estimators suppressed the grating lobe well for both target positions, and the temporal resolution has also been improved compared to DAS processing. The MMSE estimates have however some ringing, see Figure 8.17(d). This ringing behavior is most likely due to model errors caused by synchronization problems in the ALLIN array system. The PCQP estimates, shown in Figures 8.17(f) and (f), also have some extra "spikes" but the amplitude of those spikes are smaller than the oscillations in the MMSE estimates.

Evidently, both of the model based MMSE and PCQP estimators were able to accurately estimate the target strength, regardless if the wire target was located in the main or the grating lobe. This is in complete agreement with the simulations in Experiment 1 above.

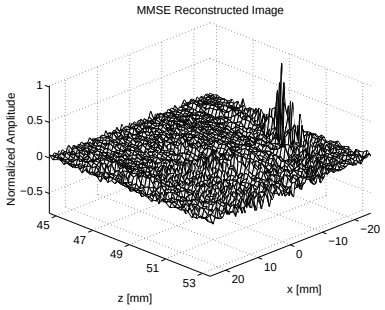
The results from the two experiments presented in this section shows that, at least for high contrast targets, the use of an exponential prior PDF for the scattering amplitudes can give impressive results with very high temporal and lateral resolution. The contrast is also very high and the noise level in the reconstructions are often very low. The remarkable performance of the exponential MAP estimator is due to the correct prior information that was incorporated into the model, which enables the estimator to effectively discriminate between noise and the target responses.



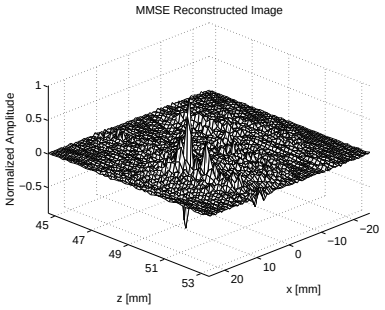
(a) DAS: Target in main lobe.



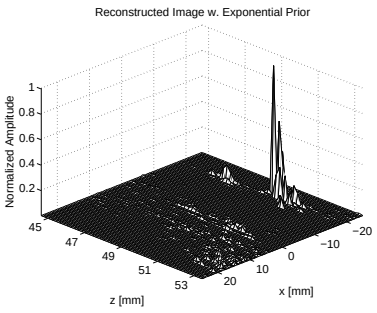
(b) DAS: Target in grating lobe.



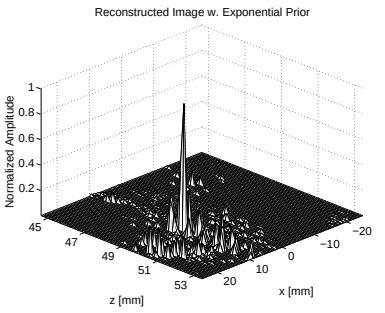
(c) MMSE: Target in main lobe.



(d) MMSE: Target in grating lobe.



(e) PCQP: Target in main lobe.



(f) PCQP: Target in grating lobe.

Figure 8.17: Reconstructions from measured data using a wire target in the main lobe (Figs. (a), (c), and (e)) and a wire target in the grating lobe (Figs. (b), (d), and (f)). The data were acquired using the ALLIN system and the concave Imasonic array. The transmit beam was steered at -20 deg and focused at 50 mm using 16 array elements.

The most significant drawback of the PCQP method is that the estimate must be obtain by an iterative optimization algorithm. This makes the method most suitable for offline processing where real-time constraints are of little concern.

8.4.1 Remarks

Focusing phase errors: The transmit focusing delays in ultrasonic array imaging systems are often subject to quantization errors. This results in a slightly different transmit beampattern then the desired. In classical beamforming this will result in increased sidelobe amplitudes [16, 90]. However, the model based estimators discussed here can compensate for these errors given that the true focusing delays are known.

Correlated scattering: In the experiments discussed in this chapter it was assumed that the scattering amplitudes at all image points were uncorrelated. In some applications the scattering objects can, for example, have a layered structure (e.g., in seismic exploration) resulting in correlations between the scatterers. This knowledge can be easily introduced in the linear MMSE estimator by means of the covariance matrix $\mathbf{C}_{\text{o}}$.

Priors: In the experiments performed with the ALLIN system presented in this chapter the *a priori* mean and variance of the scattering strength were in reality unknown. Despite this fact, we have seen in the results from both the linear MMSE and the non-linear PCQP method that the exact mean or variance are not required. Our experience is that the reconstruction results for both these model based methods are not sensitive to the particular value of the variance or the mean. The variance or mean can typically be changed by a factor of 10 without significantly affect the reconstruction performance. Furthermore, it is possible to design an experiment, with a known set of scatterers, to estimate the unknown quantities. However, this is a topic for future research.

8.5 Summary

In this chapter we have evaluated performance of the linear MMSE estimator and the PCQP estimator used for the reconstruction of ultrasonic data in

monostatic configuration as well as in multistatic parallel array imaging.

It was demonstrated that both estimators were capable of reconstructing multiple targets with high accuracy in the experiments even if the ultrasonic data were sparsely sampled. However, the PCQP estimator were more immune to the corrupting noise.

The MMSE estimator, which in the simplest case results in a spatio-temporal filter with only one scalar tuning parameter, clearly outperformed traditionally used DAS beamforming algorithms in terms of higher resolution both for the targets in water and in an immersed solid. In the parallel receive imaging, the MMSE algorithm clearly minimized target leakage from the grating lobes and could be successfully used for correct reconstruction of targets located in the grating lobes.

However, when our prior information tells us that the scattering strength always is positive the best performance was achieved using the nonlinear PCQP estimator developed for the exponential prior. This estimator clearly outperformed the linear MMSE estimators in terms of high immunity to the measurement noise and its ability to reconstruct multiple targets in the ROI.

The performance of the proposed estimators is achieved at the price of the increased computational complexity, especially, for the PCQP algorithm which takes the form of an iterative optimization algorithm.

CHAPTER 9

Performance of the Optimal Linear Estimator

IN this chapter we turn the attention to the influence that the array system design has on reconstruction performance and, in particular, how we can evaluate different array layouts. Traditionally, array systems are designed with classical beamforming in mind, meaning that the design of the array system must respect the limitations that classical beamfoming imposes. Namely, sub-wavelength array pitches and small element sizes in order to avoid artifacts and resolution loss. The Bayesian estimators, presented earlier in this thesis, do however not not impose these restrictions and a more flexible array design should therefore be possible.

The aim of this chapter is to show that Bayesian analysis offer a useful tool for designing and evaluating ultrasonic array systems. This is due to the fact that these methods provides a natural measure of the accuracy of the estimates. According to the discussion in Section 7.1, the performance of the optimal linear estimator is given by the error covariance matrix, $\mathbf{C}_\epsilon$, in (7.11), which can be analyzed for different array layouts.

In this chapter we will show that, for the linear MMSE estimator, the expected mean-squared error can give us valuable information regarding, in particular, the choice of suitable focusing strategies. In agreement with common sence, it will be demonstrated that focused beams are needed if the noise is high to obtain a sufficiently low expected error at least in the limited region of the main beam.

This chapter is organized as follows: In Section 9.1, we present an ex-

pression for the normalized expected error which is then used in Section 9.2, to evaluate the performance for three parallel array imaging setups.

9.1 The Normalized Expected Error

Recall from Chapter 7 that, for the linear MMSE estimator (7.7), the distribution for the reconstruction error, $\boldsymbol{\epsilon} = \hat{\mathbf{o}}_{\text{MMSE}} - \mathbf{o}$, is Gaussian if the measurement noise and the prior are both Gaussian. The expected mean-squared error at each observation point are given by the diagonal elements of the error covariance matrix $\mathbf{C}_\epsilon$. The performance of the MMSE estimator can then be evaluated by organizing the diagonal elements $\text{diag}\{\mathbf{C}_\epsilon\}$ into an $M \times N$ error image $\mathbf{E}_\epsilon$. Below, we present a normalized error covariance matrix for the purpose of facilitating the comparison of different transmit focusing methods.

The error covariance matrix is given by (cf. Appendix 7.A.2)

$$\begin{aligned} \mathbf{C}_\epsilon &= \mathbf{C}_{\text{o}} - \mathbf{C}_{\text{o}}\mathbf{P}^T(\mathbf{P}\mathbf{C}_{\text{o}}\mathbf{P}^T + \mathbf{C}_{\text{e}})^{-1}\mathbf{P}\mathbf{C}_{\text{o}} \\ &= (\mathbf{C}_{\text{o}}^{-\mathbf{1}} + \mathbf{P}^T\mathbf{C}_{\text{e}}^{-\mathbf{1}}\mathbf{P})^{-1}. \end{aligned} \tag{9.1}$$

Let us consider a special case where $\mathbf{C}_{\text{o}} = \sigma_{\text{o}}^2\mathbf{I}$ and $\mathbf{C}_{\text{e}} = \sigma_{\text{e}}^2\mathbf{I}$. Then the error covariance matrix $\mathbf{C}_\epsilon$ becomes

$$\begin{aligned} \mathbf{C}_\epsilon &= \sigma_{\text{o}}^2\mathbf{I} - \sigma_{\text{o}}^2\mathbf{P}^T(\sigma_{\text{o}}^2\mathbf{P}\mathbf{P}^T + \sigma_{\text{e}}^2\mathbf{I})^{-1}\mathbf{P}\sigma_{\text{o}}^2 \\ &= (\mathbf{I}/\sigma_{\text{o}}^2 + \mathbf{P}^T\mathbf{P}/\sigma_{\text{e}}^2)^{-1}. \end{aligned} \tag{9.2}$$

If we now normalize (7.67) with respect to σ_{o}^2, then the *normalized* error covariance matrix becomes

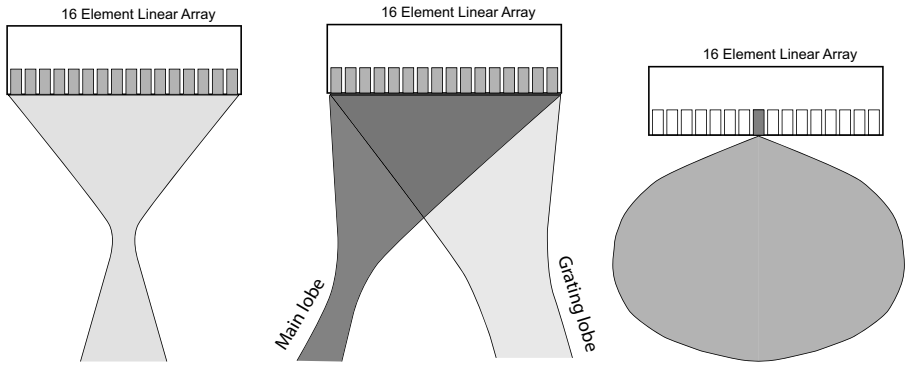
$$\begin{aligned} \tilde{\mathbf{C}}_\epsilon &= \frac{1}{\sigma_{\text{o}}^2}\mathbf{C}_\epsilon \\ &= \mathbf{I} - \mathbf{P}^T(\sigma_{\text{o}}^2\mathbf{P}\mathbf{P}^T + \sigma_{\text{e}}^2\mathbf{I})^{-1}\mathbf{P}\sigma_{\text{o}}^2 \\ &= (\mathbf{I} + \mathbf{P}^T\mathbf{P}\sigma_{\text{o}}^2/\sigma_{\text{e}}^2)^{-1}. \end{aligned} \tag{9.3}$$

If σ_{e}^2 is very large compared to $\mathbf{P}^T\mathbf{P}\sigma_{\text{o}}^2$, then the normalized error covariance matrix will reduce to the identity matrix which is the worst case scenario where the data holds no new information. A region in the normalized error image, $\tilde{\mathbf{E}}_\epsilon$ consisting of the diagonal elements of $\tilde{\mathbf{C}}_\epsilon$ with a value close to unity will then indicate poor reconstruction performance in that region.

In the next section we will utilize $\tilde{\mathbf{E}}_\epsilon$ to evaluate the performance of the linear MMSE estimator for a simulated phased array.[1]

9.2 Parallel Array Imaging Performance

To illustrate the performance of the parallel MMSE estimator (8.13), with respect to focusing method and SNR, three example array setups are considered. In all tree setups the Imasonic array described in Section 8.1 was simulated and 16 array elements are used in receive mode. The transmit beam is focused in tree different ways as illustrated in Figure 9.1.



(a) 16 transmit elements focused array.

(b) 16 transmit elements steered array.

(c) One transmit element.

Figure 9.1: Illustration of the transmit focusing methods for the simulations presented in Figures 9.2–9.4.

1. The beam is focused at $z = 50$ mm with no beam steering.

2. The beam is focused and steered -20 deg resulting in a focal point at $x = -18$, $z = 50$ mm.

3. One single element was used in transmit mode.

[1]The array was simulated using the DREAM toolbox.

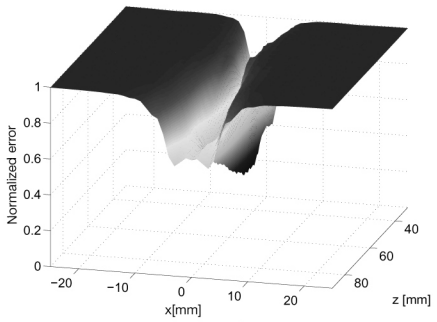
The first two setups, where the array is focused, result in regions with both high and low acoustic energy. The third setup, where a single element is used, results in a wider beam illuminating more of the ROI but the total transmitted acoustic energy will be lower. The reason for the lower transmit power is that by only using one element the active area of the transmit aperture is reduced by a factor 1/16 compared to the other two setups.

Normalized error images, obtained using (9.3) and the three focusing methods, are shown in Figures 9.2, 9.3, and 9.4, respectively. The noise variance σ_{e}^2 is highest in Figure 9.2 and lowest in Figure 9.4 and the ratio, $\sigma_{\mathrm{e}}^2/\sigma_{\mathrm{o}}^2$, in the respective figures are $2.5 \cdot 10^{-3}$, $2.5 \cdot 10^{-5}$, and $2.5 \cdot 10^{-8}$.

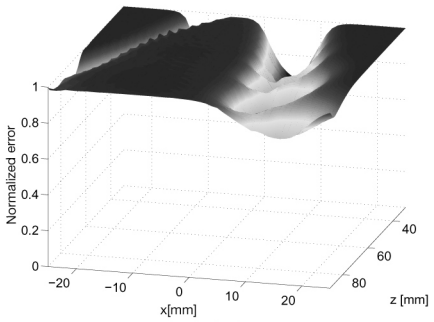
The error images presented in Figure 9.2 correspond to a very low SNR. It can clearly be seen in Figures 9.2(a) and (b) that the normalized error will only deviate from unity at those image points that are in the vicinity of a focal point or a grating lobe, that is, at those points where most of the energy is transmitted. It can also be seen that the error at the focal point is lower for the array that is not steered, even though the same number of elements are used in transmit mode for the steered array. This can be explained by considering the beam directivity of an individual array element as discussed in Section 6.2. An array element of the Imasonic array has a width of approximately 2λ at the array's center frequency, and the element's beam directivity can therefore not be neglected. The array element will transmit most energy in the normal direction of its active surface area where the SIRs are shortest, as discussed in Chapter 3. At high angles, the SIRs will attenuate the signal, especially for the higher frequencies. Elements that are horizontally closer to the focal point will, therefore, contribute more than the ones further away from the focal point. Steering at high angles will then essentially reduce the effective number of transmit elements due to the beam directivity of the array elements. In fact, it can be seen in Figure 9.2(b) that the error is of the same order both in the main- and grating lobe for the steered array. The reason for this behavior is that the grating lobe is closer to the array center than the main lobe and more elements will therefore contribute to the acoustic field.

It can also be seen in Figure 9.2(c) that the acoustic power from one element was not sufficient at this SNR since the normalized error was close to unity at all image points.

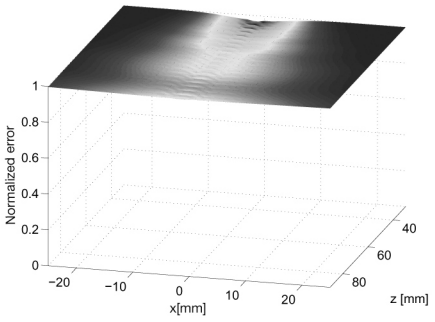
If the noise level is decreased then the estimation error will also decrease. The noise level in Figure 9.3 is 20 dB lower than in Figure 9.2 and the normalized error has been substantially decreased for all three trans-



(a) 16 transmit elements focused at $z = 50$ mm.



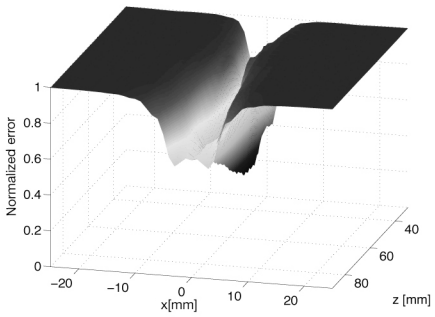
(b) 16 transmit elements focused at z = 50 mm and steered -20 deg.



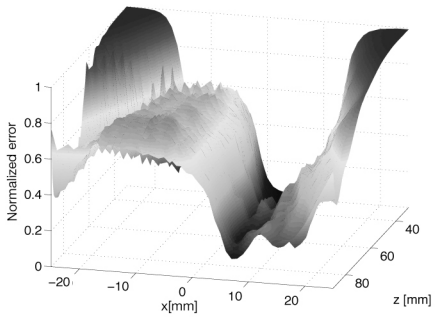
(c) One transmit element.

Figure 9.2: Normalized error images for the MMSE estimator using a 16 element 1D array. The ratio σ_e^2/σ_o^2 *is* $2.5 \cdot 10^{-3}$.

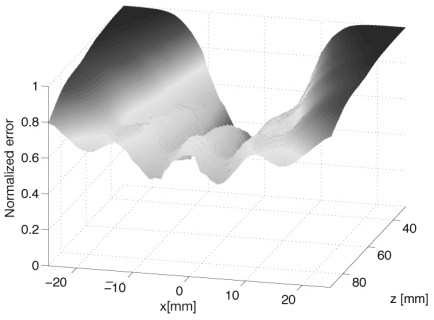
mit methods. The effects of the array element's beam directivity can now clearly be seen in Figure 9.3(c). The normalized error has been substantially decreased at image points close to center axis while the error is still rather large for $x \gtrsim 10$ mm. In Figure 9.3(c) it can also be seen that the error along the center line increases when the distance from the array increases. This is natural since the amplitude of the acoustic field decreases as the distance from the array increases.



(a) 16 transmit elements focused at $z = 50$ mm.



(b) 16 transmit elements focused at $z = 50$ mm and steered -20 deg.

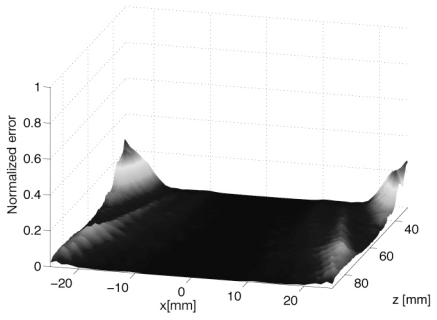


(c) One transmit element.

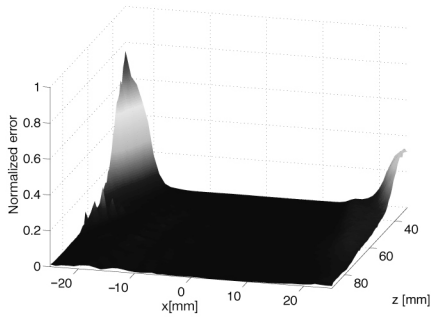
Figure 9.3: Normalized error images for the MMSE estimator using a 16 element 1D array. The ratio σ_e^2/σ_o^2 is $2.5 \cdot 10^{-5}$.

If the noise level is decreased further the error will be low also in regions outside the focused beams. In Figure 9.4 the noise level is decreased by another 20 dB and the normalized error is now very small for large regions in the ROI for all three focusing methods. It is only at image points located at high angles from the center of the array ($\gtrsim 40$ deg) that the error is still large due to the insufficient acoustic energy at those regions.

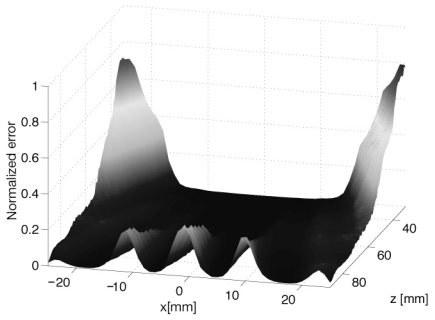
The conclusion from this investigation is that it is mainly the SNR that determines the reconstruction performance of the MMSE estimator. In other



(a) 16 transmit elements focused at $z = 50$ mm.



(b) 16 transmit elements focused at $z = 50$ mm and steered -20 deg.



(c) One transmit element.

Figure 9.4: Normalized error images for the MMSE estimator using a 16 element 1D array. The ratio σ_e^2/σ_o^2 is $2.5 \cdot 10^{-8}$.

words, provided that the SNR is sufficiently high the MMSE method can reconstruct a scattering object regardless of how the array is focused.[2] This is in sharp contrast to DAS based methods where, as seen in previous chapters, for example grating lobes can generate ghost responses.

[2]Remember that modeling errors are included in the noise term and even if the thermal measurement noise goes to zero the reconstruction error may still not go to zero due to these errors.

CHAPTER 10

Concluding Remarks and Future Research

The classical approach to ultrasonic imaging using arrays is inspired by geometrical optics and consists in designing discrete versions of lens systems capable of focusing and steering ultrasonic beams.

In the new approach to ultrasonic imaging presented in this work the ultrasonic image reconstruction was formulated as a linearized inverse-scattering problem taking into account both the discrete nature of the imaging setup as well as the modeling errors. This formulation has enabled global optimization of the imaging performance given the a priori information on the imaging system parameters and the PDF's of errors and scatterers present in ROI. Although the optimization in this work was performed in terms of resolution, given the parameters, such as, transducer size, its electro-mechanical characteristics and the exciting waveform it is also possible to include these parameters in the optimization process. This creates the possibility of optimizing transducer's bandwidth and the form of its exciting signal according to the imaging performance to achieve a global optimum of the imaging process.

The estimation was performed by means of the tools from Bayesian estimation theory using the specially developed for this purpose linear discrete matrix model of the imaging system. It is known from statistics that the MAP approach leads to analytic solutions only for a very few PDF families, in most cases the optimization of the a posteriori probability has to be performed numerically.

In the work the measurement errors were assumed to be Gaussian, and two special prior distributions of scatterers were considered, Gaussian and exponential.

It was shown that for Gaussian PDF's the MAP approach results in the MMSE criterion. The linear MMSE solution takes the form of a linear spatio-temporal filter which deconvolves the distortion caused by the diffraction effects of the transducer used for imaging using the information contained in the transducer's SIRs.

For the exponential PDF of scatterers in the ROI the solution was formulated as a nonlinear optimization problem with constrains which was solved numerically.

The imaging algorithms resulting from the proposed approach were verified using simulated and real ultrasonic data. The algorithms are of general character and can be applied both to the synthetic aperture imaging (monostatic and bistatic) as well as to the parallel beamforming.

It was demonstrated that the linear MMSE and the nonlinear PCQP estimators were superior to the classical beamforming in terms of resolution and sidelobe level.

It should be noted however, that the successive application of the proposed approach relies on the reliable information concerning the imaging setup, transducer's SIRs as well as the statistics of the measurement noise and the scatterers present in ROI. In our experiments the SIRs calculated numerically for given setup and the transducer geometry were successfully used. The statistics were not estimated explicitly but found using the trial-and-error method for the particular ultrasonic data.

Future Research

Model size reduction: The propagation matrix $\mathbf{P}$ in (4.23) is inherently large if the ROI is densely sampled. In many applications there are only a few scatterers, or the scatterers are confined to a localized area. In such situations many of the observation points will not contribute to the measurement $\mathbf{y}$. If this is known in advance the matrix $\mathbf{P}$ can be truncated resulting in a much smaller problem. If the location of the targets is not known in advance one needs a "zooming" strategy to find areas where the sampling needs to be dense.

Prior: Assumptions concerning the prior are needed for the model based

imaging algorithms. Since little is known about scatterers' statistics experiments with a set of known objects yielding reliable estimates of the mean and variance should be performed.

Optimal input signal design: The choice of input signal driving the ultrasonic array elements is fundamental in the sense that it may determine the reconstruction accuracy [91, 92]. The input signals commonly used in ultrasonic imaging, such as, tone-burst or impulse excitation may not be the optimal choice with respect to reconstruction performance. Since the developed linear model takes into account the input signal, the reconstruction algorithms based on the model will automatically compensate for the chosen input signal. In effect the reconstruction performance can be optimized with respect to the input signal. Therefore developing feasible methods for designing suitable input signals for the particular ultrasonic array imaging applications is of great interest [93].

Parallel implementations: The matrix based models are inherently large and therefore, the image reconstruction, that generally results in solving linear equation systems or solving (constrained) optimization problems, also results in large equation sets. To be able to scale up these problems and improve computation time it is desirable to distribute the computation on parallel computers.

Sparse arrays: Sparse arrays are used in 3D medical applications. Performance of the proposed algorithms should be experimentally verified for such arrays.

BIBLIOGRAPHY

[1] S.L. Bridal, J-M. Correras, A. Saïed, and P. Laugier. Milestones on the road to higher resolution, quantative, and functional imaging. *Proceedings of the IEEE*, 91(10):1543–1561, 2003.

[2] O.T. von Ramm and S.W. Smith. Beam steering with linear arrays. *IEEE Transactions on Biomedical Engeering*, 30:438–452, August 1983.

[3] K.E. Thomenious. Evaluation of ultrasound beamformers. In *IEEE Ultrasonics Symposium*, 1996.

[4] H. Krim and M. Viberg. Two decades of array signal processing research. *IEEE Signal Processing Magasine*, pages 67–94, July 1996.

[5] J.A. Seydel. Ultrasonic synthetic-aperture focusing techniques in NDT. Research Techniques for Nondestructive Testing. Academic Press, 1982.

[6] S.R. Doctor, T.E. Hall, and L.D. Reid. SAFT—the evolution of a signal processing technology for ultrasonic testing. *NDT International*, 19(3):163–172, June 1986.

[7] K.J. Langenberg, M. Berger, Th. Kreutter, and V. Mayer, K.and Schmitz. Synthetic aperture focusing technique signal processing. *NDT International*, (3):177–189, June 1986.

[8] M. Karman and M. O'Donnell. Synthetic aperture imaging for small scale systems. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 42(3):429–442, May 1995.

[9] G.S. Kino. *Acoustic Waves: Devices, Imaging and Analog Signal Processing*, volume 6 of *Prentice-Hall Signal Processing Series*. Prentice-Hall, 1987.

[10] L.W. Schmerr, Jr. *Fundamentals of Ultrasonic Nondestructive Evaluation: A Modeling Approach*. Plenum Press, 1998.

[11] J. Krautkrämer and K. Krautkrämer. *Ultrasonic Testing of Materials*. Springer-Verlag, 1990.

[12] J.A. Ogilvy and J.A.G. Temple. Diffraction of elastic waves by cracks: application to time-of-flight inspection. *Ultrasonics*, pages 259–269, Nowember 1983.

[13] B.R. Tittmann. Scattering of elastic waves from simple defects in solids, a review. *Wave Motion*, 1983.

[14] M.G. Silk. Changes in ultrasonic defect location and sizing. *NDT International*, 20(1), February 1987.

[15] P.D. Corl, G.S. Kino, C.S. DeSilets, and P.M. Grant. A digital synthetic focus acoustic imaging system. *Acoustical Imaging*, pages 39–53, 1980.

[16] B.D. Steinberg. Digital beamforming in ultrasound. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 39(6):716–721, November 1992.

[17] J.W. Goodman. *Introduction to Fourier Optics*. McGraw-Hill, second edition, 1996.

[18] R.N. Thomson. Transverse and longitudinal resolution of the synthetic aperture focusing technique. *Ultrasonics*, 22(1):9–15, January 1984.

[19] H. Lee and G. Wade. Constructing an imaging operator to enhance resolution. *Journal of the Acoustical Society of America*, 75(2):499–504, February 1984.

[20] H. Lee. Resolution enhancement by wavefield extrapolation. *IEEE Transactions on Sonics and Ultrasonics*, (6):642–645, 1984.

[21] R. Stoughton. Optimal near-field measurement in the prescence of noise and reflections. *Journal of the Acoustical Society of America*, 92(2):831–840, August 1992.

[22] R. Stoughton. Source imaging with minimum mean-squared error. *Journal of the Acoustical Society of America*, 94(2):827–834, Aug 1993.

[23] S. Strait and R. Stoughton. Source imaging with minimum mean-squared error. *Journal of the Acoustical Society of America*, 94(2):827–834, August 1993.

[24] S.K. Lehman and A.J. Devaney. Transmission mode time-reversal super-resolution imaging. *Journal of the Acoustical Society of America*, 113(5):2742–2753, May 2003.

[25] P. Blomgren and G. Papanicolaou. Super-resolution in time-reversal acoustics. *Journal of the Acoustical Society of America*, 111(1):230–248, January 2002.

[26] A.A. Aydiner and W.C. Chew. On the nature of super-resolution in inverse scattering. In *IEEE Antennas and Propagation Society, AP-S International Symposium*, pages 507–510, 2003.

[27] A.J. Willis and R. De Mello Koch. A linear reconstructive approach to super-resolution array processing. *Journal of the Acoustical Society of America*, 97(5):3002–3005, May 1995.

[28] C. Prada and J-L. Thomas. Experimantal subwavelength localization of scatterers by decomposition of the time reversal operator intepreted as a covariance matrix. *Journal of the Acoustical Society of America*, 114(1):235–243, 2003.

[29] E. T. Jaynes. Detection of extra-solar-system planets. In *Maximum-Entropy and Bayesian Methods in Science and Engineering*. Kluwer, Dordrecht, 1988.

[30] K.L. Gammelmark and J.A. Jensen. Multielement synthetic transmit aperture imaging using temporal encoding. *IEEE Transactions on Medical Imaging*, 22(4):552–563, April 2003.

[31] G. Demoment. Image reconctruction and restoration: Overview of common estimation structures and problems. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 42:1024–2036, Dec 1989.

[32] P.N.T. Wells. Ultrasonics: a window into biomedical science. *Ultrasonics*, 30(1):3–7, 1992.

[33] M. Soumekh. *Fourier Array Imaging*. Prentice Hall, 1994.

[34] B.R. Mahafza. *Radar Systems Analysis and Design Using MATLAB.* Chapman & Hall/CRC, 2000.

[35] P. de Heering. Alternate schemes in synthetic aperture sonar processing. *IEEE Journal of Oceanic Engineering*, 9(4):277–280, October 1984.

[36] P.T. Gough and D.W. Hawkins. Imaging algorithms for strip-map synthetic aperture sonar: Minimizing the effects of aperure errors and aperture undersampling. *IEEE Journal of Oceanic Engineering*, 22:27–39, January 1997.

[37] B. Haider, P.A. Levin, and K. Thomenious. Pulse elongation and deconvolution filtering for medical ultrasonic imaging. In *IEEE Ultrasonics Symposium*, pages 1303–1308, 1995.

[38] T. Olofsson. *Maximum A Posteriori Deconvolution of Ultrasonic Data with Applications in Nondestructive Testing.* PhD thesis, Uppsala University, Uppsala, Sweden, 2000.

[39] M. Fatemi and A.C. Kak. Ultrasonic B-scan imaging: theory of image formation and a technique for restoration. *Ultrasonic Imaging*, 2:1–47, 1980.

[40] S.F. Burch. Comparison of SAFT and two-dimensional deconvolution methods for the improvement of resolution in ultrasonic B-scan images. *Ultrasonics*, 25(4):259–266, 1987.

[41] A. Herment, J.P. Guglielmi, P. Dumee, P. Peronneau, and P. Delouche. Limitations of ultrasound imaging and image restoration. *Ultrasonics*, 25(5):267–273, September 1987.

[42] T. Taxt. Restoration of medical ultrasound images using two-dimensional homomorphic deconvolution. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 42(4):543–544, July 1995.

[43] M. Tanter, J-L. Thomas, and M. Fink. Time reversal and the inverse filter. *Journal of the Acoustical Society of America*, pages 223–234, July 2000.

[44] M. Tanter, J.-F. Aubry, J. Gerber, J.-L. Thomas, and M. Fink. Optimal focusing by spatio-temporal inverse filter. I. Basic principles. *Journal of the Acoustical Society of America*, pages 37–47, July 2001.

[45] O.S. Haddadin and E.S. Ebbini. Ultrasonic focusing through inhomogeneous media by application of the inverse scattering problem. *Journal of the Acoustical Society of America*, 104(1):313–325, July 1998.

[46] P.M. Morse and K.U. Ingard. *Theoretical Acoustics*. 1968.

[47] J.D. Logan. *Applied Mathematics*. Wiley, second edition, 1997.

[48] P.M. Morse and H. Feshbach. *Methods of Theoretical Physics: Part I*. McGraw-Hill, 1953.

[49] H. Lasota, R. Salamon, and B. Delannoy. Acoustic diffraction analysis by the impulse response method: A line impulse response approach. *Journal of the Acoustical Society of America*, 76(1):280–290, July 1984.

[50] J.W.S. Rayleigh. *The Theory of Sound*, volume 2. New York: Dover, 1945.

[51] B. Delannoy, H. Lasota, C. Bruneel, R. Torguet, and E. Bridoux. The infinite planar baffles problem in acoustic radiation and its experimental verification. *Journal of Applied Physics*, 50(8):5189–5195, August 1970.

[52] B. Piwakowski and B. Delannoy. Method for computing spatial pulse response: Time-domain approach. *Journal of the Acoustical Society of America*, 86(6):2422–32, Dec 1989.

[53] B. Piwakowski and K. Sbai. A new approach to calculate the field radiated from arbitrarily structured transducer arrays. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 46(2):422–40, 1999.

[54] A. Lhémery. Impulse-response method to predict echo-responses from targets of complex geometry. Part I: Theory. *Journal of the Acoustical Society of America*, pages 2799–2807, November 1991.

[55] H. Lasota and R. Salamon. Application of time-space impulse responses to calculations of acoustic field in imaging systems. *Acoustic Imaging*, pages 493–512, 1980.

[56] P.R. Stepanishen. Transient radiation fom pistons in an infinite planar baffle. *Journal of the Acoustical Society of America*, 49:1629–38, 1971.

[57] J.C. Lockwood and J.G. Willette. High-speed method for computing the exact solution for the pressure variations in the nearfield of a baffled

piston. *Journal of the Acoustical Society of America*, 53(3):735–741, March 1973.

[58] C.A. Wiley. Synthetic aperture radars. *IEEE Transactions on Aerospace and Electronic Systems*, 21:440–443, 1985.

[59] C.W. Sherwin, J.P. Ruina, and R.D. Rawcliffe. Some early developments in synthetic aperture radar systems. *IRE Transactions Military Electronics*, 6:111–115, April 1962.

[60] K.D. Rolt and H. Schmidt. Azimultal ambiguities in synthetic aperture sonar and synthetic radar imagery. *IEEE Journal of Oceanic Engineering*, 17(1):73–79, January 1992.

[61] S.I. Nikolov and J.A. Jensen. 3D synthetic aperture imaging using a virtual source element in the elevation plane. In *IEEE Ultrasonics Symposium*, 2000.

[62] C.H. Frazier and Jr. W.D. O'Brien. Synthetic aperture techniques with a virtual source element. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 45(1):196–207, January 1998.

[63] M-H. Bae and M-K. Jeong. A study of synthetic-aperture imaging with virtual source elements in B-mode ultrasound imaging systems. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 47(6):1510–1519, November 2000.

[64] D. Vray, C. Haas, T. Rastello, M. Krueger, E. Busseau, K. Schroeder, G. Gimenez, and H. Ermert. Synthetic aperture-based beam compression for intravascular ultrasound imaging. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 48(1):189–201, January 2001.

[65] C. Passman and H. Ermert. A 100MHz ultrasound imaging system for dermatologic and ophthalmologic diagnostics. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 43(4):545–552, July 1996.

[66] M. O'Donnell and L.J. Thomas. Efficient synthetic aperture imaging from a circular aperture with possible application to catheter-based imaging. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 39(3), May 1992.

[67] K. Nagai. A new synthetic-aperture focusing method for ultrasonic B-scan imaging by the Fourier transform. *IEEE Transactions on Sonics and Ultrasonics*, 32(4):531–536, July 1985.

[68] K. Mayer, R. Marklein, K.J. Langenberg, and T. Kreutter. Three-dimensional imaging system based on fourier transform synthetic aperture focusing technique. *Ultrasonics*, 28:241–255, July 1990.

[69] M. Lorenz. *Ultrasonic Imaging for the Characterization of Defects in Steel Components*. PhD thesis, Delft University of Technology, 1993.

[70] R. Hannemann. *Modeling and Imaging of Elastodynamic Wave Fields in Homogenous Anisotropic Media*. PhD thesis, Kassel University, 2001.

[71] S. Nikolov. *Synthetic Aperture Tissue and Flow Ultrasonic Imaging*. PhD thesis, Technical University of Denmark, 2001.

[72] J. Chatillon, M.E. Bouhier, and M.E. Zakharia. Synthetic aperture sonar for seabed imaging: relative merits of narrow-band and wide-band approached. *IEEE Journal of Oceanic Engineering*, 17(1):95–105, Jan 1992.

[73] J.A. Cadzow. Matched filters. In *Foundations of Digital Signal Processing and Data Analysis*, pages 442–461. Macmillan, NY, 1987.

[74] G.L. Turin. An introduction to matched filters. *IRE Transactions on Information Theory*, 6(3):311–329, 1960.

[75] C. Dorme and M. Fink. Focusing in the transmit-receive mode through inhomogeneous media: The matched filter approach. *Journal of the Acoustical Society of America*, 98:1155–1161, 1995.

[76] M. Fink. Time-reversal mirrors. *Journal of Applied Physics*, 26:1333–1350, 1993.

[77] J. Shen and E. Ebbini. A new coded-excitation ultrasonic imaging system—Part I: Basic principles. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 43(1), January 1996.

[78] S.-C. Wooh and Y. Shi. Influence of phased array element size on beam steering behavior. *Ultrasonics*, 36:737–749, 1998.

[79] D.P. Shattuck, M.D. Weinshenker, S.W. Smith, and O.T. von Ramm. Explososcan: a parallel processing technique for high speed ultrasound

imaging with linear phased arrays. *Journal of the Acoustical Society of America*, 75(4):1273–1282, April 1984.

[80] O.T. von Ram, S.W. Smith, and H.G. Pavy Jr. High-speed ultrasound volumetric imaging system. II. Parallel processing and image display. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 38(2):109–15, March 1991.

[81] T.M. Cover and J.A. Thomas. *Elements of Information Theory*. Wiley, 1991.

[82] E.T. Jaynes. *Probability Theory: The Logic of Science*. Cambridge Universiy Press, 1st edition, 2003.

[83] W.R. Gilks, S. Richardson, and D.J. Spieglhalter. *Markov Chain Monte Carlo in Practice*. Chapman & Hall/CRC, 1996.

[84] S.M. Kay. *Fundamantals of Statistical Signal Processing: Estimation Theory*, volume I. Prentice Hall, 1993.

[85] J. Nocetal and S.J. Wright. *Numerical Optimization*. Springer-Verlag, 1999.

[86] G. Strang. *Linear Algebra and its Applications*. Sounders HBJ, 1986.

[87] A. Tikonov and V. Arsenin. *Solutions of Ill-Posed Problems*. Winston Washington DC, 1977.

[88] T. Söderström. *Discrete-Time Stochastic Systems: Estimation & Control*. Prentice Hall International, 1994.

[89] B. Delannoy, R. Toutguet, C. Bruneel, E. Bridoux, J.M. Rouvaen, and H. Lasota. Acoustical image reconstruction in parallel-processing analog electronic systems. *Journal of Applied Physics*, 50(5):3153–3159, May 1979.

[90] S. Holm. Analysis of worst-case phase quantization sidelobes in focused beamforming. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 39(5):593–599, September 1992.

[91] R.K. Mehra. Optimal input signals for parameter estimation in dynamical systems—survey and new results. *IEEE Transactions on Automatic Control*, (6):753–768, Dec 1974.

[92] S.M. Sowelam and A.H. Tewfik. Waveform selection in radar target classification. *Transactions on Information Theory*, 46(3):1014–1029, May 2000.

[93] B. Haider, P.A. Levin, and K. Thomenious. Pulse elongation and deconvolution filtering for medical ultrasonic imaging. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 45(1):98–113, January 1998.