Methods for determining focal point and delay for ultrasound arrays and multichannel electronics

Svante Rosenlind

Master's thesis in Biomedical Engineering

Supervisor Fraunhofer IBMT: Marc Fournelle
Supervisors LTH: Magnus Cinthio, Tobias Erlöv
Examiner: Monica Almqvist

LUND
UNIVERSITY

Faculty of Engineering, LTH
Department of Biomedical Engineering
Lund University
Sweden

Acknowledgements

The majority of this thesis was carried out at the Fraunhofer Institut für Biomedizinische Technik (IBMT) in St. Ingbert in Saarland, Germany, and I am very grateful for the opportunity to have done so. The entire team's continuous help and patience both with my sometimes dubious time management and my even more dubious technical German vocabulary was very helpful throughout the entire process, and especially the contributions of my supervisor Marc Fournelle.

I would also like to thank my supervisors Magnus Cinthio and Tobias Erlöv at the department of biomedical engineering (BME) at LTH for helpful and quick feedback whenever I had questions.

Finally, I want to thank the many friends I made in Germany during my time there. The pandemic was taxing, and being able to have a social circle of supporting friends in a foreign country is not something I take for given.

Abstract

High intensity focused ultrasound is a growing technique for tissue ablation, among other uses, and given its destructive capabilities, there is a need for control of where the energy is delivered. There exist a number of methods for focusing such ultrasound arrays, but these often assume prior knowledge of the impulse response, or require extensive full-system simulations. This thesis explores schemes for calculating impulse and frequency response of simple but still nonhomogeneous media, and implements different focusing methods, the spatiotemporal inverse filter, the Gerchberg-Saxton algorithm, and gradient descent, to test them.

With a 128-channel transducer operating at 5 MHz, these techniques are carried out in a simulated 2D setting on water and concrete with first a straight edge and then an oblique one between the two media. With a focus depth of 5 cm, the techniques are able to clearly outperform the uncompensated results, and were able to produce feasible foci even for offset or multiple simultaneous foci locations.

Although the optimization-based method did fail to produce adequate results for parts of the test, the overall investigation was seen as a successful venture, and that extension of the techniques to more complex media and 3D settings would be needed before any practical value can be realized.

Keywords: Focusing algorithms, k-wave, Spatiotemporal inverse filter, Gerchberg-Saxton algorithm, Focused ultrasound

Contents

Chapter 1

Introduction

Sound, in its most general definition, is simply mechanical waves passing through any medium, be it solid, liquid, or gaseous. These waves can be of any frequency, but are generally divided into three bands:

- Infrasound, with frequencies lower than 20 Hz. These vibrations are too low to be heard by the human ear, but depending on intensity, they can be felt by the body. Sources range from earthquakes to whale communication, to subwoofers for music.
- Acoustic sound, with frequencies in the band between 20 Hz and 20 KHz. This is the band which are detectable to the human ear, but the exact limits vary, mainly based on age but also on other factors. Sources include the human vocal cords, musical instruments, and generally everything.
- Ultrasound, with frequencies above 20 KHz. These waves are inaudible to humans, although longer exposure can result in health risks, so they certainly do interact with the human body. Some animals use ultrasound for navigation, such as bats, and dog whistles produce similar frequencies.

It is worth noting that there is nothing fundamentally separating these three categories in terms of physical characteristics.

Ultrasound today has a wide range of uses, and although medicine might be the most well-known and prominent, it is hardly the only one. In nondestructive examination (NDE) industrial products are screened for internal cracks by using ultrasound. Any cavity would cause rippling and unpredictable effects in the propagating sound, and it is therefore possible to detect such anomalies without cutting up and exposing the area in question.

The high-frequent vibrations can also be used for cleaning items such as jewelry or parts of watches. The items are submerged in a liquid, and the sound waves give rise to small jets, which efficiently clean the item, without causing the damage a traditional cleaning method might have caused.

SONAR or similar detection techniques are found in animals such as bats, and also onboard submarines and other boats today. They emit a sound pulse, and listen

for an echo afterwards. If the time between the original signal and the echo is short, the object they are aiming at is close nearby, and if it is long, the distance is greater.

Finally, although this list is anything but complete, there are also great opportunities for ultrasound within medicine. The harmless nature of these waves, and the relatively low costs of equipment make ultrasound a useful first screening tool when locating cancer tumors or when checking the growth of a fetus in a pregnant woman. Here, the principle is the same as with echolocation, in that a signal is sent, and then the nature of the echo tells the radiologist about the physical characteristics of the tissue present.

Ultrasound can also be used for therapeutic purposes, since it can elicit vibrations in otherwise hard-to-reach tissue. This can raise temperatures which could kill cells, it could also increase transmissibility of the blood-brain-barrier, which would then increase uptake of substances, such as medicine, in the brain.

In such cases it is important that energy is only delivered to the desired regions. Raising the temperature enough to kill cells in different parts of the brain could have catastrophic consequences, and the need for an accurate focusing of the ultrasound is needed. In the simplest case, it is sufficient to send pulses staggered in such a way that they all arrive in the designated focus at the same time, but given the sometimes complicated geometry, this is not always enough. In such cases, more powerful focusing algorithms are needed, capable of adjusting for variable sound speeds in different regions of the medium, and the scattering that occurs at the borders between these regions. The aim of this thesis is the investigation of such methods, capable of compensating for the scattering caused by the cranium, or similarly sharp-edged inhomogeneities.

Workflow overview

Given that the thesis was carried out at a university-independent institute, the workflow was more fluid than the usual phases of a thesis normally found at a university. Before beginning, other minor tasks were done, in order to build familiarity with the software and methodology that was to be used. Then, work started with implementing all of the methods presented in this thesis more or less simultaneously, based on their respective original sources. When additional theory was needed, this was searched for in literature, with aid from my supervisor. There was thus no formal literature review phase. This work was evaluated in 2-week cycles, up to a point where the results were deemed satisfactory, and the writing process began.

Chapter 2

Theory

The first step towards developing models for how a phenomenon works, and to be able to predict its results, is to understand the physics behind. In this chapter, the wave equation is derived using a simple physical model, and this is then specified to the problem at hand.

2.1 Physical background

A simple model for the interconnectedness of a medium is that it consists of point masses connected with springs. In one dimension, this takes the form of a long array of springs, as seen in figure 2.1.

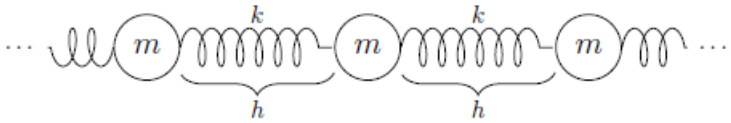

Figure 2.1: An illustration of the model of matter used for the derivation of the wave equation. Point masses m are connected by massless springs, each with spring constant k and length h.

Now, let the function $u(x)$ denote the longitudinal disturbance from the equilibrium for weight in point x, towards the right. Thus, if the mass at point x_0 is moved some distance α to the right, then $u(x_0) = \alpha$. The forces acting on the weight in an arbitrary point x can then be calculated as follows, with right being the positive direction

$$F_{prev} = k(u(x-h) - u(x)), \quad F_{next} = k(u(x+h) - u(x))$$

where subscripts indicate the forces corresponding to the previous and next springs, correspondingly.

Newton's first law then states

$$F = ma \quad \Rightarrow \quad k(u(x\text{-}h)\text{-}u(x)) + k(u(x+h)\text{-}u(x)) = m\frac{d^2u}{dt^2}$$

Rearranging gives

$$\frac{d^2u}{dt^2} = \frac{k}{m}\Big(u(x-h) - 2u(x) + u(x+h)\Big) \tag{2.1}$$

Now, the entire object is considered, with a total of N point masses. The total mass becomes $M = Nm$ and the total spring coefficient becomes $K = k/N$, as the total equivalent spring coefficient becomes smaller when multiple spring are placed in series[1]. Substitution of these new quantities into the previous expression yields

$$\frac{d^2u}{dt^2} = \frac{KN^2}{M}\Big(u(x-h) - 2u(x) + u(x+h)\Big) \tag{2.2}$$

Finally, the number of point masses N can be rewritten as L/h, where h is the total length of the object in question. Inserting this gives an expression of the well-known form

$$\frac{d^2u}{dt^2} = \frac{KL^2}{M}\frac{u(x-h) - 2u(x) + u(x+h)}{h^2} \tag{2.3}$$

Matter, at the scale that is relevant for this thesis, can be considered a continuous distribution of mass, and therefore the limit $h \to 0$ needs to be considered. Recognizing the definition of the second derivative, the final expression

$$\frac{\partial^2u}{\partial t^2} = c^2\frac{\partial^2u}{\partial x^2} \tag{2.4}$$

is reached. The constants KL^2/M are collected into one, c^2, for simplicity. The constant c, as it so happens, is the propagation speed of waves in this medium. With the inclusion of the spatial derivatives, the partial derivative notation is now also used.

This equation is known as the wave equation, and is the basis for the simulation tools used in this thesis.

2.2 Modelling ultrasound transducers as linear systems

The wave equation in the previous section is the natural way to completely describe any general sound propagation, but algorithms built upon it can be slow, especially as demands for precision and accuracy become higher. A simpler way of modelling the process is as a linear system, where the input consists of the sound that is sent out through each element in the transducer, and the output is the resulting pressure in our designated points, in practice also recorded with a transducer. These points would include the focus, but also points where the pressure ideally should be 0.

A core part of the model is that it is linear, which means that if two different inputs are entered at the same time, the output will be the sum of their respective outputs.

In addition, each sender affects each receiver, and a model like the one in figure 2.2 is the result.

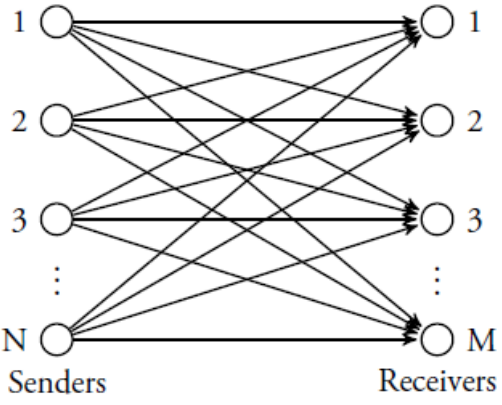

Figure 2.2: A schematic of the model. Inputs are entered into the senders on the left, which then travel along the arrows to the receivers and form the output. Each arrow symbolizes one $g_{m,n}$, a connection between a sender and a receiver.

Mathematically, this means the output f in a receiver m can be written as

$$f_m = \sum_{n=1}^{N} g_{m,n}(\nu_n) \tag{2.5}$$

which a sum of the contributions from each individual sender. The function $g_{m,n}$ is the connection linking the input ν_n in sender n with receiver m, and would in the simplest case include just delay corresponding to the travelling time, and a loss in amplitude due to dispersion, but could also include bouncing and transmission at medium edges, for instance.

The linearity allows not only for the splitting of the input into the different senders, but also into each infinitesimal time instant. The way a system responds to an input which is limited to one such time point is called the system's impulse response, and characterizes that entire part of the system. Summing these impulse responses back together is done via an operation called convolution, which for one receiver would become

$$f_m = \sum_{n=1}^{N} h_{m,n} \times \nu_n \tag{2.6}$$

with the cross indeed denoting the standard convolution operator. The time-dependence has been omitted for notational simplicity. This equation is the grounds on which the majority of the modelling in this thesis rests.

2.3 Modelling harmonic oscillations with complex numbers

Equation 2.6 in the previous section can be made simpler if the input is restricted to harmonic oscillations with some certain frequency ω. Then, assuming all transient processes caused by the input being switched on have died down, the output will also be an oscillation with the same frequency ω [2]. This can be succinctly expressed by using complex numbers. First, let the input be defined as

$$\nu_n = a_n e^{2\pi i\omega t}, \quad a_n = R_n e^{i\theta_n} \tag{2.7}$$

where R_n is the (real) amplitude and θ_n the phase delay of input ν_n. The variable a_n thus contains both the amplitude and phase delay, as its absolute value and argument, respectively. Then, the output becomes

$$f_m = b_m e^{2\pi i\omega t}, \quad b_m = W_m e^{i\psi_m} \tag{2.8}$$

where the b_m acts similarly to the a_n, but for the output oscillations. As this signal travels through the medium, the phase and amplitude changes, which can be captured in a complex variable $H_{m,n}$, known as the frequency response. The complete output in one point therefore becomes

$$f_m = \sum_{n=1}^{N} H_{m,n}\, \nu_n \tag{2.9}$$

Due to the multiplication being more allowing than the previous convolution, this can be collected and written as a matrix multiplication. With matrices defined as

$$F = \begin{pmatrix} f_1 \\ f_2 \\ f_3 \\ \vdots \\ f_m \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_m \end{pmatrix} e^{2\pi i\omega t} = Be^{2\pi i\omega t}, \quad E = \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \\ \vdots \\ \nu_m \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_m \end{pmatrix} e^{2\pi i\omega t} = Ae^{2\pi i\omega t} \tag{2.10}$$

and H as the matrix logically consisting of the elements $\{H_{m,n}\}$, the following formulation is possible:

$$F = HE \iff B = HA \tag{2.11}$$

where the second form is possible because the factor $e^{2\pi i\omega t}$ can be factored out. Logically, for two oscillations of the same frequency to be equal, all that is required is that their respective amplitudes and phases are equal, which is exactly what the second form describes. Because of its simplicity, this is the form that will be used throughout the later parts of the thesis.

Chapter 3

Methods

3.1 Calculation of the impulse response

In order to develop methods for the calculation of an optimal focus, knowledge of the impulse response is required. It is the link from input in the sender elements to output in the receiver elements, and our methods will then later aim to invert this connection, to instead go from a desired output to the input that would then be required.

3.1.1 Homogeneous media

The simplest case is when there are no obstacles or other sources of inhomogeneity in the medium. For this case it is quite simple to envision the appearance an impulse response would take, which makes it useful to create methods, that can then be expanded to situations where intuition becomes harder.

A simple, seemingly obvious, solution would be to simulate these using the same software that is then used for the verification. The main issue that arises with this strategy, is that a true delta-spike is impossible to emulate in a time- and space-discretized grid, and that the results therefore may vary heavily from what is expected, and furthermore might depend on simulation-related parameters.

A second option, that is also discarded, is to solve the wave equation analytically. When the input is a delta-impulse, the output is, quite trivially, the Green function of the wave equation. In two dimensions, this is

$$G(t,r) = \frac{1}{2\pi c\sqrt{c^2t^2 - r^2}}\delta(t - r/c) \tag{3.1}$$

and in three dimensions

$$G(t,r) = \frac{1}{1\pi r}\delta(t - r/c) \tag{3.2}$$

where r is the distance between sender and receiver. While this might be the most accurate solution in this case, it is impossible to generalize to heterogeneous media, where the wave equation becomes impossible to solve analytically.

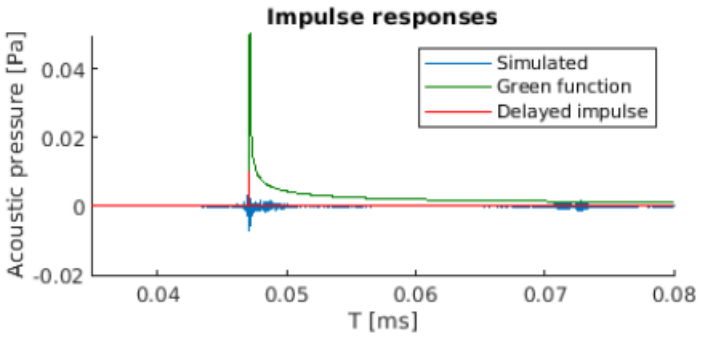

Figure 3.1: A comparison between the three described strategies for calculating the impulse response.

The final option, which is the one chosen, is the simplest: to only use a delayed impulse scaled by the dispersion factor. Sound intensity obeys an inverse square law, which means that the pressure, which is the root of the intensity, scales as $1/\sqrt{r}$ in two dimensions and as $1/r$ in three. Due to the unpredictable nature of a delta impulse in a discretized setting, the proportionality constant is simply determined with simulations.

A comparison of these three methods can be seen in figure 3.1, where $h_{64,64}$, the impulse response from sender 64 to receiver 64, was calculated.

In the figure, it can be seen that the simulated impulse response is quite different from what would be expected from theory, and has burst of noise at what would seem to be quite arbitrary points, for instance at t = 0.07 ms.

The theoretical approach, however, predicts a maximal amplitude far greater, than the one obtained in the simulations, whereas the delayed impulse obviously matches it accurately, given that it is scaled to do exactly that. Given the observations and the reasoning regarding scalability to more complex media, the delayed impulse becomes the strategy of choice moving forward.

3.1.2 Inhomogeneous media with straight borders

The case when the sound speed is constant in the entire region is not of any notable challenge and does not warrant any of the methods developed here, alone. The case when two different media are present however, becomes more interesting. The simplest case of this is pictured in figure 3.2, with two equally large regions of different sound speed.

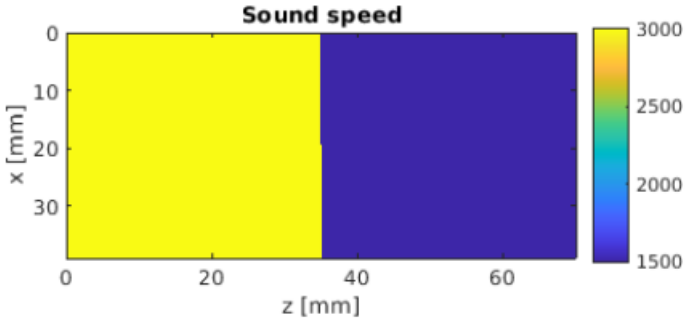

Figure 3.2: The sound speed in the first inhomogeneous medium considered.

Here, two different methods are devised, both utilizing the spherically emanating nature of the sound wave.

Firstly, a standard raytracing method, where a certain amount of straight rays are sent out from each sender element. If these rays miss the central border, they are disregarded, but if they hit, they give rise to a similar cascade of equally spread rays on the other side. When these then hit the receiver plate, the impact coordinate is rounded to the nearest receiver element, and is considered to have hit there. The total time is computed as the distance travelled in each region divided by the sound speed there.

The second method can be seen as a form of importance sampling. In Monte Carlo simulations, this would mean sampling not from the true distribution, but from another, which places higher probability on the outcomes that are considered important. These outcomes are then scaled down by their true probability.

In this way, the second strategy instead sends rays to points equally spaced on the border. From there, these are then sent to each receiver element. In both steps, the amplitudes are scaled down to match the true amplitudes that would come, if the wave was emanating spherically. An illustration of this can be seen in figure 3.3. It is easy to see that this factor becomes the area of the receiving element, as seen from the source, making the weights

$$W_\theta = \cos\theta \tag{3.3}$$

where θ is the angle of deviation from the center.

An example of how the resulting impulse responses for the homogeneous case, and then the two methods for the inhomogeneous case look, can be seen in figure 3.4.

3.1.3 Oblique region borders

The final case is when the region border, that separates the areas of different sound speeds, is oblique with some angle . There is no fundamental change to the structure of the methods used, but the calculation of the incident angle now has to also include the angle of the medium border. This changes the weighting factor to

$$W_\theta = \cos(\theta + \alpha) \tag{3.4}$$

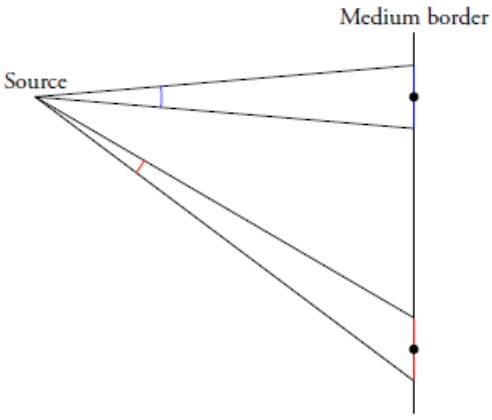

Figure 3.3: An illustrations showing the effects that cause larger angles of deviation to be weighted less. The red angle to the more distant border element is smaller, and it therefore constitutes a smaller part of the spherically emanating wave, even though the corresponding border elements are of equal size. When just considering the wave as rays travelling to the points, this effect is lost.

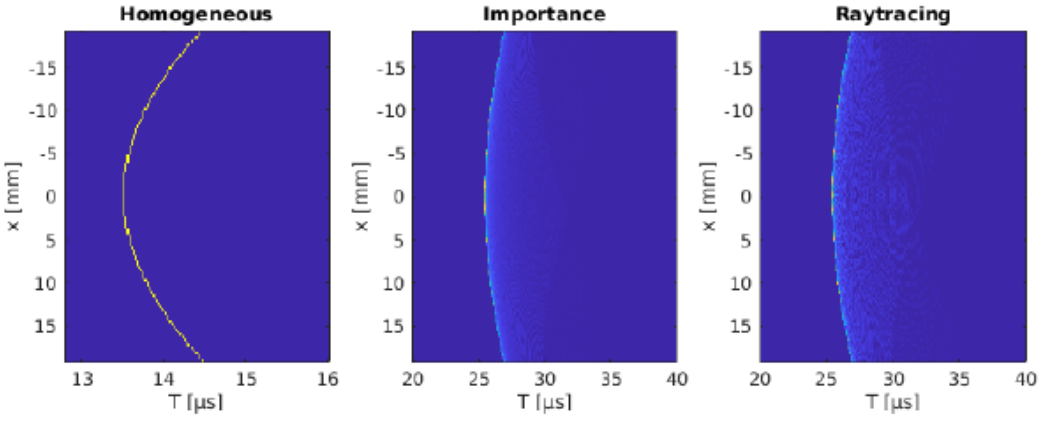

Figure 3.4: A comparison of the impulse responses for sender 64, in the middle of the sender array. In the first subplot, the impulse response for the homogeneous case is shown, and in the second and third, the response for medium 2, with the two different methods of calculation.

For the raytracing method, calculating the collision point on the edge now requires a small equation system to be solved, but it is hardly a difficulty.

3.2 Spatiotemporal inverse filter

With knowledge of how to calculate the impulse responses, a focusing algorithm can now be created. The fundamental equation that describes the system when the input is not necessarily harmonic is

$$f_m = \sum_{n=1}^{N} h_{m,n} \times \nu_n \tag{3.5}$$

By usage of the Fourier transform, the convolution becomes a simple multiplication

$$F_m\omega = \sum_{n=1}^{N} H_{m,n}(\omega)E_n(\omega) \tag{3.6}$$

where the dependence on ω as argument has been added for clarity. This can of course be written as a matrix multiplication

$$F(\omega) = H(\omega)E(\omega) \tag{3.7}$$

where, at least theoretically, only a matrix pseudoinverse is needed to compute E:

$$E = (H^T H)^{-1} H^T F \tag{3.8}$$

In fact, as the pseudoinverse essentially is a projection, this would result in an optimal E in terms of least squares. In practice, however, the matrix H tends to be ill-conditioned, and virtually rank-deficient. In other words, it has some amount of very small singular values. When inverting, these becomes very large, and what was once small errors in H, now becomes major deviations from the expected results.

3.2.1 Inverting the H matrix

Instead of the naïve approach with the pseudoinverse, a singular value decomposition is performed:

$$H = UDV^* \tag{3.9}$$

Here, U and V are unitary matrices of size $M \times M$ and $N \times N$ respectively, * denotes the Hermitian conjugate, and D is a matrix of the same size as H, $M \times N$, containing the singular values of H on the diagonal.

Now, all but the P largest singular values are set to 0, in order to avoid the ill-conditioned problem. P will be determined in the next section. This new matrix $\tilde{D}$ can be symbolically inverted by inverting each remaining singular value:

$$D = \begin{pmatrix} d_1 & 0 & 0 & 0 & \cdots \\ 0 & d_2 & 0 & 0 & \cdots \\ 0 & 0 & d_3 & 0 & \cdots \\ 0 & 0 & 0 & d_4 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \Rightarrow \tilde{D}^{-1} = \begin{pmatrix} 1/d_1 & 0 & 0 & 0 & \cdots \\ 0 & 1/d_2 & 0 & 0 & \cdots \\ 0 & 0 & 1/d_3 & 0 & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{3.10}$$

Here, P was 3, but this value will of course vary, and in general, be substantially larger.

Now that the small singular values that previously gave rise to unpredictable results have been removed, the new inverse $\tilde{H}^{-1}$ can be constructed:

$$\tilde{H}^{-1} = V\tilde{D}^{-1}U^* \tag{3.11}$$

That allows us to determine the optimal E:

$$E = \tilde{H}^{-1}F \tag{3.12}$$

3.2.2 Determining the number of relevant singular values

During the singular value decomposition, all but the largest P singular values were removed, but how this value should be determined has not yet been discussed. Three different strategies were tried:

For the first option, the matrix D represents all the independent ways in which the output can be varied in the specified frequency. This has been shown to correlate with the amount of possible sidelobes[3], or secondary maxima, capable of being emitted from the sender array, at that frequency:

$$P = \frac{2D_R}{\lambda} \sin\left(\arctan\frac{D_S}{2Z}\right) \tag{3.13}$$

The parameters D_R, D_S and Z relate to the geometry of the simulation as in figure 3.5.

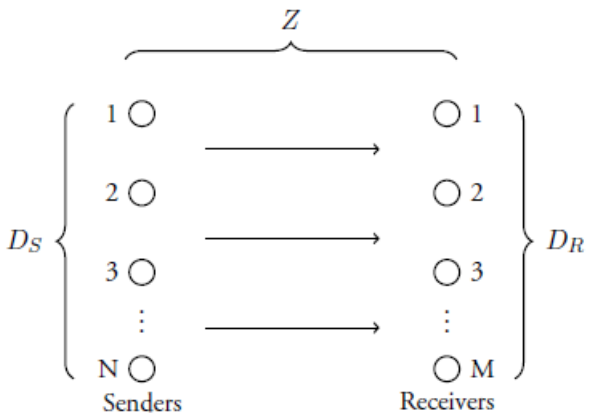

Figure 3.5: An illustration explaining the various geometrical parameters introduced in order to calculate the number of relevant singular values by the first proposed method. Note that in all of the cases investigated in this thesis, $DS = DR$.

To recover the value of λ, an investigation of the Fourier transform is needed. For H, the calculation was

$$H_{m,n}(\omega) = \sum_{t=0}^{\Omega-1} h_{m,n} e^{2\pi i \omega/\Omega} \tag{3.14}$$

Ω is here the total amount of timesteps in the simulation, and the amount of distinct frequencies that are part of the spectrum. The normalized frequency in this case becomes ω/Ω. This represents the number of cycles per sample, with aliasing when are above the Nyquist frequency. We can recover the actual frequency as

$$f = \frac{\omega}{\Omega} f_s = \frac{\omega}{\Omega \Delta t} \tag{3.15}$$

The spatial frequency is then controlled by the speed of the wave:

$$f_s = f/c_0 \Rightarrow \lambda = c_0/f \tag{3.16}$$

The second strategy is based on a simple threshold; all singular values that are above a certain magnitude are kept, and the rest are discarded. This has proven to be quite cumbersome in practice, since this threshold has to be determined for each case, and changing any geometrical parameter would require a recalibration of this value.

Thirdly, the energy of the matrix, which is the sum of the singular values, can be calculated. Then the smallest values are removed, until only some certain fraction of the original energy remains, for instance 95%.

3.3 Calculating the frequency response

Before moving on to the two remaining focusing methods, the matter of the frequency response must first be dealt with. Given that waves propagating are harmonic, some additional methods are possible, but the same method that was used for the impulse response also works here.

3.3.1 Raytracing methods

This is an modification of the techniques used for calculating the impulse response, which worked by sending out rays of sound, and seeing how they bounced, dispersed and eventually reached a receiver element. The same of course works for harmonic waves. In fact, attentive readers will notice that the impulse response can be used to calculate the output to any input, using a convolution. Therefore, the frequency response $H^{\omega}_{m,n}$ satisfies the following:

$$H^{\omega}_{m,n} \sin 2\,\pi\omega t \tag{3.17}$$

The reasoning can be made even simpler if only the phase delay and change of amplitude is considered. Then, if the input from some sender element n is

$$e_n = e^{2\pi\omega i t} \tag{3.18}$$

then in the homogeneous case, the output caused by this becomes

$$f_m = \frac{k e^{2\pi i \frac{d}{\lambda}}}{\sqrt{d}} e^{2\pi\omega i t} = H^{\omega}_{m,n} e^{2\pi\omega i t} \tag{3.19}$$

following previous reasoning about dispersion.

Extending this to inhomogeneous media is also quite simple. When there are no inhomogeneities, there is essentially only one ray that is relevant. With a border between sender and receiver, all that changes is that more rays are capable of producing outputs. Adding all of them gives an expression for the total frequency response.

3.3.2 Angular Spectrum

For media with either no inhomogeneities or with borders that are parallel with the sender array and thus at a right angle with the propagation direction, it is possible to calculate the frequency responses using the angular spectrum approach. This method consists of three steps:

Firstly, a spatial Fourier transform is done in the plane which is being propagated, perpendicular to the direction in which the waves are going. This produces what is known as the angular spectrum U:

$$U(k_x, 0) = \int_{-\infty}^{\infty} A(x, 0) e^{-ik_x x}\, dx \tag{3.20}$$

A is here, just as in the theory chapter, the complex vector containing phase and magnitude of the field, in this case at $z = 0$ which is at the sender transducer.

Then, this spectrum is multiplied with a propagation kernel

$$T = e^{iz}\sqrt{k_z^2 - k_x^2} \tag{3.21}$$

where z is the distance at which the wave front should be propagated, and the k are the wavenumbers in the respective directions, which is a scalar for z, and dependent on the frequencies in the Fourier transform for x. With this kernel, the angular spectrum can be evaluated at any z:

$$U(k_x, z) = \int_{-\infty}^{\infty} A(x, 0)e^{-ik_x x}e^{iz}\sqrt{k_z^2 - k_x^2}\, dx \tag{3.22}$$

All that remains is an inverse transformation to recover the field:

$$A(x, z) = \int_{-\infty}^{\infty} U(k_x, z)e^{ik_x x}\, dx \tag{3.23}$$

Now, this has some drawbacks that need to be combatted pertaining to the spatial Fourier transformation. It assumes the function is periodic, and thus, any wave exiting the simulation area through one of the sides, will reenter through the opposite side. This can be solved with either sufficient zeropadding to the point where the errors caused by this effect are negligible, or through a mirroring of the simulation area, with the new sources being negative. This would ensure perfect destructive interference on the borders and also remove this effect.

More egregious is, however, the inability to process oblique borders. Because the wave front needs to be parallel with the sender array and is propagated assuming equal sound speed in the entire front, it is simply not possible without major extensions to the method.

3.4 The Gerchberg-Saxton algorithm

In some cases, it is necessary to impose restrictions on the input, most commonly because of equipment limitations. For the case when the input needs to be harmonic, with some preset frequency and amplitudes, the Gerchberg-Saxton algorithm can be used to calculate the phase distribution that will give rise to some desired output.

As in section 2.3, the input is set to

$$E = \begin{pmatrix} a_1 e^{2\pi i f t} \\ a_2 e^{2\pi i f t} \\ a_3 e^{2\pi i f t} \\ \vdots \\ a_n e^{2\pi i f t} \end{pmatrix} = A e^{2\pi i f t} \tag{3.24}$$

The output then, as previously, becomes $F = Be^{2\pi i f t}$, with $B = HA$.

Now, only the amplitudes in A are actually relevant, since the phases are to be determined. Therefore, the phases are replaced with uniformly random values on the interval $[0, 2\pi]$. Then, this quantity is transformed using H, in order to determine the resulting output, changing both the amplitudes and phases.

The amplitudes are discarded and replaced with the desired output amplitudes, but the phase content is kept. This is then transformed back to the input plate using H^{-1}, the calculation of which will be discussed later. To now finish the iteration, the amplitudes are replaced with the preset input amplitudes $|A|$, and the process is repeated until some convergence criterion is met, in this implementation when the difference $||(n+1)\text{-}(n)||$ between one iteration and the next was under some threshold, after being normalized so that the first in each vector has delay 0. Given that the output phase is never specified, the result is only the relative phase distribution in the input plate, and the normalization step is therefore needed.

A flowchart of the algorithm can be seen in figure 3.6.

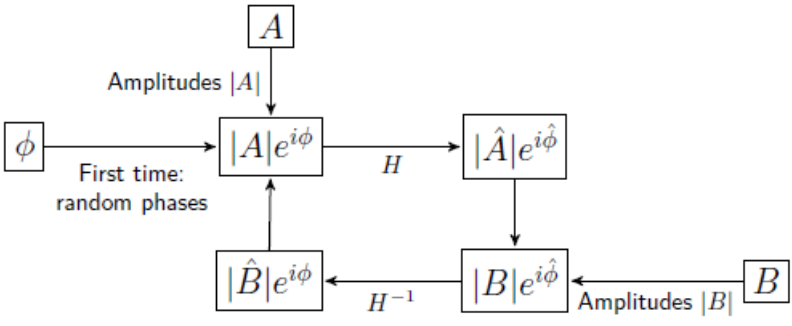

Figure 3.6: A flowchart detailing the Gerchberg-Saxton algorithm.

In the original implementation of the Gerchberg-Saxton algorithm, light scattering was studied, and there, the transform connecting the nearfield with the far-field

was instead the spatial Fourier transform $\mathcal{F}$ [4], for which there is well-studied inverse $\mathcal{F}^{-1}$. Given the geometrical differences between this case and theirs, the transform matrix H instead has to be used here, for which the inverse is not as obvious. Similarly to when the method for the inverse spatiotemporal filter was developed, the H matrix is often ill-conditioned, and the solution is the same now as previously:

1. Compute a singular value decomposition $H = U^*\Sigma V$
2. Threshold the singular values and set the sufficiently small ones to 0
3. Invert the nonzero ones and reform the inverse $H^{-1} = V^*\Sigma^{-1}U$

For non-quadratic matrices H, which arise when the number of senders N and receivers M are different, a least-squares problem will need to be solved in one of the directions. This can be done with the pseudo-inverse [5] and a similar as used here technique for circumventing the low condition number.

3.5 Optimization methods

An optimization-based approach is, at least in theory, very general. In this setting, it could be limited to three different cases:

- The inputs and outputs are not necessarily harmonic, similar to the problem solved by the spatiotemporal inverse filter.
- The system is harmonic, but both amplitude and phase of the inputs can be controlled. Only the amplitude of the output is relevant, not the phase.
- The system is harmonic, and only the input phases can be controlled. The output phases are once again, discarded.

The cases when the phases of the outputs are considered relevant could also be solved by the methods in this section, and would in fact make the mathematical derivations simpler, but for most practical purposes and actual applications they are not relevant. Therefore, their inclusion would only constrain the algorithm unnecessarily, and cause an inferior result.

In the first case, the spatiotemporal filter is already optimal in the least squares sense, and therefore, what will likely be a computationally more intensive optimization algorithm simply serves no purpose. Likewise, in the last case, the already developed GerchbergSaxton algorithm solves the problem quite quickly and with adequate results, although it is hard to reason if they are optimal or not[4]. Thus, the case with harmonic inputs with variable phases and amplitudes will be the focus of this chapter. The model will therefore be based on complex-number modelling and frequency responses as developed previously.

As before, the system can be summarized with the equation

$$B = HA \tag{3.25}$$

with absolute value and argument of the vectors A and B detailing the magnitude and phase of the sound field at the input and output, respectively. Here, however, the desired B is not a fully complex vector, but rather just real values corresponding to the wanted output magnitudes. Therefore, comparing B with HA makes little sense, and instead, an optimizing function should be constructed as

$$f(A) = g(|HA| - B) \tag{3.26}$$

The corresponding quadratic form is the standard choice, and is also selected in this case:

$$f(A) = (|HA| - B)^T W (|HA| - B) \tag{3.27}$$

The weighting matrix W is a square matrix with elements w_k only on the diagonal corresponding to the importance given to individual output elements.

3.5.1 Calculating the gradient of the penalty function

A prerequisite for many methods is knowledge of the gradient, which first requires the rewriting of the optimizing function as

$$f(A) = \sum_{m=1}^{M} w_m \Big(|\sum_{n=1}^{N} H_{m,n} a_m| - b_m \Big)^2 \tag{3.28}$$

Next, calculating the gradient with respect to am will not produce the wanted results, given its complex nature. Instead, the separation

$$a_n = \theta_n + i\gamma_n, \quad \theta_n, \gamma_n \in \mathbb{R} \tag{3.29}$$

is necessary. Now, the first step of the calculation can be carried out, using the chain rule for derivatives:

$$\frac{\partial f}{\partial \theta_k} = 2 \sum_{m=1}^{M} \left(w_m \Big(\Big| \sum_{n=1}^{N} H_{m,n} a_m \Big| - b_m \Big) \frac{\partial}{\partial \theta_k} \Big| \underbrace{\sum_{n=1}^{N} H_{m,n} a_m}_{S} \Big| \right) \tag{3.30}$$

The inner derivative requires some more attention, and can be computed by imagining it more geometrically. In the complex plane, the entire expression within the absolute value is simply a point, and as θ_k changes, this point moves in the direction of $H_{m,k}$. This can be seen in figure 3.7.

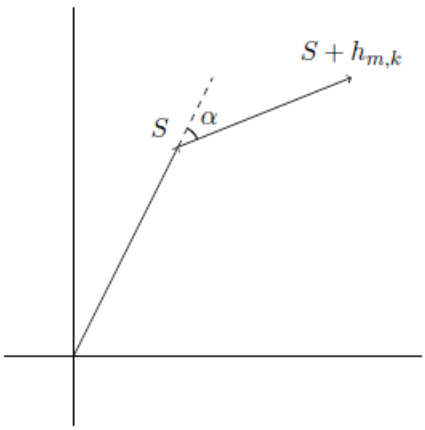

Figure 3.7: A geometrical illustration of the derivative calculation.

The movement of the total sum S as the differentiating variable θ_k changes can be split into two parts, one parallel with the vector S, and one perpendicular to it. The ratio of these parts is controlled by the angle α as seen in figure 3.7.

The perpendicular part just rotates the vector, and in the limit as this movement becomes infinitely small, this of course causes no change to the absolute value. Instead,

the parallel part is responsible for the change in absolute value of the vector S, which then logically ends up having the magnitude $|h_{m,k}| \cos \alpha$.

Thus, the total derivative becomes

$$\frac{\partial f}{\partial \theta_k} = 2 \sum_{m=1}^{M} \left(w_m \Big(\Big| \sum_{n=1}^{N} H_{m,n} a_m \Big| - b_m \Big) |h_{m,k}| \cos \alpha \right) \tag{3.31}$$

In the case of the γ_k, the calculation is very much the same, except there is an additional i to consider. As γ_k increases, the sum moves in the direction of $ih_{m,k}$, and with β denoting the angle between S and $ih_{m,k}$, the derivative can be written as

$$\frac{\partial f}{\partial \theta_k} = 2 \sum_{m=1}^{M} \left(w_m \Big(\Big| \sum_{n=1}^{N} H_{m,n} a_m \Big| - b_m \Big) |h_{m,k}| \cos \beta \right) \tag{3.32}$$

where the multiplication with i of course did not change anything in the absolute value $|h_{m,k}|$.

3.5.2 Gradient descent

As the aim of this section is not to develop the best possible optimization method, but rather, to verify its feasibility, the perhaps simplest method is chosen: gradient descent. It uses the prior knowledge of the derivative to continuously update an initial guess, by always stepping in the direction in which the function decreases the sharpest. One iteration can be written as

$$x_{n+1} = x_n - \eta \nabla f(x_n) \tag{3.33}$$

where η is a set value called the step size, indicating how far the method should go in the direction of the gradient. Because minimization is the goal, the step is always in the direction of the negative gradient. Various step sizes were tested without any major changes in result, and the final used was $\eta = 0.01$.

3.6 Testing procedure

The main results were compared for three different cases:

- **Focus 1**: a focus at distance $Z = 5$ cm from the transducer, centrally in front of the sender.
- **Focus 2**: a focus also at distance Z = 5 cm, but at x = –1.5 cm, i.e. not centrally in front of the transducer.
- **Focus 3**: two foci at Z = 5 cm, at $x = 1.5$ cm and x = –1.5 cm, respectively.

Each of these foci were tested for three different media:

- **Medium 1**: homogeneous water, with c = 1500 m/s and ρ = 1000 kg/m^3.
- **Medium 2**:a focus also at distance Z = 5 cm, but at x = –1.5 cm, i.e. not centrally in front of the transducer.
- **Medium 3**: two foci at Z = 5 cm, at $x = 1.5$ cm and x = –1.5 cm, respectively.

The full simulation parameters can be seen in table 3.1.

Variable	Value	Explanation
N	128	Number of sender elements
M	128	Number of receiver elements
s	300µm	Spacing between elements
Z	70mm	Distance between sender and receiver plates
Δx	100µm	Spatial discretization
Δt	50 ns	Temporal discretization
T	80 µs	Total simulation time

Table 3.1: A summary of the parameter values that were chosen for the simulations.

All tests were carried out using the open-source simulation software *k-wave*[6] version 1.3, and the methods were implemented in Matlab. The simulations were run on a HP laptop with an Intel i77500U processor at 4 × 2.70 GHz. The operating system was Ubuntu 20.04.

Chapter 4

Results

4.1 Spatiotemporal inverse filter

For the spatiotemporal inverse filter, the method for determining how many singular values to remove first had to be chosen, which was done using the first medium. The optimal output was considered to be a geometrically focused 2-sinus-burst in homogeneous water, and the reconstruction capabilities of the different variations on the filter were tested. Multiple frequencies between 1 and 20 MHz were tested to ensure that the higher performance of one method over another was not simply the cause of the specific parameters.

Then, all results for the three media, using the chosen technique for removing the smaller singular values, are presented. For the third focus, with two different foci, the respective inputs were simply added together.

The general testing procedure was here to first simulate a geometrically focused sender in homogeneous water, and then to use these results as the target for all future simulations. This process is described more thoroughly below. Then, if the filter could produce an equally good result in the presence of inhomogeneities, it could be said that it successfully compensates for them.

4.1.1 Comparison of singular value removal methods

For this first test case, only the first focus was used, and the distances to this element from each element in the sender array was calculated, as well as the corresponding travel times. The chosen input was this time a 2-sinus-burst at a variable frequency ω, or in other words the function

$$f(t) = \sin 2\pi\omega t, \quad 0 \leq t \leq \frac{2}{\omega} \tag{4.1}$$

This input was then staggered, so that all pulses would arrive in the focus at the same time, with the first inputs being sent at $t = 0$. This can be seen in figure 4.1.

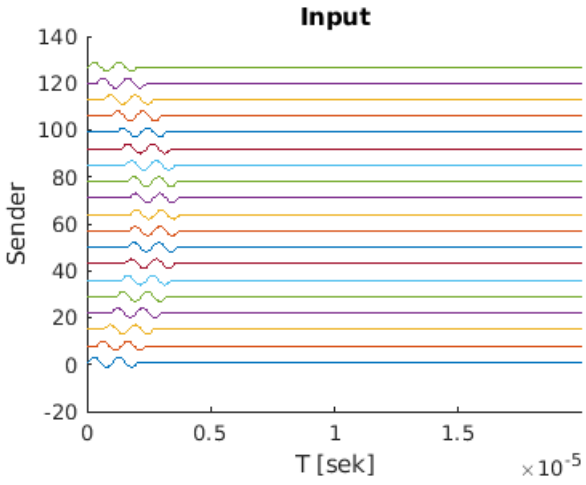

Figure 4.1: A plot showing the staggering of the input from the different transducer elements, in order to have them all arrive in receiver element 64 at the same time. The middle element travels the shortest distance, which results in the corresponding pulse being sent the last. Not all element inputs are shown, and the scale of the inputs has been chosen for visual clarity.

This simulation was then carried out using the MATLAB package *k-wave*[6], and the results can be seen in figure 4.2.

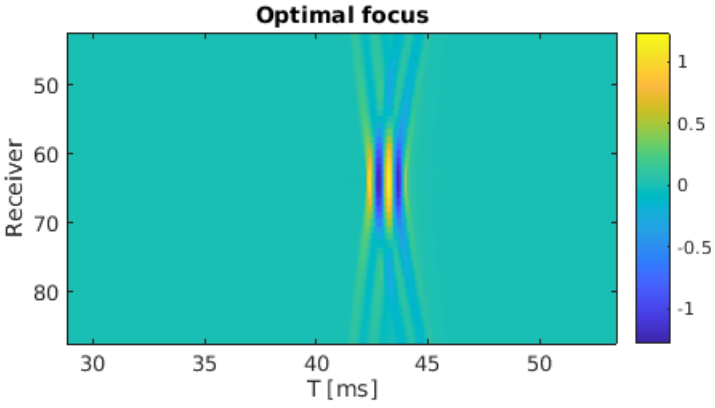

Figure 4.2: The results of the forward simulation with inputs as in figure 4.1.

The test was now to use this as the desired output and recreate the original input. The results of this can be seen in figure 4.3. The threshold used was 1, but multiple values were tested, and this was regarded the best in this case. Similarly for the energy-based method, multiple values were tested, and keeping 95 % was deemed optimal.

4.1.2 Inspection of the singular values

Before passing judgment on the three methods, a closer look at the actual singular values would also be prudent. In the ideal case, every frequency contains some amount of clearly nonzero singular values, and then some amount of negligibly small ones, and these are then, without ambiguity, separated by the algorithm of choice.

The distribution of singular values for two different frequencies can be seen in figure 4.4. They are based on the impulse responses from the test case used in the previous section.

With an eye to the results produced by the different strategies, and also the downside of the threshold having to be chosen specifically for each case, the energy-based approach is chosen. It is intuitive, universal, and easy to scale to more complex geometries or to three dimensions, which, for instance, would require a reworking of the frequency-dependent scheme.

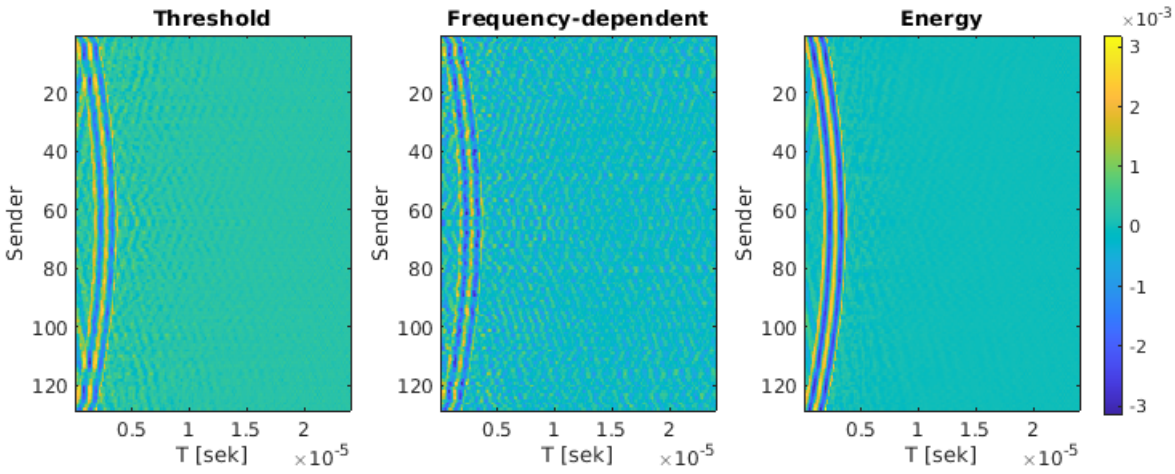

Figure 4.3: The result of each method for removal of small singular values. The third option is clearly the best at capturing the original input, with no major deviations during the initial parts, and no signals at all during the later parts of the process. The actual values, as indicated by the colorbar on the right do not take dispersion into account and are therefore erroneous on their own, but still useful for the comparison.

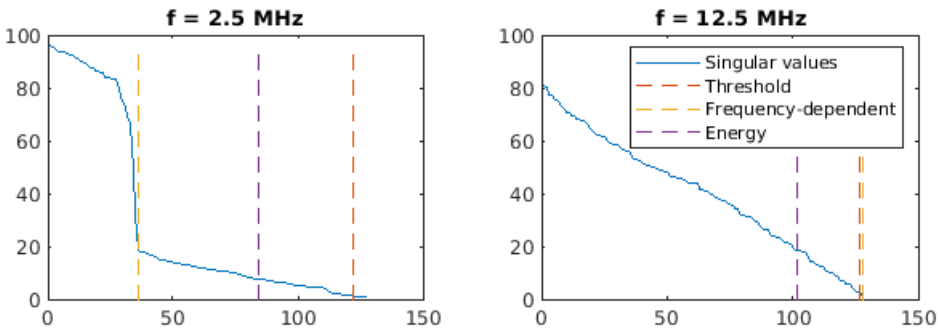

Figure 4.4: A comparison between the size and amount of singular values that the different methods include. In the second diagram, the two first essentially overlap, and both include all available singular values.

4.1.3 Medium 1

The first medium is simply the homogeneous case, so it is impossible to argue based on these results if the method can compensate for inhomegeneities in the medium. Nevertheless, they prove a valuable tool for validating that the methods work at all.

Furthermore, the different methods for computing the impulse response essentially become the same in this case, since there is no edge around which any kind of scattering could occur.

The results for the three foci can be seen in figures 4.5, 4.6, and 4.7, and it is clear that the geometrical foci are well reconstructed, and there are no major deviations, neither in calculated input or resulting output.

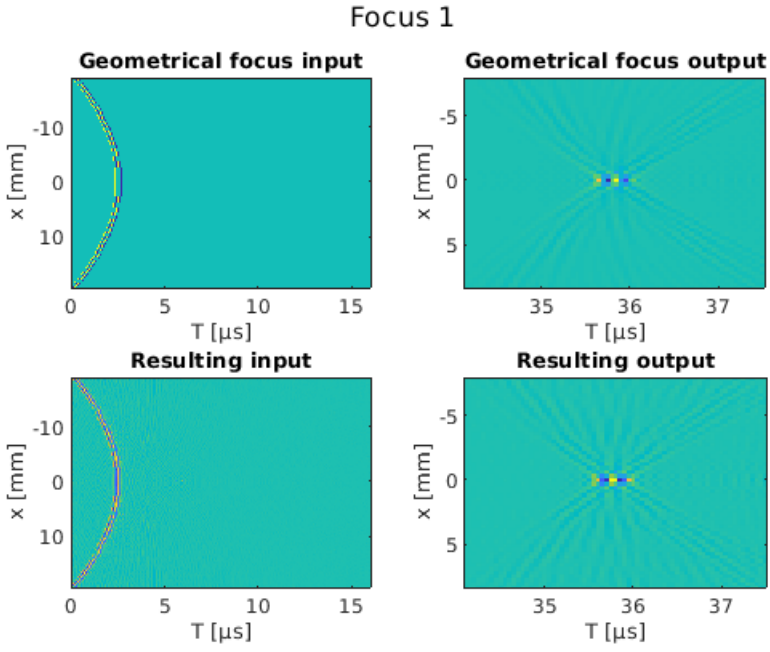

Figure 4.5: The results of the spatiotemporal inverse filter for focus 1 in medium 1.

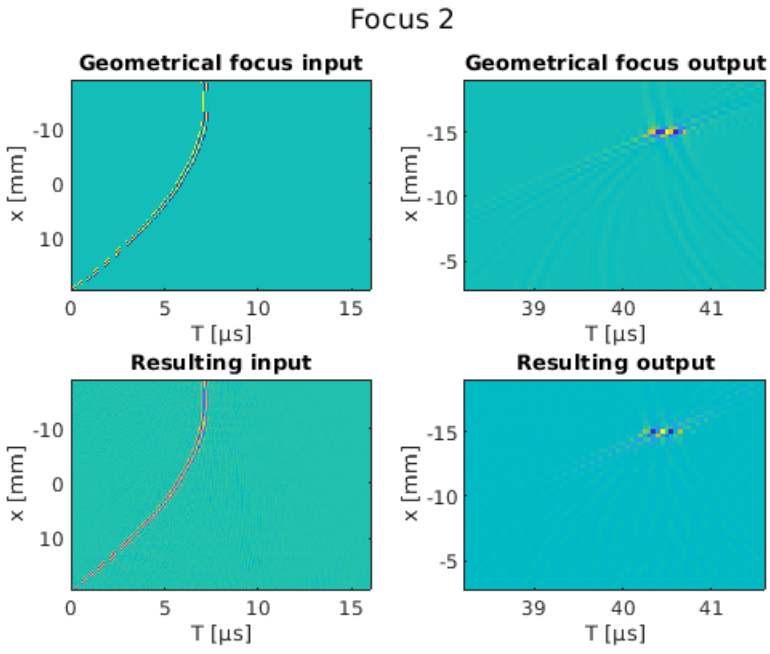

Figure 4.6: The results of the spatiotemporal inverse filter for focus 2 in medium 1.

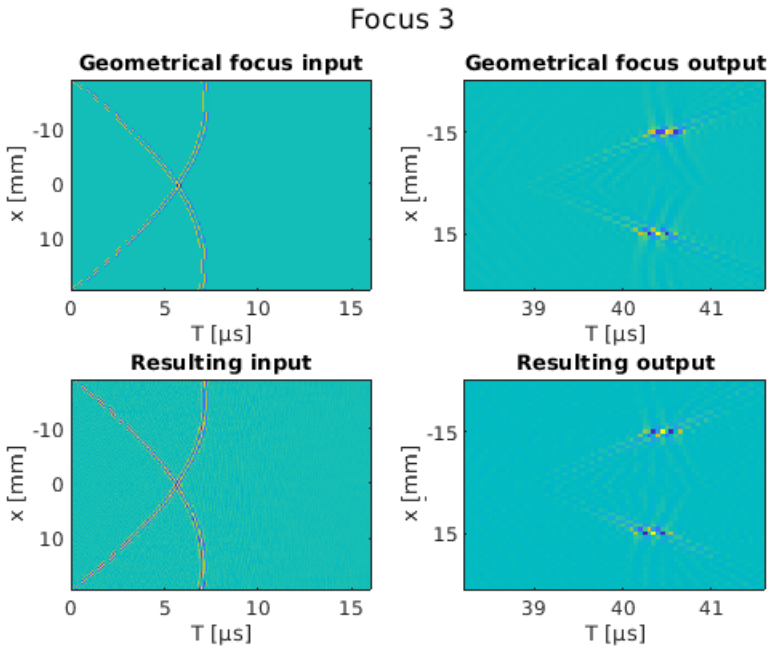

Figure 4.7: The results of the spatiotemporal inverse filter for focus 3 in medium 1. The rightmost plots have had the xaxis cut so that both foci are visible and zoomed in.

4.1.4 Medium 2

For the second medium, there are two schemes for calculating the impulse response to compare, the one based on importance sampling, labelled *importance* in the figures, and the raytracing-based method. These are visually compared for the different foci in figures 4.8, 4.9, and 4.10 below. In order to prove that they are indeed an improvement, the result from the geometric focus is also included, labelled *without compensation*.

As can be seen, there are no major differences between the two methods for this medium, but they are both improvements to the method, compared to simply disregarding the changing medium. Numerical examination similar using errors from the desired focus also did not manage to separate them noticeably.

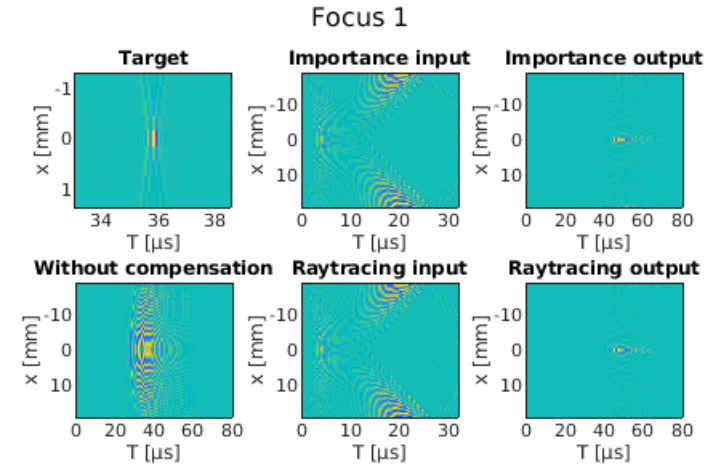

Figure 4.8: The results of the spatiotemporal inverse filter for focus 1 in medium 2, with the two different methods for calculating the impulse response pictured. The top-left target indicates the ideal focusing, that the methods are trying to replicate, while the bottom left shows the results when not taking the inhomogeneities into account.

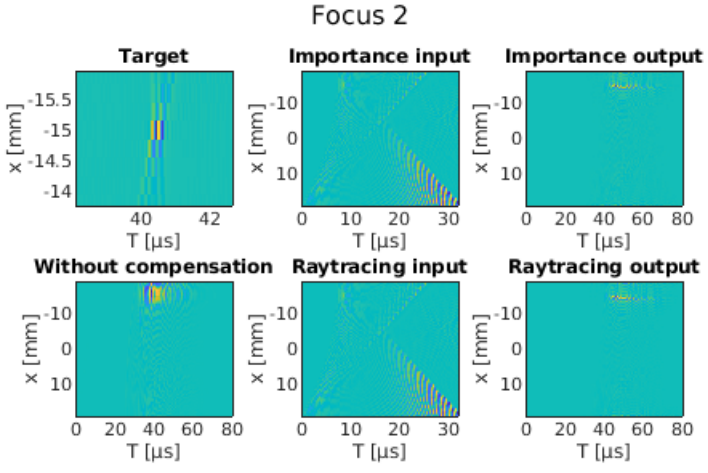

Figure 4.9: The results of the spatiotemporal inverse filter for focus 2 in medium 2, with the two different methods for calculating the impulse response pictured.

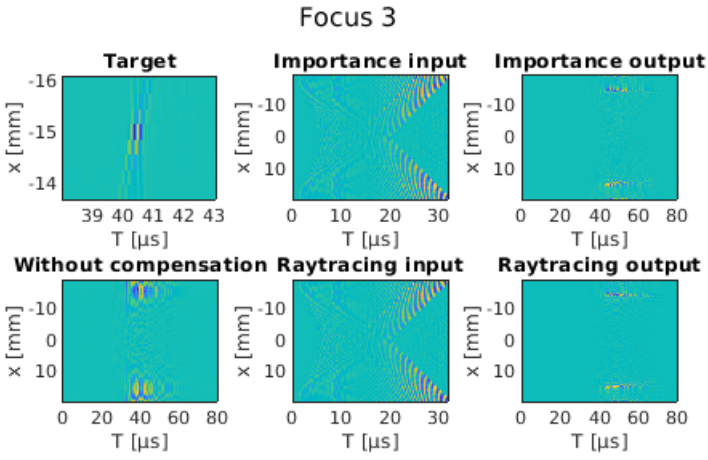

Figure 4.10: The results of the spatiotemporal inverse filter for focus 3 in medium 2, with the two different methods for calculating the impulse response pictured.

4.1.5 Medium 3

For the third medium, the two schemes for calculating the impulse response are once again compared, as with the previous section. These comparisons are shown in figures 4.11, 4.12, and 4.13 below, for the respective focus situations.

Once more, it proves difficult to discern the two methods, and visually, it is impossible to rank them.

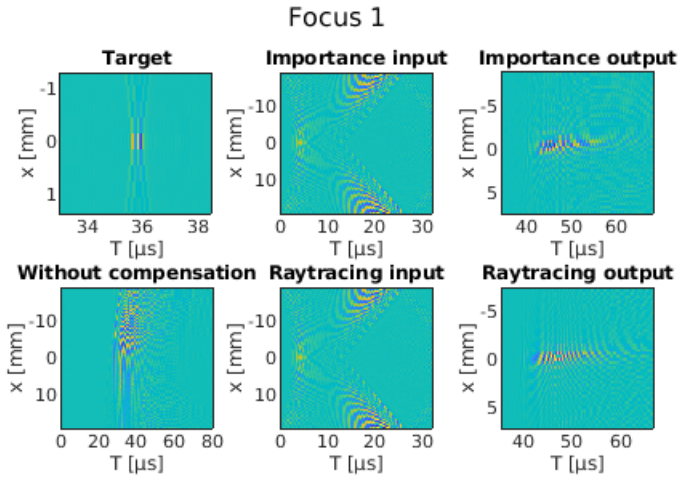

Figure 4.11: The results of the spatiotemporal inverse filter for focus 1 in medium 3, with the two different methods for calculating the impulse response pictured.

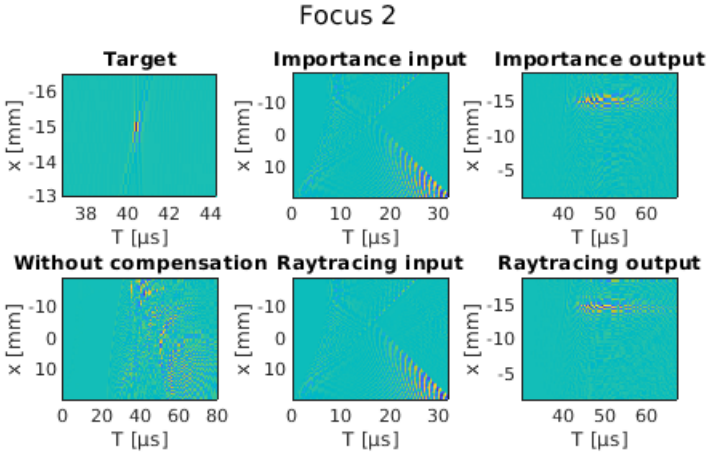

Figure 4.12: The results of the spatiotemporal inverse filter for focus 2 in medium 3, with the two different methods for calculating the impulse response pictured.

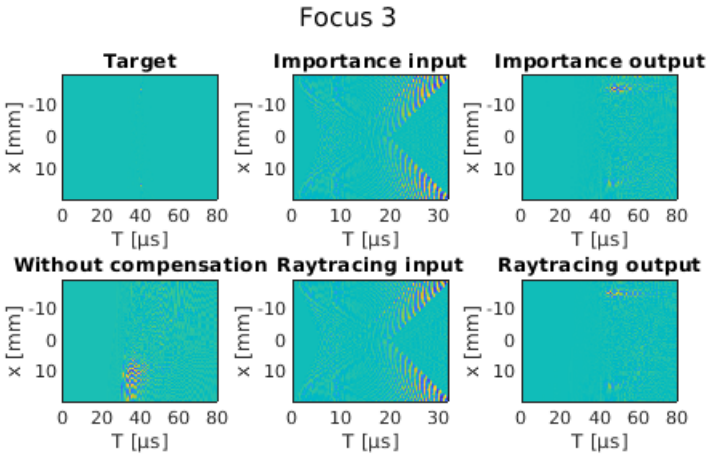

Figure 4.13: The results of the spatiotemporal inverse filter for focus 3 in medium 3, with the two different methods for calculating the impulse response pictured.

4.2 The Gerchberg-Saxton algorithm

With this method, there is no time-dependence, but rather, an infinitely stretching oscillation as input gives rise to a similarly infinite oscillation as output. Therefore, only the amplitudes are shown in the following graphs. The frequency was 5 MHz and there were no other changes compared to the previous section.

Furthermore, the target was once more set to be the result of a geometric focus, where every sender was delayed so that the oscillations reach the focus at the same time. In the third case, the inputs were first calculated for each of the two foci, and then the average was used.

As can be seen, the differences are quite small between the different variations, so the following performance measure e, the error in the 2-norm, was also used:

$$e = ||f_{target} - f_{method}||_2 = \sum_{n=1}^{N} \left(f_{target}^{(n)} - f_{method}^{(n)} \right)^2 \tag{4.2}$$

All the results were normalized by the condition $||f|| = 1$ prior to calculating the errors.

4.2.1 Medium 1

The first medium is completely homogenous, and therefore the result here just shows the ability of the method to recreate a focus that is completely attainable. Because there is no scattering, there is no need to calculate the frequency response matrix in any other way than the direct. As can be seen in figure 4.14, the target is recreated

well, with only some additional energy being sent to the middle when using the third focus.

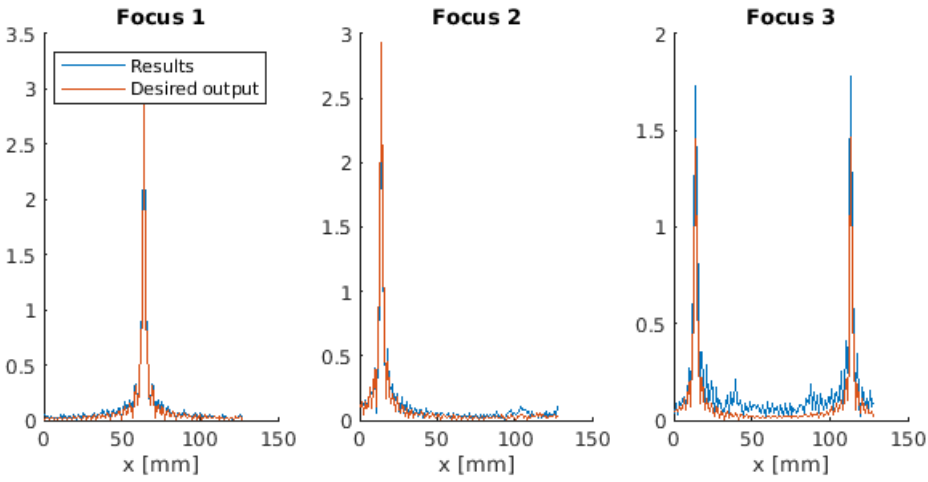

Figure 4.14: The amplitudes of the oscillations in the output plate for medium 1 after a focus had been calculated using the Gerchberg-Saxton method.

4.2.2 Medium 2

For the second medium, there are three methods for determining the frequency response to compare: the importance sampling method, denoted *importance*, the one based on raytracing and the method built on propagating the angular spectrum. These are visually compared in figure 4.15.

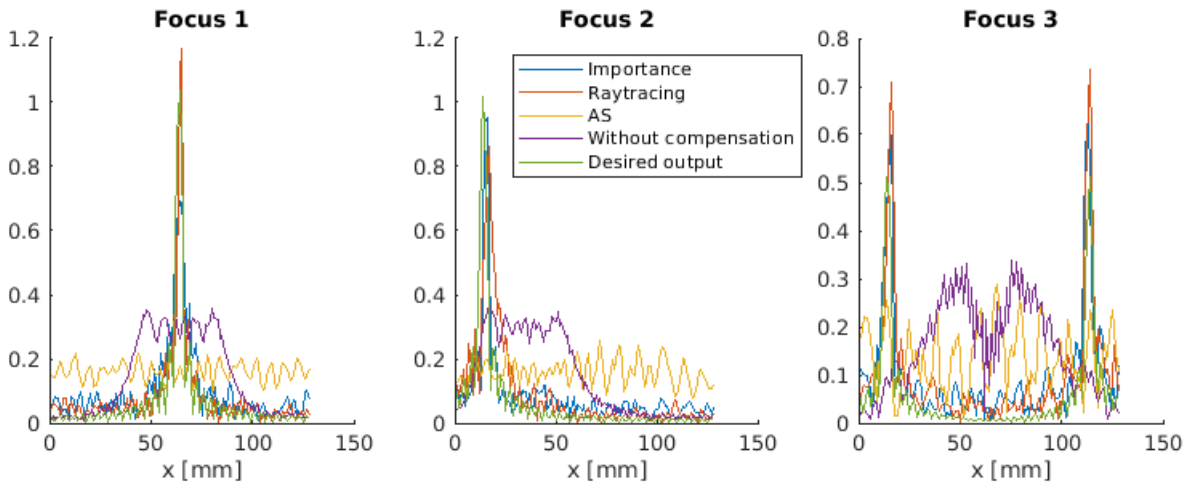

Figure 4.15: The amplitudes of the oscillations in the output plate for medium 2 after a focus had been calculated using the Gerchberg-Saxton method.

In table 4.1 below, the squared errors can be seen.

	Importance	Raytracing	AS	Uncompensated
Focus 1	0.45	0.40	1.10	0.95
Focus 2	0.63	0.77	1.10	0.99
Focus 3	0.44	0.45	0.98	1.21

Table 4.1: A table of the squared errors produced by the different methods for calculating the frequency response, when used with the GerchbergSaxton algorithm for medium 2.

4.2.3 Medium 3

For the third medium, there are only two methods for determining the frequency response to compare: the importance sampling method, and the one based on raytracing, given that the angular spectrum approach only works when the edges are perpendicular to the propagation direction. The two methods are visually compared in figure 4.16. In table 4.2 below, the squared errors can be seen.

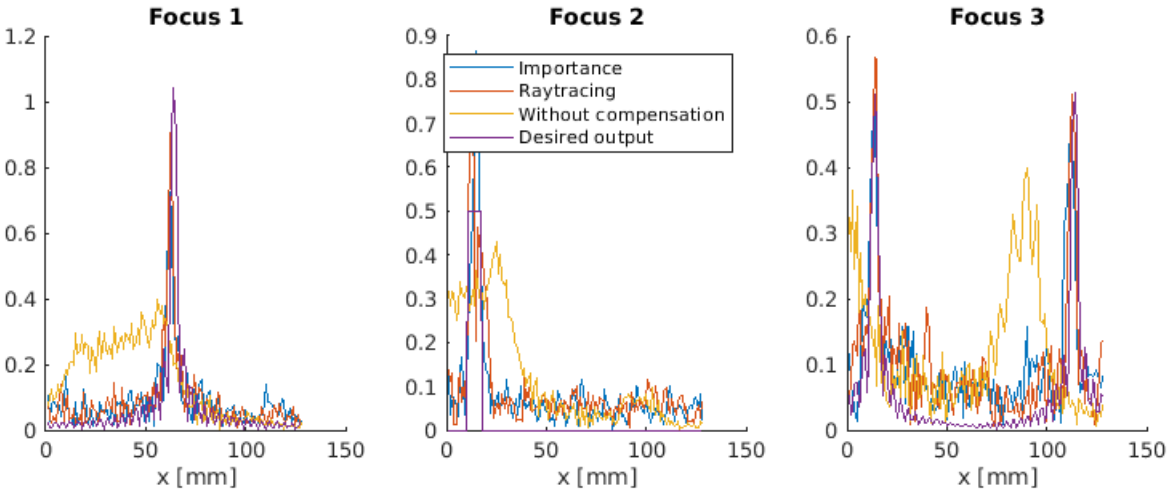

Figure 4.16: The amplitudes of the oscillations in the output plate for medium 3 after a focus had been calculated using the Gerchberg-Saxton method.

	Importance	Raytracing	AS	Uncompensated
Focus 1	0.55	0.49	1.11	0.95
Focus 2	0.65	0.65	1.10	1.02
Focus 3	0.55	0.56	0.98	1.01

Table 4.2: A table of the squared errors produced by the different methods for calculating the frequency response, when used with the Gerchberg-Saxton algorithm for medium 2.

4.3 Optimization-based methods

The results for the optimization-based methods are displayed similarly to those from the Gerchberg-Saxton algorithm, given that they are also harmonic, and that the amplitude is the only important output.

4.3.1 Medium 1

As can be seen in figure 4.17, the method is capable of recreating the focus in homogeneous media.

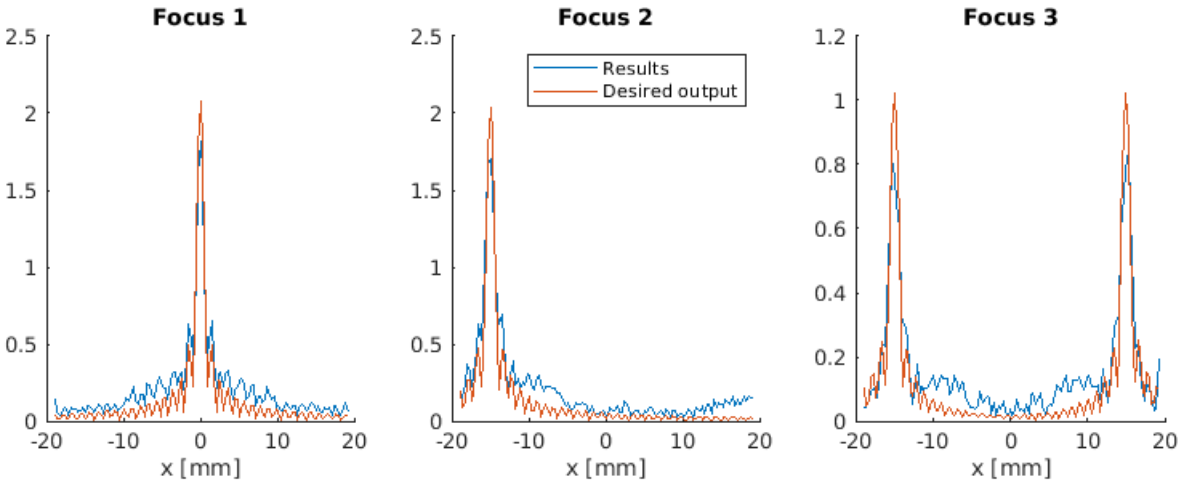

Figure 4.17: The amplitudes of the oscillations in the output plate after a focus had been calculated using the optimization method, for medium 1.

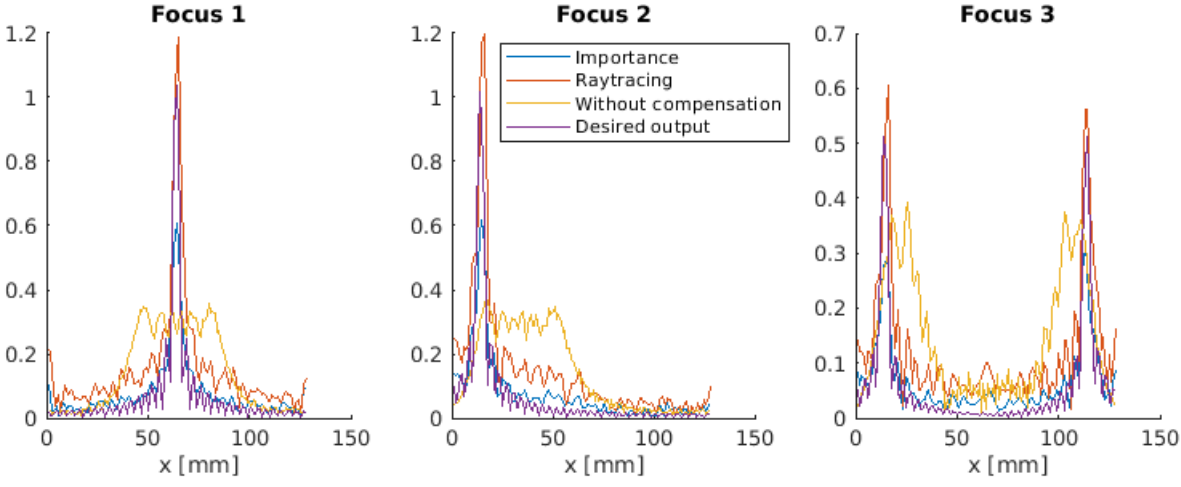

Figure 4.18: The amplitudes of the oscillations in the output plate after a focus had been calculated using the optimization method, for medium 2.

	Importance	Raytracing	AS	Uncompensated
Focus 1	0.40	0.42	1.08	0.95
Focus 2	0.40	0.50	0.93	0.99
Focus 3	0.43	0.46	1.03	0.86

Table 4.3: A table of the squared errors produced by the different methods for calculating the frequency response, when used with the optimizationbased algorithm for medium 2.

4.3.2 Medium 3

For the final medium, this approach struggled to even converge. With the first focus, all methods converged to a result, but it could hardly be described as accurate. For the second two the importance sampling method increased exponentially for the entire process, reaching values on the order 1010 during the allotted 100 iterations of the optimization loop. It has therefore been omitted from figure 4.19, where the others have been plotted.

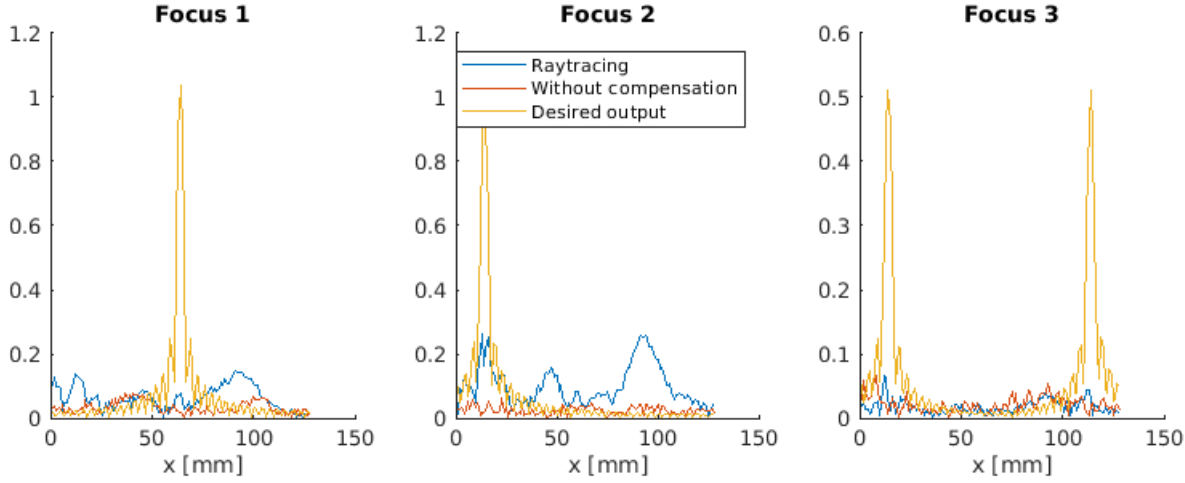

Figure 4.19: The amplitudes of the oscillations in the output plate after a focus had been calculated using the optimization method, for medium 3.

Chapter 5

Discussion

For one homogeneous medium, and two inhomogeneous ones, the presented methods were tested. Both the different schemes for generating impulse and frequency responses, and the subsequent focusing algorithms were tested in combination. For the impulse response, no major differences were found between the methods, and the spatiotemporal inverse filter using these was a significant improvement compared to not compensating for the inhomogeneities. For the frequency response, the angular spectrum method failed to produce useful results, but using the other methods, the Gerchberg-Saxton and the optimization algorithm was able to compensate for inhomogeneities. The optimization method failed to converge in one case.

5.1 Comments on results

At a cursory glance, all methods seem to produce results that at least at a general level follows the desired focus. Compared to not including any compensation, there is a significant improvement, justifying the use of these methods.

5.1.1 Spatiotemporal inverse filter

For the first focus, an essentially perfect recreation was achieved, and warrants no further attention. For the second medium, there was significant smearing of energy timewise, causing a longer pulse than intended. This is clearly not optimal, but in the case of focused ultrasound for killing cells, this can be compensated for with a lower amplitude[7], however given the nonlinear terms that do arise when modelling sound propagation in tissue[8], this might cause further errors in real experiments. The spatial accuracy is also lower than the desired focus, which carries with it more significant biological consequences, and limits the usage of this technique to cases where the target is 2 mm or larger.

For the third medium, the trend continues, and the focus is smeared timewise, causing an intended duration of less than 1 μs to last more than 10 times that. Here, the spatial focusing also worsens.

One likely explanation for the resulting outputs lasting longer than intended is the more smeared impulse responses. This can be seen in figure 3.4. Because the waves send from each sender hits the same edge centrally in the medium, the duration of these impulse responses are of the same order, and it is thus difficult to perfectly match them so that they cancel out in the nondesireable area in time and space.

In general, there is no major difference between the importance sampling method, and the raytracing method, and numerical investigation also failed to distinguish them particularly. Given that, as the number of rays increase, and the spatial discretization becomes finer, the two methods converge towards the same, it seems unlikely to matter. As can be seen in figure 3.4 there is still some difference with the parameter values used for this study, but they are minimal.

5.1.2 Gerchberg-Saxton and optimization

These methods might, at first sight, appear to perform better than the spatiotemporal inverse filter, but that is simply because there are no errors in timing, only in space. In general, they are able to reproduce the peaks well, but leak some more energy into the other regions than the ideal focus. This is, of course, excluding medium 3, where the optimization method did not produce a feasable result. For medium 2, however, the numerical errors were slightly lower for the optimization approach, but this can be explained by it optimizing over both amplitude and phase, and not just the latter.

For the first medium, there are no major differences, and both methods accurately capture the indended focus. For the second and third media, the need for compensation becomes evident. The uncompensated version produces a focus that is far wider than either method, and in the third case it is also misaligned.

Similarly to the spatiotemporal inverse filter, there does not seem to be any meaningful difference between the importance sampling method and the raytracing one. The method based on the angular spectrum, however, fails to produce any sensical results and is not better than the uncompensated result. This is likely because of the wrapping caused by the spatial Fourier transform, causing unwanted paths around the medium to the receiver. Even though this could be combatted with zero-padding, different amounts of zero-padding gave different results. In a normal simulation, there would be a limit where the waves simply would not reach around in time, but because the results are the steady-state solutions, this is not the case here. The mirroring approach also did not solve this issue completely.

Putting the results into context, the focus recreated by the Gerchberg-Saxton is approximately 1 cm wide, if the borders are put where the intensity drops below 10% of its peak. This is the scale at which brain tumours do appear[9], making it a potentially useful method for that application. This focus is a limitation of the transducer, and in order to accomodate a better focus, either more channels would be needed, or the distance to the focus would need to be decreased.

5.2 Possible improvements

Although the results are promising, the modelling is quite crude, and the possibilities for improvements are large. In general, these were seen as beyond the scope of this thesis, and in some cases, expensive or impossible to implement practically.

5.2.1 Improvements to impulse and frequency response calculation

Sound propagation and transmission at surfaces is quite complicated, and there are many factors that were not considered. Firstly, there are nonlinear terms when sound is travelling through tissue that would essentially invalidate the entire model.

Barring that, there are some parts of the sound transmission past the medium edges that are also not considered. A wave passing an medium border may give rise to surface waves, propagating on the edge[10], which of course direct some energy away from the main direction of propagation.

Otherwise, the modelling approach based on rays is a relatively sound one, and certainly an approach that is used in many other fields, even when the simulated field does not actually contain any rays or travelling particles, such as geophysics[11]. It has also been used for calculating specifically impulse responses in acoustics[12], albeit in a slightly more advanced framework.

The naïve solution of simulating the responses would in theory work well, but numerical errors cause the solutions to become quite inexact, as can be seen in the earlier figure 3.1. This then cascades to the calculation of the optimal inputs, and the end result becomes worse than the other methods. Simulating $N = 128$ senders is also time-consuming, as one simulation on an average computer took approximately 5 minutes, making the entire process last more than 10 hours.

The ultimate solution is of course to conduct physical experiments to determine the responses of the medium. This has been done [13] for human craniums, but even then, they had to fill the empty space with water, which admittedly does mimic most non-bone tissues relatively well in terms of sound speed, but certainly is not equal. Furthermore, one strength of a simulation approach is its generality; it is possible to perturb the model slightly, whereas it is impossible to test on different geometries than specifically those crania that are available.

5.2.2 Improvements to the focusing methods

The spatiotemporal inverse filter

The spatiotemporal inverse filter is, at its core, just solving a linear equation system for each frequency. In a least squares sense, this is then solved optimally, which makes the room for improvement zero. But, by use of zeropadding for the impulse responses, the resolution in the frequency domain could be made smaller, which could lead to a more accurate result. This is perhaps more so the case when the impulse responses are recorded through physical experimentation though, as the signals there would have energy in every frequency, in the general case. In a simulated setting, this can be

controlled to a greater extent, and everything would also be subject to the limit posed by the time resolution.

The different schemes for discarding the lowest singular values also deserves some attention. Previous research did provide a formula[14] for calculating the rank of the matrices that appears in the equation systems for each frequency when using the filter, and it does indeed correlate with a sharp drop at least for the first frequency, but numerical investigations showed that there was some energy left behind for others, as pictured in figure 4.4. This might however be a result of the idealized conditions in this thesis. The fact that the energy-based method produced a visually better result, as evidenced by figure 4.3, does speak for its usage at least in this setting though.

The Gerchberg-Saxton algorithm

The main problem with the method, as it is currently used, is that there is no proof of convergence. When using the Fourier transform, one can show that the squared error does not increase[4], but no similar proof exists for a general matrix transform. The fact that the optimization space is bounded to $[0, 2\pi]^N$ does mean that there is a upper bound for the error, but there is no guarantee that a solution is found, or that the method even approaches it. For the cases tested in this thesis, the method did however perform quite well.

Optimization

It is obvious that the optimization method could be improved, which is evidenced not only by the non-viable results for the last medium, but also by the fact that it consists of perhaps the simplest optimization routine possible. Any more modern optimization scheme would likely outperform this one in terms of speed, and most pre-written or proprietary algorithms would have measures against the divergence seen in the last medium.

5.2.3 Other improvements

All validation simulation were carried out using the open-source software *k-wave*[6] without any hardware acceleration or parallelization at all, as this was seen to be outside the scope of the thesis. Coupled with the fact that simulation times were approximately 5 minutes, this means that the spatial and temporal discretization could have been made much finer, which possibly could have resulted in a better result. As a means of comparing the methods used against each other, however, it seems sufficient.

With some parameter values, the simulations became unstable. This happened in particular with changes in density, which was not a problem for the three designated test media.

5.3 Parameter choices

In terms of parameters, all values are reasonable from a real-world perspective. A focusing distance of 5 cm does cover some of the applications of an ultrasound system, and

128 channels is quite common in modern equipment, even though larger arrays with 256 or more exist[15]. The transducer dimensions were also taken from equipment in use today.

A wide range of frequencies are used, varying with the application, but 5 MHz falls into the spectrum of commonly used frequencies. For focused ultrasound with tissue ablation as goal, frequencies of 0.5 - 1.5 MHz are more common though[7]. The frequency mainly impacts the result through nonlinear effects[8] though, which were not considered in this thesis. Therefore, the choice of frequency only affects the needed sampling rate and not the results directly.

5.4 Choice of medium

The choice of water as medium for ultrasound studies is a quite common one, as human tissue is mainly comprised of water, which means that they share approximate physical characteristics such as sound speed and density. Coupled with the facts that is a very practical and inexpensive material, and that a large amount of ultrasound literature uses the same medium, it was a sensible choice.

Choosing concrete as secondary medium, however, is slightly less standard and not as common in literate, in particular in the biomedical field. When using ultrasound for nondestructive testing though, concrete certainly has its place, and studies have shown that ultrasound is an effective resource for detecting subsurface cracks[16], for example. It is, similar to water, also a very practical and inexpensive material that can be cast or cut into any shape, making future experimental validations easier.

The speed of sound through the cranium is dependent on many factors, including bone density, degree of mineralization and of course general geometry. Previous studies have reported values ranging from 2200 to 3100 m/s[17], and there have also been attempts at modelling the speed of sound based on the bone density as detected by CT scans[18], with similar sound speeds reached for standard bone density values. This should be compared to the speed of sound in concrete used in this thesis, 2300 m/s. There are of course many difference in porosity between the two materials, and the water content present in bone is also expected to introduce some nonlinear viscoelastic effects. At the level of modelling in this thesis, however, the two materials are fairly similar.

5.5 Potential areas for further research

This thesis has been a study mainly of methods for focusing, and calculating system responses for simpler geometries. The largest step needed in order to approach real-world conditions would probably be moving into the third dimension. All three of the methods presented are capable of being extended to 3D, and the calculation of the system responses can also be done analogously. In general, it is possible to stack all sender points in one vector and all similarly with the receiver points, foregoing the perhaps natural matrix formulation given the 2D structure, in order to use the methods without any change. The verification simulations would likely take more

time, and the exponent in the dispersion term would increase by one, but otherwise there would be no major changes.

Even without leaving two dimensions, it would have been interesting to extend the receiver points into points not only on a line, but in the entire plane. Currently, when focusing, the regions behind and in front of the designated focus are not considered at all, and in cases where a tumor situated in sensitive tissue is to be ablated, this would certainly be cause concern. Like in the previous paragraph, the methods themselves do not need to be extended in any way to allow this, though.

More complex geometries could also be tackled with similar methods, by simply raytracing, and allowing the rays to scatter when colliding with an edge, regardless of the actual edge geometry, direction or curvature. With more advanced models, porous edges could be considered, and viscoelastic media could for instance also be simulated, further bridging the gap towards realistic models.

Chapter 6

Conclusion

In general, the results presented in the thesis were considered a promising first step towards more concretely applicable methods. In particular, the spatiotemporal inverse filter was able to compensate for the investigated inhomogeneities with arbitrary inputs, and the Gerchberg-Saxton algorithm was able to do so with time-harmonic inputs. The optimization approach failed in some cases, but this was assumed to mainly be because of the simplistic algorithm used. The test cases were admittedly simple, but proved that there is merit to using these algorithms, and it was argued that the way towards a more generalized approach is straightforward and requires no fundamental changes to the methods. Therefore, handling more complex geometries similar to those present in biomedical applications could be possible.

Bibliography

[1] Juan D Serna and Amitabh Joshi. "Studying springs in series using a single spring". In: *Physics Education* 46.1 (2010), pp. 33–40. DOI: 10.1088/0031-9120/46/1/003. URL: https://doi.org/10.1088%2F0031-9120%2F46%2F1%2F003.

[2] Sven Spanne. *Lineära System*. 1997.

[3] M. Tanter et al. "Optimal focusing by spatio-temporal inverse filter. I. Basic principles". In: *The Journal of the Acoustical Society of America* 110.1 (2001), pp. 37–47. DOI: 10.1121/1.1377051. eprint: https://doi.org/10.1121/1.1377051. URL: https://doi.org/10.1121/1.1377051.

[4] R. W. Gerchberg and W. O. Saxton. "A practical algorithm for the determination of the phase from image and diffraction plane pictures". In: *Optik* (1972).

[5] E. H. Moore. "On the reciprocal of the general algebraic matrix". In: *Bulletin of the American Mathematical Society* 26 (1920).

[6] URL: http://www.k-wave.org/.

[7] Jessica Foley et al. "Imageguided focused ultrasound: State of the technology and the challenges that lie ahead". In: *Imaging in medicine* 5 (2013), pp. 1190–1203.

[8] F.A. Duck. *Physical Properties of Tissue: A Comprehensive Reference Book*. Academic Press, 1990.

[9] Steven A. Goldman. *Overview of intracranial tumors neurologic disorders*. URL: https://www.merckmanuals.com/professional/%20neurologic-disorders/intracranial-and-spinal-tumors/%20overview-of-intracranial-tumors.

[10] Lord Rayleigh. *On Waves Propagated along the Plane Surface of an Elastic Solid*. 1885.

[11] N Rawlinson, Juerg Hauser, and Malcolm Sambridge. "Seismic ray tracing and wavefront tracking in laterally heterogeneous media". In: *Advances in Geophysics* 49 (Apr. 2008), pp. 203–273. DOI: 10.1016/S0065-2687(07)49003-3.

[12] Adil Alpkocak and Kemal Sis. "Computing Impulse Response of Room Acoustics Using the Ray-Tracing Method in Time Domain". In: *Archives of Acoustics* 35 (Dec. 2010). DOI: 10.2478/v10168-010-0039-8.

[13] Thomas Bancel et al. “Comparison Between Ray-Tracing and Full-Wave Simulation for Transcranial Ultrasound Focusing on a Clinical System Using the Transfer Matrix Formalism”. In: *IEEE transactions on ultrasonics, ferroelectrics, and frequency control* PP (Mar. 2021). DOI: 10.1109/TUFFC.2021.3063055.

[14] Z. Kopal, ed. *Proceedings of the Symposium on Astronomical Optics*. 1956.

[15] Christoph Risser. “High channel count ultrasound beamformer system with external multiplexer support for ultrafast 3D/4D ultrasound”. In: *IEEE International Ultrasonics Symposium* (2016).

[16] E. Leonidou D. Aggelis and T. Matikas. “Subsurface crack determination by onesided ultrasonic measurements”. In: *Cement and Concrete Composites* 34 (2012).

[17] Francis A. Duck. “Acoustic Properties of Tissue at Ultrasonic Frequencies”. In: 1990.

[18] Christopher Connor, Greg Clement, and Kullervo Hynynen. “A unified model for the speed of sound in cranial bone based on genetic algorithm optimization”. In: *Physics in medicine and biology* 47 (Dec. 2002), pp. 3925–44. DOI: 10.1088/0031-9155/47/22/302.