
3-D recursive extrapolation operators

Jan Thorbecke ¹

1.1 Introduction

To visualize the 3-D subsurface of the earth 3-D migration algorithms are needed which give accurate results within a reasonable computation time. In this Chapter several *recursive depth migration* algorithms are discussed and compared with each other. The backbone of every recursive depth migration algorithm is a 3-D extrapolation algorithm. In lateral homogeneous media the extrapolation algorithm can be a simple multiplication in the wavenumber domain, but extrapolation through 3-Dimensional inhomogeneous media is a more computation intensive operation and requires a space-variant spatial convolution. Recently various authors (Holberg (1988); Blacqui ere (1989); Hale (1991b); Soubaras (1992); Sollid and Arntsen (1994); Gaiser (1994); Biondi and Palacharla (1994) and Kao et al (1994)) have published articles which pay attention to an optimized calculation and efficient implementation of 3-Dimensional extrapolation operators in a recursive depth migration. This Chapter will give an overview of the existing methods and introduces several efficient optimization and implementation methods that have not yet been discussed in the Geophysical literature. The computation times of the different algorithms are compared with each other and the performance of the extrapolation algorithm is checked with the aid of a simple synthetic experiment.

1.2 Wave field extrapolation in the space-frequency domain

In laterally homogeneous media the recursive one-way extrapolation operator in the k_x, k_y - ω (wavenumber-frequency) domain can be well represented by the phase shift operator (Gazdag, 1978):

$$\tilde{F}(k_x, k_y, \omega, \Delta z) = \exp\left(-j\sqrt{\frac{\omega^2}{c^2} - (k_x^2 + k_y^2)}\Delta z\right) \quad (1.1)$$

¹E-mail: J.W.Thorbecke@ctg.tudelft.nl

with Δz being a small extrapolation step. The advantage of computation in the $k_x, k_y-\omega$ domain is that the desired result is obtained by *multiplication* of the data with the phase shift operator. However, to allow laterally varying medium functions a space variant *convolution* operator in the $x, y-\omega$ (space-frequency) domain should be used. When the spatial extrapolation operator is used in an explicit recursive depth migration algorithm it must be calculated in such a way that it gives reliable and stable results within a reasonable computation time. To arrive at this goal two steps must be taken; the first step is an optimum design of the spatial operator and the second step deals with a fast implementation of the spatial convolution. It turns out that the most efficient algorithms combine these two steps and design a spatial operator in such a way that it can be implemented in a fast way.

A **first** subdivision between the different optimization methods can be made with respect to the type of expansion of the analytical phase shift operator (equation (1.1)) in the wavenumber domain. This expansion can be a (Taylor) series expansion of the analytical phase shift operator with respect to $k_z = \sqrt{k^2 - (k_x^2 + k_y^2)}$ (this Chapter), an expansion with respect to $k_r^2 = k_x^2 + k_y^2$ ((Berkhout, 1982), (Soubaras, 1992), (Sollid and Arntsen, 1994), (Hoff, 1995) and this Chapter) or an expansion with respect to the cosine terms of the 1-D Fourier transformation (Hale, 1991a). In equations (1.4) to (1.6) these different approximation to the phase shift operator are shown

$$\tilde{F}_0(k_x, k_y) = \exp(-jk_z \Delta z) \quad (1.2)$$

$$\approx \sum_{m=0}^M \sum_{n=0}^N F_{mn} \cos(k_x m \Delta x) \cos(k_y n \Delta y) \quad (1.3)$$

$$\approx \sum_{m=0}^M F_m \cos(k_r m \Delta x) \quad (1.4)$$

$$\approx \sum_{m=0}^M a_m [k_x^2 + k_y^2]^m \quad (1.5)$$

$$\approx \sum_{m=0}^M b_m [k_z]^m \quad (1.6)$$

Equation (1.3) is the direct method which calculates the 2-D extrapolation operator F_{mn} by a direct optimization method and can be regarded as a weighted expansion in 2-Dimensional cosine terms. The direct method requires a full 2-Dimensional spatial convolution. Equation (1.4) is a reduction of the 2-Dimensional filter problem to a 1-Dimensional filter problem by using the circular symmetry of the 2-D operator and is represented by an expansion in 1-Dimensional cosine terms. The 1-Dimensional filter problem, to obtain F_m , can be solved with a preferred 1-D optimization method (Thorbecke and Rietveld, 1994). The 1-D cosine terms are approximated by short 2-Dimensional convolution filters.

Equation (1.5) and (1.6) are expansions of the operator in non-spectral polynomials. The terms a_m and b_m in the series expansions can be obtained by calculating the coefficients from a Taylor expansion (Berkhout, 1982), or optimizing the coefficients with an error definition in a preferred norm (for example the L_2 or L_∞ norm, see Appendix C). The basic polynomials $k_x^2 + k_y^2$ and

expansion	optimization	implementation
$\cos(k_x) \cos(k_y)$ $\cos(\sqrt{k_x^2 + k_y^2})$ $\sqrt{k^2 - (k_x^2 + k_y^2)}$ $k_x^2 + k_y^2$	Least Squares Weighted Least Squares Non-Linear Remez Exchange	2-D convolution 1-D convolution Chebyshev Series

Table 1.1 Three criteria which are used to discriminate between different 3-Dimensional extrapolation algorithms. Note that in principle many combinations between the elements in the three blocks are possible.

k_z which occur in equations (1.5) and (1.6) are approximated by short and accurate convolution operators.

A **second** subdivision between the different extrapolation algorithms can be made with respect to the kind of optimization method used to obtain the spatial convolution operator. The type of implementation of the spatial convolution is a **third** criterion to discriminate between the different methods. Table 1.1 gives an overview of the different techniques which can be used in the expansion, optimization and implementation. Note that in principle many combinations between the elements in the three blocks are possible. For example Holberg (1988) and Blacquièrè (1989) use a non-linear optimization technique for the operator optimization in a 2-Dimensional cosine series (weighted inverse Fourier transformation according to equation (1.3)) and have implemented this operator as a full 2-D convolution. Hale (1991a) makes use of the McClellan transformation in equation (1.4) and uses the coefficients of a 1-D convolution operators in a Chebyshev recursion scheme. Soubaras (1992) uses the Remez exchange algorithm in the optimization of the 1-D convolution operators and in the optimization of the expansion factors (with respect to powers of $k_x^2 + k_y^2$) of the phase shift operator in equation (1.5). In this Chapter the weighted least squares optimization method is introduced as a fast alternative method for the optimization of the 2-D convolution operators and in the optimization of the factors in the series expansions. The McClellan method is discussed in detail and several schemes are given which optimize the original McClellan method. The series expansions with respect to k_z and $k_x^2 + k_y^2$ given in equation (1.5) and (1.6) are worked out in detail and compared with the other extrapolation methods.

1.3 Direct method

The most straightforward method which does not make use of any series expansion is called, after Berkhout (1982), the direct method. In the direct method the optimization for the convolution operator is defined by the Fourier transformation from wavenumber to spatial domain and implemented by means of an optimization scheme (see Holberg (1988) and Blacquièrè (1989) for the results with a non-linear optimization scheme). The advantage of the direct method is the uncomplicated optimization of the operator and the simple implementation. A disadvantage of this direct method is that in the space-frequency domain the full 2-D convolution has to be carried out for every spatial position. For an operator with a typical operator size of 25×25

points this means 625 complex multiplications and summations for every grid point! By using the even symmetry in the operator the number of multiplications can be reduced by a factor of 4 by folding the data into a quarter and application of the convolution to this folded part. However the number of flops remains high, especially when we take into account that this convolution has to be carried out for every grid point, for every frequency of interest and for several depth steps. In this Chapter the results obtained with the direct method are used as a reference result for comparison with the other, non-direct, methods. In the remainder of this section three different methods, to calculate the 2-D convolution operators are discussed with respect to accuracy and efficiency and a synthetic migration experiment, to test the different extrapolation algorithms, is introduced.

1.3.1 Weighted Least Squares optimization

The most simple way to obtain space-frequency operators is an inverse Fourier transformation of the exact operators from the wavenumber-frequency domain back to the space-frequency domain. Despite of its simple form this solution is not very efficient because the spatial convolution operator obtained in this way must be very long to give stable and accurate results. Tapering the spatial operator gives some improvements (Nautiyal et al., 1993) but for accurate extrapolation results tapering cannot be used as pointed out by Thorbecke and Rietveld (1994). The aim in the design of the operator is a short convolution operator with a wavenumber-frequency spectrum which is, over a desired wavenumber band, equal or close to the phase shift operator in the $k_x, k_y-\omega$ domain. The starting point in the analysis of this optimization problem is the inverse Fourier transformation which is defined as

$$\tilde{F}(k_x, k_y) = \int \int_{-\infty}^{\infty} F(x, y) \exp(jk_x x) \exp(jk_y y) dx dy \quad (1.7)$$

Using the discrete version of the Fourier integral and the even symmetry in the phase shift operator equation (1.7) is rewritten in a discrete equation (after Blacquièrè (1989))

$$\tilde{F}(k_x, k_y) \approx \sum_{m=0}^M \sum_{n=0}^N S_{mn} F(m\Delta x, n\Delta y) \cos(k_x m\Delta x) \cos(k_y n\Delta y) \quad (1.8)$$

with S_{mn} defined as

$$S_{mn} = \begin{cases} 1 & \text{for } m = n = 0, \\ 2 & \text{for } m = 0 \vee n = 0, \\ 4 & \text{for } n \neq 0 \wedge m \neq 0 \end{cases} \quad (1.9)$$

Using the Circular symmetry in the operator by interchanging n and m and the fact that $\sum_{m=0}^M \sum_{n=0}^m = \sum_{n=0}^N \sum_{m=n}^M$ the number of equations can be further reduced to 1/8 of the original number of equations (this reduction is only possible if $\Delta x = \Delta y$)

$$\tilde{F}(k_x, k_y) \approx \sum_{m=0}^M \sum_{n=0}^m F(m\Delta x, n\Delta y) [S_{mn} \cos(k_x m\Delta x) \cos(k_y n\Delta y)]$$

$$+ S'_{nm} \cos(k_x n \Delta x) \cos(k_y m \Delta y)] \quad (1.10)$$

with S'_{nm} defined as

$$\begin{cases} 0 & \text{for } n = m, \\ 2 & \text{for } n = 0 \vee m = 0, \\ 4 & \text{for } m \neq 0 \wedge n \neq 0 \end{cases} \quad (1.11)$$

and $M \times N$ being the user specified size of the desired short operator. Using the shorter matrix notation equation (1.10) can be rewritten as

$$\tilde{\mathbf{F}}' = \mathbf{\Gamma} \mathbf{F} \quad (1.12)$$

with \mathbf{F} the desired short operator and $\tilde{\mathbf{F}}'$ being its spatial Fourier transform, yielding an approximation of the exact phase shift operator. Equation (1.8) with the quarter, or equation (1.10) with the octal part has to be solved for the unknown operator coefficients $F_{mn} = F(m\Delta x, n\Delta y)$ for all wavenumbers (k_x, k_y) of interest. Therefore a weighted error function $\tilde{\varepsilon}$ is defined as

$$\tilde{\varepsilon} = \tilde{\mathbf{E}}^h \tilde{\mathbf{\Lambda}} \tilde{\mathbf{E}} \quad (1.13)$$

with

$$\tilde{\mathbf{E}} = \mathbf{\Gamma} \mathbf{F} - \tilde{\mathbf{F}} \quad (1.14)$$

and $\tilde{\mathbf{\Lambda}}$ a diagonal matrix containing a weighting function on its diagonal. The introduced weighting function gives a good control over the desired functionality of the space-frequency operators. The least-squares solution of equation (1.13) is given by

$$\mathbf{F} = [\mathbf{\Gamma}^h \tilde{\mathbf{\Lambda}} \mathbf{\Gamma}]^{-1} \mathbf{\Gamma}^h \tilde{\mathbf{\Lambda}} \tilde{\mathbf{F}} \quad (1.15)$$

where $\mathbf{\Gamma}^h \tilde{\mathbf{\Lambda}} \mathbf{\Gamma}$ is a square matrix. The weighted least-squares method (abbreviated as WLSQ) can be used in the calculation of short (2-D) spatial convolution operators but also in the calculation of series expansion factors. In DELPHI Volume V (1994), Appendix A, the WLSQ method was used to solve the 1-D optimization problem. For the 1-D optimization problem the WLSQ method has another advantage; the matrix $\mathbf{\Gamma}^h \tilde{\mathbf{\Lambda}} \mathbf{\Gamma}$ which has to be inverted has a Toeplitz structure and can be inverted fast by using the Levinson scheme. For the 2-Dimensional problem standard LINPACK routines are used to calculate a QR decomposition of the matrix $\mathbf{\Gamma}^h \tilde{\mathbf{\Lambda}} \mathbf{\Gamma}$ and with this decomposition the solution of matrix equation (1.15).

In Figure 1.1 the wavenumber spectrum of a WLSQ optimized, 19×19 points spatial convolution operator is given for 128×128 k_x, k_y points with $c = 1000 \text{ m.s}^{-1}$, $f = 25 \text{ Hz}$, $\Delta x = \Delta y = \Delta z = 10 \text{ m}$ and a maximum angle of interest (α_{max}) of 65° . In the remainder of this Chapter these parameters will be used in all further examples which represent a phase shift operator. The WLSQ method gives an accurate operator which has a wavenumber spectrum close to the exact phase shift operator as shown in Figure 1.1. Note that due to the optimization on a rectangular

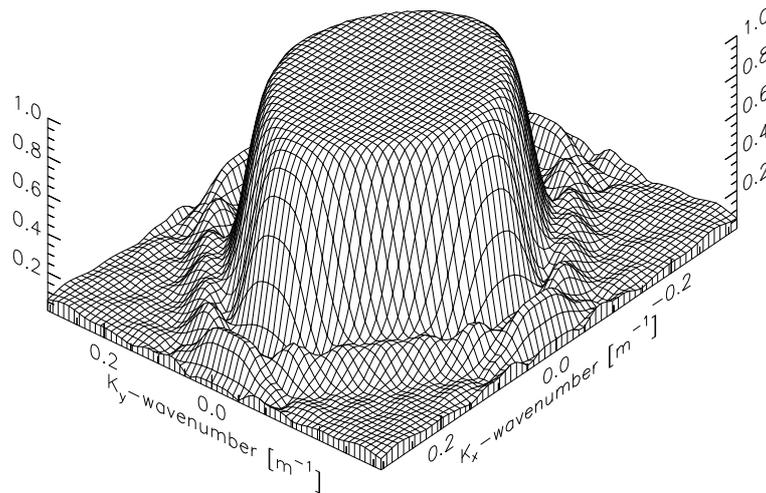


Fig. 1.1 The wavenumber spectrum of a WLSQ optimized operator with 19×19 spatial points, 128×128 k_x, k_y points with $c = 1000\text{ms}^{-1}$, $f = 25\text{Hz}$, $\Delta x = \Delta y = \Delta z = 10\text{m}$ and a maximum angle of interest of 65° .

grid the operator has a somewhat square symmetrical shape. If $\Delta x = \Delta y$ the matrix definition which uses the octal symmetry, given in equation (1.10), can be used in the implementation of the WLSQ solution. This scheme reduces the operator calculation time with a factor of 2 in comparison with equation (1.8) and the matrix problem contains less degrees of freedom so the unknown parameters are better defined.

Impulse response of an extrapolation operator

An impulse response experiment is used to test the behavior of the extrapolation operator in an explicit finite-difference migration algorithm. In the middle of a spatial limited homogeneous medium a point source is defined with the source signature shown in Figure 1.2. The zero-phase Ricker wavelet is centered at 0.512 seconds. The constructed data set is transformed to the frequency domain and extrapolated to deeper depth levels for every frequency of interest. At every depth level an imaging step is carried out and the depth image is stored in memory. At the end of the calculation for all frequencies the final depth image is written to disk. The block-scheme of this algorithm is shown in Figure 1.3. For the other extrapolation algorithms discussed in

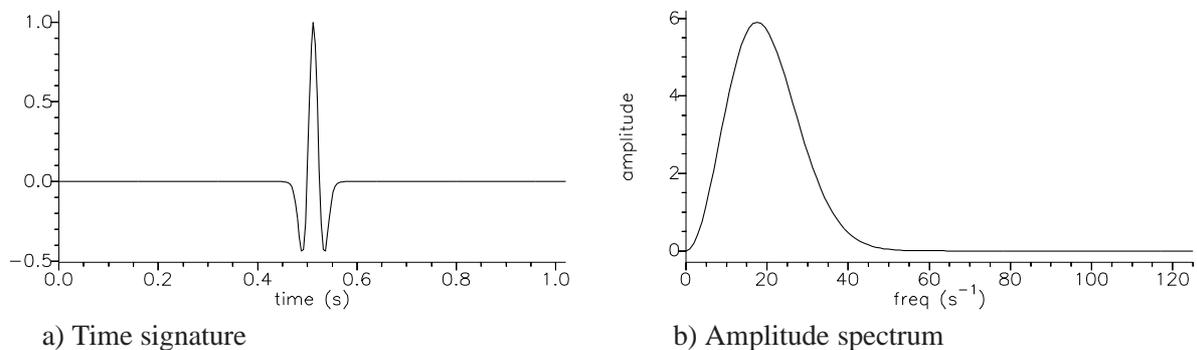


Fig. 1.2 a) Time signature and b) Amplitude spectrum of the wavelet used in the migration experiments.

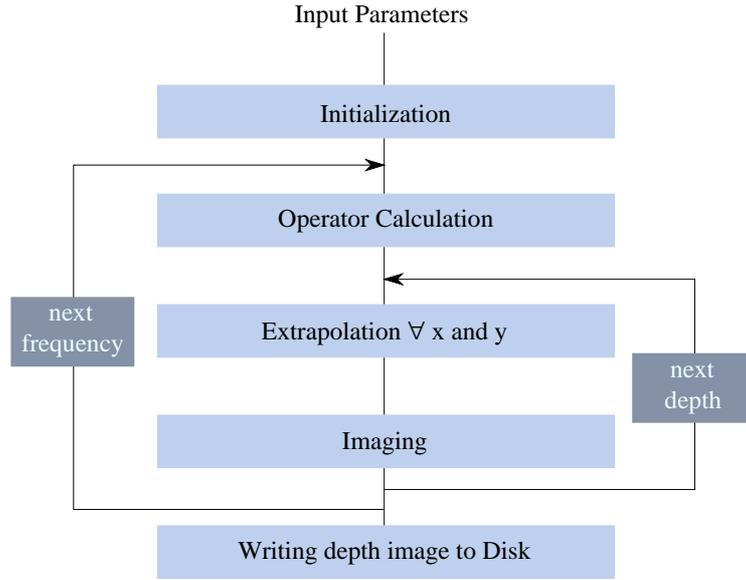


Fig. 1.3 Processing scheme for the impulse response experiment. Note that the extrapolation block in the scheme is different for every different extrapolation implementation.

this Chapter the extrapolation block in Figure 1.3 is replaced with the extrapolation algorithm of interest, everything else in the scheme remains the same. For the synthetic experiment the following parameters are used: $c = 1000\text{ms}^{-1}$, $f_{min} = 5\text{Hz}$ and $f_{max} = 45\text{Hz}$, $\Delta x = \Delta y = \Delta z = 10\text{m}$, $\Delta t = 0.004\text{s}$ and 55 depth steps are taken on a x,y grid of 111×111 samples wide. Note that the maximum frequency is positioned in the wavenumber domain at $0.9 * \frac{\pi}{\Delta x} (= k_N)$.

A reference output of this experiment can be calculated by using the exact expression of the extrapolation operator in the spatial domain. The exact spatial operator is defined by the dipole pulse response which is given by

$$G_0(r, k, \phi) = \frac{1 + jkr}{r^2} \cos \phi \exp(-jkr) \quad (1.16)$$

with $k = \frac{\omega}{c}$, $\cos \phi = \frac{z}{r}$ and $r = \sqrt{z^2 + x^2 + y^2}$. Using the complex conjugate G_0^* of the dipole response in a non recursive version of the scheme given in Figure 1.3 a reference impulse response can be calculated for the synthetic model described above.

How the zero-offset depth image is built up is shown in Figure 1.4. In this Figure the time responses for several depth steps are shown together in one picture. The imaging step at a certain depth level is equivalent with selecting the zero time value for all x- and y-positions. The lowest event in Figure 1.4 is the time response of the pulse after an inverse extrapolation step of 100 m, every higher event represents a depth level 100 m deeper. We see that for the deeper event the crossing with $t=0$ is converging to the $x=0$ position and will finally disappear if the depth exceeds 512 m. From this figure we can also derive that every depth slice corresponds to a certain dip angle. For example for a depth slice at 200 m the dip angle is given by $\cos \phi = z(ct_0)^{-1} \Rightarrow \phi \approx 67^\circ$. In Figure 1.5 three cross sections out of the 3-D depth image of the reference experiment are shown; the right pictures in Figure 1.5 show a vertical cross section for $x=0$

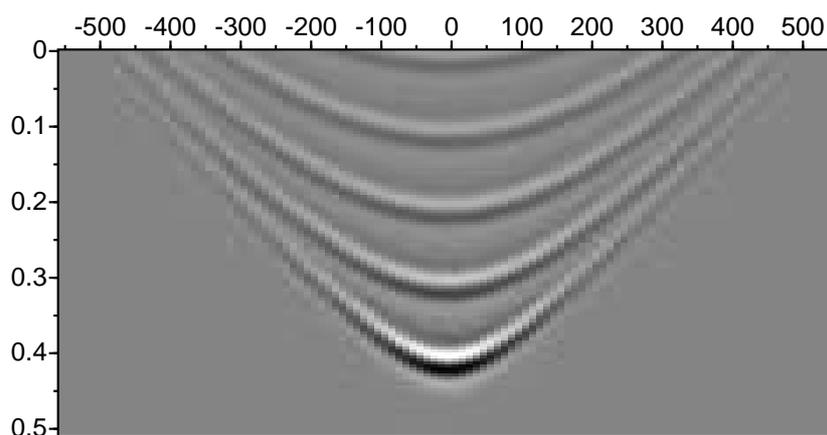


Fig. 1.4 Time responses for several depth steps. The lowest event is the time response of the pulse after an inverse extrapolation of 100 m every higher event represents a depth level 100 m deeper.

(top) and $x=y$ (bottom), the left hand side picture in Figure 1.5 shows a horizontal cross section at a depth of 220 m which corresponds to a reflector dip of 65 degrees. Note that everywhere in this chapter the presentation of impulse responses will be the same as in Figure 1.5.

Using 2 Dimensional 19×19 convolution operators obtained with the introduced WLSQ method gives the depth image shown in Figure 1.6a. In the spatial convolution scheme the even symmetry in the operator is used explicitly by folding the data into common operator point parts, which reduces the number of multiplications significantly. In the calculation of the convolution operator only 1/8 th of the total spectrum is used by making use of the circular symmetry and the fact that $\Delta x = \Delta y$. In Figure 1.6a we see that the artefacts in the depth image consists of inner 'circular' events at the higher angles which have a square structure. This square structure is due

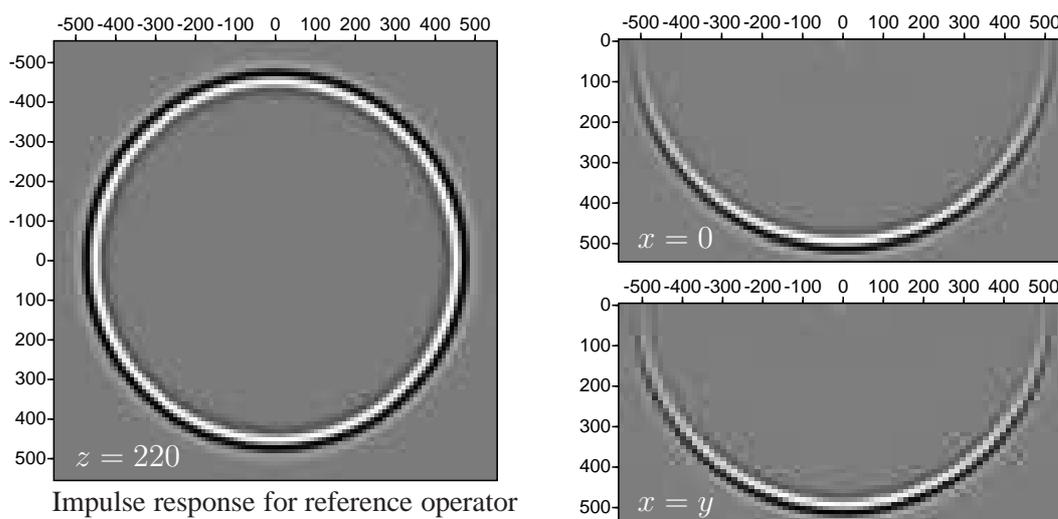


Fig. 1.5 Reference output for the migration experiment with left the depth slice at $z = 220$ m. Top right shows a vertical slice for $x=0$ and bottom right a vertical slice for $x=y$. Note the perfect circular shape and the accuracy at the higher angles.

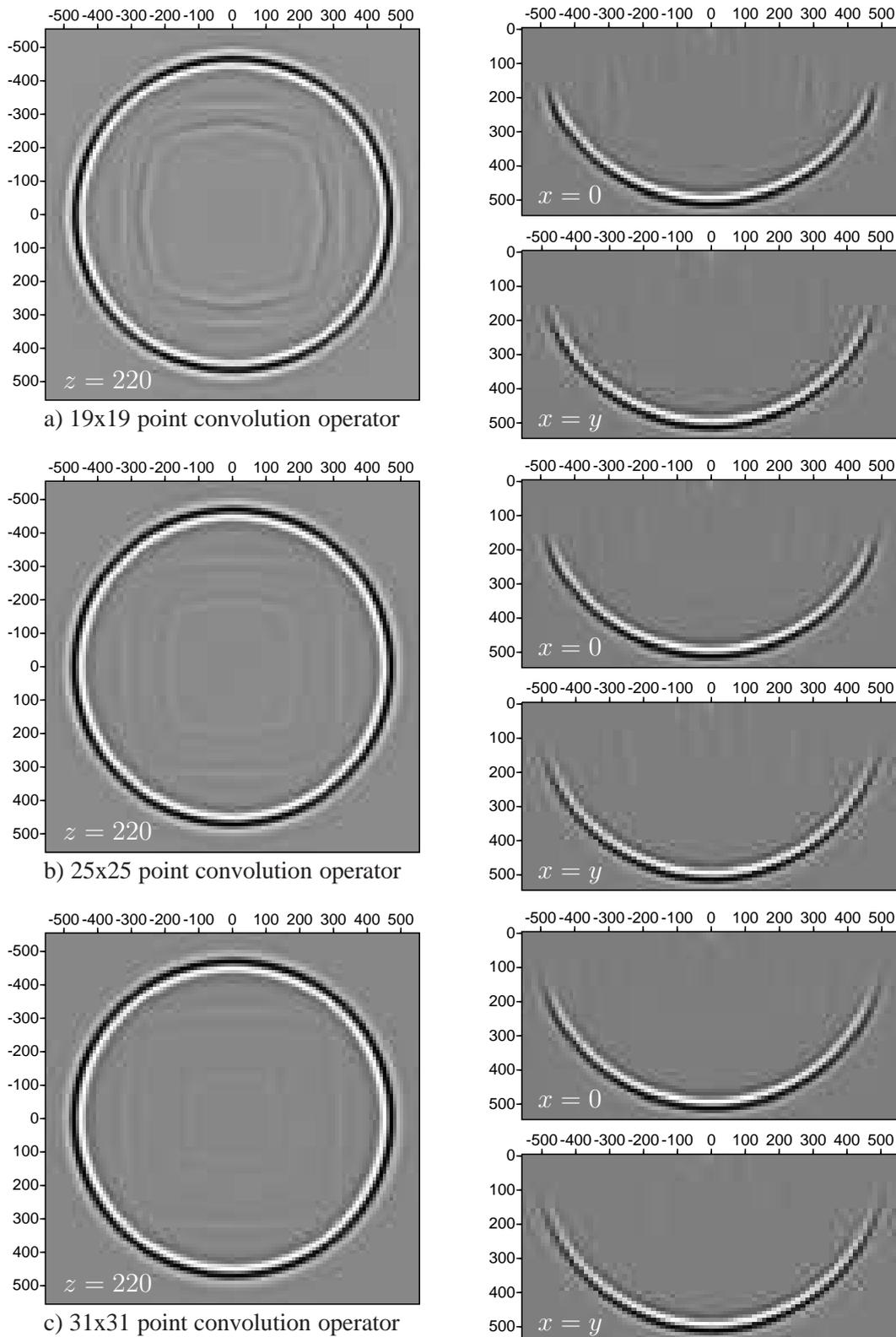


Fig. 1.6 Results of the WLSQ optimized operators for different operator sizes. Note that the artefacts which are present in the result for the 19x19 operator disappear for the larger operator sizes.

to the fact that the solution of the optimization problem is calculated on a rectangular grid. In the presentation of the paper of Kao et al (1994) similar features were observed. Using a longer 2-D convolution operator as shown in Figure 1.6b for a 25×25 and in Figure 1.6c for a 31×31 points operator these rectangular artefacts have vanished. A more detailed discussion of the errors in the extrapolation operators is given in the last subsection of this section. In Appendix A the computation times for the direct convolution method is given for different operator sizes on the different machines used in our research group.

1.3.2 Hankel Transformations

A disadvantage of the direct optimization method discussed in the previous section is that the use of a rectangular grid is displayed in the results and the circular symmetry of the operator is not used to its limits. By using the circular symmetry in the phase shift operator the operator optimization problem is better defined by using less equation to solve the unknowns. This may reduce the artefacts caused by the use of a rectangular grid and will consume less computation time. The circular optimization problem can be derived by rewriting the continuous Fourier transform pair

$$\hat{F}(k_x, k_y) = \int \int_{-\infty}^{\infty} F(x, y) \exp(jk_x x) \exp(jk_y y) dx dy \quad (1.17)$$

$$F(x, y) = \frac{1}{4\pi^2} \int \int_{-\infty}^{\infty} \hat{F}(k_x, k_y) \exp(-jk_x x) \exp(-jk_y y) dk_x dk_y \quad (1.18)$$

with the aid of polar coordinates. For a circular symmetric function in the wavenumber domain with $k_r = \sqrt{k_x^2 + k_y^2}$, $k_x = k_r \cos \theta$, $k_y = k_r \sin \theta$ and the Jacobian $dk_x dk_y = k_r dk_r d\theta$ the continuous inverse Fourier transform can be rewritten to

$$F(x, y) = \frac{1}{4\pi^2} \int_0^{2\pi} \int_0^{\infty} \hat{F}(k_r) \exp(-jk_r x \cos \theta) \exp(-jk_r y \sin \theta) k_r d\theta dk_r \quad (1.19)$$

$$F(x, y) = \frac{1}{4\pi^2} \int_0^{\infty} \hat{F}(k_r) k_r dk_r \int_0^{2\pi} \exp(-jk_r (x \cos \theta + y \sin \theta)) d\theta \quad (1.20)$$

Introducing $r = x \cos \varphi + y \sin \varphi = \sqrt{x^2 + y^2}$ ($x = r \cos \varphi$, $y = r \sin \varphi$) gives

$$F(r, \varphi) = \frac{1}{2\pi} \int_0^{\infty} \hat{F}(k_r) k_r dk_r \frac{1}{2\pi} \int_0^{2\pi} \exp(-jk_r r \cos(\theta - \varphi)) d\theta \quad (1.21)$$

With the definition of the zero order Bessel function as (Abramowitz and Stegun 9.1.18 (1968))

$$J_0(\tau) = \frac{1}{\pi} \int_0^{\pi} \cos(\tau \cos(\theta)) d\theta \quad (1.22)$$

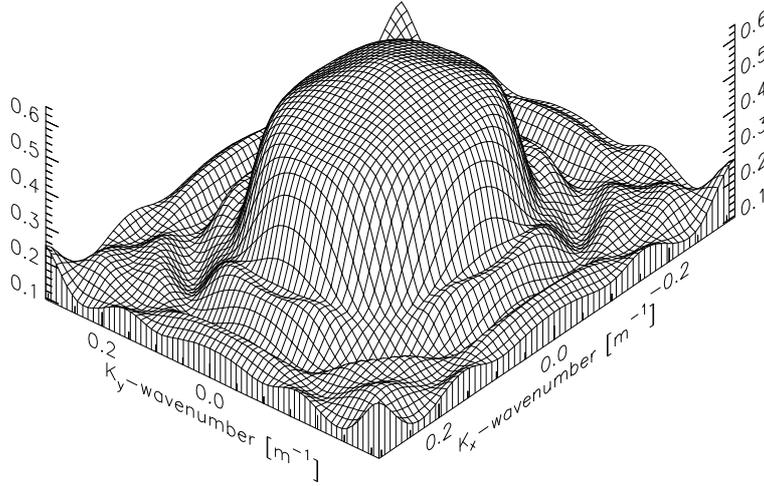


Fig. 1.7 The wavenumber spectrum of a circular WLSQ optimized operator with 19×19 spatial points and 128×128 k_x, k_y points with $c = 1000 \text{ms}^{-1}$, $f = 25 \text{Hz}$, $\Delta x = \Delta y = \Delta z = 10 \text{m}$ and a maximum angle of interest of 65° .

$$J_0(\tau) = \frac{1}{2\pi} \int_0^\pi \exp(-j\tau \cos(\theta)) + \exp(j\tau \cos(\theta)) d\theta \quad (1.23)$$

$$J_0(\tau) = \frac{1}{2\pi} \int_0^{2\pi} \exp(-j\tau \cos(\theta)) d\theta \quad (1.24)$$

substituted into equation (1.21) gives

$$F(r) = \frac{1}{2\pi} \int_0^\infty \hat{F}(k_r) J_0(k_r r) k_r dk_r \quad (1.25)$$

So the spatial convolution operator and its Fourier transformation are both circular symmetric and are related by the Hankel transform. Which is illustrated by the following formulations in polar coordinates

$$F(r) = \frac{1}{2\pi} \int_0^\infty \hat{F}(k_r) J_0(k_r r) k_r dk_r \quad (1.26)$$

$$\hat{F}(k_r) = 2\pi \int_0^\infty F(r) J_0(k_r r) r dr \quad (1.27)$$

The spectral limited and discrete version of the Hankel transform is given by

$$F(p\Delta r) = \frac{1}{N\Delta r} \sum_{n=0}^N \hat{F}(n\Delta k_r) J_0(n\Delta k_r p\Delta r) n\Delta k_r \quad (1.28)$$

with $\Delta k_r = \frac{2\pi}{N\Delta r}$. This last equation is implemented in the WLSQ optimization scheme of equation (1.15) in which the matrix $\underline{\Gamma}$ is defined by the zero order Bessel function $J_0(k_r r)$ in place of the cosine terms of the Fourier transform. The solution of this problem is an optimized short operator as function of r . In the optimization problem we can choose the points r in such a way

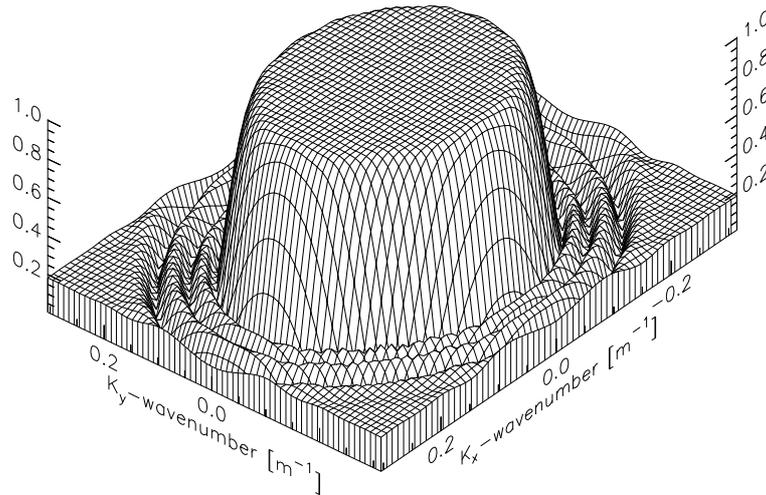


Fig. 1.8 The wavenumber spectrum of a rotated Fourier reconstructed operator with 19×19 spatial point, 128×128 k_x, k_y points with $c = 1000\text{ms}^{-1}$, $f = 25\text{Hz}$, $\Delta x = \Delta y = \Delta z = 10\text{m}$ and a maximum angle of interest of 65° .

that they coincide with the spatial grid. The wavenumber spectrum of the calculated solution is shown in Figure 1.7 where the same parameters are used as in Figure 1.1. The spectrum shown is far from good and cannot be used in an extrapolation algorithm. The problem with the Hankel transformation is that for the spatial position $r = 0$ it is not possible to define a suitable value. If we make it zero we get a singular matrix and making it non-zero is a random choice. So the Hankel transformation cannot be used directly to design circular symmetric convolution operators.

1.3.3 Rotated Fourier reconstruction

The idea of rotating a 1 dimensional operator as described in the previous section can also be used in the wavenumber domain. In the previous section it was shown that the desired 2-D convolution operator must have a circular symmetric frequency response. The projection onto a line oriented with an angle ϕ from one of the spatial axes is identical with an optimal 1-D convolution operator (Kato and Matsumoto, 1982). A slice along the same orientation in the Fourier domain corresponds to the Fourier transform of the 1-D convolution operator. This is known as the projection slice theorem. Thus the circularly symmetric frequency response is exactly described by one single projection. The problem of obtaining the circular 2-D convolution operator from the spectrum of the optimized 1-D convolution operator can be solved in several ways. The McClellan transformation is one of them and described in section 1.4. Another method is the Fourier reconstruction method which is described in this section.

The 2-D circular convolution operator is obtained by taking the inverse Fourier transform of the rotated 1-D spectrum of the optimized 1-D spatial operator with length N . To eliminate the Gibbs phenomenon the outer region of the circle is filled with the Nyquist value of the 1-D spectrum. Note that the rotated wavenumber spectrum has the circular symmetry only in the circular cen-

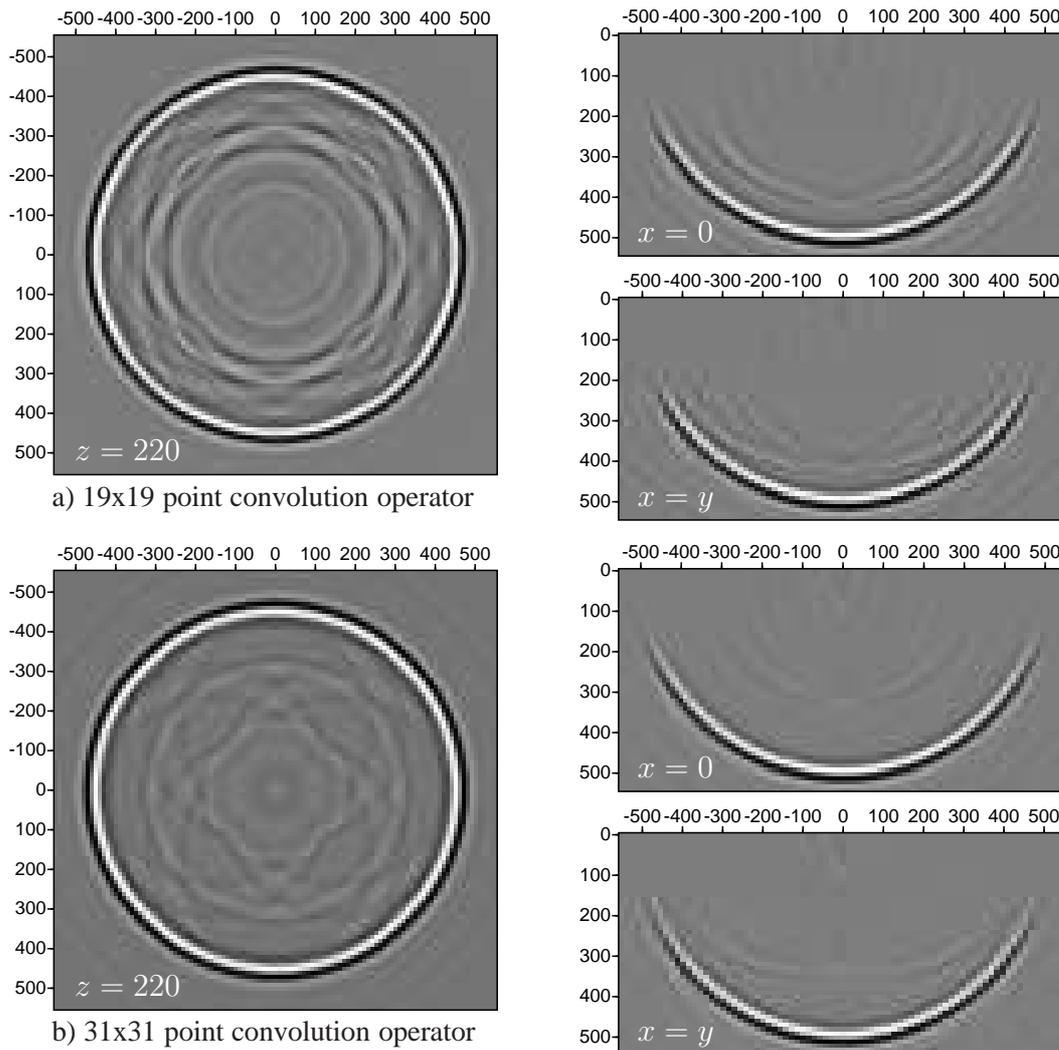


Fig. 1.9 Results of the rotated Fourier reconstructed operators for 19x19 and 31x31 point convolution operators. Note that some of the artefacts in the result for the 19x19 operator disappear for the longer operator.

trum, therefore the projection of the 2-D spectrum is only identical with the 1-D spectrum along the vertical and horizontal directions. The projection along any other direction is not exactly identical with the 1-D spectrum but gives a good approximation. After the inverse Fourier transform the result is truncated to the original length of the 1-D convolution operator ($N \times N$). Since the projection is of finite length N , the obtained convolution operator has a nearly finite support. Therefore the rectangular windowing distorts the circular frequency response only slightly.

So basically the rotated Fourier reconstruction method consists of designing a 1-D optimized spatial operator, computing and rotating its frequency response in the 2-D wavenumber plane, filling the undetermined region with the Nyquist value of the 1-D wavenumber response, performing the 2-D inverse Fourier transformation and then windowing the result. The wavenumber spectrum of an operator obtained in this way is shown in Figure 1.8.

The impulse responses for the operator sizes 19×19 and 31×31 are shown in Figure 1.9.

Note that the depth image has a perfect circular symmetry but in the middle of the circle there are some irregular artefacts visible. In the vertical cross-section a ghost event is observed after and before the main event. In the future we will try to reduce these artefacts and will use this very simple and attractive method to calculate 2D operators in a more sophisticated way. For example it is possible to use another interpolation method or a smooth window, to truncate the operator in the spatial domain.

1.3.4 Error analysis

From an engineering point of view it is interesting to investigate how the different parameters in the optimization procedure must be chosen to obtain efficient operators which are accurate up to a desired maximum angle. For this analysis it is necessary to define accuracy in a useful way. In Appendix C the most common used definitions of accuracy are given. In this section we will use the L_2 and the L_∞ norms in a certain domain of interest. The domain of interest is defined by $k_r < k \sin(\alpha_{max}) (= k_{r,max})$ with $k = \frac{\omega}{c}$.

The following L_2 and L_∞ error norms are defined over the domain of interest

$$\varepsilon_2 = \left[\frac{\int_{\phi=0}^{\frac{\pi}{4}} \int_{k_r=0}^{k_{r,max}} \|F(k_r) - \hat{F}(k_r, \phi)\|^2 k_r dk_r d\phi}{\int_{\phi=0}^{\frac{\pi}{4}} \int_{k_r=0}^{k_{r,max}} \|F(k_r)\|^2 k_r dk_r d\phi} \right]^{\frac{1}{2}} \quad (1.29)$$

$$\begin{aligned} \varepsilon_\infty^a = & \max_{0 \leq k_r \leq k_{r,max}} | \|F(k_r)\| - \|\hat{F}(k_r)\| | \\ & + \max_{k_{r,max} < k_r \leq k_N} | 1 - \|\hat{F}(k_r)\| | \quad \{\text{if } \|\hat{F}(k_r)\| > 1.0\} \end{aligned} \quad (1.30)$$

$$\varepsilon^p = \left[\int_{\phi=0}^{\frac{\pi}{4}} \int_{k_r=0}^{k_{r,max}} \left\| \frac{\partial E_p}{\partial k_r} k_r \right\|^2 dk_r d\phi \right]^{\frac{1}{2}} \quad (1.31)$$

with

$$E_p = \arg F(k_r) - \arg \hat{F}(k_r, \phi)$$

where $\hat{F}(k_r)$ is an approximation to the true function $F(k_r)$, ε_2 is the normalized least squares error, ε_∞^a the maximum amplitude error and ε^p a measurement for the derivative of the phase error with respect to the polar distance k_r . The normalized ε_2 error is a global error and is related to the accuracy of the operator. The amplitude error gives an indication of the stability of the operators in a recursive extrapolation scheme. Note that $\|F(k_r)\|^2 = 1$ in the domain of interest. Included in the ε_∞^a error is a stability measurement for $k \sin(\alpha_{max}) < k_r \leq k_N$. If the amplitude of the operator in this domain is higher than 1.0 then it contributes to the ε_∞^a error. The ε^p error is defined in such a way that it is sensitive to errors in the circular symmetry of the operator. In the ideal case the ε^p error should be zero because of the circular symmetry of the operator. To compute the derivative with respect to k_r a three point finite difference operator is used to compute the derivative with respect to k_x and k_y . With these derivatives the $\frac{\partial}{\partial k_r} = \frac{\partial}{\partial k_x} \frac{\partial k_x}{\partial k_r} + \frac{\partial}{\partial k_y} \frac{\partial k_y}{\partial k_r}$ is calculated. If the phase error is large and the derivative with respect

Operator weight	5 Hz			20 Hz			40 Hz		
	ε_2	ε_∞^a	ε^p	ε_2	ε_∞^a	ε^p	ε_2	ε_∞^a	ε^p
1e-5	1.5e-3	1.9e-2	4.4e-4	5.6e-4	2.8e-2	8.4e-4	3.9e-4	1.3e-2	1.1e-3
5e-5	2.9e-3	2.6e-3	8.4e-4	1.4e-3	2.3e-3	1.6e-3	8.3e-4	4.2e-3	1.4e-3
1e-4	4.1e-3	3.8e-3	1.2e-3	1.7e-3	3.6e-3	2.1e-3	1.3e-3	5.3e-3	1.5e-3
5e-4	6.9e-3	7.4e-3	1.9e-3	2.8e-3	7.3e-3	2.6e-3	1.9e-3	8.1e-3	2.5e-3
1e-3	7.9e-3	9.1e-3	2.0e-3	3.4e-3	9.4e-3	2.7e-3	2.4e-3	9.2e-3	3.2e-3
5e-3	1.0e-2	1.6e-2	2.1e-3	6.0e-3	1.7e-2	2.8e-3	4.0e-3	1.1e-2	7.6e-3
1e-2	1.2e-2	2.0e-2	2.1e-3	8.0e-3	2.4e-2	3.6e-3	5.4e-3	1.3e-2	1.1e-2
5e-2	2.1e-2	3.6e-2	1.6e-3	1.5e-2	3.5e-2	9.4e-4	1.1e-2	2.3e-2	2.2e-2

Table 1.2 Error analysis for different weighting factors with a constant operator size (19×19) and maximum angle of interest (60°).

to k_r , is also large then the operator will have a detectable non-circular character.

To determine the errors due to the recursive use of the operator in a homogeneous medium (which is a worst case situation) the difference with respect to the reference impulse response is calculated for every depth slice according to

$$\varepsilon_s(z) = \left[\frac{\int_{x=0}^{x_{max}} \int_{y=0}^{y_{max}} \|F(x, y, z) - \hat{F}(x, y, z)\|^2 dx dy}{\int_{x=0}^{x_{max}} \int_{y=0}^{y_{max}} \|F(x, y, z)\|^2 dx dy} \right]^{\frac{1}{2}} \quad (1.32)$$

This spatial error will be presented in an error curve as function of the angle (=depth).

The calculation of the wavenumber errors in equation (1.29), (1.30) and (1.31) gives three values for one operator defined for one frequency. For a more useful definition three frequencies, within the frequency range of interest, are analyzed: one at a low frequency (in our example 5 Hz.), a central frequency (20 Hz) and at a high frequency (40 Hz.). To have a better idea how the different errors in the operators are exposed in the depth image a number of experiments is carried out where the error in the operators is changing. The error in the operator can be varied by changing the weighting function, the operator length and the maximum angle of interest. From these experiments it is possible to derive an error criterion for the calculated operator errors which can be used as a measure of accuracy for the extrapolation result. These experiments are done with different WLSQ optimized operators with a varying weighting function, a change in operator size and a varying maximum angle of interest. All other parameters remain the same.

The error for three characterizing frequencies is given in Table 1.2 for different weighting factors, which are given in the first column. The weighting factor is defined as the value of the box-shaped weight function outside the domain of interest. Inside this domain the weight function is given the value 1.0. A practical limit of the weighting factor is 5e-5, because smaller factors gives unstable operators (reflected in the ε_∞^a error). The error for a varying operator size is given in Table 1.3 and for the results for different maximum angles is shown in Table 1.4.

- *changing weight factor*

In Figure 1.10 two impulse responses are displayed which differ with respect to the used weight-

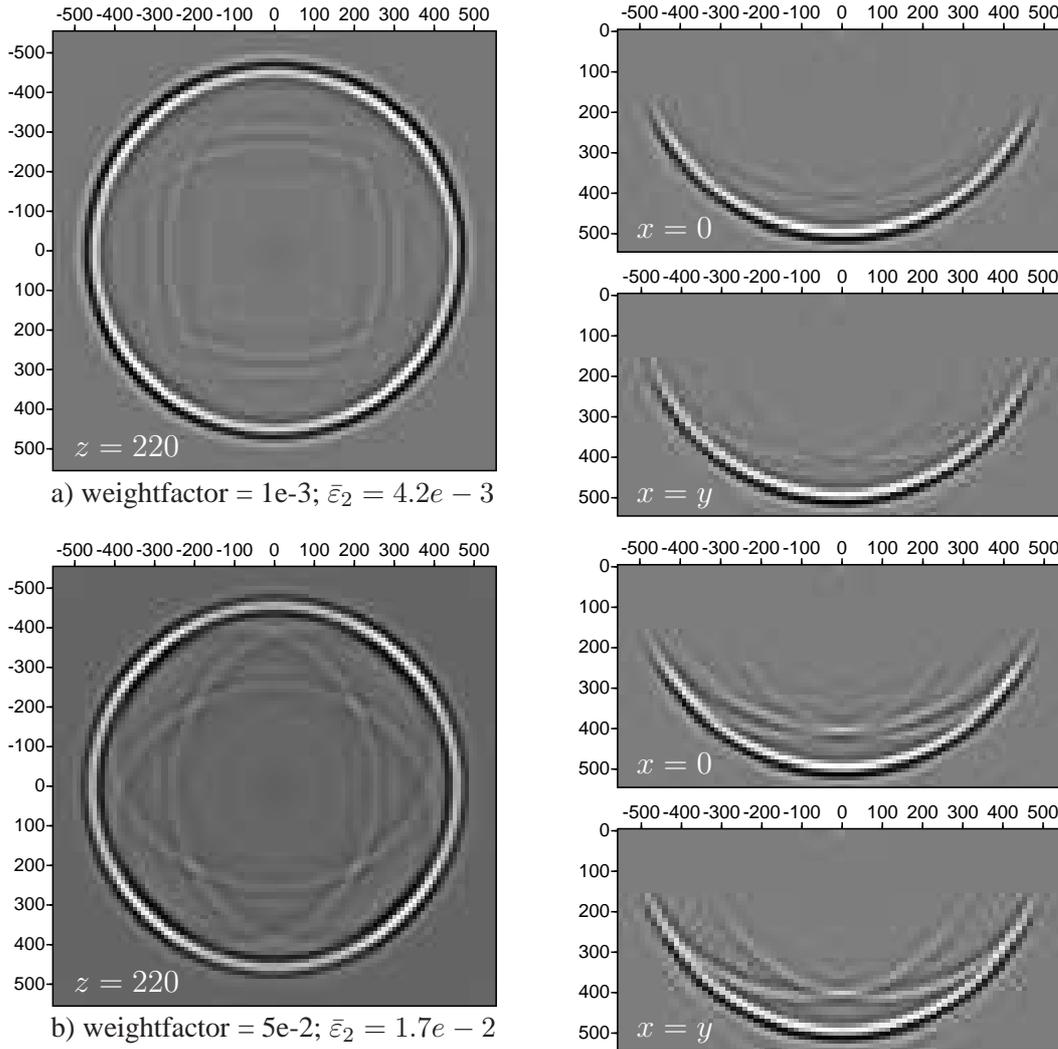


Fig. 1.10 The impulse response as function of the weighting factor. The depth cross section is equivalent with an angle of 65° . Note that for larger weighting factors the error grows and the result contains more artefacts.

ing factor in the operator calculation. The $\bar{\varepsilon}_2$ error given in the Figures is the average error over the three characterizing frequencies. From this Figure and Table 1.2 the following conclusions can be drawn;

- (1) For the stable weight factors the ε_2 and the ε_∞^a errors are increasing if the weighting factor increases. The unstable weight factor ($1e-5$) is only reflected in the stability part (second equation on the right hand side of equation (1.30)) of the ε_∞^a error. The ε^p and the ε_2 error are not sensitive for instabilities outside the domain of interest. The best weighting factor is therefore that factor which gives an operator which remains just stable. This factor is easily to determine because it can be chosen constant for all frequencies for a fixed operator size.
- (2) An average ε_2 error (averaged over all frequencies of interest, indicated by $\bar{\varepsilon}_2$) smaller than $2e-3$ gives an accurate depth image up to the desired maximum angle. A larger $\bar{\varepsilon}_2$ error gives artefacts 'inside' the main event as observed in Figure 1.10.
- (3) If the ε_2 error is small than the other errors are not by definition small too. This fact can

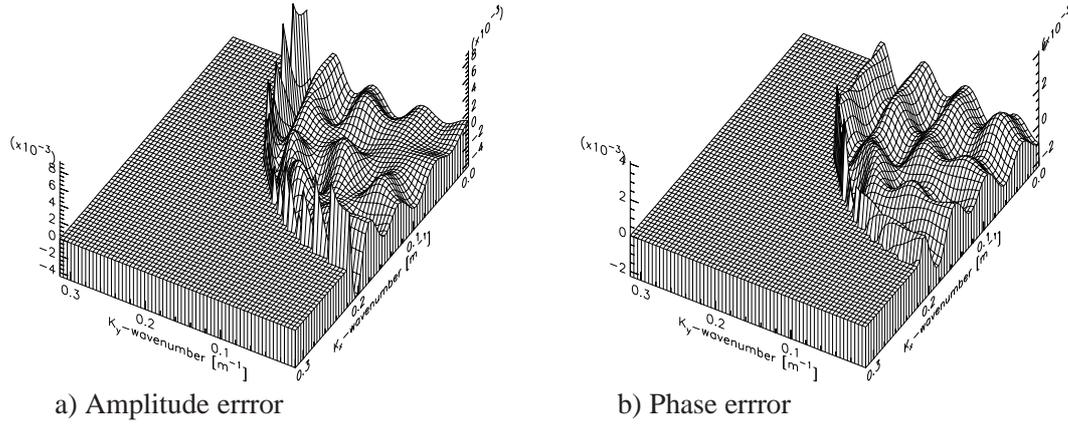


Fig. 1.11 Amplitude and phase errors for an operator at 40 Hz with a weight factor of $1e-3$ and an operator size of 19×19 points. Note the high error values at the edges of the domain of interest.

be explained by looking at the amplitude and phase errors which are shown in Figure 1.11 for an operator at 40 Hz with a weight factor of $1e-3$. These error functions have big peaks at the higher angles. In the calculation of the global ε_2 error these peaks are averaged out. In the depth images of Figure 1.10 these peaks in the error function are also not visible, but one can imagine that if these peaks become too big the recursion scheme can become unstable and inaccurate.

- *changing operator size*

In Figure 1.12 two different operators are displayed which differ with respect to the operator size. From this Figure, Figure 1.6 and Table 1.3 the following conclusions can be drawn;

(1) A larger operator size will give more accurate results, but for a certain accuracy (which is reached for this problem at an operator size of 25×25 points with an average ε_2 error in the order of $1e-3$) the improvement, by using a larger operator size, on the result is little. So there exists an optimum efficient operator size.

(2) The large ε_∞^a error for 40 Hz operator with size 37×37 is due to the fact that the operator has a little amplitude jump after 60° on the diagonal $k_x = k_y$ which is taken into account by the stability part in the ε_∞^a error. This effect can be detected by the small ε_2 and ε^p error. Choosing a slightly bigger maximum angle (65°) will give a stable operator. The jumps at the edges of

Operator size	5 Hz			20 Hz			40 Hz		
	ε_2	ε_∞^a	ε^p	ε_2	ε_∞^a	ε^p	ε_2	ε_∞^a	ε^p
13x13	6.1e-3	3.7e-3	1.9e-3	3.1e-3	3.1e-3	4.2e-3	2.7e-3	4.4e-3	6.6e-3
19x19	2.9e-3	2.6e-3	8.4e-4	1.3e-3	2.3e-3	1.6e-3	8.3e-4	4.2e-3	1.4e-3
25x25	1.7e-3	1.8e-3	4.4e-4	6.1e-4	1.6e-3	6.3e-4	4.1e-4	2.0e-3	6.8e-4
31x31	1.3e-3	1.0e-3	2.7e-4	3.5e-4	1.0e-3	3.1e-4	2.5e-4	7.6e-4	5.8e-4
37x37	8.5e-4	1.1e-3	1.6e-4	2.4e-4	6.8e-4	2.7e-4	1.8e-4	2.3e-2	4.7e-4

Table 1.3 Error analysis for different operator sizes with a constant weighting factor ($5e-5$) and maximum angle of interest (60°).

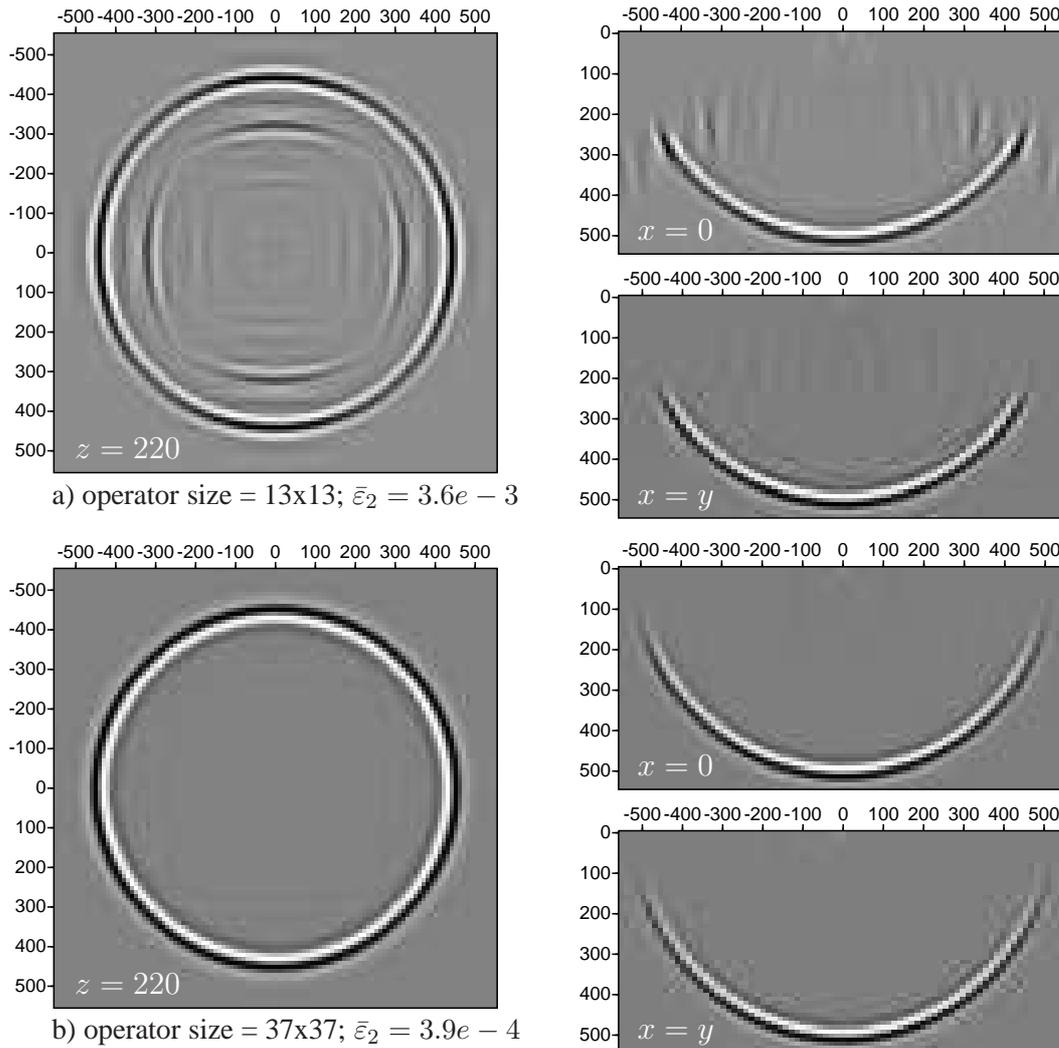


Fig. 1.12 The impulse response as function of the operator size. The depth cross section is equivalent with an angle of 60° . Note that a small operator size gives problems at higher angles.

the domain of interest are typically for least squares design methods. These peaks can be suppressed by changing the weighting function at the edges of the domain or by using an additional optimization step which uses the results of the first step.

(3) Using a small operator and defining a relative large maximum angle gives errors which are typically of the form as shown in the vertical cross section of Figure 1.12a). These artefacts are due to errors at the higher angles and are most clearly represented by the ε_2 error. For the 13×13 operator the average $\bar{\varepsilon}_2$ error is $4e-3$, for the other, larger, operators this error is smaller than $2e-3$. From these experiments we can conclude that an $\bar{\varepsilon}_2$ error smaller than $2e-3$ will give no visible artefacts in the depth image. The artefacts are due to the peaks at the edges of the domain of interest. In these critical examples the phase error is for angles larger than the maximum angle already large while the amplitude error is still small. So the large phase error is not suppressed by a small amplitude error and gives artefacts as observed in the Figures. In our research group there are some methods developed to suppress these artefacts (Hoff, 1995).

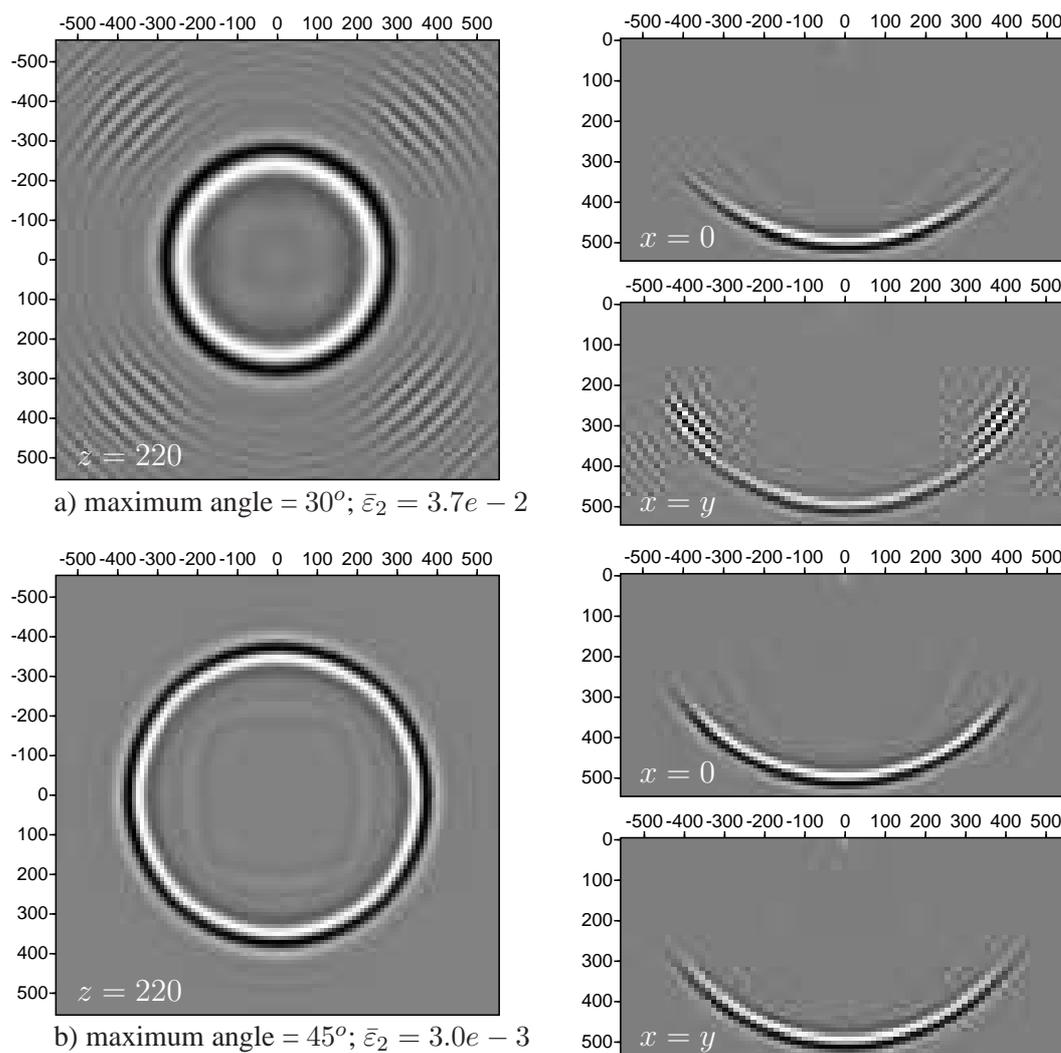


Fig. 1.13 The impulse response as function of the maximum angle of interest with an operator size of 13×13 . The depth cross section is taken at the maximum angle of interest. Note that the very small angles give problems.

- *changing maximum angle*

In Figure 1.13 two different depth images are displayed which differ with respect to the maximum angle of interest in the operator. The operator size is chosen fixed at 13×13 . The result with a maximum angle of 75° is unstable. From this Figure, Figure 1.12a and Table 1.4. The following conclusions can be drawn;

- (1) A smaller maximum angle does not automatically give a better performance for an operator with the same size. From the ε_∞^a error in Table 1.4 it is observed that there exists an optimum angle which is, for the 13×13 operator with the chosen weight factor ($5e-5$), an angle between 30 and 45 degrees. This optimum is most clearly observed in the ε_∞^a error.
- (2) The ε_2 and ε^p are of little use because these are not defined for angles outside the domain of interest.
- (3) Note that we kept the weighting factor constant throughout the different experiments, by changing the weight factor it is possible to make a very short operator which is also stable out-

Operator angle	5 Hz			20 Hz			40 Hz		
	ε_2	ε_∞^a	ε^p	ε_2	ε_∞^a	ε^p	ε_2	ε_∞^a	ε^p
15	2.1e-3	3.3e-3	-	7.6e-4	5.4e-2	2.0e-5	6.2e-4	9.6e-2	4.9e-5
30	1.2e-3	1.5e-3	4.1e-5	7.2e-4	1.1e-3	1.2e-4	5.3e-4	6.6e-2	1.5e-4
45	2.9e-3	3.0e-3	4.7e-4	1.0e-3	3.2e-3	3.1e-4	8.0e-4	3.0e-3	4.8e-4
60	6.1e-3	3.7e-3	1.9e-3	3.1e-3	3.1e-3	4.2e-3	2.7e-3	4.4e-3	6.6e-3
75	1.2e-2	3.1e-3	6.5e-3	7.8e-3	3.9e-1	2.4e-2	6.14e-3	1.3e-0	3.8e-2

Table 1.4 Error analysis for different maximum angles with a constant weighting factor (5e-5) and a fixed operator size of 13×13 .

side the domain of interest (see below).

From the experiments described above we can develop a criterion which can be used to determine if a certain operator calculation method gives stable and accurate results in a recursive extrapolation algorithm. To use this criterion the wavenumber spectrum of the operator must be calculated for three characterizing frequencies and the ε_2 , ε_∞^a and ε^p errors have to be calculated for every frequency. These errors must obey the following relations;

- $\bar{\varepsilon}_2 \leq 2e^{-3}$ accuracy measurement
- $\varepsilon_\infty^a \leq 3e^{-3}$ stability measurement
- $\varepsilon^p \leq 1e^{-2}$ circularity measurement

The ε^p error is not tested very well in this section, but in the section about the McClellan method the ε^p error turns out to be very useful and there we will define the ε^p criterion better.

Operator size	5 Hz			20 Hz			40 Hz		
	ε_2	ε_∞^a	ε^p	ε_2	ε_∞^a	ε^p	ε_2	ε_∞^a	ε^p
19x19	2.4e-2	3.9e-2	4.0e-3	5.7e-3	1.3e-2	6.2e-3	4.6e-3	1.0e-2	6.7e-3
25x25	1.8e-2	3.2e-2	2.4e-3	4.8e-3	1.3e-2	3.8e-3	2.3e-3	1.0e-2	1.6e-3
31x31	1.9e-2	3.6e-2	2.5e-3	3.7e-3	1.3e-2	2.2e-3	1.6e-3	5.1e-3	2.8e-3

Table 1.5 Fourier reconstructed operators which are accurate up to a maximum angle of 60° . The 1-D operators are obtained by using the 1D WLSQ method.

The performance of the rotated Fourier reconstruction method can be analyzed with the defined error criteria. In Table 1.5 the errors are given for different operator sizes. The 1-D operators are obtained by using the Remez exchange algorithm (similar results were obtained by using a WLSQ operator). The large ε_2 errors indicate that the overall spectrum of the operator is inaccurate. A better interpolation method (better than linear) or a smoother window in the spatial domain may improve the result. With the used linear interpolation method a larger operator size does not improve the result significantly.

Upto now we have calculated the errors in the wavenumber domain and interpreted these errors in the spatial domain. But with the aid of the reference result and equation (1.32) it is possible to calculate an error directly in the spatial domain. In Figure 1.14 this error is shown for the 19×19 and 31×31 operators which impulse response is displayed in Figure 1.6. The vertical

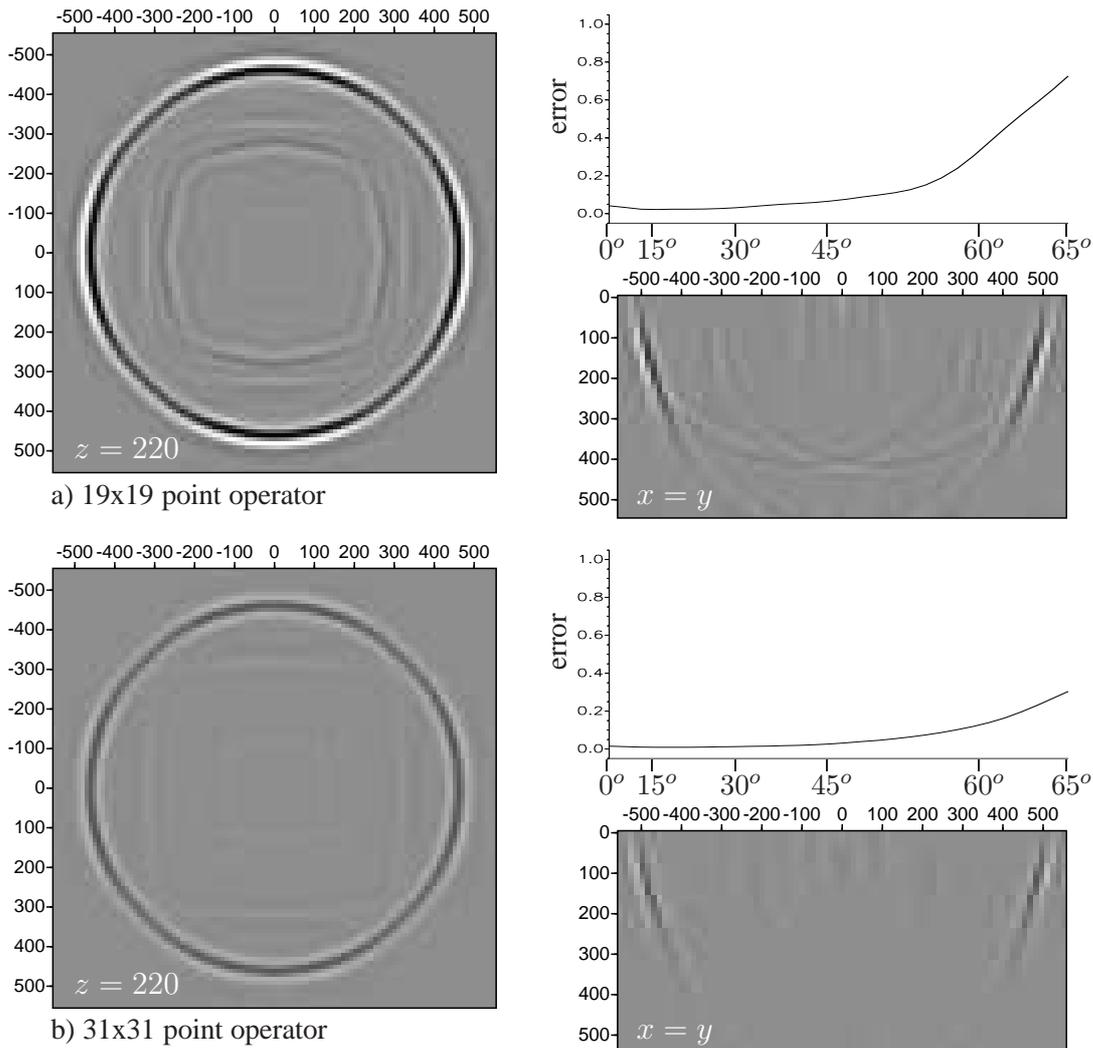


Fig. 1.14 The spatial error as function of angle (=depth) together with a horizontal (for 65°) and vertical (for $x = y$) cross section of the error. Note that the all cross sections are displayed with the same scaling factor.

cross section is displayed for the diagonal $x = y$ and the horizontal cross section for an angle of 65° . The top picture on the right hand side shows the error as function of the angle (=depth). From these errors the following observations are made:

- (1) The increasing error line for higher angles as displayed in Figure 1.14 is due to amplitude errors and artefacts for the higher angles.
- (2) Increasing the operator size increases the accuracy; less artefacts and a better amplitude.
- (3) From the impulse response alone it is difficult to interpret the accuracy of the operator, comparing it with the reference operator gives a good indication of the errors in the operator and the influence of the recursive application of the operators.

Given the error criteria we can also determine how the weight factor and the operator size must be chosen for a maximum angle of interest. The results of these experiments are summarized in Table 1.6. The small angles 15° and 30° are difficult to optimize for the given maximum angle, but by choosing a slightly bigger angle the operator can become stable and accurate for

the smallest operator size possible. For example to get the operator for 30° a maximum design angle of 40° degrees has to be chosen. For the higher angles this problem does not occur. If one wants to design operators with small maximum angles of interest and suppression of all the higher angles a larger operator size must be chosen than the one given in Table 1.6. The larger $\bar{\varepsilon}^p$ error in the 75° operators is not as bad as it looks, the largest error peaks are positioned at the edges of the domain of interest.

Operator			20 Hz			Average		
angle	size	weight	ε_2	ε_∞^a	ε^p	$\bar{\varepsilon}_2$	$\bar{\varepsilon}_\infty^a$	$\bar{\varepsilon}^p$
15	5x5	1e-5	1.6e-3	3.2e-3	1.6e-5	2.4e-3	2.2e-3	7.1e-5
30	9x9	2e-5	1.4e-3	3.0e-3	5.6e-4	1.9e-3	2.9e-3	2.7e-4
45	13x13	4e-5	1.0e-3	3.3e-3	3.5e-4	1.5e-3	3.1e-3	3.9e-4
60	19x19	4e-5	1.2e-3	2.4e-3	1.6e-3	1.5e-3	2.9e-3	1.2e-3
75	31x31	6e-5	1.7e-3	1.2e-3	4.8e-3	1.9e-3	1.4e-3	4.3e-3

Table 1.6 Optimum operators which are accurate up to a maximum angle of interest. Note that for small angles these operators are stable but don't suppress all higher angles.

Note that for small operator sizes the higher frequencies are most sensitive to errors, for the larger operators the lower frequencies are more sensitive to errors. This behavior is related to the WLSQ optimization method. A very small operator has a limited number of 'error' peaks in the frequency domain due to the limited number of contributing wavenumber components. The WLSQ optimization method with a limited number of wavenumber components cannot have very large peaks (Berkhout, 1984). If there are more wavenumber components the WLSQ method can build up large peaks in the error function (Gibbs phenomenon).

1.4 McClellan transformation

The McClellan transformation transforms a 1-D convolution operator to a 2-D convolution operator with a certain symmetry. This transformation is of interest because the implementation is simple and the computation of the transformation coefficients can be done efficiently. Hale (1991a) introduced the McClellan transformation into the Geophysical world and described two related techniques which can replace the direct 2-Dimensional spatial convolution: (1) transformation of the *non-recursive* 1-Dimensional symmetrical filter in a 1-Dimensional *recursive* filter by using the Chebyshev recursion formula (see Appendix B), (2) the McClellan transformation of a 1-Dimensional filter to a circular symmetric 2-Dimensional filter. We will first discuss the transformation from a 1-Dimensional filter to a 2-Dimensional filter. Next the Chebyshev recursion formula is explained and at the end of this section several methods are discussed which optimize the steps and coefficients used in the McClellan transformation.

McClellan Transformation form 1-D to 2-D

If the operator has a circular symmetry it is possible to reduce the computation time of the 2-Dimensional filter by means of a McClellan transform. The McClellan transform (McClellan,

1973) defines a mapping from a 1-D wavenumber axis to the 2-D wavenumber domain. The change of variables to be described depends on the fact that both the operator approximations in 1 and 2-Dimensions can be written as sums of cosine functions. The 1-Dimensional filter problem of an even symmetrical operator can be rewritten as (note the similarity with equation (1.8))

$$\tilde{F}(k_r) \approx F(0) + 2 \sum_{m=1}^M F(m\Delta x) \cos(k_r m\Delta x) \quad (1.33)$$

$$\tilde{F}(k_r) \approx \sum_{m=0}^M F'(m\Delta x) \cos(k_r m\Delta x) \quad (1.34)$$

$$\tilde{F}(k_r) \approx \sum_{m=0}^M \hat{F}(m\Delta x) \cos(k_r \Delta x)^m \quad (1.35)$$

with the choice of a suitable set of coefficients $\hat{F}(m\Delta x)$ that approximate the 1-D extrapolation operator $\tilde{F}(k_r)$. In equation (1.34) $F'_m = 2F_m$ for $m = 1, \dots, M$ and $F'_m = F_m$ for $m = 0$. The last step from equation (1.34) to equation (1.35) can be seen by letting $\phi = \cos(k_r \Delta x)$. Then $\cos(k_r m\Delta x) = \cos(m \arccos(\phi)) = T_m(\phi)$, where $T_m(\phi)$ is the Chebyshev polynomial of order m (see Appendix B). Each cosine term in equation (1.34) may then be expressed in the form

$$\cos(k_r m\Delta x) = \sum_{n=0}^N \alpha_{m,n} \cos(k_r \Delta x)^n \quad (1.36)$$

where the $\alpha_{m,n}$ are real coefficients and easily obtained with the Chebyshev recursion formula. Equation (1.34) reduces then further to

$$\tilde{F}(k_r) \approx \sum_{m=0}^M F'(m\Delta x) T_m(\phi) = \sum_{m=0}^M \hat{F}(m\Delta x) \phi^m \quad (1.37)$$

where both right-sides of the equation are now polynomials in ϕ .

The cosine terms in equation (1.35) can be approximated by a 2-Dimensional filter (assuming $\Delta x = \Delta y$)

$$\cos(k_r \Delta x) \approx \sum_{p=0}^P \sum_{q=0}^Q c_{pq} \cos(k_x p\Delta x) \cos(k_y q\Delta y) \quad (1.38)$$

where c_{pq} are called the McClellan factors (McClellan, 1973). By making the substitution of equation (1.38) into equation (1.35) it reduces to

$$\tilde{F}(k_r) \approx \sum_{m=0}^M \sum_{n=0}^N \ddot{F}(m\Delta x, n\Delta y) \cos(k_x \Delta x)^m \cos(k_y \Delta y)^n \quad (1.39)$$

which can be put in the form (using Chebyshev's recursion formula again)

$$\tilde{F}(k_r) \approx \sum_{m=0}^M \sum_{n=0}^N \check{F}(m\Delta x, n\Delta y) \cos(k_x m\Delta x) \cos(k_y n\Delta y) \quad (1.40)$$

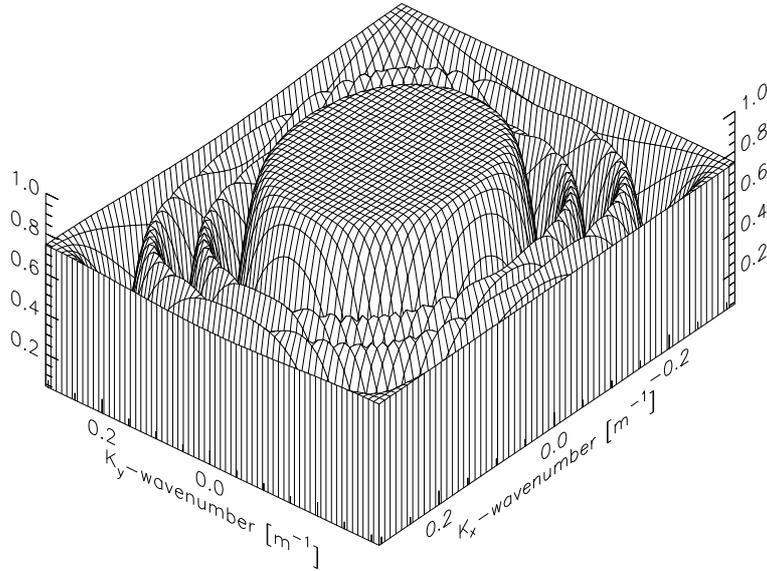


Fig. 1.15 The wavenumber spectrum of a first order McClellan operator with a 1-D operator of 19 spatial points, 128×128 k_x, k_y points with $c = 1000ms^{-1}$, $f = 25Hz$, $\Delta x = \Delta y = \Delta z = 10m$ and a maximum angle of interest of 65° .

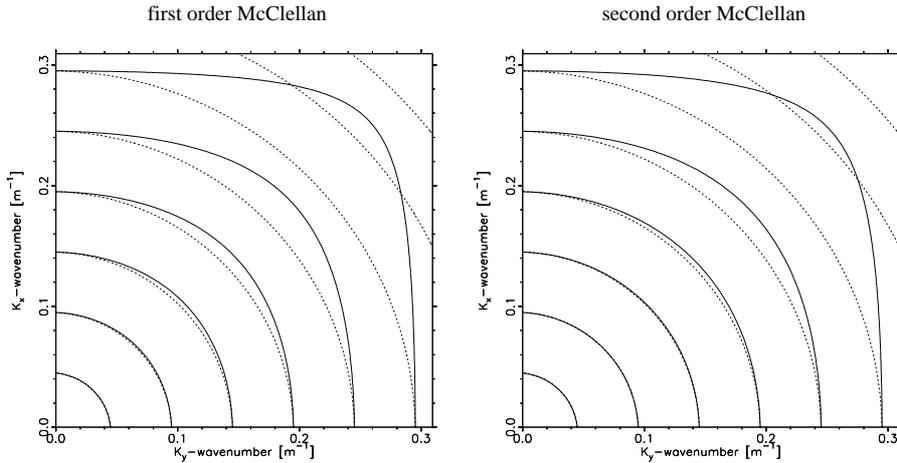


Fig. 1.16 Contour plots for the first and second order McClellan transformation. Note that for higher wavenumbers both approximations deviate from the ideal circular (dashed) line.

which is the desired form for a 2-Dimensional filter which was already shown in equation (1.8) (Note that \check{F} and $\check{\check{F}}$ are scaled versions of F). For example for $P = Q = 1$ the transformation for circular symmetry reduces to a 9-term McClellan convolution operator (also called a first order approximation) which is given by Hale (1991a) where $-c_{00} = c_{10} = c_{01} = c_{11} = 0.5$ and

$$\cos(k_r \Delta x) \approx -1 + 0.5(1 + \cos(k_x \Delta x))(1 + \cos(k_y \Delta y)). \tag{1.41}$$

In Figure 1.15 an extrapolation operator is shown which is designed with the first order McClellan-

lan transformation of equation (1.41) and a 1-D operator $F(m\Delta x)$ of 10 points (19 point full operator length). Note the square shape of the operator for the wavenumbers near the Nyquist wavenumber (the edges of the figure). $P = Q = 2$ gives a 17-term McClellan transform (second order approximation) which is also given by Hale (1991a)

$$\begin{aligned} \cos(k_r \Delta x) \approx & -1 + 0.5(1 + \cos(k_x \Delta x))(1 + \cos(k_y \Delta y)) \\ & - 0.5c(1 - \cos(2k_x \Delta x))(1 - \cos(2k_y \Delta y)). \end{aligned} \quad (1.42)$$

with $c = 0.0255$. The McClellan factors c_{pq} can be derived by defining points in the k_x, k_y plane which map to a point on the k -axes of the 1-D operator such that all coefficients are uniquely defined (for the first order McClellan transform 4 points are needed). The problem with the McClellan transform, given the original McClellan factors in equations (1.41) and (1.42), is that for higher angles the transformation deviates from the ideal circular shape. The contour plots shown in Figure 1.16 represent the contours of the first order McClellan ($P = Q = 1$) and the second order McClellan ($P = Q = 2$) transformation. In the contour plots the deviation for the higher wavenumbers is observed clearly. The second order transformation reduces the deviation a little but remains still significant.

Chebyshev recursion formula

The second improvement in the computation scheme is the transformation of the non-recursive 1-Dimensional filter to a 1-Dimensional recursive filter derived from the recursive formula of the Chebyshev polynomials

$$\cos(m\phi) = 2 \cos(\phi) \cos((m-1)\phi) - \cos((m-2)\phi) \quad (1.43)$$

This Chebyshev filter structure is not useful for 1-Dimensional filters. Direct convolution is both simpler and more efficient. The Chebyshev structure is more advantageous for 2-Dimensional operators with an even symmetry, such as the circular symmetric extrapolation operators. Writing equation (1.43) for the first four terms in the 1-Dimensional case gives

$$\begin{aligned} \tilde{F}(k_x) &= F_0 + 2 \sum_{m=1}^M F_m \cos(k_x m) \\ &= F_0 \end{aligned} \quad (1.44)$$

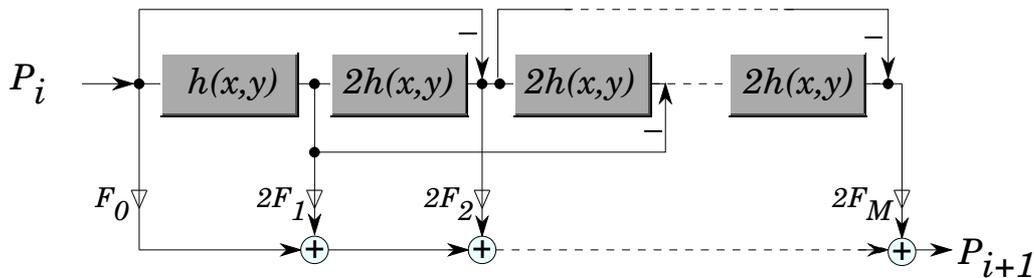


Fig. 1.17 Chebyshev recursion scheme. The $h(x, y)$ boxes represent the 2-D McClellan transformation of $\cos(k_r)$, F_m represents the coefficients of the 1-D convolution operator.

$$\begin{aligned}
 &+2F_1[\cos(k_x)] \\
 &+2F_2[2 \cos(k_x) \cos(k_x) - 1] \\
 &+2F_3[2 \cos(k_x) \left((2 \cos(k_x) \cos(k_x) - 1) \right) - \cos(k_x)] \\
 &+2F_4[2 \cos(k_x) \left(2 \cos(k_x) \left((2 \cos(k_x) \cos(k_x) - 1) \right) - \cos(k_x) \right) - (2 \cos(k_x) \cos(k_x) - 1)]
 \end{aligned}$$

This recursive scheme can be implemented in the computer without much effort. McClellan and Chan (1977) have analyzed this so called Chebyshev structure in detail and observed that the scheme requires the minimum number of multiplications in comparison with the direct scheme and it is the most stable scheme with respect to the round off noise. In Figure 1.17 a flow diagram for the Chebyshev recursion scheme is given. Note that only the coefficients of the 1-D operator are involved. Hence the number of computations depends linearly on the length N of the 1-D operator and not on N^2 as in the implementation of a direct 2-D convolution. The computation times for several 1-D operator lengths and different operators T are given in Appendix A.

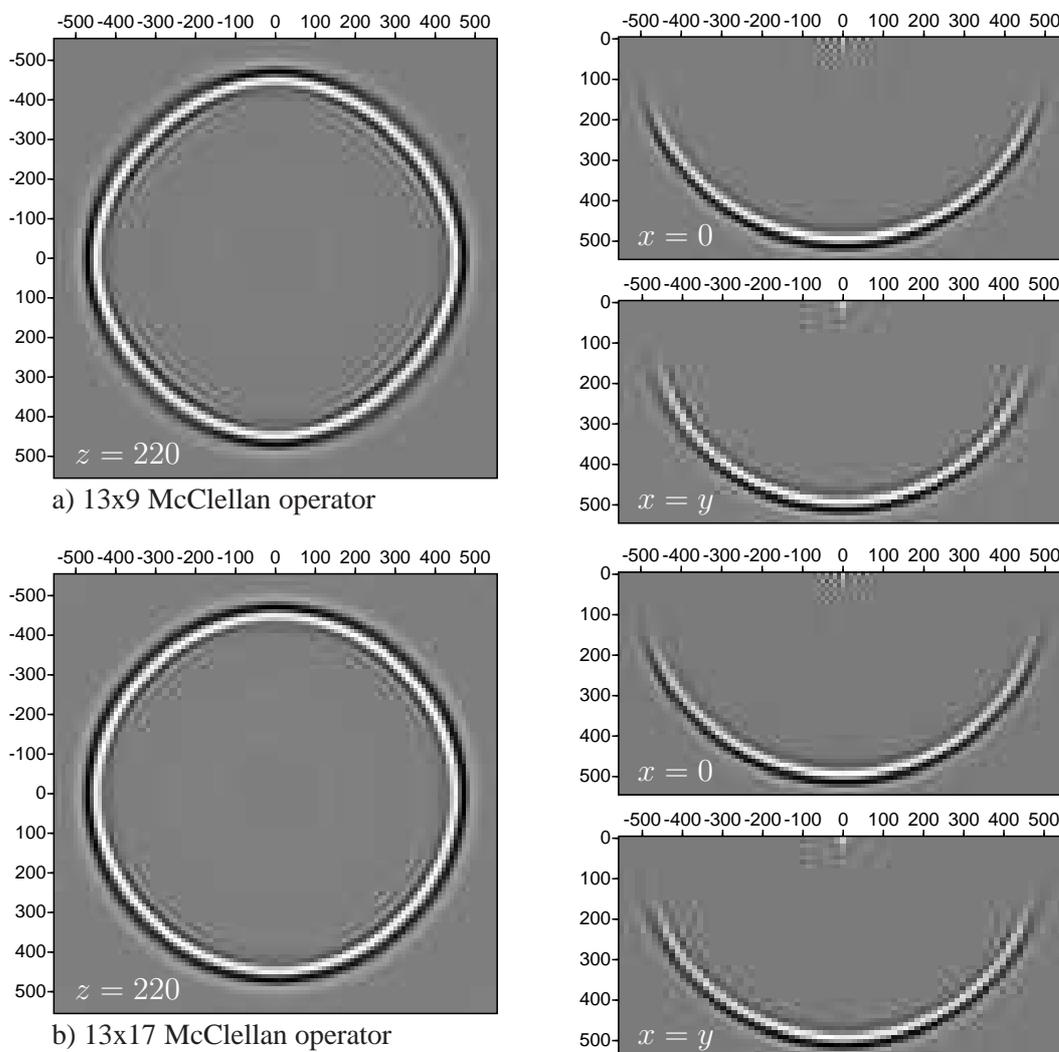


Fig. 1.18 Depth images obtained with the original McClellan transformation. For the higher angles both operators deviate from the ideal circular shape. Note the deviation at the higher angles in the diagonal slice.

In Figure 1.18 the migration results of the McClellan transformation combined with the Chebyshev recursion scheme are shown for both the first (a) and second (b) order McClellan transformation. The optimized 1-D convolution operator has a full length of 25 points and is obtained with the Remez exchange algorithm (see Appendix C for a detailed discussion on the Remez algorithm). The Chebyshev recursion scheme makes use of the even symmetry in the operator so only 13 points of the 1-D operator are needed as expansion terms in the recursive scheme. The notation of the operator size used in Figure 1.18 gives first the number of terms in the expansion and second the size of one single term. For example the notation 13×9 means that the expansion is done in 13 terms where every term consists of an operator with 9 points. How many multiplications and additions are actually needed is explained in Appendix A.

The cross-sections in Figure 1.18 give a good view of how the McClellan transformation handles the higher angles. Note that the deviation of the ideal circle for the second order (17 term) is only a little less than for the first order (9 term) McClellan transformation. The noise around the source position in the x-slice is an artefact of the used 1-D operator. The 1-D Remez optimized operator is chosen because it gives, in some way, the best 1-D operator and the even symmetry of the 1-D operator is used explicitly in the optimization scheme.

Despite the deviation at the higher angles the McClellan transformation combined with the Chebyshev recursion scheme is a very powerful and useful approach. The performance at higher angles can be improved in several ways. In the subsections 1.4.1, 1.4.2, 1.4.3 and 1.4.4 four improvements are discussed.

1.4.1 Hazra and Reddy Coefficients

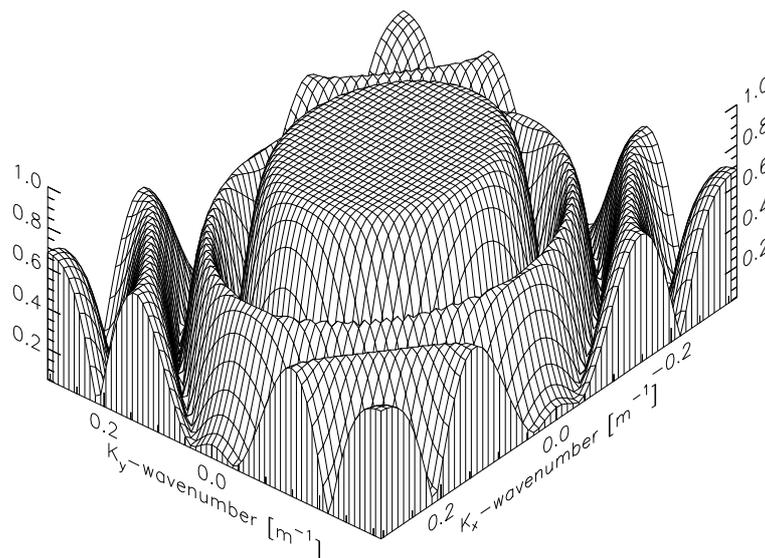


Fig. 1.19 The wavenumber spectrum of a first order McClellan operator with the optimized Hazra and Reddy coefficients and a 1-D operator of 19 spatial points with 128×128 k_x, k_y points with $c = 1000\text{m.s}^{-1}$, $f = 25\text{Hz}$, $\Delta x = \Delta y = \Delta z = 10\text{m}$ and a maximum angle of interest of 65° .

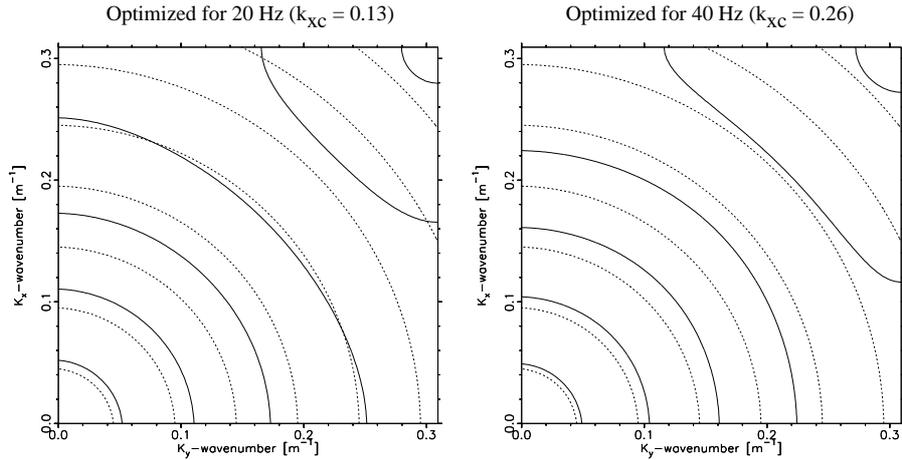


Fig. 1.20 Contour plots of the optimized Hazra and Reddy transformation for two different frequencies. Note the improvement in the circular shape in comparison with the original McClellan transformation.

Optimizing the design of the McClellan factors c_{pq} in equation (1.38) is a good way to improve the performance of the McClellan transformation. The first order filter is of special interest because it is a small and therefore fast operator. The aim of the technique proposed by Hazra and Reddy (1986) is to make the maximum contour of interest of the 2-D operator approximate a circle with a high degree of accuracy. This better approximation is achieved by mapping an additional point of the cut-off boundary of the 2-D operator onto the cut-off boundary of the 1-D filter (the cut-off boundary is defined by the maximum wavenumber of interest). This mapping of the additional point is obtained by making the cut-off wavenumber of the 1-D filter as one of the design parameters. A consequence of this is that the wavenumber of the cut-off boundary of the 2-D operator on the k_x -axis and k_y -axis may be different from the cut-off wavenumber of the 1-D filter.

The first order original McClellan transformation maps the origin $(0, 0)$ in the (k_x, k_y) plane onto the point $k_x = 0$ of the k_x -axis in the wavenumber response of the 1-D operator. The points $(k_{x,N}, 0)$, $(0, k_{y,N})$ and $(k_{x,N}, k_{y,N})$ from the (k_x, k_y) plane all map onto $k_x = k_{x,N}$. With the definition of these four points the coefficients in the first order McClellan transform are uniquely determined. This mapping has the following properties: (1) the contours of the McClellan transformation are approximately circular for low values of k_r and deviates considerably from circular contours as k_r increases and is square at $k_r = 1$, (2) the original McClellan transformation makes the frequency response of the 2-D operator along the k_x -axis and along the k_y -axis identical to the frequency response of the original 1-D operator. The contour plots in Figure 1.16 show that the deviation from the circular contour is maximum near the neighborhood of the diagonal joining the points $(0, 0)$ and $(k_{x,N}, k_{y,N})$ in the (k_x, k_y) plane. It is possible, for a given maximum k_r , to improve the contour by forcing an appropriate point on this diagonal to be on the circular contour. With this mapping of an extra point on the circular contour, it is not possible, to make the frequency response of the 1-D operator identical to the frequency response of

the original 1-D operator along the k_x -axis and along the k_y -axis. Thus the mapping of the extra point on the diagonal is only possible when the cut-off wavenumber along the k_x axes is one of the design parameters. The McClellan factors according to Hazra and Reddy are dependent on the maximum wavenumber of interest and are given by

$$c_{11} = \frac{ab - 2ac}{2bc^2}, \quad c_{01} = c_{10} = \frac{1}{g} - c_{11}, \quad c_{00} = 1 - c_{01} - c_{10} - c_{11}$$

where

$$b = \sin^2\left(\frac{k_{xc}}{2}\right), \quad c = \sin^2\left(\frac{k_{xc}}{2\sqrt{2}}\right), \quad g = 2 + \frac{b - 2c}{c^2}, \quad a = \frac{b}{g} \quad (1.45)$$

$$k_c = 2 \sin^{-1}(\sqrt{a})$$

and where k_{xc} is the maximum wavenumber of interest of the circular symmetric 2-D filter on the k_x -axis and k_c is the maximum wavenumber of interest of the 1-D operator. For a more detailed discussion on the derivation of the parameters in equation (1.46) the reader is referred to Hazra and Reddy (1986). In Figure 1.20 two contour plots are shown for two different frequencies. Up to the desired maximum value these contours are circular, outside the desired value the contours are not circular shaped.

In the extrapolation algorithm first the McClellan factors are calculated according to 1.46 for a given k_{xc} which gives besides the optimized McClellan factors also a k_c for the 1-D operator. With this calculated k_c value the 1-D operator is designed. Note that k_c is always smaller than k_{xc} . Due to the choice of the coefficients this 1-D operator is stretched to a correct 2-D operator. To compensate for this stretch the 1-D convolution operator must be calculated with a scaled Δz . This can be explained by regarding the effect of the optimized Hazra and Reddy factors as a scaling of the k_x -axes. The 1-D phase shift operator is then given by

$$\tilde{F}(k_x) = \exp(-j [k^2 - (\alpha k_x)^2]^{\frac{1}{2}} \Delta z)$$

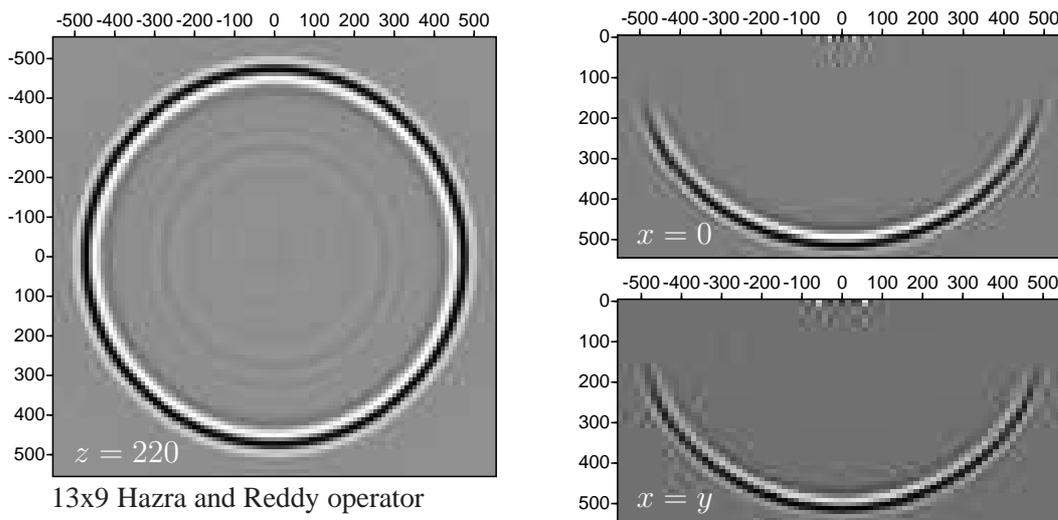


Fig. 1.21 Depth images obtained with the optimized Hazra and Reddy transformation. Note the improvement in the circular shape in the depth slice in comparison with the original McClellan transformation.

$$= \exp(-j [(\frac{k}{\alpha})^2 - k_x^2]^{\frac{1}{2}} \alpha \Delta z) \quad (1.46)$$

where $\frac{k}{\alpha} = k_c$ and $\alpha \Delta z$ is the scaled depth step.

The migration results for the Hazra and Reddy optimized coefficients are shown in Figure 1.21. The circular shape of the depth slice is good and the artefacts at the diagonals, which were present in the original McClellan transformation, are absent. Note that the computation time of the optimized coefficients is very small, only a few multiplications and additions per frequency, so the same computation effort is required as for the original McClellan transformation which makes the method very attractive. It will be interesting to investigate if it is possible to adjust the 17-points McClellan transformation with the same method.

1.4.2 Optimized McClellan factors

The aim in optimizing the McClellan transformation is to choose the McClellan factors c_{pq} in equation (1.38) such that the contours produced by the transformation have some desired shape. For some examples it is sufficient to control the shape of one single contour. In other problems it is necessary to design the shape of the contours in a specific part of the wavenumber domain (Mersereau et al., 1976, (Mersereau et al., 1976)). The error function which has to be optimized for a circular contour design is given by

$$E = M(c_{pq}, k_x, k_y) - \cos \sqrt{k_x^2 + k_y^2} \quad \text{with} \\ M(c_{pq}, k_x, k_y) = \sum_{p=0}^P \sum_{q=0}^Q c_{pq} \cos(k_x) \cos(k_y) \quad (1.47)$$

Equation (1.47) is a non-linear function of the unknown parameters, so a computation intensive non-linear optimization scheme must be used for the minimization. However, it is possible to reformulate the problem as a linear approximation problem to arrive at a sub optimum solution (Mersereau et al., 1976). In the example shown in this section a non-linear optimization scheme (CFSQP, written by C.T. Lawrence, J.L. Zhou and A.L. Tits, Version 2.0, february 1994) is used. With this scheme we try to optimize several contours within the band of interest and put the constraint

$$\left| \sum_{p=0}^P \sum_{q=0}^Q c_{pq} \cos(k_x p \Delta x) \cos(k_y q \Delta y) \right| \leq 1 \quad (1.48) \\ 0 \leq k_x \leq k_{x,N} (= \frac{\pi}{\Delta x}), 0 \leq k_y \leq k_{y,N} (= \frac{\pi}{\Delta y})$$

for all points of the mapping in the (k_x, k_y) plane. The contours to be optimized in the objective function of equation (1.47) are defined by the maximum wavenumber value of interest. With this definition of the optimization problem the first order McClellan transformation cannot be optimized any further, but the second order transformation (with an expansion to 25 points, which means that all cross terms within the second order are used) can be improved. In the implementation of the McClellan operators we make explicitly use of the circular symmetry in McClellan

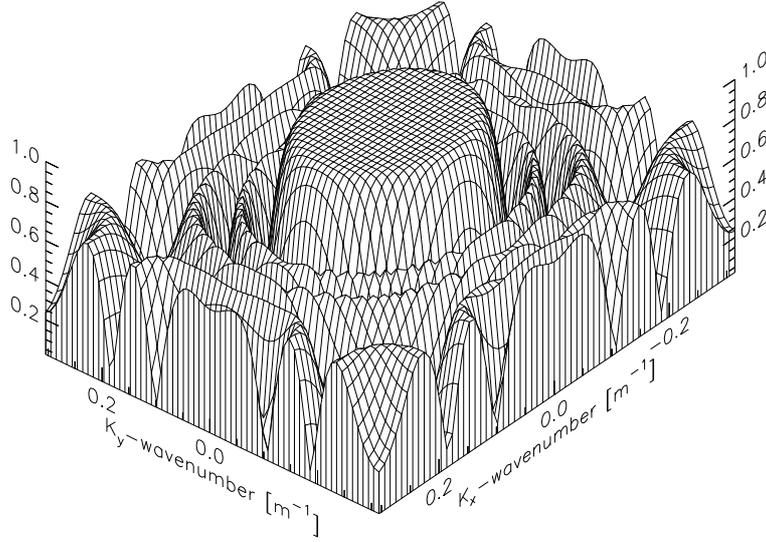


Fig. 1.22 The wavenumber spectrum of a second order McClellan operator with optimized coefficients and a 1-D operator of 19 spatial points with 128×128 k_x, k_y points with $c = 1000 \text{ m s}^{-1}$, $f = 20 \text{ Hz}$, $\Delta x = \Delta y = \Delta z = 10 \text{ m}$ and a maximum angle of interest of 65° .

operator if $\Delta x = \Delta y$. In Figure 1.23 the migration puls response is shown for the optimized 25-point McClellan operator with frequency dependent coefficients and a 1-D operator of 25 points (13 terms). The non-linear computation time can be reduced by calculating the optimized coefficients for a wavenumber range instead of every wavenumber. In the shown example we have used only four sub domains in the total wavenumber domain of interest which keeps the time to compute the McClellan factors small. It also possible to optimize the coefficients independent of the wavenumber frequency or for different shaped McClellan filters (Blacquière, 1991).

1.4.3 Rotated Coefficients

Biondi and Palacharla (1994) describe a method which reduces the error of the deviation of the circle at the diagonal in the (k_x, k_y) plane by using a rotated McClellan operator, the rotation angle being $\frac{\pi}{4}$. In the downward extrapolation scheme the rotated McClellan operator and the original McClellan filters are alternately used as convolution operator. The convolution with the rotated McClellan operator can be implemented in an efficient way. For a more detailed discussion the reader is referred to Biondi and Palacharla (1994).

1.4.4 Series expansion in $\cos(k_r \Delta x)$

All improvements described thus far make use of the Chebyshev recursion scheme, but it is also possible to use a direct expansion in $\cos(k_r \Delta x)$. To see the difference between the two schemes the Chebyshev recursion scheme and the direct scheme are given

$$\tilde{F}(k_r) \approx \sum_{m=0}^M F_m T_m(\cos(k_r \Delta x)) \quad (1.49)$$

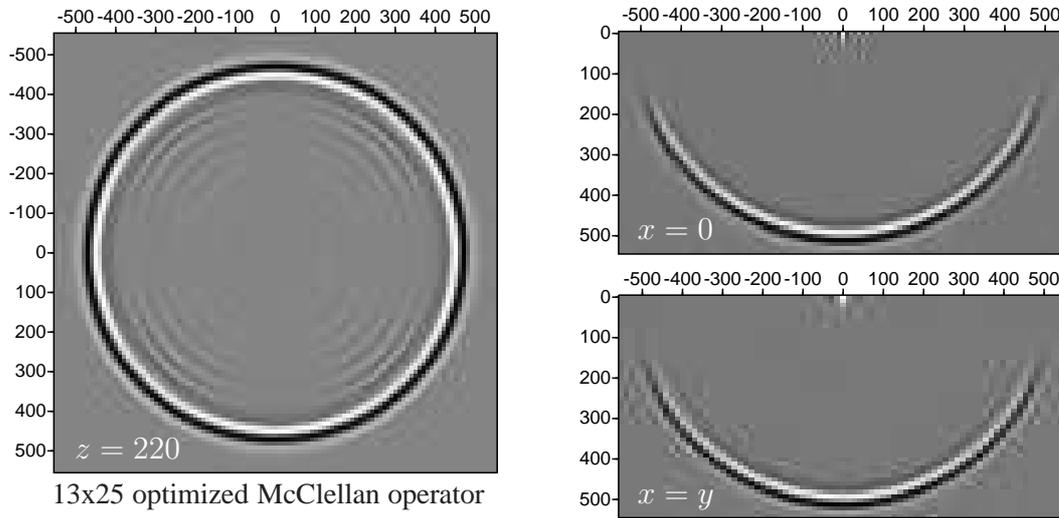


Fig. 1.23 Depth slice at $z = 220$ m, a vertical slice for $x=0$ (top right) and bottom right a vertical slice for $x=y$ with optimized McClellan coefficients. Note the circular shape and the small artefacts.

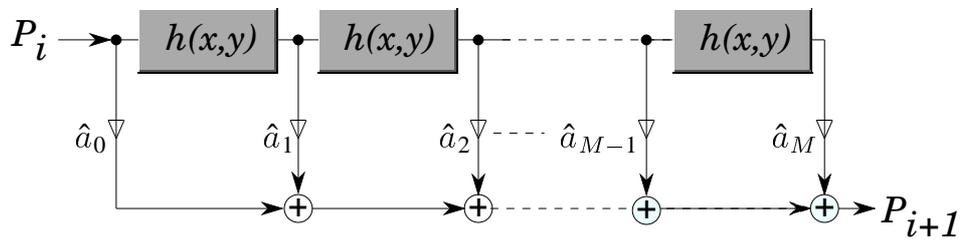


Fig. 1.24 Direct scheme for series expansion in $\cos(k_r \Delta x)$. Note the simple structure in comparison with the Chebychev recursion scheme.

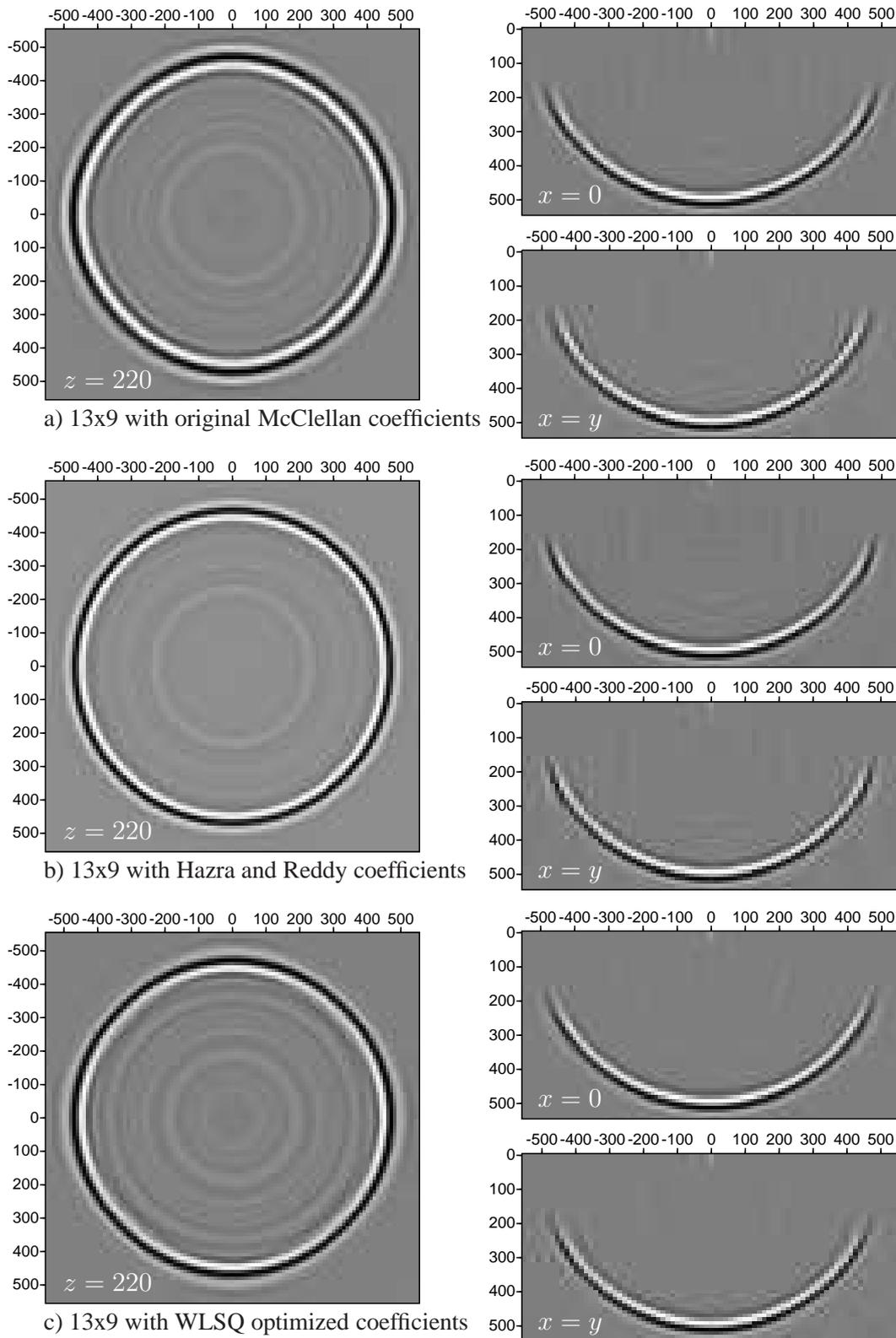


Fig. 1.25 Impulse response of the expansion in $\cos(k_r \Delta x)$ with original McClellan, Hazra and Reddy and WLSQ optimized coefficients for $\cos(k_r \Delta x)$. The series coefficients in all examples are optimized by using the WLSQ method.

$$\approx \sum_{m=0}^M a_m \cos(k_r \Delta x)^m \quad (1.50)$$

$$\approx \sum_{m=0}^M \hat{a}_m \tilde{h}^m(k_x, k_y) \quad (1.51)$$

with

$$\tilde{h}(k_x, k_y) \approx \cos(k_r \Delta x) \quad (1.52)$$

In equation (1.51) \tilde{h} (defined in equation (1.52)) is optimized first and with this approximation to $\cos(k_r \Delta x)$ \hat{a}_m is optimized. In equation (1.49), the Chebychev recursion scheme, the optimization of $\cos(k_r \Delta x)$ is independent of the optimization of F_m . The scheme of equation (1.51) a better operator can be designed, because the optimization of the series coefficients is dependent on the approximation to $\cos(k_r \Delta x)$. To optimize the \hat{a}_m the WLSQ method is used. To approximate $\cos(k_r \Delta x)$ the McClellan or Hazra and Reddy transformation can be used, but it is also possible to use the WLSQ optimization method described in section 1.3.1.

The recursive convolution scheme of equation (1.51) is given in Figure 1.24. This scheme is less complicated to implement in the computer and more important it can be optimized better by the compiler (see Appendix A). In Figure 1.25a the impulse response is shown with the McClellan coefficients of equation (1.41) and in Figure 1.25b the impulse response is shown with the Hazra and Reddy optimized coefficients for the approximation to $\cos(k_r \Delta x)$. The number of terms is in both schemes equal to 13. Note that with this method the results are better circular and contain only small artefacts even with the small (9) operator. In Figure 1.25c the series expansion in $\cos(k_r \Delta x)$ is done with an approximation to $\cos(k_r \Delta x)$ obtained with the WLSQ method introduced in section 1.3.1. Note that with a 9 point approximation to $\cos(k_r \Delta x)$ the result is accurate upto the higher angles. In the next subsection a detailed error analysis is given for all discussed McClellan methods.

1.4.5 Error analysis

Using the analysis technique, which was introduced by the direct method, it is possible to analyze the performance of the different McClellan transformations. In Table 1.7 the errors are given for five types of McClellan transformations; the original McClellan transformation in first (*x9) and second (*x17) order, the optimized Hazra & Reddy factors (HR) in first order, the non-linear optimized factors in the expanded second order (*x25) and the series expansion in $\cos(k_r \Delta x)$ with the Hazra and Reddy coefficients and the WLSQ operators. From the results in the Table and Figures 1.18, 1.21, 1.23, 1.25 the following remarks can be made;

- (1) For higher frequencies the ε^p error in the original McClellan transformation increases significantly for both the first and second order approximation.
- (2) The differences in pulse responses between the methods with, the original McClellan factors, the Hazra & Reddy factors, the non-linear optimized factors or the series expansion can be determined from the ε^p error. In Figure 1.26 the phase error and the $\frac{\partial}{\partial k_r}$ of the phase error for the different methods are displayed at a frequency of 40 Hz mximum, a maximum angle of 60°

Operator	5 Hz			20 Hz			40 Hz		
	ε_2	ε_∞^a	ε^p	ε_2	ε_∞^a	ε^p	ε_2	ε_∞^a	ε^p
McClellan									
10x9	1.7e-3	1.0e-3	4.5e-4	6.6e-3	2.2e-3	4.8e-3	5.1e-2	1.5e-3	3.8e-2
13x9	1.4e-3	1.2e-3	3.2e-4	6.3e-3	1.7e-3	3.9e-3	5.2e-2	1.4e-3	3.8e-2
16x9	1.1e-3	1.3e-3	2.3e-4	6.3e-3	1.1e-3	3.7e-3	5.1e-2	1.6e-3	3.9e-2
10x17	1.7e-3	1.0e-3	4.4e-4	1.8e-3	2.4e-3	2.7e-3	2.4e-2	1.5e-3	2.2e-2
13x17	1.3e-3	1.2e-3	3.1e-4	1.1e-3	2.0e-3	1.4e-3	2.4e-2	1.6e-3	2.3e-2
16x17	1.1e-3	1.3e-3	2.3e-4	7.8e-4	1.4e-3	9.7e-4	2.4e-2	1.9e-3	2.3e-2
HR									
10x9	2.6e-3	1.2e-3	8.1e-4	7.2e-3	1.7e-3	2.8e-3	8.2e-2	7.7e-3	1.8e-2
13x9	1.2e-3	6.7e-4	2.9e-4	7.1e-3	2.7e-3	2.3e-3	8.2e-2	1.6e-2	1.8e-2
16x9	1.4e-3	1.4e-3	3.2e-4	7.2e-3	1.7e-3	1.9e-3	8.2e-2	2.7e-2	1.8e-2
Optimized									
10x25	1.7e-3	1.0e-3	4.4e-4	1.8e-3	2.4e-3	2.6e-3	7.2e-3	1.5e-3	9.4e-3
13x25	1.4e-3	1.2e-3	3.2e-4	1.2e-3	2.0e-3	1.4e-3	7.2e-3	1.9e-3	8.6e-3
16x25	1.1e-3	1.3e-3	2.3e-4	9.2e-4	1.4e-3	1.0e-3	7.2e-3	2.3e-3	8.5e-3
Series HR									
10x9	5.2e-3	2.5e-3	1.6e-3	1.7e-3	1.8e-3	2.5e-3	8.4e-3	2.3e-3	6.7e-3
13x9	2.0e-3	1.5e-3	5.3e-4	9.8e-4	1.3e-3	1.4e-3	8.4e-3	2.4e-3	5.6e-3
16x9	1.5e-3	7.7e-4	3.6e-4	6.4e-4	1.2e-3	8.4e-4	8.4e-3	2.2e-3	5.2e-3
Series WLSQ									
10x9	2.8e-3	1.8e-3	8.2e-4	5.1e-3	1.7e-3	3.8e-3	2.2e-3	2.3e-3	4.5e-3
13x9	1.3e-3	6.2e-4	3.4e-4	5.0e-3	1.4e-3	3.4e-3	2.0e-3	1.6e-3	3.6e-3
16x9	5.3e-4	3.4e-4	1.1e-4	5.0e-4	4.1e-4	3.2e-3	1.9e-3	4.2e-4	3.4e-3
10x25	2.2e-3	1.2e-3	6.4e-4	8.3e-4	8.4e-4	1.2e-3	1.4e-3	1.6e-3	2.8e-3
13x25	1.3e-3	4.1e-4	3.2e-4	4.3e-4	6.9e-4	5.9e-4	1.2e-3	1.2e-3	1.5e-3
16x25	5.7e-4	2.7e-4	1.2e-4	1.9e-4	2.9e-4	2.6e-4	1.1e-3	3.4e-3	1.4e-3

Table 1.7 Errors in the extrapolation operators for; the original McClellan transformation in first (*x9) and second (*x17) order, the optimized Hazra & Reddy factors (HR) in first order, non-linear optimized factors in the expanded second order and the series expansion in $\cos(k_r)$ with the Hazra and Reddy coefficients and WLSQ operators. The maximum angle of interest is 60° .

and 13 terms in the expansion. The second order method with the original McClellan factors give a rapidly increasing phase error where the largest errors occurs at the diagonal from $(0, 0)$ to $(k_{x,N}, k_{y,N})$. The phase error of the Hazra & Reddy method is less rapidly increasing and the smallest error is positioned at the diagonal from $(0, 0)$ to $(k_{x,N}, k_{y,N})$. The non-linear optimized method gives error peaks at the edges of the domain of interest and has the largest error at the diagonal from $(0, 0)$ to $(k_{x,N}, k_{y,N})$. The series expansion method with a WLSQ approximation to $\cos(k_r \Delta x)$ gives the smallest errors with error peaks at the edges on k_x and k_y axes.

(4) The approximation to $\cos(k_r \Delta x)$ can be done with many different methods. Crucial in the

Fig. 1.26 *Phase errors and the radial derivative of the phase error for the different McClellan methods.
Note the different scales on the vertical axes.*

performance of the operator is that the coefficients in the expansion (Chebyshev or series) are optimized by using the approximation to $\cos(k_r \Delta x)$.

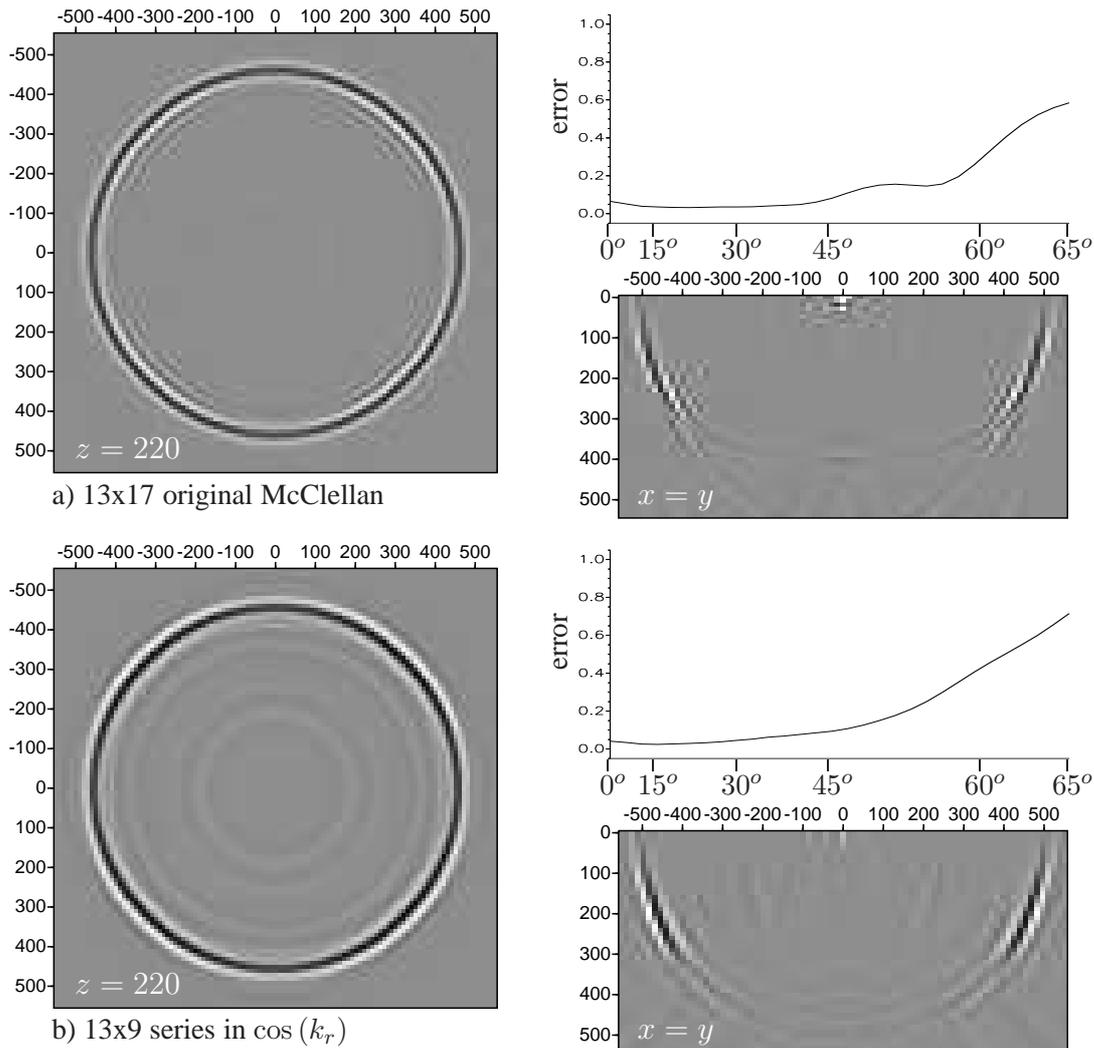


Fig. 1.27 Error in the impulse response of the Chebychev expansion in $\cos(k_r \Delta x)$ with original McClellan (a) and for a series expansion in $\cos(k_r \Delta x)$ with WLSQ optimized coefficients (b).

In Figure 1.27 the spatial error is given for a horizontal cross section at 65° and a vertical cross section at $x = y$ for both the original McClellan transformation (13x17) and the series expansion in $\cos(k_r \Delta x)$ (13x9). The original McClellan scheme gives dispersive artefacts for the higher angles. The expansion in $\cos(k_r \Delta x)$ (with less coefficients in the approximation to $\cos(k_r \Delta x)$) does not have these artefacts but is less accurate in amplitude for the higher angles.

In Table 1.8 the shortest accurate operator is given as function of the maximum angle of interest. For small angles the original first order McClellan transformation in Chebyshev series gives already good results. For intermediate angles the second order McClellan scheme or the first order Hazra & Reddy factors are sufficient, for higher angles the series expansion with WLSQ optimized series coefficients gives the best results. The Hazra & Reddy factors have a large ε_2 for the higher frequencies due to stretching of the operator. For the 75° angle a 5x5 approximation to $\cos(k_r \Delta x)$ is needed.

1.5 Single series expansion

In the previous sections the direct convolution and the McClellan transformation were explained and impulse responses were shown for different operators. The McClellan transformation uses the coefficients of a 1-D convolution operator and approximates the 2-D Fourier components with an optimum filter. However, it is also possible to approximate the phase shift operator with an expansion other than the cosine terms of the Fourier transformation. Writing space-frequency wave field extrapolation in an operator notation (Berkhout, 1982).

$$P^+(z_{m+1}) = W^+(z_{m+1}, z_m) * P^+(z_m) \quad (1.53)$$

where $W^+(z_{m+1}, z_m)$ is the propagation operator and $P^+(z_m)$ is the down going wave field at depth level z_m . In this notation the most simple approximation to finite difference wave field extrapolation is made by a single Taylor series expansion. In the spatial domain with $\Delta z = z_{m+1} - z_m$ this approximation is given by

$$P^+(z_{m+1}) \approx P^+(z_m) + \frac{\Delta z}{1!} \frac{\partial P^+(z_m)}{\partial z_m} + \frac{\Delta z^2}{2!} \frac{\partial^2 P^+(z_m)}{\partial z_m^2} + \frac{\Delta z^3}{3!} \frac{\partial^3 P^+(z_m)}{\partial z_m^3} + \dots \quad (1.54)$$

The extrapolation scheme given in (1.54) can be divided into two parts; one part deals with the **estimation** of the derivatives $\frac{\partial^i}{\partial z_m^i}$ with respect to z_m and the other part deals with the **prediction** with the aid of the Taylor series. In the wavenumber domain this approximation, with $\frac{\partial \tilde{P}^+}{\partial z_m} =$

Operator			40 Hz			Average		
angle	size	method	ε_2	ε_∞^a	ε^p	$\bar{\varepsilon}_2$	$\bar{\varepsilon}_\infty^a$	$\bar{\varepsilon}^p$
15	4x9	McC	1.7e-3	2.9e-3	1.5e-4	1.8e-3	2.0e-3	1.0e-4
30	5x9	McC	3.9e-3	2.0e-3	1.0e-3	3.1e-3	2.0e-3	5.4e-4
45	7x17	McC	5.2e-3	2.5e-3	3.3e-3	3.1e-3	2.0e-3	1.5e-3
60	10x9	Series	2.2e-3	2.3e-3	4.5e-3	3.4e-3	1.9e-3	3.0e-3
75	18x25	Series	6.9e-3	2.6e-3	2.6e-2	4.7e-3	1.9e-3	1.2e-2

Table 1.8 Optimum operators which are accurate up to a maximum angle of interest.

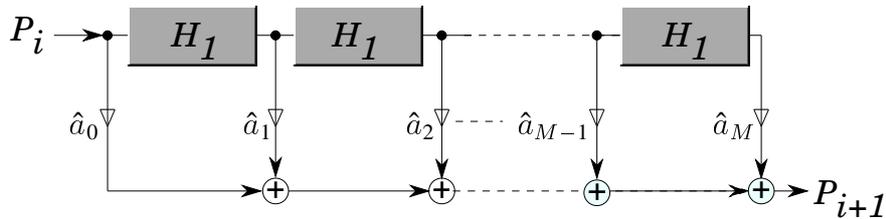


Fig. 1.28 Series expansion in terms of k_z . The H_1 boxes represent the 2-D convolution with the optimized spatial $k_z \approx \tilde{H}_1$ operator; the \hat{a}_m represent the optimized coefficients in the series expansion.

$-jk_z\tilde{P}^+$, is defined as

$$\tilde{P}^+(z_{m+1}) \approx \tilde{P}^+(z_m) - \frac{jk_z\Delta z}{1!}\tilde{P}^+(z_m) - \frac{(k_z\Delta z)^2}{2!}\tilde{P}^+(z_m) + j\frac{(k_z\Delta z)^3}{3!}\tilde{P}^+(z_m) + \dots \quad (1.55)$$

This truncation of the series expansion is an approximation of the phase shift operator by a polynomial in k_z , according to

$$\exp(-jk_z\Delta z) \approx 1 - j\Delta zk_z + \frac{(j\Delta z)^2}{2!}(k_z)^2 + O(k_z)^3 \quad (1.56)$$

The coefficients in the series expansion can be obtained by using the constants in the Taylor series as given in equation (1.56) or by using an optimization technique. In the next two subsections two different optimization methods are used; the WLSQ method with the L_2 norm and the Remez exchange method with the L_∞ norm.

1.5.1 Expansion in k_z with L_2 -norm

The advantage of an expansion in k_z is that if the k_z operator can be approximated by a short spatial convolution operator and the number of terms in the series expansion of the phase shift operator remains small the computation time can be reduced in comparison with the direct 2-D convolution. To arrive at the direct spatial convolution scheme, which is given in Figure 1.28, the k_z operator is transformed in an optimum way to the space domain and applied several times to the data. Every time a 2-D convolution (indicated by the box H_1) with the spatial k_z operator is carried out on the data, a new term is added to the series expansion. The scheme given in Figure 1.28 is more sensitive to numerical errors than the Chebyshev recursion scheme but if the number of terms remains small the Chebyshev structure will not improve the result (in the next subsection we will use the Chebyshev recursion scheme). It is interesting to note that an exact

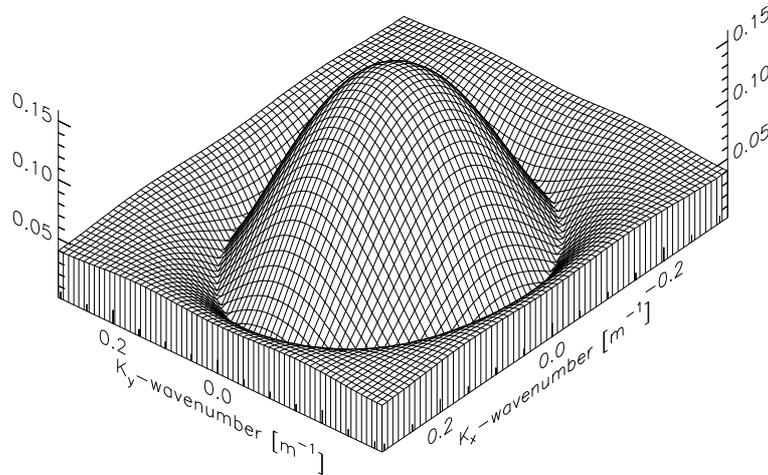


Fig. 1.29 Wavenumber spectrum of a circular Fourier reconstructed spatial k_z convolution operator (5x5).

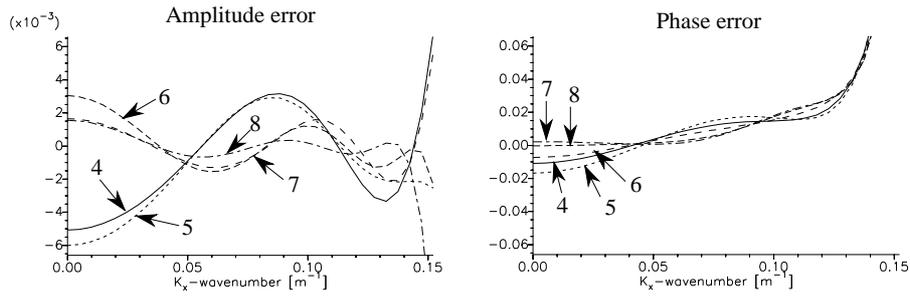


Fig. 1.30 Error in wavenumber and phase spectrum for $k_y = 0$ with 4,5,6,7 and 8 terms in the series expansion and WLSQ optimized coefficients for the circular Fourier reconstructed k_z operator of size 5×5 .

analytical expression for $H_1(x, y)$ ($H_1 \iff \tilde{H}_1 = k_z$) can be derived (see Berkhout (1982), appendix E). However, a weighted least-squares version yields a shorter operator.

The factors \hat{a}_m of the series expansion in Figure 1.28 are obtained by a least-squares optimization method with respect to the wavenumber spectrum of the *optimized* H_1 operator. The circular Fourier reconstruction of the k_z operator results in a circular short spatial convolution operator, the spectrum of an operator is shown in Figure 1.29 for a 5×5 operator. Using this 2-D convolution operator the computational effort can be reduced in comparison with the direct method if the number of terms in the series expansion remains small. It is therefore interesting to determine how the different parameters must be chosen for a stable and accurate operator. The error in the wavenumber and phase spectrum and for $k_y = 0$ is shown in Figure 1.30 with 4,5,6,7 and 8 terms in the series expansion and WLSQ optimized coefficients for the circular Fourier reconstructed k_z operator of size 5×5 . Note that for more terms in the series expansion both

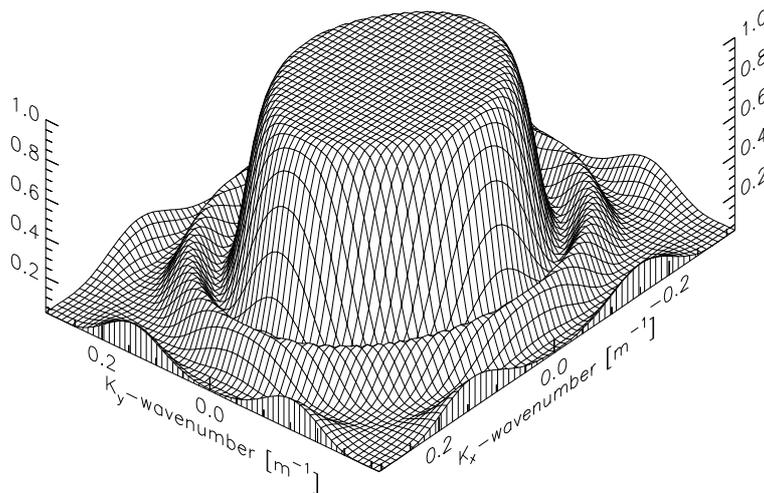


Fig. 1.31 Wavenumber spectrum of the approximated phase shift operator with 9 terms in the series expansion, 128×128 k_x, k_y points with $c = 1000\text{ms}^{-1}$, $f = 25\text{Hz}$, $\Delta x = \Delta y = \Delta z = 10\text{m}$ and a maximum angle of interest of 65° .

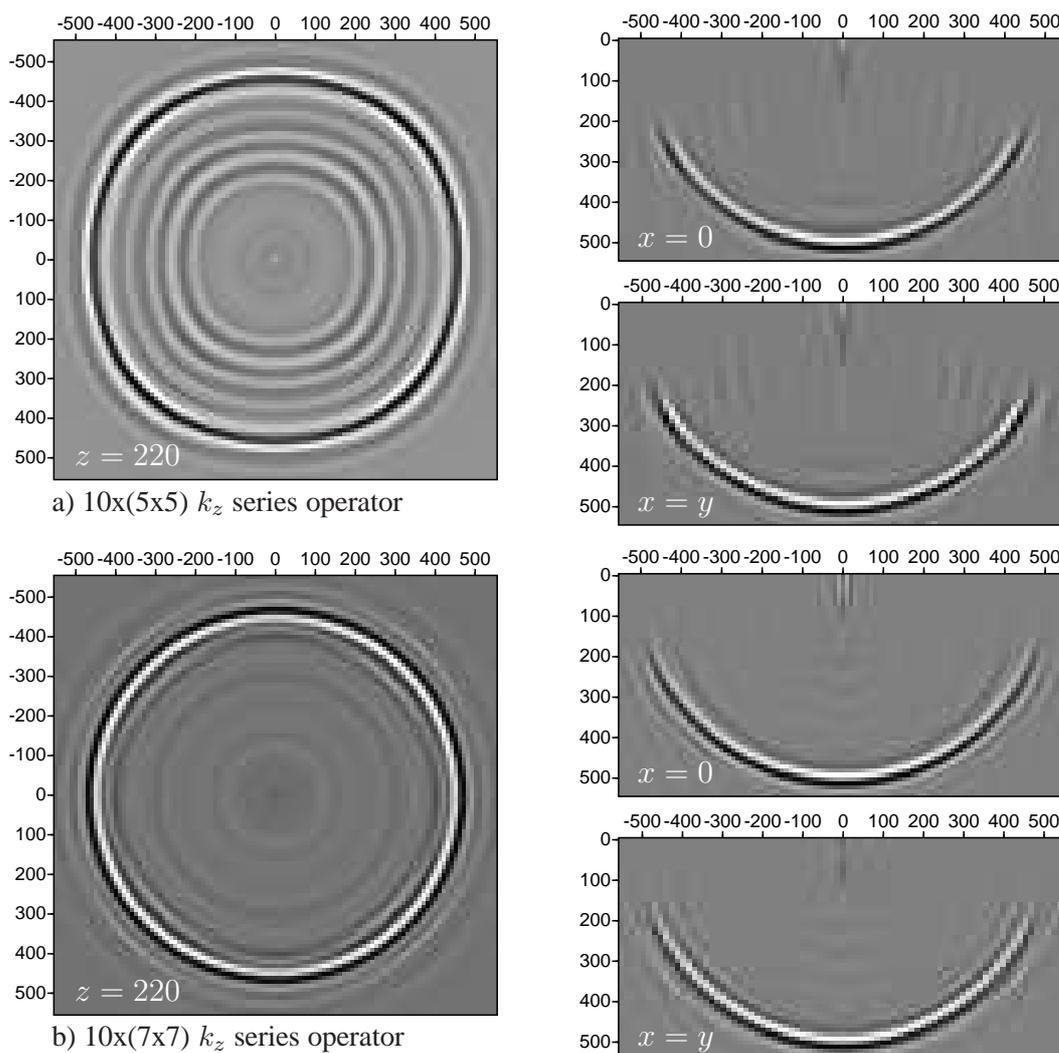


Fig. 1.32 Depth images of pulse responses obtained with the series expansion in k_z . a) shows a short basis operator (5×5) with 10 terms in the series expansion in b) a longer basis operator (7×7) is used which gives a better circular shape and less artefacts.

the amplitude and phase error decrease. Beyond a certain number of terms the error remains the same while the number of terms is increased. In that case there is not a better approximation possible with the optimized basis function of H_1 . For a better approximation a better basis function must be chosen.

In Figure 1.31 the wavenumber spectrum of the approximated phase shift operator is shown for a nine order series expansion with optimized coefficients for the *approximated* k_z operator with a spatial size of 5×5 points (the same parameters are used as in the example for the direct method). Outside the band of interest ($\alpha_{max} = 65^\circ$) the exact k_z operator (which is used as object function in the optimization) is tapered to zero. Note that with only nine terms there is already a good match with the analytical spectrum.

In Figure 1.32 two pulse responses are shown for different parameters in the single series expansion method. The convolution with the short basis function of 5×5 points gives an image

with a non-circular depth slice. Increasing the number of terms in the series expansion from 10 to 15 will reduce the artefacts around the event, but the non-circular behavior remains the same. The longer basis function with 7×7 points gives a better circular slice, but requires more computation time.

1.5.2 Expansion in k_z with L_∞ -norm

Using an optimization method which makes use of the L_∞ norm and the Chebyshev recursion structure for the implementation of the convolution may improve the result. Optimization for the series expansion terms with the L_∞ norm can be done by reducing the polynomial synthesis to symmetrical spectral synthesis with the aid of a simple transformation (see Appendix C). In this optimization only the extreme values of the wavenumber spectrum of the H_1 operator are used and not, as with the L_2 optimization, the whole spectrum of the H_1 operator. The transformation from polynomials to spectral synthesis reduces the 2-D optimization problem to a 1-D optimization problem which can be solved with the Remez exchange algorithm in a fast way. The terms in the series expansion and the spatial k_z operators are calculated for every frequency.

In Figure 1.33 one impulse responses is shown for a 5 point basis convolution operator with 10 terms in the expansion. The basis operator is not accurate enough to define the circular shape properly. A 7 point operator with the same number of terms will give a better result (see the error analysis in the next subsection). The difference with L_2 optimization is that due to the equiripple character of the L_∞ solution the error is smeared over the whole wavenumber domain. In the L_2 optimization the biggest error occurs at the edges of the domain of interest (defined by $k = \frac{\omega}{c}$ and α) which gives the artefacts as shown in Figure 1.32. Note that more terms

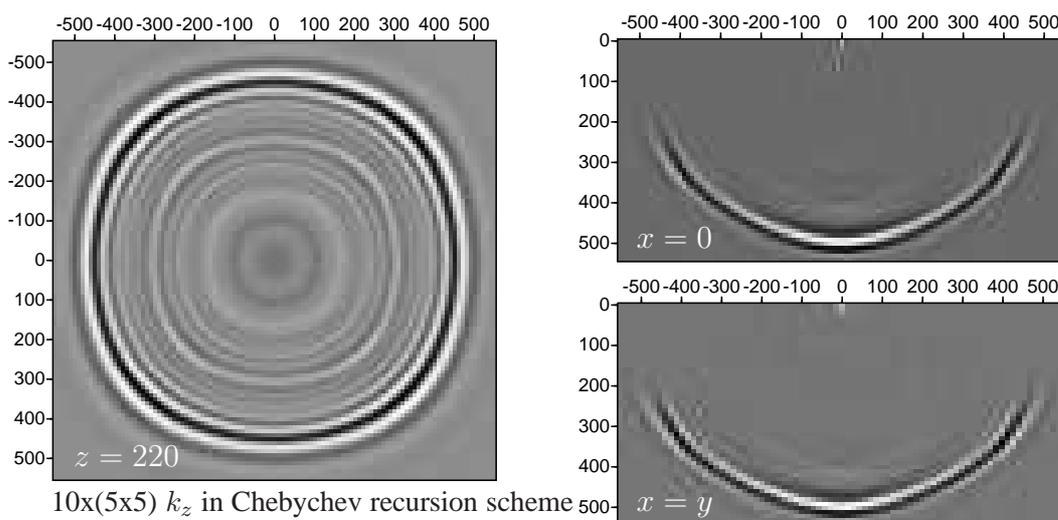


Fig. 1.33 Depth image of an impulse response obtained with the Chebyshev recursion scheme and a L_∞ optimization for the coefficients in the expansion in k_z , a short basis operator (5×5) with 10 terms is used.

in the series expansion with the same basis operator will reduce the artefacts but it will not improve the circular shape.

1.5.3 Error analysis

In Table 1.9 the errors are given for the series expansion in k_z with the direct implementation and the Chebyshev recursion scheme. From the results in the Table the following remarks can be made;

- (1) The expansion in k_z gives large ε_2 errors for most frequencies. So the operator is not very accurate which can be observed in the artefacts in the impulse responses.
- (2) Increasing the number of terms with the same size of the basis operator does not improve the result. Increasing the size of the basis operator with the same number of terms gives an improvement. This means that the approximation to the basis operator is the most important factor in the performance of the operator.
- (3) The ε^p error indicates that the used basis operator of size 5×5 in the single series expansion does not have a good circular shape. The larger operator of size 7×7 gives an improvement but is still inaccurate.
- (4) Optimization of the series coefficients with the L_∞ norm and the use of the Chebyshev recursion scheme gives a better result than L_2 optimization and a direct recursion scheme.

size	5 Hz			20 Hz			40 Hz		
	ε_2	ε_∞^a	ε^p	ε_2	ε_∞^a	ε^p	ε_2	ε_∞^a	ε^p
10x(5x5)	1.8e-3	1.8e-3	4.3e-4	1.5e-2	3.6e-3	7.5e-3	2.0e-2	5.6e-3	7.5e-3
15x(5x5)	7.4e-4	6.3e-4	1.4e-4	1.5e-2	3.9e-3	7.6e-3	2.0e-2	5.2e-3	7.1e-3
10x(7x7)	5.8e-4	1.2e-3	1.1e-4	7.6e-3	1.8e-3	3.5e-3	1.2e-2	1.9e-3	8.9e-3
15x(7x7)	4.7e-4	1.2e-3	6.5e-5	7.6e-3	3.8e-3	3.5e-3	1.2e-2	4.0e-3	8.6e-3
Chebyshev	5 Hz			20 Hz			40 Hz		
10x(5x5)	1.9e-2	2.1e-3	2.8e-3	2.1e-2	2.8e-3	9.1e-3	3.5e-2	2.5e-3	5.2e-2
15x(5x5)	2.0e-2	2.2e-3	2.7e-3	2.1e-2	4.3e-3	8.9e-3	3.5e-2	2.5e-3	5.3e-2
10x(7x7)	9.2e-2	2.5e-3	2.6e-3	3.6e-2	1.1e-2	1.4e-2	1.4e-2	2.7e-3	2.1e-2
15x(7x7)	9.2e-3	2.1e-3	2.7e-3	3.6e-2	1.5e-2	1.4e-2	1.3e-2	2.4e-3	2.1e-2

Table 1.9 Errors in the extrapolation operators for the direct series expansion in k_z and a Chebyshev recursion scheme in k_z . The maximum angle of interest is 60° .

The shortest accurate operator as function of the maximum angle of interest is not worked out for the expansion in k_z because this method is cannot be designed accurate enough within a reasonable computation effort.

1.6 Double series expansion

By using an additional series expansion, i.e. k_z is expanded in terms of $k_x^2 + k_y^2 (= k_r^2)$ (see equation (1.57)),

$$k_z = k \sqrt{1 - k_r^2} \approx k \left(1 - \frac{k_r^2}{2} - \frac{(k_r^2)^2}{8} - \frac{(k_r^2)^3}{16} - \frac{5(k_r^2)^4}{128} - \frac{7(k_r^2)^5}{256} + O((k_r^2)^6) \right) \quad (1.57)$$

there is an extra advantage (Berkhout (1982), chapter 10). The basic spatial convolution operators are reduced to the simple 1-D convolution operators: $d_2(x)$ and $d_2(y)$. The double series expansion in $k_x^2 + k_y^2$ is given in equation (1.58)

$$\begin{aligned} \exp(jk_z \Delta z) &\approx 1 - \\ &\frac{j\Delta z}{2k} (k_x^2 + k_y^2) - \\ &\frac{j\Delta z}{8k^3} (1 - jk\Delta z) (k_x^2 + k_y^2)^2 - \\ &\frac{j\Delta z}{16k^5} \left(1 - jk\Delta z + \frac{(jk\Delta z)^2}{3} \right) (k_x^2 + k_y^2)^3 + \\ &O((k_x^2 + k_y^2)^4) \end{aligned} \quad (1.58)$$

where the terms of the series expansion are derived from the Taylor series, but this is not an optimum choice (Hoff, 1995) This same expansion can also be regarded as an approximation to the cosine terms in equation (1.35). So there are two different ways to look at the double series expansion

$$\tilde{F}_0(k_x, k_y) = \exp(-jk_z \Delta z) \quad (1.59)$$

$$\approx \sum_{m=0}^M a_m T_m(\cos(k_r)) \quad (1.60)$$

$$\approx \sum_{m=0}^M b_m T_m(k_r) \quad (1.61)$$

$$\approx \sum_{m=0}^M c_m k_r^m \quad (1.62)$$

where T_m is a Chebyshev polynomial of the m 'th order. In subsection 1.6.1 equation (1.62) is regarded with the series expansion scheme. In subsection 1.6.2 equation (1.61) is discussed with the Chebyshev recursion scheme.

1.6.1 Expansion in $k_x^2 + k_y^2$ with L_2 -norm

The same techniques as discussed in the section with the series expansion in k_z can be used again: the spatial versions of k_x^2 and k_y^2 , i.e. $d_2(x)$ and $d_2(y)$, are determined by a weighted least-squares process. In Figure 1.34 two wavenumber spectra are shown for short spatial convolution operators, with operator lengths of 5 and 7 points, which represent the second order differentiation. Note that for these short operator the approximation to the exact function within the band of interest is within a reasonable error. The convolution scheme is given in Figure 1.35 where

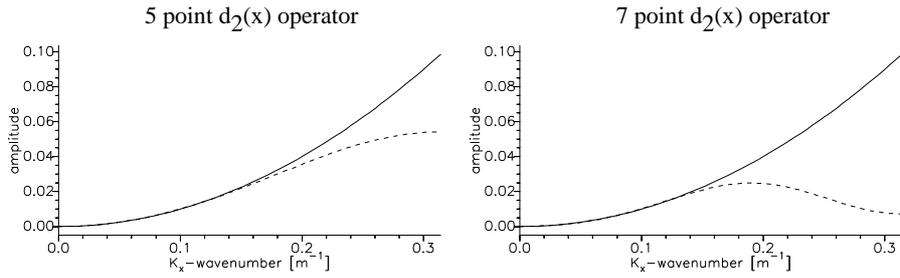


Fig. 1.34 Spectrum of two differentiation operator for 5 and 7 points. The maximum wavenumber of interest is given by $k_{max} = \frac{2\pi 25}{1000} \approx 0.16$.

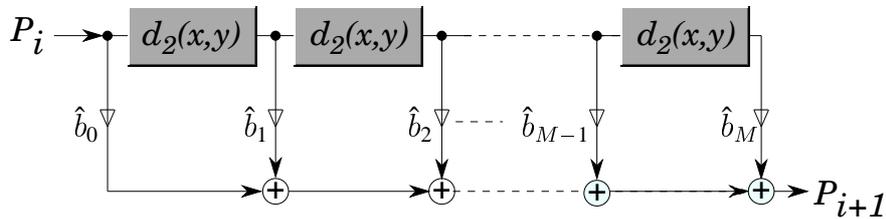


Fig. 1.35 Series expansion in terms of $k_x^2 + k_y^2 = \tilde{L}$. The $d_2(x, y)$ boxes represent the two 1-D convolutions with the optimized Laplacian operator, the \hat{b}_m represent the optimized coefficients in the series expansion.

$d_2(x, y)$ stands for the spatial Laplacian operator $d_2(x) + d_2(y)$. The factors \hat{b}_m of the series expansion in Figure 1.35 are obtained by a weighted least-squares optimization method with respect to the wavenumber spectrum of the *optimized* $k_x^2 + k_y^2$ operator. The length of the 1-D convolution operators depends on the maximum wavenumber of interest (Hoff, 1995). Hoff (1995) also showed that the value of the coefficients in the series expansion grows rapidly with increasing order. For the higher order terms values in the order of $1e19$ are normal. This means that after a certain number of terms the accuracy cannot be improved any further.

The scheme proposed in this section is similar to that of Sollid and Arntsen (1994), but there are some small differences; to obtain the series coefficients the whole wavenumber spectrum is used while Sollid and Arntsen use only the diagonal $(0, 0)$ to $(k_{x,N}, k_{y,N})$. For the least squares optimization we use a (fast) WLSQ scheme while Sollid and Arntsen use a non-linear scheme and finally the implementation is done in a series expansion and not in an Chebychev recursion scheme.

In Figure 1.36 the impulse responses is shown for a 3,5,7,9,11 combination with 15 coefficients. The impulse response shows artefacts at the higher angles which are due to edge effects of the used WLSQ method. In WLSQ design the edges of the domain of interest contain relative large error peaks. It may therefore be better to use an L_∞ norm in the design of the series coefficients. In the next subsection this second approach, after Soubaras, with a Chebychev recursion scheme is explained and results are shown.

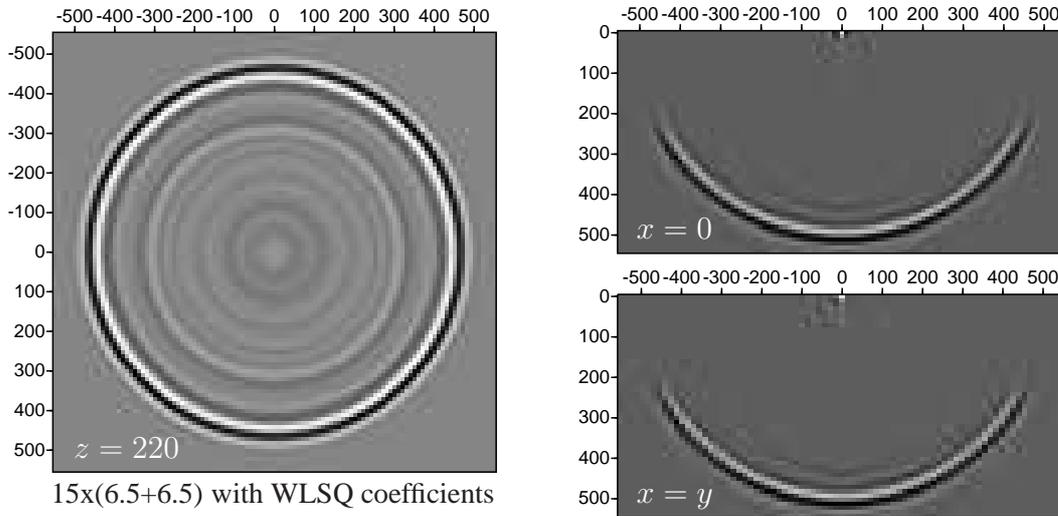


Fig. 1.36 Depth images of pulse responses obtained with the series expansion in $k_x^2 + k_y^2$, with a basis operator with optimized operator lengths with 15 terms in the series expansion (example made by Jochum Hoff).

1.6.2 Expansion in $k_x^2 + k_y^2$ with L_∞ -norm

Soubaras (1992) used the same type of expansion in $k_x^2 + k_y^2$, but in his method the optimization technique for both the terms in the series expansion and the convolution operators is the Remez exchange algorithm with the L_∞ norm. The advantage of the Chebyshev recursion scheme, given in Figure 1.37 over the series expansion, which was discussed in the previous subsection, is that the coefficients in the Chebyshev expansion are less sensitive to numerical errors. A disadvantage are the increasing number of additions needed.

Optimization of the differentiation operators with the L_∞ norm gives equiripple operators. Optimization for the series expansion terms with the L_∞ norm can be done by reducing the polynomial synthesis to symmetrical spectral synthesis with the aid of a simple transformation (see

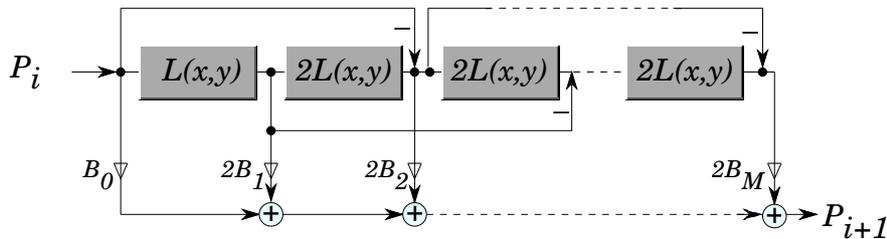


Fig. 1.37 Expansion in terms of $k_x^2 + k_y^2 = \tilde{L}$. The $d_2(x, y)$ boxes represent the two 1-D convolutions with the optimized L operator, the B_m represent the optimized coefficients in the series expansion.

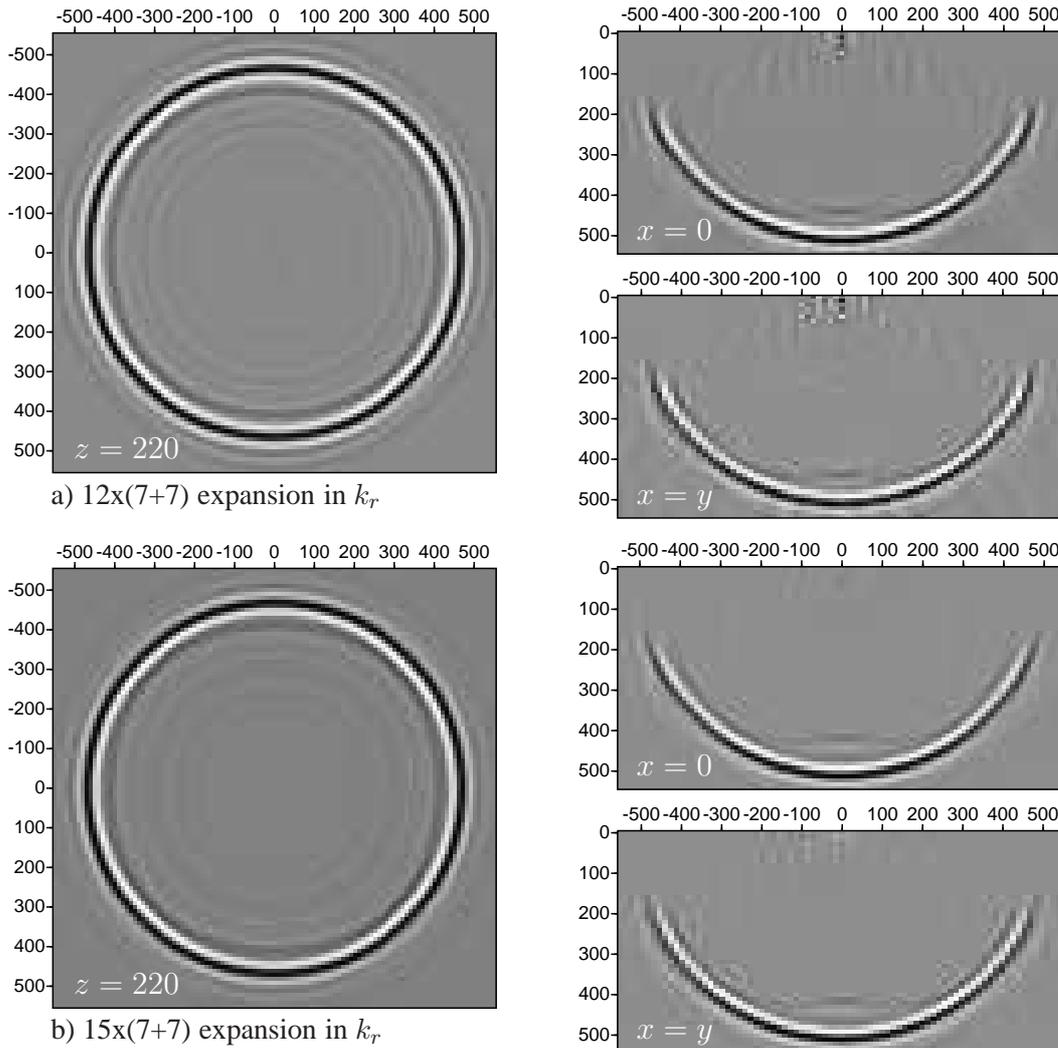


Fig. 1.38 Depth images of pulse responses obtained with the series expansion in $k_x^2 + k_y^2$ with L_∞ optimization; a) shows a 7 point 1D convolution operator with 12 terms in the series expansion and b) with the same basis operator and 15 terms in the series expansion. Note that more terms in the series expansion gives a better result.

Appendix C). In this transformation only the extreme values of the wavenumber spectrum of the differentiation operator are used in the optimization and not, as with the L_2 optimization the whole spectrum of the differentiation operator. The transformation from polynomials to spectral synthesis reduces the 2-D optimization problem to a 1-D optimization problem which can be solved with the Remez exchange algorithm in a fast way. In the extrapolation scheme the differentiation operator remains the same for all frequencies. However, the terms in the series expansion are calculated for every frequency.

In Figure 1.38 two pulse responses are shown for a 7+7 point operator with 12 and 15 terms in the expansion. The 7 point operator with 12 terms in the recursion scheme is shown in Figure 1.38a and has a circular response with only small artefacts. Using 15 terms with the same 7 point operator gives an even better result with less artefacts.

1.6.3 Error analysis

In Table 1.10 the errors are given for the series expansion in $k_x^2 + k_y^2$ implemented in a direct recursion scheme with L_2 optimization and the Chebyshev recursion scheme with L_∞ optimization. From the results in the Table the following remarks can be made;

- (1) In both optimization methods the number of terms is less essential to the accuracy than the length of the d_2 operator. Increasing the number of terms improves the result only a little, while increasing the length of the d_2 operator gives a significant improvement on the result.
- (2) The artefacts present in the impulse response with the series expansion scheme are due to instabilities at the higher angles. This effect is possibly due to the use of the L_2 norm optimization (Hoff 1994, (Hoff, 1995)). In the L_∞ norm optimization unstable error peaks are not likely to occur. The artefacts can be removed with an additional optimization step in the L_2 optimization.
- (3) Although the ε_2 errors for the lower frequencies in the L_∞ norm optimization are larger than the errors in the L_2 norm optimization the impulse response contains less artefacts.

The difference with the reference error in the spatial domain is given in Figure 1.39. In this Figure the series expansion method with 15x(3,5,7) terms and L_2 optimization can be compared with the Chebyshev recursion scheme of 15x(7+7) terms and L_∞ optimization. From this comparison we see that the errors in the L_∞ optimization are smeared out over the whole wavenumber range while the L_2 optimization has error peaks. The amplitude and phase accuracy is better for the L_2 optimization, but it suffers from artefacts at the higher angles. Choosing larger bass operator and more terms will solve this problem.

In Table 1.11 the shortest accurate operator in L_∞ optimization is given as function of the maximum angle of interest. The ε_2 error is the most sensitive error in the L_∞ optimization. For 15° and 30° angle the same number of terms must be used, a lower number of terms leads to unacceptable ε_2 errors for the low and middle frequencies. The optimum operator size can be found by trying to make the ε_2 as small as possible by choosing the number of terms high. The optimum number of terms is then found by lowering the number of terms until the ε_2 is changing

size	5 Hz			20 Hz			40 Hz		
	ε_2	ε_∞^a	ε^p	ε_2	ε_∞^a	ε^p	ε_2	ε_∞^a	ε^p
Direct									
10x(3,5,7)	5.8e-3	1.9e-3	1.8e-3	3.2e-3	3.7e-3	4.6e-3	3.4e-3	3.4e-3	6.4e-3
12x(3,5,7)	6.1e-3	2.8e-3	1.9e-3	3.0e-3	3.2e-3	4.1e-3	3.1e-3	3.4e-3	5.1e-3
15x(3,5,7)	6.1e-3	3.5e-3	1.9e-3	2.4e-3	3.2e-3	3.3e-3	3.0e-3	3.9e-3	4.8e-3
Chebyshev									
10x(7+7)	4.0e-2	3.7e-3	2.3e-3	7.1e-3	3.0e-3	4.5e-3	4.3e-3	1.9e-3	6.5e-3
12x(7+7)	4.0e-2	7.6e-3	2.1e-3	7.1e-3	5.0e-3	4.4e-3	4.1e-3	4.9e-3	5.8e-3
15x(7+7)	4.0e-2	1.2e-2	2.4e-3	6.8e-3	5.8e-3	3.4e-3	4.1e-3	6.5e-3	5.0e-3

Table 1.10 Errors in the extrapolation operators for the direct series expansion in $k_x^2 + k_y^2$ and a Chebyshev recursion scheme in $k_x^2 + k_y^2$. The maximum angle of interest is 60°.

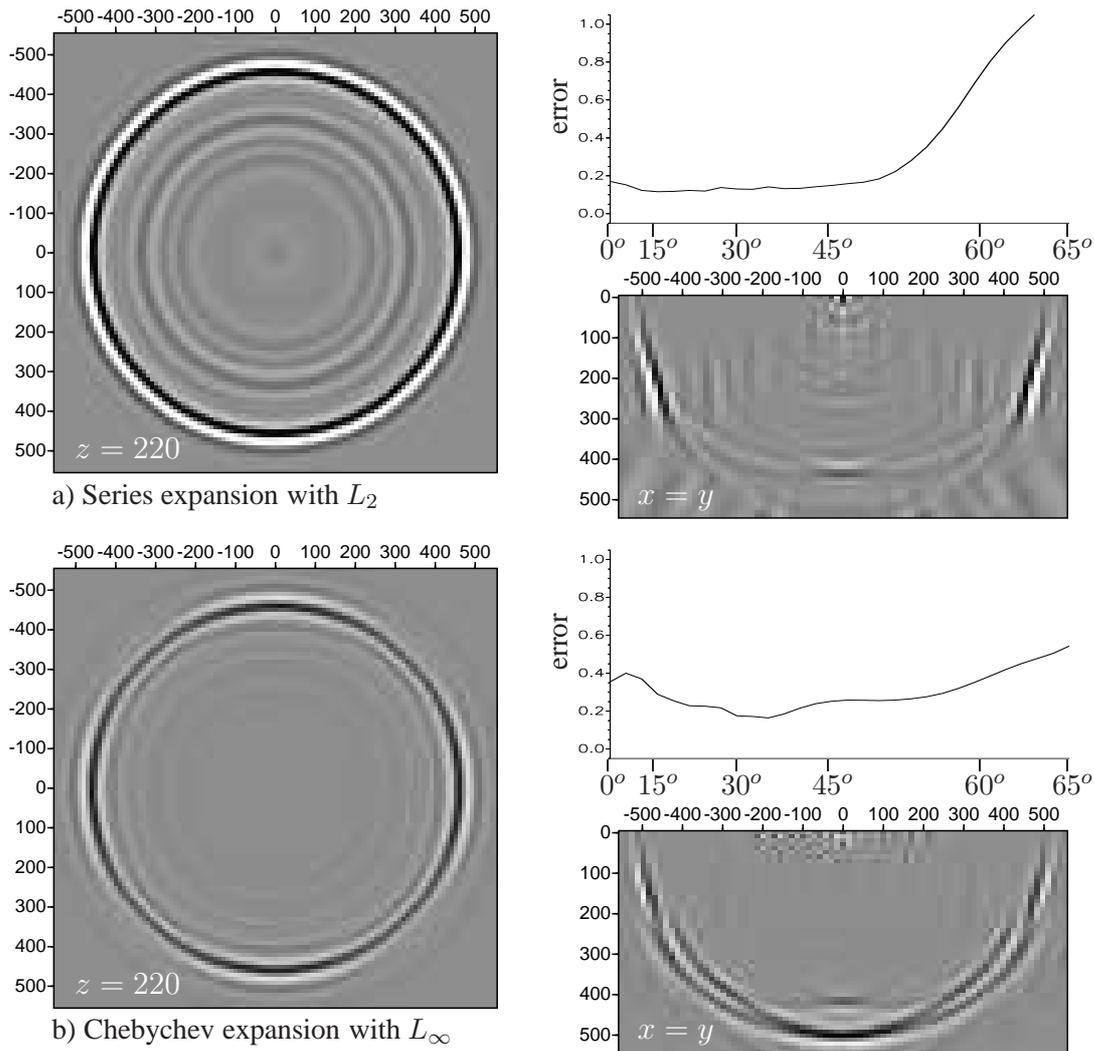


Fig. 1.39 Error in the impulse response of the Chebychev expansion in $\cos(k_r \Delta x)$ with original McClellan (a) and for a series expansion in $\cos(k_r \Delta x)$ with WLSQ optimized coefficients (b).

significant. The high average errors are due to the high errors at 5 Hz. For a maximum angle of 75° it is not possible to obtain stable operators.

Operator		20 Hz			Average		
angle	size	ε_2	ε_∞^a	ε^p	$\bar{\varepsilon}_2$	$\bar{\varepsilon}_\infty^a$	$\bar{\varepsilon}^p$
15	5x(7+7)	8.6e-3	5.1e-3	1.7e-4	9.6e-3	9.7e-3	5.0e-4
30	5x(7+7)	3.0e-3	1.9e-3	3.9e-3	1.2e-2	6.1e-3	3.7e-3
45	7x(7+7)	5.0e-3	4.5e-3	2.0e-3	1.5e-2	5.8e-3	1.0e-2
60	9x(7+7)	8.3e-3	1.8e-3	6.3e-2	1.7e-2	2.3e-3	4.0e-2

Table 1.11 Optimum operators which are accurate up to a maximum angle of interest.

1.7 Discussion and Conclusions

The 3-D extrapolation algorithm that is used in recursive depth migration can be implemented in several ways. In this Chapter the direct method, the McClellan transformation and three series expansion methods have been discussed. For the direct method a 2-Dimensional convolution operator is needed. The proposed weighted least-squares optimization method is an efficient procedure which gives stable and accurate convolution operators (Thorbecke and Rietveld, 1994). This method can be further improved by a second optimization step; for example, the Lawson algorithm (Rise and Usow, 1968 (Rice and Usow, 1968)), which will adjust the weight function in such a way that after several steps it will converge to a Chebyshev-norm solution see for example Algazi et. al 1986 (Algazi et al., 1986). The Fourier reconstruction method is a fast and simple method to obtain 2-D circular convolution operators but must be improved further to give accurate results.

The McClellan scheme which makes use of the 1-D optimized operator coefficients is attractive with respect to the computation effort and by using optimized McClellan factors the accuracy for the higher angles can be improved significantly without much effort. Using a series expansion of the phase shift operator also reduces the computation time in comparison with a direct 2-Dimensional convolution. Two expansions were discussed in this Chapter one in k_z and one in $k_x^2 + k_y^2$. These different approaches to the phase shift operator can be summarized in the following equations

$$\tilde{F}_0(k_x, k_y) = \exp(-jk_z \Delta z) \quad (1.63)$$

$$\approx \sum_{m=0}^M \sum_{n=0}^N F_{mn} \cos(k_x m \Delta x) \cos(k_y n \Delta y) \quad (1.64)$$

$$\tilde{F}_0(k_x, k_y) \approx \sum_{m=0}^M F_m T_m(\cos(\sqrt{k_x^2 + k_y^2} \Delta x)) \quad (1.65)$$

$$\approx \sum_{m=0}^M a_m \cos^m(\sqrt{k_x^2 + k_y^2} \Delta x) \quad (1.66)$$

$$\tilde{F}_0(k_x, k_y) \approx \sum_{m=0}^M B_m T_m(k_x^2 + k_y^2) \quad (1.67)$$

$$\approx \sum_{m=0}^M b_m (k_x^2 + k_y^2)^m \quad (1.68)$$

$$\tilde{F}_0(k_x, k_y) \approx \sum_{m=0}^M C_m T_m(k_z) \quad (1.69)$$

$$\approx \sum_{m=0}^M c_m k_z^m \quad (1.70)$$

with

$$\cos(\sqrt{k_x^2 + k_y^2}) \approx \sum_{p=0}^P \sum_{q=0}^Q c_{pq} \cos(pk_x) \cos(qk_y) \quad (1.71)$$

$$\cos(\sqrt{k_x^2 + k_y^2}) \approx a_0 + b_1 \sqrt{k_x^2 + k_y^2} \quad (1.72)$$

Equation (1.64) represents the direct method, equation (1.65) combined with equation (1.71) is the McClellan approach with the Chebyshev recursion scheme. Equation (1.66) is the series expansion in $\cos(k_r \Delta x)$ with a dependent optimization between the series coefficients and the approximation to $\cos(k_r \Delta x)$. Equation (1.67) with equation (1.72) is the expansion in $k_x^2 + k_y^2$ with the Chebyshev recursion scheme. Equation (1.68) is the series expansion in $k_x^2 + k_y^2$; the use of this series expansion in recursive migration was already proposed by Berkhout (1982). Equation (1.69) represents the expansion in k_z in a Chebyshev recursion scheme and Equation (1.69) the series expansion in k_z .

From an efficiency point of view the expansion in $k_x^2 + k_y^2$ and the use of the McClellan transformation are the most interesting schemes. An extra advantage of the expansion in $k_x^2 + k_y^2$ is that short 1-Dimensional convolution operators can be used. A disadvantage of these schemes is that it is not possible to write the algorithms in computer 'friendly' way due to the recursive structure in the scheme. This fact is displayed in the computation times given in Appendix A. The most accurate extrapolation is the direct convolution scheme. Another advantage of the direct scheme is that the algorithm can be designed in an efficient way. A disadvantage of the direct scheme is the intensive computation of the 2-D convolution operators.

The approximation to $\cos(k_r \Delta x)$ can be done with many different methods. Crucial in the performance of the extrapolation operator is that the coefficients in the expansion (Chebyshev or series) are optimized by using the approximation to $\cos(k_r \Delta x)$.

Method	accuracy	stable	circ	operator	simple	vector	scalar
Direct	++	+	++	-	++	+	+
McC1	o	++	o	+	o	+	++
McC2	o	++	o	+	o	+	++
McC2+	+	+	+	+	-	+	+
$\cos(k_r \Delta x)$	+ / ++	+	++	o	+	+	++
k_z	-	o	o	o	+	-	o
$k_x^2 + k_y^2 L_2$	o	o	+	+	o	o	+
$k_x^2 + k_y^2 L_\infty$	o	+	+	++	-	o	+

Table 1.12 Comparison of the different extrapolation methods with respect to computation effort and stability. Note that in the McClellan schemes the optimized McClellan factors are used.

In this table the different columns have the following meaning:

- *accuracy*: the average ε_2 error over the whole frequency range.
- *stable*: the stability of the method over the whole frequency range for all wavenumbers (ε_∞^a error). A o means that some wavenumber components canm become unstable.
- *circ*: the circularity of the impulse response (ε^p error). The McClellan in Chebyshev expansion and the k_z scheme have problems with the circulariry.
- *operator*: the effort to compute all the coefficients which are needed in the convolution scheme. For example in the direct method a 2-D convolution operator must computed, in the McClellan

scheme a 1-D convolution operator and the (optimized) McClellan factors are needed. In the table ++ means a minimum computation effort to compute the coefficients. Note that the operator coefficients can be calculated in advance and stored in an operator table.

- *simple*: the simplicity of the implementation of the convolution. The recursive schemes require more complex algorithms, so the compilers have to be good in optimization to make these schemes fast. In the recursive schemes it is difficult to make the program faster by changing the algorithm in a more computer 'friendly' way. A direct convolution requires more multiplications and additions but the algorithm can be made very efficient. This fact explains the fast computation time of the direct scheme in comparison with the other schemes (see Appendix A).
- *vector*: the performance of the scheme on a vector computer. The direct scheme is the only scheme which can be implemented in a vector efficient way.
- *scalar*: the performance of the scheme on a modern scalar computer. Note that some scalar computers may have an architecture which can be more advantageous for some implementations.

In conclusion taking into account the *computation time* of the different methods, the *simplicity* of the algorithms and most important the *accuracy* of the result then the **direct method** (1.64) is the best method for 3D extrapolation. The 2-D convolution operators should be stored in an efficient way, by using the even symmetry of the operator (one octant need be stored only), in an operator table that can be calculated in advance. Note that our algorithm of the direct scheme can be made faster by also taking the circular symmetry (if $\Delta x = \Delta y$) of the operator into account.

If a series expansion version is used we prefer the expansion in $\cos(k_r \Delta x)$ (1.66). It is our opinion that Chebychev recursion is not an advantage.

Appendix A: Computation times

The computation times of the different 3-D extrapolation methods in the space-frequency domain discussed in this chapter are given in table 1.13. The given time represent 55 recursive depth steps for one frequency (20 Hz.) with $c = 1000ms^{-1}$, $\Delta x = \Delta y = \Delta z = 10m$ on a x,y grid of 111×111 samples wide. All routines which are used are written in C and Fortran and are translated with the same type of compiler options without using options for parallel computation (see Table 1.14). However, parallel processing is easily implemented on the main frequency loop in the extrapolation algorithm. On the Convex (C-220) the -O2 option is used for vectorization of the loops. It was not possible to vectorize the C-code with specific compiler di-

	Direct			McClellan 1			McClellan 2		
Machine	19x19	25x25	31x31	10x9	13x9	16x9	10x17	13x17	16x17
SUN C	63.5	104.6	157.3	24.5	31.1	38.7	33.4	44.8	55.3
SUN F	52.6	88.0	144.8	20.3	27.0	31.7	28.8	39.0	47.8
Convex F	34.7	61.9	85.9	12.4	17.4	21.6	21.1	27.6	33.5
DEC C	28.3	46.9	71.6	9.4	13.7	15.7	13.4	17.8	21.5
DEC F	11.5	19.9	28.0	5.6	7.2	8.8	6.7	8.6	10.5
HP C	18.6	30.0	44.7	10.0	13.2	16.5	20.5	23.5	27.7
HP F	6.8	10.6	15.1	6.6	8.6	10.6	19.2	20.8	24.5
	McClellan 2+			$\cos(k_r \Delta x) \times 9$			$\cos(k_r \Delta x) \times 25$		
Machine	10x25	13x25	16x25	10x9	13x9	16x9	10x25	13x25	16x25
SUN C	41.9	55.0	67.2	18.0	23.5	29.2	35.0	45.9	59.9
SUN F	36.8	48.0	58.3	16.8	21.1	25.7	32.6	42.9	54.3
Convex F	33.4	45.5	54.1	11.6	16.2	18.5	34.0	42.3	53.9
DEC C	14.6	18.5	24.3	7.0	9.3	10.8	12.1	15.6	19.2
DEC F	7.5	9.6	11.8	4.4	5.6	6.7	6.2	7.9	9.6
HP C	23.3	26.4	30.4	7.9	10.3	12.3	13.4	17.0	20.8
HP F	15.1	18.2	22.0	4.0	5.1	6.2	12.2	14.6	17.4
	k_z			$k_x^2 + k_y^2 L_2$			$k_x^2 + k_y^2 L_\infty$		
Machine	10x(5x5)	15x(5x5)	10x(7x7)	12x5	12x7	15x7	12x5	12x7	15x7
SUN C	36.3	54.3	61.2	25.7	32.8	41.2	33.4	41.9	51.7
SUN F	33.2	49.0	57.2	21.2	26.2	35.5	27.9	32.0	39.9
Convex F	31.5	49.6	71.1	15.1	19.7	24.3	19.0	24.2	30.0
DEC C	11.8	17.6	26.3	8.6	9.6	11.9	13.6	15.2	19.2
DEC F	6.2	9.1	8.5	5.6	5.6	6.8	7.0	7.7	9.8
HP C	12.8	18.8	18.5	15.9	21.4	27.4	14.6	18.1	22.3
HP F	9.5	14.7	20.1	13.5	20.0	25.3	11.8	15.8	18.3

Table 1.13 2-D convolution computation time (in seconds) for one frequency on different machines for different operator sizes and extrapolation methods.

rectives, therefore the convolution schemes were written in Fortran code which vectorizes well. Note that the Fortran compilers are better in optimization than the C compilers.

In the direct implementation of the 2-D convolution the even symmetry in the convolution operator is used. This implementation is designed to work fast on a Vector computer. In the implementation of the McClellan transformation and the k_z expansion the circular symmetry in the basis operators is used by first adding the common terms to each other and then multiplication with the appropriate operator point. This reduces the number of multiplications with a factor 8 in comparison with a full convolution. The computation times given in Table 1.13 are real-time computation times measured during the calculation. The time needed to calculate the operators is not included in this time.

Due to the use of the even symmetry in the operator the computation time for the direct convolution is a real challenge for the other methods (Note that the direct scheme can be made even faster when $\Delta x = \Delta y$ and the circular symmetry is used). The first and second order McClellan implementations and the series expansion in $\cos(k_r \Delta x)$ are the fastest algorithms on all machines. The series expansion in k_z and in $k_x^2 + k_y^2$ are comparable with the McClellan transformation. The difference between the $k_x^2 + k_y^2$ expansion in L_2 and L_∞ is that in the L_2 scheme the direct series expansion is used and for the L_∞ the Chebyshev recursion scheme is used. From the table it is clear that the hardware design of the computer system can be optimum for some specific implementation. For example the HP has a good performance on the direct scheme and less on the McClellan schemes. The DEC has a good performance on all expansion schemes.

Which scheme is preferred depends also on the desired accuracy of the result. If one uses the extrapolation only to get a first idea of the subsurface, or to estimate the macro model, a first order McClellan can be used. For a higher accuracy the direct convolution or a series expansion method with a high number of terms can be used.

Machine	type	RAM	bits	C	F	C-opt	F-opt
SUN	10/514	256	32	gcc	f77	-O2	-O2
Convex	C-220	256	32	cc	fc	-O2	-O2
DEC	3000-500	96	64	cc	f77	-O2	-O3
HP	9000-735	144	32	cc	fort77	+O4	+O4

Table 1.14 Specification for the used machines. Note that the optimization is done for one CPU and parallel processing is not used.

In Table 1.14 a detailed specification for the different machines is given. On the SUN the gcc compiler is used because it produces faster code than the standard cc compiler delivered by SUN. Note that for the optimization options only the most common used options are chosen, it may therefore be possible that by choosing another option the scheme will perform better as described in this Appendix (suggestions for better options are welcome).

In Table 1.15 the computation effort of the different methods is calculated. The values in this table are calculated for a spatial dimension of 111×111 samples with only one depth step for one frequency. The computation times are for the HP with source code written in C. Add(f) and Mul(f) are respectively floating point additions and multiplications, (i) gives integer additions

Method	Size	Add(f)	Mul(f)	Assign	Add(i)	Mult(i)	Time
Direct V	19×19	7.68	4.92	5.70	0.32	0.13	0.34
Direct V	25×25	12.89	8.33	9.32	0.34	0.14	0.55
Direct V	31×31	19.43	12.62	13.85	0.35	0.14	0.80
Direct 8	19×19	10.25	2.71	2.84	17.00	0.05	0.56
Direct 8	25×25	17.64	4.48	4.62	29.72	0.05	0.94
Direct 8	31×31	27.03	6.70	6.84	45.98	0.05	1.42
k_z Se	15×25	8.99	2.81	3.60	14.05	0.10	0.31
k_z Se	15×49	15.89	4.19	5.00	29.23	0.10	0.57
k_z Ch	15×25	9.66	3.47	4.29	14.05	0.10	0.39
McC1	10×9	2.44	1.36	2.11	2.91	0.10	0.19
McC1	13×9	3.25	1.80	2.78	3.80	0.10	0.25
McC1	16×9	4.07	2.24	3.45	4.68	0.10	0.30
McC2	10×17	4.21	1.80	2.57	5.57	0.10	0.24
McC2	13×17	5.62	2.39	3.40	7.34	0.10	0.31
McC2	16×17	7.02	2.98	4.22	9.12	0.10	0.39
McC2+	10×25	5.99	2.02	2.80	9.12	0.10	0.26
McC2+	13×25	7.98	2.69	3.69	12.07	0.10	0.34
McC2+	16×25	9.98	3.35	4.59	15.03	0.10	0.42
$k_x^2 + k_y^2$	$15 \times (3 + 3)$	2.44	2.12	1.82	1.63	0.10	0.18
$k_x^2 + k_y^2$	$15 \times (5 + 5)$	3.82	2.81	2.57	3.01	0.10	0.37
$k_x^2 + k_y^2$	$15 \times (7 + 7)$	5.20	3.50	2.93	4.39	0.10	0.62
$k_x^2 + k_y^2$	$15 \times (9 + 9)$	6.58	4.19	3.29	5.77	0.10	1.84
$k_x^2 + k_y^2$	$10 \times (5 + 5)$	2.46	1.82	1.70	2.02	0.10	0.15
$k_x^2 + k_y^2$	$12 \times (5 + 5)$	3.01	2.22	2.04	2.41	0.10	0.18
$k_x^2 + k_y^2$ Ch	$10 \times (5 + 5)$	2.88	2.24	1.93	2.02	0.10	0.20
$k_x^2 + k_y^2$ Ch	$12 \times (5 + 5)$	3.52	2.74	2.34	2.41	0.10	0.25
$k_x^2 + k_y^2$ Ch	$15 \times (5 + 5)$	4.48	3.47	2.94	3.01	0.10	0.31
$k_x^2 + k_y^2$ Ch	$10 \times (7 + 7)$	3.77	2.69	2.17	2.90	0.10	0.25
$k_x^2 + k_y^2$ Ch	$12 \times (7 + 7)$	4.61	3.28	2.63	3.50	0.10	0.29
$k_x^2 + k_y^2$ Ch	$15 \times (7 + 7)$	5.86	4.16	3.31	4.38	0.10	0.36
$\cos(k_r)$	10×9	2.24	1.16	1.69	2.91	0.10	0.14
$\cos(k_r)$	13×9	2.98	1.53	2.21	3.79	0.10	0.19
$\cos(k_r)$	16×9	3.72	1.90	2.73	4.68	0.10	0.20
$\cos(k_r)$	10×25	5.79	1.82	2.36	9.12	0.10	0.21
$\cos(k_r)$	13×25	7.71	2.42	3.11	12.07	0.10	0.27
$\cos(k_r)$	16×25	9.64	3.00	3.85	15.03	0.10	0.34

Table 1.15 Computation effort according to the number of additions, multiplications and assignments ($1e+6$). The spatial size is chosen fixed at 111×111 samples.

Method	Operator	55x55	111x111	155x155	223x223
Direct	19×19	3.0	11.5	23.1	48.0
McC1	10×9	1.9	9.5	19.6	41.6
McC2	10×17	2.3	10.1	21.2	44.9
k_z	$10 \times (7 \times 7)$	2.4	10.6	21.0	46.9
$k_x^2 + k_y^2 L_2$	$12 \times (7 + 7)$	1.9	8.5	18.2	41.0
$k_x^2 + k_y^2 L_\infty$	$12 \times (7 + 7)$	2.9	12.9	25.9	58.0

Table 1.16 Computation time on the DEC Alpha for different model sizes with different operators.

and multiplications. Integer additions are used to indicate an array element, so a large number of integer additions indicates a lot of switching between the elements in an array. Assign indicates the number of floating point assignments. Note that complex multiplications are counted as 2 floating additions and 4 floating multiplications (just as it is implemented in C).

The Direct V scheme is designed optimum for Vector computers, the Direct 8 scheme makes use of the circular symmetry of the operator. The differences between these two direct schemes are clear from the table; the Vector scheme takes more floating additions multiplications and assignments but only a few integer additions, the Direct 8 scheme uses a lot of integer additions a little more floating additions than the Vector scheme and half the number of floating multiplications. Nevertheless the Direct 8 scheme is slower than the Vector scheme due to the inefficient method to obtain the array elements. Taking into account the computation effort and the performance on the machines the series expansion in $\cos(k_r \Delta x)$ is the most flexible scheme which has a good computation performance on all machines.

In Table 1.16 the computation time on the DEC Alpha for the different methods is given as function of the size of the model. From these results one can conclude that the direct method on the DEC is not necessary slower than the other methods. Note that the Chebyshev recursion scheme implemented in the $k_x^2 + k_y^2$ expansion is slower than the power series expansion.

Appendix B: Chebyshev Polynomial Approximation

The Chebyshev polynomial is a powerful function in approximation theory Parks, (Parks and Burrus, 1987), because of the special properties of the Chebyshev polynomials. The Chebyshev approximation $P(x)$ of a real function $F(x)$ is defined as

$$F(x) \approx P(x) = \frac{a_0}{2} + \sum_{m=1}^M a_m T_m(x) \quad (1.73)$$

for $-1 \leq x \leq 1$ and with $T_m(x)$ the Chebyshev polynomial of order m and a_m the expansion coefficients (see for example Johnson and Riess 1977, (1977); Ralston 1967, (1967); Kogbetliantz 1960, (1960) Powell 1981, (1981)). The Chebyshev polynomials $T_m(x)$ are usually defined in terms of trigonometric functions by:

$$\begin{aligned} T_m(x) &= \cos(m \arccos(x)) & \text{for } |x| \leq 1 \\ T_m(x) &= \cosh(m \cosh^{-1}(x)) & \text{for } |x| > 1 \end{aligned} \quad (1.74)$$

Using the variable substitution $x = \cos(\phi)$ the Chebyshev polynomials can be rewritten as

$$T_m(x) = \cos(m\phi) \quad (1.75)$$

With this variable substitution some properties of the Chebyshev polynomials can easily be derived

$$2T_m(x)T_n(x) = T_{n+m}(x) + T_{n-m}(x) \quad (1.76)$$

$$T_{m+1}(x) = \cos((m+1)\phi) = 2 \cos(\phi) \cos(m\phi) - \cos((m-1)\phi) \quad (1.77)$$

$$T_{m+1}(x) = 2xT_m(x) - T_{m-1}(x) \quad (1.78)$$

$$T_2(x) = 2x^2 - 1 \quad (1.79)$$

$$T_3(x) = 4x^3 - 3x \text{ etc.} \quad (1.80)$$

with $T_0(x) = 1$ and $T_1(x) = x$. The most important property is the recurrence relation given in equation (1.78). Equations (1.79) and (1.80) show that $T_m(x)$ can be written as $\sum b_m x^m$, a polynomial in x . The polynomial $T_m(x)$ has m zero's in the interval $[-1, 1]$ which are located at the points

$$x_i = \cos\left(\frac{\pi(2i+1)}{2m}\right) \quad (1.81)$$

The Chebyshev polynomials are also orthogonal with respect to the weight function $(1-x^2)^{\frac{1}{2}} = \sin(\phi)$. So

$$\int_{-1}^1 (1-x^2)^{-\frac{1}{2}} T_i(x) T_j(x) dx =$$

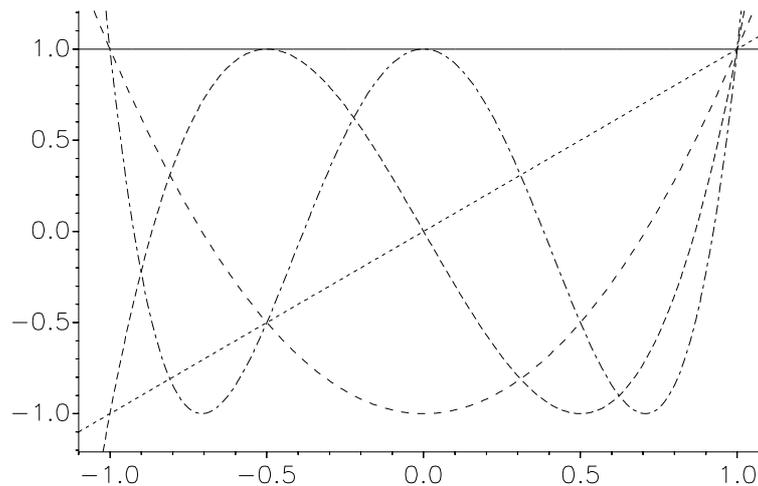


Fig. 1.40 Chebyshev polynomials for the 0,1,2, and 3th order within the interval $[-1, 1]$. Note that outside this interval the polynomials of orders ≥ 2 go to $\pm\infty$.

$$\int_0^\pi \cos(i\phi) \cos(j\phi) d\phi \begin{cases} 0 & \text{for } i \neq j \\ \frac{\pi}{2} & \text{for } i = j \neq 0 \\ \pi & \text{for } i = j = 0 \end{cases} \quad (1.82)$$

By using this orthogonal property it can be proven (Powell (1981), 1988, p. 144) that the Chebyshev approximation $P(x)$ in equation (1.73) is very nearly the *minimax polynomial*, which (among all polynomials of the same degree) has the smallest maximum deviation for the true function $F(x)$. For a fixed M equation (1.73) is a polynomial in x which approximates the function $F(x)$ in the interval $[-1, 1]$. This particular approximating polynomial is not necessarily more accurate than some other approximating polynomial of the same order M (for some specified definition of "accurate"), but equation (1.73) can be truncated to a polynomial of lower degree $n \leq M$ that *does* yield the "most accurate" approximation of degree n . So the accuracy of the approximation improves when the number of terms is increased. Since the $T_m(x)$ are all bounded between ± 1 , the difference between the truncated(n) and the larger polynomial(M) can be no larger than the sum of the neglected terms. In fact if the terms are rapidly decreasing, then the error is dominated by $c_n T_n(x)$ an oscillatory function with $(n+1)$ equal extreme distributed smoothly over the interval $[-1, 1]$. This smooth spreading out of the error is a very important property (Press et al., 1992).

The Chebyshev polynomials have a remarkable behavior. They oscillate between $+1$ and -1 for $-1 \leq x \leq 1$ and go monotonically to $\pm\infty$ outside that domain. All N of their zeros are real and fall in the domain $-1 \leq x \leq 1$. In Figure 1.40 the Chebyshev polynomials for the 0,1,2, and 3th order are plotted. Another important property of the Chebyshev polynomials is that they satisfy the "Haar condition". This condition is used in the Remez exchange algorithm to prove uniqueness of the best minimax approximation (for a more detailed discussion see Appendix C).

Appendix C: Optimization with the L_∞ norm

Introduction

For the design of a spatial 1-D convolution operator an optimal solution is desired. The optimized result must be designed accurate within the band of interest and stable outside this band. In order to solve this problem the notion accuracy must be defined in a mathematical way. The most common definitions are explained in the next subsection about the L_p norms. After defining the accuracy in a desired way a computation method must be available to calculate the solution of the approximation problem. By using the L_∞ norm the minimization problem is usually not solvable explicit in terms of formulas. However, by formulating the L_∞ optimization problem as a Chebyshev approximation problem a set of conditions (stated in the alternation theorem) is provided which completely characterize the optimal filter. To assure the convergence and uniqueness of the solution in the L_∞ norm some theorems are needed which are explained in the section about the Haar condition. In the section about the Remez exchange algorithm, which calculates the unique solution in the L_∞ norm, an iterative algorithm which calculates the L_∞ solution is explained in an intuitive manner.

The Parks-McClellan (1972) algorithm is based on the Remez exchange algorithm and Chebyshev approximation theory to design convolution operators with optimal fits between the desired and actual frequency responses. The operators are optimal in the sense that the maximum error between the desired frequency response and the actual frequency response is minimized for the given weighting function. Operators designed this way exhibit an equiripple behavior in their frequency response, and hence are sometimes called equiripple filters.

The L_p norms

A weighted error function, which measures the deviation from the true function, is defined as

$$\|\varepsilon(x)\| = \|W(x)|F(x) - P(x)|\| \quad (1.83)$$

There are several possible choices available for the norm function ($\|\cdot\|$) in equation (1.83). The most widely used are the L_∞ (also called the minimax, or Chebyshev) norm and the L_2 (least-squares) norm. The L_∞ norm is appropriate when specifications are stated in terms of minimum allowed stop-band attenuation or maximum allowed pass-band error. The L_2 error measure is appropriate when specifications are in terms of signal energy. Statistical considerations show that the L_2 norm is the most appropriate choice for data fitting, when the errors in the data have a normal distribution. The calculation of the best approximation in the L_2 norm reduces to a system of linear equations (the normal equations), which can be solved straightforward by using efficient algorithms. The L_1 norm is the least used norm and can be used for fitting a discrete data set with some gross errors in it due to blunders. The L_1 , L_2 , L_p and L_∞ norm are respectively defined as

$$\|F(x)\|_1 \equiv \int_b^a |F(x)| dx \quad (1.84)$$

$$\|F(x)\|_2 \equiv \left[\int_b^a |F(x)|^2 dx \right]^{\frac{1}{2}} \quad (1.85)$$

$$\|F(x)\|_p \equiv \left[\int_b^a |F(x)|^p dx \right]^{\frac{1}{p}} \quad (1.86)$$

$$\|F(x)\|_\infty \equiv \max_{a \leq x \leq b} |F(x)| \quad (1.87)$$

Peaks and overshoots in the frequency domain are typical of frequency sampling and least-squares designs. Windowing techniques are attempts to reduce the peaks in the error function (Parks and Burrus, 1987). But how far can the maximum error be reduced? The theory of Chebyshev approximation provides algorithms to find the coefficients of a filter with the minimum value for the maximum error. Filters that have the minimum value of the maximum error exhibit an equiripple behavior in their frequency response. A practical reason for using the L_∞ norm is that when in computer calculations a complicated mathematical function is estimated by one that is easy to calculate then it is usually necessary to ensure that the greatest value of the error function ε is less than a fixed amount. This is just the required accuracy of the approximation which is a condition on the norm $\|F - P\|_\infty$.

If there is an approximation such that the L_∞ norm error function is small then the L_2 norm and L_1 norm error functions are small also. This can be proven with the Gauchy-Schwarz inequality (Powell, 1981) that with error function $\varepsilon = \{|F(x) - P(x)|; a \leq x \leq b\}$

$$\|\varepsilon\|_1 \leq (b - a)^{\frac{1}{2}} \|\varepsilon\|_2 \leq (b - a) \|\varepsilon\|_\infty \quad (1.88)$$

The different norm functions can be presented in a graphical way shown in Figure 1.41 (after Powell 1981, (1981)). In this Figure the point (2, 1) is approximated by one point from the line $y = x$ represented by A, for the different norms. The L_1 norm solution to this problem is not unique and is given by all points lying on both A and the rotated box indicated by L_1 (the solution is the minimum of $\{|2 - x| + |1 - x| \forall x\}$). These points range from (1, 1) to (2, 2) and the minimum distance is 1. Non-minimum L_1 solutions are represented by the same box but on a larger scale. The L_2 norm is represented by a circle and has one unique solution, (1.5, 1.5) with distance 0.5, and is the position where the circle touches the line A (the solution is the minimum of $\{(|2 - x|^2 + |1 - x|^2)^{\frac{1}{2}} \forall x\}$). Other non-minimal L_2 solutions can be represented by circular contours with the same midpoint but with a different radius. The L_∞ norm is the small cube with the dashed line, which for this simple example gives the same solution as the L_2 norm (the solution is the minimum of $\max(|2 - x|, |1 - x|) \forall x$).

The Haar condition

The best minimax approximation from a set of basis functions in A, polynomials of the order n (P_n), to the function F in

$$\|\varepsilon(x)\| = \max_{0 \leq x \leq x_{max}} W(x) |F(x) - P(x)| \quad (1.89)$$

is that element of A that minimizes equation (1.89). It can be proven (Powell 1981, p.74) that in order to discover if a trial approximation is optimal one only need to consider the extreme values of the error function. Let P^* be a non-optimal trial approximation. The purpose is to improve

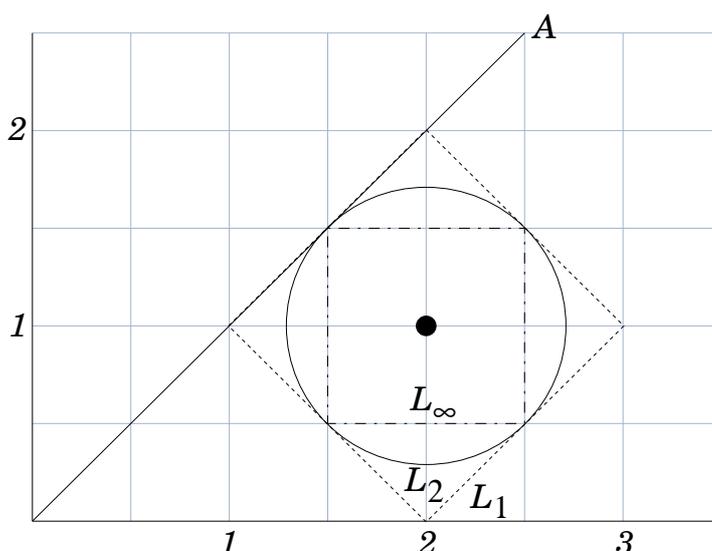


Fig. 1.41 The L_1 , L_2 and L_∞ norms displayed for a simple approximation problem (after Powell, 1981).

the approximation by satisfying equation (1.89). The set of points where the error function

$$\varepsilon^*(x) = F(x) - P^*(x) \quad (1.90)$$

takes its extreme values is called E . This set of extreme points is characterized by the condition

$$|\varepsilon^*(x)| = \|\varepsilon^*\|_\infty \quad (1.91)$$

Let $(P^* + \theta P)$ be a best approximation. Then equation (1.89) is obtained and the inequality

$$|\varepsilon^*(x) - \theta P(x)| < |\varepsilon^*| \quad (1.92)$$

is satisfied for all points in E . Assuming that θ is positive in equation (1.92) it is easy to see that for the points E the sign of $\varepsilon^*(x)$ is the same as the sign of $P(x)$. From this it follows (Powell, 1981) that if P^* is a best minimax approximation from A to F then there is no function P in A that satisfies the condition

$$[F(x) - P^*(x)]P(x) > 0 \quad (1.93)$$

for all values of x where the error function takes its extreme values. If inequality (1.93) is true for some P then there exists a positive value θ that gives a better approximation. Whether condition (1.93) is obtained can be tested easily if the set of approximating functions A satisfies the Haar condition. When A satisfies the 'Haar condition' it can also be proven that the best approximation is unique. The Haar conditions provides also an excellent method for calculating the best approximations, called the (Remez) exchange algorithm.

The Haar condition, stated in its most usual definition, is given by

Condition 1 (Haar) *If an element of P_n has more than n zeros, then it is identically zero.*

Thus A satisfies the Haar condition if and only if, for every non-zero P in A , the number of roots of the equation $\{P(x) = 0; a \leq x \leq b\}$ is less than the dimension of A . It can also be proven that

Theorem 1 (Uniqueness for L_∞) *Let A be a linear subspace $[a, b]$ that satisfies the Haar condition. Then for any f in $[a, b]$ there is just one best minimax approximation from A to F .*

If A does not satisfy the Haar condition then there are functions f in $[a, b]$ that have several best approximations in A .

Another property of the class of minimax approximation problems may be used to obtain a characterization of the solution to equation (1.89). The alternation and characterization theorems give the necessary and sufficient conditions for the best minimax approximation (Rabiner et al., 1975):

Theorem 2 (Alternation) *if $P(x)$ is a linear combination of M cosine functions i.e.,*

$$P(x) = \sum_{m=0}^{M-1} \alpha(m) \cos(2\pi mx) \quad (1.94)$$

then a necessary and sufficient condition that $P(x)$ be the unique best weighted Chebyshev approximation to a continuous function $F(x)$ on K (wavenumber domain) is that the weighted error function $\varepsilon(x) = W(x)|F(x) - P(x)|$ exhibits at least $M+1$ extremal frequencies in K .

Theorem 3 (Characterization) *$P(x)$ is the best minimax approximation if and only if there exist a set of extremal frequencies of $M+1$ points $\{\xi_i\}$, $i = 1, 2, \dots, M + 1$ such that $\xi_1 < \xi_2 \dots < \xi_M < \xi_{M+1}$, with $\varepsilon(\xi_i) = -\varepsilon(\xi_{i+1})$, $i = 1, 2, \dots, M$ and $|\varepsilon(\xi_i)| = \max_{x \in \xi} |\varepsilon(x)|$.*

For example a function in A (algebraic polynomials of order n) has at most n sign changes. If the error function $|F(x) - P_n^*(x)|$ changes sign more than n times as x ranges over the linear subspace E then according to the statement in equation (1.93) P^* is a best approximation. Conversely, if the number of sign changes does not exceed n , then it is possible to choose the zeros of a polynomial in A so that condition (1.93) is satisfied. Note that the Chebyshev polynomials satisfy the Haar condition on the interval $[-1, 1]$.

The alternation theorem means that the best Chebyshev approximation must necessarily have an equiripple error function. It also states that there is a unique best approximation for a given set of frequencies, filter length N and weight function $W(x)$. The phrase "at least $M+1$ extremal frequencies" needs some explanation. Since the best approximation for a given set of specifications is unique there will **not** be one filter with $M+1$ extremals and another filter with $M+2$ extremals for the same specifications. For a given set of specifications, the unique best filter may have more than $M+1$ extremal frequencies. If for example the optimal filter has $M+3$ extremal frequencies, then by the uniqueness property, there cannot be a filter with only $M+1$ extremals for this given set of specifications.

Theorem 3 characterizes the optimum solution so that it can be recognized, but it does not directly show how to choose the filter coefficients. If the extremal frequencies are known, the impulse-response coefficients can be found easily by solving an interpolation problem by a frequency-sampling technique. The problem of designing the filter has been reduced to the problem of finding the extremal frequencies. The Remez exchange algorithm is an algorithm which solves the Chebyshev approximation by searching for these extremal frequencies of the best approximation.

Remez exchange algorithm

The Remez Exchange algorithm is to the Chebyshev approximation as the normal equations are to the minimum least-squares solution. In both cases a set of expansion coefficients a_i is calculated that "best fit" a set of basis functions to the data. If the definition of "best fit" is to minimize the sum of squares then get the normal equations are used to calculate the a_i 's. If the "best fit" is defined to minimize the maximum error (min-max or Chebyshev fit) then the Remez Exchange algorithm gives the answer.

The Remez exchange algorithm for Chebyshev approximation is designed to make the error function of the filter satisfy the set of necessary and sufficient conditions for the optimal solution as stated in the alternation theorem. If the Haar condition holds (which is true for Chebyshev polynomials) then convergence of the exchange algorithm is obtained from *any* initial reference.

The exchange algorithm calculates the element of a $n + 1$ dimensional subspace A that minimize the approximation $P(x)$ in

$$\|\varepsilon(x)\|_{\infty} = \max_{0 \leq x \leq k_N} W(x)|F(x) - P(x)| \quad (1.95)$$

Instead of trying to reduce the error of each trial approximation, the algorithm adjust a reference set of extremal frequencies $\{\xi_i : i = 0, 1, \dots, n + 1\}$, so that it converges to a point set $\{\xi_i^* : i = 0, 1, \dots, n + 1\}$ that satisfies the characterization theorem. The adjustments are made by an iterative procedure.

In order to begin the calculation, an initial reference is chosen. It can be any set of points that satisfies the condition

$$a \leq \xi_0 < \xi_1 < \dots < \xi_{n+1} \leq b \quad (1.96)$$

At the start of each next iteration a reference is available that is different from the references of all previous iterations. The calculations of each iteration are as follows. Let $\{\xi_i : i = 0, 1, \dots, n + 1\}$ be the reference at the start of an iteration. First the function $P(x)$ in A that minimizes the expression

$$\max_{i=0,1,2,\dots,n+1} W(\xi_i)|F(\xi_i) - P(\xi_i)| \quad (1.97)$$

is calculated. The coefficients c_i in $P(x) = \sum_i c_i Q_i(x)$ may be found (Powell, page 79 and 85) by solving the linear system of equations

$$W(\xi_i)[F(\xi_i) - P(\xi_i)] = (-1)^i h \quad i = 0, 1, \dots, n + 1 \quad (1.98)$$

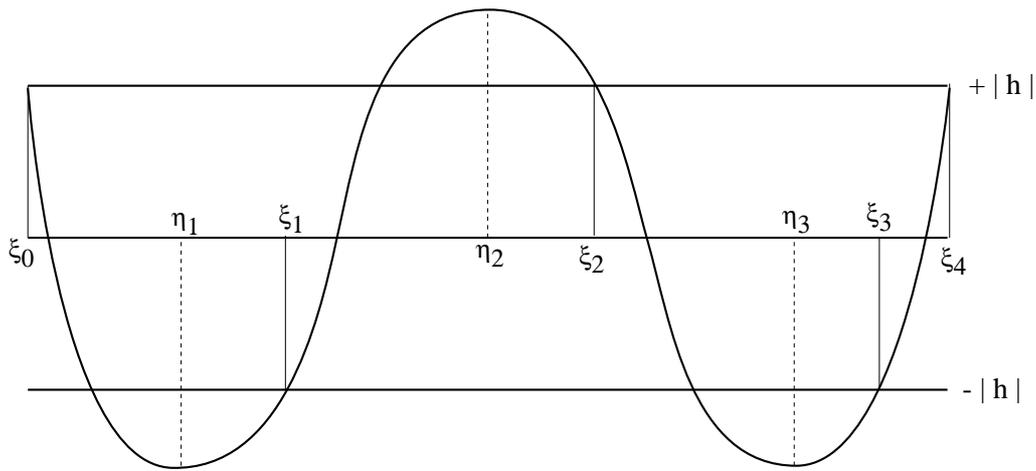


Fig. 1.42 An error function of the exchange algorithm (after Powell).

where, h is defined by the linear system and can be calculated analytically. For a given set of $n + 2$ frequencies this requires the solution of the $n + 2$ equation as given in equation (1.98). A fast method for solving this problem is given in McClellan and Parks (1972) and Rabiner et al. (1975). It follows from Powell (1981) that the bounds

$$|h| \leq \|F - P^*\|_\infty \leq \|F - P\|_\infty \quad (1.99)$$

are satisfied, where P^* is the required best approximation from A to F . In order to make use of the right-hand bound, and in order to obtain a suitable change for the reference, the error function

$$\varepsilon(x) = W(x)(F(x) - P(x)) \quad a \leq x \leq b \quad (1.100)$$

is considered. A typical error function for $n=3$ is shown in Figure 1.42. We see that equation (1.98) is satisfied, and that consequently ε has at least n turning points. The positions of the extrema, which are called η_i in the figure, are estimated by evaluating the error function at several points within $[a, b]$. It is necessary in practice to obtain these points automatically in an efficient way. Let η be a point that satisfies the equation

$$|F(\eta) - P(\eta)| = \|F - P\|_\infty \quad (1.101)$$

The calculation finishes if the difference

$$\delta = |F(\eta) - P(\eta)| - h \quad (1.102)$$

is sufficiently small, because inequality (1.99) implies the bound

$$\|F - P\|_\infty \leq \|F - P^*\|_\infty + \delta \quad (1.103)$$

Otherwise the reference is changed in order to begin another iteration. The most important property of the changes of reference is that the quantity $|h|$, which is called the levelled reference error, increases strictly monotonically from iteration to iteration. It is helpful to take the point of view that the purpose of the change of reference is to increase the value of $h(\xi_0, \xi_1, \dots, \xi_{n+1}) = |h|$. Because expression (1.102) is small only if the levelled reference error is close to the bound $\|F - P\|_\infty$ of inequality (1.99), it is advantageous to make $h(\xi_0, \xi_1, \dots, \xi_{n+1})$ as large as possible. Thus the exchange algorithm is a method of solving a maximization problem, where the variables are the points of reference. The structure of $h(\xi_0, \xi_1, \dots, \xi_{n+1})$, however, is that it is inefficient to use one of the super linearly convergent algorithms that are available in subroutine libraries.

Adjustment of the reference

The method of choosing the new reference must imply the increase

$$h(\xi_0^+, \xi_1^+, \dots, \xi_{n+1}^+) > h(\xi_0, \xi_1, \dots, \xi_{n+1}) \quad (1.104)$$

in the levelled reference error where ξ^+ is the next reference set. It can be shown that it is sufficient if the conditions

$$\text{sign}[F(\xi_{i+1}^+) - P(\xi_{i+1}^+)] = -\text{sign}[F(\xi_i^+) - P(\xi_i^+)] \quad (1.105)$$

for $i = 0, 1, \dots, n$ and

$$|F(\xi_i^+) - P(\xi_i^+)| \geq |h| \quad (1.106)$$

are satisfied. Provided that at least one of the numbers $\{|F(\xi_i^+) - P(\xi_i^+)| : i = 0, 1, \dots, n+1\}$ is greater than $|h|$. Hence, several ways of obtaining an increase in the levelled reference error are suggested by Figure 1.42. One method is to let each point of the new reference be an extremum of the error function (1.100). In this case the error curve of Figure 1.42 gives the reference $\{\xi_0, \eta_1, \eta_2, \eta_3, \xi_4\}$ and the conditions 1.105 and 1.106 are obtained. Methods that can change every reference point on every iteration are usually more efficient than a one-point exchange algorithm, in the sense that fewer iterations are required to reduce the number (1.102) to less than a prescribed tolerance. In the excellent book by Powell (1981) the one-point exchange method is discussed in further detail.

A very nice property of the exchange algorithm is that if the Haar condition holds, then convergence is obtained from any initial reference. However, some initial references are better than others. Powell gives a suitable set of reference points on the interval $[a, b]$, obtained from the properties of the Chebyshev polynomials, which is in some way the best starting point

$$\xi_i = \frac{1}{2}(a + b) + \frac{1}{2}(b - a) \cos\left(\frac{(n + 1 - i)\pi}{n + 1}\right) \quad (1.107)$$

for $i = 0, 1, \dots, n + 1$.

Optimizing 2-D extrapolation operators with L_∞ norm.

The phase shift extrapolation operator is given by

$$\tilde{F}_0(k_x) = \exp(-j\sqrt{k^2 - k_x^2}\Delta z) \quad \text{for } |k_x| < k \quad (1.108)$$

Splitting the operator in real (\Re) and imaginary (\Im) parts for $|k_x| \leq k$ gives

$$\Re\{\tilde{F}_0(k_x)\} = \cos(\sqrt{k^2 - k_x^2}\Delta z) \quad (1.109)$$

$$\Im\{\tilde{F}_0(k_x)\} = -\sin(\sqrt{k^2 - k_x^2}\Delta z) \quad (1.110)$$

The inverse discrete Fourier transformation for an even function $F(m\Delta x) = F(-m\Delta x)$ is given by

$$\tilde{F}(k_x) = F(0) + 2 \sum_{m=1}^M F(m\Delta x) \cos(k_x m\Delta x) \quad (1.111)$$

Splitting real and imaginary parts in equation (1.111) gives two independent equations

$$\Re\{\tilde{F}(k_x)\} = \Re\{F(0)\} + 2 \sum_{m=1}^M \Re\{F(m\Delta x)\} \cos(k_x m\Delta x) \quad (1.112)$$

$$\Im\{\tilde{F}(k_x)\} = \Im\{F(0)\} + 2 \sum_{m=1}^M \Im\{F(m\Delta x)\} \cos(k_x m\Delta x) \quad (1.113)$$

substituting equations (1.109) and (1.110) into (1.112) and (1.113) gives

$$\cos(\sqrt{k^2 - k_x^2}\Delta z) \approx \Re\{F(0)\} + 2 \sum_{m=1}^M \Re\{F(m\Delta x)\} \cos(k_x m\Delta x) \quad (1.114)$$

$$-\sin(\sqrt{k^2 - k_x^2}\Delta z) \approx \Im\{F(0)\} + 2 \sum_{m=1}^M \Im\{F(m\Delta x)\} \cos(k_x m\Delta x) \quad (1.115)$$

For designing a spatial convolution operator an optimal solution for the Real and Imag $F(m\Delta x)$ is desired. The wavenumber-domain approximation problem given in equations (1.114) and (1.115) can be written in the more general form

$$\tilde{F}(k_x) = F(0) + 2 \sum_{m=1}^M F(m\Delta x) \cos(m\Delta x k_x) \quad (1.116)$$

The Chebyshev polynomial approximation problem given in equation (1.73)

$$P(k_x) = \frac{a_0}{2} + \sum_{m=1}^M a_m \cos(m \arccos(k_x)) \quad (1.117)$$

can be shown to be equivalent with equation (1.116) by using the change of variables

$$\zeta = \cos(k_x) \quad (1.118)$$

or

$$k_x = \arccos(\zeta) \quad (1.119)$$

$$\cos(m\Delta x k_x) = \cos(m\Delta x \arccos(\zeta)) \quad (1.120)$$

which results in

$$\tilde{F}(\arccos(\zeta)) = \tilde{F}(\zeta) = F(0) + 2 \sum_{m=1}^M F(m\Delta x) \cos(m\Delta x \arccos(\zeta)) \quad (1.121)$$

With the aid of these changes equations (1.114) and (1.115) can be solved by using the Remez exchange algorithm to find the extremal frequencies with an error in the L_∞ norm. Having found the extremal frequencies k_e for $e = 1, 2, \dots, M$ of the error function the F_m , the impulse response coefficients, have to be calculated from it. The coefficients F_m in the approximation (IDFT)

$$F(m\Delta x) = \frac{1}{N} \sum_{n=0}^{N-1} \tilde{P}(n\Delta k_x) \exp(jn\Delta k_x m\Delta x) \quad (1.122)$$

can be found by solving an interpolation problem of fitting the function $P(k_x)$ in equation (1.122) to the N known equally spaced frequency values in equation (1.122). This procedure amounts to solving the following set of linear equations:

$$\tilde{P}(k_e) = \sum_{m=0}^M F_m \cos(jk_e m\Delta x) = \tilde{F}(k_e) \pm \delta \quad (1.123)$$

for $e = 1, 2, \dots, M$.

Polynomial synthesis reduced to symmetrical spectral synthesis

McClellan and Chan (1977) have shown that the Chebyshev structure can be used for every arbitrary polynomial $P_n(x) = T_n(P_1(x))$, where $T_n(P_1(x))$ is defined by the Chebyshev recursion formula. The approximating of a function by polynomials is given by

$$F(x) = F(P(x)) = \sum_{m=0}^M a_m P^m(x) \quad (1.124)$$

with $P^m(x)$ the m 'th order polynomial in x given over the domain $P(x) \in [a, b]$. For example $P(x)$ can be k_z or $k_x^2 + k_y^2$ for the approximation of the 2-D extrapolation operator given in equations (1.67) and (1.69). The polynomial synthesis can be reduced to symmetrical spectral synthesis by performing the following mapping (Soubaras (1992), Rabiner and Gold (1975) page 151) :

$$P(x) = g(\phi) = 0.5(a - b) \cos(\phi) + 0.5(a + b) \quad (1.125)$$

which maps the domain $P(x) \in [a, b]$ to $\phi \in [0, \pi]$.

The advantage of reducing the polynomial synthesis to a spectral synthesis is that the same technique as in the 1-Dimensional operator problem can be used; the Remez exchange algorithm with the L_∞ norm. The inverse function of $P(x) = g(\phi)$ is given by

$$\phi = g^{-1}(P(x)) = \arccos\left(\frac{(2P(x) - a - b)}{(a - b)}\right) \quad (1.126)$$

Substitution of equation (1.125) into equation (1.124) gives the desired spectral synthesis

$$F(g(\phi)) = G(\phi) = \sum_{m=0}^M \hat{a}_m \cos(m\phi) \quad (1.127)$$

Having found the solution for the \hat{a}_m in the L_∞ norm for a number of terms M we have to compute the original approximation $F(x)$ from $G(\phi)$ by using $F(x) = G(g^{-1}(P(x)))$.

$$F(P(x)) = \sum_{m=0}^M \hat{a}_m \cos\left(m \arccos\left(\frac{(2P(x) - a - b)}{(a - b)}\right)\right) \quad (1.128)$$

This equation can be rewritten by using Chebyshev recursion formula and writing $\zeta = \arccos\left(\frac{(2P(x) - a - b)}{(a - b)}\right)$

$$F(P(x)) = \sum_{m=0}^M \hat{a}_m \cos(m\zeta) \quad (1.129)$$

$$= \sum_{m=0}^M \check{a}_m [\cos(\zeta)]^m \quad (1.130)$$

$$= \sum_{m=0}^M \check{a}_m \left[\frac{2P(x) - a - b}{a - b}\right]^m \quad (1.131)$$

where $F(x)$ is a polynomial in $P(x)$. Because the L_∞ norm is invariant by any mapping we have that $G(\phi)$ is an L_∞ approximation of $F_0(g(\phi))$ and $F(P(x))$ is a L_∞ approximation of $F_0(P(x))$.

The method of polynomial synthesis is more preferable than a Taylor series truncation, because the Remez algorithm minimizes the L_∞ norm on a user defined domain $[a, b]$, while Taylor series truncation minimizes the error on a $[-\varepsilon, \varepsilon]$ interval, where ε is infinitely small. The difference with the L_2 optimization is that the WLSQ approximation takes the total spectrum of the optimized operator into account and not just its extreme values as in the L_∞ norm.

1.8 References

- Abramowitz, M., and Stegun, I. A., 1968, Handbook of mathematical functions: Dover publications Inc., New York.
- Algazi, V. R., Suk, M., and Rim, C. S., 1986, Design of almost minimax fir filters in one and two dimensions by wls techniques: IEEE Trans. Circuits and Systems, **CAS-33**, 590–596.
- Berkhout, A. J., 1982, Imaging of acoustic energy by wave field extrapolation (2nd edition): Elsevier Amsterdam.
- Berkhout, A. J., 1984, Seismic resolution; resolving power of acoustical echo techniques (volume 12): Geophysical Press.
- Biondi, B., and Palacharla, G., 1994, 3d depth migration using rotated mccllellan filters: 56th Mtg. Vienna, EAEG, Expanded Abstracts of Papers, B45.
- Blacquièrè, G., Debeye, H. W. J., Wapenaar, C. P. A., and Berkhout, A. J., 1989, 3d table-driven migration: Geophysical Prospecting, **37**, 925–958.
- Blacquièrè, G., 1989, 3d wave field extrapolation in seismic depth migration: Ph.D. thesis, Delft University of Technology.
- Blacquièrè, G., 1991, Optimized mccllellan transformation filters applied in one-pass 3-d depth migration.: 61th Ann. Internat. Mtg., SEG, Expanded Abstracts, 1126–1129.
- Gaiser, J. E., 1994, Optimum transformation and filter structure for explicit finite-difference 3d migration: 56th Mtg. Vienna, EAEG, Expanded Abstracts of Papers, B42.
- Gazdag, J., 1978, Wave equation migration with the phase-shift method: Geophysics, **43**, 1342–1351.
- Hale, D., 1991a, 3-d depth migration via mccllellan transformations: Geophysics, **56**, 1778–1785.
- 1991b, Stable explicit extrapolation of seismic wavefields: Geophysics, **56**, 1770–1777.
- Hazra, S. N., and Reddy, M. S., 1986, Design of circularly symmetric low-pass two-dimensional fir digital filters using transformation: IEEE Trans. Circuits and Systems, **CAS-33**, 1022–1026.
- Hoff, J., 1995, 3d one-way wave field extrapolation by using 1d convolution operators: Ph.D. thesis, Delft University of Technology.
- Holberg, O., 1988, Towards optimum one-way wave propagation: Geophysical Prospecting, **36**, 99–114.
- Johnson, L. W., and Riess, R. D., 1977, Numerical analysis: Addison-Wesley.

- Kao, J., G. Li, and C. Yang, 1994, Optimized 2d filters for 3d wavefield extrapolation on the massively parallel processors: 56th Mtg. Vienna, EAEG, Expanded Abstracts of Papers, B48.
- Kato, H., and Matsumoto, G., 1982, Design of circularly symmetric 2-d fir digital filters using fourier reconstruction technique: IEEE Trans. Acoustics, Speech and Signal Processing, **ASSP-30**, 505–508.
- Kogbetliantz, E. G., 1960, Generation of elementary functions:, *in* Mathematical Methods for Digital Computers New York: John Wiley & Sons, Selected Reprint, 1, 7–35.
- McClellan, J. H., and Chan, D. S. K., 1977, A 2-d fir filter structure derived from the chebychev recursion: IEEE Trans. Circuits and Systems, **CAS-24**, 372–378.
- McClellan, J. H., and Parks, T. W., 1972, Equiripple approximation of fan filters: Geophysics, **37**, 573–583.
- McClellan, J. H., 1973, The design of two-dimensional digital filters by transformations: 7th Ann.Mtg, Princeton, Princeton Conf. on Inform. Sci. and Sys., 247–251.
- Mersereau, R. M., Mecklenbräuker, W. F. G., and Quatieri, T. F., 1976, McClellan transformations for two-dimensional digital filtering: I-design: IEEE Trans. Circuits and Systems, **CAS-23**, 405–413.
- Nautiyal, A., Gray, S. H., Whitmore, N. D., and Garing, J. D., 1993, Stability versus accuracy for an explicit wavefield extrapolation operator: Geophysics, **58**, 277–283.
- Parks, T. W., and Burrus, C. S., 1987, Digital filter design: New York: John Wiley & Sons.
- Powell, M. J. D., 1981, Approximations theory and methods: Cambridge University Press.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P., 1992, Numerical recipes in c: Cambridge University Press.
- Rabiner, L. R., and Gold, B., 1975, Theory and application of digital signal processing: Prentice Hall.
- Rabiner, L. R., McClellan, J. H., and Parks, T. W., 1975, Fir digital filter design techniques using weighted chebyshev approximations:, *in* Digital Signal Processing New York: IEEE press, Selected Reprint, 2, 81–96.
- Ralston, A., 1967, Rational chebyshev approximation:, *in* Mathematical Methods for Digital Computers New York: John Wiley & Sons, Selected Reprint, 2, 264–285.
- Rice, J. R., and Usow, K. H., 1968, The lawson algorithm and extensions: Math. Comp., **22**, 118–127.
- Sollid, A., and Arntsen, B., 1994, Cost-effective 3d one-pass depth migration: Geophysical Prospecting, **42**, 755–776.

Soubaras, R., 1992, Explicit 3-d migration using equiripple polynomial expansion and laplacian synthesis: 62th Ann. Internat. Mtg., SEG, Expanded Abstracts, 905–908.

Thorbecke, J. W., and Rietveld, W., 1994, Optimum extrapolation operators: a comparison: 56th Ann. Internat. Mtg., EAEG, Expanded Abstracts, P105.

Contents

1	3-D recursive extrapolation operators	1
1.1	Introduction	1
1.2	Wave field extrapolation in the space-frequency domain	1
1.3	Direct method	3
1.3.1	Weighted Least Squares optimization	4
1.3.2	Hankel Transformations	10
1.3.3	Rotated Fourier reconstruction	12
1.3.4	Error analysis	14
1.4	McClellan transformation	22
1.4.1	Hazra and Reddy Coefficients	27
1.4.2	Optimized McClellan factors	30
1.4.3	Rotated Coefficients	31
1.4.4	Series expansion in $\cos(k_r \Delta x)$	31
1.4.5	Error analysis	34
1.5	Single series expansion	38
1.5.1	Expansion in k_z with L_2 -norm	39
1.5.2	Expansion in k_z with L_∞ -norm	42
1.5.3	Error analysis	43
1.6	Double series expansion	44
1.6.1	Expansion in $k_x^2 + k_y^2$ with L_2 -norm	44
1.6.2	Expansion in $k_x^2 + k_y^2$ with L_∞ -norm	46
1.6.3	Error analysis	48
1.7	Discussion and Conclusions	50

1.8 References 69