Fundamentals of the discrete wavelet transform for seismic data processing

Jack K. Cohen and Tong Chen

ABSTRACT

The discrete wavelet transform has been an exciting topic of mathematical research for about 10 years now. Some of the early wavelet research had seismic applications explicitly in mind, but fields as diverse as quantum physics and voice coding have also provided insights leading to the development of the modern theory. Our purpose is to explain and illustrate the effect of the discrete wavelet transform on seismic data, thus providing the information necessary for researchers to assess its possible use in their areas of data processing.

An analysis of the discrete wavelet transform of dipping segments with a signal of given frequency band leads to a quantitative explanation of the known division of the two-dimensional wavelet transform into horizontal, vertical and diagonal emphasis panels. The results must be understood in a “fuzzy” sense: since wavelet mirror filters overlap, the results stated can be slightly violated with violation tending to increase with shortness of the wavelet chosen.

The specific angles that delimit the three wavelet panels can be stated in terms of the Nyquist frequency $F_l$ in the first dimension and the Nyquist frequency $F_s$ in the second dimension. The angle $\theta_{HV} = \arctan(F_l/F_s)$ approximately separates the dips $\theta < \theta_{HV}$ that appear in the horizontal panel from those greater dips that appear in the vertical panel. Similarly, the angles $\arctan(F_l/2F_s), \arctan(2F_l/F_s)$ approximately bound the dips appearing in the diagonal panel. These results are simple and probably have been observed by other researchers, but we haven’t found a prior reference for them. As an example of seismic processing using the wavelet transform, some simple examples of de-aliasing spatially aliased data are discussed.

INTRODUCTION

Imagine that the Fourier transform was only recently receiving attention from the signal processing community. Some talks are mentioning tentative applications of this new tool. And someone you know has just gotten access to a computer code for the Fourier transform that can be run on seismic data. You transform a seismic section and begin to look at the result. At first, it looks like a mess—there is no resemblance to the original data. But soon conclusions about spatial and temporal frequency content of reflections are made and eventually this tool is in common use in the seismic community.

This paper provides a look at what a new tool—the discrete wavelet transform—does
to data. One doesn’t have to know the complete theory behind the prime factor algorithm for the Fourier transform to use that tool and understand the results. On the other hand, a basic knowledge of Nyquist limits, aliasing, etc., is necessary for intelligent use of the Fourier transform. Correspondingly, we’ll tell you some basic facts about wavelets and wavelet transforms, but will not provide a complete mathematical development. We will make some suggestions about applications, but the main thrust of the paper is to provide the information necessary to allow others to decide if wavelet methods offer advantages over existing tools in targeted areas of seismic processing.

Since the term “wavelet” is already used in seismology, there is potential for confusion with its differing use in the theory of the wavelet transform. In connection with the wavelet transform, this terminology is only loosely suggestive of seismic wavelets such as source signatures and propagating waveforms. The rationale for the use of the term in the theory of the wavelet transform is simple:

1. the wavelet function is “oscillatory” in the sense that it takes on positive and negative values due to a fundamental property dictating that the integral of the wavelet function is zero, and

2. the wavelet function is localized in the sense that it’s magnitude falls off with distance from its “main lobe”.

Hence we arrive at the nomenclature “wavelet”—small wave.

Thus, the mathematical wavelets and the seismic wavelets do share some properties. However, in the mathematical theory, the term wavelet is used in a more technical sense (e.g., a seismic modeling wavelet need not integrate to exactly zero, even though it does oscillate). In particular, it turns out that some of the mathematical wavelets don’t look much like source signatures and the seismic wavelets commonly used for modeling sources do not meet the technical criteria set down for the mathematical wavelets. Below, “wavelet” will always refer to the mathematical wavelets, not the seismic wavelets.

A seismic trace gives the time history of a geophone response. Its Fourier transform reveals the frequency content of this response. Often, what is really desired is the time evolution of the frequency response. For instance, a dominant feature of most seismic traces is the loss of high frequencies at later time. The wavelet transform of a trace provides a simultaneous display of its time-frequency content—therein lies a principal reason for its study. The short-time or windowed Fourier transform made popular by the work of Dennis Gabor (Gabor, 1946) also localizes simultaneously in time and frequency, using a uniform time window for all frequencies. In contrast, the wavelet transform adapts the window, according high frequencies the fine resolution they require and low frequencies the long windows needed to encompass them.

This paper is a user’s outline introduction to the one- and two-dimensional discrete wavelet transforms, including examples of wavelets, wavelet filters and their properties. To set the stage for seismic applications, we give a detailed analysis of the effect of the wavelet transform on sloping reflections with a signal of given frequency band. This analysis reveals a semi-quantitative explanation of the well-known division of the two-dimensional wavelet transform into horizontal, vertical and diagonal emphasis panels. The result is simple and probably has been noted by other researchers, but we haven’t found a prior reference for
it. In any case, it provides an important insight for use of the two-dimensional wavelet transform in seismic processing.

Extensive bibliographies of the wavelet transform literature appear in Daubechies (1992) and the more recent paper by Jawerth and Sweldens (1993). Thus in our own bibliography, we list only the papers naturally cited, even though this means not listing some major papers that have played an important role in our education.

**THE ONE-DIMENSIONAL DISCRETE WAVELET TRANSFORM**

To get a feel for what happens to data under the discrete one-dimensional wavelet transform, consider the effect on a single data vector \( x \) of length \( N = 2^n \) and sampling interval \( \Delta t \). At the first stage, the wavelet transform splits this data into two sub-vectors \( a^1, d^1 \) each of length \( N/2 \) and sample interval \( 2\Delta t \). (Just how the sub-vectors are defined is immaterial for the moment.) The vector \( d^1 \) is retained as the first portion of the wavelet transform of \( x \), while the splitting process is repeated on \( a^1 \) to produce \( a^2 \) and \( d^2 \). This process can be repeated until the \( n^{th} \) stage, when \( d^n \) and \( a^n \) each consist of a single point and the data vector \( x \) is replaced by the vectors \( d^1, d^2, \ldots, d^n, \) and \( a^n \) (see Figure 1). At each stage, the sum of lengths of the two sub-vectors is the same as the length of the parent vector, thus the sum of the lengths of the \( d \) vectors plus the final single point \( a^n \) is the same as the length \( N \) of the original data. Since the \( x \)'s play the role of the \( a \)'s at the first stage, it is sometimes convenient to set \( x = a^0 \) as shown in Figure 1. In practice, the wavelet decomposition is often carried out to fewer than \( n \) levels so that the final \( d \) and \( a \) contain more than a single datum.

The details of this *fast wavelet transform* (FWT) algorithm, including the precise methodology for creating the sub-vectors, are given in Daubechies (1992) and code for the algorithm may be found in Press, et al. (1992). Here, only the general properties of the FWT are discussed. As in Figure 1, denote the transformation from the parent \( a \) to the child \( a \) by \( A \) and the transformation from the parent \( a \) to the child \( d \) by \( D \). The most important fact about the transformations \( A \) and \( D \) is that they are orthogonal. Thus, the decomposition of the data is a change of basis to an orthogonal basis—the *wavelet* basis. The \( A \) phase of the decomposition is essentially a “blurring” or *averaging* of the data at the previous

---

**Figure 1.** 1D wavelet decomposition.
stage. In frequency domain, this is equivalent to a low pass filter on the previous stage. The $D$ phase of the decomposition is essentially the difference of the data at the previous stage and the averaged data at the current stage. In frequency domain, this is equivalent to a high pass filter on the previous stage. As we will see, the $A$ and $D$ operations are not “merely” averaging and differencing of adjacent data points, but these notions direct the intuition appropriately.

Since the $A$ and $D$ transformations are orthogonal, the inverse or reconstruction algorithm is well conditioned. It consists of applying transposed matrices to the current $a$ and $d$ vectors to climb the ladder in Figure 1 reconstructing the (previously discarded) parent $a$. The reconstruction algorithm is again carefully explained in Daubechies (1992) with code given in Press, et al. (1992).

The discussion to this point is analogous to having given an outline of the algorithm for the Fast Fourier Transform (FFT) without mentioning the exponential functions, $e^{i\omega t}$, that lie behind it! There are, indeed, analogs to the exponentials underlying the wavelet transform. In fact, there are many wavelet transforms depending on the “wavelet” chosen and some of the more popular varieties are introduced in the next sub-section and illustrated in appendix A.

**Wavelets and scaling functions**

The function $\phi(t)$ underlying the $A$ transformation is called the scaling function, and the function $\psi(t)$ underlying the $D$ transformation is called the “mother” or “analyzing” wavelet. Scaling functions and analyzing wavelets occur in linked pairs. The mathematical theory constructs the wavelet $\psi(t)$ from a given scaling function $\phi(t)$. The properties required of legitimate scaling functions and the construction of the analyzing wavelet associated with a particular scaling function are described in Chapter 5 of Daubechies (1992). Note, in particular, that the required property

$$\int \phi(t) \, dt = 1$$

imposed on a legitimate scaling function is consistent with the averaging interpretation and that the required property

$$\int \psi(t) \, dt = 0$$

imposed on a legitimate analyzing wavelet is consistent with the differencing interpretation. Now turn to a brief description of some of the fundamental properties of analyzing wavelets and scaling functions.

The scaling functions $\phi(t)$ and analyzing wavelets $\psi(t)$ are localized in both time and frequency. While it is impossible for a given function $s(t)$ and its Fourier transform $\tilde{s}(f)$ to simultaneously have finite (or “compact”) support, it is not impossible for both of them to have negligible amplitude outside finite intervals in their respective domains (i.e., time $t$ and frequency $f$). Scaling functions and analyzing wavelets used in practice have faster decay than $1/t$ in time and simultaneously have faster decay than $1/f$ in frequency.

An important auxiliary notion is that of multiresolution analysis: the shifts

$$\phi(t - k), \quad k = \ldots, -2, -1, 0, 1, 2, \ldots$$
of the scaling function are assumed to form an orthonormal basis for the data space (let one unit in \( k \) correspond to \( \Delta t \)). It then can be shown that the shifts

\[
\phi_{jk}(t) = 2^{-j/2} \phi(2^{-j} t - k) = 2^{-j/2} \phi(2^{-j}(t - k2^j)), \quad j, k = \ldots, -2, -1, 0, 1, 2, \ldots
\]

with translation step of size \( 2^j \) form a basis for the \( j \)th averaged (or blurred) space. Although the shifted \( \phi \)‘s are orthonormal within each level (i.e., for fixed \( j \)), they are not orthonormal between levels. To get a basis that is orthonormal between levels, use the differencing function—the wavelet. Indeed, the scaled shifts of the analyzing wavelet \( \psi(t) \),

\[
\psi_{jk}(t) = 2^{-j/2} \psi(2^{-j} t - k), \quad j, k = \ldots, -2, -1, 0, 1, 2, \ldots
\]

form an orthonormal basis of \( L^2 \). In a formal mathematical sense, these orthogonality properties are written as

\[
\int \phi_{jk}(t) \phi_{j'k'}(t) \, dt = \delta_{k,k'},
\]

\[
\int \psi_{jk}(t) \psi_{j'k'}(t) \, dt = \delta_{j,j'} \delta_{k,k'}.
\]

Notice that, in general, negative scales are allowed. However, in the seismic context, assume that the original data can be represented at level 0, so that the highest resolution available is at level 0 and \( j \geq 0 \). Notice also that once the localization properties of the analyzing wavelet \( \psi(t) \) are known, the localization properties of \( \psi_{jk}(t) \) are determined by simply translating by \( k2^j \) and scaling by \( 1/2^j \). Exactly the same is true for the scaling function \( \phi(t) \).

Thus, we arrive at a most important insight: the discrete wavelet transform provides a systematic mechanism for analysing data at different scales (octave by octave). Indeed, since the translation step \( 2^j \) is linked to the level (or octave), the high frequency events (\( j \) near 0) are treated with relatively small translation steps, so they get the fine resolution they require; while low frequency events (\( j \gg 0 \)) get the larger time windows needed to encompass them. See Figure 2 for a schematic of the equal area tiling of the time-frequency plane by a wavelet basis. The concepts of multiresolution analysis are fully explicated in Mallat (1989a), Mallat (1989b) and Daubechies (1992).

Since the components of the wavelet transform are indexed by \( j \)—which gives the scale or frequency level and \( k \)—which gives the location, it follows that the wavelet transform simultaneously displays data in time-frequency domain. It also means that a one-dimensional array is, in a sense, transformed to two dimensions. This type of transformation should be compared and contrasted with the Gabor or short time Fourier transform. There again, time is transformed to time-frequency domain by windowing the time domain and Fourier transforming each window. But the Gabor transform does not adapt the frequency band to the temporal band as does the wavelet transform, see Daubechies (1992) and the classic paper by Gabor (1946).

Recall from equation (2) that the integral of the analyzing wavelet vanishes. The analyzing wavelets used in practical applications will actually have \( \text{several} \) (say \( M \)) vanishing moments: \( \int t^m \psi(t) \, dt = 0 \), \( m = 0, 1, 2, \ldots, M - 1 \). A larger number of vanishing moments is associated with increased smoothness of the wavelet and its associated scaling function (the number of derivatives is proportional to \( M \), but the proportionality constant may be
less than one). In addition, use of a wavelet with several vanishing moments is crucial to wavelet operator compression applications (Beylkin, 1992; Beylkin et al., 1991; Beylkin et al., 1992).

The vanishing moment property of the analyzing wavelet can be used to better understand the $D$ differencing phase of the FWT discussed earlier. To make this connection, recall that although the FFT is used algorithmically to compute the coefficients $s_n$ of a function $s(t)$ in the Fourier basis, these coefficients are more fundamentally defined by the integration,

$$s_n = \frac{1}{T} \int_0^T s(t)e^{2\pi i n t/T} \, dt.$$ 

Analogously the coefficients $c_{jk}$ of a function in the wavelet basis are algorithmically computed by the FWT outlined above, but are fundamentally defined by integrating the product of the function and the wavelet, as

$$c_{jk} = \int s(t) \psi_{jk}(t) \, dt. \tag{3}$$

With this fact in hand, the intuitive notion that $D$ is a differencing operation can be replaced by a precise statement: the portions of the data at level $j$ that can be well approximated by polynomials of degree less than $M$ (the number of vanishing moments) are removed in the $D$ operation.

Finally, a technical note. The assumption made above that the data can be represented at level 0 is truly an assumption about the function $s(t)$. In seismic applications, only the samples $s_k = s(k\Delta t)$ are available. Strictly speaking, these samples should be transformed according to the wavelet basis used before the wavelet decomposition is performed. This can be done efficiently, but this step is often ignored in practice. See Daubechies (1992, 166) for the transformation algorithm.

A taxonomy of known scaling functions and analyzing wavelets includes those with finite support in time and rapid decay in frequency (the Daubchies wavelets), those with finite support in frequency and rapid fall-off in time (the Meyer wavelets) and those with infinite...
support and rapid fall-off in both domains (e.g. the Battle-Lemarié wavelets). These three families and others are discussed in Daubechies (1992).

The simplest example of a scaling function is $\phi(t) = 1$, $0 < t < 1$ and zero elsewhere. This function and its associated analyzing wavelet (also with support on $0 < t < 1$) are shown in Figure 3. The corresponding wavelet basis is known as the Haar basis. This wavelet basis has been known since the turn of the century and can be considered the first member of the Daubechies family as well as the first member of the Battle-Lemarié family. However, the sharp edges imply poor (1/f) fall-off in frequency domain. On the other hand, the Haar scaling function integrated against a function literally gives its average on the given scale, while the Haar analyzing wavelet is likewise a literal differencing operator. Thus, while the Haar example is not used for practical applications, it does provide a useful guide to the intuition that is not afforded by less accessible wavelets. Further discussion and plots of the Daubechies, Meyer and Battle-Lemarié scaling functions and analyzing wavelets are given in appendix A.

**Filters associated with scaling functions and analyzing wavelets**

As has already been mentioned, the A operation is a low-pass filtering operation. This can be given concrete expression in terms of a filter $m_0$ associated with each scaling function $\phi(t)$ according to the fundamental relation,

$$\hat{\phi}(\xi) = m_0(\xi/2)\hat{\phi}(\xi/2).$$

As before, the notation, $\hat{\phi}$, is used for the Fourier transform of $\phi$. Figure 4 shows the $m_0$ filter for a Meyer wavelet with fourth order smoothness (see appendix A for an explanation of the order $n$ of a Meyer wavelet). Similarly, the D operation is a high-pass filter $m_1$ satisfying

$$\hat{\psi}(\xi) = m_1(\xi/2)\hat{\phi}(\xi/2).$$

Figure 5 shows the amplitude of the $m_1$ filter for the same Meyer wavelet. Notice that the scaling function is present on the right side of both equation (4) and equation (5). This is consistent with the previous averaged segment giving rise to both the next wavelet segment $d^j$ and the next averaged segment $a^j$ in Figure 1.

A remarkable feature of the low- and high-pass filters is their symmetry of their amplitudes in the quarter-period point as shown in Figure 6. For this reason, these filters

![Haar Scaling Function](image1)
![Haar Wavelet](image2)

**Fig. 3.** Haar scaling function and wavelet.
Fig. 4. Meyer $m_0$ low-pass filter with 4th order smoothness.

Fig. 5. Meyer $m_1$ high-pass filter with 4th order smoothness.

are called “mirror filters” in the literature. These mirror filters also satisfy the “power conservation” law,

$$|m_0(\xi)|^2 + |m_1(\xi)|^2 = 1,$$

an equation that plays a crucial role in the mathematical development (cf. Daubechies (1992), page 132 ff) of wavelet theory. (Personal note: it is so fundamental that my (Cohen) copy of the Daubechies book falls open to this page.) Examples of Daubechies, Meyer and Battle-Lemarié $m_0, m_1$ filters are shown in appendix B.

As in the example shown (Figure 6), the wavelet low/high-pass filters always overlap—they are not strictly low and high pass! The overlap means that bandpass interpretations of wavelet filtering must always allow for an intrinsic amount of fuzziness near the boundary. The overlap can be lessened by the choice of wavelets, but it can’t be totally eliminated.

Filtering with wavelets

Observe that after the high-pass operation $D$, the resulting data have the same doubled sampling interval as the averaged or blurred segment. How can its high frequencies be present with a halved Nyquist frequency? The answer is that the frequencies are “folded”

Fig. 6. Meyer $m_0$ and $m_1$ as mirror filters.
over, just as if the data were aliased—see Figure 7. However, since only (well, almost only) the high frequencies are present in the D half-size segment, there is no confounding of information—that is, there is no true aliasing (well, almost none).

Before After

0 \[ F_1 \] 0 \[ F_{1/2} \]

0 \[ F_1 \] 0 \[ F_{1/2} \]

**Fig. 7.** Representation of frequencies in D.

The parenthetical comments in the previous sentence are again an acknowledgement of the inherent fuzziness of wavelet filtering operations. That is, the frequency overlap in the wavelet transform implies a degree of aliasing in the decomposition step. If the reconstruction algorithm is applied, then this aliasing is properly taken into account and the reconstruction is perfect. However, most applications will involve another operation interposed between decomposition and reconstruction (e.g. compression or migration). In such a case, the aliasing is not perfectly removed during reconstruction. With proper choice of wavelet, the aliasing distortion will be negligible for most applications.

Applications most suited to wavelet transform analysis are those that effectively use the simultaneous localization in time and frequency. The coefficients \( c_{jk} \) defined in equation (3) give a direct method for bringing this tool to bear on seismic data. In the theoretical form,

\[
s(t) = \sum_{jk} c_{jk} \psi_{jk}(t),
\]

of the wavelet transform decomposition, the effect of such an application could be summarized:

1. **Localize in time and frequency by choosing a subset of the \( j \)'s and \( k \)'s.**
2. **Alter the \( c_{jk} \)'s in the chosen subset.**

To make the discussion more concrete, consider a trace recorded with sampling interval \( \Delta t \), so that the Nyquist frequency is \( F_t = \Delta t \). Then the \( d^j \) sub-vector in the decomposition (cf. Figure 1) has approximate frequency content \([F_t/2^j, F_t/2^{j-1}] \) and its \( k^{th} \) component is temporally centered at \( t = k 2^j \Delta t \) (taking the time origin at the first point of the original trace). A process based on the discrete wavelet transform would be applied only to a subset of the octaves \([F_t/2^j, F_t/2^{j-1}] \), that is, one would select only the corresponding \( d^j \)'s. Then, within each selected octave, the process would operate on only certain components of the \( d^j \) corresponding to the targeted time window for that octave.

There are applications, notably data compression and filter design, where an octave is too gross a unit. One methodology for obtaining finer control in frequency domain (at the
expense of decreased temporal resolution) is to repeat the decomposition algorithm on one or more of the \( d \) vectors. This procedure is called forming \textit{wavelet packets}, see Wickerhauser (1992), Daubechies (1992). Another methodology for overcoming the octave limitation is the use of the \textit{continuous} wavelet transform, see Daubechies (1992), Rioul and Vetterli (1991) for an overview and see Niüßuma, et al. (1993) for an interesting application to quantification of shear-wave splitting. Yet a third approach to obtaining finer frequency resolution is to use “dilation factors” larger than 2, so that the fundamental frequency unit is smaller than an octave, see Auscher (1992), Daubechies (1992).

Finally, observe that everything said in this section applies without change to the case of spatial variations \( x \) and, indeed, there are sure to be interesting applications of the one-dimensional wavelet transform applied to such variables as the offset or midpoint of a seismic section. And, perhaps, the most interesting applications will involve more than one dimension, so turn now to a discussion of the two-dimensional wavelet transform.

\textbf{THE TWO-DIMENSIONAL DISCRETE WAVELET TRANSFORM}

The simplest way to obtain a wavelet basis for a two-dimensional space, say, \( t \) and \( x \), is to multiply the one-dimensional bases:

\[ \psi_{j\ell; k\ell'}(t,x) = \psi_{j\ell}(t) \psi_{k\ell'}(x). \]  

(6)

This is analogous to replacing the one-dimensional Fourier kernel \( e^{i\omega t} \) with \( e^{i\omega t} e^{-ikx} \) and does, indeed, provide a two-dimensional complete orthonormal basis in the wavelet case just as it does for Fourier transform. The disadvantage of this basis is the mixing of the scales \( j \) and \( j' \). It turns out that it is possible to construct a basis using only a \textit{single} scale \( j \) at the expense of having \textit{three} wavelets (basis functions) at each \( j \) level. These three functions are

\[ \begin{align*}
\psi_{j\ell; k\ell'}^H(t,x) &= \psi_{j\ell}(t) \phi_{j\ell'}(x), \\
\psi_{j\ell; k\ell'}^V(t,x) &= \phi_{j\ell}(t) \psi_{j\ell'}(x), \\
\psi_{j\ell; k\ell'}^D(t,x) &= \psi_{j\ell}(t) \psi_{j\ell'}(x).
\end{align*} \]  

(7)

This is the basis used in the remainder of this paper and also in many of the existing applications of the two-dimensional wavelet transform (Beylkin, 1992; Beylkin et al., 1991; Kleinert, 1992; Mallat, 1980a; Mallat 1989b).

The averaging function at level \( j \) is simply the product of the one-dimensional scaling functions:

\[ \phi_{j\ell; k\ell'}(t,x) = \phi_{j\ell}(t) \phi_{j\ell'}(x). \]

It is convenient to think of the result of the wavelet transform with this basis as being displayed in four panels as in Figure 8. The \( A \) symbol for the averaged (in each direction) panel is the analog of the averaged portion occurring at each stage of the one dimensional transform, while the three other panels correspond to the three wavelet components at level \( j \) as indicated by the respective products shown at level 1 of the figure. Just as in one dimension, each successive level has half the number of points in both time and space as the previous level and correspondingly has double the sampling intervals as the previous level.
Fig. 8. 2D data decomposition.
As indicated in Figure 8, the superscript symbols on the three wavelet portions of the basis stand for: **H**-horizontal emphasis panel, **V**-vertical emphasis panel and **D**-diagonal emphasis panel. The **V** panel is created by “averaging” in \( t \) and “differencing” in \( x \). Hence a horizontal reflector would tend to be eliminated from the \( \mathbf{V} \) panels. From such extreme examples, one can understand the observation that \( \mathbf{V} \) panel does emphasize steeper dip events, and analogously for the other panels—see the figures in Daubechies (1992) and Mallat (1989a, 1989b) for additional confirmation.

![Diagram](image)

**Fig. 9.** An example of first-level 2-D wavelet decomposition showing that the **H**, **V** and **D** panels respectively emphasize horizontal, vertical and diagonal dips.

![Wavelet Decomposition](image)

**Fig. 10.** A seismic “common offset” section and two-dimensional wavelet decomposition.

A more precise analysis of the contents of the four panels in terms of dip angle is contained in the next subsection. As simple examples, consider Figure 9 which shows the first-level
decomposition of a simple model composed of a square with a diagonal and Figure 10 which shows a seismic “common offset” section and its first level decomposition.

Just as in the one dimensional wavelet transform, “folding” occurs in the frequencies contained in each of the H, V and D panels. Specifically, the temporal frequencies are folded in H; the spatial frequencies are folded in V and both are folded in D. Again, this folding does not confound data (except for the filter overlaps), so should not be considered as true aliasing.

Finally, note that the relative position of the A, H, V and D panels in Figure 8 is arbitrary—each of these panels contains points from the entire t-x domain at its scale j. Thus, they are positioned variously in the literature, cf. Daubechies (1992) and Mallat (1989a, 1989b) with Beylkin, Coifman and Rokhlin (1991).

Analysis of sloping reflections

To develop an intuition for the nature of the wavelet transform, consider the decomposition of the reflection data from a constant slope segment. Letting \( \theta \) denote the time dip (or apparent dip or slope), the situation is shown in Figure 11. The plane is characterized by the equation \( t = t_0 + x \tan \theta \) in space-time and by the equation

\[
f_x = f_t \tan \theta
\]

in Fourier domain. This result is established by taking the two-dimensional Fourier transform of \( \delta(t - t_0 - x \tan \theta) \) using \( 2\pi f_t \) as the conjugate variable to \( t \) and \(-2\pi f_x \) as the conjugate variable to \( x \) (Chun and Jacewitz (1981)). A plane of time dip \( \theta \) in space-time and arbitrary intersect time \( t_0 \) maps to a line segment whose extension passes through the origin in Fourier domain at the complementary angle. All location information is lost in the latter domain and the extent of the segment depends on the frequency band of the time dip signal.

![Fig. 11. Planes of time dip \( \theta \) in space-time and image in Fourier domain.](image)

Indeed, seismic data are characterized by their frequency band in addition to their location in time-space. To include this vital aspect in the model, assume that the sampling interval in time is \( \Delta t \), with associated Nyquist limit \( F_t = 1/2\Delta t \), and similarly introduce \( \Delta x \) and \( F_x = 1/2\Delta x \) in space. Suppose that the seismic signature along the sloping reflection has temporal band \([f^-, f^+]\). Equation (8) shows that the corresponding spatial band is \([f^- \tan \theta, f^+ \tan \theta]\). Assume that \( f^+ \leq F_t \) to avoid temporal aliasing and also that \( f^+ \leq F_x \cot \theta \) to avoid spatial aliasing. To avoid repeated mention of a trivial case, assume that \( f^- < f^+ \) so that the signal band is not empty. Also to give explicit acknowledgement of
the inherent fuzziness in wavelet filtering, introduce the symbol $\lesssim$ to denote an inequality that in practice may be slightly violated.

**The spectrum in $A^j$.**—Observe from Figure 12 that the frequencies $f_t$ within the band $[f^-, f^+]$ mapped to $A^j$ must satisfy the inequalities $0 < f_t \lesssim F_1/2^j, 0 < f_t \lesssim F_x \cot \theta/2^j$. Here, equation (8) has been used to derive the second inequality. Thus, the $A^j$ band is $[f_{A^j}^-, f_{A^j}^+]$, where

$$f_{A^j}^- = f^-$$
$$f_{A^j}^+ = \min \left(f^+, \frac{F_1}{2^j}, \frac{F_x \cot \theta}{2^j} \right)$$

$$\begin{align*}
\begin{bmatrix}
0, & F_x \\
2^j & 2^j-
\end{bmatrix} & \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \
The $H^j$ band is empty if $f_{H^j}^+ \geq f_{H^j}^-$. From equation (11), derive three nontrivial conditions involving the particular signal band:

\[
\frac{F_1}{2^j} > f^+ \implies H^j \cong \emptyset
\]

\[
f^- > \frac{F_1}{2^{j-1}} \implies H^j \cong \emptyset
\]

\[
f^- > \frac{F_x \cot \theta}{2^j} \implies H^j \cong \emptyset
\]  

(12)

However, for the $H^j$ band, there is a fourth condition arising from equation (11) that is independent of the signal:

\[
F_t > F_x \cot \theta \implies H^j \cong \emptyset.
\]

With the definition,

\[
\theta_{HV} \overset{\text{def}}{=} \arctan \left( \frac{F_x}{F_t} \right) = \arctan \left( \frac{\Delta t}{\Delta x} \right),
\]

this condition becomes

\[
\theta > \theta_{HV} \implies H^j \cong \emptyset
\]  

(13)

Not only is this condition independent of the signal band limits, it is also independent of the level $j$. Equation (14) gives precise analytic meaning to the notion that the $H^j$ panels emphasize the horizontal aspects of the signal:

\[
\text{No } H^j \text{ can contain information from dips significantly larger than } \theta_{HV}.
\]

The spectrum in $V^j$.—The $V^j$ band is denoted by $[f_{V^j}^-, f_{V^j}^+]$, where

\[
f_{V^j}^- = \max \left( f^-, \frac{F_x \cot \theta}{2^j} \right)
\]

\[
f_{V^j}^+ = \min \left( f^+, \frac{F_1}{2^j}, \frac{F_x \cot \theta}{2^{j-1}} \right)
\]  

(15)

The $V^j$ band is empty if $f_{V^j}^- \geq f_{V^j}^+$. From equation (15), derive the three nontrivial conditions involving the particular signal band:

\[
f^- > \frac{F_1}{2^j} \implies V^j \cong \emptyset
\]

\[
f^- > \frac{F_x \cot \theta}{2^{j-1}} \implies V^j \cong \emptyset
\]

\[
\frac{F_x \cot \theta}{2^j} > f^+ \implies V^j \cong \emptyset
\]  

(16)

and the fourth condition independent of the signal and level:

\[
F_x \cot \theta > F_t \implies V^j \cong \emptyset.
\]

Thus,

\[
\theta < \theta_{HV} \implies V^j \cong \emptyset.
\]  

(17)
Equation (17) gives precise analytic meaning to the notion that the $V^j$ panels emphasize the vertical aspects of the signal:

\[
\text{No } V^j \text{ can contain information from dips significantly smaller than } \theta_{HV}.
\]

The angle $\theta_{HV}$ that approximately separates dips that appear in $V$ from those that appear in $H$ has the direct geometric meaning shown in Figure 13.

**Fig. 13. Geometrical meaning of $\theta_{HV}$ in both physical and frequency domains.**

**The spectrum in $D^j$.**—The $D^j$ band is denoted by $[f_{D^j}^-, f_{D^j}^+]$, where

\[
\begin{align*}
    f_{D^j}^- &= \max \left( f^-; \frac{F_t}{2^j}, \frac{F_x \cot \theta}{2^j} \right) \\
    f_{D^j}^+ &= \min \left( f^+; \frac{F_t}{2^{j-1}}, \frac{F_x \cot \theta}{2^{j-1}} \right)
\end{align*}
\]

The $D^j$ band is empty if $f_{V^j}^- > f_{V^j}^+$. From equation (18), derive four nontrivial conditions involving the particular signal band:

\[
\begin{align*}
    f^- > \frac{F_t}{2^j} & \Rightarrow D^j \equiv \emptyset \\
    \frac{F_t}{2^j} > f^+ & \Rightarrow D^j \equiv \emptyset \\
    f^- > \frac{F_x \cot \theta}{2^j} & \Rightarrow D^j \equiv \emptyset \\
    \frac{F_x \cot \theta}{2^j} > f^+ & \Rightarrow D^j \equiv \emptyset
\end{align*}
\]

For $D^j$ there are two conditions independent of the signal and level:

\[
\frac{F_t}{2} > F_x \cot \theta \Rightarrow D^j \equiv \emptyset
\]

and

\[
\frac{F_x \cot \theta}{2} > F_t \Rightarrow D^j \equiv \emptyset
\]

With the definitions

\[
\theta_{D_{\text{min}}} \overset{\text{def}}{=} \arctan \left( \frac{F_x}{2F_t} \right) = \arctan \left( \frac{\Delta t}{2\Delta x} \right),
\]

(20)
\[ \theta_{D_{\text{max}}} \overset{\text{def}}{=} \arctan \left( \frac{2F_x}{F_t} \right) = \arctan \left( \frac{2\Delta t}{\Delta x} \right), \]

(21)

describes these conditions can be expressed as

\[ \theta < \theta_{D_{\text{min}}} \Rightarrow D^j \cong \emptyset. \tag{22} \]

and

\[ \theta > \theta_{D_{\text{max}}} \Rightarrow D^j \cong \emptyset. \tag{23} \]

Equations 22 and 23 gives precise analytic meaning to the notion that the \( D^j \) panels emphasize the diagonal aspects of the signal:

\( D^j \) can not contain information from dips significantly outside the angle bounds \( \theta_{D_{\text{min}}}, \theta_{D_{\text{max}}} \).

The angles \( \theta_{D_{\text{min}}} \) and \( \theta_{D_{\text{max}}} \) that approximately bound the dips that appear in \( D \) have the direct geometric meanings shown in Figure 14.

**Fig. 14.** Geometrical meaning of \( \theta_{D_{\text{min}}} \) and \( \theta_{D_{\text{max}}} \) in both physical and frequency domains.

**Displays in Fourier domain.**—Since the negative frequency ranges are suppressed, Figures 11, 13 and 14 are only schematic diagrams of the frequency plane. In the sequel it will be necessary to truly display the amplitude of two-dimensional Fourier transforms. Denote the Fourier transform of a seismic section \( s(t, x) \) by \( \hat{s}(f_t, f_x) \). Since \( s \) is real,

\[ \hat{s}(-f_t, -f_x) = \hat{s}^*(f_t, f_x), \]

where \( \hat{s}^* \) denotes the complex conjugate of \( \hat{s} \). Hence

\[ |\hat{s}(-f_t, -f_x)| = |\hat{s}(f_t, f_x)|. \]

Thus, it would be redundant to display all four quadrants of the amplitude of \( \hat{s} \). Usually the negative \( f_t \) half-plane is suppressed resulting in Figure 15. The labels \( A^1, H^1, D^1, \) and \( V^1 \) in Figure 15 indicate which portions of the frequency domain are mapped to each of these level 1 panels. If the Nyquist frequencies \( F_t \) and \( F_x \) shown in the figure were relabeled as respectively \( F_t/2^j \) and \( F_x/2^j \), one would correspondingly obtain a figure showing which portions of the frequency domain are mapped to each of the level \( j \) panels. That is, the two panels labeled with \( A^1 \) in Figure 15 are subdivided to yield the portions of the frequency domain that map to the level 2 panels and so on.
In a later section, spatial aliasing will be discussed. Since the finite transform is periodic and since the above display demands that $-F_x \leq f_x \leq F_x$, frequencies from a positively dipping plane that are spatially aliased ($f_x > F_x$) will appear in the negative frequency regime; see Figure 16.

**Examples: Decomposition of some sloping reflections**

**Level one as a function of dip angle.**—Throughout this section, set $F_t = F_x = 125$ Hz and set the signal band at 0–30 Hz. These assumptions imply that the boundary angle for the $H$ and $V$ panels is

$$\theta_{HV} = \tan^{-1}(1) = 45^\circ;$$

and that the bounding angles for the $D$ panels are

$$\theta_{D_{\text{min}}} = \tan^{-1}(1/2) \approx 20^\circ, \quad \theta_{D_{\text{max}}} = \tan^{-1}(2) \approx 61^\circ;$$

and finally that the angle where spatial aliasing begins is

$$\theta_{S_A} = \tan^{-1}(125/30) \approx 76.5^\circ.$$

Start by considering the first level decomposition of planes at various $t$-$x$ dip angles. Figure 17 shows the predicted contents of the level one panels as a function of dip angle according to the equations developed in the last section. The darker curve is the maximum value in the band and the lighter curve is the minimum. When the lighter curve lies
above the darker curve, the band in that panel is empty (≈ 0). Thus, for example, D^1 and H^1 are empty for all angles—this is a reflection of the fact that the maximum frequency \( f^+ = 30 \text{ Hz} \), while half-Nyquist is 62.5 Hz. Moreover since these values are well separated, no visible leakage into these panels is expected for a reasonable wavelet. As the lower left panel of Figure 17 indicates, the inequalities developed earlier imply that the V^1 panel remains (approximately) empty until an angle of about 64°.

Now check these theoretical predictions against the decomposition of synthetic data. Data with 30° dip is shown in Figure 18 and the first level decomposition using the Daubechies \( n = 4 \) wavelet is shown in Figure 19. As expected, all panels are empty except A^1. Figure 20 tells the same story for the level 1 decomposition of a dip of 60°. Finally, Figure 21 shows the level 1 decomposition of a dip of 75°. Since this angle exceeds the limiting angle of 64° (cf. the prediction of Figure 17), the V^1 panel is no longer empty.

**Higher level decomposition for a fixed dip angle.**—Now fix the dip angle at 30° and study the higher level decompositions. Immediately, the question of display arises. At high levels, only a few points will remain and a raw display would appear “blocky”. For this reason, one seldom carries the wavelet transform all the way to the \( n^{th} \) level. Typically, even by the third level, the data begins to have a “blocky” appearance. Moreover, notice that the level-one four-partite display in Figure 19 has panels 1/4 the size of the original data in Figure 18. Level two, presented in scaled size, would be 1/16 the size of the original data and therefore it would be hard to see details. And all the more so for level three’s 1/64th-sized panels. The solution to the display problem chosen here is to run the algorithm backwards (i.e., use the reconstruction algorithm) for each panel at each level until we arrive at a panel with the same number of points as the original. To do this, use zero data in all the panels needed in the reconstruction except for the one being investigated. A uniform lower and upper clip has been applied to all the figures discussed in this subsection to correctly
Fig. 18. 30° dip data with 0–30 Hz band.

Fig. 19. Level 1 decomposition of 30° dip data with 0–30 Hz band.
Fig. 20. Level 1 decomposition of 60° dip data with 0–30 Hz band.

Fig. 21. Level 1 decomposition of 75° dip data with 0–30 Hz band.
display the relative amplitudes on the panels at different levels.

For the 30° data in Figure 18, the only panel of interest at level one is $A^1$ (cf. Figure 19 and also the prediction in Figure 17). A reconstructed, full sized version of $A^1$ is shown in Figure 33. This panel is virtually identical to the original data and corresponds to the fact that the original data were oversampled by a factor of 2 in both time and space at this dip.

The bandlimits predicted by the theory of the previous section for levels two and three are given graphically in Figures 22 and 23. Since the given dip of 30° is considerably less than $\theta_{HV} = 45°$ (relative to the wavelet overlap for the Daubechies $n = 4$ wavelet used here), all the $V$-panels should be empty. This is verified by Figures 24, 25 and 26. On the other hand, the $H$ panels will contain information for sufficiently high level. If the angle bounds derived earlier were strict, then, for the case at hand, $H^2$ would be empty. However, as Figure 22 indicates, the lower bound is only slightly higher than the upper bound (the precise values are $f^+_{H^2} = 30$ Hz and $f^-_{H^2} = 31.25$ Hz), and so it is not too surprising that Figure 28 shows some “leakage” into $H^2$ caused by the imprecision of wavelet filtering. See the next subsection for further discussion. Figure 23 and Table 1 indicate that $H^3$ should contain a substantial amount of information relative to $H^2$, and Figure 29 bears out this prediction. Figures 22 and 23 and Table 2 indicate that $D^1$, $D^2$ should be empty, while $D^3$ contains significant information. Figures 30, 31 and 32 bear out this prediction. Finally, Figures 33, 34, and 35 show the averaged panels at these three levels.

Choosing the wavelet to control “leakage”.—Return to Figure 28. The wavelet used for this and the other figures in the previous subsection was the Daubechies $n = 4$ wavelet. Recall that this panel was theoretically empty, but because of the inherent imprecision of wavelet filtering some visible “leakage” into this panel occurred. As discussed earlier, since the band is within 1.25 Hz of the boundary, a degree of leakage is inevitable. However, a wavelet with less overlap should diminish this effect. For convenience Figure 28 made
**Fig. 23.** Predicted third level decomposition.

<table>
<thead>
<tr>
<th>$j$</th>
<th>$f_{\text{H}}^-$</th>
<th>$f_{\text{H}}^+$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>62.5</td>
<td>30.</td>
</tr>
<tr>
<td>2</td>
<td>31.25</td>
<td>30.</td>
</tr>
<tr>
<td>3</td>
<td>15.625</td>
<td>27.0633</td>
</tr>
<tr>
<td>4</td>
<td>7.8125</td>
<td>13.5316</td>
</tr>
<tr>
<td>5</td>
<td>3.90625</td>
<td>6.76582</td>
</tr>
</tbody>
</table>

Table 1. Theoretical frequency bounds for $\mathbf{H}^j$ for levels $j = 1$ to 5

<table>
<thead>
<tr>
<th>$j$</th>
<th>$f_{\text{D}}^-$</th>
<th>$f_{\text{D}}^+$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>108.253</td>
<td>30.</td>
</tr>
<tr>
<td>2</td>
<td>54.1266</td>
<td>30.</td>
</tr>
<tr>
<td>3</td>
<td>27.0633</td>
<td>30.</td>
</tr>
<tr>
<td>4</td>
<td>13.5316</td>
<td>15.625</td>
</tr>
<tr>
<td>5</td>
<td>6.76582</td>
<td>7.8125</td>
</tr>
</tbody>
</table>

Table 2. Theoretical frequency bounds for $\mathbf{D}^j$ for levels $j = 1$ to 5
Fig. 24. Reconstructed $V^1$ panel for 30° dip data.

Fig. 25. Reconstructed $V^2$ panel for 30° dip data.
Fig. 26. Reconstructed $\mathbf{V}^3$ panel for 30° dip data.

Fig. 27. Reconstructed $\mathbf{H}^1$ panel for 30° dip data.
FIG. 28. Reconstructed $H^2$ panel for 30° dip data.

FIG. 29. Reconstructed $H^3$ panel for 30° dip data.
Fig. 30. Reconstructed $D^1$ panel for 30° dip data.

Fig. 31. Reconstructed $D^2$ panel for 30° dip data.
Fig. 32. Reconstructed $D^3$ panel for $30^\circ$ dip data.

Fig. 33. Reconstructed $A^1$ panel for $30^\circ$ dip data.
Fig. 34. Reconstructed $A^2$ panel for 30° dip data.

Fig. 35. Reconstructed $A^3$ panel for 30° dip data.
with the \( n = 4 \) Daubechies wavelet is repeated as Figure 36, along with the corresponding result for the \( n = 10 \) Daubechies wavelet (Figure 37). Although a degree of “leakage” due to overlap remains for this longer wavelet, the comparison makes the point that the overlap phenomena can be ameliorated by the choice of wavelet. However, using a longer wavelet is not a panacea—in particular, for operator compression, shortness of the wavelet is an over-riding concern. On the other hand, it is important to fit the wavelet to the task and, as the figures in Appendix B show, Meyer wavelets have markedly better filtering properties than the commonly used Daubechies wavelets. This fact will play a key role in the next section.

**Processing spatially aliased data**

An an illustration of some features of wavelet filtering already discussed theoretically, consider the problem of spatial aliasing. Note that no assertion of the appropriateness of this application of the wavelet transform is being made. Assume then, that the data with dip \( \theta \) and frequency \( f_t \) are spatially aliased

\[
f_x > F_x \quad \text{or} \quad f_t > F_x \cot \theta.
\]

As is typical of field seismic sections, assume also that the data are not temporally aliased, so that

\[
f_t < F_t.
\]

The last two equations imply that

\[
\theta > \arctan(F_x/F_t)
\]

for such data. Thus \( \theta > \theta_{HV} \) for such spatially aliased seismic data.

The situation is not as favorable in the other wavelet decomposition panels, \( D^j \) and \( V^j \). Worse—these are the panels that are typically contaminated by spatial aliasing. Consider the effect of spatial aliasing: the range \([F_x, 3F_x/2]\) is folded down to the range \([F_x/2, F_x]\), that is, into \( D^1 \) and \( V^1 \). Most spatially aliased data will have frequencies in this range of “slightly” aliased data and this critical regime is discussed further below. Next, the range \([3F_x/2, 2F_x]\) is folded into \([0, F_x/2]\), that is, into \( H^1 \) and \( A^1 \). The portion folded to \( A^1 \) later contaminates higher level panels in the wavelet decomposition. Then, the range \([2F_x, 5F_x/2]\) containing doubly aliased data is again folded into \([F_x/2, F_x]\) and the pattern continues in the obvious way.

A full cure for spatial aliasing probably doesn’t exist without making some special assumptions. A radical treatment is to low-pass filter the entire data set to eliminate the spatial aliasing. But this entails losing resolution on the shallow dips when typically only the steep ones are spatially aliased. A proposal that seems natural for wavelet decomposed data is to apply filtering only to the panels that need it.

As an experiment in de-aliasing data, consider a case where only the \( V^1 \) panel is contaminated with spatial aliasing. Continue using the parameter values, \( F_t = F_x = 125 \text{ Hz} \), of the last section for a single dipping plane with dip angle 75° and frequency band 0–40 Hz. In this case, the dip is significantly larger than \( \theta_{D_{\text{max}}} \approx 61^\circ \), so that unaliased data will appear only in the \( A \) and \( V \) panels. Moreover, \( f_t^+ = 40 \text{ Hz} \) and \( f_x^+ = 149 \text{ Hz} \), so that the
Fig. 36. Reconstructed $\mathbf{H}^2$ panel for 30° dip data with the $n = 4$ Daubechies wavelet.

Fig. 37. Reconstructed $\mathbf{H}^2$ panel for 30° dip data with the $n = 10$ Daubechies wavelet.
Fig. 38. Synthetic section with a single dipping plane that is spatially aliased.

aliasing (appearing as negative spatial frequencies, cf. Figures 16 and 38) is well within the $V^1$ panel. This is a favorable case for de-aliasing by just low-pass filtering the $V^1$ panel while leaving the other panels untouched. Figure 39 shows the result of

1. Decomposing to the first level with a Daubechies $n = 4$ wavelet.
2. Applying a smooth low-pass filter to the $V^1$ panel to remove the aliased frequencies.
3. Reconstructing to obtain the de-aliased version of the data.

Indeed, the extension of the dipping plane on the negative frequency side has been removed, but the reconstruction contains several artifacts and the highest frequencies on the reflector have been noticeably damaged. Although the new artifacts are weak compared to the spatial aliased segment successfully removed, they are disturbing because they are coherent.

The explanation for this problem lies in the low-pass filtering of the overlap region of the mirror filter, see Figure B-1. This filtering of $V^1$ has removed the overlaps that are needed in the reconstruction process to remove the intrinsic wavelet filter aliasing. Thus, the low-pass filter has replaced one kind of aliasing with another. This explains the weak, but coherent, artifacts in Figure 39.

One way to improve the result is use a bandpass filter in an attempt to avoid damaging the overlap region in $V^1$. The result of one such attempt is shown in Figure 40. This approach has preserved more of the unaliased part of the plane (cf. Figure 39), but still shows some overlap aliasing. Using a wavelet with large overlap, such as the Daubechies
Fig. 39. Reconstructed 75° dip data after low pass filtering the $V^1$ with the Daubechies $n = 4$ wavelet. Note the appearance of artifacts.

$n = 4$ makes it impossible to find a pass band that removes the spatial aliasing while avoiding the overlap region.

A more fundamental way to obtain a satisfactory result is to use a wavelet with smaller overlap regions. Figure 41 shows the result of low-pass filtering just as in Figure 39, except that here the decomposition used the Meyer wavelet of order $n = 10$, which has very small overlap (cf. Figure B-2). Now the low-pass filtering of $V^1$ and recomposition has successfully removed the spatial aliasing from the data without creating new artifacts.

An alternate way of exploiting a small-overlap filter is to use a bandpass filter that removes the aliased frequencies but avoids filtering in the overlap region. In general, one would be trading off incomplete removal of the spatially aliased frequencies against removal of the intrinsic overlap aliasing (thus creating overlap aliasing artifacts in the reconstruction). Fortunately, with the $n = 10$ Meyer wavelet, such fine considerations seem unnecessary.

Now turn to the more challenging problem of de-aliasing a section containing two dipping planes. Figure 42 is a synthetic section consisting of two dipping planes, one with dip angle 70° and frequency band 0–60 Hz, the other with dip angle -35° and frequency band 0–85 Hz. Figure 43 is the f-k amplitude spectrum of this section. Clearly the 70° dipping plane is spatially aliased. But since $f_{l} = 60$ Hz and $f_{h} = 165$ Hz, the aliasing occurs only in the $V^1$ panel. The -35° dipping plane, on the other hand, is confined to the $A^1$, $H^1$ and $D^1$ panels, because its dip is significantly smaller than $\theta_{HV} = 45°$. So after the first stage of wavelet transform, the aliased information is separated from the good information. Thus low-pass filtering the $V^1$ panel will remove only the aliased portion. Figure 44 shows
Fig. 40. Reconstructed 75° dip data after band pass filtering the $V^1$ panel constructed with the Daubechies $n = 4$ wavelet. Note that the wavelet filtering artifacts are diminished.

the f-k amplitude spectrum after the three-step process of wavelet decomposition, low-pass filtering on only the $V^1$ panel, and reconstruction. The process is a complete success: the aliasing is removed and the shallow dipping plane is untouched—retaining its full resolution as indicated in Figure 45.

In less favorable cases, one may have to also devise appropriate filters for $D^1$, $H^1$ and the higher level panels.

Filtering with two-dimensional wavelets

The above discussion of spatial aliasing made a good test bed for demonstrating various aspects of the two-dimensional wavelet transform. These included exploiting the division into horizontal, vertical and diagonal panels implied by the basis given in equation (7) and controlling the intrinsic aliasing by choice of analyzing wavelet. However, the applications shown could have been done with pure Fourier methodology because the spatial-temporal localization of the wavelet basis wasn’t brought to bear. A more suitable problem would be the treatment of spatial aliasing—or some other phenomenon—occurring in a local region of space-time, $x$-$t$, as well as a local region of the frequency plane, $f_x$-$f_t$.

That is, just as in the case of the one-dimensional transform, applications most suited to wavelet transform analysis are those that effectively use the simultaneous localization in data-domain and frequency-domain. The wavelet basis given in equation (7) implies the
Fig. 41. Reconstructed 75° dip data after low-pass filtering the $V^1$ panel constructed with the Meyer $n = 10$ wavelet. Note that there are no wavelet filtering artifacts.

Fig. 42. Synthetic section with two dipping planes. The steeper dip is spatially aliased.
Fig. 43. The f-k amplitude spectrum of the two dipping planes.

Fig. 44. The f-k spectrum after de-aliasing.
Fig. 45. Difference of the original data and the de-aliased data. Notice that the shallow event does not appear, indicating success in de-aliasing the steep dip without harming resolution on the shallow dip.

following theoretical expansion of the data:

\[ s(t, x) = \sum_{j, k'} c_{jk'}^H \psi_{jk'}^H(t, x) + \sum_{j, k'} c_{jk'}^D \psi_{jk'}^D(t, x) + \sum_{j, k'} c_{jk'}^V \psi_{jk'}^V(t, x). \]  

(24)

Notice that in this “non-standard” two-dimensional basis, there is still only one frequency parameter \((j)\). The natural two-dimensional product basis implied by equation (6) would afford separate localization in \(f_x\) and \(f_t\), but this has been traded away in favor of the dip localization afforded by the use of the three basis functions, \(\psi_{jk'}^H\), \(\psi_{jk'}^D\) and \(\psi_{jk'}^V\) at level-\(j\) in place of the single function \(\psi_{jk'}(x, t) = \psi_{jk}(t)\psi_{jk'}(x)\) appearing at level-\(j\), \(j'\) in the product basis.

In the non-standard basis, frequency localization by octave (in both \(f_t\) and \(f_x\)) is attained by selecting \(j\)-values. Then selection of panel \(H\), \(D\) or \(V\) further confines the frequency to a half-octave (in both \(f_t\) and \(f_x\)) and also confines the apparent dips. Within a panel, temporal localization is attained by selecting \(k\)-values and spatial localization is attained by selecting \(k'\)-values.

And, just as in one dimension, wavelet packets in \(t\) and/or \(x\) or the continuous wavelet transform in \(t\) and/or \(x\) may be superior tools for applications requiring finer frequency control than stepping by octaves.
ACKNOWLEDGEMENTS

Thanks to Gregory Beylkin, both for his patient theoretical support and for the use of his proprietary two-dimensional decomposition software. We thank Bee Bednar for sharing his knowledge and codes and John E. Anderson for helping us with his extensive knowledge of practical issues in seismic data processing. We also are grateful for important discussions with many others of our colleagues at the Colorado School of Mines, especially John W. Stockwell, Norman Bleistein, Ilya Tsankin and Frank G. Hagin. We are also indebted to the excellent computing research facilities at the Center for Wave Phenomena and the Center for Geoscientific Computing which includes RS6000 workstations, the NeXT software environment, and the Center for Wave Phenomena Seismic Unix data processing environment. Finally, we gratefully acknowledge the support of the members of the Consortium Project on Seismic Inverse Methods for Complex Structures at the Center for Wave Phenomena, Colorado School of Mines.

REFERENCES


APPENDIX A: SAMPLE WAVELETS

The Daubechies family of wavelets is indexed by the number $n$ of vanishing moments possessed by the analyzing wavelet. $n = 1$ is the Haar example. Figure A-1 shows some higher-order examples.

![Daubechies wavelets](image)

**FIG. A-1.** Some Daubechies scaling functions and wavelets.

The Daubechies wavelets are defined recursively, not pointwise. That is, there is no explicit function formula for these wavelets. Instead, there is an algorithm that first computes their values at the integers and then refines to the half-integers, quarter-integers, etc. The plots shown indicate the recursion level reached by stating the increment $dx$ between successive computed points. Observe that the Daubechies wavelets have finite support on $[0, 2n - 1]$. Thus, the Fourier transforms must have non-finite support. But, as stated earlier, this is a theoretical result only. As can be seen in the amplitude plots of Figure A-2, the Daubechies wavelets fall off rapidly in frequency domain and so “effectively” have finite support there as well. Daubechies has also constructed other families of finitely supported wavelets. The ones shown here are the “minimum phase” members of the family of Daubechies finite-support wavelets. There is also a corresponding set of maximum symmetry wavelets (zero-phase is not attainable for compact support wavelets) and a set of wavelets known as **coiflets** that possess especially desirable properties for operator compression. See Daubechies (1992) for more discussion and additional plots.
Fig. A-2. Fourier transforms of Daubechies scaling functions and wavelets.
The Meyer wavelet family is parametrized by an auxiliary taper function that rises from 0 to 1 over the unit interval. Typically, polynomial taper functions are used and then the Meyer family can be indexed by the order of smoothness at the endpoints of the taper function. Figure A-3 shows several of these scaling functions and wavelets. The corresponding Fourier transforms (amplitudes) are shown in Figure A-4. Notice that the situation for the Meyer wavelets is just the opposite of that for the Daubechies wavelets. Here, the Fourier transforms are strictly of finite support, while the wavelets and scaling functions themselves are merely of “effective” finite support.

The Battle-Lemarié wavelets are closely related to B-splines and are indexed by the degree of the corresponding spline. Figure A-5 shows the scaling functions and wavelets for four cases. The corresponding Fourier transforms (amplitudes) are shown in Figure A-6. Here, neither the functions nor the Fourier transforms are strictly of finite support. The figures show the trade-off between size of support in each regime.
Fig. A-4. Fourier transforms of Meyer scaling functions and wavelets.
Fig. A-5. Battle-Lemarié scaling functions and wavelets.
Fig. A-6. Fourier transforms of Battle-Lemarié scaling functions and wavelets.
APPENDIX B: SAMPLE WAVELET FILTERS

Turn now to the low/high-pass filters for the three wavelet families discussed in appendix A. The figures in this appendix show the positive-frequency portion of these mirror filters with the low-pass filter $m_0$ displayed dark and the high-pass filter $m_1$ displayed lighter. In each case, the overlap region about the quarter-period point indicates the “fuzziness” of the frequency aspect of the wavelet transform at each of its levels. Figures B-1, B-2, and B-3, respectively, show several members of the Daubechies, Meyer and Battle-Lemarié families. In each case, the figure at the upper right may be regarded as the “typical” member of the family.

**Fig. B-1.** Daubechies m0, m1 filters.

**Fig. B-2.** Meyer m0, m1 filters.
Fig. B-3. Battle-Lemarié $m_0$, $m_1$ filters.