Noise suppression in large wavenumber Fourier imaging

Lan Wang & Norman Bleistein
Center for Wave Phenomena, Colorado School of Mines

ABSTRACT

For the band-limited inverse problem, the inversion formulas developed by Bleistein *et al.* generate a reflector map as well as an estimate of reflection coefficients from the discontinuities of the medium parameters. Here, we describe the use of a multiscale smoothing operator to suppress the noise while preserving the features of the original processing formula. This is accomplished by applying a simple multiplier in wavenumber domain to the output of the inversion formalism. A comparison of this approach to wavelet based edge detection processing shows that this continuous multiscale operator allows more flexibility for changing the length of the smoothing operator.

Key words: multiscale smoothing operator, noise suppression, Fourier inversion, imaging, wavelet transform

Introduction

Rapid changes in the medium parameters characterize the structure of the subsurface; they are the reflectors in the earth. In situations where the medium parameter is related to a velocity field, or the perturbation in the velocity field, it is typical to think in terms of piecewise smooth functions whose discontinuity surfaces represent reflectors. The object velocity fields discussed in this report all belong to this class of step-like piecewise smooth functions.

The inversion formulas developed by Bleistein *et al.*(1996) provide a tool for obtaining correct locations of interfaces as well as model-consistent specular reflection coefficients. However, because of the band-limited nature of seismic data, the step-like wavespeed perturbation functions are not well reconstructed using this formulation. By modifying the inversion operator (Bleistein *et al.*, 1996), the output can be transformed into the singular functions of the discontinuity surfaces of the original function, scaled by the jump in the original function at each point of the discontinuity surface. This modification is a normal derivative operator in the spatial domain, achieved through an appropriate multiplier in the Fourier domain derived from asymptotic analysis.

The derivative operation is sensitive to the noise present in the wavefield. In this report, a multiscale smoothing procedure is introduced to stabilize the inversion algorithm and improve the images. The multiscale operator is based on the Fourier transform. The algorithm extracts the rapid changes in the velocity field as a function of location and scale, while noise is largely suppressed.

More specifically, we introduce a scale, s, in the convolution operator. It is known (Mallat & Hwang, 1992), that the effect of the scaling operator on noise, $N(\mathbf{x})$, is to produce an output $N_s(\mathbf{x})$ having the property that

$$|N_s(\mathbf{x})| = \mathbf{O}(\mathbf{s}^{\nu}), \quad \nu < \mathbf{0}. \tag{1}$$

That is, $N_s(\mathbf{x})$ decays with increasing s. On the other hand, for a piecewise smooth signal, $I(\mathbf{x})$, we know that the scaled outputs satisfy

$$|I_s(\mathbf{x})| = \mathbf{O}(\mathbf{s}^{\mathbf{0}}). \tag{2}$$

Thus, its peak amplitude does not decay with increasing s. However, when $I(\mathbf{x})$ is a delta function, $I_s(\mathbf{x})$ is a bandlimited delta function whose resolution decreases with increasing s. Thus, it is a "race" between noise suppression and resolution. Our tests with parameters typical of seismic data suggest that we win the race in this application.

The Bleistein-Cohen (BC) inversion formulas have the structure of Fourier transform-like integrals. One fundamental concept of the BC inversion approach is applying the underlying relationship between the observations and the Fourier transform of the function(s) that characterize(s) the unknown medium. Thus, in the following sections, attention will be primarily focused on a multiscale smoothing operator applied to Fourier imaging as a test of its ability to suppress noise. Thereafter, we address the applications of this method to the BC inversion.

This discussion adds the structure of scaling analysis to the fairly standard noise reduction processes currently in use.

Fourier Imaging

In seismic applications such as filtering, imaging, and migration/inversion, one basic concept is the reconstruction of the data by performing forward and inverse Fourier transforms with proper "focusing" filters that depend on both domains. We are concerned with aperture-limited Fourier-like transforms of data $f(\mathbf{x}')$. Such an output can be written as a cascade of forward and inverse Fourier transforms acting on $f(\mathbf{x}')$,

$$I(\mathbf{x}) = \frac{1}{(2\pi)^n}\int_{D_k} d^n k \int_{D_{x'}} d^n x' \; a(\mathbf{x}', \mathbf{k}) e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')} f(\mathbf{x}'), \quad (3)$$

where n is the dimension of interest; $\mathbf{x}$ and $\mathbf{x}'$ are n component vectors; $\mathbf{k}$ is the wave vector. $D_{x'}$ and D_k are the supports of $f(\mathbf{x}')$ in space and wavenumber domain respectively; that is, they define the apertures. In seismic data processing, D_k is constrained to a "large wavenumber" domain; the nature of the D_k is essential to the reconstruction of $f(\mathbf{x}')$. In the applications to BC inversion, f is related to the wavespeed perturbation of the subsurface. It represents a piecewise smooth function, from which it follows that the domain of the integration can be decomposed into separate domains whose boundaries include all of the discontinuities of f. The amplitude of the integrand $a(\mathbf{x}', \mathbf{k})$ is the "focusing" operator required to reconstruct the image from the data; it is allowed to depend on both $\mathbf{k}$ and $\mathbf{x}'$. The appropriate filter arises from properties of aperture-limited large wavenumber Fourier transforms. To simplify the problem and focus our attention on some insights of Fourier inversion, we start our discussion with $a = 1$ in (3).

Bleistein *et al.*(1996), have shown that the high wavenumber aperture-limited Fourier inversion of the piecewise smooth function $f(\mathbf{x}')$ is dominated by the values on the discontinuity surfaces. The inversion is approximately a band-limited step function of normal distance from each point on the discontinuity surface. The amplitude of the step function is proportional to the jump in the function across the surface at the point in question.

For better imaging of the discontinuity surfaces, the following filter in wavenumber domain is applied (Bleistein *et al.*, 1996):

$$ik\mathrm{sgn}(\hat{\mathbf{u}} \cdot \hat{\mathbf{k}}). \quad (4)$$

That is, the filter $a(\mathbf{x}', \mathbf{k})$ is replaced by the above multiplier. Here, $k = |\mathbf{k}|$ is the magnitude of vector $\mathbf{k}$, $\hat{\mathbf{k}}$ is the unit vector and $\hat{\mathbf{u}}$ is a constant vector. In seismic applications, there is usually little or no information about the plane, $k_z = 0$, so $\hat{\mathbf{u}}$ can be chosen to be $(0, \pm 1)$ in $2D$ or $(0, 0, \pm 1)$ in $3D$, respectively, so that $\mathrm{sgn}(\hat{\mathbf{u}}\cdot\hat{\mathbf{k}}) = \pm\mathrm{sgn}k_z$. The multiplication by $ik\mathrm{sgn}(\hat{\mathbf{u}}\cdot\hat{\mathbf{k}})$ replaces the Fourier inversion of $f(\mathbf{x}')$ by the singular function(s) of its boundary surface(s), with the amplitude proportional to the jump in $f(\mathbf{x}')$ at each point on the surface. The method breaks down at points where $\hat{\mathbf{u}} \cdot \hat{\mathbf{n}}$ is nearly zero, with $\hat{\mathbf{n}}$ being the unit normal to the reflector. In the application to inversion, $ik\mathrm{sgn}(\hat{\mathbf{u}} \cdot \hat{\mathbf{k}})$ is replaced by $i\omega|\bigtriangledown\phi|$, with ϕ being traveltime. In this case, no direction $\hat{\mathbf{u}}$ is distinguished; all directions are treated the same. Bleistein *et al.*(1996) have shown that this filter behaves asymptotically like a normal derivative operator in the spatial domain, even though the normal direction is not known *a priori*.

Multiscale Operator for Noise Suppression

This directional derivative operator is sensitive to noise. It enhances noise in the data at large wavenumbers more than at small wavenumbers. In order to extract the important object boundaries from the data, we need to find the local maxima representing the peaks of the singular functions while minimizing the effect of noise. A multiscale smoothing operation can do the trick. It extracts the most rapid changes in the neighborhood defined by the scale parameter while largely suppressing the noise.

The smoothing process can be viewed as a convolution operation,

$$\Theta(\mathbf{x}') * f(\mathbf{x}'). \quad (5)$$

Define $\Theta_s(\mathbf{x}')$ as the multiscale transform, given by,

$$\Theta_s(\mathbf{x}') = \Theta(\mathbf{x}'/s). \quad (6)$$

Here, s is the scale parameter. It is used to control the size of the neighborhood where the local maxima of derivatives are computed. At coarse scales - large s - a large neighborhood is encompassed, whereas at fine scales - small s - small neighborhoods contribute to the convolution.

In this report, we choose the smoothing operator $\Theta(\mathbf{x}')$ to be a Gaussian, i.e.,

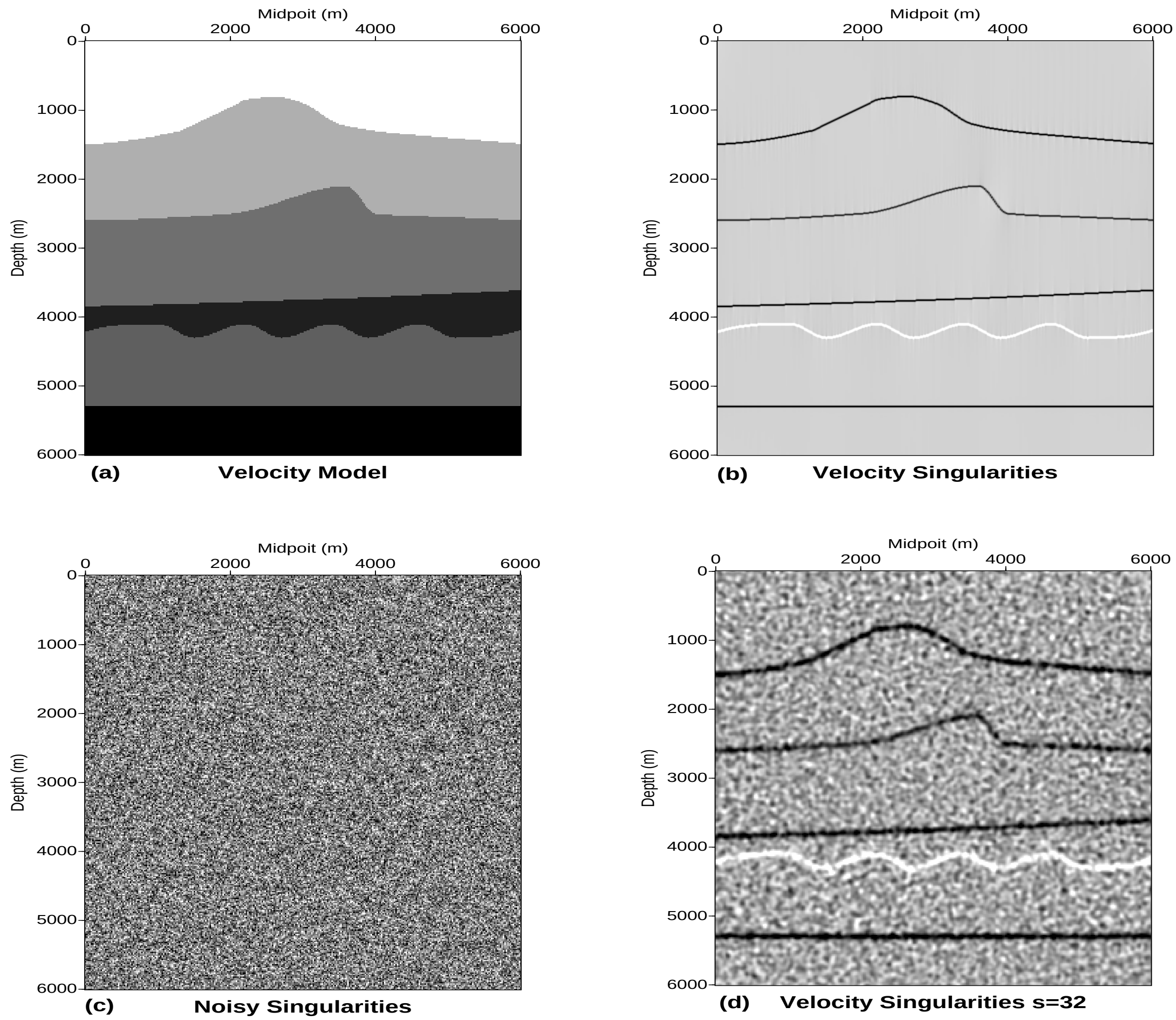

Figure 1. (a) Velocity model, (b) velocity singularities, (c) velocity singularities of noisy data with signal to noise ratio of 4/3, and (d) image of (c) by applying the smoothing operator.

$$\Theta(\mathbf{x}') = \exp\left(-\mathbf{x}' \cdot \mathbf{x}'/2\right). \tag{7}$$

The Gaussian filtering kernel function is infinitely differentiable, and has the same form in wavenumber domain. Thus, the operation is just the following multiplication,

$$\check{\Theta}_s(\mathbf{k}) = \sqrt{2\pi}s \exp\left(-s^2 \mathbf{k} \cdot \mathbf{k}/2\right), \tag{8}$$

in the wavenumber domain before inverting the transform.

By combining the multiscale smoothing operator with the directional derivative operator in the previous section, we obtain the multiscale operator,

$$\Psi_s(\mathbf{k}) = ik\,\mathrm{sgn}(\hat{\mathbf{u}} \cdot \hat{\mathbf{k}})\check{\Theta}_s(\mathbf{k}). \tag{9}$$

Replacing $a(\mathbf{x}', \mathbf{k})$ in (3) with (9), we propose the following processing formula for detection of the discontinuities of $f(\mathbf{x}')$, while simultaneously suppressing the noise:

$$\bar{I}_s(\mathbf{x}) = \frac{1}{(2\pi)^n} \int_{D_k} d^n k \ \Psi_s(\mathbf{k}) \int_{D_{x'}} d^n x' \ f(x') e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')}. \tag{10}$$

Figure 1 is an example of applying this multiscale singularity detection procedure in two dimensions. Figure 1-a shows a $2D$ earth model with constant velocity in layers and strong discontinuities. Rapid changes in the medium parameters are clearly indicated in Figure 1-b. Here, we have applied the normal derivative operator (4)

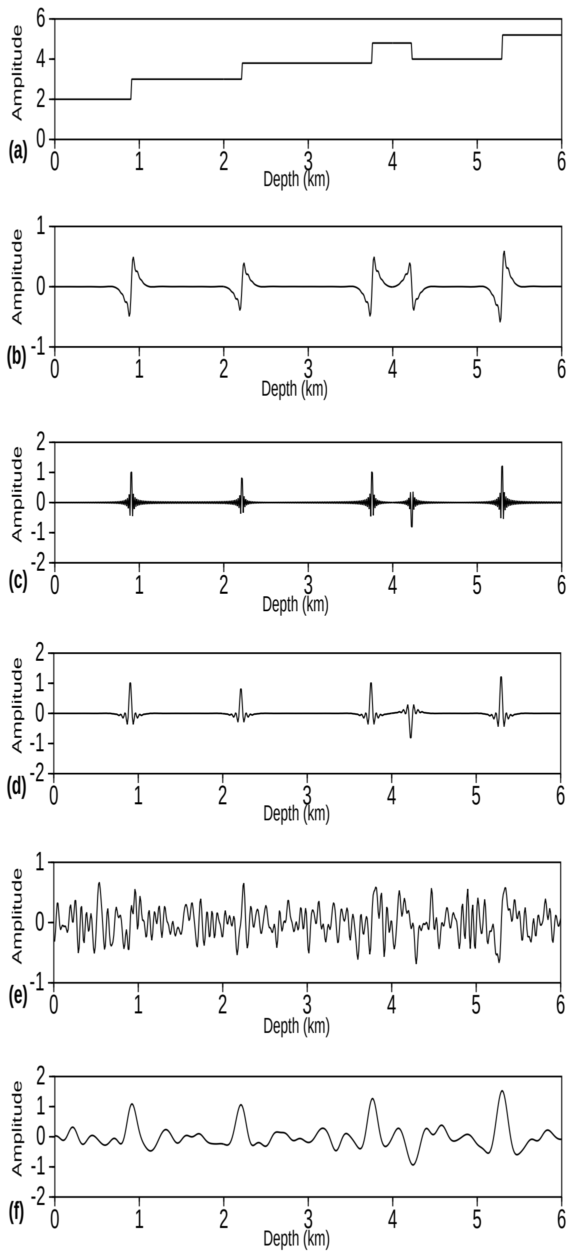

Figure 2. (a) Single trace of full-band data, (b) band-limited data, (c) singularities of full-band data (a), (d) singularities of band-limited dada (b), (e) band-limited noisy data and (f) image of (e).

to the noise-free data. When noise is added to the original model at a signal to noise ratio of 4/3, the directional derivative operator (4) enhances the noise as is shown in Figure 1-c. The structural information is seemingly lost, however, the rapid changes can still be recovered at a certain scale s, using (8-10). This is shown in Figure 1-d where we have computed (9) with $s = 32$. The scaling behavior of the noise is such that it is largely suppressed when compared with Figure 1-c. This result compares favorably with Figure 1-b.

There are various algorithms based on wavelet transforms that address this goal of noise suppression (Song & He, 1996), (Dessing *et al.*, 1996). Basically, all the algorithms consist of an operator having a behavior roughly comparable to a (band-limited) differentiation and smoothing operation followed by extraction of the local maxima of the derivative of the data. The local maxima occurs at the singular surfaces and in the direction perpendicular to them. The direction of the preference is not known *a priori*. Thus, the wavelet operation is often carried out by computation of the modulus function from mutually perpendicular wavelet transforms. The modulus function is essentially the magnitude of the gradient operator. Thus, it can extract the local maximum changes within the region defined by the scale parameter; however, it fails to indicate the sign of the data changes. On the other hand, the multiplier (4) will produce a normal derivative except when that normal is itself perpendicular to $\hat{\mathbf{u}}^{\star}$. Because seismic reflectors are rarely absolutely vertical, we can choose $\hat{\mathbf{u}}$ as a unit vector in the vertical direction and image the normal derivative in a practical range of normal directions. An advantage of this operator is that it reveals the magnitude of the step jumps by its spike amplitude. The white sinusoid in Figure 1(d) shows that it has a negative amplitude, which is consistent with the decrease of the wavespeed in the original model.

For further illustration, we process one trace of the output in Figure 1. Figure 2-a shows the trace. The Nyquist wavenumber is $50/km$. Figure 2-b shows the degradation of the bandlimited $2 - 16/km$ data when compared to the full-band data (Figure 2-a). The effect of losing low wavenumber components is the flattening of the data away from the discontinuities; if the data at zero frequency is zero, then the average of the inverse wavenumbers must be zero. Further low wavenumber filtering transforms the steps into doublets. The high wavenumber filtering only affects the width of the doublets; $16/km$ causes insignificant loss of resolution. The steps are undeterminable although the locations of the discontinuities of the data are recognizable. Figure 2-c shows the singularities of the full-band noise-free data (Figure 2-a); it is the output of applying the same procedure as shown

* However, by the observation in the first section, that, in application in our inversion technique, there is no distinguished direction, $\hat{\mathbf{u}}$.

in Figure 1-b. Figure 2-d is the corresponding result for band-limited data 2-b. A comparison of Figure 2-c and 2-d implies that the bandlimiting effect on the normal derivative operator (4) is only the insignificant differences on the widths of resolutions determined by the bandwidths. The degradation of bandlimiting on the original data has been overcomed by employing the normal derivative multiplier (4). Figure 2-e is the band-limited noisy data by adding noise to the band-limited data 2-b at a signal to noise ratio of 4/3. Figure 2-f shows the result of applying the noise filtering derivative operator proposed in the paper to these data. By comparing the peak locations here with those in Figure 2-b, we see that the discontinuities are still well recovered from the band-limited data, even though the low wavenumber components that contain information about the steps of the original function are missing. The peaks of the Gaussians clearly reveal the locations of the discontinuities, while their heights indicate the magnitude of the step scaled by the area under the filter applied to the data. The widths of the Gaussians are a manifestation of the high pass filtering effect of setting $s = 32$; this choice of scale was determined as a compromise between higher values through which further resolution was lost and lower values for which the noise level was not acceptable.

Singularity Reconstruction and Amplitude Preservation

In this section, we will analyze some features of the multiscale Gaussian operator (9) as applied to singularity detection. We will show that this operator reconstructs the discontinuities of the function by Gaussians. The location of each discontinuity is indicated by the peak of the Gaussian, while the step height proportional to the peak amplitude. The amplitude remains unchanged across different values of the scale s.

We consider one specific piecewise smooth function of $f(\mathbf{x}')$ in (10): $f(\mathbf{x}) = AH(\hat{v}\cdot(\mathbf{x}-\mathbf{x}_0))$, where $H(\mathbf{x})$ is a step function having discontinuities at the set of $\mathbf{x}_0$; $\hat{v}$ denotes the normal at each point on the discontinuous surface; and A is a scalar. When $D_{x'}$ in (10) includes all the discontinuities, and D_k is of infinite extent, the calculation can be carried out in a similar way as in the corresponding aperture-limited Fourier inversion examples in (Bleistein *et al.*, 1996). Here we simply state the solution,

$$\bar{I}_s(\mathbf{x}) = A\exp\{-\frac{[\hat{v}\cdot(\mathbf{x}-\mathbf{x}_0)]^2}{2s^2}\}. \tag{11}$$

This implies that the full-band multiscale Fourier inversion characterizes the singularities by the Gaussian. The peak of Gaussian, when $\mathbf{x} = \mathbf{x}_0$, indicates the location of the singularity, and its peak amplitude is just the magnitude of the step jump. It will not change with varying s.

Let us now consider the effect of bandlimiting on these results. We assume that D_k is a region symmetric about $\mathbf{k} = \mathbf{0}$, so that we can then consider symmetric intervals in each $k_j, j = 1, 2, ...n$. The solution can be rewritten as,

$$\begin{aligned}\bar{I}_s(\mathbf{x}) \quad &= \quad A\frac{s}{\sqrt{2\pi}}\exp\{-\frac{[\hat{v}\cdot(\mathbf{x}-\mathbf{x}_0)]^2}{2s^2}\}\cdot \\ &\left[\int_{\kappa_{v_-}}^{\kappa_{v_+}}\exp(-\frac{s^2\kappa_v^2}{2})d\kappa_v \right.\\ &\left. +\eta(s,\mathbf{x}-\mathbf{x}_0,\kappa_{v_-},\kappa_{v_+})\right],\end{aligned} \tag{12}$$

where κ_v is the wavenumber in the $\hat{v}$ direction; κ_{v_-} and κ_{v_+} are the two end points of κ_v; $\eta(s,\mathbf{x}-\mathbf{x}_0,\kappa_{v_-},\kappa_{v_+})$ is a lower order term and equals 0 at $\mathbf{x} = \mathbf{x}_0$ where the maximum of $|\bar{I}_s(\mathbf{x})|$ occurs.

Equation (12) indicates the effect of band-limited inversion compared to the full-band Fourier image. The difference between (11) and (12) at their peak values is a difference of error functions governed by the limits of integration in κ_v. For appropriate choices of $k_{v\pm}$ and s, the bandlimited output in (12) is indistinguishable from the full bandwidth version in (11). Again, the above equation implies that the peak amplitude is the step jump scaled by a factor that is asymptotically constant in s, namely 1.

The Scale Parameter and Characterization of the Singular Behavior of Functions

The behavior of the multiscale operator (9) is such that it characterizes the velocity singularities within the region defined by the scale parameter while smoothing the region. In mathematics, singularities are generally characterized by their Lipschitz exponents. A function $f(x)$ has Lipschitz exponent of ν at x_0 if and only if there exists a constant B such that for all x in a neighborhood of x_0, we have

$$|f(x) - f(x_0)| \leq B|x-x_0|^{\tilde{\nu}}, \quad \tilde{\nu} \geq \nu. \tag{13}$$

That is, ν is a lower bound on the powers of differences in x that bound differences in the function values. The function, $f(x)$, is uniformly Lipschitz ν over an open interval (a, b) if and only if the constant B holds for any $x, x_0 \in (a, b)$. For example, if $f(x)$ is a step function discontinuous at x_0, then its Lipschitz exponent is $\nu = 0$. The isolated singularity is the worst singularity inside a region that contains the isolated point. The uniform

Lipschitz regularity of a function over a certain region is then equal to the pointwise Lipschitz regularity at the isolated point. (Song & He, 1996).

When noise is present, the observed data becomes $f(\mathbf{x}) + n(\mathbf{x})$, and the image $\bar{I}_s(\mathbf{x})$ in (10) becomes $\tilde{I}_s(\mathbf{x})$; that is,

$$\tilde{I}_s(\mathbf{x}) = \bar{I}_s(\mathbf{x}) + N_s(\mathbf{x}). \tag{14}$$

If $n(\mathbf{x})$ is white noise, the Lipschitz exponent ν of $N_s(\mathbf{x})$ is less than zero (Song & He, 1996), (Mallat & Hwang, 1992). That is, there exists a constant B, such that

$$|N_s(\mathbf{x})| \leq Bs^{\nu}, \quad \nu < 0. \tag{15}$$

This inequality implies that $N_s(\mathbf{x})$ should decrease when the scale s increases. On the other hand, the peak values of $\bar{I}_s$ remains unchanged with s increasing according to equations (11) and (12). Hence, depending upon the differences between the signal singularities and the noise singularities, we can take the advantage of the multiscale operator (9) to suppress the noise in the data at a faster rate than the loss of resolution of the derivative at discontinuities. We also remark that the scale s should stay bounded above for seismic data, because of the spatial resolution being lost for increasing scale.

The scale parameter determines the size of the window for optimal peak detection combined with noise suppression. In different problems, we can consider changing the window by simply varying the scale parameter within a reasonable range. In the pure discrete wavelet approach, scaling is usually discretized to powers of 2 to facilitate synthesis of the processed signal from its wavelet transform. In contrast the multiscale operator has the advantage that we search s values over a continuous range to optimize our choice for simultaneous singal detection and noise suppression. In this application of scaling analysis, there is not need for synthesis – we will extract all information from the transform of the original data. Hence, we need not restrict our scales to powers of 2.

Applications to BC Inversion

The multiscale smoothing operator is essentially a smooth low-pass filter in wavenumber domain. It facilitates the numerical identification of the discontinuity surface(s) and the amplitude of the jump at each point on the surface(s), while suppressing the noise in a large scale. Thus, we can apply this procedure to Bleistein's inversion formula,

$$\begin{aligned}\alpha(\mathbf{x}) \quad &= \quad \int_{\Omega} d^2\xi B(\mathbf{x},\xi) \\ &\cdot \int d\omega \exp\{-i\omega\varphi(\mathbf{x},\xi)\} u_s(\mathbf{x}_g,\mathbf{x}_s,\omega).\end{aligned} \tag{16}$$

Here, $\alpha(\mathbf{x})$ is the perturbation correction to the given reference velocity field; the spatial weighting $B(\mathbf{x},\xi)$ contains a determinant that characterizes the viability of inverting source-receiver configurations; u_s is assumed to be the observed data from a medium characterized by piecewise-smooth parameter functions with discontinuity surfaces behaving as reflectors. With the band-limited high-frequency seismic input data, the output of equation (16) is a band-limited step-like function, with its discontinuity and amplitude not very well recovered. The noise in the observed data is propagated into $\alpha(\mathbf{x})$ through the procedure. The multiscale operator can then be applied as a posterior process. The output is a series of band-limited delta functions. The locations and the reflection coefficients can be extracted, while the noise is suppressed because of the properties of this multiscale operator, as shown above.

Conclusions

We have described here a multiscale smoothing operator that simultaneously suppresses noise and accurately detects the location of discontinuities of a piecewise smooth function of the type that represents velocity models in seismic inversion. Analytical and numerical investigations confirm these results and predict that the size of the discontinuity can be estimated from the output. These properties are exactly what is needed in the implementation of this operator in the BC inversion formalism, but now, with the added feature of noise suppression.

References

Bleistein, N., Cohen, J. K., & Stockwell, Jr., J. W. 1996. *Mathematics of Multidimensional Seismic Inversion.* Colorado School of Mines.

Dessing, F. J., Hoekstra, E. V., Herrmann, F. J., & Wapenaar, C.P.A. 1996. Multiscale edge detection by means of multiscale migration. *66th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts*, 459–462.

Mallat, S., & Hwang, W. L. 1992. Singularity detection and processing with wavelets. *IEEE Trans. Inform. Theory*, **38**(2), 617–643.

Song, S. Qu, C., & He, J. 1996. Imaging of velocity singularities with multiscale operators. *J. Math. Phys.*, **37**(2), 719–731.