Tweaking minimum phase calculations

Michael P. Lamoureux and Gary F. Margrave

ABSTRACT

The minimum phase characterization of impulsive seismic sources is an essential step in the deconvolution process, to remove source signature with appropriate phase. Calculation of the minimum phase equivalent of a given signal is numerically sensitive, given the presence of logarithmic singularities in certain algorithms.

We propose a simple over-sampling in the frequency domain that accommodates the singularity, and show with a few examples that the performance is improved.

INTRODUCTION

A key step in seismic imaging, and particular in source deconvolution, is the estimation of a minimum phase wavelet representing the source signature of a dynamite blast, weight drop, or other impulsive source. A minimum phase wavelet is the unique causal signal of a given amplitude spectrum in the frequency domain that maximizes the energy near the start of the wavelet in the time domain [Robinson (1967)]. An accurate estimation of the minimum phase wavelet is critical to obtaining clear images from deconvolution [Russell (1994); Dey and Lines (1998)].

There are several methods for numerically computing the minimum phase source wavelet. Two popular methods are the Wiener-Levinson double inverse method, and the Kolmogorov or Hilbert transform methods. Both begin with an estimation of the source's amplitude spectrum based on an autocorrelation of the seismic record, and then apply a numerical algorithm to estimate the phase spectrum. Details of these algorithms are in White and O'Brien (1974), Claerbout (1976), and Lines and Ulrych (1977).

There is also a close connection between minimum phase signals and outer functions that arise in the theory of complex analysis, as described in Lamoureux and Margrave (2007) and Gibson et al. (2011). The key observation is that the Fourier spectrum of a causal signal extends to an analytic function on the unit disk in the complex plane. When the signal is minimum phase, this analytic function is given explicitly by the formula

$$F(z) = \exp\left(\int_0^1 \frac{e^{2\pi i\theta} + z}{e^{2\pi i\theta} - z} \log|\hat{f}(e^{2\pi i\theta})|\, d\theta \right), \text{ for } z \text{ in the unit disk,} \tag{1}$$

where $\hat{f}$ is the Fourier spectrum of the signal in question. Details on complex analysis and outer functions in particular are available in Hoffman (1982) and Helson (1983).

What is important to notice in this integral is that there is a logarithm involved, which is very sensitive to zeros in the amplitude spectrum $|\hat{f}|$. In particular, the logarithm evaluates to minus infinity at a spectral zero which will cause problems in the numerical evaluation of the integral. This is a hint that the numerical calculation of a minimum phase wavelet might be sensitive to zeros in the spectrum – even when the spectrum is known exactly.

Examining the details of the Kolmogorov method, one finds that there is again an integral involving the log amplitude spectrum. Specifically, the log amplitude spectrum and phase spectrum of a minimum phase wavelet forms a Hilbert transform pair; thus, the phase is computed by calculating the Hilbert transform (an integral transform) of the log amplitude spectrum. The Levinson double-inverse computation, although not explicitly involving logarithms, also has numerical instabilities which we will see in the examples below.

EXAMINING THE LOGARITHMIC SINGULARITY

At the heart of the Wiener-Kolmogorov minimum phase calculation is the evaluation of a Hilbert integral transform of the log amplitude spectrum,

$$\mathcal{H}(\log|\hat{f}|)(e^{2\pi i\omega}) = p.v. \int_0^1 \cot\pi(\omega-\theta)\log|\hat{f}(e^{2\pi i\theta})|\,d\theta, \tag{2}$$

where f is our seismic source signal, $\hat{f}$ its Fourier transform and $|\hat{f}(e^{2\pi i\theta})|$ the amplitude spectrum at normalized frequency θ. Similarly, the outer function calculation in Equation (1) also involves an integral across the log amplitude spectrum. In both cases, a numerical instability may be introduced because the spectrum could have a zero in it, or a value close to zero. For this situation, the logarithm diverges to minus infinity which will cause problems when attempting a numerical integration across the singularity.

To illustrate the problem, consider the simpler case of integrating the function $\log(x)$ near its singularity at $x = 0$. The Hilbert transform is a convolution across the log while the outer function calculation is a weighted averaging across this log. So the simple $\log(x)$ example demonstrates the numerical problem in its simplest form.

The integral

$$\int_0^1 \log(x)\,dx = -1 \tag{3}$$

has a logarithmic singularity at $x = 0$, but this is such a mild singularity that the integral itself gives a finite value. Indeed, since the antiderivative of the log is given as $\int\log(x)\,dx = x\log(x) - x$, evaluating at the left endpoint $x = 0$ gives a zero value for the antiderivative while at the right endpoint the values -1 is obtained. Thus the integral is finite.

However, our usual numerical methods don't "know" how to identify the singularity and integrate exactly. Instead, a numerical approximation to the integral is obtained by sampling the interval $[0, 1]$ at N subintervals of equal length $= 1/N$ and taking the average. This sum (called a Riemann sum) will give a better and better approximation to the integral as N gets larger and larger. Unfortunately, the convergence to an exact answer is quite slow.

To see this analytically, one can use Stirling's formula to approximate this average sum:

$$\int_0^1 \log(x)\,dx \approx \frac{1}{N}\sum_{n=1}^{N}\log(\frac{n}{N}) = \frac{1}{N}(\log N! - \log(N^N)) \quad (4)$$

$$\approx \frac{1}{N}[N\log(N) - N + 0.5\log(2\pi N) - N\log(N)], \quad (5)$$

which reduces to

$$-1 + 0.5\frac{\log(2\pi N)}{N}. \quad (6)$$

This shows the sum adds up to the exact answer -1, plus a small error term. From a direct calculation, the error $0.5\frac{\log(2\pi N)}{N}$ is equal to about 0.2 (or 20% relative error) for $N = 10$ and about 0.03 (3% relative error) for $N = 100$.

STABILITY FACTOR

One common approach to stabilizing the Hilbert transform is to add a small constant to the amplitude spectrum, in the form

$$\mathcal{H}(\log(|\hat{f}(\omega)| + \epsilon)), \quad (7)$$

where ϵ is an operator-selected small parameter fixed in the range of 10^{-6} times the max amplitude spectrum. It corresponds to a white noise floor many decibels below the main signal, but chosen so the log of zero never occurs. The aim is to stabilize the logarithmic singularity by removing the singularity altogether.

To see how this is an effective method, one computes the stabilized log integral as follows:

$$\int_0^1 \log(x+\epsilon)\,dx = \int_\epsilon^{1+\epsilon} \log(x)\,dx = -1+(1+\epsilon)\log(1+\epsilon)-\epsilon\log(\epsilon) \approx -1+\epsilon-6\epsilon, \quad (8)$$

where this last 6ϵ appears because $\epsilon\log(10^{-6}) \approx -6\epsilon$. Thus we see that the error term for the stabilized logarithmic integral is on the order of a small constant times ϵ.

Comparatively speaking, this ϵ error is much smaller than the Riemann sum approximation error calculated in the previous section. Thus, although the stability factor removes the singularity, there is still a numerical challenge in computing the integral via the summation method.

In the next section, we examine some typical errors to get an idea of how many terms (how large an N) we need in the sum for accuracy when a stability term is included.

NUMERICAL EXPERIMENTS

We begin by assuming there is a logarithmic singularity at some point δ near zero, include a stability factor ϵ, and integrate across the singularity. This corresponds to a more realistic case where the numerical algorithm does not know exactly where the singularity is, and a stability term is included to avoid infinities. Thus, we are interested in computing a numerical approximation to the exact integral

$$\int_{-1}^{1} \log(|x-\delta|+\epsilon)\,dx = -2 + O(\delta,\epsilon). \tag{9}$$

For these experiments, we choose the stability term ϵ on the order of 10^{-4} to 10^{-8} which is typical in seismic processing. The location $\delta \approx 10^{-2}$ is randomly selected to reflect the fact that in physical examples we don't know exactly where the singularity is.

We ran a number of simulations using the Riemann sum method with N samples to get a sense of the error introduced. Table 1 is a typical result that shows the results are much more sensitive to the value of N than the value of ϵ. Indeed, to get errors well below 1%, one needs to choose N around a thousand or so. Moving ϵ from 10^{-4} to 10^{-6} reduces the percentage error by up to about a half, but further decreases in ϵ don't seem to make much difference.

Table 2 demonstrates that even when the location of the singularity is known exactly ($\delta = 0$), it is the number of samples N that is most relevant in reducing the error.

Table 1. Percentage error for various parameter values ($\delta = 0.0127$).

	$\epsilon = 10^{-4}$	$\epsilon = 10^{-6}$	$\epsilon = 10^{-8}$
$N = 20$	8.2	8.2	8.2
$N = 200$	0.73	0.66	0.66
$N = 2000$	0.16	0.066	0.065

Table 2. Percentage error for various parameter values ($\delta = 0.0$).

	$\epsilon = 10^{-4}$	$\epsilon = 10^{-6}$	$\epsilon = 10^{-8}$
$N = 20$	8.6	8.6	8.6
$N = 200$	0.91	0.85	0.85
$N = 2000$	0.17	0.086	0.085

TWEAKING WIENER-KOLMOGOROV

The numerical tests above suggests that the logarithm singularity can be computed accurately by taking many samples in the sum approximating the integral. This suggests one can oversample the amplitude spectrum $\hat{f}(\omega)$ in the parameter ω to make a more stable computation in the transforms.

A simple way to oversample is to append a sequence of zeros to the sampled signal $f(n), n = 0, 1, 2, \ldots, m$ and take the FFT of the zero-padded signal to get an interpolation of the Fourier spectrum. This is used as input to the Hilbert transform to compute the appropriate phase, then the iFFT is applied to recover the minimum phase wavelet (with zeros appended). The numerical experiments in the last section suggest oversampling by a factor of 100 to 1000 is necessary to keep errors small. In our tests, we found that a factor of 128 was good enough without degrading the performance of the FFT and Hilbert transforms. Indeed, the algorithm relies on the FFT for most of the intense calculation, and thus is an order $N \log N$ computation. The oversampled computation is still easily done in less than 1/100th of a second on an average-powered laptop.

This modification is an oversampled version of the Wiener-Kolmogorov method.

TESTS OF THE OVERSAMPLED MINIMUM PHASE CALCULATION

We apply this oversampled Wiener-Kolmogorov method to some useful sampled waveforms to see how the performance changes.

First, we test the oversample algorithm on a Ricker wavelet, shown in Figure 1. The shape looks familiar and seems in line with what one typically expects. A sample of the computations from the double inverse method and Kolmogorov methods is shown in in Figure 2. It appears the oversampled method is performing better than double inverse, as it removes what appears to be an extraneous bump at $t = 0.1$. However, without an exact analytical solution, we can't say that this bump shouldn't be there. We do note the Kolmogorov method also removes the bump.

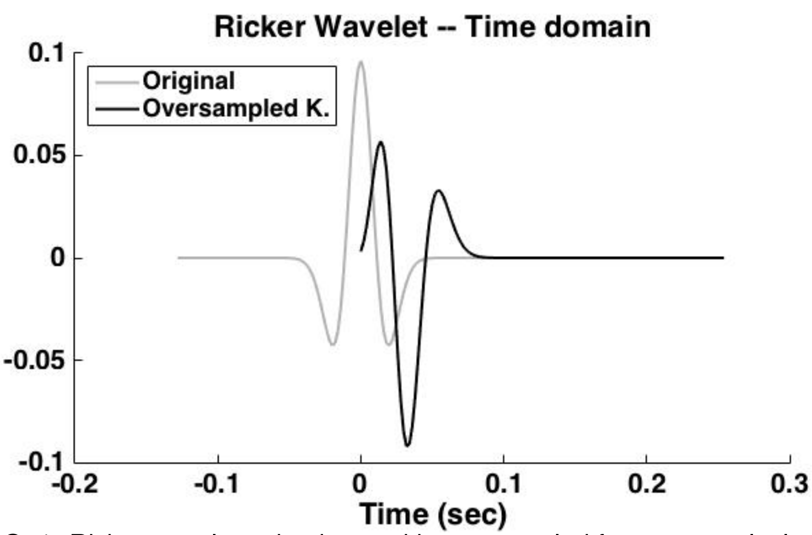

FIG. 1. Ricker wavelet, min phase with oversampled frequency calculation.

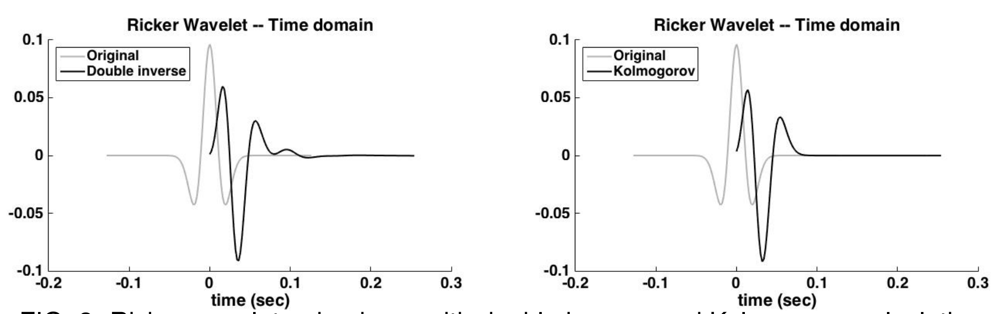

FIG. 2. Ricker wavelet, min phase with double inverse and Kolmogorov calculation.

A more careful direct comparison of the double inverse result with the oversampled frequency calculation shown in Figure 3. Again, the oversample method appears to be performing better – it moves the waveform slightly forward in time while removing the extraneous wiggle at time $t = 0.1$ in the double inverse result.

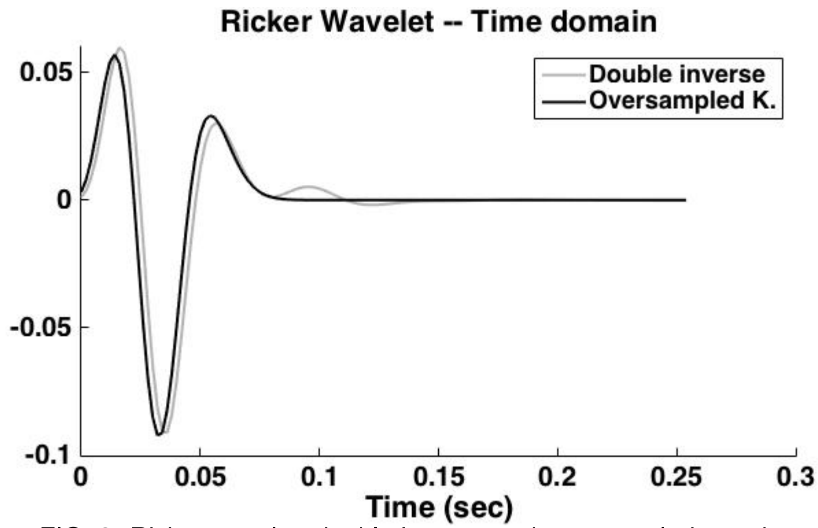

FIG. 3. Ricker wavelet, double inverse and oversampled together.

Similarly, a one-to-one comparison of the oversampled method to the Kolmogorov method, as shown in Figure 4, demonstrates the oversampled method is again slightly advanced. Notice in this case, the Kolmogorov method has no extra "wiggle" as in the double inverse result.

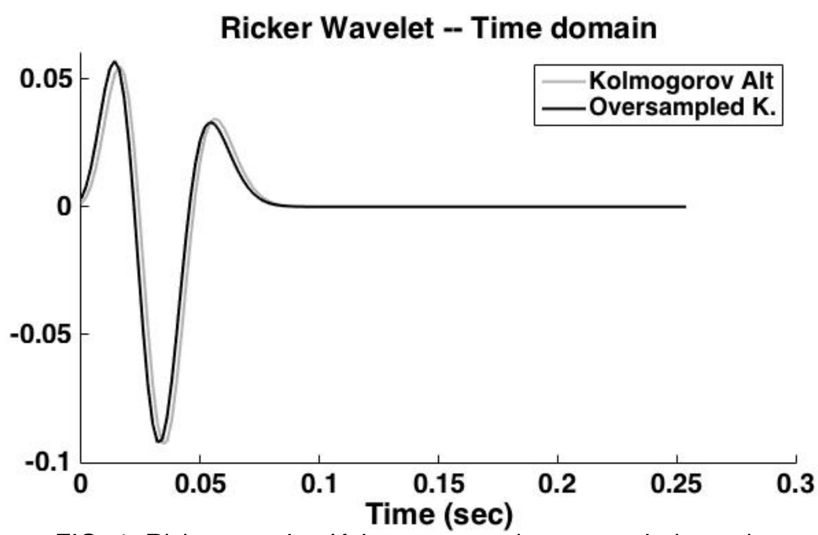

FIG. 4. Ricker wavelet, Kolmogorov and oversampled together.

To verify whether the oversampled method is indeed better, it is useful to compare the performance on exact waveforms where the mathematically precise minimum phase waveform is known. For instance, a boxcar function will be minimum phase once it has been shifted to time zero. The new oversampled methods shows a very sharp shifted boxcar in Figure 5. This gives us some confidence that we have properly stabilized this numerical calculation. We see in Figure 6 that the double inverse and Kolmogorov methods give somewhat choppy approximations to a shifted boxcar.

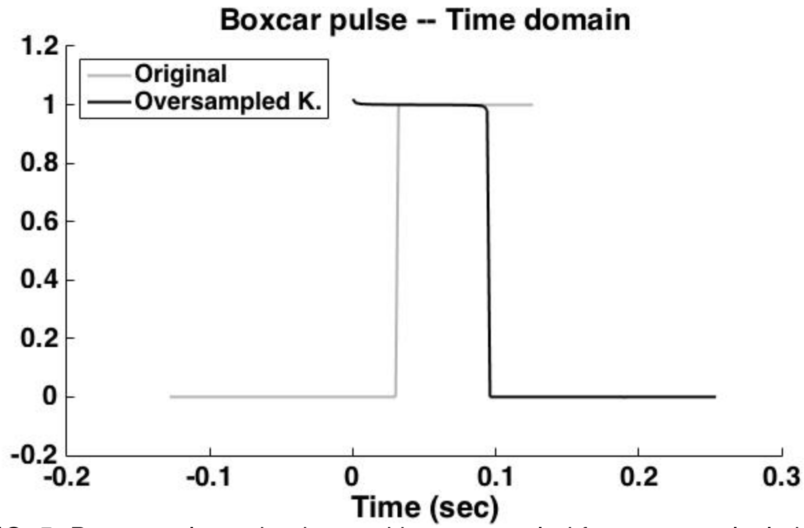

FIG. 5. Boxcar pulse, min phase with oversampled frequency calculation.

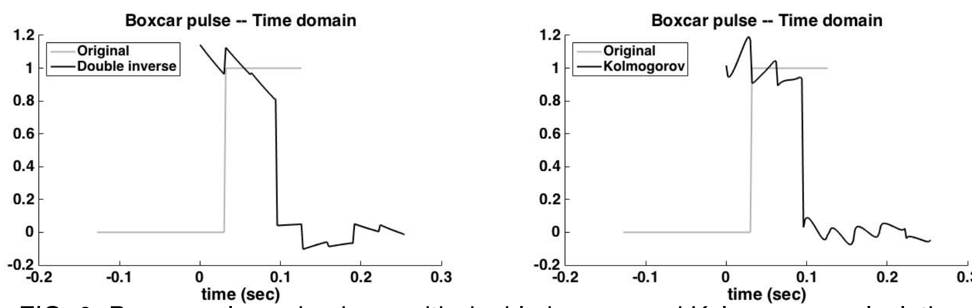

FIG. 6. Boxcar pulse, min phase with double inverse and Kolmogorov calculation.

Similarly, a rising exponential ramp also has an exact minimum phase version. It will be a decreasing exponential ramp, shifted to time $t = 0$. Again, we test on the various versions of the numerical algorithms. Figure 7 shows the result of the oversampled method which demonstrates a nearly perfect mirror image of the original exponential ramp. Again, this is reassuring evidence that the oversampling algorithm is performing in an accurate, stable manner. Figure 8 shows the results for the double inverse and Kolmogorov methods – both approximate a decreasing exponential, but with some asymmetry and artifacts near the jumps in the ramp. These are significantly different than the desired exact solution.

FUTURE WORK

We need to test this on real data and determine if there is any significant improvement in deconvolved seismic images. There is some concern that the new algorithm moves the minimum phase peak somewhat forward in time – which may give location results that are slightly different than what we are used to seeing in images. Another question is to determine how to reduce phase errors that occur due to the finite time and frequency windows occurring in practice that restrict how much spectral information we have about the real wavelet. We will include the oversampled method as a standard routine in the CREWES toolbox.

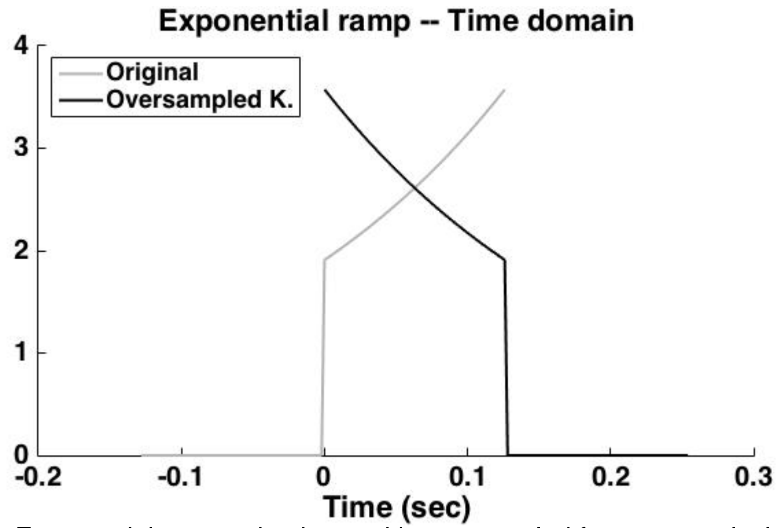

FIG. 7. Exponential ramp, min phase with oversampled frequency calculation.

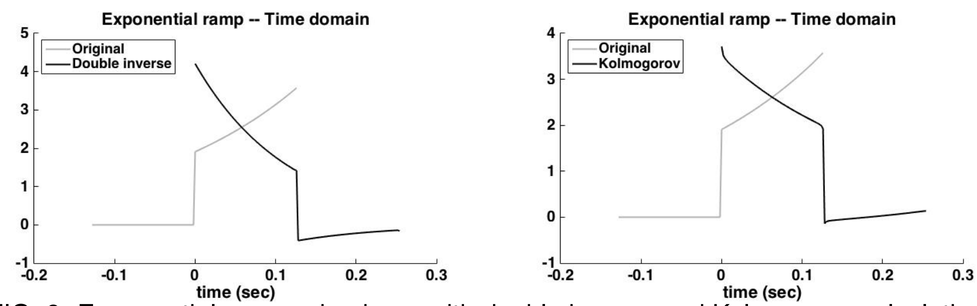

FIG. 8. Exponential ramp, min phase with double inverse and Kolmogorov calculation.

CONCLUSION

The minimum phase calculation is unstable in part because of possible logarithmic singularities in the log amplitude spectrum that arise in the integral transforms. Numerical calculations are made more accurate by oversampling in the frequency domain. Sample tests with Ricker wavelets, boxcars, and exponential ramps show improvements in the modified algorithm. More work is needed to verify that there is any improvement in imaging.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the financial support of the industrial sponsors of the CREWES project at the University of Calgary. This work was funded by CREWES industrial sponsors and NSERC (Natural Science and Engineering Research Council of Canada) through the grant CRDPJ 461179-13 and the authors' Discovery grants.

APPENDIX - CODE

Some sample code demonstrating the errors introduced in a numerical integration of a log singularity.

```
% numerical integration of a log singularity
```

```
% we integrate log|x| from -1 to 1.

for N=[20, 200, 2000]
    N  % number of step in Riemann sum
    for eps = [10^(-4), 10^(-6), 10^(-8)]
        eps  % stability factor
        del = 0*.0127; % a slight shift on the x axis

        x = linspace(-1,1,N);
        y = 2*sum(log(abs(x-del)+eps))/N;
    100*(y+2)/2  % percentage error
    end
end
```

REFERENCES

Claerbout, J., 1976, Fundamentals of seismic data processing: McGraw-Hill, N.Y.

Dey, A. K., and Lines, L. R., 1998, Seismic source wavelet estimation and the random reflectivity assumption: CREWES Research Report, **10**, 21.1–21.28.

Gibson, P. C., Lamoureux, M. P., and Margrave, G. F., 2011, Outer preserving linear operators: J. Funct. Anal., **261**, No. 9, 2656–2668.
URL http://dx.doi.org/10.1016/j.jfa.2011.07.003

Helson, H., 1983, Harmonic Analysis: Addison-Wesley.

Hoffman, K., 1982, Banach Spaces of Analytic Functions: Prentice Hall.

Lamoureux, M. P., and Margrave, G. F., 2007, An analytic approach to minimum phase signals: CREWES Research Report, **19**.

Lines, L. R., and Ulrych, T. J., 1977, The old and the new in seismic deconvolution and wavelet estimation: Geophysical Prospecting, **25**, 512–540.

Robinson, E. A., 1967, Predictive decomposition of time series with applications to seismic exploration: Geophysics 32, **32**, 418–484.

Russell, B., 1994, Seismic inversion: SEG course notes.

White, R. E., and O'Brien, P. N. E., 1974, Estimation of the primary seismic pulse: Geophysical Prospecting, **22**, 627–651.