Multiwavelet Seismic-Wave Gradiometry
by Christian Poppeliers

Abstract  A new technique to determine uncertainty estimates for wave parameters obtained by seismic-wave gradiometry is established using the multiwavelet transform. Wave gradiometry uses spatial gradients as measured by a small-scale seismic array to estimate the wave parameters of propagation azimuth, slowness, geometrical spreading, and radiation pattern. The advantage of wave gradiometry is that the wave parameters are estimated on a point-by-point basis for the entire seismogram. Previous work on wave gradiometry employed the continuous wavelet transform to decompose the wave field into narrow-band realizations prior to gradiometric analysis so that the wave parameters are resolved as functions of both time and frequency. To build on this approach, I incorporated the multiwavelet transform. The multiwavelet transform decomposes the wave field into a series of mutually orthogonal wavelet coefficients, which can then be analyzed by wave gradiometry. Each result can then be treated as a sample from a statistical population. The method is tested on synthetic data, with and without noise. A relatively low degree of uncorrelated noise can have a significant impact on the quality of the results. Finally, multiwavelet gradiometry is applied to a single earthquake recorded by a small subset of USArray stations and shows that the method accurately and robustly determines the wave parameters.

Introduction

Seismic-wave gradiometry (WG) is a method of seismic array data processing that can be used to determine wave-field attributes such as propagation azimuth, geometrical spreading, and radiation pattern. Being able to estimate these parameters can be helpful for estimating site response in seismic hazard analysis or instrument calibration. The method is based on using spatial gradients of displacement seismograms as recorded by small-scale, single- or multicomponent seismic arrays (Langston, 2007a, 2007b, 2007c; Liang and Langston, 2009). In a previous paper (Poppeliers, 2010), WG was modified to include the wavelet transform in order to decompose the wave field into bandlimited wavelet coefficients prior to wave gradiometric analysis to extract time-dependent wave attributes as a function of frequency.

Decomposing the wave field into bandlimited wavelet coefficients prior to WG provides a very rich picture of the frequency and time-dependent wave attributes. However, there is currently no method to determine the quality of the results. That is, there has been no discussion on the formal uncertainty estimates on the resolved wave attributes, all of which are functions of time and frequency. This paper attempts to mitigate this by introducing a method to estimate formal uncertainty estimates for wave attributes obtained by wave gradiometry. The method uses multiwavelets, which are orthogonal wavelets used to decompose the wave field into orthogonal, bandlimited wavelet coefficients. The coefficients can then be analyzed to yield statistically independent estimates of a given wave parameter (e.g. Lilly and Park, 1995; Bear and Pavlis, 1997, 1999; Bear et al., 1999).

Because wave propagation attributes such as back azimuth and slowness are highly frequency dependent, Poppeliers (2010) introduced the use of the continuous wavelet transform to decompose the seismograms into bandlimited versions of the original data prior to wave gradiometric analysis. The wavelet transform is a convenient strategy for time-frequency decomposition of time series (Daubechies, 1992; Mallat, 1999; Press et al., 2001; Addison, 2002). The work presented here builds on this approach by using the multiwavelet transform to decompose the wave field into a series of bandlimited orthogonal wavelet coefficients prior to WG analysis. The purpose of this approach is twofold. First, the use of multiwavelets allows the data to be decomposed into approximately orthogonal, bandlimited wavelet coefficients, which can then be used to estimate wave attributes. The point of the multiple estimates of the wave attributes at a given time–frequency location is that standard statistical tools can be employed to construct formal uncertainty estimates or confidence intervals. Second, by retaining the time–frequency decomposition of a wavelet transform, it is possible to retain the ability to analyze the data for all times and frequencies, which is important to reveal time- and frequency-dependent wave propagation attributes.
Methodology

Seismic-Wave Gradiometry

The basic method of WG builds on concepts introduced by Langston (2007a, 2007b, 2007c) and outlined in Poppeliers (2010). The fundamental assumption is that the propagating wave contains geometric spreading and a (potential) change in slowness as a function of distance:

\[ s(x, y, t) = G(x, y)f(t - xp_x - yp_y), \]  

(1)

where \( s(t, x, y) \) is the scalar displacement at location \( (x, y) \) with slowness \( (p_x, p_y) \), \( G(x, y) \) represents the amplitude variation across space, and \( f(t, x, y) \) is the phase variation as a function of time and location. Equation (1) is differentiated and solved as two one-dimensional problems to yield the wave parameters

\[ A_\zeta = [A_x, A_y] \]  

(2)

and

\[ B_\zeta = [B_x, B_y]. \]  

(3)

To economize notation, the two spatial dimensions are indicated by \( \zeta = [x, y] \). Thus the variable, for example, \( A_\zeta \) indicates the variables \( A_x \) and \( A_y \). In other words, in the equations that follow, the subscript \( \zeta \) can mean either \( x \) or \( y \), depending on the dimension considered.

The wave parameter \( A_\zeta \) is the normalized spatial gradients of the wave amplitude in either the \( x \) or \( y \) direction, and the parameter \( B_\zeta \) is proportional to the wave’s velocity across the array (Langston and Liang, 2008). Given a two-dimensional array of instruments, the spatial gradient of the displacement seisograms

\[ \frac{dS(\zeta, t)}{d\zeta} = \left[ \frac{\partial S(x, y, t)}{\partial x}, \frac{\partial S(x, y, t)}{\partial y} \right] \]  

(4)

is used to calculate the wave parameters as

\[ A_\zeta = \frac{|\frac{d}{d\zeta} S(\zeta, t)|}{|S(\zeta, t)|} \cos(\psi - \phi) \]

\[ - \frac{1}{\omega(t)} \frac{|\frac{d}{d\zeta} S(\zeta, t)|}{|S(\zeta, t)|} \frac{\partial |S(\zeta, t)|}{\partial t} \sin(\psi - \phi) \]  

(5)

and

\[ B_\zeta = \frac{1}{\omega(t)} \frac{|\frac{d}{d\zeta} S(\zeta, t)|}{|S(\zeta, t)|} \sin(\psi - \phi), \]  

(6)

where

\[ S(\zeta, t) = s(\zeta, t) - iH[s(\zeta, t)] \]  

(7)

is the analytic signal of the displacement seismogram at location \( [x, y] \), \( H[s(\zeta, t)] \) denotes the Hilbert transform of the seismogram \( s(\zeta, t) \).

\[ \omega(t) = \frac{1}{|S(\zeta, t)|^2} \left\{ \frac{d}{dt} s(\zeta, t)H[s(\zeta, t)] \right\}, \]  

(8)

and \( \psi \) and \( \phi \) are the instantaneous phases of the displacement-derivative analytic signal and the displacement analytic signal, respectively (Langston, 2007b).

The four wave parameters, propagation azimuth, geometrical spreading, slowness, and radiation pattern are obtained from the quantities \( A_\zeta \) and \( B_\zeta \), as will be shown later in this paper. However, it is important to realize that \( A_\zeta \) and \( B_\zeta \) are functions of time, which means the estimated wave parameters will also be functions of time.

Multiwavelets

The complex wavelet transform is an integral transform given by

\[ W[s(u, M)] = \int_{-\infty}^{\infty} s(t)\Psi_{M,0}(t)dt + i \int_{-\infty}^{\infty} s(t)\Psi_{M,0}(t)dt, \]  

(9)

where the integration kernels \( \Psi_{M,e} \) and \( \Psi_{M,o} \) are even and odd wavelet functions, respectively, with a scale length \( M \) and position \( u \) (Bear and Pavlis, 1997). For the purpose of this paper, the subscripts \( e \) and \( o \) denote even and odd functions. The map from scale to frequency depends on the wavelet used. Generally, a given wavelet at a given scale contains a range of frequencies with a well-defined center frequency and a rolloff that depends on the wavelet. A simple method of determining the scale-to-frequency map is to simply take the Fourier amplitude spectrum of a wavelet at different scales and then interpolate–extrapolate for other scales (Addison, 2002; Mallat, 1999).

One of the advantages to a wavelet transform is that it naturally optimizes the time-frequency tradeoff by using a scaled wavelet to analyze the (arbitrary-length) signal at different frequencies. Because of this attribute, wavelet transforms are much better suited to estimate the time-varying frequency content of a signal. This is in direct contrast to the commonly used spectrogram approach, in which Fourier transforms are applied to a signal that is windowed into equal-length time segments without any regard given to time–frequency resolution. Another advantage to using wavelet transforms is that the actual form of the analyzing wavelet can be arbitrary. For the work here, I capitalize on this flexibility and use a wavelet type developed by Slepian (1983) and Thompson (1982).

In principle, any orthogonal wavelet family could be used for the multiwavelet method described here. However, I chose to use Slepian wavelets because this wavelet family has been shown to work in previous seismic applications (e.g., Lilly and Park, 1995; Bear and Pavlis, 1997, 1999; Bear et al., 1999). Slepian wavelets are constructed as a family of orthogonal, even–odd wavelet pairs. For a given
scale, a suite of orthogonal wavelet pairs are constructed for which each wavelet pair has the same center frequency and bandwidth as all the other wavelet pairs (e.g., fig. 1 of Bear and Pavlis, 1997). Although the bandwidth and center frequency of all the wavelet pairs are similar, each wavelet pair emphasizes a different portion of the amplitude spectrum within the given bandwidth. These functions have been applied to the analysis of signals and are referred to as multitaper or multiwindow methods (e.g., Park et al. 1987). Lilly and Park (1995) generalized these methods by developing a series of real-valued, discrete time series \((w_m)\) with \(M\) samples and sampling rate \(\Delta t\). The wavelets are designed to concentrate energy within a frequency band of interest defined by center frequency \(f_c\) and bandwidth \(2f_w\), where \(f_w \leq f_c\). The center frequency of the wavelets are controlled by the parameter \(p\), where
\[
p = f_w M \Delta t,
\]
\(f_w\) is the bandwidth, \(M\) is the wavelet length, and \(\Delta t\) is the wavelet's sample interval. The bandwidth of the wavelet is controlled by the parameter \(p_c\), where
\[
p_c = f_c M \Delta t.
\]
The terms \(p\) and \(p_c\) are referred to as the time–bandwidth product and the time–bandcenter product (Lilly and Park, 1995). In practice, the parameters \(p\) and \(p_c\) are held fixed, and the wavelets' center frequency is varied by changing the time length \(M\) of the wavelet. For the work here, the values of \(p\) and \(p_c\) are held fixed throughout the analysis at \(p = 2.5\) and \(p_c = 3.0\).

The wavelets developed by Slepian and generalized for seismological applications by Lilly and Park are real valued, thus energy in the frequency domain appears in both the positive and negative frequencies. Therefore any frequency band of interest is defined by \(|f \pm f_c| \leq f_w\). The fraction of the total energy contained within this frequency interval is
\[
\lambda = \frac{\int_{-(f_c+f_w)}^{+(f_c+f_w)} |W(f)|^2 df - \int_{-(f_c-f_w)}^{+(f_c-f_w)} |W(f)|^2 df}{\int_{-(1/2\Delta t)}^{+(1/2\Delta t)} |W(f)|^2 df},
\]
where \(W(f)\) is the discrete Fourier transform of the wavelet \(w_m\) of length \(M\),
\[
W(f) = \Delta t \sum_{m=-P}^{R} w_m e^{-i2\pi fm\Delta t},
\]
\(P\) is the closest integer \(\geq M/2\), and \(R\) is the closest integer \(\leq M/2\).

Wavelets can be calculated by recasting equation (13) as an eigenvalue problem of the form
\[
Aw = \lambda w,
\]
where
\[
A_{mn} = \frac{\sin[2\pi (f_c + f_w) \Delta T(m-n)]}{\pi (m-n)} - \frac{\sin[2\pi (f_c - f_w) \Delta T(m-n)]}{\pi (m-n)}
\]
and solving for \(w\) (Lilly and Park, 1995; Bear and Pavlis, 1997). The eigenvalue problem has \(N\) orthogonal solutions of eigenvectors \(w^{[k]}\) and associated eigenvalues \(\lambda_k\). The wavelets are formed by sorting the eigenvalues (and their associated eigenvectors) from greatest magnitude to the least and then normalizing the wavelets such that
\[
\sum_{m=1}^{M} (w^{[k]}_m)^2 = 1.
\]

Note that each wavelet is normalized with respect to itself, and thus all the \(k\) wavelets have roughly the same amplitude. Because the wavelets occur as even–odd pairs, each wavelet within a pair is \(\pi/2\) radians out of phase. Thus, they can be combined into a complex function of the form
\[
w^{[k]} = w^{[k]}_{f,e} + iw^{[k]}_{f,o},
\]
where \(f\) is the center frequency, \(w^{[k]}_{f,e}\) is the \(k\)-th even wavelet, and \(w^{[k]}_{f,o}\) is the \(k\)-th odd wavelet. Also, each set of wavelet pairs is orthogonal to all other wavelet pairs. The complex function \(w^{[k]}\) can be used as the integration kernel in the multiwavelet transform
\[
W^{[k]}[s(t,f)] = \int_{t-T/2}^{t+T/2} s(\xi) w^{[k]}_{f,e}(\xi - \tau) d\xi + i \int_{t-T/2}^{t+T/2} s(\xi) w^{[k]}_{f,o}(\xi - \tau) d\xi.
\]
An appealing aspect of the multiwavelet transform is that using the \(k\) wavelet pairs described in equation (17) yields \(k\) wavelet transforms of the data. Because of the orthogonality of the \(k\) wavelet pairs, it can be argued that the \(k\) wavelet transforms of the data are statistically independent. Indeed, Bear and Pavlis (1997) showed that each complex wavelet transform yields a time series that is statistically independent when applied to Gaussian noise. They used this property to show that the same statistical independence was approximately valid when applied to actual seismic data. Thus, for a population of measurements corresponding to \(k = 1, 2, 3, \ldots n\) wavelets \(w^{[k]}\), standard statistical methods can be employed to obtain estimates of uncertainty. One important aspect of this analysis is that there is a practical limit to the number of wavelet pairs that can be formed. Generally, limiting the number of wavelet pairs to eight offers a good trade-off between having an adequate number of samples for statistical analysis and avoiding numerical instabilities associated with higher-order wavelet pairs.

Wave Gradiometry Using the Multiwavelet Transform

Multiwavelet gradiometry is similar to the wavelet gradiometry introduced in Poppelliers (2010). The primary
difference is that, rather than using a conventional wavelet transform prior to gradiometric analysis, the multiwavelet transform is used. Specifically, I start with equation (18) and form \( k \) estimates of even pairs and odd pairs of wavelet transformed data for each seismic station. However, rather than use the analytic form of the transformed signal (Langston 2007b; Poppeliers, 2010), I take the same approach taken by Bear and Pavlis (1997) and form an analogous analytic signal as

\[
W_{i,t}^{[k]}[S(t)] = w_{i,t}^{[k]}[s(t)] - iw_{i,t}^{[k],o}[s(t)],
\]

(19)

where the variables \( f_i \) denote the center frequency of a given wavelet, \( i \) is the \( i \)-th seismic station, \( k \) is the \( k \)-th wavelet pair, and the \( W[.] \) notation indicates the multiwavelet transform of signal \( s(t) \). The definition given in equation (19) is substituted into equations (5) and (6) to obtain

\[
A_{i,t;f}^{[k]} = \left( \frac{d}{df} W_{i,f}^{[k]}[S(t)] \right) \cos(\psi - \phi)
\]

\[
- \left( \frac{1}{\omega(t)} \frac{d}{df} W_{i,f}^{[k]}[S(t)] \frac{\partial}{\partial t} W_{i,f}^{[k]}[S(t)] \right) \sin(\psi - \phi)
\]

(20)

and

\[
B_{i,t;f}^{[k]} = \left[ \frac{1}{\omega(t)} \frac{d}{df} W_{i,f}^{[k]}[S(t)] \right] \sin(\psi - \phi).
\]

(21)

The gradiometry parameters \( A_{i,t;f}^{[k]} \) and \( B_{i,t;f}^{[k]} \) are now localized in both time and frequency, whereas equations (5) and (6) were only localized in time. The parameters \( A_{i,t;f}^{[k]} \) and \( B_{i,t;f}^{[k]} \) are used to obtain the wave parameters for each wavelet pair. Because \( p_x = -B_y \), it follows that \(-B_x = p_x \) and \(-B_y = p_y \). Therefore the back azimuth is given by

\[
\theta_{b,t;f}^{[k]} = \tan^{-1} \left( \frac{\text{Re}(B_{i,t;f}^{[k]})}{\text{Re}(A_{i,t;f}^{[k]})} \right).
\]

(22)

and the magnitude slowness is

\[
p_{b,t;f}^{[k]} = \sqrt{(p_{x,t;f}^{[k]})^2 + (p_{y,t;f}^{[k]})^2}.
\]

(23)

Similarly, the back-azimuth-dependent geometrical spreading is given by

\[
A_{i,t;f}^{[k]} = \text{Re}[A_{i,t;f}^{[k]} \sin(\theta) - A_{i,t;f}^{[k]} \cos(\theta)],
\]

(24)

where the \( \theta \) term is given by equation (22).

**Statistical Considerations**

The primary advantage of using multiwavelet gradiometry is that it enables us to produce multiple estimates of the wave attributes with equations (22)–(24), with one estimate for each complex wavelet transform (as a function of a given frequency and time point). For example, for eight wavelet pairs, there are eight estimates of a given wave attribute. Each of these estimates can be thought of as a sample drawn from a population with an unknown distribution. The goal of this paper is to find a parameter \( \Theta \) that defines the center, or mean, of the population, and then determine the distribution about that mean. A straightforward approach would be to compute the average of each estimate and determine the standard error. However, because the sample size is limited and it is not certain that the samples are normally distributed, I propose using the jackknife statistical method, which does not require \textit{a priori} assumptions about a population to determine its variance. The jackknife statistics, as well as uncertainties, are estimated using the approach outlined by Thompson and Chave (1991) and Efron and Stein (1981), as applied by Bear and Pavlis (1997).

Jackknife statistics are useful for determining the center and variance of a population in a one-dimensional sense. As such, jackknife statistics are well suited to estimate wave parameters such as geometrical spreading or slowness in the \( x \) or \( y \) directions. However, slowness magnitude is determined by a combination of two one-dimensional quantities: \( p_{x,t;f}^{[k]} \) and \( p_{y,t;f}^{[k]} \). This is also true for determining wave azimuth in the case of nearly horizontally propagating waves (e.g., surface waves). A straightforward approach to estimate slowness magnitude and azimuth would be to simply use the mean values of \( p_{x,t;f}^{[k]} \) and \( p_{y,t;f}^{[k]} \), as well as their corresponding confidence intervals, with equation (22). However, the estimates of these two parameters are further refined by using robust \( M \) estimators to protect the results from being severely biased by outlying values.

For waves that are propagating nearly horizontally (e.g., surface waves), it is appropriate to determine wave azimuth using the parameters \( p_{x,t;f}^{[k]} \) and \( p_{y,t;f}^{[k]} \). Ideally, for a given time–frequency location, all the estimates of \( p_{x,t;f}^{[k]} \) and \( p_{y,t;f}^{[k]} \) will be identical for all wavelet pairs. However, for actual data, there is slight variation for the different estimates of \( p \) for the different wavelet pairs. Regardless, plotting the values \( p_{x,t;f}^{[k]} \) and \( p_{y,t;f}^{[k]} \) onto the \( p_xp_y \) plane yields a cluster of points, the center of which should correspond to the mean wave azimuth for a given time–frequency point. Each of these \( p_xp_y \) pairs corresponds to an estimate of the wave azimuth for a given wavelet pair. I produce a robust estimate of the azimuth by fitting a line through the points where \( M \) estimators iteratively downweight outlying points (e.g., Press et al., 2001). The wave azimuth \( \hat{\theta}_{t;f}^{[k]} \) is determined from the robust least-squares line fit illustrated in Figure 1.

To estimate the confidence intervals for the propagation azimuth, the data are rotated such that the robust least-squares line lies on the \( p_x \) axis. Then jackknife statistics are used to estimate the width of the confidence interval \( \delta \theta_{t;f} \), using

\[
\hat{\theta}_{t;f} \pm s_{0.975},
\]

(25)

where \( s_{0.975} \) is the 0.975 quartile of the Student \( t \) distribution with \( J - 1 \) degrees of freedom and \( s \) is the variance (see equations 21–23 in Bear and Pavlis, 1997).
Note that the actual azimuth is determined using a robust estimator, which minimizes the bias of outliers. However, the confidence intervals are determined by all the data points, giving full weight to the outliers. This method is a compromise between minimizing outliers to determine the center of the distribution (here, the azimuth), but it then fully accounts for all the data when reporting on the confidence of the estimate.

The magnitude slowness is obtained by a similar multi-step process: after, the \( p_{x,k,tf} \) and \( p_{y,k,tf} \) points are rotated onto the \( p_x \) axis, the center of the distribution in the \( p_x \) direction is determined by the jackknife method, and its distance from the origin to the center of the distribution is defined as the magnitude slowness (Fig. 1). The uncertainty in slowness is determined by the using 1-D jackknife statistics on all of the \( p_{x,k,tf} \) and \( p_{y,k,tf} \) projected onto the \( p_x \) axis.

Application

For the work described here, spatial gradients of the wave field are calculated using the least-squares formulation outlined in Liang and Langston (2009). This method uses a weighting term to reduce the effects of truncation errors in the Taylor series expansion and is based on array geometry and presumed wave slowness. However, I do not include a weighting term as the truncation error is minimal as long as the phase shift of the wave field is less than 1/4 of a wavelength across the array. This condition is met either by the geometry of the array or by applying a moveout correction to the data prior to analysis (Langston 2007c). The procedure for applying a moveout correction is summarized in a later section of this paper (Analysis Method).

Synthetic Example

Two synthetic tests demonstrate the efficacy of the methods proposed here. In both cases, the synthetic two-dimensional signal has the form

\[
s_i(r, t) = \frac{1}{|r_j - r_0|} \exp\left[-\alpha(t - x_i p_x - y_i p_y)^2\right],
\]

where the vector \( r_i = (x_i, y_i) \) represents the \( i \)-th seismic station’s array coordinates, \( |r_j - r_0| \) is the distance from the master station located at \( r_0 \) to the station located at \( r_j \), \( \alpha \) is a constant that controls the period of the wave, and \( p_x \) and \( p_y \) are the horizontal slownesses in the \( x \) and \( y \) directions, respectively. This function simulates a pure plane wave traveling horizontally over an infinite half-space with constant velocity. The synthetic array is shown in Figure 2, and the data take the form shown in Figures 3a and 4a, in which two waves of differing central frequency traverse the gradiometer with differing slownesses. The two waves consist of (1) a wave with a central period of approximately 0.5 s, traversing the gradiometer with a back azimuth of 45° and arriving at ~5.0 s, and (2) a wave with a central period of approximately 0.25 s, traversing the gradiometer with a back azimuth of 225° and arriving at ~6.2 s. Note that this test is consistent with surface wave modes propagating in unconsolidated material. Both waves have a magnitude velocity of 700 m/sec. Based on the wave parameters and the geometry of the array, the smallest wave will have a wavelength greater than ten times the aperture of the array. This ensures that the phase shift between all of the array stations will be less than 1/4 of a wave cycle, which minimizes the truncation errors associated with estimating the spatial gradients.

In the first test, the synthetic data contain no noise. In the second case, noise was added to the data. The noise field was generated by filtering a vector of Gaussian-distributed random numbers to a frequency range of 1–10 Hz. The noise field was scaled to an amplitude of approximately 5% of the root mean square (rms) amplitude of the actual signal. Note that a unique noise signal is generated for each station, rendering the noise as uncorrelated between stations. For the test, multiwavelet gradiometry is used to estimate the back

![Figure 1. Robust estimation of back azimuth. (a) A robust least-squares line is fit through the \( k \)-estimates of the \((p_x, p_y)\) pairs. The average back azimuth \( \theta_{k,tf} \) is determined by the slope of the robust least-squares line. (b) The \((p_x, p_y)\) pairs are rotated such that the least-squares line lies on the positive \( p_x \) axis. The center of the distribution along the \( p_x \) direction is determined using jackknife statistics, and the center of the distribution in the \( p_y \) direction is forced to be at \( p_y = 0 \). The 95% confidence intervals are computed using jackknife statistics, and the 95% confidence interval \( \delta \theta_{k,tf} \) of the back azimuth is determined by the geometry shown on the figure. The magnitude slowness is determined by the distance to the center of the distribution.](image)

![Figure 2. The geometry of the synthetic gradiometer. (Star, master station; circles, supporting stations.)](image)
Figures 3 and 4 show the recovered wave attributes as a function of time and period. In Figures 3b and 4b, the wave attributes are displayed as constant-frequency slices and show the wave attributes with their associated 95% confidence intervals as a function of time only. Also shown are wave attributes as estimated by conventional wave gradiometry, which do not contain any estimates of uncertainty.

For the noise-free test, the wave attributes are recovered to a fairly high degree of accuracy, with only a small degree of uncertainty. Interestingly, there appears to be a higher degree of uncertainty in the recovered wave parameters for the second-arriving wave. This is likely due to two reasons. First, there is some degree of wave interference, especially at the higher periods where the filtering effects of the wavelet transform will cause the waves to overlap in time (which actually violates the assumption that one a single plane wave is traversing the gradiometer). Second, truncation errors associated with estimating the spatial gradients are more pronounced for the later arriving, higher-frequency wave. Specifically, smaller wavelengths will be more susceptible to truncation errors when estimating their spatial gradients, even if they have wavelengths greater than ten times the array aperture. Interestingly, the wave attributes estimated by conventional wave gradiometry appear less variable than those estimated by multiwavelet wave gradiometry. The reason for this is not clear.

For the test on the noisy data, the wave attributes are generally recovered, but the degree of scatter, as well as uncertainty, is much higher. For time intervals in which there is a relatively high noise amplitude relative to the actual signal (e.g., \( t < 5.0 \) s), the accuracy of the results is rather poor. However, for times greater than 5.0 s, the amplitude of the signal is much higher relative to the noise, and the accuracy of the results improves. This can be explained by the nature of the noise for this test. Specifically, the noise is incoherent from station to station. For times less than 5.0 s, the relative rms amplitude of the noise is much higher than that of the signal, so the gradiometric analysis is determining the wave attributes for the noise itself. In contrast, for later times, the relative amplitude of the signal is much higher than the noise so the gradiometric analysis essentially detects more signal than noise. This is an important factor in any signal processing procedure; however, the analysis itself cannot...
discriminate between signal and noise, so the results will necessarily be a combination of the two. As with the noise-free test, the wave parameters as estimated by conventional wave gradiometry appear to have less variability.

Application to Field Data

To demonstrate the method proposed in this paper, I apply the analysis on the $M_{8.8}$ Chilean earthquake recorded by a subset of stations deployed as part of the EarthScope USArray (Fig. 5). The $M_{8.8}$ earthquake was located at latitude $-36.12\degree$, longitude $-72.9\degree$ and occurred on 7 February 2010 ($\Delta = 81.1\degree$). A single gradiometer was created by selecting 11 stations from the USArray, as deployed on this date. The analysis was conducted on only the vertical-component surface waves to demonstrate proof of concept.

Analysis Method

It has been shown that gradiometric analysis works best when the aperture of the gradiometer is less than 10% of the smallest wave being analyzed. This is to avoid higher-order
truncation errors in the Taylor series expansion used to construct the least-squares method of estimating the spatial gradients. Given that a 40-s surface wave (velocity $\sim 3.6$ km/s) will have a wavelength of $\sim 160$ km, my gradiometer is much too large to accurately estimate the spatial gradients for analysis. However, Langston (2007b) proposed a reducing velocity method that effectively increases the phase velocity of the waves as they traverse the gradiometer. The method assumes that the medium is homogeneous within the aperture of the gradiometer and supposes that the average phase velocity $\bar{c}$ is known. I remove the phase shifts due to this average velocity by replacing $p_x$ and $p_y$ with $\bar{p}_x$ and $\bar{p}_y$, with

$$p'_x = p_x - \bar{p}_x \tag{27}$$

and

$$p'_y = p_y - \bar{p}_y \tag{28}$$

where

$$\bar{p}_x = \frac{\sin(\theta)}{\bar{c}} \tag{29}$$

and

$$\bar{p}_y = \frac{\cos(\theta)}{\bar{c}}. \tag{30}$$

This is equivalent to increasing the apparent velocity and thus lowers the truncation errors associated with estimating the spatial gradients. After gradiometric analysis, the average slowness $\bar{p}_x$ and $\bar{p}_y$ have to be added back to the slowness obtained from equation (21) (which gives $B_x = -p'_x$ and $B_y = -p'_y$). This method is implemented by shifting the waveform of the supporting station $s_i$ by $t_i = -\bar{p}_x \delta x_i - \bar{p}_y \delta y_i$ in the time domain, where $\delta x_i$ and $\delta y_i$ are the $x$ and $y$ distance of supporting station $s_i$ from the master station. Using the reducing velocity method (henceforth termed “moveout correction”), the phase moveout between the supporting stations and master station are decreased, and thus the errors in estimating the spatial gradients are reduced as well (Liang and Langston, 2009).

For the work in this paper, I focus on the analysis of the surface waves in the 30–50-s passband. There are several reasons for focusing this analysis on the surface waves in this case. First, surface waves are the lowest frequencies in the seismograms, making this phase least susceptible both to truncation errors in estimating the spatial gradients and to static errors associated with topography. Second, visualizing the results of the analysis is natural with surface waves because they travel horizontally, and thus the results can be presented appropriately in polar coordinates. Third, the size of this earthquake means that the source-time function is minutes in duration, so shorter-period body waves are likely to be highly complex with interfering arrivals. Finally, by focusing the analysis on the surface waves, only one
moveout correction needs to be applied over the analysis window. To determine the moveout correction, I used conventional beam-forming analysis on a window around the surface waves in the 30–50-s frequency band to obtain an average vector slowness of $p_x = 0.17$ s/km and $p_y = -0.21$ s/km (Fig. 6).

### Results

The results are presented in two formats. First, the mean value of three wave parameters (back azimuth, magnitude slowness, and geometrical spreading) are plotted as a function of time and period (Fig. 7). The results are presented in this format in order to visualize the time–frequency dependence of the wave attributes. In the second presentation of the results (Fig. 8), the three wave attributes are shown as a function of time for three specific frequency slices in order to show the uncertainties in the wave-attribute estimates.

The estimated back azimuth is the most stable of the estimated wave attributes in terms of variability. For the three frequency bands shown, the uncertainty is relatively small ($\pm 2–3^\circ$) throughout the entire time window. The great circle back azimuth for this event is $153.7^\circ$, whereas the back azimuth determined by the analysis ranges from $140^\circ$ to $150^\circ$. Up to five degrees of deviation of the measured back azimuth from the great circle back azimuth was observed by Liang.

**Figure 4.** Continued.
and Langston (2009) using a similar analysis. However, their analysis differed in two important ways: (1) they analyzed a much lower range of frequencies (60–150 s), and (2) their results were obtained by averaging the results over a single cycle of the highest-amplitude surface wave. In the case here, averaging the estimated back azimuth over a similar portion of the surface wave yields an estimated back azimuth of \( \sim 149^\circ \), which is approximately five degrees different from the true great circle back azimuth. This deviation is of the same magnitude as observed by Liang and Langston (2009) and is likely an actual wave propagation effect associated with geologic structure.

The time-dependent variability of the estimated back azimuth roughly correlates to the arrival of the highest-amplitude surface waves but is biased by surface-wave dispersion. Specifically, the estimated back azimuth for the 50-s surface waves shows a subtle increase at \( \sim 1800 \) s, whereas the increase in the estimated back azimuth occurs later for the higher-frequency surface waves (\( \sim 1850 \) s for the 40-s surface waves and \( \sim 1880 \) s for the 30-s surface waves). At about \( \sim 2100 \) s, the estimated back azimuth decreases, which corresponds to a notable decrease in the amplitude of the surface waves. However, this behavior of the time-dependent back azimuth decrease is frequency dependent: it is quite abrupt for the 50-s result, whereas the decrease is more gradual for the 30- and 40-s results (Fig. 8), which is also likely due to the dispersive nature of surface waves.

The estimated magnitude slowness displays quite a bit more variability than back azimuth (Fig. 8). For times prior to the Rayleigh wave arrival (\( t < \sim 1850 \) s), the magnitude slowness displays the most variability with the greatest degree of uncertainty. This portion of the seismogram is mostly miscellaneous body-wave phases and coda, all of which are interfering. Interfering phases are known to have deleterious effects on wave parameters estimated by wave gradiometry (e.g., Poppeliers, 2010). However, at roughly the onset time of the Rayleigh wave, the estimated magnitude slowness stabilizes to a value of \( \sim 0.28 \) s/km (\( \sim 3.6 \) km/s) with a very low degree of uncertainty. This behavior is seen for the entire range of frequencies analyzed here. As an independent test of value of the Rayleigh-wave magnitude slowness, data from a subset of stations that lie within \( \pm 2^\circ \) of the great circle back azimuth were collected (Fig. 9). The moveout of the Rayleigh wave on a plot of these data shows a magnitude slowness of \( 0.28 \) s/km, which agrees with the magnitude slowness as estimated by multiwavelet wave gradiometry.

An interesting effect on the analysis is a periodicity in the estimates of the back azimuth and slowness. Specifically, the estimated wave parameters display a pronounced cyclic variation that does not appear to be frequency dependent. For both the estimated back azimuth and slowness, the periodicity appears for the full length of the analysis window and is roughly constant in wavelength for all frequencies, differing only in amplitude. The periodicity in the estimated slowness shows a pronounced decrease in amplitude as the amplitude of the signal increases, corresponding to the arrival of the Rayleigh wave (Fig. 8). This suggests that the periodicity of the estimated slowness is somehow related to the signal-to-noise ratio of the signal. Conversely, the periodicity of the estimated back azimuth does not appear to coincide with the arrival of the Rayleigh wave nor does it appear to relate to the amplitude of the signal. The reason for this periodicity is not entirely clear.

In the preceding description of the results, the estimates of the geometrical spreading are generally quite small. This is expected at the teleseismic distances used here, as the surface waves are essentially planar at teleseismic distances. However, sudden jumps in the estimate of the magnitude of the geometrical spreading often occur when various waves...
interfere. Thus for the analysis here, the expected result is obtained: the geometrical spreading is, on average, close to zero, with sudden positive and negative jumps that likely correspond to frequency-dependent interference effects.

**Discussion and Conclusions**

The primary purpose of this paper was to present a method for computing formal uncertainty estimates, or confidence intervals, for wave attributes as estimated by wave gradiometry. The method builds on previously developed methods of wavelet gradiometry and the multiwavelet transform as applied to seismic beamforming. Formalizing a method to estimate confidence intervals was a necessary step in the further development of wave gradiometry, as the estimation of confidence intervals for recovered wave attributes can strengthen a given interpretation of wave type and origin. The strength of WG is that wave parameters are estimated on a point-by-point basis for the entire analysis window. However, there is some degree of smoothing of the data by the multiwavelet transform. This leads to a potential pitfall of the method here in that the smoothing of the wave field may cause individual waves to interfere in time and space. Careful windowing of the seismogram prior to the application of the multiwavelet transform may help to mitigate this; however, there will always be the potential of interfering waves.

Slepian wavelets are used in the multiwavelet transform primarily because the resolution of the passband is, to some degree, controllable by adjusting the time-bandcenter and time-bandwidth products. However, any orthogonal wavelet family could be used; the only condition is that one be able to form a series of orthogonal complex wavelets with similar center frequencies. Nominally, the estimated wave parameters and the associated uncertainty estimates would be similar regardless of the wavelet family used. A valuable future direction of the work presented here would be to test this assertion. That is, are the estimated wave parameters and the associated uncertainties similar for different wavelet families?

The method is broadband in the sense that wave attributes are estimated for all the (unaliased) frequencies recorded by the gradiometer. Specifically, the method uses the multiwavelet transform to decompose the wave field into
a range of narrowband realizations of the original data prior to wave gradiometric analysis and presents the results as a function of time and frequency. The range of analysis frequencies can be quite wide; however, it is usually limited by the array geometry and the frequency of the targeted wave types. Regardless, the result is that the analysis presents a complete picture of the time–frequency evolution of a given wave attribute; and, because the decomposition is done multiple times with mutually orthogonal wavelets, the confidence intervals can also be presented as a function of time and frequency.

When using an array as a gradiometer, it is possible that the array’s aperture is larger than 10% of the wave being analyzed. In this case, significant truncation errors associated
with estimating the spatial gradients will occur. However, truncation errors can be mitigated by the application of a moveout correction. Regardless of the aperture of the array, the data must be highly coherent across the array. To test the degree of dependence on wave coherence, I performed a test on synthetic data in which incoherent noise was added to the signal (Fig. 4). Even with a small degree of incoherent noise, the degree of uncertainty of the estimated wave parameters increased markedly. This effect varied for differing ratios of signal to noise, however noise significantly increased the scatter in the results. For actual array deployments, sources of this type of incoherent noise could arise from wind or electronic noise, but these are not viewed as a major source of noise. Rather, seismic noise such as microseism is more often apparent on recordings with a higher amplitude than random, uncorrelated noise. If microseism were the dominant source of noise in an actual gradiometer deployment, then it would be highly coherent across the gradiometer, and the analysis would resolve the wave attributes of the microseismic noise, provided that the microseismic noise was originating from a single source region.

To test multiwavelet gradiometry, I applied the method to the vertical component seismograms of the February 2010 Chilean earthquake, as recorded on a subset of USArray stations. The goal of the test was to determine whether the method would yield reasonable estimates of wave attributes for this earthquake source. Although portions of the USArray do not make an ideal gradiometer, this test was performed in order to demonstrate that even in nonideal deployments, multiwavelet gradiometry can be used to analyze various targeted wave phases. In the case shown here, the analysis focused on the Rayleigh wave, although this was done mainly for brevity. That is, the analysis could just as easily be applied to different phases, so long as the appropriate moveout correction is applied to the data. However, in the ideal case, a gradiometer would be constructed with a small enough aperture so the entire wave field can be analyzed without the need for a moveout correction.

The work in this paper applies the multiwavelet transform to wave gradiometry for only a single component of the displacement seismograms. However, there is no reason to restrict multiwavelet gradiometry to a single component: following the arguments of Langston and Liang (2008), multiwavelet transforms can be applied to the horizontal component seismograms to obtain

\[
\tan(\theta_k) = \frac{W_k^i[\frac{d}{dt} s_x(t)]}{W_k^i[\frac{d}{dy} s_x(t)]} = \frac{W_k^i[\frac{d}{dt} s_y(t)]}{W_k^i[\frac{d}{dy} s_y(t)]}, \tag{31}
\]

where \(s_x(t)\) and \(s_y(t)\) are the \(x\) and \(y\) component seismograms, respectively.

Data and Resources

All data used in this paper were collected from the Incorporated Research Institutions for Seismology (IRIS) Data Management System (http://www.iris.edu/data/passcal/exp.htm; last accessed February 2010). Data from the IRIS Data Management System are freely available to the public. All figures and maps are made using MATLAB software, available from MathWorks (www.mathworks.com). Source codes are available from the author upon request.
Acknowledgments

The author acknowledges G. L. Pavlis, Associate Editor Cleat Zeiler, and an anonymous reviewer for thoughtful and helpful critiques of this manuscript. Their comments greatly improved the quality of this paper.

References


Augusta State University
Department of Chemistry and Physics
Augusta, Georgia
cpoppei@aug.edu

Manuscript received 19 August 2010