Fields Institute Communications
Volume **00**, 0000

Why use the S-Transform?

R.G. Stockwell
Northwest Research Associates, Colorado Research Associates Division,3380 Mitchell Lane,
Boulder Colorado USA 80301

Abstract. The S-transform (ST) is a time-frequency representation known for its local spectral phase properties. A key feature of the S-transform is that it uniquely combines a frequency dependent resolution of the time-frequency space with absolutely referenced local phase information. This allows one to define the meaning of phase in a local spectrum setting, and results in many advantageous characteristics. It also exhibits a frequency invariant amplitude response, in contrast to the wavelet transform.

This manuscript outlines the derivation of both the S-transform and the efficient Discrete Orthonormal S-transform and gives a detailed description of the implementation of the algorithms. The arbitrary sampling of the time-frequency space is illustrated and the direct connection between the S-transform, the Discrete Orthonormal S-transform, and the Fourier transform is described.

Illustrations of the properties of the S-transform approach are given and a comparisons to wavelet transform is performed. The S-transform is shown to have absolutely referenced phase information, a quality that the continuous wavelet transform is lacking. In addition, the S-transform is shown to have a frequency invariant amplitude response in contrast to the continuous wavelet transform which attenuates high frequency signals relative to the low frequency signals.

1 Introduction

Spectral analysis using the Fourier transform has been a powerful tool in the analysis of geophysical data (for instance [55], [56], [57], etc.). One drawback of the technique of Fourier transforms is that it only produces the time-averaged spectrum. This is adequate for stationary time series in which the characteristics of the time series do not change with time. In the geophysical data, however, stationarity is an unrealized idealization. The spectral content of the time series changes with time, and the time-averaged amplitudes found by Fourier methods are inadequate to describe such phenomena. Thus in recent years, Fourier analysis has given way

1991 *Mathematics Subject Classification.* Primary ; Secondary.
Supported by National Science Foundation Grant ATM-0350680.

to more advanced representations known as joint time-frequency representations or TFRs.

The S-transform (ST) [22] has found application in a range of fields. It is similar to a continuous wavelet transform in having progressive resolution but unlike the wavelet transform, the S-transform retains absolutely referenced phase information and has a frequency invariant amplitude response. Absolutely referenced phase information means that the phase information given by the S-transform refers to the argument of the cosinusoid at zero time (which is the same meaning of phase given by the Fourier transform). The S-transform defines what local phase means in an intuitive way, at a peak in local spectral amplitude (indicating a quasimonochromatic signal), as well as off peak, where the rate of change of the phase leads to a channel Instantaneous Frequency analysis. The S-transform not only estimates the local power spectrum, but also the local phase spectrum. It is also applicable to the general complex valued time series.

The role of phase in wavelet analysis is not as well understood as it is for the Fourier transform, especially for orthonormal wavelet representations. Current complex wavelets such as complex Daubechies wavelets [16], dual tree complex wavelets [17, 18, 19], and complex Shannon Wavelets [20] do not have a direct relationship to the Fourier transform.

2 The Fourier Transform

It is often useful to think of the time series as a single vector in an N-dimensional vector space. The basis vectors of this time series in the time domain are the vectors (1,0,0,...,0), (0,1,0,...,0) and so on. The action of the Fourier transform is simply a change of basis on the time series, from these delta function basis vectors (time domain), to sinusoidal basis vectors (frequency domain). The time series itself, which is a defined single vector in this N-dimensional vector space, remains unchanged.

One of the reasons that the Fourier transform is ubiquitous in the analysis of geophysical phenomena is that the sinusoidal basis functions are the solution to the mathematical equations describing a small perturbation of a physical system about a stable equilibrium, and thus provides a suitable framework for studying such phenomena. Also a number of theoretical predictions concerning the evolution of such systems is easily couched in terms of Fourier theory. Thus changing the representation of the time series may present the information contained in the time series in a more easily assimilated form.

The spectrum $H(f)$ of a time series $h(t)$ is given by standard Fourier analysis as [59]:

$$H(f) = \int_{-\infty}^{\infty} h(t) e^{-\imath 2\pi f t} dt \tag{2.1}$$

and its inverse relationship is:

$$h(t) = \int_{-\infty}^{\infty} H(f) e^{\imath 2\pi f t} df \tag{2.2}$$

where the units of t and f are such that the product is dimensionless.

In dealing with a discrete time series of N points with a sampling interval of T, the discrete Fourier transform (DFT) is used (where $k = 0...N-1$ and $m = 0...N-1$)

$$H\left[\frac{n}{NT}\right] = \frac{1}{N} \sum_{k=0}^{N-1} h\left[kT\right] e^{-\frac{\imath 2\pi n k}{N}} \tag{2.3}$$

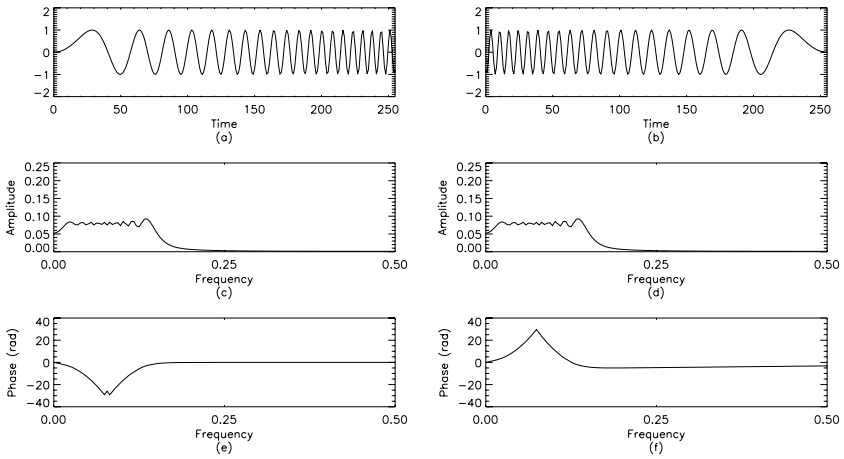

Figure 1 a) Sample 256 point time series of an linearly *in*creasing frequency chirp. The function is $h(t) = sin(\frac{2\pi}{20}\frac{t^2}{256^2})$. **b)** Sample time series of similar to (1a), except here the frequency chirp is linearly *de*creasing. **c,d)** Amplitude Spectra. **e,f)** Phase spectra.

and its inverse relationship is:

$$h\left[kT\right] = \sum_{n=0}^{N-1} H\left[\frac{n}{NT}\right] e^{\frac{\imath 2\pi nk}{N}} \tag{2.4}$$

The Fourier transform is limited in that, although it can determine the spectral components of a time series, it is lacking in information on *when* these components occurred.

2.1 A Motivating Example. A weakness in the Fourier transform representation of data is that the temporal information is not readily apparent; it is obscured in the phase spectrum of the Fourier transform. Very different data, with very different underlying physics, can have similar power spectra as shown in Figure 1 with a simple chirp function (a sinusoid whose frequency increases (plot a) or decreases (plot b) linearly in time).

As shown in Figure 1c,d the amplitude spectra are identical for both time series, indicating that there is no information about when the different frequencies existed in the time series. The time-local information is hidden in the value of the phases of the spectrum in Figure 1e,f, but this information is not readily accessible. From this point of view, the purpose of time-frequency representations is to unfold this temporal information into a more easily digestible form.

3 The S-transform

The S-transform produces a time-frequency representation of a time series. It uniquely combines a frequency dependent resolution with simultaneously localizing the real and imaginary spectra. It was first published in 1996 [22], and since has seen several interesting applications (see [69] and references therein).

The continuous S-transform [22] of a function $h(t)$ is

$$S(\tau, f) = \int_{-\infty}^{\infty} h(t) \frac{|f|}{\sqrt{2\pi}} e^{-\frac{(\tau-t)^2 f^2}{2}} e^{-i2\pi f t} dt. \tag{3.1}$$

A "voice" $S(\tau, f_0)$ is defined as a one dimensional function of time for a constant frequency f_0, which shows how the amplitude and phase for this exact frequency changes over time. A "local spectrum" $S(\tau_0, f)$ is a one dimensional function of frequency for a constant time t_0.

Why the Gaussian Window? The Gaussian window is chosen for several reasons: 1) it uniquely minimizes the quadratic time-frequency moment about a time-frequency point [70], 2) it is symmetric in time and frequency - the Fourier transform of a Gaussian is a Gaussian, 3) there are no sidelobes in a Gaussian function (a local maxima in the absolute value of the S-transform is not an artifact). However, as is the case with Power Spectral Estimation, any desired window or apodizing function may be employed.

3.1 Derivation of S-Transform from STFT:. If the time series $h(t)$ is windowed (or multiplied point by point) with a window function $g(t)$ then the resulting spectrum is

$$H(f) = \int_{-\infty}^{\infty} h(t) g(t) e^{-\imath 2\pi f t} dt \tag{3.2}$$

The S-transform can be found by first defining a particular window function, a normalized Gaussian

$$g(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}} \tag{3.3}$$

and then allowing the Gaussian to be a function of translation τ and dilation (or window width) σ.

$$S^*(\tau, f, \sigma) = \int_{-\infty}^{\infty} h(t) \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\tau)^2}{2\sigma^2}} e^{-\imath 2\pi f t} dt \tag{3.4}$$

which, with a particular value of σ, is similar in definition to the STFT. This is a special case of the Multiresolution Fourier transform. Because this is a function of three independent variables, it is impractical as a tool for analysis. Simplification can be achieved by adding the constraint restricting the width of the window σ to be proportional to the period (or inverse of the frequency)

$$\sigma(f) = \frac{1}{|\, f \,|} \tag{3.5}$$

Thus one has the S-transform of equation 3.1.

3.2 ST as a Convolution. The S-transform can be written as a convolution of two functions over the variable t

$$S(\tau, f) = \int_{-\infty}^{\infty} p(t, f) g(\tau - t, f) dt \tag{3.6}$$

or

$$S(\tau, f) = p(\tau, f) * g(\tau, f) \tag{3.7}$$

where

$$p(\tau, f) = h(\tau) e^{-\imath 2\pi f \tau} \tag{3.8}$$

and

$$g(\tau, f) = \frac{\mid f \mid}{\sqrt{2\pi}} e^{-\frac{\tau^2 f^2}{2}} \tag{3.9}$$

Let $B(\alpha, f)$ be the Fourier transform (from τ to α) of the S-transform $S(\tau, f)$. By the convolution theorem the convolution in the τ (time) domain becomes a multiplication in the α (frequency) domain:

$$B(\alpha, f) = P(\alpha, f)G(\alpha, f) \tag{3.10}$$

(Likewise, $P(\alpha, f)$ and $G(\alpha, f)$ are the Fourier transform of $p(\tau, f)$ and $g(\tau, f)$.) Explicitly,

$$B(\alpha, f) = H(\alpha + f)e^{-\frac{2\pi^2\alpha^2}{f^2}} \tag{3.11}$$

where $H(\alpha + f)$ is the Fourier transform of (3.8), and the exponential term is the Fourier transform of the Gaussian function (3.9). Thus the S-transform is the inverse Fourier transform (from α to τ) of the above equation (for $f \neq 0$).

$$S(\tau, f) = \int_{-\infty}^{\infty} H(\alpha + f)e^{-\frac{2\pi^2\alpha^2}{f^2}} e^{\imath 2\pi\alpha\tau} d\alpha \tag{3.12}$$

The exponential function in Eq. (3.12) is the frequency dependent localizing window and is called the *Voice Gaussian*. This window is centered around the zero frequency and thus plays the role of a low pass filter for each particular voice. This is in contrast to a wavelet or band pass filtered approach to calculating voices of a time-frequency representation.

3.3 Derivation of the S-transform from the Wavelet Transform. The following derivation demonstrates the relationship between the S-transform and the continuous wavelet transform. The continuous wavelet transform can be defined as a series of correlations of the time series with a function called a wavelet:

$$W(\tau, d) = \int_{-\infty}^{\infty} h(t) \frac{1}{\sqrt{d}} \psi(\frac{t - \tau}{d}) dt \tag{3.13}$$

The S-transform of a function $h(t)$ can be defined as a CWT with a specific mother wavelet multiplied by a correction factor and by replacing dilation d with the inverse of frequency f:

$$S(\tau, f) = \sqrt{\frac{f}{2\pi}} e^{-i2\pi f\tau} W(\tau, f) \tag{3.14}$$

where the mother wavelet is defined as

$$\psi((t - \tau)f) = e^{-\frac{(t-\tau)^2 f^2}{2}} e^{-\imath 2\pi f(t-\tau)} \tag{3.15}$$

The wavelet in (3.15) does not satisfy the admissibility condition of having a zero mean, and therefore (3.14) is not strictly a CWT. Writing out (3.14) explicitly gives the S-transform as seen in (3.1).

There are two vital terms in the S-transform definition: 1) the phase function $e^{-\imath 2\pi ft}$ and 2) the normalization $\frac{|f|}{\sqrt{2\pi}}$. The phase function, resulting from application of the phase factor in Eq. (3.14) deserves further discussion. It is in fact a phase correction of the definition of the Wavelet transform. It eliminates the concept of "wavelet analysis" by separating the mother wavelet into two parts, the slowly varying envelope (the Gaussian function) which localizes in time, and the oscillatory exponential kernel $e^{-\imath 2\pi ft}$ which selects the frequency being localized. It is the time localizing Gaussian that is translated while the oscillatory exponential

kernel remains stationary. By not translating the oscillatory exponential kernel, the S-transform localizes the real and the imaginary components of the spectrum independently, localizing the phase spectrum as well as the amplitude spectrum. This is referred to as *absolutely referenced phase information*. Absolutely referenced phase means that the phase information given by the S-transform is always referenced to time $t = 0$, which is also true for the phase given by the Fourier transform. This is true for each S-transform sample of the time-frequency space. The normalization factor also deserves further discussion. It normalizes the time domain localizing window (the Gaussian function) to have unit area. In doing this, the amplitude of the S-transform has the same meaning as the amplitude of the Fourier transform. This provides a frequency invariant amplitude response in contrast to the Wavelet approach. The term "frequency invariant amplitude response" means that for a sinusoid with an amplitude A_0 (h(t) = $A_0 cos(2\pi ft)$), the S-transform returns an amplitude A_0 regardless of the frequency f.

3.4 Linearity and the Effect of Noise. The S-transform is a linear operation on the time series $h(t)$. This is important for the case of additive noise in which one can model the data as $data(t) = signal(t) + noise(t)$ and thus the operation of the S-transform leads to

$$S\{data\} = S\{signal\} + S\{noise\} \tag{3.16}$$

This is an advantage over the bilinear class of TFRs (such as Cohen's Class [7]) where one finds

$$\begin{aligned} TFR\{data\} &= TFR\{signal\} + TFR\{noise\} \\ &+ 2 * TFR\{signal\} * TFR\{noise\} \end{aligned} \tag{3.17}$$

3.5 Discrete ST. In the discrete case, there are computational advantages to using the equivalent frequency domain definition of the S-transform (where $H[\frac{n}{NT}]$ is the Fourier transform of the N-point time series $h[kT]$)

$$S[jT, \frac{n}{NT}] = \sum_{m=-N/2}^{N/2-1} H[\frac{m+n}{NT}] e^{-\frac{2\pi^2 m^2}{n^2}} e^{\frac{i2\pi mj}{N}}, \quad n \neq 0 \tag{3.18}$$

and for the $n = 0$ voice, it is equal to the constant defined as

$$S[jT, 0] = \frac{1}{N} \sum_{k=0}^{N-1} h[kT] \tag{3.19}$$

where j, m and $n = 0, 1, ..., N-1$. The sampling of the S-transform is such that $S[jT, \frac{n}{NT}]$ has a point at each time sample and at each Fourier frequency sample. Similar to a STFT and CWT, this is redundant.

3.6 Discrete Inverse S-transform. The discrete inverse of the S-transform is performed through the intermediate step of computing the discrete Fourier transform. Summing the S-transform matrix along the voices (rows) gives ($n \neq 0$)

$$\sum_{j=0}^{N-1} S[\frac{n}{NT}, jT] = \sum_{j=0}^{N-1} \sum_{m=-N/2}^{N/2-1} H[\frac{m+n}{NT}] e^{-\frac{2\pi^2 m^2}{n^2}} e^{\frac{i2\pi mj}{N}} \tag{3.20}$$

Reordering the sequence of summation, we have

$$\sum_{j=0}^{N-1} S[\frac{n}{NT}, jT] = \sum_{m=-N/2}^{N/2-1} H[\frac{m+n}{NT}] e^{-\frac{2\pi^2 m^2}{n^2}} \sum_{j=0}^{N-1} e^{\frac{i2\pi mj}{N}} \tag{3.21}$$

By the orthogonal property, the sum over j is zero unless $m =$ zero in which case it is equal to N. Thus the average of the voices of $S(n/NT, jT)$ is

$$\sum_{j=0}^{N-1} S[\frac{n}{NT}, jT] = \sum_{m=-N/2}^{N/2-1} N\delta_{m,0} H[\frac{m+n}{NT}] e^{-\frac{2\pi^2 m^2}{n^2}}. \tag{3.22}$$

$$\frac{1}{N} \sum_{j=0}^{N-1} S[\frac{n}{NT}, jT] = H[\frac{n}{NT}]. \tag{3.23}$$

Therefore the discrete inverse of the S-transform is ($\forall n$):

$$h[kT] = \frac{1}{N} \sum_{n=0}^{N-1} \left\{ \sum_{j=0}^{N-1} S[\frac{n}{NT}, jT] \right\} e^{\frac{i2\pi nk}{N}}. \tag{3.24}$$

3.7 Implementation of S-Transform algorithm. Figure 2 shows a schematic representation of the implementation of the Discrete S-transform. The first step of the ST algorithm is to calculate the Fourier transform of the N-point time series. One then shifts the spectrum so that the voice frequency becomes the zero frequency. This shift is the phase correction of Eq. 3.14. One then multiplies the spectrum with the N-point voice Gaussian function to select the frequency range. The N-point inverse Fourier transform is applied to this product to populate the voice of the ST as indicated by the arrow labeled IFFT (each voice will have N-points). This procedure (applying the window and performing the inverse FT) is repeated until all the voices are populated. By employing the fast Fourier transform algorithm, the calculation of the voices is extremely efficient.

This has very strong parallels to the implementation of the DOST shown below.

3.8 The S-transform and the Analytic Signal. For a simple signal, that of a sinusoid with a constant frequency,

$$h(t) = \exp(\imath 2\pi f_o t) \tag{3.25}$$

Vakman [60] gives the following definition for the Instantaneous Frequency:

$$IF(t) = \frac{1}{\imath 2\pi} \frac{\partial}{\partial t} arg(t) \tag{3.26}$$

where the argument (or phase) of this function is

$$arg(t) = \imath 2\pi f_o t. \tag{3.27}$$

and the frequency is $f = f_o = \frac{arg}{\imath 2\pi t}$.

To calculate the Instantaneous Frequency function the original time series has to be transformed into the following form:

$$h(t) = A(t) \exp(Arg(t)) \tag{3.28}$$

and then the $Arg(t)$ function must be extracted by use of the arctan function. If the time series is real-valued, however, then the phase has no meaning (being identically equal to zero).

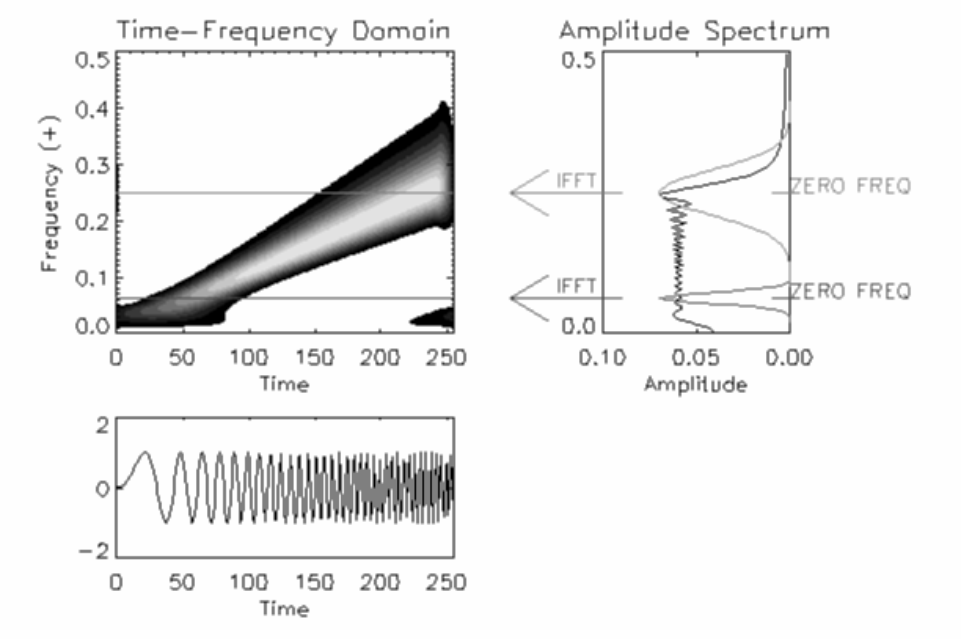

Figure 2 Schematic of the S-transform algorithm. Top left plot shows the S-transform amplitude for positive frequencies. The top right plot shows the Fourier transform amplitudes with the frequencies on the y-axis in order to line up with the frequency axis of the ST plot. The bottom left plot shows the time series with the time axis lined up with the ST plot.

Following Vakman [60] the Hilbert transform is used to extend the real-valued time series into a complex time series. The phase is then the usual definition of phase for a complex function. The Hilbert transform of a function $h(t)$ is [61]

$$X(\tau) = \frac{1}{\pi} \int \frac{h(t)}{\tau - t} dt \tag{3.29}$$

and the inverse is

$$h(t) = \frac{1}{\pi} \int \frac{X(\tau)}{\tau - t} d\tau \tag{3.30}$$

The *Analytic Signal* of a real function is defined as:

$$AS\{h(t)\} = h(t) + iX(t) \tag{3.31}$$

Therefore the Instantaneous Frequency (in radians) is the defined as [60]:

$$IF(t) = \frac{\partial}{\partial t} \arctan(\frac{\Im[AS(t)]}{\Re[AS(t)]}) \tag{3.32}$$

or

$$IF(t) = \frac{\partial}{\partial t} \arctan(\frac{X(t)}{h(t)}) \tag{3.33}$$

It can be shown that the Analytic Signal is closely related to the S-transform. The S-transform is a band pass filtered analytic signal of the original time series

$h(t)$. Starting from Eq. 3.12, one can shift the spectrum by replacing α with $\kappa = \alpha + f$, and treating f as an arbitrary parameter

$$S(\tau, f) = \int_{-\infty}^{\infty} H(\kappa) e^{-\frac{2\pi^2(\kappa-f)^2}{f^2}} e^{\imath 2\pi\alpha\tau} d\kappa \tag{3.34}$$

The negative frequencies are removed from this time series (if f is positive). Thus any voice of the S-transform corresponds to a band pass filtered Analytic Signal of the original time series.

3.8.1 *Instantaneous Frequency and the Nyquist Frequency.* Since phase is defined as:

$$\Phi(z) = \arctan(\frac{\Im[z]}{\Re[z]}) \tag{3.35}$$

there is a 2π ambiguity in the value for each value of phase resulting from the use of the arctan function. Thus when one samples the phase of the Analytic Signal, the phase function must be unrolled. For any change with a magnitude greater than π, one must add or subtract multiples of 2π so that the change is smaller than π. Since the time series is discretely sampled with a sampling interval of T, the maximum rate-of-change of the phase is π over one time unit and is the Nyquist radial frequency

$$f_{\mathrm{MAX}} = \frac{\pi}{T} \tag{3.36}$$

3.8.2 *Channel Instantaneous Frequency.* The use of Instantaneous Frequency in a local spectral representation was introduced for the S-transform [22], and independently was employed by Nelson [50] to study nonstationary multicomponent FM signals. Gardner and Magnasco [51] also used band-pass filtered Instantaneous Frequency analysis for the analysis of sound.

The S-transform approach (including the DOST described below) provides an extension of Instantaneous Frequency to broadband signals. The voice for a particular frequency ν_1 can be written as

$$S(\tau, \nu_1) = A(\tau, \nu_1) e^{i\Phi(\tau,\nu_1)}. \tag{3.37}$$

One may use the phase in Eq. 3.37 to determine the "local" Instantaneous Frequency similar to that defined by Bracewell [52].

$$GIF(\tau, \nu_1) = \frac{1}{2\pi}\frac{\partial}{\partial\tau}\{2\pi\nu_1\tau + \Phi(\tau, \nu_1)\} \tag{3.38}$$

Thus the absolutely referenced phase information leads to a generalization of the Instantaneous Frequency to broadband signals.

The measure of Instantaneous Frequency is not quantized to the sampled frequencies, and thus can achieve arbitrarily high accuracy. In essence, the application of the Generalized Instantaneous Frequency can be utilized as a local peak-finding algorithm. Again, there are ringing effects and wrap around effects that are familiar in discrete Fourier transform analysis.

3.9 Cross ST Analysis. Because of the ST phase characteristic, it can be employed in a cross spectrum analysis in a local manner. Consider two time series measured by two devices separated by a known distance. Let a sinusoidal wave propagate through the field of view of both devices. What will be seen by these devices is a sinusoid, but with a time shift between the two time series. Since the ST localizes spectral components in time, the cross correlation of specific events on two spatially separated STs will give the phase difference and hence the phase

velocity can be derived with knowledge of the frequency and the distance between measurements.

The amplitude of the CrossST indicates coincident signals. The phase of the CrossST at local maxima indicates the phase lag (or phase difference) between those two coincident signals. The Cross ST of two time series $h(t)$ and $g(t)$ is defined as:

$$\text{Cross}ST(\tau, f) = S_h(\tau, f) \left\{S_g(\tau, f)\right\}^* \tag{3.39}$$

where the phase of the CrossST is given by

$$arg(CrossST) = \Phi(\tau, f)_h - \Phi(\tau, f)_g \tag{3.40}$$

3.9.1 *The S-transform Shift Theorem.* As written above, the S-transform is defined as

$$S(\tau, f) = \frac{f}{\sqrt{2\pi}} \int h(t) e^{-\frac{(t-\tau)^2 f^2}{2}} e^{-\imath 2\pi f t} dt \tag{3.41}$$

If one translates the time series by an amount r, one can study the effect on the S-transform by making a change of variables $t \to (k - r)$

$$S(\tau, f) = \frac{f}{\sqrt{2\pi}} \int h(k-r) e^{-\frac{((k-r)-\tau)^2 f^2}{2}} e^{-\imath 2\pi f (k-r)} dk \tag{3.42}$$

Rearranging the equation gives

$$S(\tau, f) = \frac{f}{\sqrt{2\pi}} \left(\int h(k-r) e^{-\frac{(k-(\tau+r))^2 f^2}{2}} e^{-\imath 2\pi f k} dk \right) e^{\imath 2\pi f r} \tag{3.43}$$

Making another change of variables, $\tau \to (z - r)$

$$S(z-r, f) e^{-\imath 2\pi f r} = \frac{f}{\sqrt{2\pi}} \int h(k-r) e^{-\frac{(k-z)^2 f^2}{2}} e^{-\imath 2\pi f k} dk \tag{3.44}$$

Thus the S-transform Shift Theorem states that if

$$h(t) \Leftrightarrow S(\tau, f) \tag{3.45}$$

then

$$h(t-r) \Leftrightarrow S(\tau - r, f) e^{-\imath 2\pi f r} \tag{3.46}$$

3.9.2 *Time shift of a monochromatic signal.* Starting with equation 3.12

$$S(\tau, f) = \int H(\alpha + f) e^{-\frac{2\pi^2 \alpha^2}{f^2}} e^{\imath 2\pi \alpha \tau} d\alpha \tag{3.47}$$

the S-transform of a sinusoidal function $h(t)$

$$h(t) = e^{\imath 2\pi w t} \tag{3.48}$$

with a spectrum

$$H(\alpha) = \frac{1}{2} \delta(\alpha - w) \tag{3.49}$$

is

$$S(\tau, f) = \int \delta(\alpha + f - w) e^{-\frac{2\pi^2 \alpha^2}{f^2}} e^{\imath 2\pi \alpha \tau} d\alpha \tag{3.50}$$

or

$$S(\tau, f) = \frac{1}{2} e^{-\frac{(f-w)^2}{(\frac{f}{2\pi})^2}} e^{\imath 2\pi (f-w)\tau} \tag{3.51}$$

If one introduces a constant phase shift into the sinusoid,

$$g(t) = h(t) e^{i\Phi} = e^{\imath 2\pi w t + \imath \Phi} \tag{3.52}$$

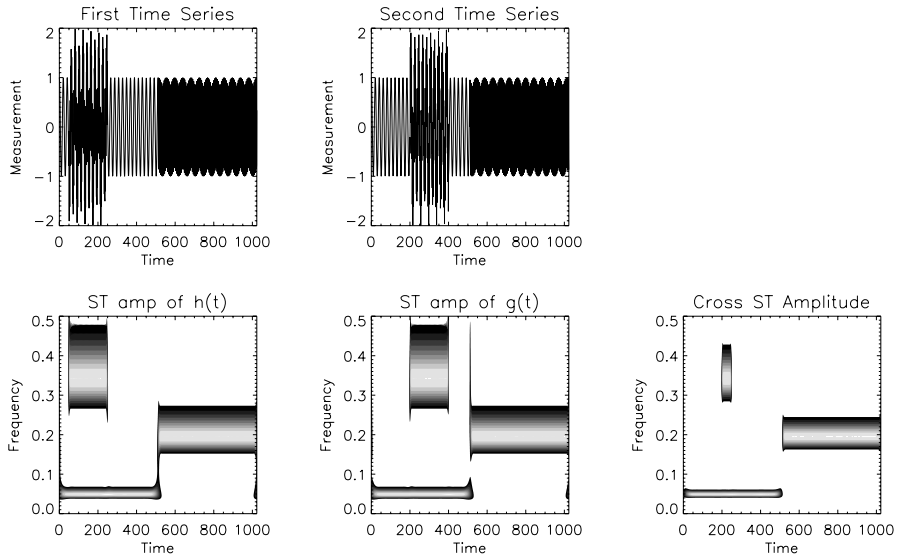

Figure 3 Illustration of the CrossST of two time series.

with a resulting spectrum

$$H(\alpha) = \frac{1}{2}\delta(\alpha - w)e^{\imath\Phi} \tag{3.53}$$

this results in a phase shift of the S-transform

$$S\{g(t)\}_{(\tau,f)} = \frac{1}{2}e^{-\frac{(f-w)^2}{(\frac{f}{2\pi})^2}}e^{\imath 2\pi(f-w)\tau}e^{\imath\Phi} \tag{3.54}$$

Thus if one performs a localized cross spectral S-transform analysis by multiplying the ST of $h(t)$ with the complex conjugate of the ST of $g(t)$

$$\begin{aligned} &S(\tau,f)_{\{h(t)\}} \cdot \left\{S(\tau,f)_{\{g(t)\}}\right\}^* \\ &= \frac{1}{2}e^{-\frac{(f-w)^2}{\frac{f}{2\pi}^2}}e^{\imath 2\pi(f-w)\tau}\frac{1}{2}e^{\frac{(f-w)^2}{\frac{f}{2\pi}^2}}e^{-\imath 2\pi(f-w)\tau}e^{-\imath\Phi} \end{aligned} \tag{3.55}$$

The phase of the Cross ST is

$$\text{Phase}\ \{S_h S_g^*\}_{(\tau,f)} = -\Phi \tag{3.56}$$

This can be used to find a localized time lag between the two time series as a function of t and f. This result is true for the general time series.

In Figure 3, two synthetic time series are created with three distinct sinusoidal components. A low frequency component for the first half, a mid frequency component for the second half, and a high frequency burst signal at the one quarter mark. For each of the two signals, the low frequency signal is in phase (i.e. both cosine function). The mid frequency is out of phase (the first signal is a cosine function, the second signal is a sine function). The high frequency burst is in phase, but temporally translated in the second functions such that there is only a brief overlap.

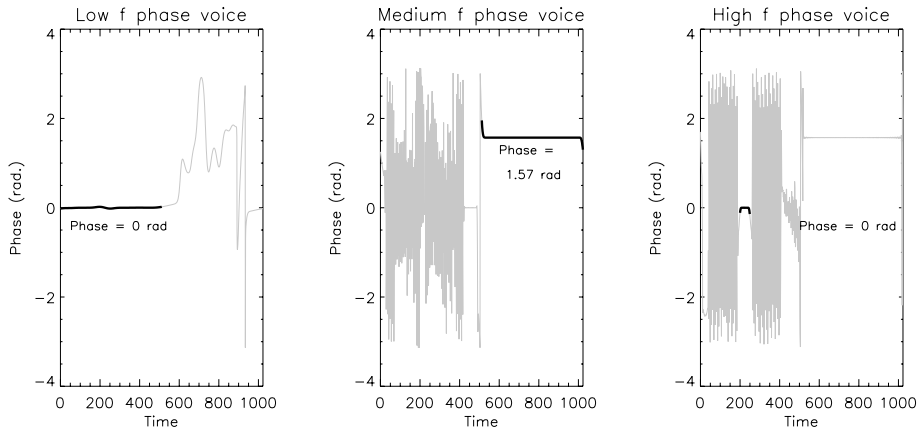

Figure 4 Illustration of the CrossST of two time series. The phase voice for each of the three components is shown, and the correct phase lag for each is indicated.

In Figure 4, individual ST phase voices of the CrossST are shown. The ST phase voice for the low frequency component correctly shows a zero phase lag between the two time series for the first half. The amplitude is negligible for this voice in the second half. For the medium frequency voice, the ST phase voice correctly shows the $\pi/2$ phase lag between the two time series for this component (the first time series had a cos(), and the second time series had a sin() for this component). For the high frequency burst, the ST phase voice in the region of overlap between time = 200 and time = 250 shows the correct phase lag of 0 radians. Other times (indicated by the grey trace) have a negligible amplitude, and therefore the phase is meaningless.

3.10 Co-ST and Quadrature ST. There is still more information that can be gleaned from the CrossST analysis. Because of the phase characteristics of the ST, one can utilize the cross ST function to analyze the in-phase and out-of-phase components in time-frequency space.

As with classical co-spectrum analysis, the real part of the cross ST function gives the in-phase components of the local spectra. The imaginary part of the cross ST function gives the in-quadrature components.

Figure 5a shows the ST amplitude of the first signal. Figure 5b shows the ST amplitude of the second signal. The ST amplitudes of the two signals are very similar, except for the temporal translation of the high frequency component. Figure 5c,d,e show the Cross ST, the Co-ST and the Quadrature ST amplitudes respectively. In the Cross ST plot, all three components appear. However the high frequency component only occurs for a brief time, indicating the overlap between the two time-shifted signals. In the co-ST, only the features that are in phase are present - those being the low frequency component and the high frequency burst. The middle frequency (which is out of phase) does not appear. In the quadrature ST plot only the out-of-phase signals appear, which is the middle frequency component.

3.11 Analysis of a Complex Time Series. Spectral analysis of a two component time series has been performed [53] on data such as the horizontal wind field

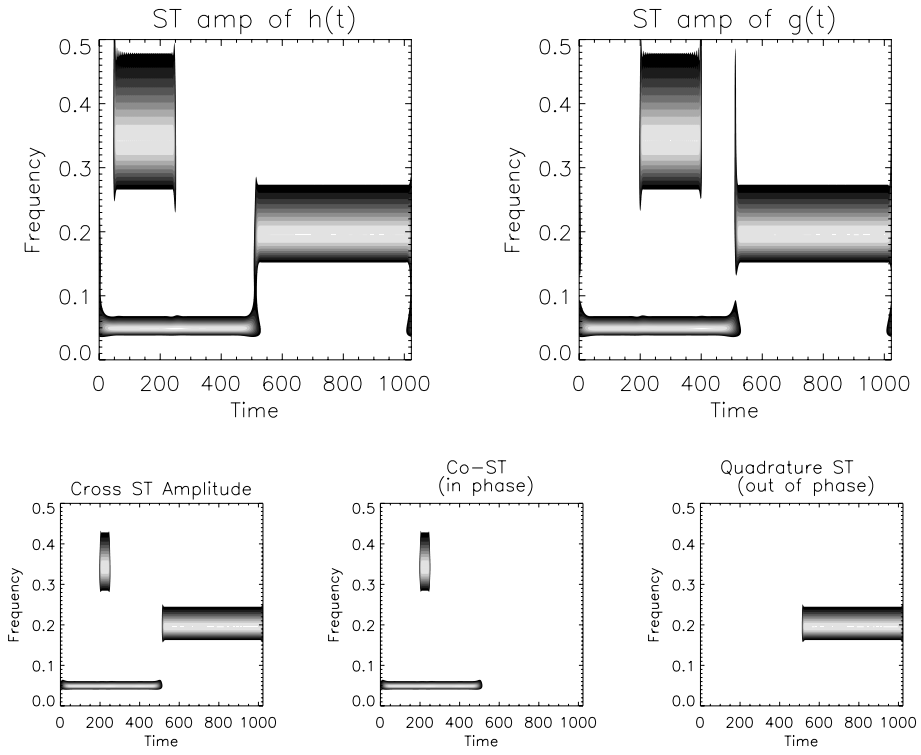

Figure 5 Illustration of the co-ST and the Quadrature ST.

[u(t),v(t)] by constructing a one dimensional complex-valued time series as follows

$$U[t] = u[t] + \imath v[t] = A(t)e^{\imath\omega t} \tag{3.57}$$

where $u[t]$ is the zonal wind (positive eastward), $v[t]$ is the meridional wind (positive northward), $\imath = \sqrt{-1}$, ω is the radial frequency of rotation, and $A(t)$ is the amplitude function. The Fourier transform of a complex valued time series requires both the positive and negative frequencies to completely characterize the wind. Such "rotary spectral" analysis will decompose the wind motions into circular (clockwise and counter-clockwise) components.

As with the Fourier transform, both the S-transform and the DOST can be applied to a 2 component vector time series such as the horizontal wind field [u(t),v(t)] by constructing a one dimensional complex valued time series as in Eq. 3.57.

By looking at the phase difference between the positive and negative frequencies it is possible to infer complete information of the generalized elliptical motion, including the sense of rotation, the major and minor axes, and ellipticity [54]. A synthetic time series illustrating this effect is shown in Figure 6. The complex time series depicts a counter-clockwise rotating surface wind vector, with an amplitude that starts small, grows in time, then decays. The time series was constructed in

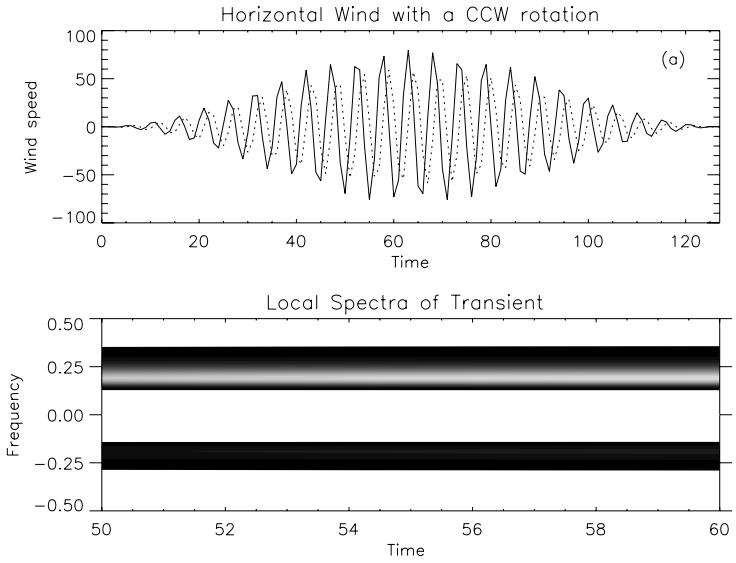

Figure 6 (a) A two component time series, (analogous to a horizontal wind measurement for instance) that initially increases in amplitude, then decreases in amplitude. Additionally, the wind is rotating counter-clockwise. (b) The S-transform amplitude. The changing amplitude can be seen. The counter-clockwise motion can be seen in that the positive frequency component is larger than the negative frequency component.

the form of Eq. 3.57 as follows $U[t] = 80\cos(2\pi t 24.4/128) + \imath 60\sin(2\pi t 24.4/128)$ where $t = 0, 1...127$ and the amplitude is varied by multiplying the time series with a hanning window. The amplitude of the S-transform of this time series is shown in Figure 6b. The white background denotes regions of negligible energy, and filled regions correspond to large amplitudes. The peak amplitude occurs at $f = 24/128$ and $t = 64$. Because this amplitude is larger than its mirror in the negative frequencies ($f = -24/128$ and $t = 64$), this represents a counter-clockwise oscillation. If the negative frequency is larger, then rotation is clockwise.

4 The Discrete Orthonormal S-transform (DOST)

There are several reasons to desire an orthonormal time-frequency version of the S-transform. An orthonormal transformation takes an N-point time series to an N-point time-frequency representation, thus achieving maximum efficiency. Also, each point of the result is independent from any other point. The transformation matrix is orthogonal, meaning that the inverse matrix is equal to the complex conjugate transpose. By being an orthonormal transformation, the vector norm is preserved. Thus a Parseval theorem applies, stating that the norm of the time series equals the norm of the DOST.

The efficient representation of the S-transform can be defined as the inner products between a time series $h[kT]$ and the basis functions defined as a function of $[kT]$, with the parameters ν (a frequency variable indicative of the center of a frequency band and analogous to the "voice" of the S-transform), β (indicating the

frequency resolution), and τ (a time variable indicating the time localization).

$$S\{h[kT]\} = S(\tau T, \frac{\nu}{NT}) = \sum_{k=0}^{k=N-1} h[kT] S_{[\nu,\beta,\tau]}[kT] \tag{4.1}$$

4.1 Derivation of the basis functions. By extending filter bank theory, in combination with the unique phase correction of the S-transform, the time domain basis functions for the S-transform are developed. Within a frequency band (i.e. a particular voice), several orthogonal basis functions are formed by a linear combination of the Fourier basis functions in that frequency band. There are β components in this operation. Thus β basis functions can be derived by applying the appropriate phase functions to the components (where each of the β basis functions is indexed by $\tau = 0, 1, ...\beta-1$). The key to creating orthogonal functions is the careful selection of the frequency shift applied to the Fourier basis functions. This action is the analog of the "phase correction" of the S-transform. This is where the absolutely referenced phase information originates, and it is what distinguishes these basis functions from wavelets. It is also the inclusion of this additional operation that is distinct from a simple filtering operation. One should note, the voices of the DOST cannot be created by a filtering operation of the original time series.

The basis functions can be derived by starting with a partitioning of the spectrum (a simple restricted sum of complex-valued Fourier basis functions) defined in the time domain (function of $[kT]$), centered at frequency ν with a bandwidth of β and applying the appropriate phase and frequency shifts:

$$S(kT)_{[\nu,\beta,\tau]} =$$

$$\frac{1}{\sqrt{\beta}} \sum_{f=\nu-\frac{\beta}{2}}^{f=\nu+\frac{\beta}{2}-1} exp(i2\pi\frac{k}{N}f) exp(-i2\pi\frac{\tau}{\beta}f) exp(-i2\pi\frac{N}{2NT}\tau T) \tag{4.2}$$

where $\frac{1}{\sqrt{\beta}}$ is a normalization factor to insure orthonormality of the basis functions.

Thus the basis function for the Discrete Orthonormal S-transform (DOST) of voice frequency ν, bandwidth β, and time index τ can be written as:

$$S(kT)_{[\nu,\beta,\tau]} = \frac{e^{-i\pi\tau}}{\sqrt{\beta}} \sum_{f=\nu-\frac{\beta}{2}}^{f=\nu+\frac{\beta}{2}-1} exp(i2\pi(\frac{k}{N} - \frac{\tau}{\beta})f) \tag{4.3}$$

Application of the identity

$$c + cx + ... + cx^{n-1} = \frac{c(1-x^n)}{1-x} \tag{4.4}$$

to Eq. 4.3 leads to the basis functions $S_{[\nu,\beta,\tau]}[kT]$ for the general case

$$S_{[\nu,\beta,\tau]}[kT] = \frac{ie^{-i\pi\tau}}{\sqrt{\beta}} \frac{\{e^{-i2\pi(\frac{k}{N}-\frac{\tau}{\beta})(\nu-\frac{\beta}{2}-\frac{1}{2})} - e^{-i2\pi(\frac{k}{N}-\frac{\tau}{\beta})(\nu+\frac{\beta}{2}-\frac{1}{2})}\}}{2sin[\pi(\frac{k}{N}-\frac{\tau}{\beta})]} \tag{4.5}$$

The limit of the basis function as $\frac{k}{N} \to \frac{\tau}{\beta}$ is well behaved and equal to:

$$\lim_{\frac{k}{N}\to\frac{\tau}{\beta}} S(kT)_{[\nu,\beta,\tau]} = \sqrt{\beta} e^{i\pi\tau} \tag{4.6}$$

At this point, the sampling of the time-frequency space has not yet been determined. Rules must be applied to the sampling of the time-frequency space to ensure orthogonality. These rules are as follows:

- Rule 1: $\tau = 0, 1, \cdots \beta - 1$.
- Rule 2: ν and β must be selected such that each Fourier frequency sample is used once and only once.

Implicit in this definition is the phase correction of the S-transform that distinguishes it from the wavelet or filter bank approach. For each voice, there are one or more local time samples (τ), this number being equal to β (see Rule 1) thus the wider the frequency resolution (large β), the greater the time resolution (large τ) which implicitly shows the Uncertainty Principle.

Distinct from a wavelet function, these basis functions have no vanishing moments (in fact the functions are normalized to unit area). These basis functions are not translations of a single function, and they are not self-similar.

One advantage of this method is that one can directly calculate any voice, without having to iterate through a series of intermediate steps. Also, there is no filter design involved nor any upsampling or downsampling algorithms required. Another advantage is that it allows one to directly apply the ideas of Power Spectrum Estimation, such as applying windows and apodizing functions, to the analysis of the local spectrum of a time series.

Note that in a departure from filter bank theory, the sum is centered on the voice frequency ν. In other words, a frequency translation has been applied. The operation of calculating the inner product of a time series with this basis function not equivalent to a simple filtering operation (in the asymptotically simple case of the time series consisting of an oscillating sinusoid, the resulting voice will be a constant, in amplitude and phase, for each time sample). This frequency shift is vital when the characteristic of absolutely referenced phase information (and cross local spectrum analysis and generalized instantaneous frequency) is described, and is the distinguishing difference between the S-transform approach, and the wavelet/filter bank approach. This shift in frequency is the key feature of the original S-transform. In Eq 11 of [22], the "n'' voice $S[jT, \frac{n}{NT}]$ has the same frequency translation applied by the shift of the spectrum $H[(m+n)/NT]$ by "n" which centers the spectrum around the n frequency in Eq. 3.18. It is precisely this property of the basis functions that provides absolutely referenced phase information, and it is also this property that implies that the basis functions are not self-similar.

It is easy to show that the basis functions are indeed orthonormal and have compact support in frequency. They are not compactly supported in time, but they are local. The property of "compact support" refers to a particular transform. An orthonormal wavelet does not have compact support under a Fourier transform. By the Uncertainty Principle, one cannot have a compactly supported function which has a compact Fourier transform. These basis functions are compact in frequency, and also local in time, while maintaining orthogonality. The requirement is that the ν and β values are chosen such that the bandwidths do not overlap, and that all discrete frequencies are sampled.

The fundamental difference between the discrete orthonormal wavelet transform (DWT) and the DOST is in the absolutely referenced phase information [69]. The phase of a voice of the DOST has the same meaning as the phase of a component of the Fourier transform.

As is well known from Fourier theory, the translation in time of the wavelet (or any function) corresponds to a phase modulation in the Fourier spectral domain. Thus, wavelet voices have a phase modulation applied to them in the spectral domain. This is not true for the DOST as there is no direct translation applied to the basis function to produce other basis functions. It is this characteristic of self-similarity that force this phase modulation (and therefore the lack of absolutely referenced phase) on all wavelet transforms.

The fundamental idea of a wavelet transform is the decomposition of a signal into deformations of a mother wavelet. In the analysis of geophysical data, often the local characteristics of the phase of the signal can lead to great insight. In order to resolve the local phase properties, it is necessary to break the analyzing wavelets from the mother wavelet, and impart an absolutely referenced phase to the basis functions. This break separates the time-localizing windowing function, from the frequency-localizing modulation. As the time-localizing window translates in time, it is essential for this formalism that the frequency modulation does not translate in time, which then forbids the possibility of self-similar basis functions. One can either have a self-similar basis (wavelets), or an absolutely referenced phase basis (DOST).

4.2 Orthonormal basis functions with Octave Sampling. In order to compare the DOST with orthonormal wavelet transforms and with the S-transform, an octave sampling of the time-frequency domain is illustrated (following [69]). This has the property of progressive resolution that the S-transform shares. Octave sampling implies that the voice bandwidth doubles for each increasing voice (as sampling allows).

By imposing specific rules on the basis functions (here octave sampling) it implies a strict definition for ν and β. By introducing a new variable p which corresponds to the octave number $p = 0, 1, 2, \cdots, log_2(N) - 1$ one can define all the parameters (ν, β, τ) of equation 4.5 in terms of p as follows:

For $p > 1$, we have:

$$p = 2, ..log_2(N) - 1 \tag{4.7}$$

$$\nu = 2^{(p-1)} + 2^{(p-2)} \tag{4.8}$$

$$\beta = 2^{(p-1)} \tag{4.9}$$

$$\tau = 0, 1, ..2^{(p-1)} - 1 \tag{4.10}$$

For the case $p = 0$, then $\nu = 0$ and $\beta = 1$ and $\tau = 0$. Also when $p = 1$, then $\nu = 1$ and $\beta = 1$ and $\tau = 0$.

Thus the DOST basis functions for octave sampling of a time series are given as follows (by application of 4.7, 4.8, 4.9 and 4.10 into 4.5):

$$S_{[p,\tau]}[kT] = \frac{ie^{-i\pi\tau}}{\sqrt{2^{(p-1)/2}}} \frac{\{e^{-i2\pi(\frac{k}{N} - \frac{\tau}{2^{p-1}})(2^{p-1} - \frac{1}{2})} - e^{-i2\pi(\frac{k}{N} - \frac{\tau}{2^{p-1}})(2^p - \frac{1}{2})}\}}{2sin[\pi(\frac{k}{N} - \frac{\tau}{2^{p-1}})]} \tag{4.11}$$

The application of the DOST with octave sampling for the case where $N = 100$ is illustrated in Figures 7 and 8 [69]. In Figure 7, octave sampled DOST basis functions are shown for a specific DOST voice ($p = 5$ and therefore $\nu = 24$ and bandwidth $\beta = 16$), for six different time samples ($\tau = 5...10$). These functions are not simple translations, as is most evident in a comparison of Figures 7d and Figure 7f.

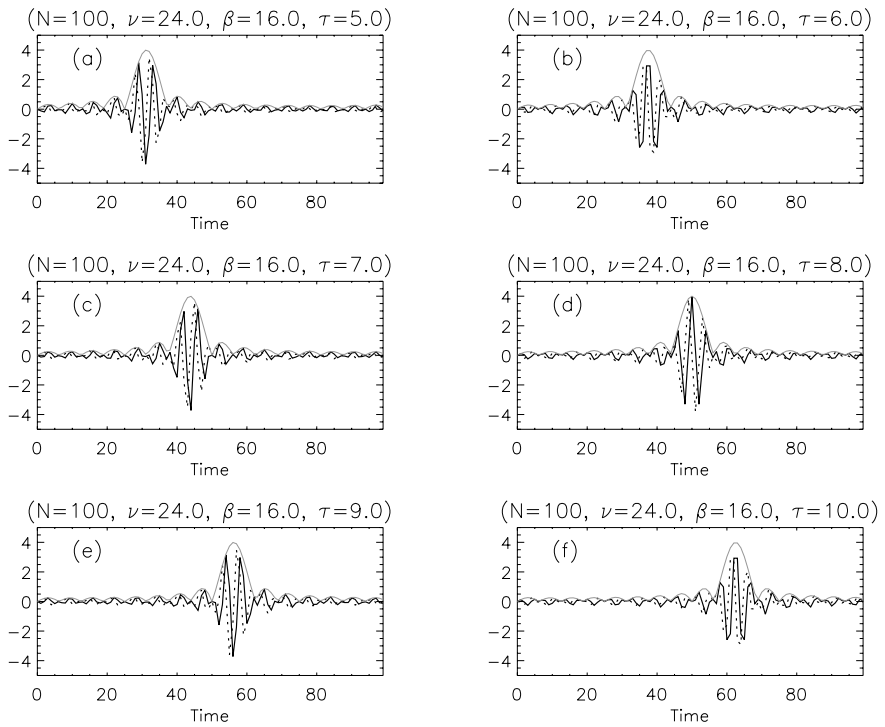

Figure 7 Selected DOST basis functions of **increasing time** with octave sampling for N = 100. Note that the DOST functions are not translates of a single function.

Figure 8 also shows six octave sampled DOST basis functions, but now the order p varies (as does voice frequency ν and the bandwidth β). The time sample τ is chosen simply to center the function. Again the lack of the self-similar property is evident. No basis function shown in Figure 8 is a dilation of any other basis function.

While octave partitioning has the familiar frequency dependent resolution seen in the S-transform and wavelets, the formalism defined in Eq. 4.5 allows for any arbitrary sampling of the time-frequency space (similar to other frame based wavelet approaches). Often the resolution of the time-frequency space is dependent on the source signal; a transient signal will have a broad spectral signature regardless of sampling. Also, while it may be inferred that the above algorithm assumes that the number of points N is a power of two, that is not a strict requirement for octave sampling (as shown in Figures 7 and 8). The octave sampling scheme can be applied to any arbitrary length time series by starting at the Nyquist frequency with

$$\beta = truncate(float(truncate(N/2) + 1)/2) \tag{4.12}$$

and for each lower frequency partition, reducing the bandwidth as follows

$$\beta_{new} = truncate(float(\beta + 1)/2) \tag{4.13}$$

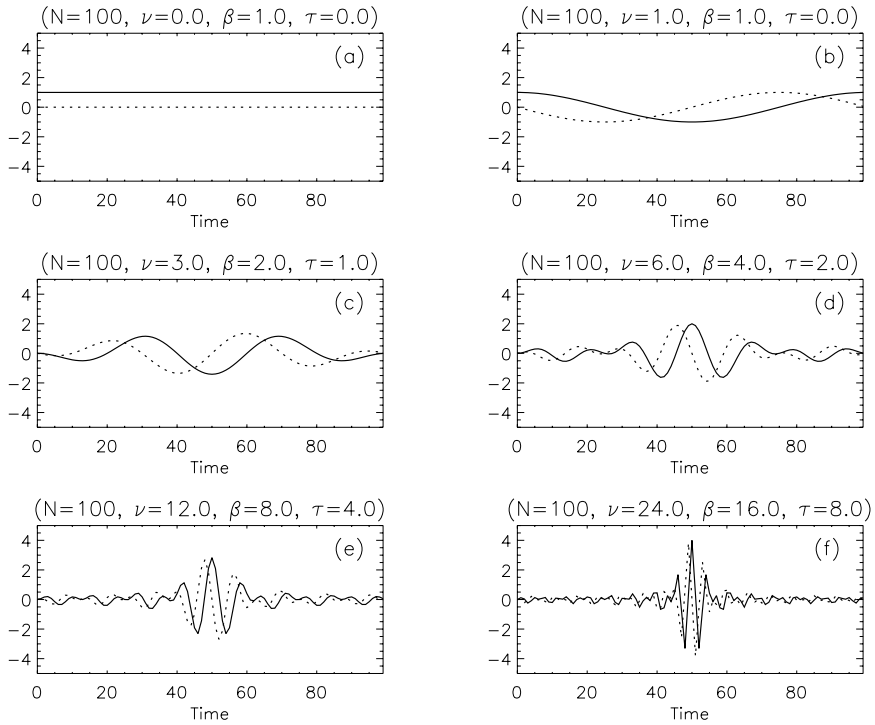

Figure 8 Selected DOST basis functions of **increasing voice frequency** with octave sampling for N = 100. Each plot shows a basis function (solid line = real, dotted line = imaginary) from a different frequency range (i.e. different voice).

For example, the octave sampling for N = 31 is {1,1,2,4,8,8,4,2,1} where the first element is the DC level, then the next four β are the positive frequencies (increasing), and the last four β are the negative frequencies (decreasing).

4.3 Implementation of the DOST. Similar to the S-transform, the implementation of the DOST is more efficiently calculated from the Fourier Spectrum of the time series. This is shown in Figure 9. The DOST begins with the DFT of the time series, indicated by the vertical list of frequency index numbers. This is partitioned into the frequency bands (ν) containing β points. The subspace inverse DFT (IDFT) is applied to to each partitioned frequency band. Note that this is not the N-point IDFT, but it is a β-point IDFT. This insures that β points are created in the voice, and that the decomposition is orthogonal.

For certain purposes it is highly desirable to window and interpolate the DOST time-frequency representation and approach the more ample sampling of the S-transform. Use of windowing and interpolation allows one to build up a smoother and less sparse representation of the time-frequency space, and in the extreme case of "fully redundant" interpolation (i.e. where a time-frequency representation point is calculated for each time sample of the original time series and for each frequency sample of the Fourier Spectrum) then one can recreate the action of the original

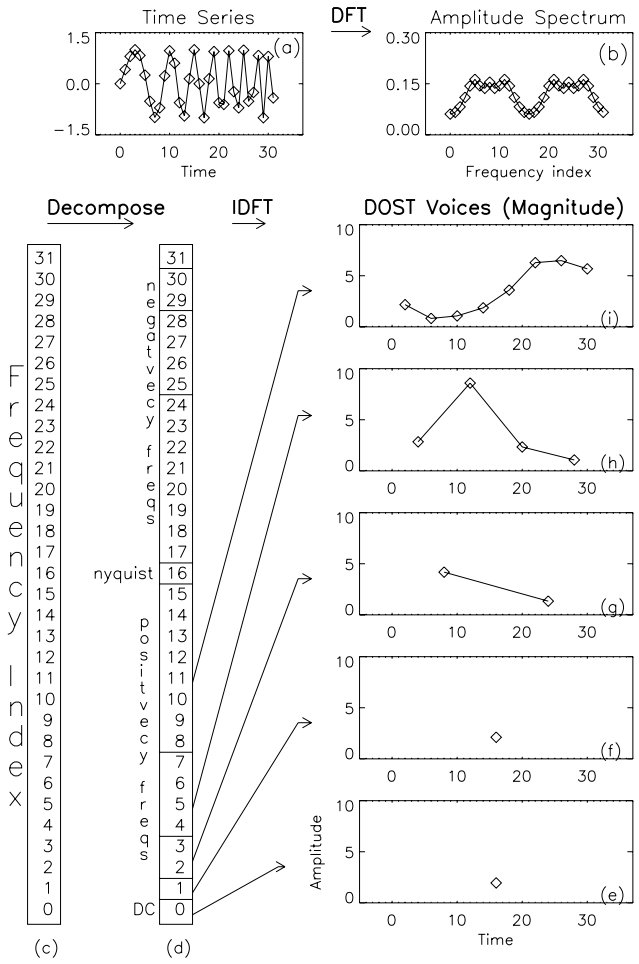

Figure 9 Schematic of the DOST algorithm.

S-transform. The DOST forms the essential skeleton of the time-frequency plane, from which the fully redundant S-transform can be derived.

Figure 10 shows a schematic of the implementation of the oversampled DOST, and the steps necessary to window and oversample the time frequency space in any arbitrary manner. In this figure, the brackets represent the application of an overlapping window functions. The choice of window function is arbitrary as with Fourier transform spectral analysis, and any window or apodizing function may be used.

The oversampling of the time-frequency space is illustrated in Figure 11. A modulated sinusoid time series is shown in Figure 11*a*. The amplitude DOST is shown in Figure 11*b*, and the smooth redundant S-transform amplitude is shown in Figure 11*e*, with intermediate steps shown in Figure 11*c* and *d*. Thus, one can achieve any arbitrary level of compromise between efficient representation (DOST) and a smooth and redundant representation by arbitrary choice of oversampling and windowing. The application of windowing to the DOST is easily done when

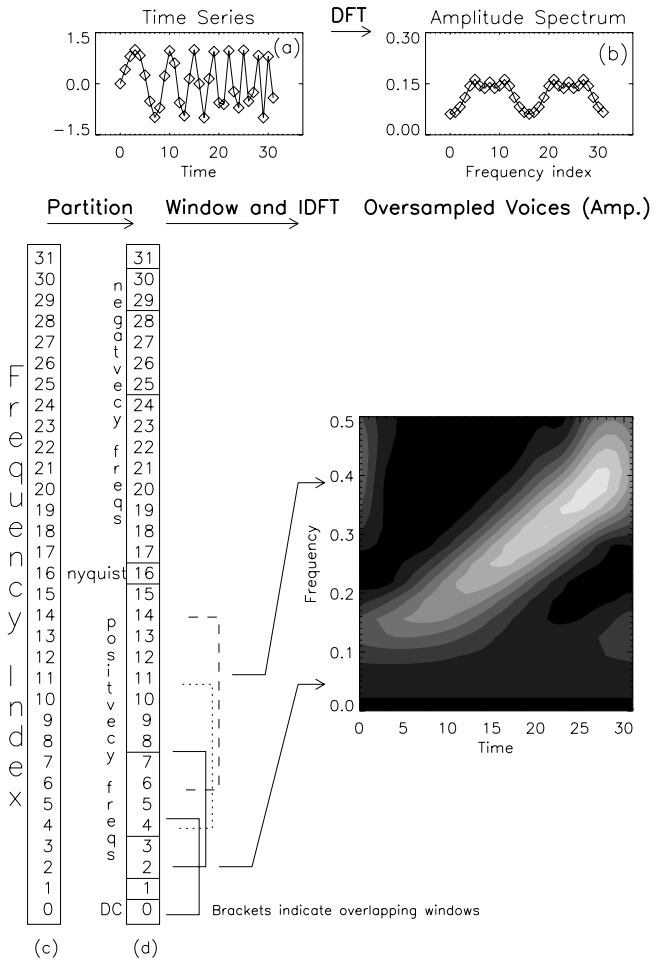

Figure 10 Schematic of the oversampling of the time-frequency space for the DOST algorithm.

one realizes that an implicit boxcar window is applied to the spectral partition in the implementation of the DOST algorithm. In fact, any window (or apodizing function) can be applied at that stage. This is called "voice windowing" and is analogous to the "voice Gaussian" of the S-transform [22]. If one applies such a window, and thus reduces one's sensitivity at the edges of the frequency bands, then it becomes desirable to interpolate and oversample in frequency. When a window is applied to the partitioned spectrum, one should expect this new band to be representative of frequencies $\nu \pm$ bandwidth of window, rather than (in the case of unwindowed DOST) $\nu \pm \beta/2$.

If one wished to interpolate the DOST in time, this can be accomplished by widening the partitions in the frequency domain so that they overlap (thus the partitioned spectrum is then extended to more higher and lower frequencies interpolating the time domain voice). As one goes to full interpolation in time (i.e. a sample at every time sample of the original time series) then there is no partitioning (each spectral partition is the original spectrum). Only the shift of the spectrum

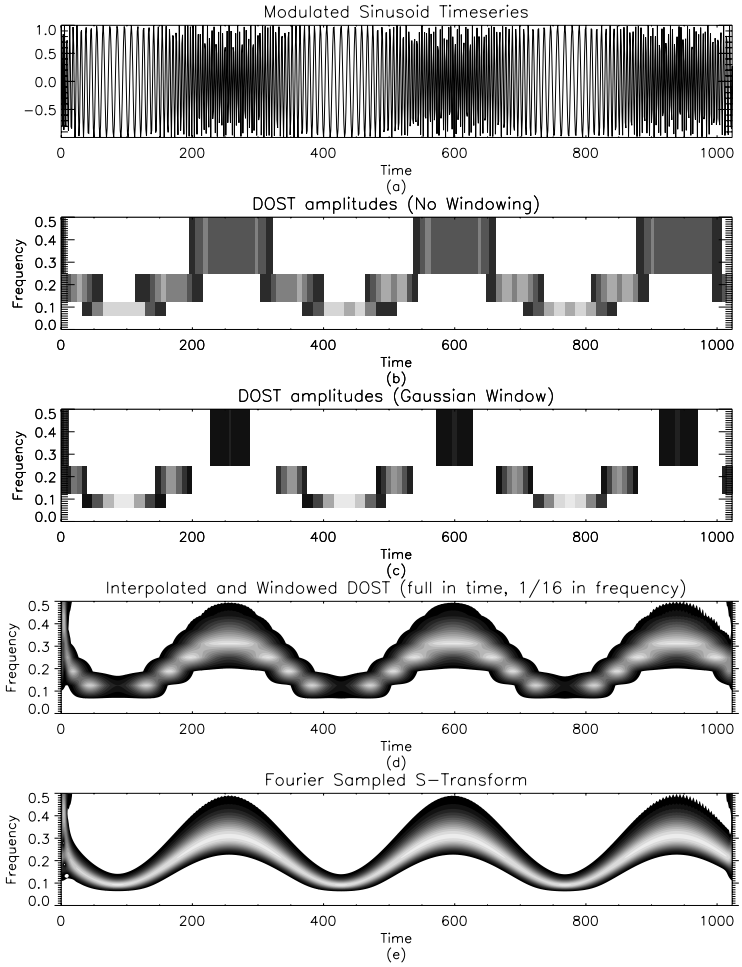

Figure 11 Illustration of the progression from the efficient DOST time-frequency representation (b) to the redundant S-transform representation (e) by increasingly interpolated (and windowed) DOSTs. Only positive frequencies and the amplitudes are shown. The time series is $h[t] = cos(2*\pi*(204+6*1024/(1+t)*cos(2*\pi*3*t/1024))*t/1024)$ where $t = 0, 1, 1023$

(centering the voice frequency ν) and the voice-windowing will differentiate different frequency ranges.

The DOST has the following properties:

- exact analytical definition of a basis for the S-transform.
- an orthogonal time-frequency transform from which the Discrete Fourier transform can be derived as a special case.
- an orthogonal time-frequency representation that collapses over the time variable to exactly give the Discrete Fourier transform spectrum.
- absolutely referenced phase, thus giving meaning to the phase of an orthonormal time-frequency representation.

- the ability to directly compare the phase of two time series in a "localized cross spectral analysis".
- the ability to employ a channel "instantaneous frequency" to each voice of the DOST.
- a general definition of a time-frequency representation to which one can apply any of the standard windows of power spectrum analysis in order to perform a localized power spectrum analysis.
- introduction of the framework to arbitrarily oversample the orthogonal time-frequency representation on all the way to the redundant S-transform, and thus showing how one can perform time-frequency analysis of almost arbitrarily large geophysical data sets with any desired sampling of the time-frequency space.
- The DOST can be extended in a straightforward method to higher dimensions for applications such as image processing, and volumetric data analysis, as has been done with the S-transform.

It should be noted that all properties of the ST described in section 3 are shared with the DOST. They are merely different yet complementary strategies for sampling and smoothing of the time-frequency space.

5 Wavelet Transform

The continuous Wavelet transform was introduced by Goupillaud, Grossmann, and Morlet in 1984 [11]. The basic philosophy is that any time series can be decomposed into a series of dilations and compressions of a mother wavelet denoted as $w(t)$. The advantage of this view is that high frequencies can be localized to a smaller time interval than low frequencies, which is a property it shares with the S-transform. The continuous wavelet transform (CWT) of $h(t)$ is given by [14]:

$$F(a,b) = \mid a \mid^{-\frac{1}{2}} \int_{-\infty}^{\infty} h(t) w^*(\frac{t-b}{a}) dt \tag{5.1}$$

where the asterisk denotes the complex conjugate. Strictly, the continuous wavelet transform gives a representation in the state space of the scale factor (or dilation) a and translation b, but with the appropriate choice of wavelet $w(t)$, it can be used to measure the power spectrum locally. Certain restrictions are placed on the function $w(t)$ in order to do a wavelet analysis. These are [66] (where $W(f)$ is the Fourier transform of $w(t)$):

- $w(t)$ must be absolutely and square integrable

$$\int_R \mid w(t) \mid^2 \frac{dt}{|t|} < \infty. \tag{5.2}$$

- $W(f)$ must be real.
- $W(f)$ must be small around $f = 0$.

These conditions ensure that the wavelet has finite energy and is band limited. The Morlet wavelet [67] is often encountered in the literature. The Morlet wavelet function consists of a complex exponential modulated by a Gaussian envelope:

$$\pi^{-1/4} s^{-1/2} \exp[ikx/s] \exp[-(x/s)^2/2] \tag{5.3}$$

where s is the wavelet scale, k is a non-dimensional parameter, and x is the position [68]. The IDL routine WV_CWT is used to calculate the CWT results in this manuscript.

The continuous wavelet transform only localizes the power spectrum as a function of time. It does not retain the absolutely referenced phase information that the S-transform contains and therefore is not directly invertible to the Fourier transform spectrum. Due to its frequency dependent normalization the CWT with Morlet wavelets diminishes the high frequency components of the time-frequency space.

6 A Comparison of the S-transform and the Continuous Wavelet Transform

What distinguishes the S-transform from the many time-frequency representations available is that the S-transform uniquely combines three characteristics: 1) progressive resolution, 2) absolutely referenced phase information, and 3) frequency invariant amplitude response. Daubechies has stated that progressive resolution gives a fundamentally more sound time-frequency representation [12].

The absolutely referenced phase of the S-transform is in contrast to a wavelet approach, where the phase of the wavelet transform is relative to the center (in time) of the analyzing wavelet. Thus as the wavelet translates, the reference point of the phase translates. This is called "locally referenced phase" to distinguish it from the phase properties of the S-transform.

From one point of view, local spectral analysis is a generalization of the global Fourier Spectrum. In fact, since no record of observations is infinite, all discrete spectral analysis ever performed on measured data have been local (i.e. restricted to the time of observation). Thus there must be a direct relationship between a local spectral representation and the global Fourier spectrum. This philosophy can be stated as the fundamental principle of S-transform analysis:

> The time average of the local spectral representation should result identically in the complex-valued global Fourier Spectrum.

$$h(t) = \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} S(\tau, f) d\tau \right\} e^{i2\pi f t} df \tag{6.1}$$

This shows that the S-transform can be thought of as a generalization of the Fourier transform to nonstationary timeseries. A similar claim can not be made for the CWT. This leads to phase and amplitude values of the local spectrum that are obvious and meaningful. Consider a signal $h(t) = Aexp(i2\pi f_o t + \phi)$. The Fourier transform spectrum of this signal at the frequency f_o would return the amplitude of A and the phase constant ϕ. In order to carry this understanding of phase into the realm of time-frequency representations and local spectra, it would require a transform that would return a voice (function of time) for the frequency f_o that had a constant amplitude A and a constant phase ϕ. Such is the case with the S-transform, and thus it can be described as a generalization of the Fourier transform to the case of nonstationary signals.

6.1 Frequency Sampling. The discrete Fourier transform has a very definite sampling of the frequencies, in order to be both complete and orthonormal (and thus provide a basis for the time series). The discrete ST has the identical sampling of the frequency space. It also has the same sampling of the time domain as the original time series. Hence there are N^2 points in the redundant ST representation.

Conversely, the CWT has a loosely defined sampling of the TFR domain (Figure 12). It is commonly employed with an octave scaling of the frequencies which

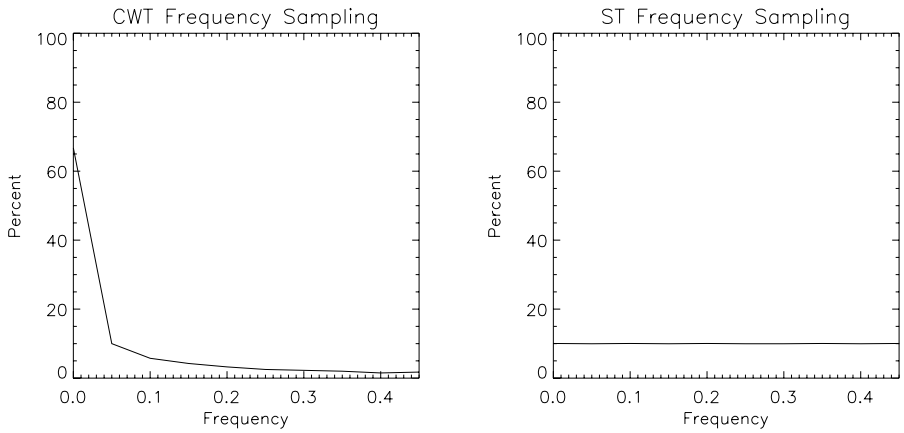

Figure 12 Frequency sampling histogram for the CWT and the ST (which is identical to the frequency sampling of the DFT). The ST is uniform in frequency, while the CWT oversamples the lower frequency region.

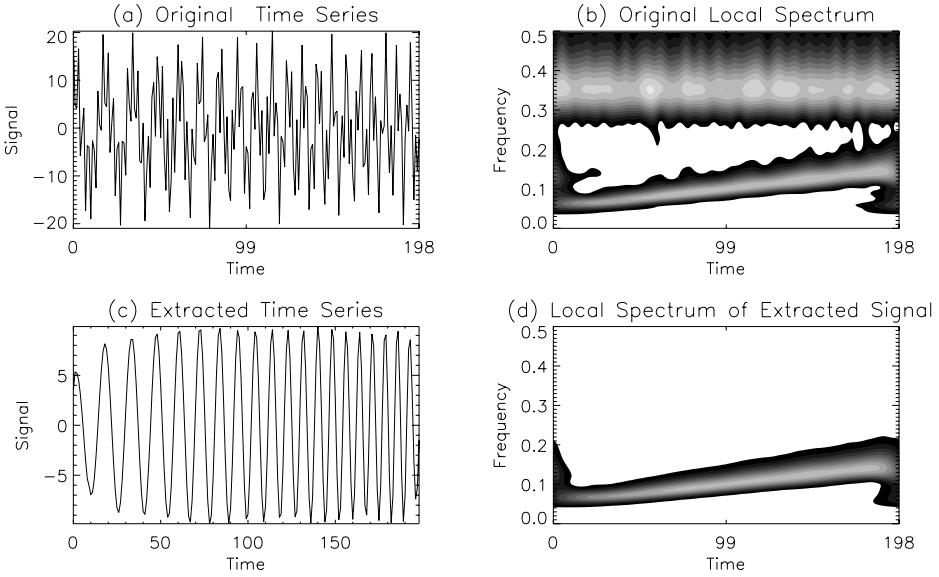

Figure 13 Illustration of the extraction of a signal from the ST.

provides an oversampled representation of the lowest frequencies, and undersamples (relative to the Fourier transform) the highest frequencies.

6.2 Direct Signal Extraction. Because of the characteristics of the S-transform one can read off the amplitude A(t) = abs(S(t,f)), the frequency f, and the phase ϕ(t) = atan(real(S(t,f))) , imaginary(S(t,f))) for each time step from the S-transform

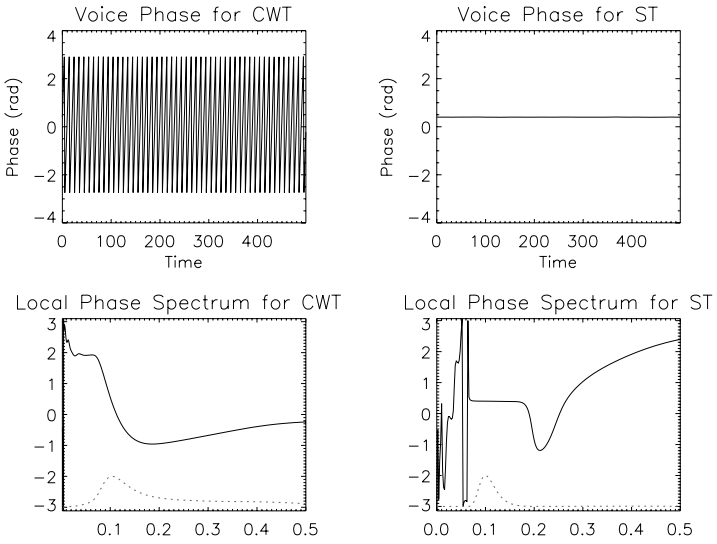

Figure 14 Illustration of the phase properties of the ST.

and directly extract that signal by reproducing it in the time domain as follows:

$$signal(t) = A(t)cos(2\pi f(t) * t + \phi(t)). \tag{6.2}$$

This direct extraction of a signal is due to the combination of absolutely referenced phase information and frequency invariant amplitude of the S-transform, and such direct extraction cannot be done with Wavelet methods. This shows that not only does the S-transform have a direct connection with the Fourier spectrum, it also has a direct connection with the original time series.

For the example shown here in Figure 13 a chirp signal is combined with a noisy high frequency signal. The nearest neighbour maxima for the chirp signal are chosen for extraction, and the resulting chirp signal is cleanly extracted (Figure 13c and d).

6.3 ST Phase vs CWT Phase. A peak in the ST is a region of stationary phase as can be seen by the definition for Generalized Instantaneous Frequency Eq. 3.38. The stationary phase of the S-transform can be seen in the phase voices of Figure 14. This figure shows a particular voice for the signal shown in Figure 3 for the mid frequency component. In the first plot, the phase of the CWT for that voice is shown, and it is rapidly varying through the full range of phase (from $-\pi$ to π). In the second plot, the phase of the ST for the mid frequency component is shown. It is constant, and is equal to $0.4 radians$ (which is the correct phase). In the third plot, the phase of local CWT spectrum is shown for time $t = 0$. The amplitude of the local CWT spectrum is overplotted as a dotted line. The phase at the peak is not equal to the phase of the signal. In the fourth plot, the phase of the S-transform local spectra is shown for time $t = 0$, and the amplitude local spectrum is overplotted as a dotted line. The phase varies quite a bit for very low frequencies, but phase is meaningless when the amplitude is negligible (as shown by the dotted line). For the region where the amplitude has a significant value, the

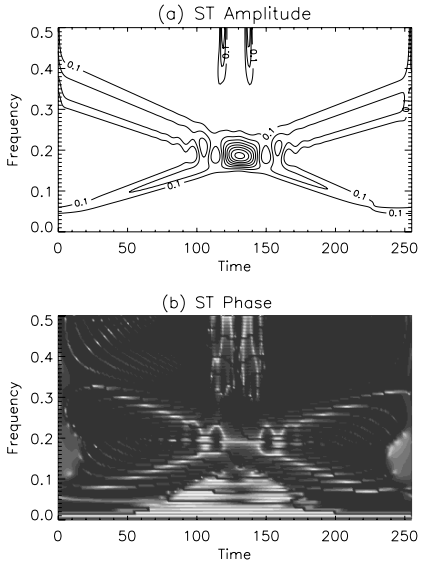

Figure 15 (a) the ST amplitude of a cross chirp signal. (b) The phase of the ST for the cross chirp signal. The chirp is apparent in the phase due to the stationary phase of the ST at local maxima. At a peak in the local spectra, the phase is constant.

phase of the ST local spectrum is constant and equal to 0.4 which is the correct value.

In Figure 15a the ST amplitude of a cross chirp signal is shown. In Figure 15b the phase of the ST for the cross chirp signal is shown. The outline of the chirp signal is readily apparent in this phase plot due to the stationary phase of the ST at local maxima. At a peak in the local spectra, the phase is constant.

6.4 ST Amplitude vs CWT Amplitude. In addition to the advantageous phase properties, the S-transform also has superior amplitude response. As mentioned above, the normalization of the S-transform is a vital distinction from the wavelet approach. The unit area localizing function (the Gaussian) preserves the amplitude response of the S-transform and ensures that the amplitude response of the ST is invariant to the frequency. In much the same way that the phase of the ST means the same as the phase of the Fourier transform, the amplitude of the ST means the same as the amplitude of the Fourier transform.

In Figure 16, a chirp function with an increasing frequency, and two high frequency bursts are shown. The amplitude is constant for all three of the components, and is equal to 10 units. Figure 16a shows the ST amplitude for this signal. The ST returns the correct amplitude for all three components, showing the high frequency bursts to have the same amplitude as the chirp signal. Figure 16b shows the CWT (with Morlet wavelets) amplitude for this signal. The amplitude is large for the lower frequencies, and diminishes as the frequency of the chirp increases. This is due to the normalization of the wavelet transform. It inherently diminishes the higher frequency components. The high frequency bursts, which have the same amplitude as the chirp, are relatively weak in the CWT representation.

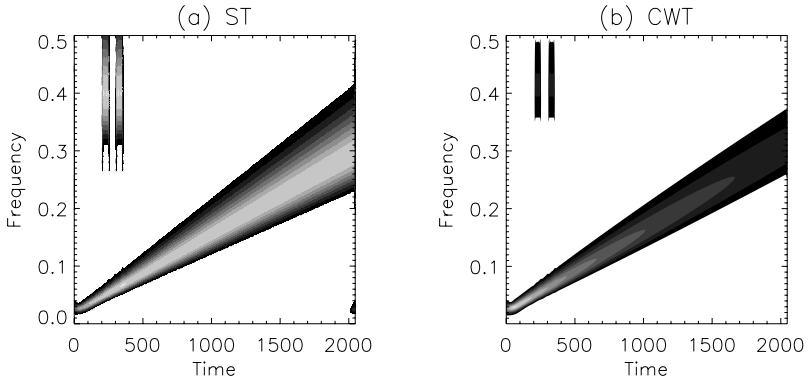

Figure 16 Illustration of the ST amplitude response vs the CWT response.

7 Summary

In summary the DOST and the S-transform approach has the following unique properties:

1. a direct connection to the Fourier transform through the inverse ST, and a direct connection to the time series through the direct signal extraction.
2. frequency invariant amplitude response.
3. absolutely referenced phase properties.
4. progressive resolution.

The ST uniquely combines the qualities these characteristics. It simultaneously estimates the local amplitude spectrum and the local phase spectrum, whereas a wavelet approach is only capable of probing the local amplitude/power spectrum. The ST fully represents the amplitude of all signals, in contrast to the CWT which attenuates high frequencies. It independently probes the positive frequency spectrum and the negative frequency spectrum, whereas many wavelet approaches are incapable of being applied to a complex time series. It is sampled at the discrete Fourier transform frequencies unlike the CWT where the sampling is arbitrary. Because of the absolutely referenced phase of the S-transform, it is possible to define a channel Instantaneous Frequency function for each voice [22]. It is possible to perform a localized cross spectral analysis as shown in [23, 24] where transient wavelike motions were detected in the time series measured by spaced receivers (physically separated by a known distance). By analyzing phase shifts in the S-transforms of the two time series, the phase speed of the wavelike motion was deduced. Such a straightforward analysis is not possible with a wavelet type approach.

8 - Acknowledgments

Supported by National Science Foundation Grant ATM-0350680. The author would like to thank Dennis Riggin for discussions on co-ST and quadrature ST.

References

[1] Aoki, T. *Calcul exponentiel des opérateurs microdifferentiels d'ordre infini.* I, Ann. Inst. Fourier (Grenoble) **33** (1983), 227–250.

[2] Brown, R. *On a conjecture of Dirichlet*, Amer. Math. Soc., Providence, RI., 1993.
[3] DeVore, R. A. *Approximation of functions*, Proc. Sympos. Appl. Math. vol. 36, Amer. Math. Soc., Providence, RI, 1986, pp. 34–56.
[4] M. Portnoff, Time-frequency representation of digital signals and systems based on short-time Fourier analysis, IEEE Transactions on Acoustics, Speech, and Signal Processing ASSP-28 (1) (1980) 55–69.
[5] D. Gabor, Theory of communication, J. Inst. Elect. Eng. 93 (3) (1946) 429–457.
[6] P. Bloomfield, Fourier Analysis of Time Series: An Introduction, Wiley. (1976)
[7] L. Cohen, Time-frequency distributions - A review, Proceedings of the IEEE 77 (7) (1989) 941–981.
[8] Y. Zhao, L. Atlas, R. Marks, The use of cone-shaped kernels for generalized time-frequency represenations of nonstationary signals, IEEE Transactions on Acoustics, Speech and Signal Processing 38 (1990) 1084–1091.
[9] H. Choi, W. Williams, Improved time-frequency representation of multicomponent signals using exponential kernels, IEEE Transactions on Acoustics, Speech and Signal Processing 37 (1989) 862–871.
[10] F. Hlawatsch, G. Boudreaux-Bartels, Linear and Quadratic Time Frequency Signal Representations, IEEE Signal Processing Magazine (1992) 21–67.
[11] P. Goupillaud, A. Grossmann, J. Morlet, Cycle-octave and related transforms in seismic analysis, Geoexploration 23 (1984) 85–102.
[12] I. Daubechies, The wavelet transform, time-frequency localization and signal analysis, IEEE Transactions on Information Theory 36 (5) (1990) 961–1005.
[13] R. Young, Wavelet Theory and its Applications, Kluwer Academic Publishers. (1993)
[14] O. Rioul, M. Vetterli, Wavelets and signal processing, IEEE Signal Processing Magazine 8 (1991) 14–38.
[15] S. Qian, D. Chen, Understanding the Nauture of Signals Whose Power Spectra Change with Time, IEEE Signal Processing Magazine March (1999) 51–67.
[16] J. Lina, Image Processing with Complex Daubechies Wavelet, Journal of Mathematical Imaging and Vision 7 (1996) 211–223.
[17] N. Kingsbury, Image processing with complex wavelets, Phil. Trans. Royal Society London A 357 (1760) (1997) 2543–2560.
[18] N. Kingsbury, The dual-tree complex wavelet transform: a new efficient tool for image restoration and enhancemen, Proc. EUSIPCO 98, Rhodes, Greece (1998) 319–322.
[19] N. Kingsbury, Shift Invariant Properties of the Dual-Tree Complex Wavelet Transform, Proceedings: International Conference on Acoustics, Speech and Signal Processing (ICASSP99) March. (1999)
[20] A. Teolis, Computational Signal Processing With Wavelets, Birkhauser. (1998)
[21] J. Lina, Complex Daubechies Wavelets: Filters Design and Applications, Proc. ISAAC Conf., Univ. of Delaware June (1997) 1–18.
[22] R. Stockwell, L. Mansinha, R. Lowe, Localization of the Complex Spectrum: The S Transform, IEEE Transactions on Signal Processing 44 (4) (1996) 998–1001.
[23] R. Stockwell, R. Lowe, Airglow imaging of gravity waves 1. Results from a small network of OH nightglow scanning imagers, Journal of Geophysical Research-Atmospheres 106 (D15) (2001) 17185–17203.
[24] R. Stockwell, R. Lowe, Airglow imaging of gravity waves 2. Critical Layer Filtering, Journal of Geophysical Research-Atmospheres 106 (D15) (2001) 17205–17220.
[25] H. Liu, R. Roble, A study of a self-generated stratospheric sudden warming and its mesospheric-lower thermospheric impacts using the coupled TIME-GCM/CCM3, Journal of Geophysical Research-Atmospheres 107 (D23) (2002) 4695.
[26] E. Merzlyakov, Y. Portnyagin, C. Jacobi, N. Mitchell, H. Muller, A. Manson, A. Fachrutdinova, W. Singer, P. Hoffmann, On the longitudinal structure of the transient day-to-day variation of the semidiurnal tide in the mid-latitude lower thermosphere - I. Winter season, Annales Geophysicae 19 (5) (2001) 545–562.
[27] D. Pancheva, E. Merzlyakov, N. Mitchell, Y. Portnyagin, A. Manson, C. Jacobi, C. Meek, Y. Luo, R. Clark, W. Hocking, J. MacDougall, H. Muller, D. Kurschner, G. Jones, R. Vincent, I. Reid, W. Singer, K. Igarashi, G. Fraser, A. Fahrutdinova, A. Stepanov, L. Poole, S. Malinga, B. Kashcheyev, A. Oleynikov, Global-scale tidal variability during the PSMOS campaign of

June-August 1999: interaction with planetary waves, Journal of Atmospheric and Solar-Terrestrial Physics 64 (17) (2002) 1865–1896.

[28] Y. Portnyagin, E. Merzlyakov, C. Jacobi, N. Mitchell, H. Muller, A. Manson, W. Singer, P. Hoffmann, A. Fachrutdinova, Some results of S-transform analysis of the transient planetary-scale wind oscillations in the lower thermosphere, Earth Planets and Space 51 (7-8) (1999) 711–717.

[29] D. Fritts, D. Riggin, B. Balsley, R. Stockwell, Recent results with an MF radar at McMurdo, Antarctica: Characteristics and variability of motions near 12-hour period in the mesosphere., Geophysical Research Letters 25 (3) (1998) 297–300.

[30] L. Mansinha, R. Stockwell, R. Lowe, M. Eramian, R. Schincariol, Local S-spectrum analysis of 1-D and 2-D data, Physics of the Earth and Planetary Interiors 103 (3-4) (1997) 329–336.

[31] J. Gao, W. Chen, Y. Li, F. Tian, Generalized S transform and seismic response analysis of thin interbeds, Chinese Journal of Geophysics-Chinese Edition 46 (4) (2003) 526–532.

[32] S. Theophanis, J. Queen, Color display of the localized spectrum, Geophysics 65 (4) (2000) 1330–1340.

[33] J. Osler, D. Chapman, Seismo-acoustic determination of the shear-wave speed of surficial clay and silt sediments on the Scotian shelf., Canadian Acoustics 24 (3) (1996) 11–22.

[34] P. Chu, A high resolution global sea surface temperature analysis and climatology, Tropical Oceans Global Atmosphere Notes 17 (1994) 5–7.

[35] P. Chu, The S-Transform for obtaining localized spectra., Journal of Marine Technological Society 29 (4) (1996) 28–38.

[36] P. Dash, B. Panigrahi, G. Panda, Power quality analysis using S-Transform, Ieee Transactions On Power Delivery 18 (2). (2003)

[37] Chilukuri, Multiresolution S-Transform based Fuzzy Recognition System for Power Quality Event, IEEE Transactions on Power Delivery 19 (1) (2003) 323–330.

[38] P. McFadden, J. Cook, L. Forster, Decomposition of gear vibration signals by the generalised S transform, Mechanical Systems and Signal Processing 13 (5) (1999) 691–707.

[39] P. McFadden, Detection of gear faults by decomposition of matched differences of vibration signals, Mechanical Systems and Signal Processing 14 (5) (2000) 805–817.

[40] C. Pinnegar, L. Mansinha, The bi-Gaussian S-transform, Siam Journal On Scientific Computing 24 (5) (2003) 1678–1692.

[41] C. Pinnegar, L. Mansinha, The S-transform with windows of arbitrary and varying shape, Geophysics 68 (1) (2003) 381–385.

[42] C. Pinnegar, L. Mansinha, Time-frequency localization with the Hartley S-transform, Signal Processing 84 (12) (2004) 2437 – 2442.

[43] C. Pinnegar, L. Mansinha, Time-local Fourier analysis with a scalable, phase-modulated analyzing function: the S-transform with a complex window, Signal Processing 84 (7) (2004) 1167 – 1176.

[44] S. Andino, R. Menendez, C. Lantz, O. Blank, C. Michel, T. Landis, Non-stationary distributed source approximation: An alternative to improve localization procedures, Human Brain Mapping 14 (2) (2001) 81–95.

[45] M. Varanini, G. Paolis, M. Emdin, A. Macerata, S. Pola, M. Cipriani, C. Marchesi, Spectral anaysis of cardiovascular time series by the S Transform., Computers in Cardiology 24 (1997) 383–386.

[46] B. Goodyear, H. Zhu, R. Brown, J. Mitchell, Removal of phase artifacts from fMRI data using a Stockwell transform filter improves brain activity detection, Magnetic Resonance in Medicine 51 (1) (2003) 16–21.

[47] S. Assous, A. Humeau, M. Tartas, P. Abraham, J. L'Huillier, Physiological effects of indomethacin and celecobix: an s-transform laser doppler flowmetry signal analysis, Phys. Med. Biol. 50 (2005) 1951–1959.

[48] X.-P. Zhang, M. Desai, Y. Peng, Orthogonal Complex Filter Banks and Wavelets: Some Properties and Design, IEEE Trans. on Signal Processing 47 (4) (1999) 1039–1048.

[49] I. Selesnick, L. Sendur, Iterated oversampled filter banks and wavelet frames, Proc. SPIE Wavelet Appl. in Signal and Image Proc. VIII (2000) 4119.

[50] D. Nelson, Cross-spectral methods for processing speech, J Acoust Soc Am. 110 (2001) 2575–2592.

[51] T. Gardner, M. Magnasco, Instantaneous frequency decomposition: An application to spectrally sparse sounds with fast frequency modulations, The Journal of the Acoustical Society of America 117 (5) (2005) 2896–2903.
[52] R. Bracewell, The Fourier Transform and Its Applications, McGraw-Hill Book Company. (1978)
[53] Y. Hayashi, Space-Time Spectral Analysis of Rotary Vector Series, Journal of the Atmospheric Sciences 36 (5) (1979) 757–766.
[54] R. Stockwell, W. G. Large, R. Milliff, Resonant inertial oscillations in moored buoy ocean surface winds, Tellus A 56 (5) (2004) 546–547.
[55] T.E. VanZandt. A universal spectrum of buoyancy waves in the atmosphere. *Geophysical Research Letters*, 9(5):575–578, 1982.
[56] E.M. Dewan, W. Pendleton, N. Grossbard, and P. Espy. Mesospheric oh airglow temperature-fluctuations - a spectral-analysis. *Geophysical Research Letters*, 19(6):597–600, 1992.
[57] D.N. Turnbull and R.P. Lowe. Temporal variations in the hydroxyl nightglow observed during aloha-90. *Geophysical Research Letters*, 18(7):1345–1348, 1991.
[58] E.M. Dewan and R.E. Good. Saturation and the "universal" spectrum for vertical profiles of horizontal scalar winds in the atmosphere. *Journal of Geophysical Research*, 91(D2):2742–2748, 1986.
[59] E.O. Brigham. The fast fourier transform. *Prentice-Hall Inc.*, 1974.
[60] D. Vakman. On the analytic signal, the teager-kaiser energy algorithm, and other methods for defining amplitude and frequency. *IEEE Transactions on Signal Processing*, 44(4):791–797, 1996.
[61] A. Papoulis. Signal analysis. *McGraw-Hill Book Company*, 1977.
[62] M. Eramian, R.A. Schincariol, R.G. Stockwell, R.P. Lowe, and L. Mansinha. Review of applications of 1-d and 2-d s transforms. *Wavelet Applications IV*, 3078, 1996.
[63] R.G. Stockwell, R.P. Lowe, and L. Mansinha. Localized cross spectral analysis with phase corrected wavelets. *Wavelet Applications III*, 2762:557–564, 1996.
[64] R.G. Stockwell, R.P. Lowe, and L. Mansinha. Instantaneous wavevector analysis. *Wavelet Applications IV*, 3078:349–358, 1996.
[65] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling. Numerical recipes in c: The art of scientific computing. *Cambridge University Press*, 1988.
[66] M. Farge. Wavelet transforms and their application to turbulence. *Annual Review of Fluid Mechanics*, 24:395–457, 1992.
[67] J. Morlet, G. Arens, E. Fourgeau, and D. Giard. Wave propagation and sampling theory 1. complex singal and scattering in multilayered media. *Geophysics*, 47(2):203–221, 1982.
[68] C. Torrence, and G.P. Compo A Practical Guide to Wavelet Analysis *Bull. Amer. Meteor. Soc.*, 79:61–78, 1998.
[69] R.G. Stockwell A basis for efficient representation of the S-transform. *Digital Signal Processing*, 17(1):371–393, 2007.
[70] A. Janssen Optimality property of the gaussian window spectrogram. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 39(1), 1991.