Wigner Distribution Representation and Analysis of Audio Signals: An Illustrated Tutorial Review*

DOUGLAS PREIS, AES Fellow

Department of Electrical Engineering and Computer Science, Tufts University, Medford, MA 02155, USA

AND

VOULA CHRIS GEORGOPOULOS**

School of Electrical Engineering and Computer Science, Ohio University, Athens, OH 45701, USA

The Wigner distribution provides a visual display of quantitative information about how a signal's energy is distributed in both time and frequency. Through its low-order moments the Wigner distribution embodies the fundamentally important concepts of both Fourier analysis and time-domain analysis. Signal energy is distributed in such a way that specific frequencies are localized in time by the group-delay time (from classical filter theory), and at specific instants in time the frequency is given by the instantaneous frequency (from classical modulation theory). The energy spectrum (energy per frequency) and the instantaneous power (energy per time) are specified by the zero-order moments of the distribution. The net positive volume of the Wigner distribution is numerically equal to the signal's total energy. While the theoretical underpinnings of the Wigner distribution are mathematically elegant and do merit in-depth study, a substantial amount of practical insight, understanding, and interpretive skill can be gained by carefully examining a wide variety of computed Wigner distributions such as those of the audio signals presented.

0 RETROSPECT

The Wigner distribution was "reintroduced" to the international digital signal processing community in 1980 at the IEEE L'Aquila Workshop by Theo Claasen and Wolfgang Mecklenbräuker [1], who shortly thereafter published a significant three-part paper on the subject [2]. While joint-domain representations like the Wigner distribution had been rather common in mathematical physics for some time [3], it was only in the early development of signal analysis that Ville [4] first alluded to time-frequency analysis in connection with his exposition of the analytic signal, and later Woodward [5] introduced the mathematically related complex ambiguity function.

There were two reasons for the broad appeal of the Wigner distribution. 1) It was shown to underlie and provide more resolution than the commonly used spectrogram, which was simply a smoothed version of the Wigner distribution, and 2) it was efficiently computed by using the fast Fourier transform (FFT) algorithm. The most controversial issue surrounding the Wigner distribution was the appearance of so-called "beats," which subsequently became known as interference terms. These terms were oscillatory and appeared because the distribution was based on the quadratic content, or energy, of the signal, but some workers found them to be seriously objectionable and arguably nonphysical because the distribution became negative, 1) implying negative energy when viewed as a distribution of energy in time and frequency and 2) not being able to be viewed as a distribution in the probabilistic sense, but able to exist where the signal had no spectral content. Although the Wigner distribution provided a fascinating link between the time domain and the frequency domain in signal analysis, its interference terms were seen by many as unwanted artifacts, which detracted significantly from its usefulness. The authors [1] correctly explained that while these cross terms must exist, they, on average, contain no signal energy.

There were, however, five undeniably attractive prop-
erties of the Wigner distribution of the analytic signal: its volume represented total signal energy, its time moments gave the energy spectral density and group delay, and its frequency moments gave instantaneous power (signal envelope squared) and instantaneous frequency. Because these moments are so useful in signal analysis, the Wigner distribution, or Wigner–Ville distribution as it became known, has continued to dominate for the past 18 years as the most popular time–frequency display (actually in the form of the spectrogram—a smoothed version of the Wigner–Ville distribution), and it has gained widespread use in fields as diverse as electrocardiogram analysis, acoustics, psychoacoustics, and speech research.

Currently the Wigner distribution is simply referred to as the Wigner transform. When applied to the complex analytic time signal, it is known to be particularly well suited to the analysis and detection of simply modulated signals such as amplitude modulation and frequency modulation, which are easily identified from envelope and instantaneous frequency moments. In addition the low-order time moments yield both the one-sided energy spectrum and group-delay time, the latter of which can be a good estimate of a signal's average time delay.

During the past 15 years an emerging trend in digital signal processing has been to use certain mathematical transformations, other than the Fourier transformation, in signal analysis. Signals such as, for example, those in medical diagnostics, seismic exploration, synthetic and natural speech, sonar, and radar require more analysis than a simple visual inspection of their time-domain waveforms. As is often the case, information about the phase of the signal or how the phase is altered by modulation or linear distortion plays an important role in revealing important underlying physical mechanisms such as modulation schemes, dispersion, or reflections. Researchers in digital signal processing have proposed, for example, using the Wigner transform, the ambiguity function, the Gabor transform, and various wavelet transforms for signal analysis. Because the Fourier transform maps, for each frequency, a complex frequency-modulated version of the signal that is averaged over the whole time interval to yield a single (complex) point in the spectrum and, second, because Fourier synthesis (inversion) used to recover or reconstruct the original signal often depends strongly on cancellation, different transformations have been sought that offer certain advantages in terms of multiresolution, time localization of frequency events, and more dependence on superposition rather than cancellation in signal synthesis. Many of these newer transformations seek to represent signals as sums of shifted (translated) and expanded (dilated) or even modulated versions (modulates) of simpler, elementary signals that occupy a short time interval. It is not that these newer signal transformations are intrinsically better, but they are thought to be more suited to specific applications. Since so much insight can be gained from a thorough examination and understanding of the various forms of phase alterations or phase changes within the general context of Fourier analysis, it is questionable how the very concept of phase manifests itself or whether phase is, in fact, lost in these newer signal transformations. One distinct advantage of the Wigner transform is the fact that it is firmly rooted in Fourier analysis. One of its moments does yield the Fourier spectral magnitude squared, and Fourier analysis phase is given indirectly in terms of its negative derivative with respect to frequency, the group-delay time. On the other hand, while the group-delay time and its dual variable, instantaneous frequency, are average quantities that are not, in general, the same as Fourier analysis time and Fourier analysis frequency, they do play an important role in identifying whether more than one frequency is present at a given instant in time or whether a certain frequency appears at different times.

Time and frequency are normally assumed to be independent variables in separate domains. The Wigner distribution joins these two domains and invites the signal analyst to inquire, for example, 1) at which instant in time a given frequency appears, or 2) what is the frequency of a signal at a given instant in time. The first-order moments of the Wigner distribution related to group-delay time and instantaneous frequency serve to answer these questions in an average sense. Indeed, it is remarkable that our very understanding of speech as well as our aesthetic impressions of music depend on our perceived answers to such questions.

1 DEFINITIONS

Instantaneous frequency is defined as [4]

\[ f_i(t) = \frac{1}{2\pi} \frac{d[\theta(t)]}{dt} \]  

(1)

where \( \theta(t) \) is the phase of the analytic signal,

\[ z(t) = |z(t)|e^{i\theta(t)} \]  

(2)

and \(|z(t)|\) is the envelope. The analytic signal of a real signal \( x(t) \) consists of the real signal as its real part and the Hilbert transform of the real signal \( \tilde{x}(t) \) as the imaginary part,

\[ z(t) = x(t) + j\tilde{x}(t) \]  

(3)

Group delay \( \tau_g \) is the delay of the envelope of the signal \( x(t) \) and is given by

\[ \tau_g = -\frac{1}{2\pi} \frac{d\phi(f)}{df} \]  

(4)

where \( \phi(f) \) is the phase of the Fourier transform of \( x(t) \).

Because the instantaneous frequency can, in practice, be difficult to estimate accurately, many different techniques have been proposed and developed [6]–[13]. A survey and unified treatment of various computational methods for estimating both instantaneous frequency and its dual variable, group-delay time, of discrete-time signals is given in [14], [15].
2 AUDIO SIGNALS

The Wigner distribution has been used extensively in audio signal analysis for spectral estimation [16], electrocardiogram analysis [17], characterization of auditory fibers [18], characterization of loudspeakers [19], and perception of phase distortion [20].

The Wigner distribution of a signal contains, in a simple way, the following four properties useful to the audio signal analyst: 1) frequency response, 2) group delay, 3) instantaneous power (the signal’s envelope squared), and 4) instantaneous frequency.

Each of these properties can be estimated visually by taking a “slice” of the elevation contours of the Wigner distribution parallel to the horizontal time axis or parallel to the vertical frequency axis in the time–frequency plane. The area under a horizontal slice at frequency \( f \) gives the numerical value of the frequency response (magnitude squared) at that frequency, whereas the center of gravity of that slice (the point at which all the area could be concentrated to produce the same moment about the vertical axis) gives the group delay time at that frequency. Similarly, the area under a vertical slice at time \( t \) gives the instantaneous power of the signal’s envelope, whereas the center of gravity of that slice equals the instantaneous frequency.

The representative signals examined here are 1) two Gaussian windowed tone bursts, 2) two Gaussian windowed chirps, 3) the impulse response of an antialiasing low-pass filter, 4) the impulse response of an all-pass filter (a filter that only alters the phase of signals applied to it), 5) a phase-shift-keyed (PSK) modulated signal, and 6) a frequency-shift-keyed (FSK) modulated signal.

2.1 Two Gaussian Windowed Tone Bursts (Fig. 1)

Two tone bursts, one at 700 Hz and the other at 1500 Hz, each 10 ms long, were windowed separately using a Gaussian window of a standard deviation of 1.2 ms. The resulting signal was 50 ms long with the 700-Hz burst centered at 15 ms and the 1500-Hz burst centered at 25 ms. The width and placement of the Gaussian windows were chosen so that the two tone bursts did not overlap in time. Fig. 1 shows the time sequence of the tone bursts (bottom), the energy spectrum (left), and the envelope squared (top), which is the instantaneous power of the signal. Because power equals energy per unit time, this curve is commonly referred to by audio engineers who assess loudspeaker performance as the ETC, or energy–time curve. It shows the arrival times of signal energy.

At right in Fig. 1 there are curves of group-delay time versus frequency and instantaneous frequency versus time. These two quantities are, mathematically, dual variables of one another and numerically equal the center of gravity of the Wigner distribution about the frequency axis and the time axis, respectively. In the center Fig. 1 shows elevation contour lines of the Wigner distribution itself as derived from the analytic signal associated with the time sequence. The analytic signal has zero Fourier spectrum for negative frequencies and is used as a mathematical means to define and extract the envelope and instantaneous frequency of the signal. The oscillatory

---

Fig. 1. Wigner–Ville distribution of two Gaussian windowed tone bursts along with its zero- and first-order moments: energy spectrum, envelope, group delay, and instantaneous frequency.

J. Audio Eng. Soc., Vol. 47, No. 12, 1999 December
terms between the signal terms in the Wigner–Ville distribution are the so-called interference terms of the Wigner–Ville distribution, which are alternatively positively and negatively valued and which contain, on average, zero energy.

2.2 Two Gaussian Windowed Chirps (Fig. 2)

Two chirps \( \cos(2\pi 700t + 50000t^2) \) and \( \cos(2\pi 1500t + 50000t^2) \) were windowed separately using a Gaussian window of a standard deviation of 2.25 ms. The resulting signal was 50 ms long with the 700-Hz chirp burst centered at 19 ms and the 1500-Hz chirp burst centered at 29 ms. Fig. 2 shows their time sequences, the FFT-based estimate of the energy spectral density, the computed envelope, the group delay and instantaneous frequency, and the Wigner–Ville distribution. In this case the signals overlap slightly in time but not significantly spectrally. The instantaneous frequency becomes oscillatory when the two signals overlap. Based on a cursory examination of the time sequence, envelope, or spectral density it is not possible to determine whether the signals are chirps or merely tone bursts. However, the instantaneous frequency, the group delay, and the orientation of the Wigner–Ville distribution clearly show the variation with time. This example illustrates important advantages of time–frequency analysis using the Wigner–Ville distribution.

2.3 Low-Pass Filter (Fig. 3)

A Wigner–Ville distribution corresponding to a 4-kHz low-pass filter (a speech-bandwidth antialias filter) is shown in Fig. 3. At left is the frequency response and below is the impulse response of the low-pass filter. On top is the instantaneous power of the envelope of the impulse response. To the right are curves of the group-delay time versus frequency and instantaneous frequency versus time. In this distribution a minor, but observable, artifact of the analytic signal used in obtaining the Wigner–Ville distribution is that the Wigner–Ville distribution as well as the energy–time curve are not causal, that is, they exist before the impulse response starts. The small, alternately positive and negative C-shaped interference terms to the right of the Wigner–Ville distribution produce the peaks and valleys in the energy–time curve but tend to cancel in the horizontal direction to yield a flat frequency response (spectral magnitude). The "ringing" in this filter's impulse response near and at 4 kHz is predicted in the upper part of the Wigner distribution by the large, horizontal positive energy concentration, which implies a damped, nearly sinusoidal oscillation in the time domain. The rising group-delay curve follows the significant values of the Wigner distribution. The instantaneous frequency rises quickly from zero, then oscillates in the region of the cutoff frequency of 4 kHz.

In qualitative terms, large concentrations in the Wigner distribution parallel to the frequency axis imply that the signal is "impulselike," whereas concentrations parallel to the time axis imply that the signal is "sinusoidalike" or oscillatory in time (see Fig. 3) [20]. From the standpoint of modulation theory, a horizontal Wigner distribution indicates "perfect" amplitude modulation or unmodulated carrier because at every instant in time there is only one frequency present (just the carrier.

Fig. 2. Wigner–Ville distribution of two Gaussian windowed chirps along with its zero- and first-order moments. Instantaneous frequency, group delay. Orientation of Wigner–Ville distribution shows variation of frequency with time.
wave). On the other hand, a vertical orientation of the Wigner distribution indicates "perfect" frequency modulation because every frequency is present at one instant in time. A change in the orientation of the Wigner distribution as time progresses can be interpreted as a conversion between frequency and amplitude modulation, for example, as Fig. 3 shows. In summary, the Wigner distribution in Fig. 3 is a graphic representation of the statement that the impulse response of a filter is, in general, an amplitude-modulated and frequency-modulated signal [21]. While these effects are always evident in the impulse response itself, they are dramatically illustrated by the orientation of the Wigner distribution.

2.4 All-Pass Filter (Fig. 4)

Fig. 4 (bottom) shows the impulse response of an all-pass filter (energy/frequency = constant) whose peak group delay equals 6.15 ms at 3.2 kHz. Examination of the impulse response waveform reveals both amplitude and frequency modulation effects. The energy–time curve (envelope power in decibels) shows the former whereas the absolute value of the instantaneous frequency indicates the latter. The envelope collapses several times, and each time this occurs, the absolute value of the instantaneous frequency simultaneously swings between 3.2 kHz and 0 kHz. The main, positive portion of the Wigner distribution has a shape similar to the group-delay curve. Between 0.1 and 2.0 ms alternating positive and negative (black and white) impulsive contributions between about 2.0 and 4.0 kHz in the Wigner distribution strongly modify the energy–time curve (envelope). The rightmost central portion of the Wigner distribution corresponds to the arrival of the "packet" of signal energy near the 6.15-ms group-delay time which, from the impulse response or instantaneous frequency, is an oscillation at approximately 3.2 kHz. The Wigner distribution contains rather small contributions that oscillate in time near the frequencies of 1.5 kHz and 5.3 kHz, which determine the "fine structure" of the envelope and instantaneous frequency.

This example and the previous one clearly show the interplay between the Wigner distribution and impulse response, frequency response, envelope power or energy–time curve, group delay, and instantaneous frequency [20].

2.5 PSK Signal (Fig. 5)

A PSK signal \(\cos(2\pi f_0 t + 2\pi i/5)\), where \(i = 1, 2, \ldots, 5\), was rectangularly windowed and sampled at 10 kHz. There are four phase shifts in the signal, each one lasting 0.01 s. Fig. 5 shows the time sequence, the FFT-based estimate of the energy spectral density, the computed envelope, the group delay and instantaneous frequency, and the Wigner–Ville distribution. The phase shifts are not all evident in the plot of the signal in the time domain. However, there are indications (impulsive) of the signal's abrupt phase shifts in the Wigner transform itself and also in the envelope and instantaneous frequency [22]. In addition, both the instantaneous frequency and the spectral magnitude reveal the frequency of the sinusoidal carrier.

2.6 FSK Signal (Fig. 6)

An FSK signal \(\cos(2\pi f_i t)\), where \(f_i\) takes on the frequency values 300, 1100, 700, and 1800 Hz sequentially...
for 0.01 s, was rectangularly windowed and sampled at 10 kHz. Fig. 6 shows the time sequence, the FFT-based estimate of the energy spectral density, the computed envelope, the group delay and instantaneous frequency, and the Wigner–Ville distribution. Inspection of the signal itself in the time domain reveals that the frequency varies with time, but it is unclear exactly where there are transitions from one frequency to the next. The energy spectral density shows the four frequencies present in the signal but, of course, does not indicate at which time

Fig. 4. Wigner–Ville distribution of all-pass filter with peak group delay of 6.15 ms at 3.2 kHz along with its zero- and first-order moments. The main, positive portion of the Wigner distribution has a shape similar to the group-delay curve.

Fig. 5. Wigner–Ville distribution of phase-shift-keyed (PSK) signal along with its zero- and first-order moments. Phase shifts are indicated by the presence of “impulses” in the Wigner–Ville distribution as well as in the envelope and instantaneous frequency.
they occur. The envelope of the signal is approximately 1 for all time, which is consistent with the presence of a sinusoidal signal, but there are "glitches" at the transition points from one frequency to the next. The higher, inner contours of the Wigner distribution (that is, those that appear like the letter I rotated 90°) indicate the frequency of the signal at given time instants. The importance of group delay and instantaneous frequency in localizing the signal in time and frequency is clearly seen [22].

3 NONNEGATIVE APPROXIMATE WIGNER DISTRIBUTIONS WITH VERY ACCURATE LOW-ORDER MOMENTS

The foregoing examples have illustrated the important role that the low-order moments, especially those related to group-delay time and instantaneous frequency, play in interpreting the Wigner–Ville distribution. As a tool for time–frequency analysis, the Wigner–Ville distribution has its shortcomings, however. Negative values of the distribution, suggesting negative-energy density, can exist. Oscillatory interference terms due to quadratic interaction can appear in regions of the time–frequency plane where the signal has no spectral energy. The Wigner–Ville distribution of causal signals has noncausal values due to the use of the analytic signal in the Wigner–Ville distribution calculation. Moment accuracy is impaired due to data windowing. Smoothing of the distribution not only reduces its resolution but also introduces further moment inaccuracies.

In the literature various techniques have appeared to reduce interference terms or result in a nonnegative distribution. However, in many instances both zero- and first-order moments of the distribution are not preserved and valuable information is lost.

Wigner showed that positive bilinear distributions satisfying the time and frequency marginals do not exist [3]. The spectrogram, for example, is manifestly positive but does not satisfy the marginals, whereas the Wigner–Ville distribution satisfies the marginals but is not manifestly positive. Claassen and Mecklenbräuker [2] have shown that the Wigner–Ville distribution for a chirp, and for that one signal only, is manifestly positive. Since our experience in everyday life is with positive energy densities [23], it is important to obtain time–frequency distributions that both are manifestly positive and satisfy the marginals. Cohen and Posch [24] have defined "proper" time–frequency distributions that satisfy positivity and the marginals and all of which can be parametrized by the function \( \rho(x, y) \). The lack of a single distribution being best for all applications has resulted in a proliferation of distributions, each corresponding to a different mapping from the signal plane to the time–frequency plane. Some of these techniques are mentioned here. It is important to emphasize that since these techniques do not in general yield good results for all classes of signals, the work described in the Appendix takes an alternative approach.

One approach is to formulate the solution as a minimum cross-entropy (MCE) optimization problem subject

---

**Fig. 6.** Wigner–Ville distribution of frequency-shift-keyed (FSK) signal along with its zero- and first-order moments. Contours of the Wigner distribution indicate the frequency of the signal at a given time instant, whereas group delay and instantaneous frequency show the localization of the signal in time and frequency.
to linear constraints, in particular, the marginal conditions [25]. In this method all the functions that satisfy the conditions are found, and then the one chosen among them is the one that maximizes the entropy, the reason being that the maximum entropy solution is the one that is unbiased. This technique was extended to satisfy an additional constraint that along a particular axis the distribution must be a specified function [26]. Another set of joint constraints has also been proposed in the entropy maximization approach, which are equivalent to the two classical marginals [27].

Reduced interference distributions (RID) have been developed by designing kernels to specifically accomplish that. For certain values of the kernels an RID approaches the Wigner distribution [28].

Another approach is to approximate positive time-frequency distributions (TFDs) through nonlinear combinations of spectrotgrams [29], where closed-form solutions for the combinations are obtained via optimization of entropy functions subject to an energy constraint. This method gives resuls superior to that achieved with individual spectrograms, and remarkably close to the positive TFDs obtained via computationally intensive methods.

Optimal-kernel time-frequency representations rely on the idea that for any given signal some TFRs are “better than others” and thus, with optimization-based kernel design, the goal is to extract the maximum possible time-frequency information. The optimal kernel may have a special binary structure, “1/0 kernel,” or it can be an additional smoothness constraint that forces the optimal kernel to be Gaussian along radial profiles [30]–[32].

Wavelets have been used to soft-threshold the Wigner–Ville distribution [33]. This technique is a potentially useful tool for Wigner–Ville spectrum estimation of unknown deterministic signals embedded in noise. For time-varying spectral analysis of noisy signals, non-linear smoothing of the Wigner–Ville distribution should offer higher performance than linear smoothing. However, this approach does not yield a distribution satisfying the marginals.

The least mean squared (LMS) adaptive algorithm has been used to simulate an exponentially smoothed pseudo Wigner–Ville distribution. This was achieved using LMS frequency-domain adaptive structures and the auto terms are separated from the cross terms in the Wigner–Ville distribution [34].

The concluding examples illustrate how the shape of the Wigner–Ville distribution shown for the low-pass filter of Fig. 3 would change if the hard constraints of nonnegativity, causality, and very high low-order moment accuracy are imposed [35]. The results shown in Figs. 7 and 8 are neither Wigner–Ville distributions nor are they unique time-frequency representations, but they do serve to illustrate the results of a simple and direct way to eliminate all of the objectionable artifacts of the Wigner–Ville distribution while preserving, or even improving, low-order moment accuracy. “Approximate” Wigner–Ville distributions such as those in Figs. 7 and 8 can be computed iteratively (see Appendix).

They are easier to interpret and may, in fact, be more useful in those applications where accuracy in the higher order moments can be compromised.

4 CONCLUSIONS

In its largest context, the problem of time–frequency analysis is to develop a two-dimensional display of one-dimensional data that can be computed efficiently, that clearly reveals those essential features to be extracted from the data, that has no distracting artifacts, and that is readily interpreted physically. This research area is important because it should provide a means to improve the human visualization and understanding of complicated data, and it could even lead to efficient, automated data or signal analysis by expert computer systems. The Wigner–Ville distribution is suited to the task because, within the framework of Fourier analysis, it combines the traditional concepts associated with the time and frequency domains into a single, so-called joint domain that contains and illustrates virtually all of the salient information about the signal in a very graphic way. Because rather prominent interference or cross terms continue to be viewed by some workers as distracting artifacts of the Wigner–Ville distribution, recent research efforts have focused on eliminating them by either modifying the distribution or seeking approximations to it.

Since the low-order moments of the Wigner–Ville distribution are so rich in information useful to the signal analyst, it is essential that they be preserved as accurately as possible when constructing any “approximate” Wigner–Ville distribution [36].

Perhaps the oldest and most common form of time–frequency analysis of audio signals is the musical score whose notes indicate the exact timing, duration, and pitch (frequency) of vocal or instrumental sounds.
To a musician, the score by itself is truly music that is seen but not heard. To the signal analyst, the visual displays of quantitative information in this paper, especially the various tones represented and analyzed in Figs. 1, 2, and 6, should, with a modest amount of study and interpretation, elicit very similar impressions related to the essential time–frequency properties of signals.

5 REFERENCES


**APPENDIX**

The Wigner distribution \( W(t, f) \) of the analytic signal \( z(t) \) is [2]

\[
W(t, f) = \int_{-\infty}^{\infty} z(t + \tau/2) z^*(t - \tau/2) e^{-j2\pi ft} d\tau.
\]

\( W(t, f) \) is usually interpreted as the signal’s energy density in time \( t \) and frequency \( f \). The double integral over time and frequency of the Wigner distribution (that is, its volume) is the total energy \( E \) of the analytic signal. Although \( E > 0 \), \( W(t, f) \) can assume positive, zero, or negative values. Here the goal is to generate a time-frequency distribution that is nonnegative but still yields the correct total energy and zero- and first-order moments of \( W(t, f) \) (envelope squared, energy spectral density, group delay, and instantaneous frequency).

**A.1 Iterative Algorithm**

The iterative algorithm used here is based on a modified version of the well-known least mean squared (LMS) algorithm developed by Widrow and Hoff in 1959 [37]. In the present application the input is the vector \( x \), and the desired output is the vector \( y \).

The update equation for the elements \( a_{ij} \) of matrix \( A \) is

\[
a_{ij}[k + 1] = a_{ij}[k] + 2c \frac{e^{ij}[k]}{\|x\|^2} e^{ij}[k] x^{ij}(6)
\]

where the error \( e^{ij}[k] \) is the difference between the \( j \)-th element of the desired output vector and the \( j \)-th element of the output vector of the algorithm on the \( k \)-th iteration, \( \|x\|^2 \) is the square of the norm of the vector \( x \) \( x^{ij} \) is the \( i \)-th element of vector \( x \), and \( c \) is a constant that controls stability and speed of convergence. A practical range for \( c \) is \( 0.05 < c < 0.5 \).

The LMS algorithm has been extended to allow adaptation of the \( a_{ij} \) by four vectors of errors, one generated for each moment. Also, a threshold-type nonlinearity is used to ensure nonnegativity.

**A.2 Iteration Steps**

An estimate of the matrix \( A \), whose elements represent time and frequency samples of the approximate Wigner distribution, can be obtained by using the following procedure.

1) Choose \( A[0] \), the initial estimate of the matrix \( A \). For example, an \( A[0] \) that satisfies exactly the total energy of the analytic signal and the zero-order moments of the Wigner distribution (square of signal envelope and energy spectral density), is given by [38]

\[
A[0] = \frac{y_s \times y_s^T}{E}
\]

where \( y_s \) is the energy spectral density vector of the analytic signal and \( y_s \) is the envelope squared vector. Other initializations [38] or the Wigner–Ville distribution itself can be used.

2) \( A[k] \), the estimate of \( A \) on the \( k \)-th iteration, is then used to generate four error vectors,

\[
e_{1}[k] = y_s - A[k] x_1
\]

\[
e_{2}[k] = y_s - A[k] x_2
\]

\[
e_{3}[k] = y_s - A[k] x_3
\]

where \( y_s \) is the instantaneous frequency vector multiplied by the envelope squared vector element by element, \( y_s \) is the group-delay vector multiplied by the energy spectral density vector element by element, and \( x_1, x_2, x_3 \) are appropriate weighting vectors used to produce the moments.

3) Based on the four error vectors, an error matrix \( e \) is defined as

\[
e[k] = e_{1}[k] \frac{x_{1}^T}{\|x_{1}\|^2} + e_{2}[k] \frac{x_{2}^T}{\|x_{2}\|^2} e_{3}[k] + e_{4}[k] \frac{x_{3}^T}{\|x_{3}\|^2} + e_{4}[k] \frac{x_{3}^T}{\|x_{3}\|^2}
\]

(12)
and the extended LMS update rule becomes
\[ a_i[k + 1] = a_i[k] + 2c e_i[k] . \] (13)

4) An additional step for this particular modification of the extended LMS algorithm is to pass each update through a threshold-type nonlinearity (with the threshold set at zero). This is necessary to ensure that the elements of matrix \( A[k] \) are nonnegative. Therefore the usual update rule is modified to become
\[ a_i[k + 1] = S[a_i[k] + 2c e_i[k]] \] (14)
where \( S(u) = u \) for \( u \geq 0 \) and \( S(u) = 0 \) otherwise. When viewed as part of a neural network, the neurons contain a threshold-type nonlinearity with the threshold set at zero.

5) Renormalize the matrix \( A[k] \) to ensure that the total energy of the distribution is exactly \( E \).

6) The algorithm then repeats steps 2–6.

---

**THE AUTHORS**

**D. Preis**

Douglas Preis was born in Chicago, IL. He received the B.S.E.E. and M.S.E.E degrees from the University of Santa Clara, CA, where he was teaching assistant, and the Ph.D. from Utah State University, Logan, as a NASA postdoctoral fellow. During the following 8 years he was postdoctoral research fellow in applied physics on the faculty of the Division of Engineering and Applied Physics at Harvard University, Cambridge, MA.

In 1978 he joined the faculty of Tufts University, Medford, MA, in the College of Engineering as assistant professor of electrical engineering and rose to the rank of professor in 1989, his current position in the Department of Electrical Engineering and Computer Science.

In 1982 he was Mellon Grant recipient at Tufts University. In 1984 he was Deutsche Forschungsgemeinschaft Professor at Ruhr Universit"at, Bochum, Germany. In 1985 he was Fulbright Professor at Technische Universität Wien, Vienna, Austria. In 1991 he was Visiting Scholar in Applied Physics, at Harvard University.

Professor Preis has published or presented more than 150 original technical papers and reports. He won the Audio Engineering Society’s Publication Award in 1978 and in 1985. He served for 5 years as associate editor of the IEEE Transaction on Acoustics, Speech and Signal Processing, and has been a member of the Review Board of the Journal of the Audio Engineering Society for more than 20 years.

He is an eminent engineer of Tau Beta Pi, a national engineering honor society, a member of both Eta Kappa Nu, a national electrical engineering society, and Sigma Xi, a national research society. In 1997 he was elected fellow of the AES for his research contributions and received the award in 1998.

**V. C. Georgopoulos**

Voula Georgopoulos was born in Lowell, MA. She holds a Diploma in electrical engineering from Democritus University of Thrace, Greece (1984), an SM in EECS (1987) and an Engineer’s Degree from the Massachusetts Institute of Technology (1988), and a Ph.D. in electrical engineering from Tufts University (1992). She received a Vinton Hayes Fellowship in communications during the academic year 1984–1985 at MIT and between 1985–1988 was a research and teaching assistant at MIT. In May 1988 she joined the MITRE Corporation, Bedford, MA, as a member of technical staff, where she worked in the Optical Communications and Photonics Specialty Group. Between April 1994 and December 1995, she worked as a research scientist in the Electrical Engineering Department of the University of Patras in Greece. In January 1996 she joined the faculty of the School of Electrical Engineering and Computer Science of Ohio University, as an assistant professor. As of September 1998 she is professor of informatics and associate department head of the Department of Speech Therapy at the Technological Educational Institute of Patras in Greece. Her current research interests are audio and speech signal analysis, perceptual modeling, multimedia speech therapy systems, neural networks, and decision-making systems.

Dr. Georgopoulos has published more than 40 publications in journals and conference proceedings in addition to 10 technical reports. She is a member of Tau Beta Pi and Sigma Xi.