Copyright 2007 Audio Engineering Society, Inc.

Wigner Distribution Analysis of Filters with Perceptible Phase Distortion*

D. PREIS** AND F. HLAWATSCH

Institut für Nachrichtentechnik, Technische Universität Wien, Vienna, Austria

AND

P. J. BLOOM AND J. A. DEER

Division of Engineering, Polytechnic of Central London, London, UK

The responses of several anti-alias and all-pass filters are displayed jointly in time and frequency using the Wigner distribution. Characteristic features of the Wigner distribution associated with perceived phase distortion are identified.

0 INTRODUCTION

The Wigner distribution [1] is a function of time t and frequency f that can be computed from the impulse response $h(t)$ or the frequency response $H(f)$ of a linear time-invariant system. It is a joint-domain or time-frequency representation of $h(t)$ or $H(f)$ and, as such, provides different insights about the system response. The commonly used engineering measures of signal and system properties, such as, for example, group delay, instantaneous frequency and power, and spectral energy density, as well as the other bilinear time–frequency representations themselves, can be derived from the Wigner distribution by appropriate averaging in time or frequency or by introducing suitable weighting functions [1].

While the Wigner distribution is a powerful tool in signal analysis that shows how the signal is distributed or spread in both time and frequency, its relevance to the human perception of sound has not yet been established. Sound perception does include some form of time–frequency analysis, however [2], [3]. In this connection it is worthwhile to examine the Wigner distributions of a variety of different signals that have been perceived as sounding identical as well as those sounding different. Of particular interest here are specific features of the Wigner distribution present when a difference is heard and this difference is due to the presence of phase distortion.

Two recent papers [4], [5] on the perception of phase distortion in anti-alias and all-pass filters provide information about the threshold of detection of group-delay distortion in frequency ranges around 2, 4, and 15 kHz. Group-delay distortion occurs when the phase shift of the filter does not decrease linearly with increasing frequency. In frequency ranges where the phase is not a linear function of frequency its (negative) derivative with respect to frequency, the group delay, deviates from a constant value. This deviation is called group-delay distortion. The impulse responses of filters with group-delay distortion were compared by auditors to those of filters with identical frequency response (magnitude) but without group-delay distortion (linear-phase filters). Results, for diotic presentation (same signal both ears) using earphones, for two-alternative forced-choice A–B experiments, indicated that a statistically significant perceptual threshold was reached when the peak value of group-delay distortion was approximately 1.5–2.0 ms in the midfrequency (2-kHz to 4-kHz) range, depending also on the bandwidth of delay distortion, and about 4.5 ms at high frequencies (15 kHz). In terms of the impulse-response duration of signals with identical spectral magnitude, a difference was usually perceived when one impulse response in the test pair lasted nearly 2 ms longer than the other,

* Presented at the 80th Convention of the Audio Engineering Society, Montreux, Switzerland, 1986 March 4–7; revised 1987 August 12.

** Permanent address: Department of Electrical Engineering, Tufts University, Medford, MA 02155, USA.

but when the durations were similar, an audible difference also could be perceived if the envelopes of the impulse responses differed significantly.

Wigner distributions related to linear transmission properties of loudspeaker systems are available in [6] and [7], and these two references include more mathematical details than the elementary introduction to the Wigner distribution included in [1] where reference to the important earlier publications can be found.

1 WIGNER DISTRIBUTIONS

Impulse responses of a variety of low-pass, anti-alias, and all-pass filters were selected from the perceptual experiments reported in [4] and [5], where more quantitative information about these filters is available. Wigner distributions can be computed directly from the impulse responses or from the so-called analytic signal associated with the impulse response (formed by mathematically subtracting the negative frequencies from the Fourier transform of the impulse response). Both forms of the Wigner distribution are shown for the 4-kHz low-pass and the 15-kHz anti-alias filters in Figs. 1–6. The minimum-phase versions (with phase distortion) appear on the left side and their linear-phase counterparts are on the right side in each of these figures. In the captions it is noted whether an audible difference existed between the original impulse responses whose Wigner distributions are on the left and right sides of each figure. Figs. 7 and 8 show Wigner distributions derived from the analytic signals of five all-pass filters. The impulse responses of these all-pass filters were compared by auditors to a pure impulse whose Wigner distribution (not shown) would be simply a vertical line at $t = 0$ along the frequency axis. Again, perceptual results are noted in the figure captions.

The Wigner distributions shown in the figures were computed and are displayed in the following manner. The pseudo-Wigner distribution was evaluated numerically from the impulse response using a 3-ms Hamming window. The nonnegative values were normalized to the range from zero to one, then these data points were compressed by a square-root mapping of the range onto itself. Ten elevation contour lines (iso lines) in the compressed range from 0.05 to 0.9 were computed and plotted in the time–frequency plane. Negative values of the Wigner distribution are not shown but do exist in between the sets of positive contour lines, however.

For the low-pass and anti-alias filters in Figs. 1–6 each individual Wigner distribution consists of a region of large signal concentration, C-shaped for minimum-

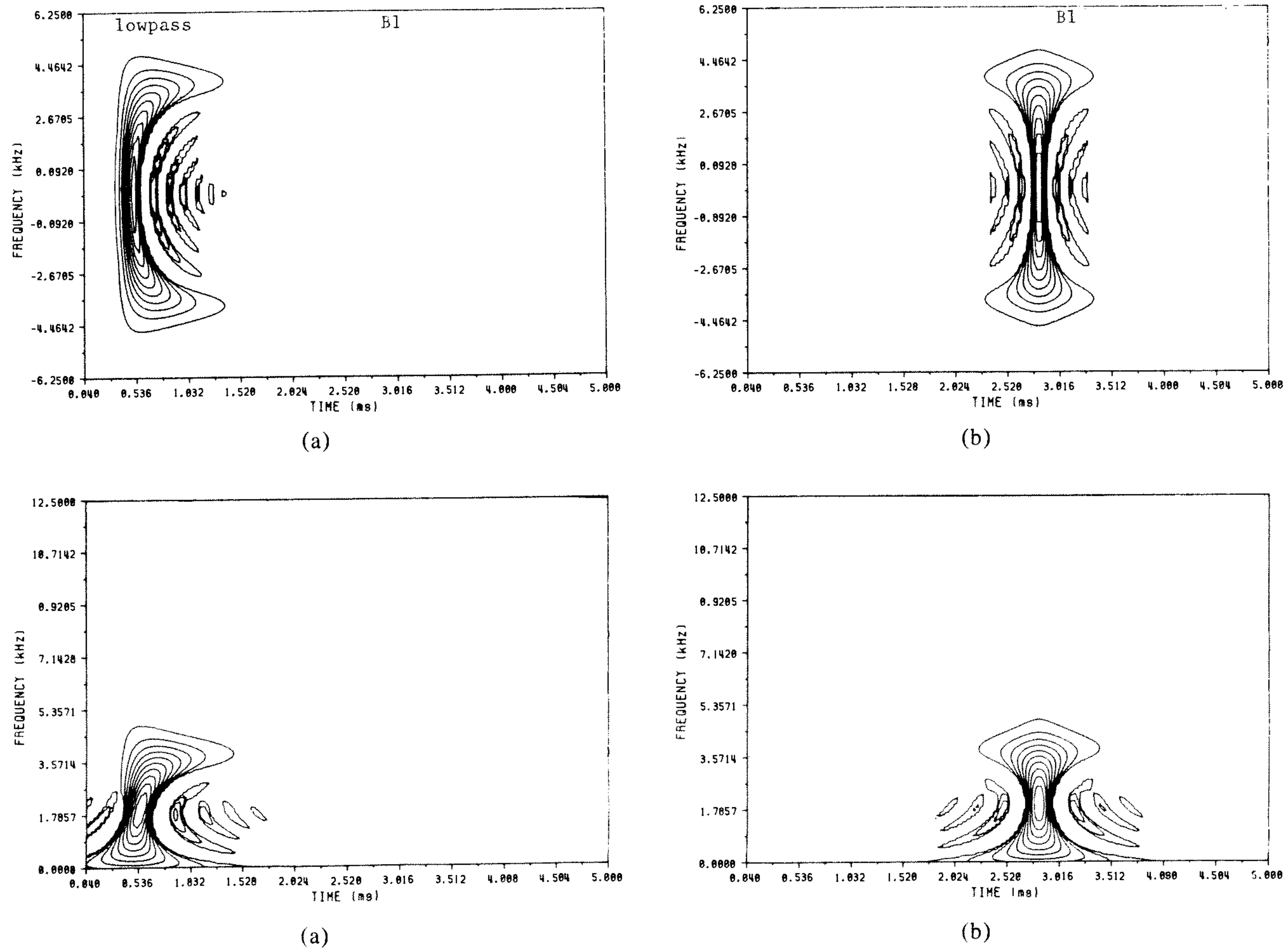


Fig. 1. Wigner distributions of one pair of 4-kHz eighth-order Butterworth low-pass filters, that is, two cascaded filters. (a) Minimum-phase version from impulse response (above) and from analytic signal (below). (b) Linear-phase version from impulse response (above) and from analytic signal (below). *No* Audible difference.

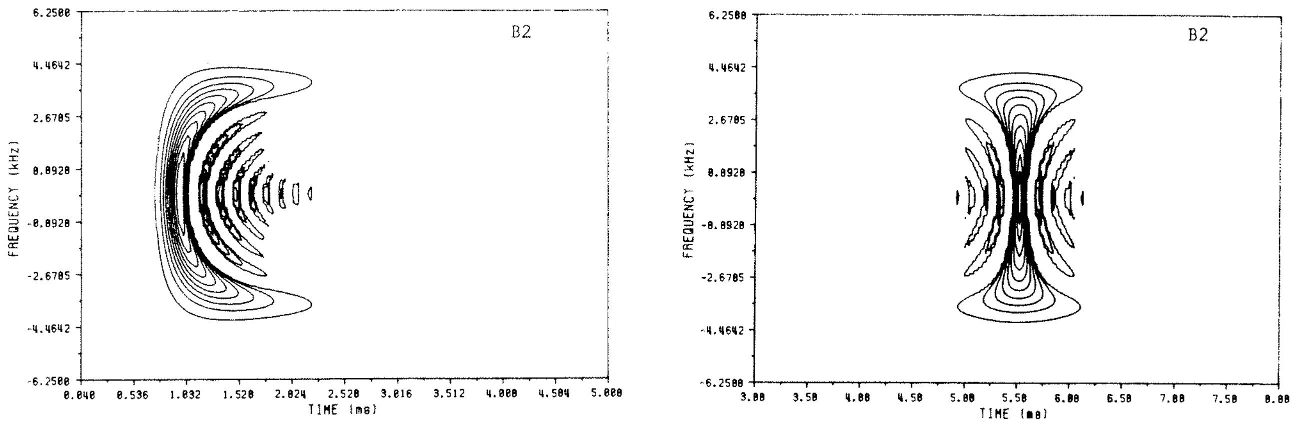
phase filters and I-shaped for linear-phase filters. The smaller cusplike, either)-shaped or (-shaped, regions are called interference terms. They are both positive and negative (only positive shown), and arise from the "quadratic" interaction of the main C- or I-shaped signal concentration.

The effect of band-limiting on the signals' spread in time and frequency is quite different for minimum-phase and linear-phase filters, even though the frequency response (magnitude) for any given pair of filters is identical. It is remarkable that the group-delay curve shape, for the minimum-phase filters, closely resembles the main C-shaped signal concentration. For the linear-phase filters the group delay is just a vertical line that bisects the distribution. The instantaneous power (energy–time) of the envelope of the analytic signal at time t_1 numerically equals the integral with respect to frequency f of the Wigner distribution along the vertical line $t = t_1$. Group delay at frequency f_1 is the "center of gravity" (normalized first moment) with respect to time t of the Wigner distribution along the horizontal line $f = f_1$. The spectral energy density at frequency f_1 is evaluated by integrating the Wigner distribution with respect to time along the line $f = f_1$.

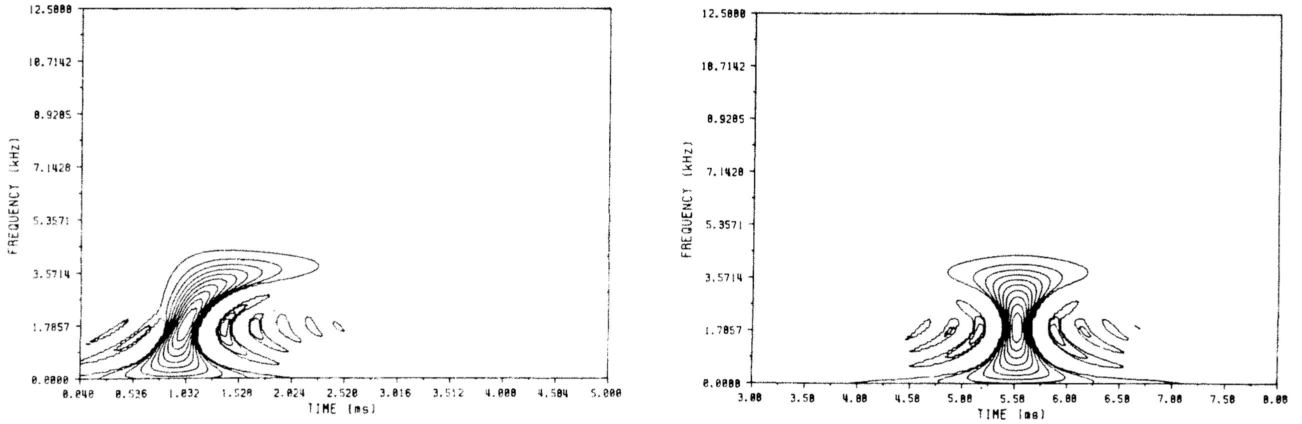
The interpretation of Wigner distributions for the all-pass filters shown in Figs. 7 and 8 is more difficult but quite interesting. The center frequency where the group delay is maximum is 2 kHz. The peak group delay for Fig. 7 is 2 ms and that for Fig. 8 is 4 ms. If the oscillating interference terms above and below the horizontal central main lobe are neglected and, further, the much stronger interference terms (alternating white and black regions) to the left and behind the main lobe, it is seen that the remaining large positive values of the Wigner distribution closely trace a curve that is very much like the group delay. This is especially clear for filters 2A in Fig. 7 and 4A in Fig. 8. The strong interference terms (positive and negative) behind the main lobe will significantly influence the shape of the envelope of the analytic signal, however.

2 CONCLUSIONS

The Wigner distribution is potentially significant in audio engineering because it is the basis for time–frequency analysis. Commonly used frequency-domain and time-domain measures of performance are easily derived from the Wigner distribution. Therefore their value is now further enhanced because they are all interrelated in a new and interesting way. Clearly, the



(a) (b)



(a) (b)

Fig. 2. Wigner distributions of two pairs of 4-kHz Butterworth low-pass filters (arranged as in Fig. 1). *Slight* audible difference.

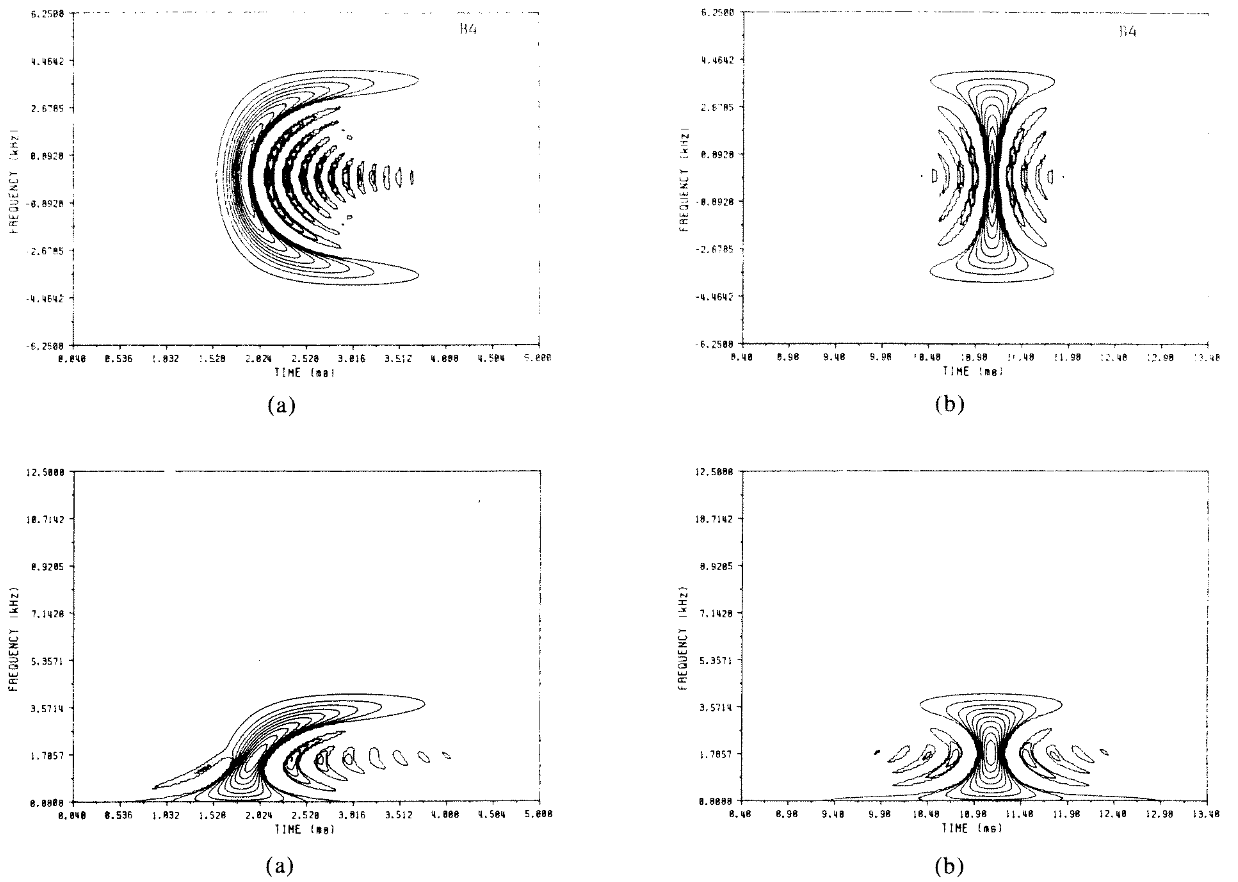


Fig. 3. Wigner distributions of four pairs of 4-kHz Butterworth low-pass filters (arranged as in Fig. 1). *Moderate* audible difference.

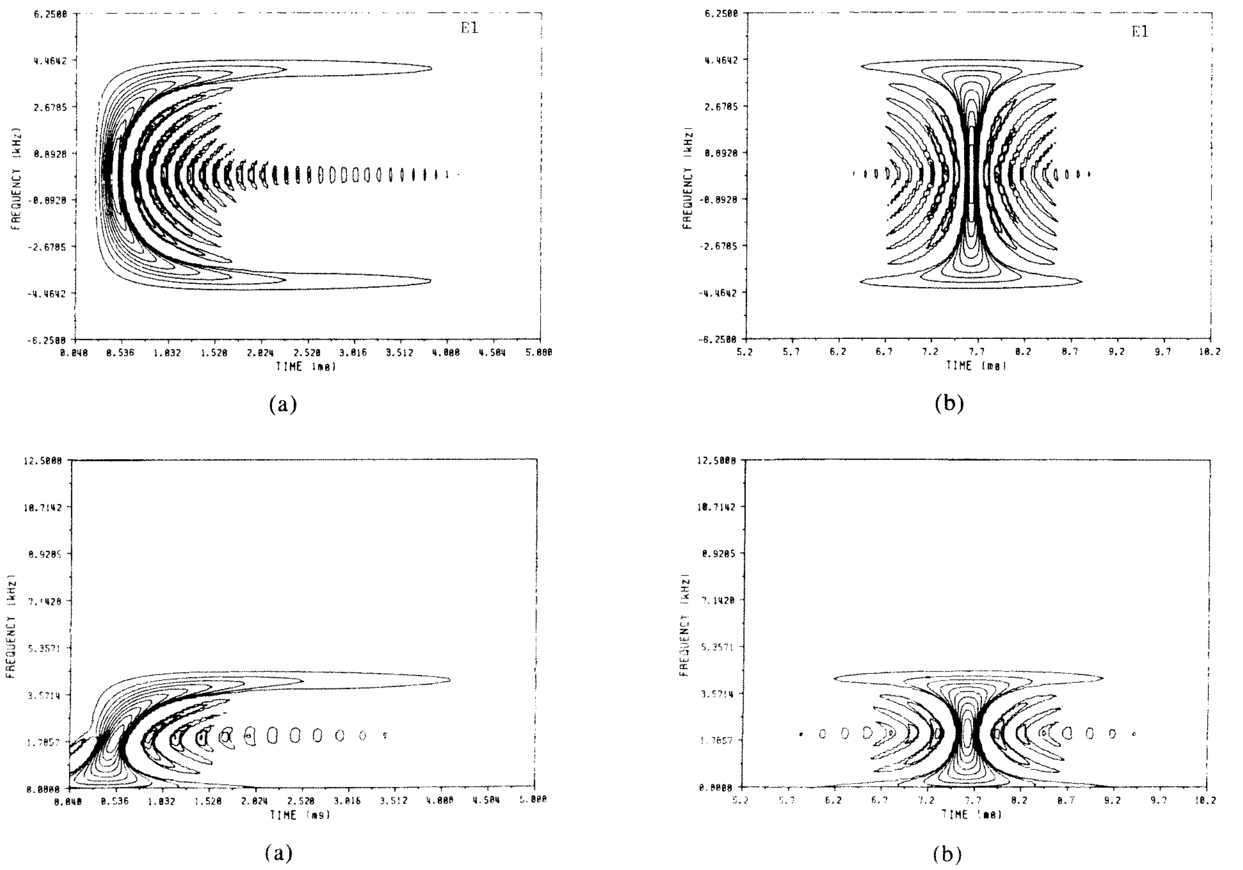


Fig. 4. Wigner distributions of one pair of 4-kHz seventh-order elliptic low-pass filters (arranged as in Fig. 1). *Strong* audible difference.

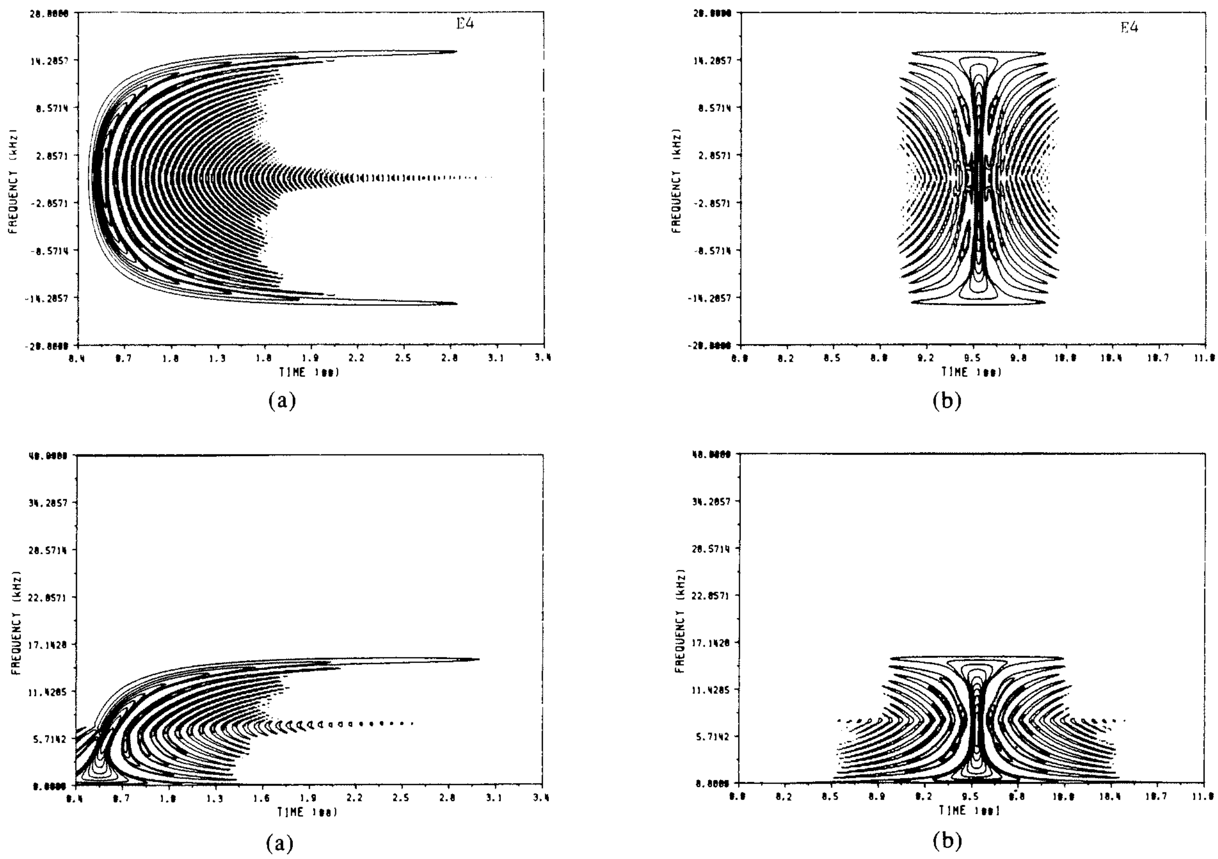


Fig. 5. Wigner distributions of four pairs of 15-kHz seventh-order elliptic anti-alias filters (arranged as in Fig. 1). *No* audible difference.

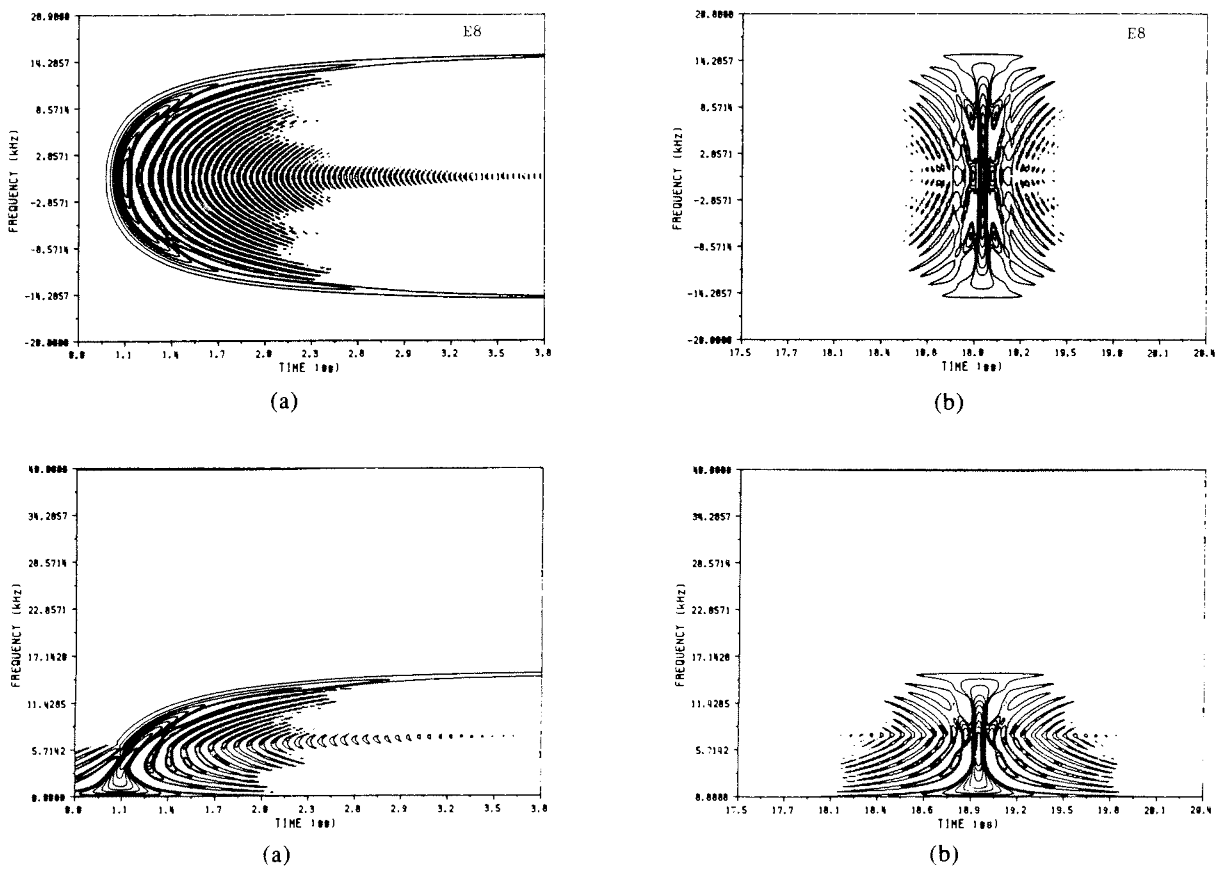


Fig. 6. Wigner distributions of eight pairs of 15-kHz anti-alias filters (arranged as in Fig. 1). *Moderate* audible difference.

Wigner distribution provides considerable new insight into and understanding of the mechanisms of linear distortions. The issue presently at hand is to assess its perceptual relevance. The 29 Wigner distributions included in this communication show, by comparison, its perceptual relevance in connection with the form of phase distortion called group-delay distortion.

3 ACKNOWLEDGMENT

This work was supported by a Fulbright grant from the Austrian-American Fulbright Commission.

4 REFERENCES

[1] D. Preis, "Phase Distortion and Phase Equalization in Audio Signal Processing—A Tutorial Review," *J. Audio Eng. Soc.*, vol. 30, pp. 774–794 (1982 Nov.).

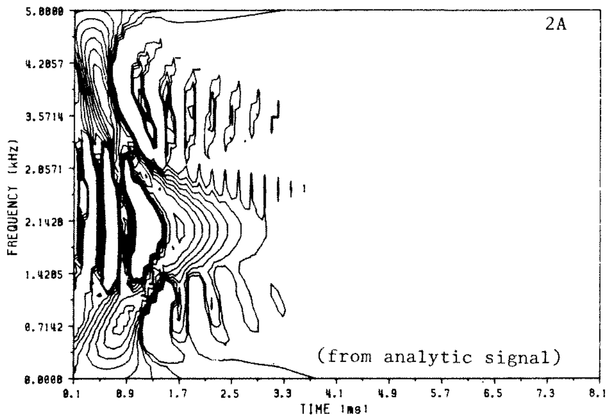
[2] J. O. Pickles, *An Introduction to the Physiology of Hearing* (Academic Press, New York, 1982).

[3] B. C. J. Moore, *An Introduction to the Psychology of Hearing* (Academic Press, New York, 1982).

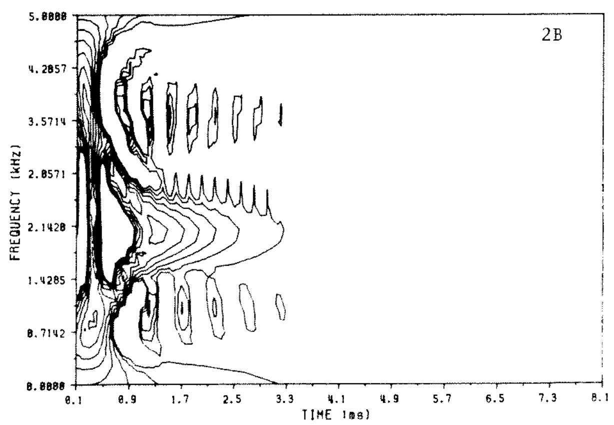
[4] D. Preis and P. J. Bloom, "Perception of Phase Distortion in Anti-Alias Filters," *J. Audio Eng. Soc.*, vol. 32, pp. 842–848 (1984 Nov.).

[5] J. A. Deer, P. J. Bloom, and D. Preis, "Perception of Phase Distortion in All-Pass Filters," *J. Audio Eng. Soc.*, vol. 33, pp. 782–786 (1985 Oct.).

[6] C. P. Janse and A. J. M. Kaizer, "Time-Frequency Distributions of Loudspeakers: The Application of the Wigner Distribution," *J. Audio Eng. Soc.*, vol. 31, pp. 198–223 (1983 Apr.).

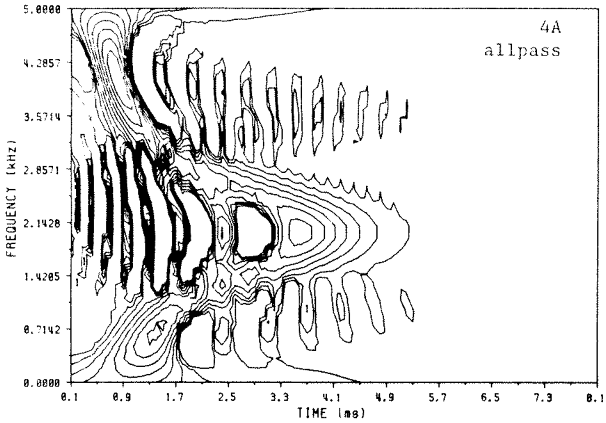


(a)

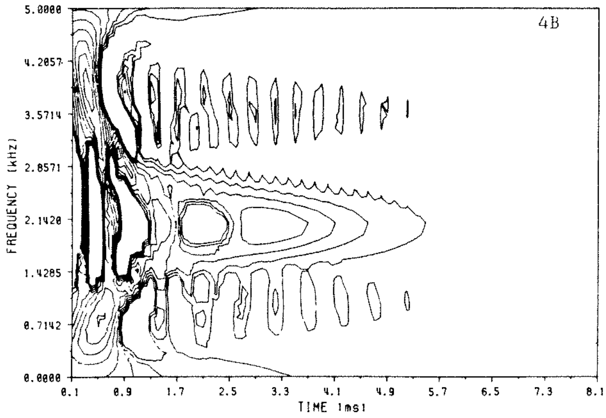


(b)

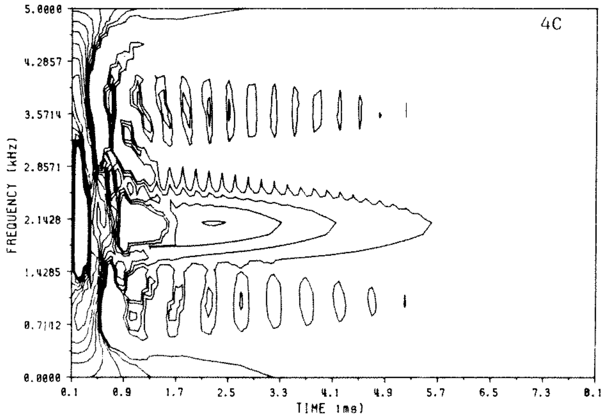
Fig. 7. Wigner distributions of two pairs of second-order all-pass filters (2-kHz center frequency) having a total of 2-ms peak group delay. (a) Broad delay bandwidth, 2 A. (b) Medium delay bandwidth, 2B. *Slight to moderate* audible difference when compared to a pure impulse.



(a)



(b)



(c)

Fig. 8. Wigner distributions of four pairs of second-order all-pass filters (2-kHz center frequency) having a total of 4 ms peak group delay. (a) Broad delay bandwidth, 4A. (b) Medium delay bandwidth, 4B. (c) Narrow delay bandwidth, 4C. *Strong* audible difference when compared to a pure impulse. No audible difference when 4A and 4B or 4B and 4C are compared. Strong audible difference between 4A and 4C.

[7] C. P. Janse and A. J. M. Kaizer, "The Wigner Distribution: A Valuable Tool for Investigating Transient Distortion," *J. Audio Eng. Soc.*, vol. 32, pp. 868–882 (1984 Nov.).

APPENDIX
INTERPRETING AND UNDERSTANDING THE WIGNER DISTRIBUTION

The Wigner distribution of a signal contains, in a simple way, the following four properties useful to audio engineers:

1) Frequency response
2) Group delay
3) Instantaneous power
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_1 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_1 gives the instantaneous power of the signal's envelope, whereas the center of gravity of that slice equals the instantaneous frequency.

As an aid for interpreting the eight figures included in this communication, Wigner distributions corresponding to a 4-kHz low-pass filter (see Fig. 3) and a 2-kHz all-pass filter are discussed here in more detail. Figs. 9 and 10 show the interplay between the Wigner distribution and impulse response, frequency response, envelope power or energy–time curve, group delay, and instantaneous frequency.

In Fig. 9 to the left is the frequency response and below is the impulse response of a 4-kHz low-pass filter. Above is the instantaneous power of the envelope of the impulse response. Because power equals energy per unit time, this curve is commonly referred to by audio engineers as the ETC or energy–time curve versus time.

To the right 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 time axis, respectively. In the center of Fig. 9 is the Wigner distribution itself as derived from the analytic signal associated with the impulse response. The analytic signal has no Fourier spectrum for negative frequencies and is used as a mathematical means to define and extract the envelope of the impulse response. A minor, but observable, artifact of the analytic signal is that its Wigner 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 in the right of the Wigner 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. 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 indicates damped, nearly sinusoidal oscillation. The rising group-delay curve follows the significant values of the Wigner distribution (the outermost contour line is 52 dB below the highest). 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 "impulsivelike," whereas concentrations parallel to the time axis imply that the signal is "sinusoidallike" or oscillatory in time (see Fig. 9). From the standpoint of modulation theory, a horizontal Wigner distribution indicates amplitude modulation (AM) because at every instant in time there is only one frequency present (just the carrier wave). On the other hand, a vertical orientation of the Wigner distribution indicates frequency modulation (FM) 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 FM and AM, for example, as Fig. 9 shows. In summary, the Wigner distribution in Fig. 9 is a graphic representation of the statement that the impulse response of a filter is, in general, an amplitude-modulated and frequency-modulated signal. While these effects are always evident in the impulse response itself, they are dramatically illustrated by the orientation of the Wigner distribution.

It is important to remark that, although instantaneous frequency indicates an FM effect, instantaneous frequency is not the same as Fourier analysis frequency. Similarly, its dual variable group-delay time is not the same as Fourier analysis time. In this context, the Wigner distribution is an important link between these two frequencies and times.

Fig. 9 (lower part) shows the impulse response of an all-pass filter (energy/frequency = constant) whose peak group delay equals 8 ms at 2 kHz. Examination of the impulse response waveform reveals both AM and FM effects. The energy–time curve (envelope power in decibels) shows the former while 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 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 1.4 and 2.9 kHz in the Wigner distribution strongly modify the energy–time curve (envelope) while, at later times, the three negative regions (nearly round

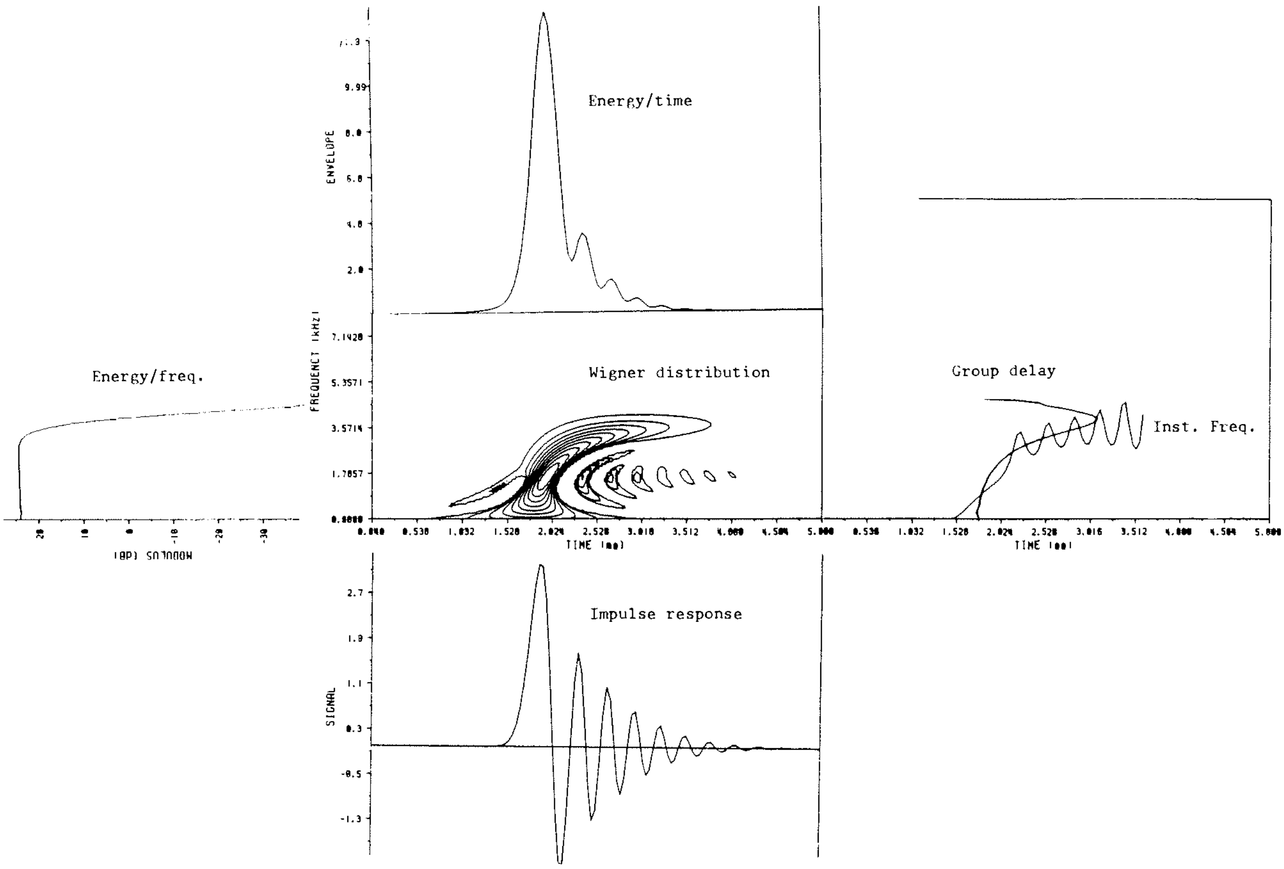


Fig. 9. 4-kHz anti-alias filter.

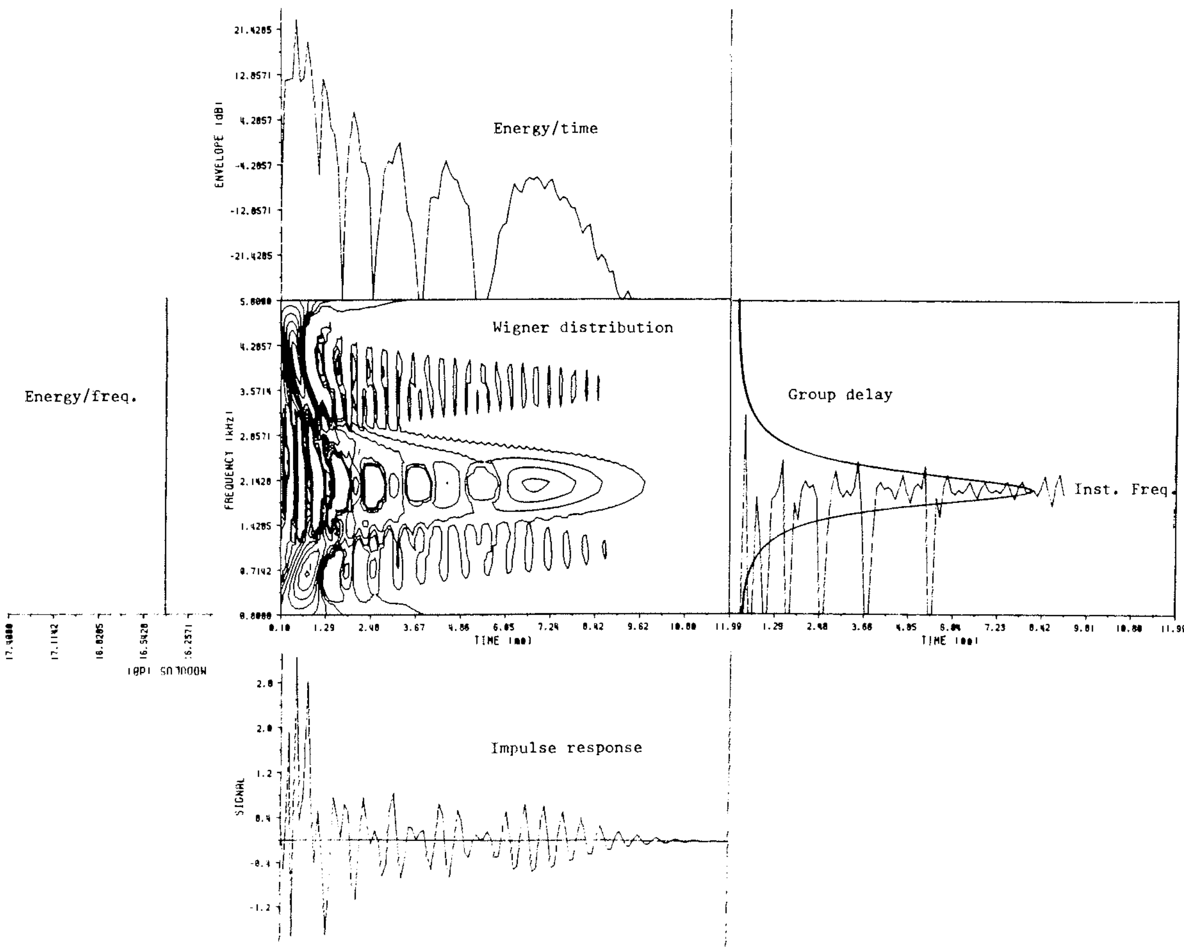


Fig. 10. All-pass filter, 2-kHz center frequency.

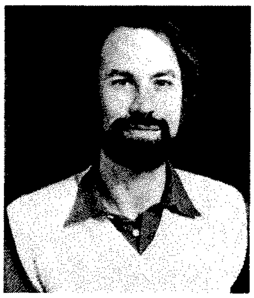
white regions) correspond to the last three collapses of the envelope. The rightmost central portion of the Wigner distribution corresponds to the arrival of the "packet" of signal energy near the 8-ms group-delay time which, from the impulse response or instantaneous frequency, is an oscillation at approximately 2 kHz. The Wigner distribution contains rather small (52 dB below maximum value) contributions that oscillate in time near the frequencies of 3.5 kHz and 0.71 kHz, which determine the "fine structure" of the envelope and instantaneous frequency.

The eight major figures contained in this communication can be interpreted and understood in a like manner.

THE AUTHORS

D. Preis

F. Hlawatsch



P. J. Bloom

J. A. Deer

Douglas Preis was born in Chicago, Illinois. He received B.S.E.E. and M.S.E.E. degrees in 1964 and 1966 from the University of Santa Clara, California. During the next three years he was a NASA fellow at Utah State University, Logan, where he received a Ph.D. degree in 1969.

Dr. Preis joined the faculty of Harvard University, Cambridge, Massachusetts, in 1970, where he was postdoctoral research fellow in the division of engineering and applied physics at the Gordon McKay Laboratory until 1978, when he joined the faculty of the department of electrical engineering of Tufts University, Medford, Massachusetts.

Professor Preis was awarded a Mellon Grant at Tufts in 1982. In 1984 he received a grant from the Deutsche Forschungsgemeinschaft to be Gastprofessor at Ruhr-Universität, Bochum, F.R.G. In 1985 he was Fulbright lecturer at Technische Universität Wien, Vienna, Austria. He has also given short courses at the universities of Barcelona and Madrid, Spain.

He has published or presented more than 90 original technical papers and reports in the field of electrical engineering. In 1978 and 1984 he received the Audio Engineering Society Publication Award for his papers, "Linear Distortion" and "Phase Distortion and Phase Equalization in Audio Signal Processing." He is the author of "Audio Spectrum and Signal Characteristics" in the forthcoming McGraw-Hill *Audio Engineering Handbook*. He has served as a member of the Technical Review Board of the *Journal of the Audio Engineering Society* since 1977, and was associate editor of the *IEEE Transactions on Acoustics, Speech, and Signal Processing* from 1979 to 1984. He is a member of Eta Kappa Nu, Sigma Xi, and an eminent engineer of Tau Beta Pi.

•

Franz Hlawatsch was born in Vienna, Austria, in 1959. He received a Diplom-Ingenieur degree in electrical engineering (with honors) from the Technical University of Vienna, Austria, in 1983. Since 1983 he has been a research and teaching assistant at the Department of Communications and High-Frequency Engineering, Technical University of Vienna, Austria. His research interests are mainly in signal theory and digital signal processing with special emphasis on time-frequency methods.

•

P. J. Bloom received a B.S. degree in physics in 1970 from Case Institute of Technology in Cleveland, Ohio. After graduation he taught science and worked as a musician, performing, writing, and recording in the Cleveland area.

In 1973 he joined the Department of Physics at University College, Cardiff, U.K., as a tutorial fellow and in 1977 completed a doctoral thesis on aspects of human sound localization. Since 1977 he has worked in the signal processing group in the engineering division of the Polytechnic of Central London, doing research in the areas of signal processing techniques for dereverberation of speech and the perception of speech. For the last few years he has been developing Wordfit, a digital signal processing system that automatically synchronizes the replacement dialog of an actor with his original dialog recorded during the shooting of a film or video.

A member of the Audio Engineering Society, Mr. Bloom also belongs to the Acoustical Society of America and the IEEE.

•

J. A. Deer received a B.Sc. degree in electronics engineering in 1984 from the Polytechnic of Central London. Currently he is studying for an M.Sc. degree in microelectronics and communication systems at the Polytechnic of Central London and is actively involved in researching the applications of digital signal processing techniques to VLSI custom and semicustom design.