

**Howard A. Gaberson**  
Energy and Utilities Development  
Division, Code ESC21  
Naval Facilities Engineering  
Service Center  
Port Hueneme, CA 93043-4328

---

# Application of Choi–Williams Reduced Interference Time Frequency Distribution to Machinery Diagnostics

*This article discusses time frequency analysis of machinery diagnostic vibration signals. The short time Fourier transform, the Wigner, and the Choi–Williams distributions are explained and illustrated with test cases. Examples of Choi–Williams analyses of machinery vibration signals are presented. The analyses detect discontinuities in the signals and their timing, amplitude and frequency modulation, and the presence of different components in a vibration signal. © 1995 John Wiley & Sons, Inc.*

---

## INTRODUCTION

Many problems in machinery diagnostics are characterized by frequency content that varies considerably and regularly with time. Signal analyzers average individual spectra to an apparent single representative spectrum, and one does not see any time frequency variations. The short time Fourier transform (STFT) or spectrogram is used to display time frequency variation in speech analysis, but does not provide sufficient resolution for machinery diagnostics.

The Wigner distribution is a fundamental concept for increased resolution in time frequency distributions, especially compared to the STFT. Forrester (1992) has shown the Wigner distribution to locate the angular position of the impact or discontinuity associated with individual gear tooth faults. However, the Wigner distribution has severe interference (cross) terms that confuse interpretation (Janse and Kaizer, 1983; Jeon and Shin, 1993; Meng and Qu, 1991; Wahl and Bolton,

1990; Yen, 1987) and require additional efforts to resolve. Most of the time frequency distributions have been shown to be members of Cohen's class of distributions (1989). Each is obtained from the general form with different kernels or selection functions. The reduced interference distributions (RIDs) employ kernels that mitigate the Wigner cross terms while preserving sharp resolution. Rohrbaugh (1993) has data that show promise in using RIDs for diagnostics. Our research has shown that the Choi and Williams distribution (1989) yields impressive detail and a significant structure in the time frequency plane. Interior impact is recognizable. This article presents an explanation of the RID calculations, with test cases explaining and illustrating the cross terms and their mitigation. Several examples of analyses of machinery diagnostic vibration signals illustrate the comprehensive structure. These methods seem to be leading to additional precise diagnostic tools for machinery and possibly structures.

---

Received March 7, 1995; Accepted June 22, 1995  
Shock and Vibration, Vol. 2, No. 6, pp. 437–444 (1995)  
© 1995 by John Wiley & Sons, Inc.

CCC 1070-9622/95/060437-08

To apply the time frequency methods to machinery vibration signals, the time frequency algorithms were programmed in Matlab (1994) script. Matlab is considered to be representative of all the high level PC signal processing software and our use of time frequency will be explained with reference to performing the calculations in Matlab on a PC.

## STFT AND WIGNER DISTRIBUTION

The most logical first step in considering a time frequency representation is to evaluate the windowed fast Fourier transform (FFT) magnitude of adjacent overlapping segments of the time signal, and see how the frequency content of these segments change. This is called the STFT or spectrogram (Boashash and Reilly, 1992). Thus, if  $x\{t\}$  is a sequence of properly digitized velocity values at evenly spaced increments of time, the STFT time frequency distribution would be given by:

$$X(t, f) = |\text{FFT}[w(\tau)x(t + \tau)]|. \quad (1)$$

Here,  $w$  is a window function, Hanning, Gaussian, or other that restricts the portion of the signal,  $x$ , being analyzed by the FFT for frequency content. The window function would only have non-zero values for half the FFT length symmetrically on either side of zero. We skip a suitable number of points and do the process again. Each FFT is the frequency content of the windowed segment of the function centered around  $t$ . Every little blip within that window contributes to the frequency content at  $t$ . If you shorten the window, you can be more accurate on the time, but you restrict the lowest frequency you can detect, because you need a wavelength within the window. This concept was coded in Matlab using skip values of nominally 10% of the length of the FFT computed.

The STFT is a popular machinery diagnostics analysis to run during startup or shutdown and is referred to by the signal analyzer manufacturers as a waterfall function. The amplitude is scaled to return the correct amplitude of an input sine wave.

The STFT does not have the time or frequency resolution of the more sophisticated approaches. A major improvement is the Wigner distribution (Janse and Kaizer, 1983). To explain this concept, consider the autocorrelation function (ACF) for a stationary and ergodic signal. It can be computed

from a single long-time segment. Because it is stationary, it will not depend on  $t$ , and the lagged product is only a function of the lag,  $\tau$ .

$$R_x(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T [x(t)x(t + \tau)] dt. \quad (2)$$

By the Wiener-Kinchine theorem (Bendat and Piersol, 1986), the Fourier transform of the ACF is the spectral density, that for practical purposes, is the average windowed value of the DFT (discrete Fourier transform) magnitudes squared. Or

$$|X|^2 = XX^* = S_{XX} = \frac{1}{2\pi} \int_{-\infty}^{\infty} R_x(\tau) e^{-j\omega\tau} d\tau. \quad (3)$$

Now going back to the ACF as expressed in Eq. (2), if we make the change of variable

$$t' = t + \frac{\tau}{2}, \quad (4)$$

it will look like a more symmetric autocorrelation function

$$R_x(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \left[ x\left(t' - \frac{\tau}{2}\right) x\left(t' + \frac{\tau}{2}\right) \right] dt'. \quad (5)$$

This still has the units of  $x^2$ , and does not depend on  $t'$ . It is evaluated over the whole infinite or at least very long record. The integration evaluates it at every  $t'$  for all  $\tau$ . Now if we had a nonstationary record where we certainly did not want to look at an average along the whole signal, but rather right at some time  $t'$ , we might take the Fourier transform of one calculation of the ACF centered about  $t'$ . We could think of some instantaneous autocorrelation at  $t'$  as

$$R_x(t', \tau) = \left[ x\left(t' - \frac{\tau}{2}\right) x\left(t' + \frac{\tau}{2}\right) \right] \quad (6)$$

and evaluate it only at this one  $t'$  for all  $\tau$ . The Fourier transform of this is certainly going to look like the magnitude of the square of the Fourier transform or spectral density of something, and that something is an instantaneous symmetrical autocorrelation of  $x$ , centered about  $t'$ . Boashash and Reilly (1992), Cohen (1989), Forrester (1992), Janse and Kaizer (1983), Jeon and Shin (1993), Meng and Qu (1991), McFadden and Wang (1992), Wahl and Bolton (1990), and Yen (1987) define

this as the Wigner transform and write it as:

$$W_x(t, \omega) = \int_{-\infty}^{\infty} e^{-j\omega\tau} x\left(t + \frac{\tau}{2}\right) x^*\left(t - \frac{\tau}{2}\right) d\tau. \quad (7)$$

Because of the complex conjugate notation in Eq. (7), Boashash and Reilly (1992), Janse and Kaizer (1983), Wahl and Bolton (1990), and Yen (1987) convert the real signal to be analyzed to a complex analytic equivalent, the real part of which is the original real signal. Matlab has a function called Hilbert that, when applied to a real signal, returns the complex equivalent analytical signal; so, for practical purposes, the requirement is easy to satisfy. To calculate, Eq. (7) has to be discretized. Again the authors give essentially the following

$$W_x(n, k) = 2 \sum_{k=-\infty}^{\infty} e^{-j2mk} x(n+m) x^*(n-m). \quad (8)$$

Note that between Eqs. (7) and (8) a “2” has crept into the exponential (because of the  $\tau/2$ ) that results in requiring twice as many samples for the same upper frequency limit as for the STFT. Again, this is not a problem because most signal processing programs offer an interpolate function that will double the sampling rate, but the computation load does become twice as long.

To avoid taking the FFT of a sequence with gross discontinuities at the sequence ends, we window the instantaneous ACF and term this a pseudo-Wigner distribution (PWD) (Janse and Kaizer, 1993). Thus, the PWD is the FFT of a term-by-term product of a window function,  $w(m)$ , and the discrete ACF or bilinear product. This can be expressed as:

$$W(n, k) = \mathcal{F}_{m \rightarrow k} [w(m)K(n, m)] \quad (9)$$

where

$$K(n, m) = x(n+m)x^*(n-m). \quad (10)$$

Adopting Boashash and Reilly's (1992) terminology,  $K(n, m)$  is the bilinear product and  $w(m)$  is the window function. To evaluate the PWD for a time  $n$ , we interpret Eq. (9) as an FFT with respect to the lag variable  $m$  of the function  $w(m)K(n, m)$ . The bilinear product has an odd number of terms. We always use an FFT, which requires an even number of points, so we add a final zero. The window,  $w(m)$ , can be a Hanning,

a Kaiser, or, as we used, a Gaussian shaped window.

To understand what to expect from the PWD, consider the bilinear product of an analytic sinusoid,

$$x = A \exp(j2\pi f \Delta t n). \quad (11)$$

Substituting Eq. (11) into (10) yields

$$K(n, m) = A \exp[j2\pi f \Delta t (n+m)] A \exp[-j2\pi f \Delta t (n-m)] \quad (12)$$

that simplifies to

$$K(n, m) = A^2 \exp(j4\pi f \Delta t m). \quad (13)$$

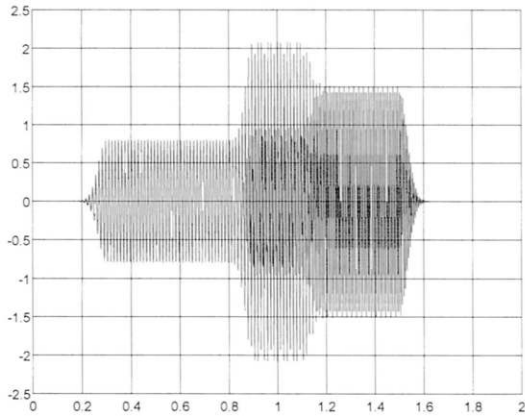
Because this is a function of  $m$ , with the square of the amplitude of Eq. (11) at twice its frequency, one can see that the FFT of this will show the amplitude squared and the frequency doubled. Thus, one can expect the Wigner transform and indeed all of the RIDs to perform similarly.

## CROSS TERMS

The difficulty with the PWD is that it generates artificial cross terms. To demonstrate the presence of the cross terms in the PWD, and their mitigation by the RID, a generated signal sequence called Twosines was used. This consists of a file of a 75-Hz sine starting at  $t = 0.2$ , rising with a Gaussian ramp to an amplitude of 0.8 at  $t = 0.3$ , remaining at constant amplitude to  $t = 1.1$ , and then dying out at  $t = 1.2$ . It also contains a 150-Hz sine, starting at  $t = 0.8$ , rising to a full amplitude of 1.5 at  $t = 0.9$ , staying constant to  $t = 1.5$ , and then dying out at  $t = 1.6$ . Our test function was sampled at 1000 samples per second and lasted 2 s. This time history is shown in Fig. 1.

The STFT analyzes this function well and is shown in Fig. 2. This was calculated with a Hanning window, using a 256-point FFT and 100 time steps, skipping 17 points between FFTs. It does not discern the 1.5 amplitude very accurately because the Hanning window can have an error of up to 16%.

To illustrate the cross term problems with the Wigner transform, it was tested on Twosines and computed in two ways. The first method was with no window, or what could be called a rectangular window. Figure 3 shows the Wigner computed



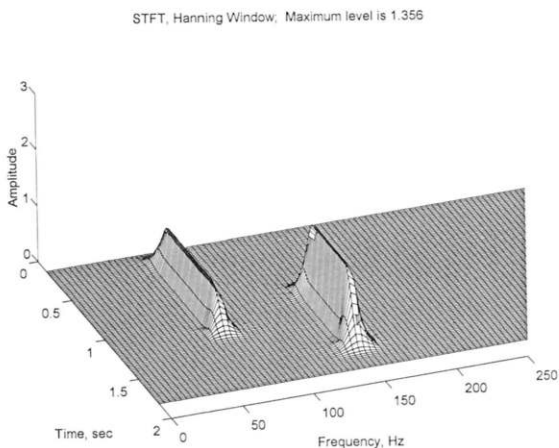
**FIGURE 1** Time history of Twosines test signal.

with a 256-point FFT and 144 time slices, skipping 12 points (or 12 ms) between each time slice. Note the terrible cross terms and the broadband ripple-like errors caused by the discontinuities at the ends of the bilinear product being transformed.

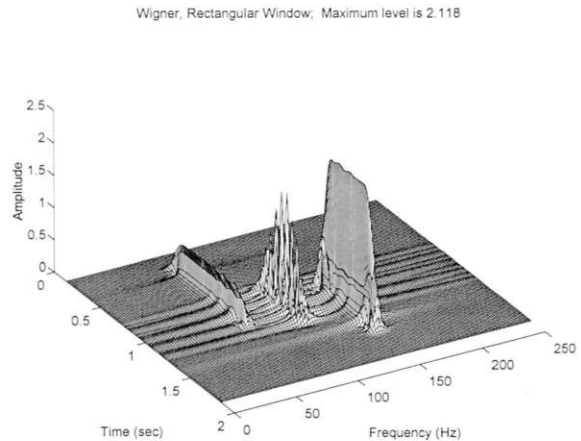
The second method, Fig. 4, shows the PWD with a Gaussian window applied to each bilinear product time slice prior to FFT'ing. Note that the ripple and broadband errors are completely gone, but the cross terms remain and, if anything, are even worse.

### RID CALCULATIONS

Because of the cross terms of the PWD, a great deal of effort has been invested in mitigating their effects. These efforts involve two-dimensional time frequency averaging. Jeon and Shin (1993),



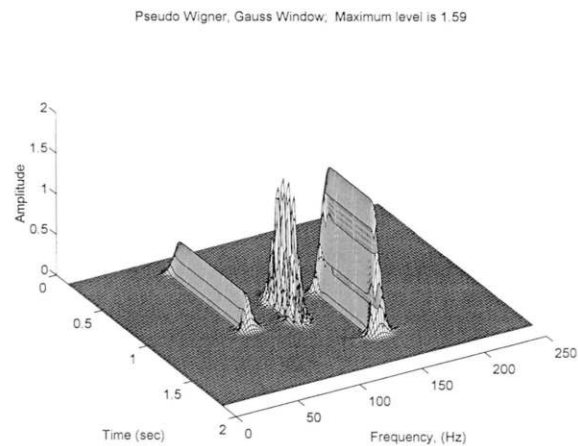
**FIGURE 2** Short time Fourier transform of Twosines.



**FIGURE 3** Pseudo-Wigner transform of Twosines with rectangular window.

Wahl and Bolton (1990), and Yen (1987) used a time frequency Gaussian smoothing function. Choi and Williams (1989) and Cohen (1989) organized these efforts and called the resulting time frequency distributions RID. Many RIDs exist, and they all involve a two-dimensional time convolution. One of the first was the Choi-Williams distribution (CWD; 1989), which is the subject here. FORTRAN code fragments for several of the RIDs were made available by Boashash and Reilly (1992), along with a description of the computations. They summarize the computation of the RID as:

1. form the bilinear product;
2. convolve it with the determining or selection function in the time dimension;



**FIGURE 4** Pseudo-Wigner transform of Twosines with Gaussian window.

3. calculate the FFT of this resulting convolved row for each time slice of the distribution.

The determining or selection function (sometimes called kernel) is what differentiates the various RIDs. Similar to Eq. (9), we write

$$\text{RID}(n, k) = \mathcal{F}_{m \rightarrow k}[G(n, m) * K(n, m)] \quad (14)$$

where  $G(n, m)$  is the selection function, and the convolution indicated is with respect to time or  $n$ . This is difficult to understand. Therefore, we shall illustrate the arrays and the convolution for

$n \downarrow$	$m \rightarrow$	-3	-2	-1	0	1	2	3	
-3		-6,0	-5,-1	-4,-2	-3,-3	-2,-4	-1,-5	0,-6	
-2		-5,1	-4,0	-3,-1	-2,-2	-1,-3	0,-4	1,-5	
-1		-4,2	-3,1	-2,0	-1,-1	0,-2	1,-3	2,-4	
0		-3,3	-2,2	-1,1	0,0	1,-1	2,-2	3,-3	(Array 1)
1		-2,4	-1,3	0,2	1,1	2,0	3,-1	4,-2	
2		-1,5	0,4	1,3	2,2	3,1	4,0	5,-1	
3		0,6	1,5	2,4	3,3	4,2	5,1	6,0	
4		1,7	2,6	3,5	4,4	5,3	6,2	7,1	
5		2,8	3,7	4,6	5,5	6,4	7,3	8,2	

The selection function,  $G(n, m)$ , is doubly symmetric; the results of the FORTRAN fragments only calculate the lower right quadrant. If we use  $G23$  to indicate  $G(2,3)$ , then  $G(n, m)$  will look like:

$G_{44}$	$G_{43}$	$G_{42}$	$G_{41}$	$G_{42}$	$G_{43}$	$G_{44}$
$G_{34}$	$G_{33}$	$G_{32}$	$G_{31}$	$G_{32}$	$G_{33}$	$G_{34}$
$G_{24}$	$G_{23}$	$G_{22}$	$G_{21}$	$G_{22}$	$G_{23}$	$G_{24}$
$G_{14}$	$G_{13}$	$G_{12}$	$G_{11}$	$G_{12}$	$G_{13}$	$G_{14}$
$G_{24}$	$G_{23}$	$G_{22}$	$G_{21}$	$G_{22}$	$G_{23}$	$G_{24}$
$G_{34}$	$G_{33}$	$G_{32}$	$G_{31}$	$G_{32}$	$G_{33}$	$G_{34}$
$G_{44}$	$G_{43}$	$G_{42}$	$G_{41}$	$G_{42}$	$G_{43}$	$G_{44}$

(Array 2)

The time convolution calculating algorithm (Boashash and Reilly, 1992) uses the template of (Array 2) as a two-dimensional moving average. Centering the  $G_{11}$  position of (Array 2) over the 0,0 position of (Array 1), multiplying the coinci-

dent members of each column term-by-term, and then adding them, results in the convolved row for time slice 0. A zero is added in column 8, and the row is FFT'd to yield a time slice of the RID at time 0.

Let us examine some actual values for the RID selection function or  $G(n, m)$ . The expression defining the CWD (1989) selection function is given by:

$$K(n, m) = x(n + m)x^*(n - m). \quad (15)$$

With that notation, the partial bilinear product, (Array 1) is:

$$G(n, m) = \frac{\sqrt{\sigma/\pi}}{2m} e^{-\sigma n^2/4m^2}. \quad (16)$$

If we consider an 8-point FFT, the code fragment will generate  $G(n, m)$  as a  $4 \times 4$  array. The algorithm computes only the lower right quadrant, that for the CWD with  $\sigma = 7$ , is:

1.0000	0.7411	0.3734	0.2556
0	0.1288	0.2411	0.2104
0	0.0007	0.0649	0.1174
0	0.0000	0.0073	0.0444

The full  $G$  matrix will actually be:

0.0444	0.0073	0.0000	0	0.0000	0.0073	0.0444
0.1174	0.0649	0.0007	0	0.0007	0.0649	0.1174
0.2104	0.2411	0.1288	0	0.1288	0.2411	0.2104
0.2556	0.3734	0.7411	1.0000	0.7411	0.3734	0.2556
0.2104	0.2411	0.1288	0	0.1288	0.2411	0.2104
0.1174	0.0649	0.0007	0	0.0007	0.0649	0.1174
0.0444	0.0073	0.0000	0	0.0000	0.0073	0.0444

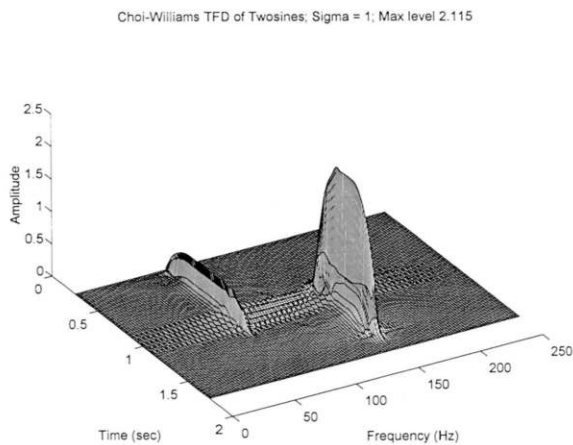
Note the “bow tie” appearance of the significant terms, which is characteristic of the RID and significantly different than the Gaussian smoother used in Jeon and Shin (1993), Wahl and Bolton (1990), and Yen (1987).

Now let us examine the performance of the CWD on the Twosines signal. The CWD contains a constant  $\sigma$ . Choi and Williams (1989) suggest a value of 1 for it, while others have suggested 7. Figure 5 is a plot of the CWD for Twosines with  $\sigma = 1$ . The cross terms are drastically reduced and spread out, and one notes a slight ripple for both low and high frequencies in the overlap region. Figure 6 gives the CWD for Twosines with  $\sigma = 7$ . The cross terms in the overlap region are slightly higher, and the ripple in the low and high frequency overlap area has disappeared. Because the bilinear kernel squares the amplitude of a sine, we might look at the square root of the amplitude, which is 2.02 for the 1.5 amplitude, 150-Hz wave, and the square root of the flat portion of the 75-Hz wave, which is 0.972 vice 0.8 in Twosines.

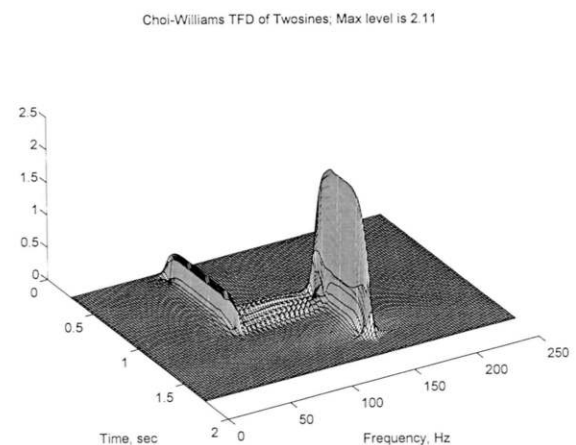
The cross term level is about 20% of the smaller 75-Hz wave level, and about 5% of the 150-Hz amplitude.

### MACHINERY VIBRATION DIAGNOSTIC SIGNAL ANALYSES

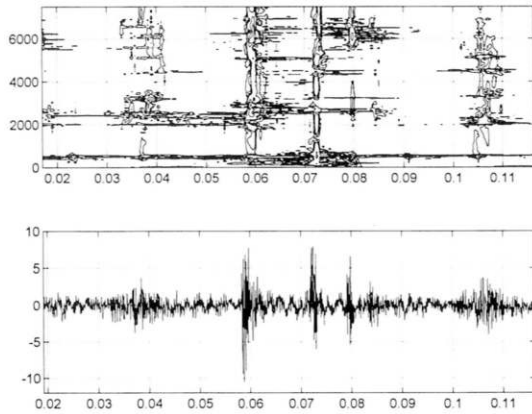
After completing the CWD testing with Twosines, we proceeded to examine the CWD performance on some machinery signals. The first example is some accelerometer data taken from the intake valve cap on the head of the high pressure cylinder of a large reciprocating compressor. Figure 7, in the lower half, gives a time history of the acceleration, and, in the upper half, a contour plot of the CWD. The burst of acceleration at 59 ms is caused by the intake valve channels slamming open. At 72 ms they close. Note that the events show a broad, impactlike high frequency content in the CWD contour plot. We consider vertical lines on the contour plot to be



**FIGURE 5** Choi-Williams distribution for Twosines with  $\sigma = 1$ .



**FIGURE 6** Choi-Williams distribution for Twosines with  $\sigma = 7$ .

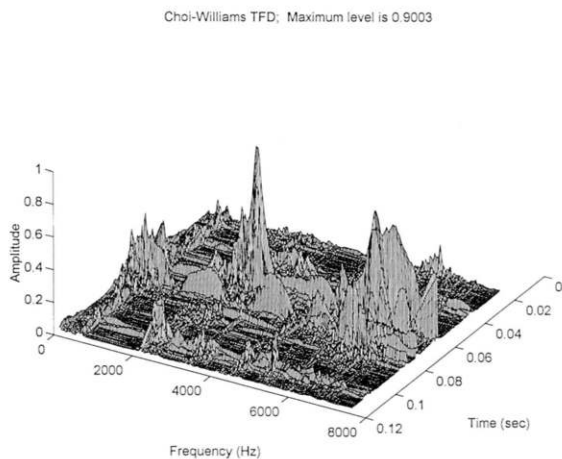


**FIGURE 7** Lower plot: acceleration–time history from the intake valve cap on a reciprocating compressor. Upper contour plot: corresponding Choi–Williams distribution.

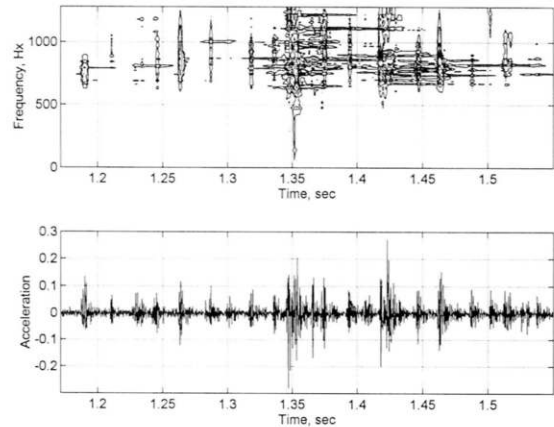
impacts, while horizontal contours are ringing or continuing vibration at that frequency. The other impact events are other valve opening and closing events, both intake and exhaust from the three cylinders.

Figure 8 is a three-dimensional mesh plot for this same CWD. The mesh plot contains too much information to grasp. We believe examination of contour plots coupled with trial and error use of specific contour levels provides the most revealing information.

The next example is an acceleration from the input bearing on a damaged and howling low speed crane gearbox. Figure 9 shows a portion of the acceleration time history and a contour plot of the CWD for the corresponding time interval.



**FIGURE 8** 3-D meshplot of Choi–Williams distribution shown in Fig. 7.



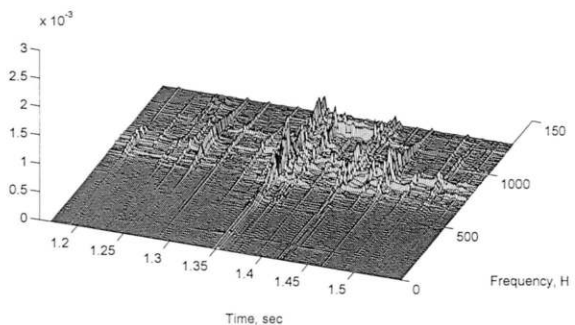
**FIGURE 9** Contour plot and acceleration–time history of CWD of crane gearbox.

Many impactlike events occur, with the most prominent and broadbanded at about 1.35 s. The rectangular structure of the contour plot with impacts and ringings has been noticed in many CWD analyses we have performed. Figure 10 shows the corresponding mesh plot.

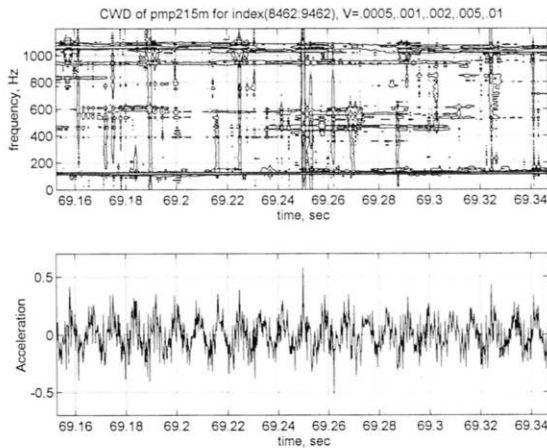
Figure 11 is a time history and a contour plot from the driven bearing on a turbine boiler feed pump that appears to indicate impacting. The bearing is probably in the first stages of spalling.

## CONCLUSION

To conclude, the CWD and the other RIDs (Boashash and Reilly, 1992) can be computed on a PC with high-level signal processing software such as Matlab. They seem to provide impact or discontinuity detection with precise timing information. They display amplitude and frequency modulation, and identify different components of the



**FIGURE 10** Mesh plot of CWD of crane gearbox signal shown in Fig. 9.



**FIGURE 11** Time history and CWD contour plot from pump bearing.

signal. We expect these methods to become valuable machinery diagnostic tools and recommend their use.

The author would like to thank the Office of Naval Research for supporting this project.

## REFERENCES

- Bendat, J. S., and Piersol, A. G., 1986, *Random Data Analysis and Measurement Procedures*, 2nd ed., Wiley, New York.
- Boashash, B., and Reilly, A., "Algorithms for Time-Frequency Signal Analysis," in B. Boashash, *Time-Frequency Signal Analysis Methods and Applications*, 1992, Wiley, New York, pp. 163-181.
- Choi, H., and Williams, W. J., 1989, "Improved Time-Frequency Representation of Multicomponent Signals Using Exponential Kernels," *IEEE Transactions in Acoustics, Speech, and Signal Processing*, Vol. 37, pp. 862-871.
- Cohen, L., 1989, "Time-Frequency Distributions—A Review," *Proceedings of the IEEE*, Vol. 77, pp. 941-981.
- Forrester, B. D., "Time-Frequency Analysis in Machine Fault Detection," in B. Boashash, *Time-Frequency Signal Analysis Methods and Applications*, 1992, Wiley, New York, pp. 406-423.
- Janse, C. P., and Kaizer, J. M., 1983, "Time-Frequency Distributions of Loudspeakers: The Application of the Wigner Distribution," *Journal of the Audio Engineering Society*, Vol. 31, pp. 198-223.
- Jeon, J. J., and Shin, Y. S., 1993, "Pseudo Wigner-Ville Distribution, Computer Program and Its Applications to Time-Frequency Domain Problems," Report NPS-ME-93-002, Naval Postgraduate School, Monterey, CA, March 1993.
- Matlab for Windows, Version 4.2, 1994, *High Performance Numeric Computation and Visualization Software*, The MathWorks, Inc., Natick, MA.
- Meng, Q., and Qu, L., 1991, "Rotating Machinery Fault Diagnosis Using Wigner Distribution," *Mechanical Systems and Signal Processing*, Vol. 5, pp. 155-166.
- McFadden, P. D., and Wang, W. J., 1992, "Analysis of Gear Vibration Signatures by the Weighted Wigner-Ville Distribution," *Institute of Mechanical Engineers*, Vol. C432/134, pp. 387-393.
- Rohrbaugh, R. A., 1993, "Advanced Time Frequency Analysis: Theory and Application to Machinery Condition Assessment," Report SADP-U93/00039-715, Ship Acoustics Department, Carderock Division, Detachment Puget Sound, Naval Surface Warfare Center, Bremerton, WA, April 1993.
- Wahl, T. J., and Bolton, J. S., 1990, "The Use of the Wigner Distribution to Analyze Structural Impulse Responses," *International Congress on Recent Development in Air- and Structure-Borne Sound and Vibration*, Auburn University, Auburn, AL, pp. 827-834.
- Yen, N., 1987, "Time and Frequency Representation of Acoustic Signals by Means of the Wigner Distribution Function: Implementation and Interpretation," *Journal of the Acoustic Society of America*, Vol. 81, pp. 1841-1850.





Hindawi

Submit your manuscripts at  
<http://www.hindawi.com>

