PHASE ERRORS IN THE CROSS SPECTRUM
ESTIMATE DUE TO MISALIGNMENT

Sverre Holm

Electrical Engineering Department
Yarmouk University, Irbid, Jordan.

ABSTRACT

When the cross spectrum is estimated where there is a delay between the two input signals, it is shown that the corresponding misalignment gives both magnitude and phase errors. The existence of the phase error is a new result. The importance of the errors depend on the algorithms used. It is shown that the misalignment may either be given by the total observation time or by the section length in the processor. In the latter case the effect of phase errors must be considered when the section length is low, that is when the cross spectrum is smooth, requiring little resolution.

INTRODUCTION

The cross spectrum is estimated as a part of the estimation of the cross correlation, using the Fourier transform relationship between the two. It is more efficient to use fast Fourier transform (FFT) methods and find the cross correlation via the cross spectrum, than to find it directly in the time domain. Besides this method also gives a straight forward way of estimating the generalized cross correlation between two signals, s(t) and r(t) [1]:

$$R_{sr}^{(g)}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Psi(f)\, G_{sr}(f) e^{j2\pi f\tau} d\tau \qquad (1)$$

In this case the cross spectrum, $G_{sr}(f)$, is weighted by a filter $\Psi(f)$ before transformation. Setting this filter equal to unity gives the cross correlation function $R_{sr}(\tau)$. The cross spectrum is also used to define the coherence between two signals:

$$\gamma_{sr}(f) = G_{sr}(f)/(G_{ss}(f)\, G_{rr}(f))^{\frac{1}{2}} \qquad (2)$$

The coherence is the cross spectrum normalized by the power spectra of the two inputs, and its magnitude is always less or equal to unity. From the coherence one can obtain information about the amount of nonlinearity that relates the two signals, or the amount of uncorrelated noise in the transmission path from s(t) to r(t).

A special case of great practical interest in bearing estimation is when there is a time delay, D, between the two signals in addition to noise. In this case the phase of the cross spectrum contains the delay information. The delay can be found from the slope of the phase as a function of frequency (assuming non-dispersive propagation) [2]. It can also be found from the generalized cross correlation, because the lag that corresponds to its peak value is the time delay.

In most practical estimation problems one has a limited, T seconds, long segment of each input signal. The delay, D, then gives a reduced overlap of the two inputs of T-D seconds. This is referred to as misalignment, and it gives an error in the cross spectrum estimate. This error was investigated by Carter [4] and Kroenert [5], but they only found the effect on the magnitude of the cross spectrum. In this paper the effect on the phase is also found.

EXPECTED VALUE OF THE CROSS SPECTRUM

The input signals are s(t) and r(t) = s(t-D) + n(t). The second signal is delayed by D seconds and has an additive noise term, n(t). The finite Fourier transforms over segments of length T (assuming rectangular time windowing) are:

$$S(f) = \int_0^T s(v) e^{-j2\pi fv} dv \qquad (3a)$$

$$R(f) = e^{-j2\pi fD} \int_{-D}^{T-D} s(u) e^{-j2\pi fu} du + \int_0^T n(u) e^{-j2\pi fu} du \qquad (3b)$$

The cross spectrum estimate is found either by letting T be large or by averaging over many segments, and it approaches the expected value:

$$E\left[G_{sr}(f)\right] = E\left[\frac{1}{T} S^*(f)\, R(f)\right]$$

$$= e^{-j2\pi fD} E\left[\frac{1}{T}\int_0^T s(v) e^{j2\pi fv} dv \int_{-D}^{T-D} s(u) e^{-j2\pi fu} du\right] \qquad (4)$$

The noise is uncorrelated with the signal and disappears. The signals are zero outside the limits of integration, and the equation can be split into four different integrals:

$$\int_0^{T-D}\left[\int_0^{T-D} + \int_{T-D}^{T}\right] \int_0^{T-D}\left[\int_0^{T-D} + \int_{-D}^{0}\right] \int_{-D}^{0}\left[\int_{T-D}^{T} \int_{-D}^{0}\right] \qquad (5)$$

where the integration kernels are as in (4). The first integral gives the following contribution after a change of variables (v=u-g):

$$e^{-j2\pi fD} \int_{-T+D}^{T-D} \frac{T-D-|g|}{T} R_{ss}(g) e^{-j2\pi fg} dg$$

CH2118-8/85/0000-0812 $1.00 © 1985 IEEE

$$= \frac{T-D}{T} e^{-j2\pi fD} \hat{G}_{ss}(f) \qquad (6)$$

This term is an estimate of the power spectrum $G_{ss}(f)$, with phase given by the delay. The misalignment weight is the same as in equation 7a in [4]. This term is the only one when there is no misalignment.

The two next integrals are similar, after substituting u=v+g in the second integral and v=u+g in the third integral of (5) we get:

$$2e^{-j2\pi fD} \frac{D}{T} \int_{-T}^{0} w(-g)\, R_{ss}(-g)\, e^{-j2\pi fg} dg \qquad (7)$$

The weight, w(g), is trapezoidal for 0<D< T/2:

$$w(g) = \begin{cases} g/D & ,\ 0 < g < D \\ 1 & \ D < g < T-D \\ (T-g)/D & \ T-D < g < T \end{cases} \qquad (8)$$

This weight plays the same role as the triangular weight of the autocorrelation in (6), but becomes trapezoidal since the range of integration in the two original integrals in (5) are different. The contribution from (7) is proportional to the delay, and it is important to notice that the transform over the autocorrelation is unsymmetrical. Thus the integral gives a complex result, and totally the phase of this term is different from that given by the delay alone. This integral will contribute to the total for all autocorrelations with non-zero terms outside lag 0.

The fourth integral in (5) can be changed to a suitable form by substituting v=u-g. As in (6) both integrals are over similar ranges, giving a triangular weighting:

$$e^{-j2\pi f(D-T)} \int_{-D}^{D} \frac{D-|g|}{T} R_{ss}(T-g) e^{-j2\pi fg} dg \qquad (9)$$

The contribution from this integral is also complex, and it depends on the autocorrelation from lag T-D to T+D. But since the segment length, T, usually is longer than the correlation time, this term will be smaller than (7).

The expected value of the cross spectrum is now the sum of equations (6), (7), and (9). With misalignment the phase of the cross spectrum is different from that given by the delay alone. To find the phase error's dependence on the delay, the segment length, and the autocorrelation the equation was discretized:

$$E\left[G_{sr}(k)\right] = W_L^{kD} \Big[\sum_{\ell=-T+D}^{T-D} (T-D-|\ell|)/T\, R_{ss}(\ell) W_L^{k\ell}$$
$$+ \frac{2D}{T} \sum_{\ell=-T}^{0} w(-\ell)\, R_{ss}(-\ell)\, W_L^{k\ell}$$
$$+ W_L^{-kT} \sum_{\ell=-D}^{D} (D-|\ell|)/T\, R_{ss}(T-\ell)\, W_L^{k\ell} \Big] \qquad (10)$$

L is the number of samples in the cross spectrum, T and D are the segment length and the delay in samples, $W_L = \exp(-j2\pi/L)$ and the window, $w(\ell)$ is still defined by equation (8).

Special cases

Some special cases are worth noting. The first is a white noise input with autocorrelation $R_{ss}(\tau) = \delta(\tau)$. In this case (6) is the only contribution giving:

$$E\left[G_{sr}(f)\right] = \frac{T-D}{T} e^{-j2\pi fD} \qquad (11)$$

and 0 if the delay is larger than the segment length. There is no phase error, only a decrease of the magnitude. This is consistent with [4], and this result is generalized in [5] to other data windows than the rectangular one implied in (3).

The second special case is a constant signal where the autocorrelation also is a constant. Inserting directly in (4) gives:

$$E\left[G_{sr}(f)\right] = 4\,T \quad \mathrm{sinc}^2(fT/2\pi) \qquad (12)$$

which is a real quantity. The delay has completely disappeared which is consistent with the fact that a delayed constant signal can not be distinguished from the signal itself. The same is true for a cosine signal whose period is the delay divided by an integer.

These special autocorrelations were used to verify that equation (10) was correct.

Low pass spectrum

A typical spectrum in underwater acoustics is one where the energy falls as the frequency increases. A spectrum which falls off with 20 dB/decade has an autocorrelation which is $R_{ss}(\tau) = \exp(-a\tau)$. The sampled autocorrelation $\exp(-\ell/10)$ was used, and the bias in the phase was found by subtracting the phase of (10) from the true phase. The bias is negative if the estimated phase corresponds to a smaller delay than the correct one, and the results are shown in figures 1 and 2.

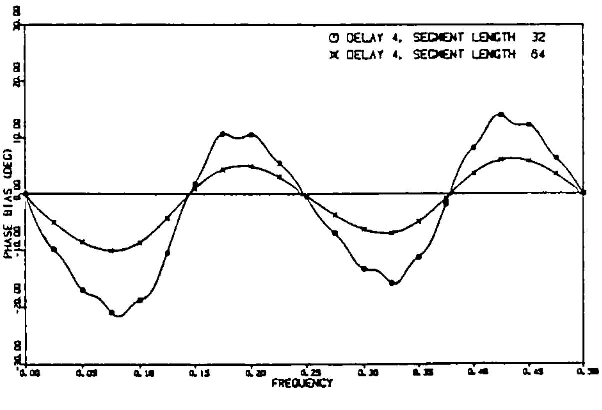

Fig. 1. Phase bias with delay 4 samples.

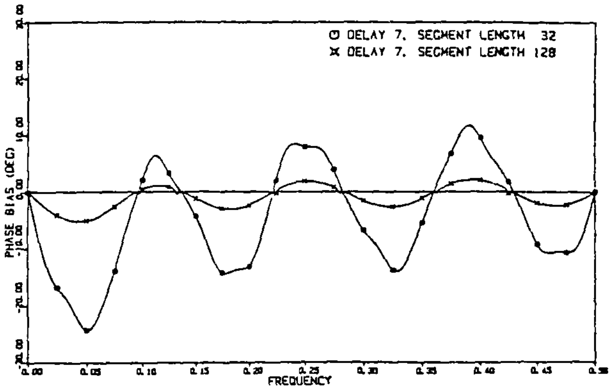

Fig. 2. Phase bias with delay 7 samples.

The phase is an oscillating function with D/2 number of periods, and the amplitude of the error falls with the segment length. The effect of the phase error is a bias in the estimated delay and also an increased variance. In the cross correlation this will be seen as a shift in the peak value and a broadening of the peak. Like for the constant signal the magnitude of the estimated phase will be too small, especially when the delay is not an even number of samples.

COMPUTATIONAL CONSIDERATIONS

The segment length in the formulas has two different interpretations depending on the procedure used for computation. It can either be the total observation time of the inputs or the section length in the processor.

Estimation of cross correlation

The most common method is to divide the N input samples in the x and y channels into M-point sections. These sections may be weighted by a smooth time window, in which case it is advantageous to use overlap. But for the estimation of cross correlation only, this is not necessary. To each M-point section, M seroes are added and the 2M-point sections are Fourier transformed, using a discrete Fourier transform (DFT). Assuming K input sections and no overlap (K=N/M) the next step is averaging, giving a biased cross spectrum:

$$V(k) = \sum_{i=1}^{K} X_i^*(k)\, Y_i(k) \qquad k=0, \ldots, 2M-1 \qquad (13)$$

The cross correlation is found by inverse transformation:

$$c_{xy}(m) = \begin{cases} v(m), & m=o, \ldots, M-1 \\ v(-m), & m=-1, \ldots, -M+1 \end{cases} \qquad (14)$$

This method requires the evaluation of 2K 2M-point DFT's (often for real inputs) and 1 2M-point inverse DFT. Often with random signals the zero-padding makes little difference and so one can use M-point DFT's instead [3]. Strictly speaking this gives an aliased estimate. This method will be referred to as method A, while the method without errors is B. For both methods the misalignment effect applies to every M sample section of the inputs. The segment length T, in formulas 3 to 10 is therefore equal to the section length M.

A more accurate method is outlined in [6]. The N input samples are again divided into K non-overlapping M-point sections. The procedure now has two steps. The first is the estimation of the cross correlation for positive lags. M zeroes are added to the x segment, and it is Fourier transformed giving $X_i(k)$ as before. But to avoid the misalignment effect in each section the y transform is obtained from input sections i and i+1 combined together in a 2M-point block. Therefore no zeroes are added for this transformation. It can be shown that the transform in the y-channel is:

$$Y_i(k) + (-1)^k\, Y_{i+1}(k) \qquad (15)$$

where $Y_i(k)$ is the transform of segment i and M zeroes. The biased cross spectrum is then accumulated:

$$V_1(k) = \sum_{i=1}^{K} X_i^*(k)\left[Y_i(k) + (-1)^k\, Y_{i+1}(k)\right] \qquad (16)$$

and the cross correlation estimate for positive lags is found from the inverse transform:

$$c_{xy}(m) = v_1(m), \quad m = o, \ldots, M-1 \qquad (17)$$

The second step is to get an estimate for negative lags and then the x and y channels must be interchanged. Form

$$V_2(k) = \sum_{i=1}^{K} \left[X_i(k) + (-1)^k\, X_{i+1}(k)\right] Y_i^*(k) \qquad (18)$$

and

$$c_{xy}(-m) = c_{yx}(m) = v_2(m), \quad m=o, \ldots, M-1 \qquad (19)$$

This method will be referred to as method C. The number of DFT's required is 2K 2M-point DFT's (possibly for real inputs) and 2 2M-point inverse transforms. The difference from method B is only one extra inverse transform, but the gain is that the misalignment is given by the total record length N=KM instead of only the section length M. The errors caused by misalignment are therefore reduced.

Estimation of cross spectrum and coherence

When the spectra are to be estimated one must consider the effect of finite Fourier transforms on sidelobe leakage and resolution. Thus the issue of windowing becomes important, but windows can be applied either on the samples (time windowing) or on the correlation estimates (lag windowing). It has been shown [7] that it is possible to transform a time window into an equivalent lag window, and this principle will be used to compare methods A, B, and C of the previous section for the estimation of spectra. To estimate the coherence, both the cross spectrum and the two auto spectra need to be estimated.

Method A, using M point segments and DFT's is the standard weighted, overlapped segment, averaging method. Usually an overlap of 50% is used together with a smooth time window, and therefore 2K-1 sections must be transformed. A total of 2(2K-1) M-point DFT's (possibly for real inputs) are needed to form the three spectral estimates. Note that the aliasing effect mentioned in the previous section only applies to the correlation estimate and not to the spectra. A program that implements this method is given in [3] and it is also referred to in [2].

Method B with zero-padding and 2M-point DFT's can be windowed either in the time or in the lag domain. No overlap is required if the window is applied to the three correlation estimates, and to find them 2K (real) 2M-point DFT's and 3 inverse DFT's are required (2 of them possibly with real results). Then a lag window obtained by transformation of a desired time window is applied, and the three correlation functions are transformed back to the frequency domain. Totally this method requires 2K+4 (possibly real) + 2 2M-point transforms. If the window is applied to the time data as in method A, 2 (2K-1) (possibly real) 2M-point transforms are required, but usually the first method is more efficient. Methods A and B will, if the windows are made to be equivalent, give the same expected values. The difference is that method B gives twice as many samples in the spectra and no aliasing of the correlation estimates.

In method C, the most accurate one, the input sections have different lengths and if they are windowed in the time domain the identity in (15) does not apply. This means that the number of DFT's will be doubled. However, the equivalent lag window can also be used here. So again get estimates of the two autocorrelations and the cross correlation for positive and negative lags, by performing 2K 2M-point real transforms and 4 inverse transforms (2 of them possibly real). Now the lag window should be applied to the autocorrelations and to the two pieces of the cross correlation joined together. They should then be inverse transformed requiring 3 transforms (2 of them possibly real) to get the cross spectrum and the two auto spectra. The total count is thus 2K+4 (real) 2M-point transforms and 3 2M-point transforms. Thanks to the equivalence between time and lag windows [7] this accurate method can be implemented with only 1 more transform than in B.

EXPERIMENTAL RESULTS

A simulation using smooth low pass signals and a limited amount of data was made. Algorithms A and B as implemented in the IEEE program collection [3] were used. Because the spectra were smooth (no sharp peaks) the resolution could be low, and the section length was between 32 and 128 samples. It was found that a delay of 8 samples gave a bias in the delay estimate (found from the generalized cross correlation) of 18% of a sample with M=32, 5% with M=64, and 1% with section length 128 samples. In all cases the magnitude of the estimated delay was less than the true delay.

CONCLUSION

It has been shown that misalignment gives both phase and amplitude errors in the estimated cross spectrum. These errors also affect the cross correlation and coherence estimates. The existence of the phase error is a new result. This error causes the estimate of delay to become smaller in magnitude than the true delay. Depending on the algorithm used for processing, the misalignment may be either the ratio of the delay and the total observation time or it may be given by the delay and the section length in the processor. The different possibilities for algorithms are outlined and due to the transformation between time and lag windows [7], it is shown that estimates with small misalignment errors may be found with almost no increase in the cost of computation compared to usual methods.

ACKNOWLEDGEMENT

The simulated results were obtained while the author was at the Electronics Research Laboratory at the Norwegian Institute of Technology, Trondheim, Norway working under contract with Bentech A/S, Tromsø, Norway. Thanks also to L. Krane of Bentech A/S for discussions that lead to this work.

REFERENCES

1. C.H. Knapp and G.C. Carter, "The generalized cross correlation method for estimation of time delay", IEEE Trans. Acoust., Speech, Sign. Proc., Vol. ASSP-24, No. 4, Aug. 1976, pp. 320-327.

2. A.G. Piersol, "Time delay estimation using phase data", IEEE Trans. Acoust., Speech, Sign. Proc., Vol. ASSP-29, No. 3, June 1981, pp. 471-477.

3. G.C. Carter and J.F. Ferrie, "A coherence and cross spectral estimation program", in Digital Signal Processing Committee (ed), Programs for digital signal processing, New York, IEEE Press, 1979.

4. G.C. Carter, "Bias in magnitude-squared coherence estimation due to misalignment", IEEE Trans., Speech, Sign. Proc., Vol. ASSP-28, No. 1, Febr. 1980, pp. 97-99.

5. J.T. Kroenert, "Some comments on bias/misalignment effects in the magnitude squared coherence estimate", IEEE Trans. Acoust., Speech, Sign. Proc., Vol. ASSP-30, No. 3, June 1982, pp. 511-513.

6. A.V. Oppenheim and R.W. Schafer, Digital signal processing, chapter 11.6.2, Prentice-Hall, Inc., New Jersey, 1975.

7. A.H. Nuttall and G.C. Carter, "Spectral estimation using combined time and lag weighting", Proc. IEEE, Vol. 70, No. 9, Sept. 1982, pp. 1115-1125.