K'X751

NUMERICAL METHODS OF NOISE REDUCTION FOR FREQUENCY DOMAIN SEM

David A. Ksienski and Thomas M. Willis*
Radiation Laboratory
Department of Electrical Engineering and Computer Science
The University of Michigan
Ann Arbor, Michigan 48109

ABSTRACT

This paper addresses the numerical problem of performing a pole residue expansion on noisy data. A method of combining dissimilar data sets to achieve an effective increase in signal-to-noise ratio is tested on computer generated data simulating scattering from a sphere and from a thin wire. The data is corrupted with white Gaussian noise and the algorithm is tested for signal-to-noise ratios ranging from -10 dB to 60 dB. An iterative version of the algorithm is also tested. Finally, the problem of filtering noise from relatively quiet data is discussed and a novel filtering algorithm is presented.

1. INTRODUCTION

This paper addresses the numerical problem of performing a pole residue expansion on noisy data. Previous investigators [1,2] have shown that even relatively low noise levels severely diminish the accuracy of the pole residue expansion. In a recent paper by Ksienski [2], a method of combining dissimilar data sets to achieve an effective increase in signal-to-noise ratio was presented and shown to be effective for sphere data at relatively low noise levels. In the present paper the method is tested on sphere data at much higher noise levels as well as on data associated with scattering from a thin wire. An iterative technique suggested in Ksienski [2] is also tested. Finally, the problem of filtering noise from relatively quiet data is discussed and a novel filtering algorithm is presented.

2. MULTIPLE DATA SETS

2.1 Formulation. Before presenting the numerical results associated with combining dissimilar data sets, a brief summary of the theoretical foundation is presented. A more detailed presentation is given by Ksienski [2]. The data will be represented using the standard pole residue expansion

$$F_k(j\omega) = \sum_{m=1}^{M} \frac{a_{km}}{j\omega - s_m} + \frac{\bar{a}_{km}}{j\omega - s_m} + N_k(j\omega) \qquad (1)$$

*AT&T Bell Laboratories Scholar.

RL-781 = RL-781

where k is an index to the various measurements and $N_k(j\omega)$ represents noise and clutter. The representation in terms of conjugate pole pairs is used to explicitly show that $F_k(j\omega)$ is an even function of the circular frequency ω and is associated with a real function of time. We wish to combine K such data sets in a manner so as to have a maximum signal-to-noise ratio in the composite data set, where the noise is assumed to have zero mean and be independently and identically distributed. Assuming a linear combination with arbitrary complex weighting coefficients, the composite data set is

$$F_{comp}(j\omega) = \sum_{k=1}^{K} w_k F_k = \sum_{k=1}^{K} w_k \left(\sum_{m=1}^{M} \left(\frac{a_{km}}{j\omega - s_m} + \frac{\bar{a}_{km}}{j\omega - \bar{s}_m} \right) + N_k(j\omega) \right) \tag{2}$$

and by interchanging the order of summation and noting the invariance of the poles s_m with measurement number k,

$$F_{comp}(j\omega) = \sum_{m=1}^{M} \left(\frac{b_m}{j\omega - s_m} + \frac{c_m}{j\omega - \bar{s}_m} \right) + N(j\omega) \tag{3}$$

where

$$b_m = \sum_{k=1}^{K} w_k a_{km} \tag{4}$$

$$c_m = \sum_{k=1}^{K} w_k \bar{a}_{km} \quad . \tag{5}$$

In general, it will only be possible to maximize one b_m or c_m at a time. This is because for different choices of m the a_{km} will not vary in unison with k. Thus, the increased signal-to-noise ratio, and hence increased accuracy, will only be obtained for the pole s_m or $\bar{s}_m$. However, since the specification of the pole is arbitrary, several poles may be obtained through successively emphasizing different b_m or c_m. For the present, we restrict ourselves to obtaining an improved estimate of s_1, which necessitates maximizing $|b_1|^2$. From (4), we may write

$$|b_1|^2 = |w^T a|^2 \tag{6}$$

where $w = [w_1, w_2, \ldots, w_K]^T$ and $a = [a_{11}, a_{12}, \ldots, a_{K1}]^T$. Without loss in generality, the weighting vector may be constrained to satisfy $\bar{w}^T w = 1$. Then, noting the dot product formulation of (6), the maximum b_1 will occur for

$$w = \frac{\bar{a}}{|a|} \tag{7}$$

and for this optimal choice of the weighting vector, $|b_1|^2 = ||a||^2$. To evaluate the energy associated with the noise $N(j\omega)$, we take the expectation of the noise squared

$$E|N(j\omega)|^2 = E\left|\sum_{k=1}^{K} w_k N_k(j\omega)\right|^2$$

$$= \sum_{k=1}^{K} E|w_k N_k(j\omega)|^2 ,$$

and assuming the variance of the noise is equal to σ^2 in each of the measurements $F_k(j\omega)$

$$E|N(j\omega)|^2 = \sigma^2 \sum_{k=1}^{K} |w_k|^2 = \sigma^2 .$$

Thus the energy associated with the noise will not increase with w normalized as in (7). If K data sets are combined each containing an approximately equal excitation of the mode associated with the desired pole, the signal-to-noise ratio will increase by about a factor of K. For any set of residues, the resulting weighting coefficients will produce the maximum possible increase in signal-to-noise ratio. However, this is dependent upon an accurate knowledge of the residues which is generally not available. Two solutions to this problem have been devised, and both focus upon increasing the accuracy of the weighting vector w. First, the error due to variation in the estimate of the real part of the pole (which often constitutes a majority of the error) may be reduced by modifying (7). Assuming that the primary effect of a pole and its associated residue is in the immediate vicinity of the pole, it is possible to get a first estimate of the residue which would have resulted had the pole been constrained to have a different real part. Since for a fixed residue the component of the signal at a point on the $j\omega$ axis closest to the pole is inversely proportional to the real part of the pole, an improved weighting vector may be obtained by normalizing each element of the weighting vector (i.e., the residue) by the real part of the pole associated with each residue. Thus (7) is used with a redefined as

$$a = \left[\frac{a_{11}}{\mathrm{Re}(s_{11})} , \frac{a_{21}}{\mathrm{Re}(s_{21})} , \ldots , \frac{a_{K1}}{\mathrm{Re}(S_{K1})}\right]^T \qquad (8)$$

where s_{i1} refers to the estimate of the pole s_1 obtained from the ith data set. The second method of improving w requires additional computation. Since the application of the algorithm should result in improved residues (using the technique described below), these improved residues may be used to construct a new and presumably more accurate weighting vector. Thus, improvements may be made through the iterative application of the algorithm. The results of the application of the original algorithm as well as the iterative version are presented in the next section.

After the composite data set is formed, it must be subjected to a pole and residue extraction procedure. The algorithm due to Levy [3] and Sanathanan and Koerner [4] assumes that the poles and residues exhibit conjugate symmetry. This is generally appropriate as it is equivalent to requiring the frequency data to correspond to a real function of time. Unfortunately, $F_{comp}(j\omega)$ does not have conjugate symmetry. This may be seen from (3), (4), and (5), where for an arbitrary complex weighting coefficient w_k, the resulting b_m and $\bar{c}_m$ will not be equal. This problem may be circumvented by using the algorithm derived in Ksienski [2], which does not force the residues to occur in conjugate pairs. The implementation of this algorithm requires negative frequency information which is obtained from the data sets $F_k(j\omega)$ by employing conjugate symmetry. After the accuracy of the pole location is improved, the next step is to compute the residues relative to the new pole estimate. Constraining a pole to a particular location is not an easy task [1], but the algorithm presented in Ksienski [2], by constraining the poles to occur in conjugate pairs without similarly restricting the residues, provides a simple alternative. Since the data set $F_k(j\omega)$ corresponds to a real function of time, its negative frequency values are given by $\bar{F}_k(j\omega) = F_k(-j\omega)$. A second set of data, call it $I(j\omega)$, is constructed so that $\bar{I}(j\omega) = -I(-j\omega)$ and thus corresponds to an imaginary function of time. $I(j\omega)$ may then be added to $F_k(j\omega)$ and the sum expanded as a series of poles which are then constrained to occur in conjugate pairs. Noting that $F_k(j\omega)$ has an even real part and an odd imaginary part while $I(j\omega)$ has an odd real part and an even imaginary part, the pole residue expansion of $F_{comp}(j\omega)$ is immediately separable into components due to $F_k(j\omega)$ and $I(j\omega)$. Specifically, the pole residue expansion of $F_k(j\omega)$ may be written as

$$\sum_{m=1}^{M} \frac{(b_m + \bar{c}_m)/2}{j\omega - s_m} + \frac{(\bar{b}_m + c_m)/2}{j\omega - \bar{s}_m}$$

while the pole residue expansion of $I(j\omega)$ is

$$\sum_{m=1}^{M} \frac{(b_m - \bar{c}_m)/2}{j\omega - s_m} - \frac{(\bar{b}_m - c_m)/2}{j\omega - \bar{s}_m}$$

Thus, the second data set may be used to constrain the poles of $F_k(j\omega)$ without corrupting the residues.

2.2 Results. The procedure was tested on two objects, a sphere and a thin wire with length-to-diameter ratio of 50 to 1. For the case of the wire, poles were extracted from computer generated data associated with current measurement on the surface of a wire. The wire was divided along its length into 50 sections to facilitate analysis by the NEC program [5]. The wire was illuminated with a plane wave electromagnetic field, with the axis of the wire parallel to the direction of the electric field. The frequency of the illumination was stepped from $(\omega L/c\pi) = 0.2$ to 4.0 in steps of 0.1. The computer simulated current measurements of the 5th, 10th, 15th, 20th, and 25th sections were then individually subjected to a pole residue expansion. A pole

common to all five data sets was located at $(sa/c) = -0.087 + i0.8855$. The estimate of the pole as obtained from the five data sets agreed to this value to within 0.2 percent. This value is also in relative agreement with that obtained by Tesche [6]. For the sphere, the pole residue expansion was performed on computer generated data corresponding to current probe measurements at θ = 0, 10, 20, and 30 degrees, with $\omega a/c$ = 0.2 to 4 by 0.1. The location of the primary pole for the sphere may be determined analytically [1], and lies at $sa/c = -0.5 + i.866$.

Since the benefit of the algorithm may be attributable to increasing the signal-to-noise ratio, the amount of improvement which might be reasonably expected can be determined by first examining the increase in accuracy which occurs when the signal-to-noise ratio is directly varied. The data was corrupted with computer generated white, Gaussian noise at levels necessary to produce signal-to-noise ratios of -19 dB to 60 dB in steps of 2.0 dB. The signal-to-noise ratio is defined as the total energy in the original uncorrupted data sets, $F_k(j\omega)$, divided by the total expected noise energy contained in these data sets. The range of errors associated with the pole as extracted from the $F_k(j\omega)$ is shown in Fig. 1 for each noise level as falling between the two horizontal marks. The fact that the pole for the wire is much more resonant than the pole for the sphere permitted the SEM to be performed at much higher noise levels for the wire than for the sphere. The errors plotted in Fig. 1 for signal-to-noise ratios less than 15 dB are those associated with the wire data, while those for signal-to-noise ratios greater than 10 dB are associated with the sphere data. For each noise level a composite data set, $F_k(j\omega)$ was created using the above procedure, and the specific case of wire data with signal-to-noise ratio of 0 dB is shown in Fig. 2 as the dashed line. The pole was extracted from $F_{comp}(j\omega)$ and the error associated with this pole is marked with a triangle. A second composite data set was created using the estimates of the poles and residues as obtained via the first $F_{comp}(j\omega)$ and the error associated with the pole extracted from this second $F_{comp}(j\omega)$ is marked with a circle. Finally, the error associated with the pole as extracted from the original data set $F_k(j\omega)$, with the strongest excitation of the dominant mode (the 25th section for the wire and the θ = 0 case for the sphere) is marked with an X.

There are several important features which are exhibited in Fig. 1. The most noticeable is the apparent effect of the pole resonance on the accuracy of extraction of the desired pole; the expansion performed on the wire data with signal-to-noise ratio of 10 dB produced more accurate results than sphere data with signal-to-noise ratio of 50 dB. Additionally, the rate of improvement in accuracy of pole extraction with increasing signal-to-noise ratio is certainly not a smooth function. For the wire, there is a large jump at -2 dB and for the sphere there is a 15 dB region centered at 40 dB where increasing the signal-to-noise ratio has a negligible effect upon the accuracy of the pole extraction. Given these peculiarities, the improvement resulting from extracting the pole from $F_{comp}(j\omega)$ produces fairly reasonable results. In regions where the slope is steep the improvement resulting from extracting the pole from the composite data set produces very nice results (remembering that the percent error is graphed on a logarithmic scale), while in regions

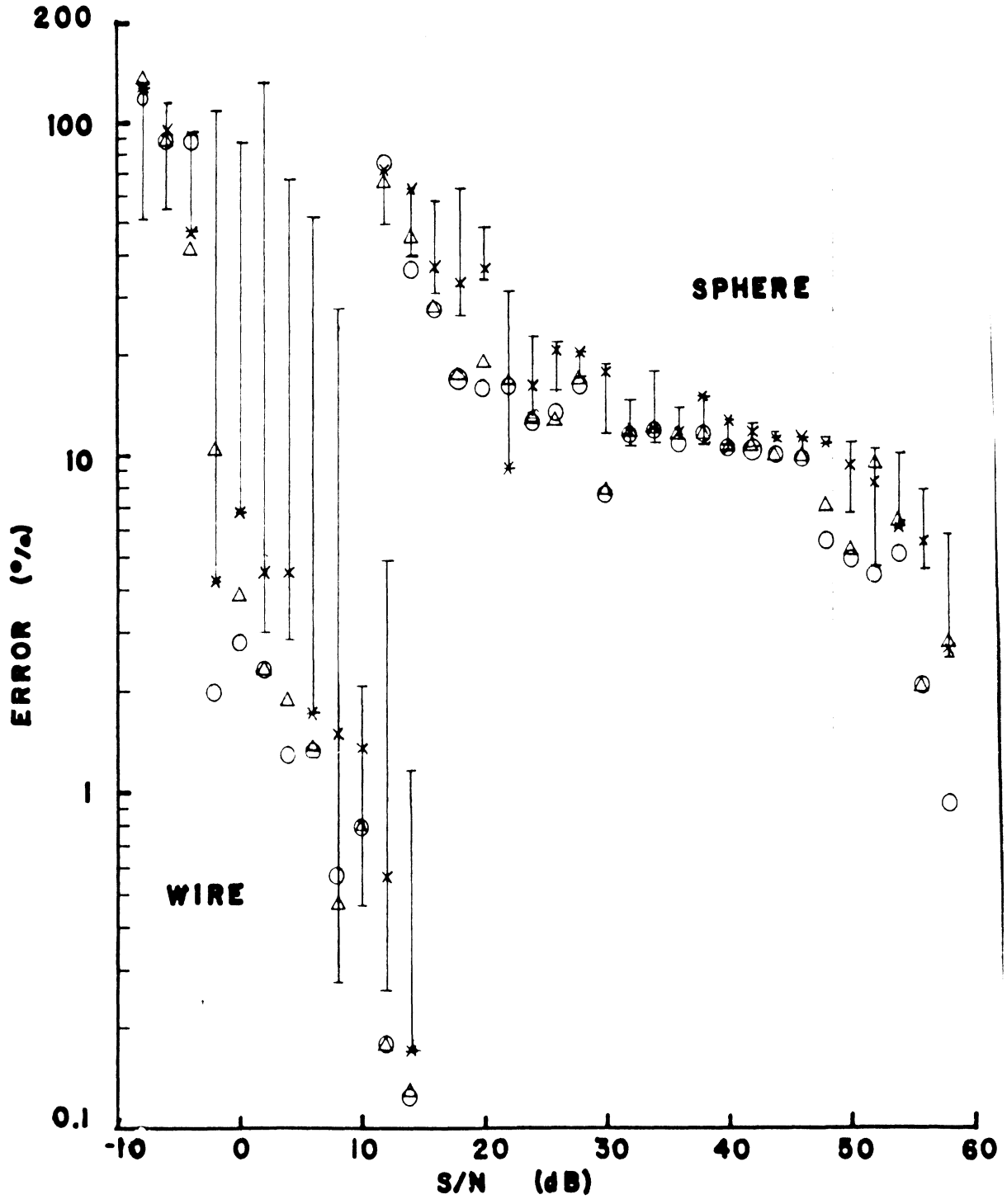

Fig. 1: Vertical bars delimit range of errors for poles extracted from $F_k(j\omega)$; X indicates error for pole extracted from $F_k(j\omega)$ with greatest S/N; triangle indicates error for pole extracted from first composite data set; circle indicates error for pole extracted from second composite data set.

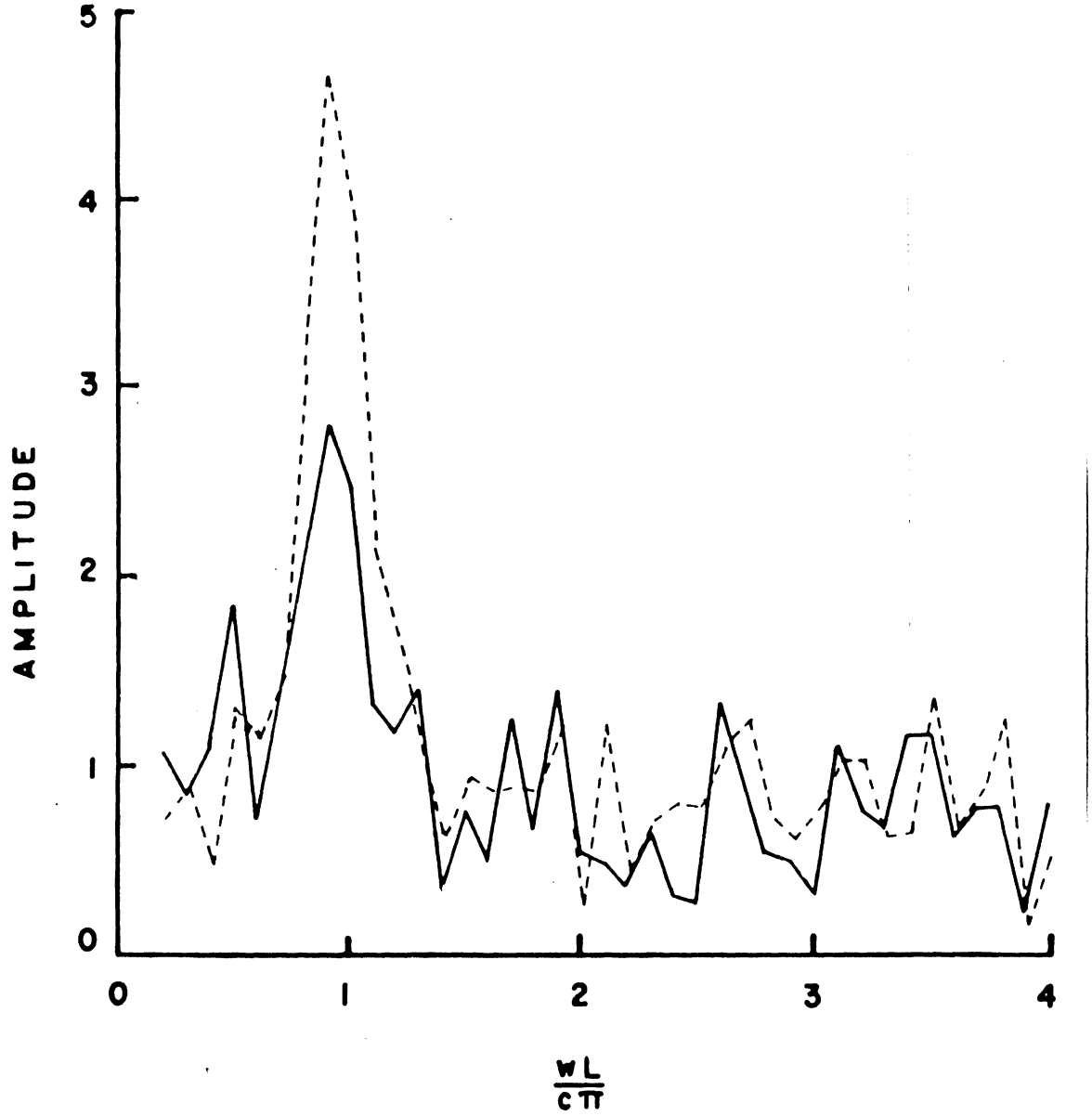

Fig. 2: Wire data for S/N = 0 dB. Solid line - 25th section; dashed line - composite data set.

where the slope is effectively zero, the improvement is negligible. Taken as a whole, the pole extracted from the composite data set is almost always better than that obtained from the data set with the greatest excitation of the desired mode, and quite often is better than any of the original estimates. This is with the exception of the region where all of the initial estimates of the pole are in error by greater than 50 percent, as this generally makes the associated residues effectively meaningless. However, when at least some of the estimates of the pole are accurate good results are still obtained. For example, at 0 dB the wire data has an estimate of the pole which is accurate to 7 percent and another estimate which has a greater than 80 percent error. Not only does the algorithm work in this case, but the iterative

application of the algorithm is able to use the improved estimate of the residues to obtain an even better estimate for the desired pole.

3. FILTERING OF DATA

3.1 Background. While experimenting with the method of Levy, Sanathanan, and Koerner [4], [5] (LSK algorithm) on sphere data, sparsely sampled data sets (39 points) and densely sampled data sets (189 points) with identical signal-to-noise ratios were found to yield poles and residues of comparable accuracies. This observation is somewhat disconcerting since the denser data set has almost five times the information of the sparse one and yet the LSK algorithm can not extract any better results. This fact has encouraged research into the possibilities of filtering the data before using it to find poles and residues. Unfortunately the use of filtering to improve the accuracy of the poles found by the LSK algorithm is somewhat more difficult than first expected. Indeed several different factors tend to make it so.

First, even if a filter is successful in improving the signal-to-noise ratio by, for example, what would otherwise be a respectable 10 dB, there is no assurance that the locations of the poles will improve. The authors found that sphere data with signal-to-noise ratios of 20 dB, 30 dB, and 40 dB all had comparable errors in the location of the poles. Only when the signal-to-noise ratio increased to 50 dB and 60 dB did the locations of the poles tend to improve. Another problem is that for sphere data very little noise is acceptable, as stated before the noise must be some 50 dB or 60 dB down from the signal for good results. This means that what ever filting scheme is used it must work on the noise without corrupting the signal. Ideally the filter should not alter the data at all in the limit as the noise tends to zero.

This places several limitations on the choice of the nature of the filter. For example, filters which exhibit a "phase shift" cannot be used since they will move the location of the resonant peaks in the data thus shifting the imaginary part of the pole. Also standard zero phase low pass filters cannot be used since they round out peaks in the data thus changing the real part of the pole to be less resonant. Additionally various types of transform filters which suffer from aliasing problems are unacceptable since they also corrupt the signal.

Despite these difficulties the authors have developed a filtering scheme which has had some success in improving the location of the poles for sphere data. This scheme is a type of finite window moving aperture filter. However, whereas the standard linear moving aperture filter [7] has a fixed weighting function, this scheme is a nonlinear filter where the weighting function is allowed to change from data point to data point.

The filtering process begins with discrete frequency domain data, $H(k)$, consisting of a signal plus noise, $S(k)$ and $N(k)$. The real part and imaginary part of the data will be filtered separately so that the analysis will concern only real quantities. A weighting function, $W_m(k)$ where integer k ranges from $-M$ to M, is used to generate filtered data, $\bar{H}(k)$ as follows:

$$\bar{H}(m) = \sum_{k=-M}^{M} W_m(k) H(m+k)$$

where M is the order of the filter. The filtered data is then used with the LSK algorithm to find poles and residues. The following section is a derivation of the weighting function $W_m(k)$.

3.2 Filter Derivation. The assumptions about the noise that will be made are as follows:

$$E[N(k)] = 0 \tag{1}$$

$$E[N(k)N(j)] = 0 \quad \text{if } k \neq j \tag{2}$$

$$E[N^2(k)] = \sigma^2 \quad . \tag{3}$$

Further, the following restrictions will apply to the weighting function

$$\sum_{k=-M}^{M} W_m(k) = 1 \tag{4}$$

$$W_m(k) = W_m(-k) \quad . \tag{5}$$

The weighting function will be an even function so that the peaks in the data will not shift in position.

It is also assumed that the signal is varying slowly enough compared to the width of the window so that it can be accurately represented by a quadratic expansion, that is:

$$S(m + k) = a_o + a_1 k + a_2 k^2 \quad ; \quad \text{for } -M \leq k \leq M \quad . \tag{6}$$

Of course the values of a_o, a_1, and a_2 will depend on the choice of m.

With this filter the estimate of S(m) will be

$$\tilde{a}_o = \sum_{k=-M}^{M} W_m(k)[S(m + k) + N(m + k)] \tag{7}$$

and so an error measure can be defined as:

$$\varepsilon = (a_o - \tilde{a}_o)^2 = \left\{ a_o - \sum_{k=-M}^{M} W_m(k)[S(m + k) + N(m + k)] \right\}^2 \quad . \tag{8}$$

Now under the previously stated assumptions the expected value of the error measure can be found

$$E(\varepsilon) = a_o^2 - 2a_o \sum_{k=-M}^{M} W_m(k)S(m + k) + \left[\sum_{k=-M}^{M} W_m(k)S(m + k) \right]^2 + \sigma^2 \sum_{k=-M}^{M} W_m^2(k)$$

$$E(\varepsilon) = a_2^2 \left[\sum_{k=-M}^{M} k^2 W_m(k) \right]^2 + \sigma^2 \sum_{k=-M}^{M} W_m^2(k)$$

$$= 4a_2^2 \left[\sum_{k=1}^{M} k^2 W_m(k) \right]^2 + 2\sigma^2 \sum_{k=1}^{M} W_m^2(k) + \sigma^2 \left[1 - 2 \sum_{k=1}^{M} W_m(k) \right]^2 . \quad (9)$$

By the proper choice of the $W_m(k)$ values it is desirable to minimize the expected value of the error measure. Therefore the values of the weighting function must satisfy

$$\frac{\partial E(\varepsilon)}{\partial W_m(J)} = 8J^2 a_2^2 \sum_{k=1}^{M} k^2 W_m(k) + 4\sigma^2 W_m(J) - 4\sigma^2 \left[1 - 2 \sum_{k=1}^{M} W_m(k) \right]$$

$$= 4\sigma^2 [W_m(J) - 1] + \sum_{k=1}^{M} W_m(k) [8\sigma^2 + 8J^2 a_2^2 k^2] = 0 \quad ;$$

$$\text{for } J = 1,2,\ldots,M \quad . \quad (10)$$

This leads to M linear equations which can be solved for the M unknowns $W_m(1)$ through $W_m(M)$. The values of $W_m(0)$ can be found from relation (4). For example, when

$$\underline{M = 1}$$

$$W_m(0) = \frac{2a_2^2 + \sigma^2}{2a_2^2 + 3\sigma^2}$$

$$W_m(1) = \frac{\sigma^2}{2a_2^2 + 3\sigma^2}$$

$$\underline{M = 2}$$

$$W_m(0) = \frac{34\, a_2^2 + \sigma^2}{5(14\, a_2^2 + \sigma^2)}$$

$$W_m(1) = \frac{24\, a_2^2 + \sigma^2}{5(14\, a_2^2 + \sigma^2)}$$

$$W_m(2) = \frac{-6\, a_2^2 + \sigma^2}{5(14\, a_2^2 + \sigma^2)}$$

Notice that if $a_2 = 0$, that is the signal has no curvature, the weighting function is uniform so that the most noise possible can be eliminated. Notice also if $\sigma^2 = 0$ then the estimate of S(m) is correct to the extent that the data can be represented by a quadratic expansion.

Therefore, to find the optimum weighting function for a given size window the variance of the noise and the a_2 coefficient for each data point must be known. Usually the noise variance is at least approximately known,

however the a_2 coefficients require some knowledge of the signal without noise. Hence the values of these a_2 coefficients can only be estimated from the signal plus noise data. For this reason the filter will only work on data with a relatively high signal-to-noise ratio (> = 20 dB).

The a_2 coefficients are found by a best curve fit with the data over the range of the window plus one point. The error between the curve fit and data is defined as:

$$\eta = \sum_{k=-M-1}^{M+1} [H(m+k) - (a_o + a_1 k + a_2 k^2 + a_3 k^3)]^2 \tag{11}$$

The coefficients a_o through a_3 are then chosen to minimize η. This leads to the following equation for the a_2 coefficients

$$a_2 = \frac{\sum k^2 \cdot \sum H(m+k) - \sum 1 \cdot \sum k^2 H(m+k)}{\left(\sum k^2\right)^2 - \sum 1 \cdot \sum k^4} \tag{12}$$

Using this relation and knowing the noise variance σ^2 weighting functions for each data point can be generated and the data filtered.

3.3 Results of Filtering. Figure 3 shows noisy sphere data (S/N = 40 dB), and the same data filtered with M = 1 and M = 2 filters. The filter was tested on various sphere data with noise. On sparsely sampled data (39 points) no improvement was made in the location of the first dominant pole for any signal-to-noise ratio. On more densely sampled data (189 points) improvement was made when the signal-to-noise ratio was between 50 dB and 60 dB. With signal-to-noise ratios greater than 60 dB the unfiltered data yielded the correct result to one percent accuracy anyway and filtering did not improve it. However, with a 55 dB signal-to-noise ratio the average error of the first dominant pole was 7 percent. After filtering with M = 1 this average error fell to 3.5 percent and with M = 2 fell to 2.5 percent. It should be noted that similar improvements can be obtained with more traditional filters, such as a uniformly weighted moving aperture, for this data. However a uniform weighted filter actually increases errors with other types of data sets, whereas our filter will not. For example, the unfiltered error of the 39 point data set with 55 dB signal-to-noise ratio averaged 7 percent, but after filtering with a uniform aperture with M = 1 the error went to 8 percent and with M = 2 climbed to 31 percent. Our filter however, had an average error of still 7 percent for both M = 1 and M = 2 filters. In conclusion the presented filter can improve the location of poles found with the LSK algorithm if both the density of the sampled data and the initial signal-to-noise ratio are sufficient. In addition this filter will not statistically increase errors of the poles and residues where more traditional filters may.

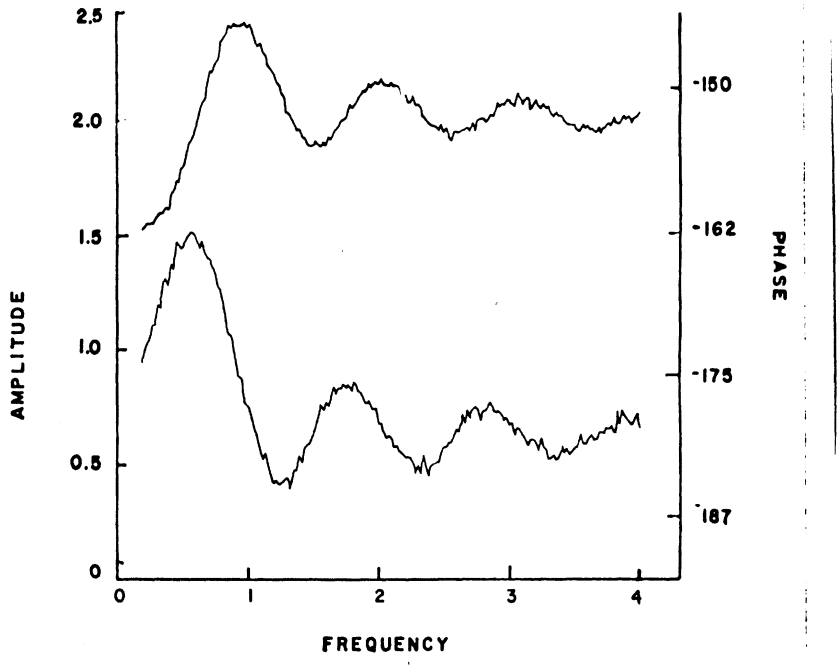

Fig. 3(a): Sphere data filtering. (Unfiltered data)

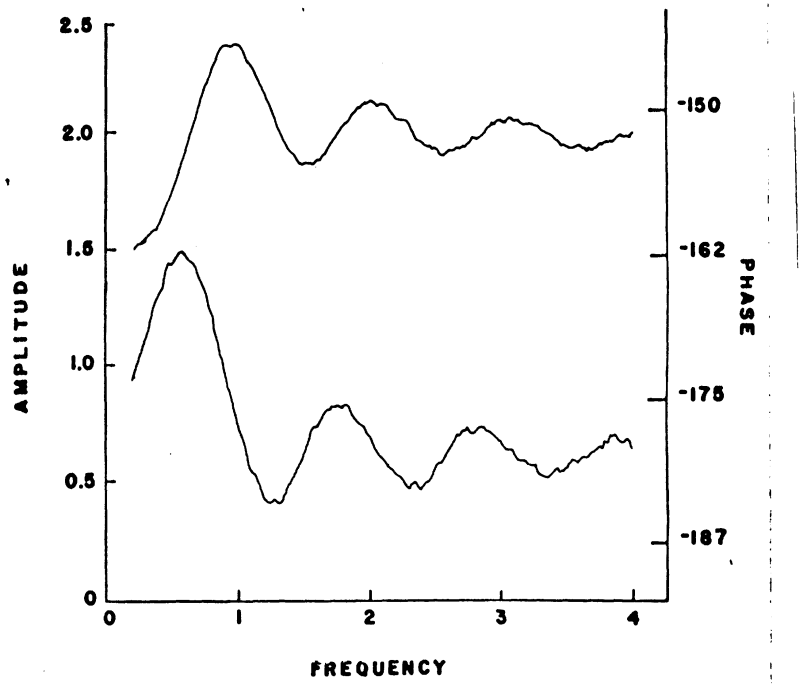

Fig. 3(b): Sphere data filtering. (Filtered data M = 1)

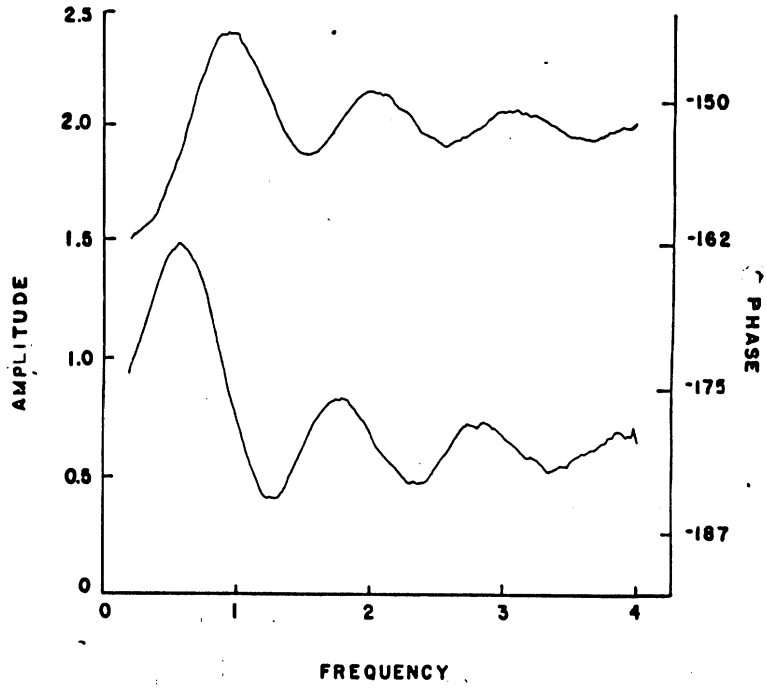

Fig. 3(c): Sphere data filtering. (Filtered data M = 2)

ACKNOWLEDGEMENT

This work was supported by the Air Force Weapons Laboratory under Contract No. F29601-82-K-0024..

REFERENCES

1. Pond, J. M. and T.B.A. Senior (1982), "Determination of SEM poles from frequency responses", Electromagnetics, Vol. 2, 55-67.

2. Ksienski, D. A. (1984), "Pole and residue extraction from measured data in the frequency domain using multiple data sets", Radio Science, Vol. 20, 13-19.

3. Levy, E. C. (1959), "Complex curve fitting", IRE Trans. on Automatic Control, Vol. AC-4, 37-44.

4. Sanathanan, C. K. and J. Koerner (1963), "Transfer function synthesis as a ratio of two complex polynomials", IEEE Trans. on Automatic Control, Vol. AC-8, 56-58.

5. Burke, G. J. and A. J. Poggio, (1977), Interaction Note 363, "Numerical Electromagnetic Code (NEC)", Air Force Weapons Laboratory, Albuquerque, N.M.

6. Tesche, F. M. (1972), "On the singularity expansion method as applied to Electromagnetic scattering from thin-wires", Interaction Note 102, Air Force Weapons Laboratory, Albuquerque, N.M.

7. Oppenheim, A. B. and R. W. Schafer (1975), Digital Signal Processing, pp. 237-269, Prentice-Hall, Englewood Cliffs, N.Y.