MATHEMATICS NOTES

NOTE 80

October 1984

Pole and Residue Extraction from Measured Data in the Frequency Domain Using Multiple Data Sets

David A. Ksienski
The Radiation Laboratory
Department of Electrical and Computer Engineering
The University of Michigan
Ann Arbor, Michigan 48109

ABSTRACT

When using the singularity expansion method on measured data, the accuracy of the results is often severely limited by the noise and clutter present in the measurement. An effective method of reducing these contaminations is repeating and averaging the measurements. The effectiveness of this method is related to the number of measurements which are combined. The present paper develops a method which increases the set of measurements which may be combined from the set of identical measurements to all measurements made on the target. The measurements are combined using an optimal weighting scheme to provide a superior estimate for a specified pole. An algorithm is described which increases the accuracy of pole locations and accurately refers the residues to these new pole locations. The algorithm is used to extract a single pole and residue, and additional poles and residues may be obtained through iterative application of the algorithm.

1. INTRODUCTION

The singularity expansion method as applied to electromagnetic scattering is a method of characterizing the response of a target in terms of transfer function parameters such as poles and residues. To this end, the poles are assumed to be invariant, while the residues vary with the particular measurement performed and the type (including direction and polarization) of illumination. At least in theory, the accurate location of a few poles could assist in determining the identity of the target, but in practice the supposedly invariant poles tend to move about with changes in the target's orientation and illumination. If the noise and clutter are nondeterministic, and if it is feasible to repeat the measurement, both the noise and clutter can be reduced by repeating the measurement several times and averaging the associated signals. If the usual assumptions about the noise and clutter (such as distributions which have zero mean and are independently and identically distributed) are satisfied, combining K measurements will result in an increase in the signal to noise ratio by a factor of K. By taking K sufficiently large, the data may be quieted enough to guarantee accurate pole and residue extraction. Of course, in practice, obtaining a large number, K, of measurements, may be somewhat expensive, or in some situations, impossible. For example, if measurements are made of a moving target such as an airplane, the requirement that the measurement be repeatable can be satisfied approximately at best, and then only for a limited amount of time. Conversely, if the measurement of a stationary object is repeated, any stray reflections from other stationary objects (or, in the case of near-field

measurements, probe-target interaction) will be repeated exactly. Under these circumstances, the clutter is deterministic and the energy associated with the clutter cannot be reduced by repeating and averaging the measurements. On the other hand, if the noise and clutter are nondeterministic, the signal to noise ratio will increase as additional data sets are incorporated into the composite data set. It is therefore desirable to include not only similar measurements, which may be combined using strict waveform averaging, but also those associated with different illuminations of the target.

This paper presents an extension of measurement averaging that eliminates the requirement of identical measurements and thus permits further increases in signal to noise ratio by increasing the number of data sets which may be properly combined. The increased signal to noise ratio is obtained for the portion of the original signals associated with a specified pole.

2. FORMULATION

Using the standard pole residue expansion, the data as a function of the circular frequency ω may be represented as

$$F_k(\mathrm{j}\omega) = \sum_{m=1}^{M} \frac{a_{km}}{\mathrm{j}\omega - s_m} + N_k(\mathrm{j}\omega) \tag{1}$$

where k is an index to the various measurements and $N_k(\mathrm{j}\omega)$ represents noise and clutter. 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. Assuming a linear combination with arbitrary complex weighting coefficients, the composite data

set is

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

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

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

where

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

In general, it will only be possible to maximize one b_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. However, since the specification of the pole is arbitrary, several poles may be obtained through successively emphasizing different b_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

$$|b_1|^2 = |\mathbf{w}^T \mathbf{a}|^2 \tag{5}$$

where $\mathbf{w} = [w_1, w_2, \ldots, w_K]^T$ and $\mathbf{a} = [a_{11}, a_{21}, \ldots, a_{K1}]^T$. Without loss in generality, the weighting vector may be constrained to $\overline{\mathbf{w}}^T \mathbf{w} = 1$. Then, noting the dot product formulation of (5), the maximum b_1 will occur for

$$\mathbf{w} = \frac{\bar{\mathbf{a}}}{|\mathbf{a}|} \tag{6}$$

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

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

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

and assuming the variance of the noise is equal to σ^2 in each of the original measurements,

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

Thus the energy associated with the noise will not increase with **w** normalized as in (6). 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.

3. IMPLEMENTATION

The above formulation assumes initial estimates of residues in order to compute the weighting coefficients. The algorithm is fairly insensitive to errors in these estimates of the residues and, in fact, a single complex error factor

associated with all of the residues will not adversely affect the weighting vector. The effect of variations in the real part of the pole is minimized by normalizing the residues by the real part of the associated poles, i.e., (6) is used with **a** redefined as

$$\mathbf{a} = \left[\frac{a_{11}}{\mathrm{Re}(s_{11})} , \frac{a_{21}}{\mathrm{Re}(s_{21})} , \ldots , \frac{a_{K1}}{\mathrm{Re}(s_{K1})} \right]^T \tag{7}$$

where s_{i1} refers to the estimate of the pole s_1 obtained from the ith data set. After the composite data set is formed, it must be subjected to a pole and residue extraction procedure. The algorithm due to *Levy* [1959] and *Sanathanan and Koerner* [1963] 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. Recently, *Tao and Zunde* [1981] extended the method of *Levy* [1969] to distributions of poles and zeroes which have no symmetry constraints. Although the algorithm of Tao and Zunde does accommodate the above data set, it is somewhat more general than required. Specifically, it does not require the poles to occur in conjugate pairs, but as may be seen from (2), $F_{comp}(j\omega)$ does preserve the conjugate symmetry of the pole pairs. To more closely match the known parameterization of the data, an algorithm was produced which relaxes the conjugate symmetry constraint on the residues while maintaining the conjugate symmetry constraint on the poles, and the derivation is given in the Appendix.

The implementation of the algorithm requires negative frequency information which is obtained from the original data sets by employing conjugate symmetry. After the accuracy of the pole location is improved, the next step is to refer the residues to the new pole estimate. Constraining a pole to a particular location is not an easy task [*Pond and Senior*, 1982], but the present algorithm provides a simple alternative. Since an original data set corresponds to a real function of time, it is an even function of frequency. A second set of data which is an odd function of frequency may be added to the original data and the sum expanded as a series of poles. Noting that the poles are constrained to occur in conjugate pairs, the component of the summation associated with the original data may be obtained by retaining the even part of the residues. Both the original data and the second set are constrained to have the same poles in the expansion, and by appropriate placement of the poles associated with the second set of data, it is possible to constrain the pole locations in the expansion of the original data. The residues are then obtained relative to these new constrained pole locations without any apparent adverse side effects. An additional advantage is that the accuracy of the location of the adjacent pole is also improved.

4. TEST OF ALGORITHM

It is generally accepted (e.g., *Van Blaricum and Mittra* [1978]) that an increased signal to noise ratio will improve the accuracy of pole extraction. In the preceding section it was shown that the present algorithm improves the signal

to noise ratio for the signal associated with a specified pole, and this is now demonstrated using nonresonant sparse data and employing only 4 data sets. Both nonresonance and sparseness severely limit the accuracy of extracted poles, and produce a worst case test of the algorithm. Additionally, using only 4 data sets limits the amount of gain in the signal to noise ratio to less than a factor of 4 or approximately 6 dB.

The data was computer generated and corresponds to a probe measurement of the current on the surface of a perfectly conducting sphere. The actual function used was the $T_2(\theta)$ function as defined in *Bowman et. al.* [1969]. Data was obtained for $\omega a/c = .2$ to 4 in .1 increments, where c is the speed of light and a is the radius of the sphere. The 4 data sets correspond to current measurements at $\theta = 0$, 10, 20, and 30 degrees. Gaussian noise was added to the data, with the variance of the noise being stepped from 10^{-2} to 10^{-7} in $10^{-0.2}$ multiplicative increments. The pole which was enhanced was the dominant pole, located at -0.5+j0.866. The error associated with extracting this pole from the $\theta = 0$ data sets is shown in Figure 1. The error is displayed versus N_p, where $N_p = 10 \times \log$ (variance of noise), so as to provide a dB scale. As can be seen, although there is improvement in the accuracy of the pole location with increasing signal to noise ratio, an increase in signal to noise ratio by even as much as 10 dB does not always produce an increase in the accuracy of the pole location, for example in the region N_p = -20 to -30. Interestingly, this is about the magnitude of the noise encountered in a laboratory measurement, and certainly illustrates the difficulty of extracting poles and residues from

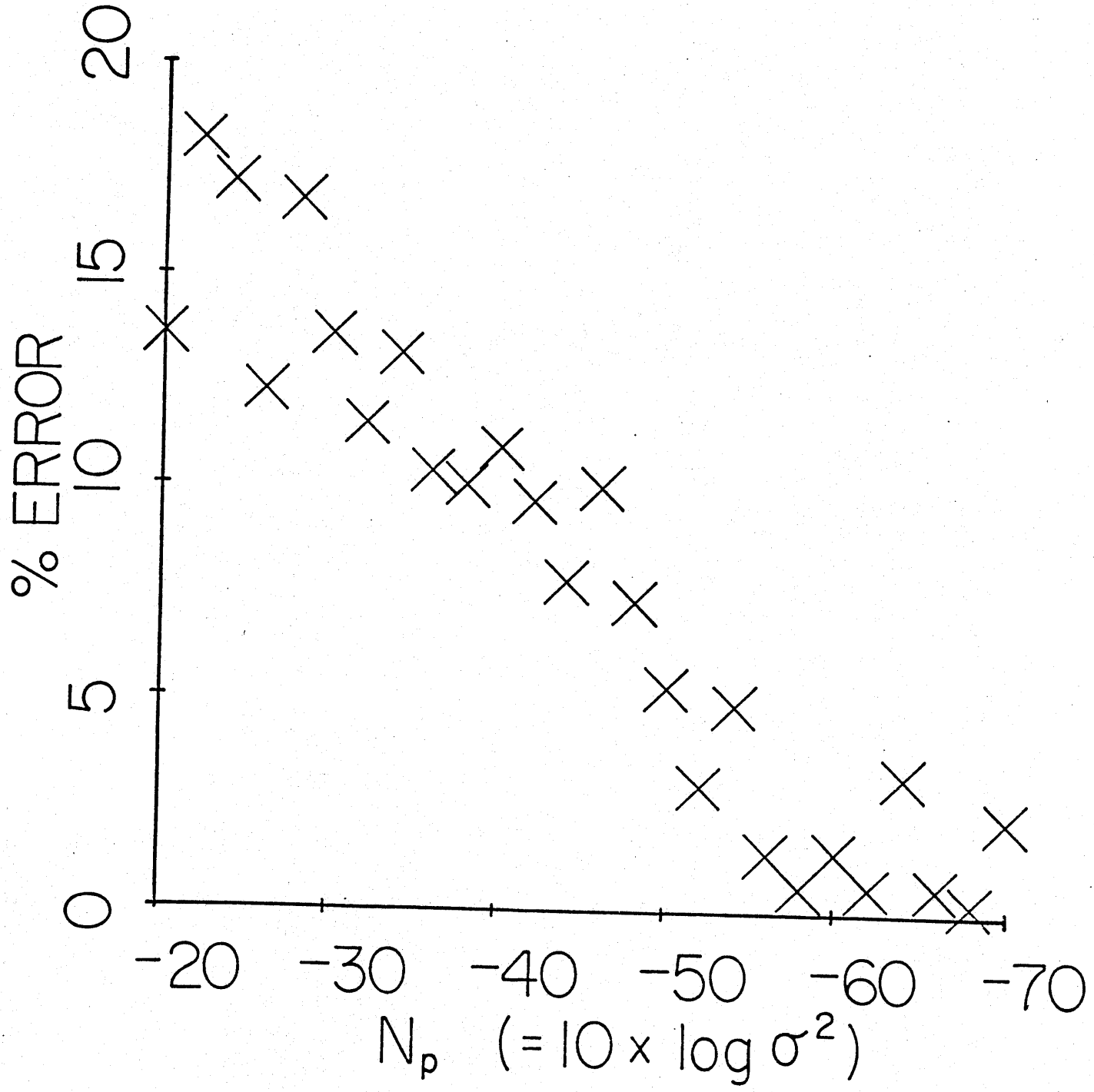

Figure 1. Error in the first pole.

experimentally measured sphere data. As stated above, this is a worst case test, and considerably better results are obtainable from dense data which is prefiltered, the latter effectively resulting in an increased signal to noise ratio. More resonant targets also offer much better results. However, for the data at hand, the region in the middle of the graph is the only area where a 6 dB improvement will be easily observable. Of course, by working with several runs and noting the average change in pole location, the 6 dB improvement would produce more accurate pole extractions anywhere on the graph, but it was desired to affect improvements which would be apparent from a single run. Thus, 4 data sets were combined, corresponding to $\theta = 0$, 10, 20, and 30 degrees, and this was performed for 20 different noise levels associated with N_p being stepped from -40 to -58 in increments of 2. The weighting coefficients were obtained from preliminary pole-residue extractions from the noisy data using (6) and (7). The improvement in the accuracy of pole location is shown in Table 1. The improvement is calculated relative to the average error of the extracted pole location obtained from the original 4 data sets. The irregularity of the improvement is fairly consistent with the trend seen in Figure 1. For comparison, the reduction in error accruing from the use of the consensus pole technique suggested by *Pearson and Roberson* [1978] is also listed. This technique involves averaging the pole location estimates produced from the initial runs, and though it always produces an improvement, the amount of improvement is seen, at least in the region tested, to be much less than that resulting from processing a composite data set. The limitation of the consensus

N_P	Percent reduction in error in the first pole	
	Using Composite Data Set	Using Consensus Pole
-40	18.13	7.78
-42	7.33	2.61
-44	31.87	3.47
-46	26.67	2.86
-48	70.24	2.81
-50	65.34	1.95
-52	54.98	0.73
-54	73.34	3.25
-56	79.73	36.47
-58	51.98	3.25

Table 1. Improvement in the pole location as extracted from the composite data set and as obtained from the consensus pole technique.

pole technique is largely a result of the bias which seems to afflict most least-squares curve fitting pole extraction algorithms. Since a bias is not reduced by averaging, the consensus pole suffers from the same bias as the initial estimates of the pole. The improvement of the composite pole over the consensus pole as well as the poles extracted from the original data sets is shown graphically in Figure 2 for the case of N_p = -50.

Finally, it was desired to refer the residues to the new pole locations in cases where a substantial improvement in pole location was obtained. The greatest improvement in pole location occurfed in the region of N_p = -48 to -58, and for these cases the residues were referred to the new pole locations. The attendant improvement in the accuracy of the residue is shown in Table 2. In addition, as a result of constraining the primary pole to a more accurate location, the accuracy of the extracted nearby pole is also improved as shown in Table 2.

5. DISCUSSION

The combination of dissimilar measurements of the same target makes it possible to reduce the energy associated with clutter and noise. An optimum method of combining the measurements has been developed and tested under worst case conditions and appears to be an effective method of increasing the accuracy of the extracted poles and residues. The method has been compared to the consensus pole technique of *Pearson and Roberson* [1978] and found superior over the range of noise levels tested. Additional improvements are presumed to be obtainable by using a more dense sampling rate, perhaps coupled with a

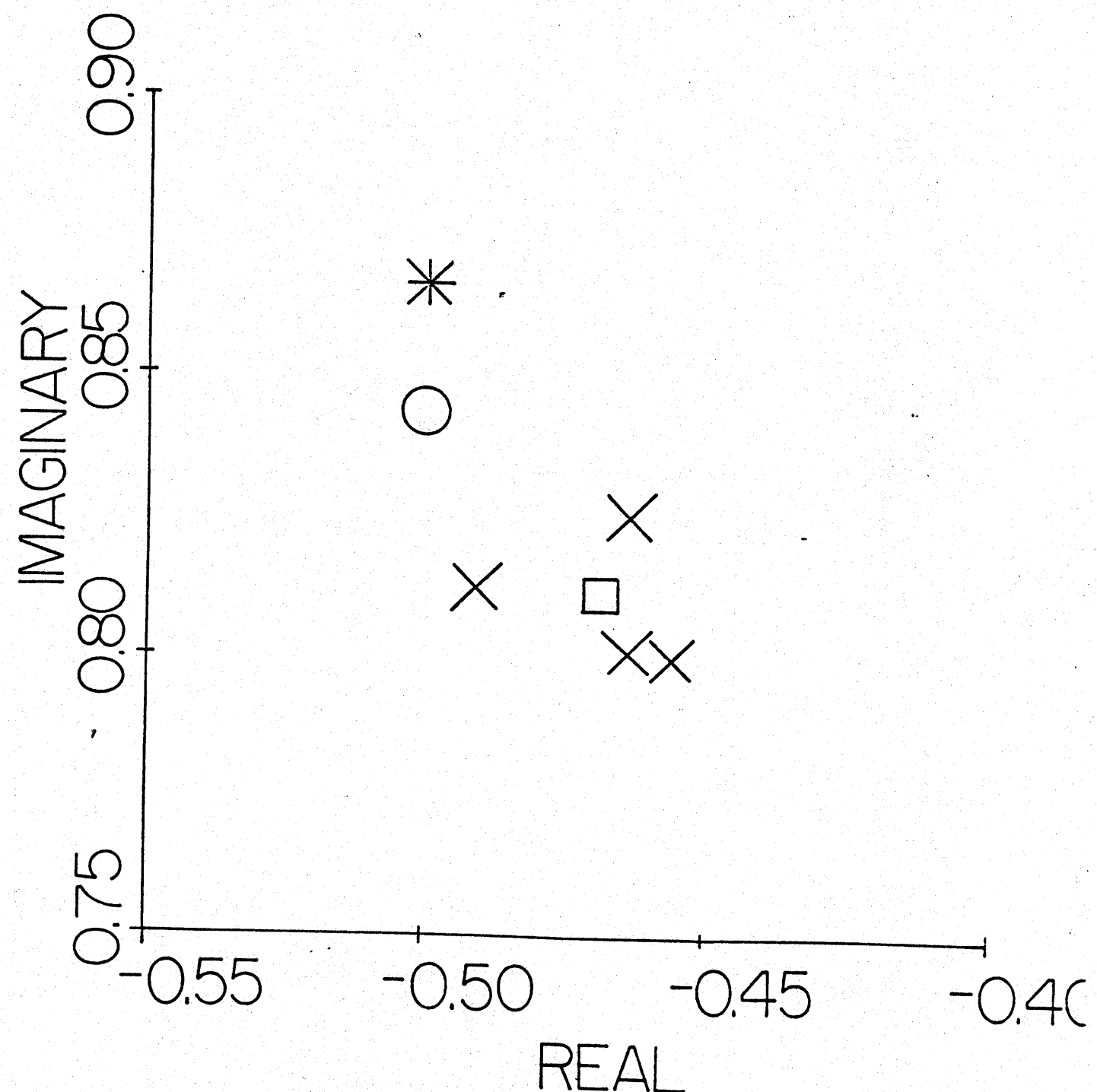

Figure 2. Locations of the first pole, N_p = -50: true pole (✳), poles extracted from original data sets (××), pole extracted from composite data set (○), and consensus pole (□).

N_p	Percent reduction in error	
	Residue of First Pole	Second Pole
-48	61.7	76.9
-50	39.1	62.4
-52	59.3	53.1
-54	74.6	55.0
-56	75.4	56.6
-58	16.1	34.1

Table 2. Improvement in the residue of the first pole and the location of the second pole resulting from constraining the location of the first pole.

prefiltering scheme, as well as by increasing the number of data sets which are combined into the composite data set. Although only one pole was specifically extracted, the nearby pole was also obtained with increased accuracy, and it is reasonable to expect further improvements from iterative application of the above algorithm. Finally, it should be noted that the algorithm is entirely linear and therefore could also be extended to time domain singularity expansion methods.

ACKNOWLEDGEMENTS

The author is grateful to Professor T. B. A. Senior and Dr. J. M. Pond for an introduction to the singularity expansion method and to Mr. T. M. Willis III for assistance with the computer programming. This work was supported by the Air Force Weapons Laboratory under Contract No. F29601-82-K-0021.

APPENDIX A

The following derivation parallels the reformulation of the work of *Levy* [1959] and *Sanathanan and Koerner* [1963] contained in *Lawrence* [1979]. Let $X(j\omega)$ and $L(j\omega)$ represent the measured data and the least-squares curve fit to the measured data, respectively, where $L(j\omega)$ is formulated as

$$L(j\omega) = \frac{(\boldsymbol{\alpha}^T + j\boldsymbol{\gamma}^T)\mathbf{p}(j\omega)}{1 + \boldsymbol{\beta}^T\mathbf{q}(j\omega)} \tag{A.1}$$

with

$$\boldsymbol{\alpha} = [a_0, a_1, \ldots, a_N]^T$$

$$\boldsymbol{\beta} = [b_1, b_2, \ldots, b_M]^T$$

$$\boldsymbol{\gamma} = [c_0, c_1, \ldots, c_N]^T$$

$$\mathbf{p}(j\omega) = [1, j\omega, (j\omega)^2, \ldots, (j\omega)^N]^T$$

$$\mathbf{q}(j\omega) = [j\omega, (j\omega)^2, (j\omega)^3, \ldots, (j\omega)^M]^T$$

and α, $\boldsymbol{\beta}$, and $\boldsymbol{\gamma}$ are real. This rational function expression of $L(j\omega)$ is necessary for the following derivation, however, $L(j\omega)$ is equivalent to an M-pole expansion. The inclusion of $\boldsymbol{\gamma}$ eliminates the conjugate symmetry constraint on the residues while maintaining the conjugate symmetry of the poles. The best fit curve, $L(j\omega)$, is obtained by minimizing $||X(j\omega)-L(j\omega)||^2 = ||\epsilon(j\omega)||^2$.

$$X(j\omega) = \frac{(\boldsymbol{\alpha}^T + j\boldsymbol{\gamma}^T)\mathbf{p}(j\omega)}{1 + \boldsymbol{\beta}^T\mathbf{q}(j\omega)} + \epsilon(j\omega) \tag{A.2}$$

$$X(j\omega)(1 + \boldsymbol{\beta}^T\mathbf{q}(j\omega)) - (\boldsymbol{\alpha}^T + j\boldsymbol{\gamma}^T)\mathbf{p}(j\omega) = \epsilon(j\omega)(1 + \boldsymbol{\beta}^T\mathbf{q}(j\omega)) \tag{A.3}$$

To minimize $\epsilon(j\omega)$ in the least-squares sense, first minimize J.

$$J = \sum_{k=1}^{K} \left|\frac{\epsilon(j\omega_k)}{d(j\omega_k)} (1 + \boldsymbol{\beta}^T\mathbf{q}(j\omega_k))\right|^2 \tag{A.4}$$

where K is the number of data points and $d(j\omega_k)$ is for the moment set to unity but with succeeding iterations converges to $(1 + \boldsymbol{\beta}^T\mathbf{q}(j\omega_k))$. Thus the initial solution will be a least-squares solution weighted by $(1 + \boldsymbol{\beta}^T\mathbf{q}(j\omega_k))$ and successive solutions will converge to an unweighted least-squares solution.

$$J = \sum_{k=1}^{K} \frac{1}{|d(j\omega_k)|^2} |X(j\omega_k)(1 + \beta^T q(j\omega_k)) - (\alpha^T + j\gamma^T)p(j\omega_k)|^2 \qquad (A.5)$$

$$= \sum_{k=1}^{K} \frac{1}{|d|^2} [|X|^2(1 + \beta^T q + \beta^T \bar{q} + \beta^T q \beta^T \bar{q})$$

$$-X(\alpha^T \bar{p} - j\gamma^T \bar{p} + \beta^T q \alpha^T \bar{p} - \beta^T q j\gamma^T \bar{p})$$

$$-\bar{X}(\alpha^T p + j\gamma^T p + \beta^T \bar{q} \alpha^T p + \beta^T \bar{q} j\gamma^T p)$$

$$+ (\alpha^T p \alpha^T \bar{p} - \alpha^T p j\gamma^T \bar{p} + j\gamma^T p \alpha^T \bar{p} + \gamma^T p \gamma^T \bar{p})]$$

where the functional dependence on $(j\omega_k)$ has been suppressed for notational economy. To minimize J,

$$\frac{\partial J}{\partial a_i} = \frac{\partial J}{\partial b_i} = \frac{\partial J}{\partial c_i} = 0 \qquad (A.6)$$

for all i, which results in a set of 2(N+1)+M linear equations with an equal number of unknowns. The simultaneous solution of these equations yields a pole-zero parameterization of the least squares curve fit to the data. To express these equations in the more compact form of a matrix equation, the scalar equations are first grouped into vector equations.

$$\frac{\partial J}{\partial \alpha} = \sum_{k=1}^{K} \frac{1}{|d|^2} [-X(\bar{p} + \bar{p}\beta^T q) - \bar{X}(p + p\beta^T \bar{q})$$

$$+ p\alpha^T \bar{p} + \bar{p}\alpha^T p - jp\gamma^T \bar{p} + j\bar{p}\gamma^T p] = 0$$

$$\frac{\partial J}{\partial \gamma} = \sum_{k=1}^{K} \frac{1}{|d|^2} [-X(-j\bar{p} - j\bar{p}\beta^T q) - \bar{X}(jp + jp\beta^T \bar{q})$$

$$+ jp\alpha^T \bar{p} - j\bar{p}\alpha^T p + p\gamma^T \bar{p} + \bar{p}\gamma^T p] = 0$$

$$\frac{\partial J}{\partial \beta} = \sum_{k=1}^{K} \frac{1}{|d|^2} [-X(-\overline{Xq} - \overline{Xq}\beta^T q) - \bar{X}(-Xq - Xq\beta^T \bar{q})$$

$$- Xq\alpha^T \bar{p} - \overline{Xq}\alpha^T p + jXq\gamma^T \bar{p} - j\overline{Xq}\gamma^T p] = 0$$

These vector equations may then be combined into

$$\frac{\partial J}{\partial v} = \sum_{k=1}^{K} \frac{1}{|d|^2} [-X(\bar{u} + \bar{u}\beta^T q) - \bar{X}(u + u\beta^T \bar{q})$$

$$+ u\alpha^T \bar{p} + \bar{u}\alpha^T p - ju\gamma^T \bar{p} + j\bar{u}\gamma^T p] = 0$$

where $u = [p^T; jp^T; -Xq^T]^T$ and $v = [\alpha^T; \gamma^T; \beta^T]^T$. Rearranging terms yields

$$\sum_{k=1}^{K} \frac{1}{|d|^2} \mathrm{Re}(X\bar{u}) = \left[\sum_{k=1}^{K} \frac{1}{|d|^2} \mathrm{Re}(u\bar{u}^T)\right] v$$

which is a matrix equation which may be solved for v. The initial solution of this equation, with d set equal to 1, represents a weighted least squares solution. If the equation is iteratively solved with d set equal to $1+\beta^T q$, the solution will converge to an unweighted least squares solution.

REFERENCES

Bowman, J. J., T. B. A. Senior and P. L. E. Uslenghi (1969), *Electromagnetic and acoustic scattering*, pp. 395-401, North-Holland, Amsterdam.

Lawrence, P. J., and G. J. Rogers (1979), Sequential transfer function synthesis from measured data, *Proc. IEEE*, Vol. 126, No. 1, 104-106.

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

Pearson, L. W. and D. R. Roberson (1978), Extrapolation of induced transient surface currents on a scatterer to general band-limited excitation conditions, *Interaction Note* 360.

Pond, J. M. and T. B. A. Senior (1982), Determination of SEM poles from frequency responses, *Electromagnetics*, Vol. 2, 55-67, also as Interaction Note 408, July 1981.

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

Tao, K. M. and P. Zunde (1981), On Levy's identification, its generalization and applications, *Int. J. Systems Science*, Vol 12, No. 11, 1317-1334.

Van Blaricum, M. L. and R. Mittra (1978), Problems and solutions associated with Prony's method for processing transient data, *IEEE Trans. Antennas Propagat.*, Vol AP-26, 174-182.