MATHEMATICS NOTES

Note 67

7 October 1977

PRONY ANALYSIS IN THE PRESENCE OF NOISE*

G. J. Scrivner**
Computer Sciences Corporation
1400 San Mateo Boulevard, S.E.
Albuquerque, New Mexico 87108

Abstract

This report studies a methodology, based on Prony's algorithm, for extracting complex frequencies and associated residues directly from transient response data. The methodology is not new and has been receiving increasing attention due to its potential usefulness in a variety of different applications in quite diverse scientific disciplines. Although this particular study was to evaluate the potential usefulness of Prony's method for analysis of System Generated Electromagnetic Pulse (SGEMP) experimental data, the results obtained are of a much broader nature. The report presents a classical derivation of the matrix equation which result if one uses Prony's algorithm in conjunction with a least squares criterion. The resulting matrix is studied with respect to the nature of its eigenvalue structure and an error analysis is developed which highlights the importance of this structure. The influence of noise on the eigenvalue structure and its impact on the determination of the number of poles in the data are examined. Optimization concepts relative to advantageous modification of the eigenvalue structure of the Prony matrix are qualitatively discussed. Difficulties associated with the two stage application of the least squares method in conventional Prony analysis are cited and an iterative method of removing this shortcoming described. Numerical calculations illustrating the majority of the above concepts are presented and discussed.

* The research in this note was performed under contract F29601-76-C-0034 with the Air Force Weapons Laboratory.

**Presently with Sandia National Laboratories

CONTENTS

TABLES

TABLES - Continued

I. INTRODUCTION

This report studies a methodology based on Prony's algorithm (ref 1) for extracting complex frequencies and associated residues directly from transient response data. The work was motivated by an interest in establishing the potential usefulnesses of Prony's method for application to experimentally generated System Generated Electromagnetic Pulse (SGEMP) data. However, the results obtained are of a much broader nature and hopefully will be of use to investigators in a variety of scientific fields.

Since SGEMP experimental data is inherently noisy, this study concentrates heavily upon this problem and its implication with regard to the Prony methodology. As a result, throughout this document it will be assumed that Prony's algorithm is being applied in conjunction with a least squares criterion.

The structure of this document is briefly outlined as follows:

In Section 2, the Prony methodology is developed and symbol convention is established. In addition, properties of the eigenvalue structure of the matrix are briefly examined.

In Section 3, an error analysis is developed which results in analytic expressions for estimating the accuracy of extracted frequencies. In addition, the nature of the results of this section provide insight for latter discussion with regard to determination of the number of poles in the data and to optimization concepts directed at advantageous modification of the structure of the Prony matrix.

In Section 4, the perturbation of the eigenvalue structure of the Prony matrix due to the inclusion of noise is studied. Analytic estimates for the expected values of all eigenvalues are obtained along with supporting bounds on their corresponding standard deviations.

---

1. R. Prony, "Essai experimental et analytique sur les lois de la dilatabilite de fluides elastiques et sur celles del la force expansive de la vapeur de l'alkool, a differentes temperatures," J. l'Ecole Polytech, (Paris), vol. 1, no. 2, pp. 24-76, 1795.

Section 5, addresses the eigenvalue structure of the unperturbed Prony matrix. Two obviously ill conditioned problems are examined with respect to the defective nature of their eigenvalue structure. In addition, potential methods of eigenvalue structure improvement by adjustment of problem structuring parameters are cited along with a qualitative description of the underly rationale.

Section 6, studies the errors introduced into the eigenvalue structure and extracted frequencies due to the presence of both random noise and a low amplitude neglected signal component sufficiently buried in the noise to be undetectable to the analyst.

Section 7, discusses an iterative technique which uses the results of a conventional Prony analysis as a starting point. The goal of this iterative technique is to remove the errors introduced by the two stage application of the least squares method associated with conventional Prony analysis.

Section 8, presents an assortment of numerical results and supportive discussion which illustrates the majority of concepts and theory developed in the earlier sections.

Finally, in Section 9 the author summarizes the results of this effort, discusses the limitations of the developed theory, and suggests potentially profitable avenues for additional study.

II. DEVELOPMENT OF PRONY'S METHOD

As has been adequately discussed in reference 2, there are many cases for which the transient response of a system can be expressed as a finite sum of exponentials. In this section, a methodology for extracting the poles and corresponding residues of each signal component directly from transient response data will be presented. The methodology is termed Prony's method because of the important role an algorithm first documented by Prony has in the development.

Prony's method can be developed in several different ways as has been discussed by Van Blaricum (ref 2). The development presented here is a classical one and is presented both for establishing a foundation upon which subsequent discussion will build and for the sake of completeness.

Suppose one is given transient response data represented as a set of ordered pairs $(y_k, t_k; k = 0, n)$. Further suppose the data is uniformly spaced in time such that

$$t_k = t_o + k\Delta \qquad k = 0, n \tag{1}$$

It is desired to represent the data in terms of a finite sum of exponentials given by

$$y(t) = \sum_{i=1}^{m} a_i e^{s_i t} \tag{2}$$

where a_i and s_i are, in general, complex.

By equation (1), one can then write

$$y(t_k) = \sum_{i=1}^{m} a_i e^{s_i(t_o + \Delta k)}$$

or

$$y(t_k) = \sum_{i=1}^{m} b_i z_i^k \qquad k = 0, n \tag{3}$$

where the quantities b_i and z_i are defined by

$$\left.\begin{aligned} b_i &= a_i e^{s_i t_o} \\ z_i &= e^{s_i \Delta} \end{aligned}\right\} \quad i = 1, m \tag{4}$$

---

2. Van, Blaricum, M.L., Mittra R., "Techniques for Extracting the Complex Resonances of a System Directly from its Transient Response," Interaction Note 301, December 1975.

One wishes to use given transient response data in conjunction with equation (3) to determine the system unknowns, b_i and z_i. To implement a solution one can use the following algorithm due to Prony. The z_i's are taken to be the m roots of the polynomial

$$z^m + \sum_{i=1}^{m} \alpha_i \, z^{m-i} = 0 \qquad (5)$$

Defining $\alpha_0 = 1$, equation (5) can be rewritten as

$$\sum_{i=0}^{m} \alpha_i \, z^{m-i} = 0 \qquad (6)$$

One then generates m+1 equations by multiplying the $(k+j)^{th}$ equation in equation (3) by α_{m-j} for j = 0, m.

Adding these m+1 equations and using (6) one obtains

$$y(t_k)\,\alpha_m + y(t_{k+1})\,\alpha_{m-1} + \ldots\ldots + y(t_{k+m})\,\alpha_o = 0$$

Since $\alpha_o = 1$, the above can be written as

$$\sum_{j=1}^{m} \alpha_j \, y(t_{k+m-j}) = - y(t_{k+m}) \qquad (7)$$

Since one can employ the same scheme for any k contained in the interval $[0, n-m]$, one has a total of n-m+1 equations in the m unknown α_j. For a unique solution one must require that

$$n+1 \geq 2m \qquad (8)$$

Throughout this document this condition will be assumed satisfied.

The actual transient response values $(y(t_k);\ k=0, n)$ of equation (7) are unknown. What is assumed available are the transient response data values $(y_k;\ k=0, n)$ which are only expected to be approximations to the corresponding $y(t_k)$ values.

Thus if one introduces given transient response data into equation (7) equality in general will not hold. Thus one is lead to define the k^{th} deviation as

$$\delta_k = \sum_{j=1}^{m} \alpha_j \, y_{k+m-j} + y_{k+m} \tag{9}$$

To achieve a solution, a classical least squares procedure will be followed. One first defines the sum of the squares of the deviation to be

$$S^2 = \sum_{k=0}^{n-m} \delta_k^{\,2} \tag{10}$$

and then requires that

$$\frac{\partial S^2}{\partial \alpha_j} = 0 \qquad j = 1, m \tag{11}$$

The resultant system of equations can easily be shown to be

$$\sum_{j=1}^{m} \alpha_j \sum_{k=0}^{n-m} y_{m+k-i}\, y_{m+k-j} = - \sum_{k=0}^{n-m} y_{m+k-i}\, y_{m+k} \qquad i=1, m \tag{12}$$

Equation (12) is a coventional matrix problem of the form

$$R\, \vec{\alpha} = \vec{b} \tag{13}$$

where

$$\left.\begin{aligned} R_{ij} &= \sum_{k=0}^{n-m} y_{m+k-i}\, y_{m+k-j} \\ b_i &= -\sum_{k=0}^{n-m} y_{m+k-i}\, y_{m+k} \end{aligned}\right\} \tag{14}$$

Examining equation (14) one observes that R is real and symmetric and hence has real eigenvalues with corresponding eigenvectors which span the m^{th} dimensional vector space. Further since $\vec{b}$ is real, one immediately concludes that if R is nonsingular then equation (13) has a unique solution $\vec{\alpha}$ which is real. Since $\vec{\alpha}$ is real, the m roots of equation (5) can in general be either real or complex. However, the complex roots must appear in conjugate pairs.

As pointed out above since R is real and symmetric it has m real eigenvalues. An even stronger conclusion with regard to R can easily be obtained.

Consider the eigenvalue problem

$$R\,\vec{\Omega}_\gamma = \lambda_\gamma\,\vec{\Omega}_\gamma \qquad (15)$$

Since the eigenvectors of R can be taken to be orthonormal one can perform an inner product on equation (15) with the vector $\vec{\Omega}_\gamma$. This yields

$$\left(\vec{\Omega}_\gamma\ ,\ R\vec{\Omega}_\gamma\right) = \lambda_\gamma\left(\vec{\Omega}_\gamma\ ,\vec{\Omega}_\gamma\right)$$

$$\left(\vec{\Omega}_\gamma\ ,\ R\vec{\Omega}_\gamma\right) = \lambda_\gamma \qquad (16)$$

Thus one computes

$$\lambda_\gamma = \sum_{i=1}^{m} \Omega_{\gamma i} \sum_{j=1}^{m} R_{ij}\ \Omega_{\gamma j}$$

$$= \sum_{k=0}^{n-m}\left(\sum_{i=1}^{m} y_{m+k-i}\,\Omega_{\gamma i}\right)\left(\sum_{j=1}^{m} y_{m+k-j}\,\Omega_{\gamma j}\right)$$

$$\lambda_\gamma = \sum_{k=0}^{n-m}\left(\sum_{i=1}^{m} y_{m+k-i}\,\Omega_{\gamma i}\right)^2 \qquad (17)$$

By equation (17) one concludes that if R is nonsingular then R has positive real numbers as eigenvalues and thus is positive definite. This property of R is a result of the fact that it is a Gramian matrix.

Returning to the development of the methodology for extracting poles and residues, the rest is rather straightforward.

One computes R and $\vec{b}$ from equation (14) using given data and solves equation (13) for the coefficients of the polynomial given in equation (5).

Equation (5) can then be solved for the m values of z_i using any one of several polynomial root-finding methods. Once the z's are obtained equation (4) enables one to compute the complex frequencies by the expression

$$s_i = \frac{\ell n\ z_i}{\Delta} \qquad (18)$$

It is important to note that one is guaranteed a unique solution for the vector $\vec{\alpha}$, and also for the m roots of the associated polynomial. However, in general, the resulting values of the complex frequencies are not unique.

This is most easily seen by observing that

$$\exp\left(s\Delta\right) = \exp\left(\left(s \pm \frac{2\pi kj}{\Delta}\right)\Delta\right) \qquad k = 0, 1, \ldots.$$

where here $j = \sqrt{-1}$. Since by equation (4)

$$z = \exp(s\Delta)$$

it is apparent that the angular frequency, corresponding to s, is uncertain by an amount

$$\pm \frac{2\pi k}{\Delta}$$

Since the computed frequency is constrained to lie within the interval $\left[0, \frac{\pi}{\Delta}\right]$ the above uncertainty implies frequencies exceeding the Nyquist criterion which requires that the highest frequency present in the data be sampled at least twice per period.

The conclusion is that within the constraints of the Nyquist criterion the complex frequencies obtained from equation (4) are also unique.

It only remains to compute the residues. This can be done in a straightforward manner by considering the values of z_i previously obtained as knowns in equation (3) and then performing another least squares solution for the unique b_i's. Once determined, the values for b_i coupled with the known values of s_i and t_o enable one to compute the a_i's by equation (4).

In the above development it has been assumed that the transient response contains only simple poles. As pointed out by Van Blaricum, the method extends quite nicely to the case of multiple poles, the only real complication being in the added complexity one encounters in formulating the second least squares problem for obtaining the residues.

Although the methodology derived above seems rather straightforward, there are numerous difficulties encountered in application.

Some of the primary ones will be discussed below.

The first problem is that one must be confident that the transient response is indeed characterized solely by contributions due to simple or multiple poles. This point is adequately treated in reference 2 and will not be reiterated here.

A second problem is that the methodology itself is quite sensitive to error in the transient response data. Previous investigators have noted signal-to-noise ratio in excess of 100 are commonly required for satisfactory results.

The third problem is perhaps the most difficult and is closely linked to the second problem mentioned above. The methodology assumes that m is known or can be determined. For relatively noiseless data the problem is trivial and simply corresponds to determining the rank of the R matrix displayed in equation (14). This can be done by successively increasing the value of m until one observes that the determinant of R vanishes or equivalently a zero is introduced as an eigenvalue. Numerically neither will actually happen due to round off errors associated with specific machine characteristics. However, except for extremely ill conditioned problems, the drop which occurs in the computation of the determinant of R or in the value of its lowest eigenvalue as one progresses from the m^{th} order to $m+1^{th}$ order R matrix is sufficiently dramatic to allow identification of the correct value of m.

On the other hand, for applications to real data the problem of determining m can be formidable. Rather modest noise levels can completely obscure any meaningful drop in either the determinant or the smallest eigenvalue of R as one progresses from the m^{th} to $m+1^{th}$ order matrix.

A fourth problem is in the theory underlying the development of the methodology. Two successive least squares problems are formulated, the second based on assuming the previously obtained frequencies are fixed and correct. Suppose such a procedure provides a solution of the form

$$\hat{y}(t) = \sum_{i=1}^{m/2} a_i e^{-\alpha_i t} \sin \omega_i t$$

Defining

$$S^2 = \sum_{k=0}^{n} \left[y_k - \hat{y}(t_k) \right]^2$$

one can easily show that the conventional Prony analysis provides

$$\frac{\partial S^2}{\partial a_i} = 0 \qquad i = 1, m/2$$

However, in general, it is not true that

$$\frac{\partial S^2}{\partial \alpha_i} = \frac{\partial S^2}{\partial \omega_i} = 0$$

Although their is no guarantee, one might expect that, if obtainable, a least square solution which simultaneously provides

$$\frac{\partial S^2}{\partial a_i} = \frac{\partial S^2}{\partial \alpha_i} = \frac{\partial S^2}{\partial \omega_i} = 0 \qquad i = 1, m/2$$

could be superior.

It is the author's feeling that most of the problems encountered in applying Prony's method are related to the four mentioned above. As previously stated, an adequate discussion of the first problem area is available elsewhere. The other three problem areas are the subject of the remainder of this document.

III. ERROR ANALYSIS

In Section 2, it was shown that to implement Prony's method one must solve a matrix equation of the form

$$R \vec{\alpha} = \vec{b} \tag{19}$$

Equation (14) of Section 2 provides the prescription for computing the matrix R and the vector $\vec{b}$ from given transient response data. In this section, the order of the system (m), will be assumed known and the given transient response data will be represented as

$$\hat{y}_k = y_k + \epsilon_k \qquad k=0, n$$

where y_k is defined to be the true transient response at the k^{th} point and ϵ_k the error associated with the given data at the k^{th} point.

The error will be assumed to consist of stationary random noise having zero mean and variance σ^2, and in addition will be assumed to be uncorrelated between points.

Equation (19) can be written as

$$(R_o + E)\vec{\alpha} = \vec{b}_o + \vec{\delta} \tag{20}$$

From equation (14) one can easily show that if all error terms are associated with E an $\vec{\delta}$, then

$$E_{ij} = \sum_{k=0}^{n-m} y_{m+k-i}\,\epsilon_{m+k-j} + y_{m+k-j}\,\epsilon_{m+k-i} + \epsilon_{m+k-j}\,\epsilon_{m+k-i} \tag{21}$$

$$\delta_i = -\sum_{k=0}^{n-m} y_{m+k-i}\,\epsilon_{m+k} + y_{m+k}\,\epsilon_{m+k-i} + \epsilon_{m+k}\,\epsilon_{m+k-i} \tag{22}$$

Assuming that E and $\vec{\delta}$ are small perturbations on the errorless quantities $\vec{R}_o$ and $\vec{b}_o$, equation (20) can be written to first order in the error terms as

$$\vec{\alpha} = \vec{\alpha}_o - R_o^{-1} (E\vec{\alpha}_o - \vec{\delta}) \tag{23}$$

where $\vec{\alpha}_o$ is defined to be the solution to the errorless matrix problem

$$\vec{\alpha}_o = R_o^{-1} \vec{b}_o \tag{24}$$

Let us define a vector $\vec{X}$ as

$$\vec{X} = E\vec{\alpha}_o - \vec{\delta} \tag{25}$$

Using the fact that $\vec{\alpha}_o$ is the solution to the errorless matrix problem described in equation (24) it follows that

$$X_i = \sum_{k=0}^{n-m} (y_{m+k-i} + \epsilon_{m+k-i}) \sum_{j=0}^{m} \alpha_j \epsilon_{m+k-j} \tag{26}$$

where $\alpha_o \equiv 1$ and α_j ($j = 1, m$) correspond to the components of the vector $\vec{\alpha}_o$.

By equations (23) and (25)

$$\vec{\alpha} - \vec{\alpha}_o = - R_o^{-1} \vec{X} \tag{27}$$

Expressing $\vec{X}$ in terms of its projection on each of the orthonormal eigenvectors of R_o one obtains

$$\vec{\alpha} - \vec{\alpha}_o = -\sum_{\gamma=1}^{m} \frac{(\vec{X}, \vec{\Omega}_\gamma)\vec{\Omega}_\gamma}{\lambda_\gamma} \tag{28}$$

where λ_γ is the eigenvalue associated with $\vec{\Omega}_\gamma$. Using equation (26) and the previously stated properties of the error, it is trivial to show that the expected value of $(\vec{X}, \vec{\Omega}_\gamma)$ is simply

$$\langle (\vec{X}, \vec{\Omega}_\gamma) \rangle = (n-m+1)\,\sigma^2\, (\vec{\alpha}_o, \vec{\Omega}_\gamma) \tag{29}$$

Combining equations (28) and (29) one obtains the expected value of the error vector

$$\left\langle \vec{\alpha} - \vec{\alpha}_o \right\rangle = -(n-m+1)\sigma^2 \sum_{\gamma=1}^{m} \frac{(\vec{\alpha}_o, \vec{\Omega}_\gamma)}{\lambda_\gamma} \vec{\Omega}_\gamma \tag{30}$$

This result that the expected value of the error vector is non zero even for the case of zero mean error is termed biasing and has long been recognized as a potential problem associated with Prony's method. However, for application to most practical problems the average error vector only has meaning if the variance of $(\vec{X}, \vec{\Omega}_\gamma)$ is small. For this reason it is of interest to compute $\left\langle (\vec{X}, \vec{\Omega}_\gamma)^2 \right\rangle$. The derivation is sufficiently lengthly to be omitted from the main text but is provided in the Appendix. The best result yet obtained by this author is that

$$\left\langle (\vec{X}, \vec{\Omega}_\gamma)^2 \right\rangle \leq \left\langle (\vec{X}, \vec{\Omega}_\gamma) \right\rangle^2 + \left(\left\langle \epsilon^4 \right\rangle - \sigma^4 \right) (n-m+1) \left(\vec{\alpha}_o, \vec{\Omega}_\gamma \right)^2$$
$$+ \sigma^2 (1 + (\vec{\alpha}_o, \vec{\alpha}_o) \left(2m(n-m+1)\sigma^2 + (m+1)\lambda_\gamma \right) \tag{31}$$

where $\left\langle \epsilon^4 \right\rangle$ is the expected value of the fourth moment of the noise at a specific point.

The conclusion is that, depending on the specifics of the problem under investigation, the expected value of the error vector given by equation (30) might provide a rather poor indication of the error to be expected from but a few statistically independent calculations.

However, one is really interested in obtaining an estimate of the accruacy of extracted roots of the polynomial

$$P(z) = \sum_{j=0}^{m} z^{m-j} \alpha_j = 0 \tag{32}$$

Differentiating the above one obtains

$$dz \approx -\left(P'(z) \right)^{-1} \sum_{j=1}^{m} d\alpha_j \; z^{m-j} \tag{33}$$

Where $P'(z)$ denotes the first derivative of $P(z)$ with respect to z. Assuming z corresponds to the unperturbed root of $P(z)$, one computes, to first order, that the expected error, $\left\langle dz \right\rangle$, is

$$\left\langle dz \right\rangle = -\left(P'(z) \right)^{-1} \sum_{j=1}^{m} \left\langle d\alpha_j \right\rangle z^{m-j} \tag{34}$$

An obvious application of Schwarz's inequality results in the bound

$$\left|\langle dz \rangle\right| \le \left|P'(z)\right| \left|\frac{|z|^{2m} - 1}{|z|^{2} - 1}\right|^{1/2} \left\| \langle \vec{\alpha} - \vec{\alpha}_o \rangle \right\| \tag{35}$$

However, again this bound is only of practical value if one or several computations of $\|\vec{\alpha} - \vec{\alpha}_o\|$ can be expected to approximate $\|\langle \vec{\alpha} - \vec{\alpha}_o \rangle\|$ reasonably well. With this concern in mind one is lead to compute

$$\langle |dz^2| \rangle = |P'(z)|^{-2} \left\langle \left| \sum_{j=1}^{m} d\alpha_j \, z^{m-j} \right|^2 \right\rangle$$

$$\langle |dz^2| \rangle \le |P'(z)|^{-2} \left| \frac{|z|^{2m} - 1}{|z|^{2} - 1} \right| \langle \vec{\alpha} - \vec{\alpha}_o , \; \vec{\alpha} - \vec{\alpha}_o \rangle \tag{36}$$

From equations (35) and (36) one concludes that to estimate the error in the roots of equation (32) it is desirable to have knowledge concerning both $\langle \|\vec{\alpha} - \vec{\alpha}_o\| \rangle$ and $\langle (\vec{\alpha} - \vec{\alpha}_o , \; \vec{\alpha} - \vec{\alpha}_o) \rangle$.

From equation (30) one determines that

$$\|\langle \vec{\alpha} - \vec{\alpha}_o \rangle\| \le (n-m+1)\sigma^2 \|\vec{\alpha}_o\| \left(\sum_{\gamma=1}^{m} \lambda_\gamma^{-2} \right)^{1/2} \tag{37}$$

From equations (28) and (31) one computes that

$$\begin{aligned} \langle (\vec{\alpha} - \vec{\alpha}_o , \; \vec{\alpha} - \vec{\alpha}_o) \rangle &\le (n-m+1)^2 \sigma^4 \|\vec{\alpha}_o\|^2 \sum_{\gamma=1}^{m} \lambda_\gamma^{-2} \\ &+ (n-m+1)\sigma^4 \left(1 + \|\vec{\alpha}_o\|^2\right) \left(2m + \langle \epsilon^4 \rangle / \sigma^4 - 1\right) \sum_{\gamma=1}^{m} \lambda_\gamma^{-2} \\ &+ (m+1)\, \sigma^2 \left(1 + \|\vec{\alpha}_o\|^2\right) \sum_{\gamma=1}^{m} \lambda_\gamma^{-1} \end{aligned} \tag{38}$$

Thus to first order, if σ and $\langle \epsilon^4 \rangle$ are known, one can use a computed vector $\vec{\alpha}$ and the computed eigenvalues of $R_o + E$ as approximations to $\vec{\alpha}_o$ and the λ_γ's to evaluate the bounds given by equations (37) and 38). These resulting bounds in conjunction with computed root locations of equation (32) and equations (35) and (36) provide bounds by which one can estimate the accuracy of the computed root locations. The author's numerical experiments to date have only addressed the adequacy of the bounds given by equations (37) and (38). These experiments, which will be described latter in this report, have proved quite successful, at least for the sample problems studied to date. Thus, although the utility of the error analysis presented has not been completely established, preliminary results are encouraging.

As was pointed out above, the preceding error analysis assumes that the variance and the expected value of the fourth moment of the noise and in addition the correct rank of R_o are all known. The following section addresses these issues.

IV. THE PERTURBED EIGENVALUES

As discussed at the end of Section 2, for the noiseless case, the order at which the R matrix acquires a zero eigenvalue provides a criteria for establishing the correct value of m in equation (3). Further, as observed at the end of Section 3, one is hopeful that the perturbed eigenvalues of the R matrix of correct order are adequate approximations to the eigenvalues of the noiseless R matrix for use in the bounds generated in the error analysis. The objective of this section is to study the statistical properties of the perturbed eigenvalues of the R matrix.

Let us again define $R = R_o + E$ where R_o is noiseless and E is an error matrix. Suppose one studies the eigenvalue problem

$$(R_o + E)\vec{\Omega}_\gamma = \lambda_\gamma \vec{\Omega}_\gamma \tag{39}$$

Employing the first order, Rayleigh-Schrodinger perturbation method (ref 3) one obtains that

$$\lambda_\gamma = \lambda^{(o)}_\gamma + (\vec{\Omega}^{(o)}_\gamma, E\vec{\Omega}^{(o)}_\gamma) \tag{40}$$

where $\lambda^{(o)}_\gamma$ and $\vec{\Omega}^{(o)}_\gamma$ are solutions to the system

$$R_o \vec{\Omega}^{(o)}_\gamma = \lambda^{(o)}_\gamma \vec{\Omega}^{(o)}_\gamma \tag{41}$$

Throughout this section, it will be assumed that the R_o matrix is of order equal to the number of poles in the data or greater. By the properties of the R_o matrix described in Section 2, one knows that the set of unperturbed eigenvalues consists of non-negative real numbers. From the preceding section one recalls that the E matrix of equation (41) is simply

$$E_{ij} = \sum_{k=0}^{n-m} y_{m+k-i}\,\epsilon_{m+k-j} + y_{m+k-j}\,\epsilon_{m+k-i} + \epsilon_{m+k-i}\,\epsilon_{m+k-j} \tag{42}$$

3. Merzbacher, Eugen, Quantum Mechanics, John Wiley & Sons, Inc., New York, December 1960.

For convenience, let us represent the components of $\Omega_\gamma^{(o)}$ as $\Omega_i^{(o)}$, $i=1, m$. Then by equations (41) and (43) one obtains

$$\lambda_\gamma = \lambda_\gamma^{(o)} + \sum_{k=0}^{n-m} \left(\sum_{i=1}^{m} \Omega_i^{(o)} y_{m+k-i} \right) \left(\sum_{j=1}^{m} \Omega_j^{(o)} \epsilon_{m+k-j} \right)$$

$$+ \left(\sum_{j=1}^{m} \Omega_j^{(o)} y_{m+k-j} \right) \left(\sum_{j=1}^{m} \Omega_i^{(o)} \epsilon_{m+k-i} \right)$$

$$+ \sum_{j=1}^{m} \sum_{i=1}^{m} \Omega_i^{(o)} \Omega_j^{(o)} \epsilon_{m+k-i} \epsilon_{m+k-j} \tag{43}$$

Since the noise is assumed to have zero mean, one observes that the expected value of λ_γ is simply

$$\langle \lambda_\gamma \rangle = \lambda_\gamma^{(o)} + \sum_{k=0}^{n-m} \sum_{j=1}^{m} \sum_{i=1}^{m} \Omega_i^{(o)} \Omega_j^{(o)} \langle \epsilon_{m+k-i} \epsilon_{m+k-j} \rangle \tag{44}$$

Assuming the noise is uncorrelated between points implies that

$$\langle \lambda_\gamma \rangle = \lambda_\gamma^{(o)} + (n-m+1) \sigma^2 \tag{45}$$

where σ^2 is the variance of the noise.

For m one greater than the true rank of R_o and for the expected value of the perturbed zero eigenvalue of R_o $(\gamma = m)$, equation (46) implies that

$$\langle \lambda_m \rangle = (n-m+1) \sigma^2 \tag{46}$$

The above is identical to an expression derived in a different way by Van Blaricum in reference 2. In reference 2 sample problems are studied in a Monte Carlo fashion to verify equation (46). In these same studies the standard deviation of λ_m is computed and for the problems examined, shown to be reasonably small with respect to $\langle \lambda_m \rangle$. Hence for these specific problems a single computation λ_m can be expected to be reasonably close to the theoretical value predicted by equation (47). However, can this result be expected for other problems?

To answer this question, one desires to compute or bound the standard deviation associated with λ_m. Further, due to the importance of the non zero eigenvalues of R_o in the error analysis expressions, one is also concerned with their standard deviations. The best bound obtained to date by the author for the standard deviation of the γ^{th} eigenvalue is

$$\sigma_\gamma^2 \leq 4m\,\sigma^2 \lambda_\gamma^{(o)} + (n-m+1)\,\sigma^4 \left\{ \frac{\langle \epsilon^4 \rangle}{\sigma^4} + 4m - 5 \right\} \tag{47}$$

The derivation of the above expression is rather lengthy and therefore is presented in the appendix. As in the preceding section, the quantity $\langle \epsilon^4 \rangle$ is the expected value of the fourth moment of the noise.

It is informative to examine the value of $\langle \epsilon^4 \rangle / \sigma^4$ for two different type noise.

The noise utilized in reference 2 was normally distributed. From reference 4, for normal processes with zero mean it is known that

$$\langle \epsilon_1 \epsilon_2 \epsilon_3 \epsilon_4 \rangle = \langle \epsilon_1 \epsilon_2 \rangle \langle \epsilon_3 \epsilon_4 \rangle + \langle \epsilon_2 \epsilon_3 \rangle \langle \epsilon_1 \epsilon_4 \rangle + \langle \epsilon_1 \epsilon_3 \rangle \langle \epsilon_2 \epsilon_4 \rangle \tag{48}$$

Thus if all indices are allowed to be identical one obtains

$$\langle \epsilon^4 \rangle = 3\sigma^4 \tag{49}$$

or for normally distributed noise

$$\langle \epsilon^4 \rangle / \sigma^4 = 3. \tag{50}$$

The noise utilized in the present author's numerical studies was uniformly distributed and can be represented by

$$\epsilon = \hat{E}x \tag{51}$$

where x is distributed uniformly on the interval from $[-1/2, +1/2]$.

4. Crandall, Stephen H., Mark, William D., Random Vibration in Mechanical Systems, Massachusetts Institute of Technology, 1963.

Therefore one can compute

$$\sigma^2 = \hat{E}^2 \int_{-1/2}^{+1/2} x^2 \, dx$$

$$\sigma^2 = \hat{E}^2/12 \qquad (52)$$

and

$$\langle \epsilon^4 \rangle = \hat{E}^4 \int_{-1/2}^{+1/2} x^4 \, dx$$

$$\langle \epsilon^4 \rangle = \hat{E}^4/80 \qquad (53)$$

or for uniformly distributed noise.

$$\langle \epsilon^4 \rangle / \sigma^4 = 1.8 \qquad (54)$$

Returning to equation (47) one can express it in another form as

$$\sigma_\gamma \le \frac{\left(\langle \lambda_\gamma \rangle - \lambda_\gamma^{o}\right)}{(n-m+1)^{1/2}} \left\{ \frac{\langle \epsilon^4 \rangle}{\sigma^4} - 5 + \frac{4m \langle \lambda_\gamma \rangle}{\langle \lambda_\gamma \rangle - \lambda_\gamma^{o}} \right\}^{1/2} \qquad (55)$$

where from equation (45) one recalls that

$$\langle \lambda_\gamma \rangle = \lambda_\gamma^{o} + (n-m+1)\sigma^2 \qquad (56)$$

For $\lambda_\gamma^{o} = 0$, equation (55) simplifies to

$$\sigma_\gamma \le \frac{\langle \lambda_\gamma \rangle}{(n-m+1)^{1/2}} \left\{ \frac{\langle \epsilon^4 \rangle}{\sigma^4} - 5 + 4m \right\}^{1/2} \qquad (57)$$

From equation (56) and (57) one observes that the standard deviation of the perturbed zero eigenvalue of R_o increases with n by the factor $(n-m+1)^{1/2}$. Also for a single computation of λ_γ to be representative of $\langle\lambda_\gamma\rangle$ one must use a reasonably large number of data points in comparison with m.

However for sufficiently large $\lambda_\gamma^{(o)}$, equation (47) implies that

$$\sigma_\gamma^{\ 2} \lesssim 4m\sigma^2 \lambda_\gamma^{\ (o)} \tag{58}$$

and therefore the standard deviation of the higher eigenvalues is independent of n except for the variation of $\lambda_\gamma^{(o)}$ itself. In the next section, it will be shown that $\lambda_\gamma^{(o)}$ is expected to be an increasing function of n. Since from equation (58)

$$\sigma_\gamma/\lambda_\gamma^{(o)} \lesssim \ 2\sigma\sqrt{m}\left(\lambda_\gamma^{\ (o)}\right)^{-1/2}$$

Thus one observes that for all eigenvalues a single computation of λ_γ becomes a better approximation to $\langle\lambda_\gamma\rangle$ as n becomes larger.

On the other hand from equation (56) one observes that the expected value of λ_γ becomes a progressively poorer approximation to $\lambda_\gamma^{\ o}$ as n increases.

Regardless, for n much larger than m, one expects that as m is taken larger than the correct rank of R_o a chain of eigenvalues having magnitude of approximately $(n-m+1)\sigma^2$ will develop. The spread in magnitude of members of this chain should be of order

$$2\sigma^2 \ (n-m+1)^{1/2} \left(\langle\epsilon^4\rangle/\sigma^4 - 5 + 4m\right)^{1/2}$$

The problem of determination of the correct rank of the system is thus equivalent to determining when this chain begins. Since the expected value of the non zero eigenvalues is

$$\langle\lambda_\gamma\rangle = \lambda_\gamma^{\ o} + (n-m+1)\sigma^2$$

The problem of determining the correct rank becomes formidable if $\lambda_\gamma^{\ o}$ is small with respect to $(n-m+1)\sigma^2$. Further, under the above condition, one suspects that the first order perturbation analysis itself is inadequate for this eigenstate. As will become apparent in the next section there are things the analyst can do to alter the value of $\lambda_\gamma^{\ o}$ in an advantageous manner in order to facilitate the determination of m.

V. THE EIGENVALUE STRUCTURE OF R_o

From the results of Sections 3 and 4, it should be very clear to the reader that the nature of the eigenvalue structure of R_o is very important with respect to the accuracy of Prony's method and with respect to the problem of determining the correct order of the system. In this section, two obviously ill conditioned problems will first be studied with respect to their impact on the eigenstructure of R_o. These problems are ones which have been observed by the author to present numerical problems and by their very nature this should be expected.

The first ill conditioned case is one in which there exists a very weak or low intensity signal, x(t), in the response data. In this case the true response can be written as

$$\hat{y}(t) = y(t) + x(t) \tag{59}$$

where x(t) can be thought to be a small perturbation on y(t).

The second case is one for which two system poles are closely spaced.

Let us assume the system has poles at s_1 and s_2 where $|t(s_1 - s_2)| \ll 1$ over the time window to be studied.

Then defining $\hat{y}(t)$ to be the total response one can write

$$\begin{aligned}\hat{y}(t) &= a_1 e^{s_1 t} + a_2 e^{s_2 t} \\ &= e^{\hat{s}t}\left(a_1 e^{(s_1 - \hat{s})t} + a_2 e^{(s_2 - \hat{s})t}\right) \\ &\simeq e^{\hat{s}t}\left(a_1 + a_2\right) + te^{\hat{s}t}\left(a_1(s_1 - \hat{s}) + a_2(s_2 - \hat{s})\right) \\ &\quad + \frac{t^2}{2} e^{\hat{s}t}\left(a_1(s_1 - \hat{s})^2 + a_2(s_2 - \hat{s})^2\right)\end{aligned}$$

Choosing $\hat{s} = \dfrac{a_1 \hat{s}_1 + a_2 s_2}{a_1 + a_2}$, one obtains

$$\hat{y}(t) \approx e^{\hat{s}t}(a_1 + a_2) + \frac{t^2}{2} e^{\hat{s}t}\left(a_1(s_1 - \hat{s})^2 + a_2(s_2 - \hat{s})^2\right) \tag{60}$$

Therefore the response can be approximated as simply the sum of contributions due to a simple pole located at the residual weighted mean of the actual poles and a second order pole at the same location having very small amplitude.

Thus for both cases mentioned above the situation is basically the same in that the response can be reasonably well approximated at a lower value of m than is strictly correct by an expression of the form

$$\hat{y}(t) = y(t) + x(t).$$

Let us assume that the matrix R_o is constructed of the y (t) values and has order m. Thus if the R_o matrix constructed of the y (t) values of order m+1 is examined it will have a zero eigenvalue.

For x (t) sufficiently small one can again use first order perturbation theory to estimate the shift in the zero eigenvalue. Note this time the shift from zero is due to neglection of a true signal component rather than random noise in the data. As before one can write for the perturbed zero eigenvalue that

$$\lambda = \left(\vec{\Omega}^{(o)},\ E\,\vec{\Omega}^{(o)}\right) \tag{61}$$

where $\vec{\Omega}^{(o)}$ is the eigenvector associated with the zero eigenvalue of R_o.

However for this case, letting m correspond to the order of E

$$E_{ij} = \sum_{k=0}^{n-m} y_{m+k-i}\,x_{m+k-j} + y_{m+k-j}\,x_{m+k-i}$$

$$+ x_{m+k-j}\,x_{m+k-i} \tag{62}$$

Hence

$$\lambda = \sum_{k=0}^{n-m} \left(\sum_{i=1}^{m} x_{m+k-i}\,\Omega_i^{(o)} \right)^2 \tag{63}$$

Again Schwarz's inequality provides a useful bound

$$\lambda \le \sum_{k=0}^{n-m} \left(\sum_{i=1}^{m} x^2_{m+k-i} \right) \left(\sum_{i=1}^{m} \left(\Omega_i^{(o)}\right)^2 \right)$$

and since $\vec{\Omega}^o$ is normalized.

$$\lambda \leq \psi = \sum_{i=1}^{m} \sum_{k=0}^{n-m} x^2_{m+k-i} \tag{64}$$

Thus for both cases previously introduced, computation of the eigenvalues of the correct rank R matrix is expected to provide at least one eigenvalue less than ψ. If there is noise in the data and if ψ is nearly equal to or smaller than $(n-m+1)\sigma^2$, one should expect to have difficulty in determining the correct order of R_o and should expect problems in determining the true locations of poles in the data.

Because of the above discussion and the results of Sections 3 and 4, one desires to maximize the smallest non zero eigenvalue of R_o. The remainder of this section addresses optimization concepts directed at achieving this result.

Assuming that m is tentatively fixed, the parameters which influence the eigen-structure of R_o are the sample spacing Δ, the time t_o corresponding to the first data point, and the length of the data set n.

From equation (64) one suspects that the eigenvalue introduced by a low intensity signal is roughly proportional to the sum of the squares of its magnitude over the data space.

Suppose there exists a low amplitude signal

$$x(t) = a e^{-\alpha t} \sin \omega t$$

Therefore one can write

$$x(t_k) = a \exp(-\alpha t_o - \alpha \Delta k) \sin(\omega t_o + \omega \Delta k)$$

The first observation is obvious. If t_o is taken to large, the $e^{-\alpha t_o}$ term can significantly reduce the magnitude of x (t) over the data window. The second observation is of equal concern. Suppose one selects t_o such that $\sin \omega t_o \approx 0$. Then if $\omega\Delta \approx \pi$ then $\sin \omega t_k$ can remain reasonably small until k becomes quite large. However, when k becomes quite large the exponential damping may be significant. Thus one wants to avoid sampling in such a way that in the early part of the data, signal strength is low.

The sensitivity of the eigenvalue structure of R_o to sample length n and the importance of this parameter is less obvious. In conventional least squares procedures the rule of thumb is to use as much data as available. In Prony's method this can be disastrous.

As observed in the preceding section the expected value of the shift in the eigenvalues of the perturbed Prony matrix increases linearly with n. For damped signals one knows that the true eigenvalues approach a limiting values as n increases since the contribution of near zero data values to the R_o matrix is minimal. Thus if one takes n excessively large, the expected value of the eigenvalues due to noise and the expected value of the shifts in the perturbed eigenstructure increase with respect to the limiting values of the unperturbed eigenstructure.

In addition, because of the above observations, it is important to know whether the true system eigenvalues are non decreasing functions of n. Again first order perturbation theory can provide some insight.

Suppose the addition of the n+1 data point is considered to result in a small change in the value of R computed with n data points.

Thus one expects the eigenvalues to shift by an amount approximated by

$$\Delta\lambda_\gamma = (\vec{\Omega}_\gamma \, , E\vec{\Omega}_\gamma \,) \tag{65}$$

where now

$$E_{ij} = y_{n+1-i}\, y_{n+1-j} \tag{66}$$

Therefore

$$\Delta\lambda_\gamma = \left(\sum_{i=1}^{m} y_{n+1-i}\, \Omega_i^{(\gamma)} \right)^2 \geq 0 \tag{67}$$

Based on equation (67) one suspects that all the eigenvalues are non decreasing functions of n, at least for n reasonably large.

This section has attempted to point out that the eigenvalue structure of the Prony matrix is dependent upon what the author prefers to call problem structuring parameters. These are: the number of data points, n; the sample spacing, Δ; and the time corresponding to the first data point, t_o. The analyst has some flexibility with respect to selection of these parameters. If n is fixed and taken sufficiently large to control the variance in the eigenvalue distributions, the first order analysis provided in Section 4 suggests that t_o and Δ can be optimized by maximizing the lowest eigenvalue of the perturbed matrix. Also, the discussion concluding Section 4 implies that an approximate value of σ^2 is relatively easy to obtain and hence the theory of Section 4 coupled with observed changes in the perturbed eigenvalue structure for various n can be helpful for ascertaining the optimum value of n.

VI. NEGLECTED SIGNAL STRENGTH

In Section 3 and 4, the error matrix E and error vector $\vec{\alpha} - \vec{\alpha}_o$ associated with Prony analysis was assumed due only to random noise. This section will address the eigenvalue structure of R and the error vector $\vec{\alpha}-\vec{\alpha}_o$ resulting if the order of the system is taken smaller than is strictly correct. For such a value of m, the R matrix can be written as the sum of a R_o matrix generated by the contributions due to a subset of the actual poles in the data and an error matrix E due to both random noise and to neglected signal strength. With the above guidelines, the error matrix is simply

$$E_{ij} = \sum_{k=0}^{n-m} y_{m+k-i} \; v_{m+k-j} + y_{m+k-j} \; v_{m+k-i} + v_{m+k-i} \; v_{m+k-j} \tag{68}$$

where

$$v_k = x_k + \epsilon_k \tag{69}$$

In equation (69), x_k represents the neglected signal strength and ϵ_k the random error present at the k^{th} data point.

One desires to estimate the impact of the presence of x_k on the perturbation of the eigenstructure of R.

Again, to first order one writes that

$$\lambda_\gamma = \lambda_\gamma^{(o)} + (\vec{\Omega}_\gamma \, , \, E\vec{\Omega}_\gamma) \tag{70}$$

Performing the expected value operation of equation (70) yields

$$\langle \lambda_\gamma \rangle = \lambda_\gamma^{(o)} + (n-m+1)\sigma^2 + \sum_{k=0}^{n-m} \sum_{i=1}^{m} \sum_{j=1}^{m} \Omega_{\gamma i} \; \Omega_{\gamma j} \left(x_{m+k-i} \; x_{m+k-j} + 2 \; y_{m+k-i} \; x_{m+k-j} \right) \tag{71}$$

For $\lambda_\gamma^{(o)} = 0$, equation (71) implies that

$$\left\langle \lambda_\gamma \right\rangle = (n-m+1)\sigma^2 + \sum_{k=0}^{n-m} \left(\sum_{i=1}^{m} \Omega_{\gamma i} \, x_{m+k-i} \right)^2 \tag{72}$$

The contribution to $\left\langle \lambda_\gamma \right\rangle$ from the neglected signal is clearly bounded by

$$\sum_{k=0}^{n-m} \left(\sum_{i=1}^{m} \Omega_{\gamma i} \, x_{m+k-i} \right)^2 \leq \sum_{k=0}^{n-m} \sum_{i=1}^{m} x^2_{m+k-i} \tag{73}$$

Therefore, if the neglected signal is such that

$$\sum_{k=0}^{n-m} \sum_{i=1}^{m} x^2_{m+k-i} < (n-m+1)\sigma^2 \tag{74}$$

One concludes it will have a reasonably small impact on the expected value of the perturbed zero eigenvalue. The above conclusion is not true for the higher eigenvalues. From equation (71) one can write

$$\left\langle \lambda_\gamma \right\rangle = \lambda_\gamma^{(o)} + (n-m+1)\sigma^2 + \varphi_\gamma \tag{75}$$

where

$$|\phi_\gamma| \leq 2 \left(\lambda_\gamma^{(o)} \psi_\gamma \right)^{1/2} + \psi_\gamma \tag{76}$$

and

$$\psi_\gamma = \sum_{k=0}^{n-m} \left(\sum_{i=1}^{m} \Omega_{\gamma i} \, x_{m+k-i} \right)^2 \leq \sum_{k=0}^{n-m} \sum_{i=1}^{m} x^2_{m+k-i} \tag{77}$$

Thus the perturbation in the expected value of the higher eigenvalues can potentially be significantly increased by the presence of the neglected signal. However, from Section 4, one recalls that the standard deviation associated with the higher eigenvalues can also be much larger than the expected value of the eigenvalue perturbation. Thus judgment of the impact of neglected signal strength must take into consideration changes in the variance of the perturbed eigenvalues. It can be shown that the presence of neglected signal strength results in equation (48) of Section 4 being modified as given below

$$\sigma_\gamma^2 \leq 4m\sigma^2 \left(\lambda_\gamma^o + \left| \phi_\gamma \right| \right) + (n-m+1)\sigma^4 \left\{ \frac{\left\langle \epsilon^4 \right\rangle}{\sigma^4} + 4m - 5 \right\} \tag{78}$$

Where $|\varphi_\gamma|$ is bound as described by equations (76) and (77). The development of the preceding equation can be found in the Appendix. From equations (76), (77) and (78) one concludes that the neglected signal has minimal impact on the variance of the perturbed eigenvalues. Thus in practical situations in which only one or several satistically independent sets of data are available, the presence of neglected signal strength should have relatively little impact on the error of computed perturbed eigenvalues.

One must now investigate the effect of neglected signal on the error analysis presented in Section 3. With neglected signal the error vector of equation (25) is modified so that

$$(\vec{X}, \vec{\Omega}_\gamma) = \sum_{k=0}^{n-m} \sum_{i=1}^{m} \Omega_{\gamma i} \left(y_{m+k-i} + v_{m+k-i}\right) \sum_{j=0}^{m} \alpha_j v_{m+k-j} \tag{79}$$

Therefore one computes that

$$\left\langle (\vec{X}, \vec{\Omega}_\gamma) \right\rangle = \sum_{k=0}^{n-m} \sum_{i=1}^{m} \Omega_{\gamma i} (y_{m+k-i} + x_{m+k-i}) \sum_{j=0}^{m} \alpha_j x_{m+k-j}$$

$$+ (n-m+1)\sigma^2 (\vec{\alpha}_o, \vec{\Omega}_\gamma) \tag{80}$$

From the above one concludes that the neglected signal can potentially have a significant effect on $\left\langle \vec{X}, \vec{\Omega}_\gamma \right\rangle$. However, as in the case of the eigenvalue analysis, one must also understand how the expected value of the second moment of equation (79) is modified.

As is shown in the Appendix the inclusion of neglected signal results in the modification of equation (31) to the form.

$$\left\langle (\vec{X}, \vec{\Omega}_\gamma)^2 \right\rangle \le \left\langle (\vec{X}, \vec{\Omega}_\gamma) \right\rangle^2 + \left(\left\langle \epsilon^4 \right\rangle - \sigma^4\right) (n-m+1) (\vec{\alpha}_o, \vec{\Omega}_o)^2$$

$$+ \sigma^2 (1 + (\vec{\alpha}_o, \vec{\alpha}_o) \left(2m (n-m+1)\sigma^2 + (m+1) \lambda_\gamma\right)$$

$$+ 4 (m+1)\sigma^2 (1 + (\vec{\alpha}_o, \vec{\alpha}_o) \left(\psi_\gamma + (\psi_\gamma \lambda_\gamma)^{1/2}\right) \tag{81}$$

where

$$\psi_\gamma \leq \sum_{k=0}^{n-m} \sum_{j=0}^{m} x^2_{m+k-j} \tag{82}$$

Therefore, one concludes that although the variance of $(\vec{X},\vec{\Omega}_\gamma)$ is rather insensitive to neglected signal, the second moment can potentially be greatly magnified resulting in enhanced errors in the solution vector $\vec{\alpha}$.

VII. ITERATIVE IMPROVEMENT

As mentioned in the text concluding Section 2, Prony's methodology consists of two successive applications of least squares criterion. The first is in the determination of the coefficients of the polynomial whose roots correspond through a trivial transformation to the system complex frequencies. The second least squares procedure is applied in the determination of the residues and for this determination the previously computed frequencies are taken to be fixed. This section describes an alternative least squares solution criteria.

Let us assume that the result of first least squares application provides m distinct complex roots. These must necessarily correspond to m/2 complex conjugate pairs. Writing the i^{th} complex frequency in terms of its real and imaginary parts, one has

$$s_i = -\alpha_i + j\omega_i \tag{83}$$

where $j = \sqrt{-1}$. Therefore, a solution of the form

$$y(t) = \sum_{i=1}^{m/2} e^{-\alpha_i t} \left(a_i \sin \omega_i t + \hat{a}_i \cos \omega_i t \right) \tag{84}$$

is expected under the above restrictions.

Often from physics considerations, causality requires that the $\hat{a}_i = 0$, $i = 1, m/2$. In problems for which $\hat{a}_i$ is truly zero, neglect of this constraint has been found to seriously degrade the solution accuracy for noisy data.

For the remainder of this section, it will be assumed that the correct formulation of the system response is

$$y(t) = \sum_{i=1}^{m/2} a_i e^{-\alpha_i t} \sin \omega_i t \tag{85}$$

Let us define the sum of the squares of the deviations between $y(t_k)$ and y_k to be

$$S^2 = \sum_{k=0}^{n} (y(t_k) - y_k)^2 \tag{86}$$

Since the last step of the Prony methodology is computation of the a_i's by a least squares procedure applied to S^2 as defined in equation (86), one is assured that within machine accuracy

$$\frac{\partial S^2}{\partial a_q} = 0 \qquad , \; q=1, \; m/2 \tag{87}$$

However, since the complex frequencies are assumed known, it should be obvious that in general

$$\frac{\partial S^2}{\partial \alpha_q} \neq 0 \quad , \quad \frac{\partial S^2}{\partial \omega_q} \neq 0 \qquad q = 1, \; m/2 \tag{88}$$

This is indeed the case as later numerical results will adequately demonstrate.

The basis of the iteration methodology to be described is that one might expect that a solution satisfying the property that

$$\frac{\partial S^2}{\partial a_q} = \frac{\partial S^2}{\partial \alpha_q} = \frac{\partial S^2}{\partial \omega_q} = 0 \tag{89}$$

might be superior to that achieved through implementation of Prony's methodology. Numerical experiments performed to date seem to support this contention as will be seen toward the end of the next section.

This iteration methodology for establishing the property stated in equation (89) begins with the numerical computation of the partial derivations $\frac{\partial S^2}{\partial \alpha_q}$, $\frac{\partial^2 S^2}{\partial \alpha_q^2}$, $\frac{\partial S^2}{\partial \omega_q}$, and $\frac{\partial^2 S^2}{\partial \omega_q^2}$.

Symbolically letting β_q represent either α_q or ω_q, one assumes that the variation in S^2 with respect to a small change in β_q, all other parameters remaining fixed, can be approximated as

$$S^2(\beta_q + \Delta) = S^2(\beta_q) + \Delta \frac{\partial S^2}{\partial p_q} + \frac{\Delta^2}{2} \frac{\partial^2 S^2}{\partial \beta_q^2} \tag{90}$$

The quantity Δ must be restricted such that the higher order terms are indeed negligible.

Using equation (90) one determines the maximum reduction in S^2 which can be achieved by optimum selection of β_q within the restricted interval.

The above process is applied for each α_q and ω_q. The one parameter that promises the maximum reduction in S^2 is then modified.

The resultant frequencies are then assumed fixed and a least squares process is applied to evaluate updated values for a_q which is guaranteed to provide additional reduction in S^2.

The process outlined above is then systematically continued until the condition stated in equation (90) is achieved within machine limitations.

Numerical experiments with this iteration technique have been limited in scope but results to date have been quite impressive. Additional effort is needed to optimize the methodology and to marry it with the error analysis results developed in previous sections. Timing studies have not yet been performed but obviously time requirements are directly proportional to the accuracy of results obtained from the conventional Prony methodology. Numerical experiments with extremely ill conditioned problems such as two root locations close together or with a signal component of very low intensity have provided interesting results upon iteration. For the low intensity component problem, S^2 becomes relatively insensitive to this component and the minimum with respect to its frequency components becomes very weak. As a result, poorly defined starting values can result in this term being iterated out of the solution. The iteration process usually accomplishes this by continuallly increasing the damping associated with this term.

Similarly, depending on starting values the iteration process can remove from the solution one of two nearby poles using the remaining one to approximate the response contribution of both.

The reader should note, however, that the development and majority of experimentation with the iteration process preceded in time; the theoretical results and understanding of Prony limitations described in the previous sections. Result from these earlier sections indicate that problems ill conditioned in the sense described above, will result in similar problems with respect to establishing the presence or absence of such components by Prony's methodology.

From preceding sections, the reader should note that because of the expected error properties of Prony analysis, one might not employ all the data to determine the pole locations. Error characteristics might dictate that only every second or third data point be used and that a sample length significantly smaller than that available be employed. The iteration process described above, apparently does not suffer from these limitations and all available data can be incorporated. Inclusion of data points dominated by noise seems to have no adverse effect on the iteration process.

VIII. NUMERICAL RESULTS

In this section, the author presents the results of numerical experiments which either illustrate concepts rather qualitatively discussed in earlier sections or examine the appropriateness of theory developed. Because of the importance of unperturbed eigenstructure of the R_o matrix, the author has chosen to present data illustrating the sensitivity of this eigenstructure to problem structuring parameters first.

In Section 5, the problem structuring parameters were defined to be the time corresponding to the first data point, t_o; the number of data points, n + 1; and the sample spacing, Δ. Table 1 studies the sensitivity of the minimum eigenvalue of the R_o matrix to the parameter, t_o, for three different transient responses with other problem structuring parameters remaining fixed. For function 1, a simple sinusoid, the variation in λ_{min} is small since the average signal strength is constant in time and thus shifting the data window to latter time has relatively little effort. For function 2 and 3 there is a somewhat more pronounced effect due to removal of relatively intense early time data and substitution of damped later time data. Tables 2, 3, and 4 display the sensitivity of the minimum eigenvalue (and for table 4, the maximum eigenvalue also) to variations in sample length, n, with other problem structuring parameters fixed as specified.

In table 2, a simple sinusoid is studied. The staristep increases in λ_{min} with $\Delta = .5$ is a result of the sin (t (n+1) $\pi/2$) vanishing for n odd. The fact that the columns for $\Delta = .3$ and $\Delta = .7$ are identical is a result of the sets of transient response data being simply of opposite signs. For the damped sinusoid displayed in table 3, the damping destroys the abovestated equivalence between $\Delta = .3$ and $\Delta = .7$. Also, due to the damping, one can observe a weak reduction in the rate of increase of λ_{min} with n as n becomes larger especially for the case $\Delta = .7$. This latter property is somewhat more apparent in table 4 since the most weak signal component is heavily damped for n = 29. It should also be noted that the change in λ_{max} is much more dramatic and increasing since the composite signal strength remains appreciable.

Table 1. Sensitivity of Minimum Eigenvalue to t_o

FUNCTION 1: $y_1(t) = \sin \pi t$

FUNCTION 2: $y_2(t) = e^{-.05t} \sin \pi t$

FUNCTION 3: $y_3(t) = 1.0\, e^{-.2t} \sin(.5\pi t) + .50\, e^{-.3t} \sin(\pi t) + .25\, e^{-.3t} \sin(1.5\pi t) + .10\, e^{-.4t} \sin(2\pi t)$

STRUCTURE PARAMETERS:

Functions 1 and 2: $\Delta = .4$, $n = 49$, $m = 2$

Function 3 : $\Delta = .2$, $n = 49$, $m = 8$

t_o	λ_{min} of y_1	λ_{min} of y_2	λ_{min} of y_3
0	$.1675 \times 10^2$	$.7394 \times 10^1$	$.1500 \times 10^{-5}$
.1	$.1664 \times 10^2$	$.7215 \times 10^1$	$.1492 \times 10^{-5}$
.2	$.1651 \times 10^2$	$.7070 \times 10^1$	$.1489 \times 10^{-5}$
.3	$.1641 \times 10^2$	$.6979 \times 10^1$	$.1491 \times 10^{-5}$
.4	$.1637 \times 10^2$	$.6948 \times 10^1$	$.1488 \times 10^{-5}$
.5	$.1641 \times 10^2$	$.6963 \times 10^1$	$.1489 \times 10^{-5}$
.6	$.1651 \times 10^2$	$.6996 \times 10^1$	$.1483 \times 10^{-5}$
.7	$.1664 \times 10^2$	$.7007 \times 10^1$	$.1397 \times 10^{-5}$
.8	$.1675 \times 10^2$	$.6960 \times 10^1$	$.1184 \times 10^{-5}$
.9	$.1680 \times 10^2$	$.6847 \times 10^1$	$.9620 \times 10^{-6}$

Table 2. Sensitivity of Minimum Eigenvalue to n for Sinusoid

FUNCTION: $y(t) = \sin \pi t$

STRUCTURE PARAMETERS: $m=2$, $t_o=0$

n	λ_{min} with $\Delta = .3$	λ_{min} with $\Delta = .5$	λ_{min} with $\Delta = .7$
40	$.7912 \times 10^1$	$.1900 \times 10^2$	$.7912 \times 10^1$
41	$.8244 \times 10^1$	$.2000 \times 10^2$	$.8244 \times 10^1$
42	$.8567 \times 10^1$	$.2000 \times 10^2$	$.8567 \times 10^1$
43	$.8573 \times 10^1$	$.2100 \times 10^2$	$.8573 \times 10^1$
44	$.8788 \times 10^1$	$.2100 \times 10^2$	$.8788 \times 10^1$
45	$.9189 \times 10^1$	$.2200 \times 10^2$	$.9189 \times 10^1$
46	$.9265 \times 10^1$	$.2200 \times 10^2$	$.9265 \times 10^1$
47	$.9359 \times 10^1$	$.2300 \times 10^2$	$.9359 \times 10^1$
48	$.9762 \times 10^1$	$.2300 \times 10^2$	$.9762 \times 10^1$
49	$.9961 \times 10^1$	$.2400 \times 10^2$	$.9961 \times 10^1$

Table 3. Sensitivity of Minimum Eigenvalue to n for Damped Sinusoid

FUNCTION: $y(t) = e^{-.05t} \sin \pi t$

STRUCTURE PARAMETERS: $m=2$, $t_o=0$

n	λ_{min} with $\Delta = .3$	λ_{min} with $\Delta = .5$	λ_{min} with $\Delta = .7$
40	$.4698 \times 10^1$	$.8501 \times 10^1$	$.2742 \times 10^1$
41	$.4800 \times 10^1$	$.8643 \times 10^1$	$.2764 \times 10^1$
42	$.4895 \times 10^1$	$.8643 \times 10^1$	$.2782 \times 10^1$
43	$.4897 \times 10^1$	$.8772 \times 10^1$	$.2783 \times 10^1$
44	$.4958 \times 10^1$	$.8772 \times 10^1$	$.2794 \times 10^1$
45	$.5066 \times 10^1$	$.8888 \times 10^1$	$.2813 \times 10^1$
46	$.5086 \times 10^1$	$.8888 \times 10^1$	$.2816 \times 10^1$
47	$.5110 \times 10^1$	$.8994 \times 10^1$	$.2820 \times 10^1$
48	$.5211 \times 10^1$	$.8994 \times 10^1$	$.2836 \times 10^1$
49	$.5258 \times 10^1$	$.9089 \times 10^1$	$.2843 \times 10^1$

Table 4. Sensitivity of Minimum and Maximum Eigenvalues to n for a Sum of Damped Sinusoid.

FUNCTION: $y(t) = 1.0\, e^{-.2t} \sin(.5\pi t) + .50\, e^{-.3t} \sin(\pi t) + .25\, e^{-.4t} \sin(1.5\pi t) + .10\, e^{-.5t} \sin(2\pi t)$

STRUCTURE PARAMETERS: $m=8$, $t_o=0$, $\Delta=.2$

n	λ_{min}	λ_{max}
20	$.1381 \times 10^{-5}$	$.2055 \times 10^{2}$
21	$.1382 \times 10^{-5}$	$.2186 \times 10^{2}$
22	$.1411 \times 10^{-5}$	$.2307 \times 10^{2}$
23	$.1424 \times 10^{-5}$	$.2399 \times 10^{2}$
24	$.1424 \times 10^{-5}$	$.2450 \times 10^{2}$
25	$.1426 \times 10^{-5}$	$.2463 \times 10^{2}$
26	$.1427 \times 10^{-5}$	$.2464 \times 10^{2}$
27	$.1427 \times 10^{-5}$	$.2487 \times 10^{2}$
28	$.1427 \times 10^{-5}$	$.2550 \times 10^{2}$
29	$.1427 \times 10^{-5}$	$.2632 \times 10^{2}$

In general from the preceding four tables, one observes rather minor sensitivity of the minimum eigenvalue of the R_o matrix to t_o and n. Certainly if a highly damped component is present in the transient response data, the sensitivity to t_o would be magnified considerably. The sensitivity of the eigenstructure of the R_o matrix to the parameter is much more dramatic as demonstrated by table 5. Clearly for the simple and damped sinusoids the optimum displayed Δ is the same and corresponds to $\Delta = .5$ which is 50% of the Nyquist criterion. With $t_o = 0$, one is therefore sampling at the maximum, the minimum, and the zero crossings of each cycle. One should also observe that with a starting time of $t_o = 0$ and with $\Delta = 1.0$ λ_{min} of y_1 and y_2 vanish since now one is sampling only on the zero crossings. For function 3 in table 5 the Nyquist criterion corresponds to $\Delta = .5$ and a zero value of λ_{min} results due to sampling of the lowest signal component at its zero crossings. Also based on the data given, one observes that the optimum Δ is 80% of the Nyquist criteria or $\Delta = .4$. Calculations of λ_{min} for y_3 were also performed for sample spacings violating the Nyquist criteria. As one might expect, for damped waveforms, there is no apparent advantage in this ploy and in fact the zero eigenvalue at $\Delta = 1.0$ is a result of sampling two signal components only at zero crossings.

Based on the preceding tables, the author suggests that the most important problem structuring parameter is Δ and the analyst must be especially careful of its selection.

Also in Section 5, two ill conditioned problems were discussed and rather crude analytic arguments made to explain the nature of the ill condition. Tables 6 and 7 display numerical results with reinforce the comments made in Section 5. In table 6, the weak signal component problem is numerically studied as a function of weak signal strength and resulting impact on the eigenvalue structure. The minimum eigenvalue is observed to drop two orders of magnitude for a factor of ten reduction in weak signal strength. Table 7 addresses the ill conditioned problem in which two poles are closely spaced. Here one observes a drop in the smallest eigenvalue of almost four orders of magnitude as the frequency separation parameter is reduced from .1 to .01.

Table 5. Sensitivity of Minimum Eigenvalue to Δ

FUNCTION 1: $y_1(t) = \sin \pi t$

FUNCTION 2: $y_2(t) = e^{-.2t} \sin \pi t$

FUNCTION 3: $y_3(t) = 1.0\, e^{-.2t} \sin(.5\pi t) + .50\, e^{-.3t} \sin(\pi t) + .25\, e^{-.4t} \sin(1.5\pi t) + .10\, e^{-.5t} \sin(2\pi t)$

STRUCTURE PARAMETERS

Functions 1 and 2: $n=49$, $m=2$, $t_o=0$

Function 3 : $n=49$, $m=8$, $t_o=0$

Δ	λ_{min} of y_1	λ_{min} of y_2	λ_{min} of y_3
.1	$.1136 \times 10^{1}$	$.5162 \times 10^{0}$	$.1358 \times 10^{-9}$
.2	$.4531 \times 10^{1}$	$.1166 \times 10^{1}$	$.1500 \times 10^{-5}$
.3	$.9961 \times 10^{1}$	$.1709 \times 10^{1}$	$.3275 \times 10^{-3}$
.4	$.1675 \times 10^{2}$	$.2151 \times 10^{1}$	$.2794 \times 10^{-2}$
.5	$.2400 \times 10^{2}$	$.2483 \times 10^{1}$	$-.4590 \times 10^{-16}$
.6	$.1675 \times 10^{2}$	$.1428 \times 10^{1}$	$.2907 \times 10^{-6}$
.7	$.9961 \times 10^{1}$	$.7290 \times 10^{0}$	$.1028 \times 10^{-5}$
.8	$.4531 \times 10^{1}$	$.2951 \times 10^{0}$	$.1632 \times 10^{-6}$
.9	$.1136 \times 10^{1}$	$.6710 \times 10^{-1}$	$.5480 \times 10^{-4}$
1.0	$.1000 \times 10^{-17}$	$.1030 \times 10^{-18}$	$-.3760 \times 10^{-15}$

Table 6. Eigenvalue Structure -vs- Weak Signal Strength

FUNCTION: $y(t) = e^{-.05t} \sin(\pi t) + \epsilon\, e^{-.05t} \sin(\pi t/2)$				
STRUCTURE PARAMETERS: $m=4$, $n=49$, $t_o=0.$, $\Delta=.4$				
ϵ	λ_1	λ_2	λ_3	λ_4
.100	$.2468 \times 10^{2}$	$.1546 \times 10^{2}$	$.1631 \times 10^{0}$	$.5128 \times 10^{-2}$
.075	$.2469 \times 10^{2}$	$.1543 \times 10^{2}$	$.9193 \times 10^{-1}$	$.2883 \times 10^{-2}$
.050	$.2472 \times 10^{2}$	$.1542 \times 10^{2}$	$.4090 \times 10^{-1}$	$.1280 \times 10^{-2}$
.025	$.2476 \times 10^{2}$	$.1542 \times 10^{2}$	$.1023 \times 10^{-1}$	$.3194 \times 10^{-3}$
.010	$.2479 \times 10^{2}$	$.1543 \times 10^{2}$	$.1636 \times 10^{-2}$	$.5103 \times 10^{-4}$

Table 7. Eigenvalue Structure -vs- Frequency Structure

FUNCTION: $y(t) = .5\, e^{-.05t} \sin \pi t + .5\, e^{-.05t} \sin \pi (1-\epsilon) t$				
STRUCTURE PARAMETERS: $m=4$, $n=49$, $t_o=0.$, $\Delta=.4$				
ϵ	λ_1	λ_2	λ_3	λ_4
.100	$.1265 \times 10^{2}$	$.7883 \times 10^{1}$	$.7645 \times 10^{-1}$	$.9539 \times 10^{-2}$
.075	$.1346 \times 10^{2}$	$.8265 \times 10^{1}$	$.3375 \times 10^{-1}$	$.4504 \times 10^{-2}$
.050	$.1718 \times 10^{2}$	$.1052 \times 10^{2}$	$.7060 \times 10^{-2}$	$.1011 \times 10^{-2}$
.025	$.2242 \times 10^{2}$	$.1375 \times 10^{2}$	$.3898 \times 10^{-3}$	$.6467 \times 10^{-4}$
.010	$.2443 \times 10^{2}$	$.1512 \times 10^{2}$	$.9395 \times 10^{-5}$	$.1681 \times 10^{-5}$

Two of the most important messages for the analyst performing Prony analysis in the presence of noise that this document attempts to convey are the importance of proper selection of sample spacing and the danger of using an excessive number of data points that are dominated by noise. Table 8 attempts to reinforce these messages with numerical examples of recovered complex frequencies in the presence of noise. Computer generated noise was added to the given function and the perturbed problem was solved using three different sets of problem structuring parameters. The discrepancies are not as dramatic as they could be but the noise level is rather modest. From the structure parameters given in table 8, the reader should note that problem 1 utilizes a good sample spacing with a reasonable number of data points. Problem 2 is identical to problem 1 except that the number of data points has been doubled. The computed frequencies are rather similar except for those associated with the weakest signal component. Computed damping factors are rather poor for both problems with respect to the two most weak signal components. Problem 3 is identical to problem 1 except that the sample spacing is halved. Consistent with the results observed in the previous tables the deterious effect is much more pronounced.

The next six tables (tables 9 through 14) present numerical results in support of the theory of the perturbed eigenvalue structure developed in Section 4. In all these tables, the highest indexed eigenvalue was computed using the R_o matrix of order one greater than the true number of poles in the data. Twenty Monte Carlo runs were performed for this eigenvalue study. The remaining eigenvalues were computed for the R matrix of correct order and 10 Monte Carlo runs were used for their study. In each table is presented the unperturbed value of the respective eigenvalues $\lambda_\gamma^{(o)}$, the Monte Carlo generated estimate of $\langle \lambda_\gamma \rangle$, and the Monte Carlo generated estimate of the standard deviation associated with λ_γ. The column labeled theoretical mean is the value which results from evaluating equation (45) of Section 4. The theoretical bound on standard deviation is evaluated by equation (47) of Section 4. One observes that the theoretical means reasonably approximate the experimental results within the limits imposed by the observed standard deviations and number of samples generated. Further, the theoretical bound on the standard deviation in general provides an estimate that is higher by no more than a factor of two than computed results.

Table 8. Problem Structure Parameter Importance

FUNCTION: $y(t) = 1.0\, e^{-.2t} \sin(.5\pi t) + .5\, e^{-.3t} \sin(\pi t)$

$+ .25\, e^{-.4t} \sin(1.5\pi t) + .10\, e^{-.5t} \sin(2\pi t)$

STRUCTURE PARAMETERS:

Problem 1: $\Delta = .4,\ n=49,\ m=8,\ t_o=0.$

Problem 2: $\Delta = .4,\ n=99,\ m=8,\ t_o=0.$

Problem 3: $\Delta = .2,\ n=49,\ m=8,\ t_o=0.$

NOISE DESCRIPTION: E = .10 ($\sigma = 2.9 \times 10^{-2}$)

Frequency Component	Input Value	Problem 1	Problem 2	Problem 3
α_1	-.2	-.1980	-.2050	-.3401
α_2	-.3	-.4808	-.5308	-.1474
α_3	-.4	-.8495	-.6599	-1.877
α_4	-.5	-.9574	-.9848	-2.568
ω_1	1.571	1.582	1.576	1.644
ω_2	3.142	3.102	3.035	3.577
ω_3	4.712	4.883	5.008	10.13
ω_4	6.283	6.630	7.854	15.02

Table 9. Monte Carlo Study on Eigenvalue Structure of Damped Sinusoid with Low Error

FUNCTION: $y(t) = e^{-.05t} \sin(\pi t)$

STRUCTURE PARAMETERS: $n=49$, $\Delta=.4$, $t_o=0$, $m=2$ and 3

NOISE DESCRIPTION: $E = .0178$ ($\sigma = 5.1 \times 10^{-3}$)

Parameter	Unperturbed Value	Mean	Standard Deviation	Theoretical Mean	Theoretical Bound on Standard Deviation
λ_1	$.1391 \times 10^2$	$.1391 \times 10^2$	$.36 \times 10^{-1}$	$.1391 \times 10^2$	$.54 \times 10^{-1}$
λ_2	$.7394 \times 10^1$	$.7393 \times 10^1$	$.15 \times 10^{-1}$	$.7395 \times 10^1$	$.40 \times 10^{-1}$
λ_3	0	$.1255 \times 10^{-2}$	$.41 \times 10^{-3}$	$.1241 \times 10^{-2}$	$.54 \times 10^{-3}$

Table 10. Monte Carlo Study on Eigenvalue Structure of Damped Sinusoid with Medium Error

FUNCTION: $y(t) = e^{-.05t} \sin(\pi t)$

STRUCTURE PARAMETERS: $n=49$, $\Delta=.4$, $t_o=0$, $m=2$ and 3

NOISE DESCRIPTION: $E = .178$ ($\sigma = 5.1 \times 10^{-2}$)

Parameter	Unperturbed Value	Mean	Standard Deviation	Theoretical Mean	Theoretical Bound on Standard Deviation
λ_1	$.1391 \times 10^2$	$.1400 \times 10^2$	$.37 \times 10^0$	$.1404 \times 10^2$	$.54 \times 10^0$
λ_2	$.7394 \times 10^1$	$.7507 \times 10^1$	$.15 \times 10^0$	$.7521 \times 10^1$	$.40 \times 10^0$
λ_3	0	$.1254 \times 10^0$	$.41 \times 10^{-1}$	$.1241 \times 10^0$	$.54 \times 10^{-1}$

Table 11. Monte Carlo Study on Eigenvalue Structure of Damped Sinusoid with High Error

FUNCTION: $y(t) = e^{-.05t} \sin(\pi t)$

STRUCTURE PARAMETERS: $n=49$, $\Delta=.4$, $t_o=0$, $m=2$ and 3

NOISE DESCRIPTION: $E = .562$ ($\sigma = 1.6 \times 10^{-1}$)

Parameter	Unperturbed Value	Mean	Standard Deviation	Theoretical Mean	Theoretical Bound on Standard Deviation
λ_1	$.1391 \times 10^2$	$.1498 \times 10^2$	$.12 \times 10^1$	$.1517 \times 10^2$	$.18 \times 10^1$
λ_2	$.7394 \times 10^1$	$.8662 \times 10^1$	$.62 \times 10^0$	$.8657 \times 10^1$	$.13 \times 10^1$
λ_3	0	$.1246 \times 10^1$	$.40 \times 10^0$	$.1237 \times 10^1$	$.54 \times 10^0$

Table 12. Monte Carlo Study on Eigenvalue Structure of a Sum of Damped Sinusoids with Low Error

FUNCTION: $y(t) = 1.0\, e^{-.2t} \sin(.5\pi t) + .5\, e^{-.3t} \sin(\pi t) + .25\, e^{-.4t} \sin(1.5\pi t) + .10\, e^{-.5t} \sin(2\pi t)$

STRUCTURE PARAMETERS: $n=49$, $\Delta=.4$, $t_o=0$, $m=8$ and 9

NOISE DESCRIPTION: $E = .0003$ ($\sigma = 8.7 \times 10^{-5}$)

Parameter	Unperturbed Value	Mean	Standard Deviation	Theoretical Mean	Theoretical Bound on Standard Deviation
λ_1	$.1114 \times 10^{2}$	$.1114 \times 10^{2}$	$.11 \times 10^{-2}$	$.1114 \times 10^{2}$	$.16 \times 10^{-2}$
λ_2	$.5208 \times 10^{1}$	$.5208 \times 10^{1}$	$.48 \times 10^{-3}$	$.5208 \times 10^{1}$	$.11 \times 10^{-2}$
λ_3	$.9654 \times 10^{0}$	$.9654 \times 10^{0}$	$.23 \times 10^{-3}$	$.9654 \times 10^{0}$	$.48 \times 10^{-3}$
λ_4	$.7135 \times 10^{0}$	$.7135 \times 10^{0}$	$.23 \times 10^{-3}$	$.7135 \times 10^{0}$	$.41 \times 10^{-3}$
λ_5	$.1255 \times 10^{0}$	$.1254 \times 10^{0}$	$.92 \times 10^{-4}$	$.1255 \times 10^{0}$	$.17 \times 10^{-3}$
λ_6	$.7102 \times 10^{-1}$	$.7100 \times 10^{-1}$	$.82 \times 10^{-4}$	$.7102 \times 10^{-1}$	$.13 \times 10^{-3}$
λ_7	$.1229 \times 10^{-1}$	$.1230 \times 10^{-1}$	$.25 \times 10^{-4}$	$.1229 \times 10^{-1}$	$.54 \times 10^{-4}$
λ_8	$.2794 \times 10^{-2}$	$.2796 \times 10^{-2}$	$.16 \times 10^{-4}$	$.2794 \times 10^{-2}$	$.26 \times 10^{-4}$
λ_9	0	$.2856 \times 10^{-6}$	$.77 \times 10^{-7}$	$.3075 \times 10^{-6}$	$.28 \times 10^{-6}$

Table 13. Monte Carlo Study on Eigenvalue Structure of A Sum of Damped Sinusoids with Medium Error

FUNCTION: $y(t) = 1.0\, e^{-.2t} \sin(.5\pi t) + .5\, e^{-.3t} \sin(\pi t) + .25\, e^{-.4t} \sin(1.5\pi t) + .10\, e^{-.5t} \sin(2\pi t)$

STRUCTURE PARAMETERS: $n=49$, $\Delta=.4$, $t_o=0$, $m=8$ and 9

NOISE DESCRIPTION: $E = .003$ ($\sigma = 8.7 \times 10^{-4}$)

Parameter	Unperturbed Value	Mean	Standard Deviation	Theoretical Mean	Theoretical Bound on Standard Deviation
λ_1	$.1114 \times 10^{2}$	$.1113 \times 10^{2}$	$.87 \times 10^{-2}$	$.1114 \times 10^{2}$	$.16 \times 10^{-1}$
λ_2	$.5208 \times 10^{1}$	$.5207 \times 10^{1}$	$.47 \times 10^{-2}$	$.5208 \times 10^{1}$	$.11 \times 10^{-1}$
λ_3	$.9654 \times 10^{0}$	$.9653 \times 10^{0}$	$.23 \times 10^{-2}$	$.9654 \times 10^{0}$	$.48 \times 10^{-2}$
λ_4	$.7135 \times 10^{0}$	$.7132 \times 10^{0}$	$.23 \times 10^{-2}$	$.7135 \times 10^{0}$	$.41 \times 10^{-2}$
λ_5	$.1255 \times 10^{0}$	$.1253 \times 10^{0}$	$.92 \times 10^{-3}$	$.1255 \times 10^{0}$	$.17 \times 10^{-2}$
λ_6	$.7102 \times 10^{-1}$	$.7086 \times 10^{-1}$	$.82 \times 10^{-3}$	$.7105 \times 10^{-1}$	$.13 \times 10^{-2}$
λ_7	$.1229 \times 10^{-1}$	$.1243 \times 10^{-1}$	$.25 \times 10^{-3}$	$.1232 \times 10^{-1}$	$.54 \times 10^{-3}$
λ_8	$.2794 \times 10^{-2}$	$.2840 \times 10^{-2}$	$.16 \times 10^{-3}$	$.2826 \times 10^{-2}$	$.26 \times 10^{-3}$
λ_9	0	$.2856 \times 10^{-4}$	$.77 \times 10^{-5}$	$.3075 \times 10^{-4}$	$.28 \times 10^{-4}$

Table 14. Monte Carlo Study on Eignevalue Structure of a Sum of Damped Sinusoids with High Error

FUNCTION: $y(t) = 1.0\, e^{-.2t} \sin(.5\pi t) + .5\, e^{-.3t} \sin(\pi t) + .25\, e^{-.4t} \sin(1.5\pi t) + .10\, e^{-.5t} \sin(2\pi t)$

STRUCTURE PARAMETERS: $n=49$, $\Delta=.4$, $t_o=0$, $m=8$ and 9

NOISE DESCRIPTION: $E = .01$ $(\sigma = 2.9 \times 10^{-3})$

Parameter	Unperturbed Value	Mean	Standard Deviation	Theoretical Mean	Theoretical Bound on Standard Deviation
λ_1	$.1114 \times 10^{2}$	$.1113 \times 10^{2}$	$.29 \times 10^{-1}$	$.1114 \times 10^{2}$	$.55 \times 10^{-1}$
λ_2	$.5208 \times 10^{1}$	$.5204 \times 10^{1}$	$.16 \times 10^{-1}$	$.5208 \times 10^{1}$	$.37 \times 10^{-1}$
λ_3	$.9654 \times 10^{0}$	$.9652 \times 10^{0}$	$.76 \times 10^{-2}$	$.9658 \times 10^{0}$	$.16 \times 10^{-1}$
λ_4	$.7135 \times 10^{0}$	$.7127 \times 10^{0}$	$.76 \times 10^{-2}$	$.7139 \times 10^{0}$	$.14 \times 10^{-1}$
λ_5	$.1255 \times 10^{0}$	$.1253 \times 10^{0}$	$.31 \times 10^{-2}$	$.1259 \times 10^{0}$	$.58 \times 10^{-2}$
λ_6	$.7102 \times 10^{-1}$	$.7077 \times 10^{-1}$	$.27 \times 10^{-2}$	$.7137 \times 10^{-1}$	$.44 \times 10^{-2}$
λ_7	$.1229 \times 10^{-1}$	$.1300 \times 10^{-1}$	$.88 \times 10^{-3}$	$.1264 \times 10^{-1}$	$.18 \times 10^{-2}$
λ_8	$.2794 \times 10^{-2}$	$.3185 \times 10^{-2}$	$.55 \times 10^{-3}$	$.3144 \times 10^{-2}$	$.91 \times 10^{-3}$
λ_9	0	$.3158 \times 10^{-3}$	$.86 \times 10^{-4}$	$.3417 \times 10^{-3}$	$.31 \times 10^{-3}$

Table 15 summarizes all studies available to the author on the statistical behavior of the noise eigenvalue introduced by the perturbation of the single zero eigenvalue of the R_o matrix of order one greater than is correct. The first four problems in table 15 correspond to example problems previously studied with guassian noise in reference 2. The first two example problems reported in reference 2 did not include standard deviation data. The remainder of the data in this table corresponds to numerical results obtained by the present author with uniform noise as defined in Section 4. The majority of these latter results are for problems described in the immediately preceding tables. The final two problems (Extra #1 and Extra #2) are the result of the present effort for a simple sinusoid, $y = \sin \pi t$ with $t_o = 0$ and $n = 99$. For the calculations labeled Extra #1, the sample spacing was $\Delta = .5$ while Extra #2 is the equivalent problem using $\Delta = .1$. For both these latter problems eleven Monte Carlo experiments were performed. Again equation (45) of Section 4 was used to evaluate the theoretical value of the $m+1^{st}$ eigenvalue and equation (47) provided the theoretical bound on the standard deviation. For the most part, the results are very favorable and speak for themselves. However, it should be noted in all cases the data taken from reference 2 provides theoretical estimates of the mean eigenvalue lower than experimentally observed. This observation was documented in reference 2. The present author's results with a damped sinusoid are consistent with this observation. However, the results for the sum of damped sinusoids and the simple sinusoid provide the opposite situation.

The conclusion is that first order perturbation theory provides rather remarkable results for the problems examined to date. What is needed is additional studies using eigenvalue structures with near degeneracies and with higher noise levels to stress the theory into the region where it might theoretically be expected to break down.

The next six tables present numerical results directed at evaluating the error analysis developed in Section 3 of this report. In each case the Monte Carlo runs were performed at the noise level stated in the tables. The bound associated with the norm of the expected value of the error vector was computed using equation (38). The bounds generated for the individual components are simply the result of assigning all the error to one component. The adjusted bounds assume the error is uniformly distributed among the

Table 15. Summary of Monte Carlo Studies on the m+1st Eigenvalue

Problem	Standard Deviation of Noise	Mean Value of m+1st Eigenvalue	Standard Deviation	Theoretical Value of m+1st Eigenvalue	Theoretical Bound on Standard Deviation
Ex. 1 (ref 2)	$.1 \times 10^{-1}$	$.154 \times 10^{-1}$		$.150 \times 10^{-1}$	
	$.5 \times 10^{-1}$	$.386 \times 10^{0}$		$.375 \times 10^{0}$	
	$.7 \times 10^{-1}$	$.757 \times 10^{0}$		$.735 \times 10^{0}$	
	$.1 \times 10^{0}$	$.154 \times 10^{1}$		$.150 \times 10^{1}$	
	$.5 \times 10^{0}$	$.386 \times 10^{2}$		$.375 \times 10^{2}$	
Ex. 2 (ref 2)	$.1 \times 10^{-2}$	$.52 \times 10^{-4}$		$.50 \times 10^{-4}$	
	$.5 \times 10^{-2}$	$.131 \times 10^{-2}$		$.125 \times 10^{-2}$	
	$.9 \times 10^{-2}$	$.423 \times 10^{-2}$		$.405 \times 10^{-2}$	
	$.1 \times 10^{-1}$	$.52 \times 10^{-2}$		$.50 \times 10^{-2}$	
	$.2 \times 10^{-1}$	$.21 \times 10^{-1}$		$.20 \times 10^{-1}$	
	$.3 \times 10^{-1}$	$.47 \times 10^{-1}$		$.45 \times 10^{-1}$	
	$.4 \times 10^{-1}$	$.836 \times 10^{-1}$		$.80 \times 10^{-1}$	
Ex. 3 (ref 2)	$.1 \times 10^{-2}$	$.22 \times 10^{-3}$	$.492 \times 10^{-4}$	$.20 \times 10^{-3}$	$.10 \times 10^{-3}$
	$.3 \times 10^{-2}$	$.197 \times 10^{-2}$	$.443 \times 10^{-3}$	$.18 \times 10^{-2}$	$.90 \times 10^{-3}$
	$.5 \times 10^{-2}$	$.546 \times 10^{-2}$	$.123 \times 10^{-2}$	$.50 \times 10^{-2}$	$.25 \times 10^{-2}$
	$.8 \times 10^{-2}$	$.139 \times 10^{-1}$	$.315 \times 10^{-2}$	$.128 \times 10^{-1}$	$.64 \times 10^{-2}$
	$.1 \times 10^{-1}$	$.218 \times 10^{-1}$	$.492 \times 10^{-2}$	$.20 \times 10^{-1}$	$.10 \times 10^{-1}$
Ex. 4 (ref 2)	$.5 \times 10^{-3}$	$.546 \times 10^{-4}$	$.123 \times 10^{-4}$	$.50 \times 10^{-4}$	$.25 \times 10^{-4}$
	$.1 \times 10^{-2}$	$.218 \times 10^{-3}$	$.493 \times 10^{-4}$	$.20 \times 10^{-3}$	$.10 \times 10^{-3}$
	$.2 \times 10^{-2}$	$.871 \times 10^{-3}$	$.196 \times 10^{-3}$	$.80 \times 10^{-3}$	$.40 \times 10^{-3}$
Table 9	$.51 \times 10^{-2}$	$.1255 \times 10^{-2}$	$.41 \times 10^{-3}$	$.1241 \times 10^{-2}$	$.54 \times 10^{-3}$
Table 10	$.51 \times 10^{-1}$	$.1254 \times 10^{0}$	$.41 \times 10^{-1}$	$.1241 \times 10^{0}$	$.54 \times 10^{-1}$
Table 11	$.16 \times 10^{0}$	$.1246 \times 10^{1}$	$.40 \times 10^{0}$	$.1237 \times 10^{1}$	$.54 \times 10^{0}$
Table 12	$.87 \times 10^{-4}$	$.2856 \times 10^{-6}$	$.77 \times 10^{-7}$	$.3075 \times 10^{-6}$	$.28 \times 10^{-6}$

TABLE 15 - Continued

Table 13	$.87 \times 10^{-3}$	$.2856 \times 10^{-4}$	$.77 \times 10^{-5}$	$.3075 \times 10^{-4}$	$.28 \times 10^{-4}$
Table 14	$.29 \times 10^{-2}$	$.3158 \times 10^{-3}$	$.86 \times 10^{-4}$	$.3417 \times 10^{-3}$	$.31 \times 10^{-3}$
Extra #1	$.1 \times 10^{-2}$	$.7872 \times 10^{-5}$	$.11 \times 10^{-5}$	$.8083 \times 10^{-5}$	$.24 \times 10^{-5}$
	$.1 \times 10^{-1}$	$.7872 \times 10^{-3}$	$.11 \times 10^{-3}$	$.8083 \times 10^{-3}$	$.24 \times 10^{-3}$
Extra #2	$.1 \times 10^{-2}$	$.7663 \times 10^{-5}$	$.13 \times 10^{-5}$	$.8083 \times 10^{-5}$	$.24 \times 10^{-5}$
	$.1 \times 10^{-1}$	$.7663 \times 10^{-3}$	$.13 \times 10^{-3}$	$.8083 \times 10^{-3}$	$.24 \times 10^{-3}$

individual components of the error vector and thus is simply the bound value divided by m. Again all results are encouraging when viewed with respect to the implied standard deviations and number of samples used. In tables 16, 17, 19 one notes that the experimentally generated estimate of the norm of the expected value of the error vector exceeds the theoretical bound. However in each of these cases the expected value of the second moment of the norm of the error vector is relatively high.

The motivation for presenting the next four tables is independent of the previous sections of this report. In tables 22 and 23 are displayed computed frequencies for values of m one and two greater than is correct for ten Monte Carlo trials. Thus the roots labeled z_3 and z_4 are extraneous roots resulting from the inclusion of noise. It should be noted that although they vary much more than the valid roots z_1, z_2 they are by no means random. Further if the analyst had but two of the trial results, say trial numbers 7 and 8, the variation in z_3 is no more than that observed in the true root. Similar results are depicted in tables 24 and 25. Unfortunately for the examples presented the extraneous roots for the most part fall on the real axis in the z plane. This appears to be a characteristic of the sample spacing however, and the author has observed similar behavior with respect to the lack of randomness of extraneous roots at quite different phases. The fact that the extraneous roots have been observed to be quite sensitive to sample spacing can potentially assist the analyst with their identification. However, due to the sensitivity of the eigenstructure of R_o to sample spacing, extreme caution must be exercised.

In Section 4 the author pointed out that first order perturbation theory suggests that for values of m in excess of the correct value a band of eigenvalues of roughly the size $(n-m+1)\sigma^2$ would be generated. The band should have a width on the order of less than twice the standard deviation computed from equation 47. From table 26, for $m = 12$ one can compute the mean value of the four lowest eigenvalues to be 3.85×10^{-6}. The width of the band is simply $\lambda_9 - \lambda_{12} = 3.7 \times 10^{-6}$. From equation 45 of Section 4, the expected value of these extraneous eigenvalues is estimated to be 3.20×10^{-6}. Evaluating equation 47, one estimates that the band should have width less than 6.9×10^{-6} which indeed it has.

Table 16. Monte Carlo Study of Solution Vector for a Damped Sinusoid with Low Error

FUNCTION: $y(t) = e^{-.05t} \sin(\pi t)$

STRUCTURE PARAMETERS: $n=49$, $\Delta=.4$, $t_o=0$, $m=2$ and 3

NOISE DESCRIPTION: $E = .0178$ ($\sigma = 5.1 \times 10^{-3}$)

i	$(\vec{\alpha}_o)_i$	$\lvert\langle\vec{\alpha}-\vec{\alpha}_o\rangle_i\rvert$	Bound	Adjusted Bound	$\langle(\vec{\alpha}-\vec{\alpha}_o)_i^2\rangle^{1/2}$	Bound	Adjusted Bound
1	-.6058	$.20 \times 10^{-3}$	$.22 \times 10^{-3}$	$.16 \times 10^{-3}$	$.63 \times 10^{-3}$	$.61 \times 10^{-2}$	$.43 \times 10^{-2}$
2	.9608	$.50 \times 10^{-3}$	$.22 \times 10^{-3}$	$.16 \times 10^{-3}$	$.78 \times 10^{-3}$	$.61 \times 10^{-2}$	$.43 \times 10^{-2}$

$\lVert\langle\vec{\alpha}-\vec{\alpha}_o\rangle\rVert$	Bound	$\langle\lVert\vec{\alpha}-\vec{\alpha}_o\rVert^2\rangle^{1/2}$	Bound
$.54 \times 10^{-3}$	$.22 \times 10^{-3}$	$.10 \times 10^{-2}$	$.61 \times 10^{-2}$

Table 17. Monte Carlo Study of Solution Vector for a Damped Sinusoid with Medium Error

FUNCTION: $y(t) = e^{-.05t} \sin(\pi t)$

STRUCTURE PARAMETERS: $n=49$, $\Delta=.4$, $t_o=0$, $m=2$ and 3

NOISE DESCRIPTION: $E = .178$ ($\sigma = 5.1 \times 10^{-2}$)

i	$(\vec{\alpha}_o)_i$	$\lvert\langle\vec{\alpha} - \vec{\alpha}_o\rangle_i\rvert$	Bound	Adjusted Bound	$\langle(\vec{\alpha} - \vec{\alpha}_o)_i^2\rangle^{1/2}$	Bound	Adjusted Bound
1	-.6058	$.14 \times 10^{-1}$	$.22 \times 10^{-1}$	$.16 \times 10^{-1}$	$.18 \times 10^{-1}$	$.65 \times 10^{-1}$	$.46 \times 10^{-1}$
2	.9608	$.20 \times 10^{-1}$	$.22 \times 10^{-1}$	$.16 \times 10^{-1}$	$.22 \times 10^{-1}$	$.65 \times 10^{-1}$	$.46 \times 10^{-1}$

$\lVert\langle\vec{\alpha} - \vec{\alpha}_o\rangle\rVert$	Bound	$\langle\lVert\vec{\alpha} - \vec{\alpha}_o\rVert^2\rangle^{1/2}$	Bound
$.24 \times 10^{-1}$	$.22 \times 10^{-1}$	$.29 \times 10^{-1}$	$.65 \times 10^{-1}$

Table 18. Monte Carlo Study of Solution Vector for a Damped Sinusoid with High Error

FUNCTION: $y(t) = e^{-.05t} \sin(\pi t)$

STRUCTURE PARAMETERS: $n=49$, $\Delta=.4$, $t_o=0$, $m=2$ and 3

NOISE DESCRIPTION: $E = .562$ $(\sigma = 1.6 \times 10^{-1})$

i	$(\vec{\alpha}_o)_i$	$\lvert\langle\vec{\alpha}-\vec{\alpha}_o\rangle_i\rvert$	Bound	Adjusted Bound	$\langle(\vec{\alpha}-\vec{\alpha}_o)_i^2\rangle^{1/2}$	Bound	Adjusted Bound
1	-.6058	$.12 \times 10^0$	$.22 \times 10^0$	$.16 \times 10^0$	$.13 \times 10^0$	$.30 \times 10^0$	$.21 \times 10^0$
2	.9608	$.15 \times 10^0$	$.22 \times 10^0$	$.16 \times 10^0$	$.16 \times 10^0$	$.30 \times 10^0$	$.21 \times 10^0$

$\lVert\langle\vec{\alpha}-\vec{\alpha}_o\rangle\rVert$	Bound	$\langle\lVert\vec{\alpha}-\vec{\alpha}_o\rVert^2\rangle^{1/2}$	Bound
$.19 \times 10^0$	$.22 \times 10^0$	$.21 \times 10^0$	$.30 \times 10^0$

Table 19. Monte Carlo Study of Solution Vector for a Sum of Damped Sinusoids with Low Error

FUNCTION: $y(t) = 1.0\, e^{-.2t} \sin(.5\pi t) + .5\, e^{-.3t} \sin(\pi t) + .25\, e^{-.4t} \sin(1.5\pi t) + .10\, e^{-.5t} \sin(2\pi t)$

STRUCTURE PARAMETERS: $n=49$, $\Delta=.4$, $t_o=0$, $m=8$ and 9

NOISE DESCRIPTION: $E = .0003$ ($\sigma = 8.7 \times 10^{-5}$)

i	$(\vec{\alpha}_o)_i$	$\lvert\langle\vec{\alpha}-\vec{\alpha}_o\rangle_i\rvert$	Bound	Adjusted Bound	$\langle(\vec{\alpha}-\vec{\alpha}_o)_i^2\rangle^{1/2}$	Bound	Adjusted Bound
1	-.1904	0	$.13 \times 10^{-3}$	$.46 \times 10^{-4}$	$.13 \times 10^{-2}$	$.85 \times 10^{-2}$	$.30 \times 10^{-2}$
2	.7715	0	$.13 \times 10^{-3}$	$.46 \times 10^{-4}$	$.12 \times 10^{-2}$	$.85 \times 10^{-2}$	$.30 \times 10^{-2}$
3	-.0531	$.90 \times 10^{-4}$	$.13 \times 10^{-3}$	$.46 \times 10^{-4}$	$.16 \times 10^{-2}$	$.85 \times 10^{-2}$	$.30 \times 10^{-2}$
4	.5785	$.10 \times 10^{-3}$	$.13 \times 10^{-3}$	$.46 \times 10^{-4}$	$.11 \times 10^{-2}$	$.85 \times 10^{-2}$	$.30 \times 10^{-2}$
5	.0401	$.40 \times 10^{-4}$	$.13 \times 10^{-3}$	$.46 \times 10^{-4}$	$.96 \times 10^{-3}$	$.85 \times 10^{-2}$	$.30 \times 10^{-2}$
6	.4407	0	$.13 \times 10^{-3}$	$.46 \times 10^{-4}$	$.77 \times 10^{-3}$	$.85 \times 10^{-2}$	$.30 \times 10^{-2}$
7	.0822	$.50 \times 10^{-4}$	$.13 \times 10^{-3}$	$.46 \times 10^{-4}$	$.37 \times 10^{-3}$	$.85 \times 10^{-2}$	$.30 \times 10^{-2}$
8	.3263	$.10 \times 10^{-3}$	$.13 \times 10^{-3}$	$.46 \times 10^{-4}$	$.51 \times 10^{-3}$	$.85 \times 10^{-2}$	$.30 \times 10^{-2}$

$\lVert\langle\vec{\alpha}-\vec{\alpha}_o\rangle\rVert$	Bound	$\langle\lVert\vec{\alpha}-\vec{\alpha}_o\rVert^2\rangle^{1/2}$	Bound
$.18 \times 10^{-3}$	$.13 \times 10^{-3}$	$.30 \times 10^{-2}$	$.85 \times 10^{-2}$

Table 20. Monte Carlo Study of Solution Vector for a Sum of Damped Sinusoids with Medium Error

FUNCTION: $y(t) = 1.0\, e^{-.2t} \sin(.5\pi t) + .5\, e^{-.3t} \sin(\pi t) + .25\, e^{-.4t} \sin(1.5\pi t) + .10\, e^{-.5t} \sin(2\pi t)$

STRUCTURE PARAMETERS: $n = 49$, $\Delta = .4$, $t_o = 0$, $m = 8$ and 9

NOISE DESCRIPTION: $E = .003$ ($\sigma = 8.7 \times 10^{-4}$)

i	$(\vec{\alpha}_o)_i$	$\lvert\langle \vec{\alpha} - \vec{\alpha}_o \rangle_i\rvert$	Bound	Adjusted Bound	$\langle (\vec{\alpha} - \vec{\alpha}_o)_i^2 \rangle^{1/2}$	Bound	Adjusted Bound
1	-.1904	$.12 \times 10^{-2}$	$.13 \times 10^{-1}$	$.46 \times 10^{-2}$	$.13 \times 10^{-1}$	$.87 \times 10^{-1}$	$.31 \times 10^{-1}$
2	.7715	$.27 \times 10^{-2}$	$.13 \times 10^{-1}$	$.46 \times 10^{-2}$	$.12 \times 10^{-1}$	$.87 \times 10^{-1}$	$.31 \times 10^{-1}$
3	-.0531	$.26 \times 10^{-2}$	$.13 \times 10^{-1}$	$.46 \times 10^{-2}$	$.16 \times 10^{-1}$	$.87 \times 10^{-1}$	$.31 \times 10^{-1}$
4	.5785	$.25 \times 10^{-2}$	$.13 \times 10^{-1}$	$.46 \times 10^{-2}$	$.11 \times 10^{-1}$	$.87 \times 10^{-1}$	$.31 \times 10^{-1}$
5	.0401	$.23 \times 10^{-2}$	$.13 \times 10^{-1}$	$.46 \times 10^{-2}$	$.10 \times 10^{-1}$	$.87 \times 10^{-1}$	$.31 \times 10^{-1}$
6	.4407	$.19 \times 10^{-2}$	$.13 \times 10^{-1}$	$.46 \times 10^{-2}$	$.80 \times 10^{-2}$	$.87 \times 10^{-1}$	$.31 \times 10^{-1}$
7	.0822	$.11 \times 10^{-2}$	$.13 \times 10^{-1}$	$.46 \times 10^{-2}$	$.38 \times 10^{-2}$	$.87 \times 10^{-1}$	$.31 \times 10^{-1}$
8	.3263	$.60 \times 10^{-3}$	$.13 \times 10^{-1}$	$.46 \times 10^{-2}$	$.50 \times 10^{-2}$	$.87 \times 10^{-1}$	$.31 \times 10^{-1}$

$\lVert\langle \vec{\alpha} - \vec{\alpha}_o \rangle\rVert$	Bound	$\langle \lVert\vec{\alpha} - \vec{\alpha}_o\rVert^2 \rangle^{1/2}$	Bound
$.57 \times 10^{-2}$	$.13 \times 10^{-1}$	$.30 \times 10^{-1}$	$.87 \times 10^{-1}$

Table 21. Monte Carlo Study of Solution Vector for a Sum of Damped Sinusoids with High Error

FUNCTION: $y(t) = 1.0\, e^{-.2t} \sin(.5\pi t) + .5\, e^{-.3t} \sin(\pi t) + .25\, e^{-.4t} \sin(1.5\pi t) + .10\, e^{-.5t} \sin(2\pi t)$

STRUCTURE PARAMETERS: $n=49$, $\Delta=.4$, $t_o=0$, $m=8$ and 9

NOISE DESCRIPTION: $E = .01$ $(\sigma = 2.9 \times 10^{-3})$

i	$(\vec{\alpha}_o)_i$	$\lvert\langle\vec{\alpha} - \vec{\alpha}_o\rangle_i\rvert$	Bound	Adjusted Bound	$\langle(\vec{\alpha} - \vec{\alpha}_o)_i^2\rangle^{1/2}$	Bound	Adjusted Bound
1	-.1904	$.18 \times 10^{-1}$	$.15 \times 10^{0}$	$.52 \times 10^{-1}$	$.43 \times 10^{-1}$	$.34 \times 10^{0}$	$.12 \times 10^{0}$
2	.7715	$.27 \times 10^{-1}$	$.15 \times 10^{0}$	$.52 \times 10^{-1}$	$.47 \times 10^{-1}$	$.34 \times 10^{0}$	$.12 \times 10^{0}$
3	-.0531	$.33 \times 10^{-1}$	$.15 \times 10^{0}$	$.52 \times 10^{-1}$	$.62 \times 10^{-1}$	$.34 \times 10^{0}$	$.12 \times 10^{0}$
4	.5785	$.32 \times 10^{-1}$	$.15 \times 10^{0}$	$.52 \times 10^{-1}$	$.53 \times 10^{-1}$	$.34 \times 10^{0}$	$.12 \times 10^{0}$
5	.0401	$.29 \times 10^{-1}$	$.15 \times 10^{0}$	$.52 \times 10^{-1}$	$.47 \times 10^{-1}$	$.34 \times 10^{0}$	$.12 \times 10^{0}$
6	.4407	$.17 \times 10^{-1}$	$.15 \times 10^{0}$	$.52 \times 10^{-1}$	$.33 \times 10^{-1}$	$.34 \times 10^{0}$	$.12 \times 10^{0}$
7	.0822	$.81 \times 10^{-2}$	$.15 \times 10^{0}$	$.52 \times 10^{-1}$	$.15 \times 10^{-1}$	$.34 \times 10^{0}$	$.12 \times 10^{0}$
8	.3263	$.15 \times 10^{-2}$	$.15 \times 10^{0}$	$.52 \times 10^{-1}$	$.15 \times 10^{-1}$	$.34 \times 10^{0}$	$.12 \times 10^{0}$

$\lVert\langle\vec{\alpha} - \vec{\alpha}_o\rangle\rVert$	Bound	$\langle\lVert\vec{\alpha} - \vec{\alpha}_o\rVert^2\rangle^{1/2}$	Bound
$.65 \times 10^{-1}$	$.15 \times 10^{0}$	$.12 \times 10^{0}$	$.34 \times 10^{0}$

Table 22. Monte Carlo Study on First Extraneous root for Damped Sinusoid

FUNCTION: $y(t) = e^{-.05t} \sin \pi t$

STRUCTURE PARAMETERS: $n = 49$, $\Delta = .4$, $t_o = 0$, $m = 3$

NOISE DESCRIPTION: $E = .562$ ($\sigma = 2.6 \times 10^{-2}$)

Trial No.	z_1, z_2	z_3
1	.301 ± .883 j	−.355 + 0. j
2	.298 ± .893 j	−.672 + 0. j
3	.282 ± .894 j	−.312 + 0. j
4	.300 ± .887 j	−.474 + 0. j
5	.294 ± .897 j	−.721 + 0. j
6	.307 ± .890 j	−.465 + 0. j
7	.302 ± .903 j	−.514 + 0. j
8	.313 ± .897 j	−.505 + 0. j
9	.282 ± .892 j	−.425 + 0. j
10	.316 ± .907 j	−.616 + 0. j

Table 23. Monte Carlo Study on Two Extraneous Roots for Damped Sinusoids

FUNCTION: $y(t) = e^{-.05t} \sin \pi t$

STRUCTURE PARAMETERS: n=49, Δ=.4, t_o=0, m=4

NOISE DESCRIPTION: E = .562 ($\sigma = 2.6 \times 10^{-2}$)

Trial No.	z_1, z_2	z_3	z_4
1	.286 ± .913 j	.585 + 0. j	-.766 + 0. j
2	.294 ± .892 j	.146 + 0. j	-.737 + 0. j
3	.278 ± .900 j	.128 + 0. j	-.430 + 0. j
4	.287 ± .897 j	.401 + 0. j	-.732 + 0. j
5	.295 ± .899 j	-.113 + 0. j	-.651 + 0. j
6	.295 ± .914 j	.328 + 0. j	-.679 + 0. j
7	.296 ± .905 j	.283 0. j	-.691 + 0. j
8	.304 ± .901 j	.288 + 0. j	-.690 + 0. j
9	.313 ± .881 j	-.287 + .421 j	-.287 - .421 j
10	.297 ± .926 j	.581 + 0. j	-.850 + 0. j

Table 24. Monte Carlo Study on First Extraneous Root for a Sum of Damped Sinusoids

FUNCTION: $y(t) = 1.0\, e^{-.2t} \sin(.5\pi t) + .5\, e^{-.3t} \sin(\pi t) + .25\, e^{-.4t} \sin(1.5\pi t) + .10\, e^{-.5t} \sin(2\pi t)$

STRUCTURE PARAMETERS: $n=49,\ \Delta=.4,\ t_o=0,\ m=9$

NOISE DESCRIPTION: $E = .01\ (\sigma = 2.9 \times 10^{-3})$

Trial No.	z_7, z_8	z_9
1	$-.658 \pm .456\, j$	$.155 + 0.\, j$
2	$-.678 \pm .463\, j$	$-.301 + 0.\, j$
3	$-.677 \pm .486\, j$	$.316 + 0.\, j$
4	$-.670 \pm .450\, j$	$.068 + 0.\, j$
5	$-.633 \pm .439\, j$	$-.359 + 0.\, j$
6	$-.669 \pm .465\, j$	$-.074 + 0.\, j$
7	$-.675 \pm .420\, j$	$-.087 + 0.\, j$
8	$-.673 \pm .428\, j$	$-.0712 + 0.\, j$
9	$-.648 \pm .499\, j$	$-.112 + 0.\, j$
10	$-.631 \pm .471\, j$	$-.141 + 0.\, j$

Table 25. Monte Carlo Study on Two Extraneous Roots for a Sum of Damped Sinusoids

FUNCTION: $y(t) = 1.0\, e^{-.2t} \sin(.5\pi t) + .5\, e^{-.3t} \sin(\pi t) + .25\, e^{-.4t} \sin(1.5\pi t) + .10\, e^{-.5t} \sin(2\pi t)$

STRUCTURE PARAMETERS: $n=49$, $\Delta=.4$, $t_o=0$, $m=10$

NOISE DESCRIPTION: $E = .01$ ($\sigma = 2.9 \times 10^{-3}$)

Trial No.	z_7, z_8	z_9	z_{10}
1	-.658 ± .487 j	-.886 + 0. j	+.887 + 0. j
2	-.678 ± .498 j	-.890 + 0. j	+.800 + 0. j
3	-.677 ± .491 j	-.718 + 0. j	+.846 + 0. j
4	-.665 ± .470 j	-.867 + 0. j	.859 + 0. j
5	-.654 ± .473 j	-.900 + 0. j	.811 + 0. j
6	-.685 ± .475 j	-.862 + 0. j	.851 + 0. j

Table 26. Eigenvalue Structure versus m

FUNCTION: $y(t) = 1.0\, e^{-.2t} \sin(.5\pi t) + .5\, e^{-.3t} \sin(\pi t) + .25\, e^{-.4t} \sin(1.5\pi t) + .10\, e^{-.5t} \sin(2\pi t)$

STRUCTURE PARAMETERS: $n=49$, $\Delta=.4$, $t_o=0$

NOISE DESCRIPTION: $E=.001$ ($\sigma = 2.9 \times 10^{-4}$)

m	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9	λ_{10}	λ_{11}	λ_{12}
2	$.67 \times 10^{1}$	$.13 \times 10^{1}$										
3	$.81 \times 10^{1}$	$.23 \times 10^{1}$	$.20 \times 10^{0}$									
4	$.84 \times 10^{1}$	$.34 \times 10^{1}$	$.49 \times 10^{0}$	$.61 \times 10^{-1}$								
5	$.84 \times 10^{1}$	$.45 \times 10^{1}$	$.70 \times 10^{0}$	$.23 \times 10^{0}$	$.18 \times 10^{-1}$							
6	$.91 \times 10^{1}$	$.49 \times 10^{1}$	$.78 \times 10^{0}$	$.49 \times 10^{0}$	$.69 \times 10^{-1}$	$.96 \times 10^{-2}$						
7	$.10 \times 10^{2}$	$.49 \times 10^{1}$	$.79 \times 10^{0}$	$.71 \times 10^{0}$	$.12 \times 10^{0}$	$.39 \times 10^{-1}$	$.45 \times 10^{-2}$					
8	$.11 \times 10^{2}$	$.52 \times 10^{1}$	$.96 \times 10^{0}$	$.71 \times 10^{0}$	$.13 \times 10^{0}$	$.71 \times 10^{-1}$	$.12 \times 10^{-1}$	$.28 \times 10^{-2}$				
9	$.11 \times 10^{2}$	$.59 \times 10^{1}$	$.10 \times 10^{1}$	$.83 \times 10^{0}$	$.15 \times 10^{0}$	$.72 \times 10^{-1}$	$.12 \times 10^{-1}$	$.36 \times 10^{-2}$	$.47 \times 10^{-5}$			
10	$.11 \times 10^{2}$	$.68 \times 10^{1}$	$.10 \times 10^{1}$	$.93 \times 10^{0}$	$.15 \times 10^{0}$	$.86 \times 10^{-1}$	$.14 \times 10^{-1}$	$.46 \times 10^{-2}$	$.59 \times 10^{-5}$	$.34 \times 10^{-5}$		
11	$.12 \times 10^{2}$	$.70 \times 10^{1}$	$.11 \times 10^{1}$	$.93 \times 10^{0}$	$.15 \times 10^{0}$	$.86 \times 10^{-1}$	$.14 \times 10^{-1}$	$.46 \times 10^{-2}$	$.60 \times 10^{-5}$	$.34 \times 10^{-5}$	$.30 \times 10^{-5}$	
12	$.13 \times 10^{2}$	$.70 \times 10^{1}$	$.11 \times 10^{1}$	$.93 \times 10^{0}$	$.15 \times 10^{0}$	$.86 \times 10^{-1}$	$.14 \times 10^{-1}$	$.46 \times 10^{-2}$	$.60 \times 10^{-5}$	$.38 \times 10^{-5}$	$.33 \times 10^{-5}$	$.23 \times 10^{-5}$

Now suppose one doesn't know σ^2. If one tentatively assumes that m=8, then λ_9 for m=9 provides an estimate of $\sigma^2 \approx \lambda_9/41 = 1.15 \times 10^{-7}$. The mean value of the lowest eigenvalues of m=12 is estimated to be 4.36×10^{-6} which is still reasonably consistent with the actual value computed. Similarly the band width is estimated to be bound by 9.4×10^{-6} which it is.

Now suppose an analyst assumes that m = 7. Using the λ_8 value for m = 8 provides an estimate of $\sigma^2 = 6.67 \times 10^{-5}$ which implies the expected value of the noise eigenvalues at m = 12 should be 2.53×10^{-3}. The computed value is 9.23×10^{-4} which is significantly different. From the table, the band width is approximately 4.6×10^{-3}. Using the estimated value of σ^2 and equation (47) one expects it to be less than 5.5×10^{-3}.

Thus at the low noise level studied in table 26, one finds that equation (47) by itself is of little use with respect to the determination of m. Equation (45) is somewhat more useful in that the incorrect assumption on m, results in a significant discrepancy between theory and computed result. However the most obvious difference between assuming m=7 or m=8 is the nature of the distribution of the noise eigenvalues. The assumption that the correct rank is 8 results in the 4 noise eigenvalues at level m=12 being rather symmetrically placed around their experiment mean value of 3.8×10^{-6}. The assumption that the correct rank is 7, results in an obviously non symmetric distribution of the 5 noise eigenvalues at level m=12. Table 27 presents a study identical to that of table 26 except that the noise level has been increased by a factor of 25. It is now much more difficult to decide where the chain of noise eigenvalues begin. To demonstrate, suppose that m and σ are unknown. If one assumes that m=8, then λ_9 resulting for m=9 provides an estimate of σ^2.

$$\sigma^2 \approx \frac{\lambda_9}{41} = 6.59 \times 10^{-5}$$

Based on this number, the expected value of the noise eigenvalues of m=12 is simply $38\,\sigma^2 = 2.5 \times 10^{-3}$. Using the values from table 27, one computes the mean value of the four lowest eigenvalues at m=12 to be 2.4×10^{-3}. The computed width of the noise band based on table 27 and the assumption that m=8 is simply $\lambda_9 - \lambda_{12} = 2.2 \times 10^{-3}$. From equation (47) one computes that the band is expected to have width

Table 27. Eigenvalue Structure versus m

FUNCTION: $y(t) = 1.0\, e^{-.2t} \sin(.5\pi t) + .5\, e^{-.3t} \sin(\pi t) + .25\, e^{-.4t} \sin(1.5\pi t) + .10\, e^{-.5t} \sin(2\pi t)$

STRUCTURE PARAMETERS: $n=49$, $\Delta=.4$, $t_o=0$

NOISE DESCRIPTION: $E = .025$ ($\sigma = 7.2 \times 10^{-3}$)

m	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9	λ_{10}	λ_{11}	λ_{12}
2	$.67 \times 10^{1}$	$.13 \times 10^{1}$										
3	$.81 \times 10^{1}$	$.22 \times 10^{1}$	$.20 \times 10^{0}$									
4	$.83 \times 10^{1}$	$.34 \times 10^{1}$	$.48 \times 10^{0}$	$.61 \times 10^{-1}$								
5	$.84 \times 10^{1}$	$.45 \times 10^{1}$	$.69 \times 10^{0}$	$.23 \times 10^{0}$	$.21 \times 10^{-1}$							
6	$.90 \times 10^{1}$	$.49 \times 10^{1}$	$.77 \times 10^{0}$	$.49 \times 10^{0}$	$.66 \times 10^{-1}$	$.12 \times 10^{-1}$						
7	$.10 \times 10^{2}$	$.49 \times 10^{1}$	$.78 \times 10^{0}$	$.71 \times 10^{0}$	$.12 \times 10^{0}$	$.40 \times 10^{-1}$	$.74 \times 10^{-2}$					
8	$.11 \times 10^{2}$	$.51 \times 10^{1}$	$.95 \times 10^{0}$	$.71 \times 10^{0}$	$.12 \times 10^{0}$	$.69 \times 10^{-1}$	$.17 \times 10^{-1}$	$.51 \times 10^{-2}$				
9	$.11 \times 10^{2}$	$.59 \times 10^{1}$	$.99 \times 10^{0}$	$.82 \times 10^{0}$	$.14 \times 10^{0}$	$.70 \times 10^{-1}$	$.17 \times 10^{-1}$	$.69 \times 10^{-2}$	$.27 \times 10^{-2}$			
10	$.11 \times 10^{2}$	$.67 \times 10^{1}$	$.10 \times 10^{1}$	$.90 \times 10^{0}$	$.14 \times 10^{0}$	$.81 \times 10^{-1}$	$.18 \times 10^{-1}$	$.75 \times 10^{-2}$	$.36 \times 10^{-2}$	$.21 \times 10^{-2}$		
11	$.12 \times 10^{2}$	$.70 \times 10^{1}$	$.10 \times 10^{1}$	$.90 \times 10^{0}$	$.14 \times 10^{0}$	$.81 \times 10^{-1}$	$.18 \times 10^{-1}$	$.75 \times 10^{-2}$	$.36 \times 10^{-2}$	$.21 \times 10^{-2}$	$.18 \times 10^{-2}$	
12	$.13 \times 10^{2}$	$.70 \times 10^{1}$	$.10 \times 10^{1}$	$.91 \times 10^{0}$	$.14 \times 10^{0}$	$.81 \times 10^{-1}$	$.18 \times 10^{-1}$	$.76 \times 10^{-2}$	$.36 \times 10^{-2}$	$.24 \times 10^{-2}$	$.20 \times 10^{-2}$	$.14 \times 10^{-2}$

less than 5.4×10^{-3}. Thus the tentative assumption that rank of the R_o matrix is 8 appears to be consistent with theory. Now suppose one assumes that m=7, and thus estimates that

$$\sigma^2 \approx 1.21 \times 10^{-4}$$

Based on this number, the expected value of the noise eigenvalues at m=12 is simply 4.61×10^{-3}. Using the values from table 27, one computes the mean value of the five lowest eigenvalues to be 3.4×10^{-3}. The width of the band is computed from the data to be 6.2×10^{-3}. The width of the band is computed from the data to be 6.2×10^{-3}. From equation (47) one computes that the band is expected to have width less than 1.0×10^{-2} which is consistent with the value above.

Thus at the higher noise level equation 45 still provides a degree of usefulness for deciding upon the correct value of m as does the nature of the resulting distributions. However for slightly higher noise levels one expects the problem of determining m to be quite formidable.

The last two tables, table 28 and 29 present results of incorporating the iteration methodology described in Section 7 at two different noise levels. Again, for the most part, the results speak for themselves. At the higher noise level of table 28, iteration results in nearly a factor of two reduction in S^2 while in table 29 the reduction is somewhat less. The tables adequately point out that Prony results in frequency components which to not correspond to minimums in the sense described in Section 7.

Table 28. Iteration Results Using Medicre Input

FUNCTION: $y(t) = 1.0\, e^{-.2t} \sin(.5\pi t) + .5\, e^{-.3t} \sin(\pi t) + .25\, e^{-.4t} \sin(1.5\pi t) + .10\, e^{-.5t} \sin(2\pi t)$

PRONY STRUCTURE PARAMETERS: $\Delta = .4$, $n=49$, $m=8$, $t_o=0.$

ITERATION STRUCTURE PARAMETERS: $\Delta=.1$, $n=250$, $m=8$, $t_o=0.$

NOISE DESCRIPTION: $E=.10$ $(\sigma = 2.9 \times 10^{-2})$

Parameter	Input Value	Prony	After Iteration
α_1	-.2	-.1980	-.1938
α_2	-.3	-.4808	-.3081
α_3	-.4	-.8495	-.3839
α_4	-.5	-.9574	-.8699
ω_1	1.571	1.582	1.569
ω_2	3.142	3.102	3.144
ω_3	4.712	4.883	4.707
ω_4	6.283	6.630	6.425
a_1	1.0	.9644	.9763
a_2	.50	.5906	.4955
a_3	.25	.3329	.2401
a_4	.10	.06154	.1447
$\partial S^2/\partial\alpha_1$		$-.1608 \times 10^{1}$	$.5552 \times 10^{-11}$
$\partial S^2/\partial\alpha_2$		$-.6834 \times 10^{0}$	$.1423 \times 10^{-11}$
$\partial S^2/\partial\alpha_3$		$-.7055 \times 10^{-1}$	$-.3906 \times 10^{-12}$
$\partial S^2/\partial\alpha_4$		$-.3880 \times 10^{-2}$	$.7458 \times 10^{-13}$
$\partial S^2/\partial\omega_1$		$.4546 \times 10^{1}$	$-.1819 \times 10^{-11}$

TABLE 28 - Continued

Parameter	Input Value	Prony	After Iteration
$\partial S^2/\partial\omega_2$		$-.4824 \times 10^{0}$	$-.1754 \times 10^{-12}$
$\partial S^2/\partial\omega_3$		$.8058 \times 10^{-1}$	$-.1203 \times 10^{-12}$
$\partial S^2/\partial\omega_4$		$-.4350 \times 10^{-2}$	$.9049 \times 10^{-13}$
$\partial S^2/\partial a_1$		$.1662 \times 10^{-11}$	$.1897 \times 10^{-11}$
$\partial S^2/\partial a_2$		$-.4683 \times 10^{-12}$	$.1596 \times 10^{-12}$
$\partial S^2/\partial a_3$		$-.9373 \times 10^{-12}$	$-.3604 \times 10^{-12}$
$\partial S^2/\partial a_4$		$.1777 \times 10^{-12}$	$-.4429 \times 10^{-12}$
S^2		.3740	.1977

Table 29. Iteration Results Using Good Input

FUNCTION: $y(t) = 1.0\, e^{-.2t} \sin(.5\pi t) + .5\, e^{-.3t} \sin(\pi t)$

$+ .25\, e^{-.4t} \sin(1.5\pi t) + .10\, e^{-.5t} \sin(2\pi t)$

PRONY STRUCTURE PARAMETERS: $\Delta = .4$, $n = 49$, $m = 8$, $t_o = 0$

ITERATION STRUCTURE PARAMETERS: $\Delta = .1$, $n = 250$, $m = 8$, $t_o = 0$

NOISE DESCRIPTION: $E = .01$ ($\sigma = 2.9 \times 10^{-3}$)

Parameter	Input Value	Prony	After Iteration
α_1	-.2	-.1996	-.1994
α_2	-.3	-.3138	-.3008
α_3	-.4	-.4134	-.3992
α_4	-.5	-.4925	-.5158
ω_1	1.571	1.530	1.571
ω_2	3.142	3.144	3.142
ω_3	4.712	4.705	4.712
ω_4	6.283	6.275	6.291
a_1	1.0	.9969	.9977
a_2	.50	.5093	.4996
a_3	.25	.2518	.2498
a_4	.10	.09773	.1020
$\partial s^2/\partial\alpha_1$		$-.1139 \times 10^{0}$	$-.9201 \times 10^{-12}$
$\partial s^2/\partial\alpha_2$		$-.1461 \times 10^{0}$	$-.6169 \times 10^{-13}$
$\partial s^2/\partial\alpha_3$		$-.2435 \times 10^{-1}$	$-.5290 \times 10^{-13}$
$\partial s^2/\partial\alpha_4$		$.7293 \times 10^{-3}$	$.2275 \times 10^{-13}$
$\partial s^2/\partial\omega_1$		$.7621 \times 10^{0}$	$-.1051 \times 10^{-11}$

TABLE 29 - Continued

Parameter	Input Value	Prony	After Iteration
$\partial S^2/\partial\omega_2$		$.5532 \times 10^{-1}$	$-.5259 \times 10^{-13}$
$\partial S^2/\partial\omega_3$		$-.2547 \times 10^{-1}$	$.4328 \times 10^{-13}$
$\partial S^2/\partial\omega_4$		$-.3704 \times 10^{-2}$	$-.1875 \times 10^{-15}$
$\partial S^2/\partial a_1$		$.3493 \times 10^{-12}$	$.3358 \times 10^{-12}$
$\partial S^2/\partial a_2$		$-.1656 \times 10^{-11}$	$.1372 \times 10^{-11}$
$\partial S^2/\partial a_3$		$-.1654 \times 10^{-12}$	$-.5157 \times 10^{-12}$
$\partial S^2/\partial a_4$		$-.3205 \times 10^{-12}$	$-.2639 \times 10^{-12}$
S^2		$.3295 \times 10^{-2}$	$.1989 \times 10^{-2}$

IX. SUMMARY AND CONCLUSIONS

In Section 3, a first order error analysis on the solution vector $\vec{\alpha}$ was developed based on the assumption that the correct rank of the unperturbed system was known. Equation (30) of this section indicates that the expected value of the solution vector should deteriorate as the minimum eigenvalue of the unperturbed matrix approaches the noise level, $(n-m+1)\sigma^2$. However, if but several statistically independent runs are available, one must be concerned about the statistical spread of the solution vector about its expected value. Equation (38) of Section 3 provides some insight with respect to this problem. The term in equation (38) which is proportional to σ^2 dominates only if

$$\frac{(n-m+1)\sigma^2}{\lambda_{min}} < \frac{m+1}{n-m+1}$$

Assuming n is reasonably large in comparison to m, one concludes that the statistical spread in solution vectors is most significant at low noise levels with respect to the magnitude of the lowest eigenvalue of the unperturbed system. This behavior is apparent in Tables 16 through 21 of the previous section. Both theoretical bounds and experimental values display this tendency for $\left\langle \|\vec{\alpha}-\vec{\alpha}_o\|^2 \right\rangle^{1/2}$ to more nearly approximate $\|\langle\vec{\alpha}-\vec{\alpha}_o\rangle\|$ at the higher noise levels.

In Section 4, first order perturbation theory was employed to study the statistical properties of the perturbed eigenvalues. Again the perturbation was due solely to random noise. It was found by equation (45) that the expected values of all perturbed eigenvalues simply shift upward by a value of $(n-m+1)\sigma^2$. This conclusion also holds for the unperturbed zero eigenvalues. Equation (47) indicates that the statistical variation of the perturbed zero eigenvalues can be controlled through problem structuring parameters and that in general this variation can be made to be reasonably small with respect to its expected value of $(n-m+1)\sigma^2$. However equation (47) also indicates that the positive unperturbed eigenvalues have a much greater statistical spread. Again this effect has been observed in tables 9 through 14.

The results of Sections 3 and 4 seem to indicate that the behavior of the Prony methodology with respect to noise is relatively insensitive to the exact statistical character of the noise. Thus results with uniformly distributed noise should differ little from guassian noise or any other distribution so long as it is symmetric, has zero mean and the same variance. It should be expected that, if noise is somewhat correlated between adjacent data points, error sensitivity would be enhanced.

In Section 5, two ill conditioned problems were discussed with respect to the deterioration of their associated unperturbed eigenvalue structures. In addition, problem structuring parameters were defined and their expected usefulness with respect to advantageous modification of the unperturbed eigenvalue structure discussed. Numerical examples of the above topics were present early in Section 8. Toward the end of Section 5, first order perturbation theory suggested that the eigenvalues are non decreasing functions of the number of data points utilized. This conclusion holds both with respect to the perturbed and the unperturbed Prony matrix. Furthermore, although the proof will not be documented in this report, this conclusion is rigorously true independent of the assumption of first order perturbation theory. Similarly, it can be shown that the eigenvalues are non decreasing functions of the order m of the matrix, as is numerically supported by the results displayed in tables 26 and 27. Based on the results of Sections 3, 4, and 5 and the supporting numerical studies in Section 8, one suspects that the correct rank of the unperturbed Prony matrix will be extremely difficult to determine if the minimum positive eigenvalue of the unperturbed matrix is nearly equal to the noise level, $(n-m+1)\sigma^2$. Even if this rank was known, the accuracy of the corresponding solution vector might be unacceptable. When Prony's method is being applied to experimental data, the possibility will usually exist that small eigenvalues associated with the true system response are so completely obscured by noise that, regardless of problem structuring parameters, there existence can not be detected. In such cases, the analyst might well choose a value of m less than that value which is strictly correct. Hence, the error in the recovered solution vector will be due to both random noise and neglected signal strength. It is this realization which motivated the analysis presented in Section 6. The results of this section indicate that the presence of neglected signal can potentially alter the expected value of perturbed eigenvalues and the solution vector to a much greater degree than the contribution due to random noise. The variance associated with these quantities is affected but to a lesser degree. For applications to experimental data, the potential existence of neglected signal strength is a real possibility and therefore accuracy assessment should most likely be based upon the assumption that neglected signal strength is present. Numerical studies are needed to confirm that neglected signal strength has the potential pronounced effect indicated by the results of Section 6.

Section 7 described a simple iterative technique for improving the complex frequencies and residues recovered from a conventional Prony analysis. Though results to date have been very encouraging, additional numerical experiments are needed. First, it would be interesting to perform additional studies on the behavior of the iteration technique for the case where an extraneous frequency is introduced. Second, little effort was expended during this investigation relative to optimization of the iteration methodology. Furthermore, the iteration methodology seems to permit inclusion of parameters to address uncertainties in zero time, time tie correction, DC shift, and rotation. This possibility has not yet been numerically explored. Third, the iteration scheme, unlike Prony, does not require the data points to be equally spaced and does not appear to be sensitive to the inclusion of data points dominated by noise. Thus, statistically independent random samples of data points could be studied in a Monte Carlo fashion. (Dr. Carl Baum first pointed out this possibility to the author). Also especially suspicious data points can be removed.

The numerical results supplied in Section 8 are of limited scope and are all restricted to artifically generated noisy data. Additional effort seems appropriate to remove both of these limitations.

With respect to the overall research presented, heavy reliance was made upon the validity of first order perturbation theory. The author knows that this restriction can, to a degree, be removed but has had insufficient time to explore this aspect.

Several other avenues of investigation presented themselves rather late in this effort and therefore have not been evaluated.

First, suppose that the eigenvalue structure dictates that only every fourth or fifth data point be utilized. It can easily be shown that if y_k satisfies equation (7), then the sum $y_k + y_{k-1}$ also satisfies equation (7). Thus, two data points, too closely spaced to be of use in conventional Prony analysis, could be added to form a modified data base upon which Prony's method could be utilized. This scheme can result in an effective doubling of signal strength, if y_k is comparable to y_{k-1}. However, if the noise is uncorrelated between points, the standard deviation of the noise should increase by only the factor $\sqrt{2}$. The net result is an enhanced signal to noise ratio for the modified data base. Second, suppose that as above the eigenvalue structure limits the analyst to a spacing corresponding to every four or five data points. Thus one can think of the entire data base as consisting of four or five sets spaced on the optimum Δ. Rather than using each set individually, S^2 in equation (10) can be constructed as the sum of the contributions of each of these sets. The result would require the solution to the matrix problem

$$R\vec{\alpha} = \vec{b}$$

where $\vec{R}$ and $\vec{b}$ are composed of the sums of corresponding terms for each data set considered individually. Thus all available data is simultaneously utilized in the least squares procedure but at a sample spacing that is greater than the spacing between the raw data points. Intuitively one expects that such a scheme could provide enhanced accuracy but again this possibility has not yet been adequately investigated.

APPENDIX

In Sections 4 and 6 of this document, bounds are given for the variance of the perturbed eigenvalues of the R matrix. In Section 4, equation 47, the perturbation is due solely to random noise while in Section 6, equation 78, the perturbation is due both to random noise and neglected signal strength. Similarly, Section 3 (equation 31) and Section 6 (equation 81) provide bounds for $\left\langle (\vec{X}, \vec{\Omega}_\gamma)^2 \right\rangle$ with random noise and with neglected signal strength, respectively. The purpose of this appendix is to provide the details associated with the development of these bounds. The development of the bounds on the variance of the perturbed eigenvalues will first be addressed.

Let us represent the components of the unperturbed eigenvector $\vec{\Omega}_{\gamma}^{(o)}$, associated with the unperturbed eigenvalue $\lambda_\gamma^{(o)}$ as α_i, i=1, m. The eigenvector will be assumed normalized and therefore

$$\sum_{i=1}^{m} \alpha_i^2 = 1 \tag{A.1}$$

Now, to first order one knows that the perturbed eigenvalue λ_γ is given by

$$\lambda_\gamma = \lambda_\gamma^{(o)} + \left(\vec{\Omega}_\gamma^{(o)} \, E \, \vec{\Omega}_\gamma^{(o)}\right) \tag{A.2}$$

Assuming the error matrix includes contributions due to both random noise and neglected signal, one can write

$$E_{ij} = \sum_{k=0}^{n-m} y_{m+k-i} v_{m+k-j} + y_{m+k-j} v_{m+k-i} + v_{m+k-i} v_{m+k-j} \tag{A.3}$$

where

$$v_k = x_k + \epsilon_k \tag{A.4}$$

The term x_k represents the neglected signal at point k, while ϵ_k is the contribution due to random noise at point k. The noise is considered uncorrelated between data points and is taken to have zero mean and variance σ^2. It follows that

$$\langle \epsilon_k \rangle = 0 \tag{A.5}$$

$$\langle \epsilon_k \ \epsilon_{\hat{k}} \rangle = \begin{cases} 0 & k \neq \hat{k} \\ \sigma^2 & k \neq \hat{k} \end{cases} \tag{A.6}$$

Further, as the development progresses, the expected value of the fourth moment of the random noise will be encountered. Thus it is convenient to define

$$\langle \epsilon^4 \rangle = \langle \epsilon_k \ \epsilon_{k,} \ \epsilon_k \ \epsilon_k \rangle \tag{A.7}$$

The variance of λ_γ is defined to be

$$\sigma_\gamma^2 = \langle \lambda_\gamma^{\ 2} \rangle - \langle \lambda_\gamma \rangle^2 \tag{A.8}$$

From equation (A.2), one obtains that

$$\sigma_\gamma^2 = \left\langle \left(\vec{\Omega}_\gamma^{(o)}, E\vec{\Omega}_\gamma^{(o)} \right)^2 \right\rangle - \left\langle \left(\vec{\Omega}_\gamma^{(o)}, E\vec{\Omega}_\gamma^{(o)} \right) \right\rangle^2 \tag{A.9}$$

Using equations (A.3), (A.4), (A.5) and (A.6), one recognizes that

$$\left\langle \left(\vec{\Omega}_\gamma^{(o)} \ E\vec{\Omega}_\gamma^{(o)} \right) \right\rangle = 2 \sum_{k=0}^{n-m} \sum_{i=1}^{m} y_{m+k-i} \ \alpha_i \sum_{j=1}^{m} x_{m+k-j} \alpha_j + \sum_{k=0}^{n-m} \left(\sum_{i=1}^{m} x_{m+k-i} \ \alpha_i \right)^2 + (n-m+1)\,\sigma^2 \tag{A.10}$$

and that

$$\left\langle \left(\vec{\Omega}_\gamma^{(o)}, E\vec{\Omega}_\gamma^{(o)} \right)^2 \right\rangle = \left\langle \left\{ 2 \sum_{k=0}^{n-m} \sum_{i=1}^{m} y_{m+k-i} \ \alpha_i \sum_{j=1}^{m} v_{m+k-j} \alpha_j + \left(\sum_{i=1}^{m} v_{m+k-i} \ \alpha_i \right)^2 \right\}^2 \right\rangle \tag{A.11}$$

Using equation (A.10), equation (A.11) can be rewritten as

$$\left\langle \left(\vec{\Omega}_\gamma^{(o)}, E\vec{\Omega}_\gamma^{(o)} \right)^2 \right\rangle = \left\langle \left\{ \left\langle \left(\vec{\Omega}_\gamma^{(o)}, E\vec{\Omega}_\gamma^{(o)} \right) \right\rangle - (n-m+1)\,\sigma^2 \right.$$

$$+ 2\sum_{k=0}^{n-m}\sum_{i=1}^{m}(y_{m+k-i}+x_{m+k-i})\alpha_i\sum_{j=1}^{m}\epsilon_{m+k-j}\;\alpha_j$$

$$+\sum_{k=0}^{n-m}\left(\sum_{i=1}^{m}\epsilon_{m+k-i}\,\alpha_i\right)^2\Bigg\}\Bigg\rangle \qquad (A.12)$$

Assuming the noise is symmetrically distributed about its zero mean equation (A.12) simplifies to

$$\left\langle\left(\Omega_\gamma^{(o)},\, E\,\Omega_\gamma^{(o)}\right)^2\right\rangle = \left\langle\left(\Omega_\gamma^{(o)},\, E\,\Omega_\gamma^{(o)}\right)\right\rangle^2 + 4\left\langle\psi^2\right\rangle$$

$$+\left\langle\psi_o^{\;2}\right\rangle - (n-m+1)^2\,\sigma^4 \qquad (A.13)$$

where

$$\psi \equiv \sum_{k=0}^{n-m}\sum_{i=1}^{m}(y_{m+k-i}+x_{m+k-i})\,\alpha_i\sum_{j=1}^{m}\epsilon_{m+k-j}\,\alpha_j \qquad (A.14)$$

$$\psi_o \equiv \sum_{k=0}^{n-m}\left(\sum_{i=1}^{m}\epsilon_{m+k-i}\,\alpha_i\right)^2 \qquad (A.15)$$

Combining equations (A.9) and (A.13) one concludes that

$$\sigma_\gamma^{\;2} = 4\left\langle\psi^2\right\rangle + \left\langle\psi_o^{\;2}\right\rangle - (n-m+1)^2\,\sigma^4 \qquad (A.16)$$

Using equation (A.14), one can write

$$\left\langle\psi^2\right\rangle = \sum_{i=1}^{m}\sum_{p=1}^{m}\sum_{j=1}^{m}\sum_{q=1}^{m}\alpha_i\,\alpha_p\,\alpha_j\,\alpha_q\sum_{k=-j}^{n-m-j}\;\sum_{\hat{k}=-q}^{n-m-q}$$

$$\left\langle\epsilon_{m+k}\,\epsilon_{m+\hat{k}}\right\rangle\cdot\left(y_{m+k+j-i}+x_{m+k+j-i}\right)\left(y_{m+\hat{k}+q-p}\right.$$

$$\left.+\,x_{m+\hat{k}+q-p}\right) \qquad (A.17)$$

Defining

$$k_1 = \mathrm{Max}\,(-j,\,-q)$$

$$k_2 = \mathrm{Min}\,(n-m-j,\; n-m-q)$$

Equation (A.17) reduces to

$$\left\langle\psi^2\right\rangle \le \sigma^2 \sum_{j=1}^{m} \sum_{q=1}^{m} \alpha_j \alpha_q \sum_{k=k_1}^{k_2} \sum_{i=1}^{m} (y_{m+k+j-i} + x_{m+k+j-i})\alpha_i$$

$$\cdot \sum_{p=1}^{m} (y_{m+k+q-p} + x_{m+k+q-p})\alpha_p \qquad (A.18)$$

Applying Schwarz's inequality one can conclude that

$$\left\langle\psi^2\right\rangle \le \sigma^2 \sum_{j=1}^{m} \sum_{q=1}^{m} \left|\alpha_j \alpha_q\right| \sum_{k=0}^{n-m} \left(\sum_{i=1}^{m} (y_{m+k-i} + x_{m+k-i}) \alpha_i\right)^2 \qquad (A.19)$$

Recognizing that

$$\sum_{j=1}^{m} \sum_{q=1}^{m} \left|\alpha_j \alpha_q\right| \le m \qquad (A.20)$$

and that by Schwarz's inequality

$$\sum_{k=0}^{n-m} \left(\sum_{i=1}^{m} (y_{m+k-i} + x_{m+k-i})\alpha_i\right)^2 \le \lambda_\gamma^{(o)} + 2$$

$$(\lambda_\gamma^{(o)} \psi_\gamma)^{1/2} + \psi_\gamma \qquad (A.21)$$

where

$$\psi_\gamma \equiv \sum_{k=0}^{n-m} \left(\sum_{i=1}^{m} x_{m+k-i} \alpha_i\right)^2 \qquad (A.22)$$

One concludes that

$$\left\langle\psi^2\right\rangle \le m\,\sigma^2 \left(\lambda_\gamma^{(o)} + 2\,(\lambda_\gamma^{(o)} \psi_\gamma)^{1/2} + \psi_\gamma\right) \qquad (A.23)$$

Recalling equation (A.16), one still must obtain $\left\langle\psi_o^{\,2}\right\rangle$. From equation (A.15), one can write

$$\psi_o = \sum_{k=0}^{m} \sum_{i=1}^{m} \sum_{j=1}^{m} \alpha_i \alpha_j \epsilon_{m+k-i} \epsilon_{m+k-j}$$

By symmetry the above can be written as

$$\psi_o = \sum_{k=0}^{n-m} \sum_{i=1}^{m} \alpha_i^2 \epsilon^2_{m+k-i}$$

$$+ 2 \sum_{r=1}^{m-1} \sum_{i=1}^{m-r} \alpha_i \alpha_{i+r} \epsilon_{m+k-i} \epsilon_{m+k-i-r}$$

Therefore

$$\langle \psi_o^2 \rangle = \left\langle \left(\sum_{k=0}^{n-m} \sum_{i=1}^{m} \alpha_i^2 \epsilon^2_{m+k-i} \right)^2 \right\rangle$$

$$+ 4 \sum_{r=1}^{m-1} \left\langle \left(\sum_{k=0}^{n-m} \sum_{i=1}^{m-r} \alpha_i \alpha_{i+r} \epsilon_{m+k-i} \epsilon_{m+k-i-r} \right)^2 \right\rangle$$

$$= \sum_{i=1}^{m} \sum_{p=1}^{m} \alpha_i^2 \alpha_p^2 \sum_{k=-i}^{n-m-i} \sum_{\hat{k}=-p}^{n-m-p} \langle \epsilon^2_{m+k} \epsilon^2_{m+\hat{k}} \rangle$$

$$+ 4 \sum_{r=1}^{m-1} \sum_{i=1}^{m-r} \sum_{p=1}^{m-r} \alpha_i \alpha_{i+r} \alpha_p \alpha_{p+r} \sum_{k=-i}^{n-m-i} \sum_{\hat{k}=-p}^{n-m-p} \langle \epsilon_{m+k}$$

$$\epsilon_{m+k-r} \epsilon_{m+\hat{k}} \epsilon_{m+\hat{k}-r} \rangle$$

$$= (n-m+1)^2 \sigma^4 + (n-m+1) \left(\langle \epsilon^4 \rangle - \sigma^4 \right)$$

$$+ 4 \sum_{r=1}^{m-1} \sum_{i=1}^{m-r} \sum_{p=1}^{m-r} \alpha_i \alpha_{i+r} \alpha_p \alpha_{p+r} \sigma^4 \left(n-m+1 \right.$$

$$\left. - |i-p| \right) \tag{A.24}$$

From equation (A.24) it is clear that

$$\langle \psi_o^2 \rangle \leq (n-m+1)^2 \sigma^4 + (n-m+1) \left(\langle \epsilon^4 \rangle - \sigma^4 \right)$$

$$+ 4 (n-m+1) \sigma^4 \sum_{r=1}^{m-1} \left(\sum_{i=1}^{m-r} |\alpha_i \alpha_{i+r}| \right)^2$$

Again by Schwarz's inequality and equation (A.1) the above provides the bound

$$\left\langle \psi_o^{\,2} \right\rangle \le (n-m+1)^2 \sigma^4 + (n-m+1)\left(\left\langle \epsilon^4 \right\rangle - \sigma^4\right)$$

$$+ 4\,(m-1)\,(n-m+1)\sigma^4 \tag{A.25}$$

Thus combination of equations (A.25) and (A.19) with (A.16) yields the desired result that

$$\sigma_\gamma^{\,2} \le 4\,m\,\sigma^2\left(\lambda_\gamma^{(o)} + 2\left(\lambda_\gamma^{(o)}\psi_\gamma\right)^{1/2} + \psi_\gamma\right) + (n-m+1)\,\sigma^4\left\{\frac{\left\langle \epsilon^4 \right\rangle}{\sigma^4} + 4m - 5\right\} \tag{A.26}$$

Where

$$\psi_\gamma = \sum_{k=0}^{n-m}\left(\sum_{i=1}^{m} x_{m+k-i}\,\alpha_i\right)^2 \tag{A.27}$$

Therefore, one observes that equation 47 of Section 4 results from equation (A.26) by substituting $\psi_\gamma = 0$. Equation 78 of Section 6 is identical to equation (A.26). It still remains to develop the bounds presented in the text, for $\left\langle \left(\vec{X}, \vec{\Omega}_\gamma\right)^2 \right\rangle$. The error vector $\vec{X}$ will be assumed due to random noise and neglected signal. Thus one recalls from Section 6 that

$$X_i = \sum_{k=0}^{n-m}\left(y_{m+k-i} + v_{m+k-i}\right)\sum_{j=0}^{m} \alpha_j\, v_{m+k-j} \tag{A.28}$$

Where

$$v_k = x_k + \epsilon_k \tag{A.29}$$

Thus one can write

$$(\vec{X}, \vec{\Omega}_\gamma) = T_1 + T_2 \tag{A.30}$$

Where

$$T_1 \equiv \sum_{k=0}^{n-m}\sum_{i=1}^{m} y_{m+k-i}\,\Omega_i \sum_{j=0}^{m} \alpha_j\, v_{m+k-j} \tag{A.31}$$

$$T_2 = \sum_{k=0}^{n-m} \sum_{i=1}^{m} v_{m+k-i}\, \Omega_i \sum_{j=0}^{m} \alpha_j\, v_{m+k-j} \qquad (A.32)$$

By equation (A.29) and the assumed properties of the noise one observes that

$$\langle T_1 \rangle = \sum_{k=0}^{n-m} \sum_{i=1}^{m} y_{m+k-i}\, \Omega_i \sum_{j=0}^{m} \alpha_j\, x_{m+k-j} \qquad (A.33)$$

$$\langle T_2 \rangle = \sum_{k=0}^{n-m} \sum_{i=1}^{m} x_{m+k-i}\, \Omega_i \sum_{j=0}^{m} \alpha_j\, x_{m+k-j}$$

$$+ \sum_{k=0}^{n-m} \sum_{i=1}^{m} \sum_{j=0}^{m} \Omega_i\, \alpha_j \langle \epsilon_{m+k-i}\, \epsilon_{m+k-j} \rangle \qquad (A.34)$$

The second term in equation (A.34) reduces to

$$(n-m+1)\sigma^2 \left(\vec{\Omega}_\gamma ,\ \vec{\alpha}_o\right)$$

Where $\vec{\alpha}_o$ is the vector having components α_i, i=1, m. Using the above observation, equation (A.34) becomes

$$\langle T_2 \rangle = \sum_{k=0}^{n-m} \sum_{i=1}^{m} x_{m+k-i}\, \Omega_i \sum_{j=0}^{m} \alpha_j\, x_{m+k-j} + (n-m+1)\sigma^2 \left(\vec{\Omega}_\gamma , \vec{\alpha}_o\right)$$

Now from equation (A.30) one can write

$$\left\langle \left(\vec{X}, \vec{\Omega}_\gamma\right)^2 \right\rangle = \langle T_1^{\,2} \rangle + 2 \langle T_1\, T_2 \rangle + \langle T_2^{\,2} \rangle \qquad (A.36)$$

From equation (A.31) one computes that

$$\langle T_1^{\,2} \rangle = \langle T_1 \rangle^2 + \sum_{i=1}^{m} \sum_{p=1}^{m} \sum_{j=0}^{m} \sum_{q=0}^{m} \sum_{k=0}^{n-m} \sum_{\hat{k}=0}^{n-m}$$

$$\Omega_i\, \Omega_p\, \alpha_j\, \alpha_q\, y_{m+k-i}\, y_{m+\hat{k}-p} \cdot \langle \epsilon_{m+k-j}\, \epsilon_{m+\hat{k}-q} \rangle$$

As was done in the development of equation (A.18), one defines

$$k_1 = \text{Max}\,(-j,\ -q)$$

$$k_2 = \text{Min}\,(n-m-j,\ n-m-q)$$

and writes

$$\left\langle T_1^{\,2}\right\rangle = \left\langle T_1\right\rangle^2 + \sigma^2 \sum_{j=0}^{m} \sum_{q=0}^{m} \alpha_j \alpha_q \sum_{k=k_1}^{k_2} \sum_{i=1}^{m} y_{m+k+j-i}\, \Omega_i$$

$$\cdot \sum_{p=1}^{m} y_{m+k+q-p}\, \Omega_p$$

Applying Schwarz's inequality and recalling the definition of λ_γ one obtains that

$$\left\langle T_1^{\,2}\right\rangle \le \left\langle T_1\right\rangle^2 + \sigma^2 \lambda_\gamma \sum_{j=0}^{m} \sum_{q=0}^{m} \left| \alpha_j \alpha_q \right|$$

Another application of Schwarz's inequality establishes that

$$\left\langle T_1^{\,2}\right\rangle \le \left\langle T_1\right\rangle^2 + (m+1)\sigma^2 \lambda_\gamma \left(1 + (\vec{\alpha}_o, \vec{\alpha}_o)\right) \tag{A.37}$$

where the fact that $\alpha_o \equiv 1$ has been utilized.

Recalling equations (A.31) and (A.32) one notes that

$$\left\langle T_1 T_2\right\rangle = \left\langle T_2\right\rangle \left\langle T_1\right\rangle$$

$$+ \left\langle \left\{ \sum_{k=0}^{n-m} \sum_{i=1}^{m} y_{m+k-i} \Omega_i \sum_{j=0}^{m} \alpha_j \epsilon_{m+k-j} \right.\right.$$

$$\cdot \sum_{k=0}^{n-m} \sum_{i=1}^{m} x_{m+k-i}\, \Omega_i \sum_{j=0}^{m} \alpha_j \epsilon_{m+k-j}$$

$$\left.\left. + \sum_{i=1}^{m} \epsilon_{m+k-i} \Omega_i \sum_{j=0}^{m} \alpha_j x_{m+k-j} \right\} \right\rangle$$

Applying similar techniques to those used in developing equation (A.37) one obtains that

$$\left\langle T_1 T_2\right\rangle \le \left\langle T_1\right\rangle \left\langle T_2\right\rangle + (m+1)\left(1 + (\vec{\alpha}_o, \vec{\alpha}_o)\right)\left(\lambda_\gamma \sum_{k=0}^{n-m}\right.$$

$$\left.\left(\sum_{i=1}^{m} x_{m+k-i}\, \Omega_i\right)^2\right)^{1/2}$$

$$+ \sigma^2 \left[m\,(m+1)\left(1 + (\vec{\alpha}_o, \vec{\alpha}_o)\right) \lambda_\gamma \sum_{k=0}^{n-m} \left(\sum_{i=0}^{m} \alpha_i\, x_{m+k-i}\right)^2 \right]^{1/2}$$

Further use of Schwarz's inequality along with the observation that

$$\sum_{i=1}^{m} x^2_{m+k-i} \leq \sum_{i=0}^{m} x^2_{m+k-i}$$

enables one to write

$$\langle T_1 T_2 \rangle \leq \langle T_1 \rangle \langle T_2 \rangle + \left(m+1 + \sqrt{m(m+1)} \right) \left(1 + (\vec{\alpha}_o, \vec{\alpha}_o) \right)$$

$$\sigma^2 \left(\lambda_\gamma \sum_{k=0}^{n-m} \sum_{i=0}^{m} x^2_{m+k-i} \right)^{1/2} \tag{A.38}$$

Returning to equation (A.32) one can write that

$$T_2 = \sum_{k=0}^{n-m} \sum_{i=1}^{m} x_{m+k-i} \Omega_i \sum_{j=0}^{m} x_{m+k-j} \alpha_j$$

$$+ \sum_{i=1}^{m} x_{m+k-i} \Omega_i \sum_{j=0}^{m} \epsilon_{m+k-j} \alpha_j$$

$$+ \sum_{i=1}^{m} \epsilon_{m+k-i} \Omega_i \sum_{j=0}^{m} x_{m+k-j} \alpha_j$$

$$+ \sum_{i=1}^{m} \epsilon_{m+k-i} \Omega_i \sum_{j=0}^{m} \epsilon_{m+k-j} \alpha_j$$

Therefore

$$\langle T_2^2 \rangle = \left(\sum_{k=0}^{n-m} \sum_{i=1}^{m} x_{m+k-i} \Omega_i \sum_{j=0}^{m} x_{m+k-j} \alpha_j \right)^2$$

$$+ 2(n-m+1)\sigma^2 (\vec{\Omega}_\gamma, \vec{\alpha}_o) \sum_{k=0}^{n-m} \sum_{i=1}^{m} x_{m+k-i} \Omega_i \sum_{j=0}^{m} x_{m+k-j} \alpha_j$$

$$+ \left\langle \left(\sum_{k=0}^{n-m} \sum_{i=1}^{m} x_{m+k-i} \Omega_i \sum_{j=0}^{m} \epsilon_{m+k-j} \alpha_j \right)^2 \right\rangle$$

$$+ \left\langle \left(\sum_{k=0}^{n-m} \sum_{i=1}^{m} \epsilon_{m+k-i} \Omega_i \sum_{j=0}^{m} x_{m+k-j} \alpha_j \right)^2 \right\rangle$$

$$+\left\langle\left(\sum_{k=0}^{n-m}\sum_{i=1}^{m}\epsilon_{m+k-i}\,\Omega_i\sum_{j=0}^{m}\epsilon_{m+k-j}\,\alpha_j\right)^2\right\rangle$$

$$+2\left\langle\left(\sum_{k=0}^{n-m}\sum_{i=1}^{m}x_{m+k-i}\,\Omega_i\sum_{j=0}^{m}\epsilon_{m+k-j}\,\alpha_j\right)\right.$$

$$\left.\left(\sum_{k=0}^{n-m}\sum_{i=1}^{m}\epsilon_{m+k-i}\,\Omega_i\sum_{j=0}^{m}x_{m+k-j}\,\alpha_j\right)\right\rangle \qquad (A.39)$$

The majority of the terms above can be bounded by recognizing their similarity to terms encountered in the development of equation (A.38). Thus one can immediately write utilizing equations (A.39) and (A.35) that

$$\left\langle T_2^{\,2}\right\rangle \leq \left\langle T_2\right\rangle - (n-m+1)^2\,\sigma^4\left(\vec{\Omega}_\gamma,\vec{\alpha}_o\right)^2$$

$$+(m+1)\left(1+(\vec{\alpha}_o,\vec{\alpha}_o)\right)\sigma^2\sum_{k=0}^{n-m}\sum_{i=1}^{m}x^2_{m+k-i}$$

$$+m\left(1+(\vec{\alpha}_o,\vec{\alpha}_o)\right)\sigma^2\sum_{k=0}^{n-m}\sum_{i=0}^{m}x^2_{m+k-i}$$

$$+2\sqrt{m\,(m+1)}\left(1+(\vec{\alpha}_o,\vec{\alpha}_o)\right)\sigma^2\left\{\sum_{k=0}^{n-m}\sum_{i=1}^{m}x^2_{m+k-i}\right.$$

$$\left.\sum_{k=0}^{n-m}\sum_{i=0}^{m}x^2_{m+k-i}\right\}^{1/2}+\left\langle\left(\sum_{k=0}^{n-m}\sum_{i=1}^{m}\epsilon_{m+k-i}\,\Omega_i\right.\right.$$

$$\left.\left.\sum_{j=0}^{m}\epsilon_{m+k-j}\,\alpha_j\right)^2\right\rangle$$

Again using the fact that

$$\sum_{i=1}^{m}x^2_{m+k-i}\leq\sum_{i=0}^{m}x^2_{m+k-i}$$

one can write that

$$\left\langle T_2^{\,2}\right\rangle \leq \left\langle T_2\right\rangle^2 - (n-m+1)^2\sigma^4\left(\vec{\Omega}_\gamma,\vec{\alpha}_o\right)^2$$

$$+\left(m+2\sqrt{m(m+1)}+m+1\right)\left(1+(\vec{\alpha}_o,\vec{\alpha}_o)\right)\sigma^2\sum_{k=0}^{n-m}\sum_{i=0}^{m}x^2_{m+k-i}$$

$$+\left\langle\left(\sum_{k=0}^{n-m}\sum_{i=1}^{m}\epsilon_{m+k-i}\Omega_i\sum_{j=0}^{m}\epsilon_{m+k-j}\alpha_j\right)^2\right\rangle \qquad (A.40)$$

Combining the results of equations (A.37), (A.38), and (A.40) with equation (A.36) implies that

$$\left\langle(\vec{X},\vec{\Omega}_\gamma)^2\right\rangle \le \left\langle T_1\right\rangle^2 + 2\left\langle T_1\right\rangle\left\langle T_2\right\rangle + \left\langle T_2\right\rangle^2$$

$$+\left(1+(\vec{\alpha}_o,\vec{\alpha}_o)\right)\sigma^2\left[(m+1)\lambda_\gamma + 2\left(m+1+\sqrt{m(m+1)}\right)\right.$$

$$\left(\lambda_\gamma\sum_{k=0}^{n-m}\sum_{i=0}^{m}x^2_{m+k-i}\right)^{1/2} + \left(m+2\sqrt{m(m+1)}+m+1\right)\sum_{k=0}^{n-m}$$

$$\left.\sum_{i=0}^{m}x^2_{m+k-i}\right] + T_o - (n-m-1)^2\sigma^4(\vec{\Omega}_\gamma,\vec{\alpha}_o)^2 \qquad (A.41)$$

Where

$$T_o \equiv \left\langle\left(\sum_{k=0}^{n-m}\sum_{i=1}^{m}\epsilon_{m+k-i}\Omega_i\sum_{j=0}^{m}\epsilon_{m+k-j}\alpha_j\right)^2\right\rangle \qquad (A.42)$$

Recognizing that

$$\left\langle(\vec{X},\vec{\Omega}_\gamma)\right\rangle = \left\langle T_1\right\rangle + \left\langle T_2\right\rangle$$

and performing a little algebra reduces equation (A.41) to the form

$$\left\langle(\vec{X},\vec{\Omega}_\gamma)^2\right\rangle \le \left\langle(\vec{X},\vec{\Omega}_\gamma)\right\rangle^2 + \left(1+(\vec{\alpha}_o,\vec{\alpha}_o)\right)\sigma^2\left[(m+1)\lambda_\gamma\right.$$

$$+4(m+1)\sum_{k=0}^{n-m}\sum_{i=0}^{m}x^2_{m+k-i}$$

$$\left.+4(m+1)\left(\lambda_\gamma\sum_{k=0}^{n-m}\sum_{i=0}^{m}x^2_{m+k-i}\right)^{1/2}\right]$$

$$+ T_o - (n-m+1)^2 \sigma^4 (\vec{\Omega}_\gamma, \vec{\alpha}_o)^2 \qquad (A.43)$$

From equation (A.42), one can write

$$T_o = \left\langle \left(\sum_{i=1}^{m} \sum_{j=0}^{m} \Omega_i \alpha_j \sum_{k=-r}^{n-m-r} \epsilon_{m+k} \epsilon_{m+k-|i-j|} \right)^2 \right\rangle$$

Where $\qquad r = \text{Min}(i, j)$

Therefore

$$T_o = \sum_{i=1}^{m} \sum_{j=0}^{m} \sum_{p=1}^{m} \sum_{q=0}^{m} \Omega_i \alpha_j \Omega_p \alpha_q \sum_{k=-r}^{n-m-r} \sum_{\hat{k}=-s}^{n-m-s} \langle \epsilon_{m+k} \epsilon_{m+k-|i-j|}.$$

$$\epsilon_{m+\hat{k}} \epsilon_{m+\hat{k}-|p-q|} \rangle \qquad (A.44)$$

Where $s = \text{Min}(p, q)$

By the properties of the noise, the only contributions to T_o occur when

$$|i-j| = |p-q|$$

Thus equation (A.44) can be written as

$$T_o = \sum_{i=1}^{m} \sum_{p=1}^{m} \Omega_i \alpha_i \Omega_p \alpha_p \sum_{k=-i}^{n-m-i} \sum_{\hat{k}=-p}^{n-m-p} \langle \epsilon^2_{m+k} \epsilon^2_{m+\hat{k}} \rangle$$

$$+ \sum_{i=1}^{m} \sum_{\substack{j=0 \\ j \neq i}}^{m} \Omega_i \alpha_j \sum_{p=1}^{m} \sum_{\substack{q=0 \\ q \neq p}}^{m} \Omega_p \alpha_q \sum_{k=-r}^{n-m-r} \sum_{\hat{k}=-s}^{n-m-s} \langle \epsilon_{m+k} \epsilon_{m+k-|i-j|}$$

$$\epsilon_{m+\hat{k}} \epsilon_{m+\hat{k}-|p-q|} \rangle$$

$$= (n-m+1)\sigma^4 (\vec{\Omega}_\gamma, \vec{\alpha}_o)^2 + (n-m+1)$$

$$(\langle \epsilon^4 \rangle - \sigma^4) (\vec{\Omega}_\gamma, \vec{\alpha}_o)^2 + \sum_{i=1}^{m} \sum_{\substack{j=0 \\ j \neq i}}^{m} \Omega_i \alpha_j \sum_{p=1}^{m} \sum_{\substack{q=0 \\ q \neq p}}^{m} \Omega_p \alpha_q$$

$$\sum_{k=-r}^{n-m-r} \sum_{\hat{k}=-s}^{n-m-s} \langle \epsilon_{m+k} \epsilon_{m+k-|i-j|} \epsilon_{m+\hat{k}} \epsilon_{m+\hat{k}-|p-q|} \rangle$$

To bound the one remaining summation term above, one makes the following arguments. Since $|i - j| \neq 0$, the noise term is either zero or σ^4. For i, j, and p fixed there are at most two values of q which will provide non zero terms ($q_1 = p + |i - j|$ and $q_2 = p - |i - j|$

Therefore

$$T_o \leq (n-m+1)^2 \sigma^4 (\vec{\Omega}_\gamma, \vec{\alpha}_o)^2 + (n-m+1)\left(\langle \epsilon^4 \rangle - \sigma^4\right)(\vec{\Omega}_\gamma \vec{\alpha}_o)^2$$

$$+ (n-m+1)\sigma^4 \sum_{i=1}^{m} \sum_{\substack{j=0 \\ j\neq i}}^{m} \sum_{p=1}^{m} \left|\Omega_i \, \Omega_p \, \alpha_j\right| \left\{\left|\alpha_{p+|i-j|}\right| + \left|\alpha_{p-|i-j|}\right|\right\}$$

Again by Schwarz's inequality one concludes that

$$T_o \leq (n-m+1)^2 \sigma^2 (\Omega_\gamma, \alpha_o)^2 + (n-m+1)\left(\langle \epsilon^4 \rangle - \sigma^4\right)\left(\vec{\Omega}_\gamma, \vec{\alpha}_o\right)^2$$

$$+ 2\, m\, (n-m+1)\sigma^4 \left(1 + (\vec{\alpha}_o, \vec{\alpha}_o)\right) \tag{A.45}$$

Combining equation (A.45) with (A.43) and rearranging one obtains the desired result that

$$\left\langle \left(\vec{X}, \vec{\Omega}_\gamma\right)^2 \right\rangle \leq \left\langle \left(\vec{X}, \vec{\Omega}_\gamma\right) \right\rangle^2 + (n-m+1)\left(\langle \epsilon^4 \rangle - \sigma^4\right)\left(\vec{\Omega}_\gamma, \vec{\alpha}_o\right)^2$$

$$+ \sigma^2 \left(1 + (\vec{\alpha}_o, \vec{\alpha}_o)\right)\left(2\, m\, (n-m+1)\sigma^2 + (m+1)\lambda_\gamma\right)$$

$$+ 4\, (m+1)\sigma^2 \left(1 + (\vec{\alpha}_o, \vec{\alpha}_o)\right)\left[\sum_{k=0}^{n-m} \sum_{i=0}^{m} x^2_{m+k-i} + \left(\lambda_\gamma \sum_{k=0}^{n-m} \sum_{i=0}^{m} x^2_{m+k-i}\right)^{1/2}\right] \tag{A.46}$$

Equation 31 of Section 3 is easily recognized as equation (A.46) with

$$\sum_{k=0}^{n-m} \sum_{i=0}^{m} x^2_{m+k-i} = 0.$$

Equation 81 of Section 6 is equivalent to equation (A.46).