EMP 3-32

ADD

Interaction Notes

Note 301

December 1975

Techniques for Extracting the Complex Resonances of a System Directly from its Transient Response

M. L. Van Blaricum
R. Mittra

University of Illinois
Urbana, Illinois

ABSTRACT

The singularity expansion method is a technique used for writing the transient response of a structure as a sum of exponentially damped sinusoids. In order to apply this method, it is first necessary to obtain the singularities of the system being studied. The conventional approach for determining the singularities of a system is based on an iterative search procedure that seeks the zeros of the system determinant in the complex frequency plane. The alternative approach of extracting the system singularities directly from the transient response function is discussed. The method developed is based on Prony's algorithm. Some basic problems which are associated with the use of Prony's algorithm are discussed and solutions are obtained.

It is demonstrated that Prony's method is applicable to systems with multiple as well as simple pole singularities. Two techniques are presented for systematically determining the number of poles contained in a transient response.

Several numerical examples are presented to demonstrate these capabilities. The dependence of Prony's method on the noise level in the transient response is of great importance if the method is to be applied to experimental systems. This question is considered in detail and the results of several statistical tests are presented in order to get a handle on the problem.

Some of the applications of Prony's method which are discussed are: systems analysis, radar target recognition; the study of spectral characteristics, and data reduction and extrapolation. Examples are presented that demonstrate the advantages that this technique has over the conventional approaches to the above applications. The Padé approximation is suggested as an alternative to Prony's method but is demonstrated to be unsatisfactory.

ACKNOWLEDGMENT

This project was supported in part by the Air Force Office of Scientific Research Grant AFOSR 75-2861.

TABLE OF CONTENTS

		Page
1.	INTRODUCTION.	8
2.	MATHEMATICAL FORM OF THE TRANSIENT RESPONSE	11
	2.1 The Impulse Response	11
	2.2 Response Due to an Arbitrary Excitation.	14
3.	THE NUMERICAL METHOD FOR EXTRACTING POLES	19
	3.1 Prony's Method for Simple Poles.	20
	3.1.1 The classical derivation.	20
	3.1.2 Relationship to the Z-transform	27
	3.1.3 Relationship to Corrington's difference equation. .	29
	3.2 Prony's Method for Multiple Poles.	31
	3.3 Numerical Examples	39
	3.4 Guidelines	68
	3.4.1 Sampling interval	68
	3.4.2 Sampling rate	68
	3.4.3 Removal of the influence of the driving function. .	68
	3.4.4 Least squares versus conventional matrix inversion.	69
	3.5 Problems Associated with Prony's Method.	69
4.	THE DETERMINATION OF THE NUMBER OF POLES (N).	71
	4.1 Householdèr Orthogonalization Method	71
	4.2 The Eigenvalue Method for Determining N.	83
5.	NOISE AND ITS RELATIONSHIP TO PRONY'S METHOD.	92
	5.1 Least Mean Squared Error Approach.	92
	5.2 Determination of the Poles from Noisy Data	96
	5.3 Numerical Examples	97
6.	APPLICATIONS.	115
	6.1 System Analysis.	115
	6.1.1 Response to various exciting waveforms.	115
	6.1.2 Compatibility with circuit theory	117
	6.1.3 Study of system parameters.	118
	6.2 Radar Target Recognition	123
	6.3 Study of Spectral Characteristics.	128
	6.4 Data Reduction and Extrapolation	132
	6.5 Other Applications of Prony's Method	134

Page

7. AN ALTERNATIVE TO PRONY'S METHOD. 135

7.1 Padé Approximation Method. 135
7.2 Examples of the Padé Procedure 140
7.3 Conclusions Regarding the Padé Approximation 142

8. CONCLUSIONS 144

REFERENCES. 146

ADDITIONAL REFERENCES ON PRONY'S METHOD 149

LIST OF FIGURES

Figure		Page
3.1	Response containing a double pole at the origin.	37
3.2	Source current on 1.0 m dipole	41
3.3	(a) Data window used in Example 1. (b) Locations of extracted poles in Example 1	42
3.4	Comparison of extracted poles of Example 1 with those of Tesche.	43
3.5	Reconstructed transient response using extracted poles of Example 1	45
3.6	Position and amplitude of poles in the second quadrant of the s-plane for Example 1.	46
3.7	(a) Data window used in Example 2. (b) Locations of extracted poles in Example 2	48
3.8	Reconstructed transient response using extracted poles of Example 2	49
3.9	(a) Data window used in Example 3. (b) Locations of extracted poles in Example 3	50
3.10	Original response minus the reproduced response of Example 3.	52
3.11	(a) Data window used when the time step is $4\Delta t$. (b) Locations of the resulting poles.	53
3.12	(a) Data window used when the time step is $5\Delta t$. (b) Locations of the resulting poles.	54
3.13	(a) Data window starting at time step sixty-one, Example 5. (b) Resulting pole locations.	56
3.14	(a) Deconvolved transient response data of Example 6. (b) Locations of resulting poles.	58
3.15	Position and amplitude of poles in the second quadrant of the s-plane for Example 6	59
3.16	(a) Data window for Example 7 . (b) Locations of resulting poles.	61
3.17	Original response minus the reconstructed response of Example 7.	63
3.18	(a) Data window for Example 8 starting at time step sixty-one . (b) Locations of resulting poles.	64

Figure		Page
3.19	Original experimental data of Example 9.	66
3.20	Pole locations in complex frequency plane for 1.0 m experimental whip of Example 9	67
4.1	Transient response resulting from the six pole pairs of Table 4.1.	69
4.2	Value of the average of the absolute value of the vectors a and i.	80
4.3	Value of the average of the absolute value of the vector a and i for the response of the 1.0 m dipole	82
4.4	Resulting eigenvalues from response generated using poles of Table 4.2 for non-noise case and noise with $\sigma = 0.005$ . . .	87
4.5	Comparison of the resulting twenty-three and twenty-five eigenvalues for 1.0 m dipole response.	89
4.6	Pole locations for the extraction of twenty, twenty-two, and twenty-four poles from the 1.0 m dipole response	91
5.1	Transient waveform used in Example 2	102
5.2	Transient waveform used in Example 4	111
6.1	Induced current at center element of 1 m dipole with resistive loadings	120
6.2	Backscattered fields for 1 m dipole with resistive loading . .	121
6.3	Trajectory of the poles for the uniformly resistive loaded 1.0 m dipole	122
6.4	Schematic representation of the radar target recognition system	125
6.5	(a) Imaginary part of input admittance for 1.0 m dipole . (b) Real part of input admittance for 1.0 m dipole.	130
6.6	Extrapolated current response for 1.0 m monopole of Example 9, Chapter 3.	133
7.1	Model of a transient waveform using step and ramp functions. .	138

LIST OF TABLES

TABLE		Page
3.1	EXTRACTED POLES AND RESIDUES FOR THE DOUBLE POLE CASE.	38
4.1	SIX POLE PAIRS AND THE RECOVERED POLES USING HOUSEHOLDER'S ORTHOGONALIZATION METHOD	78
4.2	SIX POLE PAIRS AND THE RECOVERED POLES USING THE EIGENVALUE METHOD.	85
5.1	RESULTS FOR EXAMPLE 1: $R(t) = SIN(\Pi t)$, $\gamma = 150$, $\Delta t = 0.1$ S	100
5.2	RESULTS FOR EXAMPLE 2: $R(t) = e^{-2t} SIN(\Pi t)$, $\gamma = 50$, $\Delta t = 0.1$ S	103
5.3	AMPLITUDE OF EACH POLE COMPONENT FOR EXAMPLES 3 AND 4.	105
5.4	VALUES OF EXTRACTED POLES AND THEIR PERCENT ERROR AS A FUNCTION OF THE SIGNAL TO NOISE RATIO FOR EXAMPLE 3.	106
5.5	SIGNAL TO NOISE RATIO OF THE DIFFERENT COMPONENTS OF THE SIGNAL IN EXAMPLE 3 AS A FUNCTION OF THE STANDARD DEVIATION OF THE NOISE	107
5.6	THE THEORETICAL AND EXPERIMENTAL VALUES OF THE N + 1 EIGENVALUE FOR EXAMPLE 3 AND EXAMPLE 4	108
5.7	VALUES OF EXTRACTED POLES WITH THEIR PERCENT ERROR AND STANDARD DEVIATION AS A FUNCTION OF SIGNAL TO NOISE RATIO FOR EXAMPLE 4.	112
5.8	SIGNAL TO NOISE RATIO OF THE DIFFERENT COMPONENTS OF THE SIGNAL IN EXAMPLE 4 AS A FUNCTION OF THE STANDARD DEVIATION OF THE NOISE	113
7.1	THE RESULTING POLES USING THE PADÉ APPROXIMATION ON THE TRANSIENT RESPONSE OF EXAMPLE 1: $R(r) = e^{-2t} \sin \pi t$	141
7.2	THE RESULTING POLES USING THE PADÉ APPROXIMATION ON THE TRANSIENT RESPONSE OF EXAMPLE 2: $R(t) = 2^{-2t} \sin \pi t + e^{-3t} \sin 1.5 \pi t$.	141
7.3	THE RESULTING POLES USING THE PADÉ APPROXIMATION ON THE TRANSIENT RESPONSE OF EXAMPLE 3: $R(t) = e^{-2t} \sin \pi t + e^{-3t} \sin 1.5 \pi t + e^{-4t} \sin 2 \pi t$.	143

1. INTRODUCTION

In circuit theory, the impulse response that characterizes a linear circuit may be determined from the knowledge of the singularities of the response function in the complex frequency plane and the corresponding residues. The impulse response of the circuit is then simply a summation of all the residues multiplied by exponentially damped sinusoids. In recent years a similar approach has been used to characterize the transient phenomenon of electromagnetic scattering and antenna problems. This approach is known as the singularity expansion method (SEM) [1], [2].

The development of the singularity expansion method arose from insight into the general characteristics of typical transient response behavior observed on various electromagnetic structures. The transient response waveforms of these structures appear to be dominated by a few exponentially damped sinusoids. This observation is especially apparent if one looks at the late time response of slender or thin wire structures. The resonant frequency, damping constant, and current distribution of some of these resonant modes had been calculated for the thin wire and prolate spheroid as early as 1930 [3], [4]. However, not until the singularity expansion method came into being was it possible to determine the modal resonances and the excitation coefficient of each mode for a structure with an arbitrary incident wave.

In the singularity expansion method the transient response is written as a sum of exponentially damped sinusoids. In order to write this sum, it is necessary to first determine the location of the complex natural frequencies or pole singularities of the structure being studied. The conventional approach for determining the singularities of a system is based on an iterative search procedure that seeks the zeros of the system

determinant in the complex frequency plane. This approach has been used successfully by many people and has given extremely satisfying results. A few illustrations are given in References [5] - [7]. Since the iterative search is a slow procedure, it is usually economical to extract only a few poles. Also in the conventional techniques, the poles cannot be extracted from the time-domain formulation of the problem.

Recent development of methods for the direct production of both numerical [8], [9] and experimental [10], [11] transient electromagnetic response data has generated considerable interest in the possibility of direct extraction of the poles and their accompanying residues from given time-domain system response. This thesis presents a numerical technique for systematically determining the system singularities from the transient response data of that system. The approach that will be developed here is based on Prony's algorithm which was first published in 1795 [12] and has appeared in a few good numerical analysis texts [13], [14].

In Chapter 2 the mathematical notation is developed for the impulse response function for electromagnetic antennas and scatterers. The assumptions used in reducing the impulse response to a sum of exponentials are also presented. The necessity of removing the influence of the driving function from the transient response if the driving function is not a true delta function is discussed in detail.

Prony's method is developed in detail in Chapter 3. The method is applicable to systems containing multiple as well as simple poles. Several numerical examples are presented and the results are analyzed. These results are used to establish guidelines for the use of the method and to bring out some of the problems associated with the method.

A systematic procedure by which the number of poles inherent in the transient response can be determined is necessary in order to properly use Prony's method. Chapter 4 presents two techniques for doing this. The first technique is based on an orthogonalization procedure and the second technique is based on an eigenvalue approach. Both methods have advantages and disadvantages inherent in their implementation. These advantages and disadvantages are presented by showing several numerical examples.

It is shown in Chapter 5 that noise seriously affects the poles returned by Prony's method. In some cases the noise level is high enough to completely corrupt the results. Several statistical studies are presented which relate the standard deviation of the noise to the quality of the results obtained from Prony's method.

Chapter 6 discusses several of the applications in which the developed methods can be implemented. Some of these topics are: system analysis, radar target recognition, the study of spectral characteristics, and data reduction and extrapolation. Examples are used to show that Prony's method applied to these problems has definite numerical advantages over the conventional approaches used.

An alternative to Prony's method is presented in Chapter 7. This approach is the Padé approximation but it is extremely limited in its usefulness. Chapter 8 presents conclusions and recommendations for further study.

2. MATHEMATICAL FORM OF THE TRANSIENT RESPONSE

This chapter introduces the mathematical form of the impulse response function for an electromagnetic scatterer or antenna. The impulse response will be written in both the s-plane (Laplace transform) domain and in the time domain. The assumptions necessary for reducing the impulse response to a simple sum of exponentials are discussed. It is necessary to reduce the response to a sum of exponentials so that the methods of Chapter 3 may be applied. The form of the response functions for an arbitrary excitation is also studied.

2.1 The Impulse Response

The normalized s-plane impulse or delta function response for a finite-sized object that has only pole singularities in free space is generally written [1], [2]

$$\bar{H}(\bar{r},s) = \sum_{i=1}^{\infty} \eta_i(s_i,\bar{p})\ \nu_i(\bar{r})\ (s - s_i)^{-m_i} + \bar{W}(\bar{r},s,\bar{p}) \qquad (2.1)$$

where the above terms are defined as:

s_i - natural frequency, pole singularity, natural resonances. This is a complex frequency for which the system has a response when no forcing function is applied. The poles must appear in complex conjugate pairs or lie on the real axis. The poles also must lie in the negative half of the s-plane.

$\nu_i(\bar{r})$ - natural mode. This is the response of the system at s_i which depends on the position $\bar{r}$ on the structure and the object parameters only.

$\eta_i(s_i,\bar{p})$ - coupling coefficient.

This is the strength of the natural oscillation s_i in terms of the system and the incident wave parameters. It is independent of position.

$\bar{p}$ - polarization of incident plane wave.

$\bar{r}$ - the position vector.

This is the position on the structure at which the transient response is being measured or observed.

m_i - the multiplicity of the i^{th} pole.

The term $\bar{W}(\bar{r},s,\bar{p})$ is an entire function of s and dependent on the form of the coupling coefficient and the incident wave. In the most general case this term is required by the Mittag-Leffler Theorem [15] in order to guarantee convergence of the infinite series. It has been hypothesized that the entire function is not needed in most electromagnetics problems and could normally be neglected.

The impulse response (2.1) can be written in the time domain as

$$\bar{H}(\bar{r},t) = u(t - t_0) \sum_{i=1}^{\infty} \eta_i(s_i,\bar{p})\, \nu_i(\bar{r})\, e^{s_i t} \tag{2.2}$$

where the entire function has been neglected and all poles have been assumed to be simple. The step function $u(t - t_0)$ is present so that the response does not start until t_0, the time at which the response begins at the particular observation position $\bar{r}$ on the body. Since the entire function has been neglected, it is necessary to require that t_0 be greater than zero because a delta function source applied at $\bar{r}$ at $t = 0$ yields a delta function at $t = 0$ in the impulse response which cannot be represented by the exponential terms.

As an example, consider a perfectly conducting finite dipole driven at t = 0 by a delta function plane wave with the $\bar{E}$ field polarized in the direction of the dipole axis. The impulse response of the induced current at any position $\bar{r}$ on the dipole is of the form

$$I(\bar{r},t) = \left[\sum_{i=1}^{\infty} A_i(\bar{r})\, e^{s_i t} + B_i(\bar{r})\, \delta(t) \right] u(t) \quad . \tag{2.3}$$

If the induced current is expressed as a sum of complex exponentials it is written as

$$I(\bar{r},t) = u(t - t_0) \left[\sum_{i=1}^{\infty} A_i(\bar{r})\, e^{s_i t} \right] \tag{2.4}$$

where t_0 is set equal to 0^+, the time at which the delta function turns off. Note here that for convenience the coupling coefficient and the natural mode have been combined into one term $A_i(\bar{r})$. The s-plane version of Expression (2.3) is written

$$I(s,\bar{r}) = \sum_{i=1}^{\infty} \frac{A_i(\bar{r})}{s - s_i} + B_i(\bar{r}) \tag{2.5}$$

where $B_i(\bar{r})$ is the inverse Laplace transform of $B_i(\bar{r})\, \delta(t)$, which is a constant in s and a function of position $\bar{r}$.

In Equations (2.1) - (2.5) the series contains an infinite number of terms. In general only the first few terms of this series are needed to adequately represent the late time response of the system [1], [5]. The early time response on the other hand requires a larger number of poles for reasonable convergence to the true response. This is intuitively reasonable if one realizes that, in general, as the frequency or the imaginary part of the pole increases so does the damping constant

or real part. Thus, for early times both high and low frequency components are present and as time progresses the higher frequency components damp out and disappear until only the first few dominant resonant modes remain. Moreover, all transient response data that can be generated either experimentally or numerically are necessarily band limited and thus contain only a finite number of poles. For these reasons the series is truncated after N terms, where N is the number of resonant frequencies contained in the transient response being studied.

2.2 Response Due to an Arbitrary Excitation

Since a true impulse excitation function cannot be realized, it is of interest to determine the form of the transient response due to an arbitrary exciting waveform. A general response function $R(s,\bar{r})$ is given in the Laplace transform domain as

$$R(s,\bar{r}) = F(s,\bar{r})\, H(s,\bar{r}) \tag{2.6}$$

where $H(s,\bar{r})$ and $F(s,\bar{r})$ are the Laplace transforms of the system's impulse response and the arbitrary driving function, respectively. The system's impulse response $H(s,\bar{r})$ was written in (2.1) and, thus, $R(s,\bar{r})$ can be expressed as

$$\bar{R}(s,\bar{r}) = F(s,\bar{r})\left\{\sum_{i=1}^{N} \frac{A_i(\bar{r})}{s - s_i}\right\} \tag{2.7}$$

where the possible entire function has been neglected and the coupling coefficient and natural mode have been combined into one term, $A_i(\bar{r})$. For convenience $A_i(\bar{r})$ can be regarded as the residue for the i^{th} complex

pole. Note that the residue is a function of position on the object. If the inverse Laplace transform of (2.7) is performed, then the transient response function $R(t,\bar{r})$ is written in general form as

$$R(t,\bar{r}) = \sum_{i=1}^{N} B_i(\bar{r})\, e^{s_i t} + g(t,\bar{r}) \tag{2.8}$$

where

$$B_i(\bar{r}) = F(s_i,\bar{r})\, A_i(\bar{r})$$

and the term $g(t,\bar{r})$ is dependent on the form of the driving function.

Equation (2.8) shows that the transient response due to an arbitrary exciting waveform can, in general, be written as a sum of complex exponentials plus some added term $g(t,\bar{r})$. The desire here is to be able to express the transient response as a sum of exponentials only. Thus, the character of the term $g(t,\bar{r})$ needs to be studied to determine if it can be either removed or expressed also as a sum of exponentials.

If the exciting waveform itself has pole singularities, as in the case of a step function or a sinusoidal function, then $g(t,\bar{r})$ can be expressed as a sum of exponentials. As an example, consider a driving function which is a simple step function u(t), then, the Laplace transform of the driving function is

$$F(s) = 1/s \quad .$$

If the impulse response of the system is

$$H(s) = \sum_{i=1}^{N} \frac{A_i}{s - s_i}$$

then the response R(s) is simply

$$R(s) = F(s)\, H(s) = \frac{1}{s} \sum_{i=1}^{N} \frac{A_i}{s - s_i} \quad . \tag{2.9}$$

When the inverse Laplace transform is performed on R(s), one obtains

$$R(t) = \frac{1}{2\pi j} \int_{\sigma+j\infty}^{\sigma-j\infty} \frac{1}{s} \sum_{i=1}^{N} \frac{A_i}{s - s_i} e^{st} \, ds \tag{2.10a}$$

$$= u(t)\left[\sum_{i=1}^{N} A_i \left[\frac{e^{s_i t}}{s_i} - \frac{1}{s_i}\right]\right] \tag{2.10b}$$

$$= u(t)\left[\sum_{i=0}^{N} B_i \, e^{s_i t}\right] \tag{2.10c}$$

where

$$B_0 = - \sum_{i=1}^{N} A_i / s_i$$

$$B_i = \frac{A_i}{s_i}$$

and

$$s_0 = 0 \quad .$$

Thus, since the step driving function has a singularity at the origin then the response function is written as a simple sum of exponentials.

If the exciting waveform is of finite duration, that is, it is turned off after some time t_0, then it can be shown that the term $g(t,\bar{r})$ is identically equal to zero for time t greater than t_0. Consider, for example, the case where the driving function is a simple pulse

$$F(t) = u(t) - u(t - t_0) \quad ,$$

then

$$F(s) = \frac{1}{s} (1 - e^{-st_0}) \quad .$$

Again let the impulse response be

$$H(s) = \sum_{i=1}^{N} \frac{A_i}{s - s_i} \quad ;$$

then, the response function in the Laplace transform domain is

$$R(s) = F(s)\, H(s) = \frac{1}{s}(1 - e^{st_0}) \sum_{i=1}^{N} \frac{A_i}{s - s_i} \quad . \tag{2.11}$$

The corresponding time domain expression for the response function is

$$R(t) = \int_{\sigma-j\infty}^{\sigma+j\infty} R(s)\, e^{st}\, ds \tag{2.12a}$$

$$R(t) = \begin{cases} \sum_{i=1}^{N} \frac{A_i}{s_i} (e^{s_i t} - 1), & 0 < t < t_0 \\ \sum_{i=1}^{N} B_i\, e^{s_i t}, & t > t_0 \end{cases} \tag{2.12b}$$

where

$$B_i = \frac{1}{s_i} (1 - e^{-s_i t_0}) \quad .$$

Expression (2.12b) shows that after time t_0 the response function is simply a sum of the exponentials which characterize the body while, before t_0, the sum of exponentials contains one exponential term which is dependent on the driving function.

When the driving function is not finite in time and has no pole singularities, as in the case of the Gaussian pulse, then the $g(t,\bar{r})$ term never disappears and cannot be written as a sum of exponentials. However, this difficulty may be circumvented by simply deconvolving the

response function $R(t,\bar{r})$ in a standard manner to obtain the system's impulse response $H(t,\bar{r})$. Equation (2.6) gives for instance

$$H(s) = \frac{R(s)}{F(s)} \quad . \tag{2.13}$$

Thus, removing the driving function from the response function results in the impulse response which is assumed to be a sum of exponentials only. The inevitable presence of experimental and computational noise limits the upper frequency for which the deconvolved spectrum H(s) is accurate and, in practice, the computed spectrum must be truncated beyond this frequency. This simply sets a limit on N. Care also must be taken to exclude the time t = 0 from the deconvolution because of the presence of the delta function discussed previously.

Numerical examples of the above cases are given in Chapter 3 once the method by which the poles will be extracted has been developed.

3. THE NUMERICAL METHOD FOR EXTRACTING POLES

In the previous chapter it was shown that the impulse response and for certain cases the general transient response of a system can be expressed as a sum of a finite number of exponentials. This chapter will develop the numerical method by which the values of the poles and their corresponding residues can be obtained if the impulse or transient response of the system is given. The numerical method used is based on Prony's algorithm [12], [13], [14] which does not seem to be widely known. Prony's algorithm has been used in the fields of automatic control [16] and biological signal processing [17], but only to represent or synthesize a signal in terms of a set of exponentials which do not necessarily have any physical relationship to the system which produced the signal. The desire here, however, is to extract from a system's transient response a set of complex exponentials which are in fact the characteristic resonances of the system being studied.

In this chapter Prony's method is developed for a system with simple poles only. Prony's method is then extended to systems containing multiple poles. The derivation of Prony's method is presented in detail in several different ways so that the reader will get a good feeling for the method and the problems associated with it. An abbreviated treatment is presented in [18]. Several numerical examples are presented in order to establish some guidelines for the use of the method and to illustrate some of the problems associated with the method.

3.1 Prony's Method for Simple Poles

3.1.1 The classical derivation

Thus far it has been assumed that for a system with only simple poles the system's transient response could be expressed as

$$R(t) = \sum_{i=1}^{N} A_i e^{s_i t} \tag{3.1}$$

where the s_i are the pole singularities in the complex frequency plane and the A_i are their corresponding residues. These residues will be a function of the position of observation or measurement on the body being considered and a function of the excitation function. That is, the residues contain both the coupling coefficients and the natural modes. It should also be noted that the s_i must be in complex conjugate pairs or lie on the negative real axis in order to ensure that the response R(t) is real. Since, in practice, one almost always deals with a discrete set of sampled transient data, Equation (3.1) should be rewritten as

$$R(t_n) = R_n = \sum_{i=1}^{N} A_i e^{s_i n\Delta t} \quad ; \quad n = 0,1, \ldots, M-1 \tag{3.2}$$

where Δt is the size of the time-stepping interval used in obtaining the sampled data and $t_n = n\Delta t$. Equation (3.2) consists of M nonlinear equations in 2N unknowns. Another way of writing this set of equations is

$$R_0 = A_1 + A_2 + \ldots + A_N$$

$$R_1 = A_1 Z_1 + A_2 Z_2 + \ldots + A_N Z_N$$

$$R_2 = A_1 Z_1^2 + A_2 Z_2^2 + \ldots + A_N Z_N^2$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$R_{M-1} = A_1 Z_1^{M-1} + A_2 Z_2^{M-1} + \ldots + A_N Z_N^{M-1} \tag{3.3a}$$

where

$$Z_i = e^{s_i \Delta t} \quad . \tag{3.3b}$$

In this set of equations it is necessary to solve for both the N values of Z_i and the N values of the A_i. This solution requires that the value of M be at least equal to 2N ; however, the solution to this set of equations is nontrivial since they are nonlinear in the Z_i's.

Let $Z_1, \ldots, Z_N$ be the roots of the algebraic equation

$$\alpha_0 + \alpha_1 Z^1 + \alpha_2 Z^2 + \ldots + \alpha_N Z^N = 0 \tag{3.4a}$$

so that the left-hand side of (3.4) is equal to the product

$$(Z - Z_1)\ (Z - Z_2) \ldots (Z - Z_N)$$

that is,

$$\sum_{m=0}^{N} \alpha_m Z^m = \prod_{i=1}^{N} (Z - Z_i) = 0 \quad . \tag{3.4b}$$

The coefficients α_i may be determined as follows. Multiply the first equation in (3.3) by α_0, the second by α_1, . . ., the N + 1 by α_N and add the results. Since each of the Z_i satisfies Equation (3.4a), then the result is of the form

$$\alpha_0 R_0 + \alpha_1 R_1 + \ldots + \alpha_N R_N = 0 \quad .$$

A set of M - N - 1 additional equations is obtained in the same manner by starting successively with the second, third, . . ., $(M - N)^{th}$ equations. Thus, it is possible to obtain the M - N linear difference equations

$$\alpha_0 R_0 + \alpha_1 R_1 + \ldots \alpha_N R_N = 0$$

$$\alpha_0 R_1 + \alpha_1 R_2 + \ldots \alpha_N R_{N+1} = 0$$

$$\begin{array}{ccc} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{array}$$

$$\alpha_0 R_{M-N-1} + \alpha_2 R_{M-N} + \ldots \alpha_N R_{M-1} = 0 \qquad (3.5a)$$

which can be more simply written as

$$\sum_{p=0}^{N} \alpha_p R_{p+k} = 0; \; p + k = n = 0,1, \ldots, M - 1 \quad . \qquad (3.5b)$$

Thus, the sampled values R_i satisfy an N^{th} order linear difference equation. This difference equation will be referred to as Prony's difference equation.

In Equations (3.5) the R_i are known, hence, this set can be solved to obtain the N + 1 coefficients, α. If M = 2N then the system of equations may be solved directly by matrix inversion, and if M > 2N the system may be solved approximately by the method of least squares. Equations (3.5) are most conveniently solved by defining α_N equal to one; then, the α_p's may be obtained by solving the equations

$$\sum_{p=0}^{N-1} \alpha_p R_{p+k} = -R_{N+k}; \; p + k = n = 0,1, \ldots, M - 1 \quad . \tag{3.6}$$

If M = 2N this set of equations is written in matrix form as

$$AB = C \tag{3.7a}$$

where

$$A = \begin{bmatrix} R_0 & R_1 & R_2 & \cdots & R_{N-1} \\ R_1 & R_2 & R_3 & \cdots & \\ R_2 & R_3 & R_4 & & \\ R_3 & R_4 & R_5 & & \\ \cdot & \cdot & \cdot & & \\ \cdot & \cdot & \cdot & & \\ \cdot & \cdot & \cdot & & \\ R_{N-1} & R_N & \cdot & \cdots & R_{2N-2} \end{bmatrix} , \tag{3.7b}$$

$$B = \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \cdot \\ \cdot \\ \cdot \\ \alpha_{N-1} \end{bmatrix} , \qquad (3.7c)$$

$$C = -\begin{bmatrix} R_N \\ R_{N+1} \\ \cdot \\ \cdot \\ \cdot \\ R_{2N-1} \end{bmatrix} . \qquad (3.7d)$$

Matrix A is a square symmetric circulant matrix and thus is readily invertible. If M is greater than 2N , then the set of Equations (3.6) can be solved using a least-squares approach. This is most conveniently done by performing a pseudo-inverse to write Equations (3.6) in matrix form as

$$A^T AB = A^T C \qquad (3.8a)$$

where A is now a rectangular matrix of the form

$$A = \begin{bmatrix} R_0 & R_1 & R_2 \cdots R_{N-1} \\ R_1 & R_2 & \cdots R_N \\ \cdot & & \\ \cdot & & \\ \cdot & & \\ R_{M-N-1} & R_{M-N} & R_{M-2} \end{bmatrix} \tag{3.8b}$$

and

$$C = -\begin{bmatrix} R_N \\ R_{N+1} \\ \cdot \\ \cdot \\ \cdot \\ R_{M-1} \end{bmatrix} . \tag{3.8c}$$

The square matrix formed by multiplying matrix A by its transpose A^T is simply a Gramian matrix and is also symmetric and circulant.

Once the α's have been determined by either of the above approaches, the next step is to solve for the N values of Z_i. These Z_i are obtained by finding the roots of the polynomial (3.4a). The N roots are complex numbers and appear in complex conjugate pairs. The roots of (3.4a) may be found by using any of the several polynomial root-finding methods. The most accurate and quickest of these methods is a routine based on Muller's method [19], [20].

It is now a trivial matter to obtain the poles s_i. Since the roots of (3.4a) were defined by (3.3b) as

$$Z_i = e^{s_i \Delta t} \tag{3.3b}$$

then the poles are simply

$$s_i = \frac{\ell n\, Z_i}{\Delta t} \quad . \tag{3.9}$$

The final step in Prony's method is to determine the values of the residues A_i. To do this one simply solves the matrix equation embodied in Equations (3.3a). In matrix form this set of equations is written as

$$DE = F \quad , \tag{3.10a}$$

where

$$D = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ Z_1 & Z_2 & \cdots & Z_N \\ Z_1^2 & Z_2^2 & \cdots & Z_N^2 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ Z_1^{N-1} & Z_2^{N-1} & \cdots & Z_N^{N-1} \end{bmatrix} \quad , \tag{3.10b}$$

$$E = \begin{bmatrix} A_1 \\ A_2 \\ \cdot \\ \cdot \\ \cdot \\ A_N \end{bmatrix} \quad , \tag{3.10c}$$

$$F = \begin{bmatrix} R_0 \\ R_1 \\ \cdot \\ \cdot \\ \cdot \\ R_{N-1} \end{bmatrix} . \qquad (3.10d)$$

Note that the matrix D is in the form of a Vandermonde matrix which has an inverse which can be computed in closed form. If, however, one wishes to solve Equation (3.3a) by using a least-squares approach then the resultant matrix equation has a matrix which is only symmetric and cannot be inverted as simply as the Vandermonde matrix D.

3.1.2 Relationship to the Z-transform

The method presented in the above subsection for finding the simple poles and residues from a set of discrete transient data is Prony's method as it is usually derived and used. However, Weiss and McDonough [21] have demonstrated that Prony's method may be regarded as a Padé approximation in the Z-transform domain.

Consider the Z-transform of R(t) to be

$$R(Z) = R_0 + R_1 Z^{-1} + R_2 Z^{-2} + \ldots + R_{2N-1} Z^{-(2N-1)} + \ldots , \qquad (3.11)$$

In the Padé approximation one seeks to equate the first 2N terms of (3.11) with a function of the form

$$F(Z) = \frac{a_N Z^N + a_{N-1} Z^{N-1} + \ldots + a_1 Z}{Z^N + \alpha_{N-1} Z^{N-1} + \ldots + \alpha_1 Z + \alpha_0} . \qquad (3.12)$$

Note that the denominator in (3.12) is the same as the polynomial in (3.4a). Carrying out the Padé procedure by equating (3.11) and (3.12) and multiplying through by the denominator of (3.12) yield

$$a_N Z^N + a_{N-1} Z^{N-1} + \ldots + a_1 Z = (Z^N + \alpha_{N-1} Z^{N-1} + \ldots + \alpha_1 Z + \alpha_0)$$

$$\cdot (R_0 + R_1 Z^{-1} + R_2 Z^{-2} + \ldots + R_{2N-1} Z^{-(2N-1)} + \ldots,) \quad . \quad (3.13)$$

Equating the coefficients of like powers of Z in (3.13) yields the following set of equations.

$$\left.\begin{aligned} a_N &= R_0 \\ a_{N-1} &= R_0 \alpha_{N-1} + R_1 \\ a_{N-2} &= R_0 \alpha_{N-2} + R_1 \alpha_N + R_2 \\ &\cdot \\ &\cdot \\ &\cdot \\ a_1 &= R_0 \alpha_1 + R_1 \alpha_2 + \ldots + R_{N-1} \end{aligned}\right\} \quad (3.14a)$$

$$\left.\begin{aligned} R_0 \alpha_0 + R_1 \alpha_1 + \ldots R_{N-1} \alpha_{N-1} + R_N &= 0 \\ R_1 \alpha_0 + R_2 \alpha_1 + \ldots R_N \alpha_{N-1} + R_{N+1} &= 0 \\ &\cdot \\ &\cdot \\ &\cdot \\ R_{N-1} \alpha_0 + R_N \alpha_1 + \ldots R_{2N-2} \alpha_{N-1} + R_{2N-1} &= 0 \end{aligned}\right\} \quad (3.14b)$$

Equations (3.14b) are the same as Equations (3.5a) of the previous section.

If the Z-transform of Equation (3.1) is taken, the result is

$$R(Z) = \sum_{i=1}^{N} A_i \frac{Z}{Z - Z_i} \tag{3.15}$$

where $Z_i = e^{s_i n\Delta t}$. Hence, the α_i and the Z_i of the Padé approximation in the Z-transform domain are the same as the α_i and the Z_i of the previous section. This method allows one to solve for the A_i by a somewhat simpler method. Rather than solving the matrix Equation (3.3a), one can simply solve for the a_i in (3.14a) and perform the partial fraction expansion of $\frac{R(Z)}{Z}$ to yield

$$\frac{1}{Z} R(Z) = \sum_{i=1}^{N} \frac{A_i}{Z - Z_i} \quad . \tag{3.16}$$

3.1.3 Relationship to Corrington's difference equation

In 1965 Corrington derived a difference equation from which it is possible to extrapolate the time response of a linear lumped-constant time-invariant network to late time by knowing only discrete values of the early-time response [22]. His difference equation can be shown to be Prony's difference equation in a somewhat camouflaged form. His derivation is interesting and is presented here in a similar form. This derivation is useful when the development of the derivation for multiple poles is presented in the following section.

Equation (3.2) may be rewritten in an alternate form as

$$R[(n - r)\,\Delta t] = \sum_{i=1}^{N} A_i e^{s_i n\Delta t} e^{-s_i r\Delta t} \quad ; \; r = 0,1, \ldots, N \quad . \tag{3.17}$$

The term $A_1 e^{s_1 n\Delta t}$ can be eliminated from the above N + 1 linear equations by subtracting each equation in turn from the preceding one after multiplication by $e^{s_1 \Delta t}$. This gives the set of N linear equations

$$R[(n - r)\,\Delta t] - e^{s_1 \Delta t}\, R[(n - r - 1)\,\Delta t]$$

$$= e^{s_1 \Delta t} \sum_{i=1}^{N-1} e^{-r\Delta t s_{i+1}} \prod_{k=1}^{1} \left(e^{-s_k \Delta t} - e^{-s_{k+1} \Delta t}\right)$$

$$\cdot A_{i+1}\, e^{s_{i+1} n\Delta t}\,; \quad r = 0,1, \ldots, N - 1 \tag{3.18}$$

from which the term $A_2\, e^{s_2 n\Delta t}$ can be eliminated by the same procedure. The result is

$$R[(n - r)\,\Delta t] - R[(n - r - 1)\,\Delta t]\left(e^{s_1 \Delta t} + e^{s_2 \Delta t}\right) + R[(n - r - 2)\,\Delta t]$$

$$\cdot \left(e^{s_1 \Delta t}\, e^{s_2 \Delta t}\right) = e^{s_2 \Delta t} \sum_{i=1}^{N-2} e^{-r\Delta t s_{i+2}} \prod_{k=1}^{2}$$

$$\cdot \left(e^{-s_k \Delta t} - e^{-s_{k+2} \Delta t}\right) \cdot A_{i+2}\, e^{s_{i+2} n\Delta t},$$

$$r = 0,1,2, \ldots, N - 2 \quad . \tag{3.19}$$

If this process is continued until all N of the $A_i\, e^{s_i n\Delta t}$ are eliminated, the result is

$$\sum_{p=0}^{N} \alpha_{N-p}\, R([n - p]\,\Delta t) = 0 \tag{3.20}$$

where the N + 1 values of α_p are the exact same α_p appearing in the algebraic Equation (3.4b). Since the Z_i are defined as

$$Z_i = e^{s_i \Delta t},$$

then the α_p are sums and products of the $e^{s_i \Delta t}$. For example if N = 3, then (3.4b) is written

$$\sum_{p=0}^{3} \alpha_p Z^p = \left(Z - e^{s_1 \Delta t}\right)\left(Z - e^{s_2 \Delta t}\right)\left(Z - e^{s_2 \Delta t}\right) = 0$$

and, thus the α's are

$$\alpha_3 = 1$$

$$\alpha_2 = -e^{s_1 \Delta t} - e^{s_2 \Delta t} - e^{s_3 \Delta t}$$

$$\alpha_1 = e^{s_1 \Delta t}\, e^{s_2 \Delta t} + e^{s_1 \Delta t}\, e^{s_3 \Delta t} + e^{s_2 \Delta t}\, e^{s_3 \Delta t}$$

$$\alpha_0 = -e^{s_1 \Delta t}\, e^{s_2 \Delta t}\, e^{s_3 \Delta t} \quad .$$

Hence, Equation (3.20) is an alternative way of writing Prony's difference Equation (3.5b). Corrington writes the difference equation as

$$-\alpha_N R(n\Delta t) = \sum_{p=1}^{N} \alpha_{N-p} R[(n - p)\, \Delta t]; \quad n \geq N \tag{3.21}$$

which, if the α's are known, expresses the repsonse at some time n in terms of N previous time sampled values of the response.

3.2 Prony's Method for Multiple Poles

For the most part electromagnetic antennas and scatterers possess only simple poles. However, Tesche [23] has shown that a dipole can be resistively loaded in such a way as to make it critically damped, that is, to have a double pole on the negative real axis. Multiple poles also result in the transient response of a system if the system is driven by a signal which itself has a multiple pole. The most common example of this would be a system excited by a ramp waveform. The

ramp waveform has a double pole located at the origin. It is shown in this section that Prony's method can be used to extract these multiple poles without knowing a priori that they are present.

The general form of the transient response for a system containing both multiple and simple poles can be written as

$$R(t) = \sum_{i=1}^{N} \left\{ 1 + \sum_{j=2}^{M_i} t^{j-1} B_{ji} P(M_i) \right\} A_i e^{s_i t} \tag{3.22}$$

where

$$P(M_i) = 0, \quad \text{if } M_i < 2$$

$$P(M_i) = 1, \quad \text{if } M_i \geq 2$$

and where M_i is the multiplicity of the i^{th} pole. For example, if the i^{th} pole is a double pole, then $M_i = 2$ and $P(M_i) = 1$. In discrete data form, Equation (3.22) can be written

$$R[(n - r)\,\Delta t] = \sum_{i=1}^{N} \left\{ 1 + \sum_{j=2}^{M_i} [(n - r)\,\Delta t]^{j-1} B_{ji} P(M_i) \right\} A_i e^{s_i (n-r)\Delta t}; \quad r = 0,1,2, \ldots, L \tag{3.23}$$

where L is the total number of poles if each pole is taken M_i times. If as for the simple pole case

$$Z_i = e^{s_i \Delta t},$$

then an algebraic equation can be written

$$\prod_{i=1}^{N} (Z - Z_i)^{M_i} = \sum_{m=0}^{L} \alpha_m Z^m = 0 \quad . \tag{3.24}$$

It is shown that the L Equations (3.23) can be reduced to give Prony's difference equation

$$\sum_{m=0}^{L} \alpha_{L-m} R[(n - m)\ \Delta t] = 0 \quad . \tag{3.25}$$

This difference equation is derived in the same manner as in Section 3.1.3 except for one small difference. As before, each equation in turn is subtracted from the previous one after multiplying it by $e^{s_i \Delta t}$. This step is repeated M_i times and then $e^{s_{i+1} \Delta t}$ is used. This can be best demonstrated in the following example.

Let the response be

$$R[(n - r)\ \Delta t] = \sum_{i=1}^{2} \left\{ 1 + \sum_{j=2}^{M_i} [(n - r)\ \Delta t]^{j-1} B_{ji} P(M_i) \right\} A_i e^{s_i (n-r)\ \Delta t} ;$$

$$r = 0,1,2,3$$

and if $\Delta t = 1$, $M_1 = 2$, $M_2 = 1$, then,

$$R[n - r] = (A + Bt) e^{s_1 (n-r)} + C e^{s_2 (n-r)} ;$$

$$r = 0,1,2,3 \quad . \tag{3.26}$$

Note here that the constant B is simply the product $B_{21} A_1$. If one uses the method of Section 3.1.3, the four Equations (3.26) are reduced to the three

$$R(n - r) - e^{s_1} R(n - r - 1) = B e^{s_1 n} e^{-r s_1} + C e^{s_2 n} e^{-r s_2} \cdot (1 - e^{s_1} e^{-s_2}) \quad ,$$

$$R(n-r-1) - e^{s_1} R(n-r-2) = B\, e^{s_1 n}\, e^{-rs_1}\, e^{-s_1} + C\, e^{s_2 n}\, e^{-rs_2} \cdot (e^{-s_2} - e^{-2s_2}\, e^{s_1}) \quad ,$$

$$R(n-r-2) - e^{s_1} R(n-r-3) = B\, e^{s_1 n}\, e^{-rs_1}\, e^{-2s_1} + C\, e^{s_2 n}\, e^{-rs_2} \cdot (e^{-2s_2} - e^{-3s_2}\, d^{s_1}).$$

(3.27)

Since the pole s_1 was a double pole, the step is repeated by multiplying by e^{s_1} again. The two resulting equations are

$$R(n-r) - e^{s_1} R(n-r-1) - e^{s_1} R(n-r-1) + e^{2s_1} R(n-r-2) = C\, e^{s_2 n}\, e^{-rs_2} (1 - 2\, e^{s_2}\, e^{s_1} + e^{-2s_2}\, e^{2s_1}) \quad , \qquad (3.28a)$$

$$R(n-r-1) - e^{s_1} R(n-r-2) - e^{s_1} R(n-r-2) + e^{2s_1} R(n-r-3) = C\, e^{s_2 n}\, e^{-rs_2} (e^{-s_2} - 2\, e^{-2s_2}\, e^{s_1} + e^{-3s_2}\, e^{2s_1}) \, . \qquad (3.28b)$$

Finally, the difference equation is obtained by multiplying (3.28b) by e^{s_2} and subtracting from (3.28a). The difference equation is thus

$$R(n-r) - (e^{s_1} + e^{s_1} + e^{s_2})\, R(n-r-1) + (e^{2s_1} + 2\, e^{s_1 s_2})\, R(n-r-2) - (e^{2s_1}\, e^{s_2})\, R(n-r-3) = 0 \quad . \qquad (3.29)$$

From (3.29) the α's of (3.25) are

$$\alpha_3 = 1$$

$$\alpha_2 = -e^{s_1} - e^{s_1} - e^{s_2}$$

$$\alpha_1 = e^{2s_1} + 2\, e^{s_1 s_2}$$

$$\alpha_0 = -e^{2s_1}\, e^{s_2} \quad .$$

This example shows that although there were only two distinct poles the difference equation for the response was of order three producing four coefficients.

In the above example, it was shown that Prony's difference equation is also valid for transient responses containing multiple poles. The order of the difference equation does increase from N to L thus requiring L + 1 values of α to be solved for. Once the L + 1 values of α have been determined, then the next step is to find the L roots of the polynomial (3.24). If Muller's method is used for finding the L roots of the polynomial, it does indeed return an individual root as many times as its multiplicity requires. Therefore, after finding the roots of the polynomial, the roots are scanned to see if any appear more than once, indicating the presence of a multiple pole. Hence, it is possible to proceed with Prony's method to the point of finding the poles without knowing if there are multiple poles present or not. After the poles and their multiplicity have been determined, then the problem is to calculate the values of the residues.

The calculation of the residues is done by solving the equation

$$R(n\Delta t) = \sum_{i=1}^{N} \left\{ 1 + \sum_{j=2}^{M_i} (n\Delta t)^{j-1} B_{ji} P(M_i) \right\} A_i e^{s_i n\Delta t} \qquad n = 0,1, \ldots \quad (3.30)$$

which differs from Equation (3.2) because of the presence of the terms $(n\Delta t)^{j-1}$. This fact causes the matrix Equation (3.10) to be changed by multiplying certain columns by multiples of $(n\Delta t)$. For example, if from the previous example,

$$Z_1 = e^{s_1 \Delta t}$$

$$Z_2 = e^{s_1 \Delta t}$$

$$Z_3 = e^{s_2 \Delta t} \quad ,$$

then the resulting matrix equation to be solved in order to find the coefficients A, B, and C of (3.26) is

$$\begin{bmatrix} 1 & 0 & 1 \\ Z_1 & 1\Delta t Z_2 & Z_3 \\ Z_1^2 & 2\Delta t Z_2^2 & Z_3^2 \end{bmatrix} \begin{bmatrix} A \\ B \\ C \end{bmatrix} = \begin{bmatrix} R(0\Delta t) \\ R(1\Delta t) \\ R(2\Delta t) \end{bmatrix}$$

where column 2 has been changed because of the presence of the term Bt in (3.26). The next example shows that this necessity to alter certain columns of the matrix leads to instabilities in the solution process.

As a more complicated example of a system with a double pole consider the transient response

$$R(t) = 3.0 + 7.0\ t + 6.0\ e^{-3t} \cos(4t - \pi/6) + 4.0\ e^{-3t} \cos(6t + \pi/3) \qquad (3.31)$$

which is plotted in Figure 3.1. Note that the second term in (3.31) is a ramp response which is very dominant in the plotted transient response. This ramp term gives rise to a double pole at the origin. When Prony's method was applied to these data, the extracted poles and residues in Table 3.1 were obtained. Note that the extracted poles are all within

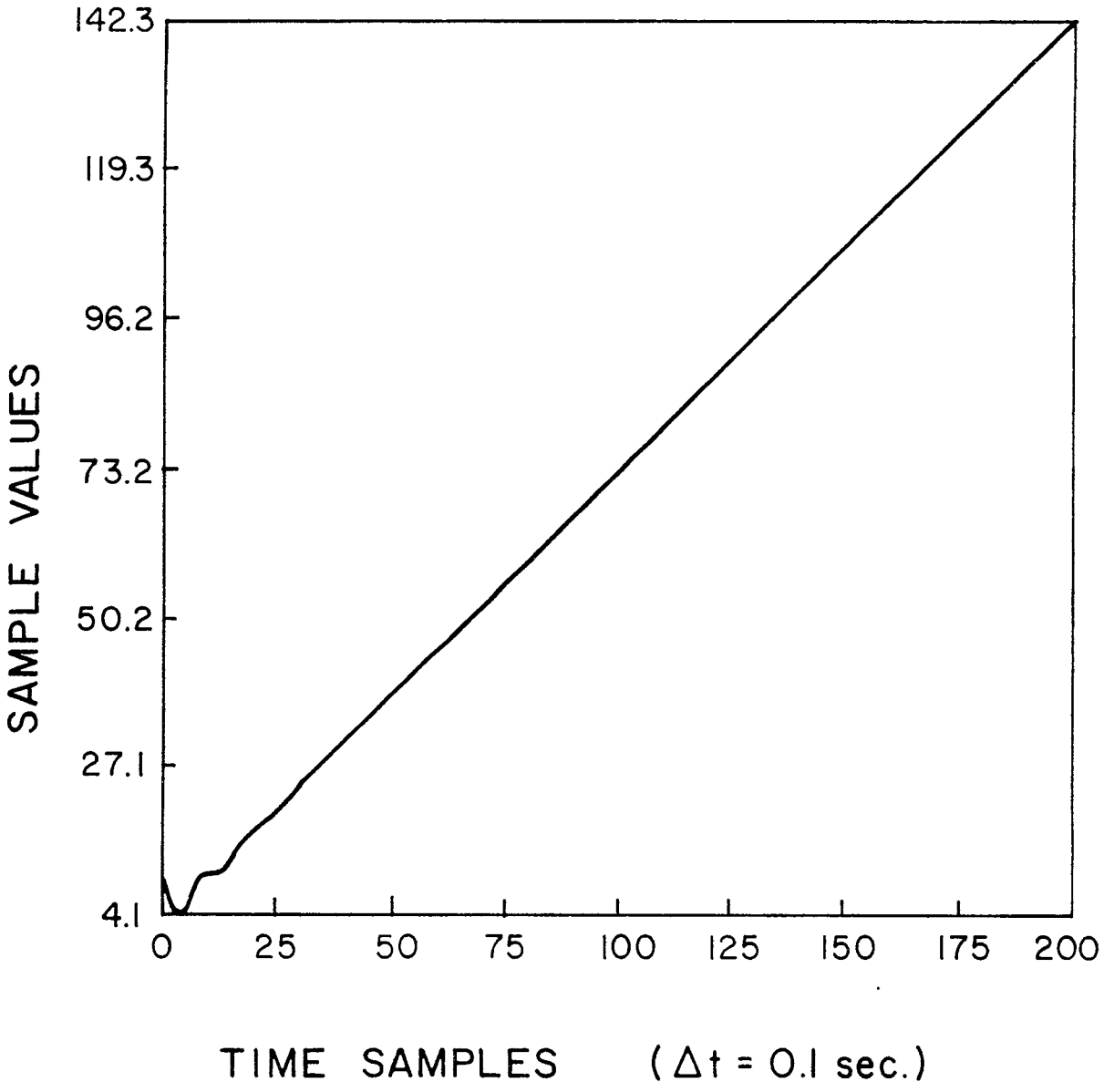


Figure 3.1. Response containing a double pole at the origin.

TABLE 3.1

EXTRACTED POLES AND RESIDUES FOR THE DOUBLE POLE CASE

	True Poles	Extracted Poles	True Residues	Extracted Residues
Real Part	0.0	0.00028	3.0	3.7758
	0.0	0.00028	7.0	6.8978
	-3.0	-3.0040	2.598	3.8069
	-3.0	-3.0040	2.598	1.2561
	-1.0	-1.0004	1.0	1.1117
	-1.0	-1.0004	1.0	0.5614
Imaginary Part	0.0	0.0	0.0	-0.5906
	0.0	0.0	0.0	0.0575
	4.0	4.0011	+1.5	0.1919
	-4.0	-4.0011	-1.5	-0.1052
	6.0	6.0009	+1.732	1.4987
	-6.0	-6.0009	-1.732	-1.2444

0.08 percent of the original poles and that the double pole was indeed found. However, the residues that were recovered do not compare with the true residues at all. The residues do not even fall in complex conjugate pairs as is required. This is undoubtedly due to the fact that the first matrix was altered, as discussed in the previous example, in such a way as to make it ill-conditioned. The important point in this example is, however, that the double pole was extracted from the transient response.

3.3 Numerical Examples

In using Prony's method, as in the use of any numerical routine, several guidelines should be followed in order for the method to work accurately and quickly. For the most part, these guidelines were obtained after running several examples on the computer and studying the results. Therefore, this section presents several numerical examples which point out the problems and guidelines which one must follow for the successful use of Prony's method.

In this section, for all but one example, the data analyzed are from the transient response of the current on a 1.0 m long dipole with a half-length-to-radius ratio of 100. The data were obtained using a numerical time-domain computer code [24]. The antenna was modeled using sixty equal-length segments and the exciting field was a Gaussian pulse which was applied across the center two segments of the antenna model. The Gaussian-pulse-time variation was $\exp(-a^2(t - t_{max})^2)$ with a, the Gaussian spread parameter, equal to 5×10^9 s^{-1} and t_{max} equal to 5.556×10^{-10} s. The induced current at the center of each of the sixty segments was calculated for 500 time steps, where the time step size was $\Delta t = 5.556 \times 10^{-11}$ s. Note that

$$\Delta t = \frac{L}{60 \times C}$$

where L is the length of the structure, 1 m, and C is the speed of light. In order to calculate the complex resonances of the structure, the current on one of the center or source segments is used. This current is plotted in Figure 3.2.

The study of Figure 3.2 reveals several facts about the dipole response. The most apparent is the damped oscillatory behavior of the current. The initial spike, which is a maximum at $t_{max} = 10\Delta t$, is the Gaussian driving function being applied. This spike is followed by an immediate negative spike which decays toward zero until time equals $60\Delta t$. The response from time 0 to $60\Delta t$ is very similar to the response of an infinite long dipole. This response is expected since the current which propagates down the antenna and gets reflected back from the end is not seen at the center of the antenna until $60\Delta t$. Thus, the current at the center of the antenna is not affected by the length of the antenna until the current has propagated to the end and back.

Example 1

The first example of the application of Prony's method is shown in Figure 3.3. Of the 500 current samples available, only eighty sampled values were actually used. These eighty samples were taken from the first 160 current samples at every second time step. Figure 3.3a shows the range of the current which was used to fill the matrix of Prony's difference equation. Prony's method was solved using the standard inversion technique. Thus, since eighty sampled values were used, forty poles and residues were obtained. Figure 3.3b shows the left half of the s-plane in which the forty extracted poles have been plotted. Figure 3.4 shows a comparison of the extracted poles with the true even poles of the dipole as calculated by Tesche [5]. As can be seen only

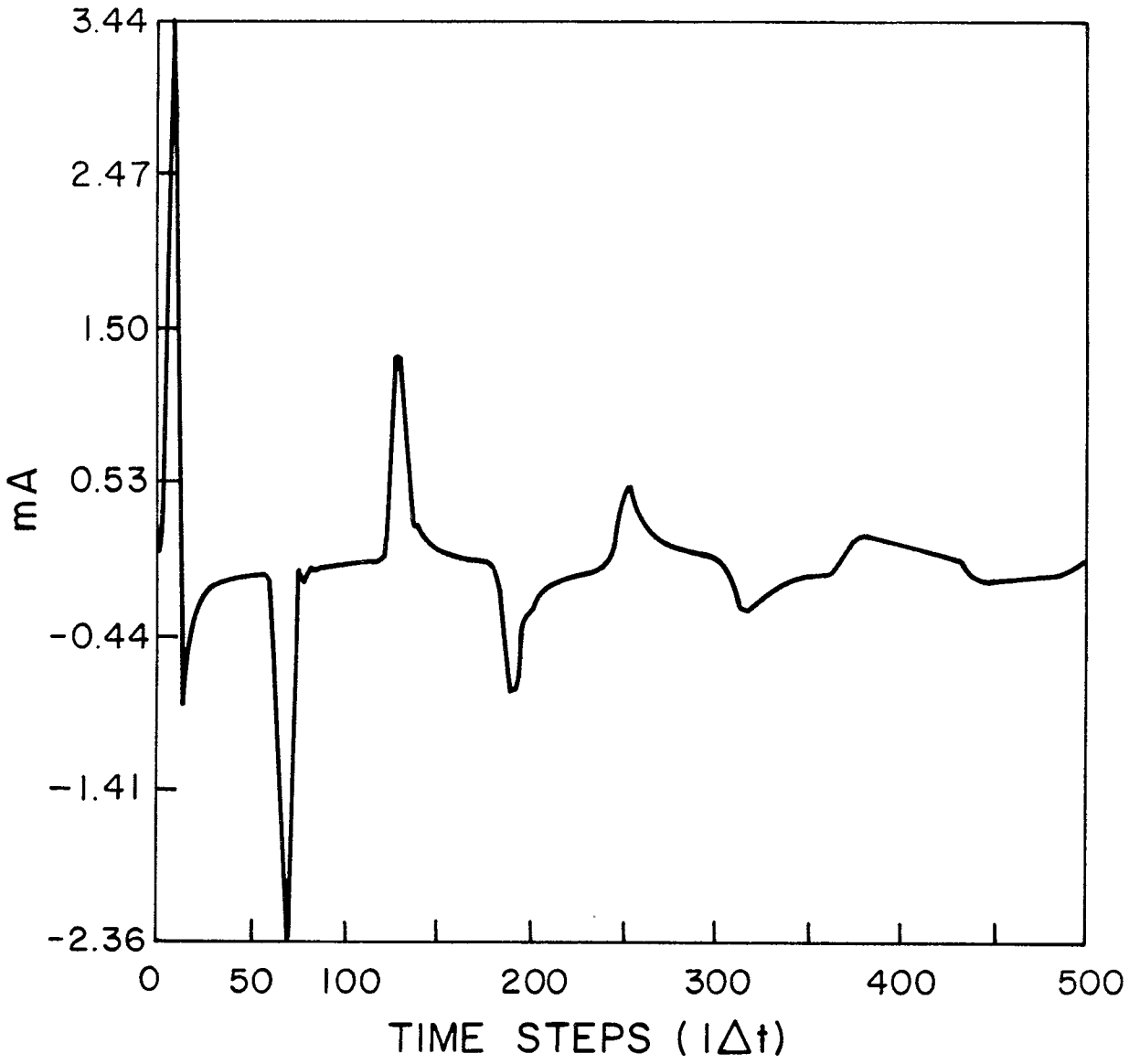
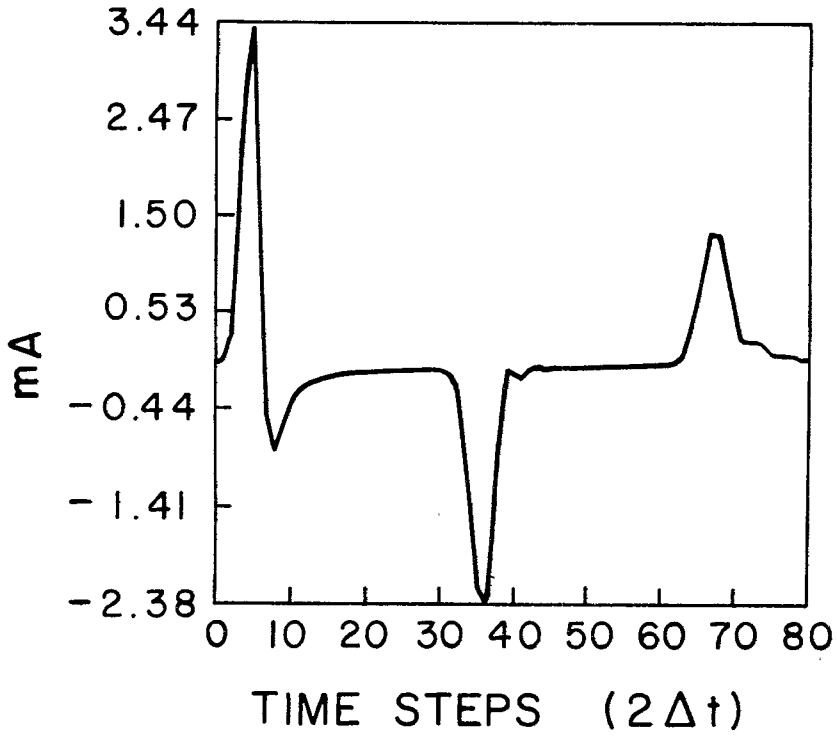
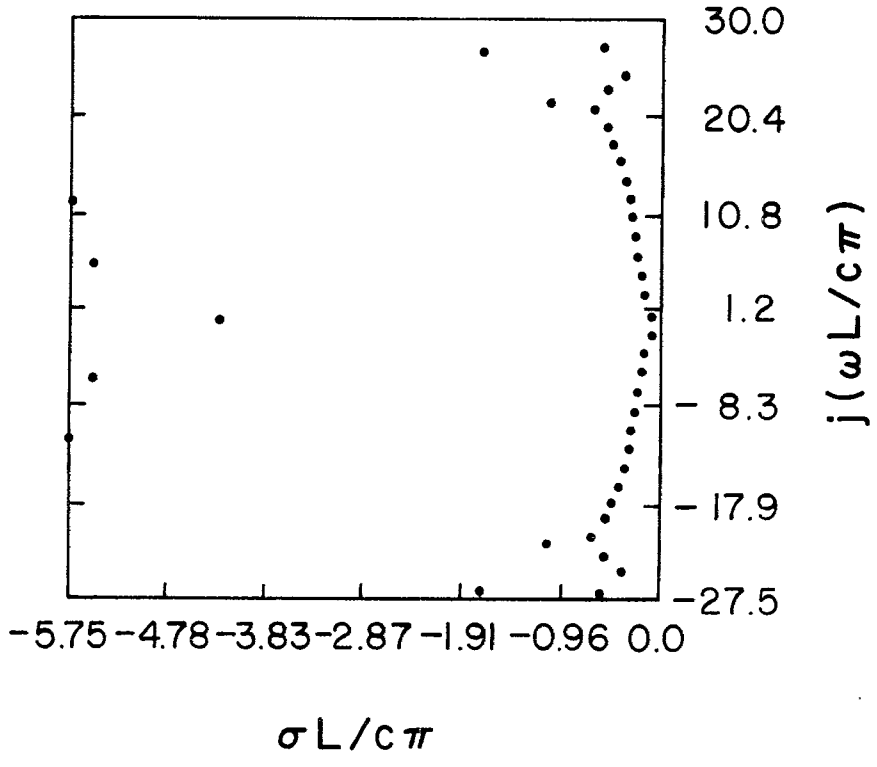


Figure 3.2. Source current on 1.0 m dipole.



(a)



(b)

Figure 3.3. (a) Data window used in Example 1.
(b) Locations of extracted poles in Example 1.

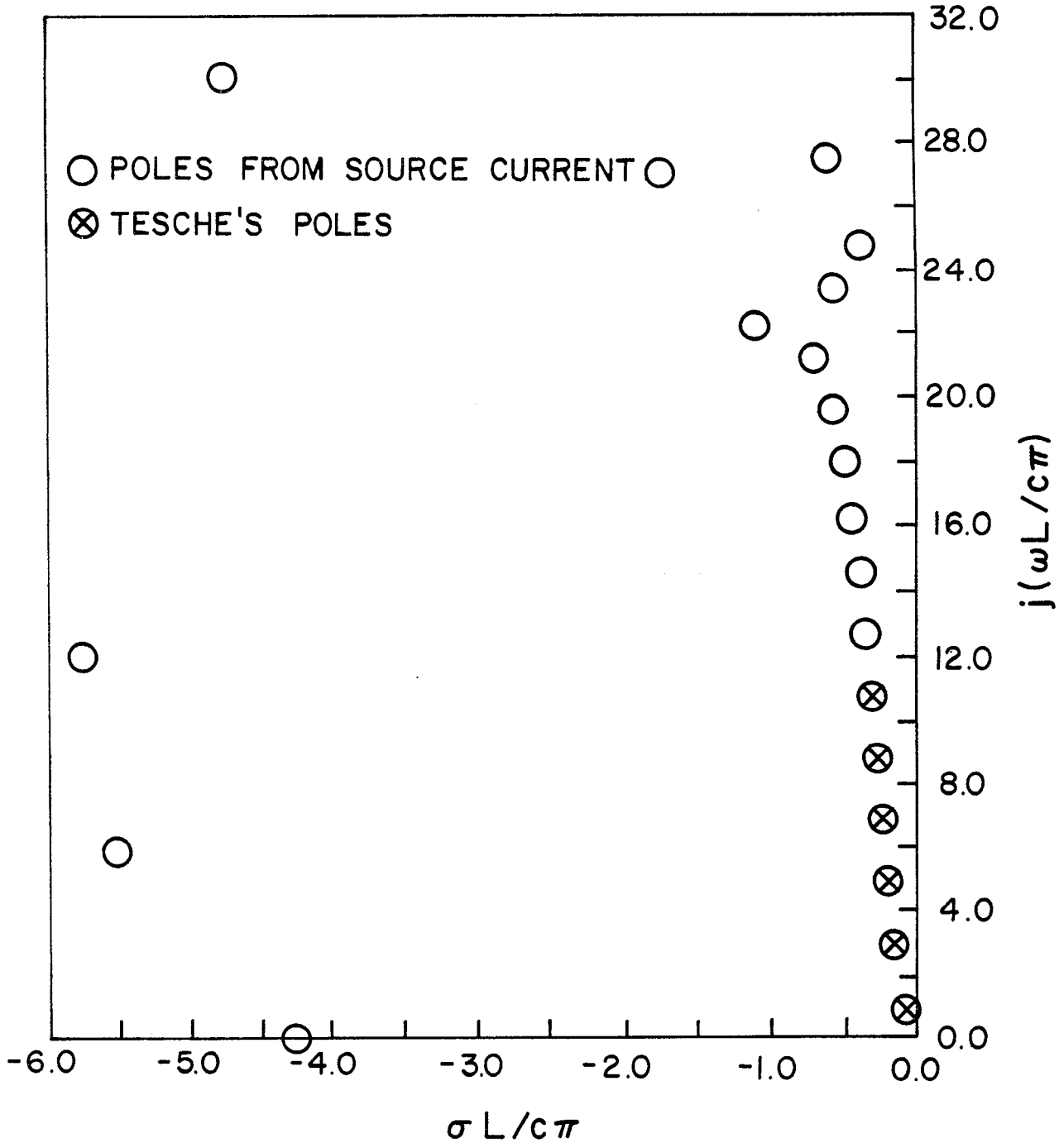


Figure 3.4. Comparison of extracted poles of Example 1 with those of Tesche.

eleven pole pairs correspond to the true poles of the system leaving eighteen poles which have no relationship to the system. The fact that only the first ten even poles were generated is not surprising if one considers the original model. The original model was a thin-wire approximation and was driven with even symmetry by a Gaussian pulse. Because of the thin-wire model and the width of the Gaussian pulse, the expected spectral response has an upper frequency limit of about 10 L/λ [25], where L equals the antenna length. The resonances for a dipole occur at approximately $\lambda/L = 1/2, 3/2, 5/2, \ldots$, thus, with the upper frequency limit of 10 L/λ not many more than the first ten even resonances can be expected. The extra eighteen poles appear for two reasons. In Prony's method, if the least squares approach is not used, when 2N sampled values are used the method returns N poles. Thus, Prony's method is forced to return more poles than are present in the system. Also, since the Gaussian-pulse driving function is only a model of a delta function, then the response function is not a true impulse response but is a response of the form of (2.8). Since the g(t) term of (2.8) was not removed, the extra poles present are needed to represent it. Figure 3.5 demonstrates how well the transient response was modeled with these forty poles and their associated residues. Note that 1000 time steps were generated when only the first 160 of the original time steps were used to generate the poles. This shows how Prony's method can be used to extrapolate late time data from a small set of early time data. Figure 3.6 is a three-dimensional plot of the second quadrant of the s-plane showing both the position and the amplitude of the poles. The length of the vertical lines represents the amplitude of the poles on a logarithmic scale. The amplitude of the true poles tapers off as the

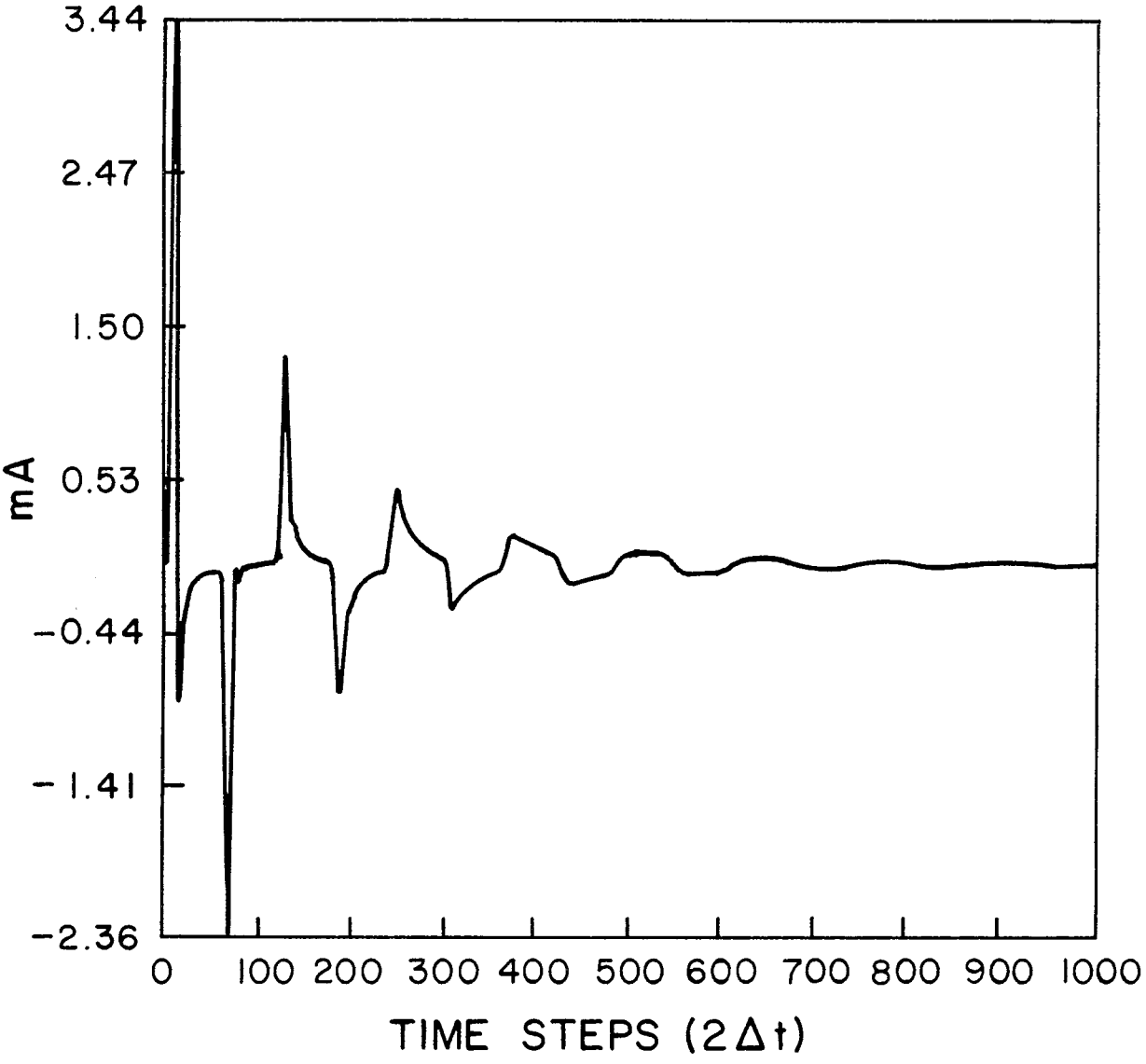


Figure 3.5. Reconstructed transient response using extracted poles of Example 1.

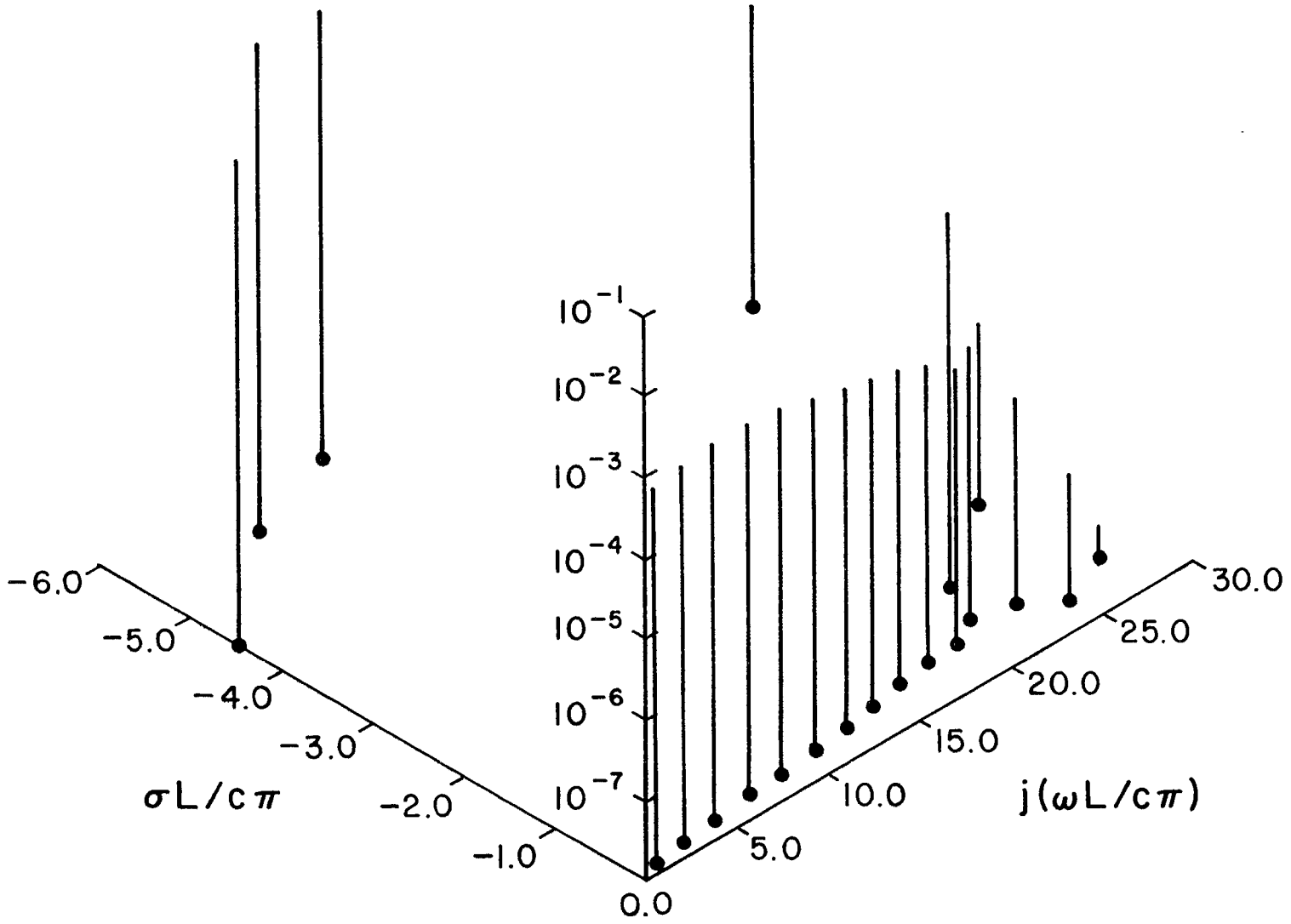


Figure 3.6. Position and amplitude of poles in the second quadrant of the s-plane for Example 1.

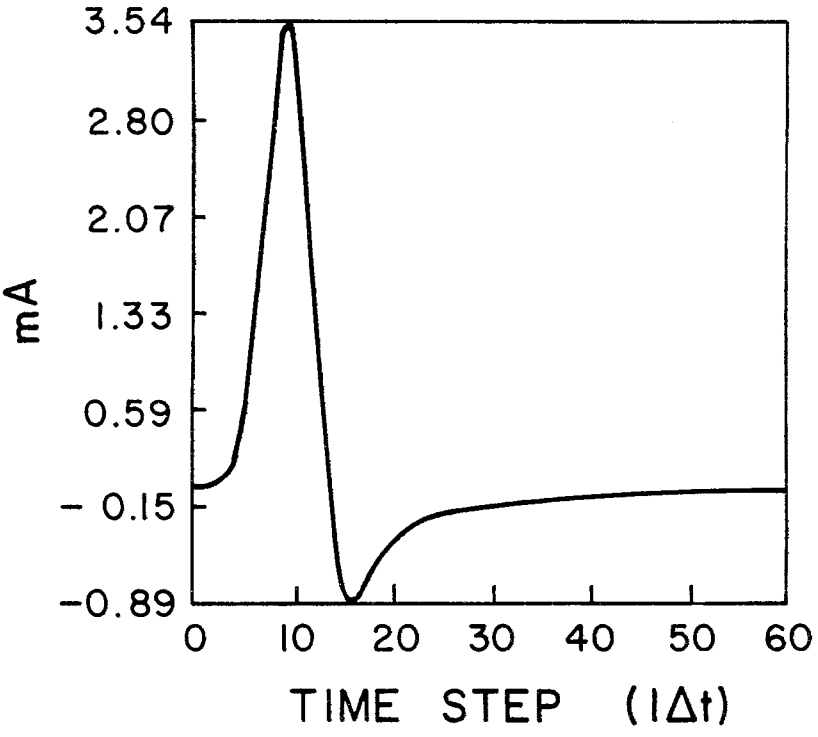
frequency increases because the amplitude of the Gaussian driving function in the spectral domain also tapers off as frequency increases. It should be noted that the amplitudes of some of the extra poles are equal to or greater in magnitude than those of the true poles. These extra poles with the large amplitude are the poles which are predominantly representing the driving function.

Example 2

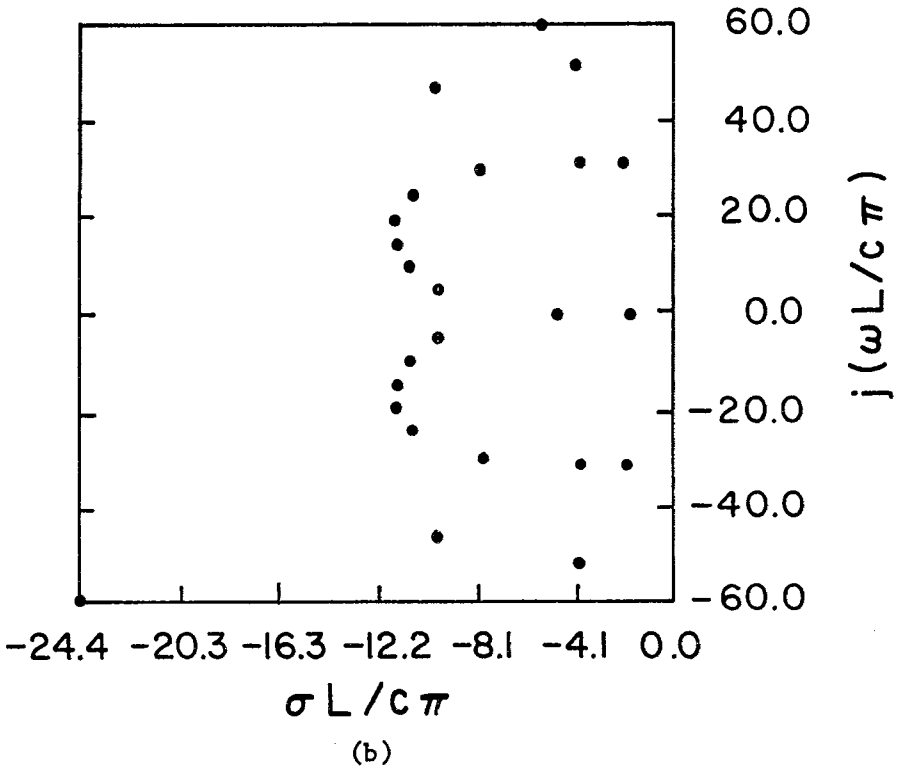
As a second example only the first sixty samples of the response are used. This sampling interval is only large enough to contain the response resembling the infinite wire, Figure 3.7a. The resulting poles are plotted in Figure 3.7b. Note that none of these poles are the true poles of the system but that the sampled portion of the response is accurately reproduced using these poles, Figure 3.8. Also, note that only twenty-five poles are plotted in Figure 3.7b when thirty were expected. This is due to the fact that the program was written so as to eliminate any right-half-plane poles that appear. In this case, five poles with positive real parts were removed before the residues were calculated. This example points out the necessity to include some portion of the oscillatory part of the response. It was found that at least one cycle of the lowest frequency present must be used. Thus, the sample interval of example 1 is the minimum which can be used to obtain the true poles of the system.

Example 3

This example also takes sixty samples but the time step size is now $3\Delta t$ thus giving a sampling interval of the required size, Figure 3.9a, and requiring the return of only thirty poles as opposed to forty in the first example. The twenty-eight poles that were found are plotted in



(a)



(b)

Figure 3.7. (a) Data window used in Example 2.
(b) Locations of extracted poles in Example 2.

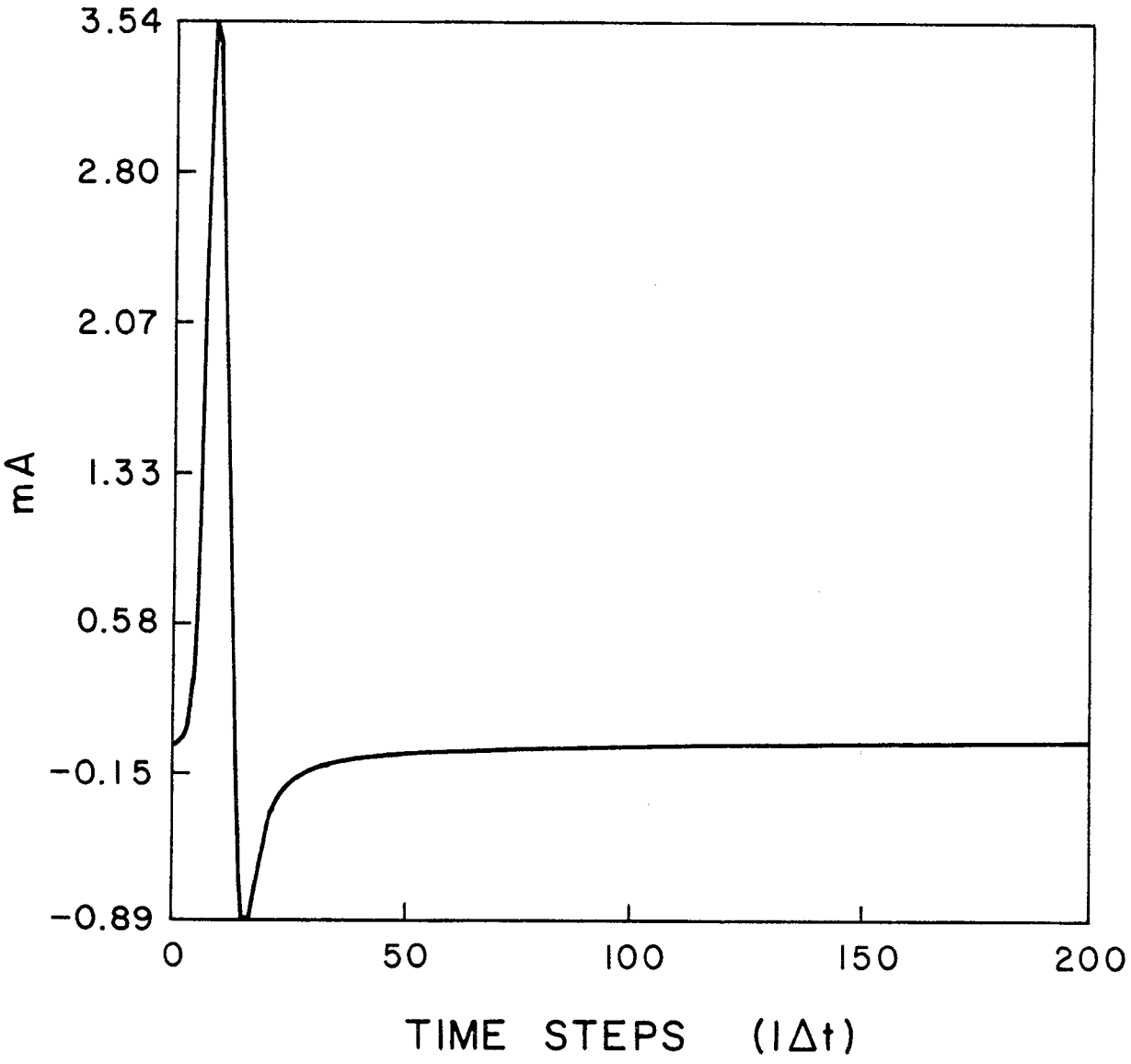
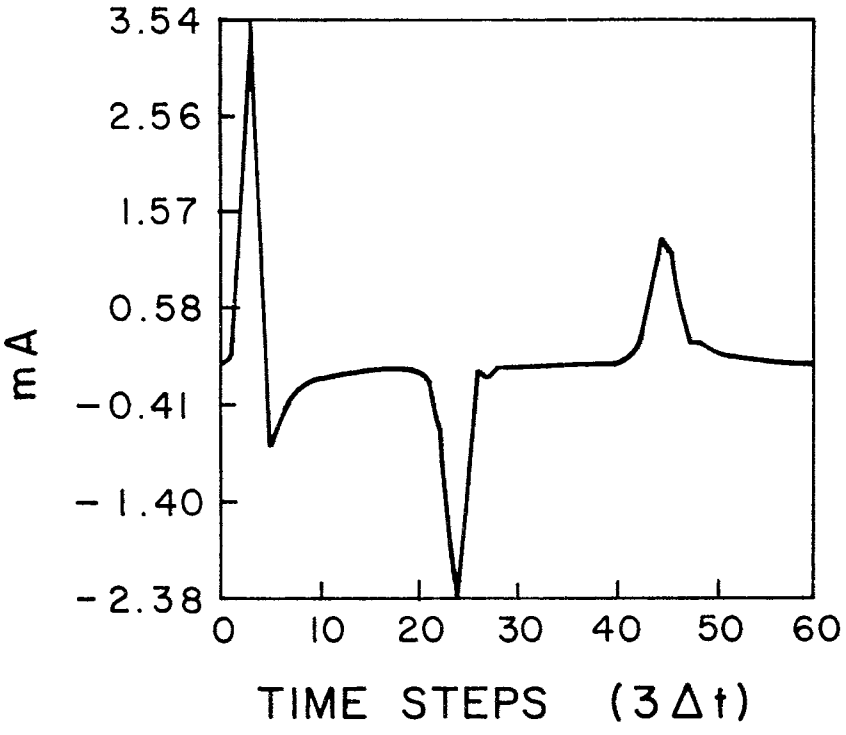
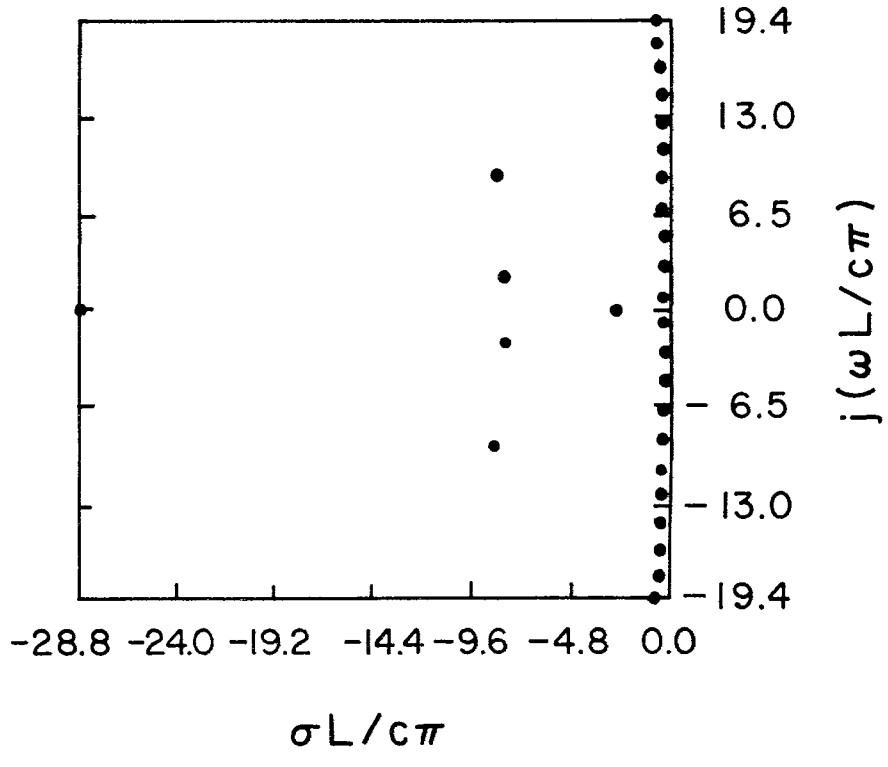


Figure 3.8. Reconstructed transient response using extracted poles of Example 2.



(a)



(b)

Figure 3.9. (a) Data window used in Example 3.
(b) Locations of extracted poles in Example 3.

Figure 3.9b. Note that the true poles appear while the extraneous poles have been reduced in number. The extra poles still must appear because of the presence of the driving function in the transient response. Figure 3.10 points out another problem inherent in the solution of the matrix equations without the use of a least squares technique. Since twenty-eight poles were found, it was necessary to solve for twenty-eight residues using the matrix Equation (3.10). Only the first twenty-eight samples of the current were needed to fill the vector F of (3.10a). Thus, the resulting response is expected to be accurate at only the first twenty-eight time steps. This is represented in the plot of Figure 3.10. The obvious correction to this problem is to use more samples to fill the vector F by using the least squares method. An example of this approach will be presented later in this section.

Example 4

The Nyquist criterion [26] states that the sampling rate of a transient response must be at least twice as fast as the highest frequency present in the data in order to be able to resolve that frequency. That is, the sampling rate Δt_N is

$$\Delta t_N \geq \frac{1}{2\delta_H}$$

where δ_H is the highest frequency desired. In the previous example the time step size was $3\Delta t = 1.6668 \times 10^{-10}$ s. Hence, the highest frequency which could be expected was $\delta_H = 2.9998 \times 10^9$ Hz or $\frac{\omega L}{c\pi} = 19.9984$ which is indeed just higher than the highest frequency pole obtained. Figures 3.11 and 3.12 also demonstrate this phenomenon. In Figure 3.11a the time step size is $4\Delta t = 2.2224 \times 10^{-10}$ s and the resulting frequency

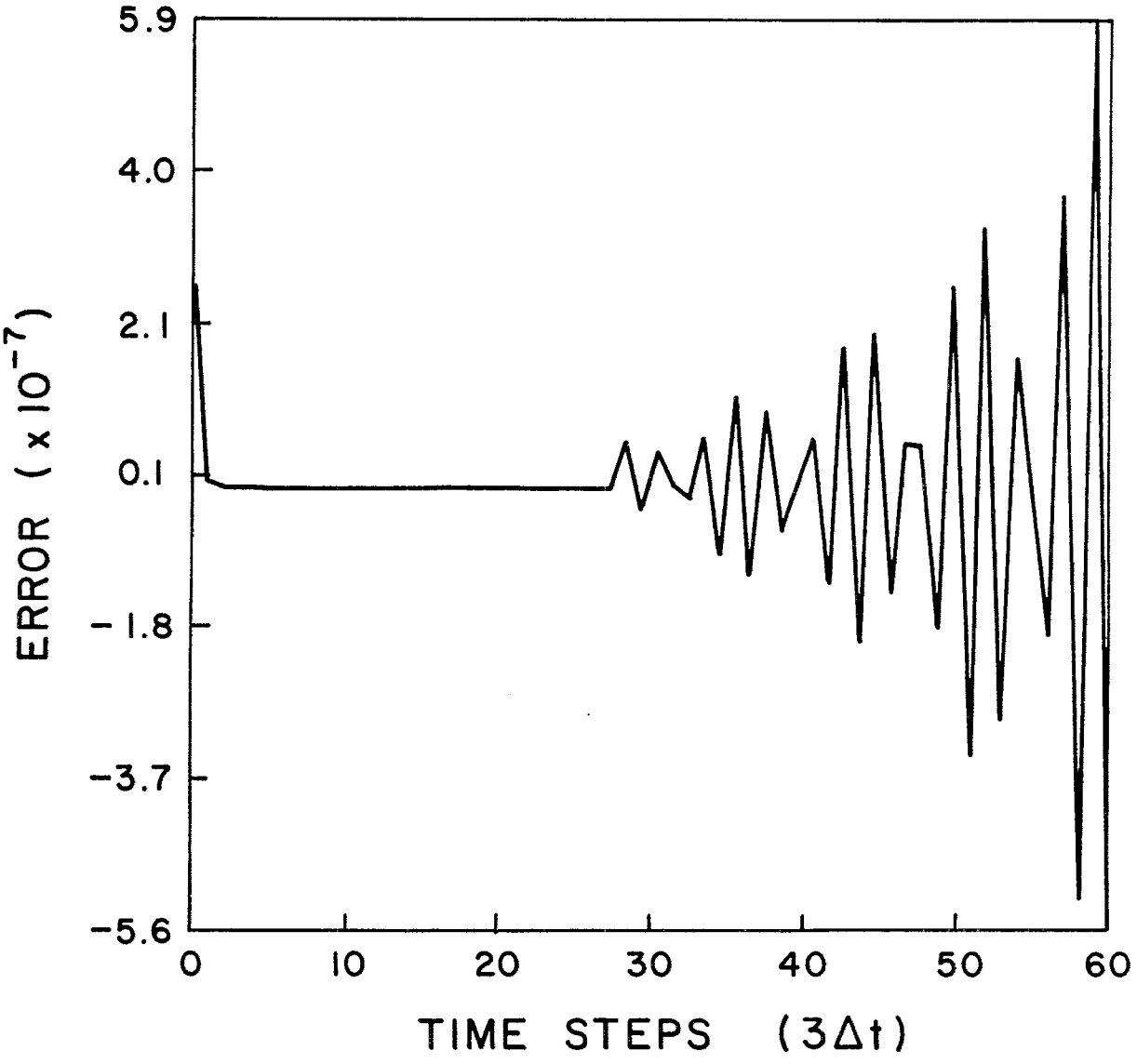
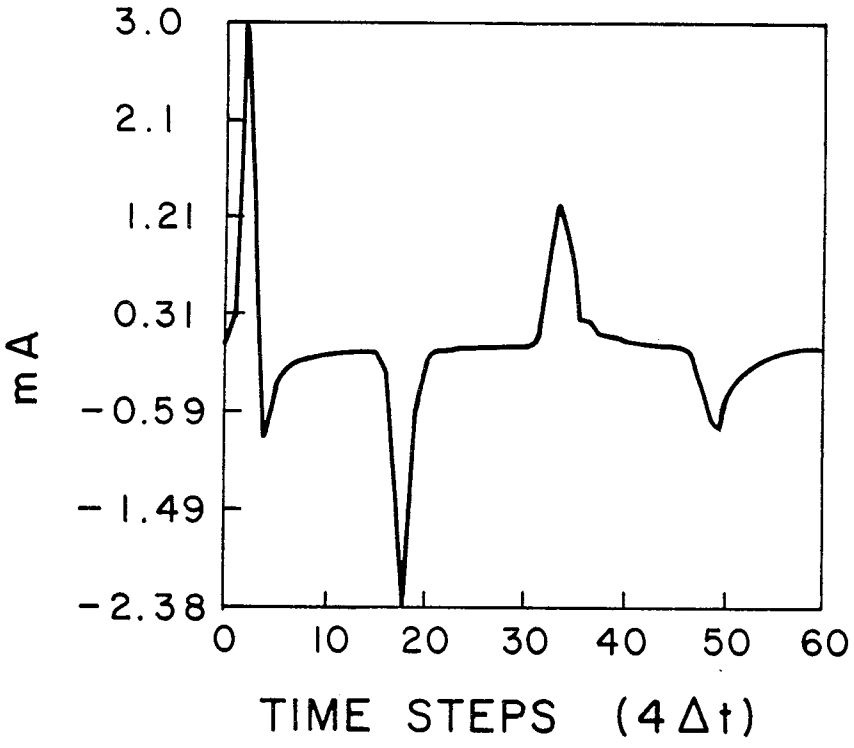
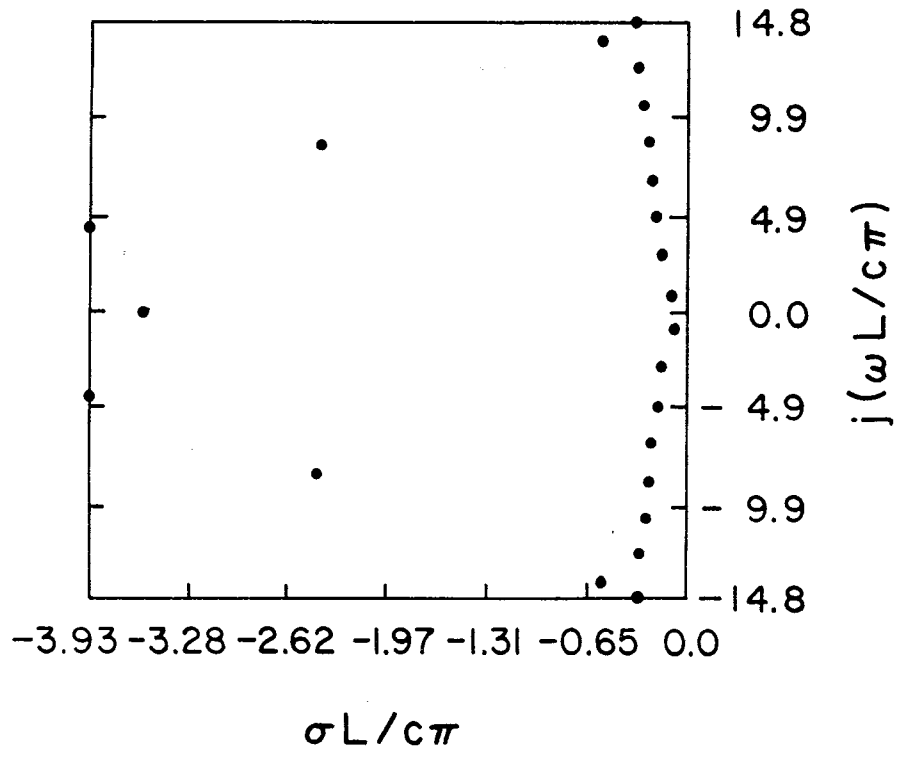


Figure 3.10. Original response minus the reproduced response of Example 3.



(a)



(b)

Figure 3.11. (a) Data window used when the time step is $4\Delta t$.
(b) Locations of the resulting poles.

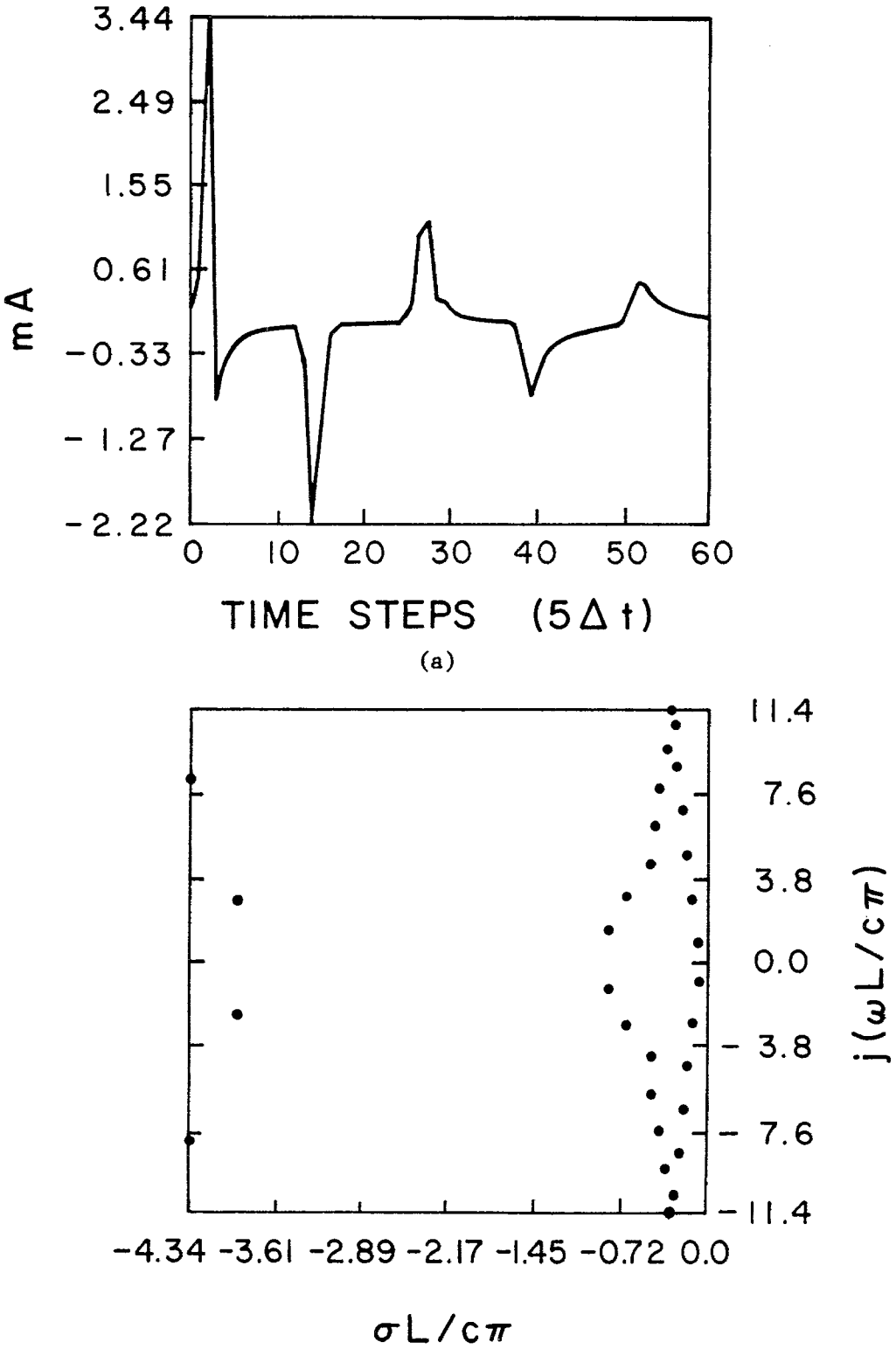


Figure 3.12. (a) Data window used when the time step is $5\Delta t$.
(b) Locations of the resulting poles.

limit is $\omega L/c\pi = 14.99$. The highest pole frequency obtained is 14.80. Likewise in Figure 3.12a the time step is $5\Delta t = 2.778 \times 10^{-10}$ s giving a frequency limit of 11.999 and the highest frequency pole obtained is 11.4. Thus, in order to obtain the highest frequency pole that is contained in the data one must sample the transient response at at least the Nyquist rate.

In all of the previous examples no attempt was made to remove the influence of the driving function, the Gaussian pulse, from the response function. Thus, poles have been present which were required to fit the g(t) term of Expression (2.8). Since the driving function is a Gaussian pulse, it cannot be modeled exactly with a finite number of poles but can be assumed to go to approximately zero for some value of time t_0. Hence, if the response function is sampled after time t_0, the g(t) term should be zero. Also, since the analytical form of the spectrum of the Gaussian pulse is known, it can be easily removed by deconvolution. The following examples demonstrate the result of using the above two approaches for removing the influence of the driving function from the response function.

Example 5

Figure 3.13a shows the response function starting at time step number sixty-one and taking sixty samples at every third time step. The starting point of $t_0 = 61\Delta t$ was selected for two reasons. First of all, by the sixty-first time step the Gaussian driving function value is well below the zero value of the computer. Also, as was mentioned earlier, the true oscillatory response of the structure does not begin until the current has propagated to the end of the structure and returned to the center at time t_0. The sampling interval of $3\Delta t$ was chosen so as to

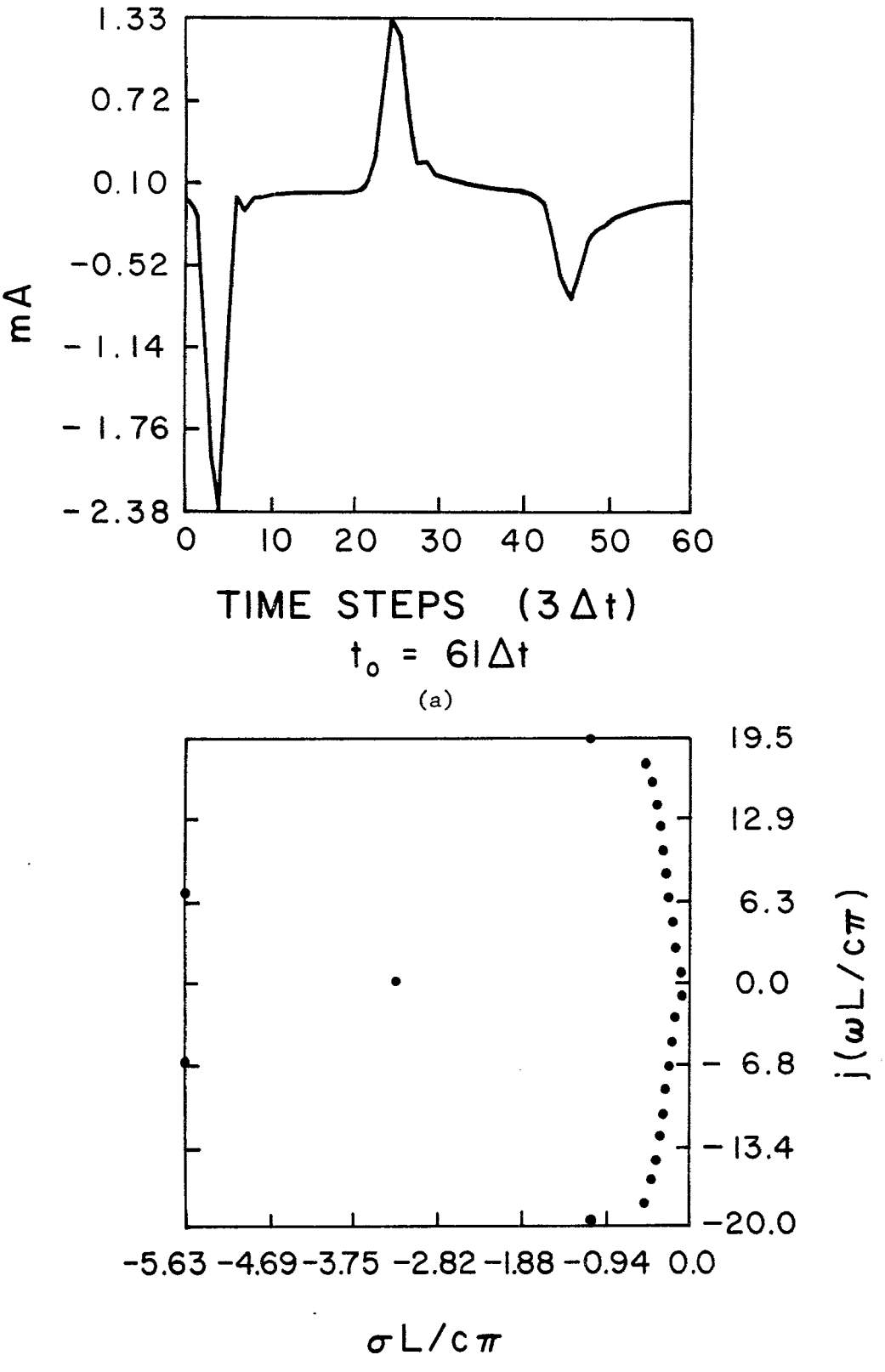
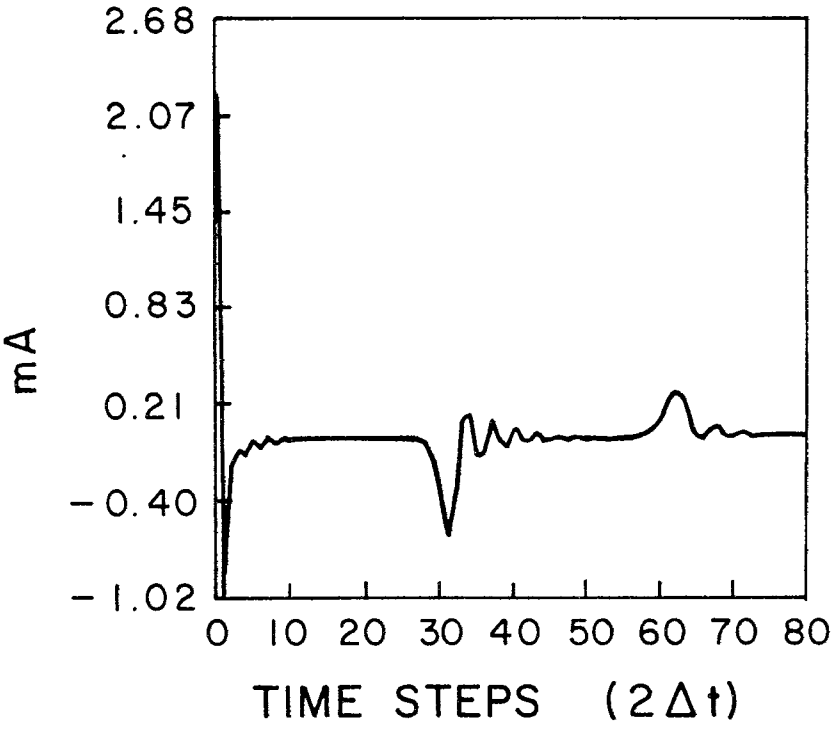


Figure 3.13. (a) Data window starting at time step sixty-one, Example 5.
(b) Resulting pole locations.

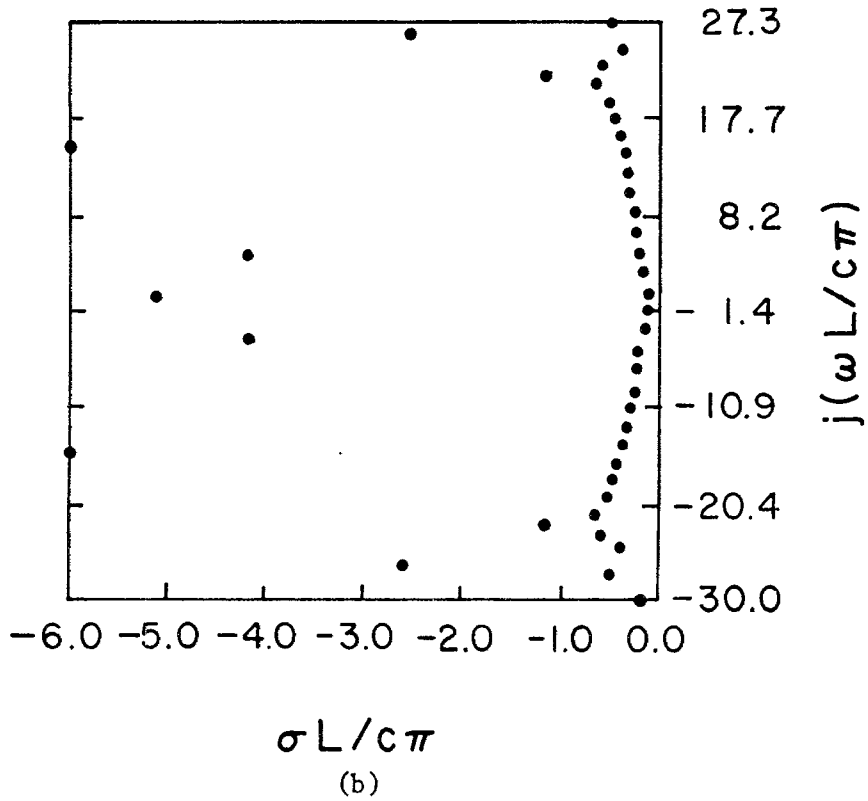
satisfy the Nyquist criterion. Since sixty samples were used, the method attempted to return thirty poles. As can be seen from Figure 3.13b, only twenty-six poles were actually recovered, four, in the right half plane, were thrown away. Of the twenty-six returned, only twenty-two are the true poles of the system. Thus, even though the driving function was removed, extraneous poles are still present due to the fact that more than the twenty-two true poles were required from the matrix inversion process. This would suggest that the best procedure would be to use a least squares approach and search for only twenty-two poles.

Example 6

The deconvolution approach is demonstrated in Figure 3.14. The 512 samples of the transient response have been transformed to the spectral domain using a fast Fourier transform routine. This spectral response was then divided by the spectral response of the Gaussian pulse and the result was transformed back to the time domain. Figure 3.14a shows the first eighty samples of the deconvolved response taken at a time step size of $2\Delta t$. Several things can be noticed from this figure. First of all, the very large spike at about $t = 0$ is the delta function which is always present in the impulse response. Also, high frequency ringing is present in the response which is not in the original transient response. This high frequency ringing is due to Gibbs' phenomenon which is inherent in the deconvolution process. The forty poles that resulted from these eighty time samples are plotted in Figure 3.14b. This pole pattern is very similar to that of Figure 3.4. The one major difference is apparent in Figure 3.15, which is a three-dimensional plot of the pole pattern showing the amplitude. Note that the amplitude of the true poles does not taper off with high frequency as it does in Figure 3.6 because the dependence



(a)



(b)

Figure 3.14. (a) Deconvolved transient response data of Example 6.
(b) Locations of resulting poles.

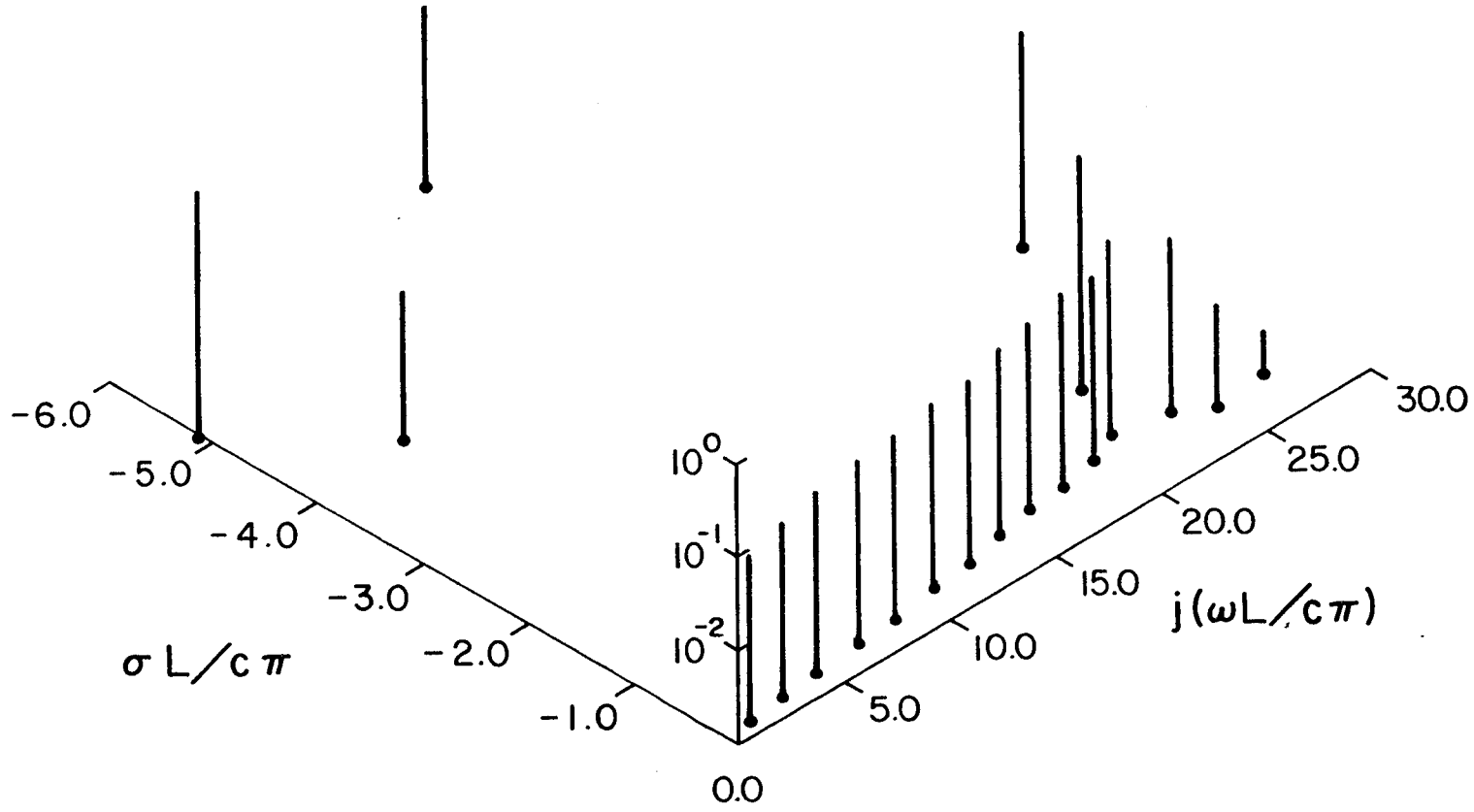


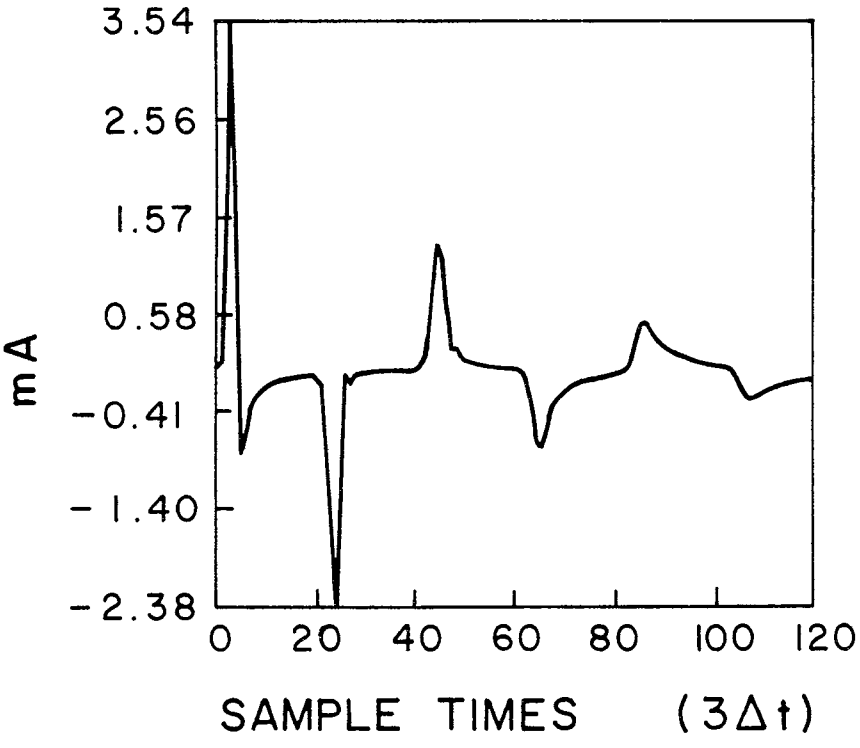
Figure 3.15. Position and amplitude of poles in the second quadrant of the s-plane for Example 6.

on the driving function has been removed by the deconvolution process. The extraneous poles still have fairly large amplitudes because the delta function at the origin was included. There are several disadvantages to performing a deconvolution to remove the dependence of the driving function. The main disadvantage is the large amount of computation time needed. Several stages of calculations must be performed which cost time and accuracy and cause errors such as the Gibb's phenomenon. In general, the deconvolution process should be avoided if at all possible.

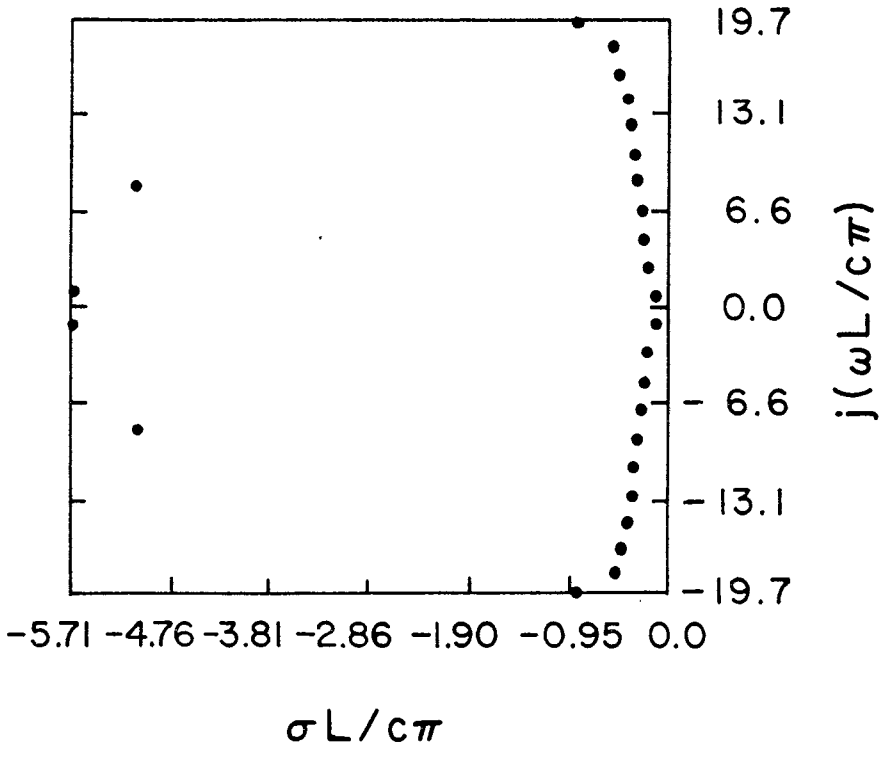
In all of the above examples the extraneous poles were present because the number of time samples that were used dictated the number of poles to be determined. Also in the above examples, errors in the reproduced signals resulted because in calculating the N residues only N time samples were used. These two problems can be overcome by using the least-squares error approach described in Section 3.1. This approach allows one to use many time samples to determine just a few poles thus permitting one to select beforehand the number of poles to be determined. Examples 7 and 8 demonstrate the advantages of the least-squares or pseudo-inverse solution over the conventional procedure.

Example 7

Figure 3.16a is a plot of the sampling interval which is used in this example. Note that 120 samples were used at a sampling rate equal to the Nyquist rate for the problem, $3\Delta t$. The first twenty samples in the interval include the driving function so the resulting poles must be expected to include poles representing the driving function. For this example only twenty-six poles were asked for, thus allowing for the twenty-two poles which were known to be the true poles of the system and



(a)



(b)

Figure 3.16. (a) Data window for Example 7.
(b) Locations of resulting poles.

for the four extra poles needed to represent the driving function's contribution. Figure 3.16b is a plot of these poles and, indeed, the eleven true pole pairs, which were expected, appeared along with two extra pole pairs. The interesting point here is that although only twenty-six poles and residues were extracted they were determined using a total of 120 samples as opposed to fifty-two samples which would have been used if the conventional approach had been employed. Figure 3.17 demonstrates how the error is equally distributed over the entire range of the 120 samples as compared to the error of Figure 3.10 which is near zero for the first twenty-eight samples and then increases for later time.

Example 8

In this example, Figure 3.18, the first sixty time samples have been neglected. Thus, the g(t) term is eliminated from the response function and only the true poles are expected. A total of 150 samples were taken at a rate of $3\Delta t$ and only twenty poles were sought. Figure 3.18a shows that the resulting poles are the true poles of the system. Note from Figure 3.18a that as frequency increases the pole pattern diverges from a sweeping curve. This indicates that there is more error in the higher frequency poles due to the lower signal level at those frequencies. Note that the frequency ω stays stable and that the real part of the pole σ is most sensitive to the noise.

Example 9

In all of the previous examples transient data, which were generated using a time-domain computer, were used. In this example actual experimental data are studied. The experimental data were generated on the transient electromagnetic measurement range at Lawrence Livermore Laboratory [11]. The response is that of a 1.0 m monopole located on

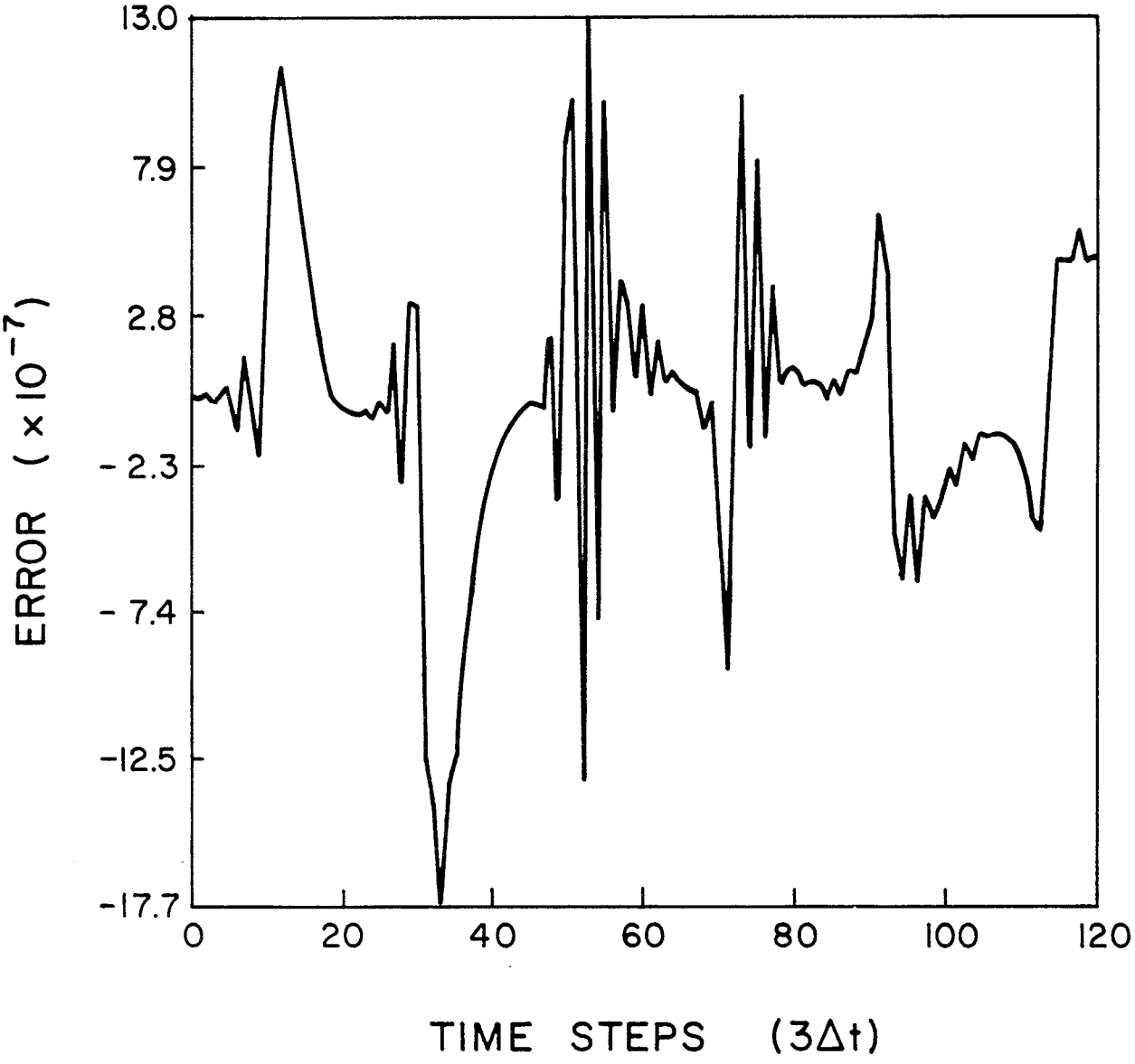
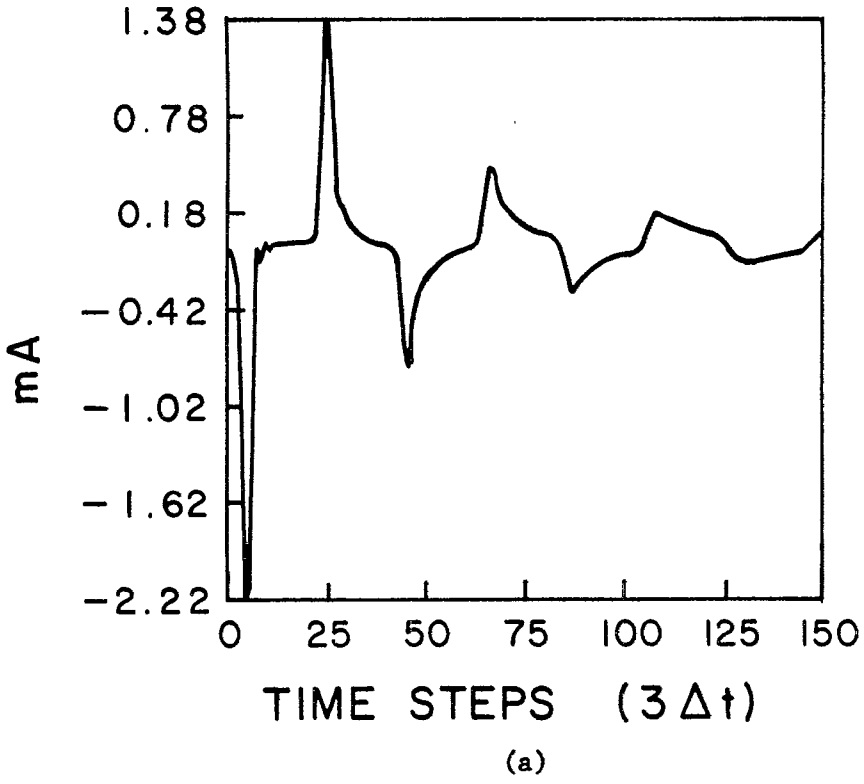
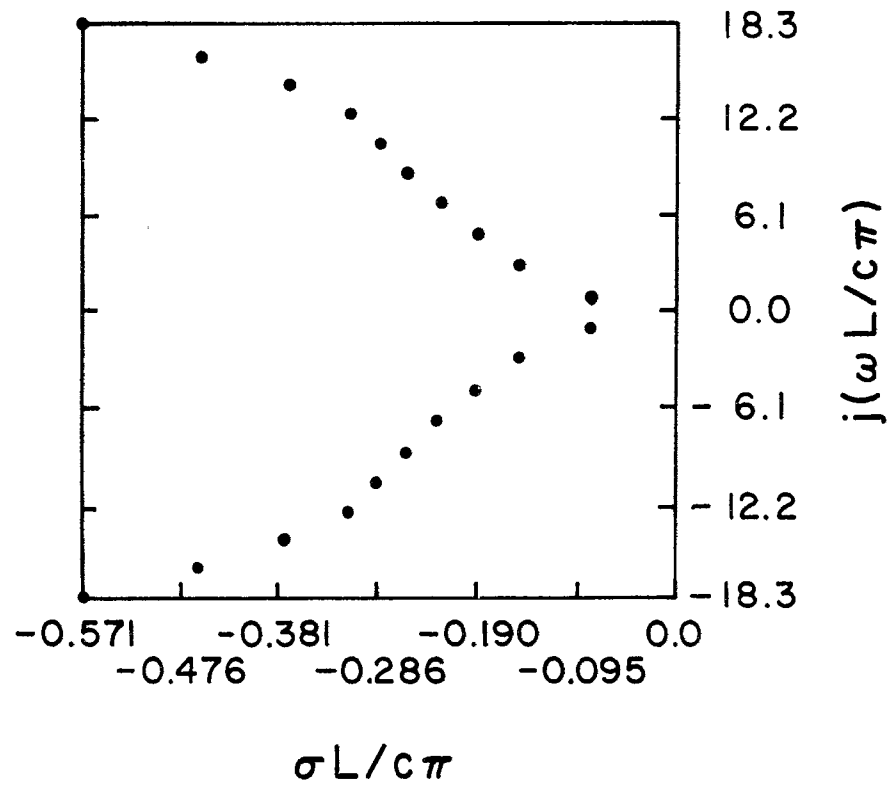


Figure 3.17. Original response minus the reconstructed response of Example 7.



(a)



(b)

Figure 3.18. (a) Data window for Example 8 starting at time step sixty-one.
(b) Locations of resulting poles.

a ground plane and excited by an approximation to a Gaussian-pulse plane wave. The diameter of the monopole is 0.3175 cm. The antenna is loaded at its base with a 50 Ω load and the voltage across this load was measured with a sampling oscilloscope. A total of 512 samples were measured at a time interval of $\Delta t = 0.4 \times 10^{-10}$ s. Of the 512 measured values only 100 samples at every fifth-time step were used. Figure 3.19 shows this measured response in terms of the current through the load. The 100 current samples were used with Prony's method and forty-one poles and residues were produced. Nine poles, which were in the right half plane, were ignored. Figure 3.20 is a plot of the generated poles found in the second quadrant of the complex plane. The first thing that is apparent about these poles is that they fall along a curve running parallel to the imaginary axis. This is typical of the pole locations for a dipole as seen from the previous examples. The frequencies of the first nine poles in this layer correspond to the first nine complex resonant frequencies of a 1 m monopole. The real parts of these poles seem to oscillate around the correct value because of the sensitivity of the real part to the noise in the data. Even though the response was measured on a sampling scope, the data were very noisy and no attempt was made at smoothing.

The fact that the remaining poles do not correspond to physical poles again relates to the fact that the pulse used did not contain frequency components higher than that of the ninth resonance. Also, since no attempt was made to remove the contribution of the driving function, these extra poles are needed to model that portion of the response. The pole sitting on the real axis close to the origin is probably due to the fact that there was a late time dc level present due to the pulser used in the measurement system.

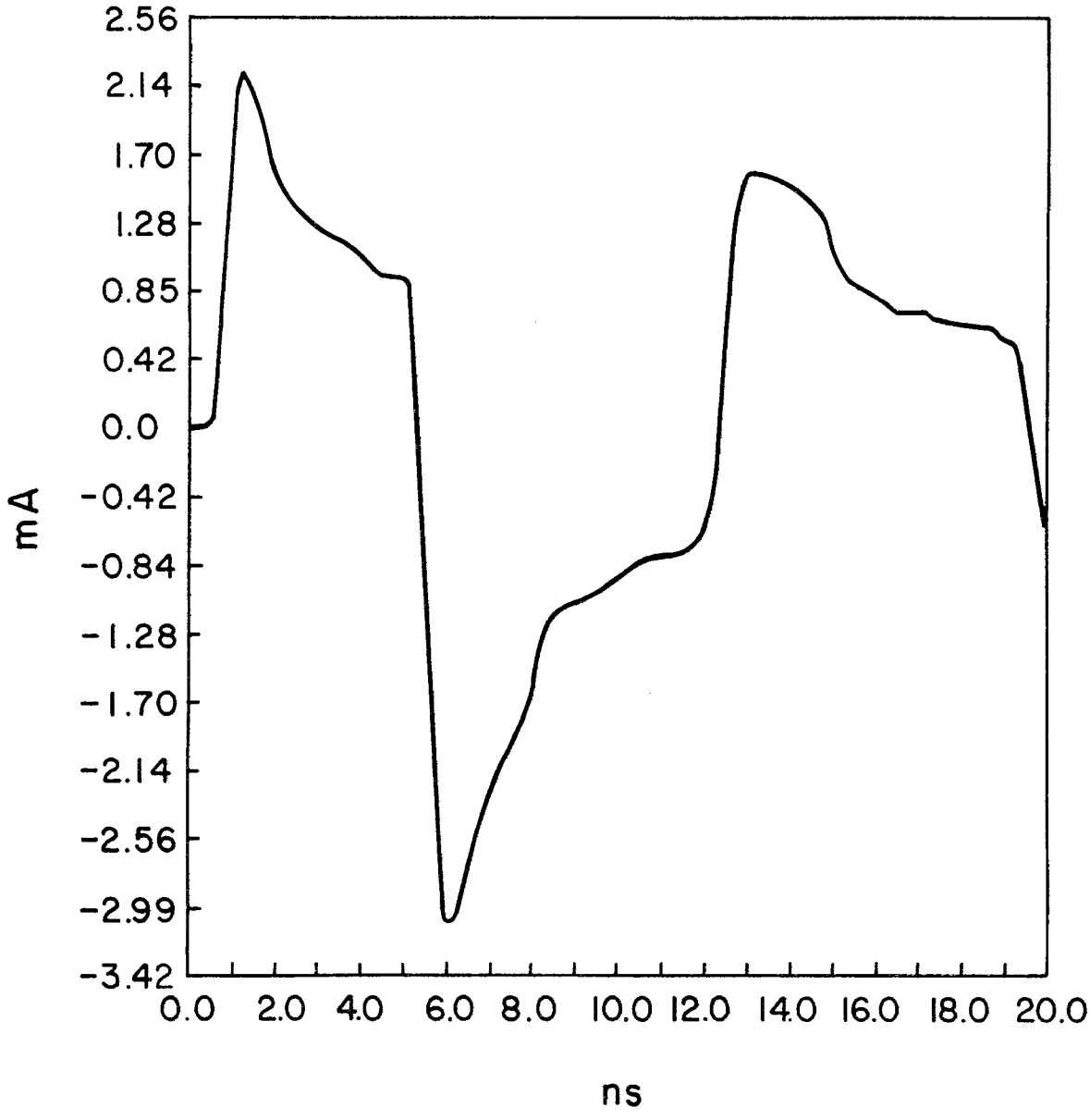


Figure 3.19. Original experimental data of Example 9.

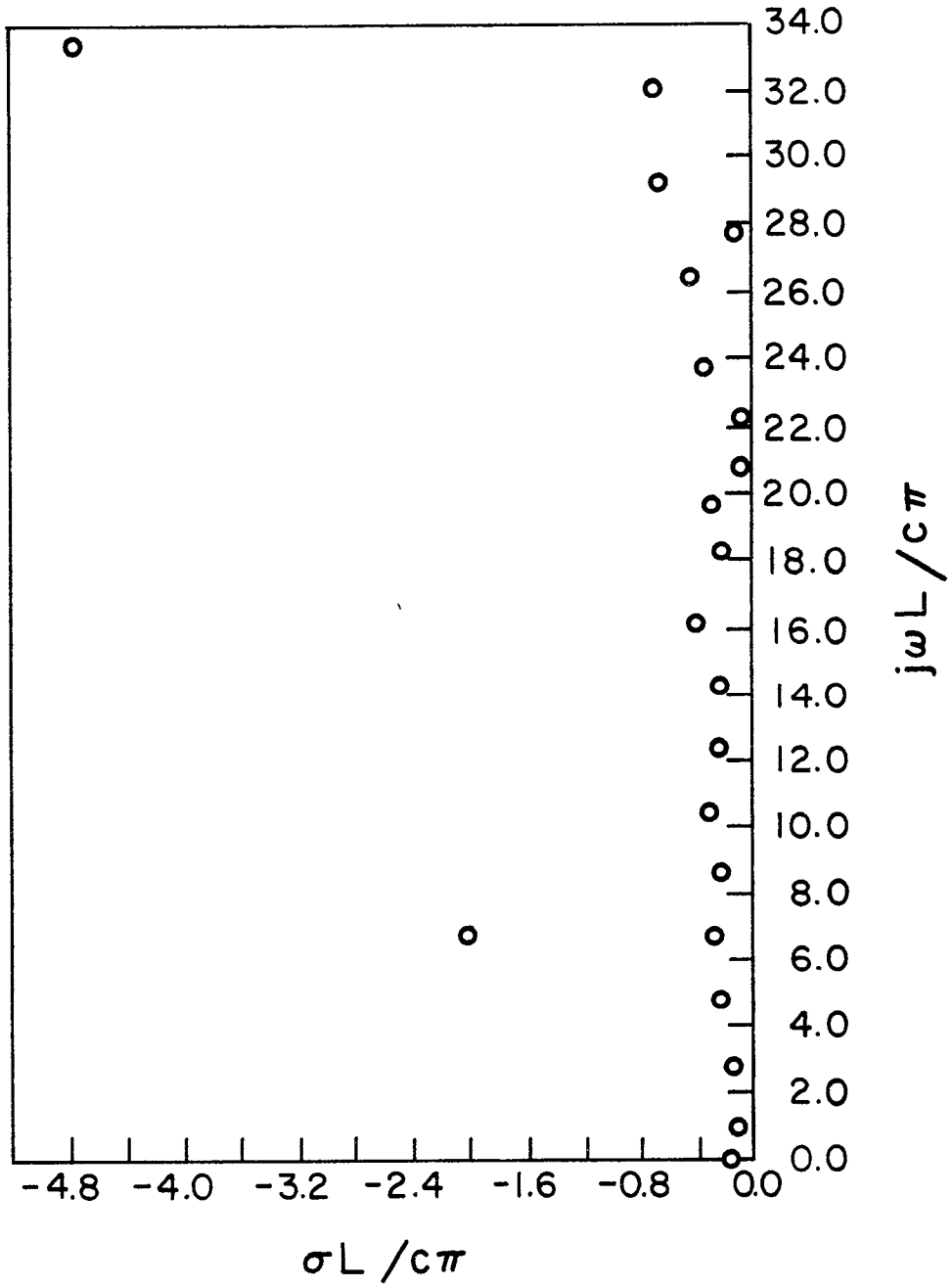


Figure 3.20. Pole locations in complex frequency plane for 1.0 m experimental whip of Example 9.

3.4 Guidelines

The previous section presented nine examples which demonstrate many of the aspects of Prony's method. Contained in those examples are guidelines which should be carefully followed when using the method. Since the guidelines were somewhat obscured by the numerous examples, they will be repeated in this section as a summary for ready reference.

3.4.1 Sampling interval

The width of the sampling interval should be selected so that it is wide enough to contain all of the characteristics of the response. For the most part, the interval should be at least as long as one period of the lowest frequency present. If a least squares method is employed, then the length of the sampling interval should be as long as is economically feasible.

3.4.2 Sampling rate

All physically obtainable transient responses will be bandlimited and thus the upper frequency limit should be determinable. The sampling rate should then be, by the Nyquist criterion, slightly more than twice the highest frequency expected.

3.4.3 Removal of the influence of the driving function

The influence of the driving function on the response function should be removed or accounted for by using one of the processes described in Section 2.2. The two easiest processes to use are to either use a driving function that can be represented by known poles or to use a driving function which is time limited. If a time-limited driving function is used, sufficient response samples must occur after the turnoff time of

the driving function so that the required number of poles may be determined. In an experimental system where this may not be practical, the driving function should be representable by a sum of sinusoids. Deconvolution should be avoided because of the expense and the increase in the error level.

3.4.4 Least squares versus conventional matrix inversion

The least-squared error or pseudo-inverse approach should always be used. This allows as many samples to be used as one desires and allows for the selection of the number of poles without being dependent on the number of samples used. This method gives better results in calculating the residues since more input samples can be used.

3.5 Problems Associated with Prony's Method

There are two major problems associated with Prony's method which need to be overcome in order for the method to be practical. The first problem is the fact that it is necessary to know the number of poles N which are contained in the transient data before Prony's algorithm can be applied. The second problem is that Prony's method is extremely sensitive to noise of any kind. These problems are discussed here and solutions to these problems are presented in the next two chapters.

The numerical examples of Section 3.3 indicate that if one asks for more poles than are effectively contained in the data then the algorithm generates a number of extraneous poles in addition to the ones that compare to the true poles. The presence of the extraneous poles causes the residues of the true poles to be inaccurate and results in unnecessary computation time. Similarly, if one underestimates the number of poles, then many of the returned poles may substantially

deviate from the true poles and, in most cases, the true poles will not be returned at all. Thus, a systematic approach for an a priori determination of N would be extremely beneficial. The examples showed that if the upper frequency limit is known and if the approximate values of the poles are known then one can make a good estimate of the value of N. Also N could be determined by trial and error but that would be very expensive. The next chapter presents two straightforward and systematic schemes for determining N.

The other problem with Prony's method is the sensitivity of the poles to a noisy set of data. The real part of the poles is extremely sensitive to noise. If the noise is too bad, then Prony's method will not return any of the true poles and will attempt to curve-fit to the noise. Therefore, it is necessary to determine the noise limitations for Prony's algorithm and to determine ways of preventing and reducing noise. This problem is handled extensively in Chapter 5.

4. THE DETERMINATION OF THE NUMBER OF POLES (N)

In the last chapter it became obvious that in order to use Prony's method optimally a systematic approach must be developed for determining the number of poles contained in the transient response. This chapter develops two techniques by which the number of poles inherent in the system can be analytically determined. The first approach discussed was developed by Householder [27]. His derivation is presented here except that it is applied to the least squares solution. The second approach is new and is shown to have several advantages over Householder's method.

4.1 Householder Orthogonalization Method

It was shown in Chapter 3 that if a set of discrete transient response data could be represented as a sum of N complex exponentials then the discrete samples I_i must satisfy a difference equation of order N. Consider now the following vectors which are filled with the data samples I_i

$$i_0 = \begin{bmatrix} I_0 \\ I_1 \\ I_2 \\ \cdot \\ \cdot \\ \cdot \\ I_\gamma \end{bmatrix}, \quad i_1 = \begin{bmatrix} I_1 \\ I_2 \\ I_3 \\ \cdot \\ \cdot \\ \cdot \\ I_{\gamma+1} \end{bmatrix}, \quad \ldots, \quad i_i = \begin{bmatrix} I_i \\ I_{i+1} \\ \\ \cdot \\ \cdot \\ \cdot \\ I_{\gamma+i} \end{bmatrix} \tag{4.1}$$

where the i^{th} vector is the $i - 1^{th}$ column of matrix (3.8b) in Chapter 3. Since the I_i satisfy Prony's difference equation of order N, (3.5b), then so must the vectors i_i. Now if $\gamma \geq N$, then any N, or fewer, of the vectors i_i will be linearly independent and any N + 1 of the vectors will be linearly dependent. Hence, the difference equation can be written

$$i_0 \alpha_0 + i_1 \alpha_1 + \ldots + i_N \alpha_N = 0 \quad . \tag{4.2}$$

The next step is to apply the Gram Schmidt or Choleski orthogonalization process to the vectors i_i.

The orthogonalization process consists of replacing each vector i_i by a linear combination of the vectors $i_0, i_1, \ldots, i_i$. The new orthogonal vectors are denoted as $a_0, a_1, \ldots, a_i$. The steps are then

$$a_0 = i_0$$

$$a_1 = i_1 - \mu_{10} a_0$$

where μ_{10} is chosen so as to make a_1 and a_0 orthogonal. That is,

$$\mu_{10} = \frac{i_1 a_0^T}{a_0 a_0^T}$$

where a_0^T is the transpose of vector a_0. Then proceeding to the second vector

$$a_2 = i_2 - \mu_{21} a_1 - \mu_{20} a_0$$

where again μ_{21} and μ_{20} are selected so that a_2 is orthogonal to both a_0 and a_1, which means that

$$\mu_{20} = \frac{a_0^T i_2}{i_0^T i_0}$$

$$\mu_{21} = \frac{a_1^T i_2}{i_1^T i_1} \quad .$$

Continuing this process, one finally obtains

$$a_{N-1} = i_{N-1} - \mu_{N-1,N-2}\, a_{N-2} + \cdots + \mu_{N-1,0}\, a_0$$

where

$$\mu_{N-1,N-2} = \frac{a_{N-2}^T i_{N-1}}{a_{N-2}^T a_{N-2}}$$

$$\cdots$$

$$\mu_{N-1,0} = \frac{a_0^T i_{N-1}}{a_0^T a_0} \quad .$$

Let V_i, A_i and M_i designate the matrices

$$V_i = (i_0, i_1, \ldots, i_i) \tag{4.3}$$

$$A_i = (a_0, a_1, \ldots, a_i) \tag{4.4}$$

$$M_i = \begin{bmatrix} 1 & \mu_{10} & \mu_{20} & \cdots & \mu_{i0} \\ 0 & 1 & \mu_{21} & & \mu_{i,1} \\ & & 1 & & \\ \cdot & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot \\ 0 & & & & 1 \end{bmatrix} \quad . \tag{4.5}$$

Then,

$$V_i = A_i M_i \tag{4.6}$$

which is the Gram-Schmidt procedure in matrix form. Since M_i is a triangular matrix, it is easily inverted so that

$$A_i = V_i M_i^{-1} \quad . \tag{4.7}$$

The difference Equation (4.2) can now be written as

$$V_{N-1} \sigma_N = -i_N \tag{4.8}$$

where

$$\sigma_N = \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \cdot \\ \cdot \\ \cdot \\ \alpha_{N-1} \end{bmatrix} \quad . \tag{4.9}$$

Equation (4.9) is equivalent to Equation (3.6) of Chapter 3. Now if a pseudo-inverse approach is used to solve Equation (4.8) then the resulting expression is

$$V_{N-1}^T V_{N-1} \sigma_N = V_{N-1}^T(-i_n) \quad . \tag{4.10}$$

From (4.6)

$$V_{N-1} = A_{N-1} M_{N-1} \tag{4.11a}$$

$$V_{N-1}^T = M_{N-1}^T A_{N-1} \tag{4.11b}$$

so that (4.10) can be written as

$$M_{N-1}^T A_{N-1}^T A_{N-1} M_{N-1} \sigma_N = M_{N-1}^T A_{N-1}^T (-i_N) \quad . \tag{4.12}$$

Let

$$A_{N-1}^T A_{N-1} = D_{N-1} \tag{4.13}$$

where the matrix D_{N-1} is a simple diagonal matrix because A_{N-1} and A_{N-1}^T are made up of orthogonal vectors. Now define the vector

$$\mu_i = \begin{bmatrix} \mu_{i,0} \\ \mu_{i,1} \\ \cdot \\ \cdot \\ \cdot \\ \mu_{i,i-1} \end{bmatrix} \tag{4.14}$$

so that the orthogonalization process can be written

$$i_{i+1} = a_{i+1} + A_i \mu_{i+1} \tag{4.15}$$

or by multiplying both sides by A_i^T

$$A_i^T i_{i+1} = A_i^T a_{i+1} + A_i^T A_i \mu_{i+1} \quad . \tag{4.16}$$

Since all columns of A_i are orthogonal to a_{i+1}, then

$$A_i^T i_{i+1} = D_i \mu_{i+1} \tag{4.17}$$

where D_i has replaced the product $A_i^T A_i$. If Equations (4.17) and (4.13) are substituted in (4.12), then

$$\left[M_{N-1}^T D_{N-1}\right] M_{N-1} \sigma_N = \left[M_{N-1}^T D_{N-1}\right] \mu_N(-1) \quad . \tag{4.18}$$

Since the bracketed terms on each side of (4.18) are nonsingular matrices, they can be removed by multiplying both sides by their inverse. Hence, (4.18) reduces to

$$M_{N-1} \sigma_N = -\mu_N \quad . \tag{4.19}$$

To summarize, the μ_{i+1} and a_{i+1} are calculated sequentially by using the Gram-Schmidt procedure, Equations (4.15) and (4.17). The vector a_{i+1} is adjoined to A_i to obtain A_{i+1} and matrix M_i is bordered by μ_{i+1} and a unit row vector to obtain M_{i+1}. Note that each vector a_{i+1} is the component of i_{i+1} that is orthogonal to the space of the previous i's. Since the vectors i_i must satisfy a difference equation of order N, then i_N is a linear combination of the preceding i's and a_N should vanish. Thus, in order to determine the value of N, the above process should be continued until one of the vectors a vanishes. However, since the time samples are subject to measurement errors and roundoff errors, it should not be expected that any vector a will vanish entirely but for some N there will be an a_N which will be negligibly small.

Once a negligibly small vector a_N has been found, it is a simple procedure to find the coefficients α of the difference equation. The coefficients α are found from (4.19). Note that the matrix M_{N-1} is a triangular matrix so that the vector σ_N may be found by back substitution, eliminating a matrix inversion. Thus, this process not only gives a method for determining the number of poles N in the system but also

allows one to actually find the N poles without having to invert a matrix. Note also that this method can use as many time samples as desired by selecting the appropriate value of γ in (4.1).

As an example of this procedure, a set of test transient data was produced using the set of six pole pairs listed in Table 4.1. For all the poles the accompanying residues were taken to be one. The resulting transient response is plotted in Figure 4.1. Householder's method was carried out exactly as outlined above. The value of γ which was used was 100. Figure 4.2a is a plot of the average of the absolute value of the first thirteen vectors i_i and the resulting orthogonal vectors a_i. Note that while the average of the first thirteen vectors i_i stays constant the first twelve values of a_i drop off slightly until at the thirteenth vector the value drops by four orders of magnitude. While this thirteenth vector did not go to zero, it did drop low enough to indicate that the thirteenth vector was linearly dependent on the preceding twelve vectors. The difference equation coefficients were calculated as per Equation (4.19) and the resulting poles are listed in Column Two of Table 4.1. Note that the resulting poles are essentially the same as the original.

As another test of the above procedure, the same transient response was used but normally distributed noise with a standard deviation of 0.005 was added to the response. Again γ was set equal to 100 and the resulting vectors a_i and i_i are plotted in Figure 4.2b. Notice now that at the thirteenth vector the average value does not drop off. This indicates that the Householder procedure is not applicable to systems with significant noise. When the poles were found for the noise case,

TABLE 4.1

SIX POLE PAIRS AND THE RECOVERED POLES USING HOUSEHOLDER'S ORTHOGONALIZATION METHOD

True Poles	Recovered Pole No Noise	Recovered Poles σ = 0.005 Noise
- 0.082 + j 0.926	- 0.0819 + j 0.9259	- 0.4953 + j 2.0169
- 0.082 - j 0.926	- 0.0819 - j 0.9259	- 0.4953 - j 2.0169
- 0.147 + j 2.874	- 0.1469 + j 2.8739	- 0.5029 + j 0.8858
- 0.147 - j 2.874	- 0.1469 - j 2.8739	- 0.5029 - j 0.8858
- 0.188 + j 4.835	- 0.1879 + j 4.8350	- 0.4834 + j 1.3269
- 0.188 - j 4.835	- 0.1879 - j 4.8350	- 0.4834 - j 1.3269
- 0.220 + j 6.800	- 0.2200 + j 6.8000	- 0.3817 + j 3.5100
- 0.220 - j 6.800	- 0.2200 - j 6.8000	- 0.3817 - j 3.5100
- 0.247 + j 8.767	- 0.2470 + j 8.7669	- 0.2915 + j 10.7418
- 0.247 - j 8.767	- 0.2470 - j 8.7669	- 0.2915 - j 10.7418
- 0.270 + j 10.733	- 0.2699 + j 10.7330	- 0.4918 + j 0.0
- 0.270 - j 10.733	- 0.2699 - j 10.7330	- 7.8833 + j 15.7079

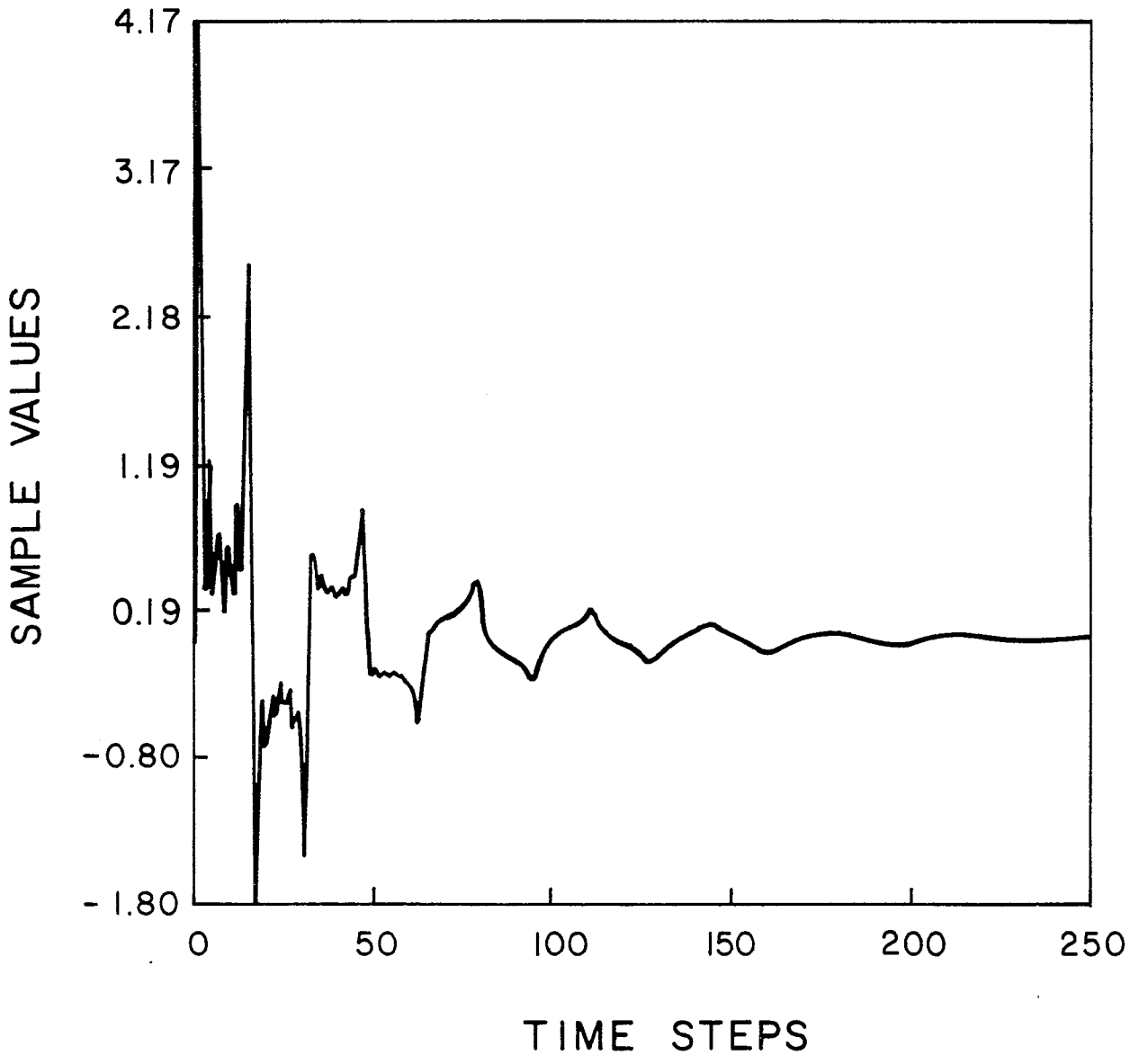
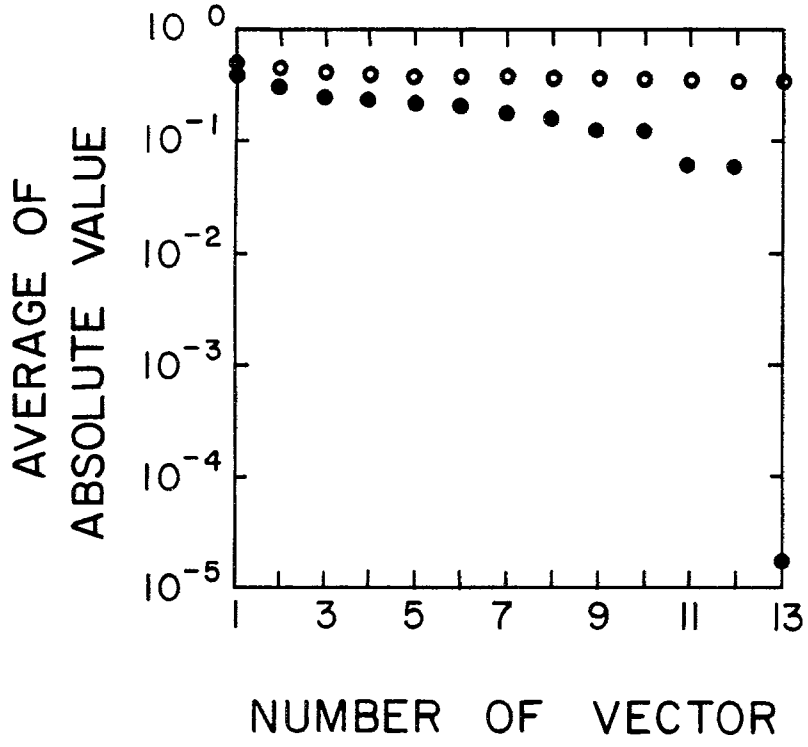
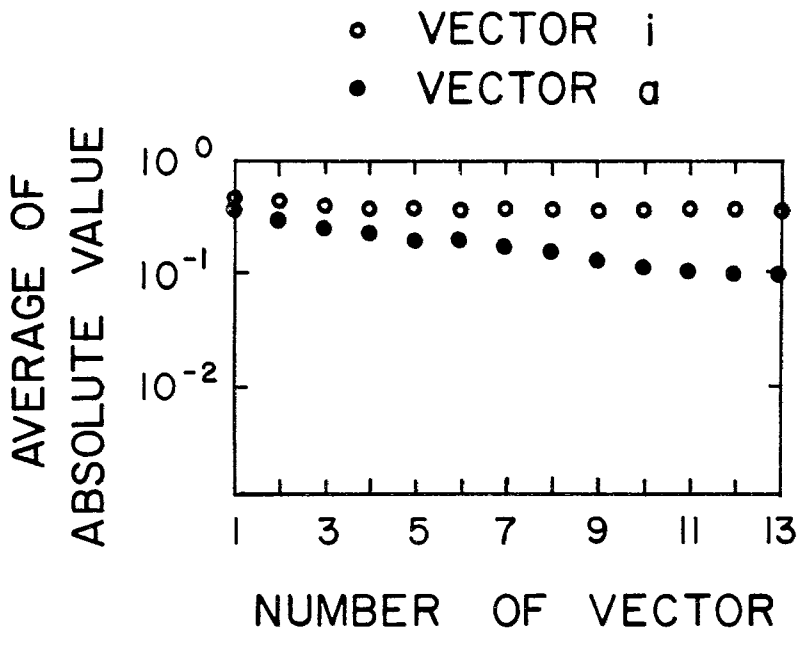


Figure 4.1. Transient response resulting from the six pole pairs of Table 4.1.



(a)



(b)

Figure 4.2. Value of the average of the absolute value of the vectors a and i.

(a) No-noise case.
(b) Noise of σ = 0.005 added to data.

they were not associated with the true poles at all, as indicated by the third column of Table 4.1. Chapter 5 shows that the alternate procedure for finding the poles, which is presented in the next section, does indeed work in the presence of noise.

As a further example the numerically generated transient response of the current on the one meter dipole discussed in Section 3.3 is used. The samples were taken at every third time step starting with time step sixty-one. Figure 4.3a shows a plot of both vectors i and the resulting orthogonal vector a. Note that while vector i has an average which is somewhat constant vector a drops off until vector number twenty-three where it stays level for a while. This is interesting since in the examples of Section 3.3 it was shown that eleven pole pairs appeared to be the optimal solution to this example. Indeed, when the process of (4.19) was applied using the twenty-third orthogonal vector the first eleven poles in Figure 3.4 were obtained. Note that the twenty-third vector a is four orders of magnitude below the twenty-third vector i. In Figure 4.2a of the previous example the thirteenth vector a was also four orders of magnitude below the thirteenth vector i. This implies that when the average of the orthogonal vector drops four orders of magnitude below its accompanying response vector then that vector is the first dependent vector.

In Figure 4.3b the vectors a and i are plotted for the dipole case considered above but since the samples start at time step one, the driving function is included. Note that none of the twenty-nine a vectors are more than one order of magnitude below the vector i. This is expected since the Gaussian pulse included in the first sixty samples cannot be written as a finite sum of exponentials.

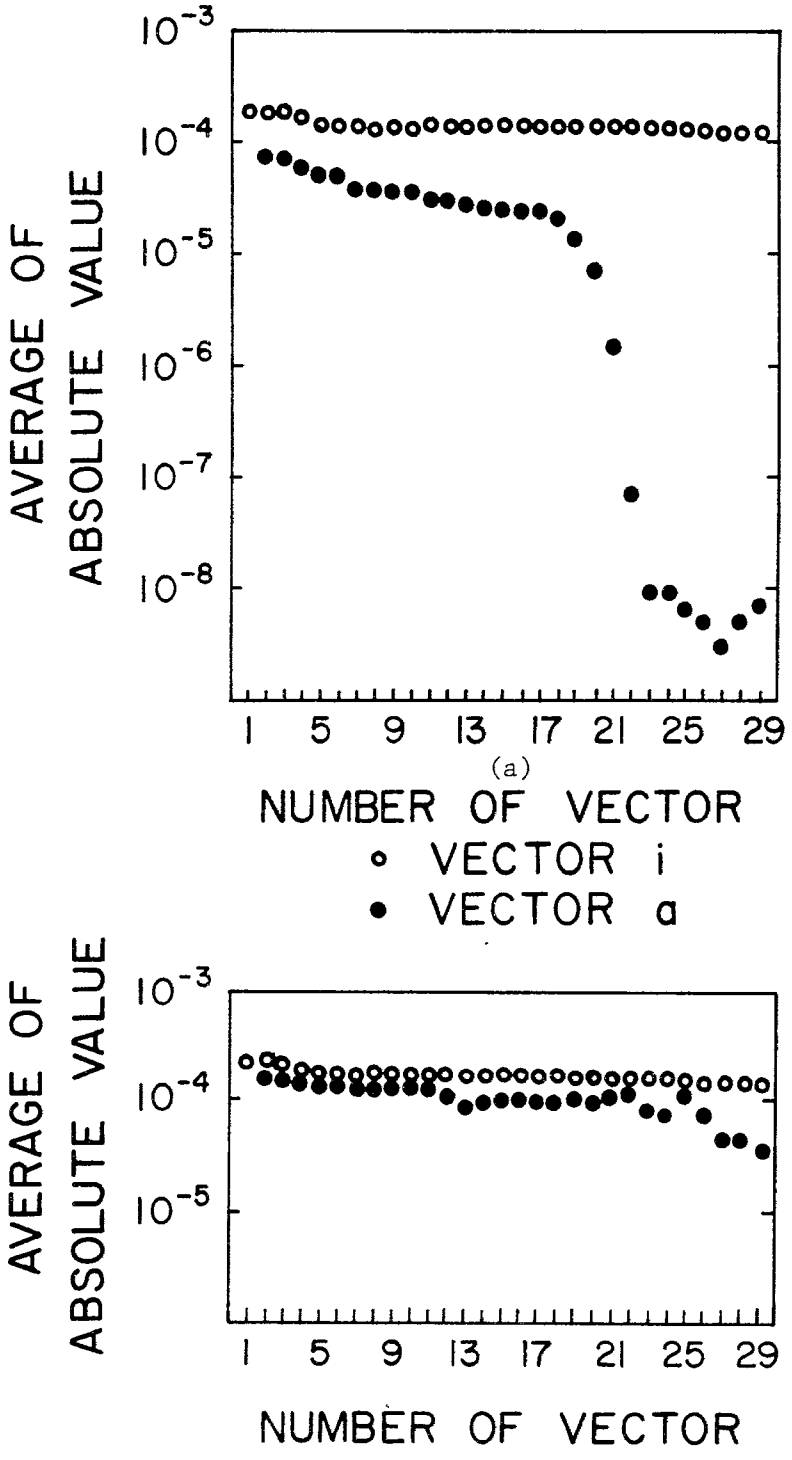


Figure 4.3. Value of the average of the absolute value of the vector a and i for the response of the 1.0 m dipole.

(a) Data window starting at time step sixty-one.
(b) Data window starting at time step one.

4.2 The Eigenvalue Method for Determining N

To start this development, it is again necessary to form the vectors v_i of Section 4.1. Note that the v_i can be written as

$$v_i = \sum_{j=1}^{N} A_j \, e^{s_j n\Delta t} \qquad n = i, i+1, \ldots, i+\gamma \tag{4.20}$$

or by separating terms

$$v_i = \sum_{j=1}^{N} (A_j \, e^{s_j i\Delta t}) \, e^{s_j n\Delta t} \qquad n = 0,1, \ldots \tag{4.21}$$

Expression (4.21) can be written as

$$v_i = \sum_{j=1}^{N} C_{j,i} \, \psi_j \tag{4.22}$$

where

$$C_{j,i} = A_j \, e^{s_j i\Delta t} \tag{4.23}$$

is the coefficient for the j^{th} mode vector ψ_j starting at the i^{th} time step. Remembering that

$$Z_j = e^{s_j \Delta t} \tag{4.24}$$

then the j^{th} mode vector is

$$\psi_j = \begin{bmatrix} Z_j^0 \\ Z_j^1 \\ Z_j^2 \\ \cdot \\ \cdot \\ \cdot \\ Z_j^\gamma \end{bmatrix} = \begin{bmatrix} 1 \\ Z_j \\ Z_j^2 \\ \cdot \\ \cdot \\ \cdot \\ Z_j^\gamma \end{bmatrix} . \tag{4.25}$$

If there are N poles in the system, then there will be N mode vectors ψ_j.

The next step is to solve the difference equation

$$\sum_{m=0}^{N} V_{m+k}\ \alpha_m = 0 \quad . \tag{4.26}$$

If a pseudo-inverse solution is used, then (4.26) would appear in matrix form as

$$\begin{bmatrix} V_0 & V_1 \cdots V_\gamma \\ V_1 & \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ V_N & \gamma_{N+\gamma} \end{bmatrix} \begin{bmatrix} V_0 & V_1 \cdots V_N \\ V_1 & \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ V_\gamma & V_{N+\gamma} \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \cdot \\ \cdot \\ \cdot \\ \alpha_N \end{bmatrix} = 0 \quad . \tag{4.26}$$

Matrix Φ is defined as the product of the above two matrices. Note that the i^{th}, j^{th} element of the Φ matrix is

$$\Phi_{i,j} = v_i^T\ v_j \tag{4.27}$$

where v_i^T is the transpose of vector v_i. Substituting (4.23) into (4.27) gives

$$\Phi_{i,j} = \sum_{\ell=1}^{N} \sum_{m=1}^{N} \left[C^*_{\ell,i}\ C_{\ell,j} \right] \psi_\ell^{T*}\ \psi_m \tag{4.28}$$

where the * indicates complex conjugate. Since Φ is a matrix of order N + 1 by N + 1, it will have one eigenvector which will be orthogonal to the N mode vectors. That is, there will be one eigenvector such that

$$\Phi\ E_{N+1} = 0 \tag{4.29}$$

where E_{N+1} is the eigenvector orthogonal to all of the modal vectors. The eigenvalue corresponding to this eigenvector is zero. If matrix Φ was made to be 2N there would be N eigenvectors orthogonal to the N modal vectors and there would be N eigenvalues equal to zero. Hence, the process for determining the N of the system is to fill matrix Φ to some dimension M by M. The corresponding M eigenvalues of the system are found and checked to see if one or more is equal to zero. If there are L eigenvalues equal to zero, then N would be equal to M - L. If L is not equal to one, then the matrix Φ is recomputed to order N + 1 by N + 1 and the eigenvalues regenerated. Equations (4.26) and (4.29) show that the eigenvector corresponding to the one zero eigenvalue is the vector of the coefficients of the difference equation. Thus, not only are the number of poles found, but the poles are found at the same time.

As the first example of this procedure, consider again the transient response of Figure 4.1 which was generated using the six pole pairs of Column One of Table 4.2. The thirteen eigenvalues of this system, which were generated using a value of $\gamma = 100$, are plotted in Figure 4.4. Notice the very sharp drop between the number twelve and the number thirteen eigenvalue. The twelve poles that resulted from the eigenvector corresponding to the thirteenth eigenvalue are listed in Column Two of Table 4.2. The resulting poles are in good agreement with the original.

As a further test of the above procedure, the same transient data were used but normally distributed noise with a standard deviation of 0.005 was added. The value of γ was set to 100 and the resulting thirteen eigenvalues are plotted in Figure 4.4. Note that the drop

TABLE 4.2

SIX POLE PAIRS AND THE RECOVERED POLES USING THE EIGENVALUE METHOD

True Poles	Recovered Poles No Noise	Recovered Poles σ = 0.005 Noise
- 0.082 + j 0.926	- 0.0819 + j 0.9259	- 0.0964 + j 0.9319
- 0.082 - j 0.926	- 0.0819 - j 0.9259	- 0.0964 - j 0.9319
- 0.147 + j 2.874	- 0.1469 + j 2.8739	- 0.1805 + j 2.8705
- 0.147 - j 2.874	- 0.1469 - j 2.8739	- 0.1805 - j 2.8705
- 0.188 + j 4.835	- 0.1879 + j 4.8350	- 0.1930 + j 4.8085
- 0.188 - j 4.835	- 0.1879 - j 4.8350	- 0.1930 - j 4.8085
- 0.220 + j 6.800	- 0.2200 + j 6.8000	- 0.2103 + j 6.8001
- 0.220 - j 6.800	- 0.2200 - j 6.8000	- 0.2103 - j 6.8001
- 0.247 + j 8.767	- 0.2470 + j 8.7669	- 0.2485 + j 8.7701
- 0.247 - j 8.767	- 0.2470 - j 8.7669	- 0.2485 - j 8.7701
- 0.270 + j 10.733	- 0.2699 + j 10.7330	- 0.2706 + j 10.7334
- 0.270 - j 10.733	- 0.2699 - j 10.7330	- 0.2706 - j 10.7334

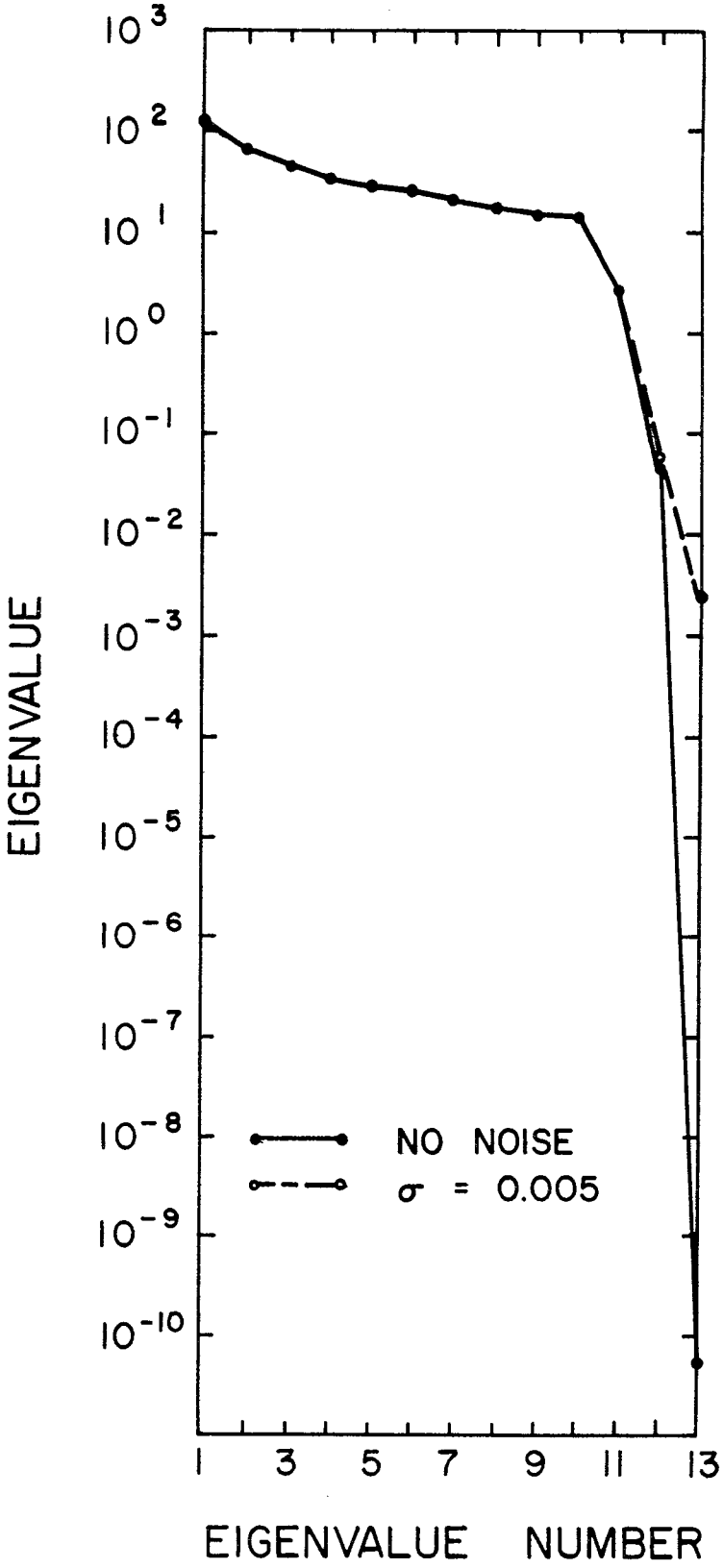


Figure 4.4. Resulting eigenvalues from response generated using poles of Table 4.2 for non-noise case and noise with $\sigma = 0.005$.

between the twelfth and thirteenth eigenvalues is only one order of magnitude, indicating that the noise has a gross effect on this procedure. The poles that were recovered from this noisy case are listed in Column Three of Table 4.2. Although the values of these poles are not precisely the true poles, the maximum percent error is only 1.5. When this fact is compared to the fact that some of the poles of Column Three of Table 4.1 are not even in complex conjugate pairs, the conclusion indicates that this procedure will work with noisy data. The next chapter will show that the eigenvalues are also a function of the noise level.

As a final example of this procedure again consider the transient response data of Figure 4.1. Samples were taken at every third time step starting at time step sixty-one. Tests were run in which twenty-one, twenty-three, and twenty-five eigenvalues were determined and the corresponding poles were extracted. Figure 4.5 shows plots of the twenty-three eigenvalues and twenty-five eigenvalues and Figure 4.6 plots the poles for the twenty, twenty-two, and twenty-four pole cases. For the case of twenty-two poles the poles correspond to the first eleven poles plotted in Figure 3.4, while for the case of twenty poles the upper frequency poles vary significantly. The case where twenty-four poles were evaluated gives two poles that do not form complex conjugate pairs and the eleventh pole pair does not correspond to the eleventh pole pair of the twenty-two pole example. This indicates that the case where twenty-three eigenvalues were generated is the proper solution and should be studied. Notice that in the plot of the twenty-three eigenvalues the eigenvalues drop off continuously until the

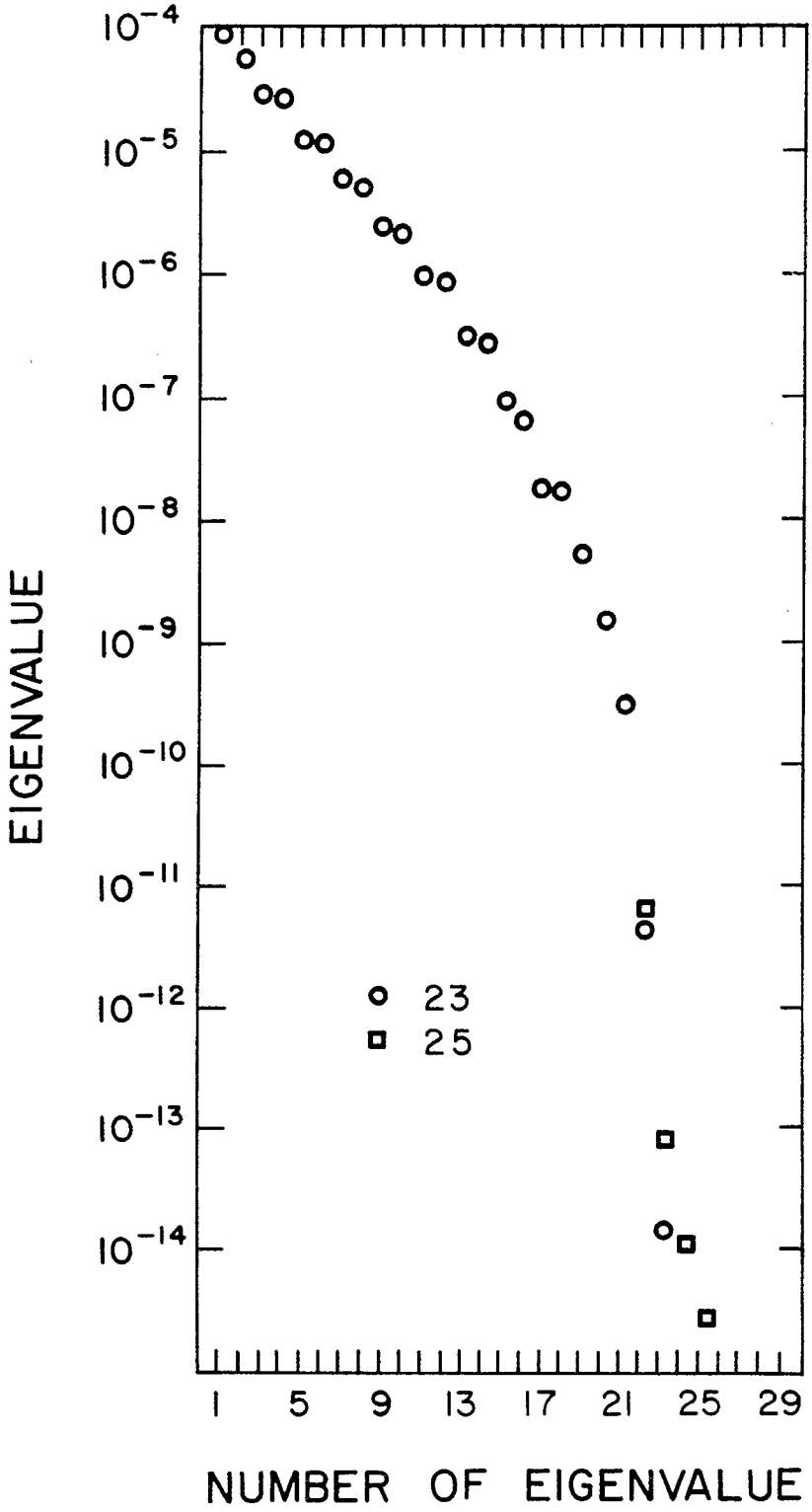


Figure 4.5. Comparison of the resulting twenty-three and twenty-five eigenvalues for 1.0 m dipole response.

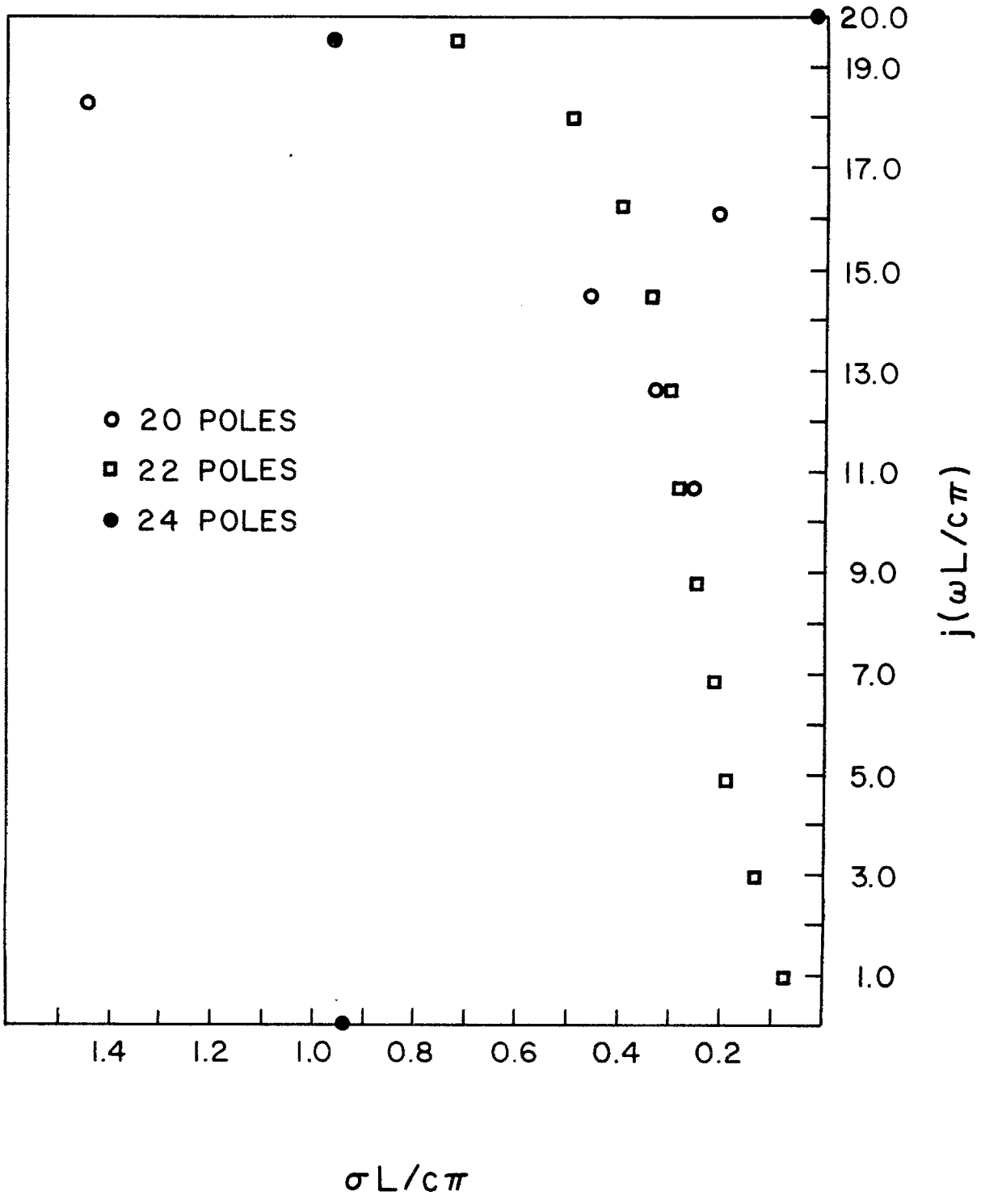


Figure 4.6. Pole locations for the extraction of twenty, twenty-two, and twenty-four poles from the 1.0 m dipole response.

twenty-third without any significant jump, except possibly between twenty-two and twenty-three. This is probably due to the fact that the residues of these eleven pole pairs drop off significantly as frequency increases. Because of this lack of a significant breakpoint, it appears that this technique is not as well suited to determining the value of N for systems which have transient responses similar to the response of this example.

5. NOISE AND ITS RELATIONSHIP TO PRONY'S METHOD

In the preceding chapters several allusions were made to the fact that noise seriously affects the values of the poles extracted by Prony's method. In Chapter 3 it was pointed out that if the noise level is high enough Prony's method will return poles that are not even remotely associated with the true poles being sought. This effect can be understood if the noise is thought of as being several arbitrary frequency components added to the signal. Thus, the true exponential nature of the signal has been corrupted. Since Prony's method is an interpolation process, it will give a set of poles which fit the noisy transient response but will not necessarily be the natural resonances. The least squares approach is used to reduce some of the error and to allow more data samples to be used. This chapter shows that the least squares approach when applied to Prony's method does not strictly give a least squares fit. If the noise level is known, it can be used to aid in the determination of the number of poles N using the eigenvalue process of Chapter 4. Statistical studies are presented which relate the noise levels to the quality of the results obtained.

5.1 Least Mean Squared Error Approach

Up to this point the least squares approach has been applied blindly by simply performing a pseudo-inverse solution. This was done primarily to allow for more transient data points to be used when solving for a set number of poles. This section studies the way in which the least mean squared error process reduces the errors that are present in the transient data.

Consider a set of transient response data which has been sampled at M equally spaced intervals Δt. These M sampled values $\dot{R}_0, \dot{R}_1, \ldots, \dot{R}_M$ are assumed to be estimates of the true response of the system $R_0, R_1, \ldots, R_M$ where

$$R_n = R(n\Delta t) = \sum_{i=1}^{N} A_i e^{s_i n\Delta t} \quad n = 0, 1, \ldots, M \quad . \tag{5.1}$$

The measured sampled values $\dot{R}_n$ differ from the true values R_n because of errors in measurements, noise, etc. It is known that the R_n satisfy Prony's difference equation of order N which may be written

$$R_{n+N} + \alpha_{N-1} R_{n+N-1} + \ldots + \alpha_1 R_{n+1} + \alpha_0 R_n = 0; \quad n = 0,1, \ldots, M \; . \tag{5.2}$$

In Chapter 3 it was shown that the pseudo-inverse or least squares solution to (5.2) could be written in matrix form as (3.8a)

$$A^T A B = A^T C \tag{5.3}$$

or

$$\Phi B = D \tag{5.4a}$$

where

$$\Phi = A^T A = \begin{bmatrix} \sum_{j=0}^{\gamma} \dot{R}_j^2 & \sum_{j=0}^{\gamma} \dot{R}_j \dot{R}_{j+1} & \cdots & \sum_{j=0}^{\gamma} \dot{R}_j R_{j+N-1} \\ \sum_{j=0}^{\gamma} \dot{R}_{j+1} \dot{R}_j & \sum \dot{R}_{j+1}^2 & \cdots & \sum_{j=0}^{\gamma} \dot{R}_{j+1} \dot{R}_{j+N-1} \\ \vdots & \vdots & & \vdots \\ \sum_{j=0}^{\gamma} \dot{R}_{j+N-1} \dot{R}_j & \sum_{j=0}^{\gamma} \dot{R}_{j+N-1} \dot{R}_{j+1} & \cdots & \sum_{j=0}^{\gamma} \dot{R}_{j+N-1} \dot{R}_{j+N-1} \end{bmatrix} \tag{5.4b}$$

$$B = \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \cdot \\ \cdot \\ \cdot \\ \alpha_{N-1} \end{bmatrix} \tag{5.4c}$$

$$D = A^T C = \begin{bmatrix} -\sum_{j=0}^{\gamma} \dot{R}_j \dot{R}_{N+j} \\ -\sum_{j=0}^{\gamma} \dot{R}_{j+1} \dot{R}_{N+j} \\ \cdot \\ \cdot \\ \cdot \\ -\sum_{j=0}^{\gamma} \dot{R}_{j+N-1} \dot{R}_{N+j} \end{bmatrix} . \tag{5.4d}$$

The γ of (5.4b) and (5.4d) must be greater than 2N - 1 and less than M - N - 1. Even though Equations (5.4) are formed in accordance with the usual least squares approach, the usual assumptions do not hold. Since the $\dot{R}_n$ are subject to errors, both the left side and the right side of (5.4a) are formed by noisy elements. In the normal least squares approach only the unknown quantities α_i are assumed subject to error. Thus, when the pseudo-inverse procedure is applied to Prony's method, it does not yield an optimum least squares approximation since the M samples $\dot{R}_n$ do not represent exactly a sum of N exponentials; that is,

$$\dot{R}_n = R_n + \varepsilon_n = \sum_{i=1}^{N} A_i e^{s_i \Delta t} + \varepsilon_n \tag{5.5}$$

where ε_n is the error in the n^{th} measured time sample. If ε_n equals zero, then the pseudo-inverse is the optimum least squares approximation.

Householder [27] developed a method by which the N parameters α_i could be determined such that the sum of the squared error

$$S = \sum w_n (R_n - \dot{R}_n)^2 \tag{5.6}$$

is minimized subject to the fulfillment of the side condition (5.2). The multipliers w_n are the statistical weights associated with the measured values $\dot{R}_n$. Householder's procedure is an iterative procedure involving Lagrange multipliers. The method is repeated several times until the results are within the accepted tolerance level. The method is very laborious and if more than just a few poles are sought it is totally unwieldy. In 1966 McBride, Schaefgen and Steighlitz [28] and in 1968 McDonough and Huggins [16] developed two different linear iterative schemes which appear quite successful for determining the poles even for large N. The methods, however, were developed for synthesizing a prescribed transient response using a sum of N complex exponentials. It was not required that the poles found be the true waveform poles or indeed it was not required that the waveform even have N poles. Hence, these methods are also not satisfactory for the requirements here, that is, finding the N true poles from a set of noisy transient response data of a system having precisely N poles.

The next section presents a method by which the poles of a system can be extracted from noisy data using an eigenvalue approach similar to that presented in Section 4.2.

5.2 Determination of the Poles from Noisy Data

In this section the approach in Section 4.2 is applied to transient responses containing noisy data. Consider the transient response vector v_i of (4.23) to include noise such that

$$v_i = \sum_{j=1}^{N} C_{j,i}\, \psi_j + \eta_i \tag{5.7}$$

where η_i is the noise vector starting at the i^{th} time step.

$$\eta_i = \begin{bmatrix} \varepsilon_i \\ \varepsilon_{i+1} \\ \cdot \\ \cdot \\ \cdot \\ \varepsilon_{\gamma+i} \end{bmatrix} . \tag{5.8}$$

Assume that the values ε_i of noise vector η_i are due to a random process and are normally distributed with zero mean and variance σ^2. The definitions of ψ_j, the j^{th} mode vector, and the coefficients $C_{j,i}$ are precisely those of (4.25) and (4.24), respectively. As before, if there are N poles in the system then there will be N mode vectors ψ_j.

The next step, as in Section 4.2, is to solve the difference equation

$$\sum_{m=0}^{N} v_{m+k}\, \alpha_m = 0 \quad . \tag{5.9}$$

If a pseudo-inverse solution of (5.9) is attempted, the set of Equations (4.26) is obtained. If the noisy transient response vectors are

substituted into the matrix Φ of (4.27), the following result is obtained for the i,j^{th} element of matrix Φ', where the prime indicates that the matrix was formed using the noise vector

$$\Phi'_{i,j} = v_i^T v_j \tag{5.10a}$$

$$\Phi'_{i,j} = \sum_{\ell=1}^{N} \sum_{m=1}^{N} [C^*_{\ell,i} C_{\ell,j}] \psi_\ell^{T*} \psi_m + \sum_{\ell=1}^{N} \psi_\ell^{T*} [C_{\ell,i} \eta_j]$$

$$+ \sum_{\ell=1}^{N} [C_{\ell,j} \eta_i^T] \psi_\ell + [\eta_i^T \eta_j] \quad . \tag{5.10b}$$

Since the term $[\eta_i^T \eta_j]$ is the product of the transpose of the i^{th} noise vector times the j^{th} noise vector and since any two noise vectors are uncorrelated, then

$$[\eta_i^T \eta_j] = 0 \quad , \quad i \neq j \tag{5.11a}$$

$$= \gamma \sigma^2, \quad i = j \tag{5.11b}$$

where γ is the length of the response vector v_i and the length of the noise vector η_i. Hence, this term gives rise to a matrix which is zero everywhere and has a value of $\gamma \sigma^2$ on the diagonal, or more precisely written,

$$H = \gamma \sigma^2 I \tag{5.12}$$

where I represents the identity matrix. The second and third terms on the right side of (5.10b) are each zero since they are the product of the noise vector and the natural mode vector which are uncorrelated and

of zero mean. The first term is left which is precisely the Φ of (4.28). This term is the product of the ℓ and n^{th} natural mode vectors. Thus,

$$\Phi' = \Phi + H \tag{5.13}$$

where Φ' is of order N + 1 by N + 1. Matrix Φ' will have N + 1 real eigenvectors of which one eigenvector will be orthogonal to the N mode vectors ψ_m. That is,

$$E_\ell^T \psi_m \neq 0 \ , \quad \ell = 1,N$$

$$= 0 \ , \quad \ell > N \tag{5.14}$$

where E_ℓ^T is the transpose of the ℓth eigenvector. Hence,

$$\Phi' E_{N+1} = \gamma \sigma^2 E_{N+1} \tag{5.15}$$

since $\Phi E_{N+1} = 0$ and matrix H can be thought of as the constant $\gamma \sigma^2$. Thus, (5.15) implies that the eigenvalue associated with the $N + 1^{th}$ eigenvector is just $\gamma \sigma^2$ which is γ times the variance of the noise. This is similar to the result of Section 4.2 where the eigenvalue associated with the $N + 1^{th}$ eigenvector is zero for the noise-free case. The procedure for determining the value of N is then the same as that outlined in Section 4.2 except that matrix Φ' is filled until it has an eigenvalue equal to $\gamma \sigma^2$ instead of zero. Likewise the $N + 1^{th}$ eigenvalue, which is orthogonal to all the mode vectors, is the vector containing the N + 1 coefficients α_i of Prony's difference equation. Thus, once the eigenvalue equal to $\gamma \sigma^2$ is found, the poles can be found from its corresponding eigenvector.

The next section applies this procedure to several examples. Different noise levels are used in order to statistically determine the poles' sensitivities to various noise variances.

5.3 Numerical Examples

Example 1

As the first example of the method developed in Section 5.2, a single undamped sinusoid is used. The transient data were produced at 200 time steps using the relation

$$R(n\Delta t) = \sin(\pi n \Delta t); \quad n = 0,1, \ldots, 199$$

where $\Delta t = 0.1$ s. Noise of different levels was added to the data samples $R(n\Delta t)$. The noise was produced by using a pseudo-random number generator on a digital computer to produce uniformly distributed noise. The uniformly distributed noise was then transformed into normally distributed noise with zero mean and a standard deviation of σ. For each different value of σ, the entire pole extracting process was repeated twenty times with twenty different sets of random numbers. That is, twenty Monte Carlo trials were performed for each standard deviation of the noise. From these twenty trials expected values of the poles were calculated along with the variance of the poles. The expected value of the N + 1 eigenvalue and its variance were also calculated. Table 5.1 shows the percent error of both the real and imaginary parts of the poles produced for different values of σ. The percent error was calculated as

$$\text{Percent error (real part)} = \frac{R\left|S_T - S_E\right|}{\sqrt{S_T^2}}$$

$$\text{Percent error (imaginary part)} = \frac{I\left|S_T - S_E\right|}{\sqrt{S_T^2}}$$

TABLE 5.1

RESULTS FOR EXAMPLE 1: $R(t) = SIN(\Pi t)$, $\gamma = 150$, $\Delta t = 0.1$ S

Standard Deviation of Noise σ	Percent Error of Pole		Standard Deviation of Pole		Theoretical N + 1 Eigenvalue	Mean Value of N + 1 Eigenvalue	Signal to Noise Level dB
	Real Part	Imaginary Part	Real Part	Imaginary Part			
0.5	3.3	3.4	0.395	0.796	37.5	38.6	2.0
0.1	0.33	0.027	0.034	0.031	1.50	1.54	16.0
0.07	0.21	0.011	0.026	0.016	0.735	0.757	19.1
0.05	0.14	0.003	0.020	0.009	0.375	0.386	22.0
0.01	0.024	0.001	0.004	0.001	0.015	0.0154	36.0

where S_T is the true value of the complex pole (in this case $S_T = 0 + j\pi$) and S_E is the calculated expected value of the extracted poles for a particular noise level. Also tabulated in Table 5.1 are the standard deviation of the extracted poles, the theoretical value of the $N + 1^{th}$ eigenvalue (i.e., $\gamma\ \sigma^2$), the mean value of the calculated $N + 1^{th}$ eigenvalue, and the signal-to-noise level in decibels. The signal-to-noise level was calculated by taking the log of the square of the average of the absolute value of the signal over the time window of length $\gamma\Delta t$ and dividing by the variance of the noise σ^2. That is

$$S/N(dB) = 10 \log \frac{\left(\sum_{i=0}^{\gamma-1} \frac{|R(i\Delta t)|}{\gamma} \right)^2}{\sigma^2} \quad .$$

For this first example the average of the absolute value of the signal is 0.63137 for all time because the signal is undamped. The values of σ for the noise used ranged from 0.5 to 0.01 or in terms of signal-to-noise level from 2.0 dB to 36.0 dB. Note that the percent error of the real part of the poles is more sensitive to noise than is the imaginary part. Note also that the theoretically predicted value of the $N + 1^{th}$ eigenvalue is extremely close to the calculated value of that eigenvalue. This example indicates that the poles can be extracted from data when the signal-to-noise level is as low as 2.0 dB. If the standard deviation of the extracted poles is studied for the 2.0 dB case, the results do not look very good. Note that for the imaginary part of the pole the standard deviation is 0.796 which is extremely high when it is remembered the value of the pole is just π. The standard deviation of the pole at the 16.0 dB level, however, is much smaller indicating that the chance of

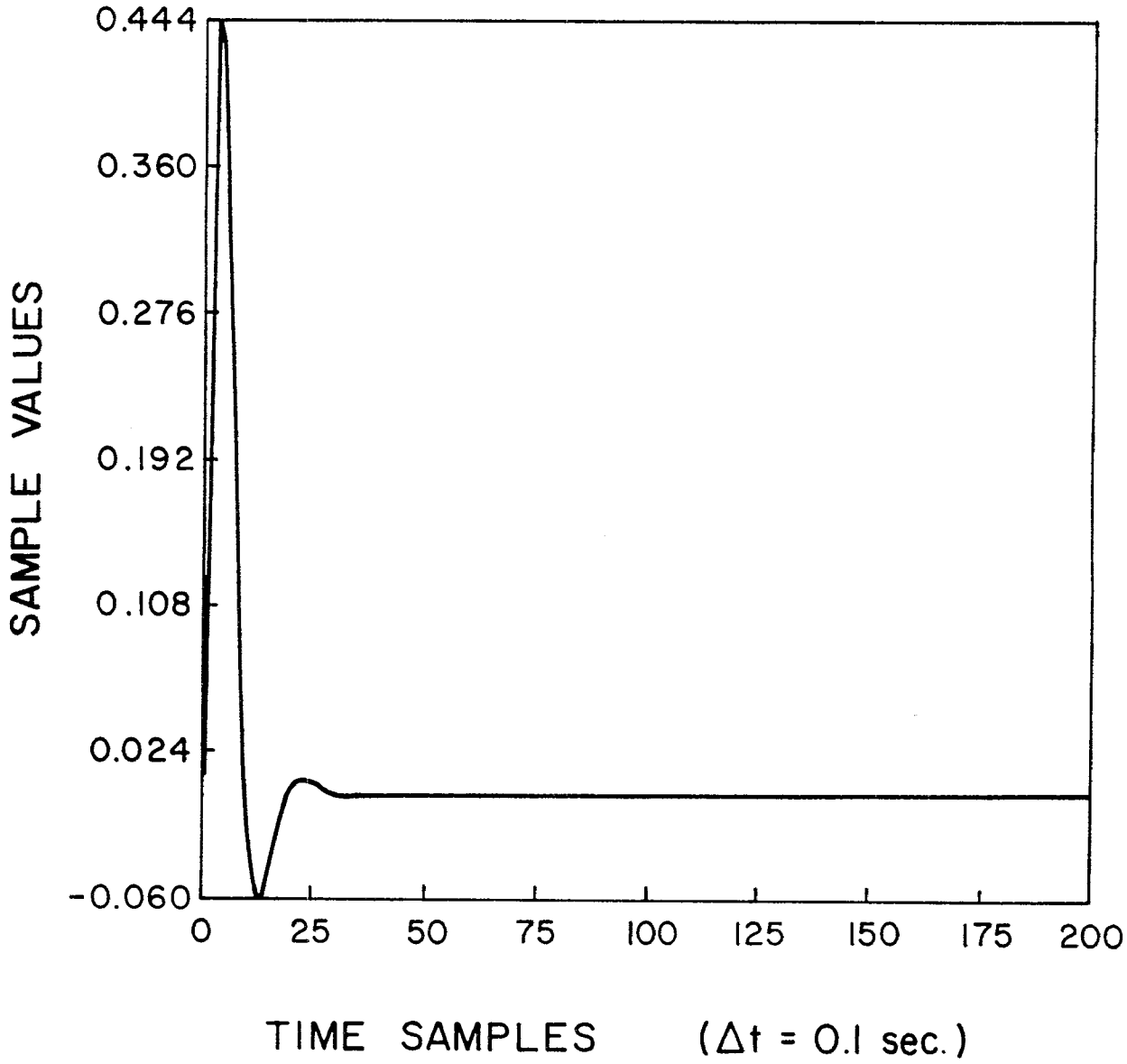


Figure 5.1. Transient waveform used in Example 2.

TABLE 5.2

RESULTS FOR EXAMPLE 2: $R(t) = e^{-2t}\ SIN(\pi t)$, $\gamma = 50$, $\Delta t = 0.1$ S

Standard Deviation of Noise σ	Percent Error of Pole		Standard Deviation of Pole		Theoretical N + 1 Eigenvalue	Mean Value of N + 1 Eigenvalue	Signal to Noise Level dB
	Real Part	Imaginary Part	Real Part	Imaginary Part			
0.04	1.80	5.90	0.800	0.566	0.08	0.0836	3.3
0.03	2.55	2.10	0.566	0.311	0.045	0.047	5.8
0.02	2.20	0.48	0.371	0.179	0.02	0.021	9.4
0.01	1.27	0.12	0.187	0.085	0.005	0.0052	15.4
0.009	1.16	0.14	0.167	0.077	0.00405	0.00423	16.3
0.005	0.67	0.14	0.094	0.043	0.00125	0.00131	21.4
0.001	0.14	0.04	0.017	0.009	0.5×10^{-4}	0.52×10^{-4}	35.4

extracting the true pole is much higher at 16.0 dB than at 2.0 dB. This result will be further substantiated in the next example.

Example 2

For the second example the same sinusoid as in Example 1 was used but now it is exponentially damped. That is,

$$R(n\Delta t) = e^{-2.0n\Delta t} \sin(\pi n\Delta t) \quad n = 0,1, \ldots, 199$$

where again $\Delta t = 0.1$ s. This waveform is plotted in Figure 5.1. A value of $\gamma = 50$ was used in this example since after fifty time steps the waveform has essentially damped to zero. Table 5.2 shows the results of several statistical tests on these data using a standard deviation of the noise ranging from 0.04 down to 0.001. It was found that if σ is greater than 0.04, a signal-to-noise level of 3.3 dB, the expected values of the two returned poles were not even complex conjugates of each other. Thus, when the signal-to-noise level was worse than 3.3 dB Prony's method was completely corrupted by the noise. Note in Table 5.2 that as the signal-to-noise level gets better the standard deviation of the poles lowers. It appears that at around 15.0 to 20.0 dB the percent error and the standard deviation of the poles are at a tolerable level, which is, of course, subject to the requirements of the problem being studied. Notice in this example, as in Example 1, that the theoretical value of the $N + 1^{th}$ eigenvalue compares closely to the calculated mean value. In all cases the theoretical value is slightly lower than the calculated mean value.

Example 3

For this example the transient response of Figure 4.1 in Chapter 4 was used. The six pole pairs and their associated amplitudes are

TABLE 5.3

AMPLITUDE OF EACH POLE COMPONENT FOR EXAMPLES 3 AND 4

Pole	Amplitude Example 3	Amplitude Example 4
-0.082 ± j 0.926	1.0	1.0
-0.147 ± j 2.874	1.0	0.5
-0.188 ± j 4.835	1.0	0.25
-0.220 ± j 6.800	1.0	0.125
-0.247 ± j 8.767	1.0	0.0625
-0.270 ± j 10.733	1.0	0.03125

TABLE 5.4

VALUES OF EXTRACTED POLES AND THEIR PERCENT ERROR AS A FUNCTION OF THE SIGNAL TO NOISE RATIO FOR EXAMPLE 3

S/N	27.4 dB	29.4 dB	33.4 dB	37.9 dB	47.4 dB	
Standard Deviation	$\sigma = 0.01$	$\sigma = 0.008$	$\sigma = 0.005$	$\sigma = 0.003$	$\sigma = 0.001$	Original Poles
Real Part of Poles	-0.1245 -0.1238 -0.1467 -0.1943 -0.2592 -0.2722	-0.0734 -0.1612 -0.1625 -0.2019 -0.2555 -0.2717	-0.0774 -0.1479 -0.1785 -0.2122 -0.2512 -0.2710	-0.0799 -0.1455 -0.1836 -0.2170 -0.2492 -0.2706	-0.0815 -0.1456 -0.1868 -0.2197 -0.2476 -0.2702	-0.082 -0.147 -0.188 -0.220 -0.247 -0.270
Imaginary Part of Poles	0.9550 2.7886 4.7921 6.8140 8.7725 10.7302	0.9214 2.8557 4.8044 6.8110 8.7712 10.7311	0.9225 2.8702 4.8233 6.8064 8.7692 10.7321	0.9238 2.8724 4.8315 6.8037 8.7681 10.7326	0.9252 2.8735 4.8350 6.8012 8.7673 10.7329	0.926 2.874 4.835 6.800 8.767 10.733
Percent Error Real Part	4.57 0.81 0.85 0.38 0.14 0.02	0.92 0.49 0.53 0.27 0.09 0.02	0.49 0.03 0.19 0.11 0.05 0.009	0.22 0.05 0.09 0.04 0.02 0.005	0.05 0.05 0.02 0.004 0.007 0.002	
Percent Error Imaginary Part	3.12 2.96 0.89 0.20 0.06 0.02	0.49 0.63 0.63 0.16 0.05 0.02	0.37 0.13 0.24 0.09 0.02 0.008	0.24 0.13 0.07 0.05 0.01 0.004	0.08 0.02 0.0 0.02 0.003 0.0	

TABLE 5.5

SIGNAL TO NOISE RATIO OF THE DIFFERENT COMPONENTS
OF THE SIGNAL IN EXAMPLE 3 AS A FUNCTION
OF THE STANDARD DEVIATION OF THE NOISE

	Standard Deviation of Noise	$\sigma = 0.01$	$\sigma = 0.008$	$\sigma = 0.005$	$\sigma = 0.003$	$\sigma = 0.001$
		Signal to Noise Ratio dB				
Portion of Signal Due to Particular Pole	Sum of Below	27.4	29.4	33.4	37.9	47.4
	$-0.082 \pm j\ 0.926$	24.9	26.8	30.9	35.4	44.9
	$-0.147 \pm j\ 2.874$	20.1	22.0	26.1	30.6	40.1
	$-0.188 \pm j\ 4.835$	17.9	19.9	23.9	28.4	37.9
	$-0.220 \pm j\ 6.800$	16.6	18.5	22.6	27.1	36.6
	$-0.247 \pm j\ 8.767$	15.6	17.5	21.6	26.1	35.6
	$-0.270 \pm j\ 10.733$	14.8	16.7	20.8	25.2	34.8

TABLE 5.6

THE THEORETICAL AND EXPERIMENTAL VALUES OF THE N + 1 EIGENVALUE FOR EXAMPLE 3 AND EXAMPLE 4

	Standard Deviation of Noise σ	Theoretical Value of N + 1 Eigenvalue	Mean Value of N + 1 Eigenvalue	Standard Deviation of N + 1 Eigenvalue
Example 3	0.01	0.02	0.0218	4.92×10^{-3}
	0.008	0.0128	0.0139	3.15×10^{-3}
	0.005	0.0050	0.00546	1.23×10^{-3}
	0.003	0.0018	0.00197	4.43×10^{-4}
	0.001	0.2×10^{-3}	0.22×10^{-3}	4.92×10^{-5}
Example 4	0.002	0.8×10^{-3}	0.871×10^{-3}	1.96×10^{-4}
	0.001	0.2×10^{-3}	0.218×10^{-3}	4.93×10^{-5}
	0.0005	0.5×10^{-4}	0.546×10^{-4}	1.23×10^{-5}

tabulated in Table 5.3. Note that for this example the amplitude of each pole is the same. Noise with standard deviation from 0.01 to 0.001, giving a signal-to-noise level from 27.4 dB to 47.4 dB, was added to the signal and statistical tests were run on the data. For all cases a value of $\gamma = 200$ was used and the number of Monte Carlo trials was twenty. Table 5.4 shows the results of these tests. For this particular signal, poles could not be extracted for levels of noise greater than $\sigma = 0.01$. That is, for noise levels greater than $\sigma = 0.01$, poles were produced but they did not appear in complex conjugate pairs and had no apparent relationship to the original poles.

Since the total waveform was made up of the sum of six damped sinusoids, it is useful to study the signal-to-noise ratio for each of the six individual components. Table 5.5 shows that for the case of $\sigma = 0.01$ with a total signal-to-noise ratio of 27.4 dB the highest frequency component has a signal-to-noise ratio of only 14.8 dB. This level is about the same as the level in Examples 1 and 2 where the poles could be extracted at a signal-to-noise level of about 2 to 3 dB, but in this example the highest signal-to-noise level of an individual component is, as stated above, 14.8 dB. Hence, indications are that summing several signals together in effect makes it necessary to have a better signal-to-noise ratio to allow extraction of the poles. This point is further demonstrated in Example 4.

The values of the $N + 1^{th}$ eigenvalues calculated in Example 3 are tabulated in Table 5.6. Note here, as in the previous two examples, that the mean value of the eigenvalues is very close but always slightly higher than the theoretically predicted values. In all cases it was found that the N^{th} eigenvalue was sufficiently distinct from the $N + 1^{th}$

eigenvalue so that proper identification of the cutoff point could be made. For example in the worst case, that of $\sigma = 0.01$, the $N + 1^{th}$ eigenvalue has a mean value of 0.0218 and the N^{th} eigenvalue has a mean value of 0.070. Thus, if the standard deviation of the noise is known in advance, there should be no problem in detecting the cutoff point which gives the value of N.

Example 4

For the final example the same six pole pairs of the previous example were used but the amplitudes of the poles were reduced as indicated in Table 5.3. The resulting transient response curve is shown in Figure 5.2. Here again a value of $\gamma = 200$ was used and twenty Monte Carlo trials were performed. Table 5.7 shows the results of noise with $\sigma = 0.002$, 0.001, and 0.0005. The noise level at which the true poles could be first detected for this example is $\sigma = 0.002$ which corresponds to a signal-to-noise level of 39.7 dB. This level is far higher than any needed in the previous examples, and is due to the fact that some of the components of the signal have very small signal levels. Table 5.8 gives the signal-to-noise level for each of the six components in this example. Note that the highest frequency has a signal-to-noise level of -1.3 dB which indicates that the noise level is higher than the average signal level. However, the signal must have been above the noise level for a significant portion of the data window in order for it to be detected. The standard deviations of the extracted poles for the $\sigma = .002$ noise level indicate that the method is not accurate enough at this level and indeed does not appear to be accurate until a noise level of $\sigma = 0.0005$ is reached. At this level the highest signal-to-noise level for any single component

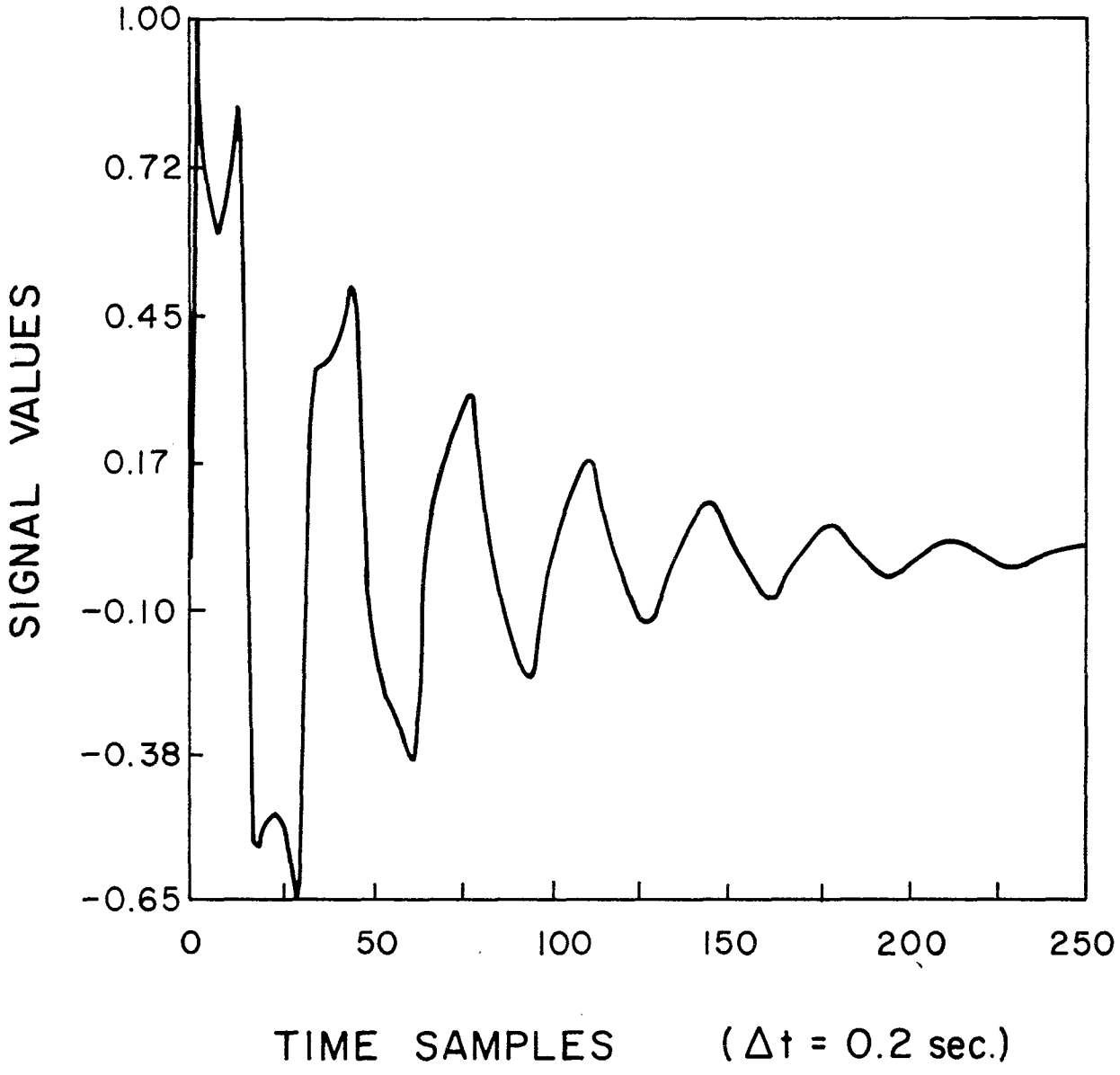


Figure 5.2. Transient waveform used in Example 4.

TABLE 5.7

VALUES OF EXTRACTED POLES WITH THEIR PERCENT ERROR AND STANDARD DEVIATION AS A FUNCTION OF SIGNAL TO NOISE RATIO FOR EXAMPLE 4

Signal to Noise Ratio	39.7 dB	45.7 dB	51.7 dB	Original Poles
Standard Deviation of Noise σ	0.002	0.001	0.0005	
Real Part of Poles	-0.0809 -0.1493 -0.3629 -0.1483 -0.3109 -0.2315	-0.0814 -0.1452 -0.1793 -0.2036 -0.2823 -0.2700	-0.0817 -0.1459 -0.1848 -0.2159 -0.2565 -0.2719	-0.082 -0.147 -0.188 -0.220 -0.247 -0.270
Imaginary Part of Poles	0.9220 2.8675 4.9708 6.7206 8.6936 10.6593	0.9251 2.8726 4.8293 6.8173 8.7688 10.7108	0.9256 2.8734 4.8342 6.8065 8.7686 10.7271	0.926 2.874 4.835 6.800 8.767 10.733
Percent Error Real Part	0.12 0.08 3.61 1.05 0.73 0.36	0.06 0.06 0.18 0.24 0.40 0.0	0.03 0.04 0.07 0.06 0.11 0.02	
Percent Error Imaginary Part	0.43 0.22 2.80 1.16 0.84 0.68	0.09 0.05 0.12 0.25 0.02 0.21	0.04 0.02 0.02 0.09 0.02 0.05	
Standard Deviation of Real Part of Poles	0.0113 0.1034 0.8509 0.1000 0.4254 0.0834	0.0054 0.0415 0.0600 0.0371 0.1871 0.0470	0.0026 0.0200 0.0300 0.0175 0.0785 0.0231	
Standard Deviation of Imaginary Part of Poles	0.0310 0.0382 0.7797 0.4604 0.3860 0.2846	0.0120 0.0120 0.0585 0.1237 0.0632 0.1225	0.0057 0.0055 0.0283 0.0616 0.0211 0.0535	

TABLE 5.8

SIGNAL TO NOISE RATIO OF THE DIFFERENT COMPONENTS OF THE SIGNAL IN EXAMPLE 4 AS A FUNCTION OF THE STANDARD DEVIATION OF THE NOISE

	Standard Deviation of Noise σ	0.002	0.001	0.0005
		Signal to Noise Ratio dB		
Portion of Signal Due to Particular Pole	Sum of Below	39.7	45.7	51.7
	$-0.082 \pm j\ 0.926$	38.9	44.9	50.9
	$-0.147 \pm j\ 2.874$	28.1	34.1	40.1
	$-0.188 \pm j\ 4.835$	19.9	25.9	31.9
	$-0.220 \pm j\ 6.800$	17.9	23.9	29.9
	$-0.247 \pm j\ 8.767$	5.5	11.5	17.5
	$-0.270 \pm j\ 10.733$	-1.3	4.7	10.7

is better than 10 dB. Table 5.6 shows that here again the mean value of the lowest eigenvalue is always slightly higher than the theoretical value.

Summary of Numerical Examples

Four different numerical examples were studied to determine the effects which noise has on Prony's method and to see if the value of N could be found from the eigenvalues. In all cases the mean value of the calculated $N + 1^{st}$ eigenvalue was slightly higher than the theoretical value of $\gamma \sigma^2$. It was also found that the $N + 1^{st}$ eigenvalue was sufficiently different from all others of lower order so that it is possible to identify it. For all four examples a different highest tolerable signal-to-noise level was obtained. However, in all the examples the best results were obtained when the signal-to-noise level was between 10 and 20 dB for the lowest level component of the signal.

The examples indicate that different tolerable noise levels will be found for different shaped waveforms. One definite conclusion is that if the noise level is higher than the signal level over the entire data window then that signal or its accompanying pole will not be detectable. These examples also indicate that further study should be done with experimentally obtained signals containing truly random noise.

6. APPLICATIONS

Now that the techniques for properly applying Prony's method have been developed it is of interest to study some of the electromagnetic applications of this method. There are four major areas to which the extraction of poles from a transient electromagnetic signal can be of great benefit. These are: system analysis, radar target recognition, the study of spectral characteristics, and data reduction and extrapolation. These four applications are discussed in some detail in this chapter.

6.1 System Analysis

6.1.1 Response to various exciting waveforms

In circuit theory after the impulse response is known, the response of the circuit to any given driving function can be determined by multiplying the Laplace transform of the driving function by the sum of poles and their corresponding residues of the impulse response. That is,

$$R(s) = F(s) \sum_{i=1}^{N} \frac{A_i}{s - s_i} \tag{6.1}$$

where R(s) is the resulting response function and F(s) is the Laplace transform of the driving function. In Chapter 2 it was shown that a similar relation is true for electromagnetic structures. The differences are that since antennas and scatterers are distributed systems the impulse response is a function of position on the structure and a function of the incident angle of the driving function. Note that the pole portion of the impulse response is independent of position but that the natural modes are functions of position and the coupling coefficients are functions of the incident angle. Thus, the general response function for an antenna or scatterer can be expressed as

$$\bar{R}(s,\bar{r},\bar{p}) = F(s,\bar{r},\bar{p})\left(\sum_{i=1}^{N} \frac{A_i(\bar{r},\bar{p})}{s - s_i}\right) \tag{6.2}$$

where the coupling coefficients $\eta_i(s_i,\bar{p})$ and the natural modes $\nu_i(\bar{r})$ have been combined into the one term, $A_i(\bar{r},\bar{p})$. Since the poles s_i are invariant of position, they may be determined by studying the impulse response at any position on the structure. After the poles have been evaluated, the residues A_i must be determined at every desired point on the structure. This however is not a difficult problem. Thus, the poles of a structure need only be extracted once and the residues must be determined at each position that the response is desired and for each angle of incidence that the driving function will be applied. As an illustration consider the following example.

Consider that the induced current at several positions, e.g., M, on a dipole due to several different time-varying broadside incident plane waves is desired. The first step is to determine the induced current at all M desired positions on the structure resulting from a broadside incident impulse plane wave. It will actually be necessary to use a narrow Gaussian plane wave as an approximation to the impulse. The true impulse response can then be obtained by using one of the methods outlined in Section 2.2. After the impulse response of the induced current is obtained, then the poles of the structure can be obtained by applying Prony's method to the current at one of the M points on the structure. Now that the poles have been determined then the M sets of residues are calculated for the M positions on the antenna,

thus giving one set of N poles and M sets of N residues $A_i(\bar{r},\bar{p})$. The induced current for an arbitrary broadside incident waveform is then obtained by applying (6.2). If an incident angle other than broadside is required, then it is necessary to recalculate the M sets of residues for an incident impulse plane wave at the new incidence angle. This seems a horrendous amount of calculations but when compared with the amount required to completely resolve the problem for each incident wave shape it is really a very small amount.

6.1.2 Compatibility with circuit theory

In practice antennas are almost always coupled to some sort of circuit or network. This presents a problem to the circuit designer because the antenna port is not classified as a lumped element. This problem is usually not severe since the network and antenna are generally designed to operate at one given frequency. In this case the antenna port can be characterized by a lumped impedance at that one frequency. If, however, the antenna is to be operated over a broad band of frequencies the circuit designer must know the antenna's terminal characteristics over that entire band. The characteristics are normally obtained from a set of graphs relating the real and imaginary parts of the input admittance to the frequency. Obviously, this is not an ideal situation. The circuit designer would like to have the antenna's input impedance given to him in the Laplace transform domain. This is a possibility if Prony's method is used. All that is necessary is to apply a time-varying voltage to the antenna terminals and to measure the resulting current as a function of time. If Prony's method is applied to the current and a set of poles and residues is obtained, then the input admittance can be expressed as

$$Y(s) = \frac{\sum_{i=1}^{N} \frac{A_i}{s - s_i}}{V(s)} \qquad (6.3)$$

where V(s) is the Laplace transform of the applied voltage. Expression (6.3) is obviously much easier for a circuit designer to use than a set of complicated graphs or tables.

6.1.3 Study of system parameters

It is quite useful to know what effect the loading of a structure has on the positions of its poles in the complex plane. Likewise it is beneficial to know what effect the positions of the poles have on the behavior of the transient response. If trajectories of the poles could be determined from a few experimental cases, it would be possible to predict pole locations as a function of loading. Once the pole locations are known then the transient waveforms can be constructed. One practical example of this problem is the resistive loading of a linear antenna in order to produce a transmitted field which simulates an electromagnetic pulse. Tesche [23] has approached this problem by determining the pole locations of a dipole which is loaded with uniform resistive loading along the antenna. He produced the trajectories of the poles by the "classical" frequency domain search procedure. An alternative approach is to obtain the time-domain solutions for the induced currents on and the scattered fields from several uniformly resistively loaded dipoles. Prony's method is applied to either the transient response of the induced currents or the scattered fields in order to determine the location of the poles. The results of such a procedure are given in the following example.

A 1.0 m dipole with a half-length-to-radius ratio of 100 was numerically modeled and excited by a broadside incident Gaussian pulse. The induced current at the center of the antenna and the back scattered electric field was calculated as a function of time for uniform resistive loading of 0, 125, 250, 500, 750 and 1683 Ω/m. These currents and fields are shown in Figures 6.1 and 6.2, respectively. The time step size used was $\Delta t = 6.9444 \times 10^{-11}$ s. The loading value of 1683 Ω/m was chosen because Tesche [23] calculated that 1683 Ω/m is the value at which the dipole would become critically damped, giving a double pole on the negative real axis. The poles that were extracted using Prony's method are plotted in Figure 6.3. The trajectories of the first seven even poles are shown. Note that for the value of 1683 Ω/m the first pole has split and moved toward the origin and toward infinity, which indicates that this value of loading does not give a critically damped dipole but produces an overdamped situation. This does not imply that Tesche's calculated value of 1683 Ω/m is incorrect. All that is indicated is that this numerical modeling procedure and Tesche's numerical modeling procedure differ. It should be noted, however, that the point at which the predicted trajectory of the first pole in Figure 6.3 splits is approximately the same value which Tesche indicates in his paper.

One very interesting point in the above example is that for even the overdamped transient response Prony's method was capable of extracting the first five poles of the system. This is a very important point since if the 1683 Ω/m case in Figure 6.1 is studied it appears that no oscillations occur. Yet Prony's method was capable of determining the first five modes including the splitting of the first pole. This example then indicates that Prony's method will work with heavily loaded structures and with fat or thick structures.

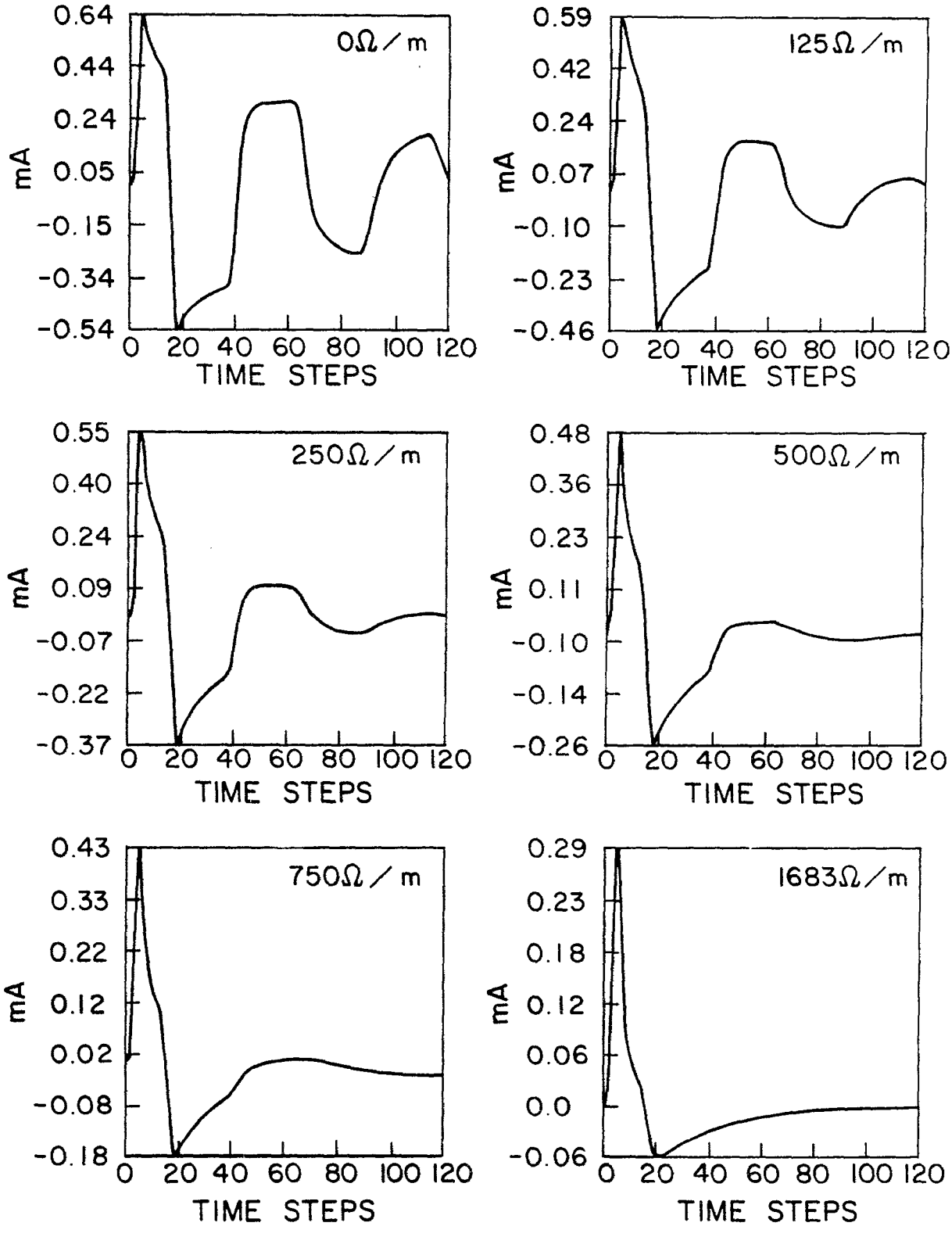


Figure 6.1. Induced current at center element of 1 m dipole with resistive loadings.

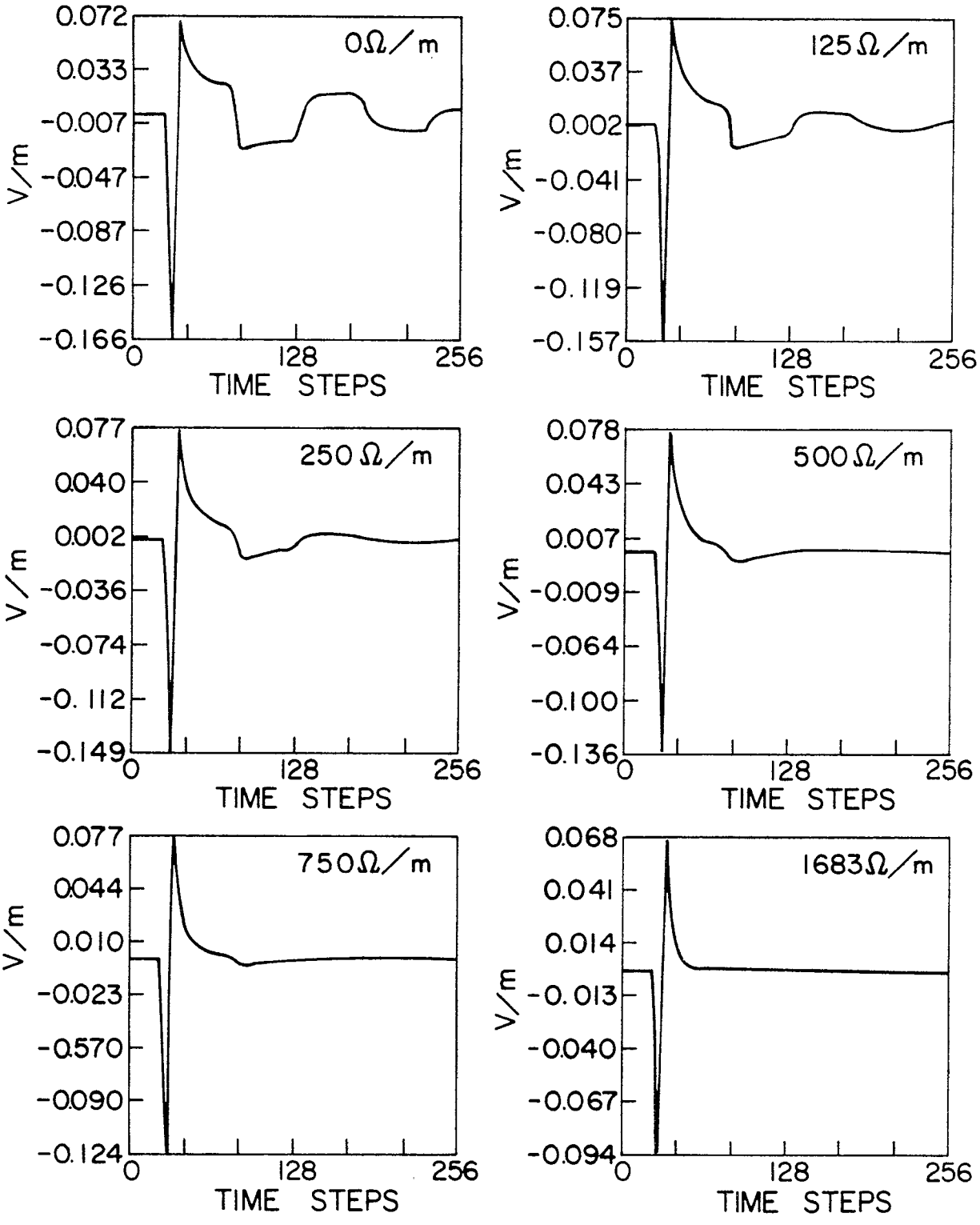


Figure 6.2. Backscattered fields for 1 m dipole with resistive loading.

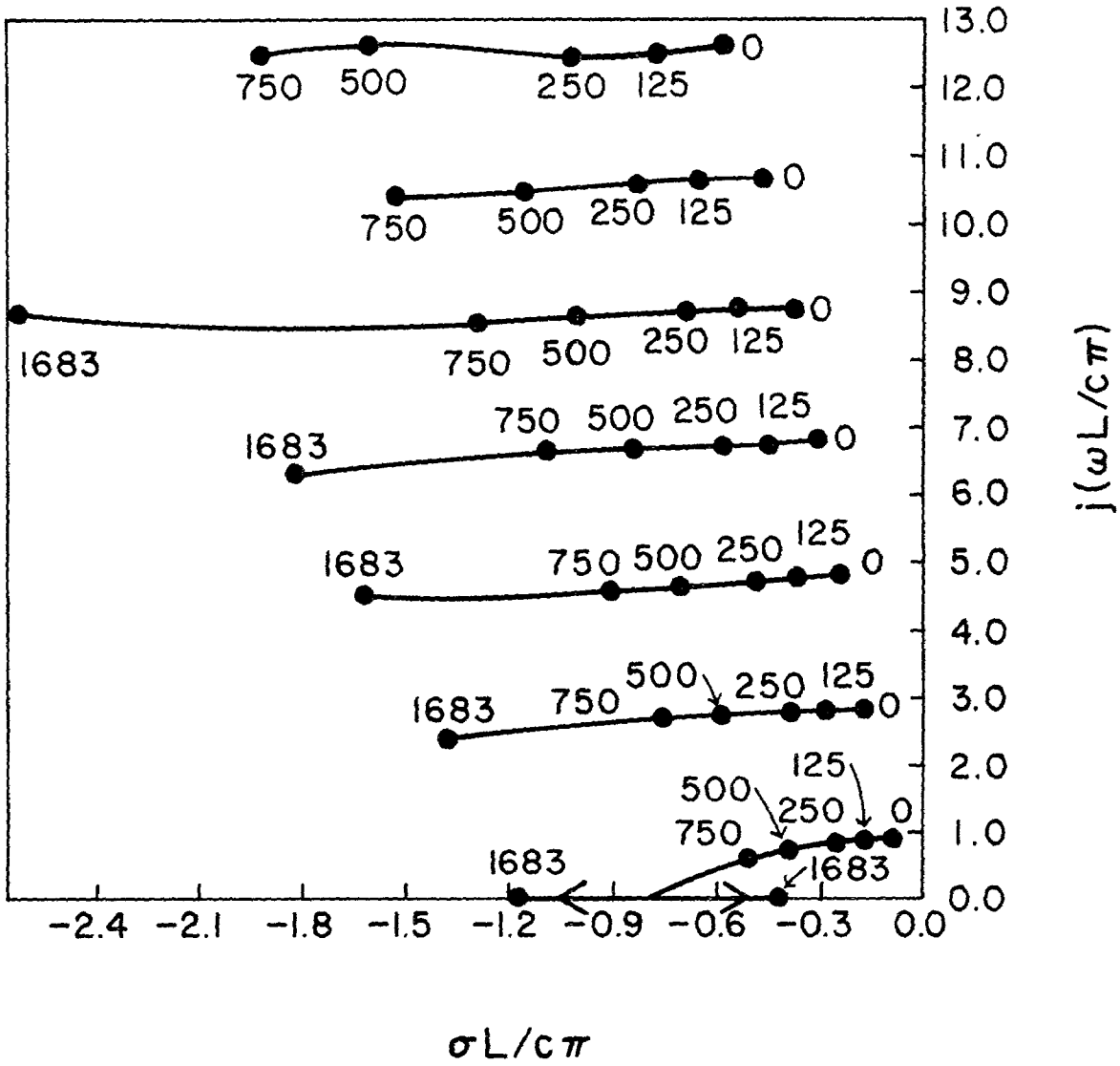


Figure 6.3. Trajectory of the poles for the uniformly resistive loaded 1.0 m dipole.

6.2 Radar Target Recognition

The problem of electromagnetic recognition of a radar target has received a great deal of attention in recent years. To recognize a target with a single-frequency radar requires, in principle, bistatic scattering cross-section information for a single transmitter aspect angle and for 4π steradians. In practice, with a priori information about the body, it is feasible to match a body to one of a known group of bodies using a somewhat smaller angular range. With multiple-frequency monostatic radar cross-section information available over a limited angular range, it is possible to identify a body as well [29]. Theoretically, the monostatic reconstruction requires information over 4π steradians and an infinite frequency range.

Recent work indicates the feasibility of using short-pulse (impulse response) excitation in a monostatic configuration for purposes of target recognition from a class of bodies [30] - [31].

In the first of these schemes, reported by Sperry Research Center, information about the body is contained in the time-history of the backscattered waveform, i.e., a time signature. This technique is complicated by the fact that the time signature is dependent on the aspect angle of the target. The result is the requirement of a rather large catalog of known signatures.

Mains and Moffatt [31] introduced the concept of using the complex natural resonances of a target as a basis for target recognition. They make use of the fact that a few natural resonances of a body are adequate to distinguish the body within a finite collection of bodies. They also make use of the fact that the natural resonances of a body as manifested in a scattered waveform are aspect independent: they do not depend on

the angular orientation of the target. They justify this fact empirically in [31]. Baum, who introduced the Singularity Expansion Method (SEM) [1] - [2], has rigorously demonstrated this aspect independence in the context of the SEM method.

Thus, this method provides a convenient means of 1) characterizing a transient scattered waveform in terms of a few coefficients and complex resonant frequencies in a series of decaying sinusoids, and 2) separating aspect-dependent characteristics of the waveform from characteristics intrinsic to the body. In particular, the complex frequencies of the sinusoids in the series are intrinsic to the body while the amplitude of the sinusoids depends on the wave shape of the excitation and on the aspect angle.

Two significant practical features accrue from the use of the SEM representation with primary attention directed to the complex frequency qualities: 1) aspect dependence is suppressed and 2) the incident wave shape need not be impulse - it needs only to have sufficient spectral content to excite several of the complex sinusoidal modes.

Since Prony's method can be used to numerically extract the weighting coefficients and the complex frequencies from a digitized backscattered time-signature, Pearson, Van Blaricum and Mittra [32] suggested that the method be used to aid in the target recognition problem. The extracted frequencies (rather than the complete time-signature) would serve as the input to a pattern-recognition algorithm. This method is believed to represent a significant improvement over direct time-history pattern recognition because the aspect-dependence is suppressed and because the frequencies comprise a smaller set of numbers than the entire sampled time-signature.

A block diagram for the recognition procedure is given in Figure 6.4. The system consists of a transmitter, a receiver capable of digitizing the transient return, and a digital processor with some convenient operator communications device such as a CRT. Of course, a tracking system and an antenna/duplexer are present.

The waveform generator must produce a transient waveform with reasonably broad spectral content. The power and frequency requirements are, of course, dictated by the application. In particular, target range requirements and receiver performance determine a power specification, while the body size determines frequency. The frequency must be such that $f = c/D$ with D a characteristic dimension of the class of targets and c the velocity of light. An alternative to a transient source and receiver is a multiple-frequency pulsed carrier scheme such as Mains and Moffatt suggest [31]. This scheme uses CW scattering returns to synthesize the transient response of the body.

The transient digitizer provides a means of detecting, sampling, and digitally expressing the transient return signal. In its simplest form, this element might be a storage oscilloscope with A-D converters on its vertical and horizontal drive signals. There are other more sophisticated systems available commercially.

The digital processor must be able to perform the following operations: a) extract the poles of the waveform by using Prony's method; b) identify (based on data designed into the system) poles due to the transmitted wave shape and to the RF system; and c) conduct a pattern recognition procedure to provide a set of "likelihood votes" associating the target with a catalog of expected targets. The first item above is self-explanatory.

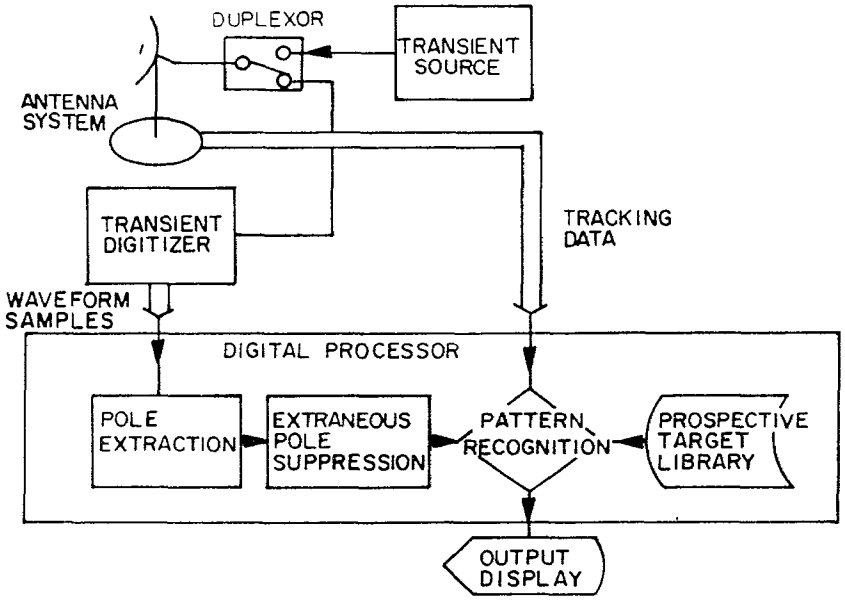


Figure 6.4. Schematic representation of the radar target recognition system.

Any waveform that can be radiated is oscillatory in character and will accordingly have associated with it at least one "waveform pole." For example, a sinusoidal segment will possess a pole at $s = j\omega_0$, where ω_0 is the radian frequency of the sinusoidal signal. In addition, it is conceivable though undesirable that the receiving antenna and RF hardware might introduce poles into the waveform representation. All of these "system" poles are known during system design, however, and it is a straightforward algorithm to locate and delete them from the set of poles extracted from the signal. An alternative approach is to numerically deconvolve the waveform spectrum and system transfer function from the received waveform.

The final process is to compare, by means of some pattern analysis algorithm, the observed pole set with a data base of known targets' poles. This algorithm would display, as its output, a "vote" or probability for a known target or targets. This information could be supplemented by an indication of recognition confidence, for example, the number of poles successfully used in the identification. A second supplementary output might be tracking data on the target.

For this method to be successful as a target identification scheme, a few poles (say less than ten) must characterize a given target among all potential targets. Mains and Moffatt [31] discuss this problem in some detail and present the pole configuration for several thin-wire geometries to indicate that a few poles do distinguish among similar objects. Also, another criterion for this system to be successful is a satisfactory signal-to-noise level of the returned time-signature so that Prony's method can properly extract the poles. Berni [33] has

suggested an identification technique, which is also based on Prony's algorithm, although Prony's method, as such, is not mentioned in his paper. His approach is to use a series of exciting pulses which then produces a series of response functions which are correlated to produce what he calls a "target correlation matrix." His correlation matrix is similar to the Φ matrix of Prony's method of Section 4.2. The difference is that his correlation matrix is filled with samples from several different sets of transient response data while that of Prony's method uses just one set of transient response data. Berni's approach may thus prove useful in reducing the noise level and some of the signal-to-noise ratio problems outlined in Section 6.3. It appears that the use of Prony's method as a tool for radar target recognition may be of great importance but more study of the problem is needed in order to satisfy many of the unanswered questions.

6.3 Study of Spectral Characteristics

It has been pointed out already that after the poles and residues of a system have been determined it is then possible to write the frequency domain version of the system's response as

$$R(j\omega) = \sum_{i=1}^{N} \frac{A_i}{-\sigma_i + j(\omega - \omega_i)} \tag{6.4}$$

where the poles s_i have been written in terms of their real and imaginary parts as

$$s_i = \sigma_i + j\,\omega_i \quad . \tag{6.5}$$

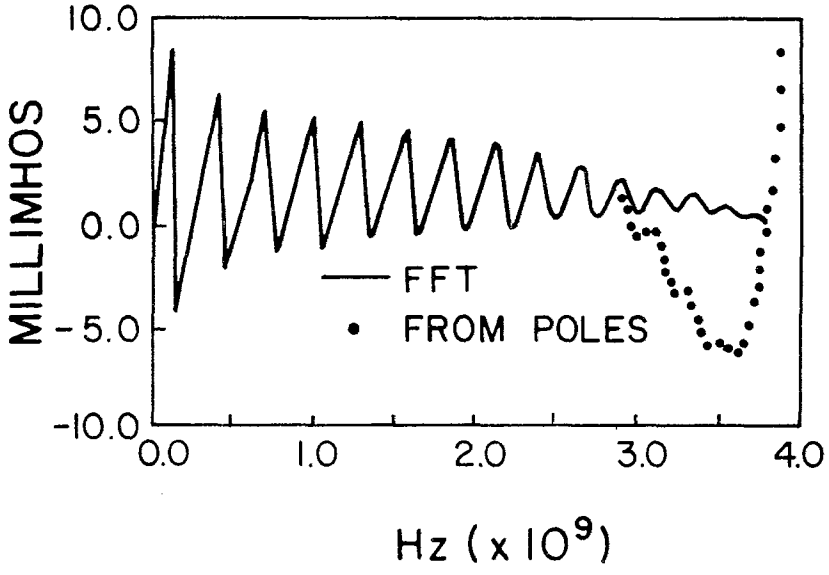
Thus, the frequency domain response can be obtained directly from the time domain response without having to perform a Fourier transform.

The real advantage is that in order to perform a Fourier transform accurately it is necessary to have a very long time history of the transient response. That is, it is usually necessary to have enough time samples so that the transient response has decayed to zero or to its steady-state value. The examples of Section 3.3 showed that Prony's method is capable of extracting the poles from a time window that is extremely short when compared to the whole response. Thus, if it is too expensive to calculate or measure more than just a few transient data samples, it is generally impossible to perform an accurate Fourier transform, but if Prony's algorithm is used, it is then possible to extract the spectral characteristics from this narrow time window.

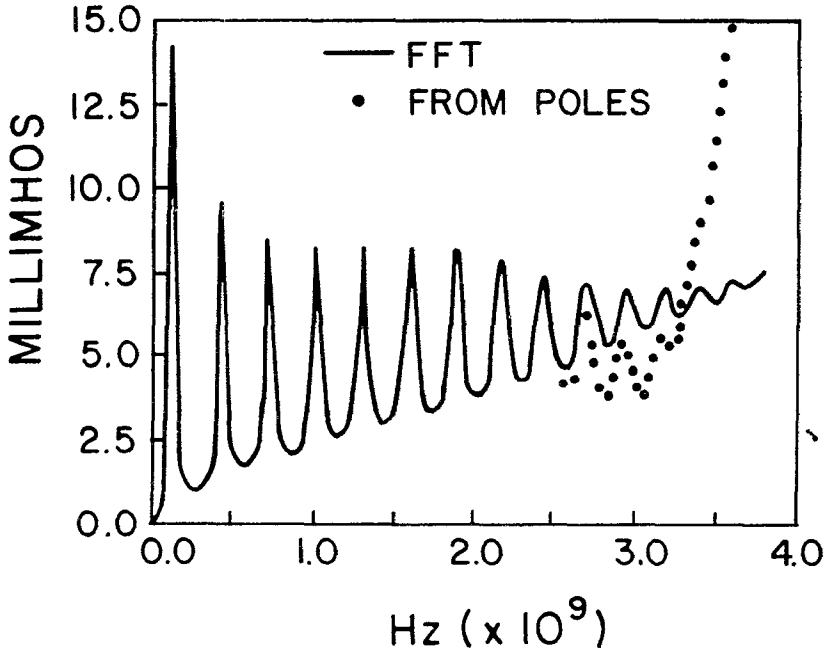
If the transient response is not the impulse response and the impulse response or the frequency domain transfer function is desired, it is possible to perform the deconvolution after the poles have been found. For example, if Prony's method gives a frequency domain response function of the form of (6.4) and if the frequency domain response of the driving function F(s) is known, it is possible to obtain the frequency domain transfer function:

$$H(s) = \frac{\sum_{i=1}^{N} \frac{A_i}{s - s_i}}{F(s)} \quad . \tag{6.5}$$

Note the similarities between (6.5) and (6.3), the expression for the input admittance of an antenna. If F(s) is the applied voltage at the driving point and the A_i and the s_i are the residues and poles for the induced current at the driving point, then H(s) is just the driving-point admittance. Figures 6.5a and 6.5b show the imaginary and real



(a)



(b)

Figure 6.5. (a) Imaginary part of input admittance for 1.0 m dipole.
(b) Real part of input admittance for 1.0 m dipole.

parts, respectively, of the input admittance of the 1.0 m dipole discussed in Section 3.3. The solid line represents the admittance obtained by using the conventional Fast Fourier Transform, and the dotted line is the admittance obtained using Expression (6.3). Since the driving function was a Gaussian pulse, it was possible to express its Laplace transform F(s) analytically. Note that the two methods give results that compare closely for all but the higher frequencies.

Similarly, if the normalized radar cross section of a scatterer is desired, it can be obtained by using the expression

$$\frac{\sigma}{\lambda^2} = \frac{\omega^2 r^2}{\pi c^2} \frac{\left|\bar{E}_{rad}(\omega)\right|^2}{\left|\bar{E}^A(\omega)\right|^2} \tag{6.6}$$

where the dependence on the incidence and observation angles and polarization has been suppressed. The $\bar{E}^A(\omega)$ term is simply the frequency domain response of the applied incident field and $\bar{E}_{rad}(\omega)$ is the frequency domain response of the scattered field measured at distance r from the scatterer. $\bar{E}_{rad}(\omega)$ can be obtained using Prony's method on the calculated or measured scattered transient response data.

When a conventional Fourier transform routine is used, the result is always a set of frequency domain data points at a specified frequency interval which is dependent on the time step size of the transient data used. Many times the study of the spectral characteristics at just a few select frequencies is of interest. If any of the analytical Expressions (6.3), (6.4), (6.5), or (6.6) are used, it is possible to calculate the spectral characteristics at any frequency desired. Thus, if Prony's method is used, it is possible to obtain as few or as many frequency domain samples at any frequency of interest, within the bandwidth of the system.

6.4 Data Reduction and Extrapolation

From the discussions and examples mentioned earlier, it is obvious that Prony's method takes a set of transient response data and reduces it to a small set of poles and residues. These poles and residues can then be used to reconstruct the transient response for all time, including those time steps past the end of the data window used. Hence, Prony's method is a valuable method for data reduction and extrapolation. Anyone who has ever generated transient response waveforms by numerical or experimental procedures soon learns that the storage problem required in keeping all of the data is very large. If Prony's method is used, it is not only possible to reduce the amount of data storage required but it is also possible to reduce the amount of transient data produced in the first place. The following example will demonstrate this adequately.

Consider again the experimentally produced data which were discussed in Section 3.3, Example 9. The experimental transient range produced 512 data samples over the time window shown in Figure 3.18. Prony's method was then applied to only 100 of these time samples at every fifth time step to produce the forty-one poles shown in Figure 3.19. The forty-one poles and residues were then used to reproduce the measured transient response and to extrapolate the measured response to very late time, 100 ns. This reproduced response is shown in Figure 6.6. Thus, using only eighty-two complex numbers, the original measured response was reproduced and extrapolated to add additional information. Remember also that only 100 of the original measured samples were needed.

Hence, a response can be measured for a fairly short period of time; then, Prony's method can be applied to the response to reduce it to a set of poles and residues. The poles and residues can then be used to

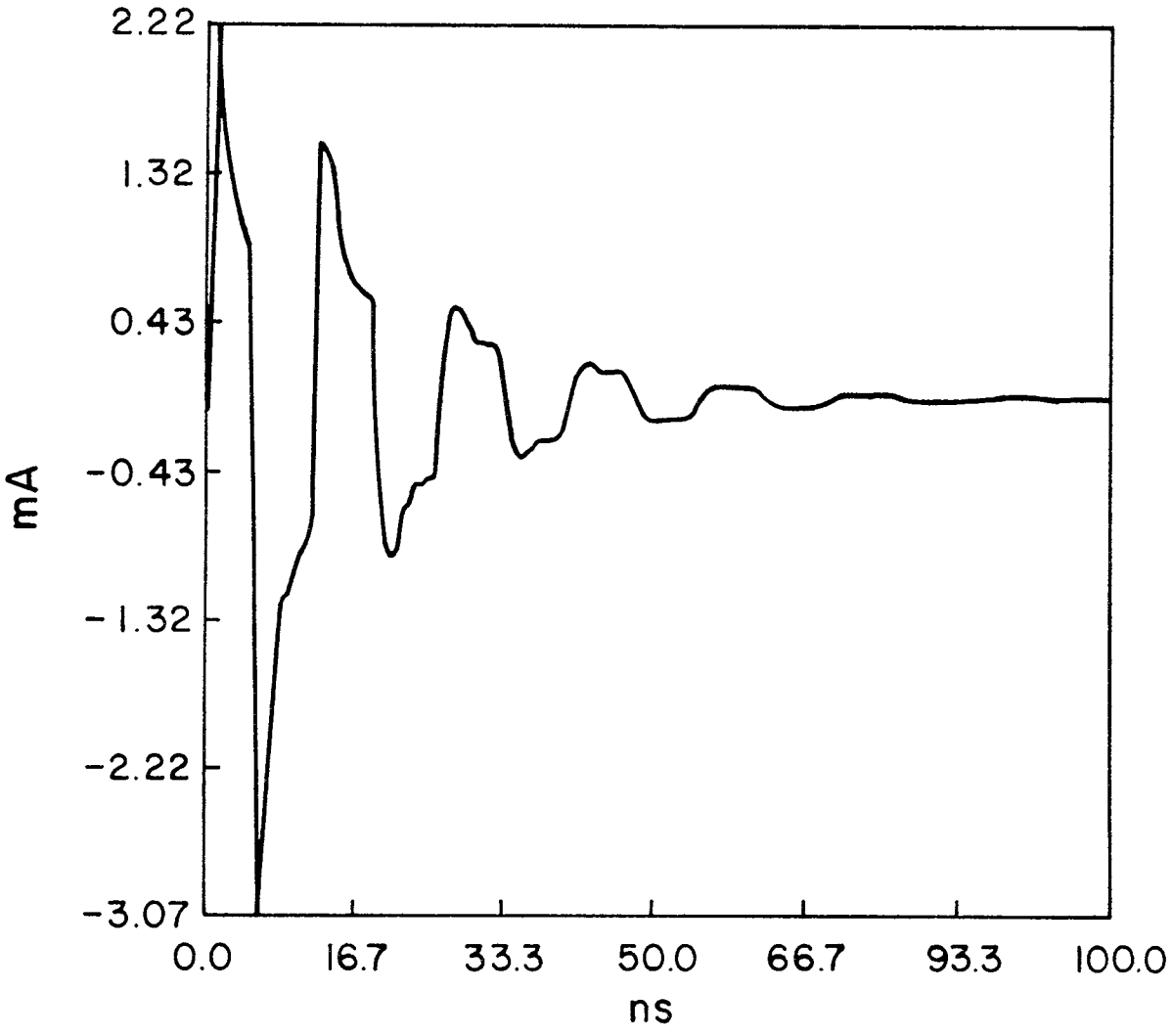


Figure 6.6. Extrapolated current response for 1.0 m monopole of Example 9, Chapter 3.

reconstruct the original data, extrapolate the original data to later time values, and produce the spectral content by using the methods of Section 6.3.

6.5 Other Applications of Prony's Method

This chapter dealt with applications of Prony's method for extracting the singularities of a system from its transient response. There are many other uses for Prony's method which have been developed in the past. These uses range from representation of electrocardiograms [17] to the measuring of the vertical angles of arrival of HF sky-wave signals [34]. These applications are very interesting and for that reason an additional bibliography is provided which lists many of the papers in which Prony's method has been used. The techniques developed in this thesis could be used to eliminate many of the problems discussed in the papers.

7. AN ALTERNATIVE TO PRONY'S METHOD

In previous chapters, Prony's method has been the only technique discussed for extracting the poles from a set of transient response data. However, another approach, known as the Padé approximation method [35], can be used. The Padé method will be discussed in this chapter and will be shown to be extremely limited in its usefulness.

7.1 Padé Approximation Method

If the Laplace transform R(s) of the transient response function R(t) is analytic at the origin, then R(s) can be represented in a Taylor's series expansion about the origin. The Taylor's series can be denoted by

$$R(s) = r_0 + r_1 s + r_2 s^2 + \ldots + r_k s^k + \sum_{i=k+1}^{\infty} r_i s^i \quad . \tag{7.1}$$

A rational function in s can always be found such that its Taylor's expansion has the same leading terms as those of (7.1). This rational function is known as the Padé approximant of R(s). The Padé approximant is usually written in the form

$$\frac{P(s)}{Q(s)} = \frac{a_0 + a_1 s + a_2 s^2 + \ldots + a_m s^m}{b_0 + b_1 s + b_2 s^2 + \ldots + b_N s^N} \quad . \tag{7.2}$$

Note that if the transient response contains exactly N poles then Expression (7.2) is an analytical expression for the Laplace transform of the sum of exponentials

$$R(t) = \sum_{i=1}^{N} A_i e^{s_i t} \quad . \tag{7.3}$$

Note also that if the coefficients b_i of the denominator of (7.2) can be obtained then it is a simple procedure to find the poles of the system. All that would need to be done is to find the roots of the N^{th} order polynomial

$$b_0 + b_1 s + b_2 s^2 + \ldots + b_N s^N = \prod_{i=1}^{N} (s - s_i) = 0 \quad . \tag{7.4}$$

The coefficients b_i of (7.3) can be obtained in the Padé approximant sense by setting the first k + 1 terms of (7.1) equal to (7.2). That is, let

$$r_0 + r_1 s + r_2 s^2 + \ldots + r_k s^k = \frac{a_0 + a_1 s + a_2 s^2 + \ldots + a_m s^m}{b_0 + b_1 s + b_2 s^2 + \ldots + b_N s^N} \tag{7.5}$$

where the value of k must be equal to m + N. If the denominator of the right side of (7.5) is multiplied by the left term and like powers of s are equated, the following set of equations is obtained.

$$\left.\begin{aligned} a_0 &= r_0 b_0 \\ a_1 &= r_1 b_0 + r_0 b_1 \\ &\vdots \\ a_m &= r_m b_0 + r_{m-1} b_1 + \ldots + r_0 b_m \end{aligned}\right\} \tag{7.6a}$$

$$\left.\begin{aligned} r_{m+1} b_0 &= r_m b_1 + \ldots + r_{m-N+2} b_{N-1} + r_{m-N+1} b_N \\ r_{m+2} b_0 &= r_{m+1} b_1 + \ldots + r_{m-N+3} b_{N-1} + r_{m-N+2} b_N \\ &\vdots \\ r_{m+N} b_0 &= r_{m+N-1} b_1 + \ldots + r_{m+1} b_{N-1} + r_m b_N \end{aligned}\right\} \quad . \tag{7.6b}$$

If in (7.6b) the b_0 terms are set equal to one, the N remaining b_i may be solved in terms of the r_i. Then the a_i in (7.6a) can be solved in terms of the b_i and the r_i.

The remaining problem is the determination of the r_i which are the coefficients of the Taylor's expansion of the transient response. These can be obtained if the transient waveform is modeled with a set of step functions $u(t - t_i)$ and a set of ramp functions, $u_{-1}(t - t_i)$. The step function and the ramp function have Laplace transforms which are known. These are

$$\mathcal{L}[u(t - t_i)] = \frac{1}{s} e^{-st_i} \tag{7.7a}$$

$$\mathcal{L}[u_{-1}(t - t_i)] = \frac{1}{s^2} e^{-st_i} \tag{7.7b}$$

where the time t_i is the turn-on time for the step and the ramp functions. Hence, if the transient response can be modeled with a set of step and ramp functions, then its Laplace transform can be expressed as a sum of the terms (7.7a) and (7.7b). As a simple example of the modeling procedure consider the portion of a transient response shown as a dotted line in Figure 7.1. This response has been modeled as a sum of step and ramp functions as indicated by the solid lines. The analytical time domain expression for the model can then be written for this example as

$$R(t) = 13\, u_{-1}(t) - 13\, u_{-1}(t - 1) + 13\, u(t - 1)$$
$$+ 3\, u_{-1}(t - 1) - 3\, u_{-1}(t - 2) - 8\, u_{-1}(t - 2) + \ldots \tag{7.8}$$

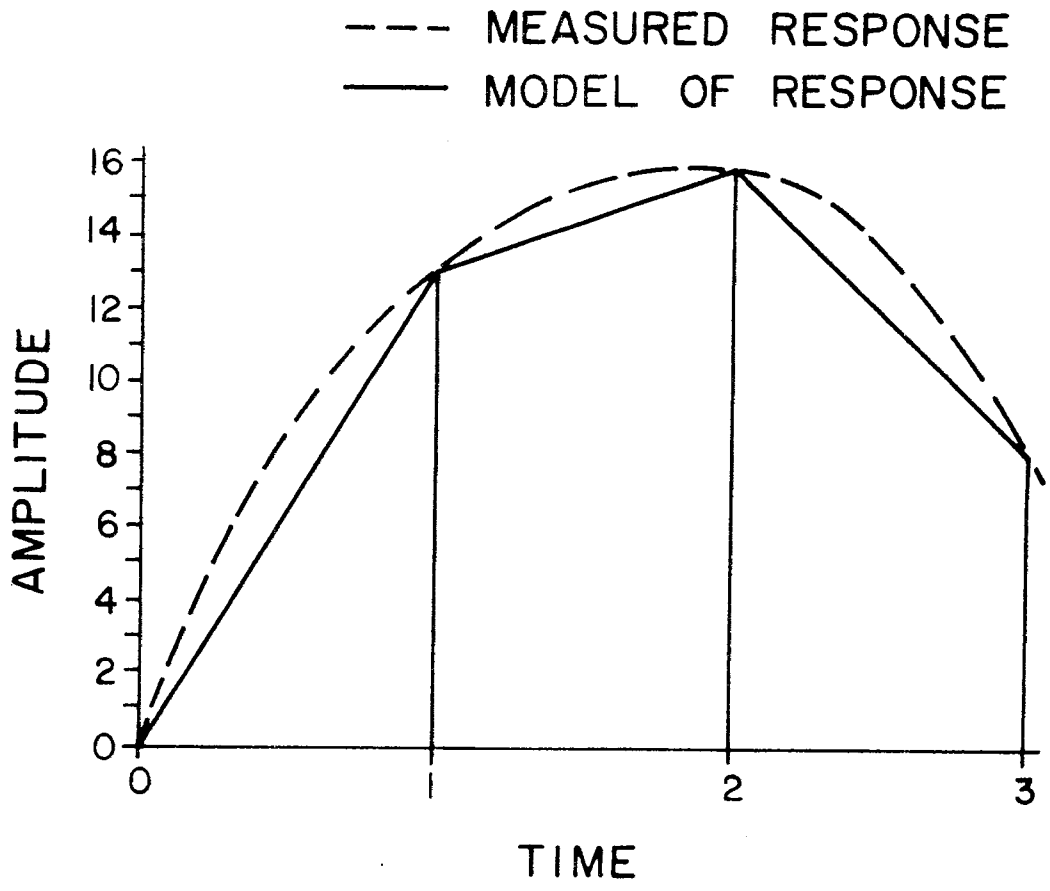


Figure 7.1. Model of a transient waveform using step and ramp functions.

Expression (7.8) then has a Laplace transform of

$$R(s) = \frac{13}{s^2} - \frac{13}{s^2} e^{-s} + \frac{13}{s} e^{-s} + \frac{3}{s^2} e^{-s}$$

$$- \frac{3}{s^2} e^{-2s} - \frac{8}{s^2} e^{-2s} + \ldots + \qquad (7.9a)$$

which reduces to

$$R(s) = \frac{13}{s^2} + e^{-s}\left(- \frac{10}{s^2} + \frac{13}{s}\right) + e^{-2s} \left(- \frac{11}{s^2}\right) + \ldots + \quad . \qquad (7.9b)$$

The exponential terms in (7.9) which result from the time shift have a Taylor series expansion of

$$e^{-t_i s} = 1 + (-t_i s) + \frac{(-t_i s)^2}{2!} + \frac{(-t_i s)^3}{3!} + \ldots + \quad . \qquad (7.10)$$

If this expansion is substituted into the expression for the Laplace transform of the pulse and ramp models of the transient response, then the response can be written in the form of (7.1). This expansion should then be truncated to include only k + 1 terms as needed to solve the set of Equations (7.6b). It is now a simple procedure to obtain the unknown coefficients b_i by solving (7.6b). The poles are then obtained as the roots of the polynomial (7.4) which has the coefficients b_i. After the a_i have been solved for from (7.6a), residues of the poles can be obtained by performing a partial fraction expansion on the rational function (7.2).

The above approach for finding the poles appears to be a simple procedure but it will be shown in the next section that because of the approximations made this technique is only useful for systems containing a couple of poles.

7.2 Examples of the Padé Procedure

Example 1

As the first example of the Padé procedure consider the response function

$$R(t) = e^{-2t} \sin \pi t \tag{7.11}$$

which has been discussed previously as Example 2 of Section 5.3 and plotted as Figure 5.1. A total of 200 samples were generated and the response was modeled for 20, 40, 80, 100, 150, and 200 of these samples using the pulse and ramp approximations discussed in Section 7.1. The results of applying the Padé approximation to determine the poles are presented in Table 7.1. Note that the results are not satisfactory until eighty samples were used. Figure 5.1 shows that the response does not damp to nearly zero until after time step forty. In order to get valid results all of the transient response must be used until the time at which the response is nearly equal to zero. The results for eighty through 200 samples are promising however.

Example 2

Here a transient response was generated using two sets of complex pole pairs. The response used is

$$r(t) = e^{-2t} \sin \pi t + e^{-3t} \sin 1.5 \pi t \quad . \tag{7.12}$$

The results of the Padé approximation procedure for this transient response are tabulated in Table 7.2. Note that in this case satisfactory results were not obtained until 200 samples were used even though the response has damped to zero at around the fiftieth time step. This indicates that a larger number of samples is needed to resolve the presence

TABLE 7.1

THE RESULTING POLES USING THE PADÉ APPROXIMATION ON THE TRANSIENT RESPONSE OF EXAMPLE 1: $R(r) = e^{-2t} \sin \pi t$

Number of Samples Used	Poles Extracted	Percent Error	
		Real Part	Imaginary Part
20	-1.5009 ± j 2.1523	13.40	26.56
40	-1.9076 ± j 2.8563	2.48	7.66
80	-1.9998 ± j 3.1405	0.004	0.004
100	-2.0001 ± j 3.1414	0.002	0.004
150	-2.0001 ± j 3.1414	0.002	0.003
200	-2.0001 ± j 3.1414	0.002	0.003

TABLE 7.2

THE RESULTING POLES USING THE PADÉ APPROXIMATION ON THE TRANSIENT RESPONSE OF EXAMPLE 2: $R(t) = 2^{-2t} \sin \pi t + e^{-3t} \sin 1.5 \pi t$

Number of Samples Used	Poles Extracted	Percent Error	
		Real Part	Imaginary Part
100	-1.5734 ± j 3.2773	11.45	3.64
	-2.1637 ± j 3.8707	14.97	15.06
200	-1.9994 ± j 3.1408	0.017	0.078
	-2.9970 ± j 3.7016	0.053	0.192

of two pole pairs than for only one pole pair. The accuracy, however, for 200 samples is still very satisfactory.

Example 3

As a final example consider the transient response made up of three pole pairs such that

$$R(t) = e^{-2t} \sin \pi t + e^{-3t} \sin 1.5 \pi t + e^{-4t} \sin 2\pi t \quad . \tag{7.13}$$

This signal is similar to those of Examples 1 and 2 in that it decays to very near zero at about the fiftieth time step. The results of applying the Padé approximation to this response are tabulated in Table 7.3. Note that satisfactory results are never really obtained for the imaginary parts of the second and third pole pairs, because the approximations used in the procedure cause the Padé method to have difficulty in extracting more than two sets of poles. This, of course, is a very limiting factor since in almost all cases one would desire to obtain at least six pole pairs.

7.3 Conclusions Regarding the Padé Approximation

It has been shown that the Padé approximation performed satisfactorily when only one or two pole pairs were desired but failed to give satisfactory poles when three pole pairs were sought. The method was also applied to signals with more than three pole pairs and which did not damp out so quickly, but again the results were unsatisfactory. For the cases of one and two pole pairs it was necessary to include essentially all of the waveform until it damped to zero. This in itself is a limiting factor. Thus, the conclusion would then be that the Padé approximation, as it was used here, is not a satisfactory method for extracting the true poles from a set of transient response data.

TABLE 7.3

THE RESULTING POLES USING THE PADÉ APPROXIMATION ON THE TRANSIENT RESPONSE OF EXAMPLE 3: $R(t) = e^{-2t} \sin \pi t + e^{-3t} \sin 1.5 \pi t + e^{-4t} \sin 2 \pi t$

Number of Samples Used	Poles Extracted	Percent Error	
		Real Part	Imaginary Part
100	-0.9671 ± j 0.2879	27.73	76.62
	-2.0313 ± j 3.1523	17.34	27.93
	-3.5884 ± j 5.3760	5.53	12.18
	-1.9991 ± j 3.1404	0.023	0.032
250	-2.9915 ± j 4.5577	0.152	2.768
	-4.0244 ± j 6.0014	0.328	3.783
	-1.9991 ± j 3.1404	0.023	0.032
400	-2.9915 ± j 4.5577	0.153	2.768
	-4.0244 ± j 6.0014	0.327	3.782

8. CONCLUSIONS

The numerical methods described in this thesis provide a technique by which the true complex natural resonances of a system may be extracted from a set of discrete transient response data of that system. The numerical procedure, known as Prony's algorithm, is applicable to systems possessing multiple poles as well as simple poles. There are two efficient and systematic methods by which it is possible to determine the number of poles inherent in the transient response. In employing both of these schemes the poles are obtained as a by-product of the calculations. The Householder orthogonalization method, which is the most systematic of the two procedures for determining the number of poles, breaks down if any significant noise is added. The eigenvalue procedure on the other hand uses the known standard deviation of the noise to aid in determining the number of poles but is not quite as systematic as the orthogonalization procedure.

The problem of noise in the transient response is a critical point. As mentioned above, the standard deviation of the noise is used to aid in determining the number of poles. However, Prony's method will not extract the true poles if the signal-to-noise ratio is below a certain level. This level appears to be in the 10 to 20 dB range for the individual signal components comprising the entire signal. The signal-to-noise level for the total signal needs to be as good as 30 dB. This fact causes Prony's method to be limited to use with extremely clean data systems. However, if several sets of noisy transient responses are measured and averaged together, the standard deviation of the noise decreases as one over the square root of the number of trials run.

This approach could be applied to laboratory test systems where a test can be repeated as many times as desired. More work should be done with actual noisy experimental data where the data are measured specifically for the application of Prony's method. The results would then allow for firm conclusions to be made about the influence of experimental noise on Prony's method.

There are several applications of Prony's method including system analysis, radar target recognition, the study of spectral characteristics, and data reduction and extrapolation. Prony's method is an invaluable tool in data reduction and in the determination of the spectral characteristics of a system. The application of the method to radar target recognition is also very exciting but seems to be limited by the noise level problem. More work should be done in applying Prony's method to the areas mentioned as well as to new areas.

The Padé approximation is an alternative to Prony's method. The study of this method shows that the Padé approximation is not applicable to systems containing more than one or two pole pairs. This restricts its usefulness for SEM related problems. No other alternatives to Prony's method were found.

The whole field of SEM is new and is growing so fast that the outlook for use of the techniques presented and developed in this thesis is very promising. It is hoped that these techniques will help simplify the study and the analysis of the more complicated electromagnetic structures.

REFERENCES

[1] C. E. Baum, "On the singularity expansion method for the solution of electromagnetic interaction problems," Interaction Note 88, December 11, 1971.

[2] C. E. Baum, "The singularity expansion method," in Transient Electromagnetic Fields, L. B. Felsen, Ed. Springer, New York: 1976, ch. 3.

[3] L. Page and N. Adams, "The electrical oscillations of a prolate spheroid, Paper 1," Phys. Rev., vol. 33, pp. 819-831, 1938.

[4] E. Hallén, "Über die elektrischen schwingungen in drohtförmigen leitern," Uppsala University Arssker., no. 1, pp. 1-102, 1930.

[5] F. M. Tesche, "On the Singularity Expansion Method as Applied to Electromagnetic Scattering from Thin-Wires", Interaction Note 102, April 1972.

[6] L. Marin, Natural-Mode Representation of Transient Scattering from Rotationally Symmetric, Perfectly Conducting Bodies and Numerical Results for a Prolate Spheroid, Interaction Note 119, September 1972.

[7] D. R. Wilton and K. R. Umashankar, "Parametric study of an L-shaped wire using the singularity expansion method," Interaction Note 152, November 1973.

[8] E. K. Miller, A. J. Poggio and G. J. Burke, "An integro-differential equation technique for the time domain analysis of thin-wire structure-the numerical method," J. Comput. Phys., vol. 12, p. 24, May 1973.

[9] C. L. Bennett, "A technique for computing approximate electromagnetic impulse response of conducting bodies," Interaction Note 222, June 22, 1968.

[10] A. M. Nicolson, C. L. Bennett, D. Lamensdorf and L. Susman, "Applications of time-domain metrology to the automation of broadband microwave measurements," IEEE Microwave Theory Tech., vol. MTT-20, no. 1, p. 3, January 1972.

[11] R. A. Anderson, J. A. Landt, F. J. Deadrich, E. K. Miller, "The LLL transient electromagnetic measurement facility: A brief description," Lawrence Livermore Laboratory, Livermore, California, UCID-16573, August 9, 1974.

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

[13] E. Wittaker and G. Robinson, The Calculus of Observations. New York: Dover, 1967, pp. 369-371.

[14] F. B. Hildebrand, Introduction to Numerical Analysis. New York: McGraw-Hill, 1956, pp. 378-382.

[15] P. Henrici, Applied and Computational Complex Analysis, vol. 1, New York: John Wiley, 1974, p. 655.

[16] R. N. McDonough and W. H. Huggins, "Best least-squares representation of signals by exponentials," IEEE Trans. Automat. Contr., vol. AC-13, no. 4, pp. 408-412, August 1968.

[17] T. Y. Young and W. H. Huggins, "On the representation of electrocardiograms," IEEE Trans. Bio-Med. Electronics, vol. BME-10, no. 3, pp. 86-95, July 1963.

[18] M. L. Van Blaricum and R. Mittra, "A technique for extracting the poles and residues of a system directly from its transient response," Interaction Note 245, February 1975.

[19] J. D. Lawrence, "Comparison of polynomial root finding methods," Lawrence Livermore Laboratory, Livermore, California, CIC Note C212-010, January 1966.

[20] J. D. Lawrence, "Polynomial root finder," Lawrence Livermore Laboratory, Livermore, California, CIC Report C212-001, December 1966.

[21] L. Weiss and R. N. McDonough, "Prony's method, z-transforms, and Padé approximations," SIAM Rev., vol. 5, no. 2, pp. 145-147, April 1963.

[22] M. S. Corrington, "Simplified calculations of transient response," Proc. IEEE, vol. 53, no. 3, pp. 287-292, March 1965.

[23] F. M. Tesche, "Application of the singularity expansion method to the analysis of impedance loaded linear antennas," Sensor and Simulation Note 177, May 1973.

[24] M. L. Van Blaricum, "A numerical technique for the time-dependent solution of thin-wire structures with multiple junctions," Electromagnetics Laboratory Report 73-15, University of Illinois, Urbana, Illinois, December 1973.

[25] E. K. Miller and M. L. Van Blaricum, "The short-pulse response of a straight wire," IEEE Trans. Antennas Propagat., vol. AP-21, no. 3, pp. 396-398, May 1973.

[26] B. P. Lathi, Signals, Systems and Communication. New York: John Wiley and Sons, 1965.

[27] A. S. Householder, "On Prony's method of fitting exponential decay curves and multiple-hit survival curves," Oak Ridge National Laboratory Report ORNL-455, Oak Ridge, Tennessee, February 1950.

[28] L. E. McBride, Jr., H. W. Schaefgen and K. Steiglitz, "Time-domain approximation by iterative methods," IEEE Trans. Circuit Theory, vol. CT13, no. 4, pp. 381-387, December 1966.

[29] W. Tabbara, "On an inverse scattering method," IEEE Trans. Antennas Propagat., vol. AP-21, no. 2, pp. 245-247, March 1973.

[30] C. L. Bennett, et al., "Space-time integral equation approach to the large body scattering problem," Interaction Note 233, January 1973.

[31] R. H. Mains and D. L. Moffatt, "Complex natural resonances of an object in detection and discrimination," TR-3424-1, Ohio State University, Electroscience Laboratory, Columbus, Ohio, June 1974.

[32] L. W. Pearson, M. L. Van Blaricum and R. Mittra, "A new method for radar target recognition based on the singularity expansion for the target," 1975 IEEE International Radar Conference, Arlington, Virginia, pp. 452-456, April 21-23.

[33] A. J. Berni, "Target identification by natural resonance estimation," IEEE Trans. Aerosp. Electron. Syst., vol. AES-11, no. 2, pp. 147-154, March 1975.

[34] J. M. Kelso, "Measuring the vertical angles of arrival of HF skywave signals with multiple modes," Radio Sci., vol. 7, no. 2, pp. 245-250, February 1972.

[35] K. L. Su, Time-Domain Synthesis of Linear Networks. Englewood Cliffs, New Jersey: Prentice Hall, 1971, pp. 23-26.

ADDITIONAL REFERENCES ON PRONY'S METHOD

R. A. Buckingham, Numerical Methods. London: Sir Isaac Pitman and Sons, 1957, pp. 329-333.

C. W. Chuang and D. L. Moffatt, "Complex natural resonances of radar targets via Prony's method," 1975 URSI Meeting, University of Illinois, Urbana, Illinois, June 3-5, pp. 67-68.

F. J. Deadrick, et al., "Object discrimination via pole extraction from transient fields," Program of the 1975 URSI Meeting, University of Illinois, Urbana, Illinois, June 3-5, p. 67.

C. D. Gardner, "Expals-fortran code for exponential approximation by least squares," UCRL-14541, Lawrence Livermore Laboratory, University of California, Livermore, California, March 1970.

R. W. Hamming, Numerical Methods for Scientists and Engineers. 2nd Edition, New York: McGraw-Hill, 1973, pp. 617-627.

D. W. Kammler, "Prony's method for completely monotonic functions," Southern Illinois University, Carbondale, Illinois, Submitted for Publication.

L. G. Kelly, Handbook of Numerical Methods and Applications. Reading, Massachusetts: Addison-Wesley, 1967, pp. 80-82.

C. Lanczos, Applied Analysis. Englewood Cliffs, New Jersey: Prentice Hall, 1956, pp. 272-280.

R. N. McDonough, "Representation and analysis of signals, Part XV, matched exponents for the representation of signals," The Johns Hopkins University, Baltimore, Maryland, AD-411-431, April 1963.

J. Makhoul, "Linear prediction: a tutorial review," Proc. IEEE, vol. 63, no. 4, pp. 561-580, April 1975.

A. Marzollo, "On the mean square approximation of a function with a linear combination of exponentials," Int. J. Contr., vol. 9, no. 1, pp. 17-26, 1969.

G. Miller, "Representation and analysis of signals; Part XXVI, least-squares approximation of functions by exponentials," Department of Electrical Engineering, The Johns Hopkins University, Baltimore, Maryland, June 1969.

R. Mittra and C. O. Stearns, "Field-strength measurements in a multipath field using linear and circular probing," Radio Sci., vol. 2, no. 1, pp. 101-110, January 1967.

A. K. Paul, "Anharmonic frequency analysis," Math. Comput., vol. 26, no. 118, pp. 437-447, April 1972.

W. Perl, "A method for curve-fitting exponential functions," Int. J. Appl. Radiation Isotopes, vol. 8, pp. 211-222, 1960.

F. R. Spitznogle and A. H. Quazi, "Representation and analysis of time-limited signals using a complex exponential algorithm," J. Acoust. Soc. Amer., vol. 47, no. 5, pp. 1150-1155, 1970.

P. Wynn, "A note on fitting of certain types of exponential data," Statistica Neerlandica, vol. 16, no. 2, pp. 143-150, 1962.

T. Y. Young, "Representation and analysis of signals, Part XVI, representation and detection of multiple-epoch signals," The Johns Hopkins University, Baltimore, Maryland, AD-405028, May 1963.