

PARAMETER ESTIMATION FROM FREQUENCY RESPONSE MEASUREMENTS USING RATIONAL FRACTION POLYNOMIALS

Mark H. Richardson & David L. Formenti
Structural Measurement Systems, Inc.
San Jose, California

ABSTRACT

This is a new formulation which overcomes many of the numerical analysis problems associated with an old least squared error parameter estimation technique. Overcoming these problems has made this technique feasible for implementation on mini-computer based measurement systems.

This technique is not only useful in modal analysis applications for identifying the modal parameters of structures, but it can also be used for identifying poles, zeros and resonances of combined electro-mechanical servo-systems.

INTRODUCTION

During the last ten years, a variety of FFT-based two channel digital spectrum analyzers have become commercially available which can make frequency response measurements. These analyzers are being used to make measurements on mechanical structures, in order to identify their mechanical resonances, or modes of vibration. Likewise, they are being used to measure the dynamic characteristics of electronic networks, and of combined electro-mechanical servo-systems.

One of the key advantages of the digital analyzers is that the measurements which they make, such as frequency response functions, are in digital form, i.e. computer words in a digital memory.

Therefore, these measurements can be further processed to identify the dynamic properties of structures and systems.

The frequency response function (FRF) is, in general, a complex valued function or waveform defined over a frequency range. (See Figure 4.) Therefore, the process of identifying parameters from this type of measurement is commonly called curve fitting, or parameter estimation.

This paper presents the results of an algorithm development effort which was begun back in 1976. At that time, we were looking for a better method for doing curve fitting in a mini-computer based modal analysis system. This type of system is used to make a series of FRF measurements on a structure, and then perform curve fitting on these measurements to identify the damped natural frequencies,

damping, and mode shapes of the predominant modes of vibration of the structure.

The three main requirements for a good curve fitting algorithm in a measurement system are 1) execution speed, 2) numerical stability, and 3) ease of use.

Previous to this development effort, we had experimented with a well known curve fitting algorithm called the Complex Exponential, or Prony algorithm. This algorithm has undergone a lot of refinement ([2], [3]) and is computationally very efficient and numerically stable in 16-bit machines. However, it curve fits the impulse response function instead of the FRF. The impulse response can be obtained by taking the Inverse Fourier transform of the FRF. When the FFT is used to obtain the impulse response from an FRF measurement, a potentially serious error can occur, which is called **wrap around error**, or **time domain leakage**. This error is caused by the truncated form (i.e. limited frequency range) of the FRF measurement, and distorts the impulse response as shown in Figure 1.

Hence, we sought to develop an algorithm with some of the same characteristics as the complex exponential method, (e.g., it is easy to use along with being numerically stable),

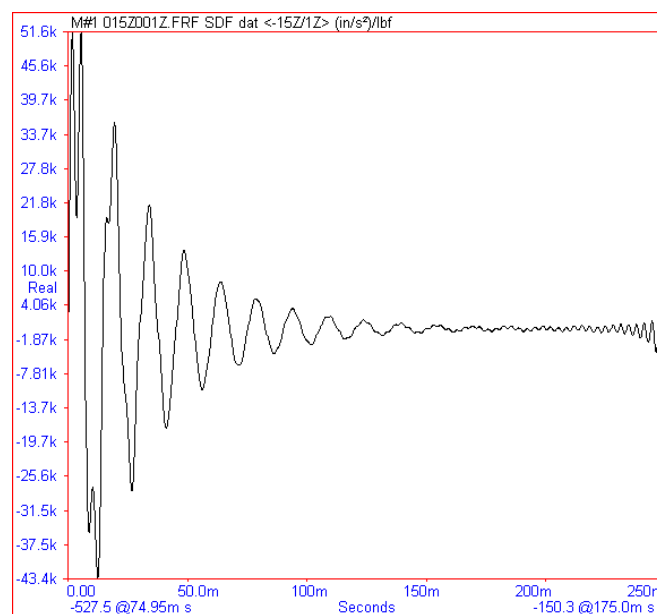


FIGURE 1. Impulse Response with Leakage.

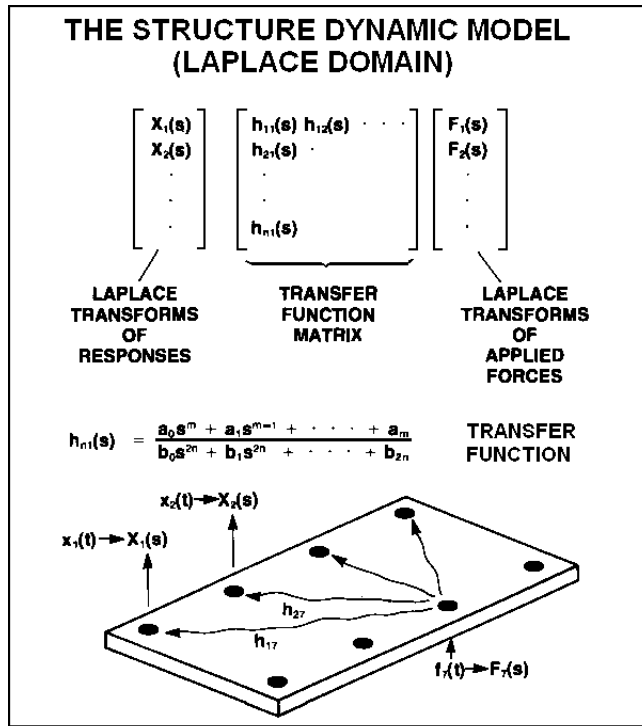


FIGURE 2

but that curve fits the FRF measurement data directly in the frequency domain.

If it is assumed that the frequency response measurement is taken from a linear, second order dynamical system, then the measurement can be represented as a ratio of two polynomials, as shown in Figure 6. In the process of curve fitting this analytical form to the measurement data, the unknown coefficients of both the numerator and denominator polynomials, $(a_k, k=0, \dots, m)$ and $(b_k, k=0, \dots, n)$, are determined. It is shown later that this curve fitting can be done in a least squared error sense by solving a set of linear equations, for the coefficients.

The greatest difficulty with curve fitting polynomials in rational fraction form is that the solution equations are ill-conditioned and hence are difficult, if not impossible to solve on a mini-computer, even for simple cases. To curve fit with an m^{th} order numerator and n^{th} order denominator polynomial, $(m+n+1)$ simultaneous equations must be solved. This is equivalent to inverting an $(m+n+1)$ matrix. Part of the problem stems from the dynamic range of the polynomial terms themselves. For instance, the highest order term in a 12^{th} order polynomial evaluated at a frequency of 1 kHz, is on the order of 10 to the 36^{th} power, which borders on the standard numerical capability of many 16 bit mini-computers.

We will see later that this problem can be minimized by re scaling the frequency values. However, this single step

doesn't change the ill-conditioned nature of the solution equations. We did find, though, that the use of orthogonal polynomials removes much of the ill-conditioning, and at the same time reduces the number equations to be solved to about half the number of equations of the ordinary polynomial case.

Much of the discussion in this paper, then, centers on the reformulation of the solution equations in terms of orthogonal polynomials, and generation of the polynomials themselves.

An alternative formulation, which yields an estimate of the characteristic polynomial from multiple measurements is also included.

Finally, some examples of the use of the curve fitter are given. Some of the problems which are common to all curve fitters, such as measurement noise, frequency resolution, and the effects of resonances which lie outside of the analysis band are discussed. companion paper (Reference [4]) discusses in more detail, these and other problems which can occur when curve fitting FRF measurements.

MODELING SYSTEM DYNAMICS IN THE LAPLACE DOMAIN

The dynamics of a mechanical structure can be modeled with a Laplace domain modal, as shown in Figure 2. In this model, the inputs and responses of the structure are represented by their Laplace transforms, Time domain derivatives (i.e. velocity and acceleration) do not appear explicitly in the Laplace domain model but are accounted for in the transfer functions of the structure. These transfer functions are contained in a transfer matrix and contain all of the information necessary to describe structural responses as functions of externally applied forces.

Using a Laplace domain model, a structure is excited by several different input forces, then its transformed response is a summation of terms, each term containing one of the transformed input forces multiplied by the transfer function between the input point (degree-of-freedom) and the desired response point.

Transfer Function of a Single Degree-of-Freedom System

The Laplace variable is a complex variable, normally denoted by the letter S . Since the transfer function is a function of the s -variable, it too is complex valued, (i.e. for every complex value of S , the transfer function value is a complex number). Plots of a typical transfer function on the S -plane are shown in Figure 3. Because it is complex,

the transfer function can be represented by its real and imaginary parts, or equivalently by its magnitude and phase. Note that the magnitude of the transfer function goes to infinity at two points in the s -plane. These singularities are called the Poles of the transfer function. These poles define resonant conditions on the structure which will "amplify" an input force. The location of these poles in the s -plane is defined by a Frequency and Damping value as shown in Figure 5.

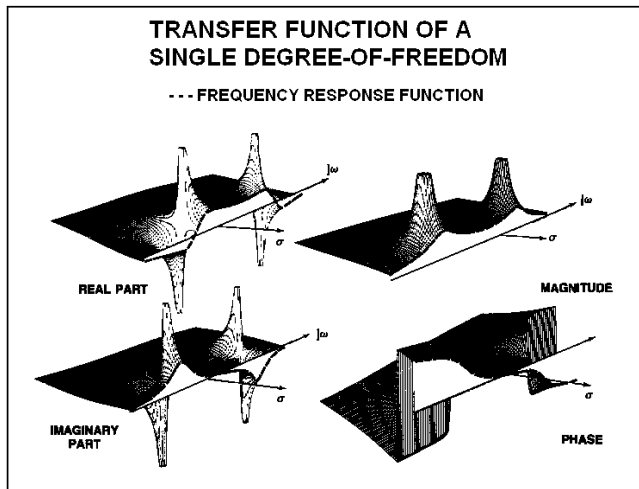


FIGURE 3

Hence the σ and $j\omega$ axes of the s -plane have become known as the damping axis and the frequency axis respectively. Note also in Figure 5 that the transfer function is only plotted over half of the s -plane, i.e. it is not plotted for any positive values of damping. This was done to give a clear picture of the transfer function values along the frequency axis.

The Frequency Response Function

In a test situation we do not actually measure the transfer function over the entire s -plane, but rather its values along the frequency axis. These values are known as the frequency response function, as shown in Figure 3. The analyzers compute the FRF by computing the Fourier transform of both the input and response signals, and then forming the ratio of response to input in the frequency domain. The resulting function is the same as evaluating the system's transfer function for $s=j\omega$. Since the transfer function is an "analytic" function, its values throughout the s -plane can be inferred from its values along the frequency axis.

A dynamic frequency domain model, similar to the Laplace domain model, can also be built using frequency response functions. The form of the frequency domain model is exactly the same as the Laplace domain model, but with frequency response functions replacing transfer functions

and Fourier transforms replacing Laplace transforms of the structural inputs and responses.

The frequency response function, being complex valued, is represented by two numbers at each frequency. Figure 4 shows some of the alternative forms in which this function is commonly plotted. The so called CoQuad plot, or real and imaginary parts vs. frequency, derives its origin from the early days of swept sine testing when the real part was referred to as the coincident waveform (that portion of the response that is in phase with the input) and the imaginary part as the quadrature waveform (that portion of the response that is 90 degrees out-of-phase with the input). The Bode plot, or log magnitude and phase vs. frequency, is named after H.W. Bode who made many contributions to the analysis of frequency response functions. (Many of Bode's techniques involved plotting these functions along a log frequency axis.)

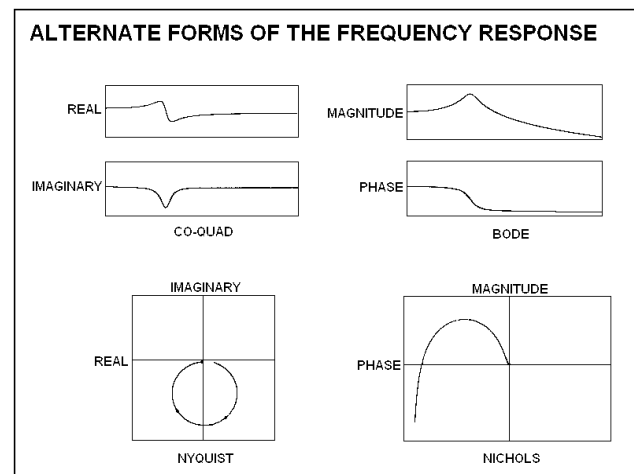


FIGURE 4

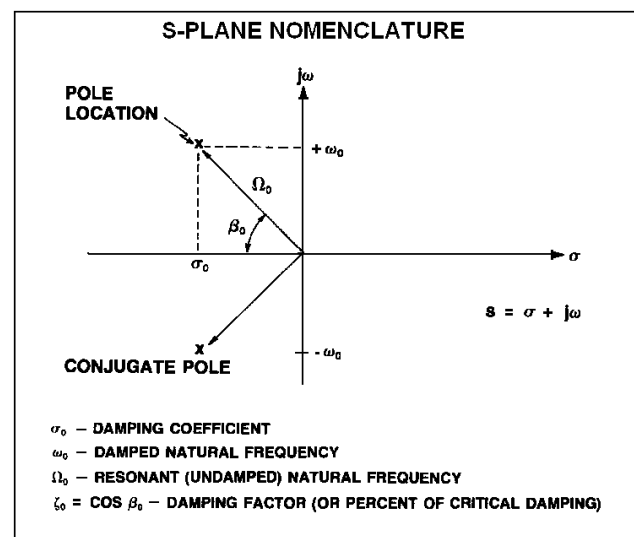


FIGURE 5

The Nyquist plot, or real vs. Imaginary part, is named after the gentleman who popularized its use for determining the stability of linear systems. The Nichols plot or log magnitude vs. phase angle is named after N.B. Nichols who used such plots to analyze servo-mechanisms.

It is important to realize that all of these different forms of the frequency response function contain exactly the same information, but are presented in different forms to emphasize certain features of the data.

ANALYTICAL FORMS OF THE FREQUENCY RESPONSE

The FRF can be represented either in rational fraction or partial fraction form, as shown in Figure 6.

Rational Fraction Form

The rational fraction form is merely the ratio of two polynomials, where in general the orders of the numerator and denominator polynomials are independent of one another. The denominator polynomial is also referred to as the characteristic polynomial of the system.

Recalling that the FRF is really the transfer function evaluated along the frequency axis, the poles of the transfer function correspond to values of the s -variable for which the characteristic polynomial is zero, i.e. the transfer function is infinite. These values of s are also called the roots of the characteristic polynomial.

Similarly, the roots of the numerator polynomial are the values of the s -variable where the transfer function is zero and are therefore called the zeros of the transfer function.

Hence, by curve fitting the analytical form in equation (1) to FRF data, and then solving for the roots of both the numerator and characteristic polynomials, the poles and zeros of the transfer function can be determined. Poles and zeros are typically used to characterize the dynamics of electronic networks and servo-systems,

Partial Fraction Form

For resonant systems, that is, systems where the poles are not located along the damping axis, the FRF can also be represented in partial fraction form. This form clearly shows the FRF in terms of the parameters which describe its pole locations. For a model with n -degrees-of-freedom, it is clear that the FRF contains n -pole pairs.

In this form, the numerator simply becomes a pair of constants, called residues, which also occur as complex conjugate pairs. Every pole has a different residue associated with it. In modal analysis, the unknown parameters of the partial fraction form, i.e. the poles and

residues, are used to characterize the dynamics of structures,

PROBLEM FORMULATION

The curve fitting problem consists of finding the unknown $(a_k, k=0, \dots, m)$ and $(b_k, k=0, \dots, n)$ such that the error between the analytical expression (1) in Figure 6 and an FRF measurement is minimized over a chosen frequency range. To begin the problem formulation, we need to define an error criterion.

**ANALYTICAL FORMS
OF THE FREQUENCY RESPONSE FUNCTION**

Rational Fraction Form

$$H(\omega) = \frac{\sum_{k=0}^m a_k s^k}{\sum_{k=0}^n b_k s^k} \bigg|_{s=j\omega} \quad (1)$$

Partial Fraction Form

$$H(\omega) = \sum_{k=1}^{n/2} \left[\frac{r_k}{s - p_k} + \frac{r_k^*}{s - p_k^*} \right] \bigg|_{s=j\omega} \quad (2)$$

$p_k = -\sigma_k + j\omega_k = k^{\text{th}} \text{ pole}$

$r_k = \text{residue for } k^{\text{th}} \text{ pole}$

FIGURE 6

Error Criterion

First we can write the error at a particular value of frequency as simply the difference between the analytical value and the measurement value of the FRF, as shown in expression (3) in Figure 7. Next, we can substitute for the analytical value in expression (5) by using expression (1). This leads to equation (4) which we will use as a measure of the error at each value of frequency.

Furthermore, we can make up an entire vector of errors, one for each frequency value where we wish to curve fit the data, as shown in expression (6). A squared error criterion can then be formed from the error vector, as shown in equation (5).

Notice that, in this case, the criterion is written for (L) frequency values, or data points. This criterion (J) will

always have a non-negative value. Therefore, we want to find values of the **a**'s and **b**'s so that the value of **J** is minimized, ideally zero.

ERROR CRITERION

Error at ith Frequency (ω_i)

$(e_i = (\text{Analytical FRF} - \text{Measured FRF}) \text{ at } \omega_i) \quad (3)$

or

$$e_i = \sum_{k=0}^m a_k (j\omega_i)^k - h_i \left[\sum_{k=0}^n b_k (j\omega_i)^k + (j\omega_i)^n \right] \quad (4)$$

where: h_i = FRF measurement data at ω_i

Squared Error Criterion

$$J = \sum_{i=1}^L e_i^* e_i = \{E^*\}^t \{E\} \quad (5)$$

Error Vector = $\{E\} = \begin{Bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{Bmatrix}$ (6)

*- denotes complex conjugate
t - denotes transpose

FIGURE 7

Minimizing the Error

Before continuing the analysis, we will re-write the error vector in a more compact vector-matrix form, by expanding each of the summations in equation (4). This yields the results shown in Figure 8.

Note that the **(m+1)** unknown numerator polynomial coefficients have now been assembled into a vector **{A}**, and that **(n)** the **(n+1)** unknown denominator coefficients have been assembled into the vector **{B}**. It is assumed from this point on that the highest order denominator coefficient is unity; (**b_n=1**). This can always be accomplished by factoring out the highest order coefficient (**b_n**) and dividing it into the other polynomial coefficients, thus yielding **b_n=1**. This accounts for the extra term (W) in expression (7).

THE ERROR VECTOR

$\{E\} = [P]\{A\} - [T]\{B\} - \{W\} \quad (7)$

where:

$$[P] = \begin{bmatrix} 1 & j\omega_1 & (j\omega_1)^2 & \cdots & (j\omega_1)^m \\ 1 & j\omega_2 & (j\omega_2)^2 & \cdots & (j\omega_2)^m \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & j\omega_L & (j\omega_L)^2 & \cdots & (j\omega_L)^m \end{bmatrix} \quad (L \times m+1)$$

$$[T] = \begin{bmatrix} h_1 & h_1(j\omega_1) & h_1(j\omega_1)^2 & \cdots & h_1(j\omega_1)^{n-1} \\ h_2 & h_2(j\omega_2) & h_2(j\omega_2)^2 & \cdots & h_2(j\omega_2)^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ h_L & h_L(j\omega_L) & h_L(j\omega_L)^2 & \cdots & h_L(j\omega_L)^{n-1} \end{bmatrix} \quad (L \times n)$$

$$\{A\} = \begin{Bmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{Bmatrix}, \quad \{B\} = \begin{Bmatrix} b_0 \\ b_1 \\ \vdots \\ b_{n-1} \end{Bmatrix}, \quad \{W\} = \begin{Bmatrix} h_1(j\omega_1)^n \\ h_2(j\omega_2)^n \\ \vdots \\ h_L(j\omega_L)^n \end{Bmatrix}$$

FIGURE 8

Using the error vector expression (7), we can now write the error criterion as shown in expression (8) in Figure 9. Note that it is now written as a function of the two unknown coefficient vectors **{A}** and **{B}**. If we could plot this function over the multi-dimensional space of **{A}** and **{B}**, it would have the shape of a bowl with a smooth rounded bottom and, most importantly, a single point at the bottom, i.e. a single minimum value.

Knowing that this criterion function has a single minimum value, we can set its derivatives (or slope) with respect to the variables **{A}** and **{B}** to zero to find the minimum point. These expressions ((9) and (10)) must be satisfied at the minimum point of the quadratic criterion function. Expression (9) is actually a set of **(m+1)** equations, and expression (10) is a set of **(n)** equations.

Since both expressions (9) and (10) contain the unknown variables **{A}** and **{B}** they must be solved together as an entire set of **(n+m+1)** equations. These equations are written in partitioned form in expression (11).

ERROR FUNCTION

$$\begin{aligned} J(\mathbf{A}, \mathbf{B}) = & \{\mathbf{A}\}^t [\mathbf{P}^*]^t [\mathbf{P}] \{\mathbf{A}\} + \{\mathbf{B}\}^t [\mathbf{T}^*]^t [\mathbf{T}] \{\mathbf{B}\} \\ & + \{\mathbf{W}^*\}^t \{\mathbf{W}\} - 2 \operatorname{Re}(\{\mathbf{A}\}^t [\mathbf{P}^*]^t [\mathbf{T}] \{\mathbf{B}\}) \\ & - 2 \operatorname{Re}(\{\mathbf{A}\}^t [\mathbf{P}^*] \{\mathbf{W}\}) - 2 \operatorname{Re}(\{\mathbf{B}\}^t [\mathbf{T}^*]^t \{\mathbf{W}\}) \end{aligned} \quad (8)$$

Necessary conditions at minimum pt of $J(\mathbf{A}, \mathbf{B})$

$$\begin{aligned} \frac{\partial J}{\partial \mathbf{A}} = & 2[\mathbf{P}^*]^t [\mathbf{P}] \{\mathbf{A}\} - 2 \operatorname{Re}([\mathbf{P}^*]^t [\mathbf{T}] \{\mathbf{B}\}) \\ & - 2 \operatorname{Re}([\mathbf{P}^*]^t \{\mathbf{W}\}) = \{0\} \end{aligned} \quad (9)$$

$$\begin{aligned} \frac{\partial J}{\partial \mathbf{B}} = & 2[\mathbf{T}^*]^t [\mathbf{T}] \{\mathbf{B}\} - 2 \operatorname{Re}([\mathbf{T}^*]^t [\mathbf{P}] \{\mathbf{A}\}) \\ & - 2 \operatorname{Re}([\mathbf{T}^*]^t \{\mathbf{W}\}) = \{0\} \end{aligned} \quad (10)$$

$\operatorname{Re}(\cdot)$ - denotes the real part of a complex no. ()

FIGURE 9**SOLUTION EQUATIONS**

$$\begin{bmatrix} \mathbf{Y} & \vdots & \mathbf{X} \\ \dots & \vdots & \dots \\ \mathbf{X}^t & \vdots & \mathbf{Z} \end{bmatrix} \begin{bmatrix} \mathbf{A} \\ \vdots \\ \mathbf{B} \end{bmatrix} = \begin{bmatrix} \mathbf{G} \\ \vdots \\ \mathbf{F} \end{bmatrix} \quad (11)$$

($m+n+1$) equations

where: $[\mathbf{X}] = -\operatorname{Re}([\mathbf{P}^*]^t [\mathbf{T}]) \quad (m+1 \times n)$

$[\mathbf{Y}] = [\mathbf{P}^*]^t [\mathbf{P}] \quad (m+1 \times m+1)$

$[\mathbf{X}] = [\mathbf{T}^*]^t [\mathbf{T}] \quad (n \times n)$

$\{\mathbf{G}\} = \operatorname{Re}([\mathbf{P}^*]^t \{\mathbf{W}\}) \quad (m+1 \text{ vector})$

$\{\mathbf{F}\} = \operatorname{Re}([\mathbf{T}^*]^t \{\mathbf{W}\}) \quad (n \text{ vector})$

FIGURE 10

Notice that the equations (11) are all real valued, which is expected since the unknowns $\{\mathbf{A}\}$ and $\{\mathbf{B}\}$ are real coefficient values.

In principle then, the least squares estimates of $\{\mathbf{A}\}$ and $\{\mathbf{B}\}$ can be obtained by solving the linear equations (11). Our experience though, and apparently that of others [5] is that these equations are generally ill-conditioned and hence

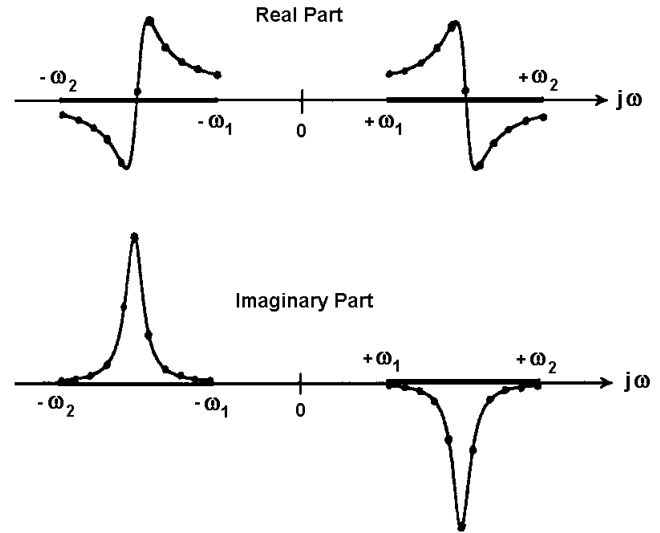
difficult to solve. We will therefore turn to a re-formulation of these equations using orthogonal polynomials.

ORTHOGONAL POLYNOMIALS

In the re-formulation of the curve fitting problem in terms of orthogonal polynomials we will take advantage of a special property of the FRF as well as the orthogonality property of the polynomials themselves to greatly simplify the calculations.

Normally, we only compute an FRF measurement in a digital analyzer for positive values of frequency, but in fact the function exists for negative frequencies as well. Furthermore, the FRF exhibits Hermitian symmetry about the origin of the frequency axis, as shown in Figure 11. That is, if we fold the real part of the FRF for positive frequencies about the origin, we obtain the real part for negative frequencies.

Similarly, if we fold the imaginary part of the FRF for positive frequencies about the origin, and change its sign, we obtain the imaginary part for negative frequencies.

HERMITIAN SYMMETRY OF THE FREQUENCY RESPONSE FUNCTION

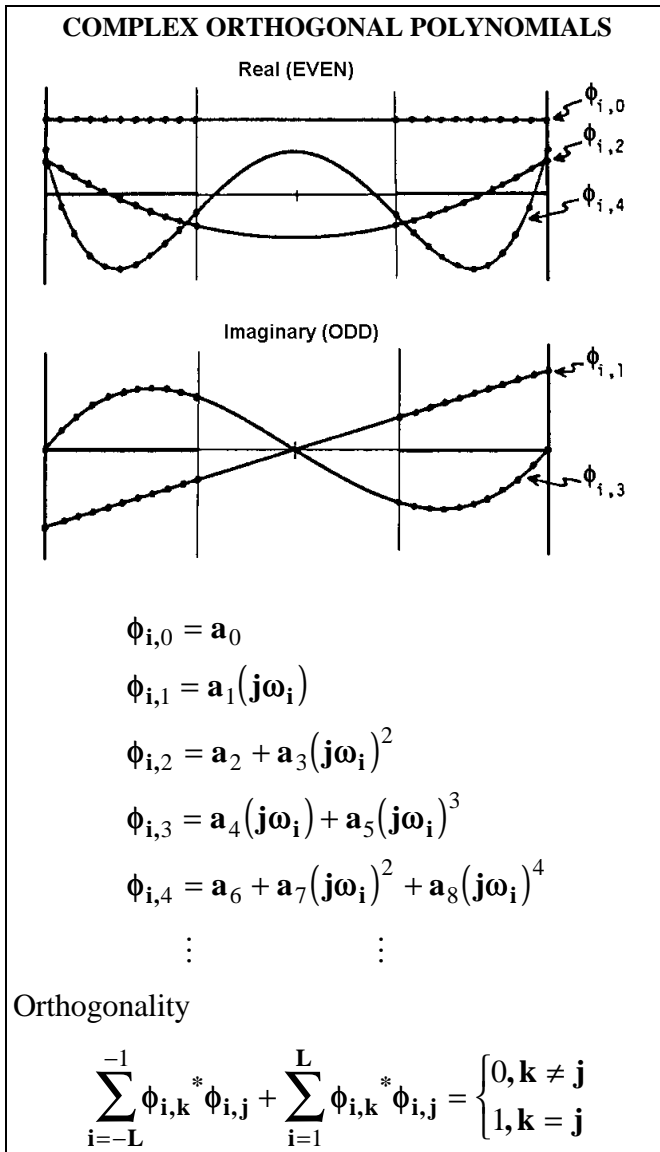
Symmetry: $\operatorname{Re}(\mathbf{h}_i) = \operatorname{Re}(\mathbf{h}_{-i})$
 $\operatorname{Im}(\mathbf{h}_i) = -\operatorname{Im}(\mathbf{h}_{-i})$

$(\mathbf{h}_i, i=1, \dots, L)$ are FRF values from $+\omega_1$ to $+\omega_2$

$(\mathbf{h}_{-i}, i=1, \dots, L)$ are FRF values from $-\omega_1$ to $-\omega_2$

FIGURE 11

The real part is also referred to as an EVEN function, and the imaginary part as an ODD function. Note that the two frequency intervals of the FRF each contain (**L**) data points, the negative frequencies represented by the indices (**-L**,...,**-1**), and the positive frequencies represented by the indices (**1**,...,**L**). In order to preserve the symmetric properties of the functions about the origin, the origin itself will be excluded from our further development of the curve fitting algorithm. This is not a serious restriction though, since FRFs are rarely made with any accuracy at D.C. (or zero frequency).

**FIGURE 12**

Complex polynomial functions can be defined as summations of EVEN and ODD functions which exhibit the same properties as the FRF. As shown in Figure 12, the

real polynomials are even functions, while the imaginary polynomials are odd functions. Furthermore we will see later on, that not only can these polynomials be generated recursively, but they can also be generated so that they satisfy the orthogonality condition in expression (12).

POLYNOMIAL HALF FUNCTIONS

$$\phi_{i,k} = \phi_{i,k}^+ + \phi_{i,k}^-, i = -L, \dots, -1, 1, \dots, L$$

$$\phi_{i,k}^+ = \text{right half function, defined for } i > 0$$

$$\phi_{i,k}^- = \text{left half function, defined for } i < 0$$

Orthogonality

$$\sum_{i=-L}^{-1} (\phi_{i,k}^-)^* \phi_{i,j}^- = \sum_{i=1}^L (\phi_{i,k}^+)^* \phi_{i,j}^+ = \begin{cases} 0, & k \neq j \\ 1, & k = j \end{cases}$$

FIGURE 13

So, instead of representing the in terms of ordinary polynomials, as in equation (1), we can represent it in terms of linear combinations of orthogonal polynomials. This is not unlike the process of representing a time domain wave form with a summation of orthogonal sine and cosine functions as is done in Fourier Series.

Orthogonal Polynomials in Terms of Half Functions

One final problem that we face is that we would like to curve fit the measurement data for only positive values of frequency and yet take advantage of the orthogonality property of the orthogonal polynomials in the solution equations. Recall that the solution equations (11) only apply for positive frequencies, but the orthogonality conditions (12) must include the polynomial values for negative frequencies as well.

However, we can again take advantage of the symmetry properties of these polynomials to write orthogonality conditions in terms of positive frequency values only. The orthogonal polynomials can be represented as the sum of two so-called half-functions as shown in equation (13). One half-function is defined only for positive frequencies and the other is defined only for negative frequencies.

The symmetric property of the polynomials is stated in expression (14) in terms of the index (**k**) of the functions themselves. That is, when (**k**) is an odd number, the **k**th function is an odd function. Likewise when (**k**) is an even number, the **k**th function is an even function.

By applying this property to the original orthogonality definition (12), a new definition of orthogonality involving

only the half-functions is possible. Equation (15) states that the positive half functions, summed only over the (L) positive frequency data points, also satisfy the orthogonality condition. Similarly, the negative half functions summed over the (L) negative frequencies satisfy the same condition.

Using this result, we are now ready to re-formulate the FRF in terms of orthogonal polynomials.

FRF in Terms of Orthogonal Polynomials

Expression (16) in Figure 14 shows the FRF in rational fraction form, but this time it is written in terms of two sets of orthogonal polynomials. The unknown coefficients in this expression are ($c_k, k=0, \dots, m$) and ($d_k, k=0, \dots, n$) and it is these values which are sought as solutions during the curve fitting process. Once the values of the c 's and d 's are known, the coefficients of the ordinary polynomials, i.e. the a 's and b 's in equation (1), can be easily recovered.

The denominator polynomials differ from the numerator polynomials in that they must satisfy a different orthogonality condition which contains a weighting function. This function is the magnitude squared of the FRF measurement data, and it will become apparent from the solution equation formulation which follows, that this orthogonality condition greatly reduces the complexity of the equations.

FREQUENCY RESPONSE FUNCTION IN TERMS OF ORTHOGONAL POLYNOMIALS

$$H(\omega_i) = \frac{\sum_{k=0}^m c_k \phi_{i,k}^+}{\sum_{k=0}^n d_k \theta_{i,k}^+} \quad i = 1, \dots, L \quad (16)$$

Orthogonality

$$\sum_{i=1}^L (\phi_{i,k}^+)^* \phi_{i,j}^+ = \begin{cases} 0, & k \neq j \\ .5, & k = j \end{cases} \quad (17)$$

$$\sum_{i=1}^L (\theta_{i,k}^+)^* |h_i|^2 \theta_{i,j}^+ = \begin{cases} 0, & k \neq j \\ .5, & k = j \end{cases} \quad (18)$$

FIGURE 14

SOLUTION EQUATIONS IN TERMS OF ORTHOGONAL POLYNOMIALS

Using the same procedure as the problem formulation for ordinary Polynomials, a new set of solution equations can be derived in terms of orthogonal polynomials. First, the error vector can be re-written in terms of orthogonal polynomials, as shown in Figure 15. (Again, it is assumed that the highest order denominator coefficient is unity; ($d_n=1$)). Then, by following the same steps as before; namely, formulation of an error criterion, and taking derivatives of the criterion with respect to the unknown polynomial coefficients, a new set of solution equations results. These equations are expressed as equations (20).

THE ERROR VECTOR IN TERMS OF ORTHOGONAL POLYNOMIALS

$$\{E\} = [P]\{C\} - [T]\{D\} - \{W\} \quad (19)$$

where:

$$[P] = \begin{bmatrix} \phi_{1,0}^+ & \phi_{1,1}^+ & \dots & \phi_{1,m}^+ \\ \phi_{2,0}^+ & \phi_{2,1}^+ & \dots & \phi_{2,m}^+ \\ \vdots & \vdots & & \vdots \\ \phi_{L,0}^+ & \phi_{L,1}^+ & \dots & \phi_{L,m}^+ \end{bmatrix} \quad (L \times m+1)$$

$$[T] = \begin{bmatrix} h_1 \theta_{1,0}^+ & h_1 \theta_{1,1}^+ & \dots & h_1 \theta_{1,n-1}^+ \\ h_2 \theta_{2,0}^+ & h_2 \theta_{2,1}^+ & \dots & h_2 \theta_{2,n-1}^+ \\ \vdots & \vdots & & \vdots \\ h_L \theta_{L,0}^+ & h_L \theta_{L,1}^+ & \dots & h_L \theta_{L,n-1}^+ \end{bmatrix} \quad (L \times n)$$

$$\{C\} = \begin{Bmatrix} c_0 \\ c_1 \\ \vdots \\ c_m \end{Bmatrix}, \quad \{D\} = \begin{Bmatrix} d_0 \\ d_1 \\ \vdots \\ d_{n-1} \end{Bmatrix},$$

$$\{W\} = \begin{Bmatrix} h_1 \theta_{1,n}^+ \\ h_2 \theta_{2,n}^+ \\ \vdots \\ h_L \theta_{L,n}^+ \end{Bmatrix}$$

FIGURE 15

Comparing equations (20) with equations (11), it is clear that using orthogonal polynomials caused two significant changes in the solution equations. First, the two matrices $[Y]$ and $[Z]$ in equation (11) are now replaced with identity matrices (diagonal ones and zeros elsewhere) in equations (20).

With these simplifications, equations (20) can be written as two sets of uncoupled equations with respect to the unknown vectors $\{C\}$ and $\{D\}$. Expression (21) is a set of (n) equations which is solved first for the denominator coefficients $\{D\}$. Then the numerator coefficients $\{C\}$ are obtained by solving equations (22).

SOLUTION EQUATIONS IN TERMS OF ORTHOGONAL POLYNOMIALS

$$\begin{bmatrix} I_1 & \vdots & X \\ \dots & \vdots & \dots \\ X^t & \vdots & I_2 \end{bmatrix} \begin{bmatrix} C \\ \dots \\ D \end{bmatrix} = \begin{bmatrix} H \\ \dots \\ 0 \end{bmatrix} \quad (20)$$

(m+n+1) equations

where: $[X] = -\text{Re}([P^*]^t [T])$ (m+1 x n)

$\{H\} = \text{Re}([P^*]^t \{W\})$ (m+1 vector)

$[I_1] = \text{Identity matrix}$ (m+1 x m+1)

$[I_2] = \text{Identity matrix}$ (n x n)

$\{0\} = \text{Zero vector}$ (n vector)

$[I - [X]^t [X]] \{D\} = -[X]^t \{H\}$ (21)

$\{C\} = \{H\} - [X] \{D\}$ (22)

FIGURE 16

GENERATING ORTHOGONAL POLYNOMIALS

In order to set up the solution equations (21) and (22) for numerical solution, a method is needed for generating numerical values of the orthogonal half functions over the interval (L) positive frequency data points in the measurement. The so-called Forsythe method has all of the computational properties we are looking for and in fact, can be simplified to handle the special types of functions we wish to compute.

Reference [6] explains the Forsythe method in detail, and contains examples and further references on the algorithm. The algorithm details will not be presented here, but we will point out two simplifications which we made in applying it to this problem.

The Forsythe equations for generating complex orthogonal polynomials are given in Figure 17, using notation similar to that in reference [6]. The weighting function for each data point is denoted by $(q_i, i=-L, \dots, L)$. In our application, the numerator polynomials are generated with a unity weighting function $(q_i=1, i=-L, \dots, L)$, while the denominator polynomials are generated with a weighting function equal to the magnitude squared of the FRF data.

COMPLEX ORTHOGONAL POLYNOMIALS BY THE FORSYTHE METHOD

$$\begin{aligned} P_{i,-1} &= 0 \\ P_{i,0} &= 1 \\ P_{i,1} &= (X_i - U_1)P_{i,0} \\ P_{i,2} &= (X_i - U_2)P_{i,1} - V_1P_{i,0} \\ &\vdots \\ P_{i,k} &= (X_i - U_k)P_{i,k-1} - V_{k-1}P_{i,k-2} \end{aligned} \quad (23)$$

$k=2,3,\dots \quad i=-L,\dots,L$

where

$$U_k = \sum_{i=-L}^L X_i |P_{i,k-1}|^2 q_i / D_{k-1} \quad (24)$$

$$V_k = \sum_{i=-L}^L X_i P_{i,k-1} P_{i,k-2}^* q_i / D_{k-1} \quad (25)$$

$$D_k = \sum_{i=-L}^L |P_{i,k}|^2 q_i \quad (26)$$

$X_i = j\omega_i$ i^{th} frequency

q_i = weighting function at the i^{th} frequency

$P_{i,k}$ = k^{th} order polynomial at the i^{th} frequency

FIGURE 17

Complex Polynomials in Terms of Real Polynomials

One computational savings which we discovered in using the Forsythe method is that the required complex polynomials could be represented in terms of real valued polynomials using, the simple rule given in equation (27), where $R_{i,k}$ represents the k^{th} real polynomial value evaluated at the i^{th} frequency. This result can be derived by substituting $(X_i=j\omega_i)$ into equations (23) through (26), and

representing the real polynomials simply as functions of the frequency (ω_i), instead of \mathbf{X}_i .

The net result, then, is that the complex polynomials $\mathbf{P}_{i,k}$ can be generated by first generating the real polynomials $\mathbf{R}_{i,k}$ and then applying equation (27).

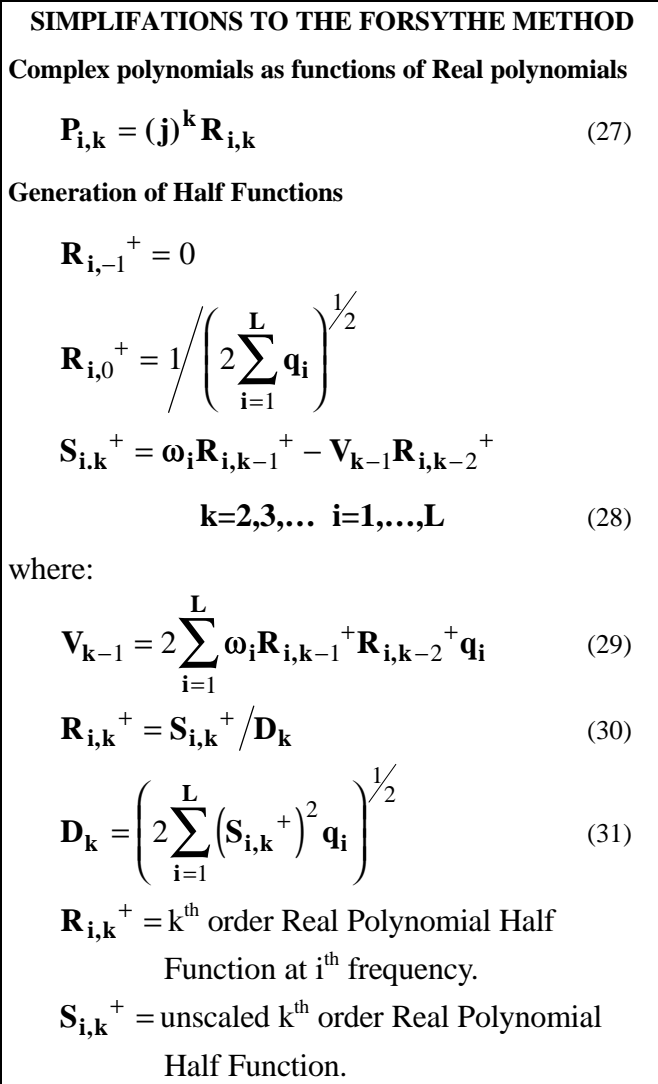


FIGURE 18

Generation of Half Functions

If the complex polynomial half function symmetry (14) and orthogonality (15) properties, together with property (27) of the complex polynomials, are applied to the Forsythe formulas, the polynomial generation equations are reduced to the forms shown in equations (28) through (31). Comparing these equations with equations (23) through (26), two computational simplifications are readily apparent. First, the variable (\mathbf{U}_k) has been removed from the calculations altogether, and secondly, all summations

are performed only for positive frequencies instead of both positive and negative frequencies.

Equations (28) through (31), together with equation (27), are used to generate all the terms of the $\{\mathbf{X}\}$ matrix and the $\{\mathbf{H}\}$ vector in solution equations (21) and (22).

Scaling the Frequency Axis

It is clear from equation (28) that orthogonal polynomial generation will also encounter numerical problems if the frequency is not scaled in some manner. The simplest scaling procedure is to scale all frequencies to the UNIT interval $t \in [-1, 1]$ by dividing all frequency values by the highest value of frequency used, (e.g., divide by ω_L Figure 11).

Conversion from Orthogonal to Normal Polynomial Coefficients

To curve fit a measurement, the solution equations (21) and (22) are solved for the unknown polynomial coefficients $\{\mathbf{C}\}$ and $\{\mathbf{D}\}$. Once these least squared error estimates are known, the curve fitting of the FRF measurement data is essentially complete. However, the coefficients of the ordinary polynomials $\{\mathbf{A}\}$ and $\{\mathbf{B}\}$ are more useful than the orthogonal polynomial coefficients, and can be computed from the $\{\mathbf{C}\}$ and $\{\mathbf{D}\}$ solutions by using some formulas given in reference [6]. These formulas need no further discussion here since they can be used directly as they are given in [6].

IDENTIFYING THE CHARACTERISTIC POLYNOMIAL FROM MULTIPLE MEASUREMENTS

Many times, it is necessary and desirable to make multiple FRF measurements on a system or structure in order to adequately identify its dynamics. When exciting modes of vibration in a structure, or in making measurements in a servo-loop, the denominators of all the measurements should contain the same characteristic polynomial. In the case of structural resonances, this is equivalent to saying that modal frequencies and damping are the same, no matter where they are measured on the structure. Alternatively, the poles of a servo-loop can be identified from measurements between any two points in the loop.

As pointed out earlier, one of the advantages of formulating the solution equations in terms of orthogonal polynomials is that the unknown characteristic polynomial coefficients $\{\mathbf{D}\}$ can be determined independently of the numerator coefficients $\{\mathbf{D}\}$. (Refer to equations (21) and (22).)

Therefore, we can repeatedly apply equation (21) to several different measurements and, assuming that each

measurement contains the same characteristic polynomial, determine multiple estimates of the polynomial coefficients. Expression (32) contains the repeated application of solution equations (21) to (**p**) different measurements. If the characteristic polynomial has (**n**) unknown coefficients, then expression (32) contains (**n x p**) equations. This is an over specified set of equations since only (**n**) equations are needed to solve for the unknown coefficients **{D}**. In this case, we can find a least squared error solution to equation (32) by solving equation (33). Expression (33) is again a set of (**n**) equations which provides a best estimate of the characteristic polynomial coefficients **{D}** by using measurement data from (**p**) different measurements.

**CHARACTERISTIC POLYNOMIAL
FROM MULTIPLE MEASUREMENTS**

$$\begin{bmatrix} U_1 \\ \vdots \\ U_2 \\ \vdots \\ U_p \end{bmatrix} \{D\} = \begin{bmatrix} V_1 \\ \vdots \\ V_2 \\ \vdots \\ V_p \end{bmatrix} \quad (32)$$

where

$$[U_k] = [I - [X_k]^t [X_k]] \quad (n \times n)$$

$$\{V_k\} = [X_k]^t \{H_k\} \quad (n\text{-vector})$$

Least Squared Error Solution

$$\sum_{k=1}^p [U_k]^2 \{B\} = \sum_{k=1}^p [U_k] \{V_k\} \quad (33)$$

FIGURE 19

Again, once the orthogonal polynomial coefficients **{D}** are known, the ordinary polynomial coefficients **{B}** can be computed, as discussed earlier.

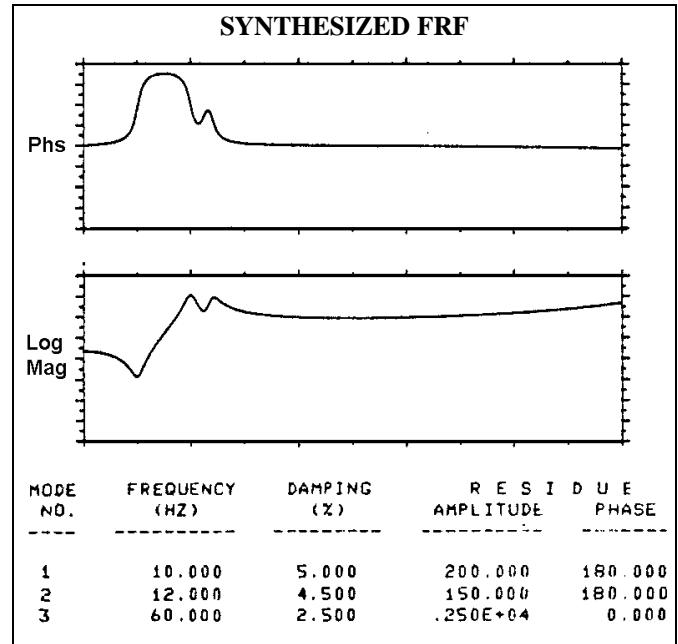
ILLUSTRATIVE EXAMPLES

Several examples of the use of the polynomial curve fitting algorithm are given here. These examples also point out some common problems which occur with all curve fitters.

Compensation for Out-of-Band Resonances

FRF measurements are always made over a limited frequency range by exciting the structure or system with some broad band signal. As a consequence, the

measurements will typically contain the residual effects of resonances which lie outside of the measurement frequency range. In addition, we normally curve fit the measurement data only in a more limited frequency range surrounding the resonance peaks. Hence, to give accurate results, all curve fitters must somehow compensate for the residual effects of resonances which lie outside of the curve fitting frequency range.

**FIGURE 20**

With this curve fitter, out-of-band effects can be approximated by specifying additional terms for either the numerator or the denominator polynomial. Figure 20 shows an ideal measurement which was synthesized using the parameters listed below it.

SINGLE DOF (or MODE) FITTING

First, we will use the polynomial fitter to identify each of the resonances, one

at a time. To model a single mode (or degree-of-freedom), the characteristic polynomial need only be a quadratic polynomial of degree **n=2**.

Normally, the numerator polynomial for a single degree-of-freedom system would have degree **m=1**. However, if we attempt to fit the data in the vicinity of each of the resonance peaks in Figure 20, the effects of the other resonances will cause large errors in the parameter estimates. So, to compensate for these out-of-band resonances, we will add four additional terms to the numerator polynomial, making **m=5**.

Figure 21 shows same results of using this single made fitter on various bands of data in the synthesized measurement. Notice that a proper choice of data around each resonance peak, (indicated by the pair of dashed cursor lines), gives reasonably good curve fit functions. (The fit function is also plotted between the cursors.) However, some of the estimates of the modal parameters are in error by substantial amounts.

MULTIPLE DOF FITTING

We can simultaneously identify the parameters of two (2) or all three (3) of the resonances by increasing the degree of the polynomials in the curve fitter.

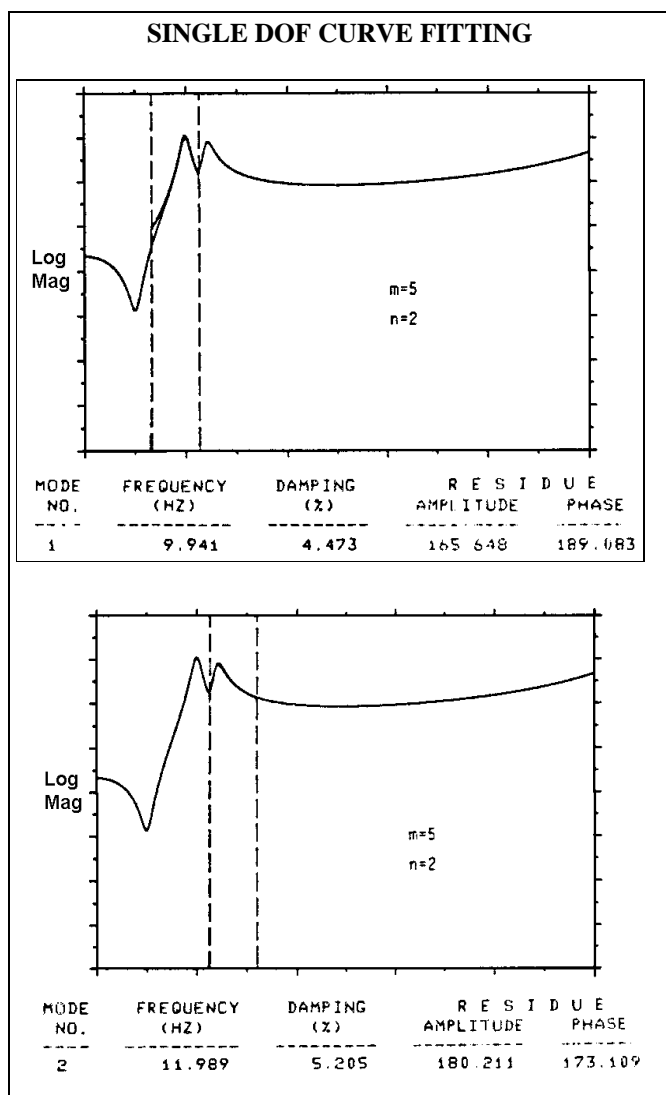


FIGURE 21

In a typical curve fitting situation we might want to identify the parameters of the 10 and 12 Hz modes by using a 2-DOF model. This means setting $n=4$, $m=3$, and then

curve fitting a band of data which includes the two resonance peaks. The results are shown in Figure 22.

It is clear that the parameters of the first two modes cannot be identified without some form of compensation for the third mode. Again, we can attempt to compensate for the effects of the third out-of-band mode by adding more terms to the numerator polynomial. Figure 22 also shows the results of this curve fit with $n=4$, $m=7$. In this case, the additional numerator terms do an excellent job of compensating for the third mode.

Finally, we can also compensate for the third out-of-band mode made by adding another DOF to the denominator. Figure 22 shows the results of the curve fit with $n=6$, $m=5$. Notice that this fit function is, as expected, a perfect match to the idealized FRF measurement, and that even the 60 Hz mode which lies far outside the curve fitting band is also correctly identified.

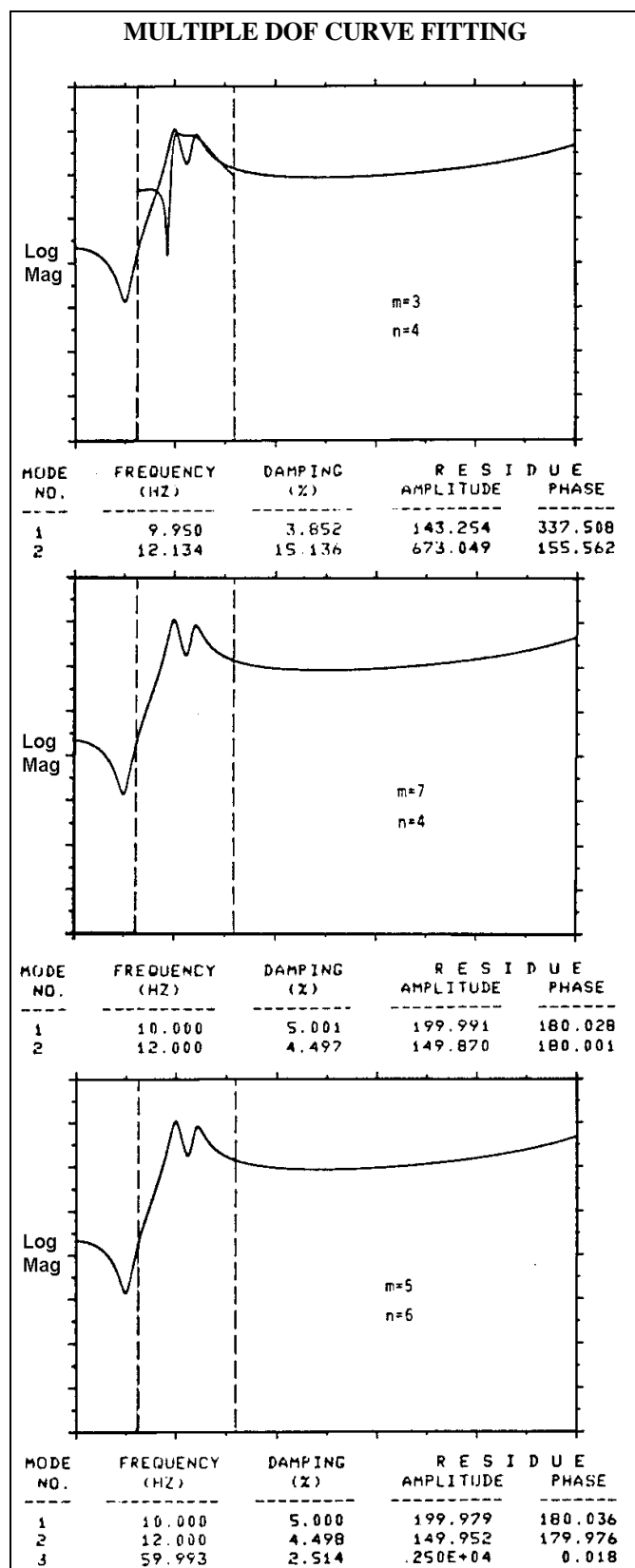
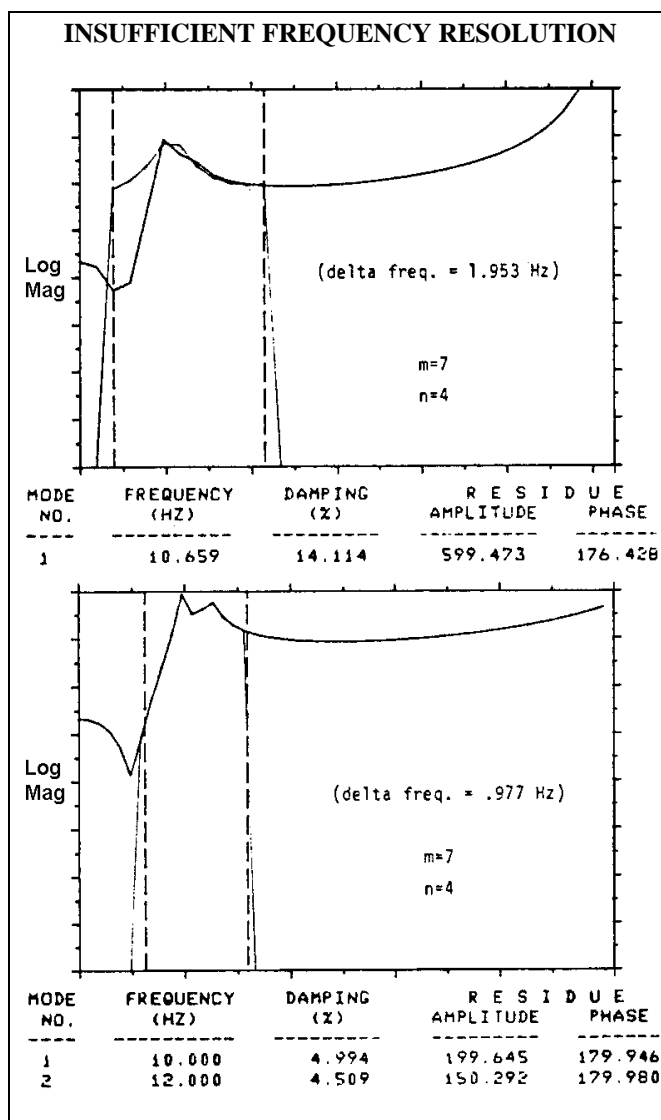
FREQUENCY RESOLUTION

All digital spectrum analyzers have finite length data memories, which means that their measurements are represented by a finite number of data points over some chosen range of frequencies. The frequency resolution of a particular measurement then is, is equal to the frequency range of the measurement divided by the number of data points, or frequency values, of the measurement. For instance, a typical measurement might be made with 256 lines, or frequency values, over a 100 Hz range giving a resolution of .39 Hz.

The polynomial curve fitter can encounter two different problems due insufficient frequency resolution,

1. There are not enough data points in the vicinity of the resonances to identify the modal parameters, or
2. The orthogonal polynomials are under sampled over the chosen frequency range, thus causing the orthogonality condition to break down.

Figure 23 shows two examples of the same synthesized FRF, as in Figure 20, but the first function has only $1/20^{\text{th}}$ as much resolution, and the second function has $1/10^{\text{th}}$ as much resolution as the function in Figure 20. The first problem caused by insufficient resolution is determining how many modes are present in the data. In the first example, one would probably say there is one mode and attempt to identify parameters for only one mode. The answer is clearly in error. In the second example, there is some slight evidence of the existence of two modes, and since the data is an ideal FRF, the curve fitter has no trouble identifying the modal parameters correctly.

**FIGURE 22****FIGURE 23****MEASUREMENTS WITH NOISE**

Since the curve fitting equations were formulated in such a manner as to yield least squared error estimates of the polynomial coefficients, the curve fitter should be able to identify modal parameters from measurements containing additive white noise.

Figure 24 shows examples of the same measurement as in Figure 20, but with various amounts of random noise added to it. When a measurement contains noise (and all real world measurements contain some noise), then it becomes more clear that curve fitting can be characterized as a data reduction process. That is, we wish to use as much measurement data as we can, especially in the vicinity of the resonance peaks, in order to identify a few system parameters. As a general rule, the more data that is used in

curve fitting, the more accurate the parameter estimates should be. The examples in Figure 24 confirm that acceptable modal parameter estimates can still be obtained from noisy measurements,

CONCLUSIONS

We have presented here a new formulation of a parameter estimation technique using rational fraction polynomials, which we have found to be practical implementation in small computers.

This technique has been implemented in several commercially available modal analysis systems, which are in wide use today. It has also been implemented in a parameter estimation package which runs on desktop computers. We have used it on a wide variety of measurements which contained various degrees of noise and modal density, and which were taken from lightly damped as well as critically damped systems.

We have found this Rational Fraction Polynomial (RFP) method to be comparable to the Complex Exponential method in terms of execution speed and accuracy. As mentioned earlier, the strongest drawback of the Complex Exponential method is that the impulse response can be severely distorted by wrap-around error, whenever FRF measurements with significant truncation, or non periodicity, are used. This distortion, plus the effects of resonances outside the curve fitting frequency band, can only be compensated for by greatly over specifying the number of modes (i.e. degree of the characteristic polynomial) during curve fitting. However, this often yields a set of unusable parameters since it is not clear which parameters should be "thrown away" and which ones should be kept.

All curve fitting algorithms can give erroneous results, though, if the degree of the curve fitting function is over specified. On the other hand, we usually use curve fitters over a limited frequency range, so the effects of resonances outside the band must be compensated for in some manner. With the RFP method, specifying additional terms for only the numerator polynomial is often sufficient for compensating for these residual effects. This allows some control over the degree of the characteristic polynomial.

In modal analysis, the number of modes in a given frequency band is usually established beforehand, and then parameters for those same modes are identified from a set of multiple measurements. The number of modes is normally determined by counting resonance peaks in the measurements. Since modes have different strengths, (or amplitudes) from one measurement to another, it is often difficult to tell from a single measurement how many modes are contained in an entire set of measurements.

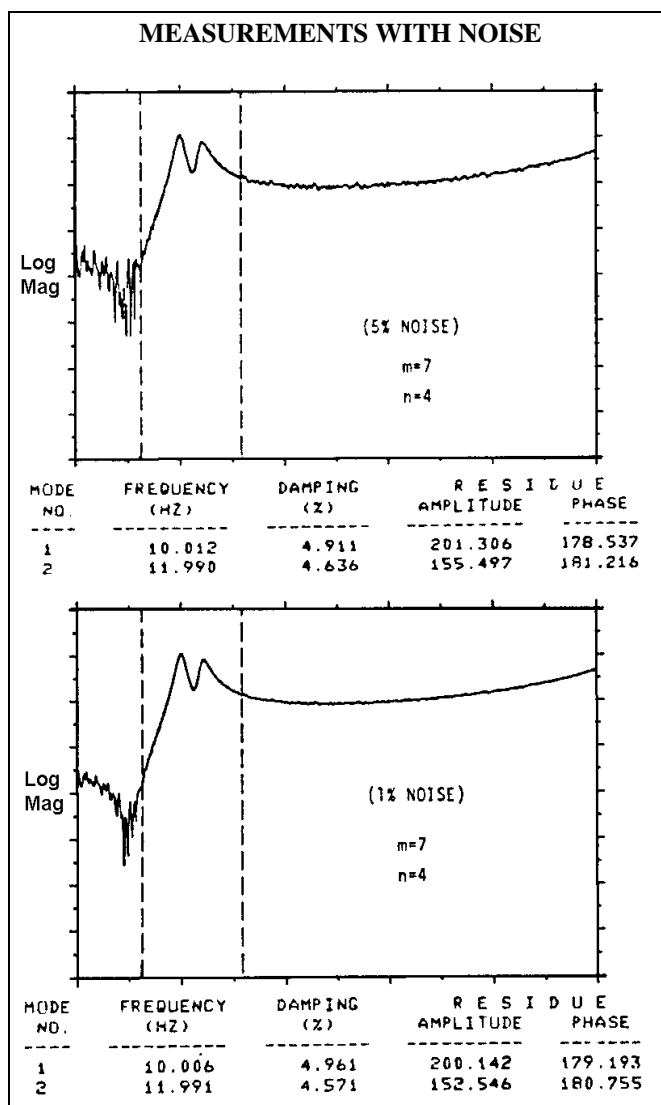


FIGURE 24

One proposed approach is to use the multiple measurement algorithm presented in this paper to determine the total number of modes, together with "global" estimates of modal frequency and damping for a structure. This procedure would require some measure of "goodness of fit" which could then be optimized as a function of the total number of modes.

In summary, we can conclude that the salient features of the parameter estimation method presented here are that it is sufficiently fast and accurate for use in desktop and mini-computers. In addition, it is non-iterative, and works directly with FRF data in the frequency domain. On the other hand, it suffers the same consequences as any other method when the measurements contain excessive noise or distortion, or they lack sufficient frequency resolution.

REFERENCES

1. Richardson, M. "Modal Analysis Using Digital Test Systems", Proceedings of Seminar on Understanding Digital Control and Analysis in Vibration Test Systems, May 1975. The Shock and Vibration Information Center, Naval Research Lab., Wash. D.C.
2. Texas Instruments Incorporated, "Representation and Analysis of Sonar Signal, Vol. I: Improvements in the Complex Exponential Signal Analysis Computation Algorithm", Office of Naval Research Contract No. N00014-69-C-90315, March 1970.
3. Spitznogle, F.P., Quazi, A.H. "Representation and Analysis's of: Time-Limited Signals Using a Complex Exponential Algorithm" J. of Acoust. Soc. of Amer., Vol 47, No 5 (Prt I), May 1970, pp. 1150-1155.
4. Formenti, D.L. "A Study of Modal Parameter Estimation Errors" to be presented at the First IMAC Conference, Orlando, Florida, Nov., 1982.
5. Miramand, N., Billand, J.F., LeLeux, F. and Kernevez, J.P. "Identification of Structural Modal Parameters by Dynamic Tests at a Single Point", France.
6. Kelly, L.G. "Handbook of Numerical Methods and Applications, Chapter 5 Curve Fitting and Data Smoothing" Addison - Wesley Pub.Co. Inc., 1967.