Aerodynamic Parameter Estimation Via Fourier Modulating Function Techniques

A. E. Pearson
Brown University • Providence, Rhode Island

April 1995

Printed copies are available from the following:

AERODYNAMIC PARAMETER ESTIMATION VIA
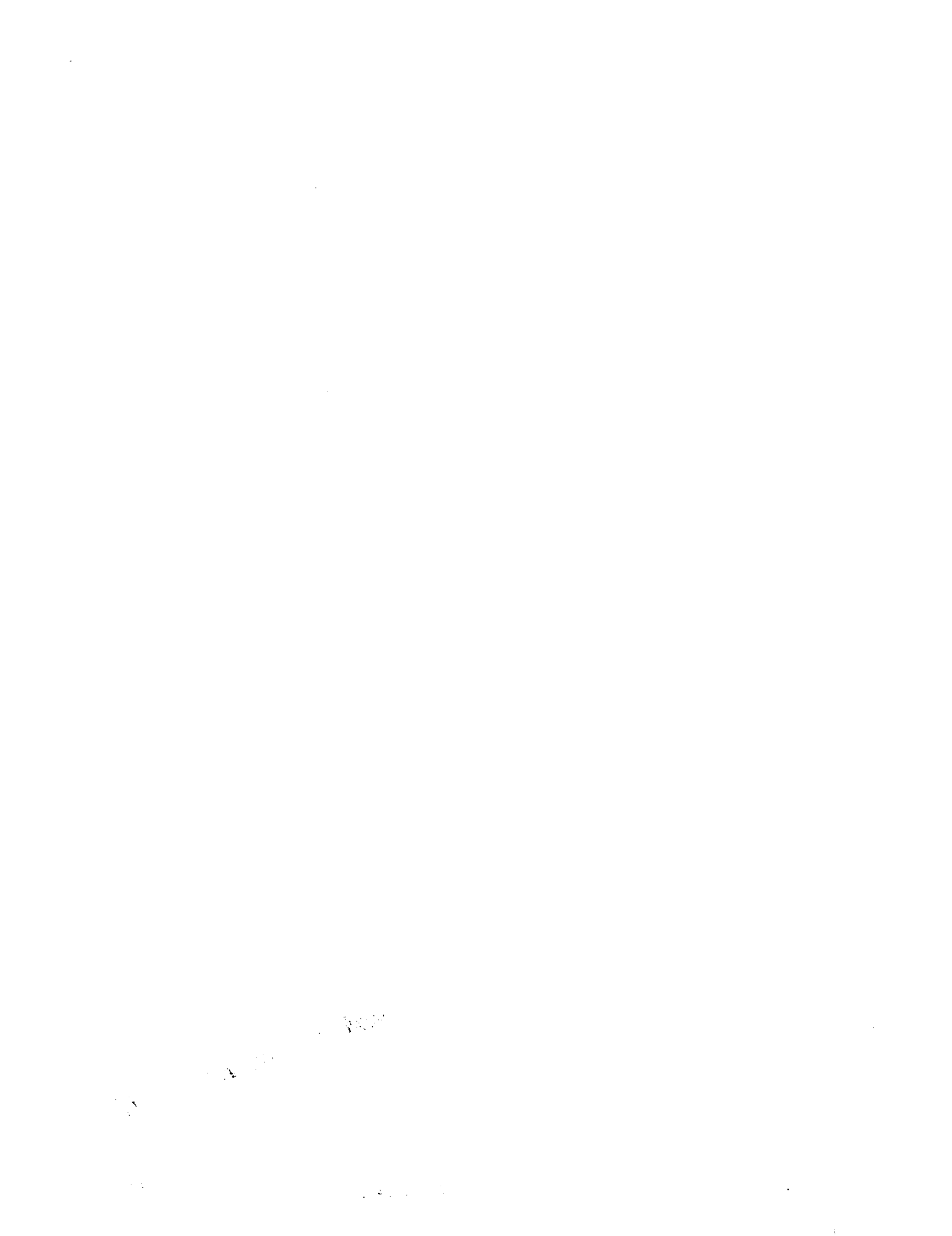# FOURIER MODULATING FUNCTION TECHNIQUES

A. E. Pearson[1]
Division of Engineering
Brown University
Providence, RI 02912

ABSTRACT

Parameter estimation algorithms are developed in the frequency domain for systems modeled by input/output ordinary differential equations. The approach is based on Shinbrot's method of moment functionals utilizing Fourier based modulating functions. Assuming white measurement noises for linear multivariable system models, an adaptive weighted least squares algorithm is developed which approximates a maximum likelihood estimate and cannot be biased by unknown initial or boundary conditions in the data owing to a special property attending Shinbrot-type modulating functions. Application is made to perturbation equation modeling of the longitudinal and lateral dynamics of a high performance aircraft using flight-test data. Comparative studies are included which demonstrate potential advantages of the algorithm relative to some well established techniques for parameter identification. Deterministic least squares extensions of the approach are made to the frequency transfer function identification problem for linear systems and to the parameter identification problem for a class of nonlinear time-varying differential system models.

---

[1] Professor of Engineering

Table of Contents

List of Figures

List of Tables

Glossary

Abbreviatons

AWLS	adaptive weighted least squares
col ( )	column vector of arguments
DFT	discrete Fourier transform
FFT	fast Fourier transform
LS	least squares
MATLAB	software package from The Mathworks
MFT	modulating function technique
MIMO	multi input multi output
NSR	noise-to-signal ratio (based upon RMS values)
PEM	prediction error method
RMS	root-mean-square (integral form relative to a time interval [0, T])
row ( )	row vector of arguments
S/E	signal-to-error ratio (in decibels based upon RMS values)
SISO	single input single output
WLS	weighted least squares

Notations

AWLS/MFT	AWLS via the MFT approach
$F(m)$	m^{th} harmonic Fourier series coefficient of a signal $f(t)$ relative to [0, T]
ω_0 or F_0	resolving frequency in rad/sec or in Hertz
ω_B or F_B	system bandwidth in rad/sec or in Hertz
Δ^n	n^{th} order finite difference operator
p	differential operator d/dt
[0, T]	time interval $0 \leq t \leq T$
$\otimes$	convolution sum

1. Introduction

Estimating aircraft parameters from flight or wind-tunnel data has been an important modeling activity for aerodynamicists over several decades. Various kinds of parametrized models exist to encompass the steady and unsteady flight conditions (Klein, 1989). Here the focus is placed on estimating the parameters of ordinary differential equation input/output models for which a number of methods exist. These include: *approximations* like the δ operator, block pulse and orthogonal function expansion methods; *moment functional techniques; filtering techniques* like the maximum likelihood/extended Kalman filter, and the *conversion* of an identified discrete-time model (obtained by a variety of methods) into a continuous-time model using, for example, the bilinear transformation. The books by Unbehauen and Rao (1987) and Johansson (1993) are good introductions to the identification of continuous-time models, and Sinha and Rao (1991) contain chapters by researchers reporting recent results on some of these methods.

The method of moment functionals, also known as the modulating function technique, is characterized by the use of integration-by-parts on an equation error model of the system to transfer the continuous-time derivatives on the input/output variables to derivatives on a set of smooth user-chosen 'modulating' functions. The result of this process is a set of algebraic equation errors which is characterized by functionals on the data that have to be computed before initiating the parameter estimation. If the modulating functions are of the Shinbrot type (Shinbrot, 1954,1957), then the chosen functions must satisfy certain end point conditions for the time interval over which data is given and, as a consequence, the resulting equation error vector will be devoid of unknown initial or boundary conditions on the data. Herein lies one advantage of Shinbrot's method in that it effectively decouples the state and parameter estimation problems for data collected under transient conditions. Utilizing Fourier based modulating functions, another advantage is that the system identification problem can be equivalently posed entirely in the frequency domain, and the functionals can be calculated by efficient DFT/FFT techniques. It is Shinbrot's method that is promulgated in this report.

Following a discussion in Section 2 of the aircraft models used to motivate the formulation and the main assumptions concerning the data, Shinbrot's method is developed in Section 3 for single input single output (SISO) linear differential systems based upon a chosen set of Fourier modulating functions. Using a second order example, a comparison is made via simulation with a commercially available prediction error algorithm which illustrates the improved accuracy accrued for the modulating function technique over a range of signal to noise ratios. The formulation of Section 3 for SISO models is extended in Section 4 to the multivariable (MIMO) aircraft models introduced in Section 2, and results are presented for a high performance aircraft based upon actual flight data. Using the same data sets, a comparison is made with the modeling results obtained via a seasoned time domain based maximum likelihood identifier. This comparison indicates good corraboration when both yield useful models, but it also supports the efficacy of the modulating function technique for data sets in which the time domain based identifier failed. The final section discusses various extensions. Included here is showing how Shinbrot's method can be developed for a frequency transfer function identification problem, and an extension of the theory to a class of nonlinear input/output models.

2. Aircraft Perturbation Models

A model for longitudinal motion can be represented by the pair of linear differential operator equations:

$$A_1(p)\alpha(t) = B_1(p)u(t) \tag{1}$$

$$A_1(p)q(t) = B_2(p)u(t) \tag{2}$$

where $(A_1(p), B_1(p), B_2(p))$ are polynomials in the differential operator $p = d/dt$ whose coefficients represent parameters to be estimated given sampled versions of the input signal $u(t)$ and the two output signals $(\alpha(t), q(t))$ over some time interval $[0,T]$. Here $(\alpha(t), q(t))$ denote respectively the angle of attack and the pitch rate. In one application $u(t)$ will be the horizontal tail deflection, denoted by $\delta_h(t)$, while in another $u(t)$ will be the longitudinal stick deflection denoted by $\eta_h(t)$. The aircraft is operating in a closed loop as indicated in Fig. 1.
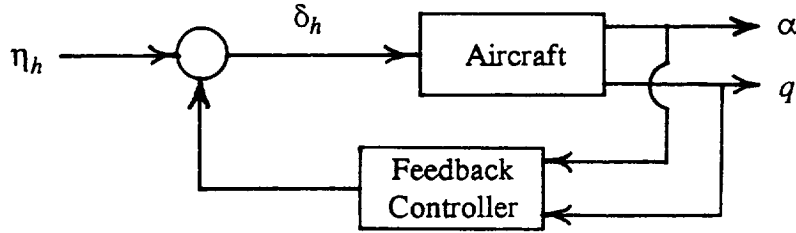
Fig. 1. Longitudinal dynamics block diagram

A model for lateral motion, presumed to be independent of the longitudinal motion, can be represented by the three differential operator equations:

$$A_2(p)\beta(t) = B_{11}(p)\delta_a(t) + B_{12}(p)\delta_r(t) \tag{3}$$

$$A_2(p)p(t) = B_{21}(p)\delta_a(t) + B_{22}(p)\delta_r(t) \tag{4}$$

$$A_2(p)r(t) = B_{31}(p)\delta_a(t) + B_{32}(p)\delta_r(t). \tag{5}$$

Here the problem is to estimate parameters comprised of the coefficients in the polynomials $(A_2(p), B_{ij}(p))$, $i=1,2,3$ and $j=1,2$, given sampled versions of the two input signals $(\delta_a(t), \delta_r(t))$, which are respectively the aileron and rudder deflections, and the three output signals $(\beta(t), p(t), r(t))$, which are respectively the sideslip angle, rolling velocity and yawing velocity, over a time interval $[0,T]$.

Underlying each set of models (1)-(2) and (3)-(5) is a state vector differential equation of the generic form:

$$\dot{x}(t) = Fx(t) + Gu(t), \quad y(t) = Hx(t). \tag{6}$$

In terms of transfer functions, the following model relation holds:

$$\frac{1}{A(s)}B(s) = H(sI - F)^{-1}G \tag{7}$$

-2-

where $A(s)=\det(sI-F)$ is designated by $A_1(s)$ for the longitudinal model (1)-(2), $A_2(s)$ for the lateral dynamics model (3)-(5), and the $B(s)$ matrix is comprised of the polynomials $B_i(s)$ or $B_{ij}(s)$ as appropriate to each set. The reason for working with the above particular model forms, i.e., the various outputs within each set are operated on by the same scalar-valued $A(p)$ operator, is a result of the ease with which the modulating function technique developed in Pearson *et al.* (1993b,c) for single input single output (SISO) systems can be extended to this form. In each set of models, the polynomial $A(p)$ is chosen monic with order fixed on the basis of physical considerations: Specifically for the perturbation equations relative to steady flight, the orders of $(A_1(s), A_2(s))$ are chosen to be (2, 4) respectively, and the polynomials comprising $B(s)$ are chosen of order one less than the corresponding $A(s)$ polynomial. This means that the model (1)-(2) possesses a total of 6 coefficient parameters, while the model (3)-(5) possesses 28 total coefficient parameters. In some applications, the lateral dynamics has included a fourth output variable - the Euler roll angle $\phi(t)$ - in which case the lateral dynamics possesses 36 total coefficient parameters. In addition, there are unknown weighting parameters that arise in the estimation procedure as will be indicated below.

2.1. Data Assumptions and Bandwidth Considerations

The major assumptions concerning the input/output data $[u(t), y(t)]$, $0 \le t \le T$, are:

(i) oversampling is employed in terms of the Nyquist sampling theorem, i.e., the sampled data $[u(kT/N), y(kT/N)]$, $k=0,1 \cdots N$, is collected in sufficient quantity and rate that the following inequality prevails:

$$2F_B << F_S = \frac{N}{T} \tag{7}$$

where F_B is the bandwidth of the system,

(ii) the various input signals are observed with negligible measurement noise and with sufficient frequency content to avoid degeneracy in the least squares estimates, and

(iii) noises corrupting the output signals are modeled as additive zero-mean independent white Gaussian noises with unknown variances. In addition, all stochastic calculus operations are presumed to be carried out in the mean square sense. This validates the various modulation processes on the differential equation models, provided all the pertinent signals are mean square differentiable of order equivalent to the model order.

The first two assumptions are quite reasonable for the aircraft modeling considered here, viz., the data is collected at a 50 Hz sampling rate, the system bandwidths are expected to be on the order of a few Hertz, and the input signals are relatively free of measurement noise while possessing sufficient frequency content for the estimation problem. The third assumption facilitates specifying a weighted-least-squares cost function in the frequency domain, via Fourier modulating functions of the Shinbrot type, whose minimization approximates a maximum likelihood estimate. This formulation for SISO systems is summarized in the following section. The basic calculation needed to set up the estimation problem is computing a finite number (specifically $M+n$ where the integers (M,n) are explained below) of Fourier series coefficients for each data variable. If $f(t)$, $0 \le t \le T$, is any such data variable and $f_k=f(kT/N)$, $k=0,1 \cdots N$, its sampled version, then the following staircase approximation to the Fourier integral is presumed to be carried out in obtaining the needed coefficients $F(m)$:

$$F(m) = \frac{1}{T}\int_0^T f(t)e^{-im\,\omega_0 t}\,dt \approx \frac{1}{N}\left[\sum_{k=0}^{N-1} f_k W_N^{mk}\right] \tag{8}$$

where $i=\sqrt{-1}$, $\omega_0=2\pi/T$, $m=0, 1, \cdots M+n$ and $W_N = e^{-i2\pi/N}$.

In the above, ω_0 is the 'resolving' frequency which is a user-chosen parameter if the data length T is selectable, n is the order of the polynomial $A(p)$ specified for the model; M is an integer that can be chosen on the basis of the number of unknown parameters or on the basis of bandwidth considerations if $T=2\pi/\omega_0$ is fixed. Thus, if n_θ denotes the number of unknown parameters ($n_\theta=6$ for the model (1)-(2) and $n_\theta=28$ for the model (3)-(5)), M can be chosen by the guideline:

$$M \approx 2n_\theta \sim 4n_\theta \tag{9a}$$

which will provide about $2n_\theta$ to $4n_\theta$ algebraic equations in the discrete-frequency domain upon which to base the least squares estimate. If the data interval $[0,T]$ is selectable, then the guideline:

$$(M+n)\omega_0 \approx \omega_B \quad \Rightarrow \quad T = 2\pi/\omega_0 \approx 2\pi(M+n)/\omega_B \tag{9b}$$

where $\omega_B=2\pi F_B$ is the system bandwidth. Hence, these two guidelines, which were upheld for the aircraft modeling carried out here, will uniquely determine the pair (M,ω_0). [2] Note that if the number of required harmonics, $M+n$, is artificially extended to equal the number of available samples, N, then the right hand side of (8) is the discrete Fourier transform of the data sequence f_k on the sampled interval, thus facilitating a DFT/FFT algorithm. Calculating the DFT's of the input/output data is a negligible computation in all the applications encountered thus far.

In terms of Hertz, the left hand side of (9b) means $(M+n)F_0 \approx F_B$, $F_0=1/T$, which together with the sampling assumption (7) implies the inequality:

$$\frac{M+n}{N/2} << 1. \tag{9c}$$

This inequality implies that only the *lowest* few percent of the available DFT harmonics will be utilized in the least squares algorithm. In particular, fewer than 10% were utilized in the aircraft modeling applications. Hence, high frequency noise rejection is inherent in the Fourier based modulating function technique if the guideline (9b) is upheld. It can be noted that this inequality is also consistent with a basic property of discrete Fourier transforms relating to the first and last halves of the DFT being complex conjugates of one another which, in turn, leads to the requirement that $(M+n) < N/2$ in order to preserve uniqueness for the Fourier coefficients. Furthermore, inequality (9c) is consistent with a fundamental tenet of identification in that the number N of available samples will ordinarily greatly exceed the

---

[2] It is advantageous in the case of high order systems to normalize the $[0,T]$ interval to $[0,2\pi]$ so that $\omega_0=1$, or some value close to unity. The reason is that derivatives in the time domain model are replaced by powers of ω_0 in the frequency domain model. Calculating powers like ω_0^n, will lead to numerically better conditioned regression equations in the frequency domain when $\omega_0\approx1$. This will be clear from the formulation in Section 3.

number n_θ of unknown parameters which from (9a) translates into the inequality $N > > M + n$.

3. Shinbrot's Method Via Fourier Modulating Functions

The basic idea of Shinbrot's method will be illustrated with respect to the problem of estimating the parameters for a second order single degree of freedom mechanical system. Given the force and displacement signals, represented respectively as input/output data $[u(t), y(t)]$ on a time interval $[0,T]$, find parameters (a_1, a_2, b_1) that best fit (in some sense) the model:

$$\ddot{y}(t) + a_1 \dot{y}(t) + a_2 y(t) = b_1 u(t), \quad 0 \le t \le T.$$

Multiplying the equation error for this model by a smooth function $\phi(t)$ and integrating over the given data interval:

$$\int_0^T \phi(t) \left[\ddot{y}(t) + a_1 \dot{y}(t) + a_2 y(t) - b_1 u(t) \right] dt = \varepsilon$$

where ε represents the accumulated equation error over $[0,T]$. Using integration-by-parts on the derivative portions of the above integral, the 'modulated' equation error ε is equivalent to

$$\varepsilon = \int_0^T \ddot{\phi}(t) y(t) dt - a_1 \int_0^T \dot{\phi}(t) y(t) dt + a_2 \int_0^T \phi(t) y(t) dt - b_1 \int_0^T \phi(t) u(t) dt + BC$$

where BC stands for 'boundary conditions'; specifically,

$$BC = \{ \phi(t) \dot{y}(t) - \dot{\phi}(t) y(t) + a_1 \phi(t) y(t) \} \Big|_0^T.$$

As introduced by Shinbrot within the context of this second-order case, $\phi(t)$ is a modulating function, herein to be called *Shinbrot-type* in order to distinguish from other types, if it satisfies the four end-point conditions:

$$\phi(0) = \phi(T) = \dot{\phi}(0) = \dot{\phi}(T) = 0.$$

Then $BC = 0$, i.e., presuming $[u(t), y(t), \dot{y}(t)]$ are bounded on $[0,T]$, and hence these conditions imply that the modulated equation error reduces to

$$\varepsilon = \int_0^T \ddot{\phi}(t) y(t) dt - a_1 \int_0^T \dot{\phi}(t) y(t) dt + a_2 \int_0^T \phi(t) y(t) dt - b_1 \int_0^T \phi(t) u(t) dt$$

$$= \gamma_0 \qquad - a_1 \gamma_1 \qquad - a_2 \gamma_2 \qquad - b_1 \gamma_3$$

which involves the linear functionals γ_i, $i = 0,1,2,3$, defined on the input/output data, but not derivatives of the data. Repeating this process with a sequence of linearly independent smooth functions $\phi_m(t)$, each satisfying the same four end-point conditions, will lead to a set of algebraic equation errors $\varepsilon(m)$, $m = 1, \cdots M$, characterized by linear functionals $\gamma_i(m)$ on the data, which can be used as a basis for estimating the coefficients (a_1, a_2, b_1). This estimate cannot be biased by unknown initial or boundary conditions in the data owing to the special end-point conditions of the Shinbrot modulating functions. Other moment functional techniques, such as the Poisson functionals (Saha *et al.* 1991), do not share this property.

Although Shinbrot's method was developed in an aeronautical setting (Shinbrot, 1954), it has remained relatively obscure in relation to the maximum likelihood and extended Kalman filtering techniques that have evolved for modern aerospace applications (Klein, 1993). There may be several reasons for this but in the main they all stem from a lack of clear guidance as to a good choice of Shinbrot-type modulating functions. Consider in this regard the following *order-n* Fourier modulating function of the Shinbrot-type which is defined and represented equivalently by [3]

$$\phi_{m,n}(t) = \frac{1}{T} e^{-im\omega_0 t}(e^{-i\omega_0 t}-1)^n, \qquad 0 \le t \le T = 2\pi/\omega_0 \tag{10a}$$

$$= \frac{1}{T}\sum_{k=0}^{n} c_k e^{-i(m+k)\omega_0 t}, \qquad c_k = (-1)^{n-k}\begin{bmatrix} n \\ k \end{bmatrix} \tag{10b}$$

$$= \frac{1}{T}\sum_{k=m}^{m+n} c_{k-m} e^{-ik\omega_0 t} \tag{10c}$$

where $i=\sqrt{-1}$, $\omega_0=2\pi/T$ is the resolving frequency, and m is any integer which shall be referred to as the 'modulating frequency index'. The family of functions defined by (10) for any specified index set $m\varepsilon\ \{m_1, \cdots m_M\}$ possesses several properties, four of which are as follows:

Property 1. For $k = 0, 1, \cdots n-1$: $\qquad p^k \phi_{m,n}(t) = 0$ at $t = 0$ and $t = T$ \qquad (11)

where p denotes the differential operator, i.e., $p=d/dt$, $p^2=d^2/dt^2$, etc.

Property 2. For a sufficiently smooth function $z(t)$ defined on $[0,T]$ and a differential operator $P(p)$ of order n, the integration of $\phi_{m,n}(t)P(p)z(t)$ over $[0,T]$ satisfies:

$$\int_0^T \phi_{m,n}(t)P(p)z(t)dt = \sum_{k=m}^{m+n} c_{k-m}P(ik\omega_0)Z(k) \tag{12a}$$

$$= \Delta^n P(im\omega_0)Z(m) \tag{12b}$$

where $Z(m)$ is the mth harmonic Fourier series coefficient of the function $z(t)$ on $[0,T]$, and Δ^n is the nth order finite difference operator, i.e., with $Q(m)=P(im\omega_0)Z(m)$,

$$\Delta^n Q(m) = \sum_{k=0}^{n}(-1)^k\begin{bmatrix} n \\ k \end{bmatrix}Q(n+m-k) = \sum_{k=m}^{m+n} c_{k-m}Q(k).$$

Property 3. For a zero mean continuous-time Gaussian white noise process $v(t)$ with covariance $Ev(t_1)v(t_2)=\sigma^2\delta(t_2-t_1)$ and a differential operator $P(p)$ of order n, the complex-valued stochastic sequence $\varepsilon(m)$, $m=0,1\cdots$, defined by

---

[3] Equation (10b) follows from (10a) by use of the binomial expansion where $\begin{bmatrix} n \\ k \end{bmatrix}$ denotes the binomial coefficient, and (10c) follows from (10b) by a change in the summation index.

$$\varepsilon(m) = \int_0^T \phi_{m,n}(t)P(p)v(t)dt = \varepsilon_R(m) + i\varepsilon_I(m)$$

is also Gaussian with the covariance relations among its real and imaginary parts given by:

$$E\varepsilon_R(m)\varepsilon_R(m+l) = \frac{\sigma^2}{2}\left[(P(0))^2\mu_0[m]\mu_0[l] + \mu_n[l]\sum_{k=m+l}^{m+n}c_{k-m}c_{k-m-l}\,|P(ik\omega_0)|^2\right]$$

$$E\varepsilon_I(m)\varepsilon_I(m+l) = \frac{\sigma^2}{2}\left[-(P(0))^2\mu_0[m]\mu_0[l] + \mu_n[l]\sum_{k=m+l}^{m+n}c_{k-m}c_{k-m-l}\,|P(ik\omega_0)|^2\right]$$

$$E\varepsilon_R(m)\varepsilon_I(m+l) = 0$$

where the notation $\mu_n[l]$ is defined by the string of unit pulses: $\mu_n[l] = \begin{cases} 1 & \text{for } 0\leq l\leq n \\ 0 & \text{otherwise} \end{cases}$

Thus, $\mu_0[m]$ and $\mu_0[l]$ are unit pulses at the origin which corresponds to the zero frequency.

Property 4. For sufficiently smooth functions $z(t)$ and $w(t)$ defined on $[0,T]$ and a differential operator $P(p)$ of order n, the integration of $\phi_{m,n}(t)w(t)P(p)z(t)$ over $[0,T]$ satisfies

$$\int_0^T \phi_{m,n}(t)w(t)P(p)z(t)dt = W(m)\otimes\Delta^n P(im\omega_0)Z(m) \tag{13}$$

where $\otimes$ stands for linear convolution in the discrete frequency domain of Fourier series coefficients, i.e., with $\tilde{Q}(m)=\Delta^n P(im\omega_0)Z(m)$,

$$W(m)\otimes\tilde{Q}(m) = \sum_{l=-\infty}^{\infty} W(l)\tilde{Q}(m-l).$$

Remarks on Verification. (i) Property 1 follows directly from the representation (10a) since $(e^{-i\omega_0 t}-1)^n$ is the nth power of a unit vector centered at -1 in the complex plane which starts at the origin and makes one complete revolution as t ranges over $[0,T]$. It is this property which qualifies $\phi_{m,n}(t)$ as a Shinbrot-type modulating function of order n.

(ii) The second property follows using the representation (10b) or (10c) while employing integration-by-parts on the left side of (12) and taking into account (11); during the process of transferring the derivatives on $z(t)$ to derivatives on the modulating function $\phi_{m,n}(t)$, all boundary point evaluations are zeroed such that

$$\int_0^T \phi_{m,n}(t)P(p)z(t)dt = \int_0^T z(t)P(-p)\phi_{m,n}(t)dt.$$

In view of the representations (10b,c), coupled with the fact that $P(-p)e^{-\lambda t}=P(\lambda)e^{-\lambda t}$, the right side of (12) is obtained which verifies Property 2. This property facilitates trading time derivatives on data-related functions with finite differences in the frequency domain unencumbered by unspecified initial/boundary conditions on the data.

(iii) The third property hinges on the presumption that all signals of interest in the

identification problem are mean square differentiable of order n, the order of the underlying model. This property is proven in Pearson and Shen (1993c) drawing upon mean square results from the stochastic calculus.

(iv) Property 4 follows using the Fourier series representation for $w(t)$ in the left hand side of (13), interchanging the summation and integration order, then applying Property 2 to each term in the sum. Property 4 is consistent with the Fourier analysis property involving the multiplication of two functions in the time domain and its corresponding relation to convolution in the frequency domain.

3.1. Parameter Estimation for SISO Models

Consider the single input single output system configuration in Fig. 2 where the transfer function $H(s)$ is represented by a ratio of polynomials $\dfrac{B(s)}{A(s)}$.
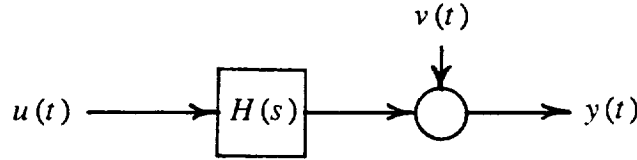

Fig. 2. System with disturbance $v(t)$

The system is modeled on $[0,T]$ by the differential operator equation:

$$p^n y(t) + \sum_{j=1}^{n} a_j p^{n-j} y(t) = \sum_{j=1}^{n} b_j p^{n-j} u(t) + e(t) \tag{14}$$

where $e(t)$ represents the equation error. For consistency between Fig. 2 and the model, $e(t)$ is defined via the relation:

$$e(t) = A(p)v(t) = \sum_{j=0}^{n} a_j p^{n-j} v(t), \quad a_0 = 1. \tag{15}$$

The problem is to estimate the (a_j, b_j) parameters, $j=1,2 \cdots n$, which are coefficients in the polynomials $(A(s), B(s))$, given the data $[u(t), y(t)]$ on $[0,T]$ and assuming the output $y(t)$ is corrupted by a zero mean white Gaussian process $v(t)$.

Modulating (14) with the Fourier modulating function $\phi_{m,n}(t)$ and integrating over $[0,T]$, the differential equation is seen to be equivalent to the following n^{th} order difference equation upon utilizing Eq. (12) of **Property 2** :

$$\Delta^n (im\omega_0)^n Y(m) + \sum_{j=1}^{n} a_j \Delta^n (im\omega_0)^{n-j} Y(m) = \sum_{j=1}^{n} b_j \Delta^n (im\omega_0)^{n-j} U(m) + \varepsilon_n(m). \tag{16}$$

In this process, the equation error $e(t)$ is transformed into the discrete frequency sequence:

$$\varepsilon_n(m) = \Delta^n A(im\omega_0)V(m) = \sum_{j=0}^{n} a_j \sum_{k=m}^{n+m} (ik\omega_0)^{n-j} c_{k-m} V(k)$$

where $V(k)$ is the kth harmonic Fourier series coefficient of the white noise process $v(t)$.

Interchanging the orders of summation, the preceding equation can be written as

$$\varepsilon_n(m) = \sum_{k=m}^{n+m} \alpha(k,m,\theta_a) V(k) \tag{17}$$

where $\theta_a = (a_1, \cdots a_n)$ and $\alpha(k,m,\theta_a)$ is a frequency-dependent function of the parameters defined by

$$\alpha(k,m,\theta_a) = c_{k-m} \sum_{j=0}^{n} a_j (ik\omega_0)^{n-j}, \quad a_0 = 1. \tag{18}$$

As can be inferred from **Property 3** , the stochastic sequence $\varepsilon_n(m)$ is Gaussian with a correlation function that possesses finite support. The implication of this is that the covariance matrix for the residuals is a banded structure, specifically it is banded by the order n of the differential operator model, which significantly simplifies the search for the appropriate weighting matrix in a weighted least squares estimation.

Defining a $2n$-column parameter vector θ as

$$\theta = (-a_1, \cdots -a_n, b_1, \cdots b_n)' = (-\theta_a{}', \theta_b{}')'$$

where prime denotes transpose, (16) can be rearranged into the following linear regression:

$$\gamma_0^y(m) = \gamma(m)\theta + \varepsilon_n(m), \quad m = 0, 1, \cdots M. \tag{19}$$

The $2n$-row vector of regressors $\gamma(m)$ is defined by:

$$\gamma(m) = (\gamma_1^y(m), \cdots \gamma_n^y(m), \gamma_1^u(m), \cdots \gamma_n^u(m)) \tag{20}$$

and the pairs ($\gamma_j^u(m)$, $\gamma_j^y(m)$) comprising $\gamma(m)$, $j=0,1 \cdots n$, are defined by:

$$\left[\gamma_j^u(m), \gamma_j^y(m) \right] = \Delta^n (im\omega_0)^{n-j} \left[U(m), Y(m) \right]. \tag{21}$$

Taking into account the fact that (19) is complex-valued, a cost function for a weighted least squares minimization is defined by:

$$J(\theta) = (Y_c - \Gamma_c \theta)' W^{-1} (Y_c - \Gamma_c \theta) \tag{22}$$

where the following notation applies:

$$Y_c = \begin{bmatrix} \mathrm{Re}\ \gamma_0^y(0) \\ \vdots \\ \mathrm{Re}\ \gamma_0^y(M) \\ \mathrm{Im}\ \gamma_0^y(0) \\ \vdots \\ \mathrm{Im}\ \gamma_0^y(M) \end{bmatrix}, \quad \Gamma_c = \begin{bmatrix} \mathrm{Re}\ \gamma(0) \\ \vdots \\ \mathrm{Re}\ \gamma(M) \\ \mathrm{Im}\ \gamma(0) \\ \vdots \\ \mathrm{Im}\ \gamma(M) \end{bmatrix} \tag{23}$$

with "Re" and "Im" meaning *real* and *imaginary* part, respectively. Assuming linearly independent regressors, minimizing (22) leads to the least squares estimate:

$$\hat{\theta} = (\Gamma_c{}' W^{-1} \Gamma_c)^{-1} \Gamma_c{}' W^{-1} Y_c. \tag{24}$$

This estimate cannot be biased by unknown initial or boundary conditions in the data owing

to the properties inherent in the modulating functions of Shinbrot type. It remains to specify the weighting matrix W^{-1} which will be done iteratively in a 'relaxation' algorithm that takes into account the parameter dependent form evident in (17) along with the white Gaussian assumption on the measurement noise $v(t)$. However, the algorithm in its simplest (deterministic) least squares format:

$$\hat{\theta}_{LS} = (\Gamma_c' \Gamma_c)^{-1}\Gamma_c' Y_c$$

i.e., the estimate (24) with $W=I$, will first be compared with a well-known prediction error algorithm in order to illustrate its performance in the most elemental terms.

3.2. A Comparison Based on Simulation

Consider the second order system over a fixed $10s$ time interval:

$$\ddot{y}(t) + 3\dot{y}(t) + 8y(t) = 5u(t), \quad 0 \le t \le T = 10s$$

which possesses the transfer function: $H(s) = 5/(s^2+3s+8)$. Shown in Fig. 3 for point of reference is: (a) the time domain step response of the system, (b) the frequency domain magnitude plot, and (c) the effect of sampling rate on the parameter estimation errors under ideal zero-noise conditions. The latter is discussed below. Since there are only 3 unknown parameters and $n=2$, M was chosen as $M=6$. This means that $(M+n)F_0=0.8$ Hz is the highest frequency that will be extracted from the data, which is roughly the bandwidth of the system as seen in Fig. 3(b). This also assures adherence to the guidelines (9a) and (9b).

TIME (s)   FREQUENCY (Hz)   SAMPLING FREQ (Hz)

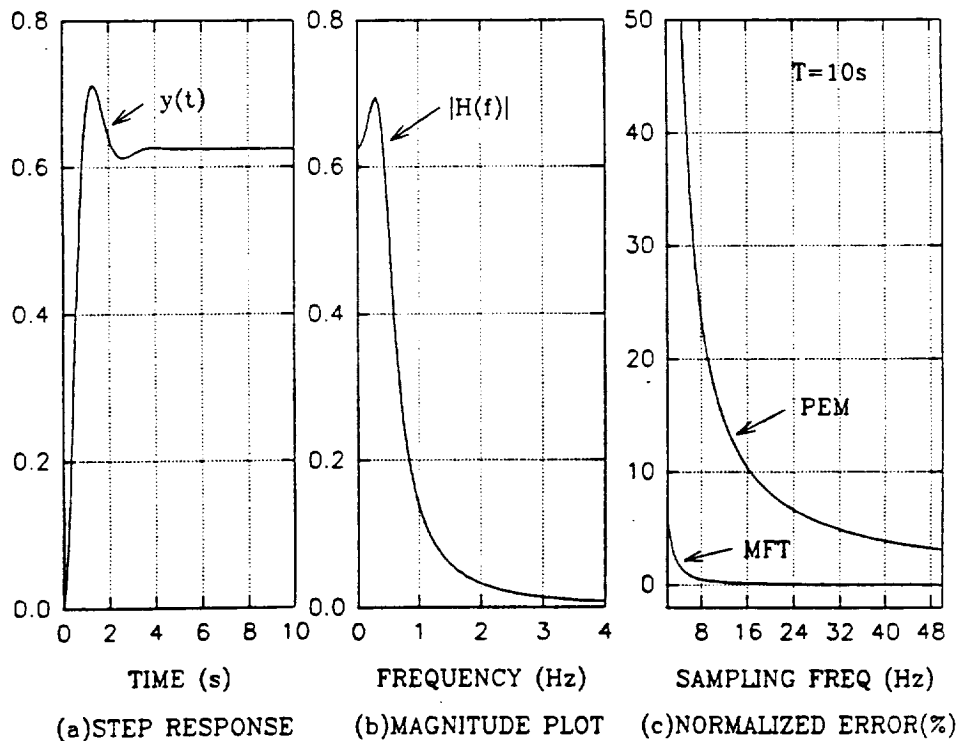(a)STEP RESPONSE   (b)MAGNITUDE PLOT   (c)NORMALIZED ERROR(%)

Fig. 3. Aspects of the simulated system

The output $y(t)$ was corrupted by additive white Gaussian measurement noise. Two hundred Monte Carlo runs were made for each of several noise-to-signal ratios (NSR) under two separate conditions: (i) the initial conditions fixed at (0,0), and (ii) the initial conditions randomized for each run, i.e., a total of 400 Monte Carlo runs at each NSR. The input was $u(t)=\sin t^2/5$ over the $T=10s$ time interval for each run. All calculations were carried out in MATLAB. The sampling rate was fixed at 25.6 Hz, thus facilitating a standard 256 point DFT for the 10s time interval.

The results are shown in Fig. 4 where "LS/MFT" means the unweighted least squares estimate based on the modulating function technique with the weighting matrix W chosen simply as the identity matrix. Also shown is the estimate based on the prediction error method (PEM) from the Identification Toolbox in MATLAB (Ljung, 1991).

(a) NORMALIZED BIAS (%)
VERSES NSR (%)

(b) NORMALIZED STD (%)
VERSES NSR (%)
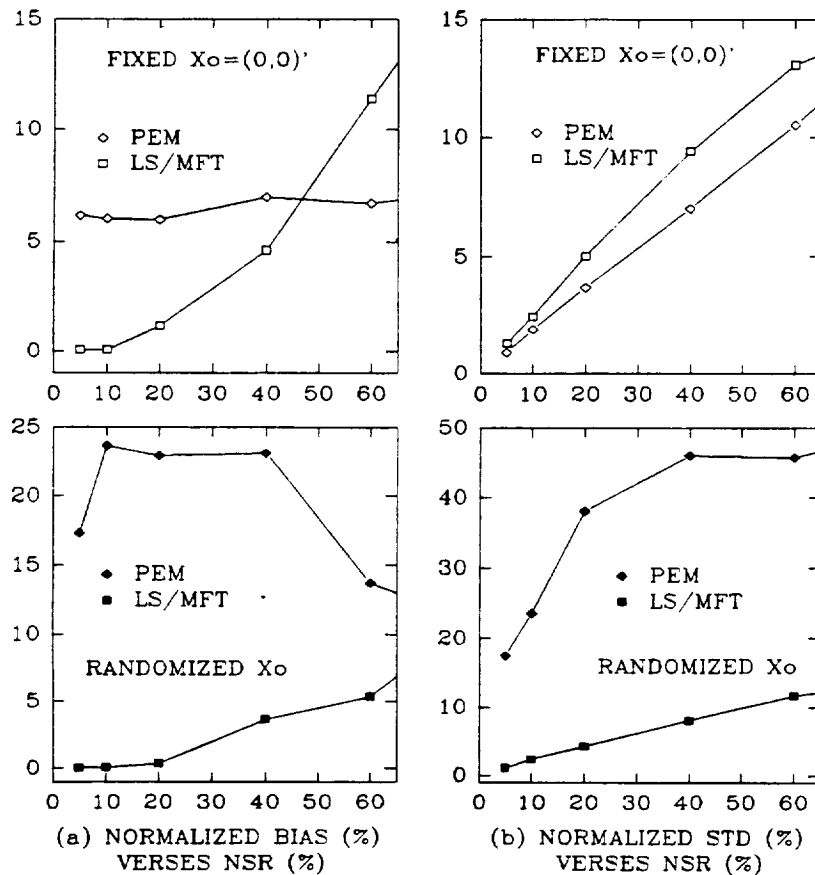
Fig. 4. Output measurement noise effects for the LS/MFT and PEM.

In the case of nonrandom (fixed) initial conditions shown in the top halves of Fig. 4, the PEM shows a fairly constant 6≈7% normalized bias error [4] as the NSR increases from 0% to 60%,

---

[4] Defined as $\left[\dfrac{1}{n_\theta}\sum_{j=1}^{n_\theta}\left[\dfrac{\hat{\theta}_j-\theta_j{}^*}{\theta_j{}^*}\right]^2\right]^{1/2}\cdot 100\%$ where $\hat{\theta}_j$ is the ensemble average of the MFT estimates for the true $\theta_j{}^*$.

whereas the bias for the MFT gradually increases from 0% to $\approx 12\%$ over the same range of NSR values. Meanwhile, the normalized standard deviations are nearly the same for this case. However, in the case of randomized initial conditions, see the lower halves of Fig. 4, both the bias and the standard deviation for the PEM show a several-fold increase and, moreover, each exhibits rather unpredictable behavior that is difficult to reproduce even with additional Monte Carlo runs. By contrast, the bias and standard deviation are very repeatable for the MFT and both have decreased at each NSR indicating that nonzero initial conditions have no deleterious effect on the MFT.

The above comparative results have been verified on other examples as well (Fullerton, 1991). One reason for the superior performance of the MFT is that the MFT does not have to estimate unknown initial conditions; another is related to the fact that the MFT is a direct identification technique for continuous-time models, while the PEM first estimates the parameters of a discrete-time model then converts this to a continuous-time model. Also, the low frequency Fourier series coefficients needed by the MFT can be computed with sufficient accuracy using modest sampling rates. Additional insight in this regard is gained by referring to Fig. 3(c) which shows the influence of sampling rate on the percent normalized error under zero measurement noise conditions. Apparently the PEM needs a much higher sampling rate for the given 10s of data, or a much longer T time interval for the given 25.6 Hz sampling rate, in order to achieve a small estimation error. The influence of sampling rate on the estimation error for the MFT is much more consistent with the Nyquist sampling theorem in view of the frequency magnitude plot in Fig. 3(b). Finally, it is interesting to note that the MFT requires substantially less computer time for each set of Monte Carlo runs (about a ten to five fold decrease depending on whether or not the PEM has to estimate the initial conditions).

4. Cost Functions for Aircraft Models

4.1. Longitudinal Dynamics

A joint cost function that reflects the equation errors for each model in (1)-(2) is specified as follows:

$$J_1(\theta_1,\theta_2) = (Y_1 - \Gamma_1\theta_1)'W^{-1}(Y_1 - \Gamma_1\theta_1) + v(Y_2 - \Gamma_2\theta_2)'W^{-1}(Y_2 - \Gamma_2\theta_2) \qquad (25)$$

where v is a scaling parameter introduced to accommodate the unknown variances in the two output measurement noises. The subscripts 1 and 2 are used to distinguish the parameters and data associated with the two equations (1) and (2) for this model. Specifically,

$$\theta_1 = \begin{bmatrix} -\theta_a \\ \theta_{b_1} \end{bmatrix}, \quad \theta_2 = \begin{bmatrix} -\theta_a \\ \theta_{b_2} \end{bmatrix} \qquad (26)$$

where $(\theta_a,\theta_{b_1},\theta_{b_2})$ are the parameter vectors comprising the coefficients of $(A_1(s), B_1(s), B_2(s))$ respectively. Likewise, the regressand vectors (Y_1, Y_2) and regression matrices (Γ_1, Γ_2) are defined akin to (23) in a way that makes the vector-matrix multiplications conformable with the defined parameter vectors. Since the parameter vector θ_a associated with $A_1(p)$ is coupled into each equation error, the regression matrices are partitioned according to

$$\Gamma_1 = [\Gamma_{a_1}, \Gamma_{b_1}], \quad \Gamma_2 = [\Gamma_{a_2}, \Gamma_{b_2}] \tag{27}$$

such that (25) can be rewritten as

$$J_1(\theta_a,\theta_{b_1},\theta_{b_2}) = (Y_1 + \Gamma_{a_1}\theta_a - \Gamma_{b_1}\theta_{b_1})'W^{-1}(Y_1 + \Gamma_{a_1}\theta_a - \Gamma_{b_1}\theta_{b_1})$$

$$+ v(Y_2 + \Gamma_{a_2}\theta_a - \Gamma_{b_2}\theta_{b_2})'W^{-1}(Y_2 + \Gamma_{a_2}\theta_a - \Gamma_{b_2}\theta_{b_2}). \tag{28}$$

4.2. Lateral Dynamics

A joint cost function that reflects a weighted composite of the equation errors for (3)-(5) is

$$J_2(\theta_1,\theta_2,\theta_3) = \sum_{i=1}^{3} v_{i-1}(Y_i - \Gamma_i\theta_i)'W^{-1}(Y_i - \Gamma_i\theta_i), \quad v_0 = 1 \tag{29}$$

where the (v_1,v_2) are scaling parameters to accommodate the unknown output measurement noise variances and permit adjustments to the minimal values of the individual equation error residuals. The parameter vectors are defined analogous to (26):

$$\theta_i = \begin{bmatrix} -\theta_a \\ \theta_{b_i} \end{bmatrix}, \quad i = 1,2,3 \tag{30}$$

where (θ_a,θ_{b_i}), $i=1,2,3$, are the parameters comprising the coefficients in the polynomials $A_2(p)$ and $(B_{i1}(p), B_{i2}(p))$, $i=1,2,3$, respectively. With a partitioning of the regressor matrices Γ_i akin to (27), the joint cost function (29) can be rewritten as

$$J_2(\theta_a,\theta_{b_1},\theta_{b_2},\theta_{b_3}) = \sum_{i=1}^{3} v_{i-1}(Y_i + \Gamma_{a_i}\theta_a - \Gamma_{b_i}\theta_{b_i})'W^{-1}(Y_i + \Gamma_{a_i}\theta_a - \Gamma_{b_i}\theta_{b_i}). \tag{31}$$

4.3. The AWLS/MFT Algorithm

Equating to zero the partial derivatives of the cost functions (28) and (31) with respect to their arguments is a necessary condition for a weighted least squares estimate of the parameters for each model set. In the case of (28), this procedure leads to the following coupled set of equations for $(\theta_a,\theta_{b_1},\theta_{b_2})$:

$$(\Gamma_{a_1}'W^{-1}\Gamma_{a_1} + v\Gamma_{a_2}'W^{-1}\Gamma_{a_2})\theta_a = -\Gamma_{a_1}'W^{-1}Y_1 - v\Gamma_{a_2}'W^{-1}Y_2$$

$$+ \Gamma_{a_1}'W^{-1}\Gamma_{b_1}\theta_{b_1} + v\Gamma_{a_2}'W^{-1}\Gamma_{b_2}\theta_{b_2} \tag{32}$$

$$\Gamma_{b_1}'W^{-1}\Gamma_{b_1}\theta_{b_1} = \Gamma_{b_1}'W^{-1}Y_1 + \Gamma_{b_1}'W^{-1}\Gamma_{a_1}\theta_a \tag{33}$$

$$\Gamma_{b_2}'W^{-1}\Gamma_{b_2}\theta_{b_2} = \Gamma_{b_2}'W^{-1}Y_2 + \Gamma_{b_2}'W^{-1}\Gamma_{a_2}\theta_a. \tag{34}$$

Here it is important to note that each equation has been arranged such that the left hand side contains the symmetric normal-type coefficient matrix, premultiplying the subparameter θ_a, θ_{b_1} or θ_{b_2}, which would be used to obtain the least squares estimate of that particular

subparameter given the other subparameter values on the right hand side and given the weighting matrix W^{-1}.

A basic question considered in Pearson and Shen (1993c) for n^{th} order SISO systems is: How to choose the weighting matrix W? Focusing on the partitions:

$$W = \begin{bmatrix} W_R & W_{RI} \\ W_{IR} & W_I \end{bmatrix}$$

where $\{W_R\}_{m,m+l} = E[\varepsilon_R(m)\cdot\varepsilon_R(m+l)]$, etc., and the subscripts R and I denote real and imaginary parts, it is shown therein that $W_{RI}=0$, $W_{IR}=0$, and the submatrices (W_R, W_I) are nearly identical, differing only in the entries corresponding to the zero (DC) frequency, cf. **Property 3**. Moreover, (W_R, W_I) are banded Toeplitz structures (the width of the band being n) that depend in a known parametric way on the θ_a parameters via the frequency dependent function $\alpha(k, m, \theta_a)$ defined in (18). Denoting this functional dependence by $W=g(\theta)$, the exact expression of which is given in Pearson and Shen (1993c) but can be gleaned from **Property 3**, the SISO "adaptive weighted least squares" algorithm iteratively seeks a solution for the pair (θ_{AWLS}, W) from the equations, cf. (24):

$$\theta_{AWLS} = (\Gamma_c' W^{-1} \Gamma_c)^{-1} \Gamma_c' W^{-1} Y_c \qquad (35a)$$

$$W = g(\theta_{AWLS}). \qquad (35b)$$

The initial guess for starting the iterations is either $W=I$, or the weighting matrix that arises in the *ideal* situation in which the equation error $e(t)$ is just a Gaussian white noise process, i.e., not dependent on the θ parameter. The resulting solution to (35a) is inserted into (35b) thereby obtaining a new estimate of W which is reinserted into (35a), etc. This "relaxation" type algorithm has been used by many researchers as a tool to achieving a weighted least squares estimate that is close to a maximum likelihood estimate under certain conditions. It almost always converges though there is no known proof of this factorage.

The following point should be emphasized in comparing the various estimates available under the MFT framework: experience shows that the AWLS estimate is far less sensitive to the chosen bandwidth F_B than the ordinary (unweighted) least squares estimate. As evidence of this statement, a fourth order SISO model was used to relate the longitudinal pilot stick command to the body axis pitch rate $q(t)$, the transfer function of which had a bandwidth around 0.5Hz. During the modeling process, a set of modulating function bandwidths ranging from 0.3Hz $\approx$ 1.0Hz (M ranging from 12 to 40) was chosen for the LS/MFT and AWLS/MFT algorithms. An in-between type algorithm WLS/MFT was also used for comparison; this is the algorithm obtained by fixing the weighting matrix W to that which corresponds to the ideal white-noise residuals case mentioned above. In this case, the weighting can be obtained in closed-form fashion from **Property 3** by substituting $P(s)=1$ and utilizing the identity:

$$\sum_{k=m+l}^{m+n} c_{k-m} c_{k-m-l} = \frac{(-1)^l (2n)!}{(n-l)!(n+l)!}.$$

The results are summarized in Table 1 where the goodness-of-fit ratio, s/e, is defined in (36) below. The 'constrained' and 'unconstrained' columns under AWLS/MFT refer to use of this algorithm in a mode whereby the estimated model is forced to have stable poles (constrained), or not (unconstrained). This constrained option for AWLS is invoked by checking the poles

of the estimated model during the iterative process and shifting any right half plane poles to the left half plane, via reflection about the imaginary axis, before continuing the iterations in case that event occurs. This can occur for aircraft modeling since the systems are frequently lightly damped. Concerning the results in Table 1, it is seen that the models for the LS/MFT and WLS/MFT are either stable with poor s/e ratios, or unstable with unacceptable s/e's. By contrast, the estimated systems for both constrained and unconstrained AWLS/MFT algorithms yield much more consistent time domain performance with no drastic variations like the LS/MFT and WLS/MFT algorithms, regardless whether the estimated models are stable or not. Hence, these results help establish the claim that the AWLS/MFT algorithm not only performs better, but is less sensitive to the pre-chosen modulating bandwidth F_B.

algorithm $F_B(Hz)/M$	AWLS constrained	AWLS unconstrained	WLS	LS
0.3/12	stable s/e = 5.84dB	stable s/e = 5.84dB	unstable s/e = - 4.78dB	stable s/e = - 6.79dB
0.4/16	stable s/e = 6.83dB	stable s/e = 6.83dB	unstable s/e = 2.09dB	unstable s/e = - 3.39dB
0.5/20	stable s/e = 7.13dB	stable s/e =7.13dB	stable s/e = 1.86dB	unstable s/e = 1.53dB
0.6/24	stable s/e = 6.91dB	unstable s/e = 7.56dB	stable s/e = 3.08dB	unstable s/e = - 7.41dB
0.7/28	stable s/e = 6.51dB	unstable s/e = 8.07dB	stable s/e = 3.32dB	unstable s/e = - 53.22dB
0.8/32	stable s/e = 6.41dB	unstable s/e = 8.10dB	unstable s/e = - 27.63dB	unstable s/e = - 41.04dB
0.9/36	stable s/e = 6.56dB	unstable s/e = 7.94dB	unstable s/e = - 33.98dB	unstable s/e = - 27.60dB
1.0/40	stable s/e = 6.47dB	unstable s/e = 8.00dB	unstable s/e = - 61.97dB	unstable s/e = - 121.1dB

Using physical flight data to build a fourth order model linking the longitudinal pilot stick movement (input) and the body pitch rate (output).

Table 1. Sensitivity of MFT algorithms to the chosen modulating bandwith

The following relaxation type algorithm is an extension of the algorithm just described for solving (35) but aimed at solving the set of equations (32) - (34), together with the equation for W that corresponds to (35b) in this case, over the parameter set $(\theta_a, \theta_{b_1}, \theta_{b_2}, W)$:

The AWLS/MFT Algorithm for the Longitudinal Dynamics Model (1)-(2)

1. Pick a scaling parameter value for ν, $\nu \geq 0$, and estimate an initial value for the pair (θ_a, W) through the model:

$$A_1(p)[\alpha(t) + q(t)] = [B_1(p) + B_2(p)]u(t)$$

    using the SISO AWLS/MFT algorithm relative to (35).
2. Substitute the values for (θ_a, W) from step 1 into (33) and (34), then solve for the pair $(\theta_{b_1}, \theta_{b_2})$.
3. Estimate a new θ_a from (32) using the values for $(\theta_{b_1}, \theta_{b_2}, W)$ from the previous step.
4. Check if the parameter value for the new θ_a has changed or not, based on a user-chosen percent change in norm. If yes, compute a new value for W from (35b) and go back to step 2, otherwise stop.
5. Check the system output-signal-to-output error ratios S/E, defined below in (36), for each of the two models (1) and (2) to see if they are in rough agreement with each other. If not, try a new value for ν and repeat steps $1 \sim 5$.

The algorithm for the lateral dynamics is similar to the above, i.e., the partial derivatives of (31) with respect to its arguments are equated to zero and arranged in a manner analogous to (32) - (34). Further details are available in Shen (1993), including a computationally efficient algorithm for inverting W (total flops of order M^3) that exploits its banded structure at each step and is very robust for large (M, n).

4.4. Modeling Results

The results of applying the algorithm to flight data for a high performance aircraft will be discussed in this section and, whenever possible, comparison will be made with the modeling results from an established maximum likelihood/output error algorithm which is based in the time domain. (See description of the latter algorithm in Klein (1993).) The figure of merit used to assess the quality of the resulting model in relation to the given data is the output-signal-to-output-error ratio, S/E, (in decibels) as defined by

$$S/E = 20 \log_{10} \frac{\text{RMS}(y)}{\text{RMS}(e)}, \quad e(t) = y(t) - \hat{y}(t), \quad 0 \leq t \leq T \tag{36}$$

where RMS(y) means the root-mean-square value of the output signal $y(t)$ over the $[0,T]$ time interval, and $\hat{y}(t)$ is the estimated output using the model.[5] Subscripts on S/E are used to distinguish a particular output among the multivariable channels.

---

[5] When required, an observable state equation model is employed and the unknown initial condition for generating $\hat{y}(t)$ is estimated from a Luenberger observer running backwards in time driven by the given input/output data over the $[0,T]$ time interval.

The data for the longitudinal case and the corresponding modeling results are given in Fig. 5 and Table 2, respectively.
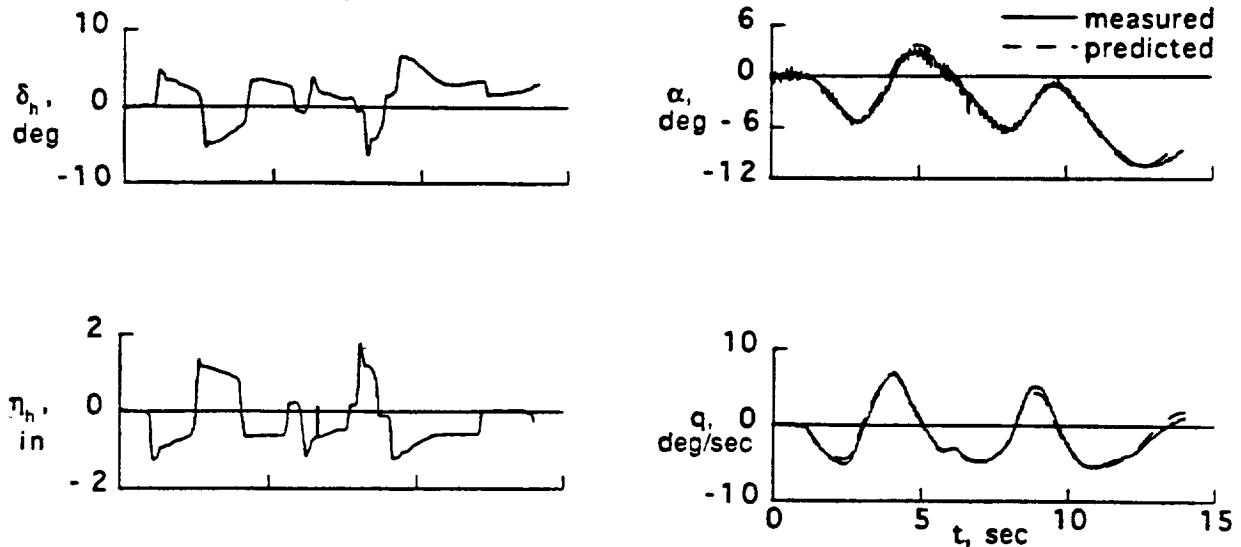

Fig. 5. Input/output data for the longitudinal dynamics, and superimposed model output responses.

The parameter symbols in Table 2 are defined by the transfer function relations:

$$\frac{\alpha(s)}{u(s)} = \frac{A_\alpha s + B_\alpha}{s^2 + K_1 s + K_0}, \qquad \frac{q(s)}{u(s)} = \frac{A_q s + B_q}{s^2 + K_1 s + K_0}$$

where u can be either δ_h or η_h depending on whether it is the transfer function obtained by using as input the signal "inside" the loop, referred to as 'Open Loop' in Table 2, or "outside" the loop, referred to as 'Closed Loop' in Table 2. (See Fig. 1.)

Parameter	Open Loop		Closed Loop
	ML	AWLS/MFT	AWLS/MFT
K_0	0.695	0.698	0.494
K_1	0.360	0.324	0.995
A_α	0.002	-0.065	-0.013
B_α	-1.558	-1.523	0.162
A_q	-1.479	-1.488	0.143
B_q	-0.328	-0.273	0.043
$(S/E)_\alpha$ (dB)	19.98	19.30	18.49
$(S/E)_q$ (dB)	18.15	17.73	12.75

Table 2. Modeling results for the longitudinal dynamics.

The symbol ML in Table 2 refers to the maximum likelihood/output error identifier described in Klein (1993). Clearly, the S/E ratios for the Open Loop case are nearly the same for each method in each of the respective output channels. The same is true for the Closed Loop case, although only the AWLS/MFT values are listed for the sake of brevity. Moreover, the measured and predicted model responses are quite close to one another, as evidenced by the right hand sides of Fig. 5.

The scaling parameter value used for the AWLS/MFT results in Table 2 was $v=0$. The results of using other v values for the Closed Loop case are depicted in Fig. 6 showing the tradeoffs that can be obtained by weighting the two output channels differently. The 'iterations' referred to in Fig. 6 count the number of loops passing through the AWLS/MFT algorithm listed in the previous section. Ordinarily one should expect the curves of the type shown in Fig. 6 to be monotonic, leveling off to some asymptotic values for each channel. This has been the case for other studies (Shen, 1993), but here it possibly indicates a straining of the modeling assumptions, e.g., linearity.
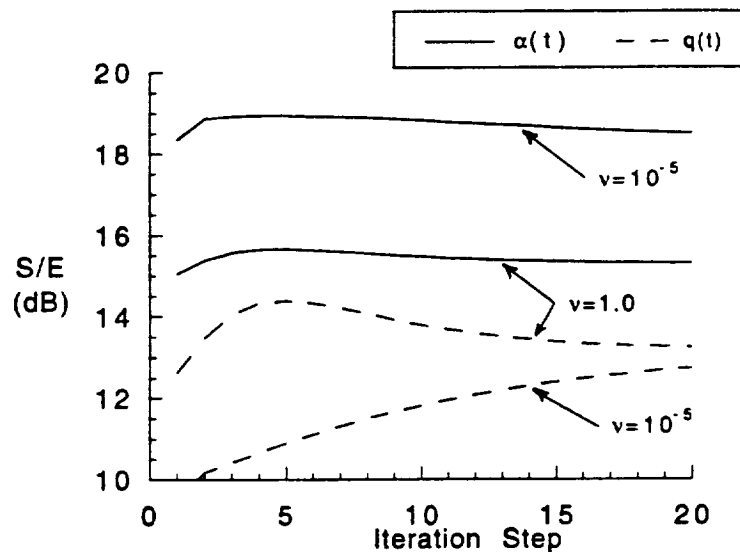

Fig. 6. Effect of different v values on S/E ratios for the longitudinal dynamics.

The data and resulting transfer functions for the lateral dynamics study are shown in Fig. 7 and Table 3, respectively. Only the AWLS/MFT technique was successful in yielding a useful result in this case, i.e., the ML technique failed. The scaling parameters were both chosen as unity for the AWLS/MFT, i.e., $v_1=v_2=1$. As seen at the bottom of Table 3, reasonable parity was achieved in the S/E ratios for each of the three output channels; hence, there is no compelling need to change the v values from unity in this case. Again, the measured and predicted model transient responses are in good agreement with each other as shown by the three plots on the right hand side of Fig. 7.
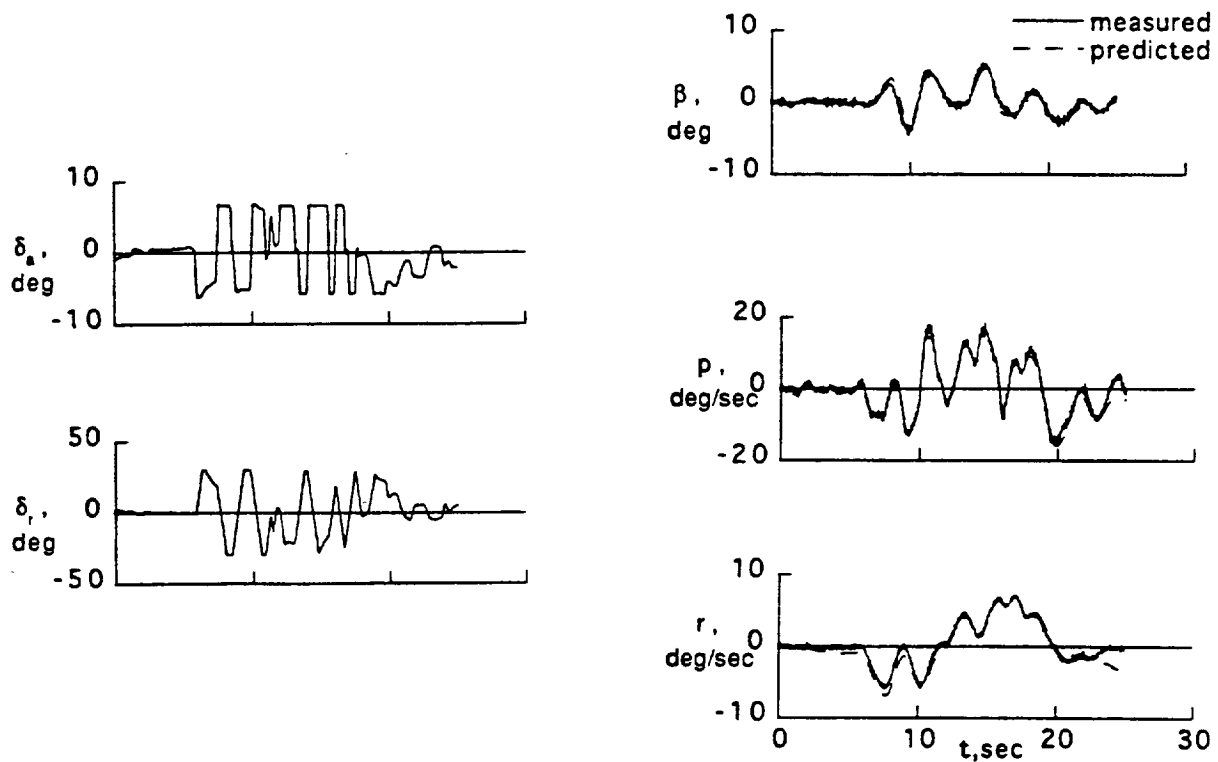
Fig. 7. Input/output data for the lateral dynamics, and superimposed model output responses.

Table 3. Modeling results for the lateral dynamics.

$$\frac{\beta}{\delta_a} = \frac{1.8131s^2 + .3099s + .2001}{s^4 + .3108s^3 + 3.2640s^2 + .2059s + .2536} \qquad \frac{\beta}{\delta_r} = \frac{.0082s^3 + .3701s^2 + .0819s + .0543}{\Delta}$$

$$\frac{p}{\delta_a} = \frac{3.0633s^3 - 2.2515s^2 - 1.2644s - .0049}{\Delta} \qquad \frac{p}{\delta_r} = \frac{.1909s^3 - .5163s^2 - 1.2768s + .1098}{\Delta}$$

$$\frac{r}{\delta_a} = \frac{-.1627s^3 - .1845s^2 - 1.3654s - .0148}{\Delta} \qquad \frac{r}{\delta_r} = \frac{-.2031s^3 - .1630s^2 - .7912s - .0962}{\Delta}$$

$$(S/E)_\beta = 9.12 \ (dB) \qquad (S/E)_p = 12.31 \ (dB) \qquad (S/E)_r = 9.96 \ (dB)$$

5. MFT Extensions

5.1. Frequency Transfer Function Analysis

Methods for determining the transfer function in the frequency domain for a stable linear system from input/output data include classical correlation and spectral analyses, as well as the direct ratio of Fourier transforms and steady state sinusoidal measurements. Each of these "nonparametric" identification techniques require either a statistical stationarity assumption on the data, or a periodic steady state condition to be established, before initiating calculations of the transfer function at pertinent frequencies (Ljung, 1987). Notwithstanding noise considerations, long data lengths may be required in order to achieve good accuracy due to the stationarity or steady state assumption, or to eliminate end point effects in case a direct ratio of Fourier transforms is used on time-limited data. By way of contrast, the modulating functions (10) will be used in this Section to extract the frequency content in short data lengths in order to set up a least squares estimation of the transfer function at selected frequencies. Since short data lengths are used there is no assumption of steady state operation or stationarity of the data, though there must be present sufficient energy content in the data at the specified frequencies in order to avoid degeneracy in the least squares estimate.

Consider a stable linear system with two inputs $(u(t),v(t))$ and a single output $y(t)$ modeled by the equation:

$$A(p)y(t) = B(p)u(t) + C(p)v(t) + e(t) \tag{37}$$

where $(A(p),B(p),C(p))$ are polynomials in the differential operator $p=d/dt$ of degree less than or equal to an a priori integer n, and $e(t)$ again represents the effect of modeling errors. Given the input-output data $[u(t),v(t),y(t)]$ over a finite set of time intervals $\{[t_r,t_r+T], r=1, \cdots N\}$, the problem considered here is to estimate the transfer functions $G(i\omega)=B(i\omega)/A(i\omega)$ and $H(i\omega)=C(i\omega)/A(i\omega)$ at a finite set of frequencies $\{k\omega_0, k=0,1 \cdots M+n\}$,[6] where ω_0 is the user selected 'resolving' frequency and M a chosen nonnegative integer. The time intervals are each chosen of length $T=2\pi/\omega_0$ and need not necessarily be disjoint.

As in the previous formulations, the model (37) is rearranged into the equation error format and modulated with $\phi_{m,n}(t)$. Thus, with respect to the time interval $[t_r,t_r+T]$:

$$\int_0^T \phi_{m,n}(t) \left[A(p)y(t+t_r) - B(p)u(t+t_r) - C(p)v(t+t_r) \right] dt = \int_0^T \phi_{m,n}(t)e(t+t_r)dt.$$

Invoking **Property 2**, Eq. (12a):

$$\sum_{k=m}^{m+n} c_{k-m} \left[A(ik\omega_0)Y_k[r] - B(ik\omega_0)U_k[r] - C(ik\omega_0)V_k[r] \right] = \sum_{k=m}^{m+n} c_{k-m}E[r] \tag{38}$$

where the following notation applies to the Fourier series coefficients in the above (Z can be U, V, Y or E):

---

[6] Thus, assuming the system bandwidth ω_B is known, the choice in (M,ω_0) such that $(M+n)\omega_0 \approx \omega_B$ covers the bandwidth at the knots $k\omega_0$, $k=0,1 \cdots M+n$.

$$Z_k[r] = \frac{1}{T}\int_0^T z(t+t_r)\cos k\,\omega_0 t\,dt - i\frac{1}{T}\int_0^T z(t+t_r)\sin k\,\omega_0 t\,dt$$

$$= \quad Z_k^c[r] \qquad - i\, Z_k^s[r]. \tag{39}$$

Define parameters $(\alpha_k{}^R, \alpha_k{}^I, \beta_k{}^R, \beta_k{}^I, \gamma_k{}^R, \gamma_k{}^I)$ for the real and imaginary parts of the polynomials $(A\,(ik\,\omega_0), B\,(ik\,\omega_0), C\,(ik\,\omega_0))$, each multiplied by c_{k-m}, e.g.,

$$c_{k-m}A\,(ik\,\omega_0) = \alpha_k{}^R + i\,\alpha_k{}^I.$$

Values of the transfer functions $G\,(i\,\omega)$ and $H\,(i\,\omega)$ are related to these parameters:

$$G\,(ik\,\omega_0) = \frac{\beta_k{}^R + i\beta_k{}^I}{\alpha_k{}^R + i\,\alpha_k{}^I}, \quad H\,(ik\,\omega_0) = \frac{\gamma_k{}^R + i\gamma_k{}^I}{\alpha_k{}^R + i\,\alpha_k{}^I}.$$

The DC value can be handled by normalizing $A\,(0)=1$ so $G\,(0)=B\,(0)$ and $H\,(0)=C\,(0)$. With this normalization and the notation of (39), the real and imaginary parts of (38) can be rearranged into the following standard linear least squares equation format represented by stages according to the values assigned to m.

Initial Stage (m=0):

$$\tilde{Y}_0[r] = \sum_{k=0}^{n} \psi_k[r]\theta_k + \varepsilon_0[r], \quad r=1,2\cdots N \tag{40}$$

where the data-related quantities are defined by the vector-matrix equations:

$$\tilde{Y}_0[r] = \begin{bmatrix} Y_0^c[r] \\ 0 \end{bmatrix}, \quad \psi_0[r] = \begin{bmatrix} U_0^c[r] & V_0^c[r] \\ 0 & 0 \end{bmatrix}$$

$$\psi_k[r] = \begin{bmatrix} Y_k^c[r] & Y_k^s[r] & U_k^c[r] & U_k^s[r] & V_k^c[r] & V_k^s[r] \\ -Y_k^s[r] & Y_k^c[r] & -U_k^s[r] & U_k^c[r] & -V_k^s[r] & V_k^c[r] \end{bmatrix} \quad (k\geq 1)$$

and the θ_k parameters are defined by

$$\theta_0 = \begin{bmatrix} G\,(0) \\ H\,(0) \end{bmatrix}, \quad \theta_k = \mathrm{col}\,(-\alpha_k^R, -\alpha_k^I, \beta_k^R, \beta_k^I, \gamma_k^R, \gamma_k^I) \quad (k\geq 1).$$

Stage m (m=1,2 $\cdots$ M):

Again, a 2 dimensional vector equation is derived from the real and imaginary parts of (38) for the case $m\geq 1$ with the same notation as above:

$$\tilde{Y}_m[r] = \psi_{n+m}[r]\theta_{n+m} + \varepsilon_m[r], \quad r=1,2\cdots N \tag{41}$$

where $\tilde{Y}_m[r]$ is defined in terms of the parameters and the data of the preceding stages by

$$\tilde{Y}_m[r] = -\sum_{k=m}^{n+m-1} \psi_k[r]\theta_k.$$

-21-

The residuals for all stages are given by

$$\varepsilon_m[r] = \sum_{k=m}^{n+m} c_{k-m} \begin{bmatrix} E_k^c[r] \\ -E_k^s[r] \end{bmatrix}.$$

Remarks. (i) Comparing (40) and (41), it is seen that the $m=0$ (initial) stage carries the major computational burden since there are many more parameters to be determined at this stage, i.e., $(2+6n)$ for (40) verses just 6 for each succeeding stage in (41). Notice that the same data can be used for each stage and that N must satisfy: $2N>(2+6n)$ in order to have more equations than unknowns. (ii) In a general multivariable situation, the transfer function representation for each output y_j, $j=1,2 \cdots m_y$:

$$y_j(s) = \sum_{k=1}^{m_u} H_{jk}(s)u_k(s) = \sum_{k=1}^{m_u} \frac{B_{jk}(s)}{A_{jk}(s)}u_k(s)$$

facilitates the following model for each j:

$$A_j(p)y_j(t) = \sum_{k=1}^{m_u} \tilde{B}_{jk}(p)u_k(t) + e(t) \tag{42}$$

where $A_j(s)=\prod_{k=1}^{m_u}A_{jk}(s)$ and $\tilde{B}_{jk}(s)$ is defined as the polynomial:

$$\tilde{B}_{jk}(s) = \frac{B_{jk}(s)}{A_{jk}(s)}A_j(s).$$

Therefore, the multi-input single output formulation developed above can be applied to each output equation (42). The fact that the pairs $(A_j(s), \tilde{B}_{jk}(s))$, $k=1, \cdots m_u$, are not generally coprime seems not to be a problem under low measurement noise conditions since it is the ratios $\tilde{B}_{jk}(i\omega)/A_j(i\omega) = B_{jk}(i\omega)/A_{jk}(i\omega)$ that are sought at the specified frequencies. This avoids a difficult issue of sufficient parametrization and minimality for a state space model which is a separate issue.

5.1.1. A MIMO Example

As an example, consider the system with the transfer function relations (notice that $H_2(s)$ is high-pass):

$$y_1(s) = G_1(s)u(s) + G_2(s)v(s)$$

$$= \frac{12s^2+487s+582}{s^3+65s^2+456s+1978}u(s) + \frac{0.7s^2+157.5s+504}{s^3+9s^2+455s+881}v(s)$$

$$y_2(s) = H_1(s)u(s) + H_2(s)v(s)$$

$$= \frac{2s+160}{s^2+20s+160}u(s) + \frac{s^2+50s+54}{s^2+40s+500}v(s).$$

The system was simulated over a 30s time interval using as inputs:

$$u(t) = \text{a random binary signal}$$

-22-

$$v(t) = \sin(25t + 0.9434) + \sin(6t) + \sin(20t^2) + \sin(5t^3).$$

Order $n = 6$ modulating functions were utilized to accommodate the degrees of the polynomials $(A_1(s), A_2(s))$ in (42) for the above transfer functions. Choosing the desired bandwidth to be covered and at what resolving frequency led to the goal of estimating the four frequency functions at the knots $\{k\omega_0, k = 0, 1 \cdots 15\}$ where $\omega_0 = 2\pi/4 = 1.57$ rad/sec, i.e., $\omega_B = 23.55$ rad/sec and T time intervals of $4s$ duration. Taking into account the number of unknowns for the initial stage, i.e., $(2+6n) = 38$, twenty-seven $[t_r, t_r + 4]$ time intervals were chosen with $1s$ overlaps, i.e., $[t_r, t_r + T] = [r, r+4]$, $r = 0, 1 \cdots 26$, which facilitates about 50% more algebraic equations as unknowns upon which to base the least squares estimate for the initial $m = 0$ stage. The sampling frequency was 200 Hz in utilizing the DFT to calculate the Fourier series coefficients of the data on each $4s$ subinterval. In the absence of noise, the estimates of the frequency functions $G_j(i\omega)$ and $H_j(i\omega)$, $j = 1, 2$, are essentially perfect at the knots $\{k\omega_0, k = 0, 1 \cdots 15\}$, $\omega_0 = 2\pi/4 = 1.57$ rad/sec, as expected. In order to test the sensitivity to noise, 50 Monte Carlo runs were made at each of several noise levels with white Gaussian additive noise at the two outputs. The ensemble mean magnitude and phase plots are shown in Fig. 8 for $G_1(i\omega)$ and $G_2(i\omega)$, and in Fig. 9 for $H_1(i\omega)$ and $H_2(i\omega)$, with magnitude on the top half and phase on the bottom half for each transfer function. Also shown for each transfer function is: (i) a low noise condition (SNR = 35 db) shown on the left hand side of each figure, and (ii) a higher noise condition (SNR = 5 db) shown on the right hand side of each figure. Generally, the estimates are better at the low frequency end of the spectrum as might be expected for white measurement noises.

5.2. Linear Time Varying Differential System Models

Consider the linear time-varying differential operator model

$$p^n y(t) + \sum_{j=1}^{n} a_j(t) p^{n-j} y(t) = \sum_{j=1}^{n_u} b_j(t) p^{n_u - j} u(t) + e(t) \tag{43}$$

where the variable coefficients are parametrized by linear combinations of smooth functions. To facilitate generality, it is convenient to adopt the row and column vector notations, "<" and ">", such that each coefficient is represented by an inner product:

$$a_{n-j}(t) = <f_j(t)\ \alpha_j>, \quad b_{n_u - j}(t) = <g_j(t)\ \beta_j> \tag{44}$$

where $<f_j(t)$ and $<g_j(t)$ are given row vector-valued smooth functions, and $(\alpha_j>, \beta_j>)$ are column vector-valued parameters. Stacking the latter columns into a pair of column vector-valued parameters $(\alpha>, \beta>)$, viz.
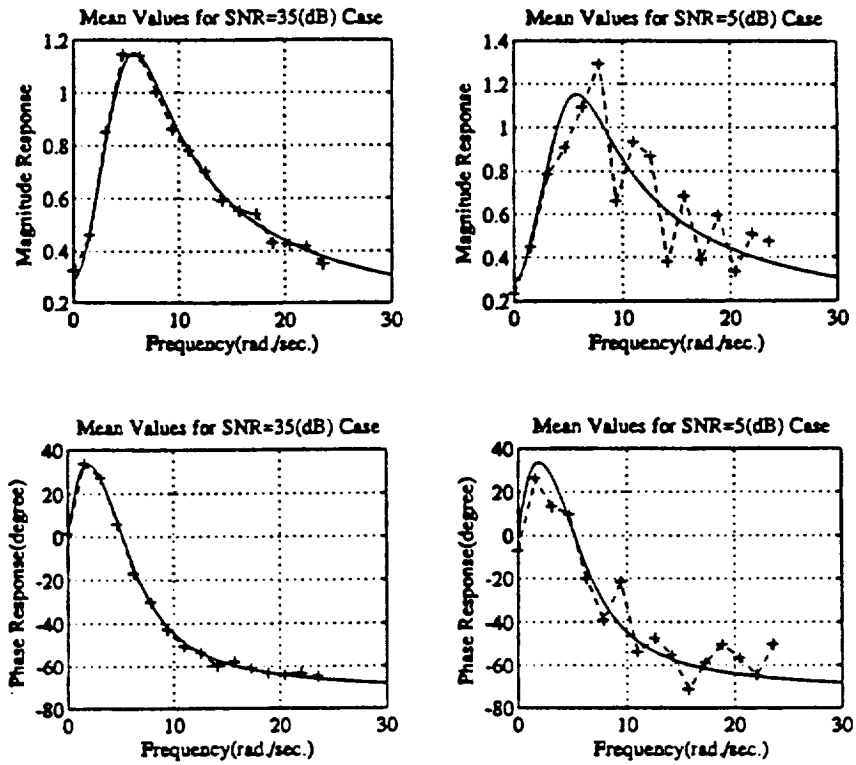
$$\alpha> = \text{col}\ (\alpha_0>, \alpha_1> \cdots \alpha_{n-1}>), \quad \beta> = \text{col}\ (\beta_0>, \beta_1> \cdots \beta_{n_u - 1}>)$$

and utilizing the Leibniz formula in order to achieve a differential operator model amenable to Property 2, it is found that (43) together with (44) is equivalent to the model:

$$p^n y(t) + \left[\sum_{k=0}^{n-1} p^k <\tilde{y}_k(t)\right] \alpha> = \left[\sum_{k=0}^{n_u - 1} p^k <\tilde{u}_k(t)\right] \beta> + e(t) \tag{45}$$

where $<\tilde{u}_k(t)$ and $<\tilde{y}_k(t)$ are row vector functions related to time-modulated versions of the input/output data, modulated in the time domain by derivatives of the given functions $<f_j(t)$
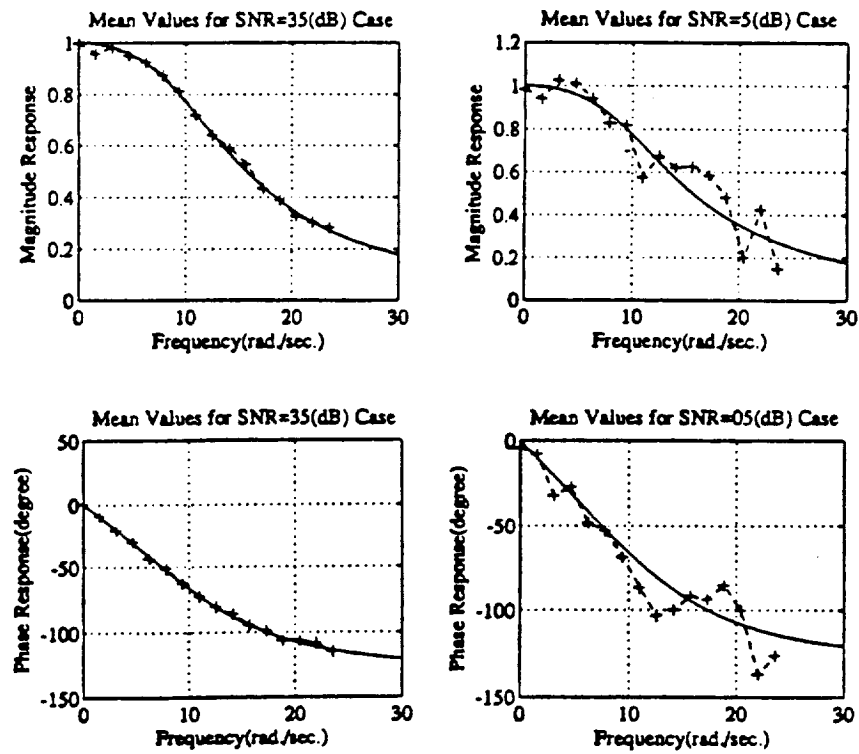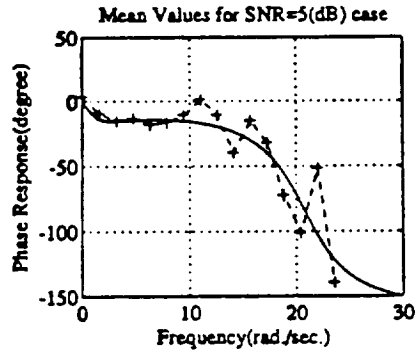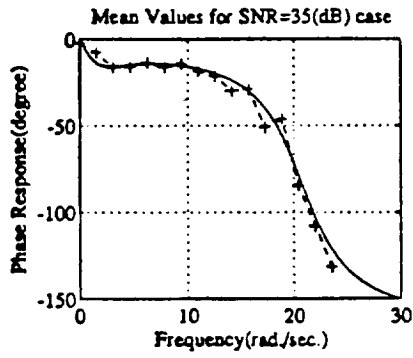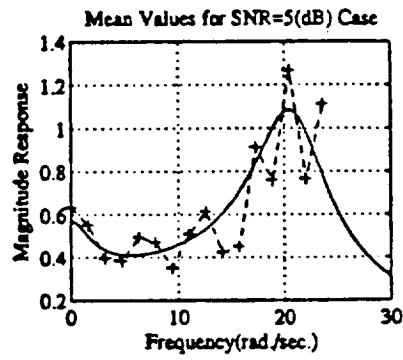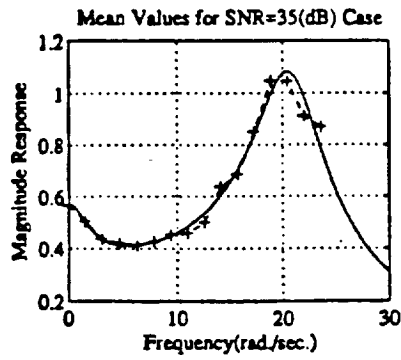
(a) $G_1(i\omega)$

(b) $G_2(i\omega)$

Fig. 8. Measurement noise effects for G_1 and G_2

(a) $H_1(i\omega)$

Mean Values for SNR=35(dB) Case

Mean Values for SNR=5(dB) Case

Mean Values for SNR=35(dB) case

Mean Values for SNR=5(dB) case

(b) $H_2(i\omega)$

Mean Values for SNR=35(dB) Case

Mean Values for SNR=5(dB) Case
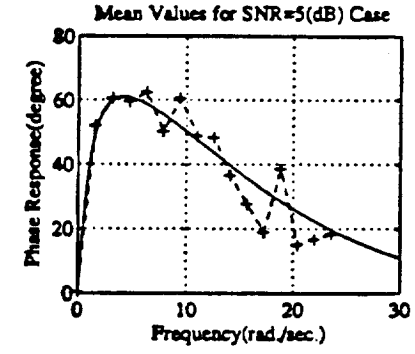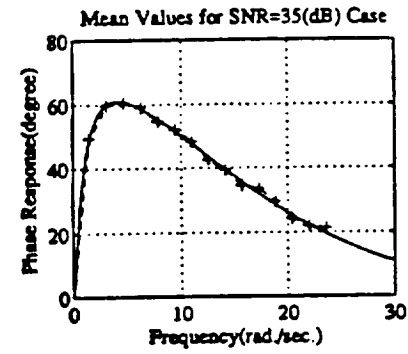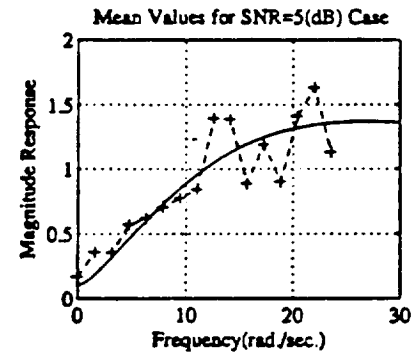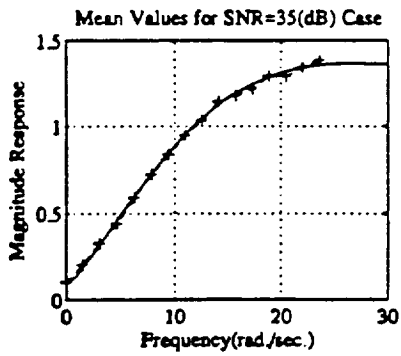
Mean Values for SNR=35(dB) Case

Mean Values for SNR=5(dB) Case

Fig. 9. Measurement noise effects for H_1 and H_2

and $<g_j(t)$, in accordance with the following definitions:

$$<\tilde{u}_k(t) = \text{row}(<v_{0k}(t), <v_{1k}(t) \cdots <v_{n_u-1,k}(t))$$

$$<\tilde{y}_k(t) = \text{row}(<z_{0k}(t), <z_{1k}(t) \cdots <z_{n-1,k}(t))$$

and

$$<v_{jk}(t) = u(t)P_{jk}(p)<g_j(t), \quad <z_{jk}(t) = y(t)P_{jk}(p)<f_j(t)$$

with $P_{jk}(p)$ defined as the differential operator:

$$P_{jk}(p) = \begin{cases} 0 & \text{for } j < k \\ (-1)^{j-k} \begin{bmatrix} j \\ k \end{bmatrix} p^{j-k} & \text{for } j \geq k \end{cases}$$

Applying **Property 2**, Eq. (12), to (45) leads directly to the regression equation (19) where

$$\gamma(m)\theta = \Delta^n \left[\sum_{k=0}^{n-1} (im\omega_0)^k <\tilde{Y}_k(m), \sum_{k=0}^{n_u-1} (im\omega_0)^k <\tilde{U}_k(m) \right] \begin{bmatrix} -\alpha> \\ \beta> \end{bmatrix}.$$

5.2.1. A Variable Damping Example

The following second order system with a parabolically varying damping term and 5 unknown parameters was simulated under a variety of conditions:

$$\ddot{y}(t) + (\zeta + c_1 t + c_2 t^2)\dot{y}(t) + ay(t) = bu(t), \quad 0 \leq t \leq T = 10s.$$

Via the Leibniz formula, this model is equivalent to

$$p^2 y(t) + \zeta p y(t) + c_1 \left[p(ty(t)) - y(t) \right] + c_2 \left[p(t^2 y(t)) - 2ty(t) \right] + ay(t) = bu(t)$$

which is an illustration of (45). The input was $u(t) = \sin t^2/3$ for each run, and the parameters (a,b,c_1,c_2) were fixed at the values: $(8, 5, -1, 0.1)$, respectively. Two cases were studied for the ζ parameter: $\zeta = 3$ in which the "frozen poles" of the system are stable for each $t \in [0,10s]$, and $\zeta = 2$ in which the "frozen poles" are unstable for approximately 50% of the 10s time interval. (The parabolas for these two cases are shown in the top halves of Figs. 11(a) and (b) respectively.) Each ζ case included: (i) 300 Monte Carlo runs at each of several noise-to-signal ratios with white additive measurement noise on the output signal $y(t)$, and (ii) structural errors in the modeling of the damping term by estimating first a linearly time-varying damping term, and then a constant damping term. The data was noise-free for this aspect of the study.

The noise effect results are summarized in Fig. 10 where the left side shows the normalized bias errors and the right side the standard deviations. The results are more or less the same for the two ζ cases and demonstrate that the estimation technique has similarly good noise-immunity properties as exists for time-invariant systems.

The effect of structural errors is shown in Fig. 11 where the left and right hand sides pertain to the $\zeta = 3$ and $\zeta = 2$ cases, respectively. In the absence of noise, the parameter estimates are essentially perfect, as expected, and yield the parabolas for the two cases as shown

in the top halves of Fig. 11. With structural modeling errors, the estimates of the damping terms are as indicated in the top halves where it is noted that modeling the damping as a constant resulted in an unstable system for the $\zeta=2$ case. Using a different input signal : $u(t)=\sin t^4/4$, i.e., different from that which was used to estimate the parameters, the resulting outputs of the estimated models are compared with one another in the bottom halves of Fig. 11. As expected, the more severe the error in structural modeling, the greater the error in the ability of the model to predict the response, though the differences are perhaps not as large as might be anticipated.
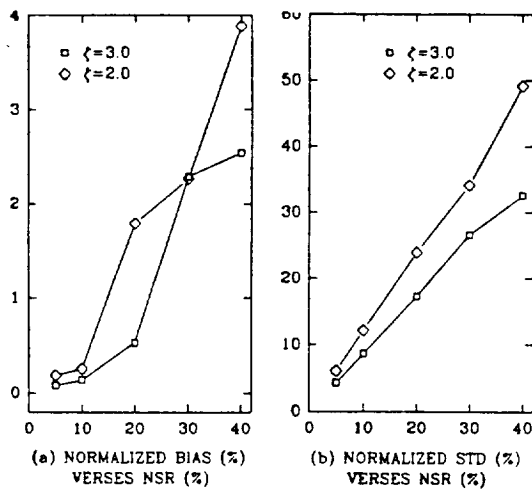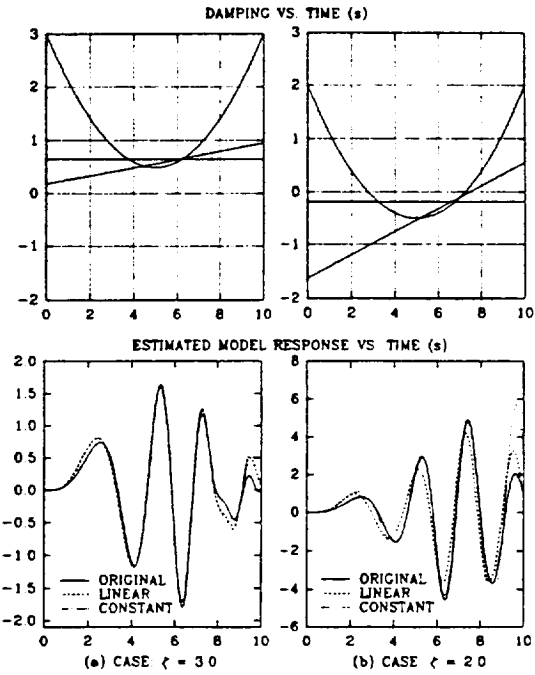

Fig. 10. Measurement noise effects.

Fig. 11. Structural error effects.

5.3. A Class of Nonlinear Input/Output Models

Parametrized state vector equations of the normal form:

$$\dot{x} = f(x,u,\theta)$$

$$y = h(x,u,\theta) \tag{46}$$

constitute the most common starting point for methods aimed towards the parameter identification of deterministic continuous-time nonlinear differential systems. Given the input/output data $[u(t),y(t)]$ on some time interval, $0\le t \le T$, together with the parametrized functions $f(x,u,\theta)$ and $h(x,u,\theta)$, the known methods include: quasilinearization, invariant embedding, state variable filters, and a variety of nonlinear filtering methods such as stochastic approximation and extended Kalman filtering. These methods are invariably *implicit* in terms of the cost function dependence on the parameter vector θ. By contrast, if an equivalent parametrized input/output differential operator equation exists in the following equation-error format $(p=d/dt)$:

$$\sum_{j=0}^{n_1} \sum_{k=1}^{n_2} g_j(\theta) F_{jk}(u,y) P_{jk}(p) E_k(u,y) = 0, \quad g_0 = 1 \qquad (47)$$

then it will be shown below that given the input/output data over $[0,T]$ it is possible to specify an *explicit* cost function of the form

$$J(\theta) = \sum_{j=0}^{n_1} \sum_{k=0}^{n_1} r_{jk} g_j(\theta) g_k(\theta) \qquad (48)$$

possessing the properties: (i) $J(\theta) \geq 0$, with $J(\theta)=0$ for any value of θ that satisfies the model (47), and (ii) the r_{jk} can be computed via integral operations on transient data over the observation time interval $[0,T]$ without the need to estimate unknown initial conditions. Hence, minimizing $J(\theta)$ in order to ameliorate modeling errors leads to a one-shot least squares estimate of the system parameters, which may or may not be unique depending on the adherence of appropriate nondegeneracy conditions on the data.

The formulation leading to (48) is a continuation of earlier developments in Pearson (1988, 1989) relating to the least squares parameter estimation for input/output models like (47) based on Shinbrot's classical method of moment functionals. The formulation in this Section is taken from Pearson (1992) and exploits the **Property 4** delineated earlier.

5.3.1. Modeling Considerations

As a first step it is assumed that the system model can be arranged into the form (47) where the $g_j(\theta)$ are given functions of the parameters θ, the $P_{jk}(p)$ are fixed polynomials of degree n in the differential operator $p = d/dt$, and the $(E_k(u,y), F_{jk}(u,y))$ are specified functions of the pair (u,y).[7] Starting from a state equation model (46), it is certainly not the case that all such models admit to an "external" differential representation (Nijmeijer and Van der Schaft, 1990) much less the special form (47). A simple example illustrating this model is the differential equation for a unit mass particle subject to a given force $u(t)$ and with drag proportional to the velocity squared:

$$\ddot{y} + \theta_1(\dot{y})^2 - \theta_2 u = 0 \qquad (49a)$$

where $y(t)$ is the displacement of the particle. Utilizing the differential identity: $p^2(y^2) = 2yp^2y + 2(py)^2$, (49a) is equivalent to the following differential operator equation which is of the form (47):

$$p^2 y + \theta_1(\tfrac{1}{2}p^2 y^2 - y p^2 y) - \theta_2 u = 0. \qquad (49b)$$

Higher order identities can be employed if a power series like $\sum_{j=1}^{N} \theta_j(\dot{y})^j$, $N>2$, is used to model the dissipation term in (49a).

In some cases we can obtain a model where all the F_{jk}'s in (47) are constants, i.e., the more restricted form:

---

[7] Sufficient smoothness and stability properties are tacitly assumed to assure the existence and uniqueness of bounded solutions $y(t)$ satisfying (47) given any admissible θ, initial conditions and bounded inputs $u(t)$ on $[0,T]$.

$$\sum_{j=0}^{n_1} \sum_{k=1}^{n_2} g_j(\theta)P_{jk}(p)E_k(u,y) = 0, \quad g_0 = 1. \tag{50}$$

In such cases the model is said to be *exact*, a term which is suggested by the notion of an "exact" or "total" differential in the calculus because the equation error for such models can be integrated exactly given the data on $[0,T]$. An example of this class is the following two compartment chemical kinetics model from Walter (1982):

$$\dot{x}_1 = -\theta_1 x_1 - \theta_2(1-\theta_3 x_2)x_1 + u$$

$$\dot{x}_2 = \theta_2(1 - \theta_3 x_2)x_1 - \theta_4 x_2$$

$$y = x_1. \tag{51a}$$

Assuming the output $y(t) > 0$ for all t, the nonmeasured state x_2 can be eliminated from the above model thereby arriving at the following equivalent input/output representation:

$$p^2(\ln y) - p(\frac{u}{y}) + \theta_1\theta_2\theta_3 y + \theta_2\theta_3(py - u) + \theta_4(p(\ln y) - \frac{u}{y}) + (\theta_1 + \theta_2)\theta_4 = 0. \tag{51b}$$

As an important modeling consideration, which applies independently of the method used for identification, it was shown in Pearson (1989) that a necessary condition for a well-posed problem involving (47) is that the vector function $(g_0(\theta), g_1(\theta) \cdots g_{n_1}(\theta))$ be an injective, i.e., single-valued, function of the parameters else the parameter identification problem will be plagued by nonuniqueness. Presuming this condition is upheld and g_0 is normalized to unity, it is, therefore, assumed that the function $g(\theta)$ defined by: $g(\theta) = (g_1(\theta), \cdots g_{n_1}(\theta))$ is injective.

5.3.2. The Cost Functions for Least Squares Minimization

Consider first the class of *exact* differential models (50). Modulating (50) with $\phi_{m,n}(t)$, integrating over $[0,T]$ and applying **Property 2** , we obtain the modulated equation error function $e(\theta,m)$ defined by

$$e(\theta,m) = \sum_{j=0}^{n_1} g_j(\theta)\gamma_j(m) \tag{52}$$

where the $\gamma_j(m)$ play the role of regressors and are defined by

$$\gamma_j(m) = \Delta^n \sum_{k=1}^{n_2} P_{jk}(im\,\omega_0)V_k(m) \tag{53}$$

and the $V_k(m)$ are defined by

$$V_k(m) = \frac{1}{T}\int_0^T E_k(u(t),y(t))e^{-im\,\omega_0 t}\,dt. \tag{54}$$

Forming the norm of $e(\theta,m)$ over the frequency range $[0,M]$, the cost function for a one-shot

least squares minimization is defined by (* denoting complex conjugate):[8]

$$J(\theta) = \sum_{j=0}^{n_1} \sum_{k=0}^{n_1} g_j(\theta)g_k(\theta)r_{jk}, \quad r_{jk} = \sum_{m=0}^{M} \gamma_j(m)\gamma_k^*(m). \tag{55}$$

Consider next the more general class of *inexact* input/output models represented by (47). Modulating (47) with $\phi_{m,n}(t)$, integrating over $[0,T]$ and employing **Property 4**, we obtain the following expression for the equation error relative to this class of *inexact* differential operator models:

$$\tilde{e}(\theta,m) = \sum_{j=0}^{n_1} g_j(\theta)\tilde{\gamma}_j(m) \tag{56}$$

where

$$\tilde{\gamma}_j(m) = \sum_{k=1}^{n_2} W_{jk}(m)\circledast\gamma_{jk}(m) \tag{57}$$

$$\gamma_{jk}(m) = \Delta^n P_{jk}(im\,\omega_0)V_k(m) \tag{58}$$

and the $(V_k(m),W_{jk}(m))$ are Fourier series coefficients of data-related functions defined by

$$V_k(m) = \frac{1}{T}\int_0^T E_k(u(t),y(t))e^{-im\,\omega_0 t}\,dt \tag{59}$$

$$W_{jk}(m) = \frac{1}{T}\int_0^T F_{jk}(u(t),y(t))e^{-im\,\omega_0 t}\,dt. \tag{60}$$

The least squares cost function for a one-shot estimate in this case is defined by, c.f. (55),

$$J(\theta) = \sum_{j=0}^{n_1} \sum_{k=0}^{n_1} g_j(\theta)g_k(\theta)\tilde{r}_{jk}, \quad \tilde{r}_{jk} = \sum_{m=0}^{M} \tilde{\gamma}_j(m)\tilde{\gamma}_k^*(m). \tag{61}$$

To illustrate the above notation for the *inexact* model (49b), the equation error for this linear regression problem is given by

$$\tilde{e}(\theta,m) = \tilde{\gamma}_0(m) + \theta_1\tilde{\gamma}_1(m) + \theta_2\tilde{\gamma}_2(m)$$

where the regressand sequence $\tilde{\gamma}_0(m)$ is defined by

$$\tilde{\gamma}_0(m) = \Delta^2(im\,\omega_0)^2 Y(m) \tag{62}$$

and the two regressor sequences $(\tilde{\gamma}_1(m),\tilde{\gamma}_2(m))$ by

$$\tilde{\gamma}_1(m) = \tfrac{1}{2}\Delta^2(im\,\omega_0)^2 Y_2(m) - Y(m)\circledast\Delta^2(im\,\omega_0)^2 Y(m), \quad \tilde{\gamma}_2(m) = -\Delta^2 U(m) \tag{63}$$

where $Y_2(m)$ is defined by

---

[8] Noting that M in (55) actually corresponds to the frequency $M\,\omega_0$ suggests that in choosing the pair (M,ω_0), their product should correspond roughly to the system bandwidth.

$$Y_2(m) = \frac{1}{T} \int_0^T y^2(t) e^{-im \,\omega_0 t} \, dt.$$

5.3.3. Implementation and Examples

The flow of calculations involved in setting up a least squares estimate for implementing the above equations is shown in Fig. 12. Assuming an accurate computation of the data-related functions $E_k(u(t),y(t))$ and $F_{jk}(u(t),y(t))$, the Fourier series coefficients in (59) and (60) can be carried out efficiently using DFT/FFT techniques as discussed in Section 2.1. Notice that for a finite M, only a finite number of such coefficients need be computed in the case of *exact* differential operator models, but a larger number is needed for *inexact* models in order to implement the linear convolutions in (57), the actual number of which depends on the degree of smoothness in the functions $(E_k(u(t),y(t)), F_{jk}(u(t),y(t)))$ on $[0,T]$.
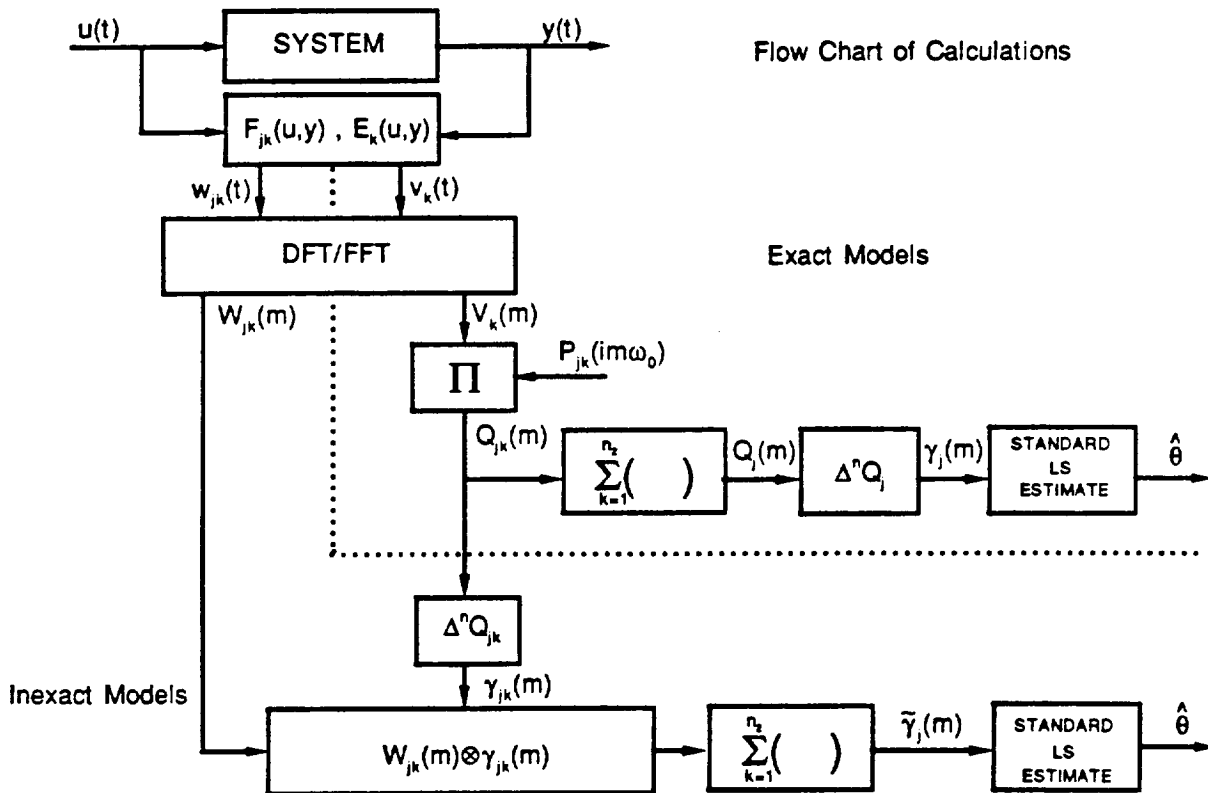

Fig. 12. An explicit least squares estimation for nonlinear systems

Example 1. A simulation study was carried out for the *inexact* model (49) using data generated by the input signal $u(t)=\sin t^2/5$ over a $T=10s$ time interval with zero initial conditions. Thus, the resolving frequency was $1/T=0.1$ Hz. The sampling rate was 25.6 Hz, and Simpson's rule (together with a 256 point DFT) was used to calculate the Fourier series coefficients of the data-related functions: $(u(t),y(t),y^2(t))$, $0\leq t\leq 10$. Zero mean white Gaussian noise of varying degrees of intensity was added to the output signal with 50 Monte Carlo runs made at each noise level. Since the parameters enter into the model linearly, the least squares problem is a linear regression once the regressand and regressor sequences are calculated from (62)-(63). With only 2 unknown parameters and noting that these sequences yield 2 values for each m, i.e., the real and imaginary parts, the sequences in (62)-(63) were computed for $m=0,1,2$ yielding 6 real equations upon which to base the least squares estimate corresponding to each input/output data pair. The mean and standard deviations of the parameter estimates for the 50 Monte Carlo runs at each of 5 noise levels are plotted in Fig. 13. These results demonstrate that the estimator gave values quite close to the true parameters on-average, but the variance for the dissipative coefficient θ_1 was significantly larger than that of the θ_2 parameter.

Figures 14 and 15 contain some additional information relating to the simulations of Example 1. Figure 14 shows output data for one run at the 12.8% noise level together with the simulated model output using the mean values of the (θ_1,θ_2) parameters estimated at this noise level for the 50 Monte Carlo runs. This shows that on-average the bias in the output signal is quite small in relation to the system output signal, which is to be expected in view of Fig. 13. Figure 15 is meant to indicate the ability of the model to predict the system output under ideal no-noise conditions, using the parameters estimated from the noisy data at the various noise levels. The figure-of-merit used to illustrate this is the same as defined in (36).

Example 2. A second simulation study was carried out for the *exact* model (51). Some conditions were the same as for Example 1, i.e., $u(t)=\sin t^2/5$, $0\leq t\leq 10s$, but the initial conditions were taken as $y(0)=\dot{y}(0)=2$, which resulted in output data well above the unity value corresponding to which difficulties with the functions $\ln y(t)$ and $u(t)/y(t)$ could be expected in view of the model (51b). Although the parameters enter nonlinearly into this model, there is basically a one-to-one map between the θ and $g(\theta)=\tilde{\theta}$ parameters so that this problem is also a linear regression in the $\tilde{\theta}$ parameters. As in Example 1, zero mean white Gaussian noise of varying degrees of intensity was added to the output signal with 50 Monte Carlo runs made at each noise level. The mean and standard deviations of the estimated values are tabulated in Tables 4 and 5 for the $\tilde{\theta}$ and θ parameters, respectively. The nonlinear mapping from $\tilde{\theta}$ back into the θ parameters accounts for the greater variability in estimating the model parameters. In spite of this greater variability, the average of the noise-free model responses using these parameter estimates exhibit reasonably good performance as shown in Fig. 16. Thus, if $\hat{\theta}(i)$, $i=1,2\cdots 50$, represent the estimated θ parameter vectors obtained from the 50 Monte Carlo runs at a particular noise level, and $y(t,\hat{\theta})$ the noise-free model output corresponding to any particular $\hat{\theta}$, then a time response in this figure was obtained from the ensemble average: $1/50\sum_{i=1}^{50}y(t,\hat{\theta}(i))$. A measure of the variance in these time responses is summarized in Fig. 17 which gives the standard deviations at each instant of time of the model responses $y(t,\hat{\theta}(i))$ over the ensemble for $i=1,2\cdots 50$.
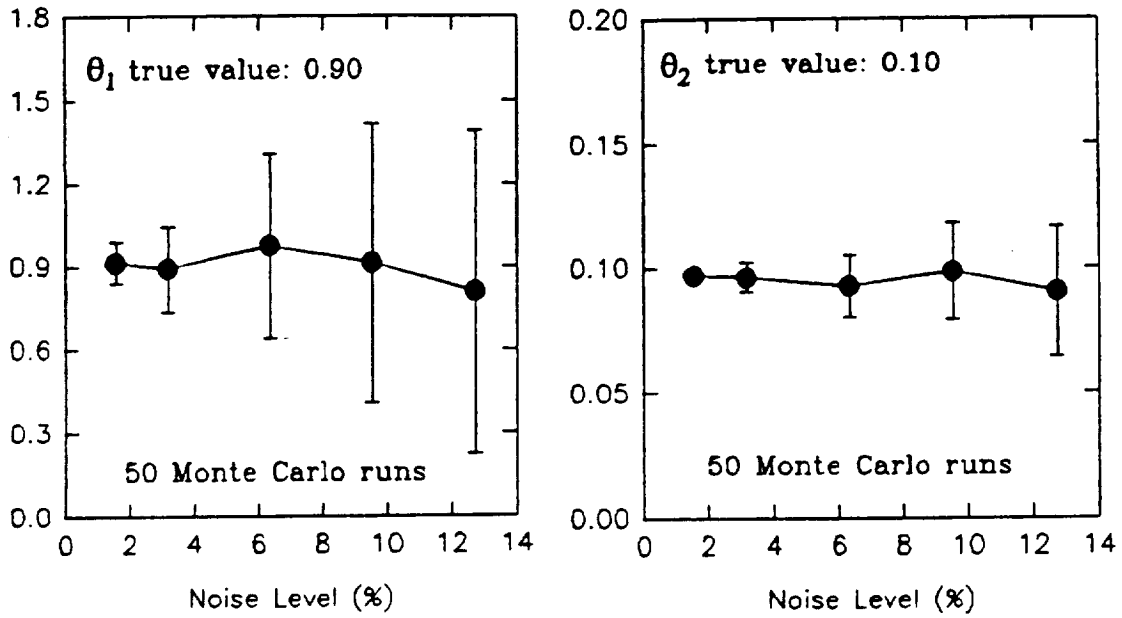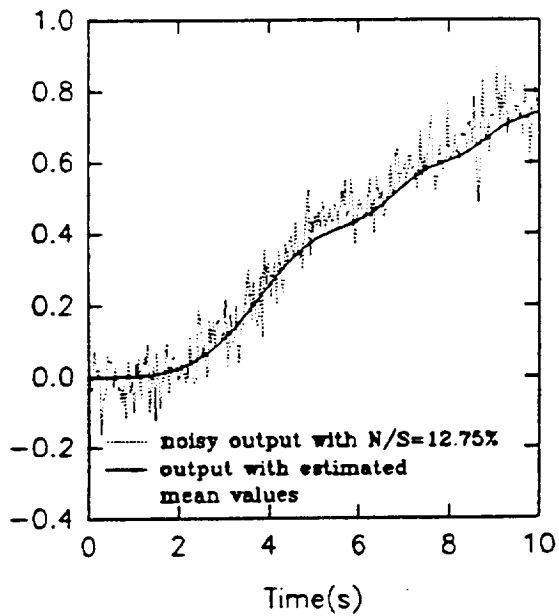
Example 1 Simulations

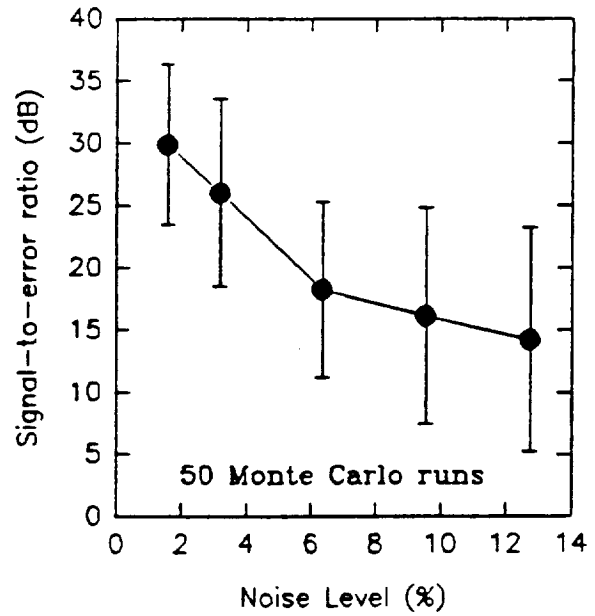Fig. 13. Parameter estimates

Fig. 14. Output data and model response

Fig. 15. S/E statistics

Example 2 Simulations

	$\theta_1\theta_2\theta_3$	$\theta_1\theta_2$	θ_4	$\theta_4(\theta_1+\theta_2)$		
true values	0.246	1.170	2.000	3.02	noise level	S/E (dB)
mean	0.247	1.164	1.974	2.988	$\|e\|_2/\|y\|_2=$	46.97
std	0.0776	0.1100	0.2477	0.1951	0.16%	9.100
mean	0.227	1.115	1.941	2.925	$\|e\|_2/\|y\|_2=$	29.29
std	0.3734	0.5309	1.1869	0.9314	0.79%	11.076
mean	0.193	0.918	1.509	2.335	$\|e\|_2/\|y\|_2=$	19.36
std	0.6291	0.8667	2.0131	1.6039	1.59%	10.164
mean	0.369	0.8831	-0.033	0.695	$\|e\|_2/\|y\|_2=$	13.03
std	0.8526	1.2365	2.5224	1.9130	3.21%	5.725

Table 4. Parameter estimates for $\tilde{\theta}$

	θ_1	θ_2	θ_3	θ_4		
true values	0.21	1.30	0.90	2.00	noise level	S/E (dB)
mean	0.207	1.319	0.881	1.974	$\|e\|_2/\|y\|_2=$	46.97
std	0.0506	0.0488	0.0575	0.2477	0.16%	9.100
mean	-0.058	2.195	0.538	1.789	$\|e\|_2/\|y\|_2=$	29.29
std	1.7710	2.8753	0.3944	1.1045	0.79%	11.076
mean	-0.041	1.159	-0.020	1.509	$\|e\|_2/\|y\|_2=$	19.36
std	2.0996	2.6653	3.6776	2.0131	1.59%	10.164
mean	1.349	-0.076	-14.424	-0.033	$\|e\|_2/\|y\|_2=$	13.03
std	3.3518	4.2351	38.3601	2.5224	3.21%	5.725

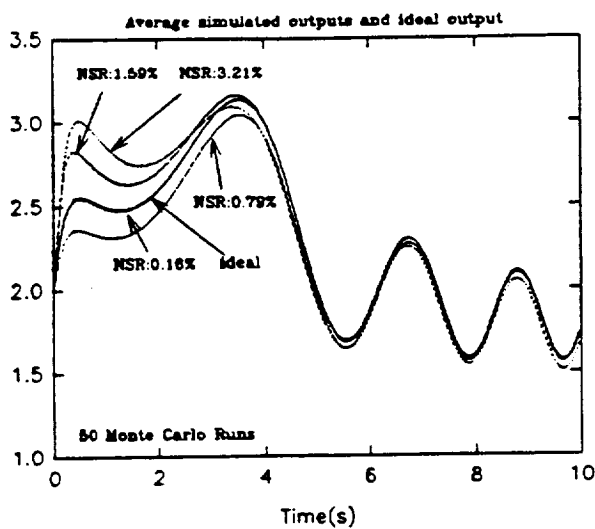Table 5. Parameter estimates for θ
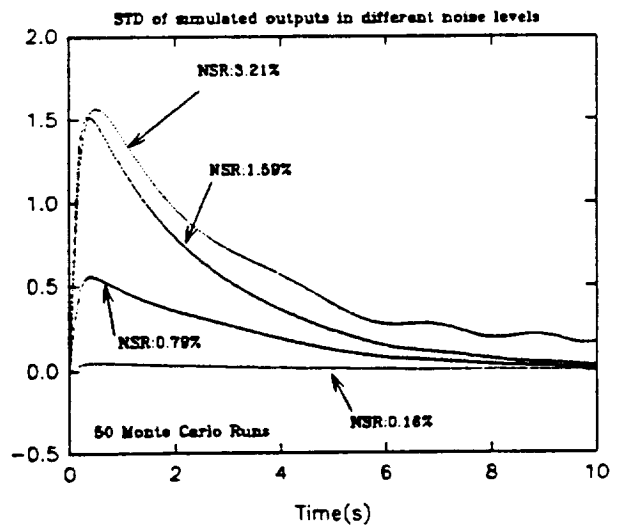

Fig. 16. Model responses

Fig. 17. Variance in model responses

6. Concluding Remarks

The adaptive weighted least squares algorithm AWLS/MFT has been developed to estimate the parameters of multivariable linear differential equation models. In addition to estimating the parameters in a least squares setting, it is aimed at finding the best weighting matrix assuming white measurement noises. As such, it can be regarded as an approximation to a maximum likelihood or Gauss-Markov estimate. However, unlike other time-domain based approaches, it is frequency-domain based - utilizing only the lowest harmonics in the discrete Fourier transform of the data and thereby automatically eliminating high frequency noise if the resolving frequency is sufficiently small, i.e., the data interval $[0,T]$ sufficiently long. Based upon the comparisons made thus far in simulations and in utilizing real data, the algorithm appears to compare favorably against some well established time-domain based identification techniques like the prediction-error method and an output-error/maximum likelihood method used for lateral and longitudinal dynamic modeling in aircraft. Applications to aircraft modeling with real data has not yet been attempted for the frequency transfer function estimation problem, nor for time-varying and nonlinear input/output models, but the extensions discussed in Section 5 suggest that the potential exists for such applications.

7. Acknowledgments

The following were graduate student advisees of the author who aided in the development of the MFT algorithms reported here: A. Fullerton, J. Q. Pan and Y. Shen. Their contribution is gratefully acknowledged.

8. References

Fullerton, A. (1991). Noise and collinearity in differential systems identification. Sc.M. Thesis, Div. of Engr., Brown Univ., Providence, RI.

Johansson, R. (1993). *System Modeling and Identification*. Prentice-Hall.

Klein, V. (1989). Estimation of aircraft aerodynamic parameters from flight data. In *Prog. Aerospace Sci.*, **26**, 1-77.

Klein, V. (1993). Application of system identification to high performance aircraft. In *Proc. of 32nd IEEE CDC*, San Antonio, TX, pp. 2253-2259.

Ljung, L. (1987). *System Identification: Theory for the User*. Prentice-Hall.

Ljung, L. (1991). *System Identification Toolbox*, Version 2.11. The MathWorks, Inc.

Nijmeijer, N. and A. J. van der Schaft (1990). *Nonlinear Dynamical Control Systems*. Springer-Verlag.

Pan, J. Q. (1992). System identification, model reduction and deconvolution filtering using Fourier based modulating signals and high order statistics. Ph.D. Thesis, Brown University, Providence, RI.

Pearson, A. E. and F. C. Lee (1985a). Parameter identification of linear differential systems via Fourier based modulating functions. *Control-Theory and Adv. Tech.*, 1, 239-266.

Pearson, A. E. and F. C. Lee (1985b). On the identification of polynomial input-output differential systems. *IEEE Trans. on Auto. Control*, AC-30, 778-782.

Pearson, A. E. (1988). Least squares parameter identification of nonlinear differential i/o models. In *Proc. of 27th IEEE Conf. on Decis. and Control*, Austin, TX, pp. 1831-1835.

Pearson, A. E. (1989). Identifiability and well-posedness in nonlinear system i/o modeling. In *Proc. of 28th IEEE Conf. on Decis. and Control*, Tampa, FL, pp. 624-625.

Pearson, A. E. (1992). Explicit parameter identification for a class of nonlinear input/output differential operator models. In *Proc. of 31st IEEE Conf. on Decis. and Control*, Tucson, AZ, pp. 3656-3660.

Pearson, A. E. (1993a). Explicit least squares system parameter identification for exact differential input/output models. Proc. of the Eighth ICMCM, X. J. R. Avula, ed. In *Math. Modelling and Sci. Computing*, **2**, pp. 101-107.

Pearson, A. E., Y. Shen and J. Q. Pan (1993b). Discrete frequency formats for linear differential system identification. In *Proc. of 12th World Congress IFAC*, Sydney, Australia, **VII**, 143-148.

Pearson, A. E. and Y. Shen (1993c). Weighted least squares/MFT algorithms for linear differential system identification. In *Proc. of 32nd IEEE CDC*, San Antonio, TX, Dec. 1993, pp. 2032-2037.

Pearson, A. E., Y. Shen and V. Klein (1994). Application of Fourier modulating functions to parameter estimation of a multivariable linear differential system. In *Proc. of IFAC SYSID'94*, Copenhagen, Denmark, vol. 3, pp. 49-54.

Saha, D. C., V. N. Bapat and B. K. Roy (1991). The Poisson moment functional technique - some new results. In Sinha, N. K. and G. P. Rao (Eds.), *Identification of Continuous-Time Systems*, Kluwer Acad. Pub., pp. 327-362.

Shen, Y. (1993). System identification and model reduction using modulating function techniques. Ph.D. Thesis, Brown University, Providence, RI.

Shinbrot, M. (1954). On the analysis of linear and nonlinear systems from transient response data. NACA Report No. TN 3288.

Shinbrot, M. (1957). On the analysis of linear and nonlinear systems. *Trans. ASME*, **79**, 547-552.

Sinha, N. K. and G. P. Rao, eds. (1991). *Identification of Continuous-Time Systems*. Kluwer Acad. Pub.

Skantze, F. P., A. E. Pearson and V. Klein (1994). Parameter identification for unsteady aerodynamic systems by the modulating function technique. In *Proc. of IFAC SYSID'94*, Copenhagen, Denmark, **2**, pp. 571-576.

Unbehauen, H. and G. P. Rao (1987). *Identification of Continuous Systems*. North-Holland.

Walter, E. (1982). *Identifiability of State Space Models*. Springer-Verlag.

REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY*(Leave blank)*	2. REPORT DATE April 1995	3. REPORT TYPE AND DATES COVERED Contractor Report

4. TITLE AND SUBTITLE
Aerodynamic Parameter Estimation Via Fourier Modulating Function Techniques

5. FUNDING NUMBERS
G NAG1-1065
WU 505-64-52-01

6. AUTHOR(S)
A. E. Pearson

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Brown University
Division of Engineering
Providence, RI 02912

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
National Aeronautics and Space Administration
Langley Research Center
Hampton, VA 23681-0001

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
NASA CR-4654

11. SUPPLEMENTARY NOTES
Langley Technical Monitor: Patrick C. Murphy

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Unclassified - Unlimited
Subject Category 05

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Parameter estimation algorithms are developed in the frequency domain for systems modeled by input/output ordinary differential equations. The approach is based on Shinbrot's method of moment functionals utilizing Fourier based modulating functions. Assuming white measurement noises for linear multivariable system models, an adaptive weighted least squares algorithm is developed which approximates a maximum likelihood estimate and cannot be biased by unknown initial or boundary conditions in the data owing to a special property attending Shinbrot-type modulating functions. Application is made to perturbation equation modeling of the longitudinal and lateral dynamics of a high performance aircraft using flight-test data. Comparative studies are included which demonstrate potential advantages of the algorithm relative to some well established techniques for parameter identification. Deterministic least squares extensions of the approach are made to the frequency transfer function identification problem for linear systems and to the parameter identification problem for a class of nonlinear time-varying differential system models.

14. SUBJECT TERMS
Aircraft system identification; Modulating function technique

15. NUMBER OF PAGES
45

16. PRICE CODE
A03

17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT