METHODS STUDIES ON SYSTEM IDENTIFICATION

FROM TRANSIENT ROTOR TESTS

Part I of Second Yearly Report under Contract NAS2-7613

Prepared for the Ames Directorate, USAAMRDL
at Ames Research Center, Moffett Field, California

by

K. H. Hohenemser

D. Banerjee

S. K. Yin

Department of Mechanical Engineering

Washington University
School of Engineering and Applied Science
St. Louis, Missouri 63130

June 1975

METHODS STUDIES ON SYSTEM IDENTIFICATION

FROM TRANSIENT ROTOR TESTS

Part I of Second Yearly Report under Contract NAS2-7613

Prepared for the Ames Directorate, USAAMRDL
at Ames Research Center, Moffett Field, California
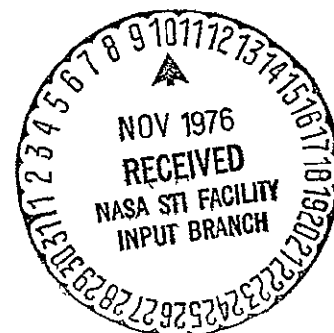
by

K. H. Hohenemser

D. Banerjee

S. K. Yin

Department of Mechanical Engineering

Washington University
School of Engineering and Applied Science
St. Louis, Missouri 63130

June 1975

CHAPTER 1

INTRODUCTION TO AIRCRAFT PARAMETER IDENTIFICATION METHODS

K. H. Hohenemser and D. Banerjee

ABSTRACT

Some of the more important methods are discussed that have been used or proposed for aircraft parameter identification. The methods are classified into two groups: Equation error or regression estimates and Bayesian estimates and their derivatives that are based on probabilistic concepts. In both of these two groups the cost function can be optimized either globally over the entire time span of the transient, or sequentially, leading to the formulation of optimum filters. Identifiability problems and the validation of the estimates are briefly outlined, and applications to lifting rotors are discussed.

TABLE OF CONTENTS

Preface to Second Yearly Report under Contract NAS2-7613

Work under Contract NAS2-7613 started on July 1, 1973 as a continuation of research conducted under Contract NAS2-4151 since February 1, 1967. The research goals stated in Contract NAS2-7613 are

(a) Assess analytically the effects of fuselage motions on stability and random response. The problem is to develop an adequate but not overly complex flight dynamics analytical model and to study the effects of structural and electronic feedback, particularly for hingeless rotors.

(b) Study by computer and hardware experiments the feasibility of adequate perturbation models from non-linear trim conditions. The problem is to extract an adequate linear perturbation model for the purpose of stability and random motion studies. The extraction is to be performed on the basis of transient responses obtained either by computed time histories or by model tests.

(c) Extend the experimental methods to assess rotor wake-blade interactions by using a 4-bladed rotor model with the capability of progressing and regressing blade pitch excitation (cyclic pitch

stirring), by using a 4-bladed rotor model with hub tilt stirring, and by testing rotor models in sinusoidal up or side flow.

Work on research goal (a) has been reported in Part I of the First Yearly Report under subject contract titled "Methods Studies' Toward Simplified Rotor-Body Dynamics". The results were published in the paper: Hohenemser, K. H. and Yin, S. K., "On the Use of First Order Rotor Dynamics in Multiblade Coordinates" presented at the 30th Annual National Forum of the American Helicopter Society, May 1974, Preprint 831.

Initial work on research goal (b) has been reported in Part II of the First Yearly Report under subject contract titled "Computer Experiments in Preparation of System -Identification from Transient Rotor Model Tests".

Initial work on research goal (c) has been reported in Part III of the First Yearly Report under subject contract titled "Experiments with a Four-Bladed Cyclic Pitch Stirring Model Rotor."

The second Yearly Report under Contract NAS2-7613 is subdivided into two parts, whereby Parts I and II are related to the research goals (b) and (c) respectively. The authors and titles of the two parts are:

Part I   Hohenemser, K. H., Banerjee, D. and Yin, S. K., "Methods Studies on System Identification from Transient Rotor Tests."

Part II   Hohenemser, K. H. and Crews, S. T., "Additional Experiments with a Four-Bladed Cyclic Pitch Stirring Model Rotor."

Part I begins with an introduction to aircraft/identification methods, it contains the results of computer experiments with several selected methods of system identification applied to lifting rotors, and contains the development of a new method for optimal data utilization.

Part II presents extended frequency response measurements with the four-bladed model rotor including dynamic wake measurements at zero advance ratio. It describes the modifications of the model and its instrumentation for transient pitch stirring tests, and it discusses the development of software for the data processing.

Scope of Part I of Second Yearly Report

The computer experiments presented in the First Yearly Report under Contract NAS2-7613 assumed a rotor condition with .4 advance ratio. The input transients consisted of rectangular and wave shaped normal flow impulses. Noise polluted blade flapping responses represented the system output. A linear identification scheme was applied that worked well if data preprocessing by a digital filter and by a Kalman filter was adopted. Blade Lock number and collective pitch angle were assumed to be not well known and were identified. During the past year this work was extended in several directions:

1. The normal flow transients were replaced by constant acceleration pitch stirring transients, since the wind tunnel model is particularly suited to accept such transients.

2. For zero advance ratio a two parameter dynamic wake model was adopted that included a time constant. A single blade representation is now no more applicable. A multiblade identification analysis was developed.

3. The previously used identification method was modified and extended to an iterative equation error estimation with updated Kalman filter.

4. The maximum likelihood method was applied to both
   the single blade and the multiblade identification.

5. The maximum likelihood scheme, assuming the absence
   of system noise, was used to obtain optimum data
   utilization, that is to avoid inaccuracies of the
   parameter estimates from insufficient data pro-
   cessing, and to avoid excessive data processing for
   a given desired accuracy of the parameter estimates.

Since parameter identification practice is a still
developing and somewhat controversial discipline, it
was believed of interest to provide a brief introduction
to aircraft parameter identification methods in Chapter
1 of this report. Chapter 2 describes the extensions
of the work with respect to items 1 to 4, while Chapter
3 describes the work performed with respect to item 5.
The 3 chapters are written as independent self suf-
ficient reports with separate abstracts and lists of
references for each.

The application of methods for system parameter
identification from transients is a rapidly growing
field of study. There is a confusing multiplicity of
methods and of names for these methods; and contradic-
tory claims are made as to the efficiency and use-
fulness of the various methods. Widely varying
parameter identification methods are being used at
the various aircraft flight test centers. There also
is a substantial difference between what is theoretically
possible and what is practically desirable to perform
a given job with adequate accuracy and with a reasonable
computer effort. Since a comprehensive review of
aircraft parameter identification methods is not
available in the literature, it was believed useful
to provide a brief introduction to this field in Chapter
I before presenting the applications to lifting rotors in
Chapters II and III. The review of identification methods
is by no means complete. Only the most important methods are
discussed. Only rough outlines are given for the
various methods. Details of the derivations and of the
application algorithms are found in the cited literature.

ELEMENTS OF SYSTEM IDENTIFICATION FROM TRANSIENTS

System identification is the process of extracting
numerical values for system parameters and other sub-
sidiary parameters (process and measurement noise

covariances, bias, initial states, etc.) from the time
history of control or other inputs and of the resulting
system responses. A schematic for the measurements is
shown in Fig. 1. The process of system identification
involves four steps:

1. Selection of a suitable input that insures parti-
   cipation of all important modes of the system in
   the transient response.

2. Selection of sufficiently complete and accurate
   instrumentation to measure the key input and output
   variables.

3. Selection of a mathematical model that adequately
   represents the actual system characteristics.

4. Selection of an efficient criterion function and
   estimation algorithm for the identification of the
   unknown system parameters.

The concept of system identification is illustrated in
Fig. 2. The design input is fed both to the actual
system and to its mathematical model that contains the
unknown parameters. The measured response, polluted by
measurement noise, is compared with the computed
response from the mathematical model. The difference
between these two responses, the response error, is used
in the parameter estimation technique based on the

EXTERNAL DISTURBANCE

EXTERNAL INPUT → PHYSICAL SYSTEM → RESPONSE OR OUTPUT

MEASUREMENT NOISE

MEASURING DEVICES

MEASURED INPUT
OUTPUT
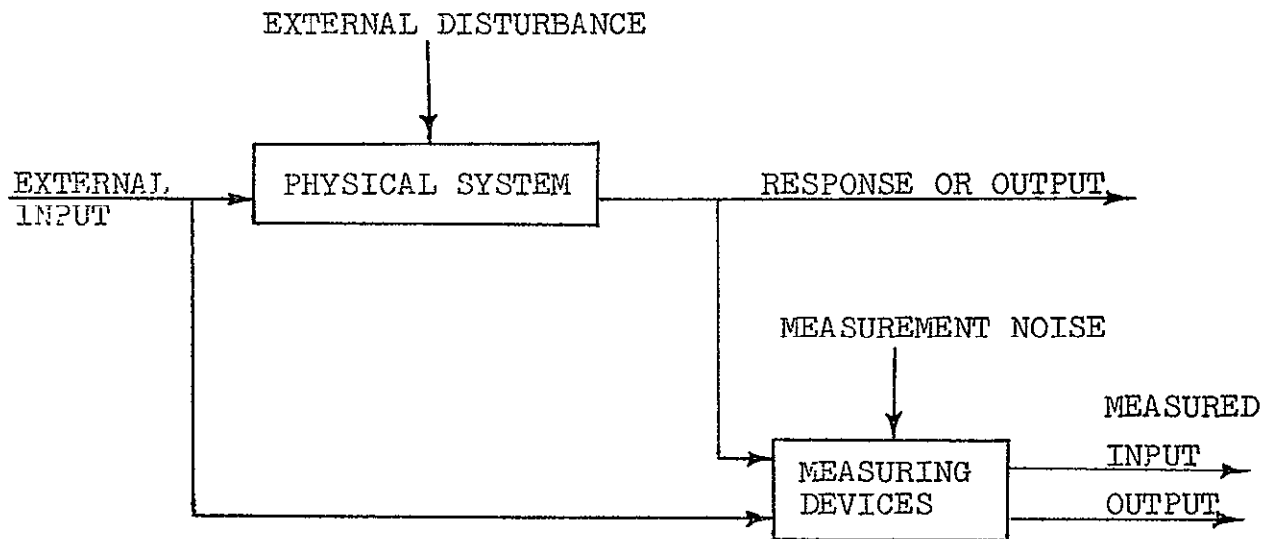
Figure 1   Schematic of Measurements for System Identification

SYSTEM NOISE   MEASUREMENT NOISE

DESIGN INPUT → SYSTEM → SYSTEM RESPONSE → + → MEASURED RESPONSE

+ −

MATHEMATICAL MODEL OF THE SYSTEM   COMPUTED RESPONSE

ESTIMATION ALGORITHM ← CRITERION FUNCTION ← RESPONSE ERROR

A PRIORI INFORMATION
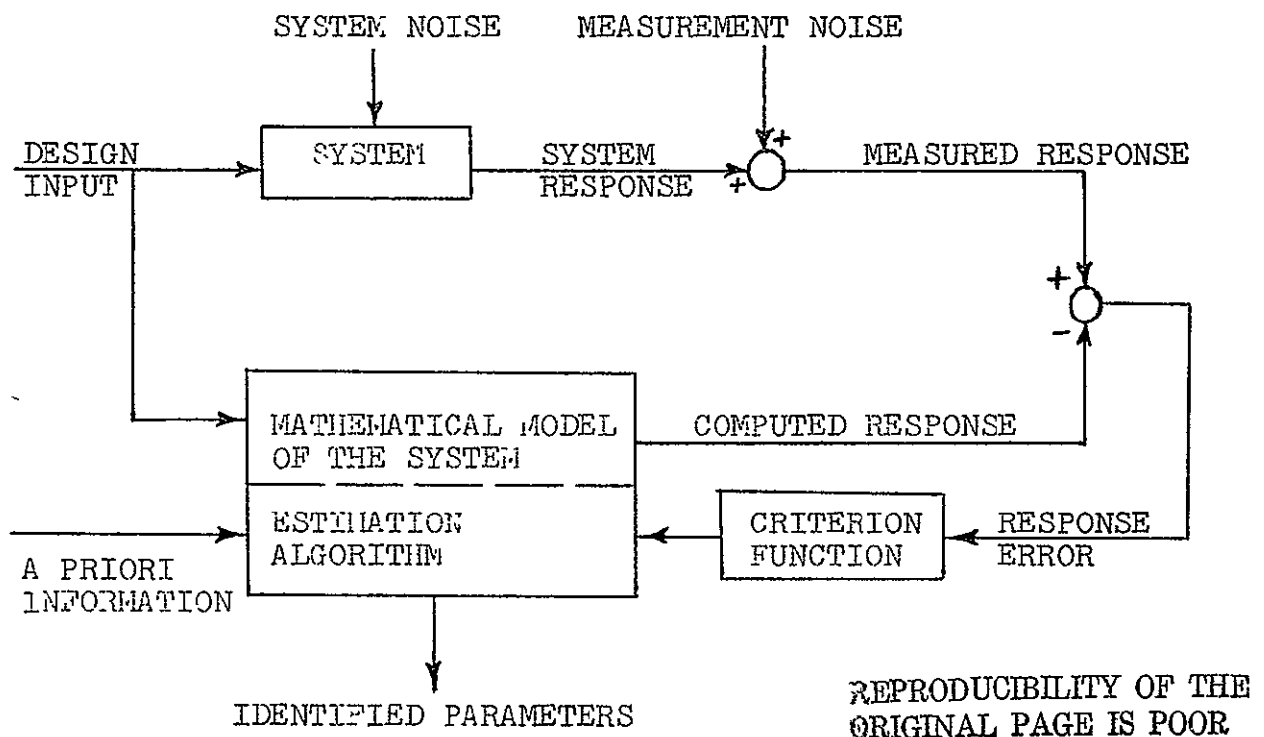
IDENTIFIED PARAMETERS

Figure 2   Illustration of System Identification

criterion function and optimizing technique. The
estimation algorithm may also use apriori information,
e.g., initial statistics of the parameters. Here we
will be mainly concerned with the fourth of the
previously listed steps, that is with the various
estimation algorithms.

The mathematical representation of the system will
be given in the non-linear case by:

System equation $\dot{x}(t) = f(x,u,t) + \Gamma(t)w(t)$ (1)

Initial condition $x(t = 0) = x_0$

Measurement Equation $y(t) = h(x,u,t) + v(t)$ (2)

where $x(t)$ n x 1   state vector

$u(t)$ p x 1   input vector

$w(t)$ q x 1   system noise vector, covariance Q

$y(t)$ r x 1   output or measurement vector

$v(t)$ r x 1   measurement noise vector,
                       covariance R

If the system is linear, equations 1 and 2 reduce to

$\dot{x}(t) = F(t)\, x(t) + G(t)\, u(t) + \Gamma(t)\, w(t)$ (3)

$y(t) = H(t)\, x(t) + D(t)\, u(t) + g + v(t)$ (4)

where g is the vector of bias errors.

CLASSIFICATION OF IDENTIFICATION ALGORITHMS

The various estimation algorithms can be classified into two groups presented in Table 1. The first group listed in Table 1 above the double line is based on statistical regression and does not admit a probabilistic interpretation. The algorithms listed in Table 1 below the double line are based on probabilistic interpretation. In the equation error estimate no measurement noise is modeled, the following 4 methods include both measurement and system noise, while in the output error estimate no system noise is modeled. The various algorithms listed in Table 1 will be discussed in the following sections.

EQUATION ERROR ESTIMATES

Equation error methods assume a performance criterion that minimizes the square of the equation error (process noise). They are least squares techniques and they require the knowledge of all response variables (states) and their derivatives. In the so called least squares method the unknown parameters are selected such that the integral over the square of the state equation error is minimized, see for example reference 1. With equation 1 we have the error function (the upper integration limit T is the time over which the measurements are taken).

Classification of Estimate	Criterion for Estimate	Solution for Estimate
Equation Error Estimate (no measurement noise is modeled)	$\hat{\Theta}$ Minimizes equation error squares and/or integrated equation error squares	Closed form global or sequential optimization by optimum filter (Method function $\delta(t-t_i)$ or e^{-st}).
Bayesian Estimate	$\hat{\Theta} = E(\Theta/Z)$	—
Quasi-Baysian Estimate with augmented state	$\hat{\Theta}, \hat{x}$ such that $\underset{\Theta,x}{\text{MAX}}\ f(\Theta, x/Z) = f(\hat{\Theta}, \hat{x}/Z)$	Sequential optimization by optimum non-linear filter (extended Kalman filter with or without local iteration and/or smoothing).
Quasi-Bayesian Estimate	$\hat{\Theta}$ such that $\underset{\Theta}{\text{MAX}}\ f(\Theta/Z) = f(\hat{\Theta}/Z)$	Iterative global optimization with state Kalman filter equations as constraints.
Maximum Likelihood Estimate	$\hat{\Theta}$ such that $\underset{\Theta}{\text{MAX}}\ f(Z/\Theta) = f(Z/\hat{\Theta})$	
Output Error Estimate (no system noise is modeled)	$\hat{\Theta}$ minimizes output error squares.	Iterative global optimization with system equations as constraints.

Table 1   Classification of Estimation Algorithms.

$$J = \int_0^T [\overset{\circ}{x} - f(x,u,\theta,t)]^T \, W[\overset{\bullet}{x}-f(x,u,\theta,t)]dt \qquad (5)$$

where θ is the m x 1 vector of unknown parameters. W is a positive definite weighting matrix. An appropriate choice for W would be Q^{-1} where Q is the covariance of the process noise. For the usual digital data processing, the variables $\overset{\circ}{x}$, x, u·are sampled, and only available at discrete time points t_i. Mathematically the sampling process can be expressed by multiplying the system equation with the delta function $\delta(t - t_i)$. The integral of equation 5 then becomes a sum. One can use instead of the delta function also a different "method function", for example e^{-st}, that would allow taking the Laplace transforms.

If the system is linear in the unknown parameters θ, the system equation can be written in the form

$$\overset{\bullet}{x}(t) = F(x,u,t)\theta + \Gamma(t) \, w(t) \qquad (6)$$

and the performance criterion (5) becomes

$$J = \int_0^T [\overset{\circ}{x} - F(x,u,t)\theta]^T \, W[\overset{\bullet}{x} - F(x,u,t)\theta]dt \qquad (7)$$

Since the function inside the integral has continuous derivatives with respect to θ we set

$$\partial J/\partial \theta = 0 \qquad (8)$$

thus resulting in the closed form solution

$$\hat{\Theta} = \left[\int_0^T F^T(x,u,t)W\ F(x,u,t)dt \right]^{-1} \int_0^T F^T(x,u,t)W\ \dot{x}(t)dt$$

(9)

The first factor is the covariance matrix of the estimate. If the system is non-linear in the unknown parameters, the solution equation 9 can be replaced by an interative solution where $F(x,u,t)$ is substituted by $\partial f(x,u,\Theta_k,t)/\partial\Theta$ and $\hat{\Theta}$ on the left hand side is replaced by $\hat{\Theta}_{k+1}$. It can be shown (for example reference 1) that the parameters in the nth row of $f(x,u,\Theta,t)$ are independent of all the elements of $\dot{x}(t)$ except $\dot{x}_n(t)$. This independence is one of the drawbacks of the least squares method, in that only one of the measured state derivatives is used in— determining a given row of the $f(x,u,\Theta,t)$ matrix. If one of the signals has not been measured, the least squares method does not provide an estimate of the parameters related to that signal. This independence also illustrates the fact that the estimate of one row of the $f(x,u,\Theta,t)$ matrix is obtained independently of the other rows, and no "trade-off" can be made between elements in different rows to improve the estimate.

For some applications it is practical to include the state vectors in the error minimization. In the

modified least squares method a combination of the
standard least squares with the integrated least
squares is used. The parameters obtained by this
method not only trace the derivative of the state
but also the state itself over the selected time
interval. The performance criterion now includes
in addition to the equation error also the integrated
equation error:

$$
J = \int_0^T \left|\left| \dot{x}(t) - F(x,u,t)\theta + \int_0^t \dot{x}(\tau)d\tau - \int_0^t F(x,u,\tau)\theta \; d\tau \right|\right|_W dt \qquad (10)
$$

where W is a positive definite weighting matrix and where

$$
||A||_W \; \underline{\underline{\Delta}} \; A^T \; W \; A \qquad (11)
$$

Minimizing the expression, equation 10 results in the
estimate

$$
\hat{\theta} = \left[\int_0^T \{F(x,u,t) + \int_0^t F(x,u,\tau) \; d\tau\}^T \; W\{F(x,u,\tau) + \int_0^t F(x,u,\tau) \; d\tau\} \; dt \right]^{-1} \qquad (12)
$$

$$
\left[\int_0^T \{F(x,u,t) + \int_0^t F(x,u,\tau)d\tau\}^T W\{\dot{x}(t) + \int_0^t \dot{x}(\tau) \; d\tau\}dt \right]
$$

This method has the same row independence of $f(x,u,\theta,t)$ as the standard least squares method.

Since these methods do not allow for measurement errors, they result in biased estimates when this type of error does exist. When measurement errors are small, as is increasingly the case in modern instrumentation, the equation error method is preferrable over other methods because of its simplicity. It is widely used also when measurement errors are substantial and then serves as start-up technique for the output error and other iterative methods.

In many applications, measurements of some of the responses or their derivatives are not available. If the response, but not the rate of response is measured, it is tempting to differentiate the measured response. However, the differentiation of measured data introduces additional uncertainty so that this technique is usually inaccurate. If e^{-st} is used as a methods function, Laplace transforms can be used. The estimation then reduces to an algebraic manipulation of the data that avoids their differentiation. The Laplace transform technique as a substitute of differentiating measurement data is discussed in reference 2.

BAYESIAN AND QUASI-BAYESIAN PARAMETER ESTIMATES

In the preceding methods we specified a cost criterion J that represented the "loss" resulting from an incorrect estimation of the unknown system parameters. The parameters were then selected in such a way as to minimize the loss. If a priori probabilities exist not only for the measurement errors but also for the unknown parameter vector Θ then one can define an expected loss and select the parameter vector in such a way as to minimize this expected loss. Such an estimate is called a Bayesian estimate, see for example reference 3.

The form of a Bayesian estimate depends on the form of both the loss function and of the a priori distribution of the measurement and parameter noise. For the particular case of positive semi-definite quadratic loss functions the Bayesian estimate is the mean of Θ conditioned on the observations. This is true regardless of the distribution of measurement and parameter noise. It has also been shown that for the case of unimodal symmetric a posteriori distribution of the parameters given the observations, the Bayesian estimate is the conditional mean for a considerably wider class of loss functions. For these reasons the Bayesian estimate can be defined as the conditional mean of the parameter distribution.

In order to compute the conditional mean it is first necessary to determine the conditional probability density for Θ. This density can be written from Bayes rule as (Z is the set of all observations)

$$f(\Theta/Z) = f(Z/\Theta) \, f(\Theta)/f(Z) \tag{13}$$

The denominator is a normalizing factor determined from

$$f(Z) = \int\limits_{\text{all } \Theta} f(Z/\Theta) \, f(\Theta) \, d\Theta \tag{14}$$

The optimal Bayesian estimate is now given by

$$\hat{\Theta} = (1/f(Z)) \int\limits_{\text{all } \Theta} \Theta \, f(Z/\Theta) \, f(\Theta) \, d\Theta \tag{15}$$

In general the evaluation of equations 14 and 15 would require the solution of the system equations for all possible values of the parameter vector Θ. This is a large effort, especially if the dimension of Θ is large. If $f(\Theta/Z)$ is unimodal and symmetric about its mean value, the conditional mean corressponds to the mode. Since $f(Z)$ is merely a scale factor the finding of the mode requires neither the evaluation of the integral in equation 14 nor that in equation 15. The mode $\tilde{\Theta}$ of Θ has the property

$$\max_{\Theta} f(\Theta/Z) \sim \max f(Z/\Theta) \, f(\Theta) = f(Z/\tilde{\Theta}) \, f(\tilde{\Theta}) \tag{16}$$

Even if the a priori density $f(\theta)$ is symmetric it does
not follow that the conditional density $f(\theta/Z)$ is also
symmetric since in general the observations depend
non-linearly on the parameters. Estimation according
to equation 16 is, therefore, called "quasi-Bayesian"
estimation. Another designation used for example in
reference 4 is maximum a posteriori probability (MAP)
parameter estimate. Since the logarithm is a monotonic
function of its argument we can replace equation 16
by maximizing the expression

$$J = \log f(Z/\theta) + \log f(\theta) \qquad (17)$$

If no a priori information about the parameters θ
is available, that is, if the a priori density is uniform,
$f(\theta) = $ constant, the quasi-Bayesian estimate reduces to
the "maximum likelihood" estimate which involves
finding the maximum of $f(Z/\theta)$.

ESTIMATES ASSUMING GAUSSIAN DISTRIBUTIONS

The evaluation of equation 17 becomes particularly
convenient if we assume Gaussian densities for the
parameters, for the observations and for the system
states. In linear systems and linear measurement
equations (equations 3 and 4) one needs only to assume
that the system noise $w(t)$ and the measurement noise $v(t)$
is Gaussian. It then follows that states $x(t)$ and

observations y(t) are also Gaussian. For non-linear systems with Gaussian noise f(Z/θ) tends to a Gaussian density as the sampling rate is increased (see for example reference 5 p. 29). The assumption of Gaussian densities for all variables is, therefore, a reasonable one. Since θ is a m x 1 vector we now have the a priori density

$$f(\theta) = \left|P_\theta\right|^{-1/2} (2\pi)^{-m/2} \exp - (1/2)(\theta-\bar\theta)^T P_\theta^{-1} (\theta-\bar\theta) \qquad (18)$$

P_θ is the parameter covariance matrix, $\bar\theta$ is the parameter mean. Except for a constant additive term, log f(θ) is now given by

$$\log f(\theta) = (-1/2) (\theta-\bar\theta)^T P_\theta^{-1} (\theta-\bar\theta) \qquad (19)$$

In order to obtain an expression for log f(Z/θ) in equation (17), we assume that Z consists of N consecutive observations y(1) .. y(N).

$$Z = Y_N = \{y(1), \ldots y(N)\} \qquad (20)$$

With successive application of Bayes rule we obtain

$$f(Y_N/\theta) = f(y(1), \ldots, y(N)/\theta) = F(y(N)/Y_{N-1},\theta)\, f(Y_{N-1}/\theta)$$

$$\simeq$$

$$\cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot$$

$$= \prod_{j=1}^{N} f(y(j)/Y_{j-1},\theta) \qquad (21)$$

Taking the logarithm we have

$$\log f(Y_N/\theta) = \sum_{j=1}^{N} \log f(y(j)/Y_{j-1}, \theta) \tag{22}$$

$\left(y(j)/Y_{j-1}, \theta\right)$ is the observation estimate at time j given all preceding observations and given the parameters. We denote the observation by $y(j)$ and its expected value and covariance respectively by $\hat{y}(j/j-1)$ and $B(j/j-1)$. We further denote the "innovation" by

$$y(j) - \hat{y}(j/j-1) = v(j) \tag{23}$$

Since $y(j)$ is a $r \times 1$ observation vector, its Gaussian density is

$$f(y(j)) = |B(j/j-1)|^{-1/2} (2\pi)^{-r/2} \exp\left\{-(1/2)v^T(j)\right.$$
$$\left. B^{-1}(j/j-1)v(j)\right\} \tag{24}$$

Taking the logarithm of equation 24 , summing according to equation 22 , inserting in equation 17 and inverting the sign we have now to minimize the expression (see also equation 19)

$$\sum_{j=1}^{N} \{v^T(j)B^{-1}(j/j-1)v(j) + \log|B(j/j-1)|\}$$

$$+ (\theta-\bar{\theta})^T P_\theta^{-1} (\theta-\bar{\theta}) \tag{25}$$

If no a priori information is available before taking observations, the last term in the expression 25 is constant and we then have the criterion for the maximum likelihood estimation. Bayesian or quasi-Bayesian estimation is rarely used since a priori densities for the parameters are in most applications not available.

MAXIMUM LIKELIHOOD PARAMETER ESTIMATION

According to expression 25 maximum likelihood estimation is equivalent to minimizing the so-called likelihood function

$$J(\theta) = \sum_{j=1}^{N} [\nu^T(j)B^{-1}(j/j-1)\nu(j) + \log |B(j/j-1)|] \qquad (26)$$

In the presence of system noise the minimization of the expression 26 is very difficult. When going from time j-1 to time j one first has to solve the prediction equations for the estimate of the state and for its covariance. Assuming the linear system equation 3 with zero mean Gaussian system noise w(t) the prediction is given by

$$\overset{\bullet}{x}(j/j-1) = F \hat{x}(j/j-1) + G u(t) , (j-1) \leq t \leq j \qquad (27)$$

$$\overset{\bullet}{P}(j/j-1) = F P(j/j-1) + P(j/j-1)F^T + \Gamma Q \Gamma^T \qquad (28)$$

where Q is the system noise covariance and P the state
covariance. These equations use the estimated state
and its covariance at time j-1: $\hat{x}(j-1/j-1)$ and
$P(j-1/j-1)$, to predict the state and its covariance at
time j: $\hat{x}(j/j-1)$ and $P(j/j-1)$. This is the prediction
before we know the result of the observations at time
j. After the observations $y(j)$ have been made the
optimum estimate is given by the Kalman filter
equations for the state and for its covariance:

$$\hat{x}(j/j) = \hat{x}(j/j-1) + K(j) \ (y(j) - H \ \hat{x}(j/j-1)) \qquad (29)$$

$$P(j/j) = (I - K(j) \ H) \ P(j/j-1) \qquad (30)$$

with the filter gain

$$K(j) = P(j/j-1) \ H^T \ (H \ P(j/j-1) \ H^T + R)^{-1} \qquad (31)$$

where R is the measurement noise covariance. The
covariance of the observations $B(j/j-1)$ that occurs in
the cost functional 26 is given in terms of the state
covariance before observations by

$$B(j/j-1) = H \ P(j/j-1) \ H^T + R \qquad (32)$$

Thus the terms in the expression 26 that is to be
minimized require the solution of the prediction equations
27 and 28 for each time interval and of the up-date

equations 29 and 30 at each sampling time together with

the solution of the measurement equation 4. Reference

5 gives an algorithm for the solution of the problem.

However, due to its complexity this algorithm has not

as yet been applied to a practical problem of aircraft

parameter estimation, see for example reference 6.

The problem of minimizing the expression 26 is

greatly simplified if the observation covariance

$B(j/j-1)$ can be assumed constant. This is for example

true for zero system noise, when according to

equation 32 $B(j/j-1) = R$. The problem then reduces

to minimizing the cost function

$$\underline{J(\theta)} = \sum_{j=1}^{N} \nu^{T}(j) \ R^{-1} \ \nu(j) \qquad (33)$$

where $\nu(j)$ is given by the innovation term 23. Since

equation 33 represents (according to equations 2 and 4)

the sum of the measurement error squares, the estimation

with equation 33 is also called output error method of

estimation. There are several algorithms available to

perform the optimization of $J(\theta)$ from equation 33. The

most widely used is the modified Newton-Raphson or

quasiliniearization method. It has the advantage that

the sensitivity or information matrix is obtained as a

byproduct. The inverted information matrix gives the

Cramér-Rao lower bound for the parameter covariance.
This lower bound is found in many applications to be a
more meaningful measure of the accuracy of the parameter
estimate than the parameter covariance obtained from
the equation error method (first factor in equation 9).

OUTPUT ERROR METHOD USING QUASILINEARIZATION

We use an iterative method beginning with an
initial parameter estimate $\hat{\theta} = \theta_0$. The problem is to
find a zero of the gradient of the cost function 33,
$\partial J/\partial \theta = 0$. Consider a two-term Taylor series
expansion of $\partial J/\partial \theta$ about the kth value of θ

$$(\partial J/\partial \theta)_{k+1} \simeq (\partial J/\partial \theta)_k + (\partial^2 J/\partial \theta^2)_k \, \Delta \theta_{k+1} \qquad (34)$$

where
$$\Delta \theta_{k+1} \overset{\Delta}{=} \theta_{k+1} - \theta_k$$

$(\partial^2 J/\partial \theta^2)_k$ is the second gradient of the cost function
with respect to θ at the kth iteration. If equation 34
is a sufficiently close approximation, the change in θ
for the $(k + 1)$th iteration to make $(\partial J/\partial \theta)_{k+1}$ approximately
zero is

$$\Delta \theta_{k+1} = - [(\partial^2 J/\partial \theta^2)_k]^{-1} (\partial J/\partial \theta)_k \qquad (35)$$

Using now for $v(j)$ the two term expansion

$$v(j)_k \simeq v(j)_{k-1} + \frac{\partial}{\partial \theta} (v(j))_k \, \Delta \theta_k \qquad (36)$$

one obtains for the first and second gradients of the
cost function

$$(\partial J/\partial \Theta)_k = 2 \sum_{j=1}^{N} \frac{\partial}{\partial \Theta} [(\nu(j)]_k^T R^{-1} [\nu(j)]_k \qquad (37)$$

$$M = (\partial^2 J/\partial \Theta^2)_k = 2 \sum_{j=1}^{N} \frac{\partial}{\partial \Theta} [\nu(j)]_k^T R^{-1} \frac{\partial}{\partial \Theta} [\nu(j)]_k \qquad (38)$$

We thus need solutions for $\nu(j)_k$ and $\frac{\partial}{\partial \Theta} \nu(j)_k$. For
this purpose we first solve the system and measurement
equations

$$\dot{\hat{x}} = f(\hat{x}, u, t)$$
$$\hat{y} = h(\hat{x}, u, t) \qquad (39)$$

for each iteration whereby the initial conditions are
either obtained from the measurements or are included
in the unknown parameters Θ. The innovation is now
obtained from equation 23. Next we solve the
"sensitivity equations" for each iteration

$$\partial \dot{\hat{x}}/\partial \Theta_i = \partial f/\partial \Theta_i + \partial f/\partial x \Big|_{x=\hat{x}} \partial \hat{x}/\partial \Theta_i$$

$$\partial \hat{y}/\partial \Theta_i = \partial h/\partial x \Big|_{x=\hat{x}} \partial \hat{x}/\partial \Theta_i \qquad (40)$$

The initial conditions of $\partial\hat{x}/\partial\Theta_i$ are zero except when
x(0) is identified as part of the parameters Θ. In
this case the initial partials have the value one. With
equation 23 we can now compute the first and second
gradient of the cost function, equations 37 and 38,
and then obtain the change in parameters for the next
iteration from equation 35 . This involves the
inversion of the sensitivity matrix M (equation 38),
whereby M^{-1} is the Cramér-Rao lower bound for the
covariance of the parameters.

The method is easily extended to the case with
a priori information on the parameters, equation 25.
The sensitivity matrix 38 is then augmented by the
term $2P_\Theta^{-1}$, and the gradient 37 is augmented by the
term $2P_\Theta^{-1}$ ($\Theta_0 - \Theta_{k-1}$), see reference 2.

PARAMETER ESTIMATION BY FILTERING

The parameter estimation methods discussed so far
can be denoted as "global" methods. The performance
criterion includes the test data for the entire duration
of the transient. Filtering is an important tool in
state and parameter estimation. It can be used either
in conjunction with global estimates, or it can be
used as a direct approach to state and parameter
estimation. An example of the first type of filter
applications is the prefiltering of test data before

using them in a least squares regression estimate,
see for example reference 4. The Graham digital filter
can remove high frequency noise. A Kalman filter
can be used to estimate state variables and their
rates not directly measured. It also removes the
noise in the measurements. The role of the Kalman
filter in maximum likelihood estimation has been
shown in equations 29 to 31, where it is used to
establish the innovation sequence.

In addition to applications in global estimation
methods, filters can also be used as substitutes for
global methods. The advantage of such direct filter
methods is a reduction in computer effort particularly
in cases with a large number of parameters. The dis-
advantage is that unlike the inverted information
matrix of the maximum likelihood method that provides a
lower bound on the parameter covariances, no physically
meaningful parameter covariances are obtained with the
direct filter methods. The covariance propagation
equations require initial values that are usually
impossible to obtain in any rational way. Though
improvements of the filter solution (forward time
integration) can be achieved by smoothing (backward
time integration), the final parameter covariances
remain arbitrary, since they evolve from arbitrary
initial covariance estimates.

Assuming that all state variables and their rates have been either measured or are otherwise known from manipulating the measurement data, the unknown parameters, if they occur in linear form in the state equation, can be found by application of a linear filter, see for example reference 7. The classical regression method is a special case of this direct filtering method, namely for infinite initial parameter covariances. In classical regression one obtains a single value of the error covariance matrix. The direct filter application allows the use of a finite initial error covariance matrix and it gives the evolution of this matrix as a function of time. One thus obtains an indication when to stop processing the test data after their information contents has been exhausted. As mentioned before, the absolute values of the error covariances are meaningless, since one usually does not have a rational way of establishing initial values for the parameter covariances.

A method that appears to be economic of computer time for large numbers of unknown parameters was used in reference 4 for application to helicopters. The method consists of a simultaneous identification of

states and parameters with the help of a non-linear filter. In other words, the unknown parameters are treated as additional state variables. Since there occur products of state variables and parameters, the filter is a non-linear one. The so called extended Kalman filter appears to be particularly useful for this purpose. Either non-linear filtering alone or linear filtering in combination with smoothing is performed. The absolute values of the parameter covariances are again of no physical significance since they depend on the arbitrary initial values. In the following, a brief discussion of the direct use of filters in parameter estimation is given.

LINEAR FILTER METHOD OF PARAMETER ESTIMATION

The equation error estimate based on the performance criterion of equation 5 was given in global form by equation 9, which is valid provided that the system is linear in the unknown parameters θ. The global estimate can be replaced by a filter solution. Consider a system with the process equation

$$\dot{\theta} = 0 \tag{41}$$

and with the measurement equation

$$\dot{x}(t) = f(x,u,\theta,t) + v(t) \tag{42}$$

The extended Kalman filter that minimizes sequentially

the performance criterion 5 is given by the filter

equation

$$\dot{\hat{\theta}} = P_\theta (\partial f/\partial \theta)^T \; W \; (\dot{x} - (\partial f/\partial \theta)\hat{\theta}) \tag{43}$$

and by the covariance equation

$$\dot{P}_\theta = -P_\theta (\partial f/\partial \theta)^T \; W \; (\partial f/\partial \theta) \; P_\theta \tag{44}$$

x and $\dot{x}$ are here quantities that are measured or that

have been reconstituted from measurements by prefiltering.

The weighting matrix W can be interpreted as the inverse

of the error covariance R^{-1} for $v(t)$. The sequential

estimate from the filter solution gives the optimum

based on all preceding data. One can either select

an initial estimate $\theta(0)$ to integrate equation 43,

or one can also assume a zero value for the initial

estimate. The initial error covariance $P_\theta(0)$, as

mentioned before, is usually not known. One should,

therefore, use a rather large value for the initial

covariance. Though the resulting values of $P_\theta(t)$ obtained as

function of time by solving equation 44 have no physical

meaning as far as absolute values are concerned, $P_\theta(t)$

will asymptotically approach zero, and it can be used as

a criterion when to stop processing the data. As can be seen from equation 43, no further changes in the parameters θ will occur, once P_θ has approached zero.

The covariance equation 44 has the closed form solution (see reference 7)

$$P_\theta(t) = [P_\theta^{-1}(0) + \int_0^t (\partial f/\partial\theta)^T W (\partial f/\partial\theta) \, d\tau]^{-1} \quad (45)$$

and the filter equation 43 has the closed form solution

$$\hat{\theta}(t) = P_0(t) [P_\theta^{-1}(0)\theta(0) + \int_0^t (\partial f/\partial\theta)^T W \dot{x} \, d\tau] \quad (46)$$

Whether it is computationally simpler to numerically integrate equations 43 and 44 or to evaluate equations 45 and 46 depends on how many points in time are to be covered. For infinite initial parameter covariance, $P_\theta^{-1}(0) = 0$, equation 46 becomes identical to equation 9.

BAYESIAN ESTIMATION AS A FILTERING PROBLEM

If we extend the quasi-Bayesian or maximum a posteriori probability (MAP) criterion 16 to include both the parameters θ and the states $x(t)$, we have

$$\max_{x(t),\theta} f(x(t),\theta/Z) \sim \max_{x(t),\theta} f(Z/x(t),\theta) f(x(t),\theta) \quad (47)$$

Assuming now the non-linear system and measurement equations 1 and 2, and assuming further that states and parameters have Gaussian distributions of the form of equation 18, the criterion 47 becomes (see for example references 4 and 8) one of minimizing the quadratic function

$$J = (1/2) \left[\left|\left|\theta(0)-\hat{\theta}(0)\right|\right|^2_{P_{\theta}^{-1}(0)} + \left|\left|x(0)-\hat{x}(0)\right|\right|^2_{P^{-1}(0)} \right. \tag{48}$$

$$\left. + \int_0^T \left\{ \left|\left|y(t)-h(x,u,t)\right|\right|^2_{R^{-1}} + \left|\left|w(t)\right|\right|^2_{\Gamma Q^{-1} \Gamma^T} \right\} dt \right]$$

subject to the constraint equation 1. If the system and measurement equations 1 and 2 are linearized about the current estimates $\hat{x}$ and $\hat{\theta}$ the recursive solution of the minimization problem 48 results in the extended Kalman filter equations given for the continuous case by (see reference 4)

$$\dot{x} = (\partial f/\partial x)\hat{x} + (\partial f/\partial u)u + P(\partial h/\partial x)^T R^{-1}(y-(\partial h/\partial x)\hat{x})$$

$$\dot{\theta} = P_{\theta x}(\partial h/\partial x)^T R^{-1} (y - (\partial h/\partial x)\hat{x})$$

$$\dot{P} = (\partial f/\partial x)P + P(\partial f/\partial x)^T - P(\partial h/\partial x)^T R^{-1}(\partial h/\partial x)P + Q \tag{49}$$

$$+ (\partial f/\partial\theta)P_{\theta x} + [(\partial f/\partial\theta)P_{\theta x}]^T$$

$$\dot{P}_{\theta x} = P_{\theta x}(\partial f/\partial x)^T + P'(\partial f/\partial\theta)^T - P_{\theta x}(\partial h/\partial x)^T R^{-1}(\partial h/\partial x)P$$

with $P' = - P_{\theta x}(\partial h/\partial x)^T R^{-1}(\partial h/\partial x)P_{\theta x}^T$

P is as before the state covariance, P_Θ the parameter

covariance, and $P_{\Theta x}$ is the combined covariance where

the parameters Θ are included in an augmented state.

P' is an abbreviation given in the last of equations

49. Even if the original system is linear, the

augmented system is non-linear and hence the filtering

problem must be solved by a non-linear filtering

technique. In reference 4 the raw data are preprocessed

by a digital filter and by a Kalman filter that does

not use the unknown parameters but merely makes use

of the transformation equations from a space-fixed to

a body-fixed reference system (Euler equations).

Lebacqz in reference 9 applies basically the same

method except for a discrete instead of the continuous

filter formulation. He further uses a one stage

filtering-smoothing algorithm which has the advantages

of reducing the bias due to non-linearities and of making

the algorithm less sensitive to initial conditions.

Mehra in reference 4 is critical of using an extended

Kalman filter for the augmented state including the

unknown parameters. His arguments are that the

uncertainties in the states are usually much smaller

than the uncertainties in the parameters. Therefore

the assumption of local linearization about the latest

estimate which are acceptable for state estimation with
an extended Kalman filter are generally less valid for
parameter estimation. Moreover, the filter for the
augmented state assumes knowledge of the a priori
parameter covariances which are unknown. As mentioned
before, the arbitrary a priori parameter covariance
used as initial conditions for a filter that includes
parameters as state variables gives unreliable confidence
limits on the parameter estimates. An added difficulty
of applying a filter to the augmented state is that poor
a priori estimates of the parameters make the convergence
rate slow or may even cause divergence of the filter
solution. Though improvements can be applied to the
extended Kalman filter like local smoothing and local
iteration and smoothing, the basic shortcomings of this
method appear to have been correctly described in reference 4.
Unfortunately, the application of the complete algorithm
of maximum likelihood identification given in reference
5 is for a large system much more demanding of computer
size and time than the filter solution with the
augmented state. While aircraft parameter identification
with the complete maximum likelihood algorithm of
reference 5 has not as yet been accomplished, the method
of filtering the augmented state has been applied to
several aircraft parameter identification cases, for
example in references 4 and 9.

IDENTIFIABILITY PROBLEMS

Identifiability problems can occur no matter what identification algorithm is used. They are related to the initial 3 steps involved in system identification as listed at the beginning of this chapter: the selection of a suitable input, the selection of the instrumentation, and the selection of the mathematical model. A few comments are added here to point out some difficulties that have been encountered due to these three initial steps.

If the input does not adequately excite some of the system modes, the associated parameters cannot be adequately identified. Sometimes it is practical to combine the responses to various types of inputs into a single identification run, see reference 4. While each of the single inputs excites only a limited number of modes, the combination of inputs provides an adequate excitation of all modes required for the estimation of the parameters. Efforts have also been made to design inputs on the basis of certain optimization criteria. More details on this problem are given in Chapter 3 of this report.

If there are large unaccounted for instrumentation errors non-physical parameter values may result. In reference 10 instrumentation lags and control measurement

errors were found to be most significant. Static
measurement errors and instrumentation lags can be a
much greater source of parameter inaccuracies than
white noise. A detailed analysis of the relationship
between static and dynamic measurement errors in states
and control inputs and the accuracy of the parameter
estimates is required.

If the selected mathematical model for the system
is inadequate, the parameters are forced to account for
some unmodeled effects. The estimated parameters may,
therefore, be quite different from those determined by
aerodynamic theory or wind tunnel tests would indicate.
A good example is given in reference 11 where a six
degree of freedom mathematical model for a helicopter
gave unrealistic derivatives, since it had to account
for effects of some neglected modes. A unique six
degree of freedom linear model for the helicopter flight
dynamics does actually not exist. When a nine degree of
freedom mathematical model is used, these difficulties
disappear. Modeling errors are also a major cause for
the lack of convergence of iteration procedures or of
parameter identification by filtering methods. The best
remedy against difficulties from modeling errors is the
adoption of a more suitable mathematical model. Some
other measures to improve the convergence of iteration
procedures or of filtering methods will be briefly discussed.

In the cases where a priori values of parameters, for example from theory or from wind tunnel tests, are available, one can use an a priori weighting matrix that expresses the confidence in these values and prevents the algorithm from deviating too much from the a priori values. Sometimes there exist some relationships between the parameters. These should then be used as constraints in the optimization problem to avoid non-physical parameter estimates. If parameter dependencies exist, difficulties are encountered in inverting the information matrix. An exact dependency between parameters should result in a zero eigenvalue of the information matrix. A rank deficient solution makes use of the fact that in case of near parameter dependencies there is a large spread between a set of small eigenvalues and another set of much larger eigenvalues of the information matrix.

In filter solutions, divergence because of modeling errors can occur when the covariance matrix becomes prematurely too small, thus preventing further test data to be of influence. There are several ways to prevent premature small covariances. One can provide fictitious noise input to the system or one can directly increase the parameter covariance in each time step according to some rule. One can also overweigh the most recent data thus causing the filter to reduce its memory of the data

of the more distant past. This indirectly increases the
parameter covariance matrix. Since too short data length
and too large errors in the initial parameter estimates
may also result in non-physical parameter values or in
divergence of the identification algorithm, longer
transients and better a priori parameter estimates can
lead to the avoidance of these difficulties.

VALIDATION OF ESTIMATES

Once a set of parameter estimates has been obtained
the question arises; what confidence can be associated
with this set? As mentioned before, the parameter
covariance matrix obtained by filtering the augmented
state is not a good measure of this confidence. The
inverted information matrix obtained with the maximum
likelihood method represents the Cramér-Rao lower bound
for the parameter covariances and is a better measure
of this confidence. Using the parameter estimates to
predict the transients from which the estimates have
been obtained, and computing the rms error with respect
to the measured transients, gives another confidence
measure. However, if the system is inadequately modeled,
one may obtain a small rms error despite the fact that
the parameter values are wrong in comparison to theoretical
or wind tunnel results, see reference 11. A better way
of validitation is to compare the prediction with the
results of test data not used in the identification process.
In fact, it is good practice not to use all of the

available test data for the parameter identification but
to reserve some of the runs for such a comparison.
Sometimes it is desirable to perform the parameter
identification not just with one mathematical model but
with a variety of models. In the case described in
reference 11 a mathematical model with more parameters
gave a much better identification result than a model
with fewer parameters, better in the sense of an
improved correlation with theoretically and wind tunnel
generated parameters. However, there are also cases
where mathematical models with a larger number of
parameters gave worse identification results than a
model with fewer parameters, see reference 12.
Adequate parameter estimation from transients requires
careful attention to the many contributing factors in
the input, instrumentation, mathematical modeling, and
the estimation algorithm; and the validation of this
process can only be considered complete after the rms
errors of the prediction with the estimated parameters
as compared to test data have been found acceptably
small for all types of possible transient excitations
of the system.

APPLICATIONS TO LIFTING ROTORS

Contrary to the fixed wings of airplanes, lifting rotor characteristics are not well approximated by the usual set of aerodynamic derivatives. One reason is ,blade modes that' must be considered particularly in rapid transients. Another reason is the dynamic rotor wake that is produced by the time varying rotor thrust and rotor pitching and rolling moments and that has a feedback effect on the rotor forces and moments. The omission of the blade modes, as shown in reference 11, results in non-unique and non-physical, rotorcraft derivatives. The identification is better if separate rotor degrees of freedom are introduced even in the crude form of a first order lag as was done in reference 13.

A variety of identification methods have been used with respect to lifting rotors. After preprocessing the test data with a digital filter followed by a Kalman filter that does not contain the aerodynamic derivatives (transformation or Euler equations), least squares identification is applied to rotorcraft transient flight test data in references 4 and 11. Each identification run is made with several transients simultaneously. The least squares results are then used as start-up values for an extended Kalman filter for the augmented state. It is not obvious that the extended Kalman filter

actually improves on the least squares results, though
filter convergence is achieved. In reference 13 the
output error method with quasilinearization is applied
without preprocessing the flight test data. The flight
data of both references 4 and 13 were obtained in calm
air. The equation error method in its filter form was
applied in reference 7 to simulated noisy blade flapping
and torsion measurements at high rotor advance ratio.
The simulated data were preprocessed by a Graham
digital filter, but not by a Kalman filter. Reference 7
assumed that all states and their derivatives had been
measured. In contrast reference 14 assumed that only
flapping deflections are measured but not flapping rates
or flapping_accelerations. For the dynamic wind tunnel
tests simulated in reference 14 there is no way of
applying a Kalman filter that does not contain the
unknown parameters. However, it was found that for the
cases studied, a Kalman filter with considerable errors
in the unknown parameters was useful in obtaining the
non measured flapping rates and accelerations. The
parameter identification was then performed by the
equation error method in its filter form.

In Chapter 2 of this report the same method (except
for using global estimates) is used in an iterative form.
In addition, the output error method with quasilineari-
zation is applied to the same and to more complex rotor

identification problems. As will be shown, the first method is more computer cost effective in cases where it works, while the second method is more reliable and more versatile.

REFERENCES

1. Iliff, K. W. and Taylor, L. W., "Determination of Stability
   Derivatives from Flight Data Using a Newton-Raphson
   Minimization Technique," NASA TN D-6579, March 1972.

2. Denery, D. G., "Identification of System Parameters from
   Input-Output Data with Application to Air Vehicles,"
   NASA TN D-6468, August 1971.

3. Giese, C. and McGhee, R. B., "Estimation of Non-Linear
   System States and Parameters by Regression Methods,"
   JACC, 1965, pp. 46-53.

4. Molusis, J. A., "Rotorcraft Derivative Identification from
   Analytical Models and Flight Test Data," AGARD Flight
   Mechanics Panel Specialists Meeting, NASA Langley,
   November 1974.

5. Stepner, D. E. and Mehra, R. K., "Maximum Likelihood
   Identification and Optimal Input Design for Identifying
   Aircraft Stability and Control Derivatives," NASA
   CR 2200, March 1973.

6. Rediess, H. A., "An Overview of Parameter Estimation
   Techniques and Applications in Aircraft Flight Testing,"
   NASA TN D-7647, April 1974, pp. 1-18.

7. Hohenemser, K. H. and Prelewicz, D. A., "Computer
   Experiments on Periodic Systems Identification Using
   Rotor Blade Transient Flapping-Torsion Responses at High
   Advance Ratio," Rotorcraft Dynamics, NASA SP-352,
   February 1974.

8. Rauch, H. E., Tung, F. and Striebel, C. T., "Maximum
   Likelihood Estimates of Linear Dynamic Systems," AIAA
   Journal, Vol. 3, No. 8, August 1965, pp. 1445-1450.

9. Lebacqz, J. V., "Application of a Kalman Filter
   Identification Technique to Flight Data from the X-22A
   Variable Stability V/STOL Aircraft," NASA TN D-7647,
   April 1974, pp. 149-174.

10. Bryant, W. H. and Hodge, W. F., "Effect of Flight
    Instrumentation Errors on the Estimation of Aircraft
    Stability and Control Derivatives," NASA TN D-7647,
    April 1974, pp. 261-280.

11. Molusis, J. A., "Helicopter Derivative Identification from Analytical Models and Flight Test Data," NASA TN D-7647, April 1974, pp. 175-186.

12. Taylor, L. W., "A New Criterion for Modeling Systems," NASA TN D-7647, April 1974, pp. 291-313.

13. Gould, D. G. and Hindson, W. S., "Estimates of the Stability Derivatives of Helicopter and V/STOL Aircraft from Flight Data," DME/NAE Quarterly Bulletin No. 1974 (4), Ottawa, January 1975.

14. Hohenemser, K. H. and Yin, S. K., "Computer Experiments in Preparation of System Identification from Transient Rotor Model Tests," Part II of First Yearly Report under Contract NAS2-7613, June 1974.

Some textbooks related to the subject matter of this chapter are:

15. Eykhoff, P., "System Identification," John Wiley & Sons, 1974.

16. Jazwinski, A. H., "Stochastic Processes and Filtering Theory," Academic Press, 1970.

17. Sage, A. P. and Melsa, J. L., "System Identification," Academic Press, 1971.

18. Bryson, A. E. and Ho, Y. C., "Applied Optimal Control," Ginn & Co., 1969.

Chapter 2

COMPUTER EXPERIMENTS WITH PITCH STIRRING TRANSIENTS

K. H. Hohenemser and S. K. Yin

ABSTRACT

Selected methods are applied in the form of computer experiments to two problems of lifting rotor state variable and parameter identification. The first problem refers to a rotor condition at .4 advance ratio. Cyclic pitch stirring with constant acceleration is assumed as known input. Noise polluted blade flapping measurements are assumed as the only measured output. Blade Lock number and collective pitch are the unknown parameters to be identified. The second problem refers to a rotor condition at zero advance ratio again with constant acceleration pitch sitrring. A two parameter dynamic inflow model is stipulated. The only measurements are noise polluted blade flapping responses. Blade Lock number and the two parameters of the dynamic inflow model (including a time constant) are the unknown parameters to be identified.

Two parameter identification methods are applied and their relative merits evaluated: Iterative equation error estimation with updated Kalman filter and the maximum likelihood method. The latter method, though requiring increased computer effort per iteration, was found to be more versatile.

TABLE OF CONTENTS   Page

INTRODUCTION

The identification method used in reference 1 is
based on the experience gained from references 2 and 3
where states and parameters were determined simul-
taneously with the help of an extended Kalman filter.
The filter can easily diverge unless good initial
estimates are available. Particularly in reference
3 a considerable effort was applied to obtain such
good initial estimates. The test data were first
processed with a digital filter that took out high
frequency noise without distorting the main signals.
The data were then processed with a Kalman filter based
on the Euler equations, which do not contain the
unknown parameters. Thus measurement bias was removed
and missing channels were reconstituted. Finally a
least squares algorithm was applied to obtain estimates
of the unknown parameters. The subsequent application
of the extended Kalman filter led to modified parameter
estimates, however it is not clear whether or not
these modifications represent improvements. In any
case the modifications were not large, and the initial
estimates appeared to be satisfactory approximations.

In trying to apply the experience from reference 3
to wind tunnel model transients a difficulty arises,

in that no equivalent to the Euler equations for the
aircraft exists. Thus there is no way of using a
Kalman filter which is free of the unknown parameters.
Instead, if a Kalman filter is to be applied,
estimates of the parameters must be inserted. Another
difficulty for our wind tunnel model tests is that
only flapping deflection measurements are made, while
the rates of deflection and the accelerations are not
measured. Thus the Kalman filter with the estimated
parameters is called upon to provide both rates and
accelerations. Finally, the least squares algorithm
of reference 3 was replaced in reference 1 by a
sequential linear estimator for the parameters.
This has the advantage that finite initial parameter
covariances can be used, and that the time history of
the parameter covariance provides a measure for the
time beyond which no more useful information can be
extracted from the test data. As far as digital
filtering is concerned it was found in reference 1 that
a cut-off frequency range from 2.5 to 2.9 removed the
high frequency noise without unduly distorting the
main signal.

The method used in reference 1 worked well in the
computer experiments using normal flow transients at
.4 advance ratio for the identification of the Lock

number and of the collective pitch angle. However, a
number of questions had to be answered by the sub-
sequent studies. It had been decided to begin wind
tunnel model rotor transient testing with constant
acceleration pitch stirring inputs, and later follow
up with normal flow transient testing. Therefore, it
was desirable to perform computer experiments with
pitch stirring transients. All of the identification
work reported here concerns such transients. The
linear sequential estimator used in reference 1
requires the simultaneous integration of the filter
and of the covariance differential equations. A
simpler "global" estimate requires only the inversion
of a system of linear equations for the unknown
parameters and the evaluation of a number of integrals
over the time period of the transient. Therefore, a
number of comparisons were made between these two
methods. Finally, the Kalman filter for the test data
requires estimates of the unknown parameters. The
question arises whether an iterative form of the
method is practical, where the identified parameters
are reinserted into the Kalman filter and a second
identification is performed, etc. This iteration
method was tried out in several cases.

While the substitution of the theoretical blade Lock number by an equivalent Lock number can be expected to provide a reasonably good approximation of dynamic rotor wake effects if the transient is relatively slow and does not contain high frequencies, a better dynamic wake representation is given by introduction of the L-matrix (reference 4) together with a time constant (reference 5). For zero advance ratio the L-matrix degenerates into a L-scalar, so that a two parameter dynamic wake description is obtained. A single blade representation is now no more applicable, and a multi-blade analysis is necessary. A number of computer experiments were conducted for this case whereby the blade Lock number was treated as third unknown parameter.

During the last decade the maximum likelihood method of parameter identification has been success-fully applied to airplane and helicopter transient testing. This method does not require preprocessing of the test data and also does not need complete measurements of the deflecitons, of their rates and of the accelerations. The parameter covariance estimates obtained with this method are more meaningful than those obtained with the linear sequential estimator used in reference 1. A number of cases were treated with the maximum likelihood method.

In its most general form presented in reference 6 the maximum likelihood method is capable of handling identification problems for cases with both measurement and system noise. The method then becomes computationally quite complex and has not as yet been applied to aircraft transient testing. The usual assumption is that the system is free of noise. In this case the maximum likelihood method is greatly simplified (output error method). It has been applied in this report in the simplified form. If the random system noise is measured, the simplified method is still applicable. However in the planned wind tunnel model testing the system noise will not be measured, and though it is expected to be small, the question is whether or not it will unacceptably degrade the estimates of the parameters. Therefore a few computer experiments were conducted with both simulated measurement noise and with simulated system noise, whereby the maximum likelihood method was applied in its usual simplified form.

EXCITATION OF PITCH STIRRING TRANSIENTS

For the wind tunnel experiments with pitch stirring transients the initial state of the rotor will be given by prescribing the advance ratio, the collective pitch angle, the rotor angle of attack (set at approximately

zero) and the cyclic control setting that will be zero longitudinal cyclic and $1.5°$ lateral cyclic. At the time t_o pitch stirring is initiated. Denoting by ω the angular pitch stirring speed, positive in the direction of rotor rotation, and by $\dot{\omega}$ the pitch stirring angular acceleration assumed to be constant, we have

$$\omega = \dot{\omega}(t - t_o) \tag{1}$$

For a progressing mode ω is negative, for a regressing mode ω is positive. In a rotating reference system the blade pitch angle is given by

$$\theta = \theta_o + 1.5 \cos [\omega(t-t_o) + t]$$

$$\omega = \begin{cases} 0 & \text{for } t \leq t_o \\ \dot{\omega}(t - t_o) & \text{for } t > t_o \end{cases} \tag{2}$$

In a multiblade representation the blade pitch angle of the kth blade is

$$\theta_k = \theta_o - \theta_I \sin \psi_k + \theta_{II} \cos \psi_k \tag{3}$$

where

$$\theta_I = \begin{cases} 0 & \text{for } t \leq t_o \\ 1.5 \sin \omega(t-t_o) & \text{for } t > t_o \end{cases} \tag{4}$$

$$\theta_{II} = \begin{cases} 1.5 & \text{for } t \leq t_o \\ 1.5 \cos \omega(t-t_o) & \text{for } t > t_o \end{cases} \tag{5}$$

θ_I represents forward cyclic pitch, θ_{II} represents
left cyclic pitch. The wind tunnel experiments will be
conducted with a variety of pitch stirring accelerations.
The computer experiments were mostly conducted with a
pitch stirring acceleration of

$$\dot{\omega} = - .1/\pi \qquad\qquad (6)$$

which is in the progressing sense.

Since in the non-dimensional time units used here
the time of one rotor revolution is 2π, the angular
pitch stirring velocity one rotor revolution after
initiation of pitch stirring is .2, that is one fifth
of the rotor angular speed. Figure 1 shows the time
history of blade pitch for about two rotor revolutions
($t_o = 12$, $t = 12$ to 24) in a rotating frame of reference.
Figure 2 shows the time history of blade pitch in
multiblade representation, that is θ_I and θ_{II} vs.
time t . Figures 1 and 2 refer to the progressing mode.

A convenient way of conducting the computer
experiments is to impose at time $t = 0$ a step input
of lateral cyclic control. If t_o is sufficiently large,
the transient from the step control input will have
subsided when pitch stirring begins at $t = t_o$. It
was found that for single blade identification a value
of $t_o = 12$ is adequate. In the multiblade identification

including the time delayed rotor wake this value was found to be insufficient. Instead, $t_0 = 70$ was used. Measurement noise was simulated by polluting the analytical flapping response with zero mean white Gaussian computer generated random sequences. System noise was simulated by adding to the multiblade inflow coordinates in the system equations zero mean white Gaussian sequences. Data processing extended over nearly two rotor revolutions, from t_0 to $t_0 + 12$. In Chapter 3 it will be shown that this choice leads to approximately optimal data utilization. If a shorter time period were used, large errors in the parameter estimates would occur. If a longer time period were used, the additional data processing would be unnecessary in view of the adequate accuracy of the parameter estimate using a time period of two rotor revolutions. For the wind tunnel experiments a variety of collective pitch settings Θ_0 will be tested. For the computer experiments the collective pitch setting was mostly assumed to be $\Theta_0 = 2^\circ$.

LINEAR SEQUENTIAL AND GLOBAL ESTIMATORS

In reference 7 the parameter identification is performed from a "system equation"

$$\dot{\Theta} = 0 \tag{7}$$

and a "measurement equation"

$$\zeta = h(\xi,\theta) + v \qquad (8)$$

Equation 8 is actually the system equation arranged in a form where the left hand side contains all terms that are free of the unknown parameters θ. If the system equation is linear in the state variables ξ and in the unknown parameters θ, $h(\xi,\theta)$ is a linear function of the parameters. The noise vector v refers only to the terms on the left hand side of equation 8. The state variables that are multiplied by the unknown parameters in $h(\xi,\theta)$ must be noise free. To obtain the θ both ζ and ξ must be known. If only part of the variables in ζ and ξ have been measured, Kalman filtering is required in order to reconstitute the missing terms.

Denoting by P_θ the parameter covariance matrix and by R the noise covariance matrix, assuming that v is zero mean Gaussian white noise, optimal parameter estimates $\hat{\theta}$ can be obtained by minimizing the cost function

$$J=(1/2)\{'(\theta(0)-\hat{\theta}(0))^T P_\theta^{-1}(0)(\theta(0)-\hat{\theta}(0))+ \int_0^T (\zeta-h(\xi,\theta))^T R^{-1}$$

$$(\zeta-h(\xi,\theta))dt\} \qquad (9)$$

where the a priori estimates $\theta(0)$, $P_\theta(0)$ are assumed to be given together with the noise covariance matrix R. The differential equations associated with this optimal problem are (see for example reference 8)

$$\dot{\hat{\theta}} = P_\theta (\partial h/\partial \theta)^T R^{-1} (\zeta - h(\xi,\hat{\theta})) \qquad (10)$$

$$\dot{P}_\theta = - P_\theta (\partial h/\partial \theta)^T R^{-1} (\partial h/\partial \theta) P_\theta \qquad (11)$$

Beginning with the initial a priori estimate for the parameters $\theta(0)$ and their covariance matrix $P_\theta(0)$, these equations can be integrated and result at each time t in the optimal parameter estimate given the preceding measurements. Since the initial parameter covariance is usually not known and the assumed values are rather arbitrary, the matrix P_θ from integrating equation 11 is not a useful measure of the actual parameter covariance. However, once P_θ has approached zero, the effect of any further measurements on the estimate $\hat{\theta}$ also approaches zero as is evident from equation 10. P_θ, therefore, is valuable in judging for what length of time the data should be processed. Equations 10 and 11 represent the "linear estimator" used in references 1 and 7.

Instead of the sequential estimation by integrating equations 10 and 11 with some initial estimates $\Theta(0)$ and $P_\Theta(0)$, one can also obtain a "global" estimate directly from equation 9. If one assumes that one and the same parameter estimate $\hat{\Theta}$ is valid throughout the time range from 0 to T, one obtains by setting $\partial J/\partial\Theta = 0$

$$\hat{\Theta} = [P_\Theta^{-1}(0) + \int_0^T (\partial h/\partial\Theta)^T R^{-1}(\partial h/\partial\Theta)dt]^{-1} [P_\Theta^{-1}(0)\Theta(0)$$

$$+ \int_0^T (\partial h/\partial\Theta)^T R^{-1} \xi\, dt] \tag{12}$$

see for example the appendix of reference 7. A convenient assumption is $P_\Theta^{-1}(0) = 0$, which means an infinite initial parameter covariance matrix. The initial estimate $\Theta(0)$ is then not required and the evaluation of equation 12 is reduced to the determination of fixed boundary integrals, a matrix inversion and a matrix multiplication. The parameter covariance matrix at the time T is given by the first factor of equation 12:

$$P_\Theta(T) = [P_\Theta^{-1}(0) + \int_0^T (\partial h/\partial\Theta)^T R^{-1}(\partial h/\partial\Theta)dt]^{-1} \tag{13}$$

which follows from the integration of equation 11, (see

for example the appendix of reference 7). $P_{\Theta}(T)$ from

equation 13 can again be used to judge whether or not

all the significant information contents has been

extracted from the data. A comparison between the

estimation with equations 10 and 11 and with the

"global" method of equation 12 will be given later

for a specific example.

ITERATIVE EQUATION ERROR ESTIMATION WITH UPDATED
KALMAN FILTER

When using the parameter estimation methods of the

preceding section it is necessary to first determine

from the noisy deflection measurements estimates for the

deflections, for their rates, and for the accelerations.

In reference 1 this was done by passing the noisy

deflection data through a digital filter that takes out

the noise above a certain frequency without distorting

the signal in the low frequency range. The filtered

deflections were then either differentiated twice, or

a Kalman filter was applied in order to obtain the

derivatives. Later studies showed large errors in the

parameters for too low cut-off frequency of the digital

filter. It was then decided to omit the digital filter

and instead use the Kalman filter in an iterative way.

In typical examples already the second iteration was as accurate as the result with the combined digital and Kalman filter given in reference 1.

Only simulated noisy blade flapping measurements were used in the Kalman filter. The filter provided the deflection rates and accelerations needed for the "global" parameter estimate, but not the deflections themselves. In other words, the parameter estimate was performed with the simulated noisy deflection measurements and with the rates and accelerations from the Kalman filter. In the first iteration a Kalman filter with estimated parameter values was used (typically 20% error). After updated parameter values had been obtained, a second pass with an updated Kalman filter was performed, etc. The deflection data remained the same for each iteration, but the rates of deflection and the accelerations were updated. As will be seen in the numerical examples this method worked well for the single blade identification. The reason for the good parameter identification in spite of substantial measurement noise is probably the following. In the equation of single blade flapping the flapping deflection term includes a large part that is independent of the Lock number, and only a relatively

small part that is dependent on the Lock number.
Therefore, the major part of the noisy deflection
term can be included in the left hand side of equation
8 in ζ. The method allows noise in this part. The
smaller part of the flapping deflection term that is
of aerodynamic origin and includes the blade Lock
number is part of the right hand side of equation 8
and should be noise free. Apparently the actual
noise in this relatively small term does not lead to
a substantial bias in the parameter estimate.

MAXIMUM LIKELIHOOD ALGORITHM

The maximum likelihood method for our particular
case pertains to a system equation (zero system noise)

$$\dot{x} = f(x, u, \theta) \tag{14}$$

where x is the state vector, u the input vector as-
sumed to be known. θ is the vector of unknown para-
meters that may include initial values of state
variables. The measurement equation is assumed to be
linear and of the form

$$y = H \ x + v \tag{15}$$

y is the vector of observed quantities, H is a matrix
relating the state variables to the observations, v
is the vector of random measurement errors, assumed to

be zero mean white noise with given covariance matrix R

$$E[v(t)v^T(\tau)] = R \; \delta(t, \tau) \tag{16}$$

R is assumed to be constant with time. Though the preceding equations do not show bias terms, bias errors could easily be included in the unknown parameter vector Θ.

A sample of measurements $y_1 \; y_2 \; . . . \; y_n$ is now made during the time of the transient and the parameter estimate $\hat{\Theta}$ is selected such that the conditional probability of this sample of measurements given Θ is maximized.

$$\hat{\Theta} = \max_{\Theta} p(y_1 . . . y_n/\Theta) \tag{17}$$

The following steps lead to the maximum of the likelihood function $p(y_1 . . . y_n/\Theta)$, though there is no assurance that the maximum is global. The method outlined here is called quasilinearization with the modified Newton-Raphson method. It assumes Gausian distributions of the random variables.

1 Select an initial parameter estimate $\hat{\Theta} = \Theta_0$.

2 Solve the system equation 14 with this parameter estimate

$$\dot{\hat{x}} = f(\hat{x}, u, \hat{\Theta}) \tag{18}$$

The initial conditions can either be obtained
from the measurements, or, where this is not
feasible, they can be included in the unknown
parameter vector $\hat{\theta}$.

3 Calculate for each measurement the "innovation term"

$$\nu_j = y_j - H \hat{x}_j \tag{19}$$

4 Solve the "sensitivity equations"

$$\partial \dot{\hat{x}}/\theta_k = \partial f/\partial \theta_k + F(t) \partial x/\partial \theta_k \tag{20}$$

where $F(t) = \partial F(t)/\partial x \Big|_{x = \hat{x}}$

The initial conditions of $\partial \hat{x}/\theta_k$ are zero except when
$x(0)$ is identified as part of the parameter vector $\hat{\theta}$.
In this case the initial partials have the value one.

5 The likelihood function for zero system noise is

$$J = \log p(y_1 \cdot \cdot y_n/\theta) = - (1/2) \sum_{j=1}^{N} \nu_j^T R^{-1} \nu_j \tag{21}$$

Determine now the gradient of this function with
respect to θ

$$\partial J/\partial \theta = \sum_{j=1}^{N} \nu_j^T R^{-1} \partial \nu_j/\partial \theta \tag{22}$$

where $\qquad \partial v_j / \partial \theta = - H \; \partial \; \hat{x}_j / \partial \theta$ (23)

6 Compute the information or sensitivity matrix

$$M = \partial^2 J / \partial \theta^2 = \sum_{j=1}^{N} (\partial v_j / \partial \theta)^T \; R^{-1} \; \partial v_j / \partial \theta$$ (24)

The inverse M^{-1} of the information matrix provides a lower bound for the covariance of the updated parameter estimates.

7 The updated parameter estimate is

$$\hat{\theta} = \theta_0 + \Delta \theta$$ (25)

where $\qquad \Delta \theta = -M^{-1} (\partial J / \partial \theta)^T$

8 Go now back to equation 18 with the updated parameter estimate and repeat the steps to equation 25. Reiterate until convergence of the information matrix and of the parameter vector is obtained.

SINGLE BLADE PARAMETER IDENTIFICATION

We first present a case where the sequential estimate is compared to the global estimate of the parameter γ (blade Lock number). The excitation of the progressing mode is given by

$$\theta = 1.5 \cos (\omega + 1)t$$

$$\omega = \begin{cases} 0 & \text{for } t \leqslant 12 \\ -(.1/\pi)(t-12) & \text{for } t > 12 \end{cases} \qquad (26)$$

This excitation is somewhat different from that defined by equation 2 and illustrated in Figure 1. While equation 2 defines a constantly accelerated progressive excitation, equation 26 deviates somewhat from a constant acceleration. The difference is, however, not essential. The collective pitch angle is $\theta_o = 0$. The rotor advance ratio is $\mu = .4$. The blade frequency in the rotating system is $P = 1.2$. System and measurement equations for the single blade case are given by equations 14 and 15 of reference 1.

The sequential estimate of γ and of its covariance, was found from equations 10 and 11, using as initial estimate at time $t = 12$ the value zero, and using the initial covariance $P_o = 2000$. The global estimate was determined from equation 12 setting the initial covariance to infinite, or $P_o^{-1} = 0$. Figure 3 gives the estimate $\hat{\gamma}$ and its covariance from the sequential method for the time $t = 12$ to 24 and also the value of the global estimate, $\hat{\gamma} = 5.05$. The true value of this

parameter is 5.00. From the covariance plot it is seen

that at t = 16 the covariance is quite small as

compared to the initial value so that any further data

processing will not appreciably change the estimate.

This is born out by the $\hat{\gamma}$ plot. The global estimate

for the time t = 24 agrees with the sequential

estimate at this time, as it should be, since there

is little difference between assuming the initial

covariance as P_O = 2000 or P_O = ∞. The method used

for figure 3 is the same as that used in reference 1.

The flapping response data were noise polluted by

white noise with a standard deviation of σ_β = .2.

This is a large noise since the maximum flapping

excursion is only 1.2. The polluted flapping

deflections were then passed through a digital filter

with the cut-off frequencies of 2.5 to 2.9. Rates

of deflections and accelerations were determined from

a Kalman filter with the erroneous value of γ = 4.0.

Sequential and global estimates were determined with

the flapping deflections from the digital filter output

and with the flapping rates and accelerations from

the Kalman filter output. It is seen that this method

worked very well in this case. The effect of the

digital filter cut-off frequencies is, however, substantial. The global method gave for the same case the following estimates

Cut-off Frequency		Estimate $\hat{\gamma}$ for t = 24
ω_c	ω_t	
1.7	1.9	5.27
2.5	2.9	5.05
3.5	3.9	4.89

Because of the sensitivity of the estimate with respect to the cut-off frequencies of the digital filter it was decided to omit the digital filter and to use only the Kalman filter, however in an iterative way.

The following example is again for a rotor advance ratio of μ = .4 blade frequency in the rotating system ω_1 = 1.2 and for a collective pitch setting of θ_o = 2°. Constant acceleration of progressive pitch stirring according to equation 2 is assumed, whereby t_o = 12. The polluted flapping response was processed with a Kalman filter using the initial parameter values of γ = 4.0 and δ = $\gamma\theta_o$ = 8.0 as compared to their true values of 5.0 and 10.0 respectively. The two parameters

assumed to be unknown, γ and δ, were determined by
the global estimate equation 12 with $P_\theta^{-1}(0) = 0$ and
$R = I$. For the initial conditions of the Kalman
filter $\beta(0)$ was taken from the simulated measurement
and $\dot\beta(0)$ was taken with a 20% error. The following
table gives in addition to the parameter values for
4 iterations also the diagonal terms of the
covariance matrix. There is some overshoot in both
parameters and the convergence is not very good.
Note that the covariance does not properly reflect
the actual errors in the parameters.

$$\delta_\beta = .2, \quad R = I$$

Parameter		γ	$\delta = \gamma\theta_o$	P_γ	P_δ
True value		5.00	10.00		
Initial estimate		4.00	8.00		
Iteration (1.6 CPU	1	5.66	8.95	.092	.113
seconds per iteration)	2	4.76	10.31	.082	.111
	3	4.52	10.09	.073	.111
	4	4.59	9.93	.071	.111

We now treat the same problem with the maximum
likelihood method, and first assume the correct initial
conditions $\beta(12) = -.921$ and $\dot\beta(12) = .986$. The first 3
iterations are given by

$$\sigma_\beta = .20$$

Parameter		γ	δ	$\partial J/\partial\gamma$	$\partial J/\partial\delta$
True value		5.00	10.00		
Initial estimate		4.00	8.00		
Iteration 1 (2.6 CPU seconds per iteration)	1	4.94	9.70	-6.11	-5.84
	2	4.91	9.82	.10	-.34
	3	4.91	9.82	-.002	.001

Though the computer CPU time is now per iteration higher than for the preceding method, the convergence is very rapid and the accuracy is very good. The Cramér-Rao lower bound for the parameter covariance matrix for the second and third iteration is

$$M^{-1} = \begin{bmatrix} .211 & -.036 \\ -.036 & .365 \end{bmatrix}, \quad R = I$$

In order to obtain the covariances for the actual simulated measurement noise $\sigma_\beta = .20$ one must multiply the above values by .04. Thus the lower bounds for the parameter standard deviations are $\sigma_\gamma = .092$, $\sigma_\delta = .12$ which is quite close to the errors in the second iteration.

In a second version of the same problem it is assumed that the initial conditions are unknown and must be included in the parameter identification. The following table gives the results of the first 3 iterations.

$$\sigma_\beta = .20$$

Parameter		γ	δ	$\beta(12)$	$\dot{\beta}(12)$
True value		5.00	10.00	-.921	.986
Initial estimate		4.00	8.00	-.750	1.200
Iteration 1 (4.3 CPU		4.96	9.68	-.89	1.02
seconds per iteration) 2		4.90	9.85	-.89	1.01
3		4.90	9.85	-.89	1.01

The parameter covariance lower bounds are

$$M^{-1} = \begin{bmatrix} .230 & -.059 & -.048 & .023 \\ -.059 & .405 & .058 & .022 \\ -.048 & .058 & .130 & -.066 \\ .023 & .022 & -.066 & .228 \end{bmatrix}, \quad R = I$$

The computer CPU time per iteration is once more increased, but the convergence is excellent and the accuracy of the estimate is very good. Again the inverted information matrix gives physically meaningful covariances for the unknown parameters.

One can then conclude that for single blade
identification with two unknown parameters γ and
$\delta = \gamma \Theta_0$ the iterative equation error estimation with
updated Kalman filter gives the lowest CPU time per
iteration, however because of the much faster convergence
of the maximum likelihood method and because of its
greater versatility (inclusion of initial response
values in the set of unknown parameters), it definitely
is preferable. For the following cases the maximum
likelihood method will be used.

MULTIBLADE PARAMETER IDENTIFICATION

The preceding example is now treated with a multi-
blade representation. The excitation in multiblade
coordinates is given by equations 3 to 5 with $t_0 = 12$.
The pitch input is shown in Figure 2 assuming a
constant progressing stirring acceleration given by
equation 6. The multiblade responses β_0, β_I, β_{II} are
polluted from $t = 12$ to $t = 24$ with zero mean Gaussian
noise of standard deviation .2.

We first identify the parameters γ and $\delta = \gamma \Theta_0$
with given initial values of β_0, $\dot{\beta}_0$, β_I, $\dot{\beta}_I$, β_{II}, $\dot{\beta}_{II}$
at $t = 12$. The responses are obtained with the periodic
system equations. However, for the parameter identi-
fication, constant coefficients are assumed so that

modeling errors are present. The initial states at
t = 12 are:

$$\beta_o = .880, \quad \beta_I = -.983, \quad \beta_{II} = 1.698, \quad \beta_d = .060$$

$$\dot{\beta}_o = .038, \quad \dot{\beta}_I = .034, \quad \dot{\beta}_{II} = .03, \quad \dot{\beta}_d = -.030$$

β_d refers to the reactionless mode of a four-bladed
rotor, β_o is the coning angle, β_I represents forward
tilt and β_{II} left tilt. For the constant coefficient
approximation β_d does not couple with the other
states and can be omitted. The multiblade equations
for a four-bladed rotor are given in reference 9.
The exact and the noise polluted responses β_I, β_{II}
are shown in figure 4.

Using the above initial conditions one obtains with
the maximum likelihood method the following values

Parameter		γ	δ
True Value		5.00	10.00
Initial Estimate		2.50	5.00
Iteration	1	4.97	8.46
(4 CPU seconds	2	5.00	9.62
per iteration)	3	4.99	9.65

The lower bound for the parameter covariance in the 3rd iteration is

$$M^{-1} = \begin{bmatrix} .125 & .003 \\ .003 & .269 \end{bmatrix} \text{, } R = I$$

The result as compared to the preceding single blade identification is better for γ and worse for δ. Note the assumed large errors of 50% in the initial parameter estimate.

In order to assess the effect of errors in the initial states at $t = 12$, it is now assumed that all initial states are zero. One then obtains:

$$\beta_O = \dot{\beta}_O = \beta_I = \dot{\beta}_I = \beta_{II} = \dot{\beta}_{II} = 0 \quad \text{at } t = 12$$

Parameter		γ	δ
True value		5.00	10.00
Initial estimate		2.50	5.00
Iteration	1	5.68	9.45
(4 CPU seconds per iteration)	2	4.87	10.27
	3	4.92	10.26

For R = I the inverted information matrix is

$$M^{-1} = \begin{bmatrix} .151 & .043 \\ .043 & .278 \end{bmatrix}$$

The result shows that large errors in the initial states can be tolerated.

Finally the 3 initial deflections have been assumed as unknown parameters, increasing the number of parameters to be identified from 2 to 5. The initial values for the response rates were assumed to be zero: $\dot{\beta}_0 = \dot{\beta}_I = \dot{\beta}_{II} = 0$ at t = 12. The initial values of β_0, β_I, β_{II} are given in the table:

Parameter		γ	δ	$\beta_0(12)$	$\beta_I(12)$	$\beta_{II}(12)$
True values		5.00	10.00	.880	-.983	1.698
Initial estimates		2.50	5.00	1.00	-1.00	2.00
Iteration (6 CPU seconds per iteration)	1	4.94	8.25	.979	-1.039	1.926
	2	5.01	9.97	.920	-.985	1.732
	3	5.00	9.93	.920	-.984	1.730

The inverted information matrix for γ and δ is:

$$M^{-1} = \begin{bmatrix} .133 & -.006 \\ -.006 & .304 \end{bmatrix}, \quad R = I$$

It is seen that now the convergence is much faster, the second iteration being even better than previously the 3rd iteration. However, the CPU time per iteration is increased from 4 seconds to 6 seconds.

The following table compares the results of the various methods. The last 4 rows refer to the maximum likelihood method.

	Iterations	γ	δ	Variances of γ	δ	Total CPU time, sec.
True Value		5.00	10.00			
Iterated equation error, single blade	3	4.52	10.09	.073	.111	4.8
Single blade, correct initial conditions	2	4.91	9.82	.0085	.0146	5.2
Single blade, identified initial conditions	2	4.90	9.85	.0092	.0162	8.6
Multiblade, zero initial values	2	4.87	10.27	.0060	.0111	8.0
Multiblade, zero initial rates identified initial deflections	2	5.01	9.97	.0053	.0210	12.0

The number of iterations indicated in the table is that for which convergence has been achieved. The iterated equation error estimation with updated Kalman filter

needs the lowest total computer effort, however, the
accuracy of the estimate is worst for γ. The maximum
likelihood estimation, due to faster convergence, needs
only moderately more computer effort and yields better
accuracy. For the single blade the identification of
the two initial states, $\beta(12)$ and $\dot{\beta}(12)$, results in the
same good accuracy of the parameters as when using
the correct initial conditions.

The multiblade identification - despite modeling
errors by omitting all periodic terms in the equations
of motion and despite assuming zero initial con-
ditions - also converges rapidly and provides only
slightly less accuracy of the parameters. The multi-
blade analysis with identification of the 3 initial
deflections gives the best accuracies, however with
more than twice the computer effort. Note that the
convariance estimate of the Kalman filter method is an
order of magnitude greater than for the maximum
likelihood method. The latter values have been obtained
from M^{-1} for R = I by multiplication with .04, since

$$R = \begin{bmatrix} .04 & 0 \\ 0 & .04 \end{bmatrix}$$ They represent lower bounds of the

actual parameter covariances but agree fairly well with
the errors found for the identified parameters. This
is not true for the covariance estimates of the Kalman

filter method. The latter method was attempted
also for the multiblade representation but was
found to require impractically long computer runs
per iteration combined with a lack of convergence.
The multiblade identification studies were, therefore,
limited to the maximum likelihood method.

DYNAMIC WAKE PARAMETER IDENTIFICATION FOR ZERO ADVANCE RATIO

For the case of zero advance ratio reference 10
gives a mathematical model of the dynamic rotor
wake that has 2 constants: The quasi-steady wake
number L that relates the pitching and rolling
moments of the rotor to the sine and cosine components
of the wake velocity, and a time constant τ. The
third rotor constant is

$$A = B^4 \gamma/8 \tag{27}$$

where B is the tip loss factor and γ the blade Lock
number. The inflow angle from the rotor wake is
assumed constant over the radius and represented
for the kth blade by

$$\lambda_k = \lambda_o + \lambda_I \cos \psi_k + \lambda_{II} \sin \psi_k \tag{28}$$

In a linearized analysis for zero advance ratio the
coning mode is uncoupled from the tilting modes. The
system equations of reference 10 are in matrix notation

$$\begin{bmatrix} \ddot{\beta}_I \\ \ddot{\beta}_{II} \\ \dot{\lambda}_I \\ \dot{\lambda}_{II} \end{bmatrix} = \begin{bmatrix} -(\omega_1^2-1) & -A & -A & -2 & A & 0 \\ A & 2 & -(\omega_1^2-1) & -A & 0 & A \\ 0 & AL/\tau & AL/\tau & 0 & -(AL+1)/\tau & 0 \\ -AL/\tau & 0 & 0 & AL/\tau & 0 & -(AL+1)/\tau \end{bmatrix} \begin{bmatrix} \beta_I \\ \dot{\beta}_I \\ \beta_{II} \\ \dot{\beta}_{II} \\ \lambda_I \\ \lambda_{II} \end{bmatrix}$$

$$+ \begin{bmatrix} 0 & A \\ -A & 0 \\ 0 & -AL/\tau \\ AL/\tau & 0 \end{bmatrix} \begin{bmatrix} \Theta_I \\ \Theta_{II} \end{bmatrix} \tag{29}$$

There are 6 state variables: β_I, $\dot{\beta}_I$, β_{II}, $\dot{\beta}_{II}$, λ_I, λ_{II}.

The flapping angles of the 4 blades β_1, β_2, β_3, β_4 are measured in the rotating system and combined into $\beta_1 - \beta_3$ and $\beta_2 - \beta_4$. The measurement equations - that is the relation between the measured quantities and the system state variables - are given by

$$\begin{bmatrix} \beta_{m1} \\ \beta_{m2} \end{bmatrix} = \begin{bmatrix} (\beta_1-\beta_3)/2 \\ (\beta_2-\beta_4)/2 \end{bmatrix} = \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix} \begin{bmatrix} \beta_I \\ \beta_{II} \end{bmatrix} + \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} \tag{30}$$

where η_1, η_2 are measurement noise of β_{m1} and β_{m2} respectively. Thus only 2 of the 6 state variables

are measured. The parameters to be identified are A, AL/τ, 1/τ, and the blade natural frequency in the rotating system is assumed to be given:

$\omega_1 = 1.2.$

The pitch stirring input is again described by equations 4 to 6 (progressing mode). The rotor wake causes an increased stabilization time. Therefore $t_o = 70$ was selected instead of $t_o = 12$. The measurement noise n_1, n_2 is assumed to be white and Gaussian with the standard deviations

$$\sigma_{\beta_{m1}} = \sigma_{\beta_{m2}} = .10$$

The reduction to one half the noise amplitude as compared to the preceding examples is justified by the fact that β_{m1} and β_{m2} are substantially smaller than β_I and β_{II} see figures 4 and 5.* The true values of the 3 parameters for which the response was determined are

$$A = .500 \qquad AL/\tau = .250 \qquad 1/\tau = .125$$

These values are in the range expected to be found for the model rotor. They will depend essentially on the collective pitch setting. In preliminary

---

*Note that figure 5 refers to $\sigma_{\beta m} = .05$

computer runs it was found that the initial estimate
of 1/τ should not be zero.

Using the true initial conditions one obtains
with the maximum likelihood method

Parameter		Λ	$\Lambda L/\tau$	$1/\tau$
True value		.500	.250	.125
Initial estimate		.400	.200	.250
Iteration (3.8 CPU	1	.500	.290	.086
seconds per	2	.486	.253	.125
iteration)	3	.487	.255	.128
	4	.487	.255	.128

With the inverted information matrix for the 4th
iteration

$$
M^{-1} = \begin{bmatrix} .052 & .059 & .013 \\ .059 & .094 & .023 \\ .013 & .023 & .010 \end{bmatrix}, \quad R = I
$$

Despite the fact that only 2 out of 6 state variables
are measured, the parameter identification is very
good and the second iteration has almost converged.

The following case is the same as before except
for regressing excitation (+ω) instead of progressing
excitation (-ω)

Parameter		A	AL/τ	1/τ
True Value		.500	.250	.125
Initial Estimate		.400	.200	.250
Iteration	1	.524	.262	.087
(3.8 CPU	2	.540	.264	.134
seconds per	3	.536	.264	.139
iteration)	4	.536	.264	.139

with

$$M^{-1} = \begin{bmatrix} .167 & .179 & .029 \\ .179 & .256 & .045 \\ .029 & .045 & .013 \end{bmatrix}, \quad R = I$$

Though convergence is good and the second iteration
has almost converged, the errors are now larger
than before as expressed also by the M^{-1} matrix.
The physical reason why regressing modes are less
suited for rotor wake identification is the fact
that at $\omega = .2$ the excitation is in resonance with
the regressing flapping mode. At this condition
no dynamic rotor wake exists since aerodynamic
excitation and aerodynamic damping cancel each other.
Since regressing mode transients include a frequency
region with a weak dynamic rotor wake, the identi-
fication of the wake parameters is less good than for
progressing mode transients.

Finally the same case is treated with very large errors in the initial parameter estimates. Now the 4 initial displacements are identified in addition to the 3 parameters, resulting in 7 instead of 3 unknown parameters. The initial rates are assumed to be zero: $\dot{\beta}_I = \dot{\beta}_{II} = \dot{\lambda}_I = \dot{\lambda}_{II} = 0$ at $t = 70$. Using progressing excitation one obtains

$$\sigma_{\beta m1} = \sigma_{\beta m2} = .10$$

Parameter	A	AL/τ	1/τ	$\beta_I(70)$	$\beta_{II}(70)$	$\lambda_I(70)$	$\lambda_{II}(70)$
True Value	.500	.250	.125	.497	.188	-.874	-.331
Initial Estimate	.400	.067	.083	.601	.305	-.531	-.268
Iteration 1	.443	.184	.150	.500	.211	-.864	-.309
(6.3 CPU seconds 2	.471	.233	.115	.497	.210	-.859	-.322
per iteration) 3	.474	.235	.119	.497	.211	-.870	-.327
4	.474	.236	.120	.497	.211	-.871	-.327
M^{-1} diagonal 4th iteration, R = I	.121	.184	.055	.088	.095	1.069	.796

The convergence is again good and the second iteration is almost converged. The errors in the 3 parameters are somewhat larger than before but still acceptable. The errors in the initial conditions, except for β_{II}, are also small, though the M^{-1} matrix shows larger

covariances for λ_I and λ_{II}. The initial estimate of
the initial displacements $\beta_I, \beta_{II}, \lambda_I, \lambda_{II}$ at t = 70
are not arbitrary but are determined from the
equilibrium equations before the beginning of pitch
stirring, using the initial estimates of A, AL/τ
and 1/τ. Since the true initial conditions for λ_I
and λ_{II} are unknown, their inclusion in the identi-
fication is necessary. β_I and β_{II} are measured, thus
their initial conditions are known within the accuracy
of the measurement and their inclusion in the
identification may not be required if the measuring
error is small. The computer time for identifying
all 4 initial displacements ($\beta_I, \beta_{II}, \lambda_I, \lambda_{II}$ at t = 70)
is however quite moderate, so that their identi-
fication is practical.

On the basis of the computer experiments the
method to be used for the rotor model tests at
zero advance ratio is then as follows:

1. Estimate the parameters A, AL/τ, 1/τ

2. Determine from the steady state equations 29
   with zero rates and zero accelerations the
   initial values for $\beta_I, \beta_{II}, \lambda_I, \lambda_{II}$ at time t = t_o,
   using the estimated parameters.

3. Assume zero initial rates:

   $$\dot{\beta}_I = \dot{\beta}_{II} = \dot{\lambda}_I = \dot{\lambda}_{II} = 0 \quad \text{at} \quad t = t_o$$

4. Apply the maximum likelihood algorithm to
equation 29 and determine iteratively the 7
parameters A, AL/τ, 1/τ, $\beta_I(t_o)$, $\beta_{II}(t_o)$, $\lambda_I(t_o)$,
$\lambda_{II}(t_o)$.

SIMULTANEOUS MEASUREMENT AND SYSTEM NOISE

Though the simplified maximum likelihood
algorithm as currently used in aircraft parameter
identification work and as defined here by equations
18 to 25 does not provide for system noise, such
noise is unavoidable in full scale and wind tunnel
tests. The origin of the system noise is either
in atmospheric or wind tunnel flow turbulence, or
in modeling errors. In order to assess the detri-
mental effects of the system noise on the quality
of the maximum likelihood estimates, the preceding
example was recomputed under the assumption that
equation 29 has an additional noise term on the
right hand side of

$$
\begin{bmatrix}
A & 0 \\
0 & A \\
-(AL + 1)/\tau & 0 \\
0 & -(AL + 1)/\tau
\end{bmatrix}
\cdot
\begin{bmatrix}
V_1 \\
V_2
\end{bmatrix}
\tag{29a}
$$

This means that we have noisy inflow components λ_I, λ_{II}.
For V_1, V_2 a zero mean computer generated Gaussian

sequence with standard deviations $\sigma_{V_1} = \sigma_{V_2} = .1$ was

assumed. Figure 6 shows the noise polluted inflow

components λ_I, λ_{II}. In addition noise was added

as before in the measurement equations 30, however

now with standard deviations $\sigma_{\beta m1} = \sigma_{\beta m2} = .05$ which

is one half the previously assumed value and which is

more representative of the expected measurement errors.

We first determine the effect of the lower

measurement noise on the parameter identification:

$$\sigma_{\beta m1} = \sigma_{\beta m2} = .05, \quad \sigma_{V_1} = \sigma_{V_2} = 0$$

Parameter		A	AL/τ	$1/\tau$	$\beta_I(70)$	$\beta_{II}(70)$	$\lambda_I(70)$	$\lambda_{II}(70)$
True Value		.500	.250	.125	.497	.188	-.874	-.331
Initial Estimate		.400	.067	.083	.601	.305	-.531	-.268
Iteration	1	.455	.189	.159	.504	.195	-.880	-.330
	2	.483	.238	.119	.500	.194	-.867	-.339
	3	.485	.239	.123	.501	.194	-.874	-.341
	4	.485	.239	.123	.501	.194	-.874	-.341

It is seen that the accuracy of the parameters is

improved as compared to the preceding case of

$\sigma_{\beta m1} = \sigma_{\beta m2} = .10$, as it should be.

Combining now the measurement and system noise one obtains:

$$\sigma_{\beta m1} = \sigma_{\beta m2} = .05, \quad \sigma_{V_1} = \sigma_{V_2} = .10$$

Parameter		A	AL/τ	1/τ	$\beta_I(70)$	$\beta_{II}(70)$	$\lambda_I(70)$	$\lambda_{II}(70)$
True Value		.500	.250	.125	.497	.188	-.874	-.331
Initial Estimate		.400	.067	.083	.601	.305	-.531	-.268
Iteration	1	.455	.189	.153	.505	.188	-.860	-.319
	2	.482	.236	.115	.502	.188	-.847	-.324
	3	.483	.235	.118	.502	.190	-.853	-.325
	4	.483	.236	.118	.502	.190	-.854	-.325

It is seen that the system noise indicated in figure 6 has only a small detrimental effect on the accuracy of the identified parameters. Of all computer experiments of parameter identification we have shown here only the results of one computer run. Since these results are dependent on the computer generated random sequences, we are actually dealing with identified parameters that are random variables. A repeat run will, however show only minor differences in the identified parameters, so that the presented results in comparison to the true values can be considered as typical. The parameter that suffers most in its

accuracy when system noise is added is the reciprocal

time constant $1/\tau$. Even so the error is only 5.5%

which would appear quite acceptable for all

practical purposes.

CONCLUSIONS

The computer experiments presented here can be

used to establish preferences with respect to various

alternatives in the rotor parameter identification

schemes.

1. Iterated Equation Error Estimation with Updated

Kalman Filter vs. Maximum Likelihood Method

For single blade parameter identification from

pitch stirring transients the equation error

method applied in an iterative form using a

Kalman filter with the latest parameter updates

worked well and required the least computer CPU

time. For multiblade parameter identification

this method became impractical because of slow

convergence and high computer CPU time. The

maximum likelihood method worked well both for

single blade and multiblade applications, though in

case of single blade identification it requires

somewhat more computer CPU time. The parameter

covariances from the maximum likelihood method are

clearly superior to and more meaningful than the
covariances determined with the equation error
method. The maximum likelihood method also gave
good parameter identifications in the presence
of both measurement and system noise, though
most of the computer experiments were conducted
with measurement noise only. Overall one can
conclude that the maximum likelihood method in
its simplified form in which system noise is
not modeled, is for the applications studied
superior to the equation error method and thus
will represent the method of choice for the
parameter identification from wind tunnel rotor
model tests.

2. Fixed vs. Identified Initial State

Since prior to the initiation of pitch stirring
the rotor is approximately in a steady condition
as far as the multiblade coordinates are concerned,
one can always assume their initial rates as zero.
For two parameter identification in forward flight
($\mu = .4$) the use of zero initial values for all
states gave good identification results. Including
the initial displacements in the identification -
while retaining initial zero rates of displacement -

greatly improved the convergence and the
accuracy, though at a cost of about twice the
computer CPU time per iteration. For the mul-
tiblade analysis it appears nevertheless
practical to identify the initial displacements
in order to improve convergence and accuracy
of the parameters. For the single blade analysis
the initial value of the flapping rate is not
zero and must be identified.

3. <u>Single Blade vs. Multiblade Analysis</u>

If rotor wake lag is included, only a multiblade
analysis is possible. Without rotor wake lag and
using an "equivalent" Lock number to describe
the wake effects, a single blade analysis is
possible. Initial flapping deflection and rate
of deflection should be identified. The multi-
blade analysis with zero initial rates and
identified initial deflections is more accurate
though more demanding in computer CPU time per
iteration. The multiblade analysis is thus
preferable also in those cases where a single
blade analysis is possible.

4. <u>Progressing vs. Regressing Pitch Stirring</u>

Progressing pitch stirring leads to more accurate
rotor wake parameter identification than regressing

pitch stirring, because the regressing stirring
transient passes through the flapping resonance
at which no dynamic rotor wake exists.

References

1. Hohenemser, K. H. and Yin, S. K., "Computer Experiments in Preparation of System Identification from Transient Rotor Model Tests," Part II of First Yearly Report under Contract NAS2-7613, June 1974.

2. Chen, R. T. N., Eulrich, B. J. and Lebacqz, J. V., "Development of Advanced Techniques for the Identification of V/STOL Aircraft Stability and Control Parameters,"Cornell Aeronautical Laboratory Report No. BM-2820-F-1, August 1971.

3. Molusis, J. A., "Helicopter Stability Derivative Extraction from Flight Data Using the Bayesian Approach to Estimation," Journal of the American Helicopter Society, Vol. 18, No. 2, April 1973, pp. 12-23.

4. Ormiston, R. A. and Peters, D. A., "Hingeless Rotor Response with Non-Uniform Inflow and Elastic Blade Bending," Journal of Aircraft, Vol. 9, No. 10, October 1972, pp. 730-736.

5. Peters, D. A., "Hingeless Rotor Frequency Response with Unsteady Inflow," Proceedings Specialists Meeting on Rotorcraft Dynamics, AHS/NASA Ames, Moffett Field, California, February 1974, NASA-SP-352.

6. Stepner, D. E. and Mehra, R. K., "Maximum
    Likelihood Identification and Optimal Input
    Design for Identifying Aircraft Stability and
    Control Derivatives," NASA CR-2200, March 1973.

7. Hohenemser, K. H. and Prelewicz, D. A., "Com-
    puter Experiments on Periodic Systems Identi-
    fication Using Rotor Blade Transient Flapping-
    Torsion Responses at High Advance Ratio," NASA
    SP-352, 1974.

8. Bryson, A. E. Jr. and Ho, Y. C., "Applied
    Optimal Control," Ginn & Co., 1969, Chapter 12.

9. Hohenemser, K. H. and Yin, S. K., "Some Appli-
    cations of the Method of Multiblade Coordinates,"
    Journal American Helicopter Society, Vol. 17,
    No. 3, July 1972, pp. 3-12.

10. Crews, S. T., Hohenemser, K. H. and Ormiston, R. A.,
    "An Unsteady Wake Model for a Hingeless Rotor,"
    Journal of Aircraft, Vol. 10, No. 12, Dec. 1973,
    pp. 758-760.

Chapter 2

FIGURE CAPTIONS

Figure 1   Time History of Blade Pitch, Rotating
            Reference System, Progressing Mode

Figure 2   Time History of Blade Pitch, Multiblade
            Representation, Progressing Mode

Figure 3   Estimate $\hat{\gamma}$ and its Variance for
            $P(0) = 2000$, $\hat{\gamma}(0) = 0$, $\sigma_\beta = .2$,
            $\Theta_o = 0$, $\mu = .4$, $\omega_1 = 1.2$, Progressing
            Excitation

Figure 4   Exact and Noise Polluted Multiblade
            Responses β_I, β_{II} to Inputs Shown
            in Figure 2 for $\mu = .4$, $\Theta = 2°$,
            $\gamma = 5.00$, $\omega_1 = 1.20$, $\sigma_{\beta_o} = \sigma_{\beta_I} = \sigma_{\beta_{II}} = .2$

Figure 5   Exact and Noise Polluted Multiblade
            Responses β_{m1}, β_{m2} to Inputs Shown in
            Figure 2 for $\mu = 0$, $A = .500$,
            $AL/\tau = .250$, $1/\tau = .125$, $\omega_1 = 1.20$,
            $\sigma_{\beta m1} = \sigma_{\beta m2} = .05$

Figure 6   Exact and Noise Polluted Inflow Components
            λ_I, λ_{II} (system noise) for $\mu = 0$,
            $A = .500$, $AL/\tau = .250$, $1/\tau = .125$,
            $\omega_1 = 1.20$, $\sigma_{\lambda_I} = \sigma_{\lambda_{II}} = .10$
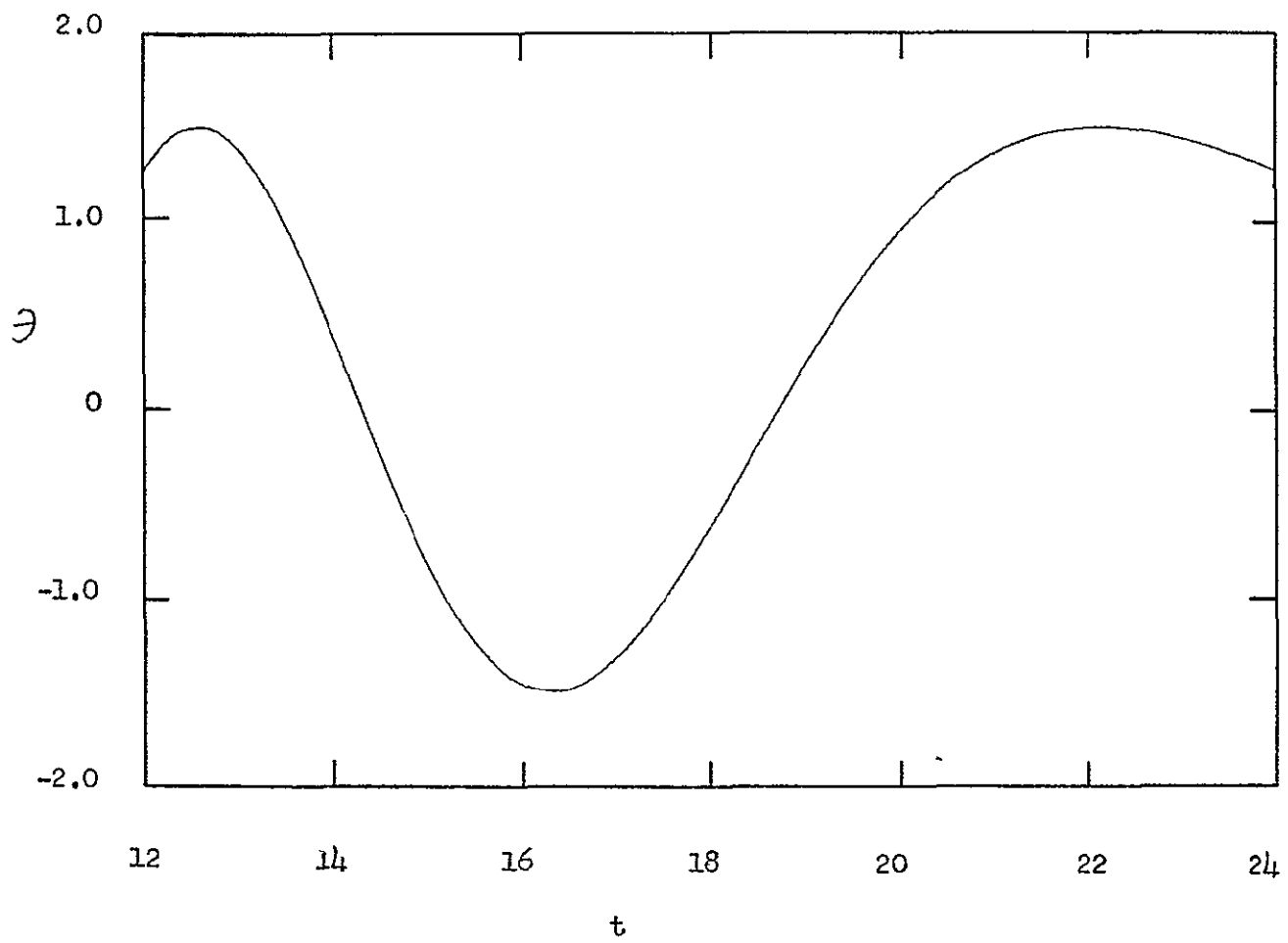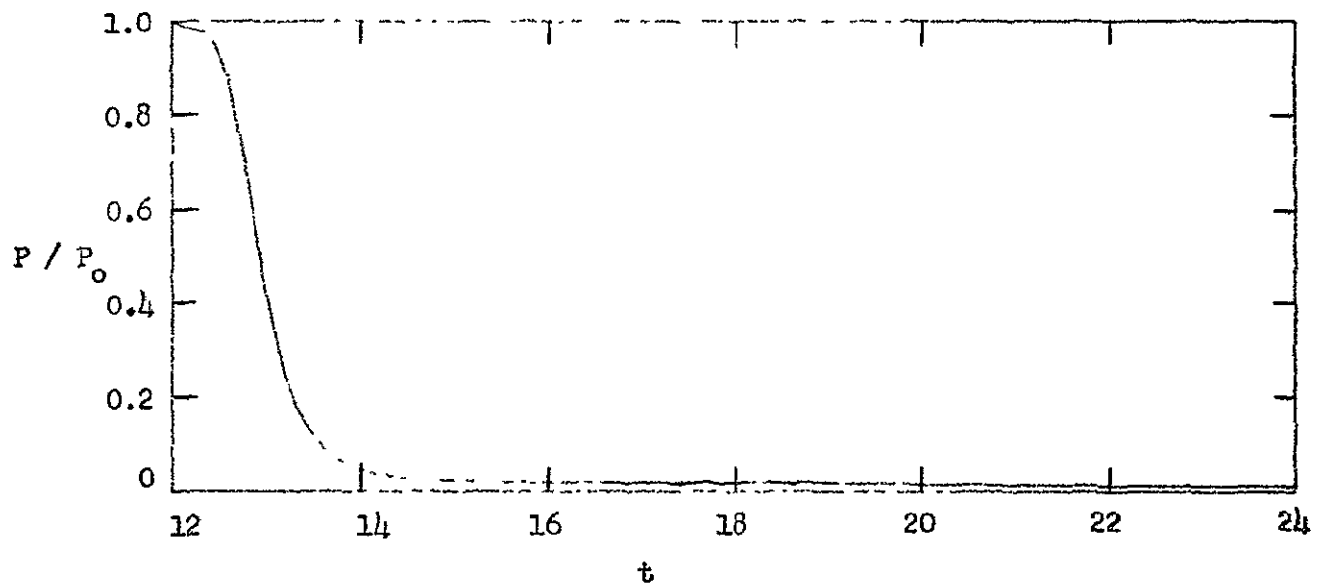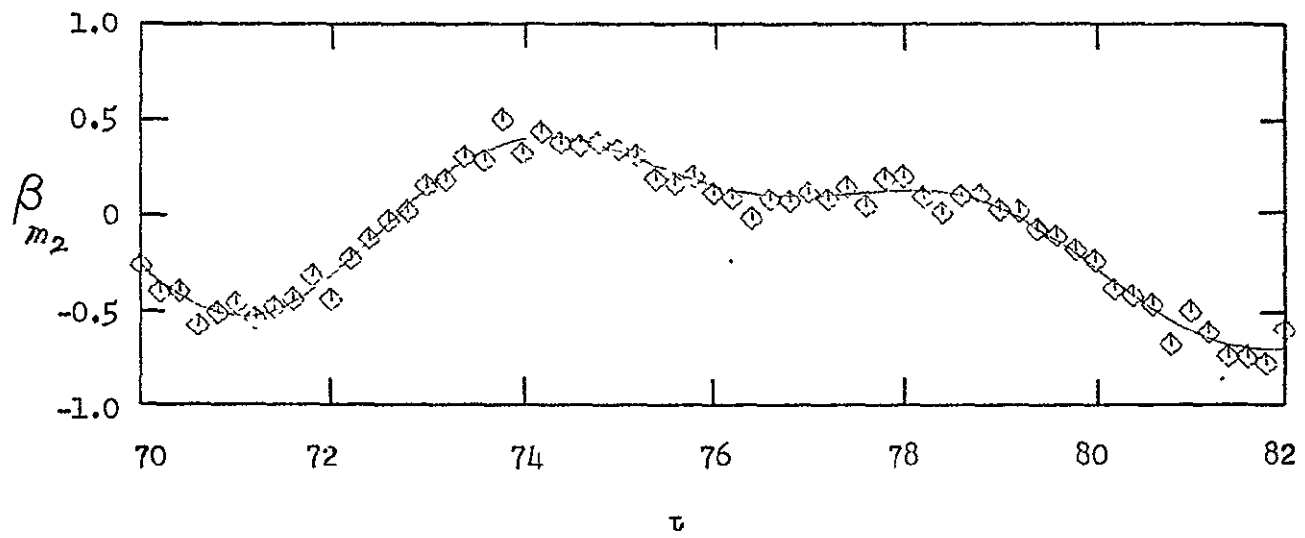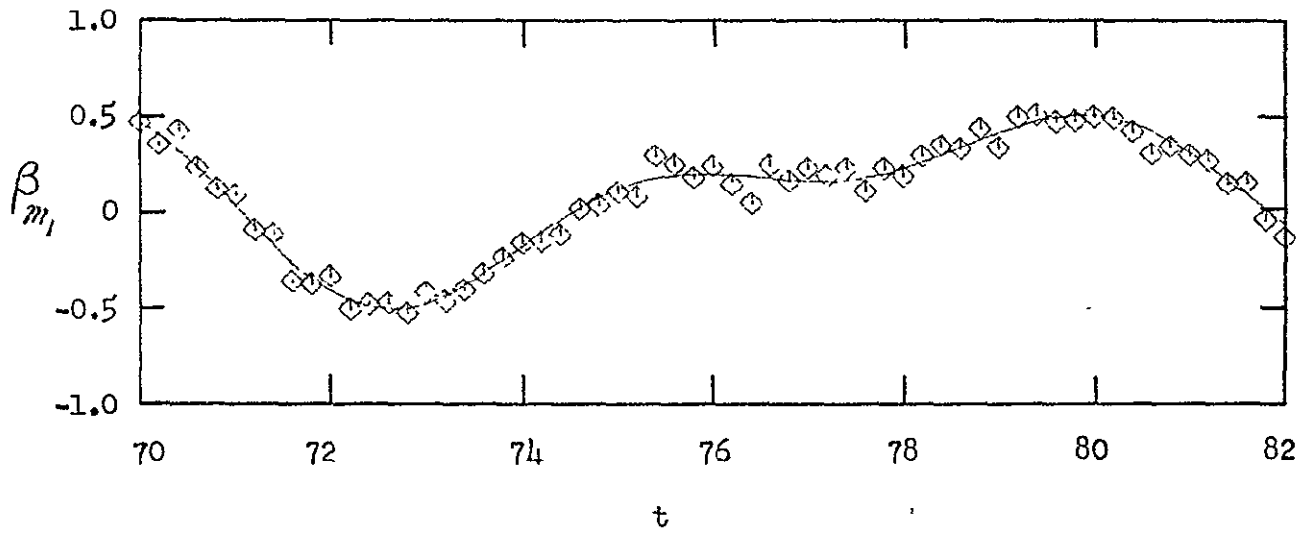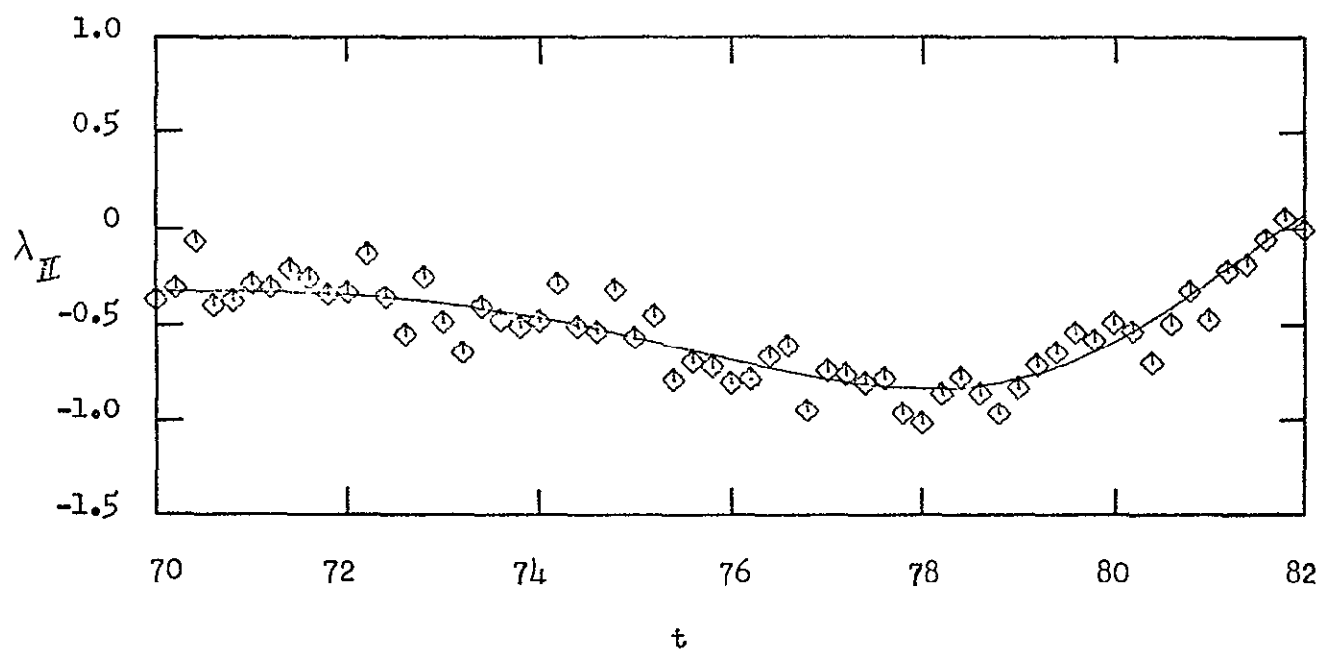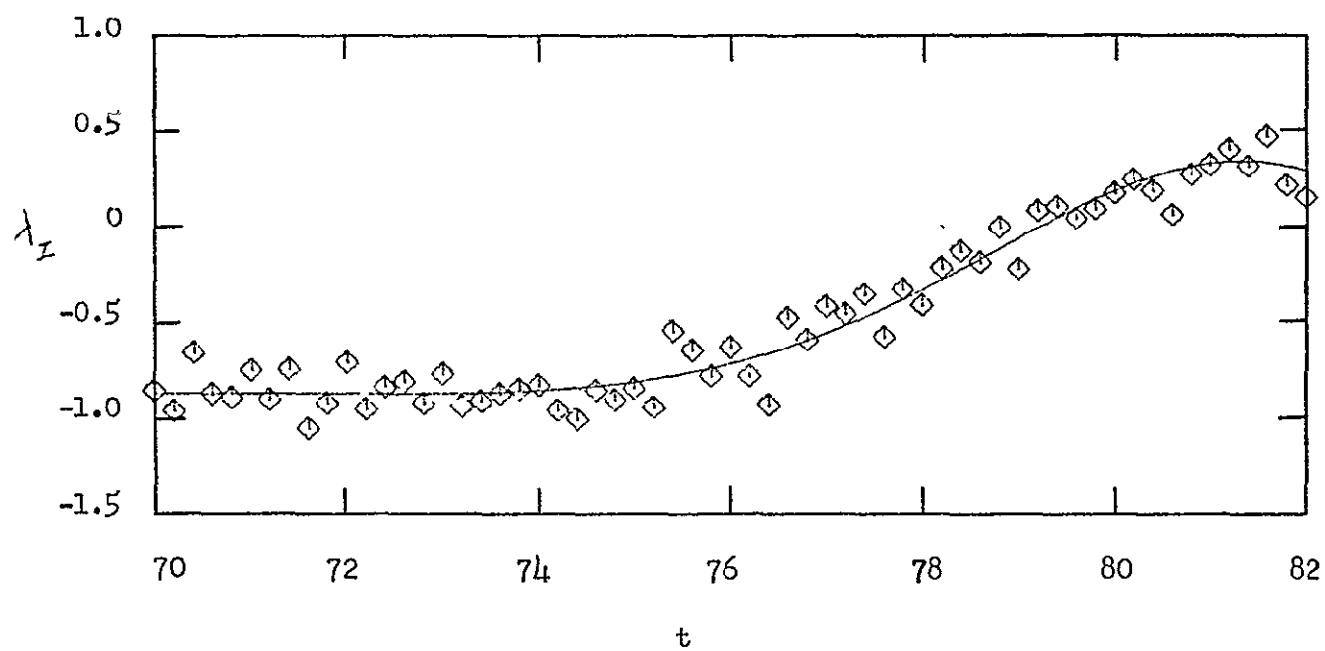
Fig. 1
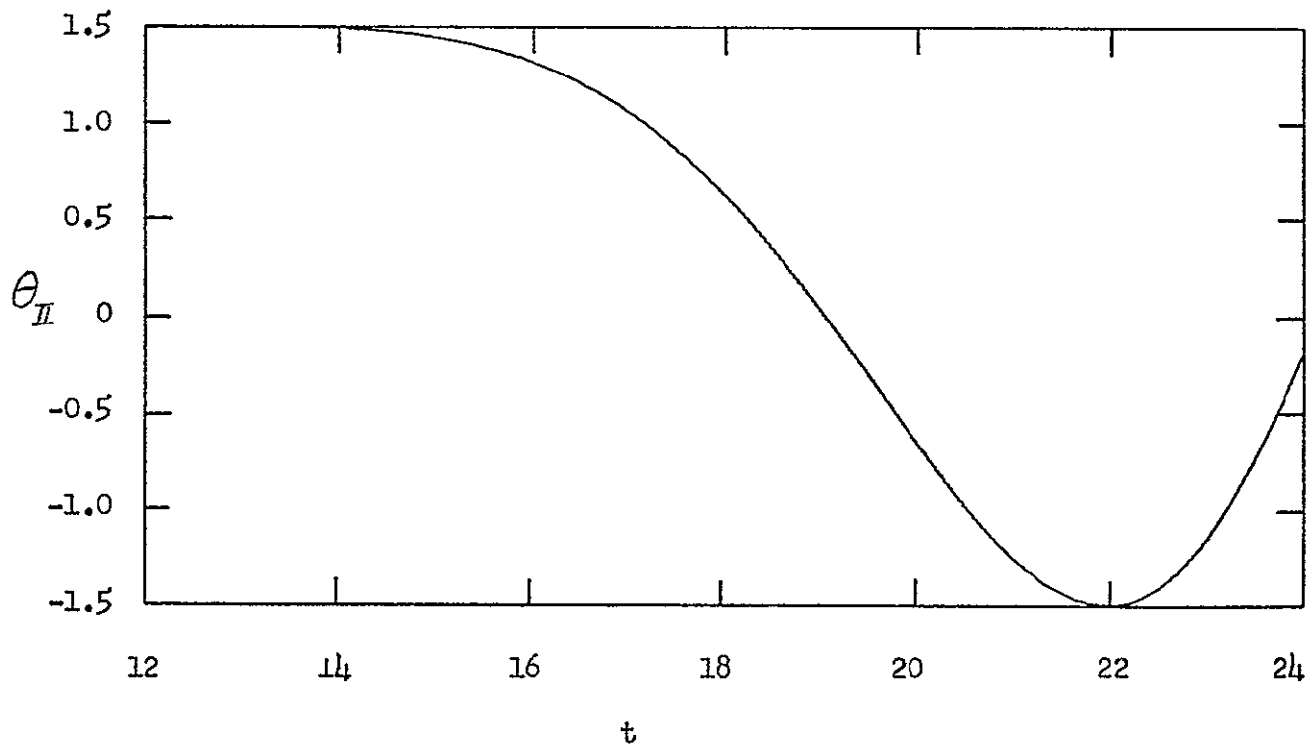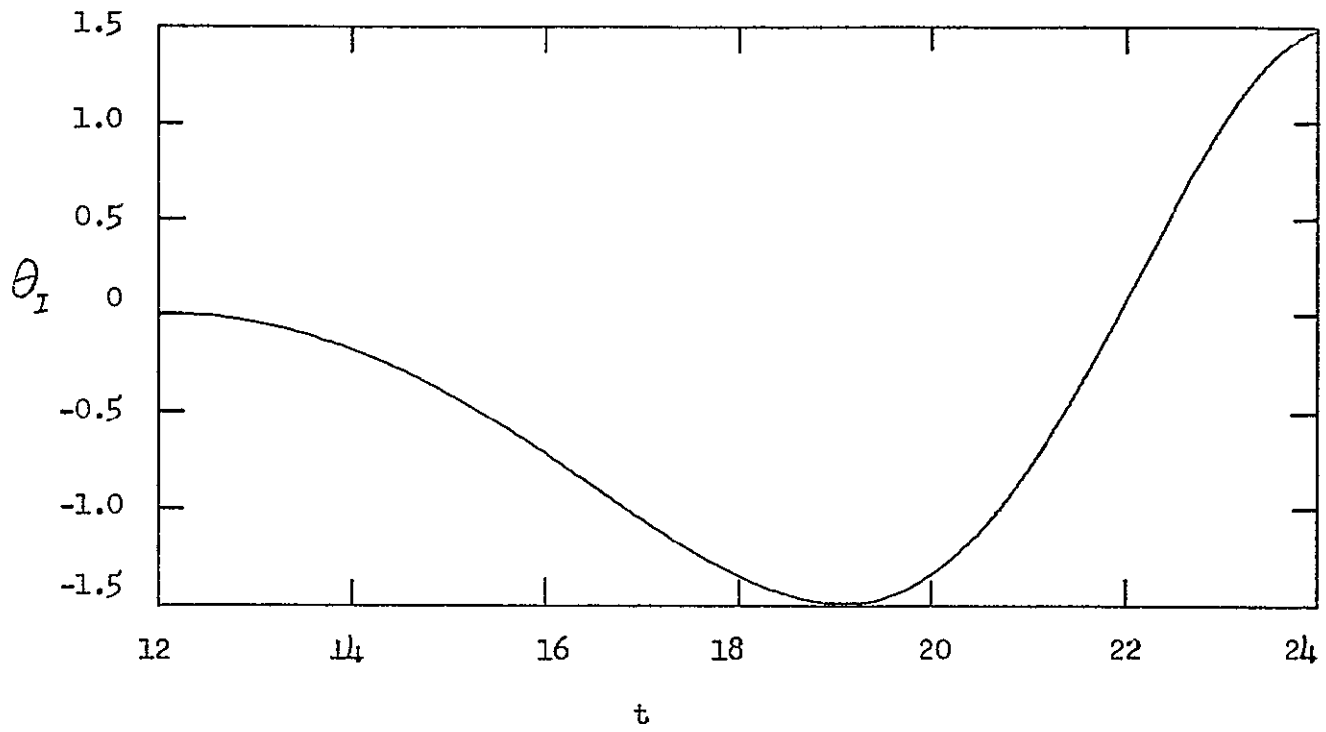
Fig. 2
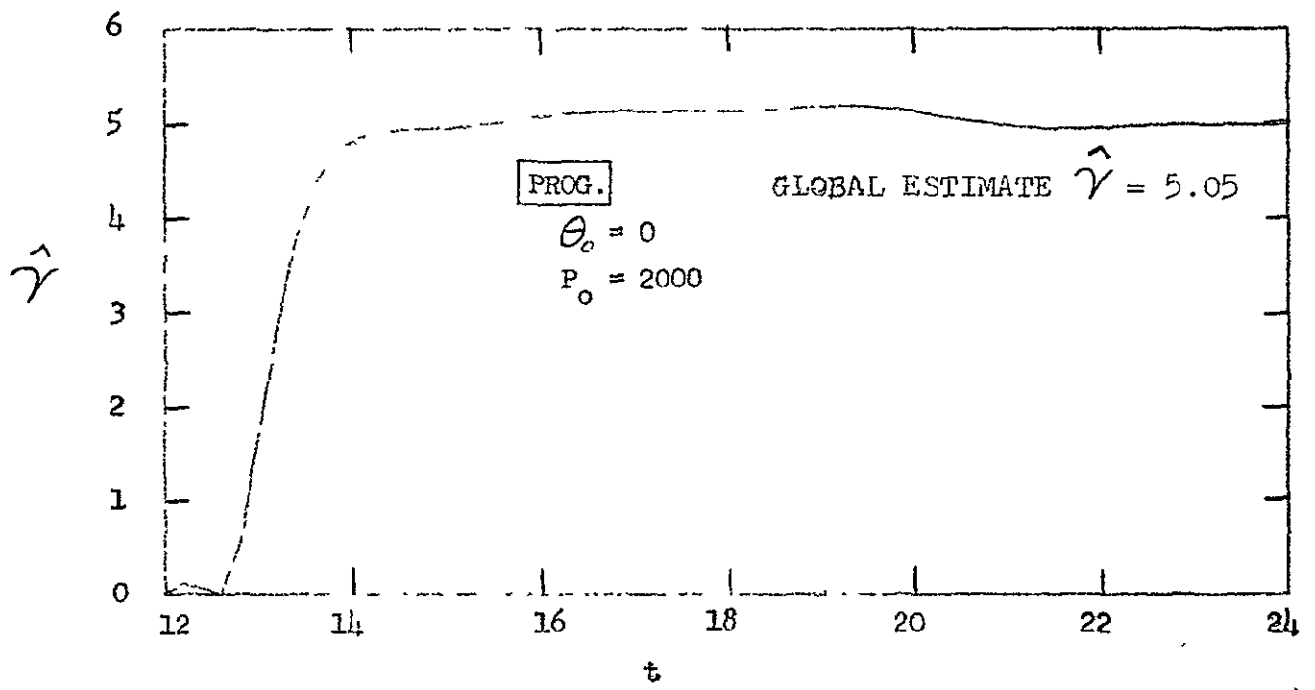
Fig. 3
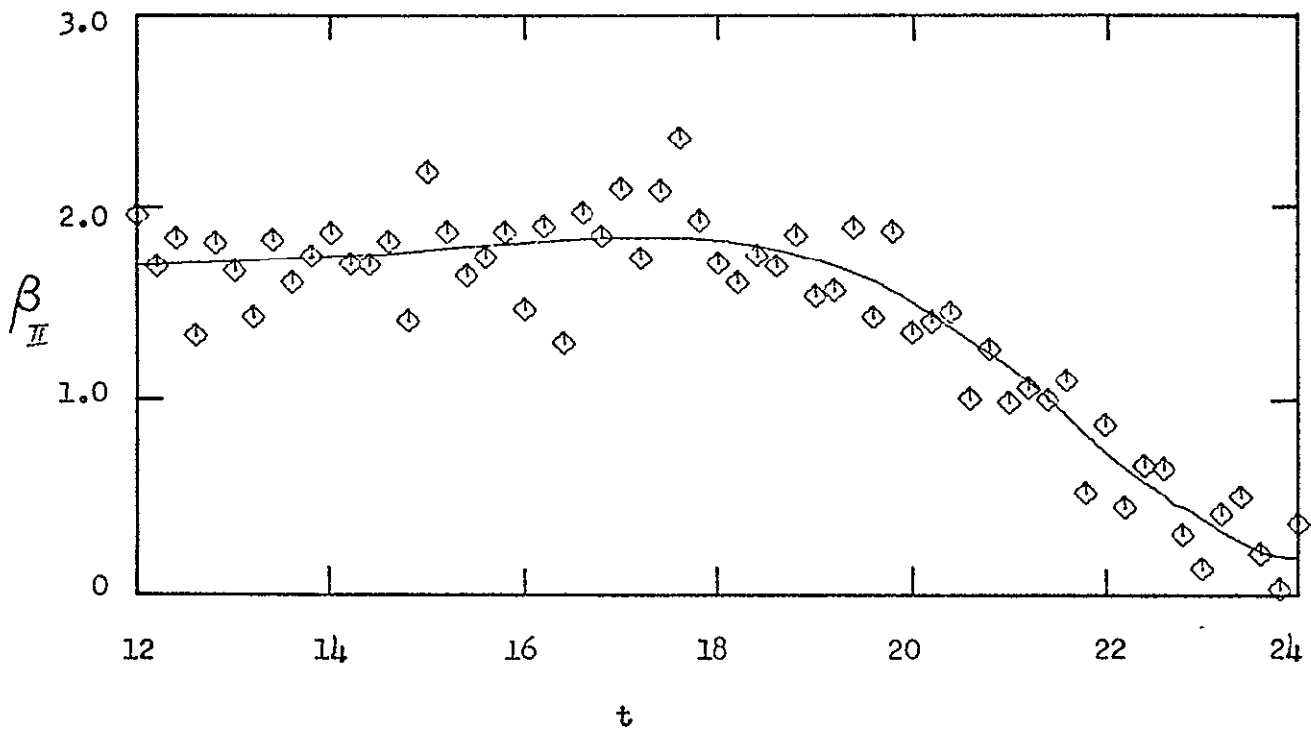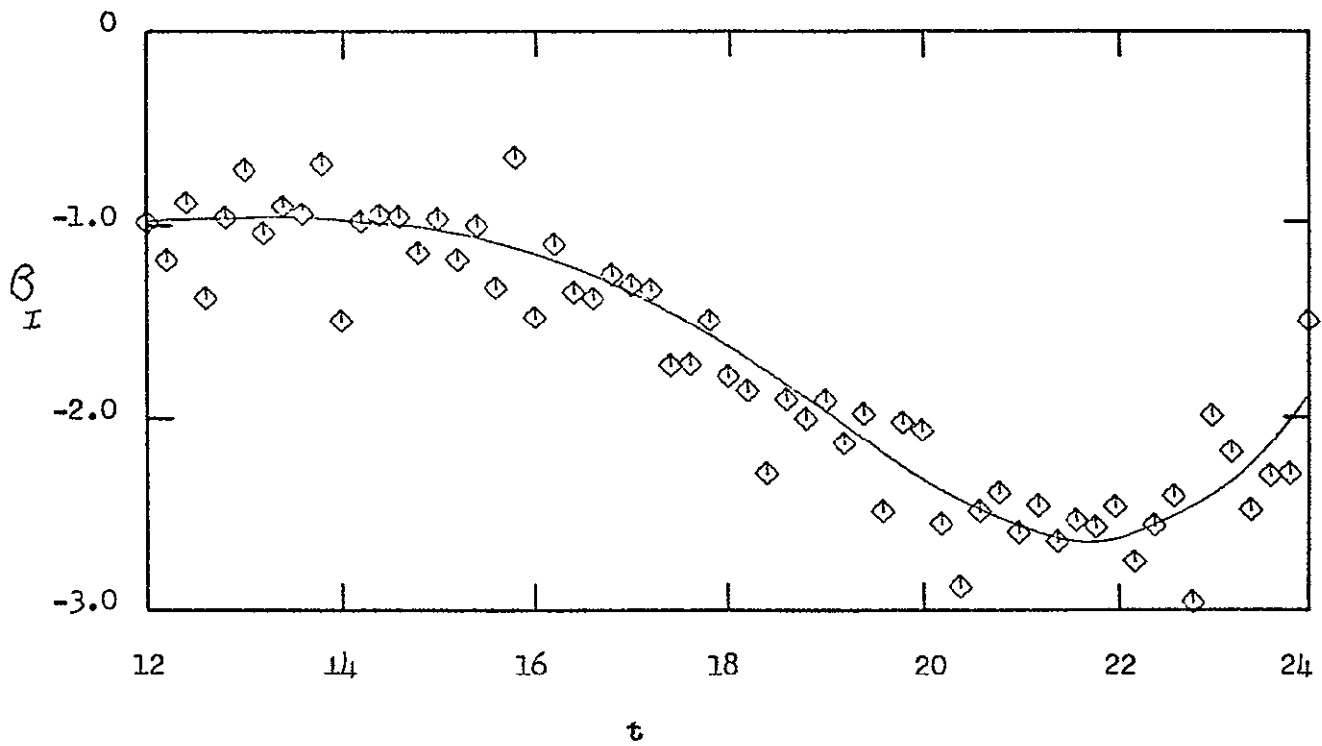
Fig. 4

Fig. 5

Fig. 6

CHAPTER 3

OPTIMAL DATA UTILIZATION FOR PARAMETER IDENTIFICATION

PROBLEMS WITH APPLICATION TO LIFTING ROTORS

K. H. Hohenemser and D. Banerjee

ABSTRACT

A method is developed for optimal data utilization
in the maximum likelihood identification of systems
without process noise, based on stipulated upper bounds
of parameter covariances. The method is applied to a
case of simulated transient wind tunnel testing of
lifting rotors.

TABLE OF CONTENTS

INTRODUCTION

Methods for state and parameter estimation from transients are widely used in aircraft testing.[1,2,3,4] The problem is to obtain optimum estimates (based on certain performance criteria) of initial states and of unknown parameters (derivatives) from noisy measurements of some inputs and response variables. In most cases of airplane parameter identification a constant system is used as an analytical model. For lifting rotor applications a periodic system model is required.[5] State and parameter identification from transients looks promising also for wind tunnel testing and may well drastically reduce the amount of test efforts as compared to e.g., frequency response testing of lifting rotors. The following study was performed in preparation for transient wind tunnel tests with a lifting rotor model.

In the most general case the identification method has to account for the following types of uncertainties:

(a)   Modeling errors which originate from the difference
      between the mathematical model and the actual
      physical phenomenon.

(b)   Uncertainties from input noise, e.g., turbulence
      (in the atmosphere or in the wind tunnel).

(c)   Uncertainties from measurement inaccuracies. This
      includes measurement bias and noise and also incom-
      plete measurements when some of the input and
      output data are missing.

Three basic methods have been applied to the air-
craft identification problem. If the measurements are
a linear function of the unknown parameters, the
classical least squares regression or equation error
method is applicable. This method was widely used in
the early stages of aircraft parameter identification.
For the application of this method the time histories of
all the states and their derivatives are needed together
with the input variables. Since the equation error
method becomes less reliable the noisier the measure-
ments are, it has been replaced in the last decade by
less restrictive methods that allow considerable
measurement noise and that work also when some states
have not been measured. One of the most widely used is
the maximum likelihood method, which, in the absence of
modeling errors and input noise, reduces to what is also

called the output error method.[1,2,4] Another way of
overcoming the limitations of the least squares
regression method is the application of the extended
Kalman filter.[3] This note is applicable to data
analysis with the output error method. This method is
not necessarily limited to smooth air testing. Atmos-
pheric or wind tunnel turbulence may be included if
appropriate input measurements are taken.[6,7]

In aircraft or wind tunnel transient testing the
question comes up as to what kind of transient should
be selected. If the transient is too short, the
parameters will be identified with inadequate accuracy.
If the transient is too long, an unnecessary amount of
data must be processed. The question we pose here for
the maximum likelihood method is: Given a required
accuracy of the parameter estimate, and given an input
function, what is the minimal quantity of measured data
necessary to achieve this accuracy? From a survey of
the pertinent literature it appears that this question
has not been asked before, much less a solution
presented. However, there are some recent studies
where certain criteria were used to define an optimum
input. We will first briefly discuss two of these
optimal input proposals, and then proceed to develop
the method of optimal data utilization for a given type
of input.

TWO PROPOSALS FOR OPTIMAL INPUT DESIGN

General questions of input design are:

(a) What type of input function should be used?

(b) For what time period should the response data be
processed to enable identification of the system
parameters with a specified accuracy? Are certain
time periods of the response particularly rich in
information contents and should they, therefore, be
preferably used?

There usually are some constraints on the input design
like amplitude constraints, smoothness constraints
(step or impulse inputs are mathematical idealizations
but often practically not realizable), instrumentation
constraints, and constraints imposed by the selected
analytical model that usually filters out the higher
frequency contents of the input.

Analytical solutions of the problem of optimal
input design require the minimization of a cost function.
Stepner and Mehra[1] use the sensitivity of the system
response to the unknown parameters as the performance
criterion for optimal input design. The time of the
transient is assumed to be fixed. Thus questions (b)
are not involved. The measurement equation is

$$y_m(t) = y(x, \theta, u, t) + v(t) \tag{1}$$

x is the state vector, θ the parameter vector, u the
input vector, v(t) the additive measurement noise with

covariance matrix R. We write the Taylor expansion

with respect to the parameter θ about the a priori

estimate θ_0 of θ and neglect higher order terms:

$$y_m(t) = y(x, \theta_0, u, t) + \frac{\partial}{\partial \theta_0} y(x, \theta_0, u, t)(\theta-\theta_0)$$

$$+ v(t) \tag{2}$$

In the output error method $(\theta - \theta_0)$ is determined by a

least squares solution of Eq. (2) for a fixed time

period (t_0, t_f). For a high degree of accuracy in

determining $(\theta - \theta_0)$ the sensitivity function $\partial y/\partial \theta$

must be large. The scalar performance index selected

in Reference 1 is

$$J = \text{Trace (WM)} \tag{3}$$

where

$$M = \int_{t_0}^{t_f} (\partial y/\partial \theta)^T \; R^{-1}(\partial y/\partial \theta)dt \tag{4}$$

Due to the introduction of R^{-1} in M the performance

criterion favors the measurements which are more accurate.

The weighting matrix W is based on the relative impor-

tance of the parameter accuracies.

Assuming now linear system and measurement equations

$$x(t) = F\ x(t) + G\ u(t) \tag{5}$$

$$y_m(t) = H\ x(t) + v(t) \tag{6}$$

together with an "energy constraint" for the input

$$E = \int_{t_o}^{t_f} u^T u \, dt \qquad (7)$$

the optimum input u can be determined as a two point
boundary value problem whereby the Hamiltonian includes
the term

$$(1/2)\mu_o \ (u^T u - E/t_f) \ .$$

μ_o is the time invariant Langrange factor (scalar) to
be evaluated from the Euler differential equations of
the optimization problem. It should be noted that the
"energy constraint" Eq. (7) has no physical signifi-
cance but is a convenient device to obtain smooth
input functions. Physically the input will usually
be limited by amplitude rather than by the quadratic
criterion (7) and quite different "optimal" inputs can
then be expected.

Chen[8] attacks the problem of optimal input design
in an entirely different way as a time-optimal control
problem by minimizing

$$J = \int_{t_o}^{t_f} dt \qquad (8)$$

under the following constraints:

System equations

$$\overset{\circ}{x} = f(x, u, \theta, t) \quad , \quad x(t_o) = x_o \tag{9}$$

Sensitivity equations

$$\partial \overset{\bullet}{x}/\partial x_o = (\partial f/\partial x)(\partial x/\partial x_o) \, , \, \partial x(t_o)/\partial x_o = I \tag{10}$$

$$\partial \overset{\bullet}{x}/\partial \theta = (\partial f/\partial x)(\partial x/\partial \theta) + \partial f/\partial \theta, \, \partial x(t_o)/\partial \theta = 0 \tag{11}$$

Information matrix equations

$$\overset{\bullet}{M}{}^{-1} = -M^{-1}(\partial \nu/\partial \theta)^T R^{-1}(\partial \nu/\partial \theta)M^{-1} \tag{12}$$

where ν is the innovation:

$$\nu = y_m - y \tag{13}$$

and where the information matrix M is given by Eq. 4.
Finally Chen assumes an amplitude constraint

$$|u| \leq V \tag{14}$$

and he prescribes the trace of the information matrix
for time t_f

$$c_{ii}(t_f) = \sigma_i^2 \tag{15}$$

One can show that for linear input u(t) into the system
equation and for an input matrix independent of any
unknown parameter, the optimal input is of the "bang-bang"
form between the amplitude constraints. The solution of
this problem requires a computer search which was not
performed in Reference 8. Rather an arbitrary set of

bang-bang inputs in the form of Walsh functions was shown to result in a specific case in lower values of $M^{-1}(t_f)$ (given t_f) than those obtained by using Mehra's "optimal input". This apparent contradiction can be explained by the differential equation (12) governing M^{-1}. For a particular value of M^{-1} the rate of decrease of M^{-1} with time is dependent on all elements of

$$(\partial y/\partial \theta)^T R^{-1} (\partial y/\partial \theta)$$

while Mehra, in his criterion (3), optimizes only the trace of WM.

While the input amplitude constraint Eq. (14) used by Chen is physically more significant than the quadratic constraint Eq. (7) used by Mehra, the actual constraints are usually still more complex. In cases of airplanes or lifting rotors one usually wishes to limit the response to the linear sub stall regime, since the analytical model to be identified is often a linear one. The stall boundary is, however, a complex function of the input and cannot be represented by an amplitude constraint for the input transient. This is particularly relevant for the lifting rotor, so that neither the Mehra nor the Chen input optimization criteria are useful for lifting rotor applications,

quite apart from the excessive computer effort involved in obtaining the optimal inputs. Furthermore, the input matrix usually contains unknown parameters. In this case Chen's optimum solution would not be of the bang-bang type and would be still more difficult to obtain. For all of these reasons it was concluded that at the present state of optimal input design methods an attempt to compare our selected inputs with an "optimum input" would not be practical. Instead, a more limited approach has been taken described in the following section.

OPTIMAL DATA UTILIZATION FOR GIVEN INPUT FUNCTION

We first point out the difference between the continuous and the discrete case. In the maximum likelihood method (output error method for zero process noise) using the Newton-Raphson approach with quasi linearization, one obtains for the parameter update increment the following expressions:

Continuous case:

$$\Delta \theta = \left[\int_{t_o}^{t_f} \left(\frac{\partial \nu}{\partial \theta}\right)^T R^{-1} \left(\frac{\partial \nu}{\partial \theta}\right) dt \right]^{-1} \int_{t_o}^{t_f} \left(\frac{\partial \nu}{\partial \theta}\right)^T R^{-1} \nu \, dt \quad (16)$$

Discrete case:

$$\Delta \theta = \left[\sum_{i=1}^{N} \left(\frac{\partial \nu}{\partial \theta}\right)_i^T R^{-1} \left(\frac{\partial \nu}{\partial \theta}\right)_i \right]^{-1} \sum_{i=1}^{N} \left(\frac{\partial \nu}{\partial \theta}\right)_i^T R^{-1} \nu_i =$$

$$= M^{-1} \left(\frac{\partial J}{\partial \theta}\right) \quad (17)$$

The Cramér-Rao lower bound has been defined only for a vector of sampled measurements and not for the continuous case.[2,9]

For high sampling rate one can define an approximate differential equation for M from Eq. (17) in the following way:

$$\text{Set} \qquad S_i \triangleq (\partial v / \partial \theta)_i \qquad (18)$$

then

$$M = \sum_{i=1}^{N} S_i^T R^{-1} S_i = \frac{1}{\Delta t} \sum_{i=1}^{N} S_i^T R^{-1} S_i \, \Delta t \qquad (19)$$

As N increases, Δt gets smaller and the right hand side of Eq. (19) can be approximated by

$$M \approx (1/\Delta t) \int_{t_o}^{t_f = N\Delta t} S^T R^{-1} S \, dt \qquad (20)$$

$$M^{-1} \approx \Delta t \left[\int_{t_o}^{t_f} S^T R^{-1} S \, dt \right]^{-1} \qquad (21)$$

Taking the derivative of M^{-1} with respect to t_f:

$$dM^{-1}/dt_f = -M^{-1}(\partial M / \partial t_f)M^{-1}$$

or with Eq. (20):

$$dM^{-1}/dt_f = -(1/\Delta t)M^{-1} S^T R^{-1} S M^{-1} \qquad (22)$$

The point is that even in a continuous formulation the
time increment Δt between samplings must occur.
Eq. (21) is the correct formulation for the Cramér-Rao
lower bound of the covariance matrix for the parameters.

We can now use the approximately valid differential
equation (22) to obtain some insight into ways of best
data utilization. Let us assume that we wish to pre-
scribe certain values for the parameter standard
deviations σ_i and that we wish to compare the Cramér-Rao
lower bound with these standard deviations. Since we
are dealing not with the unknown actual parameter
covariances but only with their lower bounds, we should
apply some conservatism to the selected σ_i, that is we
should select σ_i smaller than we really need for the
specific data processing case. We thus require

$$0 \leq M_{t_f}^{-1}(i, i) \leq \sigma_i^2 \tag{23}$$

whereby $M_{t_f}^{-1}$ is the value of M^{-1} at time t_f. For
non-zero values of S, the right hand side of Eq. (22) is
negative definite and hence $M_{t_f}^{-1}(i, i)$ are monotonically
decreasing functions of t_f. There will thus be a
minimum time for which the constraints of Eq. (23) are
satisfied.

Another way of reducing the amount of measured
data for the parameter identification is to select for

the data processing those time periods for which the components of the matrix

$$S^T R^{-1} S$$

have significant values. From Eq. (21) it follows that then the Cramér-Rao lower bound M^{-1} will be particularly small. The components of M^{-1} also decrease with decreasing time element Δt between samples.

Since it is impractical to use for the integration of Eq. (22) infinity as initial condition, it is recommended to determine M^{-1} for a small time period, say for $N = 10$, from Eq. (21) and integrate Eq. (22) with the solution to Eq. (21) as initial conditions. Since S includes parameter estimates, one needs a preliminary estimation of the unknown parameters in order to use Eq. (22).

APPLICATION TO A CASE OF LIFTING ROTOR PARAMETER IDENTIFICATION

Using the simplest analytical model of a lifting rotor, a straight blade flapping about the rotor center, one has in a rotating frame of reference for the flapping angle β the following equation.[10]

$$\ddot{\beta} + (\gamma/2)C(t)\dot{\beta} + (\gamma/2)[P^2 + K(t)]\beta = (\gamma/2)[m_\theta\ \theta + m_\lambda\ \lambda]$$

(24)

where: γ is the blade Lock number

P is the blade flapping natural frequency in the rotating system

θ is the instantaneous blade pitch angle

λ is the non-dimensional normal inflow

One rotor revolution corresponds to $t = 2\pi$. For neglected reversed flow effects, zero root cut-out and with tip loss factor B, the functions $C(t)$, $K(t)$, $m_\theta(t)$, $m_\lambda(t)$ in terms of rotor advance ratio μ are:[10]

$$C(t) = (1/4)B^4 + (1/3)B^3 \mu \sin(t) \tag{25}$$

$$K(t) = (1/3)B^3 \mu \cos(t) + (1/4)B^2 \mu^2 \sin(2t) \tag{26}$$

$$m_\theta(t) = (1/4)(B^4 + B^2\mu^2) + (2/3)B^3\mu \sin(t)$$
$$- (1/4)B^2\mu^2 \cos(2t) \tag{27}$$

$$m_\lambda(t) = (1/3)B^3 + (1/2)B^2 \mu \sin(t) \tag{28}$$

In the numerical analysis, we use $B = 0.97$. A simple improvement of this analytical model that takes into account blade bending flexibility is possible.[11] In transient conditions the inflow λ includes the dynamic rotor wake in a complicated form. As a first approximation of dynamic rotor wake effects one can use, instead of the actual blade Lock number an equivalent smaller value of γ.[12] Such an approximation can be expected to be satisfactory if the transient is relatively slow. For transients with high frequency contents, this approximation is invalid.[13]

Due to rotor induced cross flow in a wind tunnel, the inflow parameter λ will usually not be well known. In addition, the aerodynamic pitch angle θ_o, due to airfoil inaccuracies and pitch setting errors, will also not be well known. For the wind tunnel tests considered here, we assume $\lambda = 0$ and use the collective pitch setting θ_o as an unknown parameter to be determined from the blade flapping measurements. In addition we have a transient blade pitch input θ assumed to be known. The problem then is to determine from blade flapping transients caused by blade pitch inputs, the equivalent Lock number γ and the equivalent collective pitch setting θ_o.

The lifting rotor wind tunnel model described in reference 14 allows excitation of progressing and regressing flapping modes at various frequencies. By a minor modification of this model, progressing or regressing transients can be excited. One can describe such inputs as pitch stirring transients. In a helicopter, this would amount to cyclic stick stirring, whereby the amplitude of the cyclic pitch would remain constant while the frequency of the stirring motion changes. The blade pitch input for such a stirring transient is selected to satisfy the equation:

$$\theta = 1.5 \cos[\omega(t - t_o) + t] + \theta_o \qquad (29)$$

where ω is the angular stirring speed in the sense of blade
rotation

$t_o = 12$ and

$$\omega = \begin{cases} 0 & 0 \leq t < t_o \\ -(0.1/\pi)(t - t_o) & t_o \leq t \leq T \end{cases}$$

The meaning of these input equations is the fol-
lowing. At time $t = 0$, a step lateral cyclic pitch
input of 1.5 degrees is imposed. At time $t = 12$,
the response to this input is approximately stabilized.
At this time the pitch stirring acceleration of
$-(0.1/\pi)$ is introduced which leads to a progressing
flapping excitation. The identification starts at
$t = 12$ with the pitch stirring transient.

Test results with transient pitch stirring inputs
will be presented at a later time when they become
available. Here we are concerned with the problem of
designing the tests in such a way that the test data
will be sufficient to determine the two unknown para-
meters γ and θ_o with good accuracy, i.e., to determine
a suitable value of T that allows an accurate identi-
fication of parameters.

The simulated identification analysis was performed
under the assumption of a random zero mean white noise
sequence superimposed on the analytical flapping transient.

This transient was obtained for $\theta_0 = 2^\circ$, $\mu = 0.4$ and $\gamma = 5.0$. For convenience, the parameters $\delta = \gamma\theta_0$ and γ instead of θ_0 and γ were identified.

System and measurement equations corresponding to equations (5) and (6) are:

$$
\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -(\gamma/2(P^2 + K(t)) & -(\gamma/2)C(t) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ (\gamma/2)m_{\theta_0} \end{bmatrix} \theta
\tag{30}
$$

$$
y_m = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + v(t)
\tag{31}
$$

where

$$E\{v(t)\} = 0 \qquad E[v(t)\,v(\tau)] = 0.2\,\delta(t - \tau)$$

and $\quad [x_1 \quad x_2] = [\beta \quad \dot{\beta}]$

ANALYSIS OF RESULTS

We first show in Table 1 the effect of data length on the parameters and their associated $M^{-1}(i,i)$ values. The iteration of the maximum likelihood method was begun with a 20 percent error in the parameter values. It is seen that a data length of $t = 12 - 14$ is quite inadequate, a data length of $t = 12 - 18$ gives reasonably good parameters, while a data length of $t = 12 - 24$ is much better and leads to very small

lower bounds of the parameter covariance matrix.
Fig. 1 shows the correct flapping response together
with the simulated measurement data. Pitch stirring
is initiated at t = 12. Figs. 2 and 3 show $M^{-1}(\gamma)$
and $M^{-1}(\delta)$ from Eq. (22) between t = 16 and t = 24.
Two curves are plotted, one for the initial crude
estimate of the parameters (γ = 4, δ = 8), and one for
the final estimate of the parameters for t = 24,
(γ = 4.91, δ = 9.83). The two curves are in this
case not much different. Note the steep descent of
the curves to about t = 17.5. It would, therefore,
be not acceptable to use the data up to less than the
time t = 17.5. However, there is another descent to
t = 23.0, causing the improvement shown in Table 1.
From Figs. 2 and 3 it is clear that the selection of
T = 24.0 is a good one, that the use of fewer data
would result in substantial decrease in parameter
accuracy, and that the use of additional data is
unnecessary.

REFERENCES

1.   Stepner, D. E. and Mehra, R. K., "Maximum Likelihood
     Identification and Optimal Input Design for Identi-
     fying Aircraft Stability and Control Derivatives,"
     NASA CR-2200, March 1973.

2.   Iliff, R. W. and Taylor, L. W., "Determination of
     Stability Derivatives from Flight Data Using a
     Newton-Raphson Minimization Technique," NASA
     TN D-6579, March 1972.

3.   Molusis, J. A., "Rotorcraft Derivative Identification
     from Analytical Models and Flight Test Data," Pre-
     sented at AGARD Flight Mechanics Panel Specialist's
     Meeting, NASA Langley, November 1974.

4.   Gould, D. G. and Hindson, W. S., "Estimates of the
     Stability Derivatives of Helicopter and a V/STOL
     Aircraft from Flight Data," DME/NAE Quarterly
     Bulletin No. 1974(4), Ottawa, January 1975.

5.   Hohenemser, K. H. and Prelewicz, D. A., "Computer
     Experiments on Periodic System Identification Using
     Rotor Blade Transient Flapping-Torsion Response at
     High Advance Ratio," Rotorcraft Dynamics, NASA
     SP-352, AHS and NASA/Ames, Moffett Field, Cali-
     fornia, February 1974.

6.   Iliff, R. W., "Identification of Aircraft Stability
     and Control Derivatives in the Presence of Turbulence,"
     A Symposium held at Flight Research Center,
     Edwards, California, April 24-25, 1973. NASA
     TN D-7647.

7.   Wingrove, R. C., "Parameter Estimation of Powered-
     Lift STOL Aircraft Characteristics including
     Turbulence and Ground Effects," Presented at AGARD
     Flight Mechanics Specialist's Meeting, NASA-Langley,
     November 1974.

8.   Chen, R. T. N., "Input Design for Aircraft Parameter
     Identification Using Time-Optimal Control Formulation,"
     Presented at AGARD Flight Mechanics Specialist's
     Meeting, NASA-Langley, November 1974.

9.   Van Trees, A. L., "Detection, Estimation and
     Modulation Theory," Wiley, New York, 1968.

10. Sissingh, G. J., "Dynamics of Rotors Operating at High Advance Ratios," Journal American Helicopter Society, Vol. 13, No. 3, July 1968, pp. 56-63.

11. Hohenemser, K. H. and Yin, S. K., "On the Question of Adequate Hingeless Rotor Modeling in Flight Dynamics," Presented at the 29th Annual National Forum of the American Helicopter Society, Washington, D.C., May 1973, Preprint 732.

12. Curtis, H. C. Jr., "Complex Coordinates in Near Hovering Rotor Dynamics," Journal of Aircraft, Vol. 10 No. 5, May 1973, pp. 289-296.

13. Peters, D. A., "Hingeless Rotor Frequency Response with Unsteady Inflow," Rotorcraft Dynamics, NASA SP-352, 1974.

14. Hohenemser, K. H. and Crews, S. T., "Model Tests on Unsteady Rotor Wake Effects," Journal of Aircraft, Vol. 10, No. 1, January 1973, pp. 58-60.

FIGURE CAPTIONS

Fig. 1   Simulated measurements of β, σ_β = .2

Fig. 2   Cramer-Rao Lower Bound $M^{-1}(\gamma)$ vs. time

Fig. 3   Cramer-Rao Lower Bound $M^{-1}(\delta)$ vs. time

Table 1   Iteration vs. Time Variation of Parameters and Their Cramér-Rao Lower Bounds.
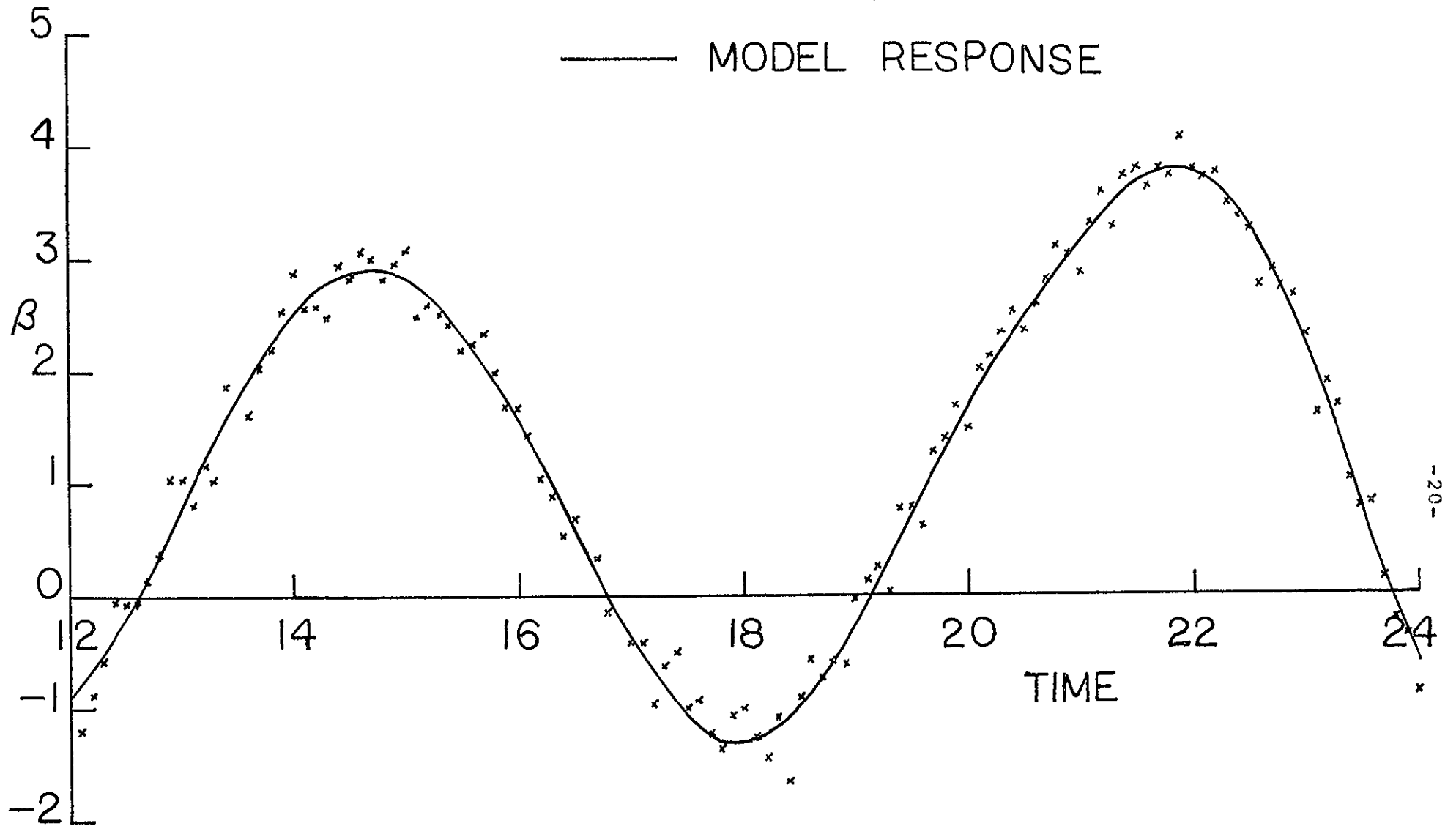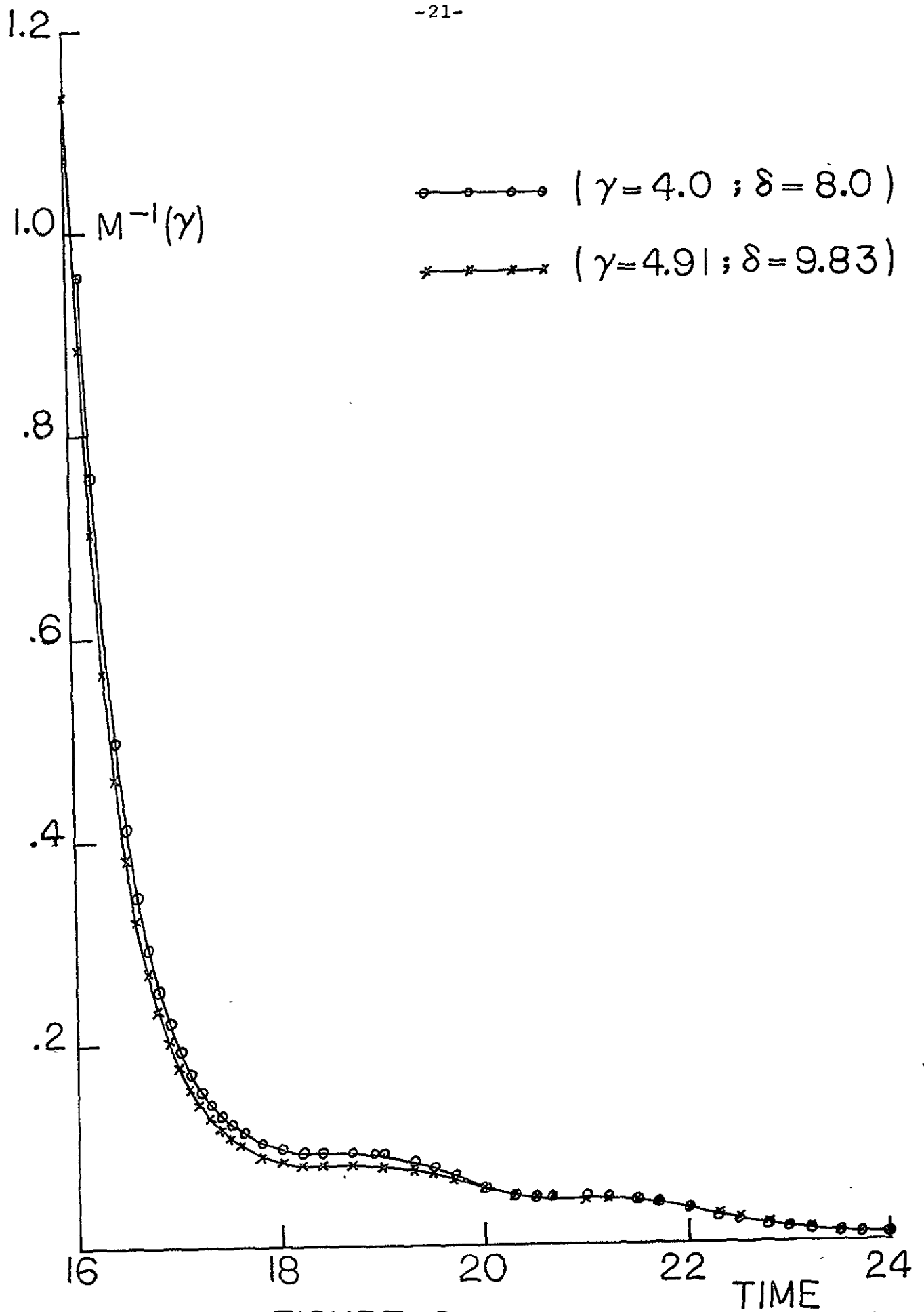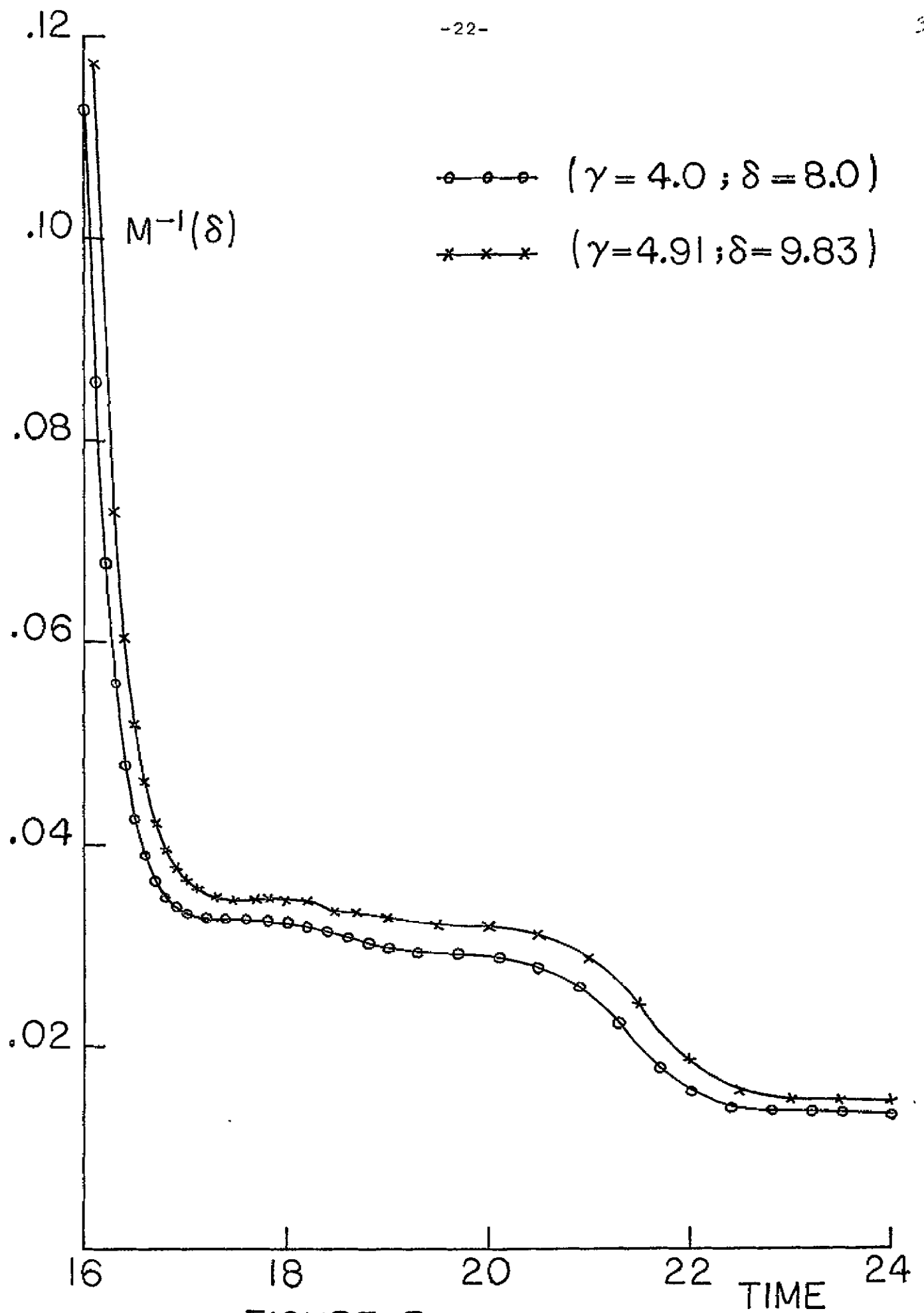
FIGURE 1

FIGURE 2

FIGURE 3

Parameter		t = 12 - 14			t = 12 - 18				t = 12 - 24			
		γ	δ	$M^{-1}(\gamma)$	$M^{-1}(\delta)$	γ	δ	$M^{-1}(\gamma)$	$M^{-1}(\delta)$	γ	δ	$M^{-1}(\gamma)$	$M^{-1}(\delta)$
True value		5.00	10.00			5.00	10.00			5.00	10.00		
Iteration	0	4.00	8.00			4.00	8.00			4.00	8.00		
	1	4.29	9.73	48	6.5	5.36	9.67	.096	.032	4.94	9.69	.007	.013
	2	4.17	9.71	37	8.1	5.23	9.73	.100	.035	4.91	9.85	.008	.015
	3	4.10	9.67	37	8.0	5.23	9.73	.094	.035	4.91	9.83	.008	.015

TABLE 1