Reduced Order Nonlinear System Identification Methodology

Peter J. Attar*

Computational Sciences Branch, Wright–Patterson Air Force Base, Ohio 45433-7913

and

Earl H. Dowell,† John R. White,‡ and Jeffrey P. Thomas§

Duke University, Durham, North Carolina 27708

DOI: 10.2514/1.16221

A new method is presented which enables the identification of a reduced order nonlinear ordinary differential equation (ODE) which can be used to model the behavior of nonlinear fluid dynamics. The method uses a harmonic balance technique and proper orthogonal decomposition to compute reduced order training data which is then used to compute the unknown coefficients of the nonlinear ODE. The method is used to compute the Euler compressible flow solutions for the supercritical two-dimensional NLR-7301 airfoil undergoing both small and large pitch oscillations at three different reduced frequencies and at a Mach number of 0.764. Steady and dynamic lift coefficient data computed using a three equation reduced order system identification model compared well with data computed using the full CFD harmonic balance solution. The system identification model accurately predicted a nonlinear trend in the lift coefficient results (steady and dynamic) for pitch oscillation magnitudes greater than 1 deg. Overall the reduction in the number of nonlinear differential equations was 5 orders of magnitude which corresponded to a 3 order of magnitude reduction in the total computational time.

I. Introduction

In recent years an increase in computational power has allowed engineers to solve increasingly complex problems in computational fluid dynamics (CFD) and computational aeroelasticity. It is still true, however, that the turnaround time needed to compute and postprocess a solution solved using a nonlinear time-domain computational fluid dynamics code is prohibitive when one considers that multiple computer runs are needed to solve the given problem. This is especially the case in computational aeroelasticity where multiple time-domain solutions are needed to compute flutter points or the consequent limit cycle oscillations that may occur. Also if optimization is to be performed hundreds or even thousands of solutions are needed before a design is optimized. In these problems, a more efficient means of computing solutions is needed to make the process tractable.

Over the last several years investigators have attempted to address the above mentioned issues using several types of reduced order methodologies [1–7]. These methods work exceedingly well when the dynamics of the system can be considered to be linear. Computational savings of 3–4 orders of magnitude are not uncommon for problems solved using these reduced order methods. Less success has been reported for the use of such models in problems where either the dynamics must be considered to be nonlinear or where a reduced

---

Nomenclature

- $A(t)$ = surface area at time $t$
- $A_i$, $B_{ij}$, $C_{ijk}$, $D_{ijkl}$, $E_{ijkl}$ = coefficients of the function $F_i$
- $\hat{a}_i$ = $i$th POD modal amplitude
- $B_i$ = linear function of the $u_i$
- $C$ = matrix of unknown coefficients
- $E_x$ = total energy
- $F$ = vector which contains $x$, $y$, and $z$ fluxes
- $F_i$ = nonlinear function of the $\hat{a}_i$
- $\bar{F}$ = $x$ component of unsteady control volume motion
- $G_{ij}$ = coefficients of the function $B_i$
- $\bar{g}$ = $y$ component of unsteady control volume motion
- $\bar{h}$ = $z$ component of unsteady control volume motion
- $N_C$ = number of unknown coefficients per equation
- $N_P$ = number of training data sets
- $N_h$ = number of harmonics kept in Fourier expansion
- $\hat{n}$ = surface normal vector
- $p$ = number of modes kept in modal expansion
- $\rho$ = fluid density
- $\phi, \omega, \bar{\omega}$ = frequency and reduced frequency
- $S$ = snapshot matrix
- $T$ = fundamental period of motion
- $TD$ = matrix of CFD training data
- $t$ = dimensional time
- $U$ = vector of conservative fluid variables
- $U_{\phi}$ = steady flow solution
- $u$ = $x$ component of fluid velocity
- $V(t)$ = control volume at time $t$
- $v$ = $y$ component of fluid velocity
- $w$ = $z$ component of fluid velocity
- $\bar{x}$ = unsteady motion of control volume
- $\alpha(t)$ = time dependent pitching motion
- $\tau = pitch magnitude$
- $\Phi$ = matrix of POD modes
- $\phi$ = $i$th POD fluid mode
- $\omega$ = frequency and reduced frequency

---

Received 21 February 2005; revision received 3 February 2006; accepted for publication 19 March 2006. Copyright © 2006 by the American Institute of Aeronautics and Astronautics, Inc. The U.S. Government has a royalty-free license to exercise all rights under the copyright claimed herein for Governmental purposes. All other rights are reserved by the copyright owner. Copies of this paper may be made for personal or internal use, on condition that the copier pay the $10.00 per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923; include the code $10.00 in correspondence with the CCC.
order model of the nonlinear static model is needed. Problems with the construction of such reduced order models which have been reported include 1) computational efficiency does not scale with the number of degrees of freedom (subspace method) [8,9] and 2) models are oftentimes numerically unstable (direct projection methods) [10].

An additional challenge with existing reduced order models is the need to modify the original CFD code in some manner to implement the method. Although this may not seem to be a large problem, one should keep in mind the large investment, both in time and money, put into the creation of these codes, and alteration of the CFD code by a user who is not the original author may take a good deal of time and effort. Also with modification of the source code comes the potential for the introduction of “bugs” into the code. Oftentimes the modification of a source code requires a revalidation which adds to the cost of building such models. Another problem is that in some cases a commercial code may be used where the user does not have access to the original source code.

To avoid the problems introduced by equation (source code) modification, researchers have in recent years looked into the construction of “black box” models which use only input and output data to create a model which can be used to predict system behavior. Many such methods, a majority of which are based on the seminal work done by Kalman [11], are used in the structural dynamics community to build linear state-space equations and to predict system parameters [12,13]. Recently similar methods have been used to build linear state-space aerodynamics models as well [14,15].

Although these system realization methods produce a compact model which relies only on input and output data, the input/output mapping produced by such methods is linear. To overcome this shortcoming, two methods have been proposed in recent years, Volterra formulations and neural networks. Although neural networks have gained popularity in many fields of science and engineering, the large amount of training data needed to build a robust, accurate model makes them less than attractive in the field of CFD. Marques and Anderson [16] computed the fluid forces on an airfoil using a neural network approach. In other work using neural networks, Denegri and Johnson [17] predicted limit cycle oscillations using a network which was trained with flight test data. Also Voitcu and Wong [18] used a neural network to predict nonlinear aeroelastic oscillations. The training sets used in this work consisted of both numerical and experimental data.

In recent years numerous papers have been published by Silva [19–25] and colleagues which address Volterra/indicial function methods and their applicability to computational fluid dynamics and aeroelasticity. Although the identification of linear Volterra kernels has proven to be very successful, the identification of nonlinear kernels is still an active research area. See the paper by Lucia and Beran [26].

In work by Cowan et al. [27], a linear autoregressive moving average (ARMA) time series method was used to identify a model which related generalized aerodynamic forces to structural modal amplitudes. Using this methodology, the authors were able to efficiently and accurately predict the linear aeroelastic stability boundary for the AGARD 445.6 wing configuration.

For a review of the current state of the art in reduced order modeling and system identification as it applies to computational fluid dynamics and aeroelasticity, the reader is referred to the comprehensive paper on the subject by Lucia et al. [28].

In the present paper, a nonlinear system identification methodology will be presented which is used to identify a set of low dimensional, nonlinear, first order ordinary differential equations (ODE). The coefficients of the terms in the ODE are found using a simple least squares method. The training data are computed using a harmonic balance solution [29,30] of the Euler equations for flow over the supercritical NLR 7301 airfoil. The dimension of these data is then reduced using proper orthogonal decomposition (POD). Training data are computed for various pitch amplitudes and reduced frequencies. Results will be presented which will compare the system identification model results and the results computed using a conventional CFD solution.

II. Theoretical Development

In this section a brief overview will be given of both the CFD method used to gather the training data for the system identification model (harmonic balance) and the method used to reduce the dimensionality of the system (POD). For further background on the harmonic balance technique the reader is referred to [29–32]. For further background on the POD methodology please refer to [6,10,28,33–35].

A. Fluid Equations of Motion

In this work, the inviscid, compressible Euler equations are used to model the fluid. The data from the solution of the discretized form of this equation for various input excitations are used as the training data needed for the system identification method. The three-dimensional Euler equations can be written in integral form in the following manner:

\[
\frac{\partial}{\partial t} \iiint_{\Omega} \mathbf{U} \, dV + \iiint_{\partial \Omega} (\mathbf{F} - \mathbf{U} \mathbf{X}) \cdot \hat{n} \, dA = 0
\]

where \( \mathbf{U} \) is a vector of the conservative fluid variables

\( \mathbf{U} = \{ \rho, \rho u, \rho v, \rho w, E_i \} \)

and \( \mathbf{F} \) consists of the \( x, y, \) and \( z \) inviscid flux vectors

\[
\mathbf{F} = \mathbf{F}_i + \mathbf{G}_j + \mathbf{H}_k
\]

The \( \mathbf{X} \) term represents the unsteady motion of the control volume and is written as

\[
\mathbf{X} = \mathbf{j} \hat{\mathbf{i}} + \hat{\mathbf{j}} \hat{\mathbf{j}} \quad \mathbf{h} \hat{\mathbf{k}}
\]

The CFD method used to discretize Eq. (1) is a variant of the standard Lax–Wendroff scheme [36,37]. A similar approach could be used with the Navier–Stokes equations, however, for an initial demonstration of the method, it is simpler to use the Euler equations. It should be noted that when the training data computed by solving the above Euler equations are used to build the reduced order model used in the current work, the first conservative variable is written as \( 1/\rho \) as opposed to \( \rho \). It can be shown that the effect of doing this on the form of the Euler equations is to replace the polynomial terms which have rational coefficients with ones which have nonrational coefficients. This step was taken to simplify the identification of the nonlinear model coefficients.

B. Generation of the Training Data: Harmonic Balance Methodology

In recent years the high dimensional harmonic balance (HDHB) technique has been used to obtain accurate and efficient solutions to problems which can be assumed to have periodic (in time) solutions [29,30,32]. The HDHB technique, developed by Hall and colleagues, is a modification of the classical harmonic balance method. In recent years, Beran and his colleagues have developed a similar method which has also been used successfully in computing periodic solutions to aeroelastic problems [38].

Note that the computation of the training solutions with a HDHB-type method is not necessary for the current identification methodology, but it does allow for an efficient and accurate means of computing periodic solutions.
If the unsteadiness of the flow is considered strictly periodic in time with a period \( T = 2\pi/\omega \), the first integral term in Eq. (1) can be expanded in a Fourier series:

\[
\int \int \int_{V(t)} \mathbf{U} \mathbf{d}V = \mathbf{Q}(t) \approx \sum_{n=1}^{N_h} \hat{\mathbf{Q}}_n e^{-j\omega t}
\]

thus the first term of Eq. (1) may be written in the following manner:

\[
\frac{\partial}{\partial t} \int \int \int_{V(t)} \mathbf{U} \mathbf{d}V \approx j\omega \sum_{n=1}^{N_h} n\hat{\mathbf{Q}}_n e^{-j\omega t}
\]

The second integral term in Eq. (1) may also be expanded in a similar manner:

\[
\int_{A(t)} (\mathbf{F} - \mathbf{U}\mathbf{X}) \cdot \mathbf{n} \mathbf{d}A = \mathbf{R}(t) \approx \sum_{n=1}^{N_h} \hat{\mathbf{R}}_n e^{-j\omega t}
\]

If the two integrals in Eq. (1) are replaced by Eqs. (6) and (7), and the resulting equation is multiplied by \( e^{-j\omega t} \) and integrated over one period for each \( m \), that is,

\[
\int_0^T \frac{1}{T} \sum_{m=-N_h}^{N_h} (j\omega \hat{\mathbf{Q}}_n + \hat{\mathbf{R}}_n) e^{-j\omega t} e^{-j\omega t} \mathbf{d}t
\]

the result is a system of nonlinear algebraic equations in the Fourier coefficients which can be represented in matrix form as

\[
\mathbf{A} \hat{\mathbf{Q}} + \hat{\mathbf{R}} = 0
\]

where \( \mathbf{A} \), \( \hat{\mathbf{Q}} \), and \( \hat{\mathbf{R}} \) are defined to be

\[
\mathbf{A} = j\omega \begin{bmatrix}
-jN_h & & \\
& \ddots & \\
& & jN_h
\end{bmatrix}, \quad \hat{\mathbf{Q}} = \begin{bmatrix}
\hat{\mathbf{Q}}_{-N_h} \\
\hat{\mathbf{Q}}_{-N_h+1} \\
\vdots \\
\hat{\mathbf{Q}}_{N_h}
\end{bmatrix}, \quad \hat{\mathbf{R}} = \begin{bmatrix}
\hat{\mathbf{R}}_{-N_h} \\
\hat{\mathbf{R}}_{-N_h+1} \\
\vdots \\
\hat{\mathbf{R}}_{N_h}
\end{bmatrix}
\]

In the classical harmonic balance methodology Eq. (9) is solved directly by combining like terms and solving the set of nonlinear equations. See, for example, [39]. However, for problems with a large number of degrees of freedom, expressing the nonlinear components in terms of the Fourier variables becomes a difficult task. To simplify this, the variables in Fourier space are transformed back to variables in the time domain (at uniformly spaced subtime levels) using the Fourier transform matrix \( \mathbf{E} \):

\[
\hat{\mathbf{Q}} = \mathbf{E} \mathbf{Q}, \quad \hat{\mathbf{R}} = \mathbf{E} \mathbf{R}
\]

where

\[
\mathbf{Q} = \begin{bmatrix}
\mathbf{Q}(t_0) \\
\mathbf{Q}(t_1) \\
\vdots \\
\mathbf{Q}(t_{N_h})
\end{bmatrix}, \quad \mathbf{R} = \begin{bmatrix}
\mathbf{R}(t_0) \\
\mathbf{R}(t_1) \\
\vdots \\
\mathbf{R}(t_{N_h})
\end{bmatrix}
\]

The time levels are defined as

\[
t_n = \frac{2\pi n}{(2N_h + 1)\omega} \quad n = 0, 1, \ldots, 2N_h
\]

and \( \mathbf{Q}(t_n) \) and \( \mathbf{R}(t_n) \) are written as

\[
\mathbf{Q}(t_n) = \int \int \int_{V(t_n)} \mathbf{U}(t_n) \mathbf{d}V, \\
\mathbf{R}(t_n) = \int \int \int_{A(t_n)} (\mathbf{F}(t_n) - \mathbf{U}(t_n) \mathbf{\hat{X}}(t_n)) \cdot \mathbf{\hat{n}}(t_n) \mathbf{d}A
\]

Substituting Eq. (11) into Eq. (9) and premultiplying the resulting equation by \( \mathbf{E}^{-1} \) results in the following matrix equation for the unknown fluid variables at \( 2N_h + 1 \) subtime levels:

\[
\mathbf{D} \hat{\mathbf{Q}} + \hat{\mathbf{R}} = 0
\]

where \( \mathbf{D} \) is defined as

\[
\mathbf{D} = \mathbf{E}^{-1} \mathbf{A} \mathbf{E}
\]

Note that Eq. (15) is a steady state equation with \( (2N_h + 1)N \) unknowns where \( N \) is the number of fluid degrees of freedom per grid point. To solve Eq. (15) a pseudotime derivative \( \frac{\partial \hat{\mathbf{Q}}}{\partial t} \) is added so that an iterative technique can be used:

\[
\frac{\partial \hat{\mathbf{Q}}}{\partial t} + \mathbf{D} \hat{\mathbf{Q}} + \hat{\mathbf{R}} = 0
\]

This is the HDHB method of Hall et al.

C. Reduced Order Modeling: Proper Orthogonal Decomposition

Although the harmonic balance method does efficiently compute accurate results to be used in training the system identification model, the spatial dimension of the data produced is still \( N \). Creating a system identification model using this full dimensional data would not be practical because a large number of terms (\( \gg N^2 \)) would need to be identified. For this reason a reduction in the dimension of the data is needed. This is accomplished in the current work using POD.

In a spectral CFD method such as POD, the CFD degrees of freedom are written as a linear combination of the product of a time varying amplitude (perturbation amplitude in this case) \( \hat{\alpha} \), and a spatially varying function (vector after the system is discretized in space) \( \phi \):

\[
\mathbf{U} = \mathbf{U}_0 + \sum_{i=1}^{P} \phi_i \hat{\alpha}_i
\]

where \( P \) is usually much less than \( N \).

In the snapshot POD method, the spatial vector is assumed to be a linear combination of the \( M \) perturbation flow solutions \( \hat{\mathbf{U}} \), where \( \mathbf{U} = \mathbf{U}_0 - \mathbf{U}_0 \). This can be written in the following way:

\[
\phi_k = \sum_{i=1}^{M} \hat{\mathbf{U}}_i \mathbf{v}_k
\]

where in Eq. (19) the vector \( \mathbf{v}_k \) contains the \( M \) contribution weights of the snapshot vectors \( \hat{\mathbf{U}}_k \) on the \( k \)th mode \( \phi_k \) and are computed such that they lie along the principal axes of the vector space spanned by the matrix of snapshot vectors \( \mathbf{S} \) defined as

\[
\mathbf{S} = [\hat{\mathbf{U}}_1, \hat{\mathbf{U}}_2, \ldots, \hat{\mathbf{U}}_M]
\]

Also the vectors \( \mathbf{v}_k \) are to be scaled such that the POD vectors \( \phi_k \) are of unit length. In order for the last two conditions to be satisfied, the following expression must be maximized:

\[
\| \mathbf{S}^T \mathbf{S} \mathbf{v}_k \|_2
\]

subject to the constraint that \( \phi_k \) is of unit length. In Eq. (21) the matrix product \( \mathbf{S}^T \mathbf{S} \) is a measure of the spatial correlation of each computed solution (snapshots) in \( \mathbf{Q} \). If a Lagrange multiplier \( \lambda_k \) is introduced, the problem of finding the \( \mathbf{v}_k \) and \( \lambda_k \) which maximizes Eq. (21) and restricts \( \phi_k \) to be of unit length, becomes a statement that the quantity \( \Pi \) is to be stationary:

\[
\Pi = \mathbf{v}_k^T \mathbf{S}^T \mathbf{S} \mathbf{v}_k - \lambda_k (\mathbf{v}_k^T \mathbf{S}^T \mathbf{S} \mathbf{v}_k - 1)
\]

If the variation of Eq. (22) is computed and set equal to 0, the resulting equation yields an eigenvalue problem for \( \mathbf{v}_k \) and \( \lambda_k \).
This eigenvalue problem can then be solved for the \( M \) eigenvectors and eigenvalues. If the eigenvectors are put into a matrix \( \mathbf{V} \):

\[
\mathbf{S}^T \mathbf{S} \mathbf{v}_k = \lambda_k \mathbf{v}_k
\]  

(23)

This eigenvalue problem can then be solved for the \( M \) eigenvectors and eigenvalues. If the eigenvectors are put into a matrix \( \mathbf{V} \):

\[
\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_M]
\]  

(24)

the matrix of POD modes \( \Phi \) can be formed by computing the matrix product of \( \mathbf{S} \) and \( \mathbf{V} \):

\[
\Phi = \mathbf{S} \mathbf{V}
\]  

(25)

There are now \( M \) POD shape vectors or basis vectors, \( \phi_i \).

**D. Identification of the Nonlinear Model Equations**

The task of any system identification methodology is to develop a model which mimics, with good accuracy, the behavior of the true system, be it experimental or computational. In the current work, the model which will be used for this purpose, is a set of first order, reduced dimension, nonlinear ordinary differential equations:

\[
\frac{d\hat{a}_i}{dt} = F_i(\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_P; t) + B_i(u_1, u_2, \ldots, u_P; t)
\]  

(26)

where the function \( F \) is a nonlinear function of the modal amplitudes \( \hat{a}_i \) and the function \( B \) is a linear function of the \( r \) inputs to the system. In this paper the \( u \) consist of rigid airfoil pitch and plunge motions. However the methodology is not limited to such inputs. Note that any coupling that may exist between \( u \) and \( \hat{a}_i \) is ignored in the present method.

The nonlinear function \( F \) is modeled using multivariable polynomials which can be written (up to the fourth order) as

\[
F_i = A_i + \sum_{j=1}^{P} B_{ij} \hat{a}_j + \sum_{j=1}^{P} \sum_{k=j}^{P} C_{ijk} \hat{a}_j \hat{a}_k + \sum_{j=1}^{P} \sum_{k=1}^{P} \sum_{l=k}^{P} D_{ijkl} \hat{a}_j \hat{a}_k \hat{a}_l + \sum_{j=1}^{P} \sum_{k=1}^{P} \sum_{l=1}^{P} E_{ijklm} \hat{a}_j \hat{a}_k \hat{a}_l \hat{a}_m
\]  

(27)

Formally this form of a multivariable polynomial can be said to exist in the space of all \( P \) dimensional polynomials of degree four denoted as \( \Pi_4(\mathbb{R}^P) \). The linear function \( B \) is expressed in the following manner:

\[
B_i = \sum_{j=1}^{r} G_{ij} u_j
\]  

(28)

In Eqs. (27) and (28) the coefficient terms \( A_i, B_{ij}, C_{ijk}, D_{ijkl}, E_{ijklm}, \) and \( G_{ij} \) are unknown and have to be determined.

The choice to use polynomials in Eq. (27) follows naturally if one thinks about a Taylor expansion of the flux terms in the Euler equations about the mean flow \( \bar{U}_i \). This will result in an equation which contains polynomials in \( U - \bar{U}_i \). The coefficients of these terms will be the Jacobians (first order and higher) of the flux vector. If the flow variables are then expressed as a linear combination of the fluid modes, as in Eq. (18), the mean flow variables will drop out of the polynomial terms and Eq. (26) is recovered. Therefore, with this in mind, the function \( F \) in Eq. (27) can be thought of as representing the Taylor expansion of the flux, written in terms of the modal degrees of freedom, up to and including the fourth order term. The omission of the constant term in the expression for \( F \) is equivalent to setting \( \bar{U}_i \) to zero.

With the form of the model given, the unknown coefficients in Eqs. (27) and (28) must be determined. The first step in doing this is obtaining training data from the CFD model for a range of inputs. Because the function \( F \) contains both linear and nonlinear terms, the range of inputs used to excite the CFD model for the generation of training data should be broad enough to initiate both linear and nonlinear behavior of the CFD model. Although it is true that computing training data which span a broad frequency spectrum is important when computing the POD modes, this is not the case when computing the training data to be used to identify the model in Eq. (26). Instead the training data which are used in computing the coefficients in Eq. (27) should cover a wide range of response (\( \hat{a} \) and \( \dot{a} \)) and input (\( u \)) levels. In a time-domain CFD code this could be accomplished using multistep inputs similar to those discussed in [27]. Because in this work the harmonic balance method is used to generate the training data, the goal of computing over a range of inputs and outputs is met by computing training data with a wide range of input frequencies and magnitudes.

Once the CFD model training data have been computed, they must be projected onto the POD modes to obtain the modal training data which are to be used to compute the coefficients. This is done formally using a least squares fit:

\[
\mathbf{T} \hat{d} = \Phi \mathbf{a}
\]  

(29)

In Eq. (29) the matrix \( \mathbf{T} \) has a dimension of \( N \times P \), where \( N \) is the number of individual solutions found using the CFD model. The matrix \( \Phi \) is the matrix of unknown modal amplitudes and has the dimension \( P \times N \). Because of the orthogonality of the POD modes the least squares solution of Eq. (29) can be simplified in the following manner:

\[
\hat{a}_j = \left( \sum_{i=1}^{N} \overline{U}_i \phi_{ij} \right) / \sum_{i=1}^{N} \phi_{ij} \]  

(30)

Because the model which is to be identified, Eq. (26), is a system of ordinary differential equations, the matrix of modal amplitude time derivatives \( \dot{\hat{a}} \) must also be computed. In the harmonic balance method this matrix can easily be computed using the matrix \( \mathbf{D} \) defined in Eq. (16). If a time-domain code is used to compute the training data, this matrix could be computed using a simple differencing scheme.

To compute the coefficients in Eqs. (27) and (28) a least squares solution of the following equation is used:

\[
\hat{a}^T = \mathbf{V} \mathbf{C}
\]  

(31)

where the matrix \( \mathbf{C} \) is the \( N_c \times P \) matrix of unknown coefficients and the \( N_c \times N_c \) matrix \( \mathbf{V} \) contains the polynomial terms from Eqs. (27) and (28), evaluated for each of the \( N_c \) training data sets. The value of \( N_c \) corresponds to the number of unknown coefficients in each one of the identified equations. There are \( N_c + P \) total unknowns.

Once the coefficients are found, the resulting system of nonlinear ordinary differential equations can be integrated using standard techniques. In this work a fourth order Runge–Kutta scheme is used. Once the time history of modal amplitudes is computed, the full CFD variable results are recovered using Eq. (18).

**III. ROM Development Outline**

An outline of the reduced order system identification methodology presented here is as follows:

1. A high fidelity CFD model is used in a “training step.” This training data should cover a wide range of response and input levels and should initiate both linear and nonlinear behavior of the CFD model. A subset (one which exhibits linear behavior of the CFD model) of this training data will be used to generate POD modes. In this work the HDHB technique was used in the generation of the training data. Note that because the model coefficients are computed using a least squares solution (see Sec. IID for details), the number of training points needed is dictated by the number of modes used in the reduced order model along with the order of the polynomial used in the ordinary differential equation for which coefficients will be identified. If necessary this step can be repeated to add additional data points to the training set.

2. A subset of the training data from step 1 is used in the computation of the fluid POD modes. In order for the identified model to be able to accurately reproduce linear CFD results, this
subset should contain data which represent linear dynamic behavior of the CFD model.

3) Choose the number of POD modes to be used in the reduced order model.

4) To transform the training data computed in step 1 into modal form, the complete set of training data from step 1 is projected onto the POD modes computed in step 2. Because of the orthogonality of the POD modes, this is done efficiently using a simplified least squares method (see Sec. II.D).

5) Compute the modal velocities. The way this is accomplished depends on the type of CFD model used in the training. If a time-domain CFD solution is used in the computation of the training data and if the time derivatives of the conservative variables are available, then the procedure used in step 4 can be used to generate modal time derivatives. If the time derivatives of conservative variables are not available then a differencing in time can be used to compute the time derivatives from the data computed in step 4. In this work because a harmonic balance method was used in the training step, the time derivatives are easily computed using the periodic in time assumption.

6) Choose the order for the polynomial used in the nonlinear ordinary differential equation whose coefficients are to be identified.

7) Compute the unknown coefficients of the nonlinear, reduced order ordinary differential equations. This step is accomplished using the data computed steps 4 and 5 along with the choice of polynomial order in step 6. The model coefficients are then computed using a least squares method.

IV. Results

The model configuration which is to be considered is the NLR 7301 constant airfoil section wing. Over the past several years, this wing has been the subject of numerous experimental and theoretical aeroelastic studies [40–45]. The CFD model of [44], which uses a harmonic balance solution of a variant of the standard Lax–Wendroff scheme, is used here to train the system identification model. The results from this model are also compared to the results generated with the current system identification model. In this paper viscous effects are not considered and so the training data are generated using a harmonic balance solution of the two-dimensional, compressible Euler equations. Two harmonics are used in the harmonic balance solution. The grid used in the generation of the training data is a C-grid structured inviscid mesh topology which consists of 49 mesh points in the radial direction and 193 mesh points in the circumferential direction.

Both forced pitch and plunge oscillations are used to compute the POD modes. The pitch oscillations are given an amplitude of 0.0001 deg and reduced frequencies which range from 0.0 to 1.20 in reduced frequency increments of 0.05. The plunge oscillations are given an amplitude of 0.0001 of the chord and the reduced frequencies ranged from 0.10 to 1.00 in increments of 0.10. A total of 34 CFD solutions (170 training points) was used in the computation of the POD modes. The mean angle of attack is taken to be zero and the Mach number for all cases is 0.764.

Here results will be shown for a system identification model which uses a three mode expansion \([P = 3\) in Eq. (18)] of the flow variables. These three modes contribute over 97% of the total energy contained in the data used to compute the POD modes.

The number of training data points \(N_{p}\) used to compute the coefficients of the terms in the functions \(F_i\) and \(B_i\) in Eq. (26) is 2080. These consist of 416 different solutions of the harmonic balance CFD model using two harmonics (five time levels) in the solution. The 416 different solutions consist of airfoil pitch oscillations which range in magnitude from 0.0001 to 14 deg and which have reduced frequencies from 0.0 to 1.2. These particular values are chosen such that the airfoil behaved in both a dynamically linear and nonlinear manner. Although both plunge and pitch motion is used to compute the POD modes, because it was determined that the initial evaluation of the methodology would be restricted to pitch motion only, plunge motion training data were not used in the identification of the unknown coefficients. The only coefficients in the expression for both \(F_i\) and \(G_i\) which would need to be identified with plunge training data would be the coefficients in \(G_i\) which multiply the plunge terms themselves. The accurate identification of the coefficients in \(F_i\) need only a broad range of \(d\) and \(d\dot{x}/dt\) which can be supplied by pitch motion alone.

Of the 450 total training solutions used to build the reduced order model, 34 of these were used in the computation of the POD modes. In a standard POD reduced order model, such as those presented in [8–10], the training portion of the model consists only of computing the data which will be used in the computation of the POD modes. In a direct comparison between the current methodology and the standard POD reduced order model, the current methodology would have an order of magnitude larger training expense than the standard POD method. However as was mentioned in the Introduction, standard POD methods often have stability (direct projection methods) or efficiency (subspace methods) issues so that it is not evident that a direct comparison such as this is applicable.

The pitching motion which is used in this study can be represented in the following manner:

\[
\alpha(t) = \alpha_0 \cos(\omega t) \tag{32}
\]

where \(\omega\) is the reduced frequency based on the airfoil chord and \(t\) is a nondimensional time. When the identified ordinary differential equations are integrated forward in time, at each time step the value of \(\alpha(t)\) is substituted into the appropriate term in the expression for \(B_i\).

In the system identification model results presented in this section, the nominal grid \((\alpha_0 = 0)\) was used in the integration of the pressure over the airfoil surface. This differs from the procedure used in the harmonic balance CFD model but at the values of \(\alpha_0\) which are studied here (12 deg or less) the effect on the lift coefficient should not be significant.

Figures 1–3 show the variation in the mean lift coefficient as a function of \(\alpha_0\) for three reduced frequencies, 0.2, 0.6, and 1.0. System identification results are shown for two different models, one which used a second order polynomial for \(F_i\) and the other which used a third order polynomial for \(F_i\). A solution computed using the full CFD model with two harmonics is also shown. At the lowest reduced frequency the system identification model which uses the second order polynomial does a better job in quantitatively matching the CFD results. However, for a reduced frequency of 0.6 the second order model is qualitatively and quantitatively much worse than the third order model as it predicts an increase in the mean lift with an increase in \(\alpha_0\). For the highest reduced frequency of 1.0 both models are similar. For all three reduced frequencies, the third order model predicts accurately the angle of attack (\(\approx 4\) deg) where the mean lift coefficient starts to deviate significantly from the steady flow mean lift coefficient. The system identification model appears to underpredict the nonlinear dependence of the mean lift coefficient on the pitch magnitude \(\alpha_0\).

Figures 4–6 show the first harmonic of the lift normalized by \(\alpha_0\) vs \(\alpha_0\) for the three reduced frequencies. Because the system identification model is a time-domain model the first harmonic value was computed using a discrete Fourier transform of the time-domain solution. In these three figures it is apparent that a second order polynomial does not adequately capture the dynamic behavior of the system. For all three reduced frequencies, the normalized first harmonic of the lift coefficient was underpredicted using the second order model. This model also underpredicted the degree of nonlinearity shown in the solution. The third order model does a better job of quantitatively and qualitatively matching the full CFD model.

A possible reason for the difference in first harmonic coefficient value for the system identification model and the harmonic balance solution is the method in which this coefficient was computed for the time-domain system identification model. For the system identification model, a discrete Fourier transform of the time history data was used to identify the first harmonic coefficient value. The coefficient data presented for the harmonic balance solution were directly taken from the solution. Of course it is also true, especially in the regions where the behavior is nonlinear, that the differences are
due to the assumptions made (number of POD modes kept and polynomial order) in building the reduced order system identification model.

Figures 7–9 show the discrete Fourier transform of the system identification model for an $\alpha_0$ of 8 deg and three reduced frequencies. Also shown are the magnitudes of the zeroth, first, and second harmonics as computed by the harmonic balance CFD solution. From the figures it is apparent that the system identification model response contains contributions from higher harmonics for this angle of attack. In comparison Fig. 10 shows the discrete Fourier transform of the system identification model for an $\alpha_0$ of 0.1 and a reduced frequency of 0.2. In this plot the only significant harmonics are the zeroth and first. For all three reduced frequencies, the order of magnitude of the harmonic components (zeroth, first and second) from the system identification model compare well with those obtained using the CFD solution.
It is interesting to note that at the higher reduced frequencies, Figs. 8 and 9, the contribution of the second and third harmonic components are of equal order whereas for the lower reduced frequency result in Fig. 7 the second harmonic is almost a full order of magnitude larger than the third harmonic. This result is not intuitive because the harmonic balance method used to generate the training data used two harmonics.

In Figs. 11–16, the maximum and peak to peak lift coefficients are plotted as function of \( \alpha_0 \) for three reduced frequencies. Overall the system identification model results compare well with the CFD model. At the lowest reduced frequency, \( \tilde{\omega} = 0.2 \), the system identification model slightly underpredicts both the maximum lift coefficient and the peak to peak lift. For the higher reduced frequencies, the current model results are slightly higher than the CFD results for angles of attack greater than 1 deg.

V. Discussion

In the Results section of this paper, coefficient of lift data is computed using the system identification method developed in the present paper and compared to the data computed using a harmonic balance CFD method. Although the results computed with the reduced order method compare well with the CFD results, there is, of course, always room for improvement. This section will address what has been tried by the current authors in the way of improving the model along with some ideas which have not been explored but appear promising.

At first glance, two obvious improvements would be the inclusion of more modes in Eq. (18) and higher order polynomials in Eq. (27). In the results presented in the previous section, the modal expansion for \( U \) was truncated at three modes. When the fourth mode was added to the analysis the results changed very little. If modes higher than the fourth were added the results actually decreased in accuracy. Although this behavior is certainly not desirable, the degradation of reduced order POD models with the addition of lower energy modes has been observed in previous studies [45]. One possible explanation for this behavior is that the higher modes possess length scales which are comparable to the CFD model grid spacing. These modes would then be inaccurate and would have the effect of adding noise to the training data. This would negatively affect the accuracy of the computed model coefficients. A more likely explanation for the behavior is that because the coefficients are computed using a least squares solution, for a fixed number of training points, increasing the number of unknown coefficients to be identified decreases the accuracy of the computed coefficients.

It should also be mentioned that a separate study was performed by the current authors which examined the convergence, with respect to the number of modes kept in the expansion, of the lift coefficient for a small value of \( \alpha_0 \) and reduced frequencies of 0.2, 0.6, and 1.0. The identified model from this study and the model used to compute the data in the Results section of this paper used the same POD modes. However the system identification model used in the convergence study was trained with linear data (i.e., data which can be thought of as being a small perturbation about \( U_0 \)) only. This study revealed that many more modes (18) could be kept in the model before the results started to decrease in accuracy. However the addition of modes greater than the third did not significantly improve these linear results.
A second way to improve upon the results would be to include higher order polynomials in the system identification model. Because the function $F$ in the current model can be thought of in terms of a Taylor series expansion about some known value, truncation at a finite number of terms will introduce an error which will grow as you move away from this value. In the context of the current paper the “value” is the airfoil steady flow solution $U_0$. The inclusion of more terms in the Taylor expansion will, in theory, allow the solution to move further away from $U_0$ and still be able to accurately represent the underlying nonlinear function. Of course the introduction of higher order polynomials can cause a new set of problems. Higher order polynomials tend to extrapolate poorly and may be less smooth between trained data points. This could cause stability problems in the identified ODE. Also the system of linear
equations used to determine the coefficients for a high order polynomial is often ill conditioned which can lead to large errors in the coefficients. Finally the inclusion of higher order polynomials also increases the number of coefficients to be computed. As was mentioned in the previous paragraph, this could also decrease the accuracy of the model for a fixed number of training points.

In traditional finite difference methods, which are based upon Taylor series expansion, the problem of moving too far away from the initial known value is dealt with using small steps in the neighborhood of the initial point and continuously reevaluating the coefficients (derivatives of the Taylor series terms). With this in mind, a possible method for improving the performance of the function \( F \) in Eq. (26) might be to allow the coefficients to be functions of some time dependent parameters. This is the method used by Chen et al. [46] to model nonlinear aeroelastic systems. In this work the coefficients of the equations of motion (mass, damping, stiffness, and aerodynamic forces) were assumed to be functions of various parameters such as system response, sensor output, and actuator input. This allowed various nonlinearities to be included without the need for solving sets of nonlinear equations.

In the current paper one could imagine that the coefficients might be functions of both \( \dot{\alpha} \) and \( \alpha \). If the Mach number was to be varied then the coefficients would also be a function of Mach number. The assumption that the coefficients are functions of both \( \dot{\alpha} \) and \( \alpha \) is equivalent to adding terms to \( F \). For example, if \( F \) was originally a cubic function of \( \dot{\alpha} \), the assumption that the coefficients of \( F \) are to be quadratic in \( \dot{\alpha} \) and linear in \( \alpha \) results in additional terms in \( F \) of order \( \dot{\alpha}^2 \), \( \dot{\alpha} \alpha \), \( \alpha^2 \), and \( \alpha^3 \). Of course we already know that the nonlinear dynamic responses of the airfoil were represented. The results were computed using the third order model for two of the three reduced frequencies (0.6 and 1.0). The third order model was able to predict accurately, for all three reduced frequencies, the nonlinear trend in the plots of the zeroth and first harmonic components of the lift coefficient versus magnitude of pitch. Quantitatively the system identification model appeared to slightly underpredict the strength of the nonlinear dependence of the lift coefficient on the magnitude of the pitch oscillation. This is especially apparent in the plots of the maximum lift coefficient versus pitch magnitude for the two highest reduced frequencies. Still, the system identification results shown Figs. 11–16 differ by no more than 15% for all of the pitch magnitudes and reduced frequencies studied.

Further work is currently in progress to extend the current ideas to viscous flows and to three-dimensional models.

VI. Concluding Remarks

In this paper a reduced order, nonlinear system identification methodology was presented in the context of computational fluid dynamics. The method was used to identify a set of nonlinear, first order ordinary differential equations for the fluid flow POD modal coordinates. The unknown coefficients in the nonlinear ODEs are computed using training data from a full CFD model and a linear least squares solution.

The technique was tested on the two-dimensional, supersonic NLR-7301 airfoil. The training data needed to identify the unknown coefficients in the identified model were computed using a harmonic balance solution of the compressible Euler equations. Results computed using the identified model were compared to those from the full CFD model for a forced pitch oscillation of the airfoil. The magnitude of these pitch oscillations was chosen such that both linear and nonlinear dynamic responses of the airfoil were represented. The results were computed for three separate reduced frequencies (0.2, 0.6, and 1.0) and for one Mach number (\( M = 0.764 \)). Results were presented for models which use second and third order polynomial expressions in the nonlinear ODE. Three POD modes were used to represent the CFD flow variables.

Overall the system identification model results compared well with the full CFD model. It was shown that the third order model was more accurate in predicting the dynamic lift coefficient behavior (maximum \( C_l \) and peak to peak \( C_l \)) for all of the reduced frequencies studied. The behavior in the mean lift coefficient as a function of the magnitude of the pitch oscillation was also predicted more accurately using the third order model for two of the three reduced frequencies (0.6 and 1.0). The third order model was able to predict accurately, for all three reduced frequencies, the nonlinear trend in the plots of the zeroth and first harmonic components of the lift coefficient versus magnitude of pitch. Quantitatively the system identification model appeared to slightly underpredict the strength of the nonlinear dependence of the lift coefficient on the magnitude of the pitch oscillation. This is especially apparent in the plots of the maximum lift coefficient versus pitch magnitude for the two highest reduced frequencies. Still, the system identification results shown Figs. 11–16 differ by no more than 15% for all of the pitch magnitudes and reduced frequencies studied.

Further work is currently in progress to extend the current ideas to viscous flows and to three-dimensional models.

Acknowledgement

This work was funded in part by AFOSR grant “A Study of Uncertainty in Aeroelastic Systems,” with Clark Allred as the program director.

References


C. Pierre
Associate Editor