On the use of modal derivatives for nonlinear model order reduction

Oliver Weeger\(^1,\,*\), Utz Wever\(^2\) and Bernd Simeon\(^3\)

\(^1\) Singapore University of Technology and Design, 8 Somapah Road, Singapore 487372, Singapore

\(^2\) Siemens AG, Corporate Technology, Otto-Hahn-Ring 6, 81739 Munich, Germany

\(^3\) TU Kaiserslautern, Department of Mathematics, P.O. Box 3049, 67653 Kaiserslautern, Germany

SUMMARY

Modal derivatives are an approach to compute a reduced basis for model order reduction of large-scale nonlinear systems that typically stem from the discretization of partial differential equations. In this way, a complex nonlinear simulation model can be integrated into an optimization problem or the design of a controller, based on the resulting small-scale state-space model. We investigate the approximation properties of modal derivatives analytically and thus lay a theoretical foundation of their use in model order reduction, which has been missing so far. Concentrating on the application field of structural mechanics and structural dynamics, we show that the concept of modal derivatives can also be applied as nonlinear extension of the Craig-Bampton family of methods for substructuring. We furthermore generalize the approach from a pure projection scheme to a novel reduced-order modeling method that replaces all nonlinear terms by quadratic expressions in the reduced state variables. This complexity reduction leads to a frequency-preserving nonlinear quadratic state-space model. Numerical examples with carefully chosen nonlinear model problems and 3-dimensional nonlinear elasticity confirm the analytical properties of the modal derivative reduction and show the potential of the proposed novel complexity reduction methods, along with the current limitations.

Copyright © 2016 John Wiley & Sons, Ltd.

KEY WORDS: model order reduction; modal derivatives; reduced basis method; nonlinear dynamics; nonlinear state-space model

1. INTRODUCTION

Model order reduction (MOR) is of increasing interest for many large-scale applications in computational engineering. Such reduction methods are applied to both stationary and transient problems and make it possibly to embed the detailed model of a component into the context of system simulation. For instance, in automotive applications, specific components like elastomer mounts are first simulated separately by means of finite elements, and then MOR is applied to generate a small-scale model that captures the essential dynamics and can be coupled with a vehicle simulation or an optimization loop. A recent survey of the state-of-the-art in MOR is given in [1].

For linear problems, model order reduction methods are quite mature. Specifically in state-space formulations, balanced truncation is a favourite method [2, 3, 4, 5, 6], since it guarantees the observability and the stability of the system [7]. For the pure reduction of linear systems, Krylov-based or modal reduction is a well-known method family, and the Craig-Bampton methods can...
be applied for substructuring of large systems [8, 9]. For structural mechanics and dynamics these are the methods of choice, because they preserve the leading eigenfrequencies and the frequency response behavior of a structure [10]. Even commercial tools such as ANSYS [11] and NX Nastran [12] offer these methods. The degree of reduction can be specified in terms of the number of leading eigenvalues.

For nonlinear equations, research is still going on. One method is, e.g., the reduced basis method, see [13, 14]. In the present paper, we follow the concept of modal derivatives (MDs) as introduced in [15, 16]. It has its merits in particular in nonlinear structural mechanics and dynamics because it preserves the leading nonlinear eigenfrequencies and has been successfully tested for several applications [17, 18, 19, 20, 21].

The background and origin of our work is nonlinear frequency analysis. In [22] we have discussed nonlinear steady-state frequency analysis for an Euler-Bernoulli beam and used the so-called harmonic balance for analysis, where in each node of the mesh a Fourier expansion is set up. When extended to 3D structures, however, harmonic balance becomes expensive, and in [23, 24] we have discussed this situation in detail. By applying a model order reduction based on modal derivatives, we succeeded in deriving a robust algorithm that can handle large deformation and nonlinear material laws.

Though being found in many applications, modal derivatives still lack a theoretical foundation. It is the purpose of this paper to close this gap by investigating their approximation properties and analytically validating their usage as reduced basis. Furthermore, we propose a novel model for complexity reduction that is based on the concept of modal derivatives and results in a quadratic representation of the reduced system in terms of the modal coordinates. Generated in a pre-processing run, this approximation serves as reduced system for the static structural mechanical problem or as a state-space model for the corresponding dynamical problem. The signature mark of this quadratic state-space model is the preservation of the leading frequencies. In our experience, it performs always better than a linear state space model and sometimes even as well as the projected system derived from modal derivatives.

The outline of this contribution is as follows: After this introductory section we give a general overview on the use of model order reduction for nonlinear dynamical systems in Section 2. Also the reduced state-space model is discussed as a standardized form for dynamical systems. Next we introduce the concept of modal derivatives in Section 3 and investigate their approximation properties. We show that the idea can also be transferred to Craig-Bampton approaches. In Section 4 we apply MDs for model order reduction and verify the convergence properties numerically for a nonlinear elasticity application. We extend the usage of modal derivatives to a novel nonlinear complexity reduction method in Section 5. This method is applied to the static nonlinear system and for generating a reduced quadratic state-space model. The new method is carefully investigated for several applications. A summary and conclusion round up the paper in Section 6.

2. MODEL ORDER REDUCTION AND STATE-SPACE MODEL

In this Section we briefly introduce projection methods for model order reduction of nonlinear dynamical systems. We also outline how these reduced order models can be translated into the context of state-space representations.

2.1. Model order reduction of nonlinear systems

For the solution of time-dependent nonlinear partial differential equation (PDE) problems, which arise in many engineering and scientific fields, usually a spatial semi-discretization is applied, for example using the finite element method. For simplicity, we consider here structural dynamics as specific application field but emphasize that our results are not restricted to this case.

The semi-discretization leads to the solution of an $n$-dimensional system of nonlinear ordinary differential equations (ODEs)

$$\mathbf{M}\ddot{\mathbf{d}}(t) + \mathbf{f}(\mathbf{d}(t)) = \mathbf{b}(t).$$

(1)
Here, \( \mathbf{d}(t) \in \mathbb{R}^n \) is the unknown vector of deformation degrees of freedom (DOFs), \( \mathbf{M} \) the mass matrix, \( \mathbf{f} \) the nonlinear internal force vector and \( \mathbf{b}(t) \) the external force vector.

To guarantee an accurate solution of the problem and capture all geometrical features of complex structures, finite element discretizations typically contain a large number of DOFs, which means that in practical engineering applications the size \( n \) of the system \( (1) \) may be around hundreds of thousands or even millions of DOFs.

In order to reduce the number of unknowns significantly and speed-up the numerical solution process, model order reduction methods can be applied. A frequent technique employed for this purpose are so-called reduced basis methods [13]. Using a suitable basis \( \mathbf{Q} \in \mathbb{R}^{n \times r} \) of an \( r \)-dimensional subspace of \( \mathbb{R}^n \), where \( r \ll n \), the time-dependent displacement vector \( \mathbf{d}(t) \) is expressed in reduced coordinates \( \mathbf{q}(t) \) as:

\[
\mathbf{d}(t) = \mathbf{Q}\mathbf{q}(t). \tag{2}
\]

Then the initial system of governing equations is projected onto the reduced basis:

\[
\begin{align*}
\mathbf{Q}^T \mathbf{M} \mathbf{Q} \ddot{\mathbf{q}}(t) + \mathbf{Q}^T \mathbf{f}(\mathbf{Q}\mathbf{q}(t)) &= \mathbf{Q}^T \mathbf{b}(t),
\end{align*}
\]

This reduced, \( r \)-dimensional ODE system is abbreviated as

\[
\ddot{\mathbf{q}}(t) + \mathbf{K}\mathbf{q}(t) = \mathbf{b}(t) \tag{4}
\]

and can then be solved instead of the full-dimensional system in order to determine the reduced coordinates \( \mathbf{q} \) and subsequently the displacements \( \mathbf{d} \).

The goal of model order reduction is to significantly decrease the computational effort for determining the displacements, but at the same time still to obtain an accurate and reliable solution by approximating \( \mathbf{d} \) as \( \mathbf{Q}\mathbf{q} \). This means that a suitable basis for projection, i.e. the projection matrix \( \mathbf{Q} \), should be small \( (r \ll n) \), easy to compute, as general and as accurate as possible.

Linear terms in \( (4) \), such as \( \mathbf{M} \), can be pre-computed offline based on the projection matrix and the original system matrices, which makes the online evaluation of linear reduced order models very efficient. However, in the nonlinear setting, terms such as the internal force vector \( \mathbf{f} \) still have to be evaluated based on the displacement of the full model \( \mathbf{d} = \mathbf{Q}\mathbf{q} \). To truly decrease the computational complexity of the model order reduction, one must also find a way to directly approximate the nonlinear force terms and directly evaluate \( \mathbf{f}(\mathbf{q}) \), see Section 5.

2.2. State-space model

Linear FEM models

\[
\mathbf{M}\ddot{\mathbf{d}}(t) + \mathbf{K}\mathbf{d}(t) = \mathbf{b}(t) \tag{5}
\]

can be derived from linear problems or as linearization of nonlinear problems such as \( (1) \) with \( \mathbf{K}\mathbf{d} = \mathbf{f}(\mathbf{d}) \). They are often used within system simulation, where the input-output behaviour of only a few state variables must be known, and therefore transformed into reduced form

\[
\ddot{\mathbf{q}}(t) + \mathbf{K}\mathbf{q}(t) = \mathbf{b}(t), \tag{6}
\]

just as outlined above for the nonlinear case, here with \( \mathbf{K} = \mathbf{Q}^T \mathbf{KQ} \). Commercial dynamical system software such as Dymola [25], MATLAB Simulink [26] or Amesim [27] offer a normalized interface for importing these models. The normalized interface is given in terms of the so-called state-space model:

\[
\begin{align*}
\dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}, \\
\mathbf{y} &= \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u}.
\end{align*}
\]

Here \( \mathbf{x} \) is the state vector, \( \mathbf{u} \) is the input vector, \( \mathbf{y} \) is the output vector, \( \mathbf{A} \) is the state or system matrix, \( \mathbf{B} \) is the input matrix, \( \mathbf{C} \) is the output matrix, and \( \mathbf{D} \) is the feedthrough matrix.
The input \( u \) is often defined as the load \( b(t) \) in equation (6) and \( y \) may be the state at some selected (critical) positions. Commercial FEM tools such as ANSYS [11] or NX Nastran [12] are able to generate state-space models [7]. The degree of reduction in terms of the number of eigenvectors is freely selectable.

The linear system (6) can be transformed into a state-space model: using the definitions
\[
\tilde{x}_1 = q, \quad \tilde{x}_2 = \dot{q},
\]
and
\[
u = b, \quad y = \begin{pmatrix} d & d & b \end{pmatrix} \begin{pmatrix} \dot{d} \\ d \\ \dot{d} \end{pmatrix}
\]
the corresponding first equation in the state-space model (7) is given by
\[
\begin{pmatrix} \dot{\tilde{x}}_1 \\ \dot{\tilde{x}}_2 \end{pmatrix} = \begin{pmatrix} 0 & -I \\ -M^{-1}K & 0 \end{pmatrix} \begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_2 \end{pmatrix} + \begin{pmatrix} 0 \\ -M^{-1}Q^T \end{pmatrix} u.
\]

The second equation should be formulated such that the output is given in the original coordinates:
\[
\begin{pmatrix} d' \\ \dot{d} \\ b \end{pmatrix} = \begin{pmatrix} Q & 0 & 0 \\ 0 & Q & Q \\ 0 & 0 & \tilde{M}^{-1} \tilde{K} \end{pmatrix} \begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_2 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ Q \tilde{M}^{-1}Q^T \end{pmatrix} u.
\]

Deriving the linear reduced state-space model as an approximation of equation (5) is state of the art. In Section 5.2 we discuss a quadratic state-space model as an approximation of the nonlinear equation (1).

3. DEFINITION AND PROPERTIES OF MODAL DERIVATIVES

Modal truncation using a reduced basis consisting of a set of linear eigenmodes (EMs) is a popular and wide-spread projection method for model order reduction of linear systems. However, it is typically not very accurate for the reduction of nonlinear systems. To obtain a suitable reduction basis for nonlinear applications, a modal basis can be enhanced by so-called modal derivatives (MDs). This approach was first presented in [15, 16] and has since been used in many applications [17, 20, 21, 23]. Similar methods have also been presented for enhancement of POD reduction in the context of fluid dynamics [28].

In the following, we want to review the derivation and definition of modal derivatives and furthermore extend the theoretical framework of their approximation properties [24], which qualifies modal derivatives as reduced basis for the solution of nonlinear PDE problems.

3.1. Eigenmodes and linear modal truncation

For a semi-discretized second order problem such as (1) the corresponding symmetric linear eigenvalue problem can be formulated as
\[
\omega_i^2 M \phi_i = K \phi_i, \quad i = 1, \ldots, n \quad \Leftrightarrow \quad M \Phi \Lambda = K \Phi.
\]

Here \((\omega_i, \phi_i), \quad i = 1, \ldots, n\) are the pairs of eigenfrequencies and eigenmodes, \(\lambda_i = \omega_i^2\) the eigenvalues and \(\Lambda = \text{diag}(\lambda_i)_{i=1,\ldots,n}\), \(\Phi = (\phi_i)_{i=1,\ldots,n}\) the matrices of eigenvalues and eigenmodes. Typically, the eigenvalues are ordered by magnitude, i.e. \(\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n\), and the eigenmodes are normalized such that
\[
\phi_i^T M \phi_j = \delta_{ij}, \quad \phi_i^T K \phi_j = \delta_{ij} \lambda_i, \quad \Leftrightarrow \quad \Phi^T M \Phi = I, \quad \Phi^T K \Phi = \Lambda.
\]
Thus the eigenmodes are an orthonormal basis of \(\mathbb{R}^n\) w.r.t. the inner product weighted by \(M\).

The discretized linear static problem
\[
Kd = b
\]  
(14)
can be transformed onto the eigenmodes using \(d = \Phi q\):
\[
\Phi^T K \Phi q = \Phi^T b \iff \Lambda q = \Phi^T b.
\]  
(15)
Then the vector of modal coordinates can easily be computed as
\[
q = \Lambda^{-1} \Phi^T b
\]  
(16)
and for the displacement it follows
\[
d = \Phi q = \Phi \Lambda^{-1} \Phi^T b.
\]  
(17)

Formulating the right hand side vector as \(b = M \Phi \Lambda p\) in terms of the basis \(\Phi\) of \(\mathbb{R}^n\) and the modal parameter vector \(p\) then yields \(q = p\). Thus the solution of the linear static problem (14) with right hand side \(b = M \Phi \Lambda p\) can be directly expressed in terms of the parameter vector as \(d = \Phi p\).

Similar, also the solution of the linear transient problem
\[
M \ddot{d}(t) + Kd(t) = b(t)
\]  
(18)
with harmonic external excitation \(b(t) = b \cos \omega t\) is \(d(t) = d \cos \omega t\) and can be projected using \(d = \Phi q\):
\[
-\omega^2 \Phi^T M \Phi q + \Phi^T K \Phi q = \Phi^T b \iff -\omega^2 I q + \Lambda q = \Phi^T b = \Lambda p.
\]  
(19)
Then this diagonalized system can be solved component-wise [29]
\[
q_i = \frac{\omega_i^2}{\omega_i^2 - \omega^2} p_i, \quad i = 1, \ldots, n \iff q = (\Lambda - \omega^2 I)^{-1} \Lambda p.
\]  
(20)
Here it can be seen easily that for \(\omega \to \omega_i\) it follows \(q_i \to \pm \infty\), i.e., resonance occurs and the amplitudes are growing to infinity.

Altogether, the diagonalization of the linear static and dynamic problems makes eigenmodes a fairly powerful and convenient basis for the projection of these problems, and thus motivates their use for reduction of linear systems: As it is not feasible to compute all eigenmodes in practical applications with \(n \approx 10^3 - 10^9\), only a subset of eigenmodes \(Q = \Phi_{\mathcal{R}} = (\phi_i)_{i \in \mathcal{R}} \in \mathbb{R}^{n \times r}\) with \(\mathcal{R} \subset \mathcal{I} = \{1, \ldots, n\}\), \(|\mathcal{R}| = r \ll n\) is selected as a reduced or truncated basis and the displacements are approximated as \(d_{\mathcal{R}} = \Phi_{\mathcal{R}} q_{\mathcal{R}}\) with \(q_{\mathcal{R}} \in \mathbb{R}^r\). This method is widely spread in engineering application and called modal truncation or modal reduction (MT/MR).

It can immediately be seen from (16) and (20) that the truncated modal basis \(\Phi_{\mathcal{R}}\) can exactly represent the solutions of (14) and (18) if \(p_i = 0 \forall i \in \mathcal{R}' = \mathcal{I} \setminus \mathcal{R}, \quad |\mathcal{R}'| = n - r\). This situation is, of course, more of academic rather than practical relevance. In general, in the static case the error of displacement is
\[
d - d_{\mathcal{R}} = \Phi q - \Phi_{\mathcal{R}} q_{\mathcal{R}} = \Phi_{\mathcal{R}}', q_{\mathcal{R}}',
\]  
(21)
with \(q_{\mathcal{R}}' \in \mathbb{R}^{n-r}\) and \(\Phi_{\mathcal{R}}' = (\phi_i)_{i \in \mathcal{R}'} \in \mathbb{R}^{n \times (n-r)}\). Thus it follows for the error in any suitable vector norm \(\| \cdot \|\):
\[
\|d - d_{\mathcal{R}}\| = \|\Phi_{\mathcal{R}}', q_{\mathcal{R}}', q_{\mathcal{R}}\| \leq \sum_{i \in \mathcal{R}'} \|\phi_i\| |q_i| \leq (n - r) \max_{i \in \mathcal{R}'} \|\phi_i\| \max_{i \in \mathcal{R}'} |q_i| = (n - r)\varepsilon\delta,
\]  
(22)
where \(\varepsilon = \max_{i \in \mathcal{R}'} |q_i|\) and \(\delta = \max_{i \in \mathcal{R}'} \|\phi_i\|\). Using the \(M\)-norm \((\|x\|_M^2 = x^T M x)\) it holds \(\delta = 1\) due to the normalization of eigenmodes.

3.2. Nonlinear problem and modal derivatives

In the previous Subsection we have seen that eigenmodes are an optimal basis for the reduction of linear problems where the right hand side is composed from a subset of eigenmodes. Extending this idea to the nonlinear regime, we are now looking for the solution of the static problem

\[ f(d) = b = M\Phi\Lambda p. \tag{23} \]

For the nonlinear force vector it shall hold that \( f \in C^2(\mathbb{R}^n, \mathbb{R}^n) \) and \( f(0) = 0 \). Moreover, \( \frac{df}{dp} = K(d) \) and \( K(0) = K \) denote the tangent stiffness matrices of the deformed and undeformed configuration respectively, which are assumed to be non-singular, i.e. \( \text{rank}(K(d)) = n \).

In general it is, of course, not possible to find a closed-form solution \( d(\mathbf{p}) \), but assuming that \( d \) is also \( C^2 \)-differentiable w.r.t. the parameter vector \( \mathbf{p} \) we can expand it into a Taylor series around \( \mathbf{p} = 0 \),

\[ d \equiv d(\mathbf{p}) = 0 + \frac{d}{d\mathbf{p}}|_0 \mathbf{p} + \frac{1}{2} \left( \frac{d^2d}{dp^2} |_0 : \mathbf{p} \right) \mathbf{p} + \mathcal{O}(\|\mathbf{p}\|^3). \tag{24} \]

Here, \( \frac{d^2d}{dp^2}(0) \) is a third-order \( n \times n \times n \)-tensor and \( (\cdot : \cdot) \) denotes the product between a third-order tensor and a vector (first-order tensor), which results in a \( n \times n \)-matrix (second-order tensor).

In order to find the derivatives in (24) we continue as in [17] by differentiating both sides of (23) with respect to the modal parameter vector \( \mathbf{p} \), which results in

\[ \frac{df}{dd} \frac{dd}{dp} = M\Phi\Lambda. \tag{25} \]

Evaluation at \( \mathbf{p} = 0 \) leads to

\[ K \frac{dd}{dp} |_0 = M\Phi\Lambda, \tag{26} \]

and from the linear eigenvalue problem (12) it follows

\[ \frac{dd}{dp} |_0 = \Phi, \tag{27} \]

i.e., the first derivatives are the linear eigenmodes, as could be expected from previous results.

Differentiating (25) once again we obtain

\[ \left( \frac{d^2f}{dd^2} : \frac{dd}{dp} \right) \frac{dd}{dp} + \frac{df}{dd} \frac{d^2d}{dp^2} = 0. \tag{28} \]

Evaluation at \( \mathbf{p} = 0 \) now yields

\[ \left( \frac{d^2f}{dd^2} |_0 : \Phi \right) \Phi + K \frac{d^2d}{dp^2} |_0 = 0, \tag{29} \]

where \( \frac{d^2f}{dd^2} |_0 = \frac{dK}{dd} |_0 =: \mathbb{H} \). Then the modal derivatives \( \frac{d^2d}{dp^2} |_0 = \frac{d\Phi}{dp} |_0 =: \Psi \) can be defined as the solutions of the system of equations

\[ K \Psi = - (\mathbb{H} : \Phi) \Phi. \tag{30} \]

Using the notation \( \Psi = (\psi_{ij})_{i,j=1,...,n} \) for the tensor of modal derivatives and exploiting that \( (\mathbb{H} : \phi_i) \) is the directional derivative of \( K \) w.r.t. \( \phi_i \), i.e.

\[ (\mathbb{H} : \phi_i) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} (K(\varepsilon\phi_i) - K) = \nabla_{\phi_i} K, \tag{31} \]

we define in vector notation:
Definition 1 (Modal derivatives)
The solution vectors $\psi_{ij} \in \mathbb{R}^n$ of
\[
K \psi_{ij} = -\nabla \phi_i K \phi_j, \quad i, j = 1, \ldots, n.
\] (32)
are called modal derivatives.

Remark 1
Originally in [15, 16] modal derivatives were actually defined by differentiating the nonlinear eigenvalue problem w.r.t. the linear eigenmodes. This yields a more complicated expression, as it also involves the derivatives of eigenfrequencies. However, [15, 16] also included the definition as given above by neglecting these derivatives of eigenfrequencies, referring to the so-called “definition without mass consideration”. Basically all authors applying modal derivatives for model reduction have been using this simpler definition, cf. [17, 18, 20, 21, 23], and [17] already used a similar procedure as above to derive it.

Lemma 1 (Symmetry of modal derivatives)
Modal derivatives are symmetric, i.e.
\[
\psi_{ij} = \psi_{ji}, \quad i, j = 1, \ldots, n.
\]

Proof
To proof the symmetry of modal derivatives we start with the symmetries of $H$. Since $f \in C^2(\mathbb{R}^n, \mathbb{R}^n)$ it holds according to Schwarz’s theorem (also called Clairaut’s theorem):
\[
H_{ijk} = H_{ikj} = H_{jik} \iff \frac{\partial^2 f_i}{\partial d_j \partial d_k} = \frac{\partial^2 f_i}{\partial d_k \partial d_j} = \frac{\partial^2 f_j}{\partial d_i \partial d_k}, \quad i, j, k = 1, \ldots, n.
\]
Then it follows for $a, b \in \mathbb{R}^n$:
\[
(\mathbb{H} : a)b = \sum_{j,k=1}^n H_{ijk}a_kb_j = \sum_{j,k=1}^n H_{ikj}a_kb_j = (\mathbb{H} : b)a,
\]
which means that the right hand side vectors in Def. 1 are symmetric w.r.t. $i$ and $j$:
\[
\nabla \phi_i K \phi_j = (\mathbb{H} : \phi_i)\phi_j = (\mathbb{H} : \phi_j)\phi_i = \nabla \phi_j K \phi_i.
\]
As $K$ is also symmetric this yields the symmetry of modal derivatives, $\psi_{ij} = \psi_{ji}$. \hfill \Box

Remark 2 (Computing modal derivatives)
In order to actually compute a set of modal derivatives based on Def. 1 one needs, of course, to solve the generalized eigenvalue problem (12) for the $r$ eigenmodes $\{\phi_i\}_{i \in \mathbb{R}}$ first.

Then $r$ directional derivatives defined in (31) have to be computed. However, assembling $\mathbb{H}$ and evaluating $(\mathbb{H} : \phi_i)$ is not feasible in practical applications, as either a closed-form representation for finite element assembly of $\mathbb{H}$ might not be available, and moreover computing time and storage requirements are very high or even unfeasible as $\mathbb{H}$ is a $n \times n \times n$-tensor. Nevertheless, we can make use of the definition as directional derivative and approximate $\nabla \phi_i K \phi_j$ as a finite difference. Using forward differences this means that $r$ nonlinear FE assemblies of the tangent stiffness matrices $K(\varepsilon \phi_i)$ have to be carried out.

Finally the linear systems of (32) can be solved. Exploiting the symmetry of MDs $r(r + 1)/2$ linear solves are required. If the FE system size $n$ is moderate such that $K$ can still be factorized using sparse QR- or LU-decomposition, these solves can be carried out efficiently since the factorization of $K$ just has to be applied to multiple right hand side vectors $-\nabla \phi_j K \phi_i$. 

Copyright © 2016 John Wiley & Sons, Ltd.

Prepared using nmeauth.cls
DOI: 10.1002/nme
3.3. Approximation properties of modal derivatives

Having defined modal derivatives we next examine their approximation properties and usage as reduced bases in more detail [24]. In particular, we show that eigenmodes \( \Phi \) and modal derivatives \( \Psi \) can be used to obtain a second-order approximation of the solution of a nonlinear static problem:

**Theorem 1** (Quadratic approximation using modal derivatives)

The representation of displacements as quadratic Taylor series expansion

\[
d_2(p) := \Phi p + \frac{1}{2} (\Psi : p) p
\]

is a second-order approximation of the solution \( d \) of the static nonlinear elasticity problem

\[
f(d) = M\Phi \Lambda p,
\]

in the neighbourhood of \( p = 0 \), i.e.

\[
\| f(d) - f(d_2(p)) \| = O(\|p\|^2).
\]

**Proof**

We start by developing \( f(d) \) as second-order Taylor series expansion around \( d = 0 \):

\[
f(d) = 0 + Kd + \frac{1}{2} (H : d)d + O(\|d\|^3).
\]

Then we substitute our ansatz for displacements

\[
d_2(p) = \Phi p + \frac{1}{2} (\Psi : p) p
\]

into the Taylor series expansion:

\[
f(d(p)) = K(\Phi p + \frac{1}{2} (\Psi : p)p) + \frac{1}{2} (H : (\Phi p + \frac{1}{2} (\Psi : p)p)) (\Phi p + \frac{1}{2} (\Psi : p)p) + O(\|d_2(p)\|^3)
\]

\[
= K\Phi p + \frac{1}{2} K(\Psi : p)p + \frac{1}{2} ((H : (\Phi p)) (\Phi p) + O(\|p\|^3)) + O(\|p\|^3)
\]

\[
= K\Phi p + \frac{1}{2} K(\Psi : p)p - \frac{1}{2} K(\Psi : p)p + O(\|p\|^3)
\]

\[
= K\Phi p + O(\|p\|^3) = M\Phi \Lambda p + O(\|p\|^3) = f(d) + O(\|p\|^3).
\]

Here we have used that \( O(\|d_2(p)\|^3) = O(\|p\|^3) \), the definition of modal derivatives Def. 1 and the linear eigenvalue problem (12).

In order to have a good approximation of \( f \) it is clear from Theorem 3.3 that it must hold \( \|p\| < 1 \), which restricts the validity of the method to right hand sides \( b = M\Phi \Lambda p \) fulfilling this condition \( (p = \Lambda^{-1} \Phi^T b) \).

As in 3.1, we now select at subset of eigenmodes \( \Phi_{\mathcal{R}} \) together with corresponding modal derivatives \( \Psi_{\mathcal{R}R} = (\psi_{ij})_{i,j \in \mathcal{R}} \). Now we can derive an error estimate for the reduced basis solution \( d_{\mathcal{R}} = d(p_{\mathcal{R}}) \) of the nonlinear problem (23):

**Theorem 2** (Error estimate for reduced basis approximation with modal derivatives)

For the error of the reduced basis representation of displacements

\[
d_{\mathcal{R}} = d(p_{\mathcal{R}}) := \Phi_{\mathcal{R}} p_{\mathcal{R}} + \frac{1}{2} (\Psi_{\mathcal{R}R} : p_{\mathcal{R}}) p_{\mathcal{R}}
\]

compared to the exact solution \( d \) of the nonlinear static problem

\[
f(d) = b = M\Phi \Lambda p
\]

with \( \|p\|_{\infty} < 1 \) it holds

\[
\| d - d_{\mathcal{R}} \| < (n-r)\varepsilon + \frac{1}{2}(n-r)^2\varepsilon^2\delta^* + r(n-r)\varepsilon\delta^* + O(\|p\|^3),
\]

with

\[
\varepsilon = \max_{i \in \mathcal{R}} |q_i| , \quad \delta = \max_{i \in \mathcal{R}} \| \Phi_i \| , \quad \delta^* = \max_{i,j \in \mathcal{R}} \| \psi_{ij} \|.
\]
Then it follows for the norm:
\[ \|d - d_R\| \leq \sum_{i \in I'} \|\Phi_i\| |p_i| + \frac{1}{2} \sum_{i \in I'} \sum_{j \in I'} \|\psi_{ij}\| |p_i| |p_j| + \sum_{i \in I'} \sum_{j \in I'} \|\psi_{ij}\| |p_i| |p_j| + O(\|p\|^3) \]
\[ < (n - r) \varepsilon + \frac{1}{2} (n - r)^2 \varepsilon^2 \delta^* + r(n - r) \varepsilon \delta^* + O(\|p\|^3). \]

In the last step we have used the condition \(\|p\|_\infty < 1\). \(\square\)

Theorem 2 shows that a reduced basis of eigenmodes together with their modal derivatives can be used to approximate the solution of a nonlinear problem with an error of order \(\varepsilon\) (as in the linear case, see Section 3.1) plus the terms of order 3 with respect to the components of \(p\). Thus the reduced basis can be considered as accurate for the nonlinear case as modal reduction is in the linear case.

The error estimates we established in Theorems 1 and 2 are only of qualitative nature and not sharp error bounds. They depend on the condition \(\|p\|_\infty < 1\), which seems very restrictive for practical applications, as well as the constants \(\varepsilon, \delta, \delta^*\). However, there is typically a problem-dependent relationship between the scales of eigenvalues \(\Lambda\), eigenmodes \(\Phi\), modal derivatives \(\Psi\), and \(p\). It could be used to determine problem-specific error bounds, which may in many cases be sharper, with an increased range of validity for \(\|p\|\). This is also indicated by the numerical studies in Section 5.3. However, for some problems with a slow decay of \(p\) the error bounds might also be meaningless.

3.4. Extension to Craig-Bampton methods

For the simulation and control of the dynamics of elastic multi-body systems [30], such as spacecraft structures, vehicles, robots or biomechanical systems, model order reduction also plays an important role [9]. The Craig-Bampton method (CBM) is a powerful and popular means of model reduction in this context [31,32].

The main idea behind CBM is to subdivide a large system into smaller components and compute reduced (modal) basis representations for these individual components and their interaction at interfaces. Focusing on one component, this means that the linear equation of motion can be partitioned into inner DOFs \(d_i \in \mathbb{R}^{n_i}\) and boundary DOFs \(d_b \in \mathbb{R}^{n_b}\) with \(n_i + n_b = n\):

\[ M \ddot{d} + Kd = b(t) \quad \iff \quad \begin{pmatrix} M_{ii} & M_{ib} \\ M_{bi} & M_{bb} \end{pmatrix} \begin{pmatrix} d_i \\ d_b \end{pmatrix} + \begin{pmatrix} K_{ii} & K_{ib} \\ K_{bi} & K_{bb} \end{pmatrix} \begin{pmatrix} d_i \\ d_b \end{pmatrix} = \begin{pmatrix} 0_i \\ b_b(t) \end{pmatrix}. \quad (38) \]

Typically the number of boundary DOFs, where the component interacts with other components of the system, is much smaller than the number of inner DOFs \((n_b \ll n_i)\).

Then the so-called Craig-Bampton transformation is applied using fixed-interface modes \(\Phi_{iR}\) and constraint modes \(\Phi_{ib}\):

\[ d = Qq \quad \iff \quad \begin{pmatrix} d_i \\ d_b \end{pmatrix} = \begin{pmatrix} \Phi_{iR} & \Phi_{ib} \\ 0_{ib} & I_{bb} \end{pmatrix} \begin{pmatrix} q_{iR} \\ q_{ib} \end{pmatrix}. \quad (39) \]
Here \( \Phi_{R^i} \subset \{1, \ldots, n_i\} \) is a subset of eigenmodes of the interior eigenvalue problem with fixed interfaces:

\[
- \omega_j^2 M_{ii} \phi_{ij} + K_{ii} \phi_{ij} = 0, \quad j = 1, \ldots, n_i,
\]

and the constraint modes are computed as

\[
\hat{\Phi}_{ib} = -K_{ii}^{-1} K_{ib}.
\]

With (39) we have thus given another basis for model order reduction, as in general already introduced in Section 2.1, which is then used to project (38).

Moving on to the nonlinear problem in partitioned form:

\[
M \ddot{d} + f(d) = b(t) \Leftrightarrow 
\begin{pmatrix}
M_{ii} & M_{ib} \\
M_{bi} & M_{bb}
\end{pmatrix}
\begin{pmatrix}
\ddot{d}_i \\
\ddot{d}_b
\end{pmatrix}
+ 
\begin{pmatrix}
f_i(d) \\
f_b(d)
\end{pmatrix}
= 
\begin{pmatrix}
o_i \\
0\end{pmatrix},
\]

we can apply the same idea as in Section 3.2 to define and compute “modal derivatives” of the modes of the Craig-Bampton basis. Expanding the solution of the static problem

\[
\begin{pmatrix}
f_i(d) \\
f_b(d)
\end{pmatrix}
= 
\begin{pmatrix}
o_i \\
I_{bb} p_b
\end{pmatrix}
\]

with \( p_b \in \mathbb{R}^{n_b} \) as quadratic Taylor series then leads to

\[
\frac{d\dot{d}}{dp_b}\bigg|_0 = \begin{pmatrix}
\hat{\Phi}_{ib} \\
I_{bb}
\end{pmatrix},
\]

and the modal derivatives can be computed from

\[
\left( \frac{d^2 f}{d\dot{d}^2} \bigg|_0 : \begin{pmatrix}
\hat{\Phi}_{ib} \\
I_{bb}
\end{pmatrix} \right) \left( \begin{pmatrix}
\hat{\Phi}_{ib} \\
I_{bb}
\end{pmatrix} \right) + K \frac{d^2 \dot{d}}{dp_b^2}\bigg|_0 = 0.
\]

Then these modal derivatives together with the Craig-Bampton eigenmodes can be used for nonlinear model order reduction in exactly the same fashion as outlined above in Section 4.1. This opens another interesting application field for the reduction with modal derivatives.

### 4. Reduced Basis Method with Modal Derivatives

The approximation results presented in the previous section show that a combination of linear eigenmodes and modal derivatives as reduced basis has fairly promising properties to capture the essential dynamics of a nonlinear system. In particular, we have seen that the quadratic approximation using linear eigenmodes and modal derivatives is able to provide similar accuracy as modal truncation in the linear case.

#### 4.1. Modal derivative reduction method (MDR)

Projection methods using a reduced basis composed of a selection of linear eigenmodes \( \Phi_{R^i} \) and modal derivatives \( \Psi_{R^s} \) have so far already been applied successfully in nonlinear analysis [17, 20, 21, 23]. The linear projection for modal derivative reduction onto a basis of \( r \) eigenmodes and \( s = r(r+1)/2 \) modal derivatives with total basis size \( r' = r + s \) is given by the projection matrix

\[
Q \in \mathbb{R}^{n \times r'} : 
Q = \left( \Phi_{R^i}, \Psi_{R^s} \right) = \left( (\phi_{i}), i \in \mathbb{R}, (\psi_{ij}), i,j \in \mathbb{R}^* \right).
\]

Due to the symmetry of modal derivatives we only have to use \( s \) MDs \( \psi_{ij} \) with \( j \geq i \). Thus we use the notation \( \mathbb{R}^* = \{ij : i,j \in \mathbb{R}, j \geq i\}, |\mathbb{R}^*| = s \) for the set of double-indices \( ij \) referring to the modal derivatives used in the matrix

\[
\Psi_{R^*} \in \mathbb{R}^{n \times s}.
\]
With the reduced coordinate vector

\[
\mathbf{q} \in \mathbb{R}^{r'} \colon \mathbf{q} = \begin{pmatrix} (q_i)_{i \in \mathcal{R}} \\ (q_{ij})_{ij \in \mathcal{R}^*} \end{pmatrix} = \begin{pmatrix} \mathbf{q}_\mathcal{R} \\ \mathbf{q}_{\mathcal{R}^*} \end{pmatrix},
\]

(47)

the reduced order representation of displacements then reads

\[
d(q) = Q\mathbf{q} = \Phi_{\mathcal{R}} \mathbf{q}_{\mathcal{R}} + \Psi_{\mathcal{R}} \mathbf{q}_{\mathcal{R}^*} = \sum_{i \in \mathcal{R}} \phi_i q_i + \sum_{ij \in \mathcal{R}^*} \psi_{ij} q_{ij},
\]

(48)

and projection of nonlinear static or dynamical systems can be carried out as in Section 2.1.

Note that for this modal derivative reduction method there are \(r(r + 1)/2\) independent coefficients \(q_{ij}\) for the modal derivatives, while in the theoretical results shown above the coefficients of MDs are expressed in terms of \(p_i\), the coefficients of the projection of the right hand side vector \(\mathbf{b}\) onto the eigenmodes \(\phi_i\). To get the same representation of \(d_{\mathcal{R}}\) as in (36) we can set

\[
q_i = p_i, \quad i \in \mathcal{R}, \quad q_{ii} = \frac{p_i^2}{2}, \quad i \in \mathcal{R}, \quad q_{ij} = p_i p_j, \quad i, j \in \mathcal{R} : j > i.
\]

(49)

Due to this additional freedom the results of MDR can be expected to be at least as accurate as indicated by Theorem 2.

**Remark 3** (Linear independence of reduction basis)

For the reduced order systems to be solvable and non-singular it is necessary that the projection matrix \(Q\) as given in (46) has full rank, i.e. that its columns are linearly independent and actually form a basis of the reduced order space. While the eigenmodes \(\phi_i\) are linearly independent by construction [29], it cannot be guaranteed that the modal derivatives \(\psi_{ij}\) are linearly independent of each other and of the eigenmodes. In order to ensure that the projection matrix only contains linearly independent vectors as columns, a singular value decomposition (SVD) can be used to obtain an admissible projection matrix with (possibly less than) \(r'\) columns.

**Remark 4** (Selection of eigenmodes on modal derivatives)

If the full solution of the linear static problem \(Kd = b\) is available, the accuracy of a reduced solution using the modes \(\Phi_{\mathcal{R}}\) can easily be checked by computing the norm of the error

\[
e_{\mathcal{R}} := d - d(\mathbf{q}_{\mathcal{R}}) = d - \Phi_{\mathcal{R}} \mathbf{q}_{\mathcal{R}} = d - \Phi_{\mathcal{R}} \Lambda_{\mathcal{R}^{-1}} \Phi_{\mathcal{R}}^T b.
\]

(50)

Thus the selection of \(\mathcal{R}\) can be controlled by \(\|e_{\mathcal{R}}\|\) and the individual contributions of modes \(q_{\mathcal{R}}\).

Since the number of modal derivatives \(s\) associated with eigenmodes grows quadratically with their number \(r\) (\(s = r(r + 1)/2\)), it might not be desirable to include all MDs denoted by \(\mathcal{R}^*\) into the reduction basis, but only a selection thereof. In [33] some criteria have been presented for the estimation of contributions of MDs to the solutions of transient problems and thus provide a guideline for the selection.

### 4.2. Application study for nonlinear mechanics

Having introduced the modal derivative reduction method, we now want to confirm its applicability for a complex problem, the mechanical deformation of a solid described by nonlinear elasticity. Here, and also later in our numerical examples in Section 5.3, we employ isogeometric finite elements for spatial discretization using the open-source library G+SMo [34]. For more details on isogeometric methods, which are based on splines, and their properties we refer the reader to [35] [36]. However, since our theoretical results are based on the (semi-) discretized form of the underlying PDE, they are independent from the spatial discretization method.

We apply the modal derivative reduction method to the static simulation of a rubber boot, as it is for instance used in cars. The complex geometry of the rubber boot is modelled as a spline volume and shown in Fig. 1(a).

For the nonlinear elasticity simulation the part is fixed at the circular back/bottom end and a surface load is applied to the front/top end (on the upper/lower row of Fig. 1). In order to study...
Figure 1. Rubber boot and its mechanical deformation resulting from nonlinear finite element analysis

Figure 2. Relative $L^2$-errors of linear analysis (lin.) and reduction methods (MT, MDR) compared to a full nonlinear analysis (NL) of the deformation of the rubber boot

The convergence behaviour of the reduction methods, the initial surface load is multiplied with a load factor in the range $[0.01, 5]$ and increased incrementally. Within this scenario we perform a linear finite element analysis (lin.), a nonlinear finite element analysis (NL), reduced linear analysis with modal truncation with first $r = 10$ eigenmodes (lin. MT), reduced nonlinear analysis with MT with first $r = 10$ eigenmodes (NL MT), and reduced nonlinear analysis with MDR with first $r = 10$ eigenmodes and corresponding MDs (NL MDR).

The results of the full nonlinear FEM simulation with $n = 3116$ DOFs and NURBS degree $p = 3$ (NL) are then considered as the reference deformation. Deformed configurations for load factors 2 and 5 are visualized in Fig. 1(b) and (c). In the last load step buckling occurs already.
In Fig. 2 we have evaluated the relative errors of the $L^2$-norms of the displacement compared to the fully nonlinear simulation. The curve lin./NL confirms that the relative error of a linear analysis compared to a nonlinear analysis is of order $p$, i.e. the absolute error is of order $p^2$. Here the difference becomes significant around $p \geq 0.5$, when the relative error becomes $\geq 4\%$. The curve lin. MT shows the relative error of modal truncation with the first $r = 10$ eigenmodes for the linear problem. As outlined in Section 3.1, the absolute error is of order $p$ and thus the relative error is constant, here it is 1.9%. NL MDR shows the relative error of MDR for reduction of the nonlinear problem. In accordance with our theoretical results from Section 3.3, for small $p$ it is comparable to lin. MT and only for larger $p$ it is growing moderately, since the error of mixed quadratic terms becomes dominant over the error of the linear modes that were left out. Already for small nonlinearity and $p \geq 0.1$ the error of a nonlinear reduction with linear eigenmodes becomes quite high, which is indicated by NL MT.

Altogether, this application shows that the approximation properties of the modal derivative reduction can be verified for demanding PDE problems such as a 3D nonlinear elasticity simulation.

5. MODAL DERIVATIVES FOR FULL COMPLEXITY REDUCTION

As we have already outlined at the end of Section 2.1, projection methods for nonlinear model reduction such as MDR decrease the computational effort for solving the reduced order systems. However, they still require the full computational complexity and effort for the assembly of the nonlinear force vector $f(d)$ and tangent stiffness matrix $K(d)$ for the argument $d(q)$ whenever the reduced order system needs to be evaluated. Therefore MOR is often combined with a complexity or dimension reduction method that allows the evaluation of $f$ or $\tilde{f} = Q^T f$ based on the reduced coordinates $q$ only.

Probably the most popular choice for such a method is the Discrete Empirical Interpolation Method (DEIM) [37, 38]. This technique is particularly powerful when used in combination with the POD method (POD-DEIM) [38]. Both POD and DEIM are general approaches where the large-scale dynamical system is evaluated to generate snapshot data for setting up the basis for the projection and for constructing the interpolation. Alternative methods that could also be used to derive a nonlinear reduced-order model are so-called moment matching [39] or Krylov subspace methods [40] for nonlinear systems.

If the specific structure of the dynamical system is available, however, we advocate to take it into account, and one means is to employ modal derivatives as first step for the projection. For the full reduction, we propose next an approach which does not require any additional function or derivative evaluations.

5.1. Complexity reduction methods

The main idea behind the complexity reduction method is to substitute the reduced ansatz for the displacements from MDR [48] into the quadratic Taylor series expansion of $f$, which we have already used in the proof of Theorem 1:

$$f(d(q)) = Kd(q) + \frac{1}{2}(H : d(q))d(q) = KQq + \frac{1}{2}(H : (Qq))(Qq)$$

$$= \sum_{i \in R} K\phi_i q_i + \sum_{ij \in R^*} K\psi_{ij} q_{ij}$$

(51)
\[ \begin{align*}
+ \frac{1}{2} \sum_{i \in \mathcal{R}} \sum_{k \in \mathcal{R}} (\mathbb{H} : \phi_i) \phi_k q_i q_k + \frac{1}{2} \sum_{i \in \mathcal{R}} \sum_{kl \in \mathcal{R}^*} (\mathbb{H} : \phi_i) \psi_{kl} q_i q_{kl} \\
+ \frac{1}{2} \sum_{ij \in \mathcal{R}^*} \sum_{k \in \mathcal{R}} (\mathbb{H} : \psi_{ij}) \phi_k q_{ij} q_k + \frac{1}{2} \sum_{ij \in \mathcal{R}^*} \sum_{kl \in \mathcal{R}^*} (\mathbb{H} : \psi_{ij}) \psi_{kl} q_{ij} q_{kl}.
\end{align*} \]

Exploiting that \((\mathbb{H} : \phi_i) \phi_k = -K \phi_{ik}\) (Def. 1) and \((\mathbb{H} : \psi_{kl}) \phi_i = (\mathbb{H} : \psi_{kl}) \phi_i\) (Lemma 1), and using the notation \((\mathbb{H} : \phi_i) = \nabla \phi_i K\), it follows:

\[ f(d(q)) = \sum_{i \in \mathcal{R}} K \phi_i q_i + \sum_{ij \in \mathcal{R}^*} K \psi_{ij} q_{ij} - \frac{1}{2} \sum_{i \in \mathcal{R}} \sum_{k \in \mathcal{R}} K \psi_{ik} q_i q_k \\
+ \sum_{i \in \mathcal{R}} \sum_{kl \in \mathcal{R}^*} \nabla \phi_i K \psi_{kl} q_i q_{kl} + \frac{1}{2} \sum_{ij \in \mathcal{R}^*} \sum_{kl \in \mathcal{R}^*} (\mathbb{H} : \psi_{ij}) \psi_{kl} q_{ij} q_{kl} \tag{52} \]

All terms in this representation can be computed explicitly using the quantities already provided for setting up the modal derivatives, except the last sum, which depends on the unknown \((\mathbb{H} : \psi_{ij})\). However, this last term is virtually a fourth-order expression as \(q_{ij} q_{kl} \sim p_i p_j p_k p_l\) and thus we omit it in the definition of the reduced-dimensional approximation \(f_R\) of \(f\):

\[ f(d(q)) \approx f_R(q) := \sum_{i \in \mathcal{R}} K \phi_i q_i + \sum_{ij \in \mathcal{R}^*} K \psi_{ij} q_{ij} - \frac{1}{2} \sum_{i \in \mathcal{R}} \sum_{k \in \mathcal{R}} K \psi_{ik} q_i q_k \\
+ \sum_{i \in \mathcal{R}} \sum_{kl \in \mathcal{R}^*} \nabla \phi_i K \psi_{kl} q_i q_{kl} \tag{53} \]

\[ = K\left( \Phi_R q_R + \Psi_{R^*} q_{R^*} - \frac{1}{2} (\Psi_{R^*} : q_{R^*}) q_{R^*} \right) + \left( \sum_{i \in \mathcal{R}} \nabla \phi_i K q_i \right) \Psi_{R^*} \cdot q_{R^*}. \tag{55} \]

Since we do not compute and store the tensor of modal derivatives \(\Psi_{R^*}\) and only use the matrix form \(\Psi_{R^*}\), we introduce the vector \(q_{R^*}^i \in \mathbb{R}^n\) defined by

\[ q_{ij}^* = \begin{cases} q_i^2 & \text{if } i = j, \\
2q_i q_j & \text{else}, \quad i, j \in \mathcal{R} : j \geq i. \end{cases} \tag{54} \]

For a more compact notation we also write the set of directional derivatives \((\nabla \phi_i K)_{i \in \mathcal{R}}\) as \(n \times n \times r\)-tensor \(\nabla \phi K\) and reformulate (55) as

\[ f_R(q) = K\left( \Phi_R q_R + \Psi_{R^*} q_{R^*} - \frac{1}{2} \Psi_{R^*} : q_{R^*}^i \right) + (\nabla \phi K : q_{R^*}) \Psi_{R^*} \cdot q_{R^*}. \tag{55} \]

5.1.1. Simple modal derivative complexity reduction (MDCs)

To be able to compute the unknown coefficient vector \(q\) of the quadratic approximation \(f_R\) of \(f\) given by (55) we proceed just like in the case of modal reduction, Section 4.1, and project \(f_R\) onto the reduced subspace using \(Q\):

\[ \tilde{f}_R(q) := Q^T f_R(q) = Q^T K \Phi_R q_R + Q^T K \Psi_{R^*} q_{R^*} - \frac{1}{2} Q^T K \Psi_{R^*} q_{R^*}^i + Q^T (\nabla \phi K : q_{R^*}) \Psi_{R^*} : q_{R^*}^i \tag{56} \]

\[ = Q^T K \Psi_{R^*} q_{R^*} - \frac{1}{2} Q^T K \Psi_{R^*} q_{R^*}^i + Q^T (\nabla \phi K : q_{R^*}) \Psi_{R^*} : q_{R^*}^i. \]

Then the nonlinear system

\[ \tilde{f}_R(q) = \tilde{b} \tag{57} \]
of size \( r' \) has been solved for the unknowns \( q \). This can be done for instance using Newton’s method, making use of the derivative matrix of \( f_R(q) \):

\[
\frac{d\tilde{f}_R}{dq}(q) = Q^T K Q - \left( Q^T K \phi_R - dq_R', O \right)
\]

\[
\quad + \left( Q^T \nabla_f K \phi_R - dq_R', O \right) = \left( \Phi_R^T K \phi_R - \Phi_R^T K \phi_R^*, \Phi_R^T \phi_R^* - \Phi_R^T \phi_R^* \right) - \left( \Phi_R^T K \phi_R - dq_R', O_R \right)
\]

\[
\quad + \left( \Psi_R^T \phi_R^* - \Psi_R^T \phi_R^*, O_R \right) = \left( \Phi_R^T \phi_R^* - \Phi_R^T \phi_R^* \right)
\]

\[
\quad + \left( \Phi_R^T K \phi_R - dq_R', O_R \right) = \left( \Phi_R^T \phi_R^* - \Phi_R^T \phi_R^* \right) + \left( \Phi_R^T K \phi_R - dq_R', O_R \right)
\]

\[ \tag{58} \]

It is important to note that the matrices appearing in \((58)\) and \((58)\), \(Q^T K Q\) and \((Q^T \nabla_f K \phi_R^*)\), are of reduced size \( r' \times r' \) and can be fully pre-computed and stored in the offline phase.

Furthermore, they are only based on the eigenmodes \( \phi_R \), the directional derivatives of the stiffness matrix \( \nabla \phi, K \), and the modal derivatives \( \psi_R^*\), which are all already necessary for the computation of MDs for pure reduction using MDR anyway. Thus the (simple) Modal Derivative Complexity Reduction does not add any additional computational cost in terms of nonlinear force vector and derivative evaluations.

As already mentioned, the MDs may not necessarily be linearly independent, but other than explained in Remark 3 for modal reduction, we cannot exclude some MDs here for complexity reduction, since they are needed to reconstruct the nonlinear force vector. Therefore \((58)\) can be singular and the small linear systems in an iterative solution process of \((58)\) have to be solved using the Moore-Penrose pseudoinverse, which can be obtained from a singular-value decomposition (SVD).

### 5.1.2. Extended modal derivative complexity reduction (MDCe)

In the steps from \((52)\) to \((55)\) we have not considered quadratic terms in \( q_R \), for the representation of \( f_R \), since products \( (H: \psi_{ij}) \psi_{kl} \) cannot be computed. However, not removing these terms, the projection of \( f_R \) onto \( Q \) would then result in products of the form \( \phi_R^T (H: \psi_{ij}) \psi_{kl} \) and \( \psi_R^T (H: \psi_{ij}) \psi_{kl} \). Making use of the symmetries of \( H \), the former of these two families of products can be transformed to \( \psi_R^T (H: \phi_m) \psi_{kl} \), which is an expression that we have used before and can actually be computed. Thus we can further enhance \( \tilde{f}_R(q) \) by means of

\[
\tilde{f}_R(q) := \tilde{f}_R(q) + \left( \frac{1}{2} \sum_{ij \in R} \sum_{kl \in R} \Phi_R^T (H: \psi_{ij}) \psi_{kl} q_{ij} q_{kl} \right)
\]

\[
\quad = \tilde{f}_R(q) + \left( \frac{1}{2} \sum_{ij \in R} \sum_{kl \in R} \psi_R^T (H: \phi_m) \psi_{kl} q_{ij} q_{kl} \right) = \tilde{f}_R(q) + \left( \frac{1}{2} \psi_R^T (\nabla_f K : q_R^*) \psi_R^* q_R^* \right) \tag{59} \]

Again this reduced nonlinear system of equations has to be solved to obtain \( q \) and subsequently \( d(q) \), which defines an extended Modal Derivative Complexity Reduction.

### 5.1.3. Direct modal derivative complexity reduction (MDCd)

MDCs and MDCe both require the solution of a small nonlinear system of equations and due to the possible lack of linear independence of modal derivatives SVD has to be employed therefore.
Thus with directly setting
\[ d_D(p_R) = \sum_{i \in R} \phi_i p_i + \sum_{i \in R} \sum_{k \in R} \psi_{ik} \frac{P_k}{2}, \]  
we can immediately compute a second-order approximation of the solution of \( f(d) = M \Phi_R A_R p_R \) from the right-hand side \( b \) and obtain a simple direct Modal Derivative Complexity Reduction method.

### 5.2. Nonlinear state-space model

Besides the stationary case, the so far presented nonlinear reduction methods based on modal derivatives are well-suited for dynamical problems. The generation of a nonlinear state-space model where all expressions are provided in a pre-processing phase is in our point of view a particularly promising feature.

Employing either of the MDC method as presented above, the projected nonlinear force term \( \tilde{f}_R(q) \) of a reduced nonlinear ODE system such as (61) can be replaced by the quadratic approximations \( \tilde{f}_R(q) \) from (56) or \( \tilde{f}_R^e(q) \) from (59). Thus we have the variants

\[
\begin{align*}
\text{MDR} & : M \ddot{q}(t) + f(q(t)) = b(t) \\
\text{MDCs} & : \hat{M} \ddot{q}(t) + \hat{f}(Q(t)) = \hat{b}(t) \\
\text{MDCe} & : \tilde{M} \ddot{q}(t) + \tilde{f}_R(q(t)) = \tilde{b}(t) \\
\text{MDce} & : \tilde{M} \ddot{q}(t) + \tilde{f}_R^e(q(t)) = \tilde{b}(t).
\end{align*}
\]

Proceeding just as in Section 2.2 for a linear state-space model, we can transform \( 63 \) to a first order representation writing \( \dot{x}_1 = q, \dot{x}_2 = \dot{q} \) as in (5) and then write the first equation of the nonlinear state-space model as:

\[
\begin{pmatrix}
\dot{x}_1 \\
\dot{x}_2 
\end{pmatrix} = \begin{pmatrix}
\tilde{M}^{-1} \tilde{f}_R(\tilde{x}_1) \\
0
\end{pmatrix} + \begin{pmatrix}
0 \\
\tilde{M}^{-1} Q^T
\end{pmatrix} u.
\]

In this way we have defined a nonlinear state-space model with a quadratic approximation of the reduced nonlinear force vector. The evaluation is not significantly more expensive than in the linear case since \( \tilde{f}_R(\tilde{x}_1) \) (or analogously \( \tilde{f}_R^e(\tilde{x}_1) \) or \( f_D(\tilde{x}_1) \)) is simply a quadratic expression in \( \tilde{x}_1 = q \). The reduced matrices and coefficients of \( \tilde{f}_R \) from any MDC method do not depend on time or states and thus may be evaluated only once in the offline pre-computing phase, which allows a fast online evaluation. Due to the use of eigenmodes and modal derivatives for the reduced order representation of the nonlinear force term and the projection with MDC, this newly derived model may be considered as a frequency-preserving quadratic state-space model.

### 5.3. Numerical study of MDC methods

The following study investigates the approximation properties of the above presented model order reduction methods numerically and verifies the theoretical properties we have derived.
We compare the proposed Modal Derivative Complexity Reduction methods (namely MDCs, MDCe and MDCd) with Modal Derivative Reduction (MDR) and the “full” solution of the original discretized nonlinear PDE problems. In order to understand the behaviour of the MDC methods in detail, we choose a family of one-dimensional nonlinear model problems that are all derived from the linear Poisson equation. While this allows us to immediately rule out bad methods, it is clear that a good performance for this problem class does not imply the same for a different setting.

The linear Poisson equation in strong form reads

\[ u : \Omega \rightarrow \mathbb{R} : \mathcal{L}[u] = -u'' = b \quad \forall x \in \Omega. \]  \hspace{1cm} (65)

Eigenvalues and eigenmodes are computed as solutions of

\[ \phi_i : \Omega \rightarrow \mathbb{R}, \omega_i \in \mathbb{R}, i \in \mathbb{N} : -\omega_i^2 \phi_i = \phi_i'' \quad \forall x \in \Omega. \]  \hspace{1cm} (66)

In all cases we use as domain the interval \( \Omega = (0, \pi) \) and set the boundary conditions \( u(0) = 0 \) and \( u'(\pi) = 0 \).

**Problem 1** (Quadratic test problem)

\[ \mathcal{L}[u] + \mathcal{N}_2[u] = -u'' - \left( \frac{1}{2} u'^2 \right)' = b. \]  \hspace{1cm} (67)

**Problem 2** (Cubic test problem)

\[ \mathcal{L}[u] + \mathcal{N}_2[u] = -u'' - \left( \frac{1}{2} u'^2 + \frac{1}{6} u^3 \right)' = b. \]  \hspace{1cm} (68)

We also vary the right hand sides and distinguish several cases, which all depend on the value of a parameter \( \hat{p} > 0 \):

(a) \( b_a = \hat{p} \omega_1^2 \phi_1 \),

(b) \( b_b = \hat{p} \omega_1^2 \phi_1 + \frac{1}{2} \hat{p} \omega_2^2 \phi_2 \),

(c) \( b_c = \hat{p} (x/\pi)^2 \).

Furthermore, as an application towards nonlinear state-space models, we also study the corresponding nonlinear wave equations, i.e. one-dimensional transient (hyperbolic) initial value problems for \( u : \Omega \times (0, T) \rightarrow \mathbb{R} \):

**Problem 3** (Transient test problem)

\[ \ddot{u} + \mathcal{L}[u] + \mathcal{N}_1[u] = b_c \sin(\omega t), \quad \forall x \in \Omega \times (0, T), \]  \hspace{1cm} (70)

\[ u(x, 0) = 0, \quad \dot{u}(x, 0) = v_0, \quad \forall x \in \Omega, \]  \hspace{1cm} (71)

with \( b_c(x) = p (x/\pi)^2 \) and \( p = 2 \) (corresponding to Problems 1 and 2). The initial velocity \( v_0 \) is chosen such that the linear problem has a periodic analytical solution.

For the discretization of these test problems we have again used isogeometric finite elements with B-Splines of order 3 with 16 elements (\( n = 19 \) DOFs) in the results shown below. However, there is no dependency of the results on the accuracy of the discretization, i.e. the accuracy of the “full” solution \( u^h \). For the numerical time integration of Problem 3 we use an implicit Euler method with a fixed step size of \( h = T_0/64 \), where \( T_0 = 2\pi/\omega \) is the length of one period of excitation. In our examples we choose \( \omega = 0.1 \) (for reference: \( \omega_1 = 0.5 \)).

In the following, we investigate the dependence of the error of the solution based on the full system \( u^h := u^h(\mathbf{d}) \) versus the solutions based on reduction methods \( u^h_\mathcal{N} := u^h(\mathbf{d}_\mathcal{N}) \) on a scalar parameter, the load factor \( p = ||\mathbf{p}||_\infty \), where \( \mathbf{p} = \mathbf{A}_\mathcal{N}^{-1}\Phi_\mathcal{N}^T \mathbf{b} \). To confirm the theoretical error estimates, we compare the curves of the numerical errors with curves of order \( \mathcal{O}(p^1), \mathcal{O}(p^2) \), etc.
We start with the quadratic problem. Here we have $p = p_1 = \hat{p}$ and the linear solution can be exactly represented using the first eigenmode. Fig. 3 shows the absolute errors $\|u^h - u^h_{\text{R}}\|_{L^2(\Omega)}$ for Problem 1(a) using (a) only the first and (b) the first 4 eigenmodes plus corresponding modal derivatives over a range of the right hand side parameter $p \in [0.01, 10]$. According to our error estimates, the reduced solutions should be able to represent the full solutions at least with an error of order 3 in $p$. This is reflected very well in Fig. 3(a) where we have only used the first eigenmode and its derivative. Since Problem 1(a) is quadratic, the MDC methods should be able to reconstruct the exact evaluation of the nonlinear force vector $f$ with an error of order 4 in $p$, which is reflected by the curves of MDCd, MDCs and MDCe. When we use a larger basis, as in Fig. 4(b) with $r = 4$, MDR becomes very accurate. The error is still of order 3, but with a small constant. The curve for MDCd is still the same as in Fig. 3(a) as $q_1 = p$ and $q_2 = q_3 = q_4 = 0$. MDC methods now even exhibit an error of order 4 in $p$, since also the terms $p_ip_k \sim p_ip_k p_l p_m$, which are virtually third order in $p$, get cancelled out. For $p > 0.1$ MDCe becomes much more accurate than MDCs. Note that for $p > 1$ convergence of the iterative methods is not assured. Here MDR, MDCs and MDCe still converged for $p < 10$ and showed the expected behaviour.

Here it holds also $p = p_1 = \hat{p}$ and the linear solution can be exactly represented using the first plus second eigenmode. Fig. 4 shows the absolute errors $\|u^h - u^h_{\text{R}}\|_{L^2(\Omega)}$ using (a) the first 2 and (b) the first 4 eigenmodes plus corresponding modal derivatives over a range of the right hand side parameter $p \in [0.01, 1]$. Using $r = 2$, which means that the eigenmodes used for construction of the right hand side are also present in the reduced solutions together with corresponding MDs, we find the expected error rates of order 3 in $p$ for all methods in Fig. 4(a). Results for $r = 4$ shown in Fig. 4(b) are also similar to previous findings. MDR becomes again much more accurate, while MDCs and MDCe are slightly more accurate and have increased error rates of order 4 in $p$. Also the difference between MDCs and MDCe becomes greater for $p > 0.1$. Here there was no convergence for all iterative methods for $p > 1$.

With Problem 1(c) we now have a right hand side and linear solution that cannot be exactly represented using a reduced number of eigenmodes. In this case it is $p = p_1 = \hat{p}/2.953$ and we also want to note that $p_5 = 3.90 \cdot 10^{-4}$ and $p_9 = 3.2 \cdot 10^{-5}$. Fig. 5 shows the absolute errors $\|u^h - u^h_{\text{R}}\|_{L^2(\Omega)}$ using (a) the first 4 and (b) the first 8 eigenmodes plus corresponding modal derivatives. First of all we notice that $p_5$ for $r = 4$ and $p_9$ for $r = 8$ correspond almost exactly to
ON THE USE OF MODAL DERIVATIVES FOR NONLINEAR MODEL ORDER REDUCTION

Figure 4. $L^2$-errors $\|u^h - u^h_R\|_{L^2(\Omega)}$ for Problem 1(b) using (a) 2 and (b) 4 EMs plus corresponding MDs.

Figure 5. $L^2$-errors $\|u^h - u^h_R\|_{L^2(\Omega)}$ for Problem 1(c) using (a) 4 and (b) 8 EMs plus corresponding MDs.

the errors of linear MT, which grow now obviously by a rate of order 1 in $p$. For small $p$ all other methods exhibit the same rate since the modal truncation error is the dominant error contribution. For increasing $p$ the rates of MDCd and MDCs change to 3 and the error grows rapidly. For MDCe the transition is much slower and, especially for $r = 8$, a good accuracy is achieved even when $p > 1$.

**Problem 2(a).** Now we move to the problem with cubic nonlinearity. Again it is $p = p_1 = \hat{p}$ and the linear solution can be exactly represented using the first eigenmode. Fig. 6 shows the absolute errors $\|u^h - u^h_R\|_{L^2(\Omega)}$ using (a) only the first and (b) the first 4 eigenmodes plus corresponding modal derivatives. In contrast to Problem 1(a) and Fig. 3 now MDR is more accurate than MDCs.
Figure 6. $L^2$-errors $\|u^h - u^h_\mathbb{R}\|_{L^2(\Omega)}$ for Problem 2(a) using (a) 1 and (b) 4 EMs plus corresponding MDs

Figure 7. $L^2$-errors $\|u^h - u^h_\mathbb{R}\|_{L^2(\Omega)}$ for Problem 2(c) using (a) 4 and (b) 8 EMs plus corresponding MDs

even for $r = 1$ and MDCe slightly more accurate than MDCs for the whole range of $p$. Furthermore, for $r = 4$ (almost) no increase in accuracy of MDCs and MDCe is recognizable and the rates remain of order 3. The reason therefore is clear, as the dominant terms of the error are now always of order 3 in $p$, since all presented reduction methods only consider quadratic terms.

Problem 2(c). Since we do not expect any additional insight from Problem 2(b) we continue directly with Problem 2(c). Fig. 7 shows the absolute errors $\|u^h - u^h_\mathbb{R}\|_{L^2(\Omega)}$ using (a) the first 4 and (b) the first 8 eigenmodes plus corresponding modal derivatives. For linear MT and MDR the dominant error is the modal truncation error and thus for both the contribution to the error is of order 1 in $p$. It is of course smaller for $r = 8$ than for $r = 4$ and again it holds that the value of the error is around $p^5$. 

Copyright © 2016 John Wiley & Sons, Ltd.

Prepared using nmeauth.cls


DOI: 10.1002/nme
ON THE USE OF MODAL DERIVATIVES FOR NONLINEAR MODEL ORDER REDUCTION

Figure 8. Displacement $u(\pi, t)$ and velocity $\dot{u}(\pi, t)$ for Problem 3.1(c) with $p = 1.0$ over two periods of integration, i.e. $t \in [0, 2T_0]$. Reduction methods using 4 EMs plus corresponding MDs ($r = 4$)

![Graph](image1)

Figure 9. $L^2$-errors $\|u^h - u^R\|_{L^2(\Omega)}$ for Problem 3.1(c) transient-hyperbolic using (a) 4 and (b) 8 EMs plus corresponding MDs at time $t = 0.25T_0$

and $p_0$ respectively. For small $p$ this holds also for MDCd and MDCe, but then the rate changes to order 3 and thus increases rapidly. Just as for Problem 2(a), MDCe is just marginally more accurate than MDCd.

**Transient Problem 3.1(c).** Next we move on to the time-integration of transient Problem 3 with quadratic nonlinearity as in Problem 1 and right-hand side $b_c$. Reduction with MDR, MDCd and MDCe was performed with $r = 4$ EM and all corresponding MDs. Displacement $u(\pi, t)$ and velocity $\dot{u}(\pi, t)$ over two periods of integration for $p = 1.0$ and $r = 4$ are shown in Fig. 8. Here the curves for fully nonlinear analysis (full NL), MDR and MDCe are visually indistinguishable, while the result of MDCd is close to the linearized problem (linear). Furthermore, the $L^2$-errors of displacements at $t = 0.25T_0$ are given in Fig. 9 for both $r = 4$ and $r = 8$. They show that MDCd results in a large error with rate $p^2$, while MDR and MDCe have a much higher accuracy with linear
convergence rate for small \( p \), just as linear MT compared to a linear analysis (MT/lin), and quadratic rate for larger \( p \) and \( r = 8 \), when the truncation error is less dominant. Quite surprisingly, for \( r = 4 \) over a wide range of \( p \) MDCe is here even more accurate than MDR. The curve MT/NL shows the error of a linear analysis with modal truncation compared to a fully nonlinear analysis and indicates that the nonlinear analysis with either MDR or MDCe reduction is still much more accurate than performing just a linear analysis (with reduction). The curve for MDCd basically follows MT/NL.

**Transient Problem 3.2(c).** For the transient problem we also investigate the cubic nonlinearity. Again displacement \( u(\pi, t) \) and velocity \( \dot{u}(\pi, t) \) over two periods of integration with \( p = 1.0 \) and \( r = 4 \) are shown in Fig. 10 In contrast to the previous example with quadratic nonlinearity, now also an error of MDCe compared to full analysis and MDR becomes visible around \( t/T_0 = 0.75 \) and \( t/T_0 = 1.75 \). However, it is still much better than a full linear analysis or MDCd. The \( L^2 \)-errors

---

**Figure 10.** Displacement \( u(\pi, t) \) and velocity \( \dot{u}(\pi, t) \) for Problem 3.2(c) with \( p = 1.0 \) over two periods of integration, i.e. \( t \in [0, 2T_0] \). Reduction methods using 4 EMs plus corresponding MDs (\( r = 4 \)).

---

**Figure 11.** \( L^2 \)-errors \( \|u_h^n - u_h^R\|_{L^2(\Omega)} \) for Problem 3.2(c) transient-hyperbolic using (a) 4 and (b) 8 EMs plus corresponding MDs at time \( t = 0.25T_0 \).
of displacements at \( t = 0.25T_0 \) are then also shown in Fig. [1] and reveal again that MDCe method doesn’t provide such a good approximation for the cubic nonlinearity since the convergence rate becomes cubic for larger \( p \). However, comparing MDCe to MT/NL shows that even for high \( p \) the error is still at least one order of magnitude lower.

**Summary of numerical results.** Based on the test problems presented above, we may conclude that the theoretical convergence properties were confirmed by the numerical convergence rates. While the simplest of the newly proposed complexity reduction methods, MDCd, has in general quite large errors and does not provide a sufficiently accurate approximation, the performance of MDCs and especially MDCe is very convincing for quadratic problems. The scheme MDCe, which we propose as the method of choice for a nonlinear complexity reduction with modal derivatives, comes close to MDR even when increasing the factor \( p \), which means a stronger influence of the nonlinear term. Though for cubic nonlinearities the MDCe scheme cannot keep up with MDR and shows weaker performance, it still provides a more accurate approximation of the solution than a purely linear analysis.

Summarizing, we conclude that a full complexity reduction using the newly proposed MDCe method provides an efficient and accurate approximation of the full solution of nonlinear static and dynamic problems, especially with quadratic nonlinearity. Furthermore, the transient results open up a promising perspective for the usage of the MDCe method for frequency-preserving, reduced nonlinear state-space models.

6. SUMMARY AND CONCLUSION

In this paper we have discussed a generalization of the state-of-the-art modal reduction technique for linear equations. The so-called modal derivatives add gradient information to the projection subspace and thus improve the approximation behaviour of the reduced system of nonlinear equations. In [23] we have used these techniques successfully for reducing a system stemming from nonlinear frequency analysis of structures undergoing large deformation.

Here we have extended the theoretical foundation of modal derivative reduction methods and investigated the general approximation properties of eigenmodes and modal derivatives for nonlinear problems. We have shown that a subspace with eigenmodes and modal derivatives is indeed well suitable for the reduction of nonlinear PDE problems using projection/reduced basis methods. Furthermore, we have generalized the idea of modal derivatives from eigenmodes to Craig-Bampton modes and methods for substructuring. The theoretical error and convergence estimates for MDR have been verified by applying it to a 3-dimensional nonlinear elasticity problem.

A further result is the extension of modal derivatives to a family of novel nonlinear complexity reduction methods. These methods result in a quadratic approximation of the nonlinear force vector and do not require any additional computational effort other than eigenmodes and modal derivative computation. They can be used for the solution of static equations and for deriving a reduced nonlinear state-space model that captures the essential dynamics. Within the method family, the scheme MDCe works extremely well for quadratic nonlinearities with general right hand sides. Our numerical results also indicate that for stronger nonlinear terms, the approximation property of the method depends on the context.

This calls for further investigations, in particular by testing the modal derivative complexity reduction methods at a broad set of nonlinear problems in computational engineering. In order to assess the full potential of the corresponding quadratic state-space model, it needs to be examined within a system simulation where the interaction with other parts or different physics is taken into account.

ACKNOWLEDGEMENTS

O. Weeger acknowledges support from the SUTD Digital Manufacturing and Design (DManD) Centre, supported by the Singapore National Research Foundation, and from the Department of Mathematics at TU Kaiserslautern.
References


10. Siemens PLM Software. NX Nastran. software.


