Model Order Reduction based on Proper Orthogonal Decomposition

Elías Cueto[*], Francisco Chinesta[†] and Antonio Huerta[‡]
[*] I3A, Universidad de Zaragoza
Maria de Luna s/n, 50018 Zaragoza, Spain
[†] ESI Group Chair on Advanced Computational Manufacturing Processes
GEM, UMR CNRS - Centrale Nantes
1 rue de la Noe, BP 92101, F-44321 Nantes cedex 3, France
[‡] Laboratori de Calcul Numeric
Universidad Politecnica de Cataluña
Jordi Girona 1, Campus Nord, C2, E-08034 Barcelona, Spain

Abstract In this chapter we review the basics of classical model order reduction techniques based on Proper Orthogonal Decomposition (POD), Principal Component Analysis (PCA) or Karhunen-Loève decompositions.

1 Introduction

Simulation has become essential in nowadays industrial processes in general, and in forming processes in particular. The appearance of the finite element method in the 50's, and the continuing growth of computer performance have allowed the industry to simulate processes and products. But today the problem remains precisely in the complexity of state-of-the-art computer simulations. These are surprisingly accurate and complex, being able to simulate complex multi-physic, multi-scale phenomena with a high degree of fidelity. Industry faces today a different challenge: to have these simulation results in a short time so as to be able to take decisions without delays. Typically, high-tech industries such as aeronautic manufacturers accept simulations able to run in one night. This makes it possible to take decisions in the morning of the following day. However, many state-of-the-art simulations take today days or even weeks.

Model order reduction (MOR) or, simply, model reduction techniques exist many decades ago. These techniques allow for a simplified representation of the evolution of physical systems, employing a very reduced number of degrees of freedom. Based on the above considerations, MOR techniques have gained a tremendous popularity in the last years due precisely to their

F. Chinesta, P. Ladevèze (Eds.), *Separated Representations and PGD-Based Model Reduction Fundamentals and Applications*, CISM International Centre for Mechanical Sciences
DOI 10.1007/ 978-3-7091-1794-1_1

ability to greatly reduce computational cost at a price of only a minimal loss of accuracy, that can be nowadays fully quantified.

MOR techniques seem, therefore, to be an appealing choice for their generalization in industrial contexts, and particularly in the simulation of forming processes. Many forming processes involve complex multi-physic phenomena in a coupled manner, with inherent strong non-linearities, uncertainty in parameter determination, etc. Thus, simulation in this context involves very complex techniques that, despite their accuracy and sound theoretical basis, are strongly CPU-consuming.

In this first chapter of the book we deal with the basis of the most popular technique for model reduction: Proper Generalized Decomposition (or Karhunen-Lo?ve transform, or Principal Component Analysis), a technique discovered and re-discovered through time in virtually all scientific speciality. We deal with the basic ingredients as well as its main associated difficulties, those responsible of the birth of a new generation of techniques such as Reduced Basis or Proper Generalized Decomposition, for instance.

2 Model reduction: extracting relevant information

The basic ingredient of POD is the need to complete some full-order problems in order to extract from them "relevant" information to be used in "similar" problems. Both "relevant" and "similar" are here quoted, since their precise meaning will be clear afterwards. It is expected that we can evaluate some general patterns in the solution of the full problems so as to employ them as basis functions for subsequent modifications of the problem, and that this relevant information is low dimensional.

Assume now that, say, a finite element approximation is constructed on top of the problem at hand. Consider a mesh having M nodes. If the problem is transient, it will be necessary to compute at each time step M values (assuming that the solution is a scalar field). For non-linear problems, this implies the solution of at least one linear algebraic system of size M at each time step. It is easy to notice the complexity of the problem for large models.

Is it possible to determine of the solution lives in a space of much lower dimensions that that spanned by the finite element mesh? In principle yes. And POD provides an efficient means to do it.

2.1 An introduction to classical Proper Orthogonal Decomposition

Assume now that the solution field $u(\mathbf{x}, t)$ is known, for the already computed full-order solutions, at the nodes $\mathbf{x}_i$ of a spatial mesh for discrete time instants $t_m = m \cdot \Delta t$, with $i \in [1, \cdots, M]$ and $m \in [0, \cdots P]$. We employ the notation $u(\mathbf{x}_i, t_m) \equiv u^m(\mathbf{x}_i) \equiv u_i^m$ and define $\mathbf{u}^m$ as the vector of nodal values u_i^m at time t_m. POD provides an efficient means to obtain the most typical or characteristic structure $\phi(\mathbf{x})$ among these $u^m(\mathbf{x})$, $\forall m$ (Ryckelynck et al., 2006). To this end, we maximize the functional

$$\alpha = \frac{\sum_{m=1}^{P} \left[\sum_{i=1}^{M} \phi(\mathbf{x}_i) u^m(\mathbf{x}_i) \right]^2}{\sum_{i=1}^{M} (\phi(\mathbf{x}_i))^2}, \tag{1}$$

which can be shown equivalent to solve the following eigenvalue problem:

$$\mathbf{c}\boldsymbol{\phi} = \alpha\boldsymbol{\phi}. \tag{2}$$

Here, the vector $\boldsymbol{\phi}$ has as the i-th component $\phi(\mathbf{x}_i)$, and $\mathbf{c}$ is known as the two-point correlation matrix

$$c_{ij} = \sum_{m=1}^{P} u^m(\mathbf{x}_i) u^m(\mathbf{x}_j) = \sum_{m=1}^{P} \mathbf{u}^m \cdot (\mathbf{u}^m)^T, \tag{3}$$

which is symmetric and positive definite. With the matrix $\mathbf{Q}$ defined as

$$\mathbf{Q} = \begin{pmatrix} u_1^1 & u_1^2 & \cdots & u_1^P \\ u_2^1 & u_2^2 & \cdots & u_2^P \\ \vdots & \vdots & \ddots & \vdots \\ u_M^1 & u_M^2 & \cdots & u_M^P \end{pmatrix}, \tag{4}$$

we have

$$\mathbf{c} = \mathbf{Q} \cdot \mathbf{Q}^T. \tag{5}$$

2.2 Reduced-order modeling based on POD

The main objective being to obtain a reduced-order model, we solve Eq. (2) and keep the N most important eigenvectors $\boldsymbol{\phi}_i$, i.e., those associated with the eigenvalues belonging, say, to the interval defined by the highest eigenvalue α_1 and α_1 divided by a large enough number (e.g. 10^8). If POD works is just because N is very often (not always!) found to be much lower than M. These N eigenfunctions $\boldsymbol{\phi}_i$ are then used as basis functions

(Ritz-like) to approximate the sought function $u^m(\mathbf{x})$, $\forall m$. Let us define the matrix $\mathbf{B} = [\boldsymbol{\phi}_1 \cdots \boldsymbol{\phi}_N]$, i.e.

$$\mathbf{B} = \begin{pmatrix} \phi_1(\mathbf{x}_1) & \phi_2(\mathbf{x}_1) & \cdots & \phi_N(\mathbf{x}_1) \\ \phi_1(\mathbf{x}_2) & \phi_2(\mathbf{x}_2) & \cdots & \phi_N(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(\mathbf{x}_M) & \phi_2(\mathbf{x}_M) & \cdots & \phi_N(\mathbf{x}_M) \end{pmatrix}. \tag{6}$$

If, for instance, an explicit time-stepping scheme is used to compute the discrete solution $\mathbf{u}^{m+1}$ at time t^{m+1} a linear algebraic system like

$$\mathbf{G}^m \ \mathbf{u}^{m+1} \ = \ \mathbf{H}^m. \tag{7}$$

must be solved at each time step.

A reduced-order model is then obtained by projecting $\mathbf{u}^{m+1}$ onto the subspace defined by the N most important eigenvectors $\boldsymbol{\phi}_i$, i.e.

$$\mathbf{u}^{m+1} \approx \sum_{i=1}^{N} \boldsymbol{\phi}_i \ \zeta_i^{m+1} \ = \ \mathbf{B} \ \boldsymbol{\zeta}^{m+1}. \tag{8}$$

Equation (7) then writes

$$\mathbf{G}^m \ \mathbf{B} \ \boldsymbol{\zeta}^{m+1} \ = \ \mathbf{H}^m, \tag{9}$$

or equivalently

$$\mathbf{B}^T \mathbf{G}^m \ \mathbf{B} \ \boldsymbol{\zeta}^{m+1} \ = \ \mathbf{B}^T \mathbf{H}^m. \tag{10}$$

Coefficients $\boldsymbol{\zeta}^{m+1}$ defining the solution of the reduced-order model are thus obtained by solving an algebraic system of size N instead of M. When $N \ll M$, as is the case in numerous applications, the solution of Eq. (10) is thus very convenient because of its very reduced size.

Remark 2.2.1. The reduced-order model Eq. (10) is built a posteriori by means of the already-computed solution to the full-order problem. This is mandatory, in principle, in standard POD techniques. To avoid this costly solution of full-order problems, some alternatives can be found in the literature (see e.g. Bialecki et al. (2005), Burkardt et al. (2006), Gunzburger et al. (2007), Maday and Ronquist (2004), Park and Cho (1996), Ryckelynck (2003), Ryckelynck (2005), Ryckelynck (2008)). The first approach consists in solving the original full-order model over a short time interval, to extract the characteristic structure that defines the reduced model, and to hope it serves well for approximating the solution at larger time intervals. The second approach consists in solving the original, full-order, model over the entire time interval, and then use the so obtained reduced approximation to solve similar problems in which, for example, slight variations in material parameters or boundary conditions are made.

2.3 A practical example

Let us now consider the following one-dimensional heat transfer problem (see Chinesta et al. (2011)), written in non-dimensional form:

$$\frac{\partial u}{\partial t} \;=\; \lambda \frac{\partial^2 u}{\partial x^2}, \tag{11}$$

with constant thermal diffusivity $\lambda = 0.01$, $t \in (0, 30]$ and $x \in (0, 1)$. The initial condition is $u(x, t = 0) = 1$ and the boundary conditions are given by $\frac{\partial u}{\partial x}\big|_{x=0,t} = q(t)$ and $\frac{\partial u}{\partial x}\big|_{x=1,t} = 0$.

Eq. (11) is solved by means of an implicit-in-time finite element method built on a mesh of $M = 100$ nodes. Linear consistency is assumed for the finite element model. Time step is set to $\Delta t = 0.1$. The resulting algebraic system of equations reads:

$$\mathbf{K}\ \mathbf{u}^{m+1} \;=\; \mathbf{M}\ \mathbf{u}^{m} \;+\; \mathbf{q}^{m+1}. \tag{12}$$

Consider first the following heat source:

$$q(t) \;=\; \begin{cases} 1 & t \leq 10 \\ 0 & t > 10 \end{cases} . \tag{13}$$

The resulting FEM temperature profiles are shown in Fig. 1 at time instants $t_m = m$, for $m = 1, 2, ..., 30$. Red curves correspond to the heating process, occurring until $t = 10$. Afterwards, blue curves represent the solution for $t > 10$.

From these 30 discrete temperature profiles, matrices $\mathbf{Q}$ and $\mathbf{c}$ can be built in order to construict the eigenvalue problem (2). The 3 largest eigenvalues are found to be $\alpha_1 = 1790$, $\alpha_2 = 1.1$, $\alpha_3 = 0.1$, being the remaining eigenvalues $\alpha_j < \alpha_1 \cdot 10^{-8}$, $4 \leq j \leq 100$. It is expected that a reduced model made after a linear combination of the mentioned 3 eigenvectors related to the first 3 largest eigenvalues should be able to approximate the solution with sufficient accuracy. Although not strictly necessary (and even if the resulting basis will not be orthogonal), to deal with the initial condition it is very convenient to include the initial condition in the approximation basis. Figure 2 shows the resulting approximation functions after normalization, i.e. $\frac{\boldsymbol{\phi}_j}{\|\boldsymbol{\phi}_j\|}$ $(j = 1, 2, 3)$ and $\frac{\mathbf{u}^0}{\|\mathbf{u}^0\|}$. Defining the matrix $\mathbf{B}$ as

$$\mathbf{B} = \left[\frac{\mathbf{u}^0}{\|\mathbf{u}^0\|} \;\; \frac{\boldsymbol{\phi}_1}{\|\boldsymbol{\phi}_1\|} \;\; \frac{\boldsymbol{\phi}_2}{\|\boldsymbol{\phi}_2\|} \;\; \frac{\boldsymbol{\phi}_3}{\|\boldsymbol{\phi}_3\|} \right], \tag{14}$$

we obtain the reduced model related to Eq. (12),

$$\mathbf{B}^T\mathbf{K}\mathbf{B}\ \boldsymbol{\zeta}^{m+1} \;=\; \mathbf{B}^T\mathbf{M}\mathbf{B}\ \boldsymbol{\zeta}^{m} \;+\; \mathbf{B}^T\mathbf{q}^{m+1}, \tag{15}$$

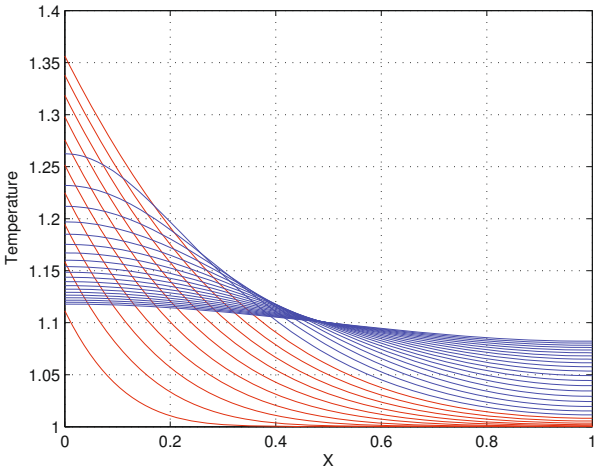


Figure 1. Temperature profiles corresponding to the problem given by Eq. (13) at time steps $t_m = m$, for $m = 1, 2, ..., 30$. Red curves correspond to the heating step, occurring until $t = 10$, while blue curves represent subsequent time step, for $t > 10$. Note the heat transfer by conduction from the warmest zones towards the cold ones.

which involves 4 degrees of freedom only. The initial condition in the reduced basis is $(\boldsymbol{\zeta}^0)^T = (1, 0, 0, 0)$.

Equation (15) and the relationship $\mathbf{u}^{m+1} = \mathbf{B}\boldsymbol{\zeta}^{m+1}$ then provide with approximate solution profiles at a very low computational cost. Results are shown in Figure 3 and are practically undistinguishable at first sight from those of the complete problem (12).

Coming back to Remark 1, assume now the usage of the reduced model (15) *as such* to solve a problem *different* from the original one. For instance, we assume now instead of (13) a substantially different heat source on the boundary:

$$q(t) = \begin{cases} \frac{t}{20} & t \leq 20 \\ \\ \frac{t-30}{5} & t > 20 \end{cases} . \tag{16}$$

Again, the solution of the reduced model is compared to that of the complete problem in Figure 4. Even though the reduced approximation basis functions are those obtained from the thermal model related to the bound-

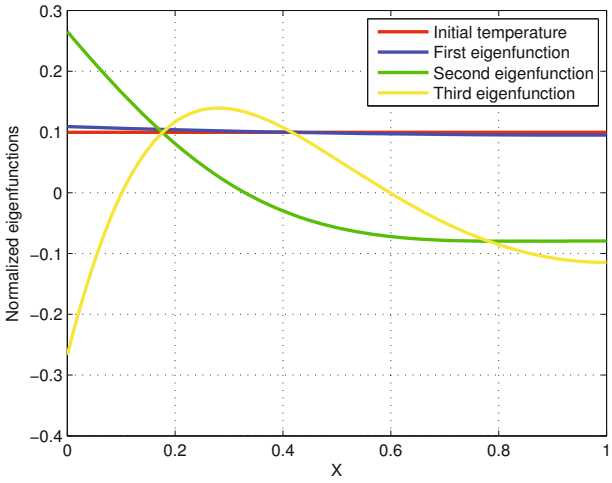


Figure 2. Reduced-order approximation basis involving the initial condition and the eigenvectors corresponding to the three largest eigenvalues.

ary condition (13), the reduced model yields a very accurate representation of the solution of this rather different problem.

2.4 Discussion

We have discussed so far the inherent advantages of model order reduction. Ideally, it would be optimal to be able to develop reduced models *a priori*, i.e. without the need for solving the complete problem. One possibility is to assess the accuracy of the reduced-order solution and, if necessary, to enrich the reduced approximation basis in order to improve accuracy (see e.g. our earlier works in the field, (Ammar et al., 2006b) and (Ryckelynck, 2005)). Proper Generalized Decomposition (PGD), which we describe in general terms in the next section, is also a very appealing alternative to these questions.

To root the origins of PGD, which will be deeply described later on, note that the previous results indicate that an accurate approximate, low-dimensional, solution to the problem can often be written as a finite sum of separated functions (here, of time and space). Indeed, when the solution evolves smoothly, as in heat transfer problems, for instance, the magnitude of the (ordered) eigenvalues α_i decreases very fast with increasing index i,

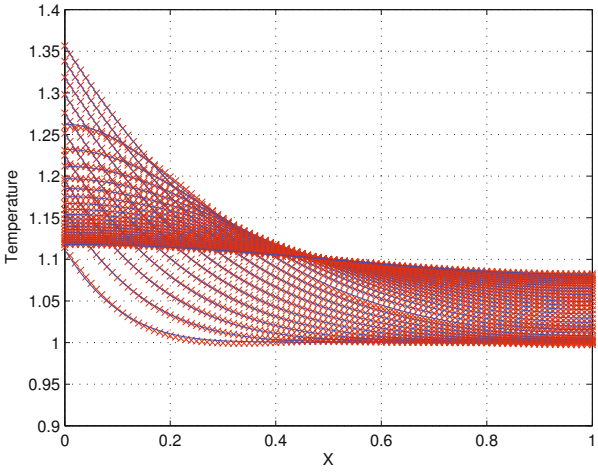


Figure 3. Global (continuous line) versus reduced-order (symbols) model solutions.

and the evolution of the field can be approximated from a very reduced number of functions in a Ritz style. Thus, if we define a threshold value ϵ (e.g. $\epsilon = 10^{-8} \cdot \alpha_1$, α_1 being the highest eigenvalue), only a small number N of modes are retained ($N \ll M$) such that $\alpha_i \geq \epsilon$, for $i \leqslant N$, and $\alpha_i < \epsilon$, for $i > N$. Equivalently:

$$u(\mathbf{x}, t) \approx \sum_{i=1}^{N} \phi_i(\mathbf{x}) \cdot T_i(t) \equiv \sum_{i=1}^{N} X_i(\mathbf{x}) \cdot T_i(t). \tag{17}$$

For the sake of clarity, the space modes $\phi_i(\mathbf{x})$ will be hereafter denoted as $X_i(\mathbf{x})$. Equation (17) rhas the form of a *separated representation*, also known as finite sum decomposition. Solutions depending on space and time can be efficiently approximated (in a great amount of cases) as a sum of a *small number* of functional products, depending on space and time, respectively. The usage of separated representations like (17) constitutes the main ingredient of PGD, the second one being the a priori nature of the method.

These functional products are determined (since we look for products of functions we face a non-linear problem) by means of iterative algorithms.The origins of this technique can be traced back to the so-called radial approximation introduced by Ladeveze (1999), Ladeveze et al. (2010), Nouy (2010)

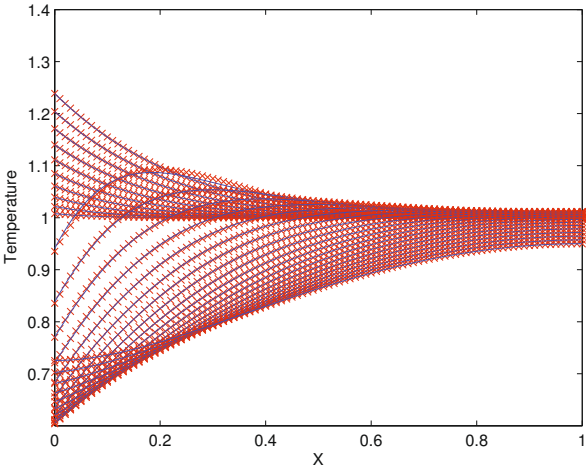


Figure 4. Global (continuous line) versus reduced-order (symbols) model solutions for the source term (16). The reduced-order approximation basis is that obtained from the solution of a different thermal problem, with the source term (13).

in the context of Computational Solid Mechanics and within the so-called LArge Time INcrement (LATIN) method.

In terms of computer cost, savings allowed by these techniques are very important. If we consider a transient problem (parabolic, hyperbolic) defined in a 3D space, the use of a standard incremental strategy with P time steps (P is often in the order of millions nowadays) requires the solution of P three-dimensional problems. On the contrary, using the a space-time separated representation (17), it is necessary to solve $N \cdot m$ three-dimensional problems to determine the spatial basis functions $X_i(\mathbf{x})$, and $N \cdot m$ more one-dimensional problems to determine the time approximation $T_i(t)$. By m we refer to the number of non-linear iterations needed for computing each term of the finite sum. Very often, it is found that $N \cdot m$ is of order 100. Computing time savings obtained by a separated representation can be of several orders of magnitude.

To this topic will be devoted the second chapter of this book. Prior to that, some aspects of POD remain to be discussed in order to fully highlight the advantages and limitations of the technique.

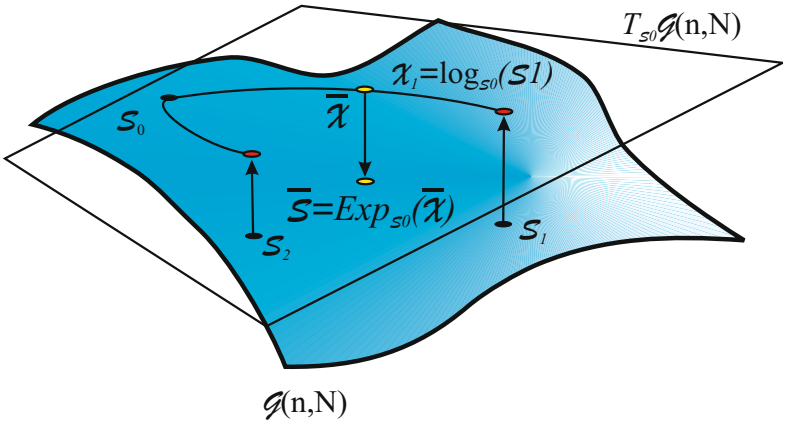


Figure 5. Schematic description of the Grassman manifold formed by the set of basis functions of all the reduced modes of a given model.

3 Interpolation of the reduced basis: a geometrical approach

One of the main limitations of POD-based reduced models is that of interpolating reduced models. Imagine that we construct a reduced approximation for a full-order problem subjected, say, to a load in a particular position. Consider now a second position of the load. Is there a relationship between these two reduced models, so that an efficient reduced-order model could be stablished for a third position of the load by interpolating between the two initial approximations?

The response is, obviously, affirmative. Amsallem and Farhat (2008) have pointed out that the set of reduced basis functions given by Eq. (2) for a given full-order model constitute the so-called Grassman manifold $\mathcal{G}(n, N)$. Therefore, in order to interpolate this set of basis functions, that obviously do not form a vector space, we must project to the tangent plane at a point of the manifold, which is a "flat" space, interpolate there, and project back to the manifold, as schematically explained in Fig. 5.

In this way, the columns of $\mathbf{B}$ constitute a basis of the subspace $\mathcal{S}_0$ of dimension n of the space $\Re^N$. At each point $\mathcal{S}$ of the manifold $\mathcal{G}(n, N)$ one can define a tangent plane of the same dimension, $\mathcal{T}_{\mathcal{S}}$, with its points defined by a matrix $\mathbf{\Gamma} \in \Re^{N\times n}$. The exponential mapping $Exp_{\mathcal{S}}$ transforms χ in an n-dimensional subspace $\mathcal{S}'$ given by a matrix $\mathbf{B}' \in \Re^{N\times n}$, such that

$$\mathbf{\Gamma} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \quad \text{(Singular value decomposition)}$$

$$\mathbf{B}' = \mathbf{B}\mathbf{V}\cos\mathbf{\Sigma} + \mathbf{U}\sin\mathbf{\Sigma}$$

Conversely, the logarithmic mapping $Log_{\mathcal{S}}$, defines a map between a point in the neighborhood of $\mathcal{S} \in \mathcal{G}(n, N)$ and the tangent plane at the origin. Thus, the image of $\mathcal{S}'$, in a neighborhood of $\mathcal{S}$, given by the logarithmic mapping, $\chi = Log_{\mathcal{S}}\mathcal{S}' \in \mathcal{T}_{\mathcal{S}}$ will be

$$(\mathbf{I} - \mathbf{B}\mathbf{B}^T)\mathbf{B}'(\mathbf{B}^T\mathbf{B}')^{-1} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \quad \text{(Singular value decomposition)}$$

$$\mathbf{\Gamma} = \mathbf{U}\tan^{-1}(\mathbf{\Sigma})\mathbf{V}^T$$

So, consider, for instance, $\mathbf{B}_0 \in \Re^{N\times n}$ and $\mathbf{B}_1 \in \Re^{N\times n}$, two matrices representing two subspaces, obtained for different parameters of the model (for instance, load positions, but the theory is completely general for other parameters of the model), s_0 and s_1. Let $\mathcal{S}_0$, $\mathcal{S}_1$ be the two subspaces originated by considering parameters s_0 and s_1. Let, in turn, $\mathcal{Y}(t)$ be the geodesic line that joins both subspaces (points in the Grassmann manifold), having $\mathcal{S}_0$ as origin. In that case, the initial derivative of the geodesic line, that belongs to the tangent plane at $\mathcal{S}_0$, will be

$$\dot{\mathcal{Y}}_0 = Log_{\mathcal{S}_0}\mathcal{S}_1$$

such that the matrix representing this initial derivative of the geodesic will be (see Amsallem and Farhat (2008) for a complete proof of this)

$$(\mathbf{I} - \mathbf{B}_0\mathbf{B}_0^T)\mathbf{B}_1(\mathbf{B}_0^T\mathbf{B}_1)^{-1} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$$

$$\mathbf{\Gamma} = \mathbf{U}\tan^{-1}(\mathbf{\Sigma})\mathbf{V}^T$$

Let $\tilde{\mathcal{S}}$ denote the point of the Grassmann manifold representing the reduced-order basis for the new value of the parameter. $\tilde{s}$. The initial derivative of the new geodesic line, joining $\mathcal{S}_0$ and the sought interpolated subspace $\tilde{\mathcal{S}}$, will be

$$\dot{\tilde{\mathcal{Y}}}_0 = \tilde{r}\dot{\mathcal{Y}}_0$$

with

$$\tilde{r} = \frac{\tilde{s} - s_0}{s_1 - s_0}$$

The computation of the singular value decomposition is not a very time-consuming task. For instance, on a PC which has 2GHz CPU, the computation of SVD(B) using Matlab, with $\mathbf{B}$ a random matrix of 5×5000 elements takes on average much less than 1 ms, still compatible with real-time constraints.

3.1 POD with interpolation (PODI)

A less rigorous, but much simpler, method to interpolate among previously computed reduced models was established in Ly and Tran (2005). Although in the standard PODI technique the POD procedure is applied to the complete set of snapshots (for different load positions, say) of the system to obtain an orthonormal basis $\mathbf{B} = [\phi_1 \cdots \phi_n]$, in this thesis we proceed by just applying the POD to each complete model (i.e., to each load position). Thus, we obtain an orthonormal basis for each system's parameter value. Basis are then interpolated for intermediate positions of the load. Although it is clear that the interpolation of orthogonal sets of functions does not yield, in general, to new orthogonal basis functions, the technique works well if the "distance" of reduced models in the Grassman manifold is not too large.

Results obtained with this technique showed to be much more efficient and accurate than those obtained with the more rigorous technique base upon interpolation on the Grassman manifold. This surprising result is still not well understood by the authors and is currently one of our research topics.

4 POD for non-linear models

One key aspect in the reduction of non-linear models is that incremental algorithms need for the evaluation of updated tangent stiffness matrices, and this is done in the full model. Therefore, one fundamental aspect of model reduction, the use of matrices with very limited size, is lost. This limitation is currently one of the most active research topics within the field of model order reduction. To fully comprehend this aspect, consider the simulation of the palpation of human cornea, considered in Niroomandi et al. (2008). In it, a non-linear hyperelastic model for the corneal tissue was considered.

The mesh consisted of 8514 nodes and 7182 hexahedral elements. A view of the geometry is shown in Fig. 6.

Usually, the corneal tissue is modeled as a fibrous-reinforcement hyperelastic medium. These fibre orientations are shown in Fig. 7.

Obviously, the first possibility is to test, for a single position of the load, the accuracy of the reduced model versus the complete one. As introduced before, once the complete model is solved, the most important eigenmodes are extracted from the computed displacement field, together with the initial tangent stiffness matrix (although other approaches can be envisaged, such as an intermediary tangent matrix, for instance). For the time being, let us consider that no updating of the tangent matrix is possible,so as to

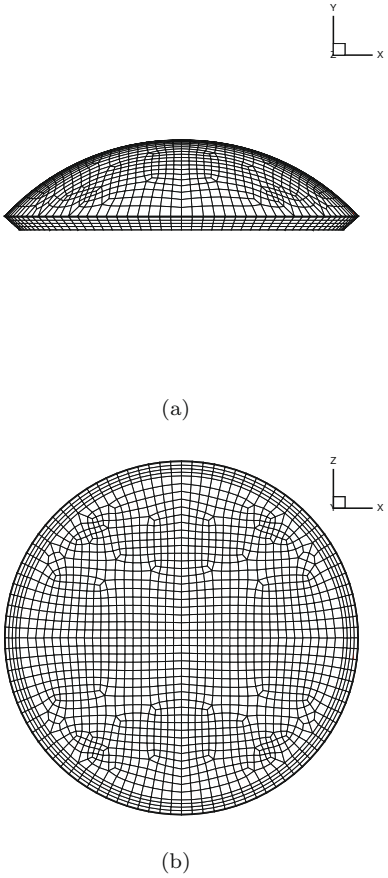
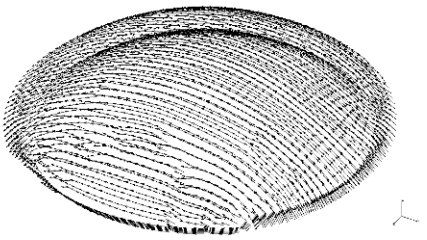


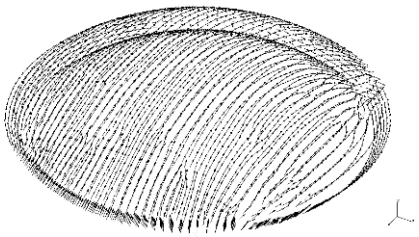
Figure 6. Geometry of the finite element model for the human cornea.

keep the computational cost as low as possible. The number of eigenmodes employed in this case was only six, leading to acceptable results. Modes are shown in Fig. 8. Their respective eigenvalues are, from the biggest to the smallest one, $9.02 \cdot 10^4$, 690, 27, 2.63, 0.221 and 0.0028. The relative influence of the eignemodes on the solution decays very fast, as can be noticed from the eigenvalues. Compare also the 8514 nodes with three degrees of freedom each, thus making 25542 degrees of freedom, of the full model, with the six (three-dimensional) degrees of freedom of the reduced model. The computational saving is enormous.

To fully compare the results, different positions of the load were con-



(a)



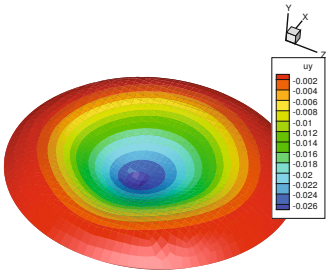
(b)

Figure 7. Fiber distribution in the finite element model of the cornea. Two families of collagen fibers, roughly perpendicular to each other, are considered.

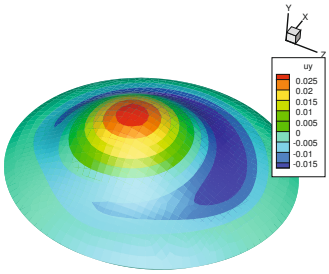
sidered. It was found that different positions do not very much influence the level of accuracy, roughly the same for any loading position. For an arbitrary location of the load, the obtained vertical displacement is shown in Fig. 9.

In Fig. 10 load-displacement curve has been depicted. In Fig. 11, however, the load was applied at a point located slightly towards the outer boundary of the model. In this case, as can be seen from Fig. 11, the displacement obtained at the point of application of the load is nearly exact, although the shape of the deformed cornea is somewhat different. This is not the case for Fig. 9, where errors of about 20% are noticed. The L_2 error norm ranged from very low values (0.08) in the early steps of the simulation, to higher values (around 0.34) for the last step. In our experience, this is a typical upper bound of the obtained error, even if very large deformations are imposed to the simulated organ, as is the case.

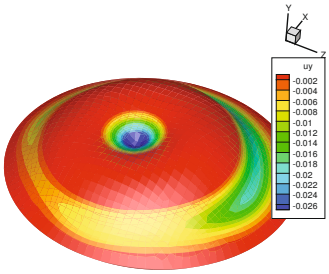
Notice that, due to the shape of shallow dome of the cornea, severe buckling phenomena are expected. This highly non-linear behavior can not be captured by a reduced model in which no tangent stiffness matrix



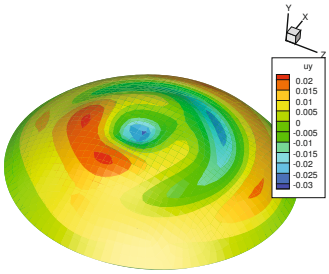
(a) $\alpha = 9.02 \cdot 10^4$



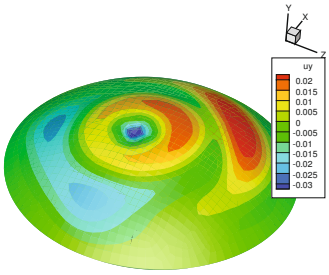
(b) $\alpha = 690$



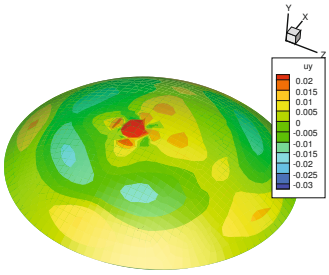
(c) $\alpha = 27$



(d) $\alpha = 2.63$



(e) $\alpha = 0.221$



(f) $\alpha = 0.0028$

Figure 8. The eigenmodes of the problem employed as global basis for the reduced model simulation.

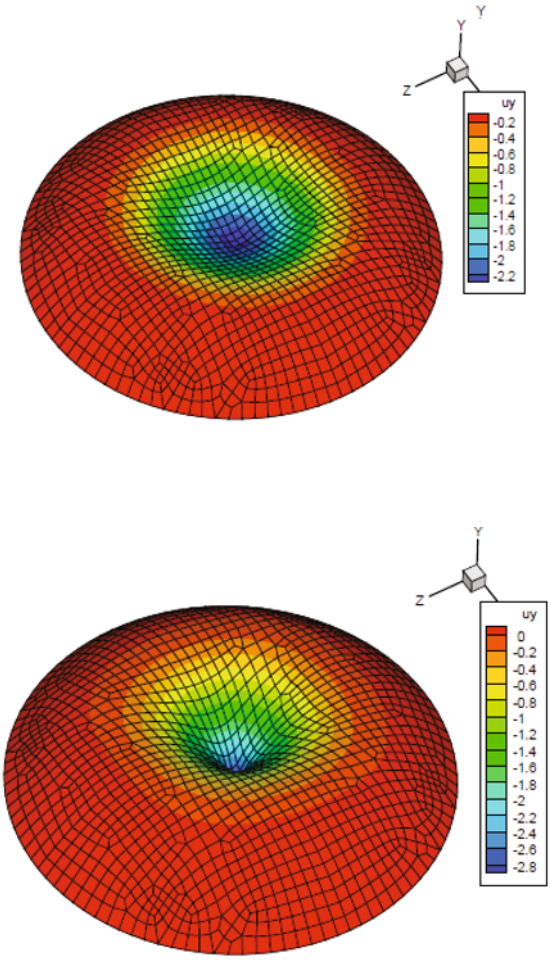


Figure 9. Vertical displacement field for a first position of the load. Complete model (up) vs. reduced model (down).

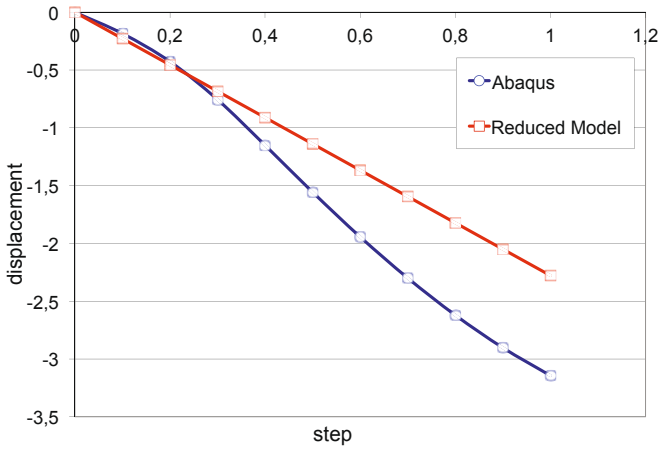


Figure 10. load-displacement curve at the vertex of the cornea.

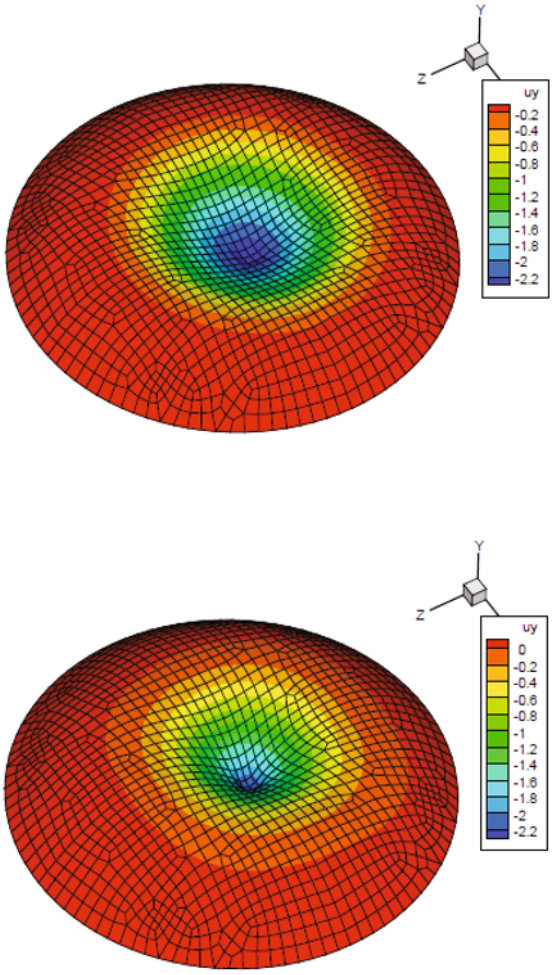


Figure 11. Vertical displacement field for a second position of the load. Complete model (up) vs. reduced model (down).

is allowed, of course. However, all the simulations presented here ran on a PC equipped with two processors (only one was employed, no parallel computing was used) AMD Quad Opteron running at 2.2 GHz and with 16 Gb RAM, under Scientific Linux. The prototype code was implemented under MATLAB.

The simulations ran at 472-483Hz, compliant with the highest requirements for haptic feedback, for instance.

4.1 Force located at an arbitrary point

Of course, one of the key aspects of the construction of a suitable reduced model based on POD is that of obtaining a good basis. Snapshots should be taken from representative states (here, load states) of the system. It is therefore of utmost importance to know the quality of the response of the system to a force located in a position whose response has not been calculated. In this work, we employ the PODI approach mentioned before (Ly and Tran, 2005). Therefore, the worst case scenario is to apply a load between two loads for which the complete, full-order, solution is known, and to compare it with a full order solution.

Under very large strains, the results of the full order simulation are shown in Fig. 12, while results for the PODI model are, as can be noticed, in good agreement, see Fig. 12(b).

The error in the prediction of the vertical displacement under the load is 27.18% at the end of the simulation (maximum of the strain). Error is computed in the $||\cdot||_2$ norm, defined as:

$$||e||_2 = \frac{1}{n}\sqrt{\sum_{I=1}^{n} e_I^2} \tag{18}$$

where e_I represents the nodal error and n the number of nodes in the model. Maximum error took a value of 29.5%, still within the limits for the best techniques available today for linear elastic materials (Lim and De, 2007).

4.2 Other approaches to the problem

Despite this classical approach to non-linear problems (simply to ignore the non-linearities), many solutions have been proposed recently in the literature of reduced order models. One of the possibilities is to employ POD in conduction with Asymptotic Numerical Methods (ANM), a technique due to M. Potier-Ferry and coworkers (Cochelin et al., 1994a; Abichou et al., 2002; Cao and Potier-Ferry, 1999; Cochelin et al., 1994b).

Let us introduce in a simple manner the ANM. Consider a very simple

non-linear problem:

$$\frac{\partial u}{\partial t} - k\Delta u = u^2 + f(\mathbf{x}, t)$$

The non linear term can be affected by a "loading" parameter λ

$$\frac{\partial u}{\partial t} - k\Delta u = \lambda \cdot u^2 + f(\mathbf{x}, t) \tag{19}$$

We denote by u_0 the solution related to $\lambda = \lambda_0 = 0$ that can be computed easily because it corresponds to the solution of the linear problem. The sought solution is the one related to $\lambda = 1$. The ANM method is based upon defining an asymptotic expansion of the unknown field u as well as of the loading parameter λ by considering the powers of a:

$$\left\{ \begin{array}{l} u = u_0 + a \cdot u_1 + a^2 \cdot u_2 \ + \cdots \\ \lambda = \lambda_0 + a \cdot \lambda_1 + a^2 \cdot \lambda_2 \ + \cdots \end{array} \right. \tag{20}$$

This non linear term can be written as:

$$u^2 = (u^2)_0 + a \cdot (u^2)_1 + a^2 \cdot (u^2)_2 \ + \cdots + a^p \cdot (u^2)_p + \cdots$$

where $(u^2)_p$ reads:

$$(u^2)_p = \sum_{i=0}^{i=p} u_i \cdot u_{p-i} = 2 \cdot u_0 \cdot u_p + \sum_{i=1}^{i=p-1} u_i \cdot u_{p-i} \tag{21}$$

Introducing (20) and (21) into (19) and identifying the different powers of a, it results:

- Power $p = 0$. Solution u_0 is assumed known because it corresponds to the solution of the linear problem resulting from $\lambda = \lambda_0 = 0$.
- Power $p = 1$. We should solve:

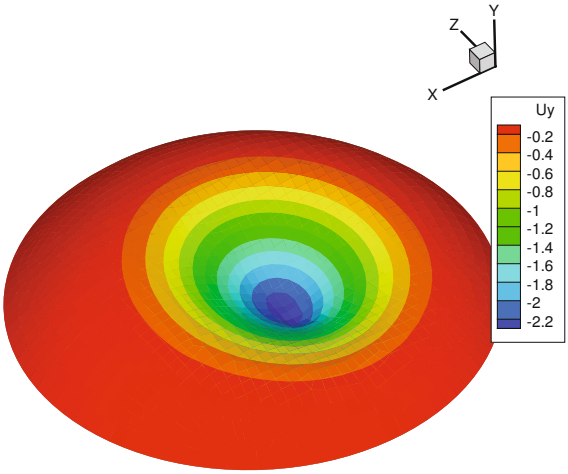
$$\frac{\partial u_1}{\partial t} - k\Delta u_1 = \lambda_0 \cdot (u^2)_1 + \lambda_1 \cdot (u^2)_0$$

and since $\lambda_0 = 0$, the evolution equation for u_1 reduces to:

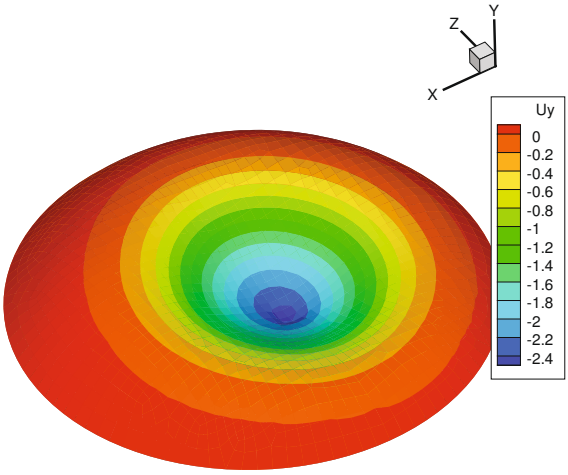
$$\frac{\partial u_1}{\partial t} - k\Delta u_1 = \lambda_1 \cdot (u^2)_0 \tag{22}$$

The previous equation contains two unknowns, u_1 and λ_1, so to obtain a solution we need for an extra equation. In Cochelin et al. (1994a) the following is considered:

$$(u - u_0, u_1) + (\lambda - \lambda_0, \lambda_1) = a \tag{23}$$



(a)



(b)

Figure 12. Vertical displacement (mm) for (a) the reduced model and (b) the complete model of a load in a position not previously computed.

where $(\cdot,\cdot)$ denotes the scalar product.

By algebraic considerations only, we arrive at an expression for the solution at power p, with $p > 1$:

$$\frac{\partial u_p}{\partial t} - k\Delta u_p = \lambda_p \cdot (u^2)_0 + \sum_{i=1}^{i=p-1} \lambda_i \cdot (u^2)_{p-i} \tag{24}$$

where we notice that the differential operator is always the same and that in the right hand member the term $\sum_{i=1}^{i=p-1} \lambda_i \cdot (u^2)_{p-i}$ has been already computed.

What is important in this method is the fact, already mentioned, that Eq. (24) contains always the same tangent operator for the term at any order of the expansion, the non-linear terms being re-directed to the force vector. Therefore, without the need to update the tangent stiffness matrix we arrive at a truly non-linear solution with high accuracy, see Niroomandi et al. (2010b).

Other approaches to this same problem are based upon the same philosophy, i.e., to interpolate the non-linear term of the equation. To do so, an appropriate set of points should be chosen so as to provide the best possible interpolated results. The so-called Empirical Interpolation Method (Chaturantabut and Sorensen, 2010) or the Best Points Interpolation Method (Nguyen et al., 2008) pertain to this same approach, although they differ in some aspects.

5 Conclusions

In this chapter a very brief overview of the most classical technique to construct reduced order models, i.e., that based upon the projection on a low dimensional subspaces obtained by Proper Orthogonal Decomposition, has been made. Undoubtedly, this is the most popular model order reduction technique, and therefore has been applied in field such as solid and fluid mechanics, chemical engineering, and many others.

However, as seen through the chapter, POD still presents two main difficulties. One os that it provides with the best (in a statistical sense) basis to perform a linear simulation (linear embedding). But if we deal with a non-linear problems, usually the price to pay is to reconstruct the reduced order matrix by updating the full order stiffness matrix of the original problem. Therefore, many of the computational savings are lost.

The second crucial aspects is that of constructing a good sample of snapshots, i.e., points at which the full order model is evaluated so as to attract the eigenmodes to be employed as basis functions in a Ritz-like framework. No definitive answer has been given to those tow big problems of POD.

Dealing with the construction of *a priori* reduced order models (i.e., those without the need for the solution of full order problems), we must cite the works by Ryckelynck et al. (2006) and, notably, the appearance of Proper Generalized Decompositions (PGD), developed independently by Ladeveze (1999) and Ammar et al. (2006a, 2007). PGD methods, that will be introduced in subsequent chapters of this book, allow for an a priori development of reduced order models without the need for the solution of complete ones, and therefore have constituted a little revolution in the field.

With respect to the other big problem, that of non-linear problems, several possibilities have been briefly overviewed in this chapter. This same problematic is common to PGD techniques. In general, either an approach based upon asymptotic expansions (Cochelin et al., 1994a) or one based on the interpolation of the non-linear terms are preferred (Chaturantabut and Sorensen, 2010; Nguyen et al., 2008).

In any case, model order reduction techniques have become a powerful technique in situations in which many-query frameworks are present (optimization, parametric problems, ...) or where real-time responses are mandatory (real-time simulation, (Niroomandi et al., 2010a), for instance).

Bibliography

H. Abichou, H. Zahrouni, and M. Potier-Ferry. Asymptotic numerical method for problems coupling several nonlinearities . *Computer Methods in Applied Mechanics and Engineering*, 191(51-52):5795–5810, 2002.

A. Ammar, B. Mokdad, F. Chinesta, and R. Keunings. A new family of solvers for some classes of multidimensional partial differential equations encountered in kinetic theory modeling of complex fluids. *J. Non-Newtonian Fluid Mech.*, 139:153–176, 2006a.

A. Ammar, D. Ryckelynck, F. Chinesta, and R. Keunings. On the reduction of kinetic theory models related to finitely extensible dumbbells. *Journal of Non-Newtonian Fluid Mechanics*, 134(1-3):136 – 147, 2006b. ISSN 0377-0257. doi: http://dx.doi.org/10.1016/j.jnnfm.2006.01.007.

A. Ammar, B. Mokdad, F. Chinesta, , and R. Keunings. A new family of solvers for some classes of multidimensional partial differential equations encountered in kinetic theory modeling of complex fluids. part ii: transient simulation using space-time separated representations. *J. Non-Newtonian Fluid Mech.*, 144:98–121, 2007.

D. Amsallem and C. Farhat. An interpolation method for adapting reduced-order models and application to aeroelasticity. *AIAA Journal*, 46:1803–1813, 2008.

R. A. Bialecki, A. J. Kassab, and A. Fic. Proper orthogonal decomposition and modal analysis for acceleration of transient fem thermal anal-

ysis. *International Journal for Numerical Methods in Engineering*, 62 (6):774–797, 2005. ISSN 1097-0207. doi: 10.1002/nme.1205. URL `http://dx.doi.org/10.1002/nme.1205`.

John Burkardt, Max Gunzburger, and Hyung-Chun Lee. {POD} and cvt-based reduced-order modeling of navier?stokes flows. *Computer Methods in Applied Mechanics and Engineering*, 196(1?3):337 – 355, 2006. ISSN 0045-7825. doi: http://dx.doi.org/10.1016/j.cma.2006.04.004.

H.-L. Cao and M. Potier-Ferry. An improved iterative method for large strain viscoplastic problems. *International Journal for Numerical Methods in Engineering*, 44:155–176, 1999.

Saifon Chaturantabut and Danny C. Sorensen. Nonlinear model reduction via discrete empirical interpolation. *SIAM Journal on Scientific Computing*, 32(5):2737–2764, 2010. ISSN 1064-8275. doi: 10.1137/090766498.

F. Chinesta, A. Ammar, A. Leygue, and R. Keunings. An overview of the proper generalized decomposition with applications in computational rheology. *Journal of Non-Newtonian Fluid Mechanics*, 166(11):578 – 592, 2011. ISSN 0377-0257. doi: http://dx.doi.org/10.1016/j.jnnfm.2010.12.012. ¡ce:title¿XVIth International Workshop on Numerical Methods for Non-Newtonian Flows¡/ce:title¿.

B. Cochelin, N. Damil, and M. Potier-Ferry. Asymptotic-numerical methods and Padé approximants for non-linear elastic structures. *International Journal for Numerical Methods in Engineering*, 37:1187–1213, 1994a.

B. Cochelin, N. Damil, and M. Potier-Ferry. The asymptotic numerical method: an efficient perturbation technique for nonlinear structural mechanics. *Revue Europeenne des Elements Finis*, 3:281–297, 1994b.

Max D. Gunzburger, Janet S. Peterson, and John N. Shadid. Reduced-order modeling of time-dependent {PDEs} with multiple parameters in the boundary data. *Computer Methods in Applied Mechanics and Engineering*, 196(4?6):1030 – 1047, 2007. ISSN 0045-7825. doi: http://dx.doi.org/10.1016/j.cma.2006.08.004.

P. Ladeveze. *Nonlinear Computational Structural Mechanics*. Springer, N.Y., 1999.

P. Ladeveze, J.-C. Passieux, and D. Neron. The {LATIN} multiscale computational method and the proper generalized decomposition. *Computer Methods in Applied Mechanics and Engineering*, 199(21?22):1287 – 1296, 2010. ISSN 0045-7825. doi: http://dx.doi.org/10.1016/j.cma.2009.06.023. ¡ce:title¿Multiscale Models and Mathematical Aspects in Solid and Fluid Mechanics¡/ce:title¿.

Yi-Je Lim and Suvranu De. Real time simulation of nonlinear tissue response in virtual surgery using the point collocation-based method of

finite spheres. *Computer Methods in Applied Mechanics and Engineering*, 196:3011–3024, 2007.

Hung V. Ly and Hien T. Tran. Modeling and control of physical processes using proper orthogonal decomposition. *Mathematical and Computer Modelling*, 33:223–236, 2005.

Y. Maday and E. Ronquist. The reduced basis element method: Application to a thermal fin problem. *SIAM Journal on Scientific Computing*, 26(1):240–258, 2004. doi: 10.1137/S1064827502419932. URL `http://epubs.siam.org/doi/abs/10.1137/S1064827502419932`.

N. C. Nguyen, A. T. Patera, and J. Peraire. A 'best points' interpolation method for efficient approximation of parametrized functions. *International Journal for Numerical Methods in Engineering*, 73(4):521–543, JAN 22 2008. ISSN 0029-5981. doi: 10.1002/nme.2086.

S. Niroomandi, I. Alfaro, E. Cueto, and F. Chinesta. Real-time deformable models of non-linear tissues by model reduction techniques. *Computer Methods and Programs in Biomedicine*, 91(3):223–231, 2008.

S. Niroomandi, I. Alfaro, E. Cueto, and F. Chinesta. Accounting for large deformations in real-time simulations of soft tissues based on reduced-order models. *Computer Methods and Programs in Biomedicine*, In Press, Corrected Proof:–, 2010a. ISSN 0169-2607. doi: DOI: 10.1016/j.cmpb.2010.06.012.

Siamak Niroomandi, Iciar Alfaro, Elias Cueto, and Francisco Chinesta. Model order reduction for hyperelastic materials. *International Journal for Numerical Methods in Engineering*, 81(9):1180–1206, 2010b. ISSN 1097-0207. doi: 10.1002/nme.2733. URL `http://dx.doi.org/10.1002/nme.2733`.

Anthony Nouy. A priori model reduction through proper generalized decomposition for solving time-dependent partial differential equations. *Computer Methods in Applied Mechanics and Engineering*, 199(23-24):1603 – 1626, 2010. ISSN 0045-7825. doi: http://dx.doi.org/10.1016/j.cma.2010.01.009.

H. M. Park and D. H. Cho. The use of the Karhunen-Loève decomposition for the modeling of distributed parameter systems. *Chemical Engineering Science*, 51(1):81–98, 1996.

D. Ryckelynck. A priori model reduction method for the optimization of complex problems. In *Workshop on Optimal Design of Materials and Structures, Ecole Polytechnique, Palaiseau, Paris (France)*, 2003.

D. Ryckelynck. A priori hyperreduction method: an adaptive approach. *Journal of Computational Physics*, 202(1):346–366, 2005.

D. Ryckelynck. Hyper-reduction of mechanical models involving internal variables. *International Journal for Numerical Methods in Engineering*, 77(1):75–89, 2008.

D. Ryckelynck, F. Chinesta, E. Cueto, and A. Ammar. On the a priori Model Reduction: Overview and recent developments. *Archives of Computational Methods in Engineering*, 12(1):91–128, 2006.