Universidad Zaragoza
1542

Tesis Doctoral

Real-Time simulation of surgery by model reduction and X-Fem techniques

Autor

Siamak Niroomandi

Director/es

Icíar Alfaro Ruiz
Elías Cueto Prendes

Centro Politécnico Superior
Departamento de Ingeniería Mecánica

2011

Real time simulation of surgery by model reduction and X-FEM techniques

Author: Siamak Niroomandi

Directors: Icíar Alfaro and Elías Cueto

Doctoral Degree in
Computational Mechanics

May, 2011. Draft version.

Universidad Zaragoza
1542

Group of Structural Mechanics and Material Modeling

Real time simulation of surgery by model reduction and X-FEM techniques

Author: Siamak Niroomandi

Directors: Icíar Alfaro and Elías Cueto

Doctoral Degree in
Computational Mechanics

May, 2011. Draft version.

Universidad Zaragoza
1542
Group of Structural Mechanics and Material Modeling

Real time simulation of surgery by model reduction and X-FEM techniques

Abstract

Interactive (that with haptic feedback) simulation of surgery needs for increasingly fast simulation techniques. Feedback rates are nowadays fixed around 1 kHz for haptic peripherals and around 30 Hz for visual display. Such tremendous rates impose very drastic limitations to the simulation procedure. In order to achieve haptic realism, several techniques have been tested in the literature. These include mass-spring systems, the use of linear elastic explicit finite element codes or, more recently, the use of Graphic Processing Units (GPU) to speedup the simulation.

In any case, in order to achieve realistic results compatible with the sense of touch, large strains must be taken into account, together with sophisticated, state of the art constitutive models for the soft tissues. This has not been achieved to date, up to the author's knowledge.

Techniques based on model order reduction (MOR) have received an increasing attention in the last years, and have been employed for interactive simulation of linear elastic solids undergoing large strains (Barbi**c** and James, 2005). In this thesis, however, a study has been made in order to discern if MOR techniques are suitable for the simulation of soft living tissues.

In chapter 2 a technique based upon Proper Orthogonal Decomposition (POD) has been employed to simulate at haptic rates the governing equations for a human cornea under a hyperelastic, fiber-reinforced constitutive model. While the technique is able to efficiently simulate these models, errors up to 20% have been noticed, still acceptable.

In chapter 3 a novel technique has been developed that allows for an accurate solution of material and geometrical non-linear hyperelastic models without the need of stiffness matrix inversions. It is based upon the combination of POD techniques and Asymptotic Numerical Methods, which provide an accurate description of the non-linear stress-strain curve of these organs in a convergence interval of sufficient width.

Finally, the use of globally supported basis functions, an essential characteristic of POD techniques, very much complicates the task of simulating surgical cutting. To that end, a multiscale method has been developed that enriches the displacement field with

an X-FEM-like, discontinuous, field that avoids the need of remeshing, impossible under such severe time restrictions.

Real time simulation of surgery by model reduction and X-FEM techniques

Resumen y Conclusiones

La simulación quirúrgica virtual presenta retos formidables para las técnicas de simulación numérica. Para aquellos sistemas en los que se busca una respuesta háptica (aquella en la que los periféricos proporcionan una respuesta en fuerza), ésta debe calcular a una frecuencia de alrededor de 1 kHz. Si lo que se pretende es un realismo visual, por ejemplo, entonces unos 30 Hz son suficientes.

Para lograr estas cifras se han empleado recientemente técnicas de muy variada índole. Los primeros intentos que se pueden encontrar en la literatura están basados en el uso de sistemas de masa-muelles que no reproducen siquiera las ecuaciones de la Elasticidad lineal. Más recientemente, diversos autores han incorporado métodos de elementos finitos explícitos o el cálculo paralelo en tarjetas gráficas (GPU).

Recientemente, Barbi**c** and James (2005) han desarrollado un método basado en la reducción de modelos que proporciona resultados muy esperanzadores para el caso de ecuaciones lineales de comportamiento sometidas a grandes deformaciones.

En esta tesis se ha planteado el desarrollo de una técnica basada en el uso de modelos reducidos que sea capaz de simular de una manera apropiada las complejas (frecuentemente, hiperelásticas reforzadas con fibras) ecuaciones de comportamiento de los tejidos vivos blandos. En el capítulo 2 se presenta una técnica capaz de ello, aunque presentando un nivel de error no despreciable, todavía compatible con los requerimientos más habituales en el campo.

En el capítulo 3 la técnica mencionada se ha generalizado, hasta ser capaz de simular materiales no lineales tanto en sus aspectos materiales como geométricos. Esta técnica, basada en la combinación de la descomposición ortogonal apropiada (POD, por sus siglas en inglés) y el método asintótico numérico, proporciona excelentes resultados en un entorno de una solución previamente calculada, sin tener que calcular actualizaciones de la matriz de rigidez tangente.

El uso de funciones de aproximación globales (Ritz), sin embargo, complica tremendamente la simulación del corte quirúrgico. A tal fin se ha desarrollado una nueva téc-

nica basada en una metodología multiescala que combina los modelos reducidos presentados en los capítulos anteriores con el X-FEM, de forma que no resulta necesario remallar para simular el corte.

Conclusiones

Las conclusiones más sobresalientes de esta tesis son las siguientes:

- Se ha desarrollado una técnica de reducción de modelos que permite, todavía con un cierto nivel de error, simular a frecuencias hápticas materiales complejos, especialmente hiperelásticos.
- Para minorar los niveles de error, las técnicas basadas en la combinación de modelos reducidos y métodos asintóticos numéricos permiten seguir complejas trayectorias no lineales in necesidad de actualizar e invertir nuevas matrices de rigidez tangentes.
- Finalmente, se ha demostrado que la combinación de las técnicas de reducción de modelos y de X-FEM permiten una simulación especialmente apropiada de los procesos de corte quirúrgico.
- En general, las técnicas de reducción de modelos se muestran como una alternativa prometedora en el campo de la simulación quirúrgica.

Contents

Contents

List of Figures

Document

Chapter 1

Introduction

1.1 Introduction

Minimally invasive surgery (MIS) has become more and more common in the last decade. A minimally invasive procedure typically involves use of laparoscopic devices and remote-control manipulation of instruments with indirect observation of the surgical field through an endoscope or similar devices. These instruments are inserted through the skin or through a body cavity or anatomical opening. This may result in less pain, less strain of the organism, small injuries (aesthetic reasons), economic gain because of shorter hospital stays, etc. On the other hand, there exist some important difficulties for the surgeon as a result of his restricted vision of the organs, difficult handling of the instruments, very restricted mobility, difficult hand-eye coordination and no tactile perception. Therefore, an important training phase is required before a surgeon acquires the skills necessary to adequately perform minimally invasive surgery. Currently, surgeons are trained to perform minimally invasive surgery by using mechanical simulators —just like plane pilots have been trained for many years—or living animals. There is some consensus on the limited realism of the mentioned simulators, see Figs. 1.1, 1.2, due to the complexity of the simulations to be carried out (Delingette and Ayache, 2004). The latter training method consists in practicing simple or complex surgical procedures on living animals (often pigs for abdominal surgery).

Because of the limitations of current training methods, there is a large interest in developing surgery simulation software, possibly with haptic feedback, for providing efficient and quantitative gesture training systems (Delingette, 2000). Haptic feedback takes advantage of a user's sense of touch by applying forces, vibrations, and/or motions to the user.

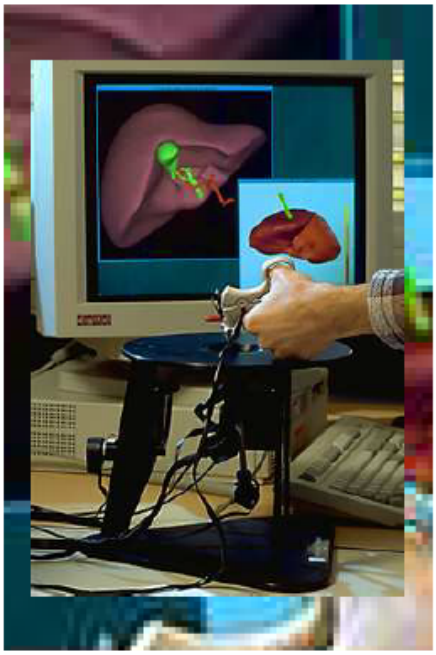


Figure 1.1. Surgical simulator at INRIA research center, France.

Figure 1.2. The only spanish surgical simulator, *insight*, made by GMV `http://www.gmv.com`.

Surgical simulators can be classified into three categories or generations (Satava, 1996). The so called first generation of the simulators is based on the anatomical information. And mainly the geometry of the organs are included in the simulator. The shape, surface, volume and morphology of the organs are considered in the simulators of this category. The second generation of simulators describes the physical properties of the human body; it enables the simulation of basic surgical gestures such as cutting or suturing. There are some simulators of this kind for the simulation of cholecystectomy, arthroscopy of the knee and hepatectomy (Delingette and Ayache, 2004). Simulators in this category are capable of computing the deformation of the soft tissue under forces and they can also take into account temperature changes. The simulators of the third generation go one step further and consider some aspects of the physiological behavior of the human body. In this way they are intended to model blood flow, organ movements and certain diseases for instance. A schematic of these three generations of simulators is shown in Fig. 1.3.

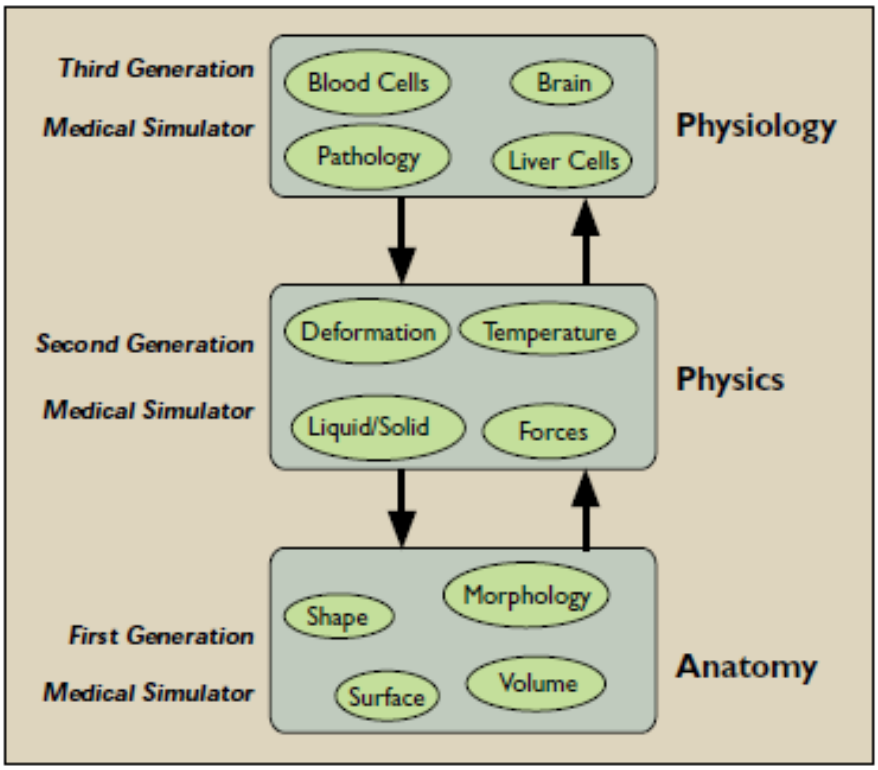


Figure 1.3. The three generations of surgical simulators. Cf. Delingette and Ayache (2004).

Although the latter category of simulators is very promising it is very difficult to take into account all those details in the simulator because of the enormous amount of com-

putation due to coupling between physiology and physics of the organs. To our knowledge no such simulator has been developed yet, although some instances are able to include tumors, for instance, in brain surgery. NeuroTouch, for instance, that is used for brain surgery, was developed with collaboration of about fifty researchers and doctors from ten research centers and funded by National Research Center of Canada (project, 2009).

In general, a surgical simulator consists of several components: input devices, the core of the simulator and output devices (Delingette and Ayache, 2004). Fig. 1.4 shows the architecture of a simulator. The input devices consist of a force-feedback (haptic) device with which the user moves the virtual surgery tools, a mouse and a keyboard to have a good view of the operation zone. The core of the simulator does the geometrical and the physical modeling. Geometrical models are obtained using medical imagery like Computerized Tomography (CT-scan) and Magnetic Resonance Imaging (MRI) after they have been converted to standard 3D graphics file formats. Physical models depend on the specific soft tissue. The core of the simulator also detects collision between virtual tools and the soft tissue and then calculates the corresponding reaction forces on the tools. The computed data is sent to the output devices which consist of a screen and a haptic device. The haptic device transforms reaction forces and moments to the user and makes him/her have the perception of virtual contact with the tissue. Fig. 1.5 displays a diagram of the processes that are performed in a surgical simulator. There are some requirements for visual and haptic feedback that make simulations difficult to perform in real time. A surgery simulator must provide a realistic visual presentation of the surgical procedure. Visual feedback is specially important in video-surgery because it helps the surgeon to acquire a tridimensional perception of the environment. If the positions, orientations and deformations of objects on the screen are updated at a rate less than about 30 times per second, users will no longer perceive the simulation as continuous motion. Haptic feedback provides the sensation of the movement to the user and therefore it significantly enhances his surgical performance. But it should have a frequency between 300 and 1000 Hz that is very difficult to achieve for complex tissue behavior and usually a trade-off between computation time and accuracy is adopted. Fig. 1.6 shows a comparison of accuracy and computation time in various disciplines.

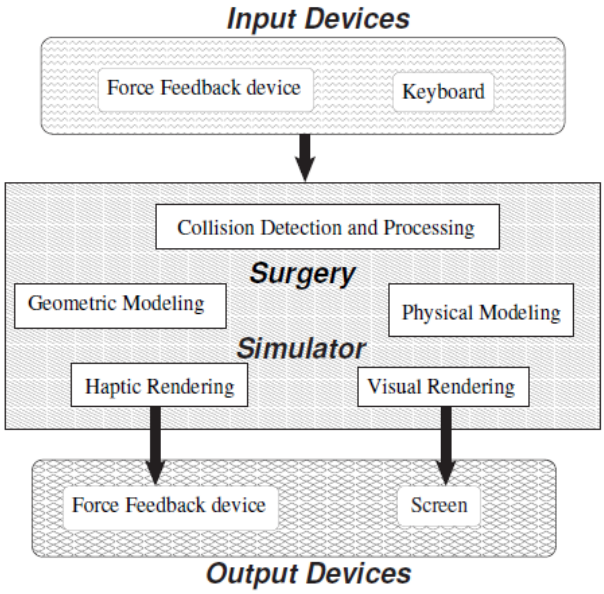


Figure 1.4. Schematic architecture of a surgical simulator (Delinguette, 1998).

1.2 State of the art

In order to comply with these requirements different methods have been used during last years often with trade-off between time and accuracy, see Fig. 1.7. Among them are mass-spring methods, surface models, finite element and boundary element methods. Spring-mass models consist of points with mass that are linked by springs and dashpots. The spring stiffness and damping constant are usually determined by experiments. They have been used extensively for simulating soft tissues (Waters, 1992), (Kenedi et al., 1975), (Delinguette et al., 1994). The main advantage of spring models is their ease of implementation and they have been used for statics and dynamics analysis. Another advantage is their ability to model cutting and suturing simply by removing or adding connections between vertices. However, there are some disadvantages with these models. The most important is that, in general, they do not reproduce the laws of linear elasticity and conservation of mass (Meier et al., 2005).

Most of the nowadays simulators are based on the laws of continuum mechanics. Among the available numerical methods for solving these models, finite element and

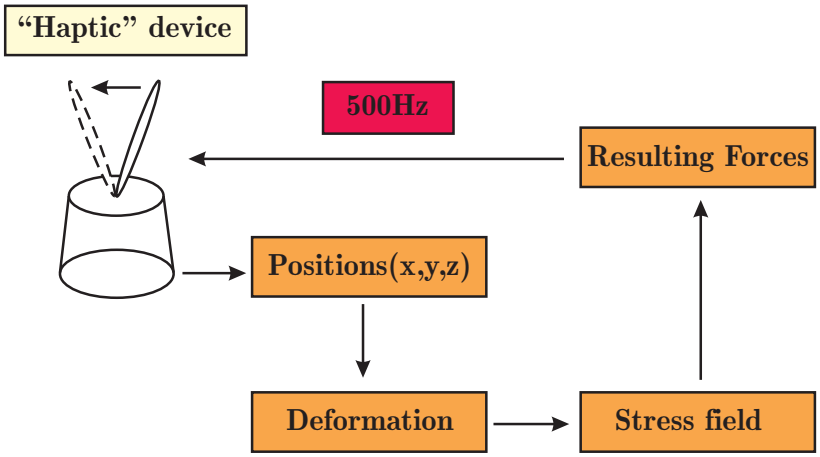


Figure 1.5. Simplified mechanism of a surgical simulator.

boundary element methods can be mentioned (Monserrat et al., 2001). A review of different soft tissue models based on continuum mechanics is given in Famaey and Stolen (2008). Many of these simulators exploit the laws of linear elasticity in which infinitesimal strain and small displacements are assumed. By doing so a lot of time is saved by pre-computation of stiffness and mass matrices. Although if the body undergoes large displacements, which is nearly always the case, the amount of error will be substantial. In order to solve this issue large displacements can be considered in the model which gives rise to the so called geometrically non-linear models. Fig. 1.8 shows a beam modeled with linear elasticity in the above figure and considering large strains in the below figure. As can be noticed, the difference is substantial, with a neat gain of volume in the former case. The result is that the surgeon perceives the motion as unrealistic.

However, it is known that soft tissues are composed of collagen fibers and include

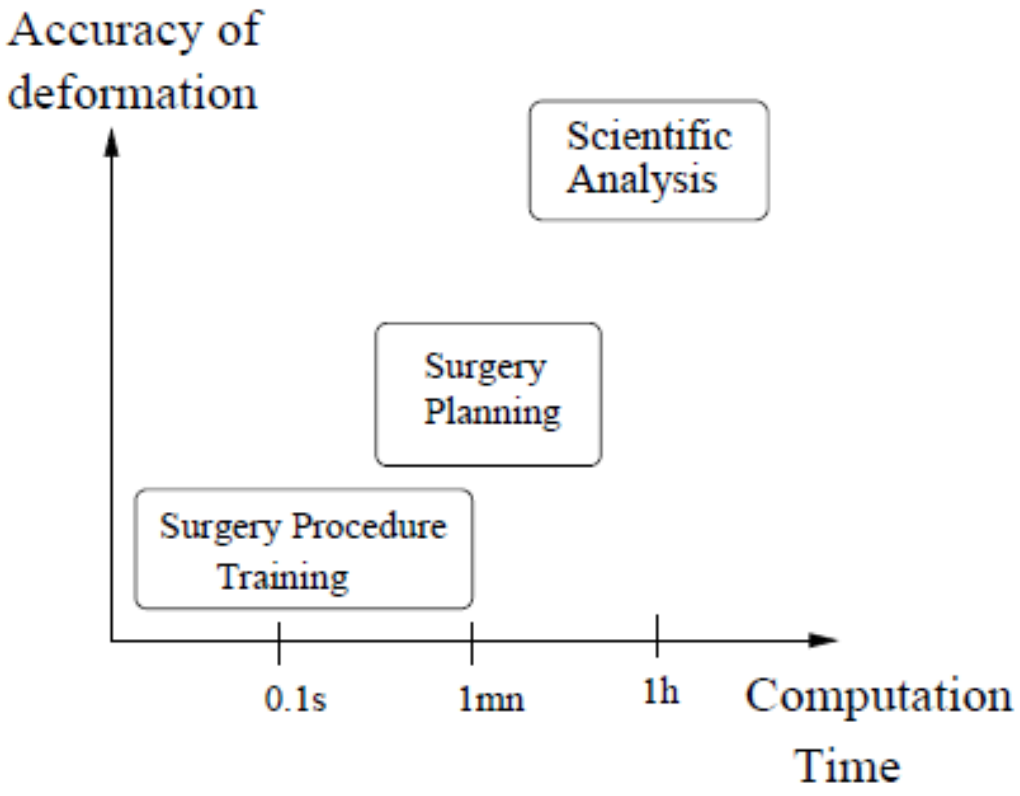


Figure 1.6. Comparison of accuracy vs. computation time in various disciplines. Cf. Delinguette (1998).

considerable amount of water (Holzapfel and Gasser, 2000). This complex composition cannot be well simulated using only geometrically non-linear models and other constitutive laws should be used for a better approximation of the behavior of these tissues. Among them the so-called Holzapfel model (Holzapfel and Gasser, 2000) and neo-Hookean strain energy functions (Bonet and Wood, 2008) are very common.

Other methods like mesh free methods have been recently used for real-time simulation of soft tissues. Lim and De (2007) applied the point collocation-based method of finite spheres (PCMFS) technique to simulate tissue deformations that are geometrically nonlinear. The technique is based on a novel combination of a multi-resolution approach coupled with a fast analysis scheme in the nonlinear model.

Capell et al. (2002) computed the equations of motion for a linearly elastic material using FEM and then by using coarse volumetric meshes to do simulations at interactive rates.

A system based upon the use of neural networks has been presented in Deo and De (2009). In it, the system is trained with a large set of possible load states in order to

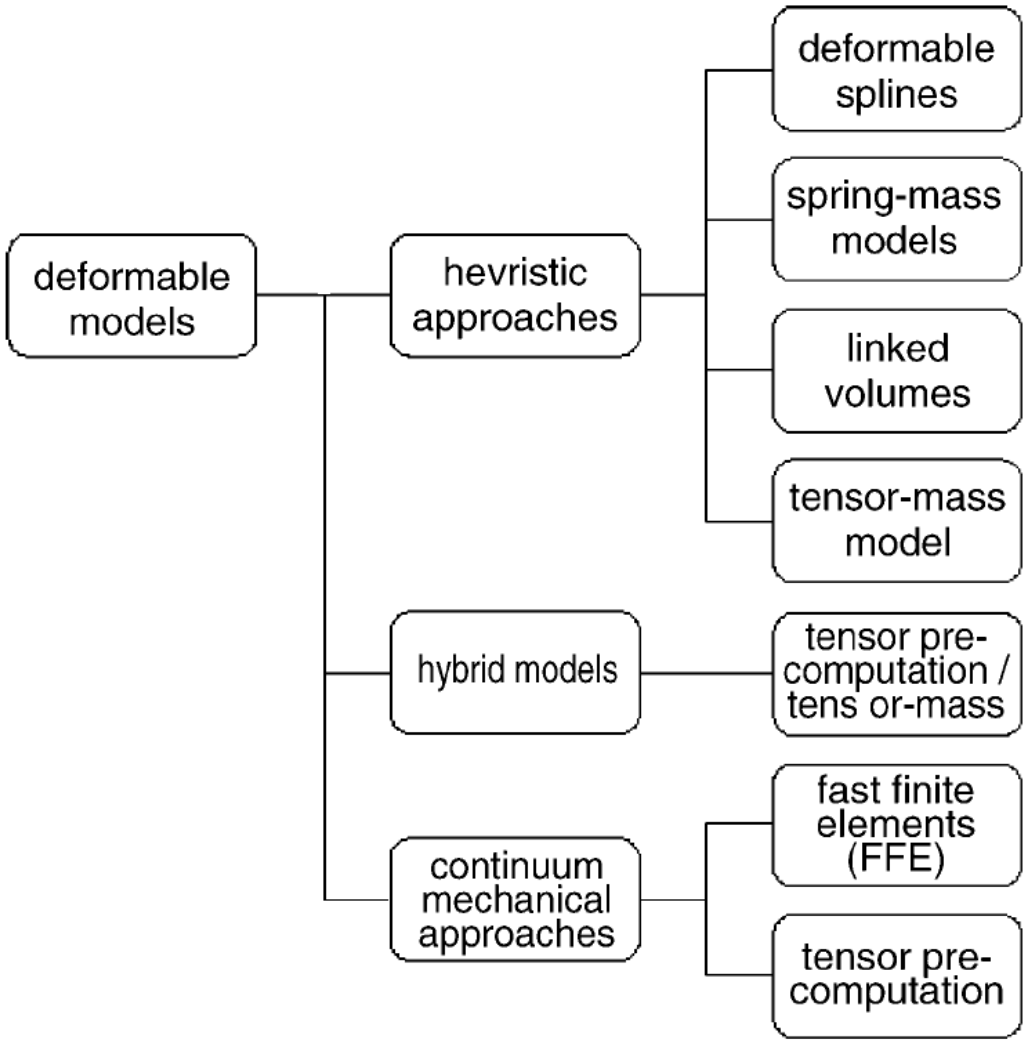


Figure 1.7. Different methods tried so far for surgery simulation, Cf. Meier et al. (2005). The list is not up to date and does not include some of the most promising attempts made with GPUs.

achieve real-time performance in the execution loop.

Very recently, General Purpose Graphic Processing Units (GPGPUs) have been incorporated into the field of surgical simulation. It is is able to perform very fast operations, and a complete non-linear, explicit, finite element code has been implemented in that work with a gain on speed of the order of 20 times (Taylor et al., 2008). This approach has allowed to implement medium-sized models (16000 tetrahedrons) including neo-Hookean behavior in an explicit total Lagrangian approach. The main drawback of this elegant approach to the problem is the conditional stability of explicit integration procedures and also the explicit evaluation of the constitutive laws.

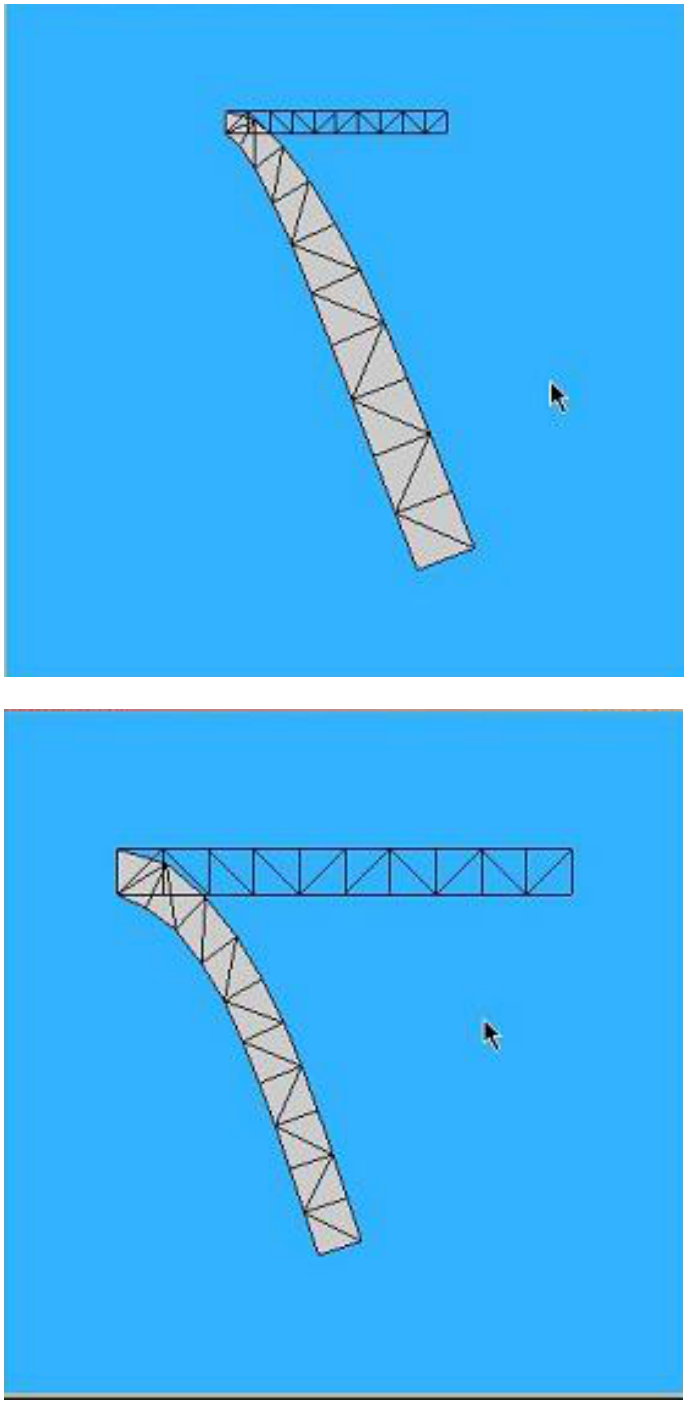


Figure 1.8. Beam modeled using linear elasticity laws (top) and geometrically non-linear strain measures (bottom).

1.3 Model Order reduction techniques

Model Order Reduction (MOR) in computational mechanics has a strong tradition, starting from the early years in which the performance of computers implied the need for models with very few degrees of freedom (Idelsohn and Cardona, 1985; Krysl et al., 2001). But similar techniques are employed in many other branches of Engineering and Applied Sciences: real-time simulation (Barbi**c** and James, 2005), Chemical Engineering (Park and Cho, 1996), turbulence (Sirovich, 1987), weather forecast (Lorenz, 1956) and many others. This method of analysis has wide application in scientific problems in which one deals with dauntingly large datasets resulting from large-scale computing. The slant of this exposition is that typically large datasets intrinsically define an optimal means for their own treatment.

This technique goes back to the beginning of the century (Pearson, 1901) and has been associated with a variety of names, since it has been employed and rediscovered in many branches of science and engineering. Maybe the most common names are Karhunen-Loève decomposition (Karhunen, 1946; Loève, 1963), proper orthogonal decomposition, POD, (Sirovich, 1987) or empirical orthogonal functions (Lorenz, 1956). Karhunen-Loève (KL) decomposition or proper orthogonal decomposition (POD) will be used throughout this text.

Recent applications of model reduction techniques for the analysis of non-linear solids and structures include the works by Krysl et al. (2001) and the works of Ryckelynck et al. for "a priori" model reduction (see Ryckelynck (2003, 2005); Ryckelynck et al. (2006)). Also, recently, new extensions of the method have been developed, giving rise to the so-called *non-linear dimensionality reduction* (Tenenbaum et al., 2000). In Krysl et al. (2001) MOR has been applied to explicit and implicit time integration of the equations of elastodynamics. As most of the computation time is devoted to solving many systems of equations, by exploiting MOR the computation time reduces by about 40 percent. Such saving is quite considerable especially in optimization and control that a great number of problems are to be solved. Although this is a general approach but computations cannot be performed in real time.

BaBarbi**c** and James (2005), however, exploited dimensional model reduction techniques to build reduced-coordinate deformable models for bodies with complex geometry using St. Venant-Kirchhoff constitutive laws. In this way they have performed simulations at real-time update rates of geometrically-nonlinear materials. Even though the method seems not to be generalized to more complex material behaviors like neo-Hookean constitutive laws, they have shown that Proper Orthogonal Decomposition,

which is not exactly their approach, is a promising tool for real-time applications.

In general, all these techniques share some common characteristics. It is necessary to have some previous data coming from previously computed, detailed models (in this frameworks, the work by Ryckelynck (2005) is an exception). These data can be obtained after numerical simulations made off-line and stored in memory. But they can be also obtained from physical experiments, for instance. For the work here presented the first option has been chosen, as will be explained later, and FE models of the organs being simulated will be considered as an "exact" to compare with. From these data the relevant information about the (non-linear) behavior of the tissues is extracted by Karhunen-Loève decompositions. Then they are employed to construct a very fast Galerkin method with very few degrees of freedom. FE methods have had a tremendous success in many branches of science and engineering because of their simplicity and good performance in many fields. They employ piece-wise polynomials as a basis to construct an interpolation of the unknown field of interest (usually the displacement field, but also velocities, pressure and stresses are unknowns in different FE formulations). These piece-wise polynomials are very simple and general functions defined over a very restricted domain (the element itself and the elements connected to that). In model reduction techniques we employ global basis functions that have the whole domain as their support; But they are of "high quality". This means that in the construction of these basis we employ available information on the problem. The more information we have, the better quality basis functions will be constructed. And, of course, the better results will be obtained.

Models that differ slightly from the original ones can thus be computed with a very important, sometimes impressive, computational saving, employing only a very limited number of degrees of freedom.

1.4 Objectives

Real time requirements pose very severe constraints in the simulation method to be employed, as commented before. One of these limitations is that, in general, there is not time enough to perform tangent stiffness matrix updates for non-linear constitutive equations of soft tissues. This is why very few works have actually dealt with material non-linearities at haptic feedback rates.

In this thesis the objective has been set to analyze if model order reduction methods, and more specifically, POD techniques, can be employed to this end. As will be

seen, POD provides with very good results (as good as existing methods, at least), but for much more sophisticated constitutive laws. This can be considered as a true advancement in the state of the art. However, this level of error is very often around 20%, and therefore can be improved.

The second objective of this thesis is the development of a method that, without tangent stiffness matrix updates, could follow complex, non-linear material paths. This will be accomplished by employing a combination of POD and asymptotic expansions of the problem at equilibrium points.

Finally, the choice of model order reduction techniques seems to allow for a much faster simulation, even for very complex and non-linear constitutive equations for soft tissues. But it very much complicates the issue of simulating surgical cutting. A cut would imply a change in the mesh, which constitutes a task not feasible at $500\ Hz$ rates. X-FEM seems to be an appealing choice to avoid remeshing, but it is not readily applicable to Ritz (globally-supported shape functions) Galerkin methods. Therefore, the third big objective of this thesis is to provide the method with an efficient means for the simulation of surgical cutting.

Krysl et al. (2001) performed a thorough analysis of the application of POD techniques to solve the equations of elastodynamics, and also some interesting conclusions can be obtained after the work of Barbi**c** and James (2005), so this thesis will restrict itself to the aspects related to material non-linearities.

1.5 Thesis overview

The organization of the thesis is as follows: the fundamentals of model reduction techniques are explained in chapter 2, where some applications of these techniques in virtual surgery are shown, demonstrating the suitability of POD for virtual surgery simulation. In chapter 3 the asymptotic numerical method (ANM) is firstly introduced and then the formulation for geometrically-nonlinear problems is reviewed. A combination of POD-ANM is then developed that allows to follow complex, non-linear, load trajectories without the need for stiffness matrix updating. Then this formulation is extended for non-linear hyperelastic materials undergoing large strains. Different examples in solid mechanics will be presented that show the performance of the proposed technique. In chapter 4 it is shown how eXtended Finite Element methods can be used in a combination with POD to model discontinuities with special attention to cutting procedure as a specific example. Finally, the conclusions obtained from this thesis and the

future work are presented in chapter 5.

Chapter 2

Real-time Simulation of Surgery based upon Proper Orthogonal Decomposition

2.1 Introduction

One of the main limitations of existing models in real-time simulation is that they do not take into account the anisotropic and highly non-linear response of virtually all soft tissues, see Meier et al. (2005) for a recent survey on the topic. Recently, geometric non-linearities have been taken into account in a work also based on model reduction, see Barbi**c** and James (2005). But in this case, only linear materials have been considered (i.e., the so-called Saint Venant-Kirchhoff model, or homogeneous isotropic linear elastic materials undergoing large strains). Most soft tissues, however, exhibit complex non-linear responses, possibly with anisotropic characteristics, and frequently show incompressible or quasi-incompressible behavior. Geometric non-linearities (those deriving from large strains) should be also considered on top of this complex material behavior. The standard simulation of these material non-linearities requires the employ of Newton-Raphson or similar techniques in an iterative framework. In other words, a tangent stiffness matrix must be inverted, possibly many times, at each time step. This makes the existing engineering FE codes impractical for real-time simulations.

While most of the existing simulation techniques in real-time, as mentioned before, are based upon Finite Element or Boundary Element techniques, we have pursued a different philosophy. Following Bro-Nielsen and Cotin (1996), "... We do not care about

the time taken for one-time pre-calculation such as setting up equations, inverting matrices, etc.".

The technique here presented is based upon existing data on the behavior of the simulated tissues. These data can be obtained after numerical simulations made off-line and stored in memory. But they can be also obtained from physical experiments, for instance. For the work here presented we have chosen the first option, and FE models of the organs being simulated will be considered as an "exact" solution to compare with. From these data the (statistically speaking) relevant information about the (non-linear) behavior of the tissues is extracted by Karhunen-Loève decomposition. Then we employ it to construct a very fast Galerkin method with very few degrees of freedom.

To this end, we employ model reduction techniques based on proper orthogonal decomposition (Karhunen, 1946), (Loève, 1963), (Lorenz, 1956), (Ryckelynck et al., 2006). FE methods have had a tremendous success in many branches of science and engineering because of their simplicity and good performance in many fields. They employ piece-wise polynomials as a basis to construct an interpolation of the unknown field of interest (usually the displacement field, but also velocities, pressure and stresses are unknowns in different FE formulations). These piece-wise polynomials are very simple and general functions defined over a very restricted domain (the element itself and the elements connected to that). In model reduction techniques we employ global basis functions that have the whole domain as their support; But they are of "high quality". This means that in the construction of these basis we employ available information on the problem. The more information we have, the better quality basis functions will be constructed. And, of course, the better results will be obtained.

To obtain the information necessary to construct these "good" basis functions, as mentioned before, we employ Proper Orthogonal Decomposition (POD) techniques. And these can be built upon finite element results (or, in general any numerical simulation results, if available) or also upon experimental results. Of course, these simulations are made off-line and their results are stored prior to starting the on-line simulation.

In order to show the performance of the method, the behavior of the human cornea is chosen to be simulated, although the technique is equally applicable to any other soft tissue. Other tissues, such as bones, that in short periods of time present almost linear response, could be also simulated with this technique, obtaining even better results. The cornea presents a highly non-linear response, with anisotropic and heterogeneous behavior due to its internal collagen fiber reinforcement. As an accurate enough model we have implemented that employed in Alastrué et al. (2006). The interested reader is

referred to that paper and the references therein for further details on the mechanical response of the cornea.

2.2 Model reduction techniques

2.2.1 Fundamentals: Karhunen-Loève's Decomposition

In this technique it is assumed that the evolution of a certain field $u(\boldsymbol{x},t)$ is known. In practical applications (assume that we have performed some numerical simulations off-line, for instance), this field is expressed in a discrete form which is known at the nodes of a spatial mesh and for some instants of time t^m. Thus, we consider that $u(\boldsymbol{x}_i,t^m) = u^m(\boldsymbol{x}_i) \equiv u_i^m$ ($t^m = m\times\Delta t$) are known. We can also write $\boldsymbol{u}^m$ for the vector containing the nodal degrees of freedom at time t^m. The main idea of the Karhunen-Loève (KL) decomposition is to obtain the most typical or characteristic structure $\phi(\boldsymbol{x})$ among these $u^m(\boldsymbol{x})$, $\forall m$. This is equivalent to obtain a function that maximizes the functional α:

$$\alpha = \frac{\sum_{m=1}^{m=M}\left[\sum_{i=1}^{i=N}\phi(\boldsymbol{x}_i)u^m(\boldsymbol{x}_i)\right]^2}{\sum_{i=1}^{i=N}(\phi(\boldsymbol{x}_i))^2} \tag{2.1}$$

where N represents the number of nodes of the complete model and M the number of computed time steps. The maximization leads to:

$$\sum_{m=1}^{m=M}\left[\left(\sum_{i=1}^{i=N}\tilde{\phi}(\boldsymbol{x}_i)u^m(\boldsymbol{x}_i)\right)\left(\sum_{j=1}^{j=N}\phi(\boldsymbol{x}_j)u^m(\boldsymbol{x}_j)\right)\right] = \alpha\sum_{i=1}^{i=N}\tilde{\phi}(\boldsymbol{x}_i)\phi(\boldsymbol{x}_i);\quad \forall\tilde{\phi} \tag{2.2}$$

which can be rewritten in the form

$$\sum_{i=1}^{i=N}\left\{\sum_{j=1}^{j=N}\left[\sum_{m=1}^{m=M}u^m(\boldsymbol{x}_i)u^m(\boldsymbol{x}_j)\phi(\boldsymbol{x}_j)\right]\tilde{\phi}(\boldsymbol{x}_i)\right\} = \alpha\sum_{i=1}^{i=N}\tilde{\phi}(\boldsymbol{x}_i)\phi(\boldsymbol{x}_i);\ \forall\tilde{\phi} \tag{2.3}$$

Defining the vector $\boldsymbol{\phi}$ such that its i-th component is $\phi(\boldsymbol{x}_i)$, Eq. (2.3) takes the following matrix form

$$\tilde{\boldsymbol{\phi}}^T\boldsymbol{c}\,\boldsymbol{\phi} = \alpha\tilde{\boldsymbol{\phi}}^T\boldsymbol{\phi};\ \forall\tilde{\boldsymbol{\phi}} \Rightarrow \boldsymbol{c}\,\boldsymbol{\phi} = \alpha\boldsymbol{\phi} \tag{2.4}$$

where the two-point correlation matrix is given by

$$c_{ij} = \sum_{m=1}^{m=M}u^m(\boldsymbol{x}_i)u^m(\boldsymbol{x}_j) \Leftrightarrow \boldsymbol{c} = \sum_{m=1}^{m=M}\boldsymbol{u}^m(\boldsymbol{u}^m)^T \tag{2.5}$$

which is symmetric and positive definite. If we define the matrix $\boldsymbol{Q}$ containing the discrete field history:

$$\boldsymbol{Q} = \begin{pmatrix} u_1^1 & u_1^2 & \cdots & u_1^M \\ u_2^1 & u_2^2 & \cdots & u_2^M \\ \vdots & \vdots & \ddots & \vdots \\ u_N^1 & u_N^2 & \cdots & u_N^M \end{pmatrix} \tag{2.6}$$

then it is easy to verify that the matrix $\boldsymbol{c}$ in Eq. (2.4) results in

$$\boldsymbol{c} = \boldsymbol{Q}\,\boldsymbol{Q}^T \tag{2.7}$$

2.2.2 A posteriori reduced modeling

If some direct simulations have been carried out, we can determine $u_i^m, \forall i \in [1, \cdots, N]$ and $\forall m \in [1, \cdots, M]$, and from these solutions the n eigenvectors related to the n-highest eigenvalues that are expected to contain the most important information about the problem solution. For this purpose we solve the eigenvalue problem defined by Eq. (2.4) retaining all the eigenvalues $\boldsymbol{\phi}_k$ belonging to the interval defined by the highest eigenvalue and that value divided by a large enough value (10^8 in our simulations). In practice n is much lower than N, and this constitutes the main advantage of the technique. Thus, we can try to use these n eigenfunctions $\boldsymbol{\phi}_k$ for approximating the solution of a problem *slightly* different to the one that has served to define u_i^m. For this purpose we need to define the matrix $\boldsymbol{A} = [\boldsymbol{\phi}_1 \cdots \boldsymbol{\phi}_n]$

$$\boldsymbol{A} = \begin{pmatrix} \phi_1(\boldsymbol{x}_1) & \phi_2(\boldsymbol{x}_1) & \cdots & \phi_n(\boldsymbol{x}_1) \\ \phi_1(\boldsymbol{x}_2) & \phi_2(\boldsymbol{x}_2) & \cdots & \phi_n(\boldsymbol{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(\boldsymbol{x}_N) & \phi_2(\boldsymbol{x}_N) & \cdots & \phi_n(\boldsymbol{x}_N) \end{pmatrix} \tag{2.8}$$

Now, if we consider the linear system of equations coming from the discretization of a generic problem, in the form:

$$\boldsymbol{K}\,\boldsymbol{u}^m \;=\; \boldsymbol{F}^{m-1} \tag{2.9}$$

where the superscript refers to the time step, then, assuming that the unknown vector contains the nodal degrees of freedom, it can be expressed as:

$$\boldsymbol{u}^m = \sum_{i=1}^{i=n} \zeta_i^m\,\boldsymbol{\phi}_i \;=\; \boldsymbol{A}\,\boldsymbol{\zeta}^m \tag{2.10}$$

from which Eq. (2.9) results

$$\boldsymbol{K}\,\boldsymbol{u}^m \;=\; \boldsymbol{F}^{m-1} \Rightarrow \boldsymbol{K}\,\boldsymbol{A}\,\boldsymbol{\zeta}^m \;=\; \boldsymbol{H}^{m-1} \tag{2.11}$$

and by multiplying both terms by $\boldsymbol{B}^T$ we obtain

$$\boldsymbol{A}^T\boldsymbol{K}\,\boldsymbol{A}\,\boldsymbol{\zeta}^m \;=\; \boldsymbol{A}^T\boldsymbol{F}^{m-1} \tag{2.12}$$

which proves that the final system of equations is of low order, i.e. the dimension of $\boldsymbol{A}^T\boldsymbol{G}\,\boldsymbol{A}$ is $n \times n$, with $n \ll N$.

2.3 Interpolation of reduced models

As it was explained before the reduced model employs the tangent stiffness matrix, $\boldsymbol{K}$, linearized from the non-linear problem formulation at a given time instant. Instead of inverting the full stiffness matrix of size $N \times N$, we employ model reduction techniques to invert the matrix $\boldsymbol{A}^T\boldsymbol{K}\,\boldsymbol{A}$, of size $r \times r$, that is much lower than the original size, as mentioned before.

However, this tangent stiffness matrix $\boldsymbol{K}$ corresponds to a given state of the model (i.e., a given load position and load value, for instance, in mechanical problems). Different load values would lead to different matrices $\boldsymbol{K}$ along the loading path (this is due to the non-linear character of the problem).

It is proposed to perform some direct simulations for a set of different positions of the surgical tool. For each tool position a reduced-order basis set can thus be computed. But the position of the tool at the on-line simulation does not need to coincide with the precomputed ones, and therefore some interpolation scheme is needed if the load is placed at positions different from the originally computed ones. Therefore, we need to interpolate the reduced basis sets, from pre-computed complete models, to an arbitrary position of the load.

How to interpolate reduced models is currently an active area of research, and different methods have been proposed so far. The most prominent two are now reviewed.

2.3.1 Interpolation of the reduced basis: a geometrical approach

Amsallem and Farhat (2008) have pointed out that the set of empirical eigenfunctions given by Eq. (2.4) for a given model forms the so-called Grassman manifold $\mathcal{G}(n, N)$.

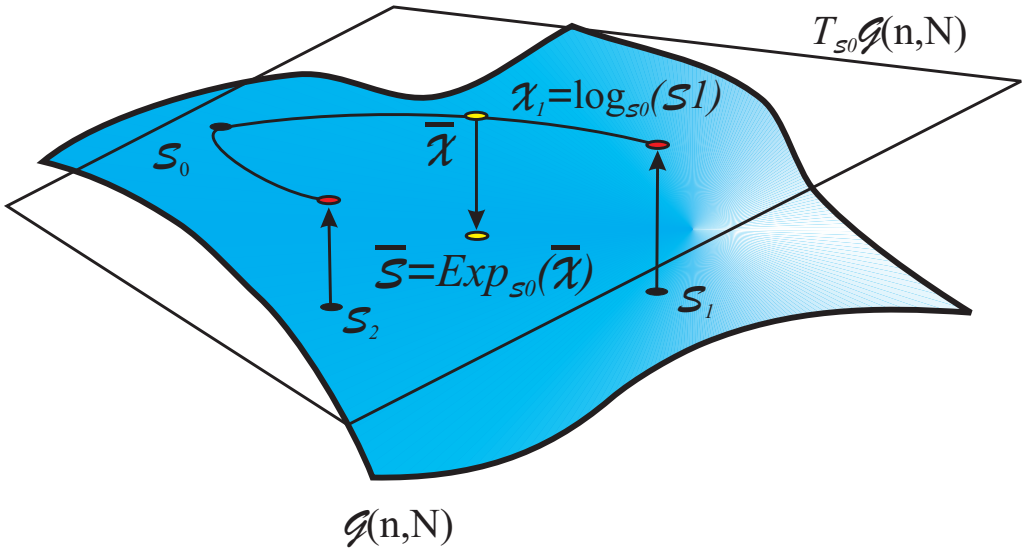


Figure 2.1. Schematic description of the Grassman manifold formed by the set of basis functions of all the reduced modes of a given model.

Therefore, in order to interpolate the set of basis functions, that consequently do not form a vector space, we must move 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. 2.1.

In this way, the columns of $\boldsymbol{A}$ 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 $\boldsymbol{\Gamma} \in \Re^{N\times n}$. The exponential mapping $Exp_{\mathcal{S}}$ transforms χ in an n-dimensional subspace $\mathcal{S}'$ given by a matrix $\boldsymbol{A}' \in \Re^{N\times n}$, such that

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

$$\boldsymbol{A}' = \boldsymbol{A}\boldsymbol{V}\cos\boldsymbol{\Sigma} + \boldsymbol{U}\sin\boldsymbol{\Sigma}$$

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

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

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

So, consider, for instance, $\boldsymbol{A}_0 \in \Re^{N\times n}$ and $\boldsymbol{A}_1 \in \Re^{N\times n}$, 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)

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

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

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(A)` using Matlab, with $\boldsymbol{A}$ a random matrix of 5×5000 elements takes on average much less than $1\ ms$, still compatible with real-time constraints.

2.3.2 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 $\boldsymbol{A} = [\boldsymbol{\phi}_1 \cdots \boldsymbol{\phi}_n]$, 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. In any case, results presented in section 2.5 were obtained with the PODI technique.

2.4 A hyperelastic model for the human cornea

As mentioned before, we have chosen the human cornea as an example of a highly non-linear tissue. This non-linearity comes from a variety of reasons, such as the internal collagen fiber reinforcement (material non-linearity) and also from the very large strains it could suffer. The human cornea is composed of a highly porous material, composed of about 80% water, and thus quasi-incompressible. Most of the cornea's thickness (around 90%) constitutes the stroma, that is composed of 300-500 plies of collagen fibers, distributed in parallel to the surface of the cornea. This microstructure induces in the corneal tissue a highly non-linear and heterogeneous behavior.

The model here employed for the simulation of the human cornea (Alastrué et al., 2006) considers the cornea as a hyperelastic material. the initial, undeformed, configuration of the cornea will be denoted by Ω_0. A continuous movement $\boldsymbol{\chi}$ translates a point $\boldsymbol{X} \in \Omega_0$ to its location at time t, $\boldsymbol{x} \in \Omega_t$. Reinforcing fibers, that move continuously together with the cornea, posses a direction $\boldsymbol{m}_0$, with $|\boldsymbol{m}_0| = 1$. After the deformation, this orientation changes to $\boldsymbol{m}(\boldsymbol{x}, t)$, always with unit modulus. The fiber stretching after the deformation will be given by

$$\lambda \boldsymbol{m}(\boldsymbol{x}, t) = \boldsymbol{F}\boldsymbol{m}_0 \tag{2.13}$$

where $\boldsymbol{F} = d\boldsymbol{x}/d\boldsymbol{X}$ represents the deformation gradient tensor. A second family of fibers, $\boldsymbol{n}_0$, is also considered as reinforcement at each point.

Due to the dependence of strain on the considered direction, the existence of a strain energy density functional, Ψ, depending on the right Cauchy-Green tensor, $\boldsymbol{C} = \boldsymbol{F}^T\boldsymbol{F}$, and the initial fiber orientations, $\boldsymbol{m}_0$ and $\boldsymbol{n}_0$, is postulated. Based on the volumetric incompressibility restrictions, this functional can be expressed as (Alastrué et al., 2006)

$$\Psi(\boldsymbol{C}) = \Psi_{vol}(J) + \bar{\Psi}(\bar{\boldsymbol{C}}, \boldsymbol{m}_0 \otimes \boldsymbol{m}_0, \boldsymbol{n}_0 \otimes \boldsymbol{n}_0) \tag{2.14}$$

where $\Psi_{vol}(J)$ describes the volumetric change and $\bar{\Psi}(\bar{\boldsymbol{C}}, \boldsymbol{m}_0 \otimes \boldsymbol{m}_0, \boldsymbol{n}_0 \otimes \boldsymbol{n}_0)$ the change in shape. Both are scalar functions of $J = det\mathbf{F}$, $\bar{\boldsymbol{C}} = \bar{\boldsymbol{F}}^T\bar{\boldsymbol{F}}$, where $\bar{\boldsymbol{F}} = J^{-1/3}\boldsymbol{F}$, $\boldsymbol{m}_0$ and $\boldsymbol{n}_0$.

Material	C_1	D	C_2	k_1	k_2	k_3	k_4
Cornea	0.1	1e-5	0.0	0.234	29.917	0.234	29.917

Table 2.1. Material properties (MPa).

Once this energy density functional is known, the second Piola-Kirchhoff stress tensor, **S**, and the fourth-order tangent constitutive tensor, **C**, can be determined by

$$\mathbf{S} = 2\frac{\partial \Psi}{\partial \mathbf{C}} \qquad \mathbf{C} = 2\frac{\partial \mathbf{S}(\mathbf{C})}{\partial \mathbf{C}} \tag{2.15}$$

A detailed derivation of the model can be obtained in Alastrué et al. (2006). The interested reader is referred to this paper for reference.

We have considered $\Psi_{vol} = \frac{1}{D}(Ln(J))^2$ to enforce the quasi-incompressible behavior of the cornea through the penalty parameter $1/D$, and to model the corneal tissue we have employed the model proposed by Holzapfel and Gasser (2000), initially developed for arterial tissue:

$$\bar{\Psi} = \frac{C_1}{2}(\overline{I}_1 - 3) + \frac{C_2}{2}(\overline{I}_2 - 3) + \frac{k_1}{2k_2}\{\exp[k_2(\overline{I}_4 - 1)^2] - 1\} + \frac{k_3}{2k_4}\{\exp[k_4(\overline{I}_6 - 1)^2] - 1\} \tag{2.16}$$

Hence, similar approaches can be employed for most soft tissues. Material characteristics are summarized in Table 2.1. $\overline{I}_1, \overline{I}_2, \overline{I}_4 \ldots, \overline{I}_9$ are different invariants of the modified symmetric right Cauchy-Green tensor (Holzapfel and Gasser, 2000). They are defined as:

$$\overline{I}_1 = tr(\overline{\boldsymbol{C}}), \quad \overline{I}_2 = \frac{1}{2}(tr(\overline{\boldsymbol{C}}))^2 - tr(\overline{\boldsymbol{C}}^2) \tag{2.17}$$

$$\overline{I}_4 = \boldsymbol{m}_0 \cdot \overline{\boldsymbol{C}}\boldsymbol{m}_0, \quad \overline{I}_5 = \boldsymbol{m}_0 \cdot \overline{\boldsymbol{C}}^2\boldsymbol{m}_0 \tag{2.18}$$

$$\overline{I}_6 = \boldsymbol{n}_0 \cdot \overline{\boldsymbol{C}}\boldsymbol{n}_0, \quad \overline{I}_7 = \boldsymbol{n}_0 \cdot \overline{\boldsymbol{C}}^2\boldsymbol{n}_0 \tag{2.19}$$

$$\overline{I}_8 = (\boldsymbol{m}_0 \cdot \boldsymbol{n}_0)\boldsymbol{m}_0 \cdot \overline{\boldsymbol{C}}\boldsymbol{n}_0, \quad \overline{I}_9 = (\boldsymbol{m}_0 \cdot \boldsymbol{n}_0)^2 \tag{2.20}$$

For this particular implementation we have chosen, based on Alastrué et al. (2006), $\overline{I}_5 = \overline{I}_7 = 0$.

2.5 Numerical results

In order to test the performance of the proposed technique, attention is focused mainly on two aspects. First, the accuracy of the results. Second, the compliance with the requirements of haptic feedback, i.e., all results must be obtained at a frequency between

300 and 1000 Hz. Aspects related to image rendering, contact detection, tissue cutting, etc, have not been addressed here. In chapter 4 a methodology to deal with tissue cutting will be discussed.

A set of tests have been accomplished, all based on the model of the human cornea presented before. Inertia effects are neglected in this problem, due to the typical slow velocity in the application of the loads in this kind of organs. Thus, we face a parametric problem like the one described in section 2.3.

The cornea was discretized with linear hexahedron finite elements. The mesh consisted of 8514 nodes and 7182 elements. A view of the geometry of the model is shown in Fig. 2.2.

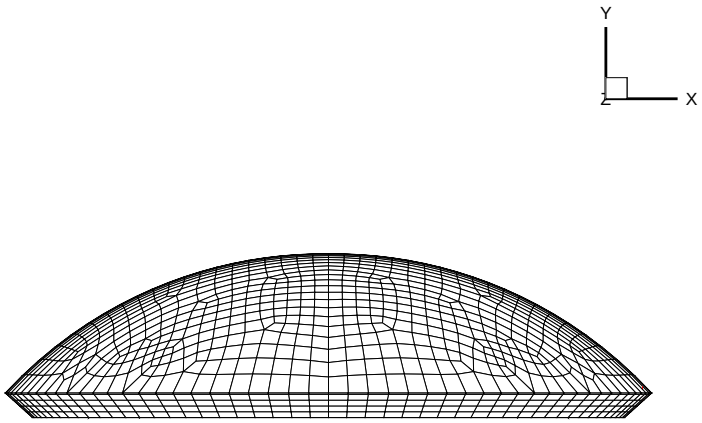
The orientation of the two families of fibers, distributed along the thickness of the cornea, is shown in Fig. 2.3.

2.5.1 Palpation of the cornea

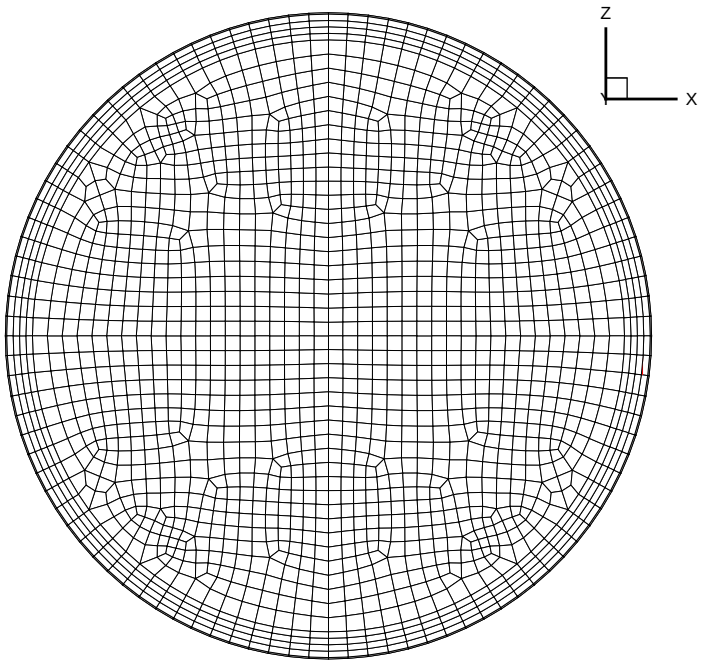
The first test for the proposed technique consists of simulating the palpation of the cornea with a surgical instrument. In order to validate the results, a load was applied to the complete FE model in the central region of the model. The obtained result was compared to the one obtained by employing the model reduction techniques presented before, for the load applied at the same location.

Once the complete model is solved, the most important eigenmodes are extracted from the computed displacement field, together with the initial tangent stiffness matrix. The number of eigenmodes employed in this case was only six, which is, in our experience, the minimum number of modes that should be employed in such a simulation. Other tangent stiffness matrices different to the initial one can also be used, perhaps with more accurate results. The modes are depicted in Fig. 2.4. The associated eigenvalues are, from the biggest to the smallest one, $9.02 \cdot 10^4$, 690, 27, 2.63, 0.221 and 0.0028. As can be noticed, the relative importance of these modes in the overall solution, measured by the associated eigenvalue, decreases very rapidly. Note that the reduced model employed only six degrees of freedom, while the complete model employed 8514 nodes with three degrees of freedom each, thus making 25542 degrees of freedom. The computational savings are obvious: instead of inverting a matrix of 25542×25542, we invert a matrix of only 6×6. Of course, if more accurate solutions are needed, a higher number of modes can be employed.

The displacement field obtained for the complete model is compared to that of the reduced model. Different positions of the load were chosen and the results compared.

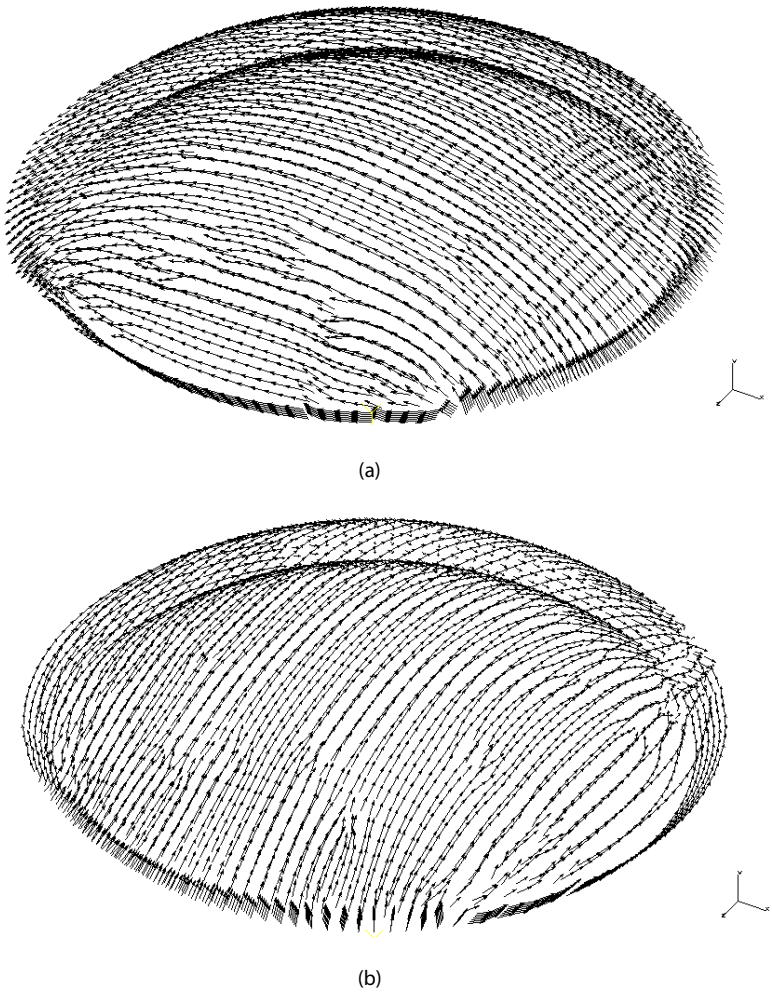


(a)



(b)

Figure 2.2. Geometry of the finite element model for the human cornea.



(a)

(b)

Figure 2.3. Fiber distribution in the finite element model of the cornea. Two families of collagen fibers, roughly perpendicular to each other, are considered.

The application of loads at different locations produced levels of error of similar values as the examples here reported. For a first location of the load, the obtained vertical displacement is shown in Fig. 2.5.

In Fig. 2.6 load-displacement curve has been depicted. In Fig. 2.7, 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. 2.7, 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. 2.5, 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. This error can be attributed to a severe

buckling phenomenon that appears in the complete model. The reduced model is not able, obviously, to capture exactly this behavior. We believe, however, that the present technique can be ameliorated in order to account for buckling phenomena.

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, which is not obviously the best solution for such type of problems. It was chosen as it was easy and fast to implement these early versions of the method. Lim and De (2007) reported very recently similar levels of error (even more, up to 30%), but for linear elastic materials undergoing large displacements (thus, only geometrically non-linear problems). They employed a meshless method called the method of Finite Spheres.

The simulations ran at 472-483Hz, which is among the limits imposed by haptic feedback realism, as mentioned before. Of course, the use of more sophisticated codes, for instance using parallel programming, could give even faster results.

2.5.2 Force prediction

The architecture of a real-time simulator requires, however, the prediction of the response force to a given displacement imposed to the model by means of the haptic device. To analyze the behavior of the proposed technique under these requirements, we studied precisely a prescribed displacement problem for the case analyzed in the previous section.

A vertical displacement was imposed to node 4144, located almost in the center of the cornea, with linearly increasing value. The response force exerted by the cornea on the tool tip was simulated with the aid of the complete model as well as the reduced one. While the complete model took around 3 hours to solve the problem, due to the large displacement imposed at the last steps of the simulation, the reduced model still runs at between 400-500 Hz. The results are summarized in Table 2.2.

As can be noticed, the predicted response is very accurate at the middle of the simulation, and gives some error both at the very beginning of the simulation and for very large strains.

u	$F_{ROM}(N)$	$F_{FEM}(N)$	$Error(\%)$
0.1	0.0045	0.0055	18
0.2	0.0091	0.0108	16
0.3	0.0136	0.0158	14
0.45	0.0204	0.0227	10
0.675	0.0307	0.0321	4
0.9	0.041	0.0405	1
1.125	0.0511	0.0482	6
1.35	0.0614	0.0555	10
1.575	0.0716	0.0628	14
1.8	0.0818	0.0702	16
2.025	0.092	0.0779	18
2.135	0.097	0.0818	19

Table 2.2. Error in the predicted force response on the tool. Reduced Order Modeling (ROM) vs. Finite Element Modeling (FEM).

2.5.3 Force located at an arbitrary point

As mentioned before, the strategy here presented includes the off-line calculation of the response of the organ to prescribed loads. Thus, a sampling strategy must be adopted to construct a basis capable of representing the overall response of the organ to virtually any load (although it is expected that a good surgeon will not make unexpected movements away from "good practice" rules in the surgery). 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. To this end, we employ the PODI approach mentioned before (Ly and Tran, 2005), by placing loads in between the loaded nodes in the complete model, in order to seek for the worst case scenario for the method.

To study this behavior, a simple test was implemented. The response of the cornea to two different loads located at different locations was computed by means of complete FE model. Then, a new load at a different position was applied, in between the two and computed the response of the system by means of PODI techniques. We assume that the three loads increase their value from zero linearly with time. The complete model for this third load was also solved in order to test the accuracy of the method. As before, this complete model is assumed as "exact" in the absence of any further knowledge on the behavior of this cornea.

The obtained results for a very large strain are shown in Fig. 2.8. The results for the PODI model are, as can be noticed, in good agreement with the results of the complete model, see Fig. 2.8(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). We have also computed the error in $||\cdot||_2$ norm, defined as:

$$||e||_2 = \frac{1}{n}\sqrt{\sum_{I=1}^{n} e_I^2} \tag{2.21}$$

where e_I represents the nodal error and n the number of nodes in the model. This 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). Of course, larger values of error can be obtained if larger strains are imposed, but they remain bounded if "expected" values are given to the loads. For situations that will not likely occur, out of surgeons' good practice, the error can of course be very large. In any case, the reduced basis of the model could be completed with more basis modes corresponding to non-expectable behavior of the surgeons.

2.6 Discussion

In this chapter a novel strategy is presented for real-time interactive simulation of non-linear anisotropic tissues. The presented technique is based on model reduction techniques and, unlike previous works (Barbi**c** and James, 2005), it allows for the consideration of both geometrical and material non-linearities.

The reduced models are constructed by employing a set of "high quality" global basis functions (as opposed to general-purpose, locally supported FE shape functions) in a Galerkin framework. These functions are constructed after some direct simulations of the organs performed by standard FE or BE techniques, for instance. These simulations are made off-line and the computed displacements and tangent stiffness matrices are stored in memory prior to beginning with the real-time simulation.

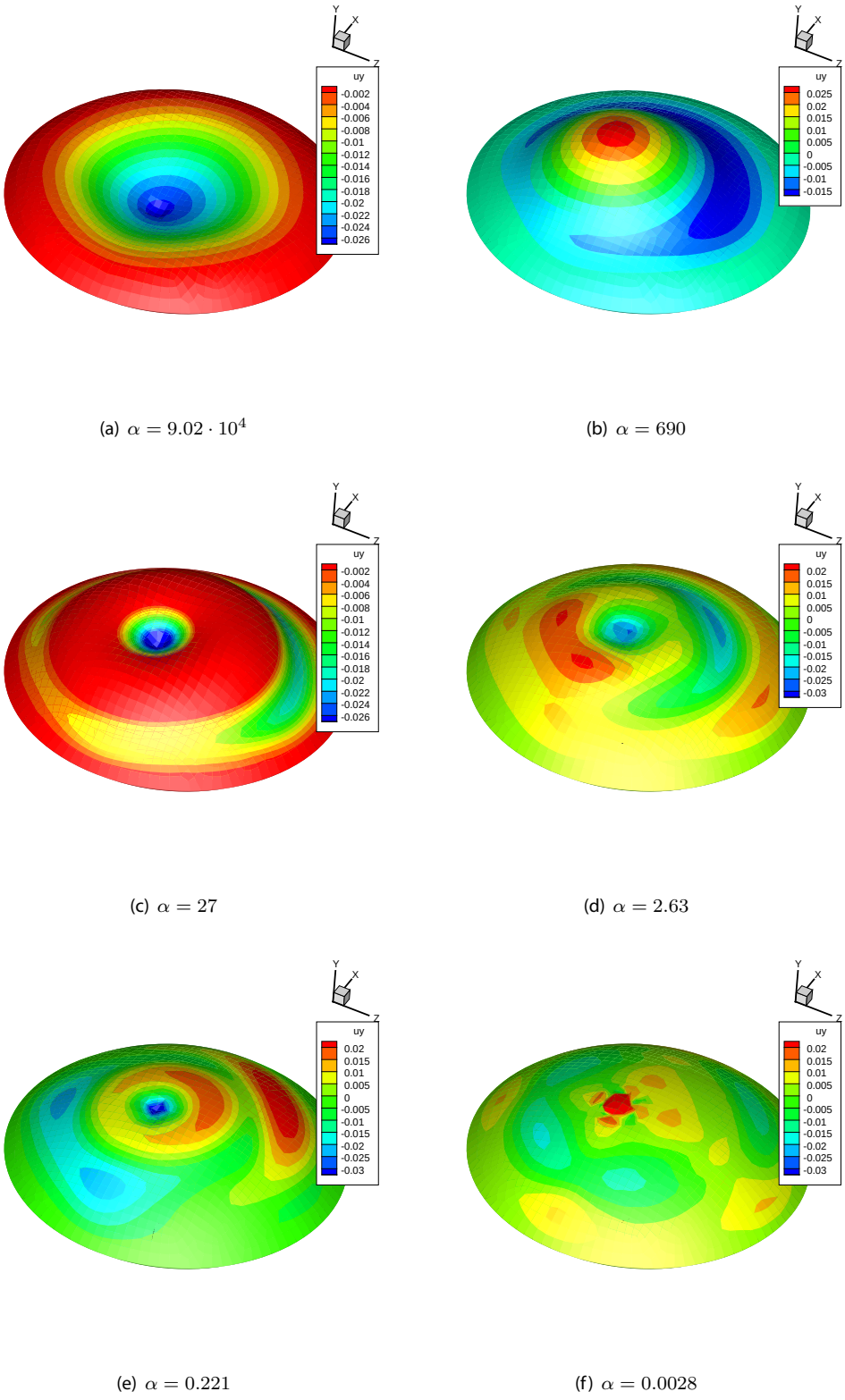
Results obtained showed good accordance with complete model results, and ran at frequencies of around 400-500 Hz, enough for real-time requirements, even for this very rude code prototypes. In sum, the technique presented constitutes in our modest opinion an alternative to standard FE simulation techniques for real-time applications involving non-linear and anisotropic materials. In the following we discuss some imitations of the applied method.

2.6.1 Limitations of standard model reduction techniques

Classical model reduction techniques, when applied to the simulation of non-linear solids and structures imply the need for frequent updating of the stiffness tangent matrix $\boldsymbol{K}$ (or, equivalently, the reduced stiffness matrix $\boldsymbol{A}^T \boldsymbol{K}\ \boldsymbol{A}$, but the matrix $\boldsymbol{A}$ is assumed constant throughout the simulation). Otherwise, the reduced model will obviously be linear as shown seen in Fig. 2.6. This implies that once the residual of the discrete, algebraic form of the problems is estimated unacceptable, the only way of searching the equilibrium is by updating the stiffness matrix of the complete problem. Modified Newton-Raphson methods could equally be employed, but in the context of reduced models their convergence is often judged too slow. Even if this strategy can be employed for reducing non-linear models the computing cost is considerable, and of course this technique is unfeasible if one is looking for quasi-real-time simulations.

A different question arises in the so-called hyper-reduction methods Ryckelynck (2008) in which only a few rows of the stiffness matrix are integrated, based on the fact of the limited number of degrees of freedom that POD techniques select for a particular problem. The problem is precisely to select appropriately in which part of the model the integration is performed.

This stiffness matrix updating is usually an expensive procedure in terms of computational cost (and noteworthy in the context of reduced models). But sometimes it is simply unaffordable. This is the case in the framework of real-time or near real-time simulations, where a frequency of 500-1000Hz in the response of the simulation is needed, for instance, for haptic realism (30 Hz would be enough for video feedback, which is too much a limitation indeed). In the following chapter a technique will be presented that could take into account the non-linearity of the problem without resorting to tangent stiffness matrix updating in the complete model.



(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 2.4. The eigenmodes of the problem employed as global basis for the reduced model simulation.

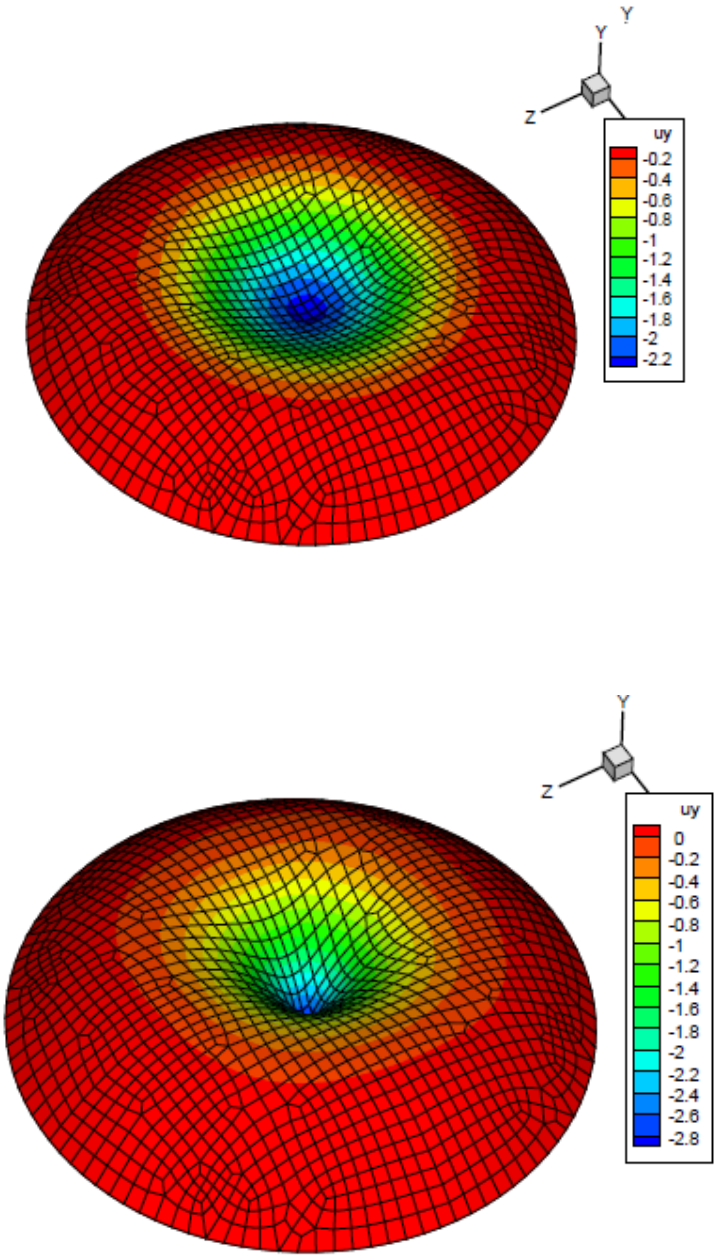


Figure 2.5. Vertical displacement field for a first position of the load. Complete model (up) vs. reduced model (down).

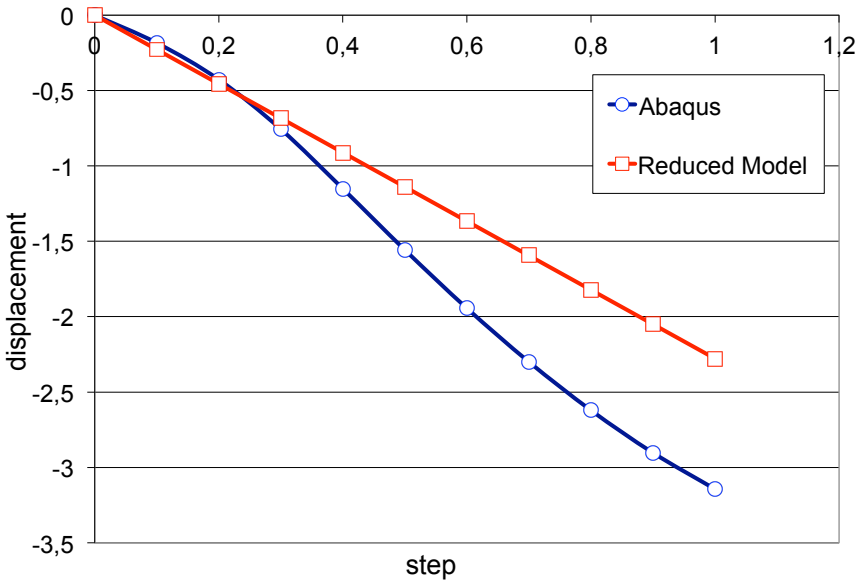


Figure 2.6. load-displacement curve at the vertex of the cornea.

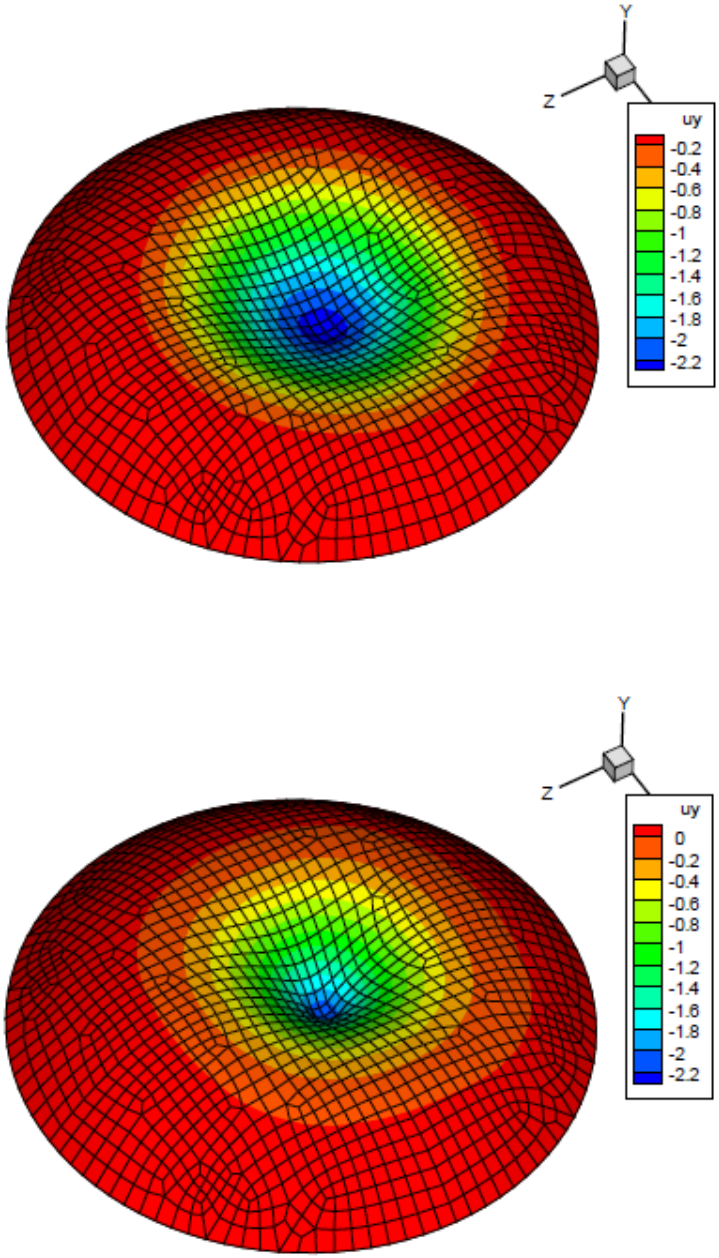
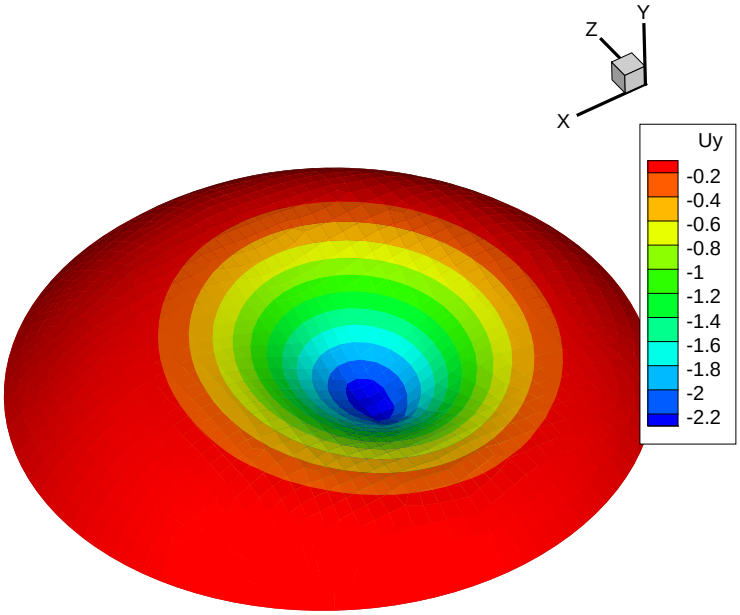
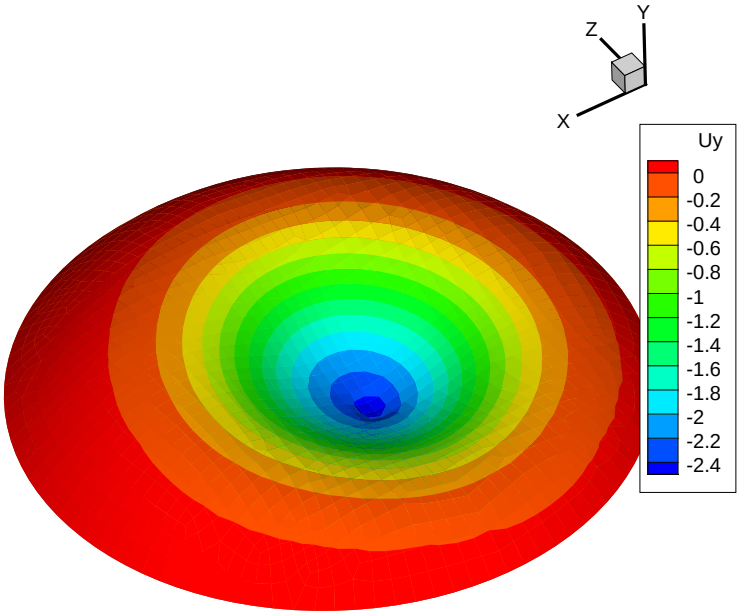


Figure 2.7. Vertical displacement field for a second position of the load. Complete model (up) vs. reduced model (down).



(a)



(b)

Figure 2.8. Vertical displacement (mm) for (a) the reduced model and (b) the complete model of a load in a position not previously computed.

Chapter 3

Real-time Simulation of Non-linear Hyperelastic Materials through Model Order Reduction

3.1 Introduction

In the preceding chapter model order reduction techniques, and particularly proper orthogonal decomposition methods, have been applied to the simulation of highly non-linear, state of the art, constitutive models for soft tissue. It has been demonstrated that POD techniques constitute an appealing means for real-time simulation of such models. However, some important limitations have been also pointed out.

Although the levels of error found in these applications can be considered high (around 20 % for some applications), it is within the best techniques published so far in the field, see for instance Lim and De (2007). The main limitation lies, therefore, in the practical impediment for tangent stiffness matrix updating due to the severe restriction posed by real time requirements.

Tangent stiffness matrix updating constitutes an essential procedure in Newton-Raphson method for the iterative solution of non-linear systems of equations. Modified Newton methods, for instance, that do not need for such an updating procedure, are unable, however, to follow buckling or post-buckling phenomena, essential in some cases, such as the before presented simulation of palpation of the human cornea.

Otherwise, the resulting simulation will be actually linear, even if it employs the best (in statistical terms) basis functions available. In the work by Ryckelynck (2005), on the contrary, Krylov subspaces are added to the reduced set of basis functions once the simulation provides an unsatisfactory norm of the residual. But this implies to come back to the last converged time step and to continue with the just enriched basis.

The approach followed in this chapter is somewhat different. In order to avoid a frequent update of the stiffness matrix, an asymptotic expansion of some variables of interest is made at the last converged time step. When this expansion is introduced into the discrete weak form of the problem, a series of problems is encountered that share the same stiffness matrix for all the terms of the expansion. Of course, the expansion has a finite radius of convergence, but it is frequently enough to perform many of the simulations considered here, as will be demonstrated.

This technique thus combines two main ingredients: the Proper Orthogonal Decomposition of the existing data coming from complete models, and an asymptotic expansion of variables of interest in the neighborhood of equilibrium points (Cochelin et al., 1994a) (Abichou et al., 2002). This technique originated n the works of Yvonnet et al. (2007) for buckling analysis of foam structures, considered linear elastic (geometrically non-linear), although it is generalized here for more complex, hyperelastic, constitutive equations including material non-linearities. This extension is not straightforward and, up to the best of our knowledge, has not been developed so far.

The structure of the chapter is as follows. In section 2 the basics of the technique proposed in Yvonnet et al. (2007) for linear elastic solids undergoing finite strains (Saint Venant-Kirchhoff models) is reviewed. In section 3 this technique is extended to account for material as well as geometrical non-linearities. The proposed technique is described in detail and, finally, in section 4 some examples of application for Saint Venant-Kirchhoff and neo-hookean models are included that show the performance of the technique and compares the computational cost with a standard FE solution.

3.2 Asymptotic Numerical Methods

Non-linear structural problems are generally solved using iterative methods such as Newton-Raphson or modified Newton schemes. Such algorithms have been successful for solving the non-linear equilibrium equations of the model. However, the computing time is usually large as compared to a linear solution because of the number of

iterations usually needed and the computation of tangent stiffness matrices. A family of asymptotic numerical methods (ANM) based on perturbation techniques and finite element methods have been proposed and intensively used by Potier-Ferry and co-workers for computing perturbed bifurcations, and applied in computing the post-buckling behavior of elastic plates and shells. Next they have extended the method to many non-linear elastic solutions, plastic deformations, etc. For a complete review the interested reader can refer to Cochelin et al. (1994a) or Abichou et al. (2002). In contrast to predictor-corrector algorithms, the non-linear equilibrium paths are determined by means of asymptotic expansions: the unknown nodal vector $\boldsymbol{U}$ and the load parameter $\overline{\lambda}$ —that represents a pseudo-time, the overline is used to avoid confusion with the Lamé coefficient—are represented by power series expansions with respect to a control parameter a. By introducing the expansions into the equilibrium equation, the non-linear problem is transformed into a sequence of linear problems in a recurrent manner and are solved by the finite element method, for instance. It is noteworthy that as all the linear problems have the same tangent stiffness matrix the method requires only one matrix inversion. Moreover one gets a continuous analytic representation of the load-displacement (response) curve which differs from the point by point representation of standard algorithms.

It is precisely this feature (the existence of only one tangent operator per equilibrium point) that makes the ANM an appealing choice to combine with model reduction techniques in order to get a very efficient technique.

3.2.1 Problem formulation for Saint Venant-Kirchhoff models

Here we review the development made in Yvonnet et al. (2007) for the buckling analysis of foam structures. We will also consider it in our numerical examples, where they will be compared with non-linear constitutive models. We use a Lagrangian description of the movement whose material coordinates are given by the vector $\boldsymbol{X}$. The solid occupies a volume Ω_0 whose boundary is denoted by Γ. Essential and natural boundary conditions are applied to the non-overlapping portions of the boundary denoted by Γ_u and Γ_t. The deformed configuration of each point is given by the vector $\boldsymbol{x}$, such that

$$\boldsymbol{x} = \boldsymbol{X} + \boldsymbol{u}, \tag{3.1}$$

where the displacement field is denoted by $\boldsymbol{u}$.

Following Cochelin et al. (1994b), we consider a linear and a non-linear term for the

Green-Lagrange strain tensor, $\boldsymbol{E}$, in the form

$$\boldsymbol{E} = \frac{1}{2}(\boldsymbol{F}^T\boldsymbol{F} - 1) = \gamma_l(\boldsymbol{u}) + \gamma_{nl}(\boldsymbol{u}, \boldsymbol{u}) \tag{3.2}$$

where $\boldsymbol{F} = \boldsymbol{\nabla u} + \boldsymbol{I}$ is the deformation gradient tensor and using the notation in Cochelin et al. (1994b),

$$\gamma_l(\boldsymbol{u}) = \frac{1}{2}(\nabla(\boldsymbol{u}^T) + \nabla(\boldsymbol{u})), \tag{3.3a}$$

$$\gamma_{nl}(\boldsymbol{u}, \boldsymbol{u}) = \frac{1}{2}\nabla(\boldsymbol{u}^T)\nabla(\boldsymbol{u}). \tag{3.3b}$$

The equilibrium equation stated in the reference configuration looks like

$$\boldsymbol{\nabla P} + \boldsymbol{B} = \boldsymbol{0} \quad \text{in } \Omega_0 \tag{3.4}$$

in which $\boldsymbol{B}$ is the body force and $\boldsymbol{P}$ is the first Piola-Kirchhoff stress tensor. The boundary conditions of the body are defined by

$$\begin{aligned} \boldsymbol{u}(\boldsymbol{X}) &= \bar{\boldsymbol{u}} \ \text{ on } \Gamma_u, \\ \boldsymbol{PN} &= \overline{\lambda}\bar{\boldsymbol{t}} \ \text{ on } \Gamma_t \end{aligned} \tag{3.5}$$

where $\boldsymbol{N}$ is the unit vector normal to Γ, $\bar{\boldsymbol{t}}$ is an applied traction and $\overline{\lambda}$ is a loading parameter, equivalent to a pseudo-time, and ranging from 0 to 1. The weak form of the problem is then given by

$$\int_{\Omega_0} \boldsymbol{S} : \delta\boldsymbol{E}\, d\Omega = \overline{\lambda} \int_{\Gamma_t} \bar{\boldsymbol{t}} \cdot \delta\boldsymbol{u} d\Gamma \ \ \forall \delta\boldsymbol{u} \in H^1(\Omega) \tag{3.6}$$

where in the above equation $\delta\boldsymbol{E}$ is expressed by

$$\delta\boldsymbol{E} = \frac{1}{2}\left[\boldsymbol{F}^T\nabla(\boldsymbol{\delta u}) + \nabla(\boldsymbol{\delta u})^T\boldsymbol{F}\right] = \gamma_l(\boldsymbol{\delta u}) + \gamma_{nl_S}(\boldsymbol{u}, \boldsymbol{\delta u}), \tag{3.7}$$

where, in turn, $\gamma_{nl_S}(\boldsymbol{u}, \delta\boldsymbol{u})$ is defined by

$$\gamma_{nl_S}(\boldsymbol{u}, \delta\boldsymbol{u}) = \gamma_{nl}(\boldsymbol{u}, \delta\boldsymbol{u}) + \gamma_{nl}(\delta\boldsymbol{u}, \boldsymbol{u}). \tag{3.8}$$

The technique is best viewed by considering the Saint Venant-Kirchhoff's model, as done in Cochelin et al. (1994b) and Yvonnet et al. (2007), since only geometric nonlinearities are present. Later on, in section 3.3, we will generalize the method for other hyperelastic models. The Saint Venant-Kirchhoff model is characterized by the energy function given by

$$\Psi = \frac{\lambda}{2}(\text{tr}(\boldsymbol{E}))^2 + \mu\boldsymbol{E} : \boldsymbol{E} \tag{3.9}$$

where λ and μ are Lamé's coefficients. The second Piola-Kirchhoff stress tensor can be obtained by

$$\boldsymbol{S} = \frac{\partial \Psi(\boldsymbol{E})}{\partial \boldsymbol{E}} = \mathbf{C} : \boldsymbol{E} \tag{3.10}$$

in which **C** is the fourth-order constitutive (elastic) tensor.

The Saint Venant-Kirchhoff's model possesses well-known limitations, particularly some instabilities when subjected to pure compression. Nevertheless, it remains to be interesting for some applications. It is noteworthy that in real-time simulation environments (see Barbi**c** and James (2005) and references therein, for instance) it is among the state-of-the-art models that can be computed under the severe limitations that real-time frameworks impose (30Hz for video feedback and 500Hz for haptic feedback).

3.2.2 Asymptotic numerical method for geometrically non-linear problems

In the ANM (Cochelin et al., 1994a), (Abichou et al., 2002) the displacement of each material point is expanded asymptotically in terms of a control parameter "a". This expansion is developed in the neighborhood of a known equilibrium point $(\boldsymbol{u}^n; \boldsymbol{S}^n; \overline{\lambda}^n)$ at step n and the series is truncated at order N. To simplify the resulting expressions, also the second Piola-Kirchhoff stress tensor and the load parameter $\overline{\lambda}$ are expanded in series prior to their introduction in the weak form of the problem:

$$\left\{ \begin{array}{c} \boldsymbol{u}^{n+1}(a) \\ \boldsymbol{S}^{n+1}(a) \\ \overline{\lambda}^{n+1}(a) \end{array} \right\} = \left\{ \begin{array}{c} \boldsymbol{u}^{n}(a) \\ \boldsymbol{S}^{n}(a) \\ \overline{\lambda}^{n}(a) \end{array} \right\} + \sum_{p=1}^{N} a^p \left\{ \begin{array}{c} \boldsymbol{u}_p \\ \boldsymbol{S}_p \\ \overline{\lambda}_p \end{array} \right\}, \tag{3.11}$$

where $(\boldsymbol{u}_p, \boldsymbol{S}_p, \overline{\lambda}_p)$ are unknown and $(\boldsymbol{u}^{n+1}(a), \boldsymbol{S}^{n+1}(a), \overline{\lambda}^{n+1}(a))$ represents the solution along a portion of the loading curve. It is noteworthy that the behavior of the solid is described continuously with respect to "a". The introduction of Eq. (3.11) into Eq. (3.6) and Eq. (3.10) leads to a series of linear problems with the same tangent operator, thus avoiding the burden associated with stiffness matrix updating in the Newton-Raphson scheme.

A general procedure for constructing quadratic forms of equations consists in developing auxiliary variables into finite series, as will be detailed in section 3.3 for neo-Hookean materials. The series expansion of $\delta \boldsymbol{E}(\boldsymbol{u})$ gives

$$\delta \boldsymbol{E}^{n+1}(a) = \gamma_l(\boldsymbol{\delta u}) + \gamma_{nl_S}(\boldsymbol{\delta u}, \boldsymbol{u^n}) + \sum_{p=1}^{n} a^p \gamma_{nl_S}(\boldsymbol{\delta u}, \boldsymbol{u}_p). \tag{3.12}$$

The series expansions of $\boldsymbol{S}$ gives in turn

$$\boldsymbol{S}^{n+1}(a) = \mathbf{C} : \boldsymbol{E}^{n+1}(a) = \\ \mathbf{C} : \left[\gamma_{nl}(\boldsymbol{u}^n, \boldsymbol{u}^n) + \gamma_l(\boldsymbol{u}^n) + \sum_{p=1}^{n} a^p \left(\gamma_l(\boldsymbol{u}_p) + \gamma_{nl_S}(\boldsymbol{u}^n, \boldsymbol{u}_p) + \sum_{i=1}^{p-1} \gamma_{nl}(\boldsymbol{u}_i, \boldsymbol{u}_{p-i}) \right) \right], \tag{3.13}$$

and at order p we obtain

$$\boldsymbol{S}_p = \mathbf{C} : \left\{ \gamma_l(\boldsymbol{u}_p) + \gamma_{nl_S}(\boldsymbol{u}^n, \boldsymbol{u}_p) + \sum_{i=1}^{p-1} \gamma_{nl}(\boldsymbol{u}_i, \boldsymbol{u}_{p-i}) \right\} \tag{3.14}$$

Introducing the asymptotic expansion into Eq. (3.10) results in

$$\int_{\Omega_0} \left\{ \left(\boldsymbol{S}^n + \sum_{p=1}^{N} a^p \boldsymbol{S}_p \right) : \left(\gamma_l(\delta\boldsymbol{u}) + \gamma_{nl_S}(\boldsymbol{u}^n, \delta\boldsymbol{u}) + \sum_{p=1}^{N} a^p \gamma_{nl_S}(\boldsymbol{u}_p, \delta\boldsymbol{u} \right) \right\} d\Omega = \\ \left(\overline{\lambda}^n + \sum_{p=1}^{N} a^p \overline{\lambda}_p \right) \Psi_{ext}(\delta\boldsymbol{u}), \tag{3.15}$$

with $\boldsymbol{\Psi}_{ext}(\delta\boldsymbol{u}) = \int_{\Gamma_t} \boldsymbol{t} \cdot \delta\boldsymbol{u} d\Gamma$. Introducing Eq. (3.14) into Eq. (3.15) and identifying terms with the same power of a results in a successive series of linear problems which at order p, $(p = 1, \ldots, N)$ takes the form

$$\mathcal{L}(\boldsymbol{\delta u}, \boldsymbol{u^n}) = \overline{\lambda}_p \boldsymbol{\Psi}_{ext}(\delta\boldsymbol{u}) + F_p^{nl}(\delta\boldsymbol{u}) \tag{3.16}$$

with

$$\mathcal{L}(\boldsymbol{\delta u}, \boldsymbol{u^n}) = \int_\Omega \{ \boldsymbol{S}^n : \gamma_{nl_S}(\boldsymbol{u}^n, \delta\boldsymbol{u}) \\ + [\gamma_l(\delta\boldsymbol{u}) + \gamma_{nl_S}(\boldsymbol{u}_p, \delta\boldsymbol{u})] : \mathbf{C} : [\gamma_l(\boldsymbol{u}_p) + \gamma_{nl_S}(\boldsymbol{u}^n, \boldsymbol{u}_p)] \} d\Omega \tag{3.17}$$

and where $\boldsymbol{F}_p^{nl}(\delta\boldsymbol{u})$ is equal to zero at order one and at order p it can be calculated as

$$F_p^{nl}(\delta\boldsymbol{u}) = -\int_\Omega \{ \sum_{i=1}^{p-1} \boldsymbol{S}_i : \gamma_{nl_S}(\boldsymbol{u}_{p-i}, \delta\boldsymbol{u}) \\ + \sum_{i=1}^{p-1} [\gamma_{nl}(\boldsymbol{u}_i, \boldsymbol{u}_{p-i})] : \mathbf{C} : [\gamma_l(\delta\boldsymbol{u}) + \gamma_{nl_S}(\boldsymbol{u}^n, \delta\boldsymbol{u})] \} d\Omega \tag{3.18}$$

Discretization of Eq. (3.16) by using finite elements leads to a sequence of linear problems in the form (Cochelin et al., 1994b)

$$\text{Order 1} \begin{cases} \boldsymbol{K}_t \boldsymbol{u}_1 = \overline{\lambda}_1 \boldsymbol{f} \\ \boldsymbol{u}_1^T \boldsymbol{u}_1 + \overline{\lambda}_1^2 = 1 \end{cases} \tag{3.19}$$

$$\text{Order } p \begin{cases} \boldsymbol{K}_t \boldsymbol{u}_p = \overline{\lambda}_p \boldsymbol{f} + \boldsymbol{f}_p^{nl}(\boldsymbol{u}_i) \quad i < p \\ \boldsymbol{u}_p^T \boldsymbol{u}_1 + \overline{\lambda}_p \lambda_1 = 0 \end{cases} \tag{3.20}$$

where $\boldsymbol{K}_t$ denotes the tangent stiffness matrix associated with Eq. (3.17), common to the problems at different orders p. It is the same as the one applied in a classical iterative algorithm like Newton-Raphson (in the first iteration). In the above, $\boldsymbol{u}_p$ is the discretized form of the displacement field at order p, $\boldsymbol{f}$ is the loading vector and $\boldsymbol{f}_p^{nl}$ represents the discretized form associated with $\boldsymbol{F}_p^{nl}(\delta \boldsymbol{u})$ in Eq. (3.18), which at order p only depends on the values of $\boldsymbol{u}_i, i < p$. The solution of these problems can be obtained as follows

$$\text{at order 1} \begin{cases} \hat{\boldsymbol{u}} = \{\boldsymbol{K}_t\}^{-1} \boldsymbol{f} \\ \overline{\lambda}_1 = \frac{1}{\sqrt{\hat{\boldsymbol{u}}^T \hat{\boldsymbol{u}} + 1}} \\ \boldsymbol{u}_1 = \overline{\lambda}_1 \hat{\boldsymbol{u}} \end{cases} \tag{3.21}$$

$$\text{at order } p \begin{cases} \boldsymbol{u}_p^{nl} = \{\boldsymbol{K}_t\}^{-1} \boldsymbol{f}_p^{nl} \\ \overline{\lambda}_p = -\overline{\lambda}_1 \{\boldsymbol{u}_p^{nl}\}^T \boldsymbol{u}_1 \\ \boldsymbol{u}_p = \frac{\overline{\lambda}_p}{\overline{\lambda}_1} \boldsymbol{u}_1 + \boldsymbol{u}_p^{nl} \end{cases} \tag{3.22}$$

3.2.3 Combined POD-ANM method

As explained before the use of the reduced order basis within the standard strategy, without updating the stiffness matrix, leads to very fast calculation of the system of equations but at the cost of some error in the results. On the other hand by using ANM we can obtain the solution of non-linear problems accurately in a neighborhood of an equilibrium point, and without the necessity of iterative procedures. In the next section this procedure is extended for solving non linear constitutive models. Here we assume that the POD basis has been calculated as explained in chapter 2. The terms of the asymptotic expansions associated with the displacements are in turn expressed as functions of POD basis as

$$\boldsymbol{u}_p = \sum_{m=1}^{M} \boldsymbol{\phi}^m \zeta_p^m = \boldsymbol{A} \boldsymbol{\zeta}_p, \tag{3.23}$$

where $\boldsymbol{\zeta}_p$ are unknowns. So the new asymptotic expansion of $\boldsymbol{u}$ is expressed by

$$\boldsymbol{u}^{n+1}(a) = \boldsymbol{A} \left(\boldsymbol{\zeta}^n + \sum_{p=1}^{N} a^p \boldsymbol{\zeta}_p \right) \tag{3.24}$$

where $(\boldsymbol{\zeta^n}, \overline{\lambda}^n)$ represents the previous converged solution. Introducing Eq. (3.24) into Eq. (3.16) results in

$$\mathcal{L}(\boldsymbol{A}\boldsymbol{\zeta}_p, \boldsymbol{A}\delta\boldsymbol{\zeta}) = \overline{\lambda}_p \delta \boldsymbol{\Psi}_{ext}(\boldsymbol{A}\delta\boldsymbol{\zeta}) + F_p^{nl}(\boldsymbol{A}\delta\boldsymbol{\zeta}). \tag{3.25}$$

After finite element discretization, and making use of the fact that the POD basis is orthonormal, we have the following sequence of linear systems of equations

$$\text{At order 1} \begin{cases} \boldsymbol{A}^T \boldsymbol{K}_t \boldsymbol{A} \boldsymbol{\zeta}_1 = \overline{\lambda}_1 \boldsymbol{A}^T \boldsymbol{f} \\ \boldsymbol{\zeta}_1^T \boldsymbol{\zeta}_1 + \overline{\lambda}_1^2 = 1 \end{cases} \tag{3.26}$$

$$\text{At order } p \begin{cases} \boldsymbol{A}^T \boldsymbol{K}_t \boldsymbol{A} \boldsymbol{\zeta}_p = \overline{\lambda}_p \boldsymbol{A}^T \boldsymbol{f} + \boldsymbol{f}_p^{nl} \\ \boldsymbol{\zeta}_p^T \boldsymbol{\zeta}_1 + \overline{\lambda}_p \overline{\lambda}_1 = 0 \end{cases} \tag{3.27}$$

The size of the above equations depend on the number of the POD basis functions, but it is very low, as explained before (typically less than ten), so they can be computed in a very short time.

The solution of these equations can be obtained as follows:

$$\text{At order 1} \begin{cases} \hat{\boldsymbol{\zeta}} = \{\boldsymbol{A}^T \boldsymbol{K}_t \boldsymbol{A}\}^{-1} \boldsymbol{A}^T \boldsymbol{f} \\ \overline{\lambda}_1 = \frac{1}{\sqrt{\hat{\boldsymbol{\zeta}}^T \hat{\boldsymbol{\zeta}} + 1}} \\ \boldsymbol{\zeta}_1 = \overline{\lambda}_1 \hat{\boldsymbol{\zeta}} \\ \hat{\boldsymbol{u}} = \boldsymbol{A} \hat{\boldsymbol{\zeta}}, \boldsymbol{u}_1 = \overline{\lambda}_1 \boldsymbol{\zeta}_1 \end{cases} \tag{3.28}$$

$$\text{At order } p \begin{cases} \boldsymbol{\zeta}_p^{nl} = \{\boldsymbol{A}^T \boldsymbol{K}_t \boldsymbol{A}\}^{-1} \boldsymbol{A}^T \boldsymbol{f}_p^{nl} \\ \overline{\lambda}_p = -\overline{\lambda}_1 \{\boldsymbol{\zeta}_p^{nl}\}^T \boldsymbol{\zeta}_1 \\ \boldsymbol{\zeta}_p = \frac{\overline{\lambda}_p}{\overline{\lambda}_1} \boldsymbol{\zeta}_1 + \boldsymbol{\zeta}_p^{nl} \\ \boldsymbol{u}_p = \boldsymbol{A} \boldsymbol{\zeta}_p \end{cases} \tag{3.29}$$

Note that, as mentioned before, in Eqs. (3.28)-(3.29) there exists only one tangent stiffness matrix per time step, shared by all the terms for any order p. The advantage, of course, comes from the fact that the radius of convergence of the method, as will be seen, and consequently the size of time steps, is considerably bigger than the time step in traditional Newton procedures.

3.3 Non-linear model reduction of hyperelastic materials

The extension of the technique explained before to other hyperelastic materials, in which non-linearities other than geometrical ones are present, is not straightforward and, up to our knowledge, it has not been done before. In this section we extend the technique to neo-hookean materials (Bonet and Wood, 2008). Extension to other hyperelastic materials would follow the same guidelines, the major difference with the

Saint Venant-Kirchhoff model is the presence of material non-linearities, in addition to the geometrical ones.

The compressible Neo-Hookean model is characterized by a strain energy function given by

$$\Psi = \frac{\mu}{2}(\text{tr}(\boldsymbol{C}) - 3) - \mu \ln J + \frac{\lambda}{2}(\ln J)^2 \tag{3.30}$$

where λ and μ are Lame's constants and $\boldsymbol{C} = \boldsymbol{I} + 2\boldsymbol{E}$ is the right Cauchy-Green strain tensor. The second Piola-Kirchhoff stress tensor can be obtained by

$$\boldsymbol{S} = \frac{\partial \Psi(\boldsymbol{E})}{\partial \boldsymbol{E}} = \mu(\boldsymbol{I} - \boldsymbol{C}^{-1}) + \lambda(\ln J)\boldsymbol{C}^{-1}. \tag{3.31}$$

In this case, an expansion similar to that in Eq. (3.11) is done, but in this case the intricate expansion procedure becomes more clear if we identify, as in Cao and Potier-Ferry (1999), the asymptotic expansion with a Taylor series of the variables of interest, denoted by $\boldsymbol{U}(a)$, in the vicinity of $a = 0$. Truncating at order N:

$$\boldsymbol{U}(a) = \boldsymbol{U}_0 + \sum_{p=1}^{N} \boldsymbol{U}_p a^p \tag{3.32}$$

where $\boldsymbol{U}_0 = \boldsymbol{U}(0)$ and

$$\boldsymbol{U}_p = \frac{1}{p!} \frac{d^p \boldsymbol{U}}{da^p}\Big|_{a=0}. \tag{3.33}$$

In this case we have selected the following variables to perform the expansion:

$$\boldsymbol{U}(a) = \begin{pmatrix} \boldsymbol{u}(a) \\ \boldsymbol{S}(a) \\ J^2\boldsymbol{C}^{-1}(a) \\ \ln \sqrt{J^2}(a) \\ \frac{1}{J^2}(a) \\ \overline{\lambda}(a) \end{pmatrix}. \tag{3.34}$$

By performing the substitution of the before mentioned variables into the weak form of the problem, see Eq. (3.6) we arrive at a problem entirely similar to that in Eqs. (3.19) and (3.20), in which the tangent stiffness matrix takes the form

$$\boldsymbol{K}_t = \int_{\Omega_0} (\boldsymbol{B}^T \boldsymbol{D} \boldsymbol{B} + \boldsymbol{G}^T \tilde{\boldsymbol{S}}_{\boldsymbol{0}} \boldsymbol{G}) d\Omega, \tag{3.35}$$

where

$$\boldsymbol{D} = \lambda \left(\frac{1}{J_0^2} \boldsymbol{C}_0^{-1} \boldsymbol{M}_0^T \right) + 2(\mu - \lambda \ln J_0) \left(\frac{1}{J_0^2} (\boldsymbol{C}_0^{-1} \boldsymbol{M}_0^T) - \tilde{\boldsymbol{C}}_0 \right) \tag{3.36}$$

Now it takes into account the material non-linearity and it has a somewhat similar appearance to the Lagrangian elastic tensor at the initial state. J_0 and $\boldsymbol{C}_0$ represent the

Jacobian and right Cauchy-Green strain tensor of the initial solution. $\boldsymbol{M}_0$ is obtained from the series expansion of the Jacobian, and contains minors of $\boldsymbol{C}_0$. Finally, $\tilde{\boldsymbol{C}}_0$ is obtained from the series expansion of $\boldsymbol{C}^{-1}$ and contains components of $\boldsymbol{C}_0$, arranged in a particular way.

The geometrical non-linearities are included in the matrices $\boldsymbol{B}$, $\boldsymbol{G}$ and $\tilde{\boldsymbol{S}}_0$. $\boldsymbol{B}$ represents the usual strain-displacement matrix, $\boldsymbol{G}$ relates the nodal displacements $\boldsymbol{u}$ and the gradient of displacement vector, and, finally, $\tilde{\boldsymbol{S}}_0$ represents a matrix that contains the initial stresses (we have chosen the same notation as in Cochelin et al. (1994b)).

In the right hand side of Eq. (3.20), the *non linear load vector* $\boldsymbol{f}_p^{nl}$ is a vector containing information of material and geometrical non-linearities of all order problems ranging from order one to $p-1$. It can be written as:

$$\boldsymbol{f}_p^{nl} = \int_{\Omega_0} (\boldsymbol{B}^T(\boldsymbol{S}_p^{nlmat} + \boldsymbol{S}_p^{nlgeom}) + \boldsymbol{G}^T\boldsymbol{S}_p^*)d\Omega \tag{3.37}$$

As in the stiffness matrix, $\boldsymbol{S}_p^{nlgeom}$ and $\boldsymbol{S}_p^*$ represent the standard matrices found in literature when ANM is used to solve geometrical non-linear problems with linear materials. $\boldsymbol{S}_p^{nlmat}$ takes into account the material behaviour:

$$\begin{aligned}\boldsymbol{S}_p^{nlmat} = (\lambda \ln J_0 - \mu)\left(\boldsymbol{CC}_0\left(RZ_p - \frac{RJ_p}{J_0^4}\right) + \frac{\boldsymbol{RCC}_p}{J_0^2} + \boldsymbol{RC}_p^{-1}\right)\\ + \lambda\left(\frac{\boldsymbol{CC}_0}{J_0^2}\left(RY_p + \frac{RJ_p}{2J_0^2}\right) + \boldsymbol{RS}_p\right)\end{aligned} \tag{3.38}$$

In this equation, $\boldsymbol{CC}_0$ represents the cofactor matrix of $\boldsymbol{C}_0$ and $\boldsymbol{RCC}_p$ is a vector containing values of C_{ij} of all problems from order one to $p-1$, obtained when the cofactor matrix of $\boldsymbol{C}$ is expanded using Taylor series:

$$\boldsymbol{CC}_p = \tilde{\boldsymbol{C}}_0\boldsymbol{C}_p + \boldsymbol{RCC}_p, \tag{3.39}$$

$$\boldsymbol{RCC}_p = \sum_{r=1}^{p-1}\begin{pmatrix} C_{22}^r C_{33}^{p-r} - C_{23}^r C_{23}^{p-r} \\ C_{11}^r C_{33}^{p-r} - C_{13}^r C_{13}^{p-r} \\ C_{11}^r C_{22}^{p-r} - C_{12}^r C_{12}^{p-r} \\ C_{13}^r C_{23}^{p-r} - C_{12}^r C_{33}^{p-r} \\ C_{13}^r C_{12}^{p-r} - C_{11}^r C_{23}^{p-r} \\ C_{12}^r C_{23}^{p-r} - C_{13}^r C_{22}^{p-r} \end{pmatrix}. \tag{3.40}$$

Here, RJ_p is a summation of products of different components of $\boldsymbol{C}_p$ and is obtained when the squared Jacobian is expanded by Taylor series:

$$(J^2)_p = \boldsymbol{M}_0^T\boldsymbol{C}_p + RJ_p. \tag{3.41}$$

$\boldsymbol{RS}_p$ collects terms concerning the expansion of $Y = \ln J$ and $\boldsymbol{C}^{-1}$:

$$\boldsymbol{RS}_p = \sum_{r=1}^{p-1} Y_r \boldsymbol{C}_{p-r}^{-1}. \tag{3.42}$$

$\boldsymbol{RC}_p^{-1}$ collects terms concerning $Z = J^{-2}$ and cofactor matrix of $\boldsymbol{C}$ expansions:

$$\boldsymbol{RC}_p^{-1} = \sum_{r=1}^{p-1} Z_r \boldsymbol{CC}_{p-r} \tag{3.43}$$

Finally, it is necessary to expand $Y = \ln J$ and $Z = J^{-2}$ by using Taylor series and the chain rule generalized to higher derivatives:

$$Y_p = \frac{1}{2J_0^2}(J^2)_p + RY_p, \text{ and } Z_p = \frac{-1}{J_0^4}(J^2)_p + RZ_p, \tag{3.44}$$

where

$$\begin{aligned}
RY_1 &= 0, \\
RY_2 &= \frac{-1}{4J_0^4}(J^2)_1^2), \\
RY_3 &= \frac{1}{6J_0^6}(J^2)_1^3 + 2\frac{-1}{4J_0^4}(J^2)_1(J^2)_2, \\
RZ_1 &= 0, \\
RZ_2 &= \frac{1}{J_0^6}(J^2)_1^2, \\
RZ_3 &= \frac{-1}{J_0^8}(J^2)_1^3 + 2\frac{1}{J_0^6}(J^2)_1(J^2)_2, \\
&\dots
\end{aligned}$$

Once the expansion of the ANM has been performed, the next step of the method consists of the use in this framework of a reduced model, as explained in the previous sections for a general case, regardless of the constitutive model chosen.

3.4 Numerical results

In order to show the performance of this method we have applied it to some problems with hyperelastic materials.

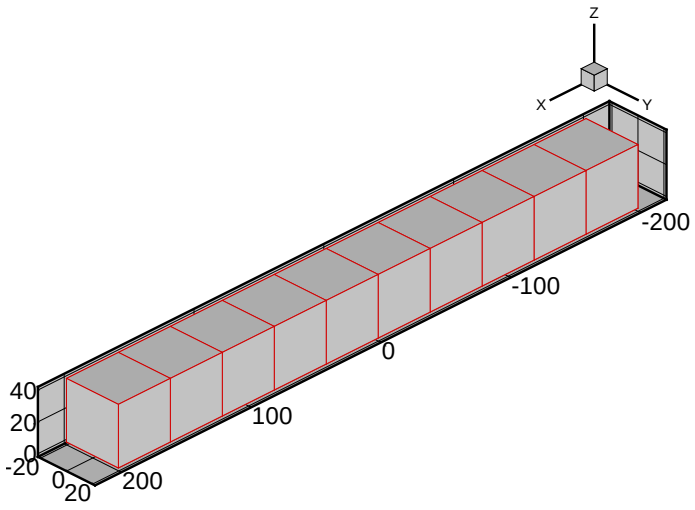


Figure 3.1. Geometry of the beam in traction.

3.4.1 Beam in traction

In this example we have applied four concentrated $100N$ forces to the left face of a beam in the axial direction. The nodes on the right face are constrained in the axial direction. The length of the beam is $400mm$ and the cross section is a square with sides of length $40mm$. The beam is shown in Figure 3.1. The material properties of the beam are $E = 1MPa$ and $\nu = 0.25$.

Saint Venant-Kirchhoff material

In this example the dimension of the basis of the reduced model is six, which are the number of modes necessary to capture the 99.9% of the energy of the system. They are depicted in Fig. 3.2. The load-displacement curve has been obtained for a node on the left face and is shown in Figure 3.3 for $p = 1, \ldots, 6$. The linear solution obtained with $p = 1$ is the same as the one that can be obtained using standard POD only, without updating the tangent stiffness matrix. In this case for higher p the load-displacement curves are clearly non-linear and in this example they coincide with the analytic solution and the one calculated using full FE Newton-Raphson equilibrium iterations (the curves are indistinguishable for $p \geq 2$).

Notice how large is the radius of convergence of the proposed technique for $p \geq 2$ (on the order of $60mm$ of tip displacement), for a beam with a total initial length of

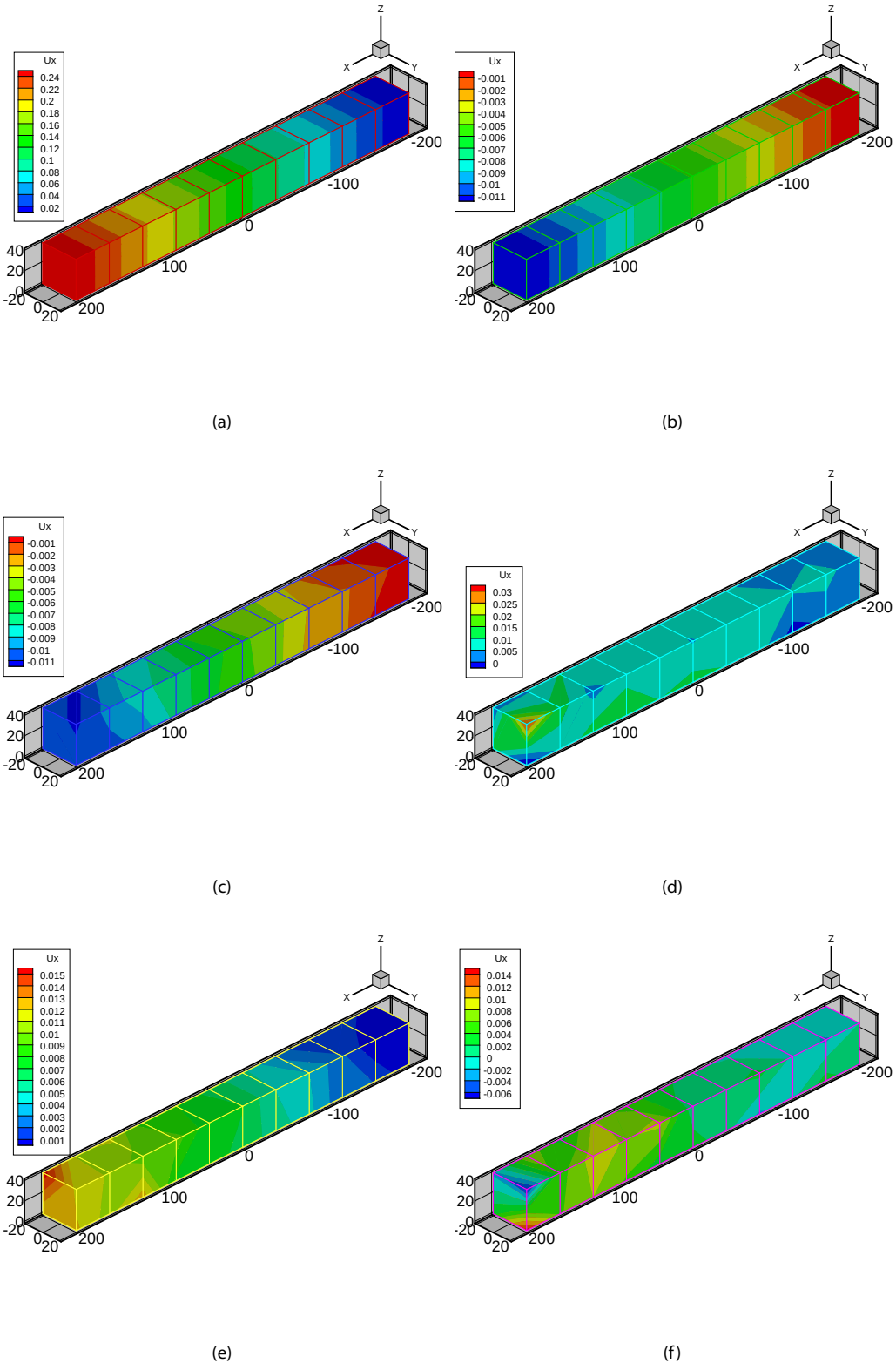


Figure 3.2. Six most important eigenmodes for the simulation of the Saint Venant-Kirchhoff beam under traction. The corresponding eigenvalues are: 38488.48 (a), 0.04 (b), $1.04E-11$ (c), $3.44E-12$ (d), $1.34E-12$ (e) and $1.28E-12$ (f).

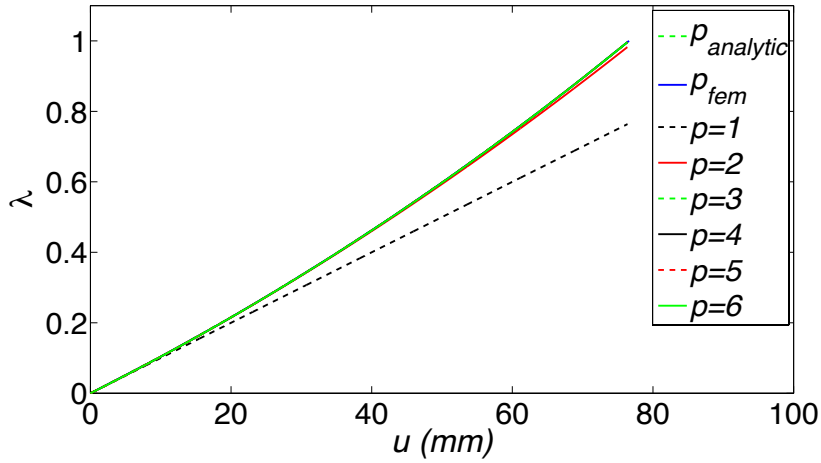


Figure 3.3. Beam in traction. Load-displacement results (in terms of $\overline{\lambda}$) for different order of approximation, p, compared with the analytical solution. Saint Venant-Kirchhoff behaviour.

$400mm$. If it is needed to follow the equilibrium path beyond this point, the method can be restarted at any equilibrium point, by a new computation of the stiffness matrix at that point.

Neo-Hookean material

For the neo-Hookean material the behavior of the proposed technique is very similar to that of Saint Venant-Kirchhoff. The load-displacement curve is depicted in Fig. 3.4, where an excellent agreement between the FEM results and the $p = 4$ reduced model has been noticed.

In Fig. 3.5 the six most important modes of the complete solution, obtained by Newton-Raphson methods, are depicted.

The excellent agreement of the results for very large beam tip displacements also deserve some comments. Note that we have plotted the solution of the complete FE model up to a tip displacement of around $120mm$. Up to this very large level of strain (the beam is 400 mm long) the reported solution by the reduced model with an order $p = 4$ is remarkable, while it seems to reproduce the "expected" solution far beyond that point.

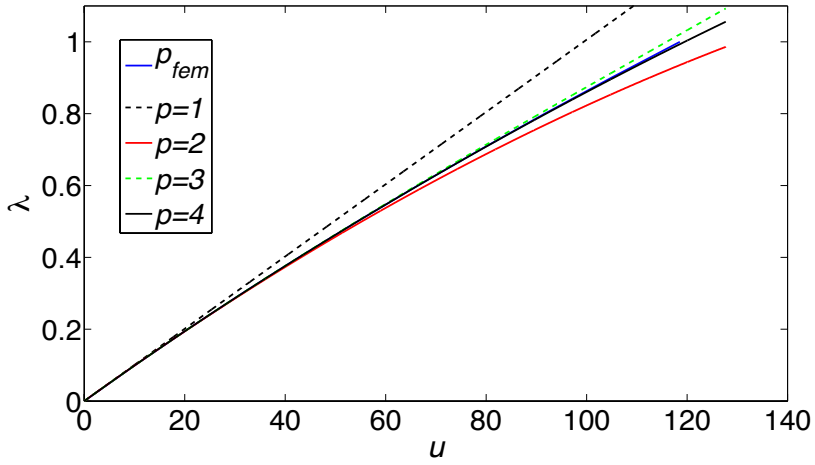


Figure 3.4. Beam in traction. Load-displacement results (in terms of $\overline{\lambda}$) for different order of approximation, p, compared with the analytical solution. Neo-hookean behaviour.

3.4.2 Pinched hemisphere

In this example a concentrated force of $30N$ is applied to the pole of a hemisphere and directed towards its centre. The bottom of the hemisphere is fixed. Its radius is $25mm$ and its Young's modulus and poison's ratio are $2MPa$ and 0.48 respectively. The mesh is shown in Figure 3.6 and it is composed of 614 nodes and 448 linear hexahedral elements. Saint Venant-Kirchhoff and neo-Hookean constitutive laws are assumed in this example.

Saint Venant-Kirchhoff material

In this example the empirical basis has ten modes, depicted in Fig. 3.8. The load-displacement curves for the node on the pole of the hemisphere for $p = 1, \ldots, 6$ have been obtained and depicted in Figure 3.7.

Again it is worth noting that the solution with $p = 1$ that is linear is the same as the solution that one could obtain using standard model reduction without stiffness matrix updating. For $p \leq 5$ it can be seen that the ANM-POD solution has a good agreement with the solution computed using full Newton-Raphson equilibrium iteration up to $u \approx -2.5mm$, which can be considered as the radius of convergence (note that this is in the order of 10% of the sphere radius). But for $p = 6$ the size of the radius of convergence is remarkable, going beyond $4mm$ (more than 20% of the radius of the sphere). Again,

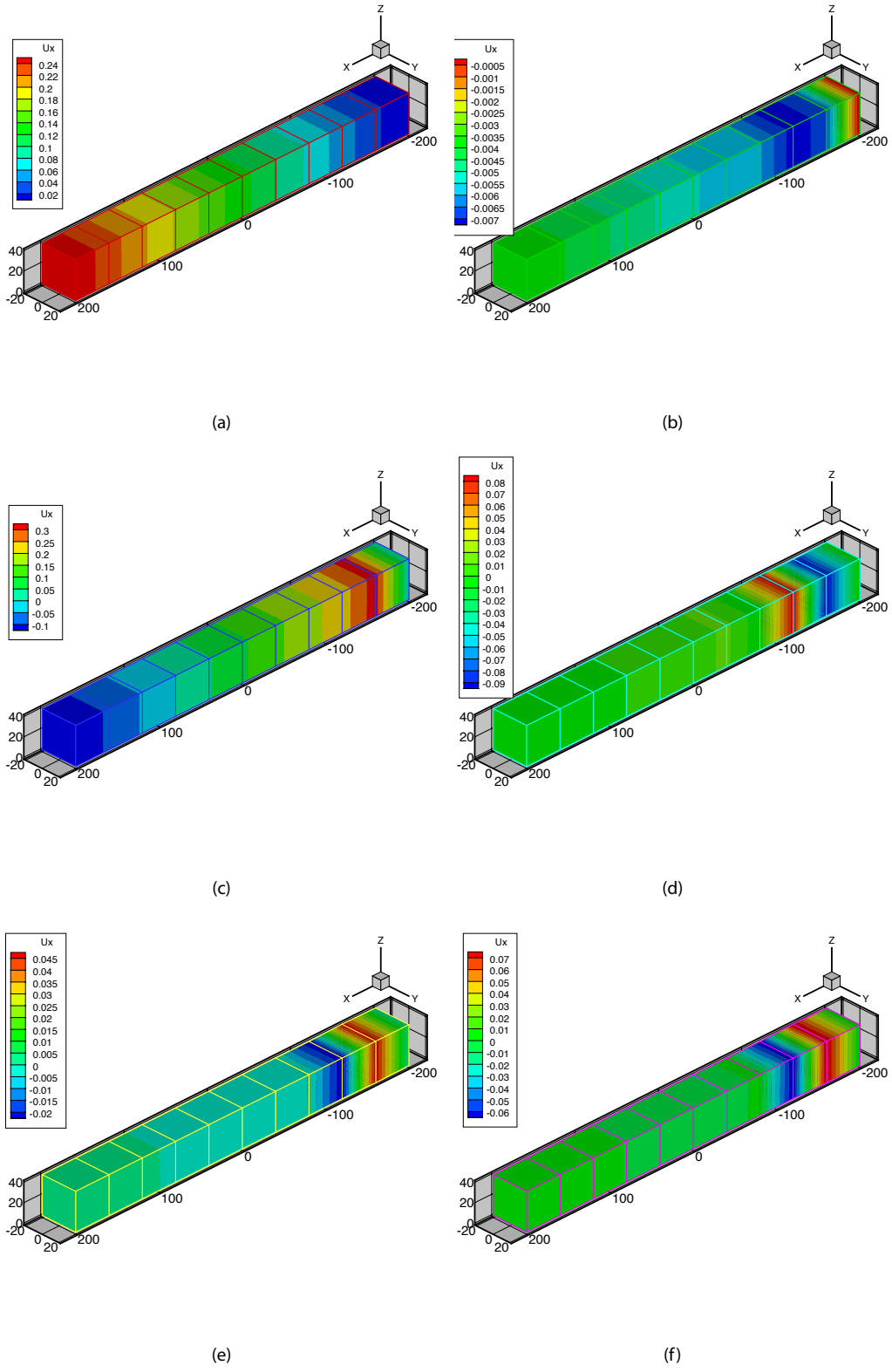


Figure 3.5. Six most important eigenmodes for the simulation of the neo-hookean beam under traction. The corresponding eigenvalues are: 38488.48 (a), 0.04 (b), $1.04E-11$ (c), $3.44E-12$ (d), $1.34E-12$ (e) and $1.28E-12$ (f).

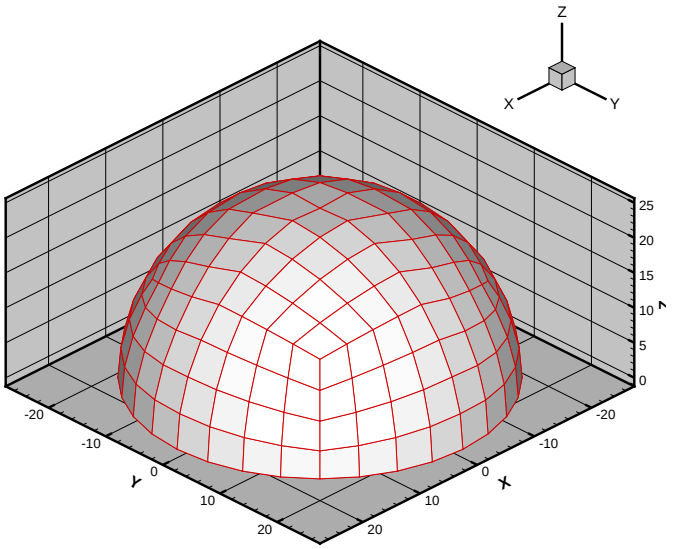


Figure 3.6. Hemisphere mesh.

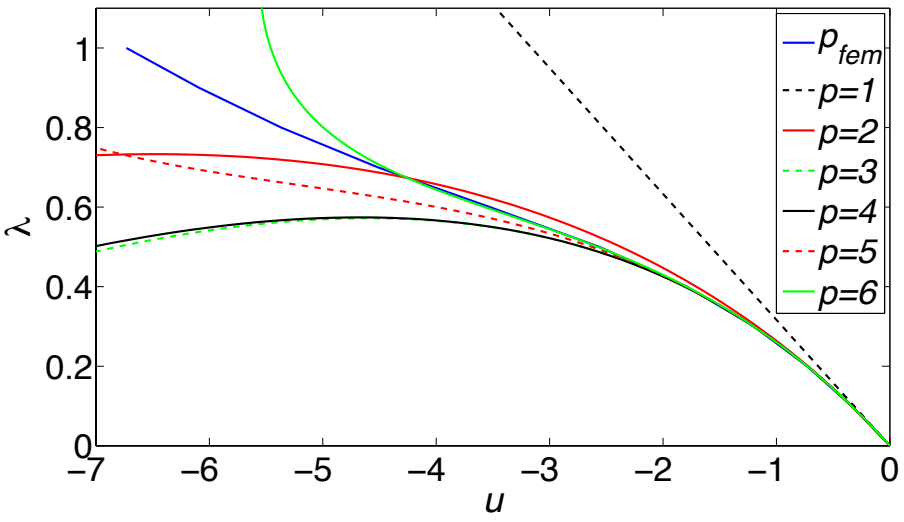


Figure 3.7. Load-displacement curves for the hemisphere problem.

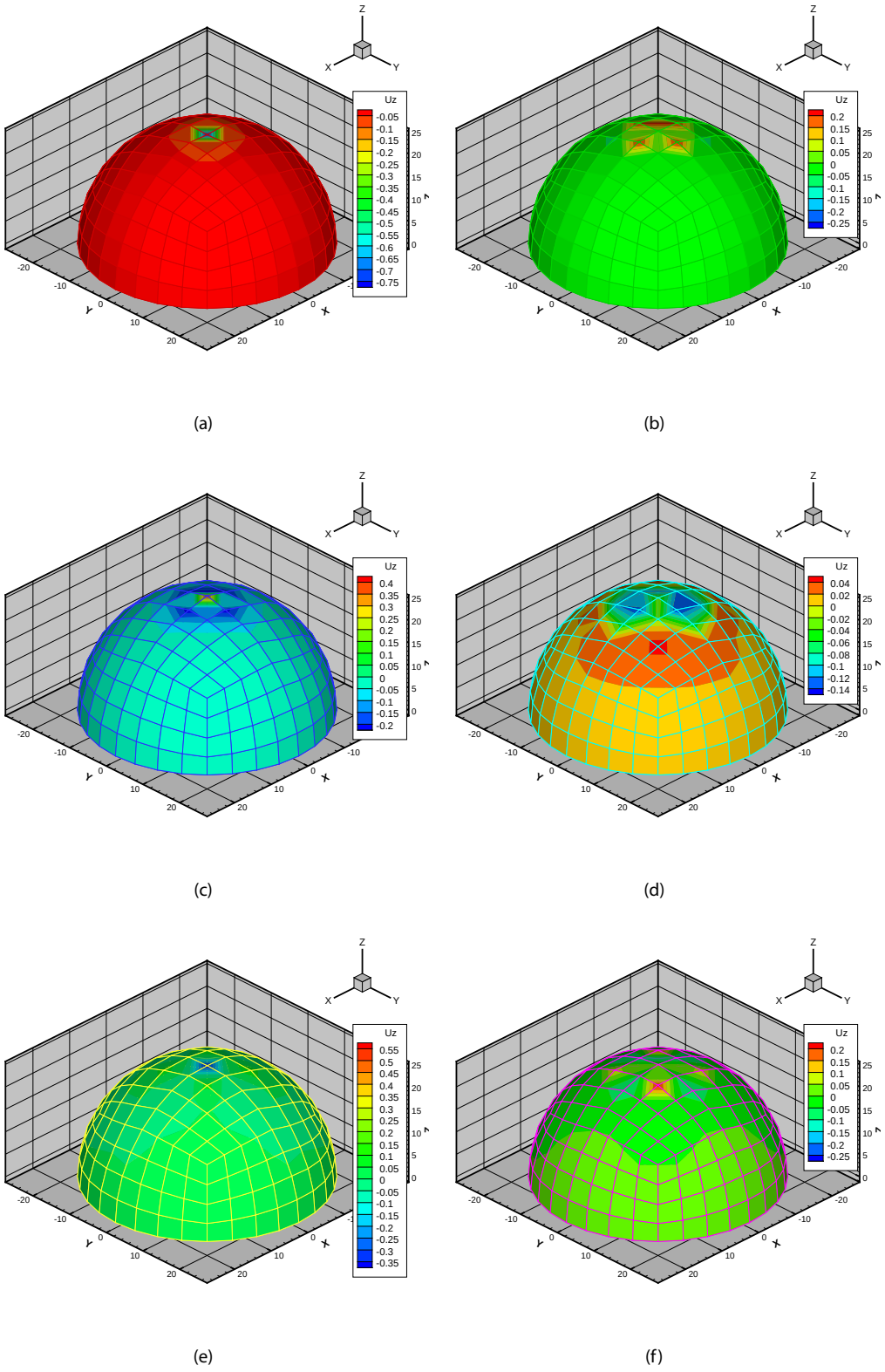


Figure 3.8. Ten most important eigenmodes employed for the simulation of the pinched hemisphere. The corresponding eigenvalues are: 25.42 (a), 0.19 (b), 0.02 (c), $4.80E-05$ (d), $6.04E-07$ (e), $1.15E-08$ (f), $2.85E-10$. Saint Venant-Kirchhoff behavior.

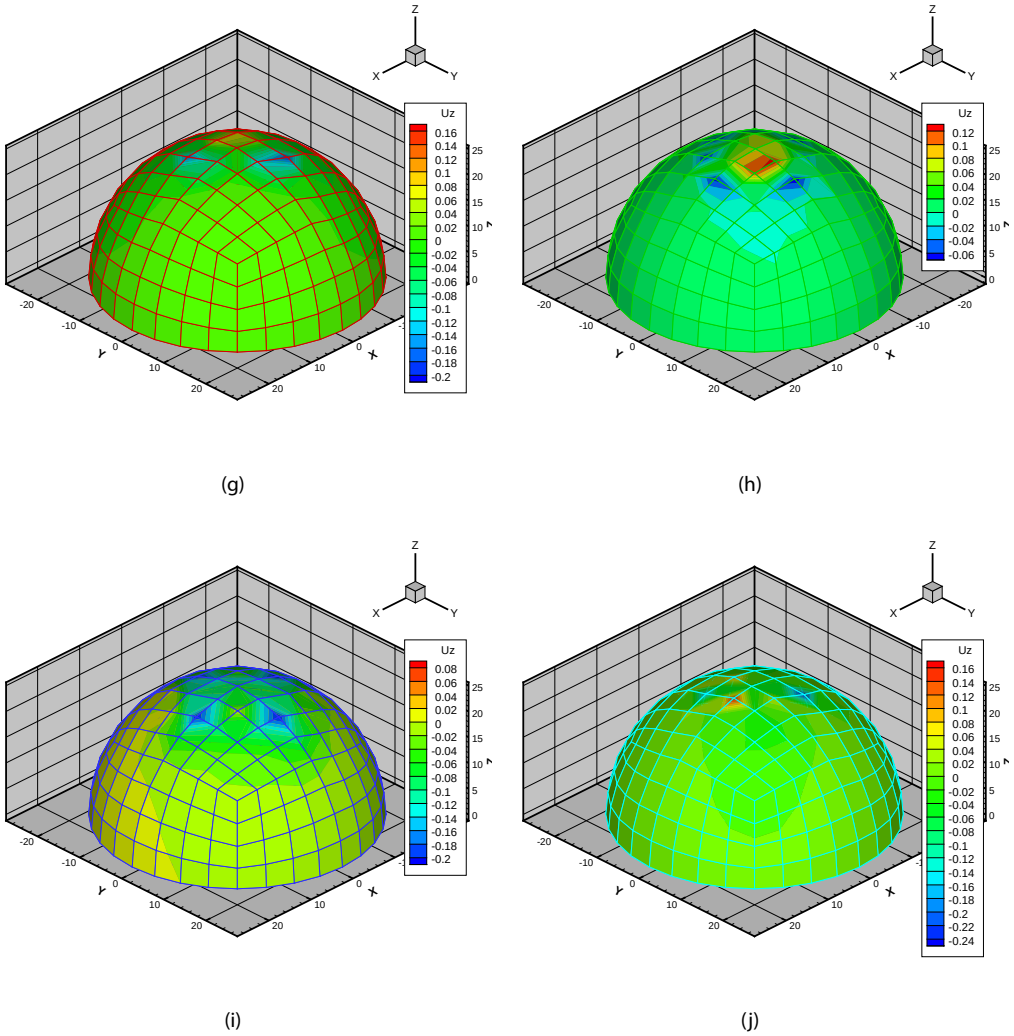


Figure 3.8. Ten most important eigenmodes employed for the simulation of the pinched hemisphere (Cont.). The corresponding eigenvalues are: (g), $4.12E-12$ (h), $6.52E-14$ (i) and $2.32E-16$ (j). Saint Venant-Kirchhoff behavior.

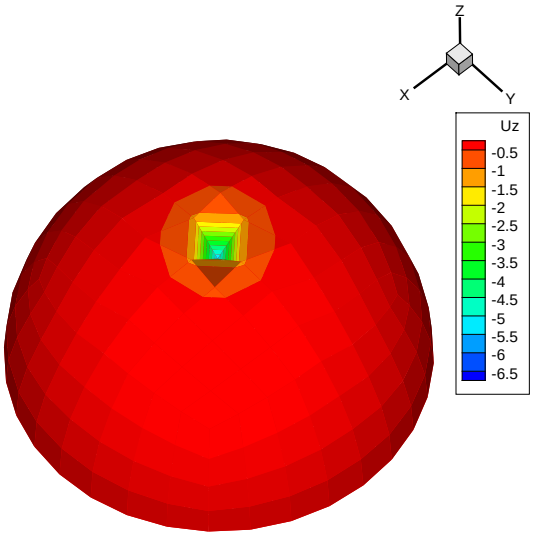


Figure 3.9. Deformed hemisphere obtained using the combined technique POD-ANM.

if the load is to be extended beyond this point, a new tangent stiffness matrix should be computed at this point, performing again the algorithm given by Eqs. (3.28)-(3.29).

The deformed hemisphere for $p = 6$ and $\overline{\lambda} = 1$ is shown in Figure 3.9 and the displacement field in z direction, u_z, obtained by Newton-Raphson procedures is shown in Figure 3.10.

Neo-Hookean behaviour

As in the previous example, results were tested also for the neo-Hookean behaviour, showing similar levels of accuracy as the Saint Venant-Kirchhoff examples. The modes employed as global basis in the example are depicted in Fig. 3.11.

Results for the complete model, solved by means of full Newton-Raphson iterations, and the reduced model, for different approximation order, are compared in Fig. 3.12. The great similitude for both models is noteworthy. Again, the load-displacement (in terms of the load factor, $\overline{\lambda}$) curve shows great similarity for $p = 2, \ldots, 4$ and a large convergence radius.

3.4.3 Pinching the human cornea

As mentioned before, real-time simulation in surgical environments is one of the fields in which very fast simulations are needed. This field arises as a natural potential appli-

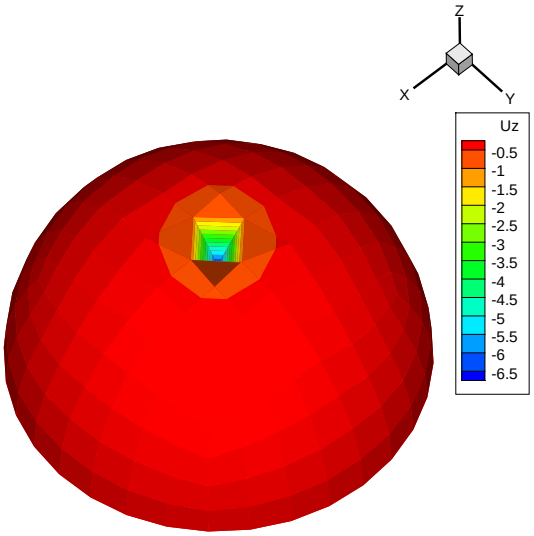


Figure 3.10. Deformed hemisphere computed by full FE Newton-Raphson equilibrium iteration.

cation of the technique here proposed, which can solve with moderate accuracy simplified models of non-linear solids without the need for tangent stiffness matrix updating nor iterative procedures.

In this example forces of $0.014N$ each have been applied to nine neighbor nodes located at the center of the cornea, thus simulating the contact of a rounded tool. The mesh data and boundary conditions are as explained in chapter 2.

Saint Venant-Kirchhoff model

The material properties of the cornea are assumed to be $E = 2MPa$ and $\nu = 0.48$, although a realistic model of corneal tissue could eventualy be used, as in chapter 2 and Niroomandi et al. (2008).

In this example nine modes were applied that provide decent approximation. The solution has been obtained using ANM-POD for $p = 1, \ldots, 6$, plotted in Fig. 3.13. Note how closely the first mode resembles the final solution. Higher modes contribute to capturing the solution near the region where the load is applied.

In order to verify the results we have computed the solution by full FE Newton-Raphson method. The loading factor ($\overline{\lambda}$) has been plotted versus the maximum (absolute value) displacement in Figure 3.14.

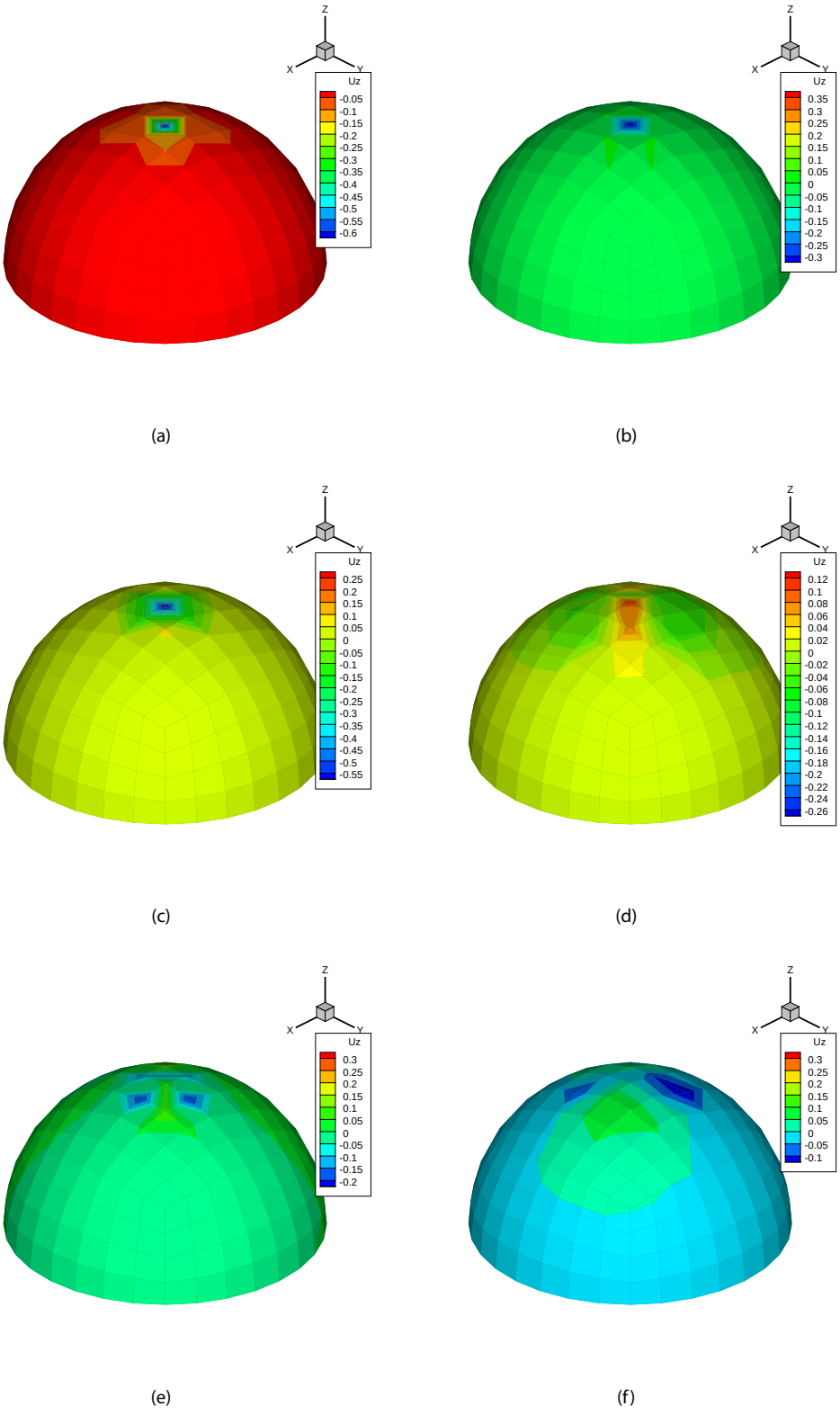


Figure 3.11. Six most important eigenmodes for the simulation of the neo-hookean pinched hemisphere. The corresponding eigenvalues are: 30.0 (a), 0.013 (b), $2.95E-05$ (c), $9.42E-08$ (d), $3.12E-10$ (e) and $4.14E-12$ (f).

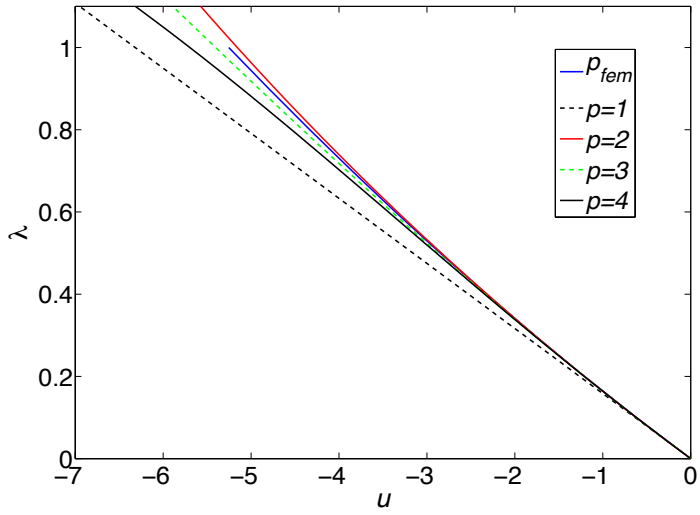


Figure 3.12. Load-displacement curves for the hemisphere problem with neo-hookean behaviour.

As it can be noticed the results have good accuracy with Newton-Raphson solution. The deformed cornea obtained using ANM-POD for $\overline{\lambda} = 1$ is shown in Figure 3.15 and the one obtained using full FE Newton-Raphson is depicted in Figure 3.16.

Note that, despite the high Poisson's ratio employed, close to 0.5, the method shows no tendency to lock in the examples studied (although it is well known that this does not constitute a valid proof of robustness).

Neo-Hookean model

In this case, only six modes were enough to simulate the cornea, thus making a stiffness matrix of size 6×6. These modes are depicted in Fig. 3.18. Again, results for the neo-hookean cornea follow the same guidelines of accuracy as those for the Saint Venant-Kirchhoff model. With an expansion of only three or four terms (see Fig. 3.17) a great accuracy is obtained along a big radius of convergence.

3.4.4 Palpation of the liver

In this example we consider the palpation of a liver with a surgical tool, assumed perfectly rigid. The liver is the biggest gland in the human body, after the skin. Liver geometry has been obtained from the SOFA project (Allard et al., 2007) and after post-processing we have obtained a mesh composed of 2853 nodes and 10519 tetrahedral

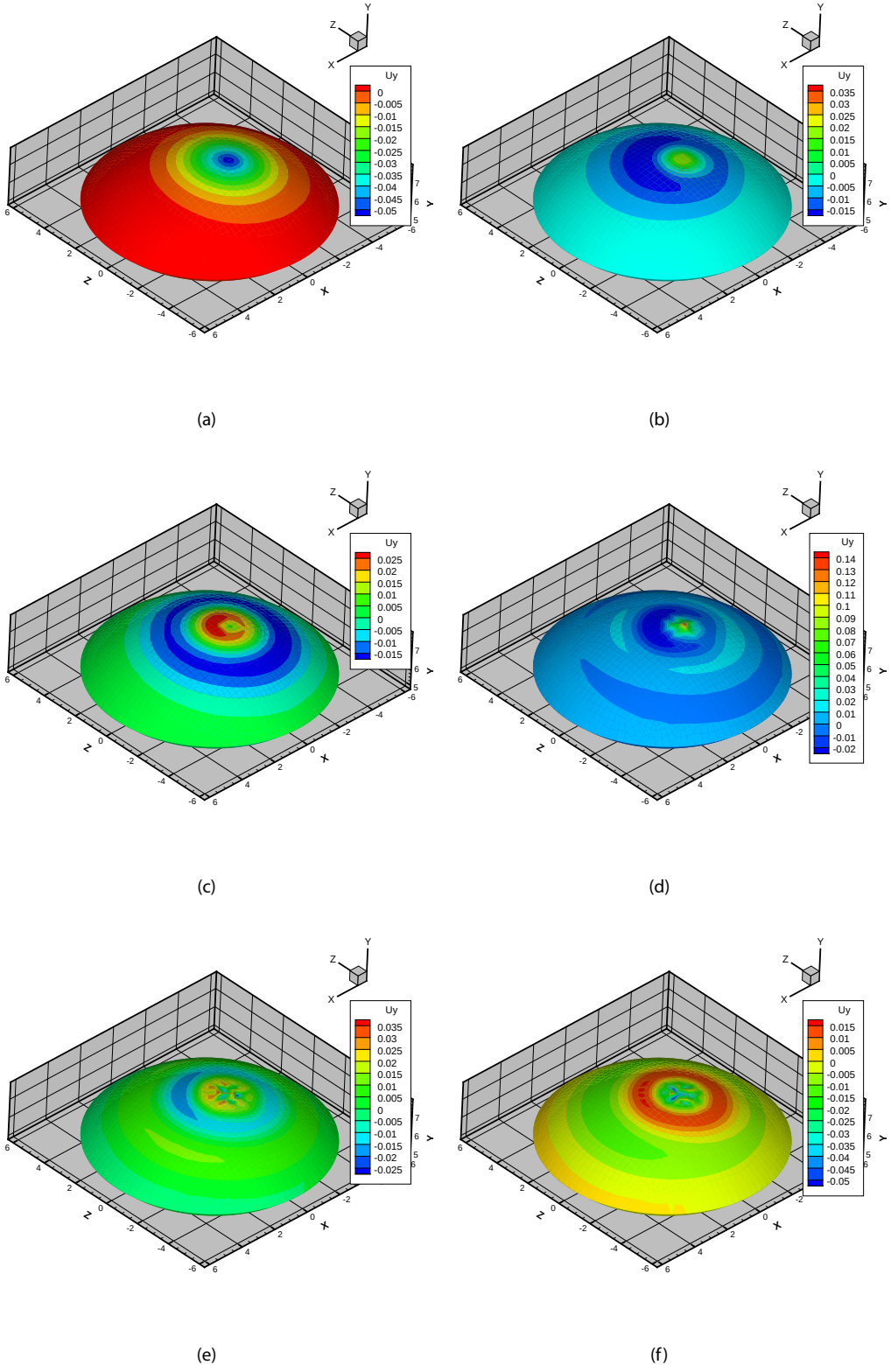


Figure 3.13. Six most important eigenmodes for the simulation of the cornea. The corresponding eigenvalues are: 67.5 (a), 0.07 (b), $5.95-4$ (c), $7.54E-7$ (d), $1.26E-8$ (e) and $3.48E-10$ (f).

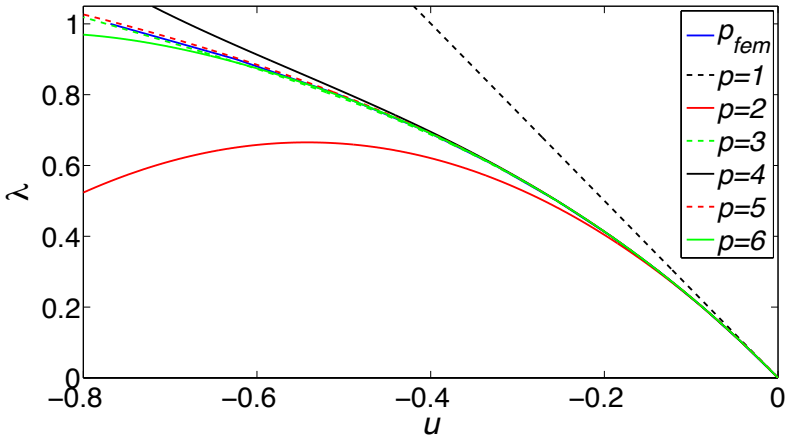


Figure 3.14. The loading factor vs. minimum displacement for the Pinched cornea. Saint Venant-Kirchhoff model.

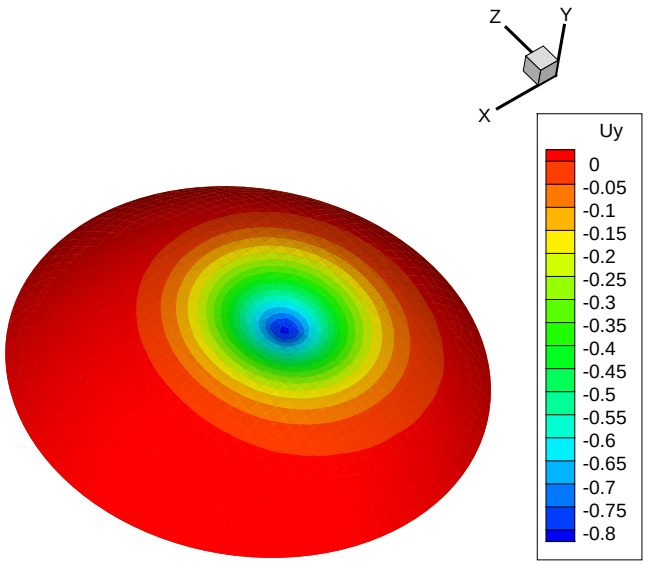


Figure 3.15. u_y-contour of the pinched cornea obtained by ANM-POD.

elements, whose geometry is shown in Fig. 3.19. The anterior surface of the liver is considered free, while the posterior one was assumed to be supported over different organs (it is connected to the diaphragm by the coronary ligament, for instance). The inferior vena cava travels along the posterior surface, and the liver is frequently assumed

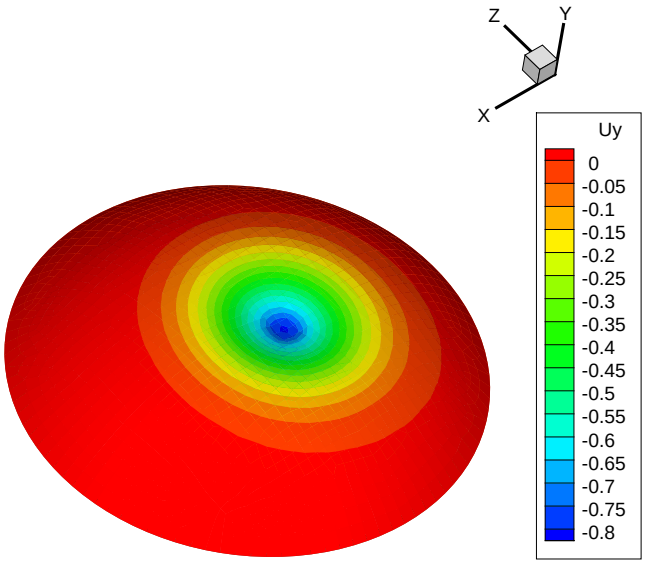


Figure 3.16. u_y-contour of the pinched cornea obtained by a full FE model with Newton-Raphson iterations.

clamped at that location. Although the assumed boundary conditions are not strictly correct from a physiological point of view, our main interest is to show that the model can be solved under real-time constraints with reasonable accuracy.

Although the literature on the mechanical properties of the liver is not very detailed, we have assumed a Young's modulus of 160 kPa, and a Poisson coefficient of 0.48, thus nearly incompressible (Delingette and Ayache, 2004).

A load of $1.2N$ has been applied at an arbitrary point on the surface of the liver and the reduced model has been constructed, composed in this case by the 9 modes depicted in Fig 3.20, that capture the 99.9% of the energy of the system.

In order to compare the results a standard Newton-Raphson algorithm was employed, without any modification such as arc-length methods or anything similar. The load (in terms of load factor $\overline{\lambda}$)-displacement curve obtained by POD-ANM techniques is depicted in Fig. 3.21. It is worth noting the high accuracy of the results for a wide range of displacements (up to $\approx 5mm$ under the tool tip) for expansions of order 5 and 6, less than 5%. This accuracy can also be noticed from Figs. 3.22 and 3.23, representing the displacements for FE complete mode and POD-ANM (sixth-order expansion)

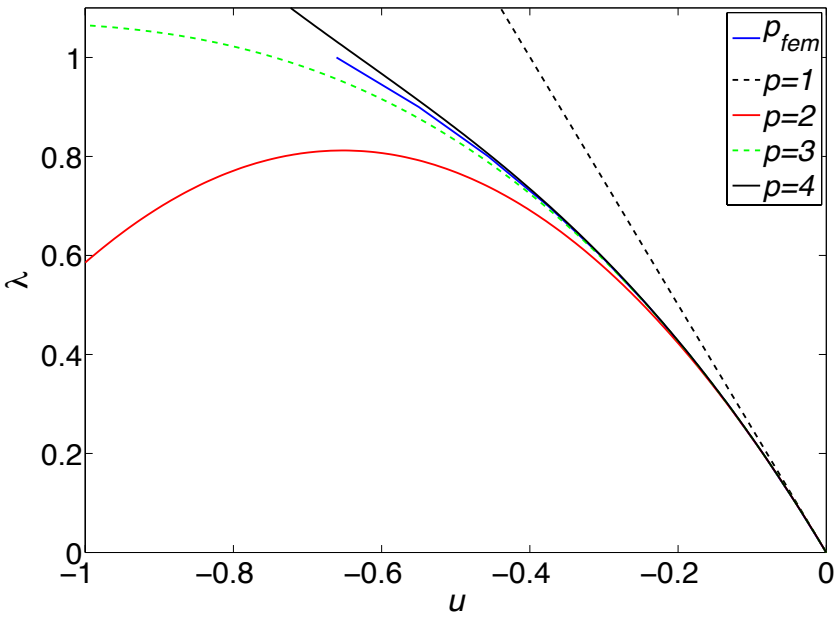


Figure 3.17. The loading factor vs. minimum displacement for the Pinched cornea. Neo-hookean model.

techniques, respectively.

3.4.5 Influence of the interpolation of reduced basis

The other big ingredient of the resulting method, namely the non-linear force term $\boldsymbol{f}_p^{nl}$ of Eqs. (3.28)-(3.29) also needs to be interpolated between pre-computed models, since it is too computationally-demanding to be evaluated in real-time. To this end, we establish a piece-wise polynomial (FE-like) interpolation among complete models. This simple procedure has rendered excellent results, as will be shown in the subsequent section.

If we consider a case in which the load is placed at a position for which a reduced model has not been stored, the interpolation procedure, as described in section 2.3, is applied. As mentioned before, the tangent stiffness matrix at origin is identical for all load states. The procedure follows by (linearly) interpolating the reduced basis set from the four nearest neighbor pre-computed states and also by interpolating the non-linear force term $\boldsymbol{f}_p^{nl}$ for each order of the expansion.

The resulting load-displacement curve is depicted in Fig. 3.24. The excellent agreement with the complete model results (denoted by p_{FEM} in the legend) is noticeable.

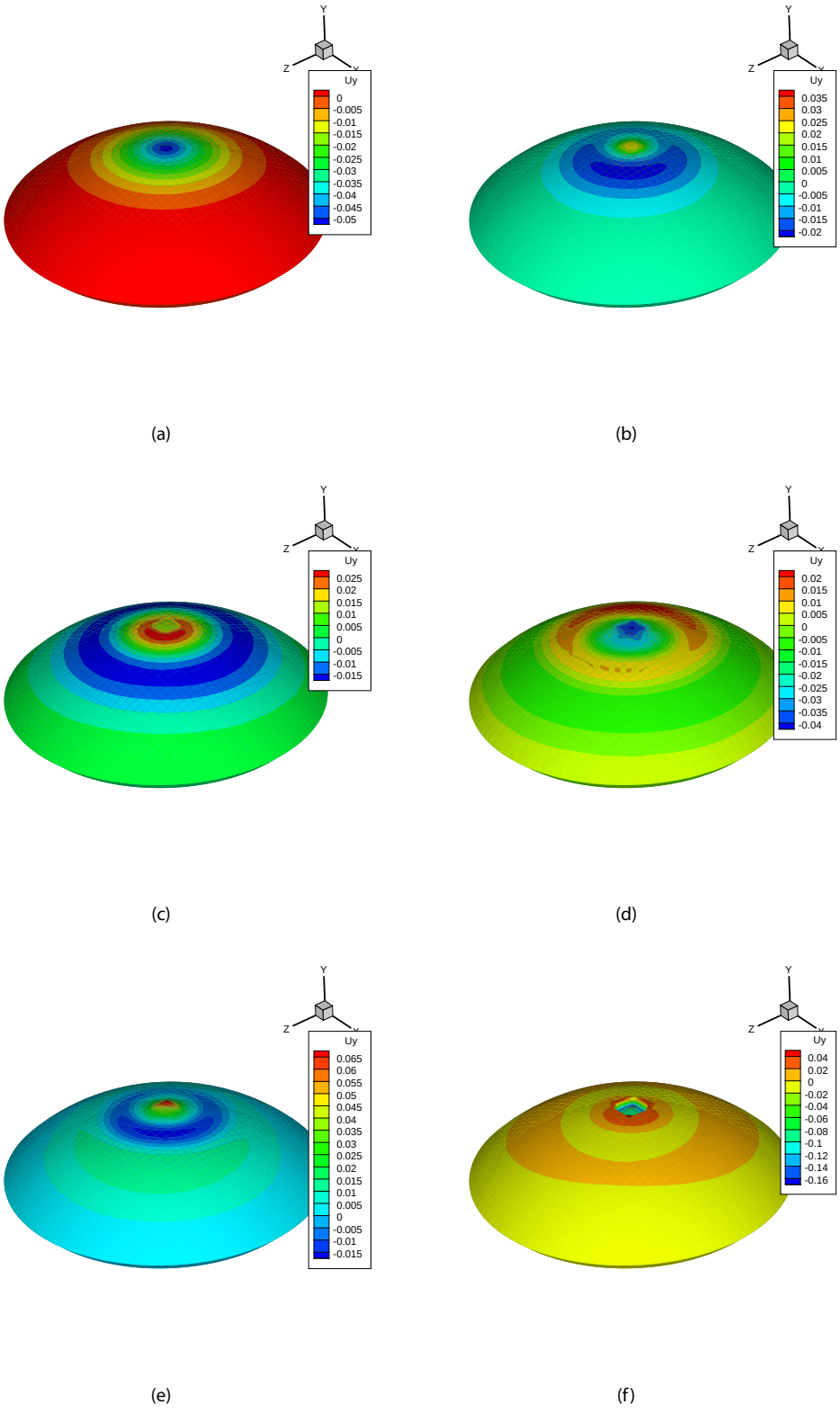


Figure 3.18. Six most important eigenmodes for the simulation of the neo-Hookean cornea. The corresponding eigenvalues are: 53.5 (a), $2.44E-02$ (b), $2.09E-04$ (c), $7.60E-08$ (d), $3.78E-10$ (e) and $7.19E-13$ (f).

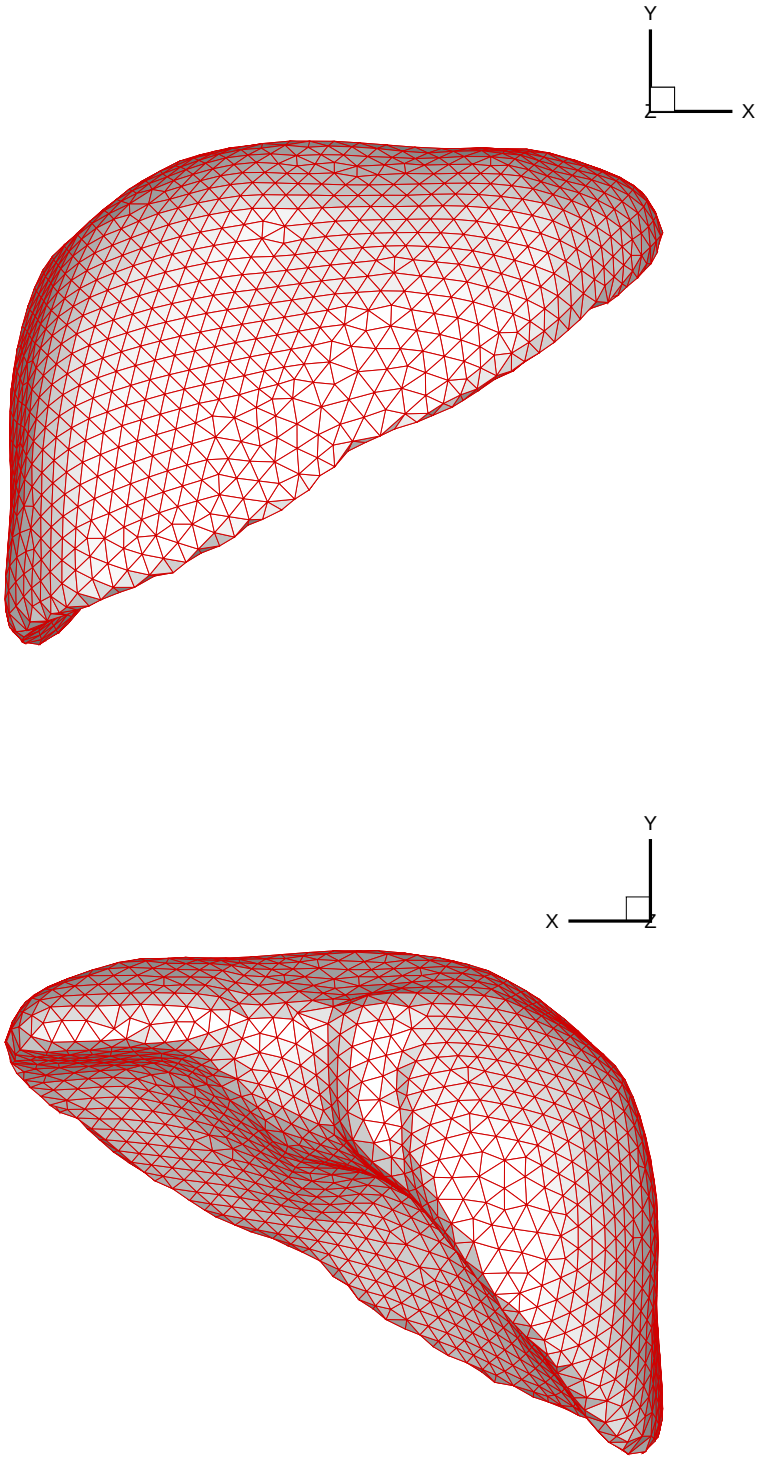


Figure 3.19. Geometry of the finite element model for the liver.

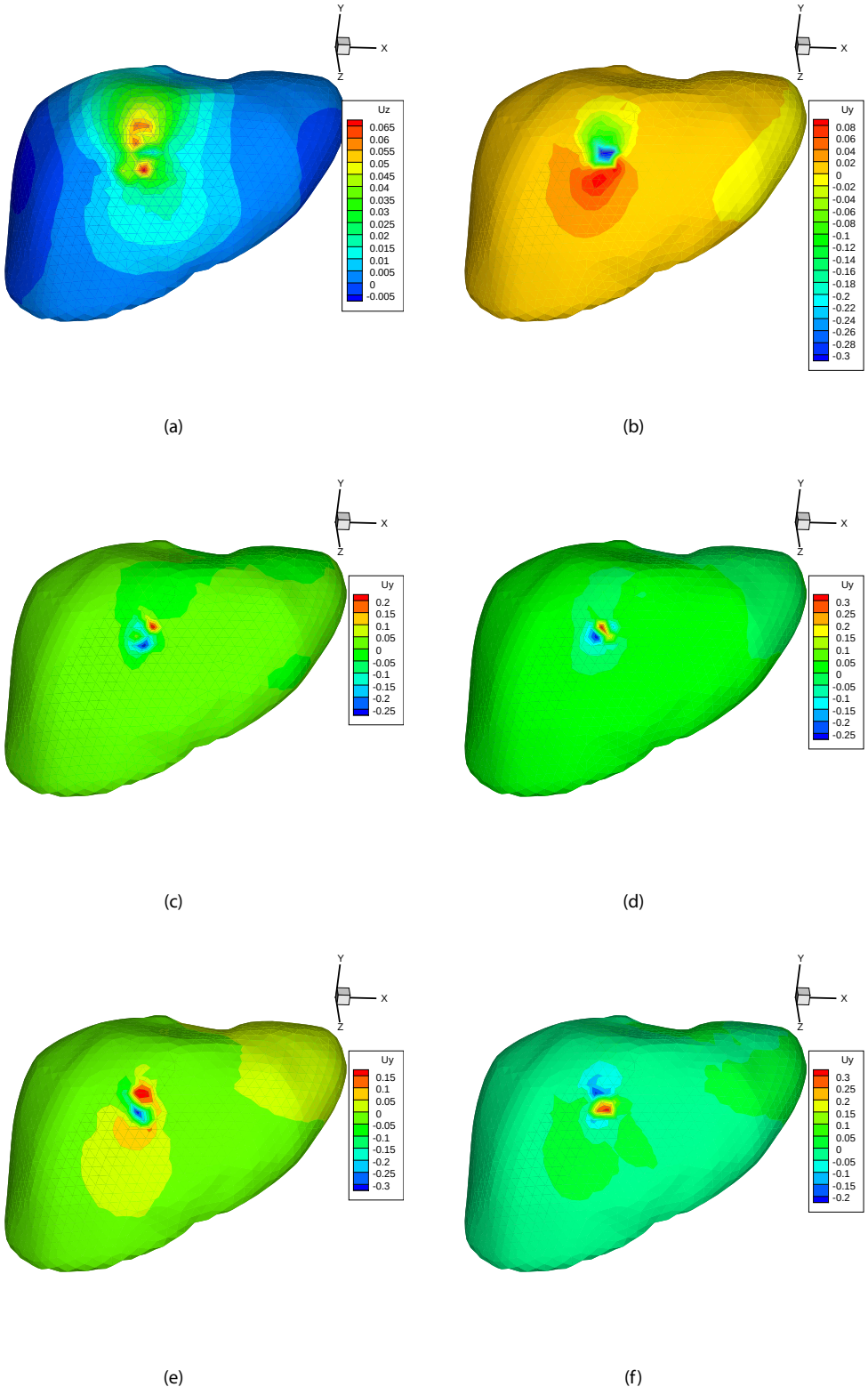


Figure 3.20. Nine most important eigenmodes for the simulation of the liver palpation. The corresponding eigenvalues are: $2.69E-03$ (a), $8.28E-05$ (b), $8.82E-07$ (c), $1.07E-08$ (d), $6.73E-11$ (e), $5.21E-13$ (f), $2.52E-15$ (g).

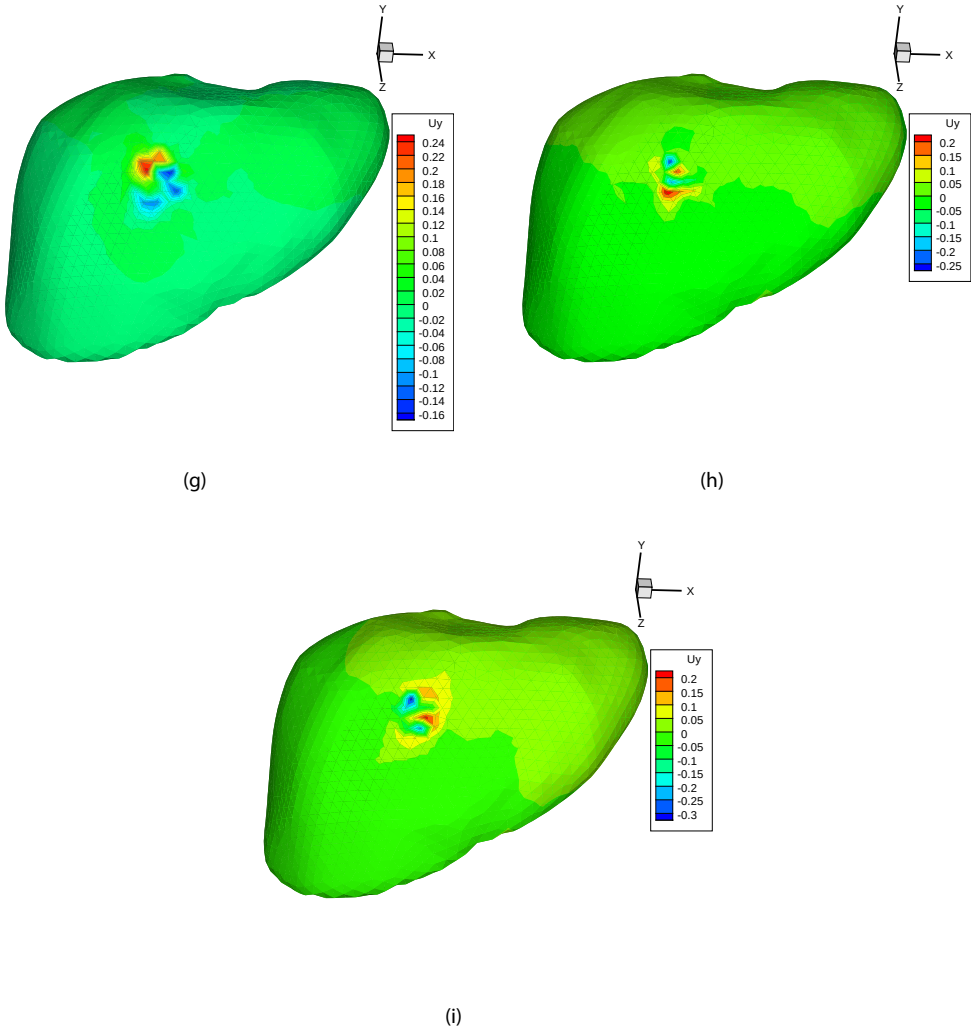


Figure 3.20. Nine most important eigenmodes for the simulation of the liver palpation (Cont.).

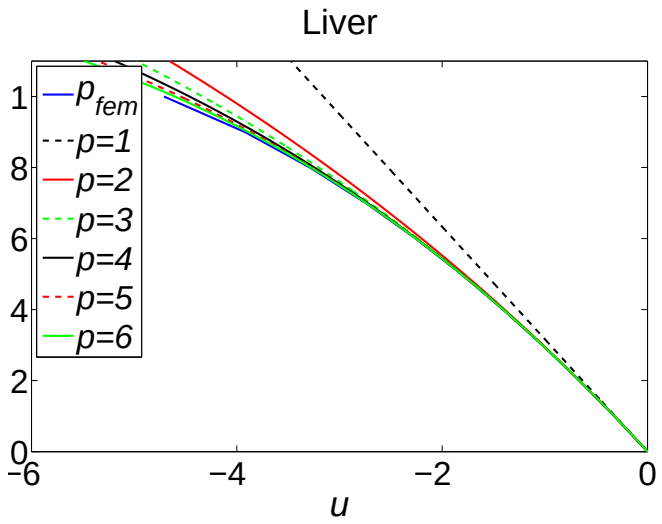


Figure 3.21. Load-displacement, $[mm]$, curve for the liver palpation for different orders of approximation. The continuous blue line represents the solution for the complete model employing Newton-Raphson method.

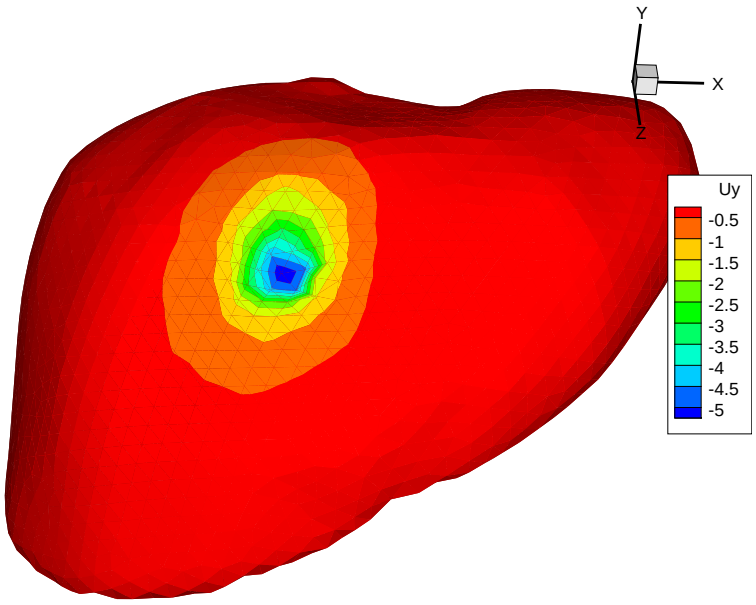


Figure 3.22. Result for the FEM model, displacement u_y, $[mm]$.

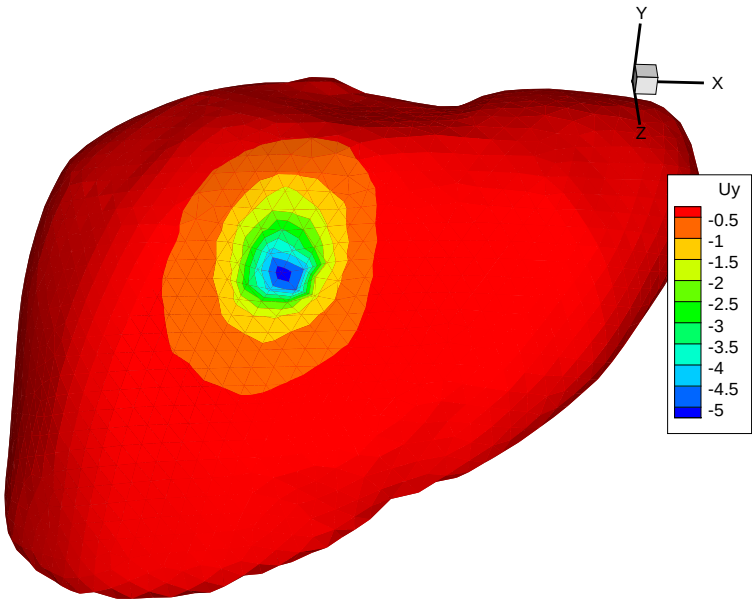


Figure 3.23. Result for the reduced-ANM model, displacement u_y, $[mm]$.

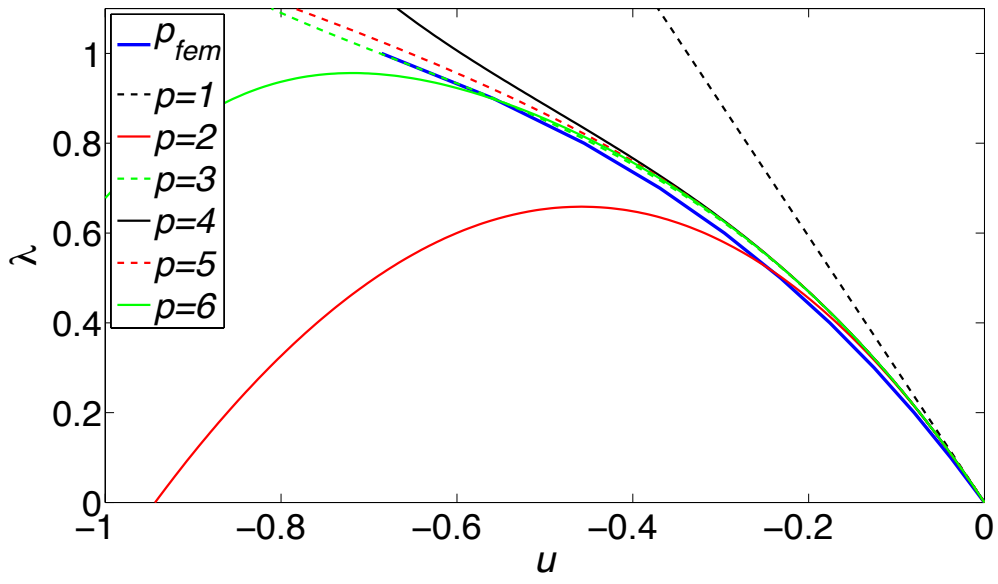


Figure 3.24. The loading factor vs minimum displacement, $[mm]$, for the Pinched cornea. The resulting behavior is interpolated among four previously stored reduced models.

3.4.6 Timing

The examples below have been performed on a laptop running Matlab on a Windows-operated computer at 2 GHz., with 2 Gb RAM. The code prototype is obviously thought for the design and testing of new algorithms. Under these circumstances, the examples ran at around 20 Hz, which is enough for visual perception, but not for haptic environments. This can be improved by considering more sophisticated programming and not a high level language such as Matlab. The summary of the computational cost for a typical simulation is as follows:

- Interpolation of the basis and the non-linear force term: $3.9 \cdot 10^{-3} s$.
- Construction of the reduced stiffness matrix $\boldsymbol{A}^T \boldsymbol{K}_t \boldsymbol{A}$ (note that the complete tangent stiffness matrix at the origin, $\boldsymbol{K}_t$, is identical for all simulations and therefore could be previously stored in memory): $4.7 \cdot 10^{-2} s$.
- Solving the resulting system of equations (6×6 in this particular example): $3.1 \cdot 10^{-4} s$.

The use of Graphics Processing Units (Taylor et al., 2008) is obviously another possibility to achieve true real-time performance, that seems to be at hand.

3.5 Discussion

In this chapter we have studied a technique for the construction of reduced models of hyperelastic solids. It is based on the use of model reduction techniques (based upon Proper Orthogonal Decomposition or Karhunen-Loève methods) and an asymptotic expansion of the solution in the neighborhood of the last converged equilibrium point. This technique has been generalized here for hyperelastic materials involving material non-linearities.

These two ingredients render a method with very attractive features, such as the use of very few degrees of freedom (up to ten in the examples showed in this chapter) and the necessity to compute only one tangent stiffness matrix within the region covered by the radius of convergence of the asymptotic expansion. This radius of convergence is often relatively large (on the order of 10% of the total dimensions of the solid for the examples showed herein). Within this radius of convergence neither iterative procedures (typical of modified Newton methods) nor the update of the stiffness matrix (typical in standard Newton-Raphson methods) are required.

The examples developed in this work showed that between six to ten eigenmodes are enough for the vast majority of the examples considered, and usually an expansion of order 4-6 renders very acceptable results, with excellent accuracy within a vast radius of convergence.

Then we have focused on various ways of interpolating the reduced basis and also non-linear forces in order to obtain near real-time rates.

Other approaches based upon reduced models exist, see Barbi**c** and James (2005), for instance. The main difference between the approach followed in this work and that of Barbic and James is that in the latter case a general-purpose reduced basis is employed. These basis functions are obtained from mass-scaled principal component analysis. These basis functions are thought to be optimal for any load state of the solid. In the present approach different sets of basis functions are obtained for different load (and possibly boundary conditions) states that can be changed according with the user experience. Basis sets for any state different from the precomputed ones are obtained after interpolation, as explained in the previous sections. In section 2.3 it has been explained how the set of reduced models for different loading states of the organ forms a manifold. Therefore, it is crucial for the method to work well that the set of complete models to be solved to feed the algorithm is chosen adequately. For the moment, the brute-force approach seems to be out of reach. In our opinion, this process should be guided by surgeons, by indicating the most probable loading states for each surgical procedure, distinguishing, for instance, between experienced surgeons and medicine students.

The method has been developed for Saint Venant-Kirchhoff constitutive laws. This kind of laws is among the best state-of-the-art existing models (Barbi**c** and James, 2005), but is judged to be poor for some applications. It is well-known that it suffers from instabilities when subjected to compression (although no instabilities have been found in the examples tested by the authors), so it would be interesting to extend the proposed technique to some more sophisticated constitutive laws, such as neo-hookean laws, for instance, or other much better suited for specific applications (Alastrué et al., 2006).

Despite the facts commented above, the proposed technique seems to be an appealing method for the simulation of linear elastic materials undergoing large strains at real time.

Chapter 4

Simulation of Surgical cutting with coupled POD and X-FEM techniques

4.1 Introduction

In the previous chapters we mentioned the challenges of real-time simulation of soft tissues. One of the sources of complexity is due to the highly non-linear behavior of soft living tissues that were investigated in chapters 2 and 3 as well as the highly demanding feedback rates imposed by the simulators (25 Hz for visual feedback and some 500 Hz if haptic feedback is to be added to the system). The third source of complexity comes from the multi-physics nature of the phenomena occurring in the actual surgery procedure: non-linear elasticity, contact, cutting, temperature, etc. While the thermal dependence of the problem is often —if not always—neglected, contact detection and cutting simulation are of utmost importance for a convincing result in terms of both visual and haptic perception.

Numerical simulation of contact mechanics is a well-known field in the computational mechanics community (see, for instance the recent book by Wriggers (2002) to acquire an overall impression of the difficulty of the topic). An accurate simulation of the process of contact between surgical tools and organs, and between organs themselves, seems to be at this moment out of reach under real-time requirements. Some simplified algorithms, however, provide acceptable results at high feedback rates, see for instance Barbi**c** and James (2007). This algorithm supports distributed contact detection between both rigid and reduced deformable models at 1KHz. The surgical tool

can be modeled as a rigid body and the soft tissue is simulated as reduced deformable model. They exploit low-dimensional deformable models using *mass-Principal Component Analysis* for gaining such high rates. And this shows that model order reduction can be a valuable tool in these kind of simulations.

As mentioned before, cutting simulation is another important source of difficulties for real-time modeling of surgery. This is so as it is necessary to modify the geometry and/or the topology of the domain and its associated mesh, and this needs to be done without penalizing the computation times of the integration of the equations of motion. A vast corps of literature has been devoted to this end. See, for instance, Courtecuisse et al. (2010); Lee et al. (2010); Meier et al. (2005); Zhang et al. (2009). All these approaches share the same spirit. All of these works propose more or less sophisticated algorithms that allow to remove finite elements from an existing mesh to create a cut or even to remove a whole part of the organ, if it is being ablated.

The X-FEM is an appealing method as it can model discontinuities within an FEM mesh without any kind of remeshing. It can be applied to complex material behaviors. This method grew out of research into meshfree methods and it has been applied by researches in computational and applied mechanics communities. The basic philosophy of the X-FEM is that features of interest in a problem, for example crack surfaces, phase boundaries, and fluid-structure interfaces, can be represented independently of the finite element mesh. As a result, simulating the evolution of these features is greatly facilitated. This is particularly true when they exhibit changes in topology of the mesh. The finite element mesh needs not to explicitly "fit" these features with the X-FEM, circumventing the need to remesh in many cases and facilitating adaptivity in others. The basic ideas behind the method are easy to understand. Most finite element approximations to the field of interest (for instance displacement or temperature) can be expressed as a linear combination of nodal shape functions. These shape functions are only able to represent discontinuities in these fields if the mesh is constructed in a particular way. For example, the classical approach to representing the jump in displacement field across a crack front is to explicitly mesh both crack faces. With the X-FEM, the classical mesh need only overlap the geometry of the crack front and does not need to be carefully aligned with it. The linear combination is then augmented with enrichment functions that capture the jump in displacement field across the crack. Crack growth can in turn be simulated through the identification of additional enriched nodes and a new construction for the enrichment function, a process that is typically much simpler than remeshing. Although the X-FEM was originally designed for linear elastic fracture mechanics, it has since been adapted to a wide range of applications.

These include the representation of complex microstructures, multi-phase flow, virtual surgery, and general fluid-structure interaction problems.

The application of X-FEM to cracks began with Belytschko and Black (1999), where they applied the partition of unity methods (see for instance Melenk and Babuska (1996)) to the problem of using finite elements with discontinuous basis functions. Moës et al. (1999) used X-FEM to create a technique for simulating crack propagation in two dimensions without remeshing the domain. The extension to three dimensions was begun by Sukumar et al. (2000a), where they used the two dimensional enrichment functions for planar cracks, and then extended in Areias and Belytschko (2005).

However, the X-FEM approach carries technical challenges: assembling the stiffness matrix requires integration of singular/discontinuous functions and implementing enrichment requires resolving material connectivity. Integrating the gradients of the X-FEM basis functions is difficult because of the singularities and discontinuities. One approach is (see, for instance, Stazi et al. (2003)) is to perform a Delaunay triangulation on the cut triangle that incorporates the crack edges, and then to use Gauss quadrature on each of the resulting triangles. This triangulation does not provide additional degrees of freedom; it is only used for integration of the basis functions. However, this approach can be difficult to implement when the geometry of the crack is complicated by branching or multiple cracks, and is generally impossible in three dimensions without introducing new vertices. Another approach is to use higher order Gauss quadrature (see Strouboulis et al. (2000)).

The use of higher order Gauss quadratures for computing the integrals, without paying attention to the position of the discontinuity has been analyzed, but it has been found that the Gibbs effect (oscillation of higher-order polynomials) greatly affects the accuracy of the results, thus being impractical.

If we resort ourselves to the field of surgical simulation, Jerabkova (2007) and Ono et al. (2009) seem to have been the first to use X-FEM for interactive cutting of deformable objects.

Although we should somehow use model order reduction techniques to obtain nearly real-time rates the use of global, Ritz, shape functions imposes additional limitations. For instance, the use of X-FEM techniques is not straightforward in this context. In this work a new method is presented that combines the features of existing methods based upon model order reduction and the ease of creating cuts and discontinuities without remeshing of the X-FEM technique. As will be noticed, the proposed technique has its origins in the s-FEM techniques by Fish (1992) and the multiscale FEM by Rank and Krause (1997).

The chapter is organized as follows. In section 4.2 we briefly review the X-FEM in the framework of standard finite elements, prior to its introduction in conjunction with reduced basis. In section 4.3 a new technique is developed that can efficiently create cuts and discontinuities in the reduced model without remeshing and with minor modifications to the problem's stiffness matrix. Finally, in Section 4.5 an analysis of its performance is done in an application to refractive surgery of the cornea.

4.2 A brief review of the eXtended Finite Element Method (X-FEM)

In order to make this work self-contained a brief review of the well-known X-FEM is included here. The interested reader is referred to Sukumar et al. (2000b), for instance, among other classical references in the field.

The basic ingredient of X-FEM methods is to consider a cut or crack Γ_d as a discontinuity in the displacement field. Therefore, by simply enriching those nodes whose shape function's support intersects the crack with a discontinuous function will suffice to obtain a conforming discretization without the need of remeshing. Thus, the new approximation of the displacement will be, in the context of X-FEM, for a single crack or cut,

$$\boldsymbol{u}^h(\boldsymbol{x}) = \sum_{i\in I} \boldsymbol{u}_i N_i(\boldsymbol{x}) + \sum_{j\in J} \boldsymbol{b}_j N_j(\boldsymbol{x}) H(\boldsymbol{x}), \tag{4.1}$$

where $J = \{j \in I : \omega_j \cap D \neq \emptyset\}$, I represents the set of all nodes in the FE mesh, $N_i(\boldsymbol{x})$ represents the i-th node shape function evaluated at $\boldsymbol{x}$, $\boldsymbol{u}_i$ represents node i's displacement vector, and $H(\boldsymbol{x})$ is discontinuous across the crack Γ_d, whose geometry is represented by D. Finally, $\boldsymbol{b}_j$ represents a new set of enriched degrees of freedom that control the magnitude of the displacement discontinuity in across the crack. Fig. 4.1 depicts the shape functions of X-FEM in one dimension. In Fig. 4.2 they are represented for two-dimensional cases and bilinear quadrilateral elements.

Classical works of X-FEM in the context of fracture mechanics such as Sukumar et al. (2000b), still add some more degrees of freedom in order to reproduce the LEFM asymptotic solution at the crack front. These are not considered here for several reasons. Firstly, for living soft tissues the form of the asymptotic solution (if any) at the crack front is in general not known. In any case, the gain in accuracy would be, in general, negligible, if compared with all the simplifying assumptions made up to now in order to reach real-time performance.

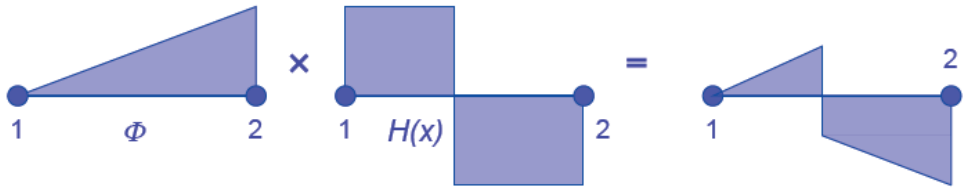


Figure 4.1. Shape functions of X-FEM. Conceptual development.

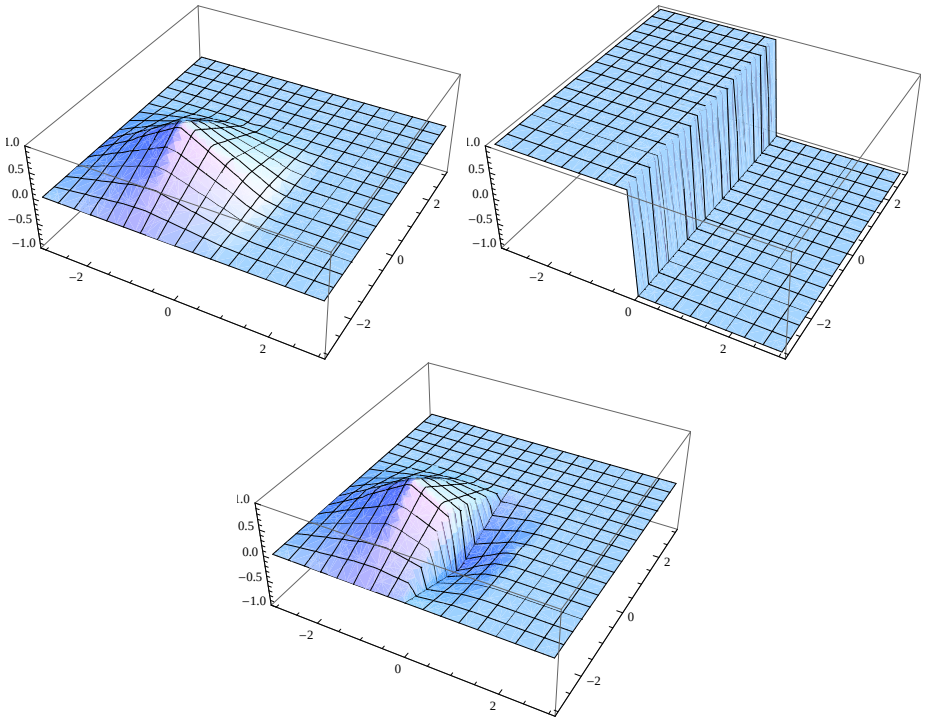


Figure 4.2. Shape functions of X-FEM in two dimensions. Top left, standard FE shape function associated to a given node. Top right, the Heaviside function. Bottom, the resulting discontinuous shape function.

4.3 Coupling reduced models and X-FEM descriptions of surgical cuts

The approach here suggested is composed by the coupling between the (Ritz) reduced basis and a superimposed patch of finite elements where the force exerted by the scalpel exceeds some limit value, thus appearing a cut in the organ. This approach can be seen as a generalization of some existing techniques in the field of finite elements, notably

the so-called s-version of the FEM by Fish (1992) and the multiscale FEM by Rank and Krause (1997), among others (Ammar et al., 2009). The basic of the method is easily understood from Fig. 4.3.

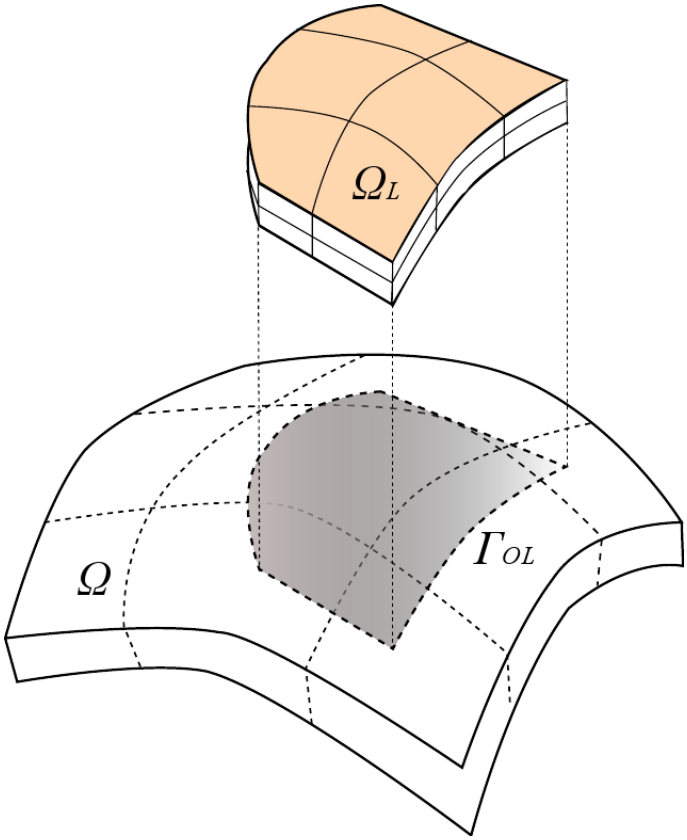


Figure 4.3. Basics of the method for reduced basis-X-FEM coupling.

The domain Ω is discretized by finite elements and a reduced-order model is constructed on top of it that employs Ritz basis functions, as explained before. Once the force exerted by the scalpel reaches a prescribed threshold (see the sections that follow) a cut is supposed to appear at the contact location. Around that region, a superimposed patch Ω_L of finite elements is placed around the scalpel tip. In Ω_L the displacement field is approximated as

$$\boldsymbol{u} = \boldsymbol{u}^0 + \boldsymbol{u}^L \text{ on } \Omega_L \tag{4.2}$$

where, to guarantee the compatibility of the displacement field, we enforce

$$\boldsymbol{u}^L = \boldsymbol{0} \text{ on } \Gamma_{0L}. \tag{4.3}$$

The reduced order model approximates the displacement everywhere, in the spirit of

Eq. (2.10),

$$\boldsymbol{u}^0(\boldsymbol{x}) = \sum_{i=1}^{i=r} \zeta_i \boldsymbol{\phi}_i(\boldsymbol{x}), \tag{4.4}$$

whereas in Ω_L a traditional fem mesh, including discontinuous enrichment, is added:

$$\boldsymbol{u}^L(\boldsymbol{x}) = \sum_{j \in \Omega_L} \boldsymbol{d}_j N_j(\boldsymbol{x}) + \sum_{k \in J} \boldsymbol{b}_k N_k(\boldsymbol{x}) H(\boldsymbol{x}). \tag{4.5}$$

The discrete equilibrium equations can be obtained after the weak form of the problem, namely, *find* $\boldsymbol{u} \in \boldsymbol{U} = \{\boldsymbol{u} : \boldsymbol{u} \in \mathcal{C}^0, \boldsymbol{u} = \overline{\boldsymbol{u}} \text{ on } \Gamma_u\}$ *such that*

$$\int_\Omega \boldsymbol{\varepsilon}^* : \boldsymbol{\sigma} d\Omega = \int_{\Gamma_t} \boldsymbol{u}^* \boldsymbol{t} d\Gamma + \int_\Omega \boldsymbol{u}^* \boldsymbol{b} d\Omega,$$
$$\forall \boldsymbol{u}^* \in \boldsymbol{U}_0 = \{\boldsymbol{u}^* : \boldsymbol{u}^* \in \mathcal{C}^0, \boldsymbol{u} = \boldsymbol{0} \text{ on } \Gamma_u\}, \tag{4.6}$$

where, as usual, Γ_u represents the Dirichlet (essential) part of the boundary of the domain, $\overline{\boldsymbol{u}}$ represents the value of the prescribed displacement at that location, $\boldsymbol{t}$ the prescribed vector of traction at the boundary, and, finally, Γ_t the Neumann (natural) part of the boundary Γ. Admissible variations of the displacement are calculated as

$$\boldsymbol{u}^* = \left(\boldsymbol{u}^0\right)^* + \left(\boldsymbol{u}^L\right)^*. \tag{4.7}$$

By substituting the displacement field approximations (4.4) and (4.5) into the weak form (4.6), and after invoking the arbitrariness of admissible displacements $\boldsymbol{u}^*$, we arrive at a discrete system of equations of the type

$$\begin{pmatrix} \boldsymbol{K}^{\zeta\zeta} & \boldsymbol{K}^{\zeta d} & \boldsymbol{K}^{\zeta b} \\ \boldsymbol{K}^{d\zeta} & \boldsymbol{K}^{dd} & \boldsymbol{K}^{db} \\ \boldsymbol{K}^{b\zeta} & \boldsymbol{K}^{bd} & \boldsymbol{K}^{bb} \end{pmatrix} \begin{pmatrix} \boldsymbol{\zeta} \\ \boldsymbol{d} \\ \boldsymbol{b} \end{pmatrix} = \begin{pmatrix} f^{\zeta} \\ f^{\boldsymbol{d}} \\ f^{\boldsymbol{b}} \end{pmatrix}, \tag{4.8}$$

where the stiffness matrices $\boldsymbol{K}^{\alpha\beta}$ are given by

$$\boldsymbol{K}_{IJ}^{\alpha\beta} = \int_{\Omega_i} (\boldsymbol{B}_I^\alpha)^T \boldsymbol{D} \boldsymbol{B}_J^\beta d\Omega \tag{4.9}$$

where matrix $\boldsymbol{B}_I^\alpha$, $\alpha = \zeta, d, b$, represent the standard shape function derivative matrix for any of the three types of approximation functions here considered, namely, the Ritz functions $\phi(\boldsymbol{x})$, the standard finite element shape functions $N(\boldsymbol{x})$ or the discontinuous enrichment shape functions, $N(\boldsymbol{x})H(\boldsymbol{x})$. Ω_i represents the domain of integration, either Ω or Ω_L, respectively. Finally, $\boldsymbol{D}$ represents the consistently linearized constitutive tensor. As mentioned before, due to the severe restrictions placed by real-time requirements in haptic environments, it is not possible, in general, to update this tangent

stiffness matrix. In this work, for the sake of simplicity in the exposition of the method, an approach similar to that applied in chapter 2 is employed, in which no updating is accounted for. This leads, of course, to higher errors in the results, still acceptable in virtual surgery environments. If more accurate results are needed that formulation that was presented in chapter 3 can be used. That method employs asymptotic expansions of the displacement field and, without any updating, allows to closely follow complex non-linear force-displacement paths.

In order to evaluate the previous integrals leading to the stiffness matrix of the problem, no special integration procedure is employed, in sharp contrast with the original works by Fish (1992), since we assume that the superimposed mesh conforms with the existing one. It is necessary, however, to perform some form of tailored numerical integration in those elements enriched by discontinuous displacements.

4.3.1 Simplified physics of the cutting procedure

The last ingredient in the method is related to the placement of the patch Ω_L during the surgery. Once contact between the scalpel and the organ has been detected by a suitable contact algorithm (see Barbi**c** and James (2007) for instance, for a valid contact criterion in reduced model settings), a criterion must be set in order to determine if cutting appears, thus generating a new boundary in the domain, or not. We follow closely the criteria set in Bielser and Gross (2000). Although greatly simplified, these criteria have demonstrated to provide realistic enough results in haptic environments. A scalpel cuts along its blade, so a decomposition of the acting force as in Fig. 4.4 is employed:

$$\boldsymbol{F}_{\text{ext}} = \boldsymbol{F}_{\perp} + \boldsymbol{F}_{\parallel} = \boldsymbol{F}_{\perp} + \boldsymbol{F}_{a} + \boldsymbol{F}_{n} \tag{4.10}$$

A threshold value of the force F_{cut} is considered such that lower modulus of the force $\boldsymbol{F}_{\parallel}$ produce static friction, but no cut. Once $\|\boldsymbol{F}_{\parallel}\|$ exceeds F_{cut}, the finite element patch Ω_L is added to the scalpel tip.

In order to simplify the process and to make it simpler and (notably) faster, once the threshold value F_{cut} is reached, a whole finite element is then cut. No cut of length smaller the the typical element size h is considered. If the finite element mesh is dense enough, this limitation does not very much affect the results. Remember that the size of the global finite element mesh, r, does not affect the size of the reduced model, N, see Eq. (2.12).

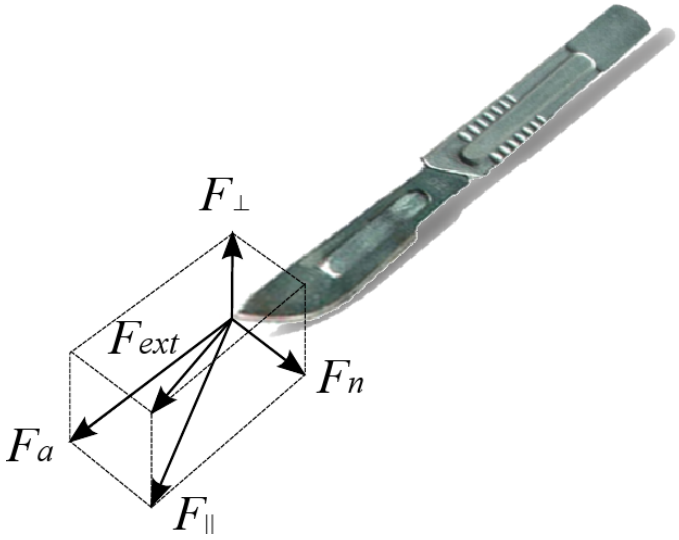


Figure 4.4. Force decomposition at the scalpel point of contact.

4.4 Finite element model of the cornea

We have used the same cornea mesh as in previous chapters. The cornea was assumed to be, without loss of generality, hyperelastic, with a Kirchhoff-Saint Venant behavior. More sophisticated material behaviors can also be efficiently tackled with this technique, as in chapter 2, where a two-families of fibers reinforced hyperelasticity model was successfully employed. We have applied nine concentrated forces on the right and left half of the cornea with about 1.5 mm eccentricity from the center of the cornea.

For a punctual load of increasing value the complete model gave the five modes that capture 99% of the energy of the model. These modes are shown in Fig. 3.13. Therefore, the reduced-order model has a stiffness matrix of 18×18 (six modes for each component of the displacement vector).

With these six modes, standard reduced order models provide an error under 20% for loads placed at different positions to the one considered in the evaluation of the modes in Fig. 4.5. This level of error is considered valid in many real-time applications as it was explained in chapter 2. If more accurate simulations are needed, the method proposed in chapter 3 provides a nearly exact simulation for displacement under the scalpel tip up to 1 mm (possibly much more than that, but at the scale of the cornea, 1 mm displacement involves very large strains and no reference finite elements results could be obtained to compare with).

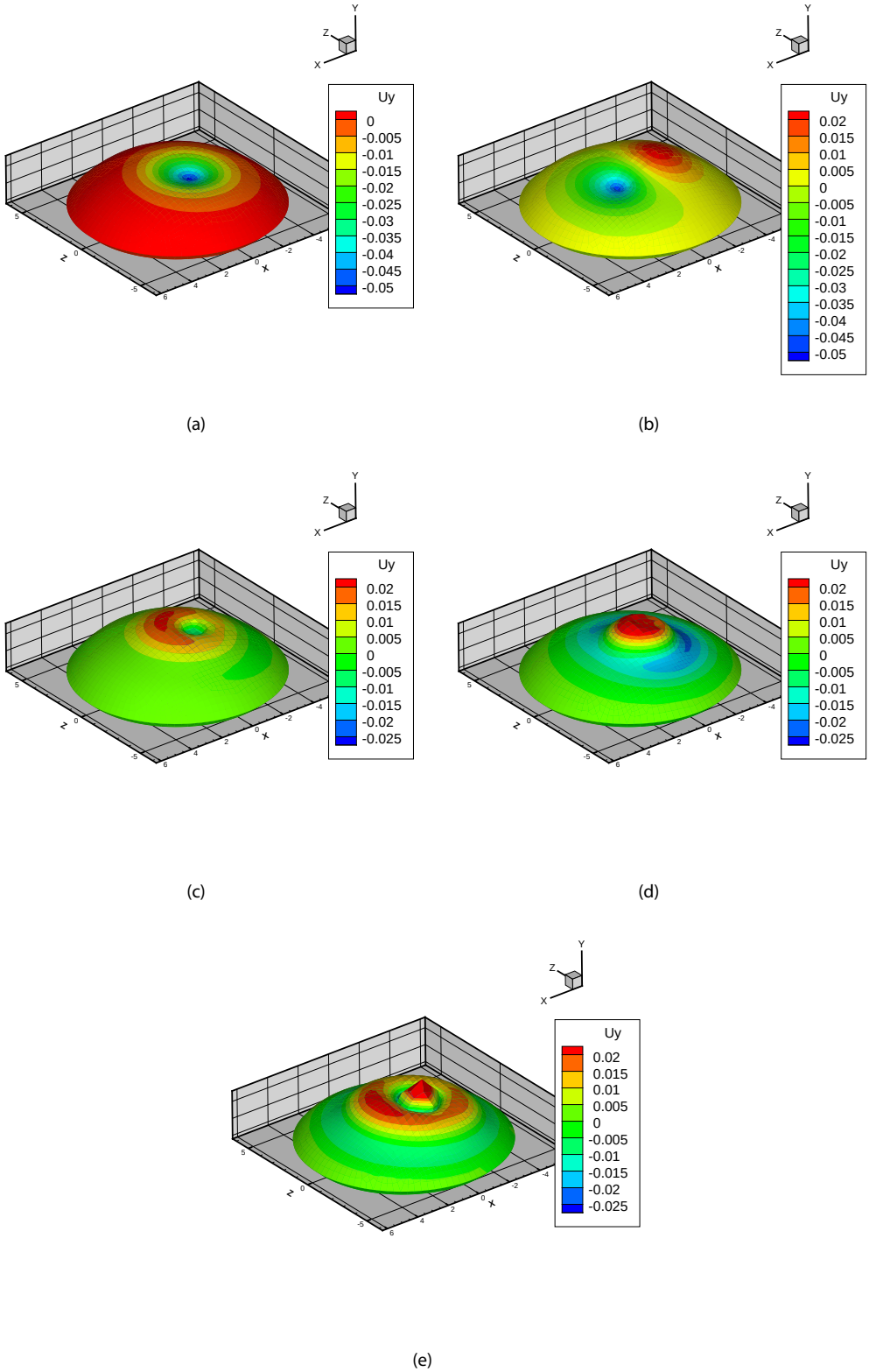


Figure 4.5. Five most important eigenmodes for the simulation of the cornea. The corresponding eigenvalues are: 10.77 (a), 0.014 (b), $2.58 \cdot 10^{-4}$ (c), $6.53 \cdot 10^{-7}$ (d), $4.43 \cdot 10^{-11}$ (e).

4.5 Application to the simulation of corneal refractive surgery

It must be highlighted at this point that the purpose of this work is not to validate the use of reduced-order models in the context of real-time applications. This has been done in the previous chapters and also in the published papers, see for instance Niroomandi et al. (2008), Niroomandi et al. (2010b), Niroomandi et al. (2010a) or Barbi**c** and James (2005), among others. Nor the validation of X-FEM technique to accurately simulate moving discontinuities in solid mechanics problems, for instance (see, among others, Sukumar et al. (2000b), Combescure et al. (2008), Meschke and Dumstorff (2007), Daux et al. (2000) or Stolarska et al. (2001)). However, the combination of both provides a unique set of features that makes the resulting technique an appealing choice. Among these features, the feasibility of real-time simulation of complex non-linear tissues (including, for instance, fiber-reinforced hyperelastic tissue, see Niroomandi et al. (2008)) combined with the possibility of simulating surgical cutting without remeshing (at haptic feedback rates, it is obvious that for realistic rendering the cut must be represented accurately, but only at some 30 Hz feedback rates not at 500 Hz). To this end the case of corneal refractive surgery is analyzed here under the framework of the proposed method.

Astigmatism is a refractive error due to the non-spherical shape of the cornea, that is, the refractive power is not uniform in all meridians. Refractive surgery techniques are used to modify the curvature in order to repair the refractive error of the eye (Alastrué et al., 2006). This defect may be corrected by making the cornea as spherical as possible, through the application of some cuts.

In addition, properties of the cornea can be quite different between patients with the same level of pathology, therefore the technique presented before can be seen as an efficient means to plan a patient-specific surgery that minimizes uncertainty in the results, providing the surgeon with the sensations and results he will obtain later in the real surgery.

The cornea was assumed to be, without loss of generality, hyperelastic, with a Saint Venant-Kirchhoff behavior. More sophisticated material behaviors can also be efficiently tackled with this technique, as in Niroomandi et al. (2008), where a two-families of fibers reinforced hyperelasticity model was successfully employed.

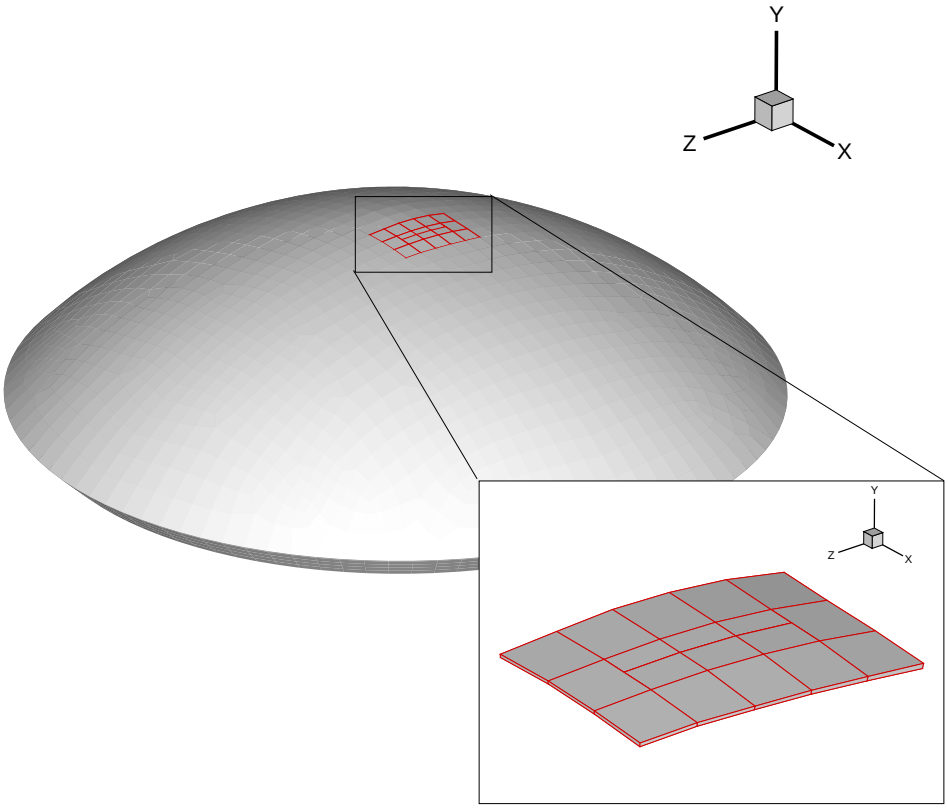


Figure 4.6. Superimposed X-FEM mesh at the position of limbal relaxing incision.

4.5.1 Simulating limbal relaxing incisions

Limbal relaxing incisions (LRI) are one f the three main types of cuts made in corneal refractive surgery. These incisions are made near the outer edge of the iris, and are used to correct minor astigmatism (typically less than 2 diopters). In this case, a cut is made roughly at this position. In Fig. 4.6 a detailed view of the deformed cornea, together with the superimposed X-FEM mesh is presented. For post-processing purposes, those elements cut by the scalpel are represented as two different finite elements in the figure, although no such elements exist in the simulation, as explained before.

The obtained displacement field is shown in Fig. 4.7. A detailed view of the x-direction displacement field is shown in Fig. 4.8, where the appearance of the cut can be noticed, in spite of the low magnitude of the displacement between crack lips.

In order to see the difference that a practitioner would see when dealing with the simulator, in Fig. 4.9 a comparison is made between the solutions obtained by employing the complete model, explicitly meshing the cut by separate hexahedral finite

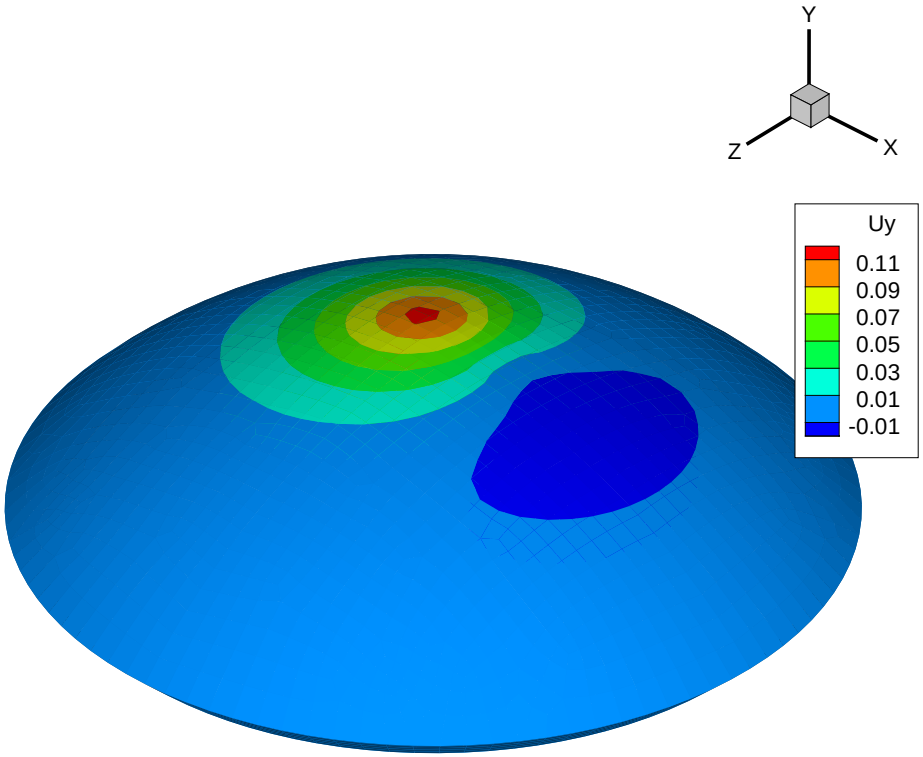


Figure 4.7. Displacement field (y-direction) in the cutting procedure.

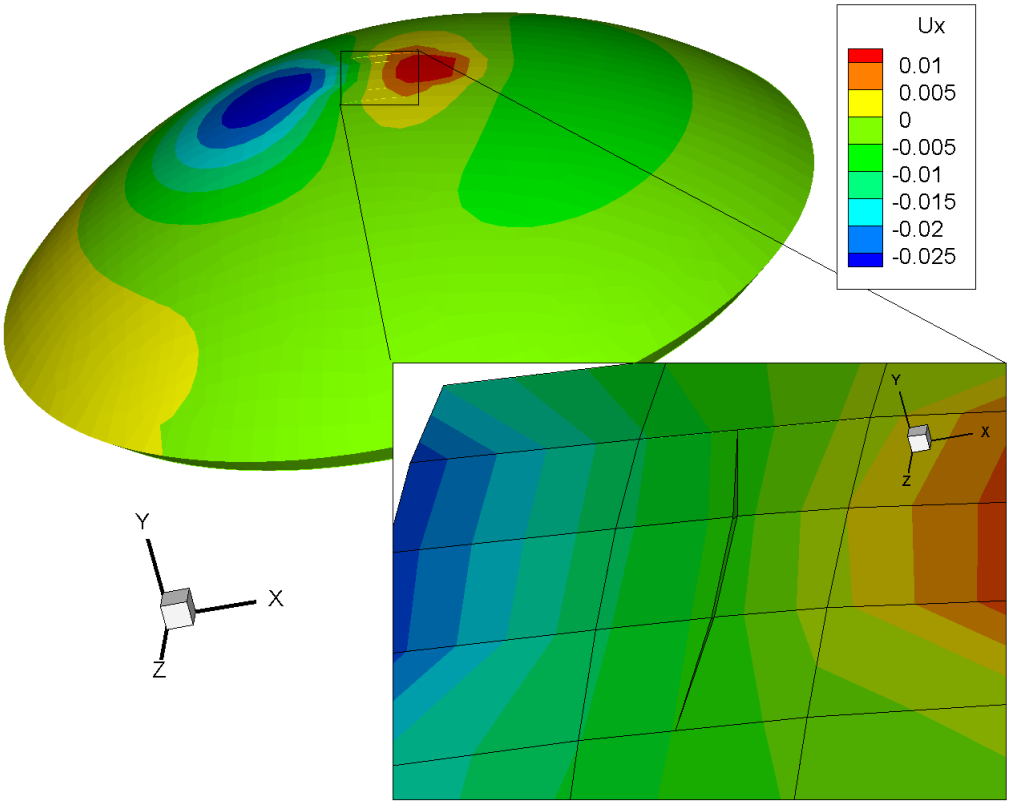


Figure 4.8. Cutting procedure. Displacement field (x-direction). Enrichment degrees of freedom are magnified 10 times to highlight the magnitude of the cut.

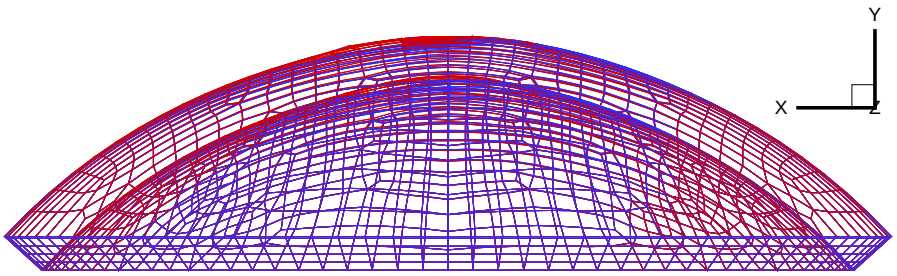


Figure 4.9. Cutting procedure. Comparison among the complete (red) and reduced (blue) models.

elements, and that obtained by the combined POD-X-FEM model. It is noticeable the practically indistinguishable displacement field that is obtained.

4.6 Conclusions

In this chapter a novel method for the real-time simulation of surgery in haptic environments has been presented. The method is based on the use of reduced order modeling. Reduced models are, to the author's knowledge, the only technique able to simulate at real-time feedback rates, highly complex constitutive models for living tissues (fibre-reinforced hyperelastic models, for instance). The problem with such an approach lies in the introduction of surgical cuts. The global character of the approximation functions (Ritz) precludes the possibility of employing standard methods in the literature that use very efficient algorithms for partitioning or eliminating the mesh in the neighboring of the cut. The approach here developed is based on the use of X-FEM techniques, coupled with the existing reduced model through a multi-scale-like method. Thus, the superimposed finite element mesh is capable of reproducing the displacement discontinuities produced by the scalpel, while the underlying reduced model is able to accurately reproduce the global behavior of the organ.

The proposed method runs at feedback rate, thus allowing to take part in a surgery simulator, together with some specialized contact detection algorithm. This opens the possibility to incorporate complex, state-of-the-art soft tissue constitutive laws into real-time simulation of surgery, thus leading the possibility of making a true second-

generation simulator at hand.

Chapter 5

Conclusion

In this thesis a novel strategy has been presented for real-time interactive simulation of non-linear anisotropic tissues. The presented technique is based on model reduction techniques and, unlike previous works, it allows for the consideration of both geometrical and material non-linearities.

The reduced models are constructed by employing a set of "high quality" global basis functions (as opposed to general-purpose, locally supported FE shape functions) in a Galerkin framework. These functions are constructed after some direct simulations of the organs to be analyzed. These simulations are made off-line and the computed displacements and tangent stiffness matrices are stored in memory prior to beginning with the on-line, real-time simulation.

The obtained results show good accordance with complete model results, and run at frequencies of around 400-500 Hz, enough for real-time requirements, even for this very rude code prototypes.

A technique for the construction of reduced models of hyperelastic solids has also been studied that is based on the use of model reduction techniques and an asymptotic numerical method. The technique, initially developed for geometrically non-linear solids was generalized here for hyperelastic materials involving material non-linearities. This allows for a simulation of complex, non-linear tissues without the need for typical (Newton-Raphson) tangent matrix updates, impossible to be done under real-time constraints.

These two ingredients render a method with very attractive features, such as the use of very few degrees of freedom (up to ten in the examples showed in this chapter) and the necessity to compute only one tangent stiffness matrix within the region covered by the radius of convergence of the asymptotic expansion. This radius of convergence is often relatively large (on the order of 10% of the total dimensions of the solid for

the examples showed herein). Within this radius of convergence neither iterative procedures (typical of modified Newton methods) nor the update of the stiffness matrix (typical in standard Newton-Raphson methods) are required.

The examples developed in this work showed that between six to ten eigenmodes are enough for the vast majority of the examples considered, and usually an expansion of order 4-6 renders very acceptable results, with excellent accuracy within a vast radius of convergence.

Then the work was focused on various ways of interpolating the reduced basis and also non-linear forces in order to obtain near real-time rates. Real-time feedback rates were obtained for visual perception using Matlab and very rude code prototypes. In the present approach different sets of basis functions are obtained for different load (and possibly boundary conditions) states that can be changed according with the user experience. Basis sets for any state different from the precomputed ones are obtained after interpolation, as explained in the previous sections. Therefore, it is crucial for the method to work well that the set of complete models to be solved to feed the algorithm is chosen adequately. For the moment, the brute-force approach seems to be out of reach. In our opinion, this process should be guided by surgeons, by indicating the most probable loading states for each surgical procedure, distinguishing, for instance, between experienced surgeons and medicine students.

The method has been developed for Saint Venant-Kirchhoff constitutive laws. This kind of laws is among the best state-of-the-art existing models in real-time applications but is judged to be poor for some applications. So it was found interesting to extend the proposed technique to some more sophisticated constitutive laws, such as neo-hookean laws, for instance, or other much better suited for specific applications. The work developed in this thesis demonstrated that the method can be further generalizable to other constitutive, hyperelastic models.

Finally, a novel method for the real-time simulation of surgery cutting in haptic environments has been presented. The method is based on the use of reduced order modeling. Reduced models are, to the author's knowledge, the only technique able to simulate at real-time feedback rates, highly complex constitutive models for living tissues. The problem with such an approach lies in the introduction of surgical cuts. The global character of the approximation functions (Ritz) precludes the possibility of employing standard methods in the literature that use very efficient algorithms for partitioning or eliminating the mesh in the neighboring of the cut. The approach here developed is based on the use of X-FEM techniques, coupled with the existing reduced model through a multi-scale-like method. Thus, the superimposed finite element mesh is ca-

pable of reproducing the displacement discontinuities produced by the scalpel, while the underlying reduced model is able to accurately reproduce the global behavior of the organ.

The proposed method runs at feedback rate, thus allowing to take part in a surgery simulator, together with some specialized contact detection algorithm. This opens the possibility to incorporate complex, state-of-the-art soft tissue constitutive laws into real-time simulation of surgery, thus leading the possibility of making a true second-generation simulator at hand.

5.1 Thesis contributions

Among the main contributions of the work done in this thesis, the following ones can be cited:

- Model Order Reduction with application to real time simulation of soft tissues has been presented in this thesis and it has been shown that it is a promising tool in this field.
- A new extension of Proper Orthogonal Decomposition and Asymptotic Numerical Method for hyperelastic materials has been obtained. The importance of this extension is that Saint Venant-Kirchhoff material is not appropriate, in general, for predicting the behavior of soft tissues. This is so, since it presents some instabilities when subjected to compressions that could eventually make it not suitable for surgery simulation.
- A new methodology for including discontinuities in the displacement field produced by surgical cutting in combination with reduced models has been proposed. The proposed method has been successfully used in interactive simulation of cutting procedure.

5.2 Future works

This thesis opens the door to new promising research lines in the group. Among the topics that have not been yet studied the following ones can be included in the real-time simulation of soft tissues:

- The POD-ANM technique could be extended to other constitutive laws that include fiber reinforcement of soft tissues. In principle there is no limitation for this generalizations, although it has not been done yet.
- The models that have been presented are sufficient for acquiring basic surgical skills but are not realistic enough to represent fully the complexity of the human anatomy and physiology. Notably, respiratory motion, bleeding and contact between organs could drastically improve the perception of the trainee. The technique here developed does not impose any limitations for the inclusion of these aspects, provided that very fast contact detection algorithms exist for reduced models.
- In this thesis Matlab has been used for the simulations. It is well known that although such a high level programming language is suitable for evaluation of the methods but it is not efficient in terms of computation time. GPU-based programming, for instance, could help in decreasing the simulation time and the results of this in other fields have been very promising.
- Recently, the so-called Proper Generalized Decomposition (PGD) method has been developed (Ammar et al., 2006), (Ammar et al., 2007), (Chinesta et al., 2010). This technique generalizes the POD method and opens promising new ways to model order reduction of the dynamics of non-linear materials such as soft tissues.

5.3 Publications

As a result of the thesis several articles have been published in journals and proceedings of national and international conferences.

Journal Publications

1. Real time simulation of surgery by model reduction and X-FEM techniques. S. Niroomandi, I. Alfaro, E. Cueto, F. Chinesta. International Journal for Numerical Methods in Biomedical Engineering, submitted, 2011.
2. Accounting for large deformations in real-time simulations of soft tissues based on reduced-order models. S. Niroomandi, I. Alfaro, E. Cueto, F. Chinesta. Computer Methods and Programs in Biomedicine, 2010, accepted. Early view.

3. Model order reduction for hyperelastic materials. S. Niroomandi, I. Alfaro, E. Cueto, F. Chinesta, International Journal for Numerical Method in Engineering, 81(9), 1180-1206, 2010.

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

Conference Papers

1. Real time simulation of surgery by combined reduced order models and X-FEM technique. S. Niroomandi, I. Alfaro, E. Cueto, F. Chinesta, IV International Conference on Computational Methods for Coupled Problems in Science and Engineering (COUPLED PROBLEMS 2011), 2011.

2. Towards the third generation of surgical simulators. S. Niroomandi, I. Alfaro, E. Cueto, F. Chinesta, IV European conference on computational mechanics- ECCM 2010, 2010.

3. Real-time simulation of living tissues by model reduction . E. Cueto, S. Niroomandi, I. Alfaro, F. Chinesta, 9th World Congress on Computational Mechanics and 4th Asian Pacific Congress on Computational Mechanics (WCCM/APCOM 2010) , 2010.

4. Simulación en tiempo real de tejidos blandos mediante reducción de modelos y métodos asintóticos numéricos. S. Niroomandi, I. Alfaro, E. Cueto, F. Chinesta, Congreso de Métodos Numéricos en Ingeniería (METNUM 2009), 2009.

5. Real-time deformable models of non-linear tissues by model reduction techniques. S. Niroomandi, I. Alfaro, E. Cueto, F. Chinesta, 8th International Sympsium on Computer Methods in Biomechanics and Biomedical Engineering , 2008.

6. On the application of model reduction techniques to real-time simulation of non-linear tissues. S. Niroomandi, I. Alfaro, E. Cueto, F. Chinesta, 4th International Symposium, ISBMS 2008, London, 2008.

7. Real-time simulation of non-linear tissues by model reduction techniques. S. Niroomandi, I. Alfaro, E. Cueto, F. Chinesta, XXI International Congress of Theoretical and Applied Mechanics,ICTAM 2008, 2008.

Appendix

Appendix A

Formulation of the Asymptotic Numerical Method for compressible neo-Hookean materials

A.1 Formulation of neo-Hookean model

The compressible neo-Hookean model is characterized by a strain energy function given by

$$\Psi = \frac{\mu}{2}(\text{tr}(\boldsymbol{C}) - 3) - \mu \ln J + \frac{\lambda}{2}(\ln J)^2 \tag{A.1}$$

where λ and μ are Lamé's constants and $\boldsymbol{C} = \boldsymbol{I} + 2\boldsymbol{\gamma}$ is the right Cauchy-Green strain tensor.

As mentioned in the body of this thesis,

$$\begin{aligned} &\boldsymbol{\gamma} = \boldsymbol{\gamma}_l(\boldsymbol{u}) + \boldsymbol{\gamma}_{nl}(\boldsymbol{u}, \boldsymbol{u}) \\ &\boldsymbol{\gamma}_l(\boldsymbol{u}) = \frac{\boldsymbol{\nabla u} + \boldsymbol{\nabla}^T \boldsymbol{u}}{2},\ \boldsymbol{\gamma}_{nl}(\boldsymbol{u}, \boldsymbol{u}) = \frac{\boldsymbol{\nabla u} \boldsymbol{\nabla}^T \boldsymbol{u}}{2} \end{aligned} \tag{A.2}$$

The second Piola-Kirchhoff stress tensor can be obtained by

$$\boldsymbol{S} = \frac{\partial \Psi(\boldsymbol{\gamma})}{\partial \boldsymbol{\gamma}} = \mu(\boldsymbol{I} - \boldsymbol{C}^{-1}) + \lambda(\ln J)\boldsymbol{C}^{-1}. \tag{A.3}$$

In the formulation developed in Cochelin et al. (1994b), the second tensor of Piola-Kirchhoff $\boldsymbol{S}$, displacement $\boldsymbol{u}$ and load factor $\overline{\lambda}$ were expanded around a point at which

the solution of the problem is known. Here, for compressible neo-Hookean material, some additional variables have been introduced in order to make the derivation less tedious. In the end the value of these additional variables will be calculated in terms of stress and displacement. These additional variables are

$$X = J^2 \tag{A.4}$$

$$Y = \ln J = \ln \sqrt{J^2} \text{ as } J > 0 \tag{A.5}$$

$$Z = \frac{1}{J^2} \tag{A.6}$$

Introducing these new variables in Eq. A.3 leads to

$$\boldsymbol{S} = \mu \boldsymbol{I} + (\lambda Y - \mu)\boldsymbol{C}^{-1} \tag{A.7}$$

A.2 Preliminary formulae

A.2.1 Expansion series

As stated in chapter 3 all variables are expanded asymptotically in terms of a control parameter "a". This expansion is done by using Taylor series. The Taylor series for a smooth function $\boldsymbol{f}(a)$ around the point $a = 0$ is as follows:

$$\boldsymbol{f}(a) = \boldsymbol{f}_0 + \sum_{p=1} \boldsymbol{f}_p a^p$$

$$\text{where } \boldsymbol{f}_0 = \boldsymbol{f}(0) \text{ and } \boldsymbol{f}_p = \frac{1}{p!} \frac{d^p \boldsymbol{f}}{da^p}\Big|_{a=0} \tag{A.8}$$

A.2.2 Taylor series for the product of two functions

In deriving the formulation it is needed to calculate the Taylor series of the product of two functions. If $C(u(a)) = A(u(a))B(u(a))$ the series will take the form

$$C_0 + C_1 a + C_2 a^2 + \cdots = (A_0 + A_1 a + A_2 a^2 + \cdots)(B_0 + B_1 a + B_2 a^2 + \cdots) \tag{A.9}$$

The coefficients C_n will be obtained as follows

$$\begin{aligned}
&\text{at order 0} \;\Rightarrow C_0 = A_0 B_0 \\
&\text{at order 1} \;\Rightarrow C_1 = A_0 B_1 + A_1 B_0 \\
&\qquad\qquad \vdots \\
&\text{at order p} \;\Rightarrow C_p = A_0 B_p + A_p B_0 + \sum_{r=1}^{p-1} A_r B_{p-r}
\end{aligned} \tag{A.10}$$

A.2.3 Composition of expanded variables

Now if $A = A(b(a))$ and both A and $b(a)$ are going to be expanded, the chain rule can be applied.

$$\begin{aligned} A &= A_0 + \sum_{p=1} A_p a^p \\ b &= b_0 + \sum_{p=1} b_p a^p \end{aligned} \tag{A.11}$$

where $A_0 = A(b_0),\ b_0 = b(0)$. A_p can be calculated by applying the chain rule for higher derivatives (Arbogast, 1800).

$$\begin{aligned} A_1 &= \frac{\partial A}{\partial a}\Big|_{a=0} = \frac{\partial A}{\partial b}\frac{db}{da}\Big|_{a=0} = \frac{\partial A}{\partial b}\Big|_{b_0} b_1, \text{ where } b_1 = \frac{db}{da}\Big|_{a=0} \\ A_2 &= \frac{1}{2}\frac{\partial^2 A}{\partial a^2}\Big|_{a=0} = \frac{1}{2}\frac{\partial}{\partial a}\left(\frac{\partial A}{\partial b}\frac{db}{da}\right) = \frac{1}{2}\frac{\partial^2 A}{\partial b^2}\Big|_{b_0} b_1^2 + \frac{\partial A}{\partial b}\Big|_{b_0} b_2 \\ &\vdots \end{aligned} \tag{A.12}$$

In the same way the rest of the coefficients can be calculated. In these expressions the following variables are used

$$\begin{aligned} b_1 &= \frac{db}{da}\Big|_{a=0} \\ b_2 &= \frac{1}{2!}\frac{d^2 b}{da^2}\Big|_{a=0} \\ &\vdots \end{aligned}$$

Again some notations will be adopted for simplicity. We now define

$$A_{b_0} = A(b_0),\ A_{b_1} = \frac{\partial A}{\partial b}\Big|_{b_0},\ A_{b_2} = \frac{1}{2!}\frac{\partial^2 A}{\partial b^2}\Big|_{b_0},\ \ldots,\ A_{b_p} = \frac{1}{p!}\frac{\partial^p A}{\partial b^p}\Big|_{b_0}$$

By using this notation the coefficients of A will be simplified as

$$\begin{aligned} A_0 &= A_{b_0} \\ A_1 &= A_{b_1} b_1 \\ A_2 &= A_{b_2} b_1^2 + A_{b_1} b_2 \\ &\vdots \end{aligned} \tag{A.13}$$

A.3 Expansion of $X = J^2$

As $X = J^2 = det(\boldsymbol{C})$, before expanding X it is necessary to expand $\boldsymbol{C}$. For that, firstly the displacement vector is expanded as

$$\boldsymbol{u}(a) = \boldsymbol{u}_0 + \sum_{p=1} \boldsymbol{u}_p a^p, \ \boldsymbol{u}_0 = \boldsymbol{u}(0), \ \boldsymbol{u}_p = \frac{1}{p!}\frac{d^p \boldsymbol{u}}{da^p}\Big|_{a=0} \tag{A.14}$$

Then the linear term of strain can be rewritten as

$$\begin{aligned} \text{at order 0} \quad \boldsymbol{\gamma}_{l_0} &= \frac{\boldsymbol{\nabla}\boldsymbol{u}_0 + \boldsymbol{\nabla}^T\boldsymbol{u}_0}{2} \\ &\vdots \\ \text{at order p} \quad \boldsymbol{\gamma}_{l_p} &= \frac{\boldsymbol{\nabla}\boldsymbol{u}_p + \boldsymbol{\nabla}^T\boldsymbol{u}_p}{2} \end{aligned} \tag{A.15}$$

If FEM is used for discretization of the weak form and the displacements are approximated by $\boldsymbol{u} = \boldsymbol{N}\boldsymbol{v_p}$, one will have

$$\begin{aligned} \boldsymbol{\gamma}_{l_0} &= \boldsymbol{B}_l \boldsymbol{v}_0 \\ &\vdots \\ \boldsymbol{\gamma}_{l_p} &= \boldsymbol{B}_l \boldsymbol{v}_p \end{aligned} \tag{A.16}$$

For the non-linear term of strain, the multiplication of expanded variables is used

$$\begin{aligned} \boldsymbol{\gamma}_{nl_0} &= \frac{\boldsymbol{\nabla}\boldsymbol{u}_0 \boldsymbol{\nabla}\boldsymbol{u}_0^T}{2} \\ &\vdots \\ \boldsymbol{\gamma}_{nl_p} &= \boldsymbol{\nabla}\boldsymbol{u}_0 \boldsymbol{\nabla}^T\boldsymbol{u}_p + \sum_{r=1}^{p-1} \frac{\boldsymbol{\nabla}\boldsymbol{u}_r \boldsymbol{\nabla}^T\boldsymbol{u}_{p-r}}{2} \end{aligned} \tag{A.17}$$

After discretization these terms take the form

$$\begin{aligned} \boldsymbol{\gamma}_{nl_0} &= \frac{1}{2}\tilde{\boldsymbol{A}}(\boldsymbol{v}_0)\boldsymbol{G}\boldsymbol{v}_0 \\ &\vdots \\ \boldsymbol{\gamma}_{nl_p} &= \tilde{\boldsymbol{A}}(\boldsymbol{v}_0)\boldsymbol{G}\boldsymbol{v}_p + \sum_{r=1}^{p-1} \frac{1}{2}\tilde{\boldsymbol{A}}(\boldsymbol{v}_r)\boldsymbol{G}\boldsymbol{v}_{p-r} \end{aligned} \tag{A.18}$$

In which $\tilde{\boldsymbol{A}}$ and $\boldsymbol{G}$ can be found in standard texts of non-linear finite element analysis as in Zienkiewicz and Taylor (1994).

$$\tilde{\boldsymbol{A}} = \begin{bmatrix} \theta_x^T & 0 & 0 \\ 0 & \theta_y^T & 0 \\ 0 & 0 & \theta_z^T \\ \theta_y^T & \theta_x^T & 0 \\ 0 & \theta_z^T & \theta_y^T \\ \theta_z^T & 0 & \theta_x^T \end{bmatrix}$$

in which

$$\theta_x = \frac{\partial u_x}{\partial x},\ \theta_y = \frac{\partial u_y}{\partial y},\ \theta_z = \frac{\partial u_z}{\partial z}$$

and

$$\boldsymbol{G} = \begin{bmatrix} \frac{\partial h_1}{\partial x} & 0 & 0 & \cdots & \frac{\partial h_8}{\partial x} & 0 & 0 \\ 0 & \frac{\partial h_1}{\partial x} & 0 & \cdots & 0 & \frac{\partial h_8}{\partial x} & 0 \\ 0 & 0 & \frac{\partial h_1}{\partial x} & \cdots & 0 & 0 & \frac{\partial h_8}{\partial x} \\ \frac{\partial h_1}{\partial y} & 0 & 0 & \cdots & \frac{\partial h_8}{\partial y} & 0 & 0 \\ 0 & \frac{\partial h_1}{\partial y} & 0 & \cdots & 0 & \frac{\partial h_8}{\partial y} & 0 \\ 0 & 0 & \frac{\partial h_1}{\partial y} & \cdots & 0 & 0 & \frac{\partial h_8}{\partial y} \\ \frac{\partial h_1}{\partial z} & 0 & 0 & \cdots & \frac{\partial h_8}{\partial z} & 0 & 0 \\ 0 & \frac{\partial h_1}{\partial z} & 0 & \cdots & 0 & \frac{\partial h_8}{\partial z} & 0 \\ 0 & 0 & \frac{\partial h_1}{\partial z} & \cdots & 0 & 0 & \frac{\partial h_8}{\partial z} \end{bmatrix}$$

Finally, the right Cauchy-Green strain tensor can then be computed as follows

$$\boldsymbol{C} = \boldsymbol{I} + 2\left(\boldsymbol{\gamma}_l(\boldsymbol{u}) + \boldsymbol{\gamma}_{nl}(\boldsymbol{u}, \boldsymbol{u})\right) \tag{A.19}$$

$$\boldsymbol{C} = \boldsymbol{C}_0 + \sum_{p=1} \boldsymbol{C}_p a^p \tag{A.20}$$

where

$$\begin{aligned} \boldsymbol{C}_0 &= \boldsymbol{I} + \left(2\boldsymbol{B}_l + \tilde{\boldsymbol{A}}(\boldsymbol{v}_0)\boldsymbol{G}\right)\boldsymbol{v}_0 \\ &\vdots \\ \boldsymbol{C}_p &= 2\left(\boldsymbol{B}_l + \tilde{\boldsymbol{A}}(\boldsymbol{v}_0)\boldsymbol{G}\right)\boldsymbol{v}_p + \sum_{r=1}^{p-1} \tilde{\boldsymbol{A}}(\boldsymbol{v}_r)\boldsymbol{G}\boldsymbol{v}_{p-r} \end{aligned} \tag{A.21}$$

Now it is possible to calculate the coefficients in the expansion of X. As stated before

$$X = J^2 = det(\boldsymbol{C}) = \begin{vmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{vmatrix} \tag{A.22}$$

In order to calculate $det\ \boldsymbol{C}$, Laplace's formula can be used along the first row. Having denoted the cofactor of entry C_{ij} by CC_{ij} one will have

$$X = C_{11}CC_{11} + C_{12}CC_{12} + C_{13}CC_{13} \tag{A.23}$$

As $\boldsymbol{C}$ is a symmetric tensor, a vector of six elements could be associated to it. In this way one will have

$$\begin{aligned} X_0 &= det\boldsymbol{C}_0 \\ &\vdots \\ X_p &= \boldsymbol{M}_0^T\boldsymbol{C}_p + RX_p \end{aligned} \tag{A.24}$$

in which

$$\begin{aligned} \boldsymbol{C}_0 &= \begin{bmatrix} C_{0_{11}} & C_{0_{22}} & C_{0_{33}} & 2C_{0_{12}} & 2C_{0_{23}} & 2C_{0_{13}} \end{bmatrix}^T \\ \boldsymbol{M}_0^T &= \begin{bmatrix} CC_{11_0} & CC_{22_0} & CC_{33_0} & \dfrac{CC_{12_0}}{2} & \dfrac{CC_{23_0}}{2} & \dfrac{CC_{13_0}}{2} \end{bmatrix} \end{aligned} \tag{A.25}$$

$$\begin{aligned} RX_p = &\sum_{r=1}^{p-1} C_{11}^r CC_{11}^{p-r} + C_{12}^r CC_{12}^{p-r} + C_{13}^r CC_{13}^{p-r} \\ &+ C_{11}^0 \left(C_{22}^r C_{33}^{p-r} - C_{23}^r C_{23}^{p-r}\right) \\ &+ C_{12}^0 \left(C_{13}^r C_{23}^{p-r} - C_{12}^r C_{33}^{p-r}\right) \\ &+ C_{13}^0 \left(C_{12}^r C_{23}^{p-r} - C_{22}^r C_{13}^{p-r}\right) \end{aligned} \tag{A.26}$$

Note that, for example,

$$\begin{aligned} C_{11}^{p-r} = &\left(C_{22}^0 C_{33}^{p-r} - C_{23}^0 C_{23}^{p-r}\right) + \left(C_{22}^{p-r} C_{33}^0 - C_{23}^{p-r} C_{23}^0\right) \\ &+ \sum_{i=1}^{p-r-1} C_{22}^i C_{33}^{p-r-i} - C_{23}^i C_{23}^{p-r-i} \end{aligned} \tag{A.27}$$

Substituting C_p by its expansion given in A.21, the final expression for X_p is obtained as

$$X_p = 2\boldsymbol{M}_0^T\boldsymbol{B}(\boldsymbol{v}_0)\boldsymbol{v}_p + \underbrace{\boldsymbol{M}_0^T \underbrace{\sum_{r=1}^{p-1} \tilde{\boldsymbol{A}}(\boldsymbol{v}_r)\boldsymbol{G}\boldsymbol{v}_{p-r}}_{\boldsymbol{R}\boldsymbol{C}_p} + RX_p}_{RT_{X_p}} \tag{A.28}$$

where $\boldsymbol{B}(\boldsymbol{v_0}) = \boldsymbol{B}_l + \tilde{\boldsymbol{A}}(\boldsymbol{v_0})\boldsymbol{G}$

A.4 Expansion of $Y = \ln J$

The Taylor series of Y can be written as follows

$$Y(a) = Y_0 + \sum_{p=1} Y_p a^p,\ Y_0 = Y(0),\ Y_p = \frac{1}{p!}\frac{d^p Y}{da^p}\Big|_{a=0} \tag{A.29}$$

As $Y = \ln\sqrt{X}$ the chain rule could be applied which results in

$$\begin{aligned}
\frac{\partial Y}{\partial X} &= \frac{\partial \ln\sqrt{X}}{\partial X} = \frac{1}{\ln\sqrt{X}}\frac{1}{2\ln\sqrt{X}} = \frac{1}{2X}\\
\frac{\partial^2 Y}{\partial X^2} &= \frac{\partial}{\partial X}\frac{1}{2X} = \frac{-1}{2X^2}\\
&\vdots\\
\frac{\partial^p Y}{\partial X^p} &= \frac{(-1)^{p-1}(p-1)!}{2X_p}
\end{aligned} \tag{A.30}$$

We now define

$$\begin{aligned}
Y_{X_0} &= \ln\sqrt{X_0} = Y_0\\
Y_{X_1} &= \frac{\partial Y}{\partial X}\Big|_{X_0} = \frac{1}{2X_0}\\
Y_{X_2} &= \frac{1}{2!}\frac{\partial^2 Y}{\partial X^2}\Big|_{X_0} = \frac{-1}{4X_0^2}\\
&\vdots\\
Y_{X_p} &= \frac{(-1)^{p-1}}{2pX_0^p}
\end{aligned} \tag{A.31}$$

The coefficients in the Eq. A.29 are obtained as

$$\begin{aligned}
Y_0 &= \ln\sqrt{X_0}\\
Y_1 &= Y_{X_1}X_1 = \frac{X_1}{2X_0}\\
Y_2 &= Y_{X_2}X_1^2 + Y_{X_1}X_2 = \underbrace{\frac{-X_1^2}{4X_0^2}}_{R_{Y_2}} + \frac{X_2}{2X_0}\\
&\vdots\\
Y_p &= \frac{X_p}{2X_0} + R_{Y_p}
\end{aligned} \tag{A.32}$$

Using equation A.28 one will get

$$Y_p = \frac{1}{X_0}\boldsymbol{M}_0^T\boldsymbol{B}(\boldsymbol{v}_0)\boldsymbol{v}_p + \underbrace{\frac{1}{2X_0}\left(\boldsymbol{M}_0^T\boldsymbol{R}\boldsymbol{C}_p + R_{X_p}\right) + R_{Y_p}}_{RT_{Y_p}} \tag{A.33}$$

A.5 Expansion of $Z = \frac{1}{J^2}$

The Taylor series of Z can be written as follows

$$Z(a) = Z_0 + \sum_{p=1} Z_p a^p,\ Z_0 = Z(0),\ Z_p = \frac{1}{p!}\frac{d^p Z}{da^p}\Big|_{a=0} \tag{A.34}$$

Considering that $Z = \frac{1}{J^2} = \frac{1}{X}$, the coefficients in the series can be obtained as

$$\begin{aligned}
\frac{\partial Z}{\partial X} &= \frac{-1}{X^2} \\
\frac{\partial^2 Z}{\partial X^2} &= \frac{2}{X^3} \\
&\vdots \\
\frac{\partial^p Z}{\partial X^p} &= \frac{(-1)^p (p)!}{X^{p+1}}
\end{aligned} \tag{A.35}$$

We now define

$$\begin{aligned}
Z_{X_0} &= \frac{1}{det(C_0)} = \frac{1}{X_0)} = Z_0 \\
Z_{X_1} &= \frac{\partial Z}{\partial X}|_{X_0} = \frac{-1}{X_0^2} \\
Z_{X_2} &= \frac{1}{2!}\frac{\partial^2 Z}{\partial X^2}\Big|_{X_0} = \frac{1}{X_0^3} \\
&\vdots \\
Z_{X_p} &= \frac{(-1)^{p-1}}{X_0^(p+1)}
\end{aligned} \tag{A.36}$$

The coefficients in Eq. A.34 can be obtained as

$$\begin{aligned}
Z_0 &= Z_{X_0} = \frac{1}{det(C_0)} \\
Z_1 &= Z_{X_1} X_1 = \frac{-X_1}{X_0^2} \\
Z_2 &= Z_{X_2} X_1^2 + Z_{X_1} X_2 = \underbrace{\frac{X_1^2}{X_0^3}}_{R_{Z_2}} + \frac{-X_2}{X_0^2} \\
&\vdots \\
Z_p &= \frac{-X_p}{X_0^2} + R_{Z_p}
\end{aligned} \tag{A.37}$$

After replacing these variables with matrices and vectors that have been defined so far Z_p will take the form

$$Z_p = \frac{-2}{X_0^2}\boldsymbol{M}_0^T\boldsymbol{B}(\boldsymbol{v}_0)\boldsymbol{v}_p + \underbrace{\frac{-1}{X_0^2}\left(\boldsymbol{M}_0^T\boldsymbol{R}\boldsymbol{C}_p + R_{X_p}\right) + R_{Z_p}}_{RT_{Z_p}} \tag{A.38}$$

A.6 Expansion of C^{-1}

In order to calculate the inverse of $\boldsymbol{C}$, its adjugate will be calculated and then divided by its determinant. If the transpose of the cofactor matrix of $\boldsymbol{C}$ is denoted by $\boldsymbol{A}$, the inverse of $\boldsymbol{C}$ will be computed as follows

$$\boldsymbol{C}^{-1} = \frac{\boldsymbol{A}}{\det\ \boldsymbol{C}} = \frac{1}{X}\boldsymbol{A} = Z\boldsymbol{A} \tag{A.39}$$

where

$$\boldsymbol{A} = \begin{pmatrix} CC_{11} & CC_{12} & CC_{13} \\ CC_{21} & CC_{22} & CC_{23} \\ CC_{31} & CC_{32} & CC_{33} \end{pmatrix} \tag{A.40}$$

The expansion of the adjugate matrix CC^T is obtained as follows

$$\begin{aligned} \boldsymbol{A}_0 &= \boldsymbol{C}\boldsymbol{C}_0{}^T \\ &\vdots \\ \boldsymbol{A}_p &= \tilde{\boldsymbol{C}}_0\boldsymbol{C}_p + \boldsymbol{R}\boldsymbol{A}_p \end{aligned} \tag{A.41}$$

in which

$$\tilde{\boldsymbol{C}}_0 = \begin{bmatrix} 0 & C_{0_{33}} & C_{0_{22}} & 0 & -C_{0_{23}} & 0 \\ C_{0_{33}} & 0 & C_{0_{11}} & 0 & 0 & -C_{0_{13}} \\ C_{0_{22}} & C_{0_{11}} & 0 & -C_{0_{12}} & 0 & 0 \\ 0 & 0 & -C_{0_{12}} & -\frac{1}{2}C_{0_{23}} & \frac{1}{2}C_{0_{13}} & \frac{1}{2}C_{0_{23}} \\ -C_{0_{23}} & 0 & 0 & \frac{1}{2}C_{0_{13}} & -\frac{1}{2}C_{0_{11}} & \frac{1}{2}C_{0_{12}} \\ 0 & -C_{0_{13}} & 0 & \frac{1}{2}C_{0_{23}} & \frac{1}{2}C_{0_{12}} & -\frac{1}{2}C_{0_{22}} \end{bmatrix} \tag{A.42}$$

$$\boldsymbol{R}\boldsymbol{A}_p = \sum_{r=1}^{p-1} \begin{pmatrix} C_{22}^r C_{33}^{p-r} - C_{23}^r C_{23}^{p-r} \\ C_{11}^r C_{33}^{p-r} - C_{13}^r C_{13}^{p-r} \\ C_{11}^r C_{22}^{p-r} - C_{12}^r C_{12}^{p-r} \\ C_{13}^r C_{23}^{p-r} - C_{12}^r C_{33}^{p-r} \\ C_{13}^r C_{12}^{p-r} - C_{11}^r C_{23}^{p-r} \\ C_{12}^r C_{23}^{p-r} - C_{13}^r C_{22}^{p-r} \end{pmatrix}. \tag{A.43}$$

After replacing these variables with matrices and vectors that have been defined so far $\boldsymbol{A}_p$ will take the form

$$\boldsymbol{A}_p = 2\tilde{\boldsymbol{C}}_0\boldsymbol{B}(\boldsymbol{v}_0)\boldsymbol{v}_p + \underbrace{\tilde{\boldsymbol{C}}_0\boldsymbol{R}\boldsymbol{C}_p + \boldsymbol{R}\boldsymbol{A}_p}_{\boldsymbol{R}\boldsymbol{T}\boldsymbol{A}_p} \tag{A.44}$$

Now the inverse of Right Cauchy-Green strain tensor can be computed as

$$\begin{aligned}
\boldsymbol{C}_0^{-1} &= \boldsymbol{A}_0 Z_0 = Inv(\boldsymbol{C}_0) \\
&\vdots \\
\boldsymbol{C}_p^{-1} &= \boldsymbol{A}_p Z_0 + \boldsymbol{A}_0 Z_p + \underbrace{\sum_{r=1}^{p-1} \boldsymbol{A}_r Z_{p-r}}_{\boldsymbol{R}\boldsymbol{C}_p^{-1}}.
\end{aligned} \tag{A.45}$$

After replacing these variables with matrices and vectors that have been defined so far $\boldsymbol{C}_p^{-1}$ will take the form

$$\boldsymbol{C}_p^{-1} = 2\left(\tilde{\boldsymbol{C}}_0 Z_0 - \frac{-1}{X_0^2}\boldsymbol{A}_0\boldsymbol{M}_0^T\right)\boldsymbol{B}(\boldsymbol{v}_0)\boldsymbol{v}_p + \underbrace{\boldsymbol{A}_0 R T_{Z_p} + \boldsymbol{R}\boldsymbol{T}\boldsymbol{A}_p Z_0 + \boldsymbol{R}\boldsymbol{C}_p^{-1}}_{\boldsymbol{R}\boldsymbol{T}\boldsymbol{C}_p^{-1}} \tag{A.46}$$

A.7 Expansion of S

Replacing the obtained values so far in the stress formulation for a neo-Hookean compressible material from Eq. A.3 will result in

$$\begin{aligned}
S_0 &= \mu\boldsymbol{I} - \mu\boldsymbol{C}_0^{-1} + \lambda Y_0\boldsymbol{C}_0^{-1} \\
S_1 &= -\mu\boldsymbol{C}_1^{-1} + \lambda\left(Y_0\boldsymbol{C}_1^{-1} + Y_1\boldsymbol{C}_0^{-1}\right) \\
&\vdots \\
S_p &= -\mu\boldsymbol{C}_p^{-1} + \lambda\left(Y_0\boldsymbol{C}_p^{-1} + Y_p\boldsymbol{C}_0^{-1} + \underbrace{\sum_{r=1}^{p-1} Y_r\boldsymbol{C}_{p-r}^{-1}}_{R\boldsymbol{S}_p}\right)
\end{aligned} \tag{A.47}$$

After replacing these variables with matrices and vectors that have been defined so

far $\boldsymbol{S}_p$ will take the form

$$\boldsymbol{S}_p = \underbrace{\left(2(\lambda Y_0 - \mu)\left(\tilde{\boldsymbol{C}}_0 Z_0 - \frac{1}{X_0}\boldsymbol{C}_0^{-1}\boldsymbol{M}_0^T\right) + \frac{\lambda}{X_0}\boldsymbol{C}_0^{-1}\boldsymbol{M}_0^T\right)}_{\boldsymbol{D}} \boldsymbol{B}(\boldsymbol{v}_0)\boldsymbol{v}_p + \underbrace{(\lambda Y_0 - \mu)\boldsymbol{R}\boldsymbol{T}\boldsymbol{C}_p^{-1} + \lambda \boldsymbol{C}_0^{-1} R T_{Y_p} + \lambda \sum_{r=1}^{p-1} Y_r \boldsymbol{C}_{p-r}^{-1}}_{\boldsymbol{R}\boldsymbol{T}\boldsymbol{S}_p = \boldsymbol{S}_p^{nl}} \tag{A.48}$$

Now $\boldsymbol{S_p}$ can be written in the following form

$$\boldsymbol{S_p} = \boldsymbol{D}\ \boldsymbol{B}(\boldsymbol{v}_0)\boldsymbol{v}_p + \boldsymbol{S}_p^{nl} \tag{A.49}$$

Non-linear stress can be expressed in terms of geometrical and material non-linear stresses.

$$\boldsymbol{S}_p^{nl} = \boldsymbol{S}_p^{nlmat} + \boldsymbol{S}_p^{nlgeom} \tag{A.50}$$

In which

$$\boldsymbol{S}_p^{nlgeom} = \frac{1}{2}\boldsymbol{D}\boldsymbol{R}\boldsymbol{C}_p = \boldsymbol{D}\sum_{r=1}^{p-1}\frac{1}{2}\tilde{\boldsymbol{A}}(\boldsymbol{v}_r)\boldsymbol{\theta}_{p-r} \tag{A.51}$$

$$\begin{aligned}\boldsymbol{S}_p^{nlmat} = {} & (\lambda Y_0 - \mu)\left(\boldsymbol{A}_0\left(RZ_p - \frac{RX_p}{J_0^4}\right) + \frac{\boldsymbol{R}\boldsymbol{A}_p}{J_0^2} + \boldsymbol{R}\boldsymbol{C}_p^{-1}\right) \\ & + \lambda\left(\frac{\boldsymbol{A}_0}{J_0^2}\left(RY_p + \frac{RX_p}{2J_0^2}\right) + \boldsymbol{R}\boldsymbol{S}_p\right)\end{aligned} \tag{A.52}$$

A.8 Obtaining the final system of equations

The weak form of the problem is

$$\text{Find } \boldsymbol{v} \in H^1(\Omega)\ \ni \int_{\Omega_0} \boldsymbol{S} : \delta\boldsymbol{\gamma}(\boldsymbol{v})d\Omega = \lambda\left(\int_{\Omega_0}\boldsymbol{b}\cdot\delta\boldsymbol{v}\ d\Omega + \int_{\gamma_0}\bar{\boldsymbol{t}}\cdot\delta\boldsymbol{v}\ d\Gamma\right)\ \forall\delta\boldsymbol{v} \in H_0^1(\Omega) \tag{A.53}$$

In order to obtain the system of equations first the variation of Green-Lagrange strain tensor is expanded as

$$\begin{aligned}\delta\boldsymbol{\gamma}(\boldsymbol{v})_0 &= \left(\boldsymbol{B}_l + \tilde{\boldsymbol{A}}(\boldsymbol{v}_0)\boldsymbol{G}\right)\delta\boldsymbol{v} = \boldsymbol{B}(\boldsymbol{v}_0)\delta\boldsymbol{v} \\ &\vdots \\ \delta\boldsymbol{\gamma}(\boldsymbol{v})_p &= \boldsymbol{A}(\boldsymbol{v}_p)\boldsymbol{G}\delta\boldsymbol{v}\end{aligned} \tag{A.54}$$

Also the load factor λ has to be expanded.

$$\lambda(a) = \lambda_0 + \sum_{p=1} \lambda_p a^p, \ \lambda_0 = \lambda(0), \ \lambda_p = \frac{1}{p!} \frac{d^p \lambda}{da^p}\Big|_{a=0}$$

After substituting variables by their expansion and identifying terms with the same order the following system of equations will be obtained.

$$\begin{aligned} \boldsymbol{K}_T \boldsymbol{v}_0 &= \lambda_0 \boldsymbol{F} \\ &\vdots \\ \boldsymbol{K}_T \boldsymbol{v}_p &= \lambda_p \boldsymbol{F} - \boldsymbol{f}_p^{nl} \end{aligned} \tag{A.55}$$

in which

$$\boldsymbol{K}_T = \int_{\Omega_0} \boldsymbol{B}(\boldsymbol{v}_0) \ \boldsymbol{D} \ \boldsymbol{B}(\boldsymbol{v}_0) + \boldsymbol{G}^T \boldsymbol{S}_0 \boldsymbol{G} d\Omega \tag{A.56}$$

$$\boldsymbol{f}_p^{nl} = \int_{\Omega_0} (\boldsymbol{B}^T (\boldsymbol{S}_p^{nlmat} + \boldsymbol{S}_p^{nlgeom}) + \boldsymbol{G}^T \boldsymbol{S}_p^*) d\Omega \tag{A.57}$$

Bibliography

Bibliography

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

Alastrué, V., Calvo, B., na, E. P., and Doblaré, M. (2006). Biomechanical modeling of refractive corneal surgery. *Journal of Biomechanical Engineering-Trasactions of the ASME*, 128:150--160.

Allard, J., Cotin, S., Faure, F., j. Bensoussan, P., Poyer, F., Duriez, C., Delingette, H., and B, L. G. (2007). Sofa an open source framework for medical simulation. In *Medicine Meets Virtual Reality*.

Ammar, A., Mokdad, B., Chinesta, F., , and Keunings., R. (2007). 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.

Ammar, A., Mokdad, B., Chinesta, F., and Keunings, R. (2006). 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.

Ammar, A., Pruliere, E., Ferec, J., Chinesta, F., and Cueto, E. (2009). Coupling finite elements and reduced approximation bases. *European Journal of Computational Mechanics*, 18(5--6):445--463.

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

Arbogast, L. F. A. (1800). *Du calcul des derivations*. Strasbourg: Levrault.

Areias, P. M. A. and Belytschko, T. (2005). Analysis of three-dimensional crack initiation and propagation using the extended finite element method. *International Journal for Numerical Methods in Engineering*, 63(5):760--788.

Barbi**c**, J. and James, D. (2007). Time-critical distributed contact for 6-DoF haptic rendering of adaptively sampled reduced deformable models. In Metaxas, D and Popovic, J, editor, *ACM SIGGRAPH/ EUROGRAPHICS Symposium Proceedings*, pages 171--180. ACM SIGGRAPH; Eurog Assoc. Symposium on Computer Animation, San Diego, CA, AUG 03-04, 2007.

Barbi**c**, J. and James, D. L. (2005). Real-time subspace integration for St. Venant-Kirchhoff deformable models. *ACM Transactions on Graphics (SIGGRAPH 2005)*, 24(3):982--990.

Belytschko, T. and Black, T. (1999). Elastic crack growth in finite elements with minimal remeshing. *International Journal for Numerical Methods in Engineering*, 45(5):601--620.

Bielser, D. and Gross, M. H. (2000). Interactive simulation of surgical cuts. In *Proceedings of the 8th Pacific Conference on Computer Graphics and Applications*, PG '00, pages 116--, Washington, DC, USA. IEEE Computer Society.

Bonet, J. and Wood, R. D. (2008). *Nonlinear continuum mechanics for finite element analysis*. Cambridge University Press.

Bro-Nielsen, M. and Cotin, S. (1996). Real-time volumetric deformable models for surgery simulation using finite elements and condensation. *Computer Graphics Forum*, 15(3):57--66.

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

Capell, S., Green, S., Curless, B., Duchamp, T., and Z.Popovic (2002). Interactive skeleton-driven dynamic deformations. In *ACM SIGGRAPH*.

Chinesta, F., Ammar, A., and Cueto, E. (2010). Recent advances in the use of the Proper Generalized Decomposition for solving multidimensional models. *Archives of Computational Methods in Engineering*, 17(4):327--350.

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

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

Combescure, A., Gravouil, A., Gregoire, D., and Rethore, J. (2008). X-fem a good candidate for energy conservation in simulation of brittle dynamic crack propagation. *Computer Methods in Applied Mechanics and Engineering*, 197(5):309 -- 318. Enriched Simulation Methods and Related Topics.

Courtecuisse, H., Jung, H., Allard, J., Duriez, C., Lee, D. Y., and Cotin, S. (2010). Gpu-based real-time soft tissue deformation with cutting and haptic feedback. *Progress in Biophysics and Molecular Biology*, 103(2-3):159 -- 168. Special Issue on Biomechanical Modelling of Soft Tissue Motion.

Daux, C., Moes, N., Dolbow, J., Sukumar, N., and Belytschko, T. (2000). Arbitrary branched and intersecting cracks with the extended finite element method. *International Journal for Numerical Methods in Engineering*, 48(12):1741--1760.

Delingette, H. (2000). Towards realistic soft tissue modeling in medical simulation. Technical Report INRIA RR-3506.

Delingette, H. and Ayache, N. (2004). Soft tissue modeling for surgery simulation. In Ayache, N., editor, *Computational Models for the Human Body*, Handbook of Numerical Analysis (Ph. Ciarlet, Ed.), pages 453--550. Elsevier.

Delinguette, H. (1998). Towards realistic soft tissue modeling in medical simulation. In *Proceedings of the IEEE: special issue on surgery simulation*.

Delinguette, H., Subsol, G., Cotin, S., and Pignon, J. (1994). A craniofacial surgery simulation testbed. In *Visualization for Biomedical Computing*.

Deo, D. and De, S. (2009). Phyness: a physics-driven neural networks-based surgery simulation system with force feedback. In *Third Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, Salt Lake City, UT, USA*.

Famaey, N. and Stolen, J. (2008). Soft tissue modeling for applications in virtual surgery and surgical robotics. *Computer Methods in Biomechanics and Biomedical Engineering*, 11(4):351--366.

Fish, J. (1992). The s-version of the finite-element method. *Computers and Structures*, 43(3):539--547.

Holzapfel, G. A. and Gasser, T. C. (2000). A new constitutive framework for arterial wall mechanics and a comparative study of material models. *Journal of Elasticity*, 61:1--48.

Idelsohn, S. R. and Cardona, R. (1985). A reduction method for nonlinear structural dynamics analysis. *Computer Methods in Applied Mechanics and Engineering*, 49:253--279.

Jerabkova, L. (2007). Interactive cutting of finite elementsbased deformable objects in virtual environments. In *Thesis*.

Karhunen, K. (1946). Uber lineare methoden in der wahrscheinlichkeitsrechnung. *Ann. Acad. Sci. Fennicae, ser. Al. Math. Phys.*, 37.

Kenedi, R., Gibson, T., Evans, J., and Barbenel, J. (1975). Tissue mechanics. *Physics in Medicine and Biology*, 20(5):163--169.

Krysl, P., Lall, S., and Marsden, J. (2001). Dimensional model reduction in non-linear finite element dynamics of solids and structures. *Int. J. Numer. Meth. in Engng.*, 51:479--504.

Lee, B., Popescu, D. C., and Ourselin, S. (2010). Topology modification for surgical simulation using precomputed finite element models based on linear elasticity. *Progress in Biophysics and Molecular Biology*, 103(2-3):236 -- 251. Special Issue on Biomechanical Modelling of Soft Tissue Motion.

Lim, Y.-J. and De, S. (2007). 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.

Loève, M. M. (1963). *Probability theory*. The University Series in Higher Mathematics, 3rd ed. Van Nostrand, Princeton, NJ.

Lorenz, E. N. (1956). *Empirical Orthogonal Functions and Statistical Weather Prediction*. MIT, Departement of Meteorology, Scientific Report Number 1, Statistical Forecasting Project.

Ly, H. V. and Tran, H. T. (2005). Modeling and control of physical processes using proper orthogonal decomposition. *Mathematical and Computer Modelling*, 33:223--236.

Meier, U., Lopez, O., Monserrat, C., Juan, M. C., and Alcañiz, M. (2005). Real-time deformable models for surgery simulation: a survey. *Computer Methods and Programs in Biomedicine*, 77(3):183--197.

Melenk, J. M. and Babuska, I. (1996). The partition of unity finite element method: Basic theory and applications. *Computer Methods in Applied Mechanics and Engineering*, 139(1-4):289 -- 314.

Meschke, G. and Dumstorff, P. (2007). Energy-based modeling of cohesive and cohesionless cracks via x-fem. *Computer Methods in Applied Mechanics and Engineering*, 196(21-24):2338 -- 2357.

Moës, N., Dolbow, J., and Belytschko, T. (1999). A finite element method for crack growth without remeshing. *International Journal for Numerical Methods in Engineering*, 46(1):131--150.

Monserrat, C., Meier, U., Alcañiz, M., Chinesta, F., and Juan, M. C. (2001). Real-time deformable models for surgery simulation: a survey. *Computer Methods and Programs in Biomedicine*, 64(2):77--85.

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

Niroomandi, S., Alfaro, I., Cueto, E., and Chinesta, F. (2010a). 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:--.

Niroomandi, S., Alfaro, I., Cueto, E., and Chinesta, F. (2010b). Model order reduction for hyperelastic materials. *International Journal for Numerical Methods in Engineering*, 81(9):1180--1206.

Ono, K., Schillhuber, G., and Ulbrich, H. (2009). Xfem approach to real-time incision haptic feedback for surgical simulation. In *XFEM ECCOMAS conference 2009*.

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

Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, 2(6):559--572.

project, N. (2009). Surgeons practice on virtual brain.

Rank, E. and Krause, R. (1997). A multiscale finite-element method. *Computers and Structures*, 64(1-4):139 -- 144. Computational Structures Technology.

Ryckelynck, D. (2003). 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)*.

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

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

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

Satava, R. (1996). Medical virtual reality: The current status of the future. In *4th conference of Medicine Meets Virtual Reality*.

Sirovich, L. (1987). Turbulence and the dynamics of coherent structures part I: coherent structures. *Quaterly of applied mathematics*, XLV:561--571.

Stazi, F. L., Budyn, E., Chessa, J., and Belytschko, T. (2003). An extended finite element method with higher-order elements for curved cracks. *Computational Mechanics*, 31:38--48. 10.1007/s00466-002-0391-2.

Stolarska, M., Chopp, D., Moes, N., and Belytschko, T. (2001). Modelling crack growth by level sets in the extended finite element method. *International Journal for Numerical Methods in Engineering*, 51(8):943--960.

Strouboulis, T., Babuska, I., and Copps, K. (2000). The design and analysis of the generalized finite element method. *Computer Methods in Applied Mechanics and Engineering*, 181(1-3):43 -- 69.

Sukumar, N., Moës, N., Moran, B., and Belytschko, T. (2000a). Extended finite element method for three-dimensional crack modelling. *International Journal for Numerical Methods in Engineering*, 48(11):1549--1570.

Sukumar, N., Moës, N., Moran, B., and Belytschko, T. (2000b). Extended finite element method for three-dimensional crack modelling. *International Journal for Numerical Methods in Engineering*, 48:1549--1570.

Taylor, Z., Cheng, M., and Ourselin, S. (2008). High-speed nonlinear finite element analysis for surgical simulation using graphics processing units. *Medical Imaging, IEEE Transactions on*, 27(5):650 --663.

Tenenbaum, J. B., de Silva, V., and Langford, J. C. (2000). A global framework for nonlinear dimensionality reduction. *Science*, 290:2319--2323.

Waters, K. (1992). A physcial model of facial tissue and muscle articulation derived from computer tomography data. *Visualization in Biomedical Computing*, 1808:574--583.

Wriggers, P. (2002). *Computational Contact Mechanics*. Wiley.

Yvonnet, J., Zahrouni, H., and Potier-Ferry, M. (2007). A model reduction method for the post-buckling analysis of cellular microstructures . *Computer Methods in Applied Mechanics and Engineering*, 197:265--280.

Zhang, J., Gu, L., Li, X., and Fang, M. (2009). An advanced hybrid cutting method with an improved state machine for surgical simulation. *Computerized Medical Imaging and Graphics*, 33(1):63 -- 71.

Zienkiewicz, O. C. and Taylor, R. L. (1994). *El Metodo de los Elementos Finitos*. McGraw-Hill-CIMNE.