Centro Internacional de Matemática

International Conference on Mathematics and Continuum Mechanics

António Joaquim Mendes Ferreira
Isabel Maria Narra de Figueiredo
Juha Videman

Porto, Portugal,
19-22 February, 2008

30

Contents

Foreword

The *International Conference on Mathematics and Continuum Mechanics* was held under the auspices of the CIM (Centro Internacional de Matemática). It took place at the FEUP (Faculdade de Engenharia da Universidade do Porto), in Porto, Portugal, from February 19 to 22, 2008.

The event covered a selected range of interdisciplinary topics related to Applied Mathematics and Continuum Mechanics. The target audience included engineers, physicists and mathematicians, as well as graduate and PhD students interested in mathematical problems in Solid and Fluid Mechanics and Geophysics.

The conference highlighted six thematic mini-symposia on

- **Computational Methods for Advanced Composites**, organized by Pedro Camanho (Faculty of Engineering, University of Porto, Portugal);

- **Contact Mechanics**, organized by Marius Cocou (Laboratoire de Mécanique et d'Acoustique CNRS, Marseille, France);

- **Mathematics and the Atmospheric Sciences**, organized by João Teixeira (NASA CalTech-Jet Propulsion Laboratory, USA);

- **Modelling of Industrial Processes**, organized by Luísa Silva (Ecole Nationale Supérieure des Mines de Paris (ENSMP), France);

- **Numerical Analysis of Thin Structures**, organized by Lourenço Beirão da Veiga (University of Milan, Italy);

- **Ocean Dynamics**, organized by Aires dos Santos (Instituto Superior Técnico, Lisboa, Portugal).

Each symposium consisted of one forty-five minute plenary lecture and two to four invited half-hour talks. Some contributed papers were selected for short presentations. Extended abstracts of most of the invited and contributed talks have been collected to this volume.

We express our most sincere thanks to the mini-symposia organizers and to all the speakers for their valuable contribution. We are also grateful to CEMAT (Centro de Matemática e Aplicações do Instituto Superior Técnico), CMUC (Centro de Matemática da Universidade

Porto, February 2008

Organizing Committee
António Mendes Ferreira (Universidade do Porto)
Isabel Narra de Figueiredo (Universidade de Coimbra)
Juha Hans Videman (Instituto Superior Técnico, Lisboa)

Conference Information

Conference Organizers

- António Joaquim Mendes Ferreira (University of Porto, Portugal)

- Isabel Maria Narra de Figueiredo (University of Coimbra, Portugal)

- Juha Hans Videman (Instituto Superior Técnico, Lisbon, Portugal)

Minisymposia

Computational Methods for Advanced Composites

Organizer: Pedro Camanho (Faculty of Engineering, University of Porto, Portugal)

Plenary talk: *Multiscale modelling of structural composites: the development of strategies for virtual testing*, Francisco Javier Llorca Martínez (Polytechnic University of Madrid, Spain)

Invited talks:

- *On the prediction of failure in laminated composites*, Silvestre Pinho (Imperial College, UK)
- *The simulation of coupled failure mechanisms using discontinuous solid-like shell elements*, Joris Remmers (University of Delft, The Netherlands)

Contact Mechanics

Organizer: Marius Cocou (Laboratoire de Mécanique et d'Acoustique CNRS, France)

Plenary talk: *Stability of quasi-static paths of finite-dimensional systems with Coulomb friction and persistent contact*, João Martins (Instituto Superior Técnico, Lisboa, Portugal)

Invited talks:

- *Recent progresses in the modeling of material behavior*, Gianpietro Del Piero (Università di Ferrara, Italy)

- *Existence and uniqueness for quasistatic contact problems with friction*, Lars-Erik Andersson (Linköping University, Sweden)

Mathematics and the Atmospheric Sciences

Organizer: João Teixeira (NASA/CalTech Jet Propulsion Laboratory, USA)

Plenary talk: *Weather forecasting: It's about dynamics, it is not about statistics*, Kevin Judd (University of Western Australia, Australia)

Invited talks:

- *Mimetic methods: from global weather forecasting to CFD*, Luca Bonaventura (Politecnico di Milano, Italy)

- *Cloud structures and organisation*, A. Pier Siebesma (Royal Netherlands Meteorological Institute)

Modelling of Industrial Processes

Organizer: Luísa Silva (Ecole Nationale Supérieure des Mines de Paris (ENSMP), France)

Plenary talk: *Moving surfaces and interfaces calculation in material forming*, Thierry Coupez (CEMEF, Ecole des Mines de Paris, France)

Invited talks:

- *Non-local models for localization in large deformations*, José César de Sá (Faculty of Engineering, University of Porto, Portugal)

- *The fixed mesh ALE approach: basic idea and two applications*, Ramon Codina (CIMNE, University of Catalunya, Spain)

Numerical Analysis of Thin Structures

Organizer: Lourenço Beirão da Veiga (University of Milan, Italy)

Plenary talk: *Fundamental and applicative challenges in the modeling and computations of shells*, Dominique Chapelle (INRIA, France)

Invited talks:

- *A-posteriori error estimates for the Reissner-Mindlin plate problem*, Carlo Lovadina (University of Pavia, Italy)
- *Numerical shell eigen problem benchmarks*, Harri Hakula (Helsinki University of Technology, Finland)

Ocean Dynamics

Organizer: Aires dos Santos (Instituto Superior Técnico, Lisbon, Portugal)

Plenary talk: *Global ocean state estimation*, Rui Ponte (Atmospheric and Environmental Research, Inc, USA)

Invited talks:

- *Operational oceanography: zoom-in modeling for local applications*, Emanuel Ferreira Coelho (University of Southern Mississippi, USA)
- *Processes over submarine canyons*, Henrique Coelho (Hidromod, Modelação em Engenharia, Lda, Portugal)
- *Towards an hydrodynamic an biogeochemical operational model of the Portuguese coast*, Paulo Chambel Leitão (Hidromod, Modelação em Engenharia, Lda, Portugal)
- *Wind and ocean induced wake instabilities leeward of a deep-sea island*, Rui Caldeira (University of Madeira, Funchal, Portugal)

PROGRAMME

CIM – International Conference on Mathematics and Continuum Mechanics FEUP, Porto, February 19-22, 2008						
Time \ Date	February 19		February 20		February 21	February 22
8:00-9:00	Registration		Registration			
9:00-10:00	MS-1(P) Room A		MS-2(P) Room A		MS-5(P) Room A	
10:00-10:30	Coffee Break		Coffee Break		Coffee Break	Coffee Break
10:30-11:30	MS-1(I) Room A		MS-2(I) Room A		MS-5(I) Room A	MS-6(I) Room A
11:30-13:00	MS-1(C) Room A	MS-3(C) Room B	MS-2(C) Room A	MS-1(C) Room B	MS-5(C) Room A	MS-2(C) Room B	MS-6(I) Room A
13:00-14:15	Lunch		Lunch		Lunch	Lunch
14:30-15:30	MS-3(P) Room A		MS-4(P) Room A		MS-6(P) Room A	
15:30-16:00	Coffee Break		Coffee Break		Coffee Break	
16:00-17:00	MS-3(I) Room A		MS-4(I) Room A		Miscellaneous Talks Room A	
19:00-22:00	Reception				Conference Dinner	

- MS-1: Mini-Symposium Computational Methods for Advanced Composites

  MS-2: Mini-Symposium Contact Mechanics

  MS-3: Mini-Symposium Mathematics and the Atmospheric Sciences

  MS-4: Mini-Symposium Modelling of Industrial Processes

  MS-5: Mini-Symposium Numerical Analysis of Thin Structures

  MS-6: Mini-Symposium Ocean Dynamics

- MS-i(P), MS-i(I) and MS-i(C): MS-i Plenary talks (P), MS-i Invited talks (I) and MS-i Selected Contributed talks (C), for i=1,...,6.

- Room A: B032

- Room B: B033

List of Participants

Adeel, Muhammad (Beihang University, Beijing, China)

Al-Mdallal, Qasem M. (United Arab Emirates University, UAE)

Andersson, Lars-Erik (Linköping University, Sweden)

Baghdasaryan, Zorayr (Yerevan State University, Armenia)

Ballard, Patrick (Laboratoire de Mécanique et d'Acoustique CNRS, Marseille, France)

Beirão da Veiga, Lourenço (University of Milan, Italy)

Bock, Igor (FEI Slovak University of Technology, Bratislava, Slovak Republic)

Bonfanti, Giovanna (Università degli Studi di Brescia, Italy)

Bonaventura, Luca (Politecnico di Milano, Italy)

Caldeira, Rui Miguel Andrade (Universidade da Madeira, Funchal Portugal)

Camanho, Pedro (University of Porto, Portugal)

Chapelle, Dominique (INRIA, France)

Cocou, Marius (Laboratoire de Mécanique et d'Acoustique CNRS, Marseille, France)

Codina, Ramon (CIMNE, University of Catalunha, Spain)

Coelho, Celestino (Instituto Superior Técnico, Portugal)

Coelho, Emanuel Ferreira (University of Southern Mississippi, USA)

Coelho, Henrique (Hidromod, Modelação em Engenharia, Lda, Portugal)

Coupez, Thierry (CEMEF, Ecole des Mines de Paris, France)

Del Piero, Gianpietro (Università di Ferrara, Italy)

Férnandez, José R. (Universidade de Santiago de Compostela, Spain)

Ferreira, António Joaquim Mendes (University of Porto, Portugal)

Ferreira, José A. Mendes (University of Coimbra, Portugal)

Figueiredo, Isabel Maria Narra de (University of Coimbra, Portugal)

Hakimzadeh, Z. (Islamic Azad University of Iran, Iran)

Hakula, Harri (Helsinki University of Technology, Finland)

Hild, Patrick (Université de Franche-Comté, Besançon, France)

Judd, Kevin (University of Western Australia, Australia)

Kumar, Prashant (Indian Institut of Technology, Kanpur, India)

Leitão, Paulo Chambel (Hidromod, Modelação em Engenharia, Lda, Portugal)

Lopes, Cláudio (Delft University of Technology, The Netherlands and University of Porto, Portugal)

Lovadina, Carlo (University of Pavia, Italy)

Luís, Euclides (Instituto Superior Técnico, Portugal)

Madureira, Luísa R. (University of Porto, Portugal)

Martínez, Francisco Javier Llorca (Polytechnic University of Madrid, Spain)

Martínez, Rebeca (Universidade de Santiago de Compostela, Spain)

Martins, João (Instituto Superior Técnico, Portugal)

Neves, Miguel M. (Instituto Superior Técnico, Portugal)

Niemi, Antti H. (Helsinki University of Technology, Finland)

Paris, Iria (Glaizer Group, France)

Pereira, Bruno (Instituto Superior Técnico, Portugal)

Pesetskaya, Ekaterina (University of Aveiro, Portugal)

Pinho, Silvestre (Imperial College, UK)

Pinho-da-Cruz, J. (University of Aveiro, Portugal)

Policarpo, Hugo (Instituto Superior Técnico, Portugal)

Ponte, Rui (Atmospheric and Environmental Research, Inc, USA)

Pozzolini, Cédric (Laboratoire de Mécanique et d'Acoustique CNRS, Marseille, France)

Remmers, Joris (University of Delft, The Netherlands)

Roque, Carla M. C. (University of Porto, Portugal)

Sá, José César de (University of Porto, Portugal)

Santos, Aires dos (Instituto Superior Técnico, Portugal)

Siebesma, A. Pier (Royal Netherlands Meteorological Institute)

Silva, Luísa (Ecole Nationale Supérieure des Mines de Paris (ENSMP), France)

Soares, Pedro Matos (University of Lisbon, Portugal)

Syam, Muhammed I. (United Arab Emirates University, UAE)

Teixeira, João (NASA/Caltech - Jet Propulsion Laboratory, USA)

Teixeira, Miguel (University of Lisbon, Portugal)

Vicente, Manuel A. Facas (Universidade de Coimbra, Portugal)

Videman, Juha (Instituto Superior Técnico, Portugal)

Articles

Computational Methods for Advanced Composites

Analytical investigation of the effective properties of 2D composite materials

L. P. Castro[*]   E. Pesetskaya[†]

Abstract

The effective conductivity of 2D double periodic composite materials with circle disjoint inclusions is investigated. Ideal and non-ideal contact conditions on the boundary between the matrix and the inclusions are considered. In both cases, explicit formulas for the effective conductivity are obtained. These formulas contain all parameters of the considered model such as the conductivity, radii, centers of the inclusions, the conductivity of the matrix, and also special Eisenstein functions. Based on the explicit formulas, the effect of perturbation of the inclusion in the case of ideal contact conditions is also investigated. It is found that any random motion of the inclusions increase the effective conductivity in the case when the conductivity of the inclusions is greater than the conductivity of the matrix.

1   Introduction

For the investigation of composite material properties with a 2D structure, the theory of analytical and harmonic functions is widely used ([4]). With the help of this theory, it is possible to describe some characteristics of 2D models with different configurations in an explicit form, i.e., in the form of a formula which contains basic parameters of the models. This approach allows to easily analyse the behavior of the models when the parameters change.

According to the homogenization theory (see, for example, [2]), fibrous composite materials with a large number of cylindrical parallel inclusions can be geometrically represented as a 2D double periodic model with disjoint inclusions, and their properties can be studied on the basis of one representative cell which contains all material features (see Figure 1). Mathematically, a 2D model with disjoint inclusions is a multiply connected domain. Physical fields

---

[*]Research Unit *Mathematics and Applications*, Department of Mathematics, University of Aveiro, Campus Universitário de Santiago, 3810-193, Aveiro, Portugal. E-mail:`castro@ua.pt`.

[†]Research Unit *Mathematics and Applications*, Department of Mathematics, University of Aveiro, Campus Universitário de Santiago, 3810-193, Aveiro, Portugal. E-mail:`kate@ua.pt`. Supported by *Unidade de Investigação Matemática e Aplicações* of Universidade de Aveiro, and the Portuguese Science Foundation (*FCT–Fundação para a Ciência e a Tecnologia*) through the grant number SFRH/BPD/34649/2007.

in the composite components (the matrix and the inclusion) are supposed to be potential, i.e., satisfy the Laplace equation in each inner point of the considered domain. Boundary conditions are defined on the basis of mechanical assumptions. In other words, mechanical processes in composite materials can be described in terms of mixed boundary value problems for the Laplace equation in multiply connected domains.
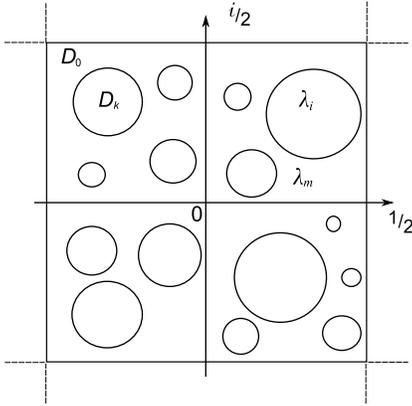
Fig.1: The representative cell $Q_{(0,0)}$.
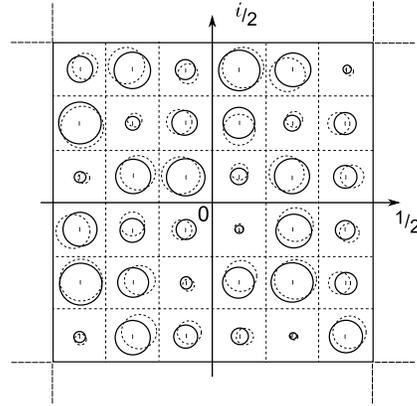
Fig.2: Perturbation effect.

Using methods of complex analysis, explicit formulas for the effective conductivity in the case of ideal and non-ideal contact conditions on the boundary between the matrix and the inclusions are obtained. These formulas contain the conductivity of components, radii, centers, the number of the inclusions, and special meromorphic double periodic Eisenstein functions ([6]). Based on the formulas for the effective conductivity in the case of ideal contact conditions, the effect of perturbation of the inclusions (see Figure 2) is investigated and it is therefore found an extremal property.

2   Statement of the problem

Let us consider a lattice L which is defined by the two fundamental translation vectors 1 and $\imath$ $(\imath^2 = -1)$ in the complex plane $\mathbb{C} \cong \mathbb{R}^2$ of the complex variable $z = x + \imath y$. The *representative cell* $Q_{(0,0)}$ is the square $\{z = t_1 + \imath t_2 \in \mathbb{C} : -\frac{1}{2} < t_p < \frac{1}{2},\ p = 1,2\}$. Let $\mathcal{E} := \bigcup_{m_1, m_2} \{m_1 + \imath m_2\}$ be the set of the lattice points, where $m_1, m_2 \in \mathbb{Z}$. The cells corresponding to the points of the lattice $\mathcal{E}$ will be denoted as $Q_{(m_1, m_2)} = Q_{(0,0)} + m_1 + \imath m_2 := \{z \in \mathbb{C} : z - m_1 - \imath m_2 \in Q_{(0,0)}\}$. It is considered the situation when mutually disjoint disks (inclusions) of different radii $D_k := \{z \in \mathbb{C} : |z - a_k| < r_k\}$ with the boundaries $T_k := \{z \in \mathbb{C} : |z - a_k| = r_k\}$ $(k = 1, 2, \ldots, N)$ are located inside the representative cell $Q_{(0,0)}$ and periodically repeated in all cells $Q_{(m_1, m_2)}$. Let us denote the connected domain $D_0 := Q_{(0,0)} \setminus (\bigcup_{k=1}^{N} D_k \cup T_k)$ obtained by removing of the inclusions from the representative cell $Q_{(0,0)}$.

We state the problem to define the effective conductivity of the double periodic composite material with matrix $D_{matrix} = \bigcup\limits_{m_1,m_2} ((D_0 \cup \partial Q_{(0,0)}) + m_1 + \imath m_2)$ and inclusions $D_{inc} = \bigcup\limits_{m_1,m_2} \bigcup\limits_{k=1}^{N} (D_k + m_1 + \imath m_2)$ occupied by materials of conductivities $\lambda_m > 0$ and $\lambda_i > 0$, respectively. This problem is equivalent to the determination of the potential of the corresponding fields, i.e., to find a function u satisfying the Laplace equation in each component of the composite material:

$$\triangle u(z) = 0, \quad z \in D_{matrix} \cup D_{inc}. \tag{1}$$

In addition, in the mechanics of composite materials, one of the widespread boundary conditions is the ideal contact conditions:

$$u^+(t) = u^-(t), \quad \lambda_m \frac{\partial u^+}{\partial n}(t) = \lambda_i \frac{\partial u^-}{\partial n}(t), \quad t \in \bigcup\limits_{m_1,m_2} T_k, \tag{2}$$

where $\frac{\partial}{\partial n} = n_1 \frac{\partial}{\partial x} + n_2 \frac{\partial}{\partial y}$ is the outward normal derivative and $u^+(t) := \lim\limits_{z \to t, z \in D_0} u(z)$, $u^-(t) := \lim\limits_{z \to t, z \in D_k} u(z)$. The first condition in (2) is the temperature balance, and the second one is the flux balance.

Other conditions on the boundary between the matrix and the inclusions are the non-ideal contact conditions:

$$\lambda_m \frac{\partial u^+}{\partial n}(t) = \lambda_i \frac{\partial u^-}{\partial n}(t), \quad \lambda_i \frac{\partial u^-}{\partial n}(t) + \gamma_k(u^+(t) - u^-(t)) = 0, \quad t \in \bigcup\limits_{m_1,m_2} T_k, \tag{3}$$

where $\gamma_k > 0$, $n = (n_1, n_2)$ is the outward unit normal vector to T_k.

We suppose that the external field is applied in the x-direction. It means that $u(z)$ has a constant jump in the x-direction: $u(z + 1) = u(z) + 1$, $u(z + \imath) = u(z)$.

Let us now consider certain complex potentials $\varphi(z)$ and $\varphi_k(z)$ which are analytic in D_0 and D_k, and continuously differentiable in the closures of D_0 and D_k, respectively. These complex analytic functions $\varphi(z)$ and $\varphi_k(z)$ are introduced in such a way that they are related with the harmonic potential $u(z)$ by means of the following equality:

$$u(z) = \begin{cases} \Re(\varphi(z) + z), & z \in D_0, \\ \frac{2\lambda_m}{\lambda_m + \lambda_i} \Re\varphi_k(z), & z \in D_k. \end{cases} \tag{4}$$

Thus, the function $\varphi(z)$ is double periodic: $\varphi(z + 1) = \varphi(z)$, $\varphi(z + \imath) = \varphi(z)$. To determine the current $\nabla u(x, y)$, we need to obtain the derivatives of the complex potentials:

$$\psi(z) := \frac{\partial \varphi}{\partial z} = \frac{\partial u^+}{\partial x} - \imath \frac{\partial u^+}{\partial y} - 1, \quad \psi_k(z) := \frac{\partial \varphi_k}{\partial z} = \frac{\lambda_m + \lambda_i}{2\lambda_m} \left(\frac{\partial u^-}{\partial x} - \imath \frac{\partial u^-}{\partial y} \right). \tag{5}$$

The boundary conditions (2) and (3) can be rewritten in terms of the complex potentials $\psi(z)$ and $\psi_k(z)$ (see [1], [3]). For the boundary conditions (2), we have

$$\psi(t) = \psi_k(t) + \rho \left(\frac{r_k}{t - a_k} \right)^2 \overline{\psi_k(t)} - 1. \tag{6}$$

15

For the boundary conditions (3), the problem turns out to be stated in the form

$$\psi(t) = (1+\mu)\psi_k(t) + \rho\left(\frac{r_k}{t-a_k}\right)^2 \overline{\psi_k(t)} + \mu(t-a_k)\psi_k'(t) - \mu\frac{r_k^4}{(t-a_k)^3}\overline{\psi_k'(t)} - 1, \quad (7)$$

where $\rho = \frac{\lambda_i - \lambda_m}{\lambda_i + \lambda_m}$ is the contrast parameter, and $\mu = \frac{\lambda_i}{2r_k\gamma_k(\rho-1)}$.

The problems (6) and (7) are solved by the method of functional equations (cf. [1] and [3]). The main idea of this method consists in the following. We have N contours T_k and N complex conjugation conditions on each contour T_k. Then, we need to find $N+1$ functions $\psi, \psi_1, \ldots, \psi_N$, i.e., we need one more condition to close up the system. To obtain this condition, we define (on the basis of functions $\psi, \psi_1, \ldots, \psi_N$) a new function Φ such that Φ is analytic in $Q_{(0,0)}$ and on $\bigcup_{k=1}^{N} D_k$ has a zero jump along each T_k, $k = 1, 2, \ldots, N$, and is double periodic on $\mathbb{C}$. Therefore, from Liouville's theorem for double periodic functions, we have that $\Phi(z) = c$ for some constant c. This gives up an additional condition on $\psi, \psi_1, \ldots, \psi_N$. It is shown (see [1]) that the constant c turns to be equal to zero. The function Φ for the conditions (6) and (7) is built in [1] and [3]. The condition $\Phi(z) = 0$ gives us the system to find $\psi, \psi_1, \ldots, \psi_N$ which is inside a formula for the effective conductivity. The functions $\psi_k(z)$ are looked for in the form of a series expansion in r_k^2: $\psi_k(z) = \sum_{s=0}^{\infty} \psi_k^{(s)}(z)\, r_k^{2s}$.

3   Results

The effective conductivity λ_e of an isotropic composite material with N inclusions is obtained in the form $\lambda_e = \lambda_m + 2\rho\lambda_m\nu\frac{1}{N}\sum_{k=1}^{N}\psi_k(a_k)$, cf. [1]. This can be rewritten in the form of a series in ν:

$$\lambda_e = \lambda_m + 2\lambda_m\rho\nu\sum_{p=0}^{s}\mathbf{A}_p\nu^p + o(\nu^{s+1}). \qquad (8)$$

The coefficients $\mathbf{A}_p$ are defined by the equalities $\mathbf{A}_p = \frac{1}{\pi^p N^{p+1}}\sum_{k=1}^{N}\psi_k^{(p)}(a_k)$, $p = 0, 1, 2, \ldots$, where the functions $\psi_k^{(p)}$ in the case of ideal contact conditions, i.e. for (6), are defined by the following recurrence relations (cf. [1]):

$$\psi_k^{(0)}(z) = 1,$$

$$\psi_k^{(1)}(z) = \rho\Big[\sum_{m\neq k}^{N}\overline{\psi_{0\,m}^{(0)}}E_2(z-a_m) + \overline{\psi_{0\,k}^{(0)}}\sigma_2(z-a_k)\Big],$$

$$\vdots \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (9)$$

$$\psi_k^{(p+1)}(z) = \rho\Big[\sum_{m\neq k}^{N}\overline{\psi_{p\,m}^{(0)}}E_{p+2}(z-a_m) + \overline{\psi_{p\,k}^{(0)}}\sigma_{p+2}(z-a_k) + \sum_{m\neq k}^{N}\overline{\psi_{p-1\,m}^{(1)}}E_{p+1}(z-a_m)$$

$$+ \overline{\psi_{p-1\,k}^{(1)}}\sigma_{p+1}(z-a_k) + \cdots + \sum_{m\neq k}^{N}\overline{\psi_{0\,m}^{(p)}}E_2(z-a_m) + \overline{\psi_{0\,k}^{(p)}}\sigma_2(z-a_k)\Big].$$

16

For non-ideal contact conditions, the $\psi_k^{(p)}$ elements are defined as follows (see also [3]):

$$\psi_k^{(0)}(z) = -\frac{\mu}{1+\mu}(z-a_k)[\psi_k^{(0)}(z)]' + \frac{1}{1+\mu},$$

$$\psi_k^{(1)}(z) = -\frac{\mu}{1+\mu}(z-a_k)[\psi_k^{(1)}(z)]' + \frac{\rho}{1+\mu}\Big[\sum_{m\neq k}\overline{\psi_{0m}^{(0)}}E_2(z-a_m) + \overline{\psi_{0k}^{(0)}}\sigma_2(z-a_k)\Big],$$

$$\vdots \tag{10}$$

$$\psi_k^{(p)}(z) = -\frac{\mu}{1+\mu}(z-a_k)[\psi_k^{(p)}(z)]' + \frac{\rho}{1+\mu}\Big[\sum_{m\neq k}\overline{\psi_{0m}^{(p-1)}}E_2(z-a_m) + \overline{\psi_{0k}^{(p-1)}}\sigma_2(z-a_k)$$

$$+ \sum_{m\neq k}\overline{\psi_{1m}^{(p-2)}}E_3(z-a_m) + \overline{\psi_{1k}^{(p-2)}}\sigma_3(z-a_k) + \cdots + \sum_{m\neq k}\overline{\psi_{p-1m}^{(0)}}E_{p+1}(z-a_m)$$

$$+ \overline{\psi_{p-1k}^{(0)}}\sigma_{p+1}(z-a_k)\Big] - \frac{\mu}{1+\mu}\Big[\sum_{m\neq k}\overline{\psi_{1m}^{(p-2)}}E_3(z-a_m) + \overline{\psi_{1k}^{(p-2)}}\sigma_3(z-a_k)$$

$$+ \sum_{m\neq k}2\,\overline{\psi_{2m}^{(p-3)}}E_4(z-a_m) + 2\,\overline{\psi_{2k}^{(p-3)}}\sigma_4(z-a_k) + \cdots$$

$$+ \sum_{m\neq k}(p-1)\overline{\psi_{p-1m}^{(0)}}E_{p+1}(z-a_m) + (p-1)\overline{\psi_{p-1k}^{(0)}}\sigma_{p+1}(z-a_k)\Big],$$

Here, $\psi_{lm}^{(s)}$ is the derivative of order l of s approximation, $E_p(z)$ is the elliptic Eisenstein function of order p ([6]), σ_p is the modified Eisenstein function (cf., e.g., [1]).

The situation of random perturbation of the inclusions in the case of ideal contact conditions is considered in what follows (cf. Figure 2 and see also [5]). For definiteness, it is assumed that $N = n^2$, with $n \in \mathbb{N}$. Let $b_1, b_2, \ldots, b_N$ be the centers which form a periodic square array in the unit cell $Q_{(0,0)}$, and identically distributed random variables $\mathbf{d_1}, \mathbf{d_2}, \ldots, \mathbf{d_N}$ have uniform distribution inside the disk $|z| < d_k$, i.e., they have the probability density

$$f_k(z) = \begin{cases} \dfrac{1}{\pi d_k^2}, & |z| < d_k, \\ 0, & |z| \geq d_k. \end{cases}$$

Let $\mathbf{a_1}, \mathbf{a_2}, \ldots, \mathbf{a_N}$ be the random variables defined by formula $\mathbf{a_k} = \mathbf{d_k} + \mathbf{b_k}$, $k = 1, \ldots, N$, where the variables $\mathbf{b_k}$ possess the constant values b_k. In what follows, a_k refers to the values of the above introduced random variables $\mathbf{a_k}$. Thus, the disks $D_k = \{z : |z - b_k| < d_k\}$ are randomly distributed inside the periodic cell. It is assumed that $d_k + r_k < \frac{1}{\sqrt{N}}$. This means that each disk lies in the prescribed part of the original cell and does not cross or touch the boundary of the corresponding part. Each coefficient $\mathbf{A}_p$ in (8) is now decomposed as $\mathbf{A}_p = A_p^{(0)} + \widetilde{\mathbf{A}_p}$, where $A_p^{(0)}$ corresponds to the deterministic periodic structure ($d_k = 0$), and $\widetilde{\mathbf{A}_p}$ is the correction due to randomness of the positions of a_k. The expected value from the

latter relation is $\langle \mathbf{A}_p \rangle = A_p^{(0)} + \langle \widetilde{\mathbf{A}_p} \rangle$. By definition

$$\langle \widetilde{\mathbf{A}_p} \rangle := \int\limits_{[Q]^N} \widetilde{\mathbf{A}_p}(a_1, \ldots, a_N) \prod_{k=1}^{N} f_k(a_k)(d\sigma)^N, \quad p = 0, 1, 2, \ldots.$$

Here, the integration is performed with respect to the complex-valued variables $a_k = x_1 + \imath x_2$ $(k = 1, 2, \ldots, N)$ which are the values of random variables $\mathbf{a_k}$, and $d\sigma := dx_1 dx_2$. Since λ_e is linear in A_p, the same decomposition for λ_e is obtained: $\lambda_e = \lambda_e^{(0)} + \widetilde{\lambda}_e$. The expected value is defined by the formula $\langle \lambda_e \rangle = \lambda_e^{(0)} + \langle \widetilde{\lambda}_e \rangle$, where

$$\langle \widetilde{\lambda}_e \rangle = 2\lambda_m \rho \nu \sum_{p=0}^{s} \langle \widetilde{\mathbf{A}_p} \rangle \nu^p + o(\nu^{s+1}).$$

It is shown in [5] that $\langle \widetilde{\mathbf{A}_p} \rangle \geq 0$, $p = 0, 1, 2, \ldots$, for $\rho > 0$. It follows from the inequality $\langle \widetilde{\mathbf{A}_p} \rangle \geq 0$ that random perturbations of the inclusions increase the effective conductivity of a composite material when the contrast parameter ρ is greater than zero. Thus, in the case when the conductivity of the inclusions is greater than the conductivity of the matrix, a periodic structure with circular inclusions has a lower effective conductivity than any non-periodic structure which consists of inclusions perturbed in an arbitrary way.

References

[1] L. Berlyand and V. Mityushev. Generalized clausius-mossotti formula for random composite with circular fibers. *J. Statist. Phys.*, 102(1/2):115–145, 2001.

[2] V. V. Jikov, S. M. Kozlov, and O. A. Olejnik. *Homogenization of Differential Operators and Integral Functionals.* Springer-Verlag, 1994.

[3] E. Pesetskaya L. P. Castro and S. Rogosin. Effective conductivity of a composite material with non-ideal contact conditions. to appear.

[4] V. Mityushev and S. Rogosin. *Constructive Methods for Linear and Nonlinear Boundary Value Problems for Analytic Functions. Theory and Applications*, volume 108 of *Monographs and Surveys in Pure and Applied Mathematics.* Chapman & Hall/CRC, New York, 1999.

[5] E. V. Pesetskaya. Effective conductivity of composite materials with random position of cylindrical inclusions. finite number of inclusions in the cell. *Applicable Analysis*, 84:843–865, 2005.

[6] A. Weil. *Elliptic Functions According to Eisenstein and Kronecker.* Springer-Verlag, 1976.

Modeling desorption in polymers

J.A. Ferreira[*]   P. de Oliveira [†]   P. da Silva[‡]

Abstract

In this paper a mathematical model for desorption phenomena in polymers in studied. The well-posedness of the model and its numerical simulation are considered.

1 Introduction

Over the past decades the study of polymers behaviour has received the attention of many theoretical and experimental researchers. The reasons for this interest lie on the very challenging mathematical models underlying the phenomena but mainly on the fact that due to its properties polymers are used in a large number of industries as pharmaceutical, equipment, clothing and sealants.

The problems reported in the literature are essentially of two different but related kinds: sorption of penetrants by dry polymeric matrices and desorption of penetrants from polymeric saturated matrices. Two main approaches can be found in the literature to model sorption phenomena. In the first one a Fickian diffusion is considered in the glassy and rubbery regions and the kinetics of the glass-rubber polymer transition is taken properly into account. Due to this kinetics these models have a locally non Fickian character ([7], [9], [12]). In the second approach globally non Fickian models are considered as Fick's law is modified by introducing in the flux a viscoelastic stress ([3], [4], [5]). Much less attention has been placed on the mathematical modeling of polymer desorption process. In the literature we find mainly mathematical models where a viscoelastic stress has been introduced in the flux ([1],[5]). Nevertheless, the numerical simulations obtained using those models exhibit sharp fronts which in some cases do not agree with the experimental data ([10], [11]).

Experimentalists consider that in a system where a front separates two different states the only obvious violation of Fick's law takes place at the front ([9],[10]).

---

[*]CMUC-Department of Mathematics,University of Coimbra . E-mail:`ferreira@mat.uc.pt`. This work was partially supported by Project:PTDC/MAT/74548/2006

[†]CMUC-Department of Mathematics, University of Coimbra. E-mail:`poliveir@mat.uc.pt`

[‡]Department of Physics and Mathematics, Instituto Superior de Engenharia de Coimbra E-mail:`pascals@mail.isec.pt`

Following this idea the authors proposed in [6] a mathematical model for desorption that uses Fick's law in the rubbery and glassy regions and a rate controlled motion of the moving front separating the glass and rubber regions. The discontinuity of the concentration in the front is considered and the glassy concentration in this point is not prescribed. Due to this fact, the proposed model is an open model in the sense that a new condition for the front speed is needed. In this paper we consider the desorption model proposed in [6] with a new condition for the front speed depending on a phenomenological parameter which complete the open model. A numerical method is designed and its simulation is considered.

2   The desorption model

The model proposed in [6] is based on Fick's law for the flux. The concentration in both parts of the polymer are described by the diffusion equation

$$
\begin{cases}
C_t^g = D_g C_{xx}^g, \ x < s(t) \\
\\
C_t^r = D_r C_{xx}^r, \ x \geq s(t)
\end{cases}
, \tag{1}
$$

where $x = s(t)$ represents the position of the front and D_g and D_r represent the molecular diffusion coefficients respectively in the glassy and rubbery regions. In (1) C^g and C^r denote the concentrations respectively in the glassy and rubbery regions. The two coupled diffusion equations (1) are linked by the interface condition

$$
s'(t)\big([C]_{s(t)} + k_1\big) = [J]_{s(t)}, \tag{2}
$$

where $[g]_{s(t)} = g(s(t)_+, t) - g(s(t)_-, t)$ represents the jump of g at $s(t)$ and $J = -D(C)C_x$, and k_1 is a small parameter which prevents the cancelation of the expressions in brackets.

The two coupled diffusion equations (1) and the front condition (2) are complemented with a boundary condition at the front which is written as

$$
C^r(s(t), t) = C_\star, t > 0, \tag{3}
$$

where $C_\star$ represents the characteristic solute concentration that distinguishes the glassy from the swollen state. In the swollen saturated region, the concentration C of the penetrant within the polymer is greater than $C_\star$, while in the glassy drier region, $C < C_\star$. A discontinuity in the concentration is admissible because we do not impose the glassy concentration, $C^g(s(t), t)$ to be equal to $C_\star$.

To complete the model of desorption we consider an initial condition

$$
C(x, 0) = C_{init}, \ x > 0. \tag{4}
$$

The far end is insulated which translates to

$$
C_x(\infty, t) = 0, t \geq 0. \tag{5}
$$

20

At $x = 0$ we consider

$$C(0, t) = C_{ext}, \ t \geq 0, \tag{6}$$

where C_{ext} stands for the external concentration. This condition corresponds to a model with infinite permeability at the outflow end.

As we mentioned before, the model (1)-(6) is an open model in the sense that a new condition for the front speed is needed. The model is completed by prescribing for the front speed $s'(t)$ the new condition

$$s'(t) = \frac{1}{k_2}[C]_{s(t)}, \tag{7}$$

where k_2 denotes a parameter. The phenomenological character of the model is given by the front conditions (2), (7). In fact (2) appears as a consequence of the decreasing behaviour of the overall solvent mass in the polymer. The motion rate (7) is similar to the one considered in [2], [8] for sorption phenomena.

3 Well-posedness of the model

In order to study the well-posedness of the initial boundary value problem (1), (3-(7) we start by analysing the behaviour of $\|C(t)\|_{L^2}$ when $C_{ext} = 0$. By $(.,.)$ and $\|.\|_{L^2}$ we denote the L^2 inner product and the L^2 norm respectively.

It can be easily shown that

$$\frac{d}{dt}\|C(t)\|_{L^2}^2 = 2\int_0^L C_t C \, dx + s'(t)\left(C(s(t)_-, t)^2 - C_\star^2\right).$$

Taking into account the boundary conditions we obtain

$$\frac{1}{2}\frac{d}{dt}\|C(t)\|_{L^2}^2 = -(DC_x, C_x) + \frac{1}{2}s'(t)\left(C(s(t), t)_-^2 - C_\star^2\right) + [CJ]_{s(t)}.$$

If

$$\frac{1}{2}s'(t)\left(C(s(t)_-, t)^2 - C_\star^2\right) + [CJ]_{s(t)} \leq 0 \tag{8}$$

then we conclude that $\frac{1}{2}\frac{d}{dt}\|C(t)\|_{L^2}^2 = -(DC_x, C_x)$ which implies

$$\|C(t)\|_{L^2} \leq \|C_{init}\|_{L^2}, t \geq 0. \tag{9}$$

We prove in what follows that (8) is consequence of the coupled conditions (2) and (7). From the transition conditions we get

$$\frac{1}{2}[J]_{s(t)}(C(s(t)_-, t) + C_\star) = \frac{k_1}{2k_2}[C]_{s(t)}(C(s(t)_-, t) + C_\star) + \frac{1}{2}s'(t)[C]_{s(t)}(C(s(t)_-, t) + C_\star)$$

which is equivalent to

$$\frac{1}{2}s'(t)(C(s(t)_-, t)^2 - C_\star^2) = -\frac{1}{2}[J]_{s(t)}(C(s(t)_-, t) + C_\star) + \frac{k_1}{2k_2}(C(s(t)_-, t)^2 - C_\star^2). \tag{10}$$

From (10), we conclude (8) if we prove that

$$-\frac{1}{2}[J]_{s(t)}(C(s(t)_-,t)+C_\star)+\frac{k_1}{2k_2}(C(s(t)_-,t)^2-C_\star^2)+[CJ]_{s(t)}\leq 0. \qquad (11)$$

It is easy to show that (11) holds provided the parameters satisfy

$$\frac{k_1}{k_2}\leq\frac{D_rC_x(s(t)_+,t)+D_gC_x(s(t)_-,t)}{C(s(t)_-,t)+C_\star}, \qquad (12)$$

for all t. As we are dealing with a desorption phenomenon, if we assume that $C_x>0$ then the condition (12) depends mainly on the values $C_\star$ and C_{ext}. In fact as we have $C_{ext}+C_\star\leq C(s(t)_-,t)+C_\star\leq 2C_\star$ if the phenomenological parameters are such that $\frac{k_1}{k_2}\leq\frac{D_rC_x(s(t)_+,t)+D_gC_x(s(t)_-,t)}{2C_\star}$, then the condition (12) also holds.

Let us consider now two solutions C and $\tilde{C}$ of the initial boundary value problem (1), (3-(7) with the initial conditions $C(x,t)=C_{init}$ and $\tilde{C}(x,0)=\tilde{C}_{init}$, respectively. Then for $w(x,t)=C(x,t)-\tilde{C}(x,t)$ holds the estimate (9) with $\|C_{init}\|_{L^2}$ replaced by $\|C_{init}-\tilde{C}_{init}\|_{L^2}$. Such estimate enable us to conclude the stability of the coupled initial boundary value problem (1), (3-(7) and the uniqueness of the solution for this problem.

4 Numerical simulation

In this section we illustrate the behaviour of the phenomenological model studied in the previous sections. Let $\{x_i=x_{i-1}+h,\ i=1,\dots,N\}$, with $x_0=0, x_N=L$, be a spatial grid in $[0,L]$ where L is large enough. In $[0,T]$, $T>0$, we define the time grid $\{t_j=t_{j-1}+\Delta t, j=1,\dots,M\}$ with $t_0=0$ and $t_M=T$.

Each time step is decomposed in two stages. In the first stage we compute a Fickian approximation which is corrected in the second stage. We describe separately the initial step and a typical step.

1 <u>Initialization.</u> **i**. First stage: "Fickian prediction":Initially the polymer is saturated. We consider that the desorption phenomena is governed by the classical diffusion equation $C_t=D_rC_{xx}$, $x\in(0,L)$, which is discretized by $\dfrac{\tilde{C}_i^{j+1}-\tilde{C}_i^j}{\Delta t}=D_rD_{2,x}\tilde{C}_i^j$, $i=1,\dots,N$, $j\geq 0$, with initial and boundary conditions $\tilde{C}_i^0=C_{init}, i=0,\dots,N$, $\tilde{C}_0^j=C_{ext}$, $\tilde{C}_{N-1}^j=\tilde{C}_{N+1}^j\ j\geq 1$, where $D_{2,x}\tilde{C}_i^j=\dfrac{\tilde{C}_{i+1}^j-2C_i^j+\tilde{C}_{i-1}^j}{h^2}$ and $x_{N+1}=L+h$.

Let t_p be the first time level such that, at some grid point, the numerical approximation is less than the transition concentration $C_\star$. Let $\tilde{s}_p$ be such that $\tilde{C}_i^p<C_*$, $x_i\leq\tilde{s}_p-h, \tilde{C}_i^p=C_\star$ for $x_i=\tilde{s}_p$. Then $\tilde{s}_p$ is the first transition point determined by the Fickian behaviour of the penetrant in the polymer and, in $[0,\tilde{s}_p)$ and $[\tilde{s}_p,L]$, at this instant, the polymer presents numerical glassy and rubbery states respectively. The computed transition points are, in what follows, referred as Fikian transition points.

ii. Second stage: "Locally non Fickian correction". The speed separating the rubber and glassy states is controlled by the transition conditions (2) and (7). Discretizing (2) we obtain

$$D_g \frac{\tilde{C}^{g,p}(\tilde{s}_p-) - \tilde{C}^{g,p}(\tilde{s}_p - h)}{h} - D_r \frac{-\tilde{C}^{r,p}(\tilde{s}_p + h)C_\star}{h} = \frac{\tilde{s}_{p+1} - \tilde{s}_p}{\Delta t}\left(k_1 + C_\star - \tilde{C}^{g,p}(\tilde{s}_p-)\right). \quad (13)$$

which leads to computation of $\tilde{C}^{g,p}(\tilde{s}_p-)$. In the initialization we consider $\tilde{s}_{p+1} = \tilde{s}_p$. In the computation of the new transition point s_p^* we consider $\tilde{s}_p = \tilde{s}_p + \frac{\Delta t}{k_2}\left(C_\star - \tilde{C}^{g,p}(\tilde{s}_p-)\right)$, and we define the transition point s_p is defined by $s_p = \min\{s_p^*, \tilde{s}_p\}$. Let i_p be such that $x_{i_p} = s_p$.

a. If the new transition point is the Fickian one, that is $s_p = \tilde{s}_p$, then the computed numerical approximations for the concentrations $\tilde{C}^r(x_i, t_p)$, $\tilde{C}^g(x_i, t_p)$ are corrected considering the finite difference equations

$$\frac{C_\ell^{r,p} - \tilde{C}_\ell^{r,p}}{\Delta t} = D_r D_{2,x}\tilde{C}_\ell^{r,p}, \; \ell = i_p + 1, \ldots, N, \quad C_{i_p}^{r,p} = C_\star, \; C_{N-1}^{r,p} = C_{N+1}^{r,p}, \quad (14)$$

$$\frac{C_\ell^{g,p} - \tilde{C}_\ell^{g,p}}{\Delta t} = D_g D_{2,x}\tilde{C}_\ell^{g,p}, \; \ell = 1, \ldots, i_p - 1, \quad C_0^{g,p} = C_{ext}, \; \tilde{C}_{i_p}^{g,p} = \tilde{C}^{g,p}(\tilde{s}_p-) \quad (15)$$

b. If the new transition point is obtained with a non-Fickian correction, that is $x_{i_p} = s_p^*$, then in (15) we consider the boundary condition $C_{i_p}^{g,p} = \tilde{C}_{i_p}^{g,p}$.

2. <u>Typical steps.</u> *i*. First stage : "Fickian prediction". Let $x_{i_q} = s_q$ be the transition point at time level t_q. The approximations for the concentrations at level t_{q+1} are computed using

$$\frac{\tilde{C}_\ell^{q+1} - C_\ell^q}{\Delta t} = D_g D_{2,x}C_\ell^q, \; \ell = 1, \ldots, i_q - 1, \quad \frac{\tilde{C}_\ell^{q+1} - C_\ell^q}{\Delta t} = D_r D_{2,x}C_\ell^q, \; i = i_q + 1, \ldots, N, \text{ and}$$

$$\frac{\tilde{C}_{i_q}^{q+1} - \frac{C_\star + C^{g,q}(s_q-)}{2}}{\Delta t} = \frac{1}{h}\left(D_r \frac{C_{i_q+1}^{r,q} - C_\star}{h} - D_g \frac{C^{g,q}(s_q-) - C_{i_q-1}^{g,q}}{h}\right).$$

We compare the concentrations $\tilde{C}_i^{q+1}, i = 1, \ldots, N$, with $C_\star$ and we compute $\tilde{s}_{q+1}$ such that $C_i^{q+1} < C_\star$, for $x_i < \tilde{s}_{q+1}$, and $C_i^{q+1} = C_\star$ for $x_i = \tilde{s}_{q+1}$.

ii. Second stage: "Locally non Fickian correction". We compute s_{q+1}^* be using $s_{q+1}^* = s_q + \frac{\Delta t}{k_2}\left(C_\star - C^{g,q}(s_q-)\right)$, where $C^{g,q}(s_q-)$ is defined by (13) with the convenient modifications. Then we define the new transition point s_{q+1} by $s_{q+1} = \min\{s_{q+1}^*, \tilde{s}_{q+1}\}$. The numerical approximations computed in the first stage are then corrected using the procedure described in the second stage of the initialization with convenient adaptations.

In the numerical experiments presented bellow we consider $L = 5$, $\Delta t = 4 \times 10^{-5}$, $h = 0.01$, $C_\star = 0.8$, $D_r = 1$, $D_g = 0.4$, $C_{ext} = 0$.

In Figure 1 (a),(b),(c) we plot the numerical approximations obtained with the method presented before for different values of k_2. The behaviour of our model for different values of k_1 is illustrated in Figure 1 (c) ($k_1 = 0, k_2 = \frac{1}{3}$) and (d) ($k_1 = \frac{1}{5}, k_2 = \frac{1}{3}$). The behaviour of the desorbed mass $M(0) - M(t)$, where $M(0) = LC_{init}$ and $M(t) = \int_0^L C(x,t)\,dx$, for different values of the parameter $k_2(k_1 = 0)$ is illustrated in Figure 1 (e). Finally, in Figure 1 (f) we plot the front position for different values of k_2. ($k_1 = 0$).

(a)　　　　　　　(b)　　　　　　　(c)

(d)　　　　　　　(e)　　　　　　　(f)
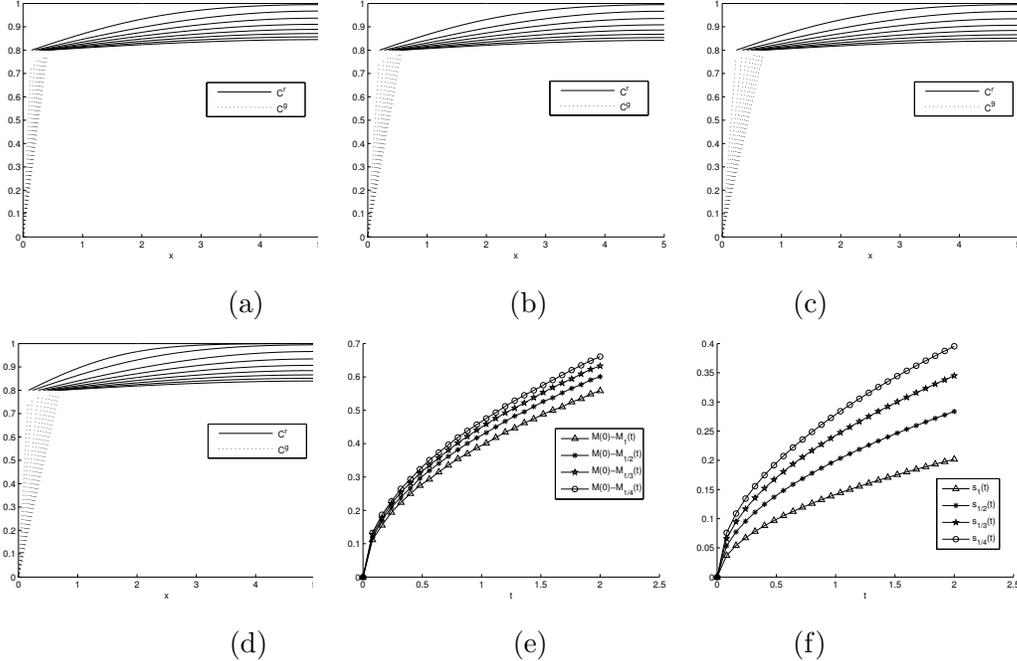
Figure 1: The C_g and C_r approximations for $k_1=0$, $t = 1, \ldots, 8$: (a) $k_2 = 1$ (b) $k_2 = 1/2$, (c) $k_2 = 1/3$; (d) The C_g and C_r approximations for $k_1 = \dfrac{1}{5}$, $k_2 = \dfrac{1}{3}$; (e)-Dependence on k_2 of the mass loss, $k_1 = 0$; (f) -Dependence on k_2 of the moving front $k_1 = 0$.

5   Conclusions

The mathematical model studied in this paper is based on the open model introduced in [6]. Prescribing a new condition for the position front depending on a phenomenological parameter, a complete mathematical model is obtained. This new condition is analogous to the one considered in the literature for sorption phenomena as for instance in [2], [8].

From the numerical simulation we conclude that the increasing of the phenomenological parameter k_2 retarded the position of the transition front and consequently implies a decreasing of the desorbed mass.

References

[1] R.A. Cairncross, L.F. Francis, and L.E. Scriven. Predicting drying in coatings that react and gel: Drying regime maps. *AIChE J.*, 42:55–67, 1996.

[2] D.S. Cohen and T. Erneux. Free boundary problems in controlled release pharmaceuticals. i: Diffusion in glassy polymers. *SIAM J. Appl. Math.*, 48:1451–1465, 1988.

[3] D.S. Cohen and A.B. White Jr. Sharp fronts due to diffusion and viscoelastic relaxation in polymers. *SIAM J. Appl. Math.*, 51:472–483, 1991.

[4] R.W. Cox and D.S. Cohen. A mathematical model for stress-driven diffusion in polymers. *J. Poly Sci.*, B:589–602, 1989.

[5] D.A. Edwards. A spatially nonlocal model for polymer-penetrant diffusion. *J. Appl. Math. Phys.*, 52:254–288, 2001.

[6] J.A. Ferreira, P. de Oliveira, P. Da Silva, and D. M. G. Comissiong. A phenomenological model for desorption in polymers. Submitted for publication.

[7] A. Friedman and G. Rossi. Phenomenological continuum equations to describe case II diffusion in polymeric materials. *Macromolecules*, 30:153–154, 1997.

[8] P. Guidotti. Diffusion in glassy polymers: Sorption of finite amount of solvent. *Nonlinear Differ. Equ. Appl.*, 6:297–317, 1999.

[9] T. Qian and P.L. Taylor. From the thomas-windle model to a phenomenological description of case II diffusion in polymers. *Polymer*, 41:7159–7163, 2000.

[10] M.A. Samus and G. Rossi. Methanol absorption in ethylene-vinyl alcohol copolymers: Relation between solvent diffusion and changes in glass transition temperature in glassy polymers. *Macromolecules*, 29:2275–2288, 1996.

[11] M. Sanopoulou, D.F. Stamatialis, and J.H. Petropoulos. Investigation of case II diffusion behavior. i. theoretical studies based on the relaxation dependent solubility model. *Macromolecules*, 35:1012–1020, 2002.

[12] N.L. Thomas and A.H. Windle. A theory of case II diffusion. *Polymer*, 23:529–542, 1982.

Numerical prediction of low-velocity impact damage resistance of composite laminates

C. Lopes[*]   P.Camanho[†]   Z. Gürdal[‡]   P. Maimí[§]

Abstract

Low-velocity impact events occur with some frequency on composite applications such as airplane components. From ground operations to unavoidable birds, there is a range of situations where an aircraft outer component may be subjected to unexpected impact loads. In most cases, the impactor has a relative low velocity and only Barely Visible Impact Damage (BVID) results. In such case the component shall be able to withstand the ultimate load for which it was designed and therefore replacement is not required. Therefore it is of major importance to design composite laminates in such a way as to maximize the range of impact events that cause only BVID as this may substantially reduce the airplane maintenance costs.

By one side, analytical predictions of the impact damage resistance of composite laminates are overly simplified and unreliable. On the other side, testing each promising design is time consuming and costly. Low-cost virtual testing by means of nonlinear Finite Element (FE) analyses can replace most of the actual impact testing of laminates. Once the dynamics of the impact phenomena and the damage mechanisms are correctly simulated, progressive failure analyses can be a valuable tool in the accurate prediction of impact damage resistance of composites.

In this research, a reliable virtual tool for the numerical simulation of the impact damage on composite laminates is proposed. A continuum material model for the simulation of intraply damage phenomena is implemented in a numerical scheme [8, 9] as a user subroutine of the commercially available ABAQUS® FE code. Each layer is modelled with a single through-the-thickness solid element. Delaminations, a critical damage phenomenon in low-velocity impact events, are simulated by making use of cohesive elements [13].

---

[*]Delft University of Technology and University of Porto. E-mail:`c.s.farialopes@tudelft.nl`. Supported by Fundação para a Ciência e Tecnologia

[†]FEUP, University of Porto E-mail:`pcamanho@fe.up.pt`

[‡]Aerospace Structures, Delft University of Technology E-mail:`z.gurdal@tudelft.nl`

[§]AMADE, University of Girona E-mail:`pere.maimi@udg.es`

1   Simulation of Low-Velocity Impact Damage

As a result of a low-velocity impact load on a plate, a dent is formed and a bending moment is generated. If the impacted region is made of a composite laminate with symmetric layup, the top layers, from the point-of-view of the impactor, are loaded under compression whilst the bottom ones are put under tension. The first damage mechanism to occur in the laminate is matrix cracking at the bottom plies, as sketched in Fig. 1(a). This is because, in general the strength of a fibre-matrix ply has the lowest value for transverse tension loading. If the impact energy is high enough, these cracks easily propagate trough-the-thickness until they reach an interface with a layer at a different angle. At this stage, delaminations may be triggered, as shown in Fig. 1(b), due to the discontinuity in the stress-displacement field at these interfaces.Further increasing the impact energy may eventually lead to fibre breakage by tension on the lower plies.

(a) Initiation of matrix cracks   (b) Onset of delaminations
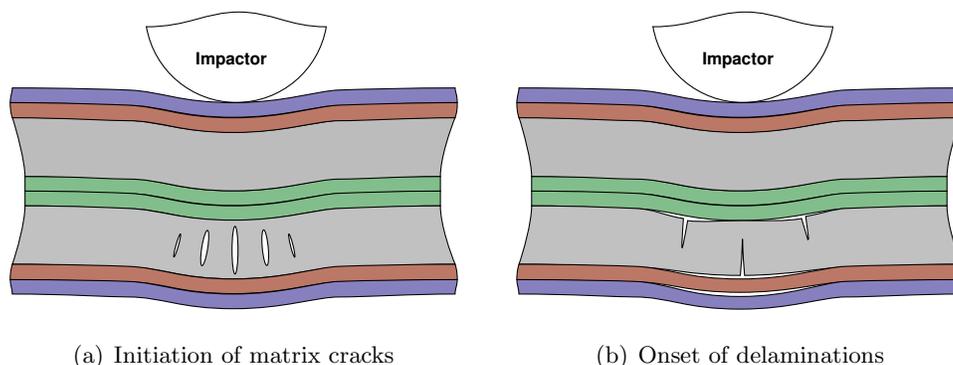
Figure 1: Illustration of damage mechanisms in a typical impact event on a symmetric composite laminate with clustered layers at the same fibre angle.

A typical progressive failure analysis starts with a nonlinear analysis to account for the geometrically nonlinear response of the structure. Using this nonlinear solution, the local lamina stresses are determined and checked against a failure criteria to determine whether any failures have occurred for this load increment. If lamina failure is not detected, the applied load is incremented and the analysis continues. When matrix or fibre fracture occur, a new stiffness tensor is calculated based on the material damage model. In a dynamic analysis, such as the simulation of impact, the load may not cause the structure to fail altogether but lead to the accumulation of substantial damage and localised material failure. In a FE simulation, the localised stiffness reduction and excessive element distortion introduce numerical convergence difficulties which may cause the analysis to abort.

1.1 Intraply Damage

The model used in this work to simulate fibre and matrix damage and failure is based in the previous work of the authors [8, 9]. Its main advantages in relation to previously developed damage models for composites are:

- The use of physically-based criteria to predict the onset of matrix cracking and fiber fracture under tensile and compression loads (Fig. 2) ; the LaRC04 failure criteria [11] which include the in-situ effects of ply thickness and surface constraining on its transverse tensile and shear strengths.

- The modification of Bažants crack band model [7] to ensure a mesh independent solution in scenarios where the fracture planes may have several orientations, and

- The simulation of crack-closure effects under load reversal cycles.

Furthermore, the majority of the material properties needed for the definition of the damage model can be measured using standard test methods that are defined at ply level. In summary, the independent material properties required are:

- Ply elastic properties (E_1, E_2, G_{12}, v_{12}) and ply strengths (X_T, X_C, Y_T, Y_C, S_L). These properties can be measured using test standards defined by the American Society for Testing and Material Standards (ASTM)[4, 2, 3].

- Four components of the fracture toughness, associated with longitudinal failure in tension and compression (G_{1+} and G_{1-}, respectively) and with transverse failure in tension and shear (G_{2+} and G_6, respectively). G_{2+} can be measured using a standard test procedure devised by the ASTM [5]. G_{1+} and G_{1-} are measured using compact tension and compact compression tests developed by Pinho et al.[12]. G_6 can be measured using the four-point bending end-notched flexure test proposed by Martin et al.[10].

The fracture toughness G_{2-} is dependent on G_6 and on the angle α_0 in the form $G_{2-} = G_6/\cos(\alpha_0)$ [8, 9]. Additionally, the model requires the input of the in-situ strengths Y_T^{is} and S_L^{is} which are functions of the independent material properties, as shown above.

The damage model was implemented by Maimí et al. [8, 9] in ABAQUS® FE code [1] as a user-written UMAT (*User MATerial*) subroutine. The subroutine calculates the material tangent stiffness tensor and updates the stress tensor associated to the elements which use the material damage model. In the current analyses, the described damage model is used to describe the constitutive behaviour of ABAQUS solid elements C3D8R and C3D8I.

1.2 Delaminations

In the FE models the bond between each two layers is simulated by means of the cohesive zone approach. In this approach, the fracture behaviour (delamination) is lumped into a single

29

a) Longitudinal tensile fracture b) Longitudinal compressive fracture

c) Transverse fracture with $\alpha=0^o$ d) Transverse fracture with $\alpha=53^o$
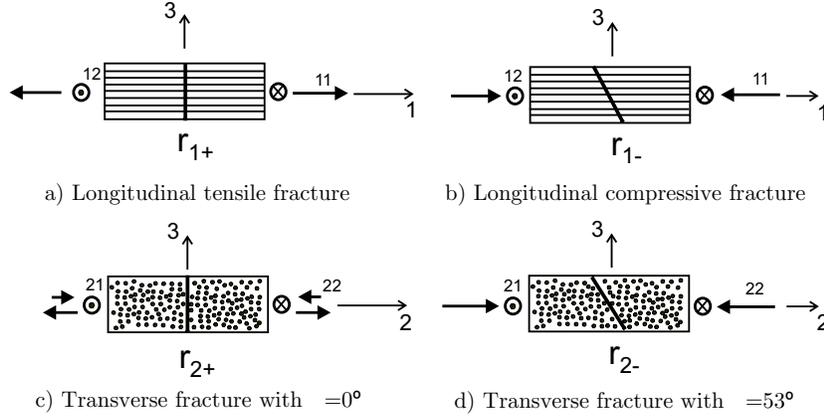
Figure 2: Ply failure modes predicted by the LaRC04 set of failure criteria.

plane, which is represented by interface elements placed between two layers. These interface elements consist of two surfaces, which are attached to the adjacent continuum elements that model the layers (Fig. 3). The relative displacement of the two surfaces is a measure for the opening of the delamination crack. The opening is controlled by means of a cohesive constitutive relation that completely characterises the delamination process.
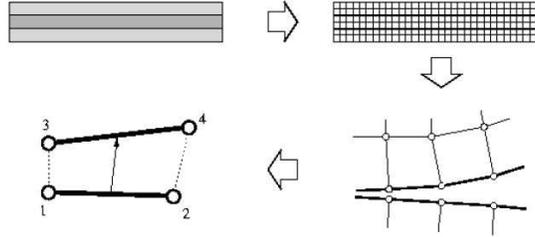
Figure 3: Simulation of delamination in a multiply material.

In this study, a bi-linear cohesive relation as proposed by Turon [13] is used. This model consists of four material parameters: mode I and mode II fracture toughnesses (G_{Ic} and G_{IIc}) and the corresponding ultimate tractions (τ_{Iu} and τ_{IIu}) at which debonding is initiated. A fifth parameter, η, is necessary to completely define the mixed-mode propagation criterion as function of G_{Ic} and G_{IIc} only. This value must be extracted from the correlation of the test data.

1.3 Virtual Test Setup

The FE model definition follows the simulation of the actual testing derived from the *standard test method for measuring the damage resistance of a fiber-reinforced polymer matrix composite to a drop-weight impact event* [6], as devised by the American Society for Testing and Materials

Standards (ASTM). The objective of the test is the evaluation of the level of damage induced by an impact event, of a predetermined energy, on a specified laminated composite material. The geometry for this test is sketched in Fig. 4.
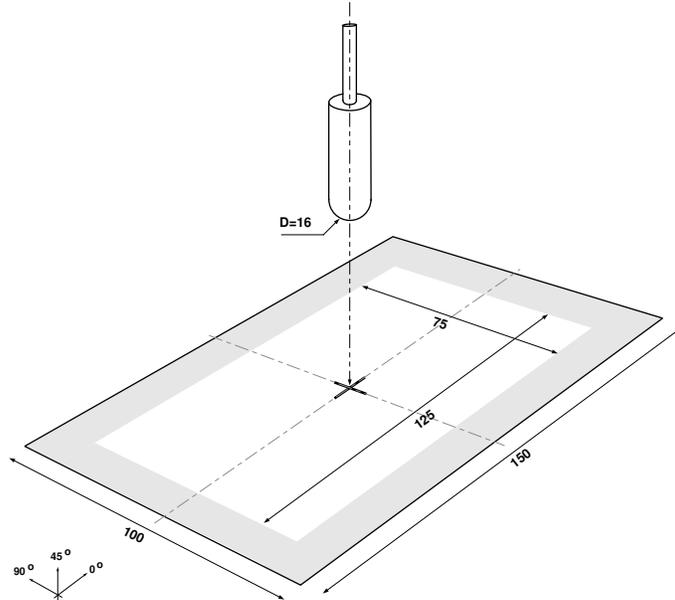
Figure 4: Standard impact test setup. Dimensions are in millimetres

The impact test consisted in dropping a spherically shaped steel dart (16mm in diameter) onto a 150x100mm composite laminate specimen clamped on all sides. In Fig.4, the clamped areas are represented by the light-shaded regions. The tested impact energies varied from 10 to 50J. Due to the availability of all the properties required by the implemented damage model, the IM7/855-2 carbon-epoxy system was chosen as ply material for the simulated composite laminate impact specimens.

2   Results

Promising preliminary results are available for the analyses proposed. A middle section cut along the longitudinal direction of a $[\pm45/90/0/45/0_4/-45/0_2]_s$ laminate specimen under a 30J impact is show in Fig. 5. The ABAQUS simulation snapshot corresponds to a moment around maximum indentation of the specimen. The colours identify the matrix damage level. The red colour represents a maximum transverse stiffness degradation of 99%.

Cohesive elements are deleted from the mesh once they delaminate. Ply interpenetration is prevented by a hard-contact kinematic constraint and sliding is constrained by friction forces. Delaminations are visible in Fig. 5 mainly between the 45° and clustered 0° layers. A superimposition of all the predicted delaminations for this test configuration is depicted in
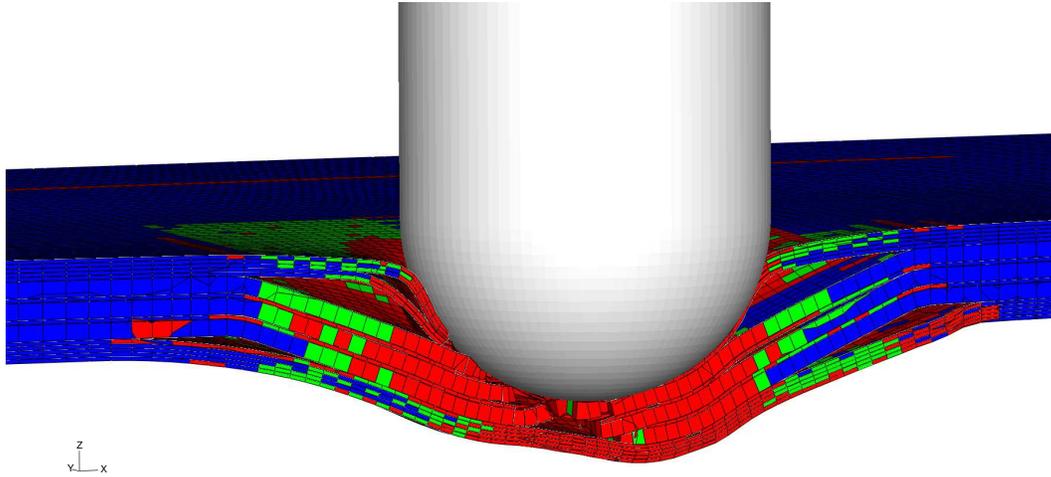
Figure 5: Fibre, matrix and interfacial damage on $[\pm 45/90/0/45/0_4/-45/0_2]_s$ laminate subjected to a 30J impact load.

Fig. 6(b) and compared against C-Scan results in Fig. 6(a). Although the exact pattern is not accurately captured, there is good correlation in terms of damaged area.

Ply elements are deleted from the FE mesh when complete failure in the fibre direction direction is identified. Fig. 5 clearly shows serious fibre breakage in the top and middle, $0°$ clustered, layers. Correlation with experimental results is not available at this point.

3   Conclusions

A high-fidelity FE analysis of the critical low-velocity impact loading events on composite laminates is proposed. The three-dimensional ABAQUS simulations take into account laminae interfacial delaminations as well as the intraply damage modes such as matrix cracking and fibre breakage.

From the preliminary results available, the physical impact event and relevant processes are being correctly predicted. Good correlation with experimental results is achieved for the delaminated area.

References

[1] ABAQUS, Inc. *ABAQUS Version 6.7 User's Manual*. Pawtucket, RI, USA, 2007.

[2] Standard test method for compressive properties of unidirectional or cross-ply fiber-resin composites. Technical report, American Society for Testing and Materials (ASTM), West Conshohocken, PA, USA, 1987. ASTM D 3410-87.
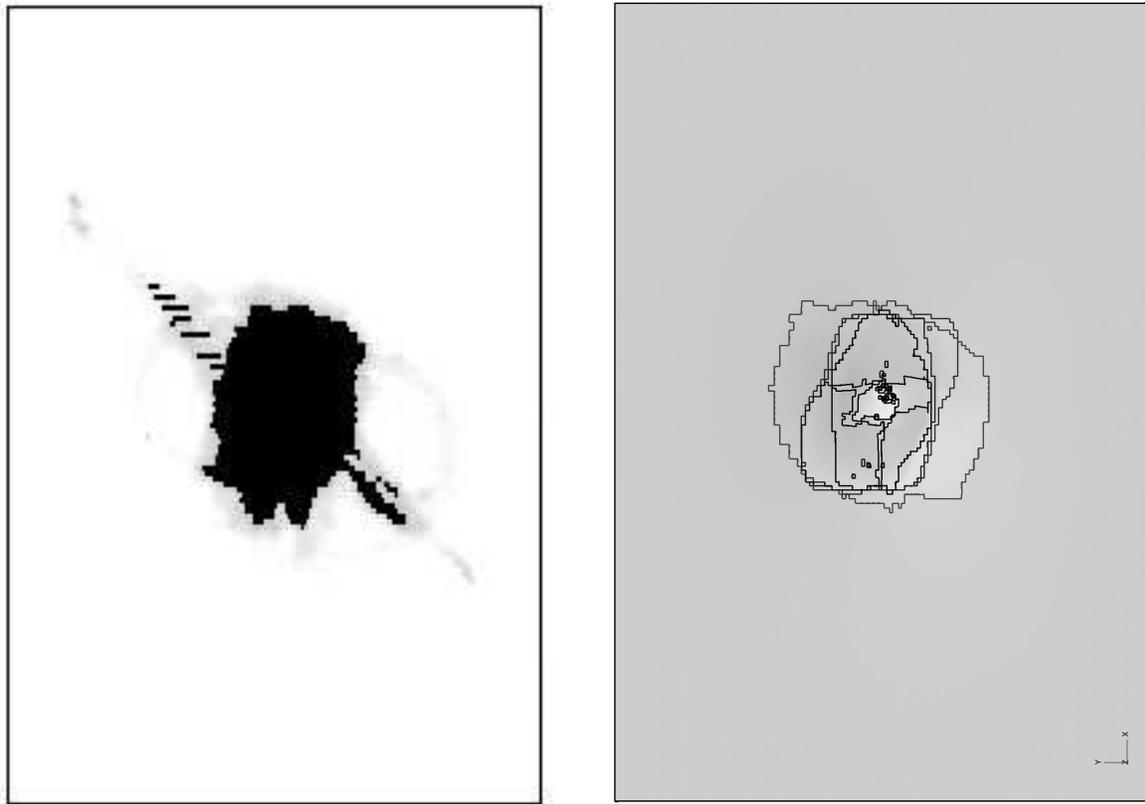
(a) C-Scan

(b) FE analysis

Figure 6: Actual and predicted delaminations on $[\pm45/90/0/45/0_4/-45/0_2]_s$ laminate subjected to a 30J impact.

[3] Standard test method for in-plane shear response of polymer matrix composite materials by test of a $\pm45\,^{\circ}$ laminate. Technical report, American Society for Testing and Materials (ASTM), West Conshohocken, PA, USA, 1994. ASTM D3518/3518M-94.

[4] Standard test method for tensile properties of polymer matrix composite materials. Technical report, American Society for Testing and Materials (ASTM), West Conshohocken, PA, USA, 2000. ASTM D 3039/D 3039M-00.

[5] Standard test methods for mode i interlaminar fracture toughness of unidirectional fiber-reinforced polymer matrix composites. Technical report, American Society for Testing and Materials (ASTM), West Conshohocken, PA, USA, 2001. ASTM D 5528-01.

[6] Standard test method for measuring the damage resistance of a fiber-reinforced polymer matrix composite to a drop-weight impact event. Technical report, American Society for

Testing and Materials (ASTM), West Conshohocken, PA, USA, 2005. ASTM D 7136/D 7136M-05.

[7] Z.P. Bažant and M. Jirásek. Nonlocal integral formulations of plasticity and damage: survey of progress. *Journal of Engineering Mechanics*, 128(11):1119–1149, 2002.

[8] P. Maimí, P.P. Camanho, J.A. Mayugo, and C.G. Dávila. A continuum damage model for composite laminates: Part i - constitutive model. *Mechanics of Materials*, 39:897–908, 2007.

[9] P. Maimí, P.P. Camanho, J.A. Mayugo, and C.G. Dávila. A continuum damage model for composite laminates: Part ii - computational implementation and validation. *Mechanics of Materials*, 39:909–919, 2007.

[10] R. Martin, T. Elms, and S. Bowron. Characterization of mode ii delamination using the 4enf. In *Proceedings of the 4th European Conference on Composites: Testing and Standardisation*, Lisbon, Portugal, 1998.

[11] S.T. Pinho, C.G. Dávila, P.P. Camanho, L. Iannucci, and P. Robinson. Failure models and criteria for FRP under in-plane or three-dimensional stress states including shear non-linearity. Technical report, NASA, Langley Research Center, Hampton, VA, February 2005. NASA/TM-2005-213530.

[12] S.T. Pinho, P. Robinson, and L. Iannucci. Fracture toughness of the tensile and compressive fibre failure modes in laminated composites. *Composites Science and Technology*, 66(13):2069–2079, 2006.

[13] A. Turon, P.P. Camanho, J. Costa, and C.G. Dávila. A damage model for the simulation of delamination in advanced composites under variable-mode loading. *Mechanics of Materials*, 38:1072–1089, November 2006.

Homogenisation of composite materials in thermoelasticity: formal mathematics and computational issues

J. Pinho-da-Cruz[*]   J.A. Oliveira[†]   F. Teixeira-Dias[‡]

Abstract

The thermomechanical modelling of the behaviour of metal matrix composites requires a deep understanding of physical phenomena that span two or more spatial length scales, since the effective overall behaviour of the composite material is affected by both microstructural details and metal matrix composite's constituent thermoelastic properties. In this context, the asymptotic expansion homogenisation method — mathematical theory of homogenisation — is a highly suitable technique to address this multiscale dependence. With this approach it is possible to estimate the effective thermoelastic properties of complex microstructures. In contrast to other homogenisation approaches, this methodology is useful due to its inherent capability to perform both the homogenisation and the localisation, by considering the microstructural details on a macrolevel analysis.

In this work, the mathematical formulation of the linear (uncoupled and quasi-static) thermoelastic homogenised problem is referred. This problem is obtained considering formal asymptotic expansions of both the temperature and the displacement fields, applying the classical multiple-scale method — widely used in Mechanics and Physics for the resolution of problems containing several small parameters describing different scalings. Computational issues related to the finite element implementation of the asymptotic expansion homogenisation method and its applications are also referred.

Finally, the asymptotic expansion homogenisation methodology is applied to the two-scale modelling of the thermomechanical behaviour of an infinite hollow tube made from aluminium matrix composite reinforced with silicon carbide. The prescription of different temperatures to its inside and outside walls, leading to a non-linear temperature field along the wall thickness, results in an expansion of the cylinder and induces stress fields in its interior. In order to illustrate one of the foremost advantages of the asymptotic expansion homogenisation methodology — the localisation, which allows the study of the microscale behaviour based on the macroscale results —, some numerical results of both the heat flux and the equivalent stress microstructural fields are presented.

[*]Department of Mechanical Engineering — University of Aveiro. E-mail:jpc@ua.pt.

[†]Department of Mechanical Engineering — University of Aveiro. E-mail:jalex@ua.pt.

[‡]Department of Mechanical Engineering — University of Aveiro. E-mail:ftd@ua.pt.

The homogenisation of physical properties is a concept studied since the 19th century [9, 6, 10]. However, the first published work about the mathematical theory of homogenisation dates back only to the end of the 20th century 60s' [14, 12, 4]. Since then, several distinct homogenisation techniques were proposed, from which the asymptotic expansion homogenisation (AEH) procedure arises as one of the most important [2, 13, 5, 1, 7]. This method is used to solve problems that involve physical phenomena on continuous media with a periodic microstructure. In particular, it's a useful technique on the study of the behaviour of structural components built with composite materials. The main advantages of this methodology lie on the fact that it allows a significant reduction of the problem size and has the capability to characterise heat flux, stress and deformation microfields, in a process usually called localisation. The localisation process consists in the opposite process of the homogenisation.

Consider a homogenous linear thermoelastic material associated to a body Ω, being its microstructure made from the spatially periodic distribution of a representative unit-cell, associated with a body Y, as shown in figure 1. Most heterogeneous materials have a small relation ε between the characteristic dimensions of the micro- and macroscale ($\varepsilon \ll 1$). Because of this, the thermomechanical loading of these materials results in periodic oscillations of the resulting microfields. These oscillations result from the periodicity of the microstructural heterogeneities and manifest themselves on a boundary of dimension ε of any point in Ω. It is therefore usual to assume the existence of two separate dimension scales, $\mathbf{x}$ and $\mathbf{y}$, related to material behaviour phenomena at the macroscale Ω and microscale Y levels, respectively (*vd.* figure 1). Thus, the variables related to these fields functionally depend on both $\mathbf{x}$ and $\mathbf{y}$, where $\mathbf{y} = \mathbf{x}/\varepsilon$.
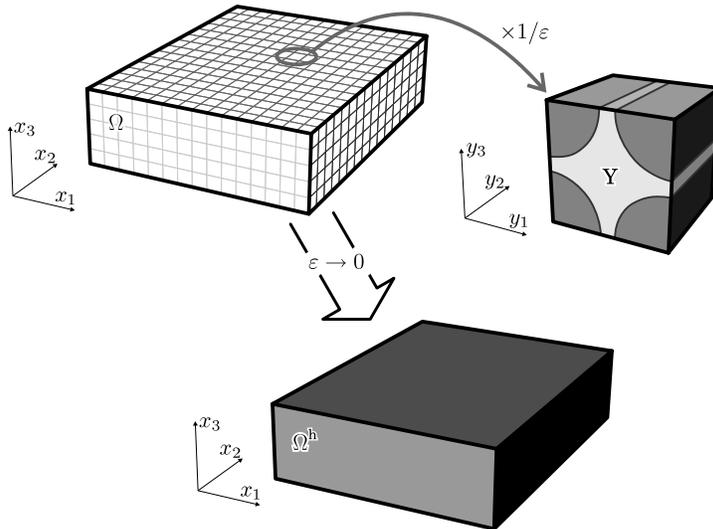

Figure 1: Schematic representation of the heterogeneous elastic material body Ω and the unit-cell Y used in the AEH, which results, with $\epsilon \to 0$, in the homogenous material Ω^{h}.

The analysis of the two distinct **x** and **y** scales comes from the possibility that, due to the periodicity of the microstructural heterogeneities, the resulting temperature and displacement fields exhibit periodic oscillations. These oscillations are superimposed on macroscopic fields that don't take into account the influence of microstructural details. Thus, these oscillations can be interpreted as fluctuations around an average macroscopic value.

In this context, the AEH method consists in an exact mathematical technique, through which one can solve a problem associated with a differential partial operator with high frequency periodic variations of its coefficients in a simpler fashion, solving, for the thermoelasticity problem, two uncoupled problems associated with differential operators with constant coefficients, which is called the homogenised thermoelasticity problem. The coefficients of the homogenised problem are determined from the solution of three problems defined on the microscale unit-cell, enforcing periodic constraints on its boundaries (a more detailed presentation and analysis of these issues can be consulted in [3]). Therefore, on a numerical sense, one of the main advantages of this method lies on a significant reduction of the number of degrees of freedom of the thermoelasticity problem. In fact, this technique allows the modelling of the microstructural details of the composite material with a single representative unit-cell, being the macroscale modeled as an equivalent homogenous body. On the other hand, another advantage of the AEH method is that it allows the characterisation of the microstructural heat flux, stress and deformation fields. In fact, unlike other usual homogenisation methods, the AEH method explicitly defines equations for this purpose in the localisation process.

In order to illustrate the application of the AEH method, let's consider the thermoelastic behaviour of an infinitely long and thick-walled cylindrical tube subjected to different temperatures at its inside and outside walls. In this situation, a nonlinear temperature field over the thickness will result. The cylinder will expand and this fact will lead to the development of a stress field. It is assumed that the thick-walled cylindrical tube is made from an aluminium matrix composite reinforced with spherical silicon carbide particles. The thermoelastic properties of the matrix (m) and of the reinforcement (r) of the composite material are presented in table 1 [3].

The cylinder with inside radius, $r_{\text{int}} = 1$ cm, and outside radius, $r_{\text{ext}} = 2$ cm, has an inside wall temperature, $T_{\text{int}} = 400$ K, and an outside wall temperature, $T_{\text{ext}} = 300$ K. A reference temperature $T_0 = 300$ K was considered. Since the problem is intrinsically axisymmetric, with a two-dimensional deformation field, the strain is symmetric with respect to the rotation axis and the shear strains are zero. Therefore, the problem is independent of the angular coordinate and all quantities only depend on the radial coordinate. The numerical modelling was performed using 20-node hexahedral finite element, considering just one-fourth of the surface and a width of 1 mm. Based on a numerical convergence study [11], the obtained mesh of finite element is illustrated in figure 2.

Table 1: Thermoelastic properties of the matrix (m) and reinforcement (r) materials of the aluminium matrix composite reinforced with spherical silicon carbide particles.

Property	Value
Matrix elastic modulus, E_{m} [GAP]	70.2
Matrix Poisson coefficient, ν_{m} [-]	0.33
Matrix thermal conductivity coefficient, k_{m} [Wa^{-1}K^{-1}]	171
Matrix thermal expansion coefficient, α_{m} [K^{-1}]	23×10^{-6}
Reinforcement elastic modulus, E_{r} [GAP]	410.0
Reinforcement Poisson coefficient, ν_{r} [-]	0.24
Reinforcement thermal conductivity coefficient, k_{r} [Wa^{-1}K^{-1}]	200
Reinforcement thermal expansion coefficient, α_{r} [K^{-1}]	$4,3 \times 10^{-6}$
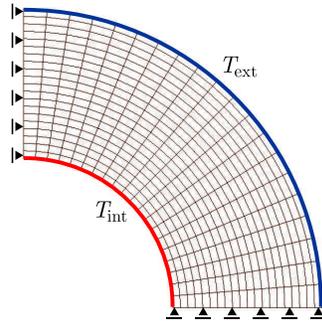
Figure 2: Numerical model of the thick-walled cylinder.

Five distinct representative unit-cells with spherical particles reinforcement fractions of 10%, 20%, 30%, 40% and 50% where generated using the SPHERECELL program [8]. Figure 3 shows the variation of the maximum values of the radial heat flux, q_{rr}, with the reinforcement fraction, f_{r}. It can be seen that the maximum values of the radial heat flux increase with the reinforcement fraction, since the thermal conductivity of the reinforcement material is greater than the one of the matrix material (cf. table 1). This kind of study can also be made in relation to the macrostructural deformation and stress fields, enabling the use of the AEH methodology to predict the thermomechanical macrostructural behaviour of the heterogeneous material. On the other hand, as previously referred, the AEH method also allows the characterisation of the microstructural heat flux, stress and deformation fields. In this context, the process of the localisation of the vector heat flux field and of the scalar equivalent stress field, performed in two distinct integration points of the macrostructural finite element mesh, are presented, along with the correspondent macrostructural fields, in figures 4(a) and 4(b), respectively. In figure 4(a), it can be seen that the localised vector heat flux fields preserve the directions associated to the macroscale field that corresponds to
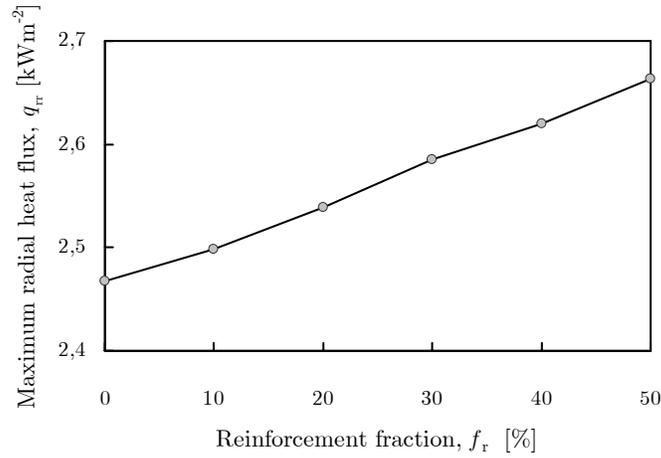
Figure 3: Variation of the maximum values of the radial heat flux, q_{rr}, with the reinforcement fraction, f_{r}.

the position of the integration points. On the other hand, in accordance with figure 4(b), the equivalent stress in the reinforcement is greater than the one in the matrix. In general, this analysis is very important since it enables the identification of the most critical zones of the microstructure, namely the zones of damage initiation and, in more general situations, of plastification.

(a)   (b)
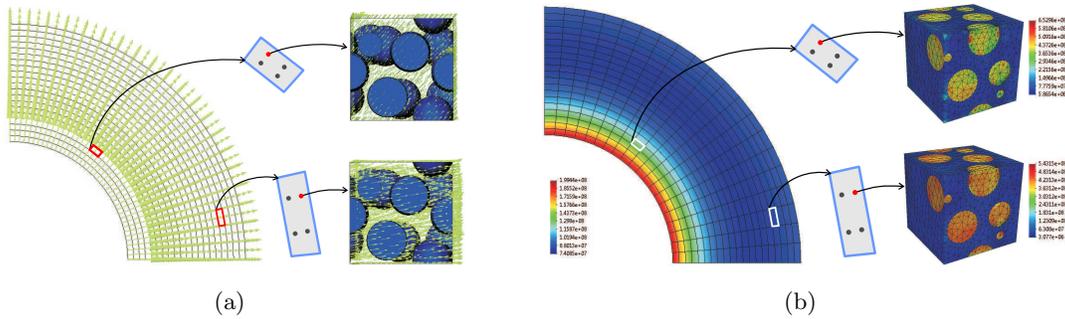
Figure 4: Schematic representation of the localisation process of (a) the vector heat flux field and (b) the scalar equivalent stress field in two distinct integration points of the macroscale finite element mesh.

References

[1] N. S. Bakhvalov and G. P. Panasenko. *Homogenization: averaging processes in periodic media*, volume 36 of *Mathematics and its Applications (Soviet Series)*. Kluwer Academic

Publishers, Dordrecht, Netherlands, 1989.

[2] A. Bensoussan, J. L. Lions, and G. Papanicolaou. *Asymptotic analysis for periodic structures.* North-Holland, Amesterdam, Netherlands, 1978.

[3] J. Pinho da Cruz. *Thermomechanical characterisation of multiphasic materials using homogenisation procedures.* PhD thesis, Department of Mechanical Engineering, University of Aveiro, Portugal, 2007. in Portuguese.

[4] E. De Giorgi and S. Spagnolo. Sulla convergenza degli integrali dell'energia per operatori ellitici del secondo ordine. *Bollettino della Unione Matematica Italiana*, 8(4):391–411, 1973.

[5] J. L. Lions. *Some methods in the mathematical analysis of systems and their control.* Gordon & Breach Science Publishers, New York, United States of America, 1981.

[6] J. C. Maxwell. *A Treatise on electricity and magnetism.* Clarendon Press, Oxford, United Kingdom, 1873.

[7] O. A. Oleĭnik, A. S. Shamaev, and G. A. Yosifian. *Mathematical problems in elasticity and homogenization*, volume 26 of *Studies in Mathematics and its Applications.* North-Holland Publishing Co., Amesterdam, Netherlands, 1992.

[8] J. A. Oliveira. Micromechanical modelling of the behaviour of aluminium matrix composites. Master's thesis, Department of Mechanical Engineering, University of Aveiro, Portugal, 2006. in Portuguese.

[9] S. D. Poisson. Second mémoire sur la théorie du magnétisme. *Académie Royale des Sciences de l'Institut de France*, 5:488–533, 1824.

[10] J. W. Rayleigh. On the influence of obstacles arranged in rectangular order upon the properties of a medium. *Philosophical Magazine*, 34(5):481–502, 1892.

[11] A. R. Rizzo. Estimating errors in FE analyses. *Mechanical Engineering*, 113(5):61–63, 1991.

[12] E. Sanchez-Palencia. Solutions périodiques par rapport aux variables d'espace et applications. *Compte-rendus de l'Académie des Sciences, série A-B*, 271:A1129–A1132, 1970.

[13] E. Sanchez-Palencia. *Non-homogeneous media and vibration theory*, volume 127 of *Lectures Notes in Physics.* Springer-Verlag, Berlin, Germany, 1980.

[14] S. Spagnolo. Sulla convergenza di soluzioni di equazioni paraboliche ed ellitiche. *Annali della Scuola Normale Superiore di Pisa*, 22(3):571–597, 1968.

An eigenfrequency approach to characterize and optimize 1D two material structures for axial vibration attenuation

H. Policarpo[*]   O. M. Silva[†]   M. M. Neves[‡]

Abstract

Bloch wave theory has been used to obtain the dispersion curves of elastic wave propagation through structures built of periodic distributions of materials. By assumption, the periodicity repetition is infinite implying the existence of pass-bands and stop-bands. These techniques have been used to design periodic structures and to extrapolate for the case of finite periodicity repetition which result in considerable attenuations instead of stop-bands. Eigenfrequency approach can be equally used, and we found, in this case, it more intuitive for finite periodicity repetition as well as for the non periodic case. We propose a new perspective of the problem choosing to plot frequency versus frequency number against dispersive curve plot and frequency response plot. We apply a numerical technique, combining finite element analysis (FEA) and optimization techniques, to design the one-dimensional two-material periodic, or non periodic, that are optimized to maximize the separation of two adjacent axial eigenfrequencies. Two optimization problems are considered here. The design variables are in the first the proportion of each material (in axial length) in the respective cell, and in the second the two-material distribution (in mixture density) in the structure for a fixed cell size. Both cases results in a significant attenuations for the case of longitudinal vibrations. Then we combine it with a study on materials selection for the frequency range of interest. Some results are presented to illustrate the potential of this method.

1 Introduction

To study structures built with infinite periodicity repetition, Bloch's theorem (see, e.g., [1]) — also related with Lyapunov-Floquet theorem (see, e.g., [6] and [7]) — can be used to obtain a characterization of longitudinal waves leading to the corresponding dispersion relation

---

[*]MSc Mechanical Eng. student at IST - TU of Lisbon - Portugal. E-mail: `hugo.policarpo@ist.utl.pt`

[†]PhD Mechanical Eng. student - GTVA - Federal University of Santa Catarina - Florianopolis - Brazil. E-mail: `olavo@lva.ufsc.br`

[‡]Professor IST, TU of Lisbon, IDMEC-Institute of Mechanical Engineering. E-mail: `maneves@dem.ist.utl.pt`

that will present a band structure analogues to those of electrons in a crystal and light in supperlattices (see, e.g., [1]). In mediums with periodic heterogeneity, there are ranges of frequencies, known as pass-bands and stop-bands, over which all incident waves are effectively allowed to propagate and effectively stopped, respectively. This frequency-banded dynamic response has interest across multiple disciplines. For structures with finite periodic repetition, researchers are still reviewing and trying to improve existing models for vibration's problems. A description of the basics can be found in, e.g., [3] and [10], while for structural optimization for longitudinal vibrations it can be found, among others, in [12], [5], [9], [8] and [13]. We propose an alternative method in which is used the relation between the eigenfrequency results, the dispersion relations and frequency response. This can be combined with a selection of materials taken for a frequency range of interest. A numerical eigenfrequency (modal) approach based on finite element method is used. To demonstrate the advantage of the proposed method, two structural optimization problems are introduced in which the design variables are in the first, the proportion of each material (in axial length) of each repetitive cell and in the second the two-material distribution in the structure for a fixed cell size. Both cases results in significant attenuations for the case of longitudinal vibrations. The models presented are limited to the elastic bar (i.e., link or rod elements) and each material is assumed to be isotropic with linear elasticity. No damping effects are considered. The numerical implementation technique used in the present work combines commercial code ANSYS for finite element analysis with optimization algorithms running on MatLab environment (see, e.g., [4]). To demonstrate its applicability at frequencies of interest, a plot of results for different materials is presented.

2   Models for structural analysis

The structural responses considered in the present study are restricted to the axial vibrations of the elastic bar, where the materials are assumed to be isotropic with linear elasticity and no damping effects are considered.

2.1   Theoretical bar model for a periodic structure with infinite repetitively

Consider the structure with parallel alternating layers of two-material differing in elastic properties as shown in the Figure 1.

The axial vibrations are governed by the diferential equation:

$$\rho_\alpha \frac{\partial^2 u^\alpha}{\partial t^2} - E_\alpha \frac{\partial^2 u^\alpha}{\partial x^2} = f, \ \alpha = 1, 2, \dots \tag{1}$$

where $E_\alpha(x)$ is the modulus of elasticity, $\rho_\alpha(x)$ is the mass density and f is the external force. The displacements $u^\alpha(x, t)$ and stresses $\sigma^\alpha(x, t)$ will be considered continuous at the layer boundaries.
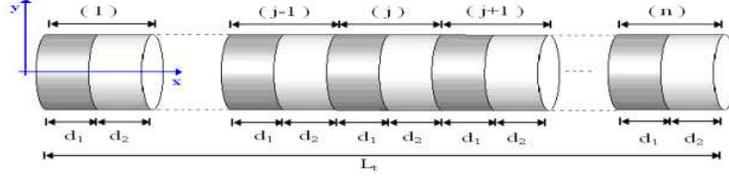
Figure 1: Periodic rod made of two different materials with different length in each repetitive cell.

2.2 *Bloch* wave technique

By virtue of the structure's periodicity, eigenmodes can be characterized by a quasi-wave number k, considering that the field in a repetitive cell takes the form,

$$u_n(x) = u_0(x - nd)e^{iknd}, \ n = 1, 2, \dots \tag{2}$$

where n is the cell's number and d the structure's period. For this case, the dispersion equation obtained from (1) by [11] determines the frequency as an implicit function of the quasi-wave number and it is given by

$$\cos(kd) = \cos(k_1 d_1)\cos(k_2 d_2) - \frac{1}{2}\left(\frac{k_1}{k_2} + \frac{k_2}{k_1}\right)\sin(k_1 d_1), \tag{3}$$

where $k_1 = \frac{\omega}{c_1}$ and $k_2 = \frac{\omega}{c_2}$ being ω the frequency and c_1 and c_2 the wave's velocity in the type 1 and 2 layers. If d_1 and d_2 represent the layer's thicknesses then, the structure's period is given by $d = d_1 + d_2$. Equation (3) allows us to compute and plot the frequency as an implicit function of the quasi-wave number for this type of structure.

2.3 Structural modal and harmonic response via finite element analysis

The finite element used is a 2D spar element with uniaxial tension — compression and two degrees of freedom at each of the two nodes: translations in the nodal x and y directions. In our case, we allow the nodes to displace only in the axial direction obtaining the following elementar stiffness $[K_e]$, mass $[M_e]$ and force $\{F_e\}$ matrices of an element with length h_e,

$$K_e = \frac{E_\alpha A}{h_e}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \quad M_e = \frac{\rho_\alpha h_e}{6}\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \quad F_e = \left\{ \begin{array}{c} F_1 \\ F_2 \end{array} \right\} e^{jwt} \tag{4}$$

where A is the element's cross-sectional area.

The structural analysis via finite element method, and after the respective matrices assemblage, is expressed by a modal problem (5) to obtain the eigenvalues (eigenfrequencies) or by a harmonic analysis (6) to obtain the frequency response, respectively:

$$\left([K] - \omega_i^2[M]\right)\{\phi_i\} = \{0\}, \tag{5}$$

$$\left([K] - \omega_{ap}^2[M]\right)\{U\} = \{F_\omega\}. \tag{6}$$

In the above expressions, ω_i is the eigenfrequency, ϕ_i the eigenmode, ω_{ap} the applied force excitation frequency, U the respective maximum displacement and F_ω the magnitude of the applied force.

3 Optimization of the separation of two adjacent eigenfrequencies

In this section, we introduce two optimization problems in which the design variables are in the first the proportion of each material (in axial length) of each repetitive cell and in the second the two-material distribution in the structure.

3.1 Structural optimization of the proportion of the material in each repetitive cell (periodic case)

Consider a periodic rod made of two different materials, where the objective is to optimize the portion of each material, in axial length, of each repetitive cell, as illustrated by Figure 1. Doing so, we maximize the separation of two specified adjacent eigenfrequencies ω_{i+1} and ω_i , i.e for a given i $(i = 1, 2, 3, 4, 5, \ldots)$. The size optimization problem is formulated as:

$$\begin{aligned}
\max_{d_2} \quad & (\omega_{i+1} - \omega_i) \\
\text{subject to} \quad & \left([K] - \omega^2[M]\right)\{\phi_i\} = \{0\} \\
& d = d_1 + d_2 \\
& n \times d \leq L_t
\end{aligned} \tag{7}$$

The design variable in the repetitive cell is d_2 where the total number of cells is n and the total length of the rod, L_t, are known.

3.2 Structural optimization of the distribution of two materials in each design variable (general case)

In this section, the objective is to find an optimized distribution of two different material components along a given rod, as the distribution illustrated by Figure 2. The distribution should maximize the separation of two specified adjacent eigenfrequencies ω_{i+1} and ω_i , i.e. for a given i $(i = 1, 2, 3, 4, 5, \ldots)$.

A topology optimization method is formulated as:

$$\begin{aligned}
\max_{t} \quad & (\omega_{i+1} - \omega_i) \\
\text{subject to} \quad & \left([K] - \omega^2[M]\right)\{\phi_i\} = \{0\} \\
& 0 \leq t_j \leq 1, \;\; j = 1, 2, \ldots, NDes.
\end{aligned} \tag{8}$$
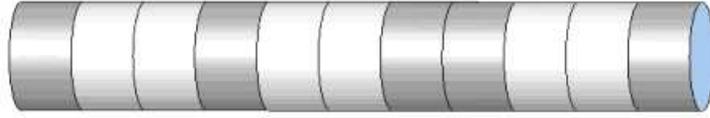
Figure 2: Example of a distribution of two material components along the rod.

The formulation follows the presented in [9] where the material interpolation proposed for this problem is different from the usual SIMP technique described by [2]. It is expressed by the following relations:

$$E = \frac{E_1}{1 + t\left(\frac{E_1}{E_2} - 1\right)} \text{ and } \rho = \rho_1 + t(\rho_2 - \rho_1). \tag{9}$$

The design variable is t_j and it has one value in the interval between $t_j = 0$ (where 0 means only material 1 is present) and $t_j = 1$ (where 1 means only material 2). To relax the problem for a continuum variation of the parameter t, mixtures between these materials are allowed and represented by intermediate values of t. Material properties are then computed by (9). Each finite element is associated with one value of t although each design variable can be assigned to several consecutive finite elements.

For the numerical optimizations we defined an iterative procedure with the following steps:

1. Initialize by giving initial design values and the order i of the lowest eigenvalue to separate;

2. Run a harmonic analysis for the initial design;

3. Start the optimization loop in Matlab by calling the `fmincon` routine. After writing the values of the current iteration design variables for a file it will call ANSYS program in batch mode and will return to the fmincon routine the current calculated values of eigenfrequencies ω_{i+1} and ω_i; and

4. After satisfying the stopping criteria, run a harmonic analysis for the final design.

4 Numerical results

Here, a given rod is optimized for the formulation of sections 3.1 and 3.2 (to maximize the separation of two adjacent eigenfrequencies). For all the presented results, it is assumed that material 1 is urethane (UR) and the material 2 is steel (ST). The rod has a uniform diameter of 0.01m, a total length of L_t=0.036m and the two material properties are: $E_{ST} = 200.0$ GPa, $\rho_{ST} = 7860$kg/m^3, $E_{UR} = 0.003$ GPa and$\rho_{UR} = 1142$kg/m^3. Bar finite elements

(element LINK1 in ANSYS) were used, observing that the number of finite elements should be adjusted taking in account the highest frequency in analysis and the finite element length. Finite element analysis are done at ANSYS® with no transversal displacements, $Uy = 0$ for all nodes, and $Ux = 0$ at $x = L_t$ (see Figure 1).

4.1 Proposed method - Design of structures based on modal analysis

In order to confirm the relation between frequency plots associated with a finite repetion and dispersion curves associated with an infinite repetion, the rod is analyzed for different cell numbers ($n = 2, \ldots, 9$). As illustrated in Figure 3 a) and 3 b) for the first and second attenuation bands, respectively, such a relation exists. Figures 3 a) and Fig. 3 b) are obtained from two different approaches. The first, a modal analysis (Equation 5), with n cells for a given L_t (see Figure 1), allow us to identify the lower and upper bounds of the first and second attenuation bands. The second, a *Bloch* wave analysis which consists on solving Equation 3, allow us to compute and plot the frequency as an implicit function of the quasi-wave number, where the stop-bands are identified for k real, positive and constant in frequency or for k imaginary and positive. This way we identify the lower and upper bounds of the first and second attenuation bands. For the second attenuation band (see Figure 3 b)) the results obtained via modal and *Bloch* analysis present a slight difference in the results for the upper bound (ω_{i+1}) that we associated to the number of finite elements used and to the move from the infinite to the finite repetion case.
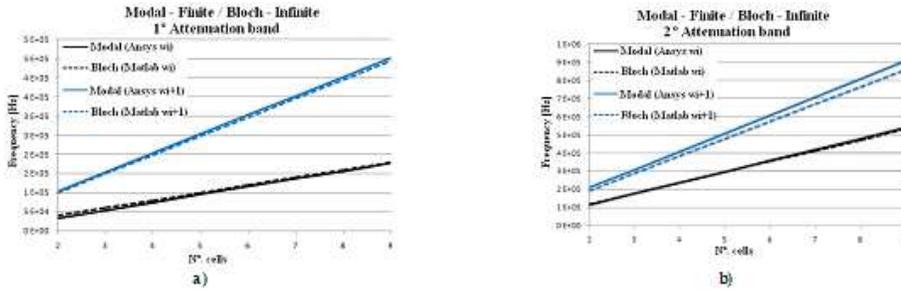
Figure 3: Limit curves of the band gap versus number of cells: a) first attenuation gap and b) second attenuation gap.

4.2 Results from the maximization of the separation of two adjacent eigen-frequencies ω_{i+1} and ω_i for a periodic material distribution (periodic case)

The number of cells n used were 3 and 7, respectively. The initial design consists in the design variables $d2$ being $L_t/2n$. Improvement in the separation of two adjacent frequencies (i and $i+1$), relatively to the initial design, i.e. from initial separation Δ_{ini} to optimized final value Δ_{fin} were obtained as shown by the values in Figures 4 a) and 5a), for $n = 3$ and 7, respectively. The dispersion curves obtained are presented in Figures 4 b) and 5 b). The frequency response of displacement at node $x = (L - (L/n))$ the node that separates left variables from the last node is presented at Figures 4 c) and 5 c) for an applied axial harmonic force with 200 N magnitude at $x = 0$ m. The attenuation bands are identified (by the gray areas) to relate it with the stop-bands given by the dispersion curves.
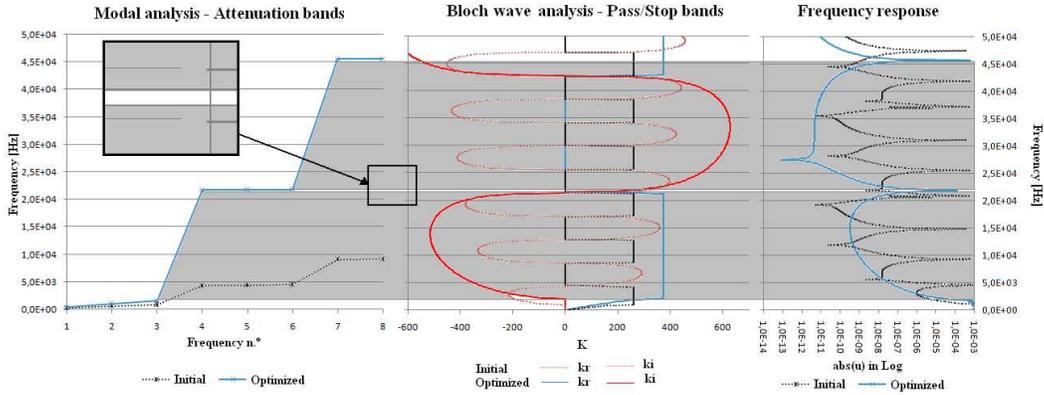
Figure 4: Results for $n = 3$ cells: a) Eigenfrequencies curve plot; b) Dispersion curves and c) Frequency response plot.

Table 1: Output results from the optimization cases with $i = 3$ and 7.

n.º of cells	3	7
d_2 initial [m]	6.00×10^{-3}	2.60×10^{-3}
d_2 Optimized [m]	1.20×10^{-3}	0.50×10^{-3}
Δ initial [Hz]	3.44×10^3	7.83×10^3
Δ optimized [Hz]	20.16×10^3	46.72×10^3

These separations of adjacent eigenvalues are evident when comparing frequency response curves from initial design (see Figures 4 c) and 5 c)) with corresponding curves of optimized designs for $i = 3$ and 7, respectively.
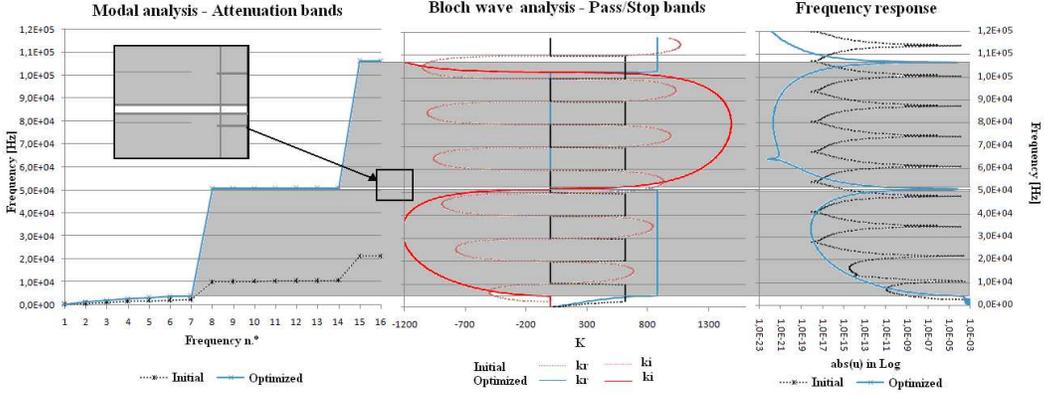
Figure 5: Results for $n = 7$ cells: a) Eigenfrequencies curve plot; b) Dispersion curves and c) Frequency response plot.

Table 2: Output results from the optimization cases with $i = 1$ and 5.

i	1	5
Δ initial [Hz]	0.507×10^3	0.512×10^3
Δ optimized [Hz]	2.40×10^3	6.32×10^3

4.3 Results from the maximization of the separation of two adjacent eigenfrequencies ω_{i+1} and ω_i by a distribution of two different materials in each design variable (general case)

The number of design variables t_n used is eleven, all with the same length. The initial design consists in all design variables t_n being 0.5, as indicated by a blue line in Figures 6 c) and 7 c). Using the optimization procedure described in Section 3.2, the optimized designs were obtained for a specific i value and presented at mentioned figures, the red line in Figures 6 c) and 7 c), where in the first it indicates a sequence of materials: UR/ST/UR/ST. Improvement in the separation of two adjacent frequencies (i and $i+1$), relatively to the initial design were obtained as shown by the values in Figures 6 a) and 7 a), for $i = 1$ and 5, respectively. The frequency response of displacement at node $x = (L - (L/11))$ in Figures 6 b) and 7 b) for an applied axial harmonic force with 200 N magnitude at $x = 0$ m.

4.4 A material selection on the frequency range of interest

For a structure in the conditions of the implemented model one can recur to Figure 8 and select the suited for a particular frequency range of interest. The limit values presented were obtained considering the model used in Section 3.1 with the total number of cells n,
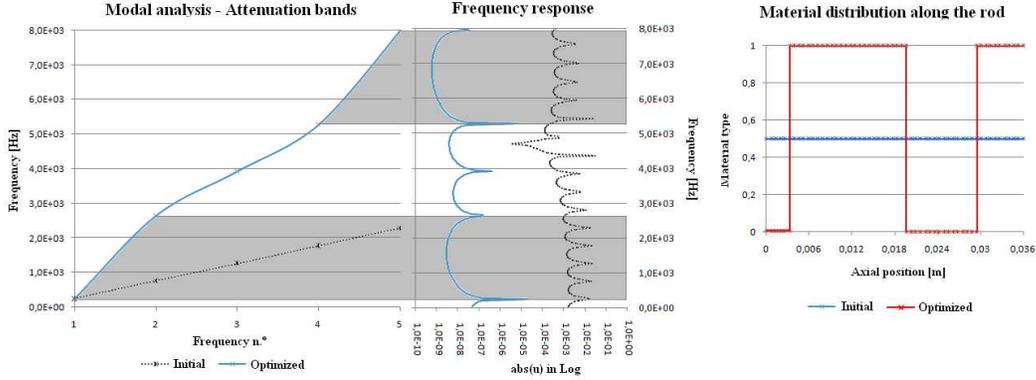
Figure 6: Results for $i = 1$: a) Frequency response plot ; b) Frequency response plot and c) Material Distribution in the rod.
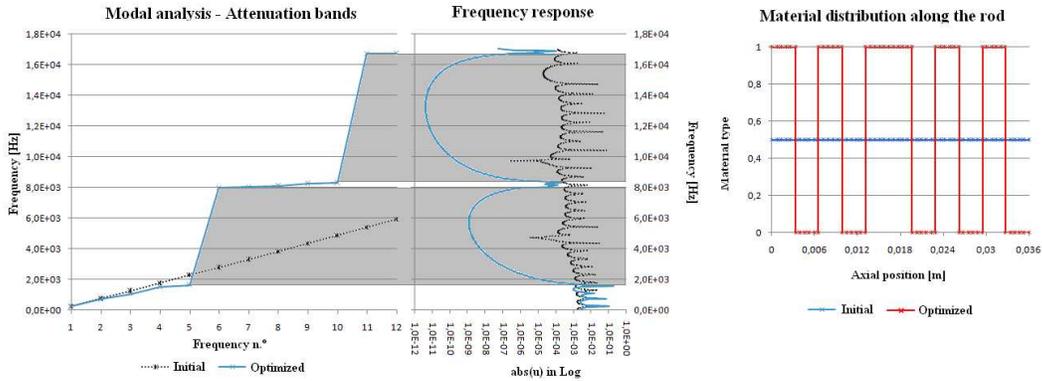
Figure 7: Results for $i = 5$: a) Frequency response plot ; b) Frequency response plot and c) Material Distribution in the rod.

$6 \leq n \leq 12$, the total length of the rod L_t, 0.036m $\leq L_t \leq$ 0.1m, the total length of the rod L_t, 0.036m $\leq L_t \leq$ 0.1m and a relative material density ρ_v, $0.5 \leq \rho_v \leq 1$.

The left points (lower bounds) of each horizontal line are calculated for $n = 6$; $L_t = 0.1$m; $\rho_v = 0.5$ and the right points (upper bounds) of each horizontal line are calculated for $n = 12$; $L_t = 0.036$m; $\rho_v = 1$. The shaded area represents the use of foam to model the material with lowest modulus of elasticity in the structure. The foam model used is: $E_{\text{foam}} = \rho_v^3 E$ and $\rho_{\text{foam}} = \rho_v \rho$.

5   Conclusions

Obtained values shown that, in this case, *Bloch* wave results can be similarly obtained by a modal (eigenfrequency) approach, even though the number of finite elements used should
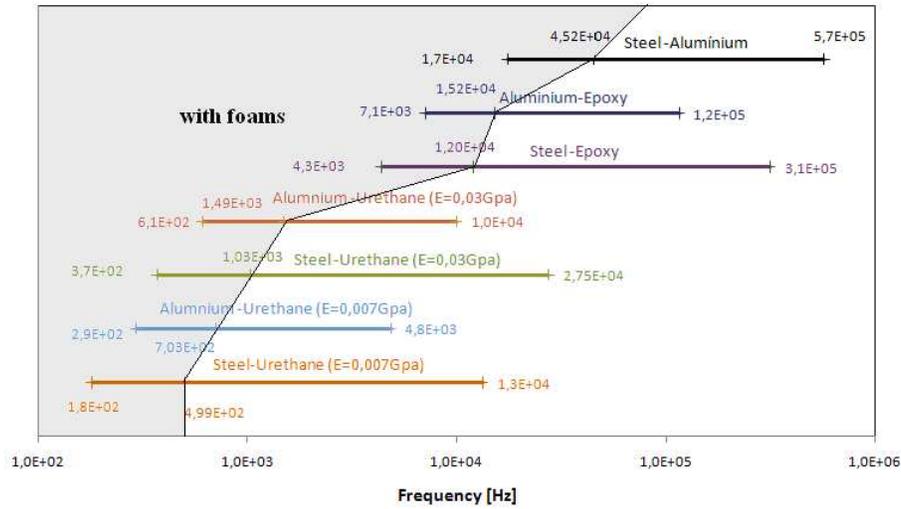
Figure 8: Results obtained for the model used in section 3.1 with different materials.

be carefully chosen. The numerical technique implementated in present work demonstrates the interest of the proposed method in which for both optimization problems was possible to get a considerable improvement in the separation of two adjacent eigenfrequencies. Eigenfrequencies curve plot show to be interesting to this kind of analysis and was not found similar in the literature. The material selection presented illustrates its dependence on the frequency range of interest for the axial excitation in which one can recur to it and select which types of materials are suited for a particular frequency range of interest. Again, we notice that the results obtained are in a frequency range of interest for applications which was not found similar in the literature. Finally, we should note the interest in future experimental verification of these results, especially with respect to the use of available foams and similar materials.

Acknowledgements

References

[1] N. W. Ashcrofat and N. D. Mermin. *Solid State Physics*. Wiley, New York, 1976.

[2] M. P. Bendsoe and O. Sigmund. *Topology Optimization – Theory, Methods and Applications*. Springer Verlag, 2003.

[3] L. Brillouin. *Wave Propagation in Periodic Structures.* Dover Phoenix Editions, 2nd edition, 1946.

[4] M. Carvalho, J. L. Vale, L. T. Pinheiro, and M. M. Neves. Techniques adopted in integrating structural analysis with numerical optimization. In *CD of Proceedings of 5th International Conference on Mechanics and Materials in Design*, Porto, 2005.

[5] A. R. Diaz, A. G. Haddow, and L. Ma. Design of band-gap grid structures. *Structural and Multidisciplinary Optimization*, 29:418–431, 2005.

[6] G. Floquet. Sur les equations differentielles lineaires à coefficients doublement periodiques. *Ann. Sci.Ecole Norm. Sup.*, 12:47–88, 1883.

[7] D. L. Goodstein. *States of a Matter.* Dover, New York, 1985.

[8] M. I. Hussein, G. M. Hulbert, and R. A. Scott. Dispersive elastodynamics of 1d banded materials and structures: analysis. *Journal of Sound and Vibration*, 307(3–5):865–893, November 6 2007.

[9] J. S. Jensen and N. L. Pedersen. On maximal eigenfrequency separation in two-material structures: The 1d and 2d scalar cases. *Journal of Sound and Vibration*, 289(4–5):967–986, 2006.

[10] D. J. Mead. Wave propagation in continuous periodic structures: research contributions from southampton: 1964-1995. *J Sound Vib*, 190(3):495–524, 1996.

[11] S. M. Rytov. Akust. zh. *Sov. Phys. Acoust. 2*, 68, 1956.

[12] O. Sigmund and J. S. Jensen. Systematic design of phononic band-gap materials and structures by topology optimization. *Phil. Trans. R. Soc. Lond.*, A 361:1001–1019, 2003.

[13] O. M. Silva, M. M. Neves, and J. L. M. Fernandes. On finite element analysis and optimization of structural properties for a maximum passive vibration attenuation. In *in Proceedings of International Conference on Modeling and Optimization of Structures, Processes and Systems, ICMOSPS'07 Cd of Proceedings*, Durban, South Africa, January 22-24 2007.

Analysis of sandwich plates by a higher-order formulation and an improved meshless method using radial basis functions

C. M. C. Roque[*]   A. J. M. Ferreira[†]   R. M. N. Jorge[‡]

Abstract

The purpose of this work is to use a meshless collocation method with multiquadric radial basis functions (RBFs), with a higher-order shear deformation theory. The multiquadric RBF method was introduced by Kansa [3, 4] for solving boundary-value problems governed by partial differential equations. The method requires from the user the indication of a shape parameter, c. This parameter can influence the quality of the solution and it's choice is not straightforward. We propose a method based on a statistical technique for choosing the shape parameter and show that this method predicts deflections of sandwich plates that agree very well with analytical solutions.

1 Introduction

The Higher Order Shell deformation Theory (HOST9) presented by Khare et al. [5] is in this paper reduced to seven degrees of freedom, and applied to the modeling of extreme flat shells (plates). The theory is briefly presented in section 4 and readers should consult reference [5] for more details. In this paper we compare the analytical solution for several sandwich plates as presented in [5] with the solution obtained with the meshless multiquadric radial basis functions method. The meshless multiquadric method, first presented by Kansa [3, 4], is well known for solving systems of partial differential equations with excellent accuracy. However, it has the problem of the choice of an adequate shape parameter for the multiquadric interpolation function. We use a cross validation technique implemented via Rippa's algorithm [6] that overcomes the choice of the shape parameter to the simple indication of a user-defined interval.

---

[*]Faculdade de Engenharia da Universidade do Porto. E-mail:`croque@fe.up.pt`.
[†]Faculdade de Engenharia da Universidade do Porto. E-mail:`ferreira@fe.up.pt`.
[‡]Faculdade de Engenharia da Universidade do Porto. E-mail:`rnatal@fe.up.pt`.

2   The meshless multiquadric method

Consider the generic Boundary Value Problem (BVP) with a domain Ω with boundary $\partial\Omega$, and some linear differential operators L and B: $Lu(x) = f(x),\ x \in \Omega \subset \mathbb{R}^n;\ Bu|_{\partial\Omega} = q$. The function $\mathbf{u}(\mathbf{x})$ is approximated by: $\mathbf{u} \simeq \bar{\mathbf{u}} = \sum_{j=1}^{N} \alpha_j g_j$, were α_j are parameters to be determined after the collocation method is applied. We consider a global collocation method where the linear operators L and B acting in domain $\Omega \backslash \partial\Omega$ and boundary $\partial\Omega$ define a set of global equations in the form

$$\begin{pmatrix} \mathbf{L}_{ii} & \mathbf{L}_{ib} \\ \mathbf{B}_{bi} & \mathbf{B}_{bb} \end{pmatrix} \begin{pmatrix} \alpha_i \\ \alpha_b \end{pmatrix} = \begin{pmatrix} f_i \\ q_b \end{pmatrix}; \qquad \text{or} \qquad \begin{bmatrix} \mathcal{L} \end{bmatrix} \begin{bmatrix} \alpha \end{bmatrix} = \begin{bmatrix} \lambda \end{bmatrix} \tag{1}$$

where i and b denote domain and boundary nodes, respectively; f_i and q_b are some external conditions in domain and boundary (in bending shells these can be external forces).

The function g represents a Radial Basis Function (RBF). In the present case, we choose the multiquadric function, defined as: $g(r,c) = (\|x - x_j\|^2 + c^2)^{\frac{1}{2}}$, were r is the euclidian distance between two nodes and c is a shape parameter that improves the function surface so that convergence gets faster [3, 4]. Other radial basis functions could be used (gaussians, splines, etc). However, multiquadrics proved to be excellent for global, smooth, boundary-value problems, like shells in bending [2].

3   An optimization technique

An optimal shape parameter c can be obtained for an interpolation problem $A\alpha = f, A = g(\|\mathbf{x}_j - \mathbf{x}_i, c\|)$, by the leave-one-out cross validation technique in regression analysis. The problem can be formulated as finding c in order to minimize a cost function given by the norm of an error vector $E(c)$ with the following components: $E_i(c) = f_i - \sum_{j=1,j\neq i}^{N} \alpha_j^{(i)} g(\|\mathbf{x}_j - \mathbf{x}_i, c\|)$. Here $\sum_{j=1,j\neq i}^{N} \alpha_j^{(i)} g(\|\mathbf{x}_j - \mathbf{x}_i, c\|)$ is the function value predicted at the i-th data point using RBF interpolation based on a set of data that excludes the i-th point.

A more efficient algorithm, from a computational point of view, is given by the following formula [6, 8]: $E_i(c) = \frac{\alpha_i}{A_{i,i}^{-1}}$, where α_i is the i-th coefficient for the full interpolation problem and $A_{i,i}^{-1}$ is the i-th diagonal element of the inverse of the corresponding interpolation matrix A. In the case of our boundary value problem, the error to be minimized is a residual error, of the form [1]:

$$E_i(c) = \lambda_i - \sum_{j=1,j\neq i}^{N} \alpha_j^{(i)} \mathcal{L} g(\|\mathbf{x}_j - \mathbf{x}_i, c\|) \tag{2}$$

Now the generalization of the cross-validation algorithm is straightforward. Our BVP is given by equation (1). We can use the following formula:

$$E_i(c) = \frac{\alpha_i}{\mathcal{L}_{i,i}^{-1}} \tag{3}$$

where α_i is the i-th coefficient for the full collocation problem (1) and $\mathcal{L}_{i,i}^{-1}$ is the i-th diagonal element of the inverse of the corresponding collocation matrix $\mathcal{L}$. Having the cost function, we use the MATLAB function `fminbnd` to find a local minimum.

4   Higher-order shear deformation theory

The equilibrium equations are derived considering the following displacement field (u, v, w) [5]:

$$u(x, y, z) = u_0(x, y) + z\phi_x(x, y) + z^3\phi_x^*(x, y) \qquad\qquad w(x, y, z) = w_0(x, y) \qquad (4)$$
$$v(x, y, z) = v_0(x, y) + z\phi_y(x, y) + z^3\phi_y^*(x, y) \qquad\qquad\qquad\qquad\qquad (5)$$

For small displacements and assuming $h/R_x, h/R_y << 1$ (being h, R_x and R_y the thickness and curvature radius, respectively), the strain-displacements relations are given by:

$$
\begin{Bmatrix} \epsilon_{xx} \\ \epsilon_{yy} \\ \gamma_{xy} \\ \gamma_{yz} \\ \gamma_{xz} \end{Bmatrix} =
\begin{Bmatrix} \epsilon_{xx}^{(0)} \\ \epsilon_{yy}^{(0)} \\ \gamma_{xy}^{(0)} \\ \gamma_{yz}^{(0)} \\ \gamma_{xz}^{(0)} \end{Bmatrix} + z
\begin{Bmatrix} \epsilon_{xx}^{(1)} \\ \epsilon_{yy}^{(1)} \\ \gamma_{xy}^{(1)} \\ \gamma_{yz}^{(1)} \\ \gamma_{xz}^{(1)} \end{Bmatrix} + z^2
\begin{Bmatrix} \epsilon_{xx}^{(2)} \\ \epsilon_{yy}^{(2)} \\ \gamma_{xy}^{(2)} \\ \gamma_{yz}^{(2)} \\ \gamma_{xz}^{(2)} \end{Bmatrix} + z^3
\begin{Bmatrix} \epsilon_{xx}^{(3)} \\ \epsilon_{yy}^{(3)} \\ \gamma_{xy}^{(3)} \\ \gamma_{yz}^{(3)} \\ \gamma_{xz}^{(3)} \end{Bmatrix} \qquad (6)
$$

$$
= \begin{Bmatrix} \frac{\partial u_0}{\partial x} + \frac{w_0}{R_x} \\ \frac{\partial v_0}{\partial y} + \frac{w_0}{R_y} \\ \frac{\partial u_0}{\partial y} + \frac{\partial v_0}{\partial x} \\ \phi_y + \frac{\partial w_0}{\partial y} - \frac{v_0}{R_x} \\ \phi_x + \frac{\partial w_0}{\partial x} - \frac{u_0}{R_y} \end{Bmatrix} + z
\begin{Bmatrix} \frac{\partial \phi_x}{\partial x} \\ \frac{\partial \phi_y}{\partial y} \\ \frac{\partial \phi_x}{\partial y} + \frac{\partial \phi_y}{\partial x} \\ -\frac{\phi_y}{R_y} \\ -\frac{\phi_x}{R_x} \end{Bmatrix} + z^2
\begin{Bmatrix} 0 \\ 0 \\ 0 \\ 3\phi_y^* \\ 3\phi_x^* \end{Bmatrix} + z^3
\begin{Bmatrix} \frac{\partial \phi_x^*}{\partial x} \\ \frac{\partial \phi_y^*}{\partial y} \\ \frac{\partial \phi_x^*}{\partial y} + \frac{\partial \phi_y^*}{\partial x} \\ -\frac{\phi_y^*}{R_y} \\ -\frac{\phi_x^*}{R_x} \end{Bmatrix} \qquad (7)
$$

By neglecting transverse normal stress, σ_z, the stress-strain relations in the local (material) cartesian system can be obtained as:

$$
\begin{Bmatrix} \sigma_1 \\ \sigma_2 \\ \tau_{12} \\ \tau_{23} \\ \tau_{31} \end{Bmatrix} =
\begin{bmatrix} Q_{11} & Q_{12} & 0 & 0 & 0 \\ Q_{12} & Q_{22} & 0 & 0 & 0 \\ 0 & 0 & Q_{33} & 0 & 0 \\ 0 & 0 & 0 & Q_{44} & 0 \\ 0 & 0 & 0 & 0 & Q_{55} \end{bmatrix}
\begin{Bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \gamma_{12} \\ \gamma_{23} \\ \gamma_{31} \end{Bmatrix} \qquad (8)
$$

where subscripts 1 and 2 are respectively the fiber and the normal to fiber inplane directions, and 3 is the direction normal to the shell. The reduced stiffness components, Q_{ij} are given by: $Q_{11} = \dfrac{E_1}{1 - \nu_{12}\nu_{21}}; Q_{22} = \frac{E_2}{1-\nu_{12}\nu_{21}}; Q_{12} = \nu_{21}Q_{11}; Q_{33} = G_{12}; Q_{44} = G_{23}; Q_{55} = G_{31}\nu_{21} = \nu_{12}\frac{E_2}{E_1}$; in which E_1, E_2, ν_{12}, G_{12}, G_{23} and G_{31} are material properties of the lamina. Note that with this higher-order theory, the use of shear correction factors is not needed.

The equilibrium equations are then obtained using the principle of virtual work, $\delta\Pi = \int_x \int_y \int_z (\sigma_{xx}\delta\epsilon_{xx} + \sigma_{yy}\delta\epsilon_{yy} + \tau_{xy}\delta\gamma_{xy} + \tau_{xz}\delta\gamma_{xz} + \tau_{yz}\delta\gamma_{yz} - q\delta w_0)dx\,dy\,dz = 0$, producing seven equations of static equilibrium:

$$\delta u_0 : \frac{\partial N_{xx}}{\partial x} + \frac{\partial N_{xy}}{\partial y} + \frac{Q_{xz}}{R_x} = 0 \qquad\qquad \delta v_0 : \frac{\partial N_{yy}}{\partial y} + \frac{\partial N_{xy}}{\partial x} + \frac{Q_{yz}}{R_y} = 0 \tag{9}$$

$$\delta w_0 : \frac{\partial Q_{xz}}{\partial x} + \frac{\partial Q_{yz}}{\partial y} - \frac{N_{xx}}{R_x} - \frac{N_{yy}}{R_y} = q \qquad \delta\phi_x : \frac{\partial M_{xx}}{\partial x} + \frac{\partial M_{xy}}{\partial y} - Q_{xz} + \frac{K_{xz}}{R_x} = 0 \tag{10}$$

$$\delta\phi_y : \frac{\partial M_{yy}}{\partial y} + \frac{\partial M_{xy}}{\partial x} - Q_{yz} + \frac{K_{yz}}{R_y} = 0 \qquad \delta\phi_x^* : \frac{\partial S_{xx}}{\partial x} + \frac{\partial S_{xy}}{\partial y} - 3T_{xz} + \frac{U_{xz}}{R_x} = 0 \tag{11}$$

$$\delta\phi_y^* : \frac{\partial S_{yy}}{\partial y} + \frac{\partial S_{xy}}{\partial x} - 3T_{yz} + \frac{U_{yz}}{R_y} = 0 \tag{12}$$

where the stress resultants are obtained by [5]

$$\begin{Bmatrix} \{N\} \\ \{M\} \\ \{S\} \end{Bmatrix} = \begin{bmatrix} [A] & [B] & [E] \\ [B] & [F] & [G] \\ [E] & [G] & [H] \end{bmatrix} \begin{Bmatrix} \{\epsilon^{(0)}\} \\ \{\epsilon^{(1)}\} \\ \{\epsilon^{(3)}\} \end{Bmatrix}; \quad \begin{Bmatrix} \{Q\} \\ \{K\} \\ \{T\} \\ \{U\} \end{Bmatrix} = \begin{bmatrix} [I] & [K] & [L] & [M] \\ [K] & [N] & [O] & [P] \\ [L] & [O] & [Q] & [R] \\ [M] & [P] & [R] & [S] \end{bmatrix} \begin{Bmatrix} \{\gamma^{(0)}\} \\ \{\gamma^{(1)}\} \\ \{\gamma^{(2)}\} \\ \{\gamma^{(3)}\} \end{Bmatrix}$$

$$(13)$$

The stiffness components in equations (13) are given by:

$(A_{ij}, B_{ij}, E_{ij}, F_{ij}, G_{ij}, H_{ij}) = \int_{-h/2}^{h/2} Q(i,j)(1, z, z^3, z^2, z^4, z^6)\,dz;\ i,j = 1,2,3$, and

$(I_{ij}, K_{ij}, L_{ij}, M_{ij}, N_{ij}, O_{ij}, P_{ij}, Q_{ij}, R_{ij}, S_{ij}) = \int_{-h/2}^{h/2} Q(i,j)(1, z, z^2, z^3, z^2, z^3, z^4, z^4, z^5, z^6)\,dz;$ $i,j = 4,5$.

The boundary conditions for a simply supported shell for a border along the x axis are:

$u_0 = 0;\ N_{yy} = 0;\ w_0 = 0;\ \phi_x = 0;\ M_{yy} = 0;\ \phi_x^* = 0;\ S_{yy} = 0$

5 Numerical examples

Sandwich plates with different h/a ratios and $R_x = R_y = 10^9$ are tested with a mechanical sinusoidal load $Q = q\sin(\frac{\pi x}{a})\sin(\frac{\pi y}{a})$. The mechanical properties of the face sheets and sandwich core are: Face sheets : $E_x = 25; E_y = 1; G_{xz} = 0.5; G_{xy} = 0.5; G_{yz} = 0.2; \nu_{xy} = 0.25; h_f = 0.1h$. Core : $E_x = 0.04; E_y = 0.04; G_{xz} = 0.06; G_{xy} = 0.016; G_{yz} = 0.06; \nu_{xy} = 0.25; h_c = 0.8h$. The results for the central deflection are normalized by $\overline{w} = \frac{100wE_y h^3}{qa^4}$, and the interval chosen for the shape parameter optimal search was [0.1-2].

The values obtained for the central normalized deflection with a shape parameter $c = 2/\sqrt{n}$ are presented in table 1 (a). The value $2/\sqrt{n}$ was previously used by the authors in the analysis of composite and sandwich plates and shells with great success [2, 7]. In table 1(b) the examples are repeated, now using an optimized shape parameter obtained with Rippa's algorithm. In this table, both normalized central deflection and optimized shape parameter

Table 1: Central normalized deformation $\overline{w}$, with $c = 2/\sqrt{n}$ and $c = c_{opt}$

c	h/a	9	11	13	15	17	19	21	HOST9 [5]
					n				
(a) $c = 2/\sqrt{n}$									
	0.01	0.9774	0.9263	0.9081	0.9001	0.8961	0.8941	0.8927	0.8910
	0.1	2.0733	2.0791	2.0818	2.0831	2.0838	2.0842	2.0844	2.0848
	0.25	7.1319	7.1435	7.1485	7.1509	7.1521	7.1528	7.1532	7.1538
(b) $c = c_{opt}$									
	0.01	0.8873	0.8929	0.8922	0.8911	0.8913	0.8922	0.8910	0.8910
	c_{opt}	1.3775	0.8598	0.7701	0.7701	0.8387	0.5221	0.5836	
	0.1	2.0834	2.0846	2.0834	2.0847	2.0847	2.0843	2.0845	2.0848
	c_{opt}	1.4365	1.4129	0.7278	0.9825	0.8034	0.4798	0.4798	
	0.25	7.1511	7.1518	7.1529	7.1529	7.1531	7.1530	7.1537	7.1538
	c_{opt}	1.3518	1.0011	0.8885	0.6925	0.6034	0.4798	0.5541	

(c_{opt}) are presented. The relative error of the normalized central deflection for both tables are plotted in figure 1. The large gray bars correspond to the relative error produced using $c = 2/\sqrt{n}$, for several a/h ratios. The superimposed white bars correspond to the relative error produced when $c = c_{opt}$ is used. The relative error for central normalized deflection is evaluated using the formula: $\text{error}(\%) = \dfrac{\text{computed value} - \text{exact value}[5]}{\text{exact value}[5]} \times 100$. Analysing figure 1 it is possible to observe an improvement for all the examples when the propose optimization technique is used, instead of $c = 2/\sqrt{n}$. This allows the method to achieve a good solution with fewer points. For example, the error produced with $c = 2/\sqrt{n}$ and $n = 15$ is the roughly the same as the one produced by c_{opt} with just $n = 9$.

6   Conclusions

The radial basis collocation method produces excellent results in the analysis of sandwich plates, when compared to analytical solutions. The cross-validation technique, implemented via Rippa's algorithm, improved all examples previously obtained with $c = 2/\sqrt{n}$. Unfortunately, the method still requires some input from the user; in this case, the interval [0.1-2] was chosen.

References

[1] A. H. D. Cheng, M. A. Golberg, E. J. Kansa, and Q. Zammito. Exponential convergence and h-c multiquadric collocation method for partial differential equations. *Numerical Methods for Partial Differential Equations*, 19(5):571–594, 2003.
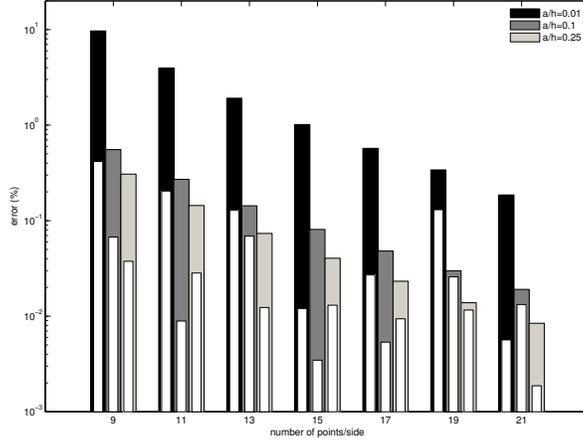
Figure 1: Relative error for $c = 2/\sqrt{n}$ and c_{opt}, with various a/h ratios and number of points/side, n. The white bars correspond to the relative error produced with $c = c_{opt}$.

[2] A. J. M. Ferreira, C. M. C. Roque, and R. M. N. Jorge. Modelling cross-ply laminated elastic shells by a higher-order theory and multiquadrics. *Computers and Structures*, 84(19-20):1288–1299, 2006.

[3] E. J. Kansa. Multiquadrics. a scattered data approximation scheme with applications to computational fluid-dynamics. i. surface approximations and partial derivative estimates. *Computers & mathematics with applications*, 19(8-9):127–145, 1990.

[4] E. J. Kansa. Multiquadrics. a scattered data approximation scheme with applications to computational fluid-dynamics. ii. solutions to parabolic, hyperbolic and elliptic partial differential equations. *Computers & mathematics with applications*, 19(8-9):147–161, 1990.

[5] R.K. Khare, T. Kant, and A.K. Garg. Closed-form thermo-mechanical solutions of higher-order theories of cross-ply laminated shallow shells. *Composite Structures*, 59(3):313–340, 2003.

[6] S. Rippa. An algorithm for selecting a good value for the parameter c in radial basis function interpolation. *Advances in Computational Mathematics*, 11(2-3):193–210, 1999.

[7] C. M. C. Roque, A. J. M. Ferreira, and R. M. N. Jorge. Modelling of composite and sandwich plates by a trigonometric layerwise deformation theory and radial basis functions. *Composites Part B: Engineering*, 36(8):559–572, 2005.

[8] B. P. Wang. Parameter optimization in multiquadric response surface approximations. *Structural and Multidisciplinary Optimization*, 26(3-4):219–223, 2004.

Contact Mechanics

Quasistatic frictional contact problems. Relations between the wedge problem and the rate problem

Lars-Erik Andersson[*]

Abstract

A frictional contact problems in elasticity consists of finding, for a given load history and an initial state, the time evolution of displacements and reaction forces. The problems may be dynamic, quasistatic or static (incremental).

It is well known that for *large* coefficients of friction there are no general results of existence and/or uniqueness for these problems, even for very simple systems.

Here some existence and uniqueness for frictional systems with finitely many spatial degrees of freedom, will be summarized.

Then we will discuss the relation between the problem of so called wedging and the problem of uniqueness for quasistatic evolution problems. In particular we will investigate the relation between critical friction bounds for wedging and for nonuniqueness respectively.

1   Notation and definitions

We consider a linearly elastic system, *e.g.*, finitely many displacement points (nodes) and connecting springs, or some FEM-approximation of a continuous elastic system. Some of the nodes are kept fixed so that no rigid body motions are possible. The free nodes are divided into two classes, the contact nodes, which may come into frictional contact with rigid obstacles, and the rest, which are called interior nodes. The interior and contact nodes are acted upon by given external forces. We assume that the obstacles are flat and that Coulomb's friction law is valid.

Let us introduce the following notation for the exterior force-field and the displacement field.

$$F = (f_I^*, f^*)^* = (f_{1I}^*, f_{2I}^*, \ldots, f_{mI}^*, f_1^*, f_2^*, \ldots, f_l^*)^* \in R^{3(m+l)}$$
$$U = (u_I^*, u^*)^* = (u_{1I}^*, u_{2I}^*, \ldots, u_{mI}^*, u_1^*, u_2^*, \ldots, u_l^*)^* \in R^{3(m+l)}$$

where f_{jI}, f_i, u_{jI}, $u_i \in R^3$, for $1 \leq j \leq m$, $1 \leq i \leq l$, and where $*$ denotes transposition.

---
[*]Department of Mathematics, University of Link´oping, Sweden, E-mail:`leand@mai.liu.se`.

f_I is the part of the force-field acting on the interior nodes and f the part acting on the contact nodes. Similarly, u_I and u are the displacement fields for interior- and contact nodes.

For the contact nodes we also introduce the field

$$r = (r_1^*, r_2^*, \ldots, r_l^*)^* \in R^{3l}$$

of reaction forces from the obstacles, $r_i \in R^3$, $1 \leq i \leq l$. We then have

$$\begin{pmatrix} f_I \\ f+r \end{pmatrix} = \begin{pmatrix} K_{II} & K_{Ic} \\ K_{cI} & K_{cc} \end{pmatrix} \begin{pmatrix} u_I \\ u \end{pmatrix}$$

where

$$\mathcal{K} = \begin{pmatrix} K_{II} & K_{Ic} \\ K_{cI} & K_{cc} \end{pmatrix}, \quad \mathcal{K} = \mathcal{K}^*,$$

is the positive definite stiffness matrix. Introducing the inverse matrix $\mathcal{L} = \mathcal{K}^{-1}$ we may also write,

$$\begin{pmatrix} u_I \\ u \end{pmatrix} = \begin{pmatrix} L_{II} & L_{I1} \\ L_{1I} & L \end{pmatrix} \begin{pmatrix} f_I \\ f+r \end{pmatrix}. \tag{1.1}$$

Consequently, for the contact node vector u we have

$$u = Lr + MF \tag{1.2}$$

where $M = (L_{1I}, L)$

Now the contact nodes are subject to restrictions

$$(u_i, n_i) \leq 0, \ 1 \leq i \leq l \quad (\text{or} \ \leq g_i)$$

with n_i denoting an exterior unit normal vector to a flat, fixed obstacle. Further $(\cdot, \cdot)$ denotes the inner product in R^3 and $|\cdot|$ the corresponding norm. We now decompose the vectors into tangential and normal components:

$$\begin{aligned} r_i &= r_{iT} + r_{iN} n_i \quad \text{with} \ (r_{iT}, n_i) = 0 \\ u_i &= u_{iT} + u_{iN} n_i \quad \text{with} \ (u_{iT}, n_i) = 0 \end{aligned}$$

for $1 \leq i \leq l$. We thus write

$$\begin{aligned} r_T &= (r_{1T}^*, r_{2T}^*, \ldots, r_{lT}^*)^* & r_N &= (r_{1N}, r_{2N}, \ldots, r_{lN})^* \\ u_T &= (u_{1T}^*, u_{2T}^*, \ldots, u_{lT}^*)^* & u_N &= (u_{1N}, u_{2N}, \ldots, u_{lN})^*. \end{aligned}$$
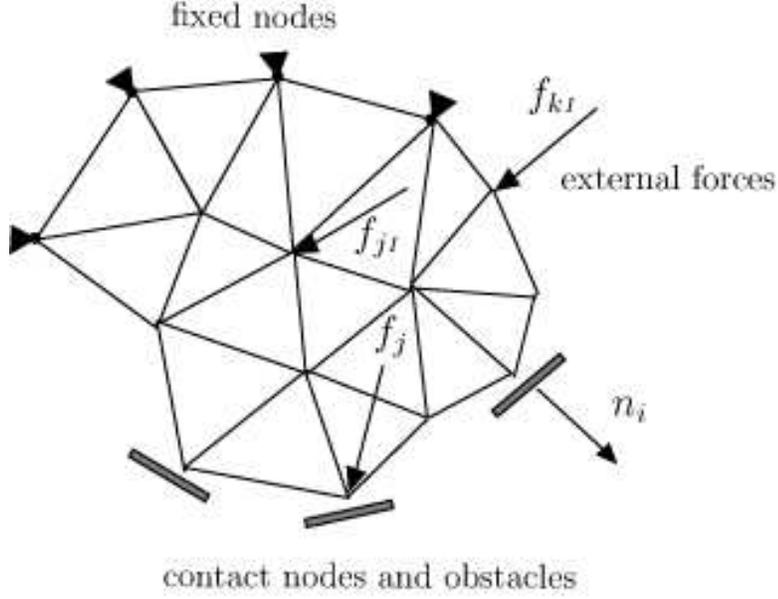
Figure 1: Elastic system with finitely many degrees of freedom

2 Existence and uniqueness for a system with finitely many degrees of freedom

Collecting the structural equation, the non-penetration contact conditions and Coulomb's friction law we then have the following space-discrete quasistatic evolution problem. Here $\dot{u}^+$ denotes the time derivative to the right.

DQP For a given exterior force vector $F = F(t)$, find displacements $u = u(t)$ and reaction forces $r = r(t)$ satisfying (1.2), the Signorini contact conditions:

$$u_{iN} \leq 0, \quad r_{iN} \leq 0, \quad u_{iN} r_{iN} = 0, \tag{2.3}$$

and Coulomb's friction law:

$$
\begin{aligned}
&u_{iN} = 0 \Rightarrow |r_{iT}| \leq -\mu_i r_{iN}, &(2.4)\\
&0 < |r_{iT}| = -\mu_i r_{iN} \Rightarrow \dot{u}_{iT}^+ = -\lambda_i(t) r_{iT}, \quad \lambda_i(t) \geq 0,\\
&|r_{iT}| < -\mu_i r_{iN} \Rightarrow \dot{u}_{iT}^+ = 0.
\end{aligned}
$$

By taking the right-hand time derivative of the inequalities and equations of **DQP**, it is possible to derive a so-called rate problem which, given an initial state, concerns finding right-hand time derivatives of unknown variables. This rate problems turns out to be a Nonlinear Complementarity Problem, which we denote **NLCP**. In this way it is possible to conclude that the full problem **DQP** can be divided into two subproblems:

- **Rate Problem:** Given r and $\dot{F}$, is there a unique solution $\dot{r}^+$ to the **NLCP**, so that we may write
$$\dot{r}^+ = \mathcal{F}(r, \dot{F}) \quad ?$$

- **Integration Problem:** If so, can we integrate this integral equation for a (unique?) solution $r(t)$ with
$$\dot{r}^+(t) = \mathcal{F}(r(t), \dot{F}(t)) \quad ?$$

2.1 A fundamental frictional parameter and existence results

In the following we give, in the case of a small friction coefficient, an existence and uniqueness result for the rate problem as well as an existence result for **DQP**. The smallness of the friction coefficient is expressed by a fundamental parameter which we first define:

Definition 2.1.
$$\tilde{\varphi}(L, n) = \max_{\substack{u \neq 0 \\ u_i \perp n_i}} \left\{ \min_{1 \leq i \leq l} \varphi_i(L, n; u) \right\},$$
$$\tilde{\mu} = \tilde{\mu}(L, n) = \cot \tilde{\varphi}(L, n),$$

where
$$\varphi_i(L, n; u) = \begin{cases} \angle(u_i, P_i r_i) & \text{if } |u_i||P_i r_i| \neq 0 \\ \pi & \text{else} \end{cases}$$

and P_i is the orthogonal projection on span $\{u_i, n_i\}$.

The parameter $\tilde{\varphi}$ has the following simple mechanical interpretation: If the system is acted upon by an arbitrary force vector r on the contact nodes $(F = 0)$ such that all displacements are tangential, then for at least one node we have

$$\varphi_i \leq \tilde{\varphi}.$$

In [1] the following result was obtained. See also [2].

Theorem 2.1. *If*
$$\mu_i < \tilde{\mu}(L, n)$$
for all contact nodes then the rate problem has a unique solution $\dot{r}^+$.

In the same report the following theorem of existence for the discrete quasistatic problem **DQP** is given.

Theorem 2.2. *If in addition $r(0)$, $u(0)$ satisfy the conditions (2.3) and (2.4), $F(t)$ is absolutely continuous and if $\dot{F} \in L^\infty(0, T)$, then there exists a solution r to the quasistatic problem so that, a.e.,*
$$\dot{r}^+(t) = \mathcal{F}(r(t), \dot{F}(t)).$$

Moreover, there exists a constant $C = C(\mu)$ *so that, a.e.,*

$$\|\dot{r}(t)\| \leq C \|\dot{F}(t)\|.$$

For a system with *one node and 2 DOF* it is possible to show that the solution given in the previous existence theorem is *unique*, see, *e.g.*, [1].

For the case that the force field is *piecewise real analytic* it is possible to obtain a partial uniqueness result for the full problem **DQP**.

3   Some frictional parameters for wedging problems

Assume that all the force fields f_I and f are zero and that system is acted upon only by the forces r on the contact nodes. In the analysis r is not necessarily to be interpreted as reaction forces from the obstacles, *i.e.*, we do not always require that $r_N \leq 0$. With the previous notation we then have

$$r = Ku \quad \text{and} \quad u = Lr$$

We introduce the following sets

$$\mathcal{D}(n) = \{u \in R^{3l} : u = Lr, \ u_{iN} \leq 0, r_{iN} \leq 0, u_{iN}r_{iN} = 0, 1 \leq i \leq l\} \tag{3.5}$$

$$\tilde{\mathcal{D}}(n) = \{u \in R^{3l} : u = Lr, u_{iN}r_{iN} = 0, 1 \leq i \leq l\} \tag{3.6}$$

$$\tilde{\mathcal{D}}_{ext}(n) = \{u \in R^{3l} : u = Lr, u_{iN}r_{iN} \leq 0, 1 \leq i \leq l\} \tag{3.7}$$

which are closed cones in R^{3l}. Further let us for each i define define functions f_i by

$$f_i(u) =) = \begin{cases} |r_{iT}|/|r_{iN}| & \text{if} \ \ |r_{iN} \neq 0 \\ +\infty & \text{if} \ \ r_{iN} = 0, \ |r_{iT}| \neq 0 \\ 0 & \text{if} \ \ r_i = 0. \end{cases} \tag{3.8}$$

Next assume that a vector $w = (w_1, w_2, \ldots, w_l)$ of positive weights w_i, is given and define

$$f_w(u) = \max_i f_i(u)/w_i \tag{3.9}$$

The functions f_i and f_w are lsc and homogeneous on the sets $\mathcal{D}(n)$, $\tilde{\mathcal{D}}(n)$ and $\tilde{\mathcal{D}}_{ext}(n)$. It follows that f_w attains a minimum on each of the sets $\mathcal{D}(n) \setminus \{0\}$, $\tilde{\mathcal{D}}(n) \setminus \{0\}$ and $\tilde{\mathcal{D}}_{ext}(n) \setminus \{0\}$. We denote these minima by

$$c_w = \min_{\mathcal{D}(n) \setminus \{0\}} f_w(u) \tag{3.10}$$

$$\tilde{c}_w = \min_{\tilde{\mathcal{D}}(n) \setminus \{0\}} f_w(u) \tag{3.11}$$

and

$$\tilde{c}_{w,ext} = \min_{\tilde{\mathcal{D}}_{ext}(n) \setminus \{0\}} f_w(u) \tag{3.12}$$

Since $\mathcal{D}(n) \subset \tilde{\mathcal{D}}(n) \subset \tilde{\mathcal{D}}_{ext}(n)$ it is clear that $\tilde{c}_{w,ext} \leq \tilde{c}_w \leq c_w$.

Lemma 3.1. $\tilde{c}_{w,ext} = \tilde{c}_w$

Lemma 3.2. *A minimizer $u \in \tilde{\mathcal{D}}(n) \setminus \{0\}$ for the function $f_w(u)$ may be chosen such that*

$$|r_{iT}| = |r_{iN}|w_i c_w, \quad \text{for all} \quad i. \tag{3.13}$$

Theorem 3.1. *If $\mu_i < w_i c_w$ for all i, then there does not exist any wedging state. If $\mu_i \geq w_i c_w$ for all i then there exists a wedging state.*

Theorem 3.2. *If $\mu_i < w_i \tilde{c}_w$ for all i, then there does not exist any double wedging state. If $\mu_i \geq w_i c_w$ for all i then there exists a double wedging state.*

Comment: By double wedging we mean wedging under bilateral contact conditions. The definition of the parameter $\tilde{\mu}(L, n)$ above may be modified to $\tilde{\mu}_w(L, n)$ by introducing a weight vector w. Then we have the following

Lemma 3.3. $\tilde{\mu}_w(L, n) \leq \tilde{c}_{w,ext} = \tilde{c}_w \leq c_w$.

Theorem 3.3. *If $\mu_i < w_i \tilde{\mu}_w(L, n)$ for all i, then the rate problem has a unique solution. If $\mu_i \geq w_i c_{wu}$ for all i then there exists an equilibrium state such that the rate problem is not uniquely solvable (or does not have a solution). If the time derivative of the applied force-field is zero, then the equilibrium is unstable.*

Theorem 3.4. *In two spatial dimensions $\tilde{\mu}_w(L, n) = \tilde{c}_w$, i.e.,*

$$c_{wu} = \tilde{c}_w \leq c_w.$$

Comment: This means that the critical limit for non-uniqueness for the rate problem is equal to the critical limit for double wedging. In general, though, not equal to the critical limit for wedging, which is dependent on the directions of the obstacle-normals.

References

[1] Lars-Erik Andersson. Quasistatic frictional contact problems with finitely many degrees of freedom. Technical Report LiTH-MAT-R-1999-22, Department of Mathematics, University of Linköping, SE-581 83 Linköping Sweden, 1999.

[2] Lars-Erik Andersson and Anders Klarbring. A review of the theory of static and quasi-static frictional contact problems in elasticity. *R. Soc. Lond. Philos. Trans. Ser. A Math. Phys. Eng. Sci.*, 359(1789):2519–2539, 2001. Non-smooth mechanics.

Frictional contact problems for elastic membranes and plates as a paradigm of Moreau's sweeping process

Patrick Ballard[*]   Cédric Pozzolini

Abstract

The frictionless equilibrium of a membrane (linearized theory) against a rigid obstacle provides a paradigm for variational inequalities and the Lions-Stampacchia theorem.

In this communication, we show that if the frictionless assumption is dropped and Coulomb friction is assumed instead, then the resulting evolution problem provides a paradigm for Moreau's sweeping process [2] in the Hilbert space H_0^1. As a consequence, the existence of a unique solution is obtained.

Then, the quasi-static evolution of a plate against a rigid obstacle with Coulomb friction is investigated in the same spirit. In this situation, the harmonic operator must be replaced by the biharmonic one. This leads to specific difficulties that are overcome by escaping from the hilbertian framework. Moreau's sweeping process is used to solve a regularized problem and a unique solution for the original problem is obtained by use of an appropriate limit procedure.

To lighten the notations, strings and beams will be considered here instead of membranes and plates.

Let H be some Hilbert space and $\mathcal{C}(t) \subset H$ a moving closed convex set. Moreau's sweeping process is the evolution process satisfied by the unknown function $u(t)$:

$$
\left|
\begin{array}{l}
-\dot{u}(t) \in \partial I_{\mathcal{C}(t)}\big[u(t)\big], \qquad \text{for a.a. } t, \\
u(0) = u_0,
\end{array}
\right.
$$

where $\partial I_{\mathcal{C}(t)}$ denotes the subdifferential of the indicatrix function of $\mathcal{C}(t)$, that is, the outward normal cone. Provided the multivocal function $\mathcal{C}(t)$ is absolutely continuous (the set of closed subset of H being equipped with the Hausdorff distance), then Moreau's sweeping process admits a unique solution $u \in W^{1,1}([0,T]\,; H)$ [2]. Originally designed for elasto-plastic systems, Moreau's sweeping process has found many applications to mechanics. However,

---

[*]Laboratoire de Mécanique et d'Acoustique, 31, chemin Joseph Aiguier, 13402 Marseille Cédex 20, FRANCE. E-mail:`ballard@lma.cnrs-mrs.fr`.

these applications concern mainly, up to now, finite freedom mechanics, and in this framework the Hilbert space is the euclidean space $\mathbb{R}^n$. As the frictionless equilibrium of a membrane against a rigid obstacle provides a paradigm of variational inequality in the Hilbert space H_0^1, it is worth noticing that the evolution problem associated with quasi-static evolution of this membrane when the contact with the obstacle is assumed to obey to Coulomb dry friction law, offers a paradigm of Moreau's sweeping process in the Hilbert space H_0^1.

1 Quasi-static evolution of a string near a rigid obstacle with Coulomb friction

The coordinate system Oxy will be used in the usual euclidean plane. Consider a string whose initial configuration is the interval $[0,1]$ of the Ox axis. This initial configuration supports some homogeneous tension $T_0 > 0$ and is an equilibrium configuration. Denoting $f\,\mathbf{e}_x + g\,\mathbf{e}_y$ the external force by length unit, and $u\,\mathbf{e}_x + v\,\mathbf{e}_y$, the displacement, the *linearized* equations which govern the equilibrium of the string, assumed to be elastic, read as :

$$\left|\begin{array}{l} k\,u''(x) + f(x) = 0, \\[2mm] u(0) = u_0^{\mathrm{p}}, \qquad u(1) = u_1^{\mathrm{p}}, \\[2mm] T_0\,v''(x) + g(x) = 0, \\[2mm] v(0) = v_0^{\mathrm{p}}, \qquad v(1) = v_1^{\mathrm{p}}, \end{array}\right.$$

where $u_0^{\mathrm{p}}\,\mathbf{e}_x + v_0^{\mathrm{p}}\,\mathbf{e}_y$ and $u_1^{\mathrm{p}}\,\mathbf{e}_x + v_1^{\mathrm{p}}\,\mathbf{e}_y$ are the prescribed displacements at the ends $x = 0, 1$ of the string.

Describing some given fixed rigid obstacle by its cartesian equation $y = \psi(x)$ and denoting $r\,\mathbf{e}_x + s\,\mathbf{e}_y$ the reaction force exerted on the string, the equations governing the quasi-static evolution of the string above the obstacle when Coulomb friction of coefficient μ is assumed, read formally as :

$$\left|\begin{array}{l} k\,\dfrac{\partial^2 u}{\partial x^2} + f + r = 0, \\[3mm] r\big(\hat{u} - \dot{u}\big) + \mu s\big(|\hat{u}| - |\dot{u}|\big) \geq 0, \qquad \forall \hat{u} \in \mathbb{R}, \\[3mm] u(x = 0) = u_0^{\mathrm{p}}, \qquad u(x = 1) = u_1^{\mathrm{p}}, \\[3mm] T_0\,\dfrac{\partial^2 v}{\partial x^2} + g + s = 0, \\[3mm] v - \psi \geq 0, \qquad s \geq 0, \qquad s(v - \psi) = 0, \\[3mm] v(x = 0) = v_0^{\mathrm{p}}, \qquad v(x = 1) = v_1^{\mathrm{p}}. \end{array}\right.$$

It should be underlined that the equations governing the normal displacement v are *not coupled* with the equations that govern the tangential displacement u. As a consequence of

this feature which is specific to the *linearized* equilibrium equations, the problem governing the normal displacement v is the same in the frictionless and frictional cases.

2 Existence and uniqueness of solutions

The problem that governs the normal displacement has to be solved first.

Theorem 2.1. *We assume* $T_0 > 0$, $g \in W^{1,1}(0,T;H^{-1})$, $\psi \in H^1(0,1;\mathbb{R})$ *and* $v_0^{\mathrm{p}}, v_1^{\mathrm{p}} \in W^{1,1}(0,T;\mathbb{R})$ *satisfying the compatibility condition :*

$$\forall t \in [0,T], \qquad \left| \begin{array}{l} v_0^{\mathrm{p}}(t) > \psi(0), \\ v_1^{\mathrm{p}}(t) > \psi(1). \end{array} \right.$$

Setting :

$$K(t) = \left\{ \hat{v} \in H^1(0,1;\mathbb{R}) \; ; \; \forall x \in \,]0,1[, \quad \hat{v}(x) \geq \psi(x), \quad and \quad \forall x \in \{0,1\}, \quad \hat{v}(x) = v_x^{\mathrm{p}}(t) \right\},$$

there exists a unique $v \in W^{1,1}(0,T;H^1)$ *such that :*

- $\forall t \in [0,T], \qquad v(t) \in K(t),$

- $\forall t \in [0,T], \quad \forall \hat{v} \in K(t), \qquad T_0 \int_0^1 \frac{\mathrm{d}v}{\mathrm{d}x}\left(\frac{\mathrm{d}\hat{v}}{\mathrm{d}x} - \frac{\mathrm{d}v}{\mathrm{d}x} \right) \geq \left\langle g, \hat{v} - v \right\rangle_{H^{-1},H_0^1}.$

Theorem 2.1 is a straightforward consequence of standard results on variational inequalities [1]. A consequence of theorem 2.1 is that the equations governing the normal component v of the displacement of the membrane determine uniquely the normal component of the reaction exerted by the obstacle :

$$s \in W^{1,1}(0,T;H^{-1}).$$

Theorem 2.2. *We assume* $k > 0$, $f \in W^{1,1}(0,T;H^{-1})$ *and* $u_0^{\mathrm{p}}, u_1^{\mathrm{p}} \in W^{1,1}(0,T;\mathbb{R})$. *We are also given some initial condition* $u_0 \in H^1(0,1;\mathbb{R})$ *satisfying the compatibility conditions :*

- $u_0(0) = u_0^{\mathrm{p}}(t=0), \qquad u_0(1) = u_1^{\mathrm{p}}(t=0),$

- $\forall \hat{u} \in H_0^1, \quad k \int_0^1 \frac{\mathrm{d}u_0}{\mathrm{d}x} \frac{\mathrm{d}\hat{u}}{\mathrm{d}x} \geq \left\langle f(t=0), \hat{u} \right\rangle_{H^{-1},H_0^1} - \left\langle \mu s(t=0), |\hat{u}| \right\rangle_{H^{-1},H_0^1}.$

Then, there exists a unique $u \in W^{1,1}(0,T;H^1)$ *such that :*

- $u(t=0) = u_0,$

- $\forall t \in [0,T], \qquad u(t,0) = u_0^{\mathrm{p}}(t), \quad u_0(t,1) = u_1^{\mathrm{p}}(t),$

- $\begin{array}{l} \textit{for a.a. } t \in [0,T], \\ \forall \hat{u} \in H_0^1, \end{array} \qquad \left\langle f + k\frac{\partial^2 u}{\partial x^2}, \hat{u} - \dot{u} \right\rangle_{H^{-1},H_0^1} \leq \left\langle \mu s, |\hat{u}| - |\dot{u}| \right\rangle_{H^{-1},H_0^1}.$

Remark. At first sight, the evolution inequality in theorem 2.2 does make sense only for $\dot{u} \in H_0^1$. However, from the assumptions of theorem 2.1, it can be proved that $\operatorname{Supp} s \subset [\alpha, \beta] \subset]0, 1[$, for all $t \in [0, 1]$. Therefore, the duality product is actually defined for $\dot{u} \in H^1$.

Proof. The proof is particularized to the specific case $u_0^{\mathrm{P}} = u_1^{\mathrm{P}} = 0$. The changes that have to be made to handle more general boundary conditions are straightforward. The function $\Phi : H_0^1 \to \mathbb{R}$ defined by :

$$\Phi(\hat{u}) = \Big\langle \mu s, |\hat{u}| \Big\rangle_{H^{-1}, H_0^1} - \Big\langle f, \hat{u} \Big\rangle_{H^{-1}, H_0^1},$$

is continuous and convex. In terms of subdifferential, the problem is now to find $u \in W^{1,1}(0, T; H^1)$ such that :

$$\left|\begin{array}{ll} -u(t) \in \partial\Phi\big[\dot{u}(t)\big], & \text{for a.a. } t, \\ u(t = 0) = u_0, & \end{array}\right.$$

where the subdifferential should be understood in the sense of the scalar product of H_0^1 induced by the elastic energy. The dual function $\Phi^\star$ is nothing but the indicatrix funtion (in the sense of convex analysis) of the bounded convex subset $\mathcal{C}(t)$ of H_0^1 defined by :

$$\mathcal{C}(t) = \left\{ u \in H_0^1 \ ; \ \forall \hat{u} \in H_0^1, \quad k \int_0^1 \frac{\mathrm{d}u}{\mathrm{d}x}\frac{\mathrm{d}\hat{u}}{\mathrm{d}x} \geq \Big\langle f, \hat{u} \Big\rangle_{H^{-1}, H_0^1} - \Big\langle \mu s, |\hat{u}| \Big\rangle_{H^{-1}, H_0^1} \right\}.$$

Now, it is readily seen that our problem has been reduced to find $u \in W^{1,1}(0, T; H^1)$ such that :

$$\left|\begin{array}{ll} \dot{u}(t) \in \partial I_{\mathcal{C}(t)}\big[-u(t)\big], & \text{for a.a. } t, \\ u(t = 0) = u_0, & \end{array}\right.$$

which is a Moreau sweeping process and our claim has become a direct consequence of Moreau's article [2]. $\qquad\qquad\qquad\square$

3   Replacing the string by a beam

In this section, a beam is considered instead of the string. The initial configuration is still the interval $[0, 1]$ of the Ox axis. This initial configuration is not stressed at all and we restrict ourselves to the circumstance where the *linearized* equations of equilibrium can be used. We shall assume that the ends of the beam are simply supported.

The equations governing the quasi-static evolution of the beam above the obstacle when

Coulomb friction of coefficient μ is assumed, read formally as :

$$\left|\begin{array}{l} k\,\dfrac{\partial^2 u}{\partial x^2} + f + r = 0, \\[2mm] r\big(\hat{u} - \dot{u}\big) + \mu s\big(|\hat{u}| - |\dot{u}|\big) \geq 0, \qquad \forall \hat{u} \in \mathbb{R}, \\[2mm] u(x = 0) = u_0^{\mathrm{p}}, \qquad u(x = 1) = u_1^{\mathrm{p}}, \\[2mm] l\,\dfrac{\partial^4 v}{\partial x^4} - g - s = 0, \\[2mm] v - \psi \geq 0, \qquad s \geq 0, \qquad s(v - \psi) = 0, \\[2mm] v(x = 0) = v_0^{\mathrm{p}}, \qquad v(x = 1) = v_1^{\mathrm{p}}, \\[2mm] \dfrac{\partial^2 v}{\partial x^2}(x = 0) = \dfrac{\partial^2 v}{\partial x^2}(x = 1) = 0, \end{array}\right.$$

where k and l are the elastic stiffnesses of the beam respectively in extension and flexion. Problems governing the normal and tangential components of the displacement are still uncoupled, but the problem governing the normal component is now of order 4.

4   Existence and uniqueness of solutions

As prevously, the "normal" problem has to be solved first.

Theorem 4.1. *We assume $l > 0$, $g \in W^{1,1}(0, T; H^{-1})$, $\psi \in H^3(0, 1; \mathbb{R})$ and $v_0^{\mathrm{p}}, v_1^{\mathrm{p}} \in W^{1,1}(0, T; \mathbb{R})$ satisfying the compatibility condition :*

$$\forall t \in [0, T], \qquad \left|\begin{array}{l} v_0^{\mathrm{p}}(t) > \psi(0), \\[2mm] v_1^{\mathrm{p}}(t) > \psi(1). \end{array}\right.$$

Setting :

$$K(t) = \Big\{ \hat{v} \in H^2(0, 1; \mathbb{R}) \ ; \ \forall x \in \,]0, 1[, \quad \hat{v}(x) \geq \psi(x), \quad and \quad \forall x \in \{0, 1\}, \quad \hat{v}(x) = v_x^{\mathrm{p}}(t) \Big\},$$

then, for all $t \in [0, T]$, there exists a unique $v(t) \in K(t)$ such that :

$$\forall \hat{v} \in K(t), \qquad l \int_0^1 \frac{\partial^2 v}{\partial x^2}\left(\frac{\mathrm{d}^2 \hat{v}}{\mathrm{d}x^2} - \frac{\partial^2 v}{\partial x^2}\right) \geq \Big\langle g, \hat{v} - v \Big\rangle_{H^{-1}, H_0^1}.$$

Moreover, the function $v : [0, T] \to H^2$ thus defined, owns the following regularity :

$$v \in W^{1,1}(0, T; H^2) \cap L^\infty(0, T; H^3).$$

As a consequence of theorem 4.1, the normal component s of the reaction force has the following regularity : $s \in W^{1,1}(0, T; H^{-2}) \cap L^\infty(0, T; H^{-1})$. It can be proved that we have actually :

$$s \in W^{1,1}(0, T; H^{-2}) \cap L^\infty(0, T; \mathcal{M}).$$

It is easy to design an example of a beam in contact with a fixed smooth obstacle for which the coincidence set reduces to one "moving" point. Such an example shows that the regularity $s \in W^{1,1}(0,T;H^{-1})$ cannot be expected in general, and also that jumps (with respect to the x variable) of the velocity $\dot{u}$ have to be allowed. Thus, the "tangential" problem is not a sweeping process and we cannot use directly theorem 2.2 to solve it.

One way to overcome this difficulty is to introduce a sequence of mollifiers $\rho_n \in C_c^\infty(\mathbb{R};\mathbb{R})$ to regularize s by convolution : $s_n = \rho_n \overset{x}{*} s$, recalling that $\operatorname{Supp} s \subset [\alpha, \beta] \subset \,]0,1[$, for all $t \in [0,1]$. Then, theorem 2.2 can be used to exhibit the solution u_n of the regularized problem and there remains only to find the appropriate estimates to go to the limit $n \to \infty$ to solve the original problem.

Theorem 4.2. *We assume $k > 0$, $f \in W^{1,1}(0,T;H^{-1})$ and $u_0^{\mathrm{p}}, u_1^{\mathrm{P}} \in W^{1,1}(0,T;\mathbb{R})$. We are also given some initial condition $u_0 \in H^1(0,1;\mathbb{R})$ satisfying the compatibility conditions :*

- $u_0(0) = u_0^{\mathrm{p}}(t=0), \qquad u_0(1) = u_1^{\mathrm{P}}(t=0),$

- $\forall\, \hat{u} \in H_0^1, \quad k \int_0^1 \dfrac{\mathrm{d}u_0}{\mathrm{d}x} \dfrac{\mathrm{d}\hat{u}}{\mathrm{d}x} \geq \Big\langle f(t=0), \hat{u} \Big\rangle_{H^{-1}, H_0^1} - \Big\langle \mu s(t=0), |\hat{u}| \Big\rangle_{H^{-1}, H_0^1}.$

Then, there exist unique $u \in L^\infty(0,T;H^1)$, $r \in L^\infty(0,T;\mathcal{M})$, $\dot{u} \in L^1(0,T;VBN([0,1];\mathbb{R}))$ such that :

- $u(t) = u_0 + \displaystyle\int_0^t \dot{u},$

- $\forall\, t \in [0,T], \qquad u(t,0) = u_0^{\mathrm{p}}(t), \quad u_0(t,1) = u_1^{\mathrm{P}}(t),$

- $k\,\dfrac{\partial^2 u}{\partial x^2} + f + r = 0,$

- $\begin{aligned}&\text{for a.a. } t \in [0,T],\\ &\forall\, \hat{u} \in C^0,\end{aligned} \qquad \displaystyle\int_{[0,1]} r\big(\hat{u} - \dot{u}\big) + \int_{[0,1]} \mu s\Big(|\hat{u}| - |\dot{u}|\Big) \geq 0.$

Here $VBN([0,1];\mathbb{R}))$ denotes the Banach space of functions $f : [0,1] \to \mathbb{R}$ that have bounded variation and satisfy in addition : $f(x) = (f(x-)+f(x+))/2$, for all $x \in \,]0,1[$. Recall that functions having bounded variation are universally integrable (that is, with respect to any measure).

Remark. At the time being, theorem 4.2 is only a conjecture since time has lacked to write down the details of the proof.

References

[1] David Kinderlehrer and Guido Stampacchia. *An Introduction to Variational Inequalities and Their Applications.* Academic Press, 1980.

[2] Jean Jacques Moreau. Evolution problem associated with a moving convex set in a Hilbert space. *Journal of Differential Equations*, 26:347–374, 1977.

Numerical analysis of a dynamic piezoelectric contact problem arising in viscoelasticity

M. Barboteu[*]   J.R. Fernández[†]   T. Raffat[‡]

Abstract

A dynamic frictionless contact problem between a viscoelastic piezoelectric body and a deformable obstacle is numerically studied in this paper. The contact is modelled using the normal compliance contact condition and the linear electro-viscoelastic constitutive law is employed to simulate the piezoelectric effects. The variational formulation is a coupled system composed of a parabolic nonlinear variational equation for the velocity field and a linear variational equation for the electric potential. An existence and uniqueness result is recalled. Then, a fully discrete scheme is introduced, based on the finite element method to approximate the spatial variable and the backward Euler scheme to discretize the time derivatives. Error estimates are derived from which the linear convergence of the algorithm is deduced under suitable regularity conditions. Finally, an academic two-dimensional example is presented to demonstrate the performance of the algorithm.

1 The mechanical model and its variational formulation

Piezoelectricity is the ability of certain cristals, like the quartz (also ceramics ($BaTiO_3$, $KNbO_3$, $LiNbO_3$, etc) and even the human mandible), to produce a voltage when they are subjected to mechanical stress. On a nanoscopic scale, the piezoelectric phenomenon arises from a nonuniform charge distribution within a crystal unit cells, and the piezoelectricity is then perceived as the electrical polarization due to mechanical input. Thus, the piezoelectric effect is characterized by the coupling between the mechanical and the electrical properties of the material: it was observed that the appearance of electric charges on some cristals was due to the action of body forces and surface tractions and, conversely, the action of the electric field generated strain or stress in the body. This kind of materials appears usually in the industry as switches in radiotronics, electroacoustics or measuring equipments (see, e.g., [4]).

[*]Laboratoire de Mathématiques, Physique et Systèmes (LAMPS). Université de Perpignan, France. E-mail:`barboteu@univ-perp.fr`.

[†]Departamento de Matematica Aplicada. Universidade de Santiago de Compostela, Spain. E-mail:`jramon@usc.es`. Supported by MEC-Spain (Project MTM2006-13981).

[‡]Laboratoire de Mathématiques, Physique et Systèmes (LAMPS). Université de Perpignan, France. E-mail:`raffat@univ-perp.fr`.

Let $\Omega \subset \mathbb{R}^d$, $d = 1, 2, 3$, denote a domain occupied by a visco-elasto-piezoelectric body with a smooth boundary $\Gamma = \partial\Omega$ decomposed either into three disjoint parts Γ_D, Γ_F, Γ_C or either into two disjoint parts Γ_A and Γ_B such that meas $(\Gamma_D) > 0$, meas $(\Gamma_A) > 0$, and $\Gamma_C \subseteq \Gamma_B$. Finally, let $[0, T]$, $T > 0$, be the time interval of interest and denote by $\boldsymbol{\nu}$ the unit outer normal vector to Γ.

Let $\boldsymbol{x} \in \Omega$ and $t \in [0, T]$ be the spatial and time variables, respectively, and, in order to simplify the writing, we do not indicate the dependence of the functions on $\boldsymbol{x}$ and t. Moreover, a dot above a variable represents the derivative with respect to the time variable.

Let $\boldsymbol{u}$ denote the displacement field, $\boldsymbol{\sigma}$ the stress tensor, $\boldsymbol{\varepsilon}(\boldsymbol{u}) = (\varepsilon_{ij}(\boldsymbol{u}))_{i,j=1}^d$ the linearized strain tensor given by $\varepsilon_{ij}(\boldsymbol{u}) = \dfrac{1}{2} \left(\dfrac{\partial u_i}{\partial x_j} + \dfrac{\partial u_j}{\partial x_i} \right)$, and φ the electric potential.

The body is assumed visco-elasto-piezoelectric and satisfying the following constitutive law (see [3]),

$$\boldsymbol{\sigma} = \mathcal{A}\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}) + \mathcal{B}\boldsymbol{\varepsilon}(\boldsymbol{u}) - \mathcal{E}^* \mathbf{E}(\varphi),$$

where $\mathcal{A}$ and $\mathcal{B}$ are the fourth-order viscosity and elastic tensors, respectively, $\mathbf{E}(\varphi) = (E_i(\varphi))_{i=1}^d$ represents the electric field defined by $E_i(\varphi) = -\dfrac{\partial\varphi}{\partial x_i}$, $i = 1, \ldots, d$, and $\mathcal{E}^* = (e_{ijk}^*)_{i,j,k=1}^d$ denotes the transpose of the third-order piezoelectric tensor $\mathcal{E} = (e_{ijk})_{i,j,k=1}^d$. Moreover, according to [2], the constitutive law for the electric potential is given by

$$\mathbf{D} = \mathcal{E}\boldsymbol{\varepsilon}(\boldsymbol{u}) + \beta\mathbf{E}(\varphi),$$

where $\mathbf{D}$ is the electric displacement field and β is the electric permittivity tensor.

We turn now to describe the boundary conditions.

On the boundary part Γ_D we assume that the body is clamped and thus the displacement field neglects there (and so $\boldsymbol{u} = \boldsymbol{0}$ on $\Gamma_D \times (0, T)$). Moreover, we assume that a density of traction forces, denoted by $\boldsymbol{f}_F$, acts on the boundary part Γ_F, i.e., $\boldsymbol{\sigma}\boldsymbol{\nu} = \boldsymbol{f}_F$ on $\Gamma_F \times (0, T)$.

On the part Γ_C the body can become in contact with a deformable insulator obstacle. Following [5] the normal compliance contact condition is written as $-\sigma_\nu = p(u_\nu - g)$ on $\Gamma_C \times (0, T)$, where $\sigma_\nu = \boldsymbol{\sigma}\boldsymbol{\nu} \cdot \boldsymbol{\nu}$ is the normal stress, $u_\nu = \boldsymbol{u} \cdot \boldsymbol{\nu}$ denotes the normal displacement, g represents the gap between the body and the obstacle measured along the normal direction $\boldsymbol{\nu}$ and p is a given constitutive function. Finally, we assume that the contact is frictionless and therefore, $\boldsymbol{\sigma}_\tau = \boldsymbol{\sigma}\boldsymbol{\nu} - \sigma_\nu\boldsymbol{\nu} = \boldsymbol{0}$.

Let Ω be subject to a prescribed electric potential φ_A on Γ_A and to a density of surface electric charges q_F on Γ_B, that is, $\varphi = \varphi_A$ on $\Gamma_A \times (0, T)$ and $\mathbf{D} \cdot \boldsymbol{\nu} = q_F$ on $\Gamma_B \times (0, T)$, where we assume that $q_F = 0$ on Γ_C.

Finally, denote by $\boldsymbol{u}_0$ and $\boldsymbol{v}_0$ the initial conditions for the displacement and the velocity fields, respectively, and assume that $\varphi_A = 0$ for the sake of simplicity. It is straightforward to extend the results presented in this section, and in the following one, to more general situations.

In order to obtain the variational formulation, let us denote by $H = [L^2(\Omega)]^d$ and construct the variational spaces V, W and Q as follows,

$$V = \{\boldsymbol{w} \in [H^1(\Omega)]^d \, ; \, \boldsymbol{w} = \boldsymbol{0} \quad \text{on} \quad \Gamma_D\}, \quad W = \{\psi \in H^1(\Omega) \, ; \, \psi = 0 \quad \text{on} \quad \Gamma_A\},$$
$$Q = \{\boldsymbol{\tau} = (\tau_{ij})_{i,j=1}^d \in [L^2(\Omega)]^{d \times d} \, ; \, \tau_{ij} = \tau_{ji}, \ i,j = 1,\dots,d\}.$$

We will make the following assumptions on the problem data (see [1] for details):

- The viscosity and the permittivity tensors are bounded, symmetric and elliptic.

- The elastic and piezoelectric tensors are bounded and symmetric.

- The normal compliance function is a monotone and Lipschitz function satisfying $p(\boldsymbol{x}, r) = 0$ for all $r \leq 0$ and $\boldsymbol{x} \in \Gamma_C$.

- The mechanical and electric forces are continuous with respect to the time variable.

- The gap function is a positive function and the initial conditions $\boldsymbol{u}_0 \in V$ and $\boldsymbol{v}_0 \in H$.

We denote by V' the dual space of V and by $\langle \cdot, \cdot \rangle_{V' \times V}$ the duality pairing between V' and V. Next, we can define the elements $\boldsymbol{f}(t) \in V'$ and $q(t) \in W$ given by

$$\langle \boldsymbol{f}(t), \boldsymbol{w} \rangle_{V' \times V} = \int_\Omega \boldsymbol{f}_0(t) \cdot \boldsymbol{w} \, d\boldsymbol{x} + \int_{\Gamma_F} \boldsymbol{f}_F(t) \cdot \boldsymbol{w} \, d\Gamma \quad \forall \boldsymbol{w} \in V,$$
$$(q(t), \psi)_W = \int_\Omega q_0(t)\psi \, d\boldsymbol{x} + \int_{\Gamma_B} q_F(t)\psi \, d\Gamma \quad \forall \psi \in W.$$

Let $v_\nu = \boldsymbol{v} \cdot \boldsymbol{\nu}$ and denote by $j : V \times V \to \mathbb{R}$ the normal compliance functional defined as

$$j(\boldsymbol{u}, \boldsymbol{v}) = \int_{\Gamma_C} p(u_\nu - g) v_\nu \, d\Gamma \quad \forall \boldsymbol{u}, \boldsymbol{v} \in V.$$

Keeping in mind that $\mathbf{D} = -\nabla \varphi$ and applying a Green's formula, we derive the following variational formulation in terms of the velocity field $\boldsymbol{v}(t) = \dot{\boldsymbol{u}}(t)$.

Problem VP. *Find a velocity field $\boldsymbol{v} : [0, T] \to V$ and an electric potential field $\varphi : [0, T] \to W$ such that $\boldsymbol{v}(0) = \boldsymbol{v}_0$ and for a.e. $t \in [0, T]$,*

$$\langle \rho \dot{\boldsymbol{v}}(t), \boldsymbol{w} \rangle_{V' \times V} + (\mathcal{A}\boldsymbol{\varepsilon}(\boldsymbol{v}(t)), \boldsymbol{\varepsilon}(\boldsymbol{w}))_Q + (\mathcal{B}\boldsymbol{\varepsilon}(\boldsymbol{u}(t)), \boldsymbol{\varepsilon}(\boldsymbol{w}))_Q$$
$$+ (\mathcal{E}\nabla\varphi(t), \boldsymbol{\varepsilon}(\boldsymbol{w}))_Q + j(\boldsymbol{u}(t), \boldsymbol{w}) = \langle \boldsymbol{f}(t), \boldsymbol{w} \rangle_{V' \times V} \quad \forall \boldsymbol{w} \in V,$$
$$(\beta \nabla\varphi(t), \nabla\psi)_H - (\mathcal{E}\boldsymbol{\varepsilon}(\boldsymbol{u}(t)), \nabla\psi)_H = (q(t), \psi)_W \quad \forall \psi \in W,$$

where the displacement field $\boldsymbol{u}(t)$ is defined by $\boldsymbol{u}(t) = \displaystyle\int_0^t \boldsymbol{v}(s) \, ds + \boldsymbol{u}_0$.

Using analogous ideas to those employed in [6], we obtain the following.

Theorem 1.1. *Assume that the previous assumptions hold. Suppose, in addition, that $\boldsymbol{f}$ lies in $H^2(0, T; V')$. Then there exists a unique solution to Problem VP such that,*

$$\boldsymbol{v} \in C^1([0, T]; H) \cap H^1([0, T]; V), \quad \varphi \in C([0, T]; W).$$

The proof of Theorem 1.1 is based on standard results for evolution variational inequalities, a priori estimates and a fixed-point theorem.

2 Fully discrete approximations: error estimates

The discretization of Problem VP is done as follows. First, we consider two finite dimensional spaces $V^h \subset V$ and $W^h \subset W$ approximating the spaces V and W, respectively. $h > 0$ denotes the spatial discretization parameter.

Secondly, the time derivatives are discretized by using a uniform partition of $[0, T]$, denoted by $0 = t_0 < t_1 < \ldots < t_N = T$. Let k be the time step size, $k = T/N$, and, for a continuous function $f(t)$, let $f_n = f(t_n)$. Finally, for a sequence $\{w_n\}_{n=0}^N$, we denote by $\delta w_n = (w_n - w_{n-1})/k$ the divided differences.

Thus, using the backward Euler scheme, the fully discrete approximation of Problem VP is the following.

Problem VP^{hk}. *Find a discrete velocity field* $\boldsymbol{v}^{hk} = \{\boldsymbol{v}_n^{hk}\}_{n=0}^N \subset V^h$ *and a discrete electric potential field* $\varphi^{hk} = \{\varphi_n^{hk}\}_{n=0}^N \subset W^h$ *such that* $\boldsymbol{v}_0^{hk} = \boldsymbol{v}_0^h$ *and for all* $n = 1, \ldots, N$,

$$(\rho \delta \boldsymbol{v}_n^{hk}, \boldsymbol{w}^h)_H + (\mathcal{A}\boldsymbol{\varepsilon}(\boldsymbol{v}_n^{hk}), \boldsymbol{\varepsilon}(\boldsymbol{w}^h))_Q + (\mathcal{B}\boldsymbol{\varepsilon}(\boldsymbol{u}_n^{hk}), \boldsymbol{\varepsilon}(\boldsymbol{w}^h))_Q$$
$$+ (\mathcal{E}^* \nabla \varphi_n^{hk}, \boldsymbol{\varepsilon}(\boldsymbol{w}^h))_Q + j(\boldsymbol{u}_n^{hk}, \boldsymbol{w}^h) = \langle \boldsymbol{f}_n, \boldsymbol{w}^h \rangle_{V' \times V} \quad \forall \boldsymbol{w}^h \in V^h,$$
$$(\beta \nabla \varphi_n^{hk}, \nabla \psi^h)_H - (\mathcal{E}\boldsymbol{\varepsilon}(\boldsymbol{u}_n^{hk}), \nabla \psi^h)_H = (q_n, \psi^h)_W \quad \forall \psi^h \in W^h,$$

where the discrete displacement field $\boldsymbol{u}^{hk} = \{\boldsymbol{u}_n^{hk}\}_{n=0}^N \subset V^h$ *is given by* $\boldsymbol{u}_n^{hk} = \boldsymbol{u}_0^h + \sum_{j=1}^n k \boldsymbol{v}_j^{hk}$,

and $\boldsymbol{u}_0^h$ *and* $\boldsymbol{v}_0^h$ *are appropriate approximations of the initial conditions* $\boldsymbol{u}_0$ *and* $\boldsymbol{v}_0$, *respectively.*

We notice that the fully discrete problem VP^{hk} can be seen as a coupled system of non-linear variational equations. Using well-known results, we easily obtain that Problem VP^{hk} admits a unique solution $\boldsymbol{v}^{hk} \subset V^h$ and $\varphi^{hk} \subset W^h$.

Our interest in this section lies in estimating the errors $\|\boldsymbol{v}_n - \boldsymbol{v}_n^{hk}\|_H$ and $\|\varphi_n - \varphi_n^{hk}\|_W$. We have the following main error estimates result (see [1]).

Theorem 2.1. *Let the assumptions of Theorem 1.1 hold. Let* $(\boldsymbol{v}, \varphi)$ *and* $(\boldsymbol{v}^{hk}, \varphi^{hk})$ *denote the solutions to problems* VP *and* VP^{hk}, *respectively. Then, there exists a positive constant* $c > 0$, *independent of the discretization parameters* h *and* k, *such that the following error estimates hold for all* $\boldsymbol{w}^h = \{\boldsymbol{w}_j^h\}_{j=1}^N \subset V^h$ *and* $\psi^h = \{\psi_j^h\}_{j=1}^N \subset W^h$,

$$\max_{1 \le n \le N} \{\|\boldsymbol{v}_n - \boldsymbol{v}_n^{hk}\|_H^2 + \|\varphi_n - \varphi_n^{hk}\|_W^2\} + \sum_{n=1}^N k \|\boldsymbol{v}_n - \boldsymbol{v}_n^{hk}\|_V^2 \le c \Big(\max_{1 \le n \le N} \|\varphi_n - \psi_n^h\|_W^2$$
$$+ \sum_{j=1}^N k \Big[\|\dot{\boldsymbol{v}}_j - \delta \boldsymbol{v}_j\|_H^2 + \|\boldsymbol{v}_j - \boldsymbol{w}_j^h\|_V^2 \Big] + \max_{1 \le n \le N} \|\boldsymbol{v}_n - \boldsymbol{w}_n^h\|_H^2 + \|\boldsymbol{u}_0 - \boldsymbol{u}_0^h\|_V^2$$
$$+ \|\boldsymbol{v}_0 - \boldsymbol{v}_0^h\|_H^2 + k^2 + \frac{1}{k} \sum_{j=1}^{N-1} \|\boldsymbol{v}_j - \boldsymbol{w}_j^h - (\boldsymbol{v}_{j+1} - \boldsymbol{w}_{j+1}^h)\|_H^2 \Big).$$

We notice that the above error estimates are the basis for the analysis of the convergence rate of the algorithm. Thus, let Ω be a polyhedral domain and denote by $\mathcal{T}^h$ a triangulation

of $\overline{\Omega}$ compatible with the partition of the boundary $\Gamma = \partial\Omega$ into Γ_D, Γ_F, Γ_C on one hand, and on Γ_A and Γ_B, on the other hand. Let V^h and W^h consist of continuous and piecewise affine functions. Assume that the discrete initial conditions $\boldsymbol{u}_0^h$ and $\boldsymbol{v}_0^h$ are obtained by

$$\boldsymbol{u}_0^h = \Pi^h \boldsymbol{u}_0, \quad \boldsymbol{v}_0^h = \Pi^h \boldsymbol{v}_0, \tag{1}$$

where $\Pi^h = (\pi^h)_{i=1}^d : [C(\overline{\Omega})]^d \to V^h$, and $\pi^h : C(\overline{\Omega}) \to B^h$ is the standard finite element interpolation operator.

Then, we have the following corollary which states the linear convergence of the algorithm under suitable regularity conditions.

Corollary 2.1. *Let the assumptions of Theorem 1.1 hold. Let $(\boldsymbol{v}, \varphi)$ and $(\boldsymbol{v}^{hk}, \varphi^{hk})$ denote the solutions to problems VP and VP^{hk}, respectively, and let the discrete initial conditions be given by (1). Under the regularity conditions $\boldsymbol{u} \in C^1([0,T]; [H^2(\Omega)]^d) \cap H^3(0,T;H)$ and $\varphi \in C([0,T]; H^2(\Omega))$, the linear convergence of the algorithm is achieved; that is, there exists a positive constant $c > 0$, independent of the discretization parameters h and k, such that*

$$\max_{1 \leq n \leq N} \left\{ \|\boldsymbol{u}_n - \boldsymbol{u}_n^{hk}\|_V + \|\varphi_n - \varphi_n^{hk}\|_W \right\} \leq c(h+k).$$

3 Numerical results in an academical example

A full description of the algorithm developed to solve Problem VP^{hk} can be seen in [1] and thus, we refer the reader there for details. This is based on a penalty approach (see [7]) to simulate the normal compliance law.

As a two-dimensional example, we consider a rectangular body $\Omega = (0,4) \times (0,2)$ with the boundary Γ which can come in contact on Γ_C with a deformable obstacle. The body is in frictionless contact with an insulator obstacle with normal compliance, in the form $p(u_\nu - g) = c_p(u_\nu - g)_+$, where $(r)_+ = \max\{r, 0\}$, with a deformability coefficient $c_p = 10^5$ and a gap function $g = 6 \times 10^{-3} m$ on $\Gamma_C = [0,4] \times \{0\}$. To fix the geometry we set the points $P_1 = (0,0)$, $P_2 = (4,0)$, $P_3 = (4,2)$, $P_4 = (0,2)$. We define $\Gamma_B = \Gamma_C = [P_1, P_2]$, $\Gamma_D = \{P_3\} \cup \{P_4\}$ and $\Gamma_F = \Gamma \setminus (\Gamma_C \cup \Gamma_D)$. Here, we use as material a viscoelastic piezoelectric body made of an isotropic piezoceramic material with hexagonal symmetry like zinc oxide (class $6mm$ in the international classification). We suppose that the body is clamped on Γ_D and we employ the following data:

$$T = 1\,s, \quad \rho = 7800\,kg/m^3, \quad \boldsymbol{f}_F = \boldsymbol{0}\,N/m^2, \quad \boldsymbol{f}_0 = \boldsymbol{0}\,N/m^3, \quad q_0 = 0\,C/m^3,$$

$$q_F = 0\,C/m^2, \quad \boldsymbol{u}_0 = \boldsymbol{0}\,m, \quad \boldsymbol{v}_0 = \boldsymbol{0}\,m/s, \quad \varphi_A = \begin{cases} -3\,V & \text{on } (P_3, P_4), \\ 0\,V & \text{on } (P_1, P_4), \\ 0\,V & \text{on } (P_2, P_3). \end{cases}$$

This physical setting allows us to observe the inverse piezoelectric effect that corresponds to the appearance of strain or stress in the body, due to the action of the electric field. This example represents a dynamic contactor stimulated by an electric field.

According to Figure 1 (left-hand side), it can be seen that the action of the difference of the electric field on (P_3, P_4) with (P_1, P_4) and (P_2, P_3) induces a deformation of the body. That results in to make come into contact the body with the obstacle on Γ_C. Indeed, we
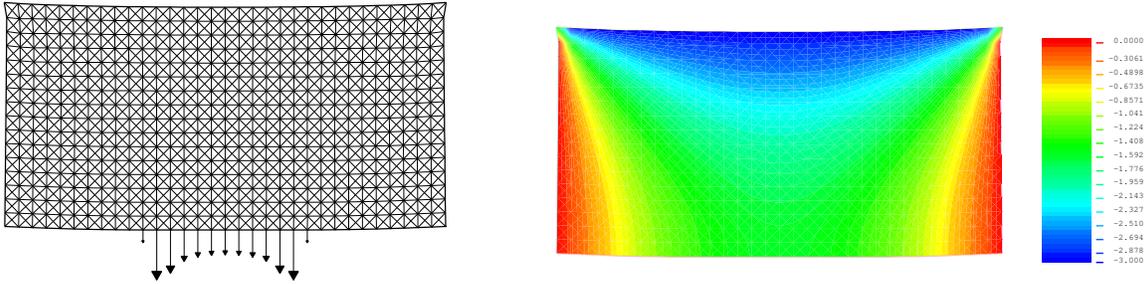

Figure 1: Amplified deformed mesh with contact interface forces (left) and electric potential (right).

remark for the discretization used ($h = 1/16$) that 13 nodes of Γ_C comes in contact with the deformable obstacle and the contact forces are produced along with the outward normal vector, $\boldsymbol{\nu}$, to Γ_C. This is due to the fact that the problem is frictionless. Moreover, the corresponding electric potential is shown on the right-hand side.

References

[1] M. Barboteu, J. R. Fernández, and T. Raffat. Numerical analysis of a dynamic piezoelectric contact problem arising in viscoelasticity. Preprint, 2007.

[2] R. C. Batra and J. S. Yang. Saint-Venant's principle in linear piezoelectricity. *J. Elasticity*, 38:209–218, 1995.

[3] G. Duvaut and J. L. Lions. *Inequalities in mechanics and physics.* Springer Verlag, Berlin, 1976.

[4] T. Ideka. *Fundamentals of piezoelectricity.* Oxford University Press, Oxford, 1990.

[5] A. Klarbring, A. Mikelić, and M. Shillor. Frictional contact problems with normal compliance. *Internat. J. Engrg. Sci.*, 26:811–832, 1988.

[6] K. L. Kuttler and M. Shillor. Regularity of solutions to a dynamic frictionless contact problem with normal compliance. *Nonlinear Anal.*, 59:1063–1075, 2004.

[7] P. Wriggers. *Computational contact mechanics.* Wiley, 2002.

A problem of adhesive contact with thermal effects

Elena Bonetti*   Giovanna Bonfanti †   Riccarda Rossi ‡

1   Introduction

In the papers [2, 3] we investigate a contact problem with adhesion between a viscoelastic body and a rigid support. We refer to a model proposed by Frémond [4, Chap. 14], which mainly combines the damage and the contact theories making use of the phase field approach.

In the phenomenon of contact with adhesion, the resistance to the tension on the contact surface is due to microbonds between the surface of the body and its support. Thus, in the model by Frémond the description of the contact surface includes the description of the state of such bonds, which can break (or get damaged) owing to microscopic motions. Actually, a phase parameter $\chi \in [0,1]$ is introduced to represent the state of damage of the bonds: in particular, the value $\chi = 0$ corresponds to completely damaged bonds, $\chi = 1$ to the undamaged case, and $\chi \in (0,1)$ to an intermediate situation. We consider, at first, an isothermal phenomenon. Thus, the other variable of the model, describing the mechanical equilibrium of the system, is the macroscopic deformation $\varepsilon(\mathbf{u})$ of the viscoelastic body (i.e. the linearized symmetric strain tensor, $\mathbf{u}$ being the vector of small displacements). The equations are recovered from a generalized version of the principle of virtual power, in which there are included the effects of the contact micro-forces and of the micro-movements breaking the bonds.

The viscoelastic body is assumed to be located in a smooth bounded domain $\Omega \subset \mathbb{R}^3$ on a time interval $(0,T)$; let its boundary be $\Gamma = \overline{\Gamma}_1 \cup \overline{\Gamma}_2 \cup \overline{\Gamma}_c$, where Γ_1, Γ_2, and Γ_c are open subset in the relative topology of Γ with a smooth boundary and disjoint one from another; moreover, we suppose that the contact surface Γ_c and the region Γ_1 have strictly positive measure.

---

*Department of Mathematics, University of Pavia, Italy. E-mail: `elena.bonetti@unipv.it`.

†Department of Mathematics, University of Brescia, Italy. E-mail: `bonfanti@ing.unibs.it`.

‡Department of Mathematics, University of Brescia, Italy. E-mail: `riccarda.rossi@unibs.it`.

Normalizing some physical constants, the system we are dealing with is the following

$$- \operatorname{div}(K\varepsilon(\mathbf{u}) + K_v\varepsilon(\mathbf{u}_t)) = \mathbf{f} \quad \text{in } \Omega \times (0, T), \tag{1.1}$$

$$\mathbf{u} = \mathbf{0} \quad \text{in } \Gamma_1 \times (0, T), \quad (K\varepsilon(\mathbf{u}) + K_v\varepsilon(\mathbf{u}_t))\mathbf{n} = \mathbf{g} \quad \text{in } \Gamma_2 \times (0, T), \tag{1.2}$$

$$(K\varepsilon(\mathbf{u}) + K_v\varepsilon(\mathbf{u}_t))\mathbf{n} + \chi\mathbf{u} + \partial I_{]-\infty,0]}(\mathbf{u} \cdot \mathbf{n})\mathbf{n} \ni \mathbf{0} \quad \text{in } \Gamma_c \times (0, T), \tag{1.3}$$

$$\chi_t - \Delta_s\chi + \partial I_{]-\infty,0]}(\chi_t) + \partial I_{[0,1]}(\chi) \ni w_s - \frac{1}{2}|\mathbf{u}|^2 \quad \text{in } \Gamma_c \times (0, T), \tag{1.4}$$

$$\partial_{\mathbf{n}_s}\chi = 0 \quad \text{in } \partial\Gamma_c \times (0, T), \tag{1.5}$$

$$\mathbf{u}(\cdot, 0) = \mathbf{u}_0 \text{ in } \Omega, \quad \chi(\cdot, 0) = \chi_0 \text{ in } \Gamma_c. \tag{1.6}$$

In (1.2)-(1.4) (and in the sequel) we omit the trace symbol for $\mathbf{u}$ and $\mathbf{u}_t$. Moreover, Δ_s is the laplacian on Γ_c, while $\mathbf{n}$ ($\mathbf{n}_s$, respectively) is the outward unit normal vector to Γ (to $\partial\Gamma_c$, resp.). Further, K is the rigidity matrix, K_v the viscosity matrix, $\mathbf{f}$ an applied volume force, $\mathbf{g}$ a traction applied on a part of the boundary, and w_s a positive constant representing the cohesion of the adhesion. In the boundary condition (1.3), the presence of the subdifferential operator $\partial I_{]-\infty,0]}(\mathbf{u} \cdot \mathbf{n})$ entails the impenetrability condition $\mathbf{u} \cdot \mathbf{n} \leq 0$, being $\partial I_{]-\infty,0]}(\mathbf{u} \cdot \mathbf{n}) = 0$ if $\mathbf{u} \cdot \mathbf{n} < 0$ and $\partial I_{]-\infty,0]}(0) = [0, +\infty[$. Analogously, the subdifferential $\partial I_{[0,1]}$ in (1.4) represents the physical constraint on χ, which is forced to take values in the interval $[0, 1]$ ($\partial I_{[0,1]}(\chi) = 0$ if $\chi \in]0, 1[$, $\partial I_{[0,1]}(0) =]-\infty, 0]$, and $\partial I_{[0,1]}(1) = [0, +\infty[$). Moreover, with $\partial I_{]-\infty,0]}(\chi_t)$ in (1.4), enforcing $\chi_t \leq 0$, we are prescribing that the evolution of the adhesion is irreversible, i.e. the bonds cannot repair themselves, once they are broken.

In the paper [2] we obtain a global in time existence result (see Theorem 1 below) for a generalized version of system (1.1)-(1.6), by means of a fixed point technique, combined with regularization tools and with a careful and non standard prolongation procedure. Moreover, we prove in [2] a uniqueness result for a special class of solutions, corresponding to a suitable choice of the initial data and of the involved nonlinearities. In the general case, uniqueness is not to be expected, essentially due to the doubly nonlinear character of (1.4).

On the contrary, dealing with the simplified physical situation of reversible phenomena of adhesion (i.e., without the term $\partial I_{(-\infty,0]}(\chi_t)$ in (1.4)), we show in [3] the well-posedness of problem (1.1)-(1.6) and study the long-time behavior of the solutions investigating the structure of the associated ω-limit set, i.e. the set of the cluster points as time goes to infinity of the solution trajectories.

In the recent work [1], we allow for the temperature influencing the evolution of the adhesion, so that we take into account thermal effects both in the body and on the contact surface, assuming that the temperature on the contact surface may be different from the trace on the boundary of the volume temperature. Entropy equations govern the evolution of the

temperatures. More precisely, we consider the absolute temperature ϑ of the body, satisfying

$$\partial_t(\log \vartheta) - \operatorname{div}(\mathbf{u}_t) - \Delta \vartheta = h \qquad \text{in } \Omega \times (0, T) \,, \tag{1.7}$$

$$\partial_n \vartheta = -k(\chi)(\vartheta - \vartheta_s) \text{ in } \Gamma_c \times (0, T) \,, \quad \partial_n \vartheta = 0 \text{ in } (\Gamma \setminus \Gamma_c) \times (0, T) \,, \tag{1.8}$$

and the absolute temperature ϑ_s of the adhesive material on Γ_c, satisfying

$$\partial_t(\log \vartheta_s) - \partial_t(\lambda(\chi)) - \Delta_s \vartheta_s = k(\chi)(\vartheta - \vartheta_s) \qquad \text{in } \Gamma_c \times (0, T) \,, \tag{1.9}$$

$$\partial_n \vartheta_s = 0 \qquad \text{in } \partial \Gamma_c \times (0, T) \,, \tag{1.10}$$

where h is an external heat source, the function $\lambda(\chi)$ represents the latent heat and $k(\chi)$ is a nonlinear function related to the adhesive properties of the material. Let us observe that, mainly due to the presence of the logarithmic nonlinearity in (1.7) and (1.9), we directly deduce the positivity of the temperatures (which implies thermodynamical consistence of the model), avoiding any application of the maximum principle. Now, we couple (1.7)–(1.10) with a momentum balance law for the macroscopic movements

$$- \operatorname{div}(K\varepsilon(\mathbf{u}) + K_v\varepsilon(\mathbf{u}_t) + \vartheta \mathbf{1}) = \mathbf{f} \qquad \text{in } \Omega \times (0, T) \,, \tag{1.11}$$

$$\mathbf{u} = \mathbf{0} \quad \text{in } \Gamma_1 \times (0, T) \,, \ (K\varepsilon(\mathbf{u}) + K_v\varepsilon(\mathbf{u}_t) + \vartheta \mathbf{1})\mathbf{n} = \mathbf{g} \quad \text{in } \Gamma_2 \times (0, T) \,, \tag{1.12}$$

$$(K\varepsilon(\mathbf{u}) + K_v\varepsilon(\mathbf{u}_t) + \vartheta \mathbf{1})\mathbf{n} + \chi\mathbf{u} + \partial I_{(-\infty,0]}(\mathbf{u} \cdot \mathbf{n})\mathbf{n} \ni \mathbf{0} \text{ in } \Gamma_c \times (0, T) \,, \tag{1.13}$$

and an equilibrium equation describing the evolution of the adhesion (in the reversible case)

$$\chi_t - \Delta_s \chi + \partial I_{[0,1]}(\chi) \ni -\lambda'(\chi)(\vartheta_s - \vartheta_{\text{eq}}) - \frac{1}{2}|\mathbf{u}|^2 \quad \text{in } \Gamma_c \times (0, T) \,, \tag{1.14}$$

$$\partial_{\mathbf{n}_s} \chi = 0 \qquad \text{in } \partial \Gamma_c \times (0, T) \,, \tag{1.15}$$

where $\mathbf{1}$ denotes the identity matrix and the positive constant ϑ_{eq} is an equilibrium temperature. We supplement (1.7)–(1.15) with the initial conditions

$$\log \vartheta(\cdot, 0) = \log \vartheta_0 \ \text{ in } \Omega \,, \quad \log \vartheta_s(\cdot, 0) = \log \vartheta_s^0 \ \text{ in } \Gamma_c \,, \tag{1.16}$$

$$\mathbf{u}(\cdot, 0) = \mathbf{u}_0 \ \text{ in } \Omega \,, \quad \chi(\cdot, 0) = \chi_0 \ \text{ in } \Gamma_c, \tag{1.17}$$

for suitable initial data (see (2.16)-(2.17) and (2.4)-(2.5) below). System (1.7)–(1.15) is highly nonlinear: the main analytical difficulties are related to the presence of multivalued operators rendering physical constraints on the variables, to the singular character of the temperature equations, and to the nonlinear coupling between the equations themselves. These problems are carefully handled in our argument for existence, which consists of the following steps. First, a generalized version of system (1.7)–(1.17) is approximated by regularizing the involved nonlinearities and by adding some vanishing viscosity terms in (1.7) and (1.9). Then, a global solution of the resulting approximating problem is obtained by means of a fixed point procedure and, finally, the passage to the limit in the regularizing parameter shows the convergence to a solution of the original problem. In this way, we prove in [1] a global in time existence result (see Theorem 2 below) for system (1.7)–(1.17). We also study the asymptotic behavior of the solutions investigating the associated ω-limit set.

81

2 Main results

We introduce here some useful notation and the set of our hypotheses. Hereafter, for the sake of simplicity, we assume that Ω is a bounded smooth set of $\mathbb{R}^3$, such that Γ_c is a smooth bounded domain of $\mathbb{R}^2$ (one may think of a flat surface). We set $W := \{\mathbf{v} \in (H^1(\Omega))^3 : \mathbf{v} = \mathbf{0} \text{ a.e. on } \Gamma_1\}$. We assume the material to be homogeneous and isotropic, so that the rigidity matrix K may be represented by $K\varepsilon(\mathbf{u}) = \lambda \mathrm{tr}\,\varepsilon(\mathbf{u})\mathbf{1} + 2\mu\varepsilon(\mathbf{u})$, where $\lambda, \mu > 0$ are the so-called Lamé constants. Hence, to introduce the variational formulation for system (1.1)-(1.6), let us consider the continuous bilinear symmetric forms $a, b : W \times W \to \mathbb{R}$ defined by $a(\mathbf{u}, \mathbf{v}) := \lambda \int_\Omega \mathrm{div}(\mathbf{u})\,\mathrm{div}(\mathbf{v}) + 2\mu \sum_{i,j=1}^3 \int_\Omega \varepsilon_{ij}(\mathbf{u})\varepsilon_{ij}(\mathbf{v})$ and $b(\mathbf{u}, \mathbf{v}) = \sum_{i,j=1}^3 \int_\Omega b_{ij}\varepsilon_{ij}(\mathbf{u})\varepsilon_{ij}(\mathbf{v})$ where (b_{ij}) is the viscosity matrix K_v. Moreover, we denote $\int_{\Gamma_c}$ - $\int_{\Gamma_2}$, resp.- for the duality pairing $_{(H^{-1/2}(\Gamma_c))^3}\langle\cdot,\cdot\rangle_{(H^{1/2}(\Gamma_c))^3}$ -for $_{(H^{-1/2}(\Gamma_2))^3}\langle\cdot,\cdot\rangle_{(H^{1/2}(\Gamma_2))^3}$, respectively-.

Since our results apply to more general graphs than the subdifferentials $\partial I_{]-\infty,0]}$ and $\partial I_{[0,1]}$, we introduce the following functionals

$$\varphi : (H^{1/2}(\Gamma_c))^3 \to [0, +\infty] \text{ proper, convex and lower semicontinuous,}$$

$$\text{with } \varphi(\mathbf{0}) = 0, \text{ such that } \alpha = \partial\varphi : (H^{1/2}(\Gamma_c))^3 \to 2^{(H^{-1/2}(\Gamma_c))^3}, \tag{2.1}$$

$$\rho : \mathbb{R} \to 2^{\mathbb{R}} \text{ maximal monotone operator with } 0 \in \rho(0), \tag{2.2}$$

$$\beta : \mathbb{R} \to 2^{\mathbb{R}} \text{ maximal monotone operator,}$$

$$\text{with } \mathrm{dom}(\beta) \subseteq [0, +\infty) \text{ and } 0 \in \beta(0). \tag{2.3}$$

Next, concerning the Cauchy conditions (1.6), we assume that

$$\mathbf{u}_0 \in W \text{ and } \mathbf{u}_0 \in \mathrm{dom}(\varphi), \tag{2.4}$$

$$\chi_0 \in H^2(\Gamma_c), \; \partial_{\mathbf{n}_s}\chi_0 = 0 \text{ a.e. in } \partial\Gamma_c, \quad \psi(\chi_0) \in L^1(\Gamma_c),$$
$$\text{and} \quad \beta^0(\chi_0) \in L^2(\Gamma_c), \tag{2.5}$$

where $\psi : \mathbb{R} \to [0, +\infty]$ is such that $\beta = \partial\psi$ and $\psi(0) = 0$, and $\beta^0(\chi_0)$ denotes the element of minimal norm in $\beta(\chi_0)$. Moreover, we prescribe that

$$\mathbf{f} \in L^2(0, T; H^3), \quad \mathbf{g} \in L^2(0, T; (H^{-1/2}(\Gamma_2))^3). \tag{2.6}$$

Here is the main result of the paper [2].

Theorem 1. *Under assumptions (2.1)–(2.6), there exist* $(\mathbf{u}, \chi, \boldsymbol{\eta}, \xi, \zeta)$ *fulfilling (1.5), (1.6),*

and

$$\mathbf{u} \in H^1(0,T;W), \tag{2.7}$$

$$\chi \in W^{1,\infty}(0,T;L^2(\Gamma_c)) \cap H^1(0,T;H^1(\Gamma_c)) \cap L^\infty(0,T;H^2(\Gamma_c)), \tag{2.8}$$

$$b(\mathbf{u}_t,\mathbf{v}) + a(\mathbf{u},\mathbf{v}) + \int_{\Gamma_c} (\chi\mathbf{u} + \boldsymbol{\eta}) \cdot \mathbf{v}$$

$$= \int_\Omega \mathbf{f}(t) \cdot \mathbf{v} + \int_{\Gamma_2} \mathbf{g}(t) \cdot \mathbf{v} \quad \forall \mathbf{v} \in W \quad a.e. \ in \ (0,T), \tag{2.9}$$

$$\boldsymbol{\eta} \in \alpha(\mathbf{u}) \ a.e. \ in \ (0,T) \qquad \boldsymbol{\eta} \in L^2(0,T;(H^{-1/2}(\Gamma_c))^3), \tag{2.10}$$

$$\chi_t - \Delta\chi + \xi + \zeta = w_s - \frac{1}{2}|\mathbf{u}|^2 \quad a.e. \ in \ \Gamma_c \times (0,T), \tag{2.11}$$

$$\xi \in \rho(\chi_t) \ a.e. \ in \ \Gamma_c \times (0,T) \qquad \xi \in L^\infty(0,T;L^2(\Gamma_c)), \tag{2.12}$$

$$\zeta \in \beta(\chi) \ a.e. \ in \ \Gamma_c \times (0,T) \qquad \zeta \in L^\infty(0,T;L^2(\Gamma_c)). \tag{2.13}$$

Now, in order to deal with the non isothermal problem (1.7)–(1.17), we require

$$k : \mathbb{R} \to [0,+\infty) \ \text{Lipschitz continuous}, \quad \lambda \in C^{1,1}(\mathbb{R}), \tag{2.14}$$

$$h \in L^2(0,T;(H^1(\Omega))'), \tag{2.15}$$

$$\vartheta_0 \in L^1(\Omega), \quad \vartheta_0 > 0 \ \text{a.e. in } \Omega \quad \text{and} \quad \log\vartheta_0 \in (H^1(\Omega))', \tag{2.16}$$

$$\vartheta_s^0 \in L^1(\Gamma_c), \quad \vartheta_s^0 > 0 \ \text{a.e. in } \Gamma_c \quad \text{and} \quad \log\vartheta_s^0 \in L^2(\Gamma_c). \tag{2.17}$$

Now, we state the main result of the paper [1].

Theorem 2. *Under the assumptions* (2.1), (2.3)–(2.6), *and* (2.14)–(2.17), *there exist* $(\vartheta, \vartheta_s, \mathbf{u}, \chi, \boldsymbol{\eta}, \zeta)$

fulfilling (1.16)–(1.17), (2.7), (2.8), (2.10), (2.13), (1.5), *and*

$$\vartheta \in L^2(0,T;H^1(\Omega)) \cap L^\infty(0,T;L^1(\Omega))\,, \tag{2.18}$$

$$\log \vartheta \in H^1(0,T;(H^1(\Omega))') \quad \vartheta > 0 \ a.e. \ in \ \Omega \times (0,T)\,, \tag{2.19}$$

$$\vartheta_s \in L^2(0,T;H^1(\Gamma_c)) \cap L^\infty(0,T;L^1(\Gamma_c))\,, \tag{2.20}$$

$$\log \vartheta_s \in L^\infty(0,T;L^2(\Gamma_c)) \cap H^1(0,T;H^1(\Gamma_c)') \tag{2.21}$$

$$\vartheta_s > 0 \ a.e. \ in \ \Gamma_c \times (0,T)\,, \tag{2.22}$$

$${}_{(H^1(\Omega))'}\langle(\log\vartheta)_t,v\rangle_{H^1(\Omega)} - \int_\Omega \mathrm{div}(\mathbf{u}_t)\,v + \int_\Omega \nabla\vartheta\,\nabla v + \int_{\Gamma_c} k(\chi)(\vartheta-\vartheta_s)v$$
$$= {}_{(H^1(\Omega))'}\langle h,v\rangle_{H^1(\Omega)} \quad \forall\,v \in H^1(\Omega) \quad a.e. \ in \ (0,T)\,, \tag{2.23}$$

$${}_{H^1(\Gamma_c)'}\langle(\log\vartheta_s)_t,v\rangle_{H^1(\Gamma_c)} - \int_{\Gamma_c} \partial_t\lambda(\chi)\,v + \int_{\Gamma_c} \nabla\vartheta_s\,\nabla v$$
$$= \int_{\Gamma_c} k(\chi)(\vartheta-\vartheta_s)v \quad \forall\,v \in H^1(\Gamma_c) \quad a.e. \ in \ (0,T)\,, \tag{2.24}$$

$$b(\mathbf{u}_t,\mathbf{v}) + a(\mathbf{u},\mathbf{v}) + \int_\Omega \vartheta\,\mathrm{div}(\mathbf{v}) + \int_{\Gamma_c}(\chi\mathbf{u}+\boldsymbol{\eta})\cdot\mathbf{v}$$
$$= \int_\Omega \mathbf{f}(t)\cdot\mathbf{v} + \int_{\Gamma_2} \mathbf{g}(t)\cdot\mathbf{v} \quad \forall\,\mathbf{v} \in W \quad a.e. \ in \ (0,T)\,, \tag{2.25}$$

$$\chi_t - \Delta\chi + \zeta = -\lambda'(\chi)(\vartheta_s-\vartheta_{\mathrm{eq}}) - \frac{1}{2}|\mathbf{u}|^2 \quad a.e. \ in \ \Gamma_c \times (0,T)\,. \tag{2.26}$$

References

[1] E. Bonetti, G. Bonfanti, and R. Rossi. Analysis of a problem of adhesive contact with thermal effects. preprint.

[2] E. Bonetti, G. Bonfanti, and R. Rossi. Global existence for a contact problem with adhesion. Math. Meth. Appl. Sci., to appear.

[3] E. Bonetti, G. Bonfanti, and R. Rossi. Well-posedness and long-time behaviour for a model of contact with adhesion. Indiana Univ. Math. J., to appear.

[4] M. Frémond. *Non-smooth Thermomechanics*. Springer-Verlag, 2002.

On the hyperbolic contact problem for a von Kármán plate

I. Bock*   J. Jarušek†

1   Problem formulation and penalization

The dynamic contact problems are not frequently solved in the framework of variational inequalities. For the elastic problems there is only a very limited amount of results available (cf. [6] and there cited literature). The aim of the present paper is to extend these results to the nonlinear von Kármán plates. The presented results also extend the research made for the quasistatic contact problems for these plates [3] and [4]. The solvability of dynamic contact problems for von Kármán plates with short and long memory has been proved in [1] and [2], respectively.

The existence of solutions is proved for an approximate penalized problem at first. The limit process to the original problem is enabled by an L_1 estimate of the penalty term and by the use of the compact imbedding theorem and by a proper use of the interpolation technique). The procedure is based on the idea introduced by K. Maruo in [8].

Let $\Omega \subset \mathbb{R}^2$ be a bounded domain with a sufficiently smooth boundary Γ and $I \equiv (0, T)$ a bounded time interval. The unit outer normal vector is denoted by $\vec{n} = (n_1, n_2), \vec{\tau} = (-n_2, n_1)$ is the unit tangent vector. The constants $E > 0$ and $\nu \in \left\langle 0, \frac{1}{2} \right)$ denote the Young modulus of elasticity and the Poisson ratio, respectively. Further we use the abbreviation

$$b = \frac{h^2}{12\varrho(1-\nu^2)},$$

where h is the the plate thickness and ϱ is the density of the material. We denote

$$[u, v] \equiv \partial_{11}u\partial_{22}v + \partial_{22}u\partial_{11}v - 2\partial_{12}u\partial_{12}v.$$

Here and in the sequel we employ the notation

$$\frac{\partial}{\partial s} \equiv \partial_s, \ \frac{\partial^2}{\partial s \partial r} \equiv \partial_{sr}, \ \partial_i = \partial_{x_i}, \ i = 1, 2,$$

$$\dot{v} = \frac{\partial v}{\partial t}, \ \ddot{v} = \frac{\partial^2 v}{\partial t^2}, \qquad Q = I \times \Omega, \ S = I \times \Gamma.$$

By $W_p^k(M)$ with $k \geq 0$ and $p \in [1, \infty]$ the Sobolev (for a noninteger k the Sobolev-Slobodetskii) spaces are denoted provided they are defined on a domain or an appropriate manifold M. If $p = 2$ we use the notations $H^k(M)$ and $H_0^k(M)$ the space with zero traces on ∂M. The anisotropic spaces are denoted by $W_p^k(M)$, $k = (k_1, k_2)$, $\in \mathbb{R}_+^2$, where k_1 is related with the time while k_2 with the space variables (with the obvious consequences for $p = 2$) provided M is a time-space domain. We set $\mathcal{H} = L_\infty(I; H^2(\Omega))$.

We consider a plate free on its boundary. Classical formulation is composed of the system

$$\left. \begin{array}{l} \ddot{u} + bE\triangle^2 u - [u, v] = f + g, \\[2mm] u \geq 0, \ g \geq 0, \ ug = 0, \\[2mm] \triangle^2 v + E[u, u] = 0 \end{array} \right\} \text{ on } Q, \tag{1}$$

the boundary conditions

$$u \geq 0, \ \Sigma(u) \geq 0, \ u\Sigma(u) = 0, \ \mathcal{M}(u) = 0, \ v = 0 \text{ and } \partial_n v = 0 \text{ on } S, \tag{2}$$

$$\begin{aligned} \mathcal{M}(u) &= bEM(u), \\ M(u) &= \triangle u + (1 - \nu)(2n_1 n_2 \partial_{12} u - n_1^2 \partial_{22} u - n_2^2 \partial_{11} u); \\ \Sigma(u) &= bEV(u), \\ V(u) &= \partial_n \triangle u + (1 - \nu)\partial_\tau[(n_1^2 - n_2^2)\partial_{12} u + n_1 n_2(\partial_{22} u - \partial_{11} u)] \end{aligned}$$

and the initial conditions

$$u(0, \cdot) = u_0 \geq 0, \ \dot{u}(0, \cdot) = u_1 \text{ on } \Omega. \tag{3}$$

For $u, y \in L_2(I; H^2(\Omega))$ we define the following bilinear form

$$A : (u, y) \mapsto b\big(\partial_{kk} u \partial_{kk} y + \nu(\partial_{11} u \partial_{22} y + \partial_{22} u \partial_{11} y) + 2(1 - \nu)\partial_{12} u \partial_{12} y\big) \tag{4}$$

almost everywhere on Q and introduce a cone $\mathcal{C}$ as

$$\mathcal{C} := \{y \in \mathcal{H}; \ y \geq 0\}. \tag{5}$$

Then the variational formulation of the problem (1–3) has the following form:

Look for $\{u, v\} \in \mathcal{C} \times L_2(I; H_0^2(\Omega))$ such that

$$\int_Q (EA(u, y_1 - u) + \ddot{u}(y_1 - u) - [u, v](y_1 - u)) \, dx \, dt \geq \int_Q f(y_1 - u) \, dx \, dt, \tag{6}$$

$$\int_\Omega \big(\triangle v \triangle y_2 + E[u, u]y_2\big) \, dx = 0 \ \forall \ (y_1, y_2) \in \mathcal{C} \times H_0^2(\Omega). \tag{7}$$

We define the bilinear operator $\Phi : H^2(\Omega)^2 \to H_0^2(\Omega)$ by means of the variational equation

$$\int_\Omega \triangle\Phi(u, v)\triangle\phi \, dx = \int_\Omega [u, v]\phi \, dx, \ \phi \in H_0^2(\Omega). \tag{8}$$

The equation (8) has a unique solution, because $[u, v] \in L_1(\Omega)$. The equation (8) has a unique solution, because $[u, v] \in L_1(\Omega) \subset H^2(\Omega)^*$. The well-defined operator Φ is evidently compact and symmetric. The domain Ω fulfils the assumptions enabling to apply Lemma 1 from [7] due to which $\Phi : H^2(\Omega)^2 \to W_p^2(\Omega)$, $2 < p < \infty$ and

$$\|\Phi(u, v)\|_{W_p^2(\Omega)} \leq c\|u\|_{H^2(\Omega)}\|v\|_{W_p^1(\Omega)} \;\; \forall u \in H^2(\Omega), \; v \in W_p^1(\Omega). \tag{9}$$

With its help we reformulate the system (6,7) into the following variational inequality:

Problem $\mathcal{P}$. *We look for $u \in \mathcal{C}$ such that $\ddot{u} \in \mathcal{H}^*$, the initial conditions (3) are satisfied in a certain generalized sense, and the inequality*

$$\langle \ddot{u}, y - u \rangle_0 + \int_Q E\big(A(u, y - u) + [u, E\Phi(u, u)](y - u)\big)\, dx\, dt$$
$$\geq \int_Q f(y - u)\, dx\, dt. \tag{10}$$

holds for any $y \in \mathcal{C}$.

Here $\langle \cdot, \cdot \rangle_0$ denotes the duality pairing between $\mathcal{H}$ and its dual as an extension of the $L_2(Q)$–duality.

For any $\eta > 0$ we define the *penalized problem* which includes the system of equations

$$\left.\begin{array}{r} \ddot{u} + bE\triangle^2 u - [u, v] = f + \eta^{-1}u^-, \\[2mm] \triangle^2 v + E[u, u] = 0 \end{array}\right\} \;\; \text{on } Q, \tag{11}$$

the boundary conditions

$$\Sigma(u) = 0, \;\; \mathcal{M}(u) = 0, \;\; v = 0 \text{ and } \partial_n v = 0 \text{ on } S \tag{12}$$

and the initial conditions (3). It has the variational formulation:

Look for $\{u, v\} \in L_\infty(I; H^2(\Omega)) \times L_2(I; H_0^2(\Omega))$ such that $\ddot{u} \in L_2(I; (H^2(\Omega))^*)$ and the following system

$$\int_Q \big(\ddot{u}z_1 + EA(u, z_1) - [u, v]\,z_1 - \eta^{-1}u^- z_1\big)\, dx\, dt = \int_Q f z_1\, dx\, dt, \tag{13}$$

$$\int_\Omega \big(\triangle v \triangle z_2 + E[u, u]\big)z_2\, dx = 0 \tag{14}$$

is satisfied for any $(z_1, z_2) \in L_2(I; H^2(\Omega)) \times H_0^2(\Omega)$ and there hold the conditions (3).

With the help of the operator Φ we get the following reformulation of (13), (14):

Problem $\mathcal{P}_\eta$.

We look for $u \in L_\infty(I, H^2(\Omega))$ such that $\ddot{u} \in L_2(I; (H^2(\Omega))^*)$, the equation

$$\int_Q \big(-\ddot{u}z + EA(u, z) + E[u, \Phi(u, u)]z - \eta^{-1}u^- z\big)\, dx\, dt = \int_Q f z\, dx\, dt \tag{15}$$

holds for any $z \in L_2(I; (H^2(\Omega))$ and the initial conditions (3) remain valid.

We verify the existence of a solution to the penalized problem.

Theorem 1.1. *Let $f \in L_2(Q)$, $u_0 \in H^2(\Omega)$, and $u_1 \in L_2(\Omega)$ Then there exists a solution u*

of the problem $\mathcal{P}_\eta$. *If* $v = -E_0\Phi(u,u)$, *then a couple* $\{u,v\}$ *is a solution of the problem* (13), (14), (3).

Proof. Let us denote by $\{w_i \in H^2(\Omega); i \in \mathbb{N}\}$ an orthonormal in $L_2(\Omega)$ basis of $H^2(\Omega)$. We construct the Galerkin approximation u_m of a solution in a form

$$u_m(t) = \sum_{i=1}^{m} \alpha_i(t)w_i, \ \alpha_i(t) \in \mathbb{R}, \ i = 1, ..., m, \ m \in \mathbb{N},$$

$$\int_\Omega \left(\ddot{u}_m(t)w_i + EA(u_m(t), w_i) + E[u_m(t), w_i]\Phi(u_m, u_m)(t)) - \eta^{-1}u_m(t)^- w_i \right) dx$$
$$= \int_\Omega f(t)w_i \, dx, \ i = 1, ..., m,$$

$$(16)$$

$$u_m(0) = u_{0m}, \ \dot{u}_m(0) = u_{1m}, \ u_{1m} \to u_1 \text{ in } L_2(\Omega), \ u_{0m} \to u_0 \text{ in } H^2(\Omega) \quad (17)$$

The system (16) can then be expressed in the form

$$\ddot{\alpha}_i = F_i(t, \dot{\alpha}_1, ..., \dot{\alpha}_m, \alpha_1, ..., \alpha_m), \ i = 1, ..., m.$$

Its right-hand side satisfies the conditions for the local existence of a solution fulfilling the initial conditions corresponding with the functions u_{0m}, u_{1m}. Hence there exists a Galerkin approximation $u_m(t)$ defined on some interval $I_m \equiv [0, t_m]$, $0 < t_m < T$. After multiplying the equation (16) by $\dot{\alpha}_i(t)$, summing up with respect to i, taking in mind

$$\int_\Omega [u, v]y \, dx = \int_\Omega [u, y]v \, dx \quad (18)$$

if at least one element of $\{u, v, y\}$ belongs to $H_0^2(\Omega)$, cf. [5] and integrating we obtain the estimates

$$\|\dot{u}_m\|^2_{L_\infty(I;L_2(\Omega))} + \|u_m\|^2_{L_\infty(I;H^2(\Omega))} + \|\Phi(u_m, u_m)\|^2_{L_\infty(I;H^2(\Omega))}$$
$$+ \eta^{-1}\|u_m^-\|^2_{L_\infty(I;L_2(\Omega))} \le c \equiv c(f, u_0, u_1).$$

$$(19)$$

and

$$\|\Phi(u_m, u_m)\|_{L_\infty(I;W_p^2(\Omega))} \le c_p \equiv c_p(f, u_0, u_1) \forall p > 2. \quad (20)$$

The estimate (20) further implies

$$[u_m, \Phi(u_m, u_m)] \in L_2(I; L_r(\Omega)), \ r = \frac{2p}{p+2},$$
$$\left\|[u_m, \Phi(u_m, u_m)]\right\|_{L_2(I;L_r(\Omega))} \le c_r \equiv c_r(f, u_0, u_1).$$

$$(21)$$

From the equation (16) we obtain straightforwardly the estimate

$$\|\ddot{u}_m\|^2_{L_2(I;V_m^*)} \le c_\eta, \ m \in \mathbb{N}, \quad (22)$$

where $V_m \subset H^2(\Omega)$ is the linear hull of $\{w_i\}_{i=1}^m$.

We proceed with the convergence of the Galerkin approximation. Applying the estimates (19-22) and the compact imbedding theorem we obtain for any $p \in [1, \infty)$ a subsequence of

$\{u_m\}$ (denoted again by $\{u_m\}$), and a function u the convergence

$$
\begin{aligned}
u_m &\rightharpoonup^* u && \text{in } \mathcal{H}, \\
\dot{u}_m &\rightharpoonup^* \dot{u} && \text{in } L_\infty(I; L_2(\Omega)), \\
\ddot{u}_m &\rightharpoonup \ddot{u} && \text{in } \left(L_2(I; H^2(\Omega))\right)^*, \\
u_m &\to u && \text{in } C(I; H^{1-\varepsilon}(\Omega)) \cap L_\infty(I; H^{2-\varepsilon}(\Omega)), \;\; \varepsilon > 0, \\
\Phi(u_m, u_m) &\to \Phi(u, u) && \text{in } L_2(I; H^2(\Omega)), \\
\Phi(u_m, u_m) &\rightharpoonup^* \Phi(u, u) && \text{in } L_\infty(I; W_p^2(\Omega))
\end{aligned}
\tag{23}
$$

The convergence process (23) and the property (18) imply that a function u fulfils

$$
\int_Q \left(\ddot{u} z_\mu + E A(u, z_\mu) + E[u, \Phi(u, u)] z_\mu - \eta^{-1} u^- z_\mu \right) dx\, dt = \int_Q f z\, dx\, dt.
$$

Functions $\{z_\mu\}$ form a dense subset of the set $L_2(I; H^2(\Omega))$, hence a function u fulfils the identity (15). The initial conditions (3) follow due to (17) and the proof of the existence of a solution is complete.

The estimate (19) implies

$$
\begin{aligned}
&\|\dot{u}_\eta\|^2_{L_\infty(I; L_2(\Omega))} + \|u_\eta\|^2_{L_\infty(I; H^2(\Omega))} + \|\phi(u_\eta, u_\eta)\|^2_{L_\infty(I; W_p^2(\Omega))} \\
&\quad + \eta^{-1} \|u_\eta^-\|^2_{L_\infty(I; L_2(\Omega))} \le c \equiv c(f, u_0, u_1).
\end{aligned}
\tag{24}
$$

with u_η a solution of the penalized problem. Since $L_2(Q)$ is dense in $\left(L_\infty(I; L_2(\Omega))\right)^*$, for $f \in \left(L_\infty(I; L_2(\Omega))\right)^*$, there is a sequence $\{f_k\} \subset L_2(Q)$ such that $f_k \to f$ in $\left(L_\infty(I; L_2(\Omega))\right)^*$. The solutions u_k of the penalized problems with f_k satisfy the same convergence as in (23). Hence for any $f \in \left(L_\infty(I; L_2(\Omega))\right)^*$ there exists a solution of the penalized problem.

2 The limit process to the original problem

We rewrite the penalized problem (11) into the operator form

$$
\ddot{u}_\eta + B(u_\eta) - \eta^{-1} u_\eta^- = f
\tag{25}
$$

with

$$
B: H^2(\Omega) \to H^2(\Omega)^*, \;\; \langle B(v), w \rangle = E \int_\Omega \left(A(v, w) + [\Phi(v, v), v] w \right) dx, \;\; w \in H^2(\Omega)
$$

and the initial conditions (3). Let us multiply the equation (11) by $z = 1$. We get

$$
0 \le \int_Q \eta^{-1} u_\eta^-\, dx\, dt = \int_\Omega \dot{u}_\eta(T, \cdot)\, dx - \int_\Omega u_1\, dx - \int_Q f\, dx\, dt \le C,
\tag{26}
$$

where C is independent of η (cf. (24)). Since $B(u_\eta)$ takes its estimate in (24) and $L_1(\Omega) \hookrightarrow L_\infty(\Omega)^* \hookrightarrow H^2(\Omega)^*$ we get the dual estimate of the acceleration term

$$
\|\ddot{u}_\eta\|_{\mathcal{H}^*} \le C
\tag{27}
$$

Hence there is a sequence $\eta_k \searrow 0$ such that for $u_k \equiv u_{\eta_k}$ the following convergence hold

$$
\begin{aligned}
u_k &\rightharpoonup^* u & &\text{in } \mathcal{H}, \\
\dot{u}_k &\rightharpoonup^* \dot{u} & &\text{in } L_\infty(I; L_2(\Omega)), \\
\ddot{u}_k &\rightharpoonup^* \ddot{u} & &\text{in } \mathcal{H}^*, \\
u_k &\to u & &\text{in } C(I; H^{1-\varepsilon}(\Omega)) \cap L_\infty(I; H^{2-\varepsilon}(\Omega)), \ \varepsilon > 0, \\
\Phi(u_k, u_k) &\to \Phi(u, u) & &\text{in } L_2(I, H^2(\Omega)), \\
\Phi(u_k, u_k) &\rightharpoonup^* \Phi(u, u) & &\text{in } L_\infty(I; W_p^2(\Omega)), \\
\eta_k^{-1} u_k^- &\rightharpoonup^* g & &\text{in } L_\infty(Q)^* \hookrightarrow \mathcal{H}^*,
\end{aligned}
\tag{28}
$$

where g is the corresponding contact force.

The performed convergence has proved that the limit u satisfies the equation

$$
\ddot{u} + B(u) = f + g \tag{29}
$$

in the dual sense in $(L_\infty(I; H^2(\Omega)))^*$. To prove $\langle g, u \rangle = 0$ we take in mind that from the just proved facts $\langle g, u \rangle = \lim\limits_{k \to +\infty} 1/\eta_k \langle u_k^-, u_k^- \rangle = 0$, because $u_k^- \to 0$ in $L_\infty(Q)$ and $\eta_k^{-1} u_k^-$ is bounded in $L_1(Q)$. With this fact it is obvious that putting $v - u$ as a test function in (29) with an arbitrary $v \in \mathcal{C}$ we get the variational inequlity (10). The initial condition for u is satisfied in the sense of a weak limit in $H^2(\Omega)$ while that for $\dot{u}$ is satisfied in the sense of the integration by parts.

Hence we have proved the following

Theorem 2.1. *Let $u_0 \in H^2(\Omega)$, $u_1 \in L_2(\Omega)$ and $f \in \left(L_\infty(I; L_2(\Omega))\right)^*$. Then there exists a solution of the contact Problem $\mathcal{P}$.*

Remark 2.1. *The idea of the proof of this theorem was substantially based on the imbedding $H^2(\Omega) \hookrightarrow C(\bar{\Omega})$. It cannot be extended to contact problems of membranes or bodies.*

References

[1] I. Bock and J. Jarušek. Unilateral dynamic contact of viscoelastic von Kármán plates. *Adv. Math. Sci. Appl.*, 16(1):175–187, 2006.

[2] I. Bock and J. Jarušek. Unilateral dynamic contact of von Kármán plates with singular memory. *Appl. Math.*, 52(6):515–527, 2007.

[3] I. Bock and J. Lovíšek. On unilaterally supported viscoelastic von Kármán plate with a long memory. *Math. Comput. Simul.*, 61:399–407, 2003.

[4] I. Bock and J. Lovíšek. On a contact problem for a viscoelastic von Kármán plate and its semidiscretization. *Appl. Math.*, 50:203–217, 2005.

[5] P. G. Ciarlet and P. Rabier. *Les équations de von Kármán.* Springer, Berlin, 1980.

[6] C. Eck, J. Jarušek, and M. Krbec. *Unilateral Contact Problems in Mechanics. Variational Methods and Existence Theorems.* Number 270 in Monographs & Textbooks in Pure & Appl. Math. Chapman & Hall/CRC (Taylor & Francis Group), Boca Raton – London – New York – Singapore, 2005.

[7] H. Koch and A. Stachel. Global existence of classical solutions to the dynamical von Kármán equations. *Math. Methods in Applied Sciences*, 16:581–586, 1993.

[8] K. Maruo. Existence of solutions of some nonlinear wave equations. *Osaka Math. J.*, 22:21–30, 1984.

Recent progresses in the modeling
of material behavior

Gianpietro Del Piero [*]

Abstract

The quasi-static evolution of a body made of an inelastic, rate-independent material is determined by a minimum principle, according to which the evolution takes place along the direction which minimizes the increment of the sum of two functionals, free energy rate and dissipation rate. Specific examples from plasticity, fracture, and adhesive contact with damage are given.

1 The minimum energy principle in elasticity

The solution of the equilibrium problems of Mechanics takes advantage from their formulation as minimum problems for the associated energy functional. For elastic bodies, the energy of a body Ω is assumed to be the sum of two parts, the elastic strain energy, defined as the integral over the body of the elastic strain energy density w, and the energy of the loads

$$E(v) = \int_\Omega w(\nabla v(x))\, dx \, - \, \ell(v)\,.$$

The solutions u of the equilibrium problem are identified with the global minimizers of E in the set of all displacement fields v which belong to an appropriate function space H, and satisfy a boundary condition prescribed on a portion Γ of $\partial\Omega$:

$$E(u) = \min\{\, E(v) \quad | \quad v \in H,\ v(x) = \bar{u}(x)\ \forall x \in \Gamma \,\}\,.$$

Unfortunately, the approach based on global minimization does not go far beyond linear elasticity. Even for finite elasticity, there are still many open problems [1]. For inelastic bodies, global minimization makes no sense because the solutions are not path-independent, in the sense that different loading processes with the same final value do not determine, in general, the same solution.

93

2   The incremental problem in small-strain plasticity

For a large class of rate-independent materials, the energy can be split into the sum of a recoverable free energy and of a dissipative part. For example, in small-strain plasticity the energy is the sum of an elastic and a plastic part. The elastic part is

$$E^e(v) = \tfrac{1}{2} \int_\Omega \mathbb{C}\nabla v^e(x) \cdot \nabla v^e(x)\, dx - \ell(v)\,,$$

where $\mathbb{C}$ is the elastic tensor and ∇v^e is the elastic part of the decomposition $\nabla v = \nabla v^e + \nabla v^p$ of the displacement gradient. The plastic part, which depends on the whole path $\lambda \mapsto u_\lambda$ and not only on its final value, id defined by

$$E^p(\lambda \mapsto u_\lambda) = \int_0^1 \int_\Omega \mathbb{C}\nabla u_\lambda^e(x) \cdot \nabla \dot{u}_\lambda^p(x)\, dx\, d\lambda\,,$$

where the superimposed dot denotes differentiation with respect to λ. It is understood that all u_λ are in equilibrium with appropriate loads ℓ_λ.

Given a configuration u which satisfies a boundary condition $\bar{u}$ on Γ and is in equilibrium with the load ℓ, the incremental equilibrium problem from u consists in finding an incremental displacement field $\dot{u}$ which satisfies a given incremental boundary condition $\dot{\bar{u}}$ on Γ and is in equilibrium with a given load increment $\dot{\ell}$. This problem is equivalent to finding the field $\dot{u}$ which minimizes the total energy increment

$$\begin{aligned}
\dot{E}(v) &= \int_\Omega \mathbb{C}\nabla u^e(x) \cdot \nabla \dot{v}^e(x)\, dx - \ell(\dot{v}) - \dot{\ell}(u) + \int_\Omega \mathbb{C}\nabla u^e(x) \cdot \nabla \dot{v}^p(x)\, dx \\
&= \int_\Omega \mathbb{C}\nabla u^e(x) \cdot \nabla \dot{v}(x)\, dx - \ell(\dot{v}) - \dot{\ell}(u)\,.
\end{aligned}$$

The term $\dot{\ell}(u)$ has no influence on the minimization, since both $\dot{\ell}$ and u are known, and the sum of the remaining terms is zero by the virtual work equation, which expresses the fact that u is in equilibrium with ℓ. Then there is nothing to minimize in the first-order energy increment, and minimization has to be made on the second-order increment

$$\ddot{E}(v) = \int_\Omega \left(\mathbb{C}\nabla \dot{v}^e(x) \cdot \nabla \dot{v}(x) + \mathbb{C}\nabla u^e(x) \cdot \nabla \ddot{v}(x) \right) dx - \ell(\ddot{v}) - 2\dot{\ell}(\dot{v}) - \ddot{\ell}(u)\,.$$

Here, too, $\ddot{\ell}(u)$ is known, and the terms involving $\ddot{v}$ vanish due to the virtual work equation. Therefore, minimizing the energy increment becomes equivalent to minimizing the functional

$$I(\dot{v}) = \tfrac{1}{2} \int_\Omega \mathbb{C}\nabla \dot{v}^e(x) \cdot \nabla \dot{v}(x)\, dx - \dot{\ell}(\dot{v})\,. \tag{1}$$

This is the *principle of maximum plastic work* of Hill [6]. Versions of this principle appropriate to more general classes of rate-independent materials have been formulated by Fedelich and Ehrlacher [4] and, more recently, by Mielke [7].

3   The evolutionary problem in fracture mechanics

There are cases in which there is not enough regularity for the incremental problem. An example is fracture mechanics, in which not only differentiability, but even spatial continuity is lost at the opening of a fracture. In Griffith's theory, the total energy is

$$E(v) = \int_\Omega w(\nabla v(x))\, dx \; - \; \ell(v) + \gamma \text{ area } S(v)\,,$$

where γ is a positive constant and $S(v)$ is the jump set of v. In [5], for given families of data $\lambda \mapsto \bar{u}_\lambda$, $\lambda \mapsto \ell_\lambda$, Francfort and Marigo look for a family $\lambda \mapsto u_\lambda$ of global minimizers, subject to the condition

$$\mu < \lambda \quad \Longrightarrow \quad S(u_\mu) \subseteq S(u_\lambda)\,.$$

This *condition of irreversibility of fracture* transforms global into evolutionary minimization, since the solutions for each λ depend on the solutions for all preceding μ. An inconvenience of this procedure is better illustrated by a one-dimensional example, in which the energy has the form

$$E(v) = \tfrac{1}{2} \int_0^l k v'^2(x)\, dx + \gamma \, \#_v\,,$$

with l the length of the bar, k an elastic constant, and $\#_v$ the number of the jumps of v. No applied loads are assumed, and the boundary conditions are $v(0) = 0$ and $v(l) = \beta l$. They are more conveniently written in the form

$$\int_0^l v'(x)\, dx + \sum_{x \in S(v)} [v](x) \; = \; \beta l\,.$$

It is easy to see that v' is constant along the bar in every local or global minimizer. Then v' can be eliminated by combining the two preceding equations, and the energy takes the form

$$E(v) = \tfrac{1}{2}\, kl \left(\beta - l^{-1} \sum_{x \in S(v)} [v](x)\right)^2 + \gamma \, \#_v\,.$$

It is to immediate to determine the minima for each fixed $\#_v$: for $\#_v = 0$ there are no jumps, and therefore $E(v)$ is equal to $\tfrac{1}{2}kl\beta^2$, and for $\#_v > 0$ the minimum is $E(v) = \gamma\,\#_v$. These minima are plotted in Fig.1a as functions of β. We see that there is a critical value β_c such that the global minimum is attained at $\#_v = 0$ for $\beta < \beta_c$ and at $\#_v = 1$ for $\beta > \beta_c$.

But it is wrong to conclude that the bar breaks at $\beta = \beta_c$. Indeed, $\beta = \beta_c$ is only a situation in which two different configurations have the same energy, which is a global minimum. But these configurations need not be accessible from each other. In this specific example, a transition from $\#_v = 0$ to $\#_v = 1$ requires the finite energy γ, while a continuation along the branch $\#_v = 0$ only requires an infinitesimal energy [3]. The conclusion is that the bar never breaks. This confirms the well known fact that Griffith's theory is unable to predict the fracture onset [5]. This also shows that an analysis based on global minimization is
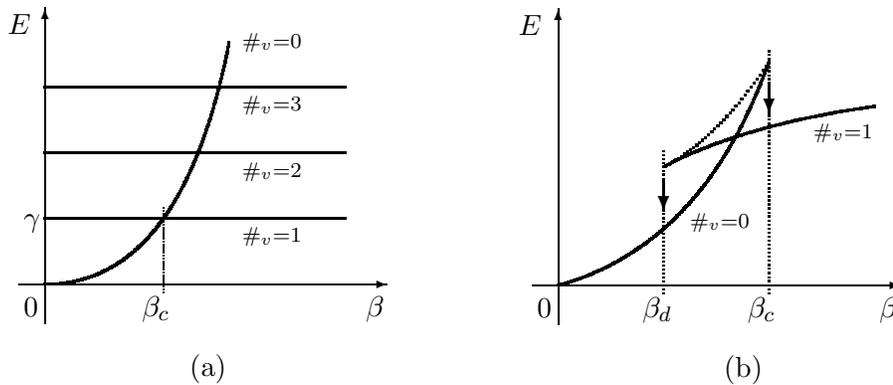
Figure 1: energy-elongation curves for the uniaxial stretching of a bar, according to Griffith's (a) and Barenblatt's (b) theories of fracture

unrealistic, and makes evident the fact that the solutions of the evolutionary problem are in general only local minimizers.

We now discuss the same example within Barenblatt's theory. In it, the energy of a jump is a concave, monotonic increasing function of the jump amplitude [3]. The result of the minimization is given in Fig.1b, where two curves, one for $\#_v = 0$ and one for $\#_v = 1$, are shown. All equilibrium configurations with $\#_v > 1$ are unstable and can be ignored.

Here, too, the model follows the curve $\#_v = 0$ even when the global minimum switches to the curve $\#_v = 1$. But now the curve $\#_v = 0$ ends at $\beta = \beta_c$, and the only equilibrium configurations for $\beta > \beta_c$ are those with $\#_v = 1$. Then one is forced to admit that a non-equilibrium transition occurs at $\beta = \beta_c$, followed by continuation along the curve $\#_v = 1$. In the figure, this transition is marked by a vertical arrow.

This type of response is not bad. Indeed, it gives a prediction about the fracture onset, which is supposed to occur at $\beta = \beta_c$. Moreover, it is not disturbing to admit that the opening of a jump is a catastrophic event which requires some non-equilibrium transition. The problem comes at unloading. For decreasing β, the model follows the curve $\#_v = 1$ backwards. When the curve ends at $\beta = \beta_d$, there is the alternative between admitting that equilibrium is impossible at loads smaller than β_d, and jumping back to the curve $\#_v = 0$, in contradiction with the assumption of the irreversibility of fracture. Both possibilities are unrealistic. This drawback of the model originates from the fact that in the applications of Barenblatt's theory the energy of a jump is usually considered as a part of the free energy. A more correct response is obtained by assuming that it is instead a dissipated energy [3]. This alternative approach is discussed in the next section.

4   An adhesive contact problem with elastic unloading

In the functional (1) associated with the incremental problem, the increment $\nabla \dot{v}^e$ of the elastic part of the deformation is defined by appropriate constitutive laws. In plasticity, the definition of $\nabla \dot{v}^e$ must account for *elastic unloading*, which is the property that, whenever an increment $\nabla \dot{v}$ has a non-null plastic part, the opposite increment $-\nabla \dot{v}$ is purely elastic. Here I show that, rather than from specific constitutive assumptions, elastic unloading can be obtained from incremental minimization. Once again, I will show this with the help of a simple one-dimensional example.

Consider two rigid bodies in contact through a planar material interface, modeled as an adhesive elastic layer [2]. The thickness u of the layer and the adhesion force σ are both supposed to be orthogonal to the plane. Given a displacement process $t \mapsto u(t)$ from the natural state $\sigma = u = 0$, the force $\sigma(t)$ is supposed to evolve in the following way. As long as the map $t \mapsto u(t)$ is increasing, the force follows the loading curve $\sigma = f(u)$ shown in Fig.2. When, after reaching a maximum u_m, the values $u(t)$ decrease, σ follows the unloading line

$$\sigma = g(u_m)\, u\,, \qquad g(u_m) = \frac{f(u_m)}{u_m}\,, \tag{2}$$

which is the straight line joining the point $(f(u_m), u_m)$ with the origin. As long as u stays smaller than u_m, σ follows back and forth the same line. When this threshold is exceeded, σ again follows the loading line, until a new maximum u_m is reached. Then σ follows the unloading line from the new u_m, and so forth. Clearly, the points (σ, u) with $\sigma > g(u_m)\, u$ are not any more accessible after the displacement u_m has been reached. Then the elastic modulus $g(u_m)$ can only decrease, and its decay is a measure of the progressive damage of the interface.

Let me show how the same response can be obtained by incremental energy minimization. For given u, u_m, let the elastic and the dissipative parts of the energy be defined by

$$E^e(u, u_m) = \tfrac{1}{2}\, g(u_m)\, u^2\,, \qquad E^d(u_m) = \int_0^{u_m} f(v)\, dv - \tfrac{1}{2}\, g(u_m)\, u_m^2\,.$$

In the figure, $E^e(u, u_m)$ is the area of the triangle with basis u and height σ, and $E^d(u_m)$ is the area of the region below the loading curve and above the unloading line. Let us minimize the increment of the total energy $E = E^e + E^d$

$$\dot{E}(u, u_m) = \tfrac{1}{2}\, g'(u_m)\, u^2 \dot{u}_m + g(u_m)\, u\dot{u} - \tfrac{1}{2}\, g'(u_m)\, u_m^2 \dot{u}_m\,.$$

for a given initial pair (u, u_m) and for a given increment $\dot{u}$ of u. The non-negativeness of the dissipation rate, which is the last term in the above equation, requires that $-g'(u_m)\dot{u}_m$ be non-negative, and because $g'(u_m)$ is assumed to be negative we have the condition

$$\dot{u}_m \geq 0\,. \tag{3}$$

Since $\dot{u}$ is known, the functional to be minimized reduces to $g'(u_m)(u^2 - u_m^2)\, \dot{u}_m$. If $u < u_m$, the functional is non-negative and the minimum is achieved for $\dot{u}_m = 0$.
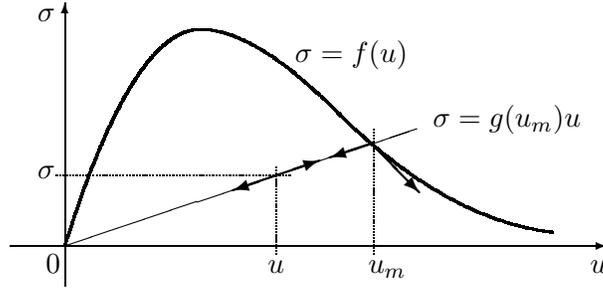
Figure 2: representation of interface behavior for contact with adhesion and damage

This means that at all points located below the loading curve the incremental response follows the unloading line $\sigma = g(u_m)\,u$. At points located on the loading curve u is equal to u_m, and therefore the functional to be minimized vanishes. Then, just as in the case examined in Section 2, one has to switch to the higher-order energy increment

$$\ddot{E}(u, u_m)\big|_{u=u_m} = -g'(u_m)\,u_m(\dot{u}_m^2 - 2\dot{u}_m\dot{u})\,.$$

Due to condition (3), the result of the minimization is

$$\dot{u}_m = \dot{u} \ \text{ if } \ \dot{u} \geq 0\,, \qquad \dot{u}_m = 0 \ \text{ if } \ \dot{u} < 0\,.$$

Thus, for points located on the loading curve a continuation follows the same curve if $\dot{u} \geq 0$, and the unloading line $\sigma = g(u_m)u$ if $\dot{u} < 0$. This is precisely the inelastic loading-elastic unloading behavior shown in Fig. 2.

References

[1] JM Ball, *Some open problems in elasticity*, CISM Course Poly-, Quasi-, and Rank-One Convexity in Applied Mechanics (2007), CISM Courses and Lectures, to appear

[2] G Del Piero, M Raous, M Schryve, *A unified model for adhesive interfaces with damage, viscosity, and friction*, in preparation

[3] G Del Piero, L Truskinovsky, *Elastic bars with decohesions*, in preparation

[4] B Fedelich, A Ehrlacher, *Sur un principe de minimum concernant des matériaux à comportement indépendant du temps physique*, C.R. Acad. Sci. Paris 308/II, 1391-1394, 1989

[5] G Francfort, J-J Marigo, *Revisiting brittle fracture as an energy minimization problem*, J. Mech. Phys. Solids 46, 1319-1342, 1998

[6] R Hill, *The Mathematical Theory of Plasticity*, Oxford University Press (1950). Reprinted in Oxford Classic Series, Clarendon Press, Oxford (1998)

[7] A Mielke, *Energetic formulation of multiplicative elasto-plasticity using dissipation distances*, Cont. Mech. Thermodyn. 15, 351-382, 2003

Numerical analysis of a contact problem including bone remodeling

José R. Fernández[*]   Rebeca Martínez[†]

Abstract

A contact problem between an elastic body and a deformable obstacle is numerically studied. The bone remodeling of the material is also taken into account in the model. The contact is modelled using the normal compliance contact condition. The variational problem is written as a nonlinear variational equation for the displacement field, coupled with a first-order ordinary differential equation to describe the physiological process of bone remodeling. A fully discrete approximations are introduced based on the finite element method to approximate the spatial variable and an Euler scheme to discretize the time derivatives. Error estimates are obtained, from which the linear convergence of the algorithm is derived under suitable regularity conditions. Finally, some two-dimensional numerical results are presented in order to demonstrate the performance of the algorithm.

1   Mechanical and variational problems

A contact problem between an elastic body and a deformable obstacle, including the bone remodeling process, is numerically studied in this paper. This bone remodeling model, derived by Cowin and Hegedus ([1, 6]), is a generalization of the nonlinear elasticity, and it is based on the fact that the living bone is continuously adapting itself to external stimuli. Since this process has an enormous effect on the overall behavior and health of the entire body, the ability of these models to predict the bone remodeling is of great importance.

During the last ten years, some papers dealt with mathematical issues of these models as the existence and uniqueness of weak solutions under some quite strong assumptions (see, e.g., [7, 8]) or the analysis of an asymptotic rod model ([3, 5, 4]). Here, our aim is to consider the d-dimensional contact problem, providing its numerical analysis and its numerical resolution.

Let us denote by $\Omega \subset \mathbb{R}^d$, $d = 1, 2, 3$, an open bounded domain and let $\Gamma = \partial\Omega$ be its outer surface which is assumed to be Lipschitz continuous and it is divided into three disjoint

parts Γ_D, Γ_N and Γ_C. The body is being acted upon by a volume force of density $\boldsymbol{f}$, it is clamped on Γ_D and surface tractions with density $\boldsymbol{g}$ act on Γ_N. Finally, we assume that the body may come in contact with a deformable obstacle on the boundary part Γ_C, which is located at a distance s, measured along the outward unit normal vector $\boldsymbol{\nu}$.

Let $\boldsymbol{u} = (u_i)_{i=1}^d$ be the displacement field, $\boldsymbol{\sigma} = (\sigma_{ij})_{i,j=1}^d$ is the stress field, $\boldsymbol{\varepsilon}(\boldsymbol{u}) = (\varepsilon_{ij}(\boldsymbol{u}))_{i,j=1}^d$ represents the linearized strain field given by $\varepsilon_{ij}(\boldsymbol{u}) = \dfrac{1}{2}\left(\dfrac{\partial u_i}{\partial x_j} + \dfrac{\partial u_j}{\partial x_i}\right)$, and denote by e the so-called bone remodeling function, which measures the change in the volume fraction from a reference configuration.

The body is assumed elastic and, according to [1, 6], the constitutive law is then written in the form $\boldsymbol{\sigma} = (\xi_0 + e)\mathcal{C}(e)\boldsymbol{\varepsilon}(\boldsymbol{u})$, where ξ_0 represents the reference volume fraction and $\mathcal{C}(e) = (C_{ijkl}(e))_{i,j,k,l=1}^d$ is a constitutive function whose properties will be described below.

Since the contact is assumed with a deformable obstacle, the well-known normal compliance contact condition is employed; that is, the normal stress $\sigma_\nu = \boldsymbol{\sigma}\boldsymbol{\nu}\cdot\boldsymbol{\nu}$ on Γ_C is given by $-\sigma_\nu = p_\nu(u_\nu - s)$, where $u_\nu = \boldsymbol{u}\cdot\boldsymbol{\nu}$ denotes the normal displacement in such a way that, when $u_\nu > s$, the difference $u_\nu - s$ represents the interpenetration of the body's asperities into those of the foundation. We also assume that the contact is frictionless, i.e. the tangential component of the stress field, denoted $\boldsymbol{\sigma}_\tau = \boldsymbol{\sigma}\boldsymbol{\nu} - \sigma_\nu\boldsymbol{\nu}$, vanishes on the contact surface.

The evolution of the bone remodeling function is obtained from the first-order ordinary differential equation (see [1, 6]), $\dot{e} = a(e) + \mathcal{A}(e) : \boldsymbol{\varepsilon}(\boldsymbol{u})$, where $a(e)$ is a constitutive function and $\mathcal{A}(e) = (A_{ij}(e))_{i,j=1}^d$ denote the bone remodeling rate coefficients. Moreover, a dot above a variable represents the time derivative.

Finally, let $\gamma > 0$ be the density of the full elastic material which is assumed constant for the sake of simplicity, e_0 denotes the initial bone remodeling function and Φ_L is the truncation operator defined as $\Phi_L(r) = r$ is $|r| \leq L$ or $\Phi_L(r) = L$ otherwise.

We obtain now the variational formulation. First, let us denote by $Y = L^2(\Omega)$ and $H = [L^2(\Omega)]^d$, and define the variational spaces V and Q as,

$$V = \{\boldsymbol{v} \in [H^1(\Omega)]^d \; ; \; \boldsymbol{v} = \boldsymbol{0} \quad \text{on} \quad \Gamma_D\},$$
$$Q = \{\boldsymbol{\tau} = (\tau_{ij})_{i,j=1}^d \in [L^2(\Omega)]^{d\times d} \; ; \; \tau_{ij} = \tau_{ji}, \quad 1 \leq i,j \leq d\}.$$

The following assumptions are done on the given data:

- The elasticity coefficients $(\xi_0 + e)C_{ijkl}$ are assumed to be bounded, symmetric, Lipschitz and V-elliptic.

- The normal compliance function p_ν is assumed to be Lipschitz, monotone and satisfying $p_\nu(\boldsymbol{x}, r) = 0$ for all $r \leq 0$ and $\boldsymbol{x} \in \Gamma_C$.

- The constitutive function $a(e)$ and the bone remodeling rate coefficients $A_{ij}(e)$ are Lipschitz and bounded functions.

- The reference volume fraction ξ_0 satisfies $\xi_0 \in C(\overline{\Omega})$ and $0 < \xi_0^m \leq \xi_0(\boldsymbol{x}) \leq 1$, for some $0 < \xi_0^m < 1$ and for all $\boldsymbol{x} \in \overline{\Omega}$.

- The density forces are continuous functions with respect to the time variable and the initial value of the bone remodeling function e_0 lies in $C(\overline{\Omega})$.

For every $e \in L^\infty(\Omega)$, let us define the following bilinear form $c(e; \cdot, \cdot) : V \times V \to \mathbb{R}$,

$$c(e; \boldsymbol{u}, \boldsymbol{v}) = \int_\Omega (\xi_0 + e) \mathcal{C}(e) \boldsymbol{\varepsilon}(\boldsymbol{u}) \,:\, \boldsymbol{\varepsilon}(\boldsymbol{v}) \, d\boldsymbol{x} \quad \forall \boldsymbol{u}, \boldsymbol{v} \in V,$$

the linear form $L(e; \cdot) : V \to \mathbb{R}$ given by

$$L(e; \boldsymbol{v}) = \int_\Omega \gamma(\xi_0 + \Phi_L(e)) \boldsymbol{f} \cdot \boldsymbol{v} \, d\boldsymbol{x} + \int_{\Gamma_N} \boldsymbol{g} \cdot \boldsymbol{v} \, da \quad \forall \boldsymbol{v} \in V,$$

and the functional $j : V \times V \to \mathbb{R}$ as $(v_\nu = \boldsymbol{v} \cdot \boldsymbol{\nu}$ for all $\boldsymbol{v} \in V)$,

$$j(\boldsymbol{u}, \boldsymbol{v}) = \int_{\Gamma_C} p_\nu(u_\nu - s) \, v_\nu \, da \qquad \forall \boldsymbol{u}, \boldsymbol{v} \in V.$$

Applying Green's formula, we then derive the following variational formulation.

Problem VP. *Find a displacement field* $\boldsymbol{u} : [0, T] \to V$ *and the bone remodeling function* $e : [0, T] \to L^\infty(\Omega)$ *such that* $e(0) = e_0$ *and for a.e.* $t \in (0, T)$,

$$c(e(t); \boldsymbol{u}(t), \boldsymbol{v}) + j(\boldsymbol{u}(t), \boldsymbol{v}) = L(e(t); \boldsymbol{v}) \quad \forall \boldsymbol{v} \in V,$$
$$\dot{e}(t) = a(e(t)) + \mathcal{A}(e(t)) \,:\, \boldsymbol{\varepsilon}(\boldsymbol{u}(t)).$$

The following result, which states the existence of a unique weak solution to Problem VP, is obtained proceeding as in [4].

Theorem 1.1. *Let the previous assumptions hold. Assume that, for a given function* $e \in C^1([0, T]; C(\overline{\Omega}))$, *the unique solution to the problem*

$$\boldsymbol{u}(t) \in V, \quad c(e(t); \boldsymbol{u}(t), \boldsymbol{v}) + j(\boldsymbol{u}(t), \boldsymbol{v}) = L(e(t); \boldsymbol{v}) \quad \forall \boldsymbol{v} \in V,$$

has the regularity $\boldsymbol{u} \in C([0, T]; [H^3(\Omega)]^d)$. *Then, there exists a unique solution to Problem VP with the regularity* $\boldsymbol{u} \in C([0, T]; [C^1(\overline{\Omega})]^d)$ *and* $e \in C^1([0, T]; C(\overline{\Omega}))$.

2 Numerical analysis of a fully discrete scheme

The discretization of problem VP is done in two steps. First, we consider two finite dimensional spaces $V^h \subset V$ and $B^h \subset L^\infty(\Omega) \subset Y$, approximating the spaces V and $L^\infty(\Omega)$, respectively. Here, $h > 0$ denotes the spatial discretization parameter.

Secondly, the time derivatives are discretized by using a uniform partition of the time interval $[0, T]$, denoted by $0 = t_0 < t_1 < \ldots < t_N = T$, and let k be the time step size, $k = T/N$. Moreover, for a continuous function $f(t)$, we let $f_n = f(t_n)$.

Using the forward Euler scheme, the fully discrete approximation of problem VP is as follows.

Problem VP^{hk}. *Find a discrete displacement field* $\boldsymbol{u}^{hk} = \{\boldsymbol{u}_n^{hk}\}_{n=0}^N \subset V^h$ *and a discrete bone remodeling function* $e^{hk} = \{e_n^{hk}\}_{n=0}^N \subset B^h$ *such that* $e_0^{hk} = e_0^h$ *and for* $n = 1, \ldots, N$,

$$c(e_n^{hk}; \boldsymbol{u}_n^{hk}, \boldsymbol{v}^h) + j(\boldsymbol{u}_n^{hk}, \boldsymbol{v}^h) = L(e_n^{hk}; \boldsymbol{v}^h) \quad \forall \boldsymbol{v}^h \in V^h,$$

$$\frac{e_n^{hk} - e_{n-1}^{hk}}{k} = a(e_{n-1}^{hk}) + \mathcal{A}(e_{n-1}^{hk})) : \boldsymbol{\varepsilon}(\boldsymbol{u}_{n-1}^{hk}),$$

where e_0^h *is an appropriate approximation of the initial condition* e_0 *and* $\boldsymbol{u}_0^{hk}$ *is defined as the solution to the following problem,*

$$\boldsymbol{u}_0^{hk} \in V^h, \quad c(e_0^h; \boldsymbol{u}_0^{hk}, \boldsymbol{v}^h) + j(\boldsymbol{u}_0^{hk}, \boldsymbol{v}^h) = L(e_0^h; \boldsymbol{v}^h) \quad \forall \boldsymbol{v}^h \in V^h.$$

Using classical results on nonlinear variational equations it is straightforward to obtain the existence and uniqueness of the discrete solution.

The aim of this section is to derive error estimates on the numerical errors $\|\boldsymbol{u}_n - \boldsymbol{u}_n^{hk}\|_V$ and $\|e_n - e_n^{hk}\|_Y$. Therefore, we have the following main error estimates result.

Theorem 2.1. *Let the assumptions of Theorem 1.1 hold. Let* $(\boldsymbol{u}, e)$ *and* $(\boldsymbol{u}^{hk}, e^{hk})$ *denote the respective solutions to problems VP and VP^{hk}. Then there exists a positive constant* $C > 0$, *independent of h and k, such that, for all* $\{\boldsymbol{v}_n^h\}_{n=0}^N \subset V^h$,

$$\max_{0 \le n \le N} \{\|e_n - e_n^{hk}\|_Y^2 + \|\boldsymbol{u}_n - \boldsymbol{u}_n^{hk}\|_V^2\} \le \|e_0 - e_0^h\|_Y^2 + C\Big(k \sum_{j=1}^N \Big[\|\dot{e}_j - \delta e_j\|_Y^2 + \|\boldsymbol{u}_j - \boldsymbol{u}_{j-1}\|_V^2\Big]$$

$$+ k^2 + \max_{1 \le n \le N} \|\boldsymbol{u}_n - \boldsymbol{v}_n^h\|_V^2 + \|\boldsymbol{u}_0 - \boldsymbol{u}_0^{hk}\|_V^2\Big),$$

where the notation $\delta e_j = (e_j - e_{j-1})/k$ *is employed.*

The above error estimates are the basis for the analysis of the convergence rate of the algorithm, which we now present.

Let Ω be a polyhedral domain and denote by $\mathcal{T}^h$ a triangulation of $\overline{\Omega}$ compatible with the partition of the boundary $\Gamma = \partial \Omega$ into Γ_D, Γ_N and Γ_C. Let V^h and B^h consist of continuous and piecewise affine functions and piecewise constant functions and let the discrete initial condition e_0^h be defined by $e_0^h = \pi^h e_0$, where $\pi^h : C(\overline{\Omega}) \to B^h$ is the standard finite element interpolation operator.

Corollary 2.1. *Let the assumptions of Theorem* 1.1 *hold. Under the additional regularity conditions* $e \in C([0,T]; H^1(\Omega)) \cap H^2(0,T;Y)$ *and* $\boldsymbol{u} \in C^1([0,T];V) \cap C([0,T];[H^2(\Omega)]^d)$,

the fully discrete scheme is linearly convergent; that is, there exists a positive constant C, independent of h and k, such that

$$\max_{0 \leq n \leq N} \left\{ \|\boldsymbol{u}_n - \boldsymbol{u}_n^{hk}\|_V + \|e_n - e_n^{hk}\|_Y \right\} \leq C(h + k).$$

3   Numerical results

In this section we present some numerical results obtained in a two-dimensional example. For details concerning the implementation of the numerical scheme provided by Problem VP^{hk} we refer the reader to [2].

As a two-dimensional example, we consider the domain $\Omega = (0, 6) \times (0, 1.2)$ which is clamped on the boundary part $\Gamma_D = \{0\} \times [0, 1.2]$. No volume forces are supposed to act in the body, a linearly increasing surface force acts on the boundary part $[0, 6] \times \{1.2\}$ and, finally, the body is supposed to be in contact with a deformable obstacle on the contact boundary $\Gamma_C = [0, 6] \times \{0\}$.

The following data were employed in this example:

$$T = 105 \, \text{days}, \quad \boldsymbol{f} = \boldsymbol{0} \, N/m^3, \quad \boldsymbol{g}(x, y, t) = (0, -5x) \, N/m^2, \quad \mathcal{C}(e) = \frac{1}{\xi_0 + e}(\mathcal{C}^0 + \mathcal{C}^1 e),$$

$$\mathcal{A}(e) = \mathcal{A}^0 + \mathcal{A}^1 e, \quad \xi_0 = 0.892, \quad \gamma = 1740 \, Kg/m^3, \quad s = 0 \, m, \quad a(e) = a_0 + a_1 e + a_2 e^2,$$

$$a_0 = -1296 \times 10^{-4} \, (100 \text{ days})^{-1}, \quad a_1 = -1296 \times 10^{-2} \, (100 \text{ days})^{-1},$$

$$a_2 = 216 \times 10^{-2} \, (100 \text{ days})^{-1}, \quad p_\nu(r) = \mu \max\{r, 0\}, \quad \mu = 10^4,$$

where the fourth-order tensors $\mathcal{C}^0 = (C_{ijkl}^0)_{i,j,k,l=1}^2$ and $\mathcal{C}^1 = (C_{ijkl}^1)_{i,j,k,l=1}^2$ and the second-order tensors $\mathcal{A}^0 = (A_{ij}^0)_{i,j=1}^2$ and $\mathcal{A}^1 = (A_{ij}^1)_{i,j=1}^2$ have the following components:

$$C_{1111}^0 = 25.69 \, GPa, \quad C_{2211}^0 = 11.67 \, GPa, \quad C_{2222}^0 = 25.69 \, GPa, \quad C_{1211}^0 = C_{1222}^0 = 0 \, GPa,$$

$$A_{22}^0 = -216 \, (100 \text{days})^{-1}, \quad C_{1111}^1 = 252.08 \, GPa, \quad C_{2211}^1 = 114.58 \, GPa, \quad C_{1212}^0 = 7 \, GPa,$$

$$C_{1211}^0 = C_{1222}^0 = 0 \, GPa, \quad C_{1212}^1 = 68.75 \, GPa, \quad A_{11}^0 = 216 \, (100 \text{days})^{-1}, \quad A_{12}^0 = A_{21}^0 = 0,$$

$$A_{12}^0 = A_{21}^0 = 0, \quad A_{11}^1 = 216 \, (100 \text{days})^{-1}, \quad A_{22}^1 = 216 \, (100 \text{days})^{-1}, \quad C_{2222}^1 = 252.08 \, GPa.$$

Moreover, we assume that the initial bone remodeling function is given by

$$e_0(x, y) = 0.01 \sin(\frac{\pi y}{3}) \quad \forall (x, y) \in (0, 6) \times (0, 1.2).$$

Taking $k = 0.01$ as the time discretization parameter, the displacements field (multiplied by 20) and the reference configuration are plotted in Figure 1 at initial time (left) and after 105 days (right). We observe that the deformation has decreased and that no penetration into the obstacle has been produced, because of the size of the deformability coefficient μ.

Moreover, in Figure 2 the bone remodeling function is shown at initial time (left) and after 105 days (right). As can be seen, the bone remodeling function is positive on the right part which causes that the stiffness increases and so the displacements decrease there. We also notice that the bone remodeling function seems to be constant through the vertical direction at final time.

Figure 1: Reference configuration and displacements field at initial time (left) and after 105 days (right).
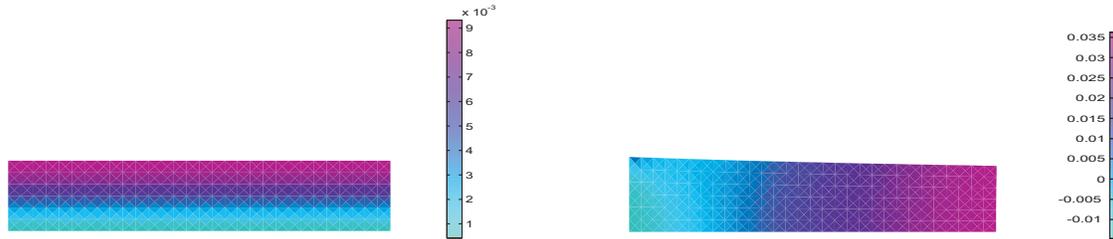
Figure 2: Bone remodeling function at initial time (left) and after 105 days (right).

References

[1] S. C. Cowin and D. H. Hegedus. Bone remodeling I: theory of adaptive elasticity. *J. Elasticity*, 6(3):313–326, 1976.

[2] J. R. Fernández and R. Martínez. Numerical analysis of a contact problem including bone remodeling. Preprint, 2007.

[3] I. N. Figueiredo. Approximation of bone remodeling models. *J. Math. Pures Appl.*, 84:1794–1812, 2005.

[4] I. N. Figueiredo, C. Leal, and C. S. Pinto. Conical differentiability for bone remodeling contact rod models. *ESAIM: Control, Optimisation and Calculus of Variations*, 11(3):382–400, 2005.

[5] I. N. Figueiredo, C. Leal, and C. S. Pinto. Shape analysis of an adaptive elastic rod model. *SIAM J. Appl. Math.*, 66(1):153–173, 2005.

[6] D. H. Hegedus and S. C. Cowin. Bone remodeling II: small strain adaptive elasticity. *J. Elasticity*, 6(4):337–352, 1976.

[7] J. Monnier and L. Trabucho. Existence and uniqueness of a solution to an adaptive elasticity model. *Math. Mech. Solids*, 3:217–228, 1998.

[8] L. Trabucho. Non-linear bone remodeling: an existence and uniqueness result. *Math. Methods Appl. Sci.*, 23:1331–134, 2000.

A posteriori error analyses for the unilateral contact problem with Coulomb friction in elastostatics

Patrick Hild[*]   Vanessa Lleras[†]

Abstract

In scientific computing, an important task consists of evaluating numerically the quality of the finite element computations by using a posteriori error estimators. The aim of the present work is to propose and study a residual error estimator for the unilateral contact problem with Coulomb friction in elastostatics. A recent uniqueness result by Y. Renard in [13] allows us to perform such a study and to obtain both upper and lower bounds of the discretization error.

1 The frictional contact problem in elasticity

We consider the deformation of an elastic body occupying, in the initial unconstrained configuration a domain Ω in $\mathbb{R}^2$ where plane strain assumptions are assumed. The Lipschitz boundary $\partial\Omega$ of Ω consists of Γ_D, Γ_N and Γ_C where the measure of Γ_D does not vanish. The body Ω is clamped on Γ_D and subjected to surface traction forces $\mathbf{F}$ on Γ_N; the body forces are denoted $\mathbf{f}$. In the initial configuration, the part Γ_C is a straight line segment considered as the candidate contact surface on a rigid foundation for the sake of simplicity which means that the contact zone cannot enlarge during the deformation process. The contact is assumed to be frictional and the stick, slip and separation zones on Γ_C are not known in advance. We denote by $\mu \geq 0$ the given friction coefficient on Γ_C. The unit outward normal and tangent vectors of $\partial\Omega$ are $\mathbf{n} = (n_1, n_2)$ and $\mathbf{t} = (-n_2, n_1)$ respectively. The frictional contact problem consists of finding the displacement field $\mathbf{u} : \Omega \to \mathbb{R}^2$ satisfying (1)–(4):

$$\mathbf{div}\,\boldsymbol{\sigma}(\mathbf{u}) + \mathbf{f} \;=\; \mathbf{0}, \quad \boldsymbol{\sigma}(\mathbf{u}) = \mathcal{C}\,\boldsymbol{\varepsilon}(\mathbf{u}) \quad \text{in } \Omega, \tag{1}$$

$$\mathbf{u} \;=\; \mathbf{0} \quad \text{on } \Gamma_D, \quad \boldsymbol{\sigma}(\mathbf{u})\mathbf{n} = \mathbf{F} \quad \text{on } \Gamma_N. \tag{2}$$

The linearized strain tensor field is $\varepsilon(\mathbf{u}) = (\boldsymbol{\nabla}\mathbf{u} + \boldsymbol{\nabla}^T\mathbf{u})/2$, $\mathcal{C}$ is the fourth order symmetric and elliptic tensor of linear elasticity an $\boldsymbol{\sigma}(\mathbf{u})$ represents the stress tensor field. Afterwards we adopt the following notation for any displacement field $\mathbf{v}$ and for any density of surface forces $\boldsymbol{\sigma}(\mathbf{v})\mathbf{n}$ defined on $\partial\Omega$: $\mathbf{v} = v_n\mathbf{n} + v_t\mathbf{t}$ and $\boldsymbol{\sigma}(\mathbf{v})\mathbf{n} = \sigma_n(\mathbf{v})\mathbf{n} + \sigma_t(\mathbf{v})\mathbf{t}$. On Γ_C, the three conditions representing unilateral contact are given by

$$u_n \leq 0, \qquad \sigma_n(\mathbf{u}) \leq 0, \qquad \sigma_n(\mathbf{u})\, u_n = 0, \tag{3}$$

and the static Coulomb friction law is summarized by the following conditions:

$$\begin{cases} u_t = 0 \implies |\sigma_t(\mathbf{u})| \leq \mu|\sigma_n(\mathbf{u})|, \\ u_t \neq 0 \implies \sigma_t(\mathbf{u}) = -\mu|\sigma_n(\mathbf{u})|\dfrac{u_t}{|u_t|}. \end{cases} \tag{4}$$

The variational formulation of problem (1)–(4) in its mixed form consists of finding $(\mathbf{u}, \boldsymbol{\lambda}) = (\mathbf{u}, \lambda_n, \lambda_t) \in \mathbf{V} \times \mathbf{M}(\mu\lambda_n)$ which satisfy:

$$\begin{cases} a(\mathbf{u},\mathbf{v}) + b(\boldsymbol{\lambda},\mathbf{v}) = L(\mathbf{v}), & \forall \mathbf{v} \in \mathbf{V}, \\ b(\boldsymbol{\nu} - \boldsymbol{\lambda},\mathbf{u}) \leq 0, & \forall \boldsymbol{\nu} = (\nu_n, \nu_t) \in \mathbf{M}(\mu\lambda_n), \end{cases} \tag{5}$$

where $\mathbf{V} = \{\mathbf{v} \in (H^1(\Omega))^2 : \mathbf{v} = \mathbf{0} \text{ on } \Gamma_D\}$, $\mathbf{M}(\mu\lambda_n) = M_n \times M_t(\mu\lambda_n)$ with $M_n = \{\nu \in X'_n : \nu \geq 0 \text{ on } \Gamma_C\}$ and for any $g \in M_n$, $M_t(g) = \{\nu \in X'_t : -g \leq \nu \leq g \text{ on } \Gamma_C\}$, where X'_n (resp. X'_t) is the dual space of of X_n (resp. X_t) with $X_n = \{v_{n|_{\Gamma_C}} : \mathbf{v} \in \mathbf{V}\}$ (resp. $X_t = \{v_{t|_{\Gamma_C}} : \mathbf{v} \in \mathbf{V}\}$). Note that $H^{1/2}_{00}(\Gamma_C) \subset X_n \subset H^{1/2}(\Gamma_C)$, $H^{1/2}_{00}(\Gamma_C) \subset X_t \subset H^{1/2}(\Gamma_C)$ and that the inequality conditions incorporated in the definitions of M_n and $M_t(g)$ have to be understood in the dual sense. In (5), we have

$$a(\mathbf{u},\mathbf{v}) = \int_\Omega (\mathcal{C}\varepsilon(\mathbf{u})) : \varepsilon(\mathbf{v})\, d\Omega, \qquad L(\mathbf{v}) = \int_\Omega \mathbf{f}.\mathbf{v}\, d\Omega + \int_{\Gamma_N} \mathbf{F}.\mathbf{v}\, d\Gamma,$$

$$b(\boldsymbol{\nu},\mathbf{v}) = \langle \nu_n, v_n \rangle_{X'_n, X_n} + \langle \nu_t, v_t \rangle_{X'_t, X_t}$$

for any $\mathbf{u}$ and $\mathbf{v}$ in $(H^1(\Omega))^2$ and $\boldsymbol{\nu} = (\nu_n, \nu_t) \in X'_n \times X'_t$.

It is easy to see that if $(\mathbf{u}, \lambda_n, \lambda_t)$ is a solution of (5), then $\lambda_n = -\sigma_n(\mathbf{u})$ and $\lambda_t = -\sigma_t(\mathbf{u})$. The existence of a weak solution to (1)–(4) has been proved for small friction coefficients, first in [11] (see also [4] and the references therein for larger bounds ensuring existence). More precisely existence holds if $\mu \leq \sqrt{3 - 4P}/(2 - 2P)$ where $0 \leq P < 1/2$ denotes Poisson's ratio. In [5, 6] some multi-solutions of the problem (1)–(4) are exhibited for triangular or quadrangular domains. These multiple solutions involve either an infinite set of slipping solutions or two isolated (stick and separation) configurations. Note that these examples of non-uniqueness involve large friction coefficients (i.e., $\mu > \sqrt{(1 - P)/P}$) and tangential displacements with a constant sign on Γ_C. The forthcoming partial uniqueness result is obtained in [13]: it defines some cases where it is possible to affirm that a solution to the

Coulomb friction problem is in fact the unique solution. More precisely, if a regular solution to the Coulomb friction problem exists (here the denomination regular means, roughly speaking, that the transition is smooth when the slip direction changes) and if the friction coefficient is small enough then this solution is the only one. We now introduce the space of multipliers M of the functions ξ defined on Γ_C such that the following norm is finite:

$$\|\xi\|_a = \sup_{\substack{v_t \in X_t \\ v_t \neq 0}} \frac{\|\xi v_t\|_{a,\Gamma_C}}{\|v_t\|_{a,\Gamma_C}} \quad \text{with} \quad \|v_t\|_{a,\Gamma_C} = \inf_{\substack{\mathbf{w} \in \mathbf{V}, \\ w_t = v_t \text{ on } \Gamma_C}} (a(\mathbf{w}, \mathbf{w}))^{1/2}.$$

Since Γ_C is assumed to be straight, M contains for any $\varepsilon > 0$ the space $H^{1/2+\varepsilon}(\Gamma_C)$. The partial uniqueness result is given assuming that $\lambda_t = \mu \lambda_n \xi$, with $\xi \in M$. It is easy to see that it implies $|\xi| \leq 1$ a.e. on the support of λ_n. More precisely, this implies that $\xi \in \mathrm{Dir}_t(u_t)$ a.e. on the support of λ_n, where $\mathrm{Dir}_t(.)$ is the subdifferential of the convex map $\mathbb{R}^2 \ni x \longmapsto |x_t|$. This means that it is possible to assume that $\xi \in \mathrm{Dir}_t(u_t)$ a.e. on Γ_C.

Proposition 1.1. *[13] Let $(\mathbf{u}, \boldsymbol{\lambda})$ be a solution to Problem (5) such that $\lambda_t = \mu \lambda_n \xi$, with $\xi \in M$, $\xi \in Dir_t(u_t)$ a.e. on Γ_C and $\mu < (C\|\xi\|_a)^{-1}$. Then $(\mathbf{u}, \boldsymbol{\lambda})$ is the unique solution to Problem (5).*

The case $\xi \equiv 1$ corresponds to an homogeneous sliding direction and the previous result is complementary to the non-uniqueness results obtained in [5, 6]. For a more precise discussion concerning the assumption $\lambda_t = \mu \lambda_n \xi$, $\xi \in M$, $\xi \in \mathrm{Dir}_t(u_t)$ and the cases where the assumption cannot be fulfilled independently of the regularity of the solution we refer the reader to [7], Remark 2.

2 Mixed finite element approximation

We approximate this problem by a standard finite element method. Namely we fix a regular family of meshes $T_h, h > 0$, made of closed triangles. For $K \in T_h$ we recall that h_K is the diameter of K and $h = \max_{K \in T_h} h_K$. Let E_K denote the set of edges of the element K and set $E_K^{int} = E_K \cap \Omega$, $E_K^N = E_K \cap \Gamma_N$, $E_K^C = E_K \cap \Gamma_C$. We will need local subdomains (also called patches): as usual, let ω_K be the union of all elements having a nonempty intersection with K. The finite element space used in Ω is then defined by

$$\mathbf{V}_h = \left\{ \mathbf{v}_h \in (C(\overline{\Omega}))^2 : \quad \forall \kappa \in T_h, \quad \mathbf{v}_h|_\kappa \in (\mathbb{P}_1(\kappa))^2, \quad \mathbf{v}_h|_{\Gamma_D} = \mathbf{0} \right\}.$$

In order to express the frictional contact constraints by using Lagrange multipliers on the contact zone, we have to introduce the space:

$$W_h = \left\{ \nu_h \in C(\overline{\Gamma_C}) : \exists \mathbf{v}_h \in \mathbf{V}_h \text{ s.t. } \mathbf{v}_h \cdot \mathbf{n} = \nu_h \text{ on } \Gamma_C \right\}.$$

The choice of the space W_h allows us to define the following closed convex cones:

$$M_{hn} = \left\{ \nu_h \in W_h : \int_{\Gamma_C} \nu_h \psi_h \, d\Gamma \geq 0, \ \forall \psi_h \in W_h, \psi_h \geq 0 \right\},$$

and, for $g \in M_{hn}$:

$$M_{ht}(g) = \left\{ \nu_h \in W_h : \left| \int_{\Gamma_C} \nu_h \psi_h \, d\Gamma \right| \leq \int_{\Gamma_C} g \psi_h \, d\Gamma, \ \forall \psi_h \in W_h, \psi_h \geq 0 \right\}.$$

The discrete formulation of the frictional contact problem is to find $(\mathbf{u}_h, \boldsymbol{\lambda}_h) = (\mathbf{u}_h, \lambda_{hn}, \lambda_{ht}) \in \mathbf{V}_h \times \mathbf{M}_h(\mu\lambda_{hn}) = \mathbf{V}_h \times M_{hn} \times M_{ht}(\mu\lambda_{hn})$ satisfying:

$$\begin{cases} a(\mathbf{u}_h, \mathbf{v}_h) + b(\boldsymbol{\lambda}_h, \mathbf{v}_h) = L(\mathbf{v}_h), & \forall \mathbf{v}_h \in \mathbf{V}_h, \\ b(\boldsymbol{\nu}_h - \boldsymbol{\lambda}_h, \mathbf{u}_h) \leq 0, & \forall \boldsymbol{\nu}_h = (\nu_{hn}, \nu_{ht}) \in \mathbf{M}_h(\mu\lambda_{hn}). \end{cases} \tag{6}$$

Using a fixed point argument it can be proven that the problem (6) admits at least a solution and that there is a unique solution when $\mu \leq C(h)$. Unfortunately one can not prove that $C(h)$ does not vanish when h vanishes.

3 An error estimator: η

The usual norm of $(H^s(D))^d$ is denoted by $\|\cdot\|_{s,D}$ and we keep the same notation when $d = 1$ or $d = 2$. For shortness the $(L^2(D))^d$-norm will be denoted by $\|\cdot\|_D$ when $d = 1$ or $d = 2$. Finally the notation $a \lesssim b$ means that there exists a positive constant C independent of a and b (and of the meshsize of the triangulation) such that $a \leq C b$.

3.1 Definition of the residual error estimator

Definition 3.1. *The local and global residual error estimators are defined by*

$$\eta = \left(\sum_{K \in T_h} \eta_K^2 \right)^{1/2}, \eta_K = \left(\sum_{i=1}^8 \eta_{iK}^2 \right)^{1/2}, \eta_{1K} = h_K \|\mathbf{f}\|_K,$$

$$\eta_{2K} = h_K^{1/2} \left(\sum_{E \in E_K^{int} \cup E_K^N} \|J_{E,n}(\mathbf{u}_h)\|_E^2 \right)^{1/2}, \eta_{3K} = h_K^{1/2} \left(\sum_{E \in E_K^C} \|\lambda_{hn} + \sigma_n(\mathbf{u}_h)\|_E^2 \right)^{1/2},$$

$$\eta_{4K} = h_K^{1/2} \left(\sum_{E \in E_K^C} \|\lambda_{ht} + \sigma_t(\mathbf{u}_h)\|_E^2 \right)^{1/2}, \eta_{5K} = \left(\sum_{E \in E_K^C} \int_E -\lambda_{hn+} u_{hn} \right)^{1/2},$$

$$\eta_{6K} = \left(\sum_{E \in E_K^C} \|\lambda_{hn-}\|_E^2 \right)^{1/2}, \eta_{7K} = \left(\sum_{E \in E_K^C} \int_E (|\lambda_{ht}| - \mu\lambda_{hn+})_- |u_{ht}| + (\lambda_{ht} u_{ht})_- \right)^{1/2},$$

$$\eta_{8K} = \left(\sum_{E \in E_K^C} \|(|\lambda_{ht}| - \mu\lambda_{hn+})_+\|_E^2 \right)^{1/2},$$

where the notations $(.)_+$ *and* $(.)_-$ *denote the positive and negative parts respectively;* $J_{E,n}(\mathbf{u}_h)$ *represents the constraint jump of* $\mathbf{u}_h$ *in the normal direction:*

$$J_{E,n}(\mathbf{u}_h) = \begin{cases} \left[\!\left[\boldsymbol{\sigma}(\mathbf{u}_h)\mathbf{n}_E \right]\!\right]_E, \forall E \in E_K^{int}, & \textit{(interior jump)} \\ \boldsymbol{\sigma}(\mathbf{u}_h)\mathbf{n}_E - \mathbf{F}, \forall E \in E_K^N & \textit{(Neumann jump)}. \end{cases}$$

Remark 3.1. *From the previous definition, we see that there are eight contributions for any local estimator* η_K*. We first mention that there are only two classical contributions (η_{1K} : equilibrium residual and η_{2K} : constraint jumps) for all the elements which do not have an edge belonging to* Γ_C*. The remaining elements on the contact area have six supplementary terms. The terms η_{3K} and η_{4K} come from the choice of a mixed finite element method, the terms η_{5K} and η_{6K} represent the nonfulfillment of the unilateral contact conditions (3) and the terms η_{7K} and η_{8K} represent the nonfulfillment of friction conditions (4).*

The main results are given next. The proofs can be found in [8].

3.2   Upper error bound

Theorem 3.1. *Let* $(\mathbf{u}, \boldsymbol{\lambda})$ *be the solution to Problem (5) such that* $\lambda_t = \mu\lambda_n\xi$*, with* $\xi \in M$*,* $\xi \in Dir_t(u_t)$ *a.e. on* Γ_C *and* $\mu < (C\|\xi\|_a)^{-1}$*. Let* $(\mathbf{u}_h, \boldsymbol{\lambda}_h)$ *be a solution to the discrete problem (6). Then*

$$\|\mathbf{u} - \mathbf{u}_h\|_{1,\Omega} + \|\boldsymbol{\lambda} - \boldsymbol{\lambda}_h\|_{-\frac{1}{2},\Gamma_C} \lesssim \eta.$$

3.3   Lower error bound

Theorem 3.2. *For all elements* K*, the following local lower error bounds hold:*

$$\begin{aligned} \eta_{1K} &\lesssim \|\mathbf{u} - \mathbf{u}_h\|_{1,K}, \\ \eta_{2K} &\lesssim \|\mathbf{u} - \mathbf{u}_h\|_{1,\omega_K}. \end{aligned}$$

Assume that $\boldsymbol{\lambda} \in (L^2(\Gamma_C))^2$*. For all elements* K *such that* $K \cap \Gamma_C \neq \emptyset$*, the following local lower error bounds hold:*

$$\eta_{iK} \lesssim \sum_{E \in E_K^C} h^{1/2}\|\boldsymbol{\lambda} - \boldsymbol{\lambda}_h\|_E + \|\mathbf{u} - \mathbf{u}_h\|_{1,K}, \ i = 3, 4,$$

$$\eta_{jK} \lesssim \sum_{E \in E_K^C} \left(\|\boldsymbol{\lambda} - \boldsymbol{\lambda}_h\|_E + \|\mathbf{u} - \mathbf{u}_h\|_E + \|\boldsymbol{\lambda} - \boldsymbol{\lambda}_h\|_E^{1/2}\|\mathbf{u}\|_E^{1/2}\|\mathbf{u} - \mathbf{u}_h\|_E^{1/2}\|\boldsymbol{\lambda}\|_E^{1/2} \right), j = 5, 7,$$

$$\eta_{lK} \lesssim \sum_{E \in E_K^C} \|\boldsymbol{\lambda} - \boldsymbol{\lambda}_h\|_E, \ l = 6, 8.$$

The study of residual estimators with a lower number of local contributions is actually under investigation in [8]. In this thesis the extension of the analysis to the XFEM method (see

[10]) is also considered. Extensions of the present work to other friction models as smoothed Coulomb friction (see [2, 3]), compliance models (see [9, 12]) or quasistatic Coulomb friction (see [1, 14]) could lead to some interesting studies.

References

[1] L.-E. Andersson. Existence results for quasistatic contact problems with coulomb friction. *Appl. Math. Optim.*, 42:169–202, 2000.

[2] M. Cocu. Existence of solutions of signorini problems with friction. *Int. J. Engrg. Sci.*, 22:567–575, 1984.

[3] G. Duvaut. Equilibre d'un solide élastique avec contact unilatéral et frottement de coulomb. *C. R. Acad. Sci. Sér. I Math.*, 290:263–265, 1980.

[4] C. Eck, J. Jarušek, and M. Krbec. *Unilateral contact problems: variational methods and existence theorems*, volume 270 of *Pure and Applied Mathematics*. CRC Press, 2005.

[5] P. Hild. Non-unique slipping in the Coulomb friction model in two-dimensional linear elasticity. *Q. Jl. Mech. Appl. Math.*, 57:225–235, 2004.

[6] P. Hild. Multiple solutions of stick and separation type in the Signorini model with Coulomb friction. *Z. Angew. Math. Mech.*, 85:673–680, 2005.

[7] P. Hild and Y. Renard. An error estimate for the Signorini problem with Coulomb friction approximated by finite elements. *SIAM J. Numer. Anal.*, 45:2012–2031, 2007.

[8] V. Lleras. Thesis. in preparation.

[9] J. A. C. Martins and J. T. Oden. Existence and uniqueness results for dynamic contact problems with nonlinear normal and friction interface laws. *Nonlinear Anal.*, 11:407–428, 1987.

[10] N. Moës, J. Dolbow, and T. Belytschko. A finite element method for cracked growth without remeshing. *Int. J. Numer. Meth. Engng.*, 46:131–150, 1999.

[11] J. Nečas, J. Jarušek, and J. Haslinger. On the solution of the variational inequality to the Signorini problem with small friction. *Bolletino U.M.I.*, 17(5):796–811, 1980.

[12] J. T. Oden and J. A. C. Martins. Models and computational methods for dynamic friction phenomena. *Comput. Methods. Appl. Mech. Engrg.*, 52:527–634, 1985.

[13] Y. Renard. A uniqueness criterion for the Signorini problem with Coulomb friction. *SIAM J. Math. Anal.*, 38:452–467, 2006.

[14] R. Rocca and M. Cocou. Existence and approximation of a solution to quasistatic Signorini problem with local friction. *Int. J. Engrg. Sci.*, 39:1233–1255, 2001.

Mathematics and the Atmospheric Sciences

Mimetic methods: from global weather forecasting to CFD

L. Bonaventura[*]   A. Abbá [†]

Abstract

Numerical methods endowed with discrete conservation properties have been proposed already in the early stages of numerical weather forecasting. More recently, a renewed attention has been devoted to these techniques and there is increasing awareness that they can lead to significant increase in accuracy also for more conventional CFD simulations. In this paper, we briefly review some of the main contributions in this area and summarize some recent results, showing how discrete vorticity preservation is indeed highly beneficial in increasing the accuracy for simulations of laminar and turbulent incompressible flow without any increase in the computational effort required.

1 Introduction

The development of numerical methods for fluid flow preserving discrete analogs of some invariants of the equations of motion (such as e.g. mass, momentum, energy, enstrophy) attracted great attention in the early phases of computational fluid dynamics. In particular, finite differencing techniques possessing such properties were developed in context of numerical models for large scale atmospheric flows, see e.g. [3], [4], [9], [15], [16]. Many of these attempts focused on simplified two dimensional systems such as the shallow water equations.

Also more recent work has devoted attention to these discrete properties, see e.g. [5], [6], [8], [10], [11], [13], [14]. The development of numerical methods with discrete conservation properties can take advantage of so called *mimetic* finite difference schemes, for which discrete analogs of continuous identies hold, such as $\nabla \times \nabla\phi = 0$, integration by parts formulae and the Helmholtz decomposition theorem. Examples of mimetic finite differences are given e.g. by [7], [12], in two and three dimensional frameworks, respectively.

The original motivation for preserving discrete invariants in atmospheric flow modelling was mainly related to the necessity of avoiding spurious trends in climate models used for very long range simulations. Furthermore, on a rotating planet a constant source of vorticity

---

[*]MOX - Politecnico di Milano. E-mail:luca.bonaventura@polimi.it.

[†]MOX - Politecnico di Milano. E-mail:antonella.abba@polimi.it.

is present, so that avoiding spurious sources of relative vorticity is assumed to lead to a more faithful reproduction of the large scale dynamics, even if relatively coarse grids are used, as it is often the case in climate models. Similar considerations apply to the case of discrete energy conservation, which in early numerical models was also sought as a guarantee of numerical stability Some evidence of the benefits of these conservative approaches is given by the consistent reproduction of energy and enstrophy spectra in long term decaying turbulence simulations, as proposed e.g. in [5], [14].

In more conventional CFD applications, the arguments supporting the use of mimetic schemes are also related to the desire of reproducing correct turbulence spectra, see e.g. the discussion in [13]. In particular, in Large Eddy Simulation (LES) approaches, spurious sources of energy and vorticity can produce undesired unphysical long term trends. Furthermore, it was suggested in [11] that apparently pathological solutions of the Euler equations may indeed be the result of spurious vorticity production, which could be avoided if vorticity preserving discretizations were employed.

In the following, we will briefly introduce some of the recent results obtained by the authors, which are presented in greater detail in [5], [1], [2], showing how discrete vorticity preservation is indeed highly beneficial in increasing the accuracy for simulations of both large scale atmospheric dynamics and laboratiry scale incompressible flow, without any increase in the computational effort required.

2 A mimetic scheme for global NWP: outline and selected results

The shallow water equations, considered as a model equation set for large scale atmospheric dynamics, are written here in vector invariant form as

$$\frac{\partial h}{\partial t} + \nabla \cdot (\mathbf{v} h) = 0, \tag{1}$$

$$\frac{\partial \mathbf{v}}{\partial t} = -\eta \mathbf{k} \times \mathbf{v} - \nabla \left[g \left(h + h^s \right) + K \right], \tag{2}$$

where $\mathbf{v}$ is the velocity vector, $\eta = \zeta + f$ is the absolute vorticity, ζ is the vertical component of the relative vorticity, h fluid depth, h^s is orographic height and K the kinetic energy per unit mass $K = |\mathbf{v}|^2/2$. The potential vorticity can be defined in this context as $q = \eta/h$, so that potential enstrophy is given by q^2.

The discretization is carried out on a Delaunay mesh whose triangular cells are denoted by i and whose edges by l. The prognostic quantities are the cell averaged free surface height h_i and the component of velocity normal to the cell edge u_l. The full velocity vector is reconstructed by means of a Raviart Thomas interpolation of degree 0, while div and δ_ν denote

appr.The vorticity is computed naturally on the dual Voronoi cell associated at each cell vertex. All the details of the discretization can be found in [5]. A semi-implicit discretization using e.g. three time level leapfrog time-stepping of these equations coupled to this mimetic discretization is given by

$$u_l^{n+1} = u_l^{n-1} + 2\Delta t \bar{\eta}_l^n v_l^n - \Delta t \left[\delta_\nu (g(h^n + h^s) + K^n) \right]_l. \tag{3}$$

$$h_i^{n+1} = h_i^{n-1} - \Delta t \ \mathrm{div}(\bar{h}^n u^{n+1})_i - \Delta t \ \mathrm{div}(\bar{h}^n u^{n-1})_i. \tag{4}$$

The advantages of the resulting method have been discussed extensively in in [5]; here, we show in figure 1 an example of the typical results in a shallow water test with an isolated mountain placed in an initially zonal flow. It can be seen how the model, in spite of the relatively coarse mesh used for the computation, is able to reproduce correctly the vorticity pattern.
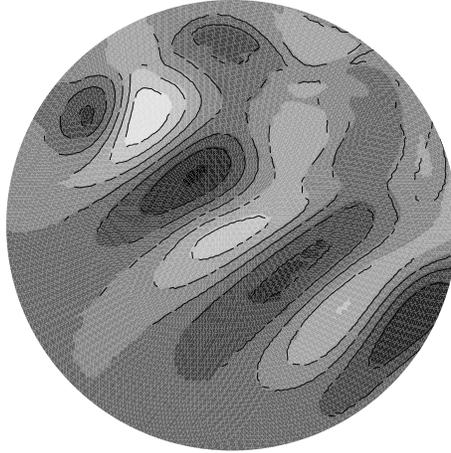

Figure 1: Relative vorticity field in test with isolated mountain: results of vorticity preserving model (shaded contours) and of reference spectral model (black lines).

3 A mimetic scheme for the Navier Stokes equations: selected results

The Navier-Stokes equations for a constant density, incompressible fluid can be formulated as

$$\frac{\partial \mathbf{u}}{\partial t} = -\boldsymbol{\omega} \times \mathbf{u} - \nabla(p + K) - \mu \nabla \times \boldsymbol{\omega} \tag{5}$$

$$\nabla \cdot \mathbf{u} = 0 \tag{6}$$

Taking the curl of the momentum equation, an evolution equation for vorticity can also be obtained

$$\frac{\partial \boldsymbol{\omega}}{\partial t} = -\nabla \times \left[\boldsymbol{\omega} \times \mathbf{u}\right] + \mu \Delta \boldsymbol{\omega}. \tag{7}$$

A mimetic discretization can be defined on the staggered cartesian meshes commonly used in the Marker and Cell (MAC) approach as

$$
\begin{aligned}
\frac{\partial}{\partial t} u_{i+\frac{1}{2},j,k} &= -\bar{\omega}^{y}_{i+\frac{1}{2},j,k} \bar{w}_{i+\frac{1}{2},j,k} + \bar{\omega}^{z}_{i+\frac{1}{2},j,k} \bar{v}_{i+\frac{1}{2},j,k} - \delta_x(p + \bar{K})_{i+\frac{1}{2},j,k} \\
&\quad + \mu\left[\delta_z(\omega^y)_{i+\frac{1}{2},j,k} - \delta_y(\omega^z)_{i+\frac{1}{2},j,k}\right]
\end{aligned} \tag{8}
$$

$$
\begin{aligned}
\frac{\partial}{\partial t} v_{i,j+\frac{1}{2},k} &= -\bar{\omega}^{z}_{i,j+\frac{1}{2},k} \bar{u}_{i,j+\frac{1}{2},k} + \bar{\omega}^{x}_{i,j+\frac{1}{2},k} \bar{w}_{i,j+\frac{1}{2},k} - \delta_y(p + \bar{K})_{i,j+\frac{1}{2},k} \\
&\quad + \mu\left[\delta_x(\omega^z)_{i,j+\frac{1}{2},k} - \delta_z(\omega^x)_{i,j+\frac{1}{2},k}\right]
\end{aligned} \tag{9}
$$

$$
\begin{aligned}
\frac{\partial}{\partial t} w_{i,j,k+\frac{1}{2}} &= -\bar{\omega}^{x}_{i,j,k+\frac{1}{2}} \bar{v}_{i,j,k+\frac{1}{2}} + \bar{\omega}^{y}_{i,j,k+\frac{1}{2}} \bar{u}_{i,j,k+\frac{1}{2}} - \delta_z(p + \bar{K})_{i,j,k+\frac{1}{2}} \\
&\quad + \mu\left[\delta_y(\omega^x)_{i,j,k+\frac{1}{2}} - \delta_x(\omega^y)_{i,j,k+\frac{1}{2}}\right]
\end{aligned} \tag{10}
$$

$$\operatorname{div}(u, v, w)_{i,j,k} = 0. \tag{11}$$

Here, the definitions of the discrete operators are, in general, those customary of the MAC discretization. Only the definition of the vorticity fluxes ω^x, ω^y and ω^z is less standard and is described in full detail in [1], [2]. The main difference between the present approach and the discretization proposed in [13] lies in the location of the velocity and vorticity points. In the present discretization, the tangential velocity components at the cell edge (for example, the terms $\bar{w}_{i+\frac{1}{2},j,k}, \bar{v}_{i+\frac{1}{2},j,k}$ in equation 8) are averaged separately at the edge midpoint, while in [13] the whole momentum advection term is computed at the cell vertex (i.e., at the discrete location where vorticity is defined).

Among the large amount of results obtained, we show here a benchmark test case consisting in the simulation of the flow around a square cylinder. At relatively low Reynolds numbers a laminar flow regime is guaranteed, while strong vorticity production takes place at the obstacle corners, along with vortex shedding in the lee. The difference between a standard and a vorticity preserving discretization is shown in figure 2, highlighting the higher vorticity values at the cylinder corner. The correctness of the latter is assured by the fact that, when discrete vorticity preservation is guaranteed, the transverse velocity component is zero to machine accuracy, while all the other schemes we have tested at the same resolution yield spurious transverse velocities of the same order of magnitude of the main flow.
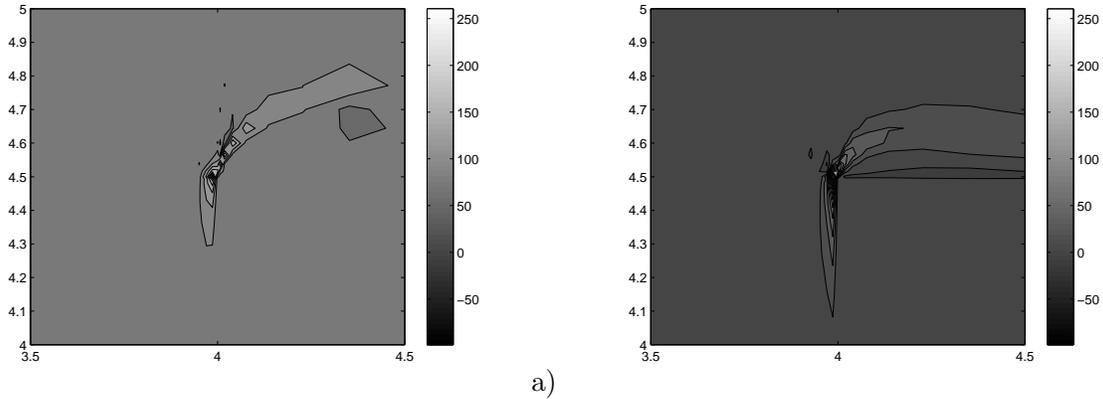
Figure 2: Vorticity field around the square cylinder at $Re = 500$, detail of cylinder corner simulated with a) centered finite differences, b) vorticity preserving method.

4 Conclusions and ongoing work

We have briefly reviewed some of the main contributions in the area of mimetic finite difference schemes for large scale weather forecasting and incompressible flow simulation. We have summarized some recent results that show the improvements in local accuracy that can result from the use of discrete vorticity preserving methods, without any increase in the computational cost. Ongoing research work is now aimed at extending the previously described method for the incompressible Navier Stokes equations to arbitrary unstructured meshes in two and three dimensions, as well as to demonstrate the importance of these techniques in improving the results of Large Eddy Simulations.

References

[1] A. Abbà and L. Bonaventura. A mimetic finite difference discretization for the incompressible Navier-Stokes equations. Technical Report Internal report n., MOX - Politecnico di Milano, 2007.

[2] A. Abbà and L. Bonaventura. A mimetic finite difference discretization for the incompressible Navier-Stokes equations. to appear. *International Journal of Numerical Methods in Fluids*, 2008.

[3] A. Arakawa and V. Lamb. A potential enstrophy and energy conserving scheme for the shallow water equations. *Monthly Weather Review*, 109:18–136, 1981.

[4] A. Arakawa and M. Suarez. Vertical differencing of the primitive equations in sigma coordinates. *Monthly Weather Review*, 111:34–45, 1983.

[5] L. Bonaventura and T. Ringler. Analysis of discrete shallow water models on geodesic Delaunay grids with C-type staggering. *Monthly Weather Review*, 133:2351–2373, 2005.

[6] J.E. Hicken, F.E. Ham, J. Militzer, and M. Koksal. A shift transformation for fully conservative methods: turbulence simulation on complex, unstructured grids. *Journal of Computational Physics*, 208:704–734, 2005.

[7] J. Hyman and M. Shashkov. The orthogonal decomposition theorems for mimetic finite difference methods. *SIAM Journal of Numerical Analysis*, 36:788–818, 1999.

[8] S.J. Lin and R.B. Rood. An explicit flux-form semi-Lagrangian shallow water model on the sphere. *Quarterly Journal of the Royal Meteorological Society*, 123:2477–2498, 1997.

[9] F. Mesinger. Horizontal advection schemes on a staggered grid: an enstrophy and energy conserving model. *Monthly Weather Review*, 109:467–478, 1981.

[10] Y. Morinishi, T.S.Lund, O.V. Vasilyev, and P. Moin. Fully conservative higher order finite difference schemes for incompressible flow. *Journal of Computational Physics*, 143:90–124, 1998.

[11] K.W. Morton and P.L. Roe. Vorticity preserving Lax-Wendroff type schemes for the system wave equation. *SIAM Journal of Scientific Computing*, 23:170–192, 2001.

[12] R.A. Nicolaides. Direct discretization of planar div-curl problems. *SIAM Journal of Numerical Analysis*, 29:32–56, 1992.

[13] B. Perot. Conservation properties of unstructured staggered mesh schemes. *Journal of Computational Physics*, 159:58–89, 2000.

[14] T.D. Ringler and D.A. Randall. A potential enstrophy and energy conserving numerical scheme for solution of the shallow-water equations a geodesic grid. *Monthly Weather Review*, 130:1397–1410, July 2002.

[15] R. Sadourny. The dynamics of finite difference models of the shallow water equations. *Journal of the Atmospheric Sciences*, 32:680–689, 1975.

[16] G. J. Tripoli. A nonhydrostatic mesoscale model designed to simulate scale interaction. *Monthly Weather Review*, 120:1343–1359, 1992.

Forecasting: its about dynamics, its not about statistics

Kevin Judd*

Abstract

In 1963 Edward Lorenz published a paper that changed the way scientists think about the prediction of geophysical systems. Two years earlier, Rudolf Kalman had published a paper that changed the way engineers thought about prediction of electronic and mechanical systems. In recent years geophysicists have become increasingly interested in Kalman filters, where as engineers have become increasingly interested in chaos. I will argue that prediction of the weather has more to do with nonlinear dynamics than it has to do with linear statistics. A position that I think both Lorenz and Kalman would likely agree with. I will attempt to show why attractors, shadowing trajectories, nonlinear filters, optimal control, and other concepts from nonlinear dynamics are important in weather prediction, by first illustrating the ideas with Lorenz's 1963 example of a chaotic system, then confirming these results by experiments with an operational weather forecasting model.

Forecasting is a process whereby a series of *observations* of *reality* is transformed into a single best-guess *state*, or an ensemble of states, of a *model*, from these states forecasts are obtained by solving the model equations. The process of obtaining model states from observations of sometimes referred to as *data assimilation*, and requires solving an *inverse problem*. The question being addressed here is: how does one best go about launching forecasts to obtain useful predictions of the future behaviour of reality, given that the observations are limited and the model is imperfect? These notes are not intended to be complete or self contained; the intention is these notes are viewed in conjunction with the seminar slides as a reminder or summary. The details of the theory and results, often just passed over here, can be found in the various referenced papers.

1   Two papers that changed forecasting

In 1963 the meteorologist Edward Lorenz had published a paper [13] that described the complex dynamics of a simple system of ordinary differential equations. This landmark paper established in minds of many the notion of chaos; in particular notions of attractors, sensitivity

---

*School of Mathematics and Statistics, The University of Western Australia.

to initial conditions, and the limits to forecasting nonlinear systems. Three years previously Rudolf Kalman had published a paper [12] concerning optimal filtering and state estimation; this too is a landmark paper. These two papers are not unrelated. In both cases the authors were looking for a better forecasting alternative to the Wiener filter, which operates in the frequency domain, and both authors adopted a state-based approach. Lorenz considered deterministic nonlinear systems; Kalman considered linear stochastic systems. When it comes to modelling and forecasting the divide between a deterministic dynamical approach of Lorenz and a linear stochastic approach of Kalman is still apparent today.

It would be difficult to say which of these papers has been more influential, but in recent years in the geophysical literature it is certain that Kalman filters have received more attention than nonlinear dynamics. Many meteorologists and geophysicists have become increasingly interested in applying Kalman filters. I will argue that this is both unfortunate and ironic, because there is more to be gained from the approach of Lorenz than there is from the approach of Kalman, and I think Kalman would agree.

2   Limits of sequential filters

Kalman understood that nonlinear systems are more complex than linear systems and would require specialised algorithms, and foresaw the day when his techniques would be superseded by better nonlinear techniques. At the time few people had much idea of the complexity of nonlinear systems or how to approach the forecasting problem for them. It was Lorenz's paper three years after Kalman's paper that revealed to many just how difficult forecasting nonlinear systems can be.

There have been many attempts to generalise the Kalman filter to nonlinear systems. Many of these variants of the extended Kalman filter use a nonlinear forecast model, but linearise the model about the current state to achieve an update of the error covariance. This approach is successful provided the errors are small relative to the nonlinearity of the model, if this is not the case, then this type of filter fails badly [2].

An essential problem of the Kalman update rule is that it requires errors to have a Gaussian distribution. The class of Gaussian distributions are invariant under linear transforms, and hence the assumption Gaussian distributions is often appropriate for linear systems. This assumption is definitely not true of nonlinear systems, because they have complex non-Gaussian distributions [8]. Another class of Kalman filter variants (often called Bayesian filters, or Particle filters) use an ensemble of states to represent the non-Gaussian distributions. Although a step forward, these methods are inferior to the techniques we advocate, because these are all *sequential filters* that fail to fully exploit the dynamics of nonlinear systems.

3   A new way forward

Some forty years after the papers of Kalman and Lorenz, a great deal of knowledge and understanding is available about nonlinear systems, although the theory is still incomplete. Nonetheless, the nonlinear techniques that Kalman envisioned are now becoming available. These new nonlinear techniques are firmly based on the properties of nonlinear deterministic systems, and Lorenz's equations, being one of the archetype nonlinear systems, provides a conceptual basis for developing these ideas. Two new tools, or concepts, are the *shadowing filter* [2] and *indistinguishable states* [8, 9], which we will discuss.

The title of this talk encapsulates its essence: forecasting is not about *statistics*, its about *dynamics*. This is not saying that in forecasting nonlinear systems statistics has no role, only that dynamics has a more important role. Dynamics plays a role in the Kalman filter, but a minor one compared to the importance of randomness. In nonlinear systems, dynamics provides structure and memory; things that are largely absent from linear stochastic processes. An alternative title might say "Its not about *randomness*, its about *geometry*." Once again this is not exclusive, rather it is a change of emphasis. Randomness does play a role in the nonlinear theory, and geometry does play a role in statistics. In nonlinear systems though, the geometric structures are rich, complex and significant.

There is another key element to forecasting that we have so far ignored, and it is something that is too often ignored, it is the problem of *model error.* How does one make a forecast knowing that the forecasting model is wrong and there is no opportunity to correct it? In linear stochastic systems model error is not an important issue, often model error is subsumed into the stochastic element. In nonlinear systems model error is not so easily swept away, and this fact is not widely recognised.

4   Lessons from Lorenz

Lorenz's equations and its butterfly shaped attractor reveal many of important properties of nonlinear systems. Everyone knows some of these properties, many are aware of other properties, but few appreciate their full significance.

One of most important properties that affects forecasting is *sensitivity to initial conditions*. Almost any two states close-by on the attractor will move apart at (typically, initially) an exponential rate. The consequences for forecasting nonlinear systems is immediately apparent: any small error identifying the "true" state leads to exponential growth of the forecast error. Many accept as a truism that forecasting errors *always* grow exponentially; this is not necessarily true. It assumes the forecast model is *perfect*, that the states are *on the attractor*, and in a region with *unstable flow* — even on the attractor there are regions where all forecast errors shrink [15].

Another important property of Lorenz's equations is that there is an attractor. Almost all

initial states quickly move onto, then remain on, the *attractor*, or more correctly an *attracting manifold*. (We will just say *attractor* in the following.) The attractor is the *physically realizable* states of the model. Of all the possible states of a system, one never expects to find a nonlinear system anywhere other than on its attractor. If we had a *perfect model* of reality, then it never makes any sense to start a forecast from anywhere other than on the attractor. Anywhere else is not a sensible state of the model, and will just move onto the attractor in any case.

If you are familiar with these ideas and are thinking ahead you might wonder whether forecasts should always start on a model's attractor when the model is *imperfect*, because the attractor might be in the wrong place, or have the wrong shape.

5   Shadowing filters and indistinguishable states

For the time being, let us stay with the *perfect model* fiction and ask: How does one obtain from observations states that are on the model's attractor? Assimilation of observations (by standard methods like optimal interpolation, variational assimilation, Kalman filters, Bayesian filters) typically result in states that are not on the model's attractor. We require some method to *project* observations onto the model attractor.

The *shadowing filter* can do this [2, 8]. The shadowing filter attempts to find a trajectory in close proximity to the observations. It is generally effective at doing so. (It is not guaranteed to be 100% effective, but it is rare for it to fail, and if it does, we can tell.) The shadowing filter achieves its goal by exploiting information in the observations revealed by model's dynamics. Statistics is secondary to the dynamics.

The important distinction of the shadowing filter compared to other techniques, is that it views data assimilation from the point of view of a dynamical system. The shadowing filter attempts to find a *trajectory* in close proximity to a whole *sequence of observations*; unlike sequential filters (Kalman filter and other Bayesian filters) that depend only on the most recent best-guess state and the present observation, which implies the process has no longer-term memory. The shadowing filter exploits dynamics to extract useful information from deeper into the past. Curiously, the shadowing filter is not concerned about obtaining a *maximum likelihood* trajectory. Its turns out that the statistical notion of maximum likelihood can be misleading when applied to nonlinear dynamical systems [3], and well as notions like Bayesian inference [1].

For nonlinear dynamical systems it appears that the appropriate concept (as an alternative to maximum likelihood and so forth) is *indistinguishable states* [8, 9], and there is a very close connection between these and shadowing trajectories. Essentially, indistinguishable states are the set of shadowing trajectories consistent with the any typical set of observations. Indistinguishable states have many useful and interesting properties, and in addition, have the

advantage that the concept generalises to the imperfect model scenario. The most important role of indistinguishable states is their use in constructing ensemble forecasts.

6   The geometry of model error

Model error plays a much more significant role in forecasting of nonlinear systems than it does in linear systems. Consequently, the techniques we describe are probably different from what both Kalman and Lorenz, and others, have envisioned.

When forecast models are formulation as stochastic models, then model error is usually just swept up into a stochastic component. Under some circumstances this is legitimate, or even the best course of action [9, 11], but often it is not [7].

The lessons of Lorenz tell us that the forecast model should have an attracting manifold, but if the model is imperfect, this attractor can be in the wrong location, can have the wrong shape, and can have the wrong direction of flow on the attractor. Recognising this enables us to resolve errors into different components, in particular, *projection errors* related to the attractor being the wrong shape or in the wrong location, and *direction errors* related to flow on the attractor being in the wrong direction. These different error components have different characteristics and different effects on forecast errors. Furthermore, the shadowing filter helps us to resolve these components. This efficiently provides useful information about the nature of model errors that can be used to improve the model. It also provides useful information about how to construct ensemble forecasts that compensate for the revealed model errors when the model cannot be immediately corrected.

7   Now and the future

Presently the methods advocated for providing forecasts of nonlinear systems have reached a stage where they can be usefully employed. We have been testing them for sometime in simple chaotic dynamical systems [8, 9, 2] and simple atmospheric models [10]. More recently we have applied the methods to an operational weather forecasting model at reduced resolution using real atmospheric observations [6, 7]. Experiments to date have established that the methods advocated are practical and useful.

However, considerable more work is still necessary. On-going theoretical analysis has established many new properties of nonlinear dynamical systems that shed light on the problems of forecasting with imperfect models and properties of proposed algorithms. These range from deriving convergence rates and robustness of algorithms, to the underlying structure of attracting manifolds and their effect on algorithms [4, 5, 10, 14].

References

[1] K. Judd. Chaotic-time-series reconstruction by the Bayesian paradigm: Right results by wrong methods. *Physics Review E*, 67:026212, 2003.

[2] K. Judd. Nonlinear state estimation, indistinguishable states and the extended Kalman filter. *Physica D*, 183:273–281, 2003.

[3] K. Judd. Failure of maximum likelihood methods for chaotic dynamical systems. *Physical Review E*, 75:036210, 2007.

[4] K. Judd. Forecasting with imperfect models, dynamically constrained inverse problems, and geometrically modified gradient descent. *Physica D*, *In press*, 2007.

[5] K. Judd. Shadowing pseudo-orbits and gradient descent noise reduction. *Journal of Nonlinear Science*, *In press*, 2007.

[6] K. Judd, C.A. Reynolds, and T. Rosmond. Toward shadowing in operational weather prediction. Technical Report NRL/MR/7530-04-18, Naval Research Laboratory, Monterey, CA, 93943-5502 USA., 2004.

[7] K. Judd, C.A. Reynolds, T.E. Rosmond, and L.A. Smith. The geometry of model error. *Journal of Atmospheric Science*, In press, 2007.

[8] K. Judd and L.A. Smith. Indistinguishable states I : perfect model scenario. *Physica D*, 151:125–141, 2001.

[9] K. Judd and L.A. Smith. Indistinguishable states II : imperfect model scenarios. *Physica D*, 196:224–242, 2004.

[10] K. Judd, L.A. Smith, and A. Weisheimer. Gradient free descent: shadowing and state estimation with limited derivative information. *Physica D*, 190:153–166, 2004.

[11] K. Judd, L.A. Smith, and A. Weisheimer. How good is an ensemble at capturing truth. *Q. J. Roy. Met. Soc.*, 2005.

[12] R.E. Kalman. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering (Series D)*, 82:35–45, 1960.

[13] E. N. Lorenz. Deterministic nonperiodic flow. *J. Atmos. Sci.*, 20:130–141, 1963.

[14] D. Ridout and K. Judd. Convergence properties of gradient descent noise reduction. *Physica D*, 165:27–48, 2001.

[15] L.A. Smith, C. Zimmerman, and K. Fraedrich. Uncertainty dynamics and predictability in chaotic systems. *Quarterly Journal of the Royal Meteorological Society*, 125:2855–2886, 1999.

Analytical treatment of critical layer filtering of mountain waves using a WKB approximation

Miguel A. C. Teixeira*   Pedro M. A. Miranda†

Abstract

The present study addresses the momentum flux profiles associated with internal gravity waves generated by mountains existing in directionally sheared flows with Richardson numbers of order one. It is shown that this problem may be treated using a WKB approximation, but the solutions must be extended to third order in the small perturbation parameter for the shear to have a correct impact on the surface value of the drag and also on wave absorption at critical levels. This approach enables the derivation of expressions for the normalized momentum flux and momentum flux divergence that are correct to second order, extending previous results. These calculations, which are valid for generic wind profiles as long as those are compatible with a WKB approximation, are made possible by the use of contour integration techniques. In particular, because the momentum flux only varies in the vertical due to critical levels (where the wave equation is singular), its variation only depends on local features of the flow at these levels.

1   Introduction

For the purpose of parametrizing the effects of internal gravity waves generated by mountains in numerical models, it is necessary to understand the dynamics of these waves. The direct impact of mountain waves on the atmospheric circulation is due to the deposition of momentum associated with the waves at critical levels or other levels where the waves break. The first process is treatable analytically within the framework of linear theory, but existing studies have considered that the waves are totally absorbed at the critical levels, since they focus on the case of high Richardson numbers, Ri. When Ri is lower, the surface drag exerted on topography may be appreciably modified, as shown previously by the authors [2], and the absorption, or rather filtering process of the associated momentum, is also altered, since critical levels are no longer perfect absorbers [3].

The present study aims to provide expressions for the momentum flux and momentum flux divergence profiles for mountain waves in sheared flows at relatively low Ri, so that existing gravity wave drag parametrizations may be improved, taking into account these effects.

2   Theory

The flow is assumed to be inviscid, steady and hydrostatic. If the equations of motion linearized with respect to an incoming reference flow, that depends only on height z, are combined, a single equation for the vertical velocity perturbation results. The Fourier transform of the vertical velocity perturbation $\hat{w}$ satisfies the Taylor-Goldstein equation:

$$\hat{w}'' + \left[\frac{N^2(k^2 + l^2)}{(Uk + Vl)^2} - \frac{U''k + V''l}{Uk + Vl} \right] \hat{w} = 0, \tag{1}$$

where N and (U, V) are the Brunt-Väisälä frequency and wind velocity of the incoming flow, (k, l) is the horizontal wavenumber vector of the waves and the primes denote differentiation with respect to z. This equation is solved subject to the boundary conditions that the flow must be tangential to the terrain at the surface

$$\hat{w}(z = 0) = i(U_0 k + V_0 l)\hat{h}, \tag{2}$$

where $(U_0, V_0) = (U, V)(z = 0)$ and $\hat{h}$ is the Fourier transform of the terrain elevation, and that the wave energy radiates upwards at infinity. A solution of the form

$$\hat{w} = \hat{w}(z = 0) \exp \left[i \int_0^z m(\varepsilon\zeta)d\zeta \right] \tag{3}$$

is adopted, where $m = m_0 + \varepsilon m_1 + \varepsilon^2 m_2 + \varepsilon^3 m_3$, the vertical wavenumber, is expanded in a power series of a small parameter ε, consistent with a 3rd-order WKB approximation. The introduction of (3) into (1) and use of (2) and the radiation boundary condition enables a complete solution of the problem to be found, yielding equations for m_0, m_1, m_2 and m_3 in terms of the incoming flow parameters.

The aim here is to obtain expressions for the momentum flux, which may be expressed as an integral in Fourier space as

$$M_x = -4\pi^2 \rho_0 \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \hat{u}^* \hat{w} \, dk dl \tag{4}$$

for the x component, where ρ_0 is a reference constant density, $\hat{u}$ is the Fourier transform of the velocity perturbation along x and the asterisk denotes complex conjugate. A similar expression is valid for M_y if $\hat{u}$ is replaced by $\hat{v}$. It should be noted that $\hat{u}$ and $\hat{w}$ may be related through

$$\hat{u} = \frac{i}{k^2 + l^2} \left(ikm - l\frac{V'k - U'l}{Uk + Vl} \right) \hat{w}. \tag{5}$$

Inserting the complex conjugate of (5) into (4), using (3) and also (2), it is possible to express the vertical momentum flux along x as

$$M_x = 4\pi^2\rho_0 \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \frac{k}{k^2+l^2}\mathrm{Re}(m)(U_0 k + V_0 l)^2|\hat{h}|^2 \mathrm{e}^{-2\mathrm{Im}(s)}\,dkdl, \tag{6}$$

where $s = \int_0^z m(\varepsilon\zeta)\,d\zeta$, Re denotes real part and Im denotes imaginary part. The momentum flux along y, M_y, only has the difference that the k appearing in the numerator of (6) is replaced by an l.

The remaining task is to give explicit expressions for $\mathrm{Re}(m)$ and $\mathrm{Im}(s)$. From the solutions previously found for m_0, m_1, m_2 and m_3, it is easy to show that

$$\mathrm{Re}(m) = \frac{N(k^2+l^2)^{1/2}}{Uk+Vl}\left[1 - \frac{1}{8}\frac{(U'k+V'l)^2}{N^2(k^2+l^2)} - \frac{1}{4}\frac{(Uk+Vl)(U''k+V''l)}{N^2(k^2+l^2)}\right]. \tag{7}$$

The calculation of s is a bit more complicated. If this quantity is expanded in a power series of ε, like m itself, it takes the form $s = s_{-1} + s_0 + s_1 + s_2$, where $s_{i-1} = \varepsilon^i \int_0^z m_i(\varepsilon\zeta)d\zeta$ and $i = 0, 1, 2, 3$. It can be shown that s_0 and s_2 may be integrated analytically, yielding

$$\mathrm{Im}(s_0) = -\frac{1}{2}\log\left|\frac{Uk+Vl}{U_0 k+V_0 l}\right|,$$

$$\begin{aligned}
\mathrm{Im}(s_2) = {} & \frac{1}{16}\frac{(U_0'k+V_0'l)^2}{N^2(k^2+l^2)} + \frac{1}{8}\frac{(U_0 k+V_0 l)(U_0''k+V_0''l)}{N^2(k^2+l^2)} \\
& - \frac{1}{16}\frac{(U'k+V'l)^2}{N^2(k^2+l^2)} - \frac{1}{8}\frac{(Uk+Vl)(U''k+V''l)}{N^2(k^2+l^2)}
\end{aligned} \tag{8}$$

(m_1 and m_3 are pure imaginary). However, since m_0 and m_2 are pure real (and could not be integrated analytically anyway), the imaginary part of their integrals s_{-1} and s_1 only comes from the singularities of these integrals. These singularities correspond to the critical levels of the flow, where $Uk + Vl = 0$. Using the residue theorem to evaluate the integrals at these singularities, it can be shown that

$$\begin{aligned}
\mathrm{Im}(s_{-1}) &= \pi\frac{N(k^2+l^2)^{1/2}}{|U_c'k+V_c'l|}H(z-z_c), \\
\mathrm{Im}(s_1) &= -\frac{\pi}{8}\frac{|U_c'k+V_c'l|}{N(k^2+l^2)^{1/2}}H(z-z_c),
\end{aligned} \tag{9}$$

where $H(z)$ is the Heaviside step function and $(U_c', V_c') = (U', V')(z = z_c)$ with z_c denoting the critical level. Obviously, these terms are only non-zero for heights $z > z_c$. Note that, in general, for 3D flow with directional shear, z_c and all functions of it (e.g. (U_c', V_c')) depend on the horizontal wavenumber (k, l).

3 Results and discussion

With the model described above, it is possible to obtain expressions for the momentum flux for generic wind profiles at relatively low Ri. If flow over a circular mountain is additionally

assumed, these expressions, when normalized by the surface drag for a constant wind, become also independent of the detailed shape of the orography. The use of a circular mountain is also convenient because it excites internal waves with horizontal wavenumbers spanning all possible directions, illustrating better how they are affected by critical levels. The normalized momentum flux along x is most easily expressed adopting polar coordinates, being given by

$$M'_x = \frac{2}{\pi} \left(\int_{\psi-\pi/2}^{\psi_0+\pi/2} I_1 d\theta - \int_{\psi_0+\pi/2}^{\psi+\pi/2} I_1 \cdot I_2 d\theta \right), \tag{10}$$

where

$$\begin{aligned} I_1 &= \cos\theta \cos(\theta - \psi_0)[1 - S(\theta, z)] \exp[S(\theta, z) - S(\theta, z = 0)], \\ I_2 &= \exp[-2\pi C(\theta)], \end{aligned} \tag{11}$$

and ψ_0 and ψ are, respectively, the lowest and highest angle spanned by the wind profile (it is assumed that the wind direction varies monotonically with height). $S(\theta, z)$ and $C(\theta)$ are defined as

$$\begin{aligned} S &= \frac{1}{8}\frac{(U'\cos\theta + V'\sin\theta)^2}{N^2} + \frac{1}{4}\frac{(U\cos\theta + V\sin\theta)(U''\cos\theta + V''\sin\theta)}{N^2}, \\ C &= = \frac{N}{|U'_c\cos\theta + V'_c\sin\theta|}\left[1 - \frac{1}{8}\frac{(U'_c\cos\theta + V'_c\sin\theta)^2}{N^2}\right]. \end{aligned} \tag{12}$$

To obtain an expression for M'_y, it suffices to multiply I_1 by $\tan\theta$ in (11). The momentum flux divergence corresponding to (10) can be obtained by differentiating this equation with respect to z, which yields

$$\frac{dM'_x}{dz} = -\frac{2}{\pi}\psi'\sin\psi\sin(\psi - \psi_0)[1 - S_\psi(z)]\exp[S_\psi(z) - S_\psi(z = 0)]\{1 + \exp[-2\pi C_\psi(z)]\}, \tag{13}$$

where

$$\begin{aligned} S_\psi(z) &= \frac{1}{8}\frac{(U'\sin\psi - V'\cos\psi)^2}{N^2} + \frac{1}{4}\frac{(U\sin\psi - V\cos\psi)(U''\sin\psi - V''\cos\psi)}{N^2}, \\ C_\psi(z) &= \frac{N}{|U'_c\sin\psi - V'_c\cos\psi|}\left[1 - \frac{1}{8}\frac{(U'_c\sin\psi - V'_c\cos\psi)^2}{N^2}\right]. \end{aligned} \tag{14}$$

An equivalent expression for dM'_y/dz would result if (13) was multiplied by $\cot\psi$.

3.1 Linear wind profile

In order to test the accuracy of the derived expressions, numerical simulations are carried out using a nonlinear and non-hydrostatic numerical model (called NH3D), but which is run here in approximately linear and hydrostatic conditions. Two wind profiles are considered. The first of them has constant shear and is given by:

$$U = U_0 - \alpha z, \quad V = U_0, \tag{15}$$
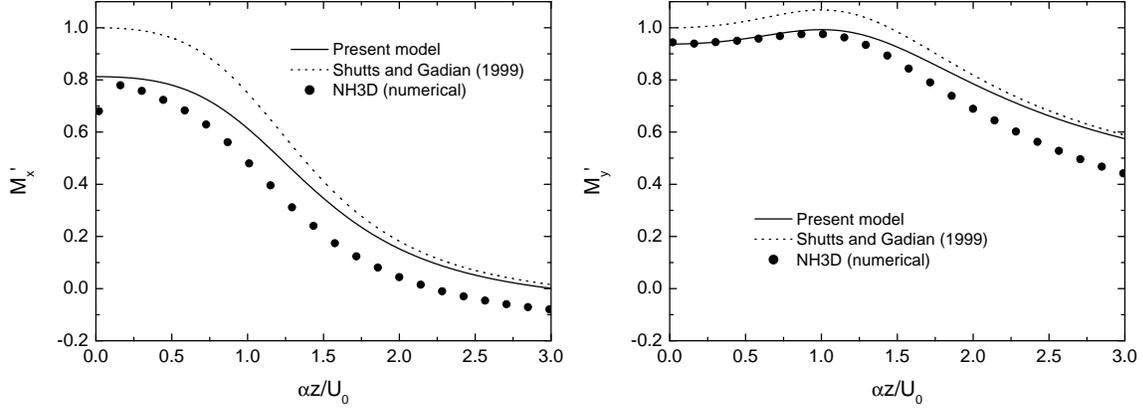
Figure 1: Normalized momentum flux (x and y components) as a function of normalized height for linear wind profile (15) with $Ri = 0.5$. See legend for details.

where $U_0 > 0$ and $\alpha > 0$ are constants. In this case,

$$I_1 = \cos\theta\cos(\theta - \pi/4)\left(1 - \frac{1}{8Ri}\cos^2\theta\right), \quad I_2 = \exp\left[-2\pi\frac{Ri^{1/2}}{|\cos\theta|}\left(1 - \frac{1}{8Ri}\cos^2\theta\right)\right], \quad (16)$$

and

$$S_\psi = \frac{1}{8Ri}\sin^2\psi, \quad C_\psi = \frac{Ri^{1/2}}{|\sin\psi|}\left(1 - \frac{1}{8Ri}\sin^2\psi\right). \quad (17)$$

Fig. 1 shows the variation of the two components of the vertical momentum flux with height. The solid line represents the model developed here, and the symbols are the numerical results, both for $Ri = 0.5$. The dashed line corresponds to the theory of [1], developed for infinite Ri. It can be seen that the present model, although overestimating slightly the numerical results, clearly improves the prediction of the theory assuming $Ri = \infty$. The faster decay of the numerical results with height may be due to spurious dissipation by the numerical scheme.

3.2 Wind that turns with height

Consider now a wind profile where the velocity turns with height at a constant rate, while maintaining its magnitude, namely:

$$U = U_0\cos(\beta z), \quad V = U_0\sin(\beta z), \quad (18)$$

where $\beta > 0$ is a constant. Then, it can be shown that

$$
\begin{aligned}
I_1 &= \cos^2\theta\left\{1 + \frac{1}{8Ri}\left[2\cos^2(\theta - \beta z) - \sin^2(\theta - \beta z)\right]\right\} \\
&\quad \times \exp\left\{\frac{1}{8Ri}\left[\sin^2(\theta - \beta z) - 2\cos^2(\theta - \beta z) - \sin^2\theta + 2\cos^2\theta\right]\right\}, \\
I_2 &= \exp\left[-2\pi Ri^{1/2}\left(1 - \frac{1}{8Ri}\right)\right],
\end{aligned}
\quad (19)
$$

133
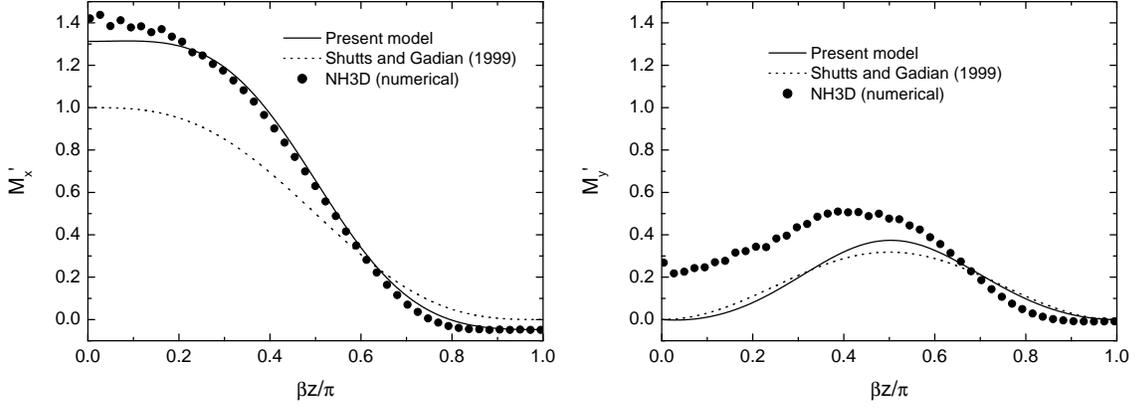
Figure 2: Same as Fig. 1, but for a wind that turns with height.

and

$$S_\psi = \frac{1}{8Ri}, \quad C_\psi = Ri^{1/2}\left(1 - \frac{1}{8Ri}\right). \tag{20}$$

Fig. 2 shows the momentum flux profiles for this case. The symbols are as in Fig. 1. Although, M_x' is slightly underestimated near the surface, and the prediction of M_y' is not particularly good, again the present model improves the results of [1], and the slope of the M_x' profile is in very good agreement with the numerical results at intermediate levels. This is particularly relevant because this slope corresponds to the force exerted by the mountain on the atmosphere (a key quantity in drag parametrization schemes).

References

[1] G. J. Shutts and A. Gadian. Numerical simulations of orographic gravity waves in flows which back with height. *Quart. J. Roy. Meteor. Soc.*, 125:2743–2765, 1999.

[2] M. A. C. Teixeira and P. M. A. Miranda. A linear model of gravity wave drag for hydrostatic sheared flow over elliptical mountains. *Quart. J. Roy. Meteor. Soc.*, 132:2439–2458, 2006.

[3] M. A. C. Teixeira, P. M. A. Miranda, and J. L. Argain. Mountain waves in two-layer sheared flows: critical level effects, wave reflection and drag enhancement. *J. Atmos. Sci., in press*, 2007.

Modelling of Industrial Processes

The fixed mesh ALE approach: basic idea and two applications

Ramon Codina[*]   Herbert Coppola-Owen[†]   Guillaume Houzeaux[‡]

Abstract

In this work we describe what we call Fixed-Mesh ALE (FM-ALE) method. The idea is to use the ALE framework but, instead of moving the mesh, to project the results onto a fixed one. This idea is explained in some detail and two applications to problems in fluid mechanics are presented.

1   Introduction

In many coupled problems of practical interest the domain of at least one of the problems evolves in time. The Arbitrary Eulerian Lagrangian (ALE) approach is a tool very often employed to cope with this domain motion. In this work we aim at describing a particular version of the ALE formulation that can be used in different coupled problems, and which we will apply to two problems in fluid mechanics.

In the classical ALE approach to solve CFD problems the mesh in which the computational domain is discretized is deformed. This is done according to a prescribed motion of part of its boundary, which is transmitted to the interior nodes in a way as smooth as possible so as to avoid mesh distortion. In this work we present an ALE-type strategy with a different motivation. Instead of assuming that the computational domain is defined by the mesh boundary, we assume that there is a level set function that defines where the flow takes place. When this level set function moves, the flow domain changes, and that must be taken into account at the moment of writing the conservation equations that govern the flow, which need to be cast in the ALE format. However, our purpose here is to explain how to use always a background *fixed* mesh. That requires a virtual motion of the mesh nodes followed by a projection of the new node positions onto the fixed mesh.

We have applied the fixed-mesh ALE idea to two coupled problems of practical interest. One is the simulation of lost foam casting. In this case the flow is coupled to the heat equation

---

[*]Universitat Politècnica de Catalunya. E-mail:`ramon.codina@upc.edu`.

[†]Universitat Politècnica de Catalunya. E-mail:`howen@cimne.upc.edu`.

[‡]Barcelona Supercomputing Center. E-mail:`guillaume.houzeaux@bsc.es`.

because of the interface evolution, which is governed by the advance of the burning front of the molten metal used in the casting. Therefore, the boundary velocity is given and the pressure in the fluid is unknown. The second problem considered is a classical free surface problem in which, once more, the free surface position is modeled by a level set function. In this case, the velocity on the free surface is unknown, but the pressure can be prescribed to zero (or to the atmospheric pressure).

The numerical ingredients, which we will not describe here, consist of a stabilized finite element method (FEM) to solve the ALE flow equations, a stabilized FEM to advance the level set function, finite difference time integration schemes and an efficient projection scheme between different meshes on the same domain.

2 FM-ALE formulation

The particularity of our ALE description is that it is used on a fixed mesh. It is called FM-ALE and has been introduced in the paper on lost foam casting [2] and further extended in [1] in the context of free surface flows.

2.1 Problem statement

Suppose we wish to solve a flow problem in a domain Ω^0, but that the fluid at time t occupies a subdomain $\Omega(t) \subset \Omega^0$. Suppose also that the boundary of $\Omega(t)$ is defined by part of $\partial\Omega^0$ and a level set function $\psi(\boldsymbol{x}, t)$, so that $\Gamma_{\text{free}}(t) = \partial\Omega(t) \setminus \partial\Omega^0 \cap \partial\Omega(t) = \{\boldsymbol{x} \in \Omega \mid \psi = 0\}$, with $\boldsymbol{x} \in \Omega$.

The incompressible Navier-Stokes formulated in $\Omega(t)$, accounting also for the motion of this domain, can be written as follows: find a velocity $\boldsymbol{u}$ and a pressure p such that

$$\rho[\partial_t \boldsymbol{u} + (\boldsymbol{u} - \boldsymbol{u}_m) \cdot \nabla \boldsymbol{u}] - \nabla \cdot [2\mu \nabla^S(\boldsymbol{u})] + \nabla p = \boldsymbol{f}$$
$$\nabla \cdot \boldsymbol{u} = 0$$

where $\boldsymbol{u}_m$ is the motion of the ALE system used to write the conservation of momentum and mass, ρ is the fluid density, μ the viscosity and $\boldsymbol{f}$ the vector of body forces.

In the applications we have considered, the boundary conditions on $\Gamma_{\text{free}}(t)$ can be of two different types: a) Free surface flows: p (or the stress) given, $\boldsymbol{u}$ unknown on Γ_{free}; b) Lost foam casting: $\boldsymbol{u}$ given, p (or the stress) unknown on Γ_{free}. On the rest of the boundary of $\Omega(t)$ the usual boundary conditions can be considered.

2.2 Steps in the algorithm

Suppose Ω^0 is meshed with a finite element mesh M^0 and that at time level t^n the domain Ω^n is meshed with a finite element mesh M^n. Let $\boldsymbol{u}^n$ be the velocity already computed on Ω^n. The purpose is to obtain the fluid region Ω^{n+1} and the velocity field $\boldsymbol{u}^{n+1}$.

The steps of the algorithm we propose are the following:

1. Define $\Gamma_{\text{free}}^{n+1}$ by updating a *level set function.*

2. Deform *virtually* the mesh M^n to M_{virt}^{n+1} using the classical ALE concepts and compute the mesh velocity $\boldsymbol{u}_m^{n+1}$.

3. Write down the ALE Navier-Stokes equations on M_{virt}^{n+1}.

4. *Split the elements* of M^0 cut by $\Gamma_{\text{free}}^{n+1}$ to define a mesh on Ω^{n+1}, M^{n+1}.

5. *Project* the ALE Navier-Stokes equations from M_{virt}^{n+1} to M^{n+1}.

6. Solve the equations on M^{n+1} to compute $\boldsymbol{u}^{n+1}$ and p^{n+1}.

In the following we describe all these steps.

2.3   Step 1. Level set function update

The boundary Γ_{free} is defined using the classical level set approach, which consists in computing a level set function solution of

$$\partial_t \psi + \boldsymbol{u} \cdot \nabla \psi = 0 \qquad \text{in } \Omega^0, \ t > 0 \tag{1}$$

$$\psi = \overline{\psi} \qquad \text{on } \Gamma_{\text{inf}}, \ t > 0 \tag{2}$$

$$\psi(\boldsymbol{x}, 0) = \psi_0(\boldsymbol{x}) \qquad \text{in } \Omega^0 \tag{3}$$

where $\Gamma_{\text{inf}} := \{\boldsymbol{x} \in \partial\Omega^0 \mid \boldsymbol{u} \cdot \boldsymbol{n} < 0\}$ is the inflow part of the boundary. Once ψ is computed, the boundary $\Gamma_{\text{free}}(t)$ is defined as its zero isovalue surface.

It is noted that (1)-(3) is solved on the whole domain Ω^0. As mentioned earlier, we approximate this problem using a stabilized finite element method. For the discrete problem it is necessary to extrapolate the velocity defined on $\Omega(t)$ to the rest of Ω^0.

2.4   Step 2. Mesh velocity

Updating ψ defines the deformation of the domain from Ω^n to Ω^{n+1}. Consequently, the mesh M^n used at time step n has to be deformed to adapt to the domain Ω^{n+1}. This mesh deformation has to be defined by means of a mesh velocity.

The mesh velocity on the boundary points can be computed from their position $\boldsymbol{x}^{n+1}$ and $\boldsymbol{x}^n$. Using the simplest first order approximation, this mesh velocity would be $\boldsymbol{u}_m = (\boldsymbol{x}^{n+1} - \boldsymbol{x}^n)/\delta t$, where δt is the time step size. Once the velocity on the nodes of Γ_{free} is known, it has to be extended to the rest of the nodes. A classical possibility is to solve the Laplace problem $\Delta \boldsymbol{u}_m = 0$. However, It is also possible to restrict $\boldsymbol{u}_m \neq \boldsymbol{0}$ to the nodes next to $\Gamma_{\text{free}}^{n+1}$, since in our approach mesh distortion does not accumulate from one time step to another. See Fig. 1 for a schematic of the mesh deformation.
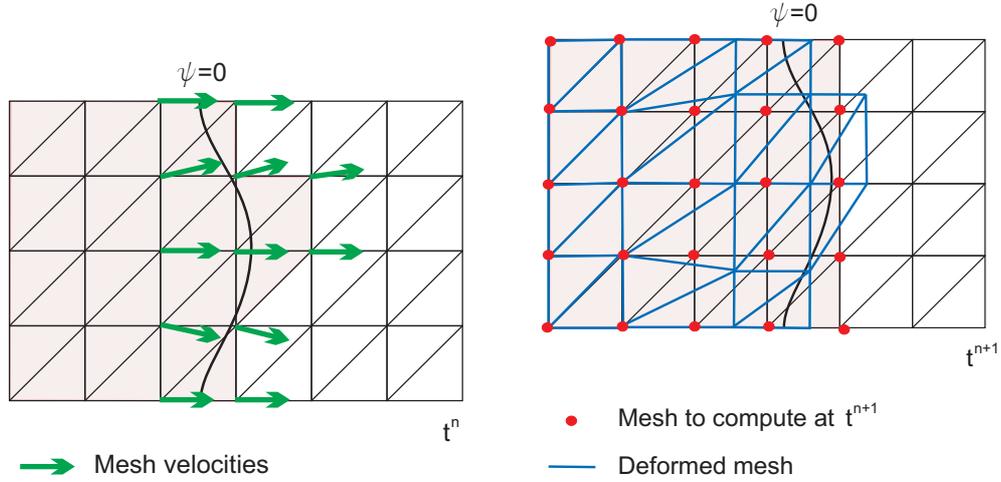
Figure 1: Two dimensional FM-ALE schematic

2.5 Step 3. Solving the flow equations I: Equations on the deformed mesh

The previous procedure defines the domain Ω^{n+1} and a mesh that we call M_{virt}^{n+1}, obtained from a deformation of the mesh M^n. The equations to be solved there, using for example the backward Euler scheme, are:

$$\rho \left[\frac{1}{\delta t}(\boldsymbol{u}^{n+1} - \boldsymbol{u}^n) + (\boldsymbol{u}^{n+1} - \boldsymbol{u}_m) \cdot \nabla \boldsymbol{u}^{n+1} \right] - \nabla \cdot [2\mu \nabla^S (\boldsymbol{u}^{n+1})] + \nabla p^{n+1} = \boldsymbol{f}$$

$$\nabla \cdot \boldsymbol{u}^{n+1} = 0$$

in Ω^{n+1} with boundary conditions on $\Gamma_{\text{free}}^{n+1}$. The mesh velocity $\boldsymbol{u}_m$ has been computed before. These equations should in principle be computed on M_{virt}^{n+1}.

2.6 Step 4. Splitting of elements

Elements on the background mesh M^0 cut by $\Gamma_{\text{free}}^{n+1}$ are split to define a mesh on M^{n+1}. This is a strictly algorithmic step that shall not be discussed here. In the case of 2D linear elements, Fig. 2 shows how the splitting can be done and the numerical integration points (red points) required in each triangle resulting from this splitting.

Both M^{n+1} and M_{virt}^{n+1} are meshes of the domain Ω^{n+1}. However, M^{n+1} *is a minor modification of the background mesh* M^0. They only differ in the subelements created after the splitting just mentioned.

2.7 Step 5. Solving the flow equations II: Equations on the background mesh

Let P^{n+1} be the projection of finite element functions defined on M_{virt}^{n+1} to M^{n+1}. To define it, for each node of M^{n+1} the element in M_{virt}^{n+1} where it is placed has to be identified. Once
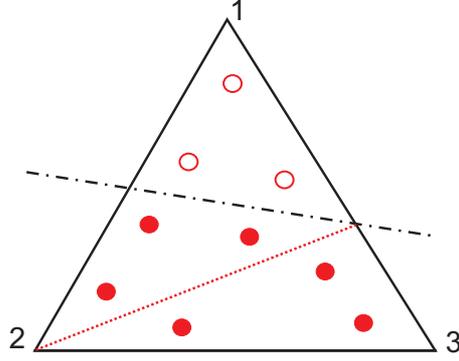
Figure 2: Splitting of elements

this is done, the value of any unknown at this node can be obtained through interpolation, possibly with restrictions.

The velocity $\boldsymbol{u}^n$ in M_{virt}^{n+1} is known because its nodal values correspond to those of mesh M^n. However, its nodal values on M^{n+1} have to be computed using the projection just described. The same happens with the mesh velocity $\boldsymbol{u}_m$.

Denoting again $\boldsymbol{u}^{n+1} \equiv P^{n+1}(\boldsymbol{u}^{n+1})$, the flow equations on M^{n+1} are

$$\rho \left[\frac{1}{\delta t}(\boldsymbol{u}^{n+1} - P^{n+1}(\boldsymbol{u}^n)) + (\boldsymbol{u}^{n+1} - P^{n+1}(\boldsymbol{u}_m)) \cdot \nabla \boldsymbol{u}^{n+1} \right] - \nabla \cdot [2\mu \nabla^S(\boldsymbol{u}^{n+1})] + \nabla p^{n+1} = \boldsymbol{f}$$

$$\nabla \cdot \boldsymbol{u}^{n+1} = 0$$

in Ω^{n+1}. Boundary conditions have to be prescribed on the whole boundary of Ω^{n+1}, including $\Gamma_{\text{free}}^{n+1}$, which is defined by the level set function.

The problem obtained can now be discretized on M^{n+1} which, as it has been said, coincides with M^0 except for the splitting of the elements crossed by the interface. Even this difference can be avoided if instead of prescribing exactly the boundary conditions an approximation is performed, for example using a Nitsche method or Lagrange multipliers. Therefore, the goal of using a fixed mesh during the whole simulation has been achieved.

2.8 Comparison with the classical ALE approach

To conclude this section, it is important to highlight the differences between our fixed-mesh ALE approach and a classical ALE formulation:

- Given a position of the fluid front on the fixed mesh, elements cut by the front are split into subelements (only for integration purposes), so that the front coincides with the edges of the subelements.

- After deforming the mesh from one time step to the other using classical ALE procedures, results are projected back to the original mesh .

- The front is represented by the level set function, and not by the position of the material points at the free surface as in a classical ALE method.

3 Application to free surface flows and lost foam casting

We have applied the strategy described to two problems of practical interest. From the conceptual point of view, the difference between them is only the boundary conditions to be prescribed on the evolving interface defined by the level set function ψ. These applications are:

Free surface flows. In this case, the pressure on the moving interface is prescribed, for example to zero, and the velocity of the fluid in unknown, determined precisely by the dynamics of the flow. A detailed description of the FM-ALE formulation in this setting can be found in [1].

Lost foam casting. In this casting process the molten metal enters the cavity to be filled by burning a foam that originally occupies it. The evolving interface is determined by the way in which this combustion is done. A simple energy budged allows to obtain a front velocity. Thus, contrary to the previous example, now the front velocity is known but the pressure of the fluid interface needs to be computed. The complete formulation we use for modeling this process can be found in [2].

References

[1] H. Coppola-Owen and R. Codina. A finite element model for free surface flows on fixed meshes. *International Journal for Numerical Methods in Fluids*, 54:1151–1171, 2007.

[2] G. Houzeaux and R. Codina. A finite element model for the simulation of lost foam casting. *International Journal for Numerical Methods in Fluids*, 46:203–226, 2004.

Improvements on CMM calibrations

M. A. F. Vicente[*]

Abstract

In this paper, we present a new calibration artefact to determine the Coordinate Measuring Machine global errors, describe a method to make its self-calibration and, finally, we present some obtained simulated results for a machine calibration.

1 Introduction

A Coordinate Measuring Machine (CMM) is a mechanical system designed to move a measuring probe to determine the coordinates of the surface of a workpiece and, as other instruments, need to be calibrated, in order to correct their measurement errors and establish their uncertainty of measurement. Calibration and numerical error correction of CMMs require methods, based on the use of artefacts, to assess the errors. Several artefact-based methods can be used to perform the calibration of CMMs: using gauge blocks, step gauges, ball bars, ball plates, hole plates, laser interferometers, etc.[1]. Attending that all those methods require expensive tools, are time consuming and need especially skilled personnel, we describe in this paper a new artefact which avoid several requirements mentioned above.

2 The artefact for global calibration of CMMs

We intend as global calibration the determination of the machine global errors, which are the errors in the determination of the sensor coordinates when measuring. For this purpose, a new kind of artefact has been introduced in [4]. The proposed artefact is similar to a tripod system[2] constituted by three links connecting, using universal joints, three fixed spheres to a mobile sphere, with the mobile sphere acting as the reference point for CMM calibration. Knowing the artefact initial geometry, the position of the mobile sphere can be determined

---

[*]Department of Mathematics, Faculty of Sciences and Technology, University of Coimbra. E-mail:vicente@mat.uc.pt. Supported by INESC-Coimbra.

[1]Additionally, there are uncalibrated devices, like the mechanical six-sphere artefact presented in [4].

[2]The artefact can be seen as a simplification of the hexapod manipulator, where the base platform is defined by the centers of the 3 fixed spheres, the mobile platform is reduced to just one point (the center of the mobile sphere) and the struts connecting the two platforms are reduced, obviously, from 6 to 3.
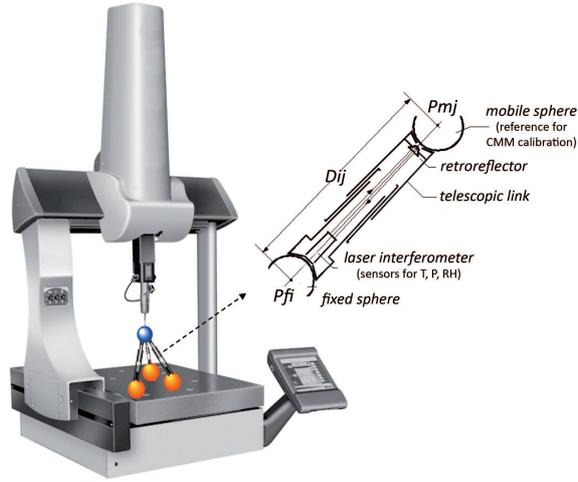
Figure 1: CMM calibration using the artefact.

from the strut lengths, using the trilateration equations. The validation of the artefact and the measurement techniques use data fusion applied to non-linear and non-observable systems. The global calibration of CMMs is done by comparing the calculated position of the moving sphere center with the corresponding measured position by the CMM in the measurement volume. Only length variations are used to obtain the calculated position and these length variations are measured by three miniature laser interferometers installed inside the telescopic links (see figure 1). The artefact uses a self-calibrated method and its modeling is based on laser trilateration. The CMM calibration can be divided into two steps:

1. *Prediction* (self-calibration of the artefact): knowledge about system geometry (localization of the fixed points) and tracking (following the mobile point) are used to obtain geometry identification. That is done by finding the initial length of the links L_i, the coordinates x_j and y_j of the mobile point Pm_j and the coordinates $x_{0,i}$ and $y_{0,i}$ of the fixed points Pf_i, using the laser interferometer measurements dL_{ij}, $i = 1, 2, 3$, $j = 1, \ldots, m$.

2. *Update* (system identification): after identifying the geometry, the remaining problem is the identification of the mobile point position $Pm_j = (x_j, y_j, z_j)$, in real time, using 3D laser trilateration based on the laser interferometric measurements dL_{ij}, $i = 1, 2, 3$, $j = 1, \ldots, m$.

The centers Pf_1, Pf_2 and Pf_3 of the 3 fixed spheres of the artefact define its coordinate system $(O, \overrightarrow{e_1}, \overrightarrow{e_2}, \overrightarrow{e_3})$: $O \equiv Pf_1$, $\overrightarrow{e_1} = \frac{\overrightarrow{Pf_1Pf_2}}{\left\|\overrightarrow{Pf_1Pf_2}\right\|}$, $\overrightarrow{e_3} = \frac{\overrightarrow{Pf_1Pf_2} \times \overrightarrow{Pf_1Pf_3}}{\left\|\overrightarrow{Pf_1Pf_2} \times \overrightarrow{Pf_1Pf_3}\right\|}$ and $\overrightarrow{e_2} = \overrightarrow{e_3} \times \overrightarrow{e_1}$. So, $Pf_1 = (0, 0, 0)$, $Pf_2 = (x_{0,2}, 0, 0)$ and $Pf_3 = (x_{0,3}, y_{0,3}, 0)$. All fixed spheres are constrained to the CMM table and, during self-calibration, also the moving sphere. Flatness table errors are neglected because their contributions to initial link length, so as for link variations, are

144
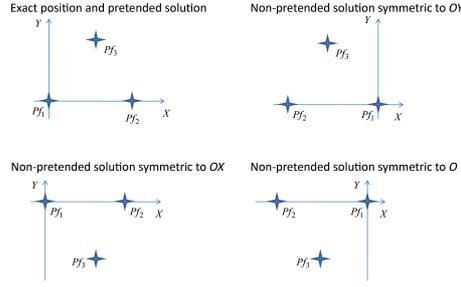
Figure 2: Possible evident solutions for self-calibration.

second order contributions (as proved in [4]). Self-calibration (geometry identification) is done by moving the mobile sphere over the plane of the fixed spheres (the plane OXY), in m positions, with $m \geq 6$, and minimizing the cost function c:

$$c = \sqrt{\sum_{j=1}^{m}(a_{1j}^2 + a_{2j}^2 + a_{3j}^2)}, \text{ with } \begin{cases} a_{1j} = \sqrt{(x_j - x_{0,1})^2 + (y_j - y_{0,1})^2} - (L_1 + dL_{1j}) \\ a_{2j} = \sqrt{(x_j - x_{0,2})^2 + (y_j - y_{0,2})^2} - (L_2 + dL_{2j}) \\ a_{3j} = \sqrt{(x_j - x_{0,3})^2 + (y_j - y_{0,3})^2} - (L_3 + dL_{3j}) \\ \qquad\qquad j = 1, \ldots, m \end{cases}, \quad (1)$$

with $x_{0,1}$, $y_{0,1}$, $x_{0,2}$, $y_{0,2}$, $x_{0,3}$ and $y_{0,3}$ the coordinates of the centers Pf_1, Pf_2 and Pf_3 of the fixed spheres (with the previous definitions, $x_{0,1} = y_{0,1} = y_{0,2} = 0$), x_j and y_j the estimated coordinates of the mobile sphere in the jth position Pm_j, L_1, L_2, L_3 the estimated lengths for the three links at the initial position ($j = 1$) and dL_{1j}, dL_{2j}, dL_{3j} the measured length variations of the three links (with $dL_{11} = dL_{21} = dL_{31} = 0$). The problem has many solutions, and the fixed spheres location enables a quick evaluation of the validity of the solution. In figure 2 we present four possible solutions for the same inputs, resulting from the evident axis and origin symmetrical solutions. Attending to the defined artefact referential, we should find the solutions in the first quadrant of OXY. So, we impose the $4j + 12$ constraints

$$0 \leq x_{0,2}, x_{0,3}, y_{0,3}, L_1, L_2, L_3, x_1, y_1, \ldots, x_j, y_j \leq Lm, \qquad (2)$$

with Lm the physical maximum values for the coordinates and struts lengths. The optimal solution search is done starting with a constrained method[3] in order to fix the solution in the adequate region and concluding with an unconstrained method[4] to obtain a quicker and better solution. The cost function value is then used to obtain the estimated uncertainty for the estimated parameters.

After the self-calibration of the artefact, the coordinates of the fixed spheres centers Pf_1, Pf_2 and Pf_3 are known as also as the initial lengths L_1, L_2 and L_3 of the links. Moving now

---

[3]Using the optimality Karush-Kuhn-Tucker conditions and search of solution by sequential quadratic programming (as defined in [3]).

[4]Using a Quasi-Newton method with BFGS actualization for the Hessian matrix (as defined in [1]).

the mobile sphere in several positions Pm_j in the CMM measurement volume (leaving the fixed spheres plane) and measuring, by laser interferometry[5], the three distances D_{ij} between the fixed spheres and the position of the moving sphere Pm_j ($D_{ij} = \overline{Pf_i Pm_j} = L_i + dL_{ij}$), its coordinates (x_j, y_j, z_j) and their uncertainties can be determined using the relations:

$$x_j = \frac{D_{1j}^2 - D_{2j}^2 + x_{0,2}^2}{2x_{0,2}}, \tag{3}$$

$$y_j = \frac{D_{1j}^2 - D_{3j}^2 + x_{0,3}^2 + y_{0,3}^2 - x_{0,3}\dfrac{D_{1j}^2 - D_{2j}^2 + x_{0,2}^2}{x_{0,2}}}{2y_{0,3}}, \tag{4}$$

$$z_j = \sqrt{D_{1j}^2 - x_j^2 - y_j^2}, \tag{5}$$

which are obtained from the general tridimensional trilateration equations. The coordinates (x_j, y_j, z_j) are then compared[6] to the correspondent coordinates measured by the CMM, in order the perform its calibration. We also use the equations (3), (4) and (5) to find, following the GUM[7], the uncertainty for the reference coordinates, knowing the uncertainties for the center of the fixed spheres and also for the length of the links.

3   Numerical simulations

The numerical simulations, using the optimization toolbox of *Matlab*, are divided into two major parts: in the first one, the purpose is to realize the self-calibration of the system and, in the second one, using the results for the adjusted parameters obtained in the first part and the trilateration equations, estimated coordinates and the correspondent uncertainty are obtained for some reference points. The initial parameters are the maximum link length, the link's length variation noise amplitude, the number of constrained steps (necessary to constrain the area where the solutions could be found) and forbidden area radius (where no points can be placed, corresponding to a physical limitation of the artefact). After the introduction

---

[5]The increments dL_{ij} in the links lengths are measured by laser interferometry and are perturbed, at least, by ambient humidity, ambient temperature and atmospheric pressure. Also, measured dL_{ij} are estimated with an uncertainty composed by the uncertainties of the measured perturbations.

[6]After the necessary coordinate transformation done in order to have all the coordinates in the same referential (the CMM coordinate system).

[7]Following the *ISO Guide to the Expression of Uncertainty in Measurement* ([2]), the uncertainty u_b for a dependent measure $b = g(a_1, a_2, \ldots, a_n)$, function of n uncorrelated measurements a_i with uncertainties u_{a_i}, is defined by $u_b = \sqrt{\sum\limits_{i=1}^{n} \left[\frac{\partial g}{\partial a_i}(P)\, u_{a_i} \right]^2}$, with $P = (a_1, a_2, \ldots, a_n)$. Taking into account that the uncertainties have small values, we substitute $\frac{\partial g}{\partial a_i}$ by finite differences $\frac{g(P + \delta a_i e_i) - g(P)}{\delta a_i}$ and identify each increment δa_i with the uncertainty of a_i, i.e., $\delta a_i = u_{a_i}$, $i = 1, \ldots, n$. Consequently, we get $u_b \simeq \sqrt{\sum\limits_{i=1}^{n} [g(P + u_{a_i} e_i) - g(P)]^2}$. For our purpose, g is given by (3), (4) and (5).
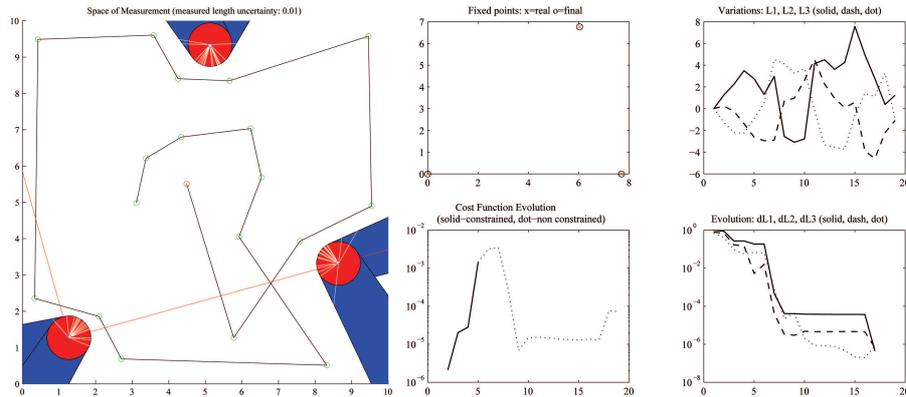
Figure 3: Example of fixed spheres position and mobile sphere trajectory during self-calibration (left) and graphical results from self-calibration (right).

of the initial parameters, the three fixed spheres positions and the various moving sphere positions are chosen (see figure 3, left). For every new mobile position, the parameters are recalculated using all the information available and the results of self-calibration are given, in a graphical and in a numerical form (see table 1). The results, in a graphical form, shows the real positions versus estimated positions of the fixed points, the variations of the links lengths, the evolution of the cost function and the evolution of the estimated values for the initial link lengths (see figure 3, right).

	reference values	estimated values	deviations
L_1, L_2, L_3	294.3829, 487.5984, 497.9873	294.3831, 487.5982, 497.9870	0.0002, -0.0002, -0.0003
$x_{0,2}, x_{0,3}, y_{0,3}$	619.0832, 487.7940, 624.2054	619.0831, 487.7945, 624.2051	-0.0001, 0.0005, -0.0003
(x_1, y_1) of Pm_1	(187.5136, 226.9360)	(187.5139, 226.9361)	(0.0003, 0.0001)
(x_2, y_2) of Pm_2	(203.2694, 367.9109)	(203.2699, 367.9109)	(0.0005, 0.0000)
...	...	...	...
$Lm = 20$ dm, forbidden area radius 0.6 dm and 5 constrained steps			

Table 1: Results, in mm, of self-calibration with links length noise amplitude 0.0005 mm.

After the self-calibration, some values of 3D reference points coordinates are compared with the correspondent estimated coordinates, obtained from the final parameters (given by the self-calibration process) using trilateration. Those results have a corresponding estimated uncertainty, function of the uncertainty of the links lengths (in table 2 we summarize some results, in mm, obtained in three simulations with different noise amplitudes).

4 Conclusions

The coordinates and their uncertainties, which are the reference values for CMMs calibration, are calculated in real time, by a method based on laser interferometer measurement of the links length variation. We presented a technique to incorporate the measured values uncertainties in order to obtain the calibration results final uncertainties (using data fusion

reference point coordinates	links length noise amplitude		
	0	0.0005	0.005
	estimated coordinates		
$x = 10$	9.999941 ± 0.000008	10.0001 ± 0.0005	9.997 ± 0.005
$y = 10$	9.999840 ± 0.000020	9.9998 ± 0.0008	10.010 ± 0.008
$z = 10$	10.000040 ± 0.000020	10.0000 ± 0.0020	9.990 ± 0.020
$x = 50$	49.999933 ± 0.000008	50.0001 ± 0.0005	50.004 ± 0.005
$y = 50$	49.999850 ± 0.000020	50.0001 ± 0.0007	49.998 ± 0.007
$z = 50$	50.000040 ± 0.000020	50.0010 ± 0.0010	50.000 ± 0.010
$x = 500$	499.999845 ± 0.000007	500.0004 ± 0.0008	500.012 ± 0.008
$y = 400$	400.000000 ± 0.000007	400.0003 ± 0.0005	400.000 ± 0.005
$z = 300$	300.000020 ± 0.000010	299.9998 ± 0.0005	299.997 ± 0.005
$x = 1000$	999.99962 ± 0.00002	1000.002 ± 0.002	1000.030 ± 0.020
$y = 1000$	1000.00022 ± 0.00002	1000.001 ± 0.002	999.980 ± 0.020
$z = 1000$	999.99998 ± 0.00002	999.998 ± 0.002	1000.000 ± 0.020

Table 2: Some trilateration and uncertainty results.

as a learning system, meaning that the optimal uncertainty estimation is constantly adapted during self-calibration of the device). However, this is not the final answer for this problem of artefact validation: different techniques in non-linear and non-observable state estimation can be compared. We think that the developed artefact enables the measurement of global errors of the CMM in a easy way, providing good accuracy, and this artefact, in addition to CMM's calibration, can be applied in many other fields, for instance to perform tests of robots, to perform robot localization, to act as components of machines or even as CMMs. Moreover, being self-calibrated, the expensive and complex calibration of the artefact is avoided. Furthermore, the results given by using the developed *Matlab* routines based on the above make us conclude that the self-calibration of the artefact is correctly done. In fact, the system geometry identification is done in a fast and reliable way, independent from the initial approximation. The modeling of the artefact is the base to develop a real artefact which should be tested in order to compare the numerical simulated results with experimental results.

References

[1] P. E. Frandsen, K. Jonasson, H. B. Nielsen, and O. Tingleff. *Unconstrained Optimization.* Informatics and Mathematical Modelling, Technical University of Denmark, 3rd edition, 2004.

[2] International Organization for Standardization. *ISO Guide to the Expression of Uncertainty in Measurements (GUM).* Geneva, Switzerland, 1995.

[3] J. Nocedal and S. J. Wright. *Numerical Optimization.* Springer Series in Operations Research. Springer, USA, 1999.

[4] M. A. F. Vicente. *Simulação Numérica de Novos Artefactos de Calibração de Máquinas de Medição Tridimensional.* PhD thesis, Departamento de Matemática da Faculdade de Ciências e Tecnologia da Universidade de Coimbra, Coimbra, Portugal, 2007.

Numerical Analysis of Thin Structures

A-posteriori error estimates for the Reissner-Mindlin plate problem

L. Beirão da Veiga[*]   C. Chinosi[†]   C. Lovadina[‡]   R. Stenberg[§]

Abstract

We present an a-posteriori error analysis for some mixed finite element methods to approximate the solution of the Reissner-Mindlin plate problem. More precisely, we focus our attention on a low-order triangular mixed element presented by Falk and Tu (see [5]). For such a scheme we introduce a suitable residual-based error estimator, and we discuss its reliability and efficiency. In particular, we show that the error estimator is robust with respect to the choice of the thickness parameter.

1   Introduction

The finite element simulation of Reissner-Mindlin plates has been an active research area, due both to its practical importance, and to the non-trivial problems to overcome. In particular, non-standard techniques need to be used to avoid the *shear-locking phenomenon*, as well as the *spurious mode occurrence* and the *boundary layer effects* (see [3], for example). We here focus on the *a-posteriori* analysis for a family of triangluar mixed element presented by Falk and Tu (cf. [5]) establishing the equivalence, uniformly in the plate thickness, between a suitable residual-based estimator and the error. To obtain this, we follow the guidelines of [6]. The proofs of all the results here stated, together with an improved *a-priori* error analysis and several numerical tests, can be found in [1]. Throughout the paper we will use standard notations for Sobolev norms and seminorms. Moreover, we will denote with C a generic constant which may take different values in different occurrences, and which is *independent* of the mesh parameter h and the plate thickness t.

---

[*]Dipartimento di Matematica, Università di Milano. E-mail: `beirao@mat.unimi.it`.

[†]Dipartimento di Scienze e Tecnologie Avanzate, Università del Piemonte Orientale. E-mail: `chinosi@mfn.unipmn.it`.

[‡]Dipartimento di Matematica, Università di Pavia. E-mail: `carlo.lovadina@unipv.it`.

[§]Institute of Mathematics, Helsinki University of Technology. E-mail: `rolf.stenberg@hut.fi`.

2 The Reissner-Mindlin problem and the finite element discretization

The Reissner-Mindlin equations for a plate with polygonal mid-plane Ω require to find $(\boldsymbol{\theta}, w, \boldsymbol{\gamma})$ such that

$$
\begin{cases}
-\operatorname{\mathbf{div}} \mathbf{C}\varepsilon(\boldsymbol{\theta}) - \boldsymbol{\gamma} = \mathbf{0} & \text{in } \Omega, \\
-\operatorname{div} \boldsymbol{\gamma} = g & \text{in } \Omega, \\
\boldsymbol{\gamma} = \mu t^{-2}(\nabla w - \boldsymbol{\theta}) & \text{in } \Omega, \\
\boldsymbol{\theta} = \mathbf{0}, \; w = 0 & \text{on } \Gamma_C, \\
\boldsymbol{\theta} \cdot \mathbf{t} = 0, \; w = 0, \; (\mathbf{C}\varepsilon(\boldsymbol{\theta})\mathbf{n}) \cdot \mathbf{n} = 0 & \text{on } \Gamma_H, \\
w = 0, \; \mathbf{C}\varepsilon(\boldsymbol{\theta})\mathbf{n} = \mathbf{0} & \text{on } \Gamma_S, \\
\mathbf{C}\varepsilon(\boldsymbol{\theta})\mathbf{n} = \mathbf{0}, \; \boldsymbol{\gamma} \cdot \mathbf{n} = 0 & \text{on } \Gamma_F.
\end{cases}
\tag{1}
$$

Here, the plate boundary $\partial\Omega$ is split into four non-overlapping parts Γ_C, Γ_H, Γ_S and Γ_F, where clamped, hard simply supported, soft simply supported and free boundary conditions are respectively imposed; we also assume that Γ_C has positive measure. Moreover, $\mathbf{C}$ is the tensor of bending moduli, $\boldsymbol{\theta}$ represents the rotations, w the transversal displacement, $\boldsymbol{\gamma}$ the scaled shear stresses and g a given transversal load. Finally, ε is the usual symmetric gradient operator, μ is the shear modulus, and t is the thickness. The classical variational formulation of problem (1) is:

$$
\begin{cases}
\text{Find } (\boldsymbol{\theta}, w, \boldsymbol{\gamma}) \in \boldsymbol{\Theta} \times W \times \boldsymbol{\Gamma} \text{ such that} \\
a(\boldsymbol{\theta}, \boldsymbol{\eta}) + (\nabla v - \boldsymbol{\eta}, \boldsymbol{\gamma}) = (g, v) & \forall (\boldsymbol{\eta}, v) \in \boldsymbol{\Theta} \times W, \\
(\nabla w - \boldsymbol{\theta}, \boldsymbol{\tau}) - \mu^{-1} t^2 (\boldsymbol{\gamma}, \boldsymbol{\tau}) = 0 & \forall \boldsymbol{\tau} \in (L^2(\Omega))^2,
\end{cases}
\tag{2}
$$

where

$$
\begin{aligned}
\boldsymbol{\Theta} &= \left\{ \boldsymbol{\eta} \in (H^1(\Omega))^2 \; : \; \boldsymbol{\eta} = \mathbf{0} \text{ on } \Gamma_C \; ; \; \boldsymbol{\eta} \cdot \mathbf{t} = 0 \text{ on } \Gamma_H \right\}, \\
W &= \left\{ v \in H^1(\Omega) \; : \; v = 0 \text{ on } \Gamma_C \cup \Gamma_S \right\},
\end{aligned}
\tag{3}
$$

$(\cdot, \cdot)$ is the inner-product in $L^2(\Omega)$ and $(L^2(\Omega))^2$. Moreover, the bilinear form $a(\cdot, \cdot)$ is defined by

$$
a(\boldsymbol{\theta}, \boldsymbol{\eta}) := \int_\Omega \mathbf{C}\varepsilon(\boldsymbol{\theta}) : \varepsilon(\boldsymbol{\eta}).
$$

Let $\{\mathcal{T}_h\}_{h>0}$ be a sequence of decompositions of Ω into triangular elements T, satisfying the usual compatibility conditions (see [2]). We also assume that the family $\{\mathcal{T}_h\}_{h>0}$ is *regular*. Moreover, given the decomposition $\mathcal{T}_h$ we will denote with $\mathcal{E}_h$ the set of all the edges E of the triangles $T \in \mathcal{T}_h$, and with $\mathcal{E}_h^I \subset \mathcal{E}_h$ the set of all internal edges. The subset of boundary edges will instead be indicated with $\mathcal{E}_h^C$, $\mathcal{E}_h^H$, $\mathcal{E}_h^S$ and $\mathcal{E}_h^F$, representing respectively the set of edges laying in Γ_C, Γ_H, Γ_S and Γ_F.

The family to be considered uses the FE spaces (see [5]), for $k \geq 1$:

$$\boldsymbol{\Theta}_h = \left\{ \boldsymbol{\eta} \in \boldsymbol{\Theta} \mid \boldsymbol{\eta}_{|T} \in [P_k(T) + B_{k+3}(T)]^2 \ \forall T \in \mathcal{T}_h \right\}, \tag{4}$$

$$W_h = \left\{ v \in W \mid v_{|T} \in P_{k+1}(T) \ \forall T \in \mathcal{T}_h \right\}, \tag{5}$$

$$\boldsymbol{\Gamma}_h = \left\{ \boldsymbol{\tau} \in (L^2(\Omega))^2 \mid \boldsymbol{\tau}_{|T} \in [P_k(T)]^2 \ \forall T \in \mathcal{T}_h \right\}, \tag{6}$$

where $P_l(T)$, $l = k, k+1$, is the space of polynomials of degree at most l defined on T, and $B_{k+3}(T) = P_{k+3}(T) \cap H_0^1(T)$ is a space of bubbles on T. Then, we consider the discrete problem:

$$\begin{cases} \text{Find } (\boldsymbol{\theta}_h, w_h; \boldsymbol{\gamma}_h) \in \boldsymbol{\Theta}_h \times W_h \times \boldsymbol{\Gamma}_h \text{ such that} \\ a(\boldsymbol{\theta}_h, \boldsymbol{\eta}) + (\boldsymbol{\gamma}_h, \nabla v - \boldsymbol{\eta}) = (g, v) & (\boldsymbol{\eta}, v) \in \boldsymbol{\Theta}_h \times W_h, \\ (\nabla w_h - \boldsymbol{\theta}_h, \boldsymbol{\tau}) - \mu^{-1} t^2 (\boldsymbol{\gamma}_h, \boldsymbol{\tau}) = 0 & \boldsymbol{\tau} \in \boldsymbol{\Gamma}_h. \end{cases} \tag{7}$$

For the above family of elements, one can develop *a priori* error estimates with respect to the norms (cf. [4], [7], [1])

$$\|(\boldsymbol{\eta}, v)\|_h^2 := \|\boldsymbol{\eta}\|_1^2 + \|v\|_1^2 + \sum_{T \in \mathcal{T}_h} \frac{1}{h_T^2 + t^2} \|\nabla v - \boldsymbol{\eta}\|_{0,T}^2 \qquad \forall (\boldsymbol{\eta}, v) \in \boldsymbol{\Theta} \times W \tag{8}$$

and

$$\|\boldsymbol{\tau}\|_{\boldsymbol{\Theta}'} + t \|\boldsymbol{\tau}\|_0 \qquad \forall \boldsymbol{\tau} \in (L^2(\Omega))^2 . \tag{9}$$

Above, $\| \cdot \|_{\boldsymbol{\Theta}'}$ denotes the norm in $\boldsymbol{\Theta}'$, the dual space of $\boldsymbol{\Theta}$.

3 A posteriori error estimates

Following the lines of [6], we introduce a suitable error estimator and we study its *reliability* and *efficiency*. To begin, we split the edge set $\mathcal{E}_h$ into

$$\mathcal{E}_h = \mathcal{E}_h^I \cup \mathcal{E}_h^F \cup \mathcal{E}_h^S \cup \mathcal{E}_h^H , \tag{10}$$

where $\mathcal{E}_h^I$ is the set of the inner edges, while $\mathcal{E}_h^F$, $\mathcal{E}_h^S$ and $\mathcal{E}_h^H$ are the sets where free, soft simply supported and hard simply supported boundary conditions are respectively imposed. For each $T \in \mathcal{T}_h$ we introduce the following quantity

$$\widetilde{\eta}_T^2 := h_T^2 \| \mathbf{div} \, \mathbf{C}\boldsymbol{\varepsilon}(\boldsymbol{\theta}_h) + \boldsymbol{\gamma}_h \|_{0,T}^2 + h_T^2 (h_T^2 + t^2) \| \mathrm{div} \, \boldsymbol{\gamma}_h + g_h \|_{0,T}^2$$
$$+ \frac{\mu^2}{h_T^2 + t^2} \| \mu^{-1} t^2 \, \boldsymbol{\gamma}_h - (\nabla w_h - \boldsymbol{\theta}_h) \|_{0,T}^2 , \tag{11}$$

where g_h is some approximation of the load g. Moreover, we set

$$\eta_E^2 := h_E \| [\![\mathbf{C}\boldsymbol{\varepsilon}(\boldsymbol{\theta}_h)\mathbf{n}]\!] \|_{0,E}^2 + h_E (h_E^2 + t^2) \| [\![\boldsymbol{\gamma}_h \cdot \mathbf{n}]\!] \|_{0,E}^2 \qquad E \in \mathcal{E}_h^I \cup \mathcal{E}_h^F ,$$
$$\eta_S^2 := h_E \| [\![\mathbf{C}\boldsymbol{\varepsilon}(\boldsymbol{\theta}_h)\mathbf{n}]\!] \|_{0,E}^2 \qquad\qquad\qquad\qquad E \in \mathcal{E}_h^S , \tag{12}$$
$$\eta_H^2 := h_E \| ([\![\mathbf{C}\boldsymbol{\varepsilon}(\boldsymbol{\theta}_h)\mathbf{n}]\!]) \cdot \mathbf{n} \|_{0,E}^2 \qquad\qquad\qquad E \in \mathcal{E}_h^H ,$$

where h_E is the length of the side E and $[\![\cdot]\!]$ denotes the jump operator. We adopt the usual notation that the jump operator on a boundary edge is equal to the restriction operator on that edge. We then define a *local* indicator η_T as

$$\eta_T := \left(\widetilde{\eta}_T^2 + \sum_{E \subset \partial T \cap (\mathcal{E}_h^I \cup \mathcal{E}_h^F)} \eta_E^2 + \sum_{E \subset \partial T \cap \mathcal{E}_h^S} \eta_S^2 + \sum_{E \subset \partial T \cap \mathcal{E}_h^H} \eta_H^2 \right)^{1/2} , \qquad (13)$$

and a *global* indicator η as

$$\eta := \left(\sum_{T \in \mathcal{T}_h} \widetilde{\eta}_T^2 + \sum_{E \in \mathcal{E}_h^I \cup \mathcal{E}_h^F} \eta_E^2 + \sum_{E \in \mathcal{E}_h^S} \eta_S^2 + \sum_{E \in \mathcal{E}_h^H} \eta_H^2 \right)^{1/2} . \qquad (14)$$

3.1 Lower bounds

As far as the *efficiency* of our error estimator is concerned, the techniques detailed [6] can be used to establish the following proposition.

Proposition 3.1. *Let $(\boldsymbol{\theta}, w; \boldsymbol{\gamma})$ (resp. $(\boldsymbol{\theta}_h, w_h; \boldsymbol{\gamma}_h)$) be the solution of the continuous (resp. discrete) problem. Given $T \in \mathcal{T}_h$, it holds*

$$\eta_T \leq C \left(\frac{1}{(h_T^2 + t^2)^{1/2}} \big|\big| \nabla(w_h - w) - (\boldsymbol{\theta}_h - \boldsymbol{\theta}) \big|\big|_{0,T} + ||\boldsymbol{\theta}_h - \boldsymbol{\theta}||_{1,\omega_T} \right.$$
$$\left. + ||\boldsymbol{\gamma}_h - \boldsymbol{\gamma}||_{\boldsymbol{\Theta}(\omega_T)'} + t\, ||\boldsymbol{\gamma}_h - \boldsymbol{\gamma}||_{0,\omega_T} + \big(\sum_{T' \subset \omega_T} h_{T'}^2 (h_{T'}^2 + t^2) ||g - g_h||_{0,T'}^2 \big)^{1/2} \right) , \qquad (15)$$

where η_T is defined by (11)–(13), and ω_T is the union of the triangles in $\mathcal{T}_h$ sharing a side with T. Moreover, $\boldsymbol{\Theta}(\omega_T)'$ is the dual of the space

$$\boldsymbol{\Theta}(\omega_T) = \{ \boldsymbol{\eta}_{|\omega_T} \ : \ \boldsymbol{\eta} \in \boldsymbol{\Theta} \ and \ \boldsymbol{\eta} = \mathbf{0} \ on \ \partial \omega_T \setminus \partial \Omega \}.$$

3.2 Upper bounds

The indicator just introduced can be used as a *reliable* error estimator. Given an integer $k \geq 1$ (and, therefore, an element in the Falk-Tu family), our upper bounds can be proved by means of a saturation assumption involving the higher order $(k+1)$-th Falk-Tu element. In order to avoid cumbersome notation, we will denote all the quantities relative to this latter element by a "tilde". We need to make the following

Saturation assumption: Let $(\boldsymbol{\theta}_h, w_h, \boldsymbol{\gamma}_h) \in \boldsymbol{\Theta}_h \times W_h \times \boldsymbol{\Gamma}_h$ (resp. $(\widetilde{\boldsymbol{\theta}}_h, \widetilde{w}_h, \widetilde{\boldsymbol{\gamma}}_h) \in \widetilde{\boldsymbol{\Theta}}_h \times W_h \times \widetilde{\boldsymbol{\Gamma}}_h$) be the discrete solution using the k-th (resp. $(k+1)$-th) Falk-Tu element. We assume that

there exists $0 < \rho < 1$ such that

$$
\begin{aligned}
\|(\boldsymbol{\theta} - \widetilde{\boldsymbol{\theta}}_h, w - \widetilde{w}_h)\|_h + ||\boldsymbol{\gamma} - \widetilde{\boldsymbol{\gamma}}_h||_{\boldsymbol{\Theta}'} + t\,||\boldsymbol{\gamma} - \widetilde{\boldsymbol{\gamma}}_h||_0 \\
\leq \rho\Big(\|(\boldsymbol{\theta} - \boldsymbol{\theta}_h, w - w_h)\|_h + ||\boldsymbol{\gamma} - \boldsymbol{\gamma}_h||_{\boldsymbol{\Theta}'} + t\,||\boldsymbol{\gamma} - \boldsymbol{\gamma}_h||_0\Big).
\end{aligned}
\tag{16}
$$

$\square$

We have the following proposition.

Proposition 3.2. *We have*

$$
\begin{aligned}
\|(\widetilde{\boldsymbol{\theta}}_h - \boldsymbol{\theta}_h, \widetilde{w}_h - w_h)\|_h + ||\widetilde{\boldsymbol{\gamma}}_h - \boldsymbol{\gamma}_h||_{\boldsymbol{\Theta}'} + t\,||\widetilde{\boldsymbol{\gamma}}_h - \boldsymbol{\gamma}_h||_0 \\
\leq C \left(\sum_{T \in \mathcal{T}_h} \left(\eta_T^2 + h_T^2(h_T^2 + t^2)||g - g_h||_{0,T}^2 \right) \right)^{1/2}.
\end{aligned}
\tag{17}
$$

3.3 A numerical result: L-shaped plate with constant load

We consider an L-shaped plate, where the domain Ω is obtained carving out the upper right quarter of a $[-1,1] \times [-1,1]$ square. The applied load is a constant function $g = 1$ and the material constants are chosen as in the previous problem. The plate is clamped along the two edges forming the re-entrant corner, while it is free on the other edges. In Figure 1 we show two meshes obtained at different stages of refinement, for both the choices $t = 10^{-2}$ and $t = 10^{-4}$. The initial decomposition is a uniform mesh of 12 triangles. Looking at the refined meshes, the error indicator succeeds in spotting the solution irregularity around the re-entrant corner.

References

[1] L. Beirão da Veiga, C. Chinosi, C. Lovadina, and R. Stenberg. A-priori and a-posteriori error analysis for a family of Reissner-Mindlin plate elements. *Submitted.*

[2] Dietrich Braess. *Finite elements.* Cambridge University Press, Cambridge, third edition, 2007. Theory, fast solvers, and applications in elasticity theory, Translated from the German by Larry L. Schumaker.

[3] Franco Brezzi and Michel Fortin. *Mixed and hybrid finite element methods*, volume 15 of *Springer Series in Computational Mathematics.* Springer-Verlag, New York, 1991.

[4] D. Chapelle and R. Stenberg. An optimal low-order locking-free finite element method for Reissner-Mindlin plates. *Math. Models Methods Appl. Sci.*, 8(3):407–430, 1998.

[5] Richard S. Falk and Tong Tu. Locking-free finite elements for the Reissner-Mindlin plate. *Math. Comp.*, 69(231):911–928, 2000.
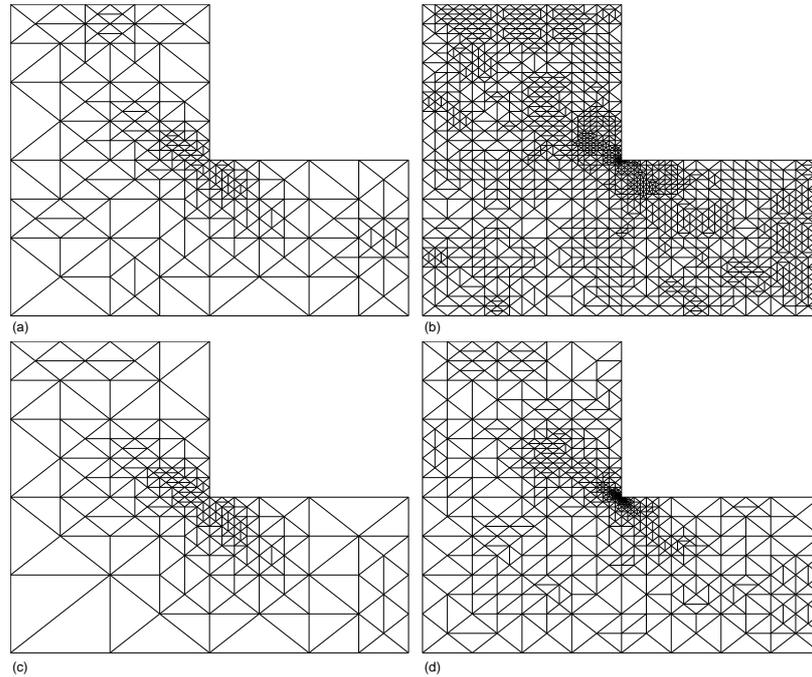
Figure 1: L-shaped plate, clamped corner: (a)-(b) fifth and eighth refinement steps N=3261, for $t = 10^{-2}$; (c)-(d) fifth and eighth refinement, for $t = 10^{-4}$.

[6] Carlo Lovadina and Rolf Stenberg. A posteriori error analysis of the linked interpolation technique for plate bending problems. *SIAM J. Numer. Anal.*, 43(5):2227–2249 (electronic), 2005.

[7] Rolf Stenberg. A new finite element formulation for the plate bending problem. In *Asymptotic methods for elastic structures (Lisbon, 1993)*, pages 209–221. de Gruyter, Berlin, 1995.

Fundamental and applicative challenges in the modeling and computations of shells: an overview

Dominique Chapelle*

As is well-known in engineering practice, shell structures may produce dramatically different responses – especially when the shell thickness is rather small compared to other characteristic dimensions – depending on their geometries and boundary conditions, in particular. The key to understanding these phenomena is to analyse their *asymptotic behaviors*, namely by considering a sequence of problems indexed by the thickness parameter that we vary while maintaining the midsurface geometry and the boundary conditions fixed. We then find that shell structures undergo two main types of very distinct asymptotic behaviors, namely, bending-dominated or membrane-dominated, see [7] and references therein.

Of course, when performing finite element analyses of shell structures we would like the discrete solutions to accurately reflect the diversity of the above behaviors. More precisely, since we only discretize the problem over the midsurface (i.e. not across the thickness), we expect an accuracy that would only depend on criteria prevailing in 2D analysis, namely, uniform convergence in the surface discretization regardless of the thickness parameter.

However, it was soon recognized in the development of structural analysis procedures that standard finite element techniques – such as displacement-based shell finite elements – fail to display uniformly converging behaviors in general, and that instead finite element approximations tend to dramatically deteriorate when the thickness of the structure decreases [1, 9]. In fact, when pursuing the above objectives one faces a dilemma which we now outline. When considering a bending-dominated structure, the numerical difficulty to deal with is *numerical locking*, since the asymptotic behavior then corresponds to a penalized formulation such as for nearly-incompressible elasticity. In order to treat locking, one is led to resorting to *mixed formulations*, which amount to relaxing the penalized constraints by modifying the numerical energy [4]. However, this modification induces a consistency error that is of major concern in membrane-dominated behaviors. Some quadrilateral shell elements representing a satisfactory compromise in this dilemma have been proposed: the MITC ('Mixed Interpolation of Tensorial Components') family [2, 5]. Unfortunately, the formulation of effective triangular shell elements is still more difficult, whereas triangular elements are much needed in practical applications where complex geometries must be handled [3].

*INRIA - Rocquencourt, B.P. 105, 78153 Le Chesnay cedex, France. E-mail:`Dominique.Chapelle@inria.fr`

In this general perspective, we will also present some results and challenges directly related to applicative concerns and numerical practice, in particular as regards general shell elements and 3D-shell elements. These are finite element procedures formulated using 3D constitutive equations and variational principles, which makes them more versatile and easier to implement in engineering practice than methods directly based on shell models – although some underlying mathematical shell models have also been identified for these 3D-based procedures [6, 8].

References

[1] K.J. Bathe. *Finite Element Procedures*. Prentice Hall, 1996.

[2] K.J. Bathe and E.N. Dvorkin. A formulation of general shell elements—the use of mixed interpolation of tensorial components. 22:697–722, 1986.

[3] L. Beirão da Veiga, D. Chapelle, and I. Paris Suarez. Towards improving the MITC6 triangular shell element. *Computers & Structures*, 85:1589–1610, 2007.

[4] F. Brezzi and M. Fortin. *Mixed and Hybrid Finite Element Methods*. Springer-Verlag, New York, 1991.

[5] M.L. Bucalem and K.J. Bathe. Higher-order MITC general shell elements. 36:3729–3754, 1993.

[6] D. Chapelle and K.J. Bathe. The mathematical shell model underlying general shell elements. *Internat. J. Numer. Methods Engrg.*, 48(2):289–313, 2000.

[7] D. Chapelle and K.J. Bathe. *The Finite Element Analysis of Shells – Fundamentals*. Springer-Verlag, 2003.

[8] D. Chapelle, A. Ferent, and K.J. Bathe. 3D-shell finite elements and their underlying model. *M3AS*, 14:105–142, 2004.

[9] T.J.R. Hughes. *The Finite Element Method : Linear Static and Dynamic Finite Element Analysis*. Prentice Hall, 1987.

Numerical shell eigenproblem benchmarks

Harri Hakula[*]

Abstract

The solutions of the thin shell problems depend on a number of problem-specific parameters such as the (dimensionless) thickness of the shell. Asymptotic analysis is concerned with the question of what happens as the thickness tends to zero. From the analysis of the static problems it is known that the shell surface geometry and the kinematic constraints play a central role in determining on what deformation state, bending or membrane, dominates in energy at the limit.

Here we discuss the issues on designing and implementing numerical benchmark computations supporting asymptotic analysis of numerical shell eigenproblems. A complete analysis of the cylindrical shell of revolution and an extension to more challenging elliptic and hyperbolic configurations is presented. Finally, on-going work on the idealized rotor blade vibration problem is considered.

1 Introduction

We consider a set of shells of revolution, whose midsurfaces are defined as follows. Let $I = [x_1, x_2] \subset \mathbf{R}$ be a bounded closed interval, and let $f(x) : I \to \mathbf{R}^+$ be a regular function. The shell midsurface is parametrised by means of the mapping

$$
\phi \; : \; I \times [0, 2\pi] \; \longrightarrow \; \mathbf{R}^3
$$
$$
\phi(\alpha_1, \alpha_2) = (\alpha_1, f(\alpha_1)\sin\alpha_2, f(\alpha_1)\cos\alpha_2) \; . \tag{1}
$$

Refer to Figure 1 for a pictorial view of the curvilinear coordinates α_1 and α_2 adopted for this description.

In what follows we will refer to the following kinds of shells of revolution, grouped in terms of Gaussian curvature which cover all the fundamental types of mid-surface geometry. All the shells are assumed to be clamped at the boundary.

1 – *Zero Gaussian curvature shells [parabolic].* $f''(\alpha_1) = 0, \quad \forall\, \alpha_1 \in I.$

2 – *Positive Gaussian curvature shells [elliptic].* $f''(\alpha_1) < 0, \quad \forall\, \alpha_1 \in I.$

3 – *Negative Gaussian curvature shells [hyperbolic].* $f''(\alpha_1) > 0, \quad \forall\, \alpha_1 \in I.$

[*]Institute of Mathematics, Helsinki University of Technology, POBox 1100, FIN-02015 TKK. E-mail:`Harri.Hakula@tkk.fi`.
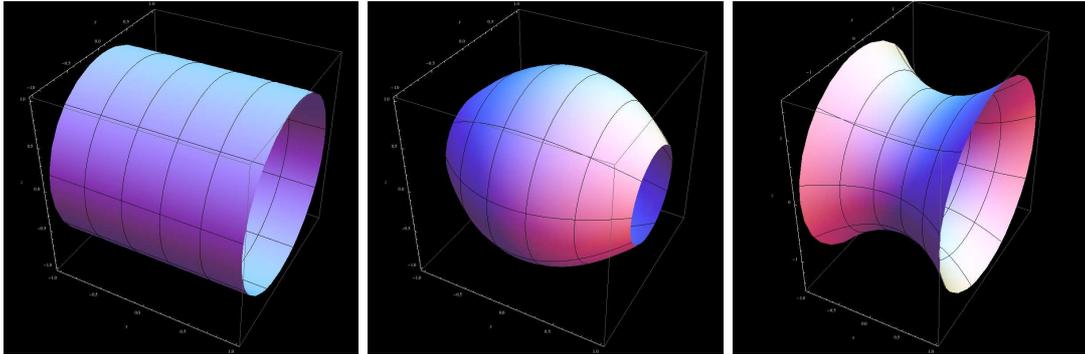
(a) Parabolic   (b) Elliptic   (c) Hyperbolic

Figure 1: Geometries

2   The shell eigenvalue problem

Throughout the paper we consider the dimensionally reduced model of Reissner and Naghdi (Seen for instance [1]). Assuming a time harmonic displacement field, the free vibration problem for a general shell of thickness t leads to the following *eigenvalue* problem

$$
\begin{cases}
\text{Find } \mathbf{u}_t \text{ and } \omega_t^2 \in \mathbf{R} \text{ such that} \\
t\mathcal{A}_M \mathbf{u}_t + t\mathcal{A}_S \mathbf{u}_t + t^3 \mathcal{A}_B \mathbf{u}_t = \omega_t^2 \mathcal{M}(t)\mathbf{u}_t \\
+ \text{ boundary conditions.}
\end{cases}
\tag{2}
$$

Above, $\mathbf{u}_t = (u_t, v_t, w_t, \theta_t, \psi_t)$ represents the shell displacement field of three translations and two rotations, while ω_t^2 represents the square of the eigenfrequency. The differential operators $\mathcal{A}_M, \mathcal{A}_S$ and $\mathcal{A}_B$ account for membrane, shear and bending potential energies, respectively and are *independent of* t. Finally, $\mathcal{M}(t)$ is the inertia operator, which in this case can be split into the sum $\mathcal{M}(t) = \rho(t\mathcal{M}^l + \frac{t^3}{12}\mathcal{M}^r)$, with displacement part $\mathcal{M}^l$ and rotation dependent part $\mathcal{M}^r$ *independent of* t and ρ denoting density.

The variational formulation of problem (2), with the space V of admissible displacements is

$$
\begin{cases}
\text{Find } (\mathbf{u}_t, \omega_t^2) \in V \times \mathbf{R} \text{ such that} \\
t\mathcal{A}_M(\mathbf{u}_t, \mathbf{v}) + t\mathcal{A}_S(\mathbf{u}_t, \mathbf{v}) + t^3 \mathcal{A}_B b(\mathbf{u}_t, \mathbf{v}) = \omega_t^2 \mathcal{M}(t)(\mathbf{u}_t, \mathbf{v}) \quad \forall \, \mathbf{v} \in V \ ,
\end{cases}
\tag{3}
$$

160

where the different energy components are given by

$$t^3 \mathcal{A}_B(\mathbf{u}, \mathbf{u}) = t^3 \cdot \int_\Omega \{\nu(\kappa_{11}(\mathbf{u}) + \kappa_{22}(\mathbf{u}))^2 + (1 - \nu) \sum_{i,j=1}^2 \kappa_{ij}(\mathbf{u})^2\} A_1 A_2 \, d\alpha_1 d\alpha_2$$

$$t \mathcal{A}_M(\mathbf{u}, \mathbf{u}) = t \cdot 12 \int_\Omega \{\nu(\beta_{11}(\mathbf{u}) + \beta_{22}(\mathbf{u}))^2 + (1 - \nu) \sum_{i,j=1}^2 \beta_{ij}(\mathbf{u})^2\} A_1 A_2 \, d\alpha_1 d\alpha_2$$

$$t \mathcal{A}_S(\mathbf{u}, \mathbf{u}) = t \cdot 6(1 - \nu)\lambda \int_\Omega \{\rho_1(\mathbf{u})^2 + \rho_2(\mathbf{u})^2\} A_1 A_2 \, d\alpha_1 d\alpha_2,$$

where further κ_{ij}, β_{ij} and ρ_i denote the bending, membrane and transverse shear strains, respectively, ν is the Poisson number of the material, and λ is the so-called shear correction factor. The integrals are calculated over the midsurface Ω of the shell which is parametrized by the (generally curvilinear) principal curvature coordinates α_i. The metric of the shell surface is given by the Lamé parameters A_i.

Remark 2.1. *Here we use the scaled thickness $t = d/L$ where L denotes the characteristic length of the shell. In the cases under study L equals unity.*

Remark 2.2. *In our model we have omitted the constant factor $D = \frac{E}{12(1-\nu^2)}$, where E is the Young modulus of the material, from the energy expressions. It is included in the numerical tests below.*

The strains are taken to be [5]

$$\beta_{11} = \frac{1}{A_1} \frac{\partial u}{\partial \alpha_1} + \frac{v}{A_1 A_2} \frac{\partial A_1}{\partial \alpha_2} + \frac{w}{R_1}, \quad \beta_{22} = \frac{1}{A_2} \frac{\partial v}{\partial \alpha_2} + \frac{u}{A_1 A_2} \frac{\partial A_2}{\partial \alpha_1} + \frac{w}{R_2}$$

$$\beta_{12} = \frac{1}{2}\Big(\frac{1}{A_1} \frac{\partial v}{\partial \alpha_1} + \frac{1}{A_2} \frac{\partial u}{\partial \alpha_2} - \frac{u}{A_1 A_2} \frac{\partial A_1}{\partial \alpha_2} - \frac{v}{A_1 A_2} \frac{\partial A_2}{\partial \alpha_1}\Big) = \beta_{21}$$

$$\kappa_{11} = \frac{1}{A_1} \frac{\partial \theta}{\partial \alpha_1} + \frac{\psi}{A_1 A_2} \frac{\partial A_2}{\partial \alpha_2}, \quad \kappa_{22} = \frac{1}{A_2} \frac{\partial \psi}{\partial \alpha_2} + \frac{\theta}{A_1 A_2} \frac{\partial A_2}{\partial \alpha_1}$$

$$\kappa_{12} = \frac{1}{2}\Big[\frac{1}{A_1} \frac{\partial \psi}{\partial \alpha_1} + \frac{1}{A_2} \frac{\partial \theta}{\partial \alpha_2} - \frac{\theta}{A_1 A_2} \frac{\partial A_1}{\partial \alpha_2} - \frac{\psi}{A_1 A_2} \frac{\partial A_2}{\partial \alpha_1}$$
$$- \frac{1}{R_1}\Big(\frac{1}{A_2} \frac{\partial u}{\partial \alpha_2} - \frac{v}{A_1 A_2} \frac{\partial A_2}{\partial \alpha_1}\Big)$$
$$- \frac{1}{R_2}\Big(\frac{1}{A_1} \frac{\partial v}{\partial \alpha_1} - \frac{u}{A_1 A_2} \frac{\partial A_1}{\partial \alpha_2}\Big)\Big] = \kappa_{21}$$

and

$$\rho_1 = \frac{1}{A_1} \frac{\partial w}{\partial \alpha_1} - \frac{u}{R_1} - \theta, \quad \rho_2 = \frac{1}{A_2} \frac{\partial w}{\partial \alpha_2} - \frac{v}{R_2} - \psi$$

where the R_i's are the principal radii of curvature of the shell at the point (α_1, α_2).

3 Results of asymptotic analysis

We now present some theoretical results on the asymptotic behavior of the shell eigenvalue problem (3) for the three types of geometry presented above. In particular, we are interested in the three functions:

$$t \longmapsto \omega_t^2 \ ;$$
$$t \longmapsto K_t \ ;$$
$$t \longmapsto R(t, \mathbf{u}_t) := \frac{t^3 \mathcal{A}_B(\mathbf{u}_t, \mathbf{u}_t)}{t^3 \mathcal{A}_B(\mathbf{u}_t, \mathbf{u}_t) + t \mathcal{A}_M(\mathbf{u}_t, \mathbf{u}_t) + t \mathcal{A}_S(\mathbf{u}_t, \mathbf{u}_t)} \ . \qquad (4)$$

We recall that ω_t^2 represents the smallest eigenvalue, and the corresponding eigenfrequency is ω_t. K_t denotes the mode or harmonic number associated with the smallest eigenvalue for a fixed value of t. For shells of revolution it is naturally always an integer valued function. $R(t, \mathbf{u}_t)$ represents the ratio of bending strain energy over total strain energy, associated with ω_t^2 and the corresponding eigenvector $\mathbf{u}_t$. We now need to introduce the following notation concerning the behavior of the function $t \to \omega_t^2$. When we write

$$\omega_t^2 \sim t^\alpha \qquad \alpha \in \mathbf{R} \ , \qquad (5)$$

we mean that

$$\forall \varepsilon > 0 \quad : \quad \lim_{t \to 0^+} \frac{\omega_t^2}{t^{\alpha - \varepsilon}} = 0 \quad \text{and} \quad \lim_{t \to 0^+} \frac{\omega_t^2}{t^{\alpha + \varepsilon}} = +\infty \ . \qquad (6)$$

Remark 3.1. *$\omega_t^2 \sim t^\alpha$ essentially means that t^α is the best power of t which asymptotically fits ω_t^2. In some sense, we are measuring the behavior of ω_t^2 by using the scale of the real power functions. We also notice that $\omega_t^2 \sim t^\alpha$ does not imply that ω_t^2 behaves exactly like t^α. For instance, we have $\omega_t^2 \sim t^\alpha$ for $\omega_t^2 = -t^\alpha \log t$, in accordance with (5)–(6).*

For fully clamped shells of revolution, the following results for ω_t^2 and $\lim_{t \to 0^+} R(t, \mathbf{u}_t)$ can be established applying the results of [2, 3]; the proof can be found in [4], the results for K_t are based on numerical results.

1 – *Zero Gaussian curvature shells [parabolic].* $\omega_t^2 \sim t$; $K_t \sim t^{-\frac{1}{4}}$; $\lim_{t \to 0^+} R(t, \mathbf{u}_t) = \frac{1}{2}$.

2 – *Positive Gaussian curvature shells [elliptic].* $\omega_t^2 \sim t^0$; $K_t \sim t^{-\frac{2}{5}}$; $\lim_{t \to 0^+} R(t, \mathbf{u}_t) = 0$.

3 – *Negative Gaussian curvature shells [hyperbolic].* $\omega_t^2 \sim t^{2/3}$; $K_t \sim t^{-\frac{1}{3}}$; $\lim_{t \to 0^+} R(t, \mathbf{u}_t) = \frac{1}{3}$.

4 Benchmarks

In what follows three families of shells will be investigated. In order to simplify the treatment we define the shell geometry in terms of one single parameter $H > 0$, assuming $\alpha = -\beta$ and $\beta = H$.

(a) Parabolic: $\omega_t^2 \sim t$ (b) Elliptic: $\omega_t^2 \sim t^0$ (c) Hyperbolic: $\omega_t^2 \sim t^{2/3}$ (d) Together

Figure 2: ω^2 vs t

(a) Parabolic: $K_t \sim t^{-\frac{1}{4}}$ (b) Elliptic: $K_t \sim t^{-\frac{2}{5}}$ (c) Hyperbolic: $K_t \sim t^{-\frac{1}{3}}$ (d) Together
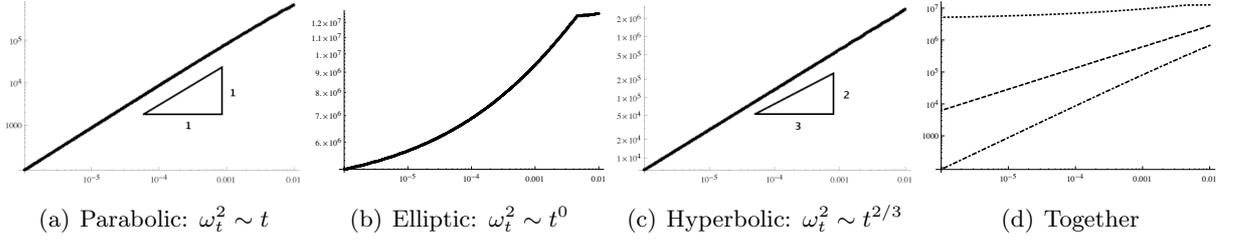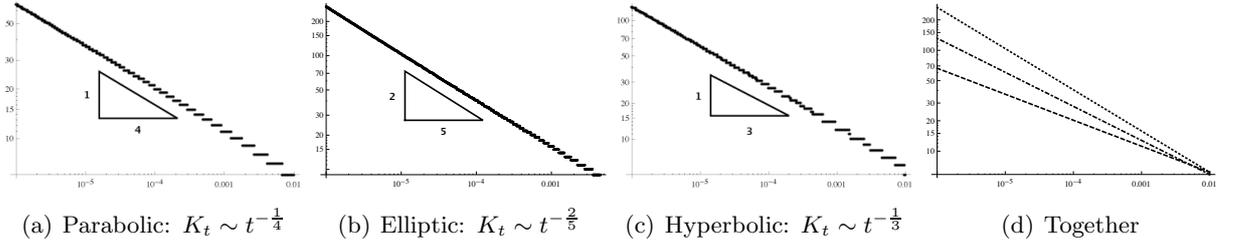
Figure 3: K vs t

1 – *Zero Gaussian curvature shells [parabolic].* $H = 1 , \quad f(x) = 1 .$

2 – *Positive Gaussian curvature shells [elliptic].* $H = 0.892668 , \quad f(x) = 1 - \frac{x^2}{2} .$

3 – *Negative Gaussian curvature shells [hyperbolic].* $H = 0.892668 , \quad f(x) = 1 + \frac{x^2}{2} .$

We impose clamped boundary conditions at both ends $x_1 = \alpha$ and $x_1 = \beta$ and compute over a set of 1601 thicknesses equally spaced on logarithmic scale chosen over the interval $[10^{-6}, 10^{-2}]$. Material constants adopted for all simulations are: $E = 2.069 \times 10^{11}$ MPa, $\nu = 0.3$, $\lambda = 5/6$ and $\rho = 7868$ Kg/m^3.

By fixing K we can reduce the original eigenproblem to a one-dimensional one. The minimal values of K_t and ω_t^2 can be found by solving over a range of K for a fixed t. We have used the p-extension of the finite element method in the overkill fashion in every individual computation.

The numerical results are summarized in Figures 2, 3, and 4. We have an excellent agreement with the theoretical results. It is interesting to note that in the elliptic case the order is correct in the sense of our notation, even though no clear trend can be seen. Numerical investigation of the slope strongly suggest behavior $\sim t^\alpha \log t$.

Note that event hough K_t is integer valued, the transition from one mode to another always occurs at a double eigenvalue. Thus, both the preceding and following dominant modes are present simultaneously at the critical value of t. In other words, an outside observer does not see any large displacements as the mode changes.

163

(a) Parabolic: $\lim R(t, \mathbf{u}_t) = \frac{1}{2}$ (b) Elliptic: $\lim R(t, \mathbf{u}_t) = 0$ (c) Hyperbolic: $\lim R(t, \mathbf{u}_t) = \frac{1}{3}$
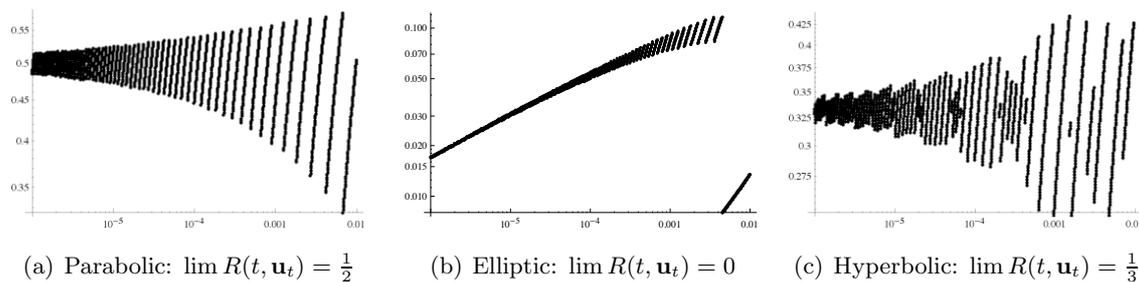
Figure 4: R vs t

5 Future work and conclusions

For the test cases considered here we have an agreement between theory and numerical experiments. For shells of full revolution it appears that the boundary layers do not play a prominent role in the eigenmode associated with the smallest eigenvalue.

We have experimental evidence that this is not the case for pieces of shells of revolution. For rotor blades or pieces of parabolic shells, e.g. $\Omega = [0,1] \times [-\omega, \omega]$, with kinematic constraints at $\alpha_1 = 0$ and/or $\alpha_1 = 1$, the boundary layers along the characteristic line at $\alpha_2 = \pm\omega$ (scale $\sim t^{\frac{1}{4}}$) must be taken into account. Any analysis of the asymptotic behavior must address the boundary layer resolution and separate it from the underlaying dominant mode.

References

[1] D. Chapelle and K.J. Bathe. *The Finite Element Analysis of Shells.* Springer–Verlag, 2003.

[2] L. Beirão da Veiga and C. Lovadina. An interpolation theory approach to shell eigenvalue problems. *To appear.*

[3] L. Beirão da Veiga and C. Lovadina. Asymptotics of shell eigenvalue problems. *C.R. Acad. Sci. Paris*, (9):707–710, 2006.

[4] H. Hakula E. Artioli, L. Beirão da Veiga and C. Lovadina. Free vibrations for some koiter shells of revolution. *submitted to Appl. Math. Letters.*

[5] Mika Malinen. On the classical shell model underlying bilinear degenerated shell finite elements: general shell geometry. *Internat. J. Numer. Methods Engrg.*, 55:629–652, 2002.

The deformation of cylindrical shells under generalized loads using a mixed formulation

Luísa R. Madureira[*]   F. J. Queirós Melo[†]

Abstract

A mixed formulation is used to obtain the solution for the ovalization of a thin walled straight pipe subjected to edge forces and/or prescribed displacements or localized loads acting on the shell surface. The solution for the displacement field is obtained from calculus of variations leading to the stationarity of functional defined displacements. The unknown functions to be evaluated are the displacements along the longitudinal direction of the pipe and they are related to a set of Fourier trigonometric expansions. The total energy resulting from the pipe deformation process is evaluated and then, by a variational procedure, its stationarity is imposed via a set of differential equations. The analytical solution is reached and various boundary conditions are considered in the examples in order to obtain the deformations caused by:

- prescribed displacements on the surface;
- prescribed displacements on the edges.

The solutions are evaluated for the system of differential equations and two and three-dimensional graphics are presented. Results are compared with other published work where computational or experimental approach is used.

Keywords: hybrid formulation, Fourier series, system of differential equations, eigenvalues, boundary conditions.

1   Introduction

Cylindrical shells represent a particular geometry of shells of revolution. Their application is quite extended as can be seen from process engineering (piping and pressure vessels) or automotive industry (storage tanks for road transport of fluids) or in aeronautics (aircraft body and engine nacelles). The efforts that are likely to occur are mainly the following:

---

[*]Departamento de Engenharia Mecânica e Gestão Industrial, Faculdade de Engenharia da Universidade do Porto, Portugal. E-mail: `luisa.madureira@fe.up.pt`.

[†]Departamento de Engenharia Mecânica, Universidade de Aveiro, Portugal.

- Distributed loads modeling the effect of internal or external pressure;

- Concentrated loads generated from the action of local attachments like pipes connected to the cylindrical shell developing radial loads or moments;

The present solution is obtained by an algorithm based on a variational method using a semi-analytic formulation, where unknown functions depending on the shell axial x-coordinate are associated with trigonometric terms along the circumferential direction. The method shows a good accuracy defining the deformed shape of the cylinder when subjected either to local loads or prescribed edge displacements. The solution is obtained after integration of a system of ordinary differential equations in the unknowns referring the auxiliary x-coordinate functions together with trigonometric terms.

2 Formulation of the problem

The total energy involved in the shell deformation process is:

$$
\begin{aligned}
E &= \frac{1}{2} \int \int A\varepsilon_{xx}^2 + S\gamma_{x\theta}^2 + Dk_{xx}^2 + Dk_{\theta\theta}^2 + \frac{D(1-\nu)}{2}k_{x\theta}^2 \\
&= \frac{1}{2} \sum_{i=2,3,4} \left[A\left(b_i' \cos i\theta\right)^2 + S\left(-\frac{i}{r}b_i - \frac{a_i'}{i}\right)^2 (\sin i\theta)^2 \right. \\
&\quad \left. + \frac{1}{r^4}D(-i^2+1)^2 \left(a_i \cos i\theta\right)^2 + \frac{D(1-\nu)}{2}\frac{\left(-2i+\frac{1}{i}\right)^2}{r^2}a_i'^2 \sin^2 i\theta + Da_i''^2 \cos^2 i\theta \right]
\end{aligned}
\tag{1}
$$

with

$$
\begin{aligned}
k_{xx} &= (a_i'' \cos i\theta) & \text{(the shell axial curvature)} \\
k_{x\theta} &= \tfrac{1}{r}\left(-2i+\tfrac{1}{i}\right)a_i' \sin i\theta & \text{(the shell surface twist)} \\
k_{\theta\theta} &= \tfrac{1}{r^2}\left(\left(-i^2+1\right)a_i \cos i\theta\right) & \text{(the circumferential curvature)} \\
\varepsilon_{xx} &= (b_i' \cos i\theta) & \text{(the axial membrane strain)} \\
\varepsilon_{\theta\theta} &= 0 & \text{(the circumferential membrane strain;} \\
& & \text{is zero as the pipe is circumferentially inextensible)} \\
\gamma_{x\theta} &= \left(-\tfrac{i}{r}b_i - \tfrac{a_i'}{i}\right)\sin i\theta & \text{(the membrane shear strain)}
\end{aligned}
\tag{2}
$$

The stationarity condition is obtained by minimizing the previous functional U separately for each of the unknowns a_i, b_i and the following differential equations are obtained:

$$\pi r D a_2^{iv} = -\frac{9\pi D}{r^3}a_2 + \pi S b_2' + \left(\frac{\pi r S}{2^2} + \left(-2\times 2 + \frac{1}{2}\right)^2 \frac{\pi D(1-\nu)}{2r}\right)a_2''$$

$$\pi r D a_3^{iv} = -\frac{64\pi D}{r^3}a_3 + \pi S b_3' + \left(\frac{\pi r S}{3^2} + \left(-2\times 3 + \frac{1}{3}\right)^2 \frac{\pi D(1-\nu)}{2r}\right)a_3''$$

$$\pi r D a_4^{iv} = -\frac{15\pi D}{r^3}a_4 + \pi S b_4' + \left(\frac{\pi r S}{4^2} + \left(-2\times 4 + \frac{1}{4}\right)^2 \frac{\pi D(1-\nu)}{2r}\right)a_4'' \qquad (3)$$

$$\pi r A b_2'' = \left(\frac{4\pi S}{r}\right)b_2 + \pi S a_2'$$

$$\pi r A b_3'' = \left(\frac{9\pi S}{r}\right)b_3 + \pi S a_3'$$

$$\pi r A b_4'' = \left(\frac{16\pi S}{r}\right)b_4 + \pi S a_4'$$

The final system of differential equations to be solved is:

$$a_2^{iv} = -\frac{9}{r^4}a_2 + \frac{S}{rD}b_2' + \left(\frac{S}{2^2 D} + \left(-2\times 2 + \frac{1}{2}\right)^2 \frac{1-\nu}{2r^2}\right)a_2''$$

$$a_3^{iv} = -\frac{64}{r^4}a_3 + \frac{S}{rD}b_3' + \left(\frac{S}{3^2 D} + \left(-2\times 3 + \frac{1}{3}\right)^2 \frac{1-\nu}{2r^2}\right)a_3''$$

$$a_4^{iv} = -\frac{15}{r^4}a_4 + \frac{S}{rD}b_4' + \left(\frac{S}{4^2 D} + \left(-2\times 4 + \frac{1}{4}\right)^2 \frac{1-\nu}{2r^2}\right)a_4'' \qquad (4)$$

$$b_2'' = \left(\frac{4S}{r^2 A}\right)b_2 + \frac{S}{rA}a_2'$$

$$b_3'' = \left(\frac{9S}{r^2 A}\right)b_3 + \frac{S}{rA}a_3'$$

$$b_4'' = \left(\frac{16S}{r^2 A}\right)b_4 + \frac{S}{rA}a_4'$$

3   Examples and numerical results

A simple example consists of prescribing a surface ovalization at a section equidistant of both edges of a carbon steel pipe.

The pipe dimensions and material properties are:

$$r = 0.05\text{m}$$
$$L = 2\text{m}$$
$$h = 0.001\text{m}$$
$$E = 210\text{GPa}$$

The solution is obtained for half of the pipe, as symmetry is considered, and the boundary conditions are:

$$a_i(L/2) = -0.001\text{m} \qquad a_i'(L/2) = 0 \qquad b_i(L/2) = 0 \qquad a_i''(L) = 0 \qquad b_i'(L) = 0 \qquad b_i''(L) = 0$$
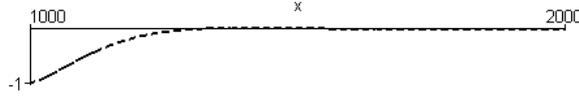
Figure 1: Propagation of ovalization along x: half-pipe symmetry.
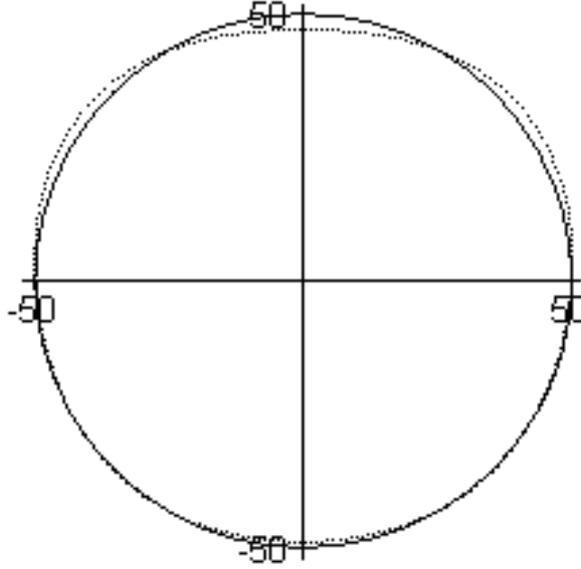
Figure 2: Ovalization of cross section at $x = 1000$mm for a pipe of length $L = 2000$mm.

The second example for the same pipe consists of prescribing an ovalization of -0.001m at both edges. However, the edges are assumed to warp and no edge bending moment is considered. The boundary conditions for these physical settings are:

$$a_i(0) = -0.001\text{m} \quad \text{and} \quad a_i(L) = -0.001\text{m} \qquad a_i''(L) = 0 \qquad a_i''(0) = 0 \qquad b_i'(L) = 0 \qquad b_i'(0) = 0$$

In the third example a shorter pipe with a free edge at $x = 0$ and a flange ring at $x = L$ is considered. The boundary conditions are:

$$a''(-L) = a'''(-L) = 0a(0) = 1 \qquad a'(0) = 0 \qquad a(L) = 0 \qquad a'(L) = 0$$

The propagation of ovalization along x is represented in Figure 5.

4 Conclusions

The solution presented has shown good accuracy and a quite low computational effort, once it deals with a closed form algorithm for a specified number of trigonometric terms. An analytic
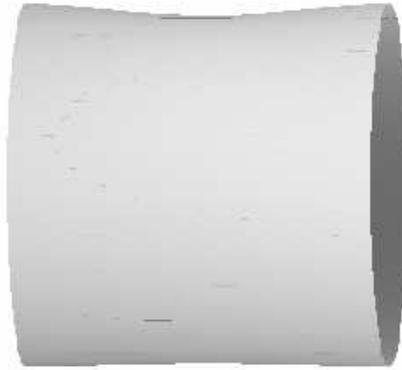
Figure 3: Detail of three dimensional deformation of the pipe.
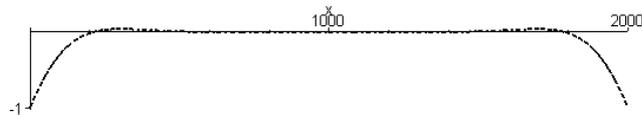
Figure 4: propagation of ovalization $(a_2(x))$ along x.

Figure 5: Propagation of ovalization at $\theta = 0$ along x for a pipe of length $L = 0.4$m.

solution using several different boundary conditions is reached in a simple way. Even for only one trigonometric term results are as good as those obtained with the finite element method.

References

[1] W. Anwen and Z. Wei. Asymptotic solution for buckling of toroidal shells. *Int. J. Press. Vessels & Piping*, 46:61–72, 1991.

[2] W. Flügge. *Thin Elastic Shells*. Springer, 2nd edition, 1973.

[3] H. Kraus. *Thin Elastic Shells*. Wiley, New York, 1967.

[4] A. Millard and R. Roche. Elementary solutions for the propagation of ovalization along straight pipes and elbows. *Int. Journ. of Press. Vessels & Piping*, 16(1):101–129, 1984.

[5] C. Wylie and L. Barret. *Advanced Engineering Mathematics*. McGraw-Hill, 6th edition, 1995.

Approximation of shell layers using bilinear elements

Antti H. Niemi*

Abstract

We study numerical locking effects in the finite element approximation of shell layers. The focus is on flat bilinear elements that are perhaps the most widely used in practical computations, but also some closely related curved elements are considered. Our approach is based on simplified reformulations of the original 3D elements in the context of a classical shallow shell model, where the deformation of the shell is described in terms of the three displacements of the middle surface and two dimensionless rotations related to transverse shear deformations. Within that model we compute finite element approximations of layers generated by smooth line segments such as the boundary line. Our numerical results show that approximation failure occurs when flat elements are used on anisotropically refined meshes and that the accuracy depends substantially on the specific geometry of the shell.

1   Introduction

In this work we bring to light the numerical locking effects associated to the bilinear finite element approximation of shell layers using flat elements on anisotropically refined meshes. In addition to a good understanding of the numerical properties of the bilinear degenerated 3D FEM, we need knowledge of the asymptotic nature of shell deformations beyond the usual membrane and inextensional theories. In the classical literature on the theory of shells (see [11]) the layers are known as 'edge effects', but their treatment there is rather limited from the modern perspective. Nevertheless, a more systematic asymptotic and numerical analysis of these phenomena has restarted fairly recently e.g. in [13, 3, 4, 10].

Here we consider layers generated by a smooth line segment S and concentrate on the following main layer modes in shell deformations:

1. The line layer that decays in the length scale $L \sim \sqrt{Rt}$ from S, where t is the thickness of the shell and R measures the total curvature of the shell. This layer is possible in all shell geometries.

*Institute of Mathematics, Helsinki University of Technology, P.O. Box 1100, FIN-02015 HUT, Finland. E-mail:`antti.h.niemi@tkk.fi`.

2. The line layer that decays in the length scale $L \sim \sqrt[3]{R^2 t}$ from S. This arises in hyperbolic shell geometry when S is a characteristic line of the middle surface of the shell.

3. The line layer that decays in the length scale $L \sim \sqrt[4]{R^3 t}$ from S. This layer is possible when S is the characteristic line of a parabolic middle surface.

Since the shell deviates only slightly from a plane in the decay length scales, the shallow shell model derived by Pitkäranta et al. in [13] serves as a natural starting point for our study. By treating S as a straight line within that model, we introduce an analytically solvable model problem where the shell is under a concentrated line load that varies smoothly along S. Our numerical models are based on the standard bilinear scheme for the 2D shell model but the strain expressions are modified in accordance with the geometric and physical assumptions of the original 3D elements.

Our computations show that the use of flat elements results in approximation failure in all of the above cases when anisotropically refined meshes are used, and that the case of hyperbolic degeneration (Case 2) appears to be the worst in this sense. On the other hand, geometrically compatible elements based on the shallow shell theory approach appear to be more efficient here provided that the membrane strains are modified very carefully. The conclusion is that when approximating shell layers, the use of flat elements, or *faceting*, is not the best choice among the possible *numerical tricks* aiming at avoiding locking.

Our computational observations can be supported also by theoretical error estimates (see [8]) showing that for flat elements the relative error in the energy norm depends on the dimensionless thickness t/R and becomes severely amplified as $t/R \to 0$. More precisely, the optimal error bound for the lowest-order FEM is magnified by factors $(R/t)^{1/2}$, $(R/t)^{2/3}$ and $(R/t)^{1/4}$ in Cases 1, 2 and 3, respectively. Note that in the conforming, geometrically compatible finite element models the corresponding factors arising from membrane locking effects would be 1, $(R/t)^{1/3}$ and $(R/t)^{1/2}$, see [13]. On the other hand, our study shows that a clever reduction of the geometrically compatible membrane strains can avoid the error amplification completely when the mesh is rectangular and aligned with the layer.

2   Model problem

In the dimensionally reduced Reissner-Naghdi model for a shell of thickness t, the strain energy is given by the quadratic functional

$$\mathcal{A}(\mathbf{u}, \mathbf{u}) = \mathcal{A}_m(\mathbf{u}, \mathbf{u}) + \mathcal{A}_s(\mathbf{u}, \mathbf{u}) + \mathcal{A}_b(\mathbf{u}, \mathbf{u}), \tag{1}$$

where $\mathbf{u} = (u, v, w, \theta, \psi)$ is the generalized displacement field defined on the middle surface Γ and the scaled membrane, transverse shear and bending energy functionals $\mathcal{A}_m, \mathcal{A}_s, \mathcal{A}_b$ are

defined by

$$\mathcal{A}_m(\mathbf{u}, \mathbf{u}) = \int_\Gamma \left[\nu(\beta_{11} + \beta_{22})^2 + (1 - \nu)(\beta_{11}^2 + 2\beta_{12}^2 + \beta_{22}^2) \right] \mathrm{d}\Gamma,$$

$$\mathcal{A}_s(\mathbf{u}, \mathbf{u}) = \frac{1 - \nu}{2} \int_\Gamma (\rho_1^2 + \rho_2^2) \, \mathrm{d}\Gamma, \tag{2}$$

$$\mathcal{A}_b(\mathbf{u}, \mathbf{u}) = \frac{t^2}{12} \int_\Gamma \left[\nu(\kappa_{11} + \kappa_{22})^2 + (1 - \nu)(\kappa_{11}^2 + 2\kappa_{12}^2 + \kappa_{22}^2) \right] \mathrm{d}\Gamma.$$

Here ν is the Poisson ratio of the material (isotropic & homogeneous) and β_{ij}, ρ_i, and κ_{ij} represent the membrane, transverse shear, and bending strains, respectively.

We consider as a model problem a shallow shell such that the middle surface occupies a planar domain $\Gamma = (-1, 1) \times (0, 2\pi)$ in the rectilinear coordinate system (x, y) and $\mathrm{d}\Gamma = \mathrm{d}x\,\mathrm{d}y$ in (2). The strains are associated to the displacement field $\mathbf{u}$ then as follows (see [13]):

$$\beta_{11} = \frac{\partial u}{\partial x} + aw, \quad \beta_{22} = \frac{\partial v}{\partial y} + bw, \quad \beta_{12} = \frac{1}{2}\left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right) + cw,$$

$$\rho_1 = \theta - \frac{\partial w}{\partial x}, \qquad \rho_2 = \psi - \frac{\partial w}{\partial y}, \tag{3}$$

$$\kappa_{11} = \frac{\partial \theta}{\partial x}, \qquad \kappa_{22} = \frac{\partial \psi}{\partial y}, \qquad \kappa_{12} = \frac{1}{2}\left(\frac{\partial \theta}{\partial y} + \frac{\partial \psi}{\partial x} \right).$$

Here the constants $a = b_{11}$, $b = b_{22}$, and $c = b_{12} = b_{21}$ are the components of the curvature tensor of the shell so that the shell is elliptic when $ab - c^2 > 0$, parabolic when $ab - c^2 = 0$, and hyperbolic when $ab - c^2 < 0$.

In the benchmark problem the shell is assumed to be loaded by a normal concentrated line load

$$f(x, y) = \delta(x) \cos(ky), \tag{4}$$

where the wave number k is an integer representing the load variation along the line $x = 0$. The corresponding load functional is defined by

$$\mathcal{L}(\mathbf{u}) = \int_0^{2\pi} \cos(ky) w(0, y) \, \mathrm{d}y$$

and the resulting deformation of the shell is obtained by minimizing the total energy

$$\mathcal{F}(\mathbf{u}) = \frac{1}{2}\mathcal{A}(\mathbf{u}, \mathbf{u}) - \mathcal{L}(\mathbf{u}) \tag{5}$$

in the energy space $\mathcal{U}$, where we impose the kinematic constraints $u = v = w = \theta = \psi = 0$ at $x = \pm 1$ along with periodic boundary conditions at $y = 0, 2\pi$.

Concerning the geometry of the shell, the three main cases described above are obtained with the following choices:

$$\begin{aligned} \text{Case 1:} \quad & a = c = 0, \quad b \neq 0, \\ \text{Case 2:} \quad & a = b = 0, \quad c \neq 0, \\ \text{Case 3:} \quad & b = c = 0, \quad a \neq 0. \end{aligned}$$

173

Case 1 represents the most common occasion where the curvature along the line $x = 0$ does not vanish. Cases 2 and 3, respectively, are examples of hyperbolic and parabolic degeneration, where the curvature along the line $x = 0$ is zero, i.e., the line $x = 0$ is a characteristic line of the middle surface.

Due to the specific shape of the load (4), we have a one-dimensional problem only. The solution is of the form $\mathbf{u}(x, y) = \mathbf{U}(x) \otimes \boldsymbol{\varphi}_k(y)$, where

$$\boldsymbol{\varphi}_k(y) = \begin{cases} (\cos(ky), \sin(ky), \cos(ky), \cos(ky), \sin(ky)) & \text{(Cases 1,3)}, \\ (\sin(ky), \cos(ky), \cos(ky), \cos(ky), \sin(ky)) & \text{(Case 2)}. \end{cases} \tag{6}$$

and $\mathbf{U}$ is of the *boundary layer form* (see [8])

$$\mathbf{U}(x) \sim e^{-\alpha x/L} \left(\mathbf{A} \cos \frac{\beta x}{L} + \mathbf{B} \sin \frac{\beta x}{L} \right), \quad \alpha, \beta = \mathcal{O}(1), \quad \alpha, x > 0, \tag{7}$$

where L is the characteristic length scale of exponential decay. Assuming that $a, b, c, k \sim R^{-1}$, we will have $L \sim \sqrt[n]{R^{n-1}t}$, where $n = 2$ in Case 1, $n = 3$ in Case 2, and $n = 4$ in Case 3.

3 Benchmark computations

In this section, the model problem is solved approximately by the finite element method without exploiting the trigonometric variation of the solution. As in [12], the computational domain is restricted to $\Gamma = (0, 1) \times (0, \pi/4)$ by imposing symmetry/antisymmetry boundary conditions along the lines $x = 0$, $y = 0$ and $y = \pi/4$. The finite element models are based on rectangular subdivisions of Γ with piecewise uniform $(M + Q) \times N$ grids, see Figure 1. Here M denotes the number of mesh intervals in the x-direction on the zone $0 < x < 4L$ where the layer solutions have a significant amplitude.
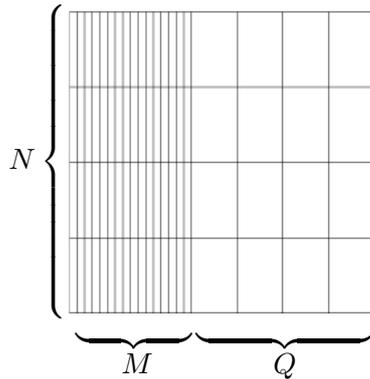

Figure 1: An anisotropically refined rectangular mesh used in numerical computations.

In the following we approximate the five displacement components independently by bi-linear elements and modify the strain energy functional numerically in order to avoid locking

effects. Concerning the transverse shear strains, the wide engineering literature supports the modifications (see e.g. [5, 2, 14, 1])

$$\rho_1 \hookrightarrow \Pi_x \rho_1, \quad \rho_2 \hookrightarrow \Pi_y \rho_2, \tag{8}$$

where Π_x and Π_y are elementwise averaging operators in the coordinate direction indicated by the subscript. On the other hand, the local replacement of the shell middle surface by its isoparametric bilinear approximation corresponds to the modification of the membrane strains as follows (see [7]):

$$
\begin{aligned}
\beta_{11} &\hookrightarrow \frac{\partial u}{\partial x} + a \Pi_x w + c R_y w, \\
\beta_{22} &\hookrightarrow \frac{\partial v}{\partial y} + b \Pi_y w + c R_x w, \\
\beta_{12} &\hookrightarrow \frac{1}{2} \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} + c \left(\Pi_x w + \Pi_y w \right) + a R_x w + b R_y w \right).
\end{aligned}
\tag{9}
$$

Here Π_x, Π_y are the averaging operators defined above whereas the interpretation of R_x, R_y is more subtle. The analysis in [7] anyway shows that if the displacement field is inextensional with $\beta_{11} = \beta_{22} = 0$, then the modification (9) amounts to averaging the membrane strains as

$$\beta_{11} \hookrightarrow \Pi_x \beta_{11}, \quad \beta_{22} \hookrightarrow \Pi_y \beta_{22}, \quad \beta_{12} \hookrightarrow \Pi_{xy} \beta_{12}, \tag{10}$$

where Π_{xy} is the elementwise averaging operator.

Concerning other kind of deformation states and especially boundary layers, it appears that the modification (10) is a rather favorable interpretation of the flat element procedure (9), see [6]. In the following, we analyze both formulations and label the element as MITC4-F when using the modification (9) and as MITC4-S when using the modification (10), cf. [10, 9]. In addition we consider a variant of these formulations, labeled as MITC4-R. This was introduced in [9] as a reformulation of MITC4-S so as to allow more general quadrilateral element shapes.

We turn now to numerical experiments by choosing $k = 2$, $\nu = 1/3$, $t = 1/1000$ and setting the non-vanishing curvature parameter to unity in each case. Figures 2–4 show the transverse deflection profile on the zone $0 < x < 4L$ for the three benchmark cases, as computed with the different bilinear elements with $M = 16$ and $N = Q = 4$. The analytical reference solution is visualized by a solid line in all cases.

In Case 1, the performance of MITC4-F and MITC4-R is about equal but clearly inferior to MITC4-S which performs fairly well. In Case 2, the flat element fails almost completely. MITC4-R performs better but still gives a poor result compared with MITC4-S whose error is negligible. Finally in Case 3, all elements give good results but MITC4-S is still slightly ahead of MITC4-F and MITC4-R.
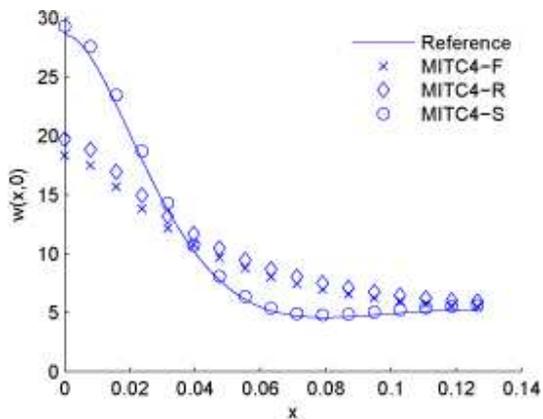
Figure 2: Transverse deflection profile using different elements: Case 1
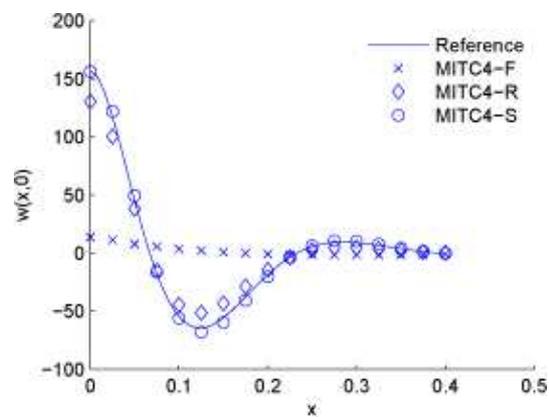

Figure 3: Transverse deflection profile using different elements: Case 2
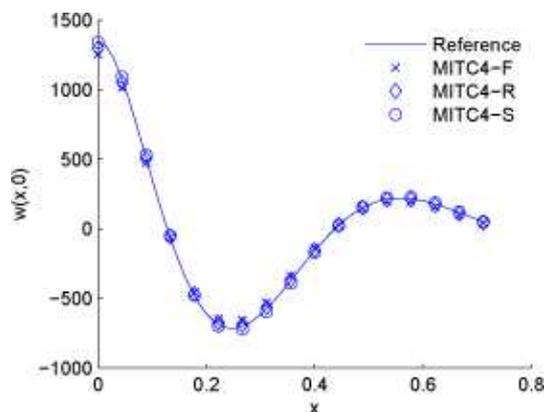
Figure 4: Transverse deflection profile using different elements: Case 3

References

[1] Klaus-Jürgen Bathe and Eduardo N. Dvorkin. A four-node plate bending element based on Mindlin/Reissner plate theory and a mixed interpolation. *Int. J. Numer. Meth. Engng*, 21:367–383, 1985.

[2] T. J. R. Hughes and T. E. Tezduyar. Finite Elements Based Upon Mindlin Plate Theory With Particular Reference to the Four-Node Bilinear Isoparametric Element. *J. Appl. Mech.*, 48:587–596, 1981.

[3] P. Karamian-Surville, J. Sanchez-Hubert, and E. Sanchez-Palencia. Propagation of singularities and structure of layers in shells. hyperbolic case. *Comput. Struct.*, 80:747–768, 2002.

[4] L. Beirão da Veiga. Asymptotic energy behavior of two classical intermediate benchmark shell problems. *Math. Models Methods Appl. Sci. (M^3AS)*, 13:1279–1302, 2003.

[5] Richard H. MacNeal. A simple quadrilateral shell element. *Comput. Struct.*, 8:175–183, 1978.

[6] Mika Malinen. On the classical shell model underlying bilinear degenerated shell finite elements. *Int. J. Numer. Meth. Engng*, 52:389–416, 2001.

[7] Mika Malinen. On the classical shell model underlying bilinear degenerated shell finite elements: general shell geometry. *Int. J. Numer. Meth. Engng*, 55:629–652, 2002.

[8] Antti H. Niemi. Approximation of shell layers using bilinear elements on anisotropically refined rectangular meshes. *Preprint submitted to Comput. Methods Appl. Mech. Engrg*, 2007.

[9] Antti H. Niemi and Juhani Pitkäranta. Bilinear finite element for shells: isoparametric quadrilaterals. *Accepted for publication in Int. J. Numer. Meth. Engng*, DOI: 10.1002/nme.2252, 2007.

[10] Antti H. Niemi, Juhani Pitkäranta, and Harri Hakula. Benchmark computations on point-loaded shallow shells: Fourier vs. FEM. *Comput. Methods Appl. Mech. Engrg.*, 196:894–907, 2007.

[11] V. V. Novozhilov. *The Theory of Thin Shells*. P. Noordhoff, Ltd. Groningen, The Netherlands, 1959.

[12] Juhani Pitkäranta, Yrjö Leino, Otso Ovaskainen, and Jyrki Piila. Shell deformation states and the finite element method: a benchmark study of cylindrical shells. *Comput. Methods Appl. Mech. Engrg.*, 128:81–121, 1995.

[13] Juhani Pitkäranta, A.-M. Matache, and C. Schwab. Fourier mode analysis of layers in shallow shell deformations. *Comput. Methods Appl. Mech. Engrg.*, 190:2943–2975, 2001.

[14] G. Wempner, D. Talaslidis, and C.-M. Hwank. A simple and efficient approximation of shells via finite quadrilateral elements. *J. Appl. Mech.*, 49:115–120, 1982.

On reliable triangular MITC6 shell elements

Iria Paris Suarez*

Abstract

The main difficult when developing reliable shell finite elements is to reduce the locking phenomenon arising in bending dominated situations without destroying the ability of the procedure to accurately capture membrane dominated behaviors. A simple numerical patch-test aimed at detecting membrane locking and the existence of "membrane spurious modes" for MITC membrane tying schemes has been proposed in [5].

The quadrilateral elements of the MITC family have been shown to be quite reliable for general asymptotic behaviors [1], but an optimal triangular MITC shell finite element – more adequate for complex geometries – remains to be found. The 6-node MITC6a element – formulated by Bathe and Lee [6] – seems to be the most promising candidate of the MITC6 family, though it features parasitic modes that might considerably deteriorate the displacement graphs in membrane dominated frameworks as it is shown in [3]. A richer shear interpolation – MITC6rs – and a remedy based on a stabilized bilinear form by adding an unreduced shear term have been proposed in [3, 7]. In view of a detailed assessment, the use of some stabilization in the MITC6a formulation is advocated.

1   Introduction

In engineering practice shell structures are frequently encountered, hence shell finite elements are of utmost interest. As is now well established [4], one of the great challenges in the design of shell elements is to ensure reliability of the numerical procedures for all types of asymptotic behaviors. The main difficulty in this respect consists in circumventing the locking phenomena arising in bending dominated asymptotic behaviors without destroying the ability of the procedure to accurately represent membrane-dominated and mixed behaviors.

General isoparametric shell finite elements show a good behavior for membrane dominated shell problems, but they clearly lock in bending dominated frameworks independently of the degree of interpolation [1]. Some shell finite element procedures have been shown to be quite reliable for general asymptotic behaviors, and in particular the quadrilateral elements of the "MITC" (Mixed Interpolation of Tensorial Components) family, namely the MITC4, MITC9 and MITC16 shell elements, see [1, 2].

---

*Glaizer Group – Innovation Department, 32 Rue Guy Môquet, 92240 - Malakoff (France). Email: iria.paris@glaizer.com.

However, in practice complex geometries must be handled, and *triangular* elements are often more adequate – and even necessary – in this respect. The vertex numbering isotropy discussed in [6] must be ensured, and thus two additional strain fields in the hypotenuse direction of the right-angled triangle in the reference local coordinate system must be considered (refer to Figure 1).

Some triangular MITC6 shell elements have been proposed in [5, 6, 7], with detailed assessments leading to the conclusion that the MITC6a element – in particular – performs well in the various test problems considered, although some attenuated locking was observed through bending dominated problems, especially for very small values of the thickness.

2 MITC membrane spurious modes

We consider a shell of uniform thickness t and overall characteristic length L. We define the relative thickness $\varepsilon = t/L$. Throughout this paper, isotropic linearized elasticity is considered, and we express the general isoparametric shell variational formulation as

$$
\begin{aligned}
&\text{Find } \vec{U}_h \in \mathcal{V}_h \text{ such that}\\
&\varepsilon^3 A_b(\vec{U}_h, \vec{V}) + \varepsilon A_m(\vec{U}_h, \vec{V}) = F^\varepsilon(\vec{V}), \quad \forall \vec{V} \in \mathcal{V}_h
\end{aligned}
\tag{1}
$$

in the notation of [4], where $\vec{U}_h$ denotes the unknown finite element solution sought in the space $\mathcal{V}_h$ and consisting of a displacement field $(\vec{u}_h)$ and a rotation field $(\vec{\theta}_h)$ – with the corresponding arbitrary test function $\vec{V} = (\vec{v}, \vec{\eta})$ –, F^ε represents the external virtual work, and A_m denotes the scaled bilinear form corresponding to the discrete membrane and shear strains, whereas the scaled bilinear form A_b contains all other higher terms in ε, namely, essentially the bending energy. We henceforth need to clearly distinguish between membrane and shear energies, hence, we now denote by A_{me} and A_s the scaled membrane and shear bilinear form, respectively, so that

$$
A_m = A_{me} + A_s.
\tag{2}
$$

As is well known, the great challenge in the design of shell elements is to obtain a discrete space of inextensional displacements – which is characterized by vanishing membrane and shear strains – sufficiently rich in bending dominated situations, and null in membrane dominated cases [4]. The numerical locking phenomenon is directly related to the failure of the first condition, and classical P_k displacement-based elements suffer from this pathology regardless of the displacement interpolation order k (see [4] and references therein).

The essence of the MITC procedure is to avoid numerical locking by relaxing the constraint of vanishing membrane and shear strains, so that a larger discrete inextensional space is obtained. Though shell finite elements formulation clearly differs from plate elements, the shear treatment for plates suitably works for shells, and the condition of zero membrane

strains constitutes the main obstacle when developing effective shell finite elements (see [1, 4] and references therein).

Following the MITC approach, strains and displacements are separately interpolated, being both interpolations connected at midsurface *tying points* which must be suitably chosen. We express the corresponding MITC variational formulation as:

$$\text{Find } \vec{U}_h \in \mathcal{V}_h \text{ such that}$$
$$\varepsilon^3 A_b^h(\vec{U}_h, \vec{V}) + \varepsilon A_m^h(\vec{U}_h, \vec{V}) = F^\varepsilon(\vec{V}) \quad \forall \vec{V} \in \mathcal{V}_h \tag{3}$$

where $\varepsilon^3 A_b^h + \varepsilon A_m^h$ represents the bilinear form taking into account the mixed interpolation of strain components: the scaled bilinear form A_m^h contains the reduced membrane and shear strains, whereas the scaled bilinear form A_b^h essentially contains the reduced bending strains. We will henceforth denote by A_{me}^h and A_s^h the scaled reduced membrane and shear bilinear forms so that:

$$A_m^h = A_{me}^h + A_s^h. \tag{4}$$

We will call "spurious membrane mode" a non-zero discrete displacement field $\vec{v}$ such that

$$A_{me}^h(\vec{v}, \vec{v}) = 0 \tag{5}$$

A detailed study in [7] has shown that for some combinations of midsurface geometry and boundary conditions, the amplitude of these parasitic modes may dominate that of the underlying "correct" solution in membrane dominated situations, giving erroneous displacement graphs. Furthermore, the same study shows that these modes may substantially deteriorate the performance of the MITC element as they seem to be only controlled by the shear and bending energies introduced by the rotations, which is non-physical in a membrane dominated situation. In addition, we point out that these modes cannot be well detected by an error indicator using reduced strains – such as the s-norm – since they have near-zero membrane reduced strains, and error measures provided by the unreduced membrane and shear energies are necessary to assess the numerical solutions.

A "membrane patch-test" aimed at disclosing the reliability of MITC tying schemes has been developed [5]: it is based in a matrix-formulation inferred from the system of partial differential equations given by vanishing membrane strains. This test revealed some promising MITC6 membrane-tying schemes, which have been further assessed by membrane and bending dominated benchmarks in [5, 7], being the second order Raviart Thomas space used for shear interpolation (see Figure 1 on the right). These results show that the MITC6a element is a most promising candidate for engineering practice, though it features some non-physical membrane spurious modes.

Considering that membrane shell problems are, by far, the most frequently encountered situations in engineering practice, the presence of these parasitic modes is of concern for the applicability of the MITC6a element.

$$e_{qq} = \frac{e_{rr} + e_{ss}}{2} - e_{rs} \qquad\qquad e_{qz} = \frac{1}{\sqrt{2}}\left(e_{sz} - e_{rz}\right)$$
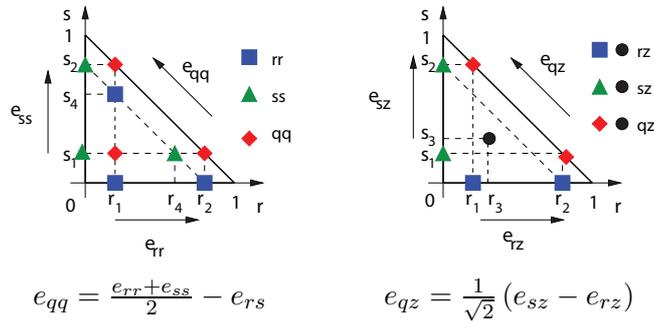
Figure 1: MITC6a tying points for in-plane strains (left) and transverse shear strains (right) in the reference element $\hat{T}$ ($r_1 = s_1 = \frac{1}{2} - \frac{1}{2\sqrt{3}}$, $r_2 = s_2 = \frac{1}{2} + \frac{1}{2\sqrt{3}}$, $r_3 = s_3 = \frac{1}{3}$, $r_4 = s_4 = \frac{1}{\sqrt{3}}$).

3   Towards improving the MITC6a element

It has been observed in [3, 7] that difficulties arise in membrane dominated situations due to the existence of discrete displacements such that

$$A_{me}(\vec{V}, \vec{V}) \gg A_{me}^h(\vec{V}, \vec{V}) , \quad \vec{V} \in \mathcal{V}_h \tag{6}$$

The main difficulty – in the attempt of curing this phenomenon – is that the same characterization (6) is essentially shared by all the discrete inextensional modes in bending dominated problems. In other words, in bending dominated situations almost all discrete inextensional displacements are expected to behave as in equation (6); this is exactly what distinguishes the unlocked MITC6a element from the "fully locking" P_2 displacement-based shell elements (see [7]). But condition (6) is essentially the only general characterization at our disposal.

It is observed in [7] that when membrane spurious modes are present in membrane dominated situations, there is an amplification by several orders of magnitude when comparing the reduced and unreduced shear energies, while the amplification is limited in the membrane energy. Furthermore, the same study shows that whenever the relative thickness is sufficiently small, the unreduced shear energy is dominant over the total unreduced energy. This phenomenon is in complete contrast with the P_2 displacement-based results, where the shear part of the energy in the "correct" solution should be negligeable. These observations indicate that the nature of the spurious part of the solution is mainly reflected in the shear energy. Moreover, it is well known that in typical bending dominated situations the main source of locking is the membrane energy, and not the shear energy [4]. The combination of these two arguments suggests to modify only the shear part of the MITC6a energy in order to reach our aforementioned goal.

The most natural modification is to propose an MITC6a formulation in which the parasitic modes are handled by adding an unreduced weighted shear part to the discrete method, and this modification must be made locally (see [3]). After a detailed analysis, the MITC6a

stabilized formulation proposed in [3] uses the bilinear form:

$$\varepsilon^3 A_b^h(\vec{U}_h, \vec{V}) + \varepsilon A_{me}^h(\vec{U}_h, \vec{V}) + \varepsilon \sum_{T \in \mathcal{T}_h} \left[\left(1 - C \frac{h_T^2}{L^2} \right) A_{s,T}^h(\vec{U}_h, \vec{V}) + C \frac{h_T^2}{L^2} A_{s,T}(\vec{U}_h, \vec{V}) \right] \quad (7)$$

for $\vec{V} \in \mathcal{V}_h$, where C is a positive scalar, T represents a general triangle of the mesh $\mathcal{T}_h$, h_T the diameter of T, and $A_{s,T}$ and $A_{s,T}^h$ respectively denote the local unreduced and reduced shear bilinear forms integrated over the element. Note that we include no dependence of C on ε in order to have a method which, in principle, provides an improved stability also in the membrane limit problem (as $\varepsilon \to 0$). Note further that the global energy consistency is not perturbed with the added term, since the corresponding fraction of the reduced shear energy is substracted.

Adopting a richer interpolation space for the transverse shear strains may provide another treatment. The MITC6 family uses essentially the second order rotated Raviart Thomas space for the transverse shear strains in the reference element, which has dimension 8. In analogy with the techniques adopted in plate elements, we now introduce the following richer space for the transverse shear strains on the reference element $\hat{T}$:

$$\hat{S}_h = \left\{ \rho \in [P_2(\hat{T})]^2 \ : \ \rho|_{\hat{e}} \cdot \tau_{\hat{e}} \in P_1(\hat{e}) \ \forall \hat{e} \text{ edge in } \partial \hat{T} \right\} , \quad (8)$$

where $\tau_{\hat{e}}$ represents the anticlockwise tangent unit vector to edge $\hat{e}$. The space $\hat{S}_h$ has dimension 9 and contains the second order rotated Raviart Thomas space. Therefore, as starting degrees of freedom for the $\hat{S}_h$ space, we select the same isotropic tying points as in the MITC6a transverse shear interpolation (see again Figure 1). As a ninth and last degree of freedom, and in order to keep the vertex numbering isotropy discussed in [6], we make the following construction: let $\hat{e}_i$, $i = 1, 2, 3$, denote the three edges of the reference element $\hat{T}$, and $\hat{m}_i$ the respective midpoints. Moreover, let $\hat{\eta}_i = \hat{\psi}_i / \|\hat{\psi}_i\|$, where $\hat{\psi}_i$ is the vector connecting each midpoint $\hat{m}_i$ with the opposite vertex. Then, the last degree of freedom R_9 for the $\hat{S}_h$ space is given by

$$R_9(e_{rz}, e_{sz}) = \sum_{i=1}^{3} e_{\hat{\eta}_i z}(m_i) . \quad (9)$$

The value of the transverse shear component $e_{\hat{\eta}_i z}$ can be clearly obtained from e_{rz}, e_{sz}. It is easy to check that the evaluation rule (9) is isotropic with respect to vertex permutations by construction. Moreover, together with the MITC6a tying evaluations in Figure 1, (9) constitutes a set of linearly independent and unisolvent degrees of freedom for $\hat{S}_h$. We therefore propose a new element, named as MITC6rs, which is similar to the MITC6a element, except for the transverse shear strain space and tying procedure which is the one introduced above.

4 Conclusion

A detailed assessment in [3, 7] shows that the MITC6rs and the stabilized formulation presented here constitute an effective cure for the MITC6a shell finite element in membrane dominated shell problems, whereas the MITC performance is not severely hindered for the bending dominated and penalized shell problems. Namely, *some locking* still occurs for small relative thickness and coarse meshes, but the stabilized results do not differ very much from the original MITC6a ones and are substantially improved as the mesh is refined. Furthermore, the membrane spurious modes are substantially filtered out for membrane dominated shell problems and almost optimal and uniform convergence is achieved.

From a practical point of view, two constants must be set to compute the proposed stabilized formulation: the characteristic length L and the stabilization coefficient C (without dimension). Of course, a characteristic length is somewhat arbitrary, and may be difficult to choose for complex geometries. Nevertheless, a detailed assessment in [3, 7] shows that the performance of the method is quite robust with respect to the stabilization coefficient C, hence also with respect to L.

References

[1] K.J. Bathe. *Finite Element Procedures.* Prentice Hall, Englewood Cliffs, 1996.

[2] K.J. Bathe, A. Iosilevich, and D. Chapelle. An evaluation of the MITC shell elements. *Comput. & Structures*, 75(1):1–30, 2000.

[3] L. Beirão da Veiga, D. Chapelle, and I. Paris Suarez. Towards improving the MITC6 triangular shell element. *Computers & Structures*, Article in press.

[4] D. Chapelle and K.J. Bathe. *The Finite Element Analysis of Shells – Fundamentals.* Springer-Verlag, 2003.

[5] D. Chapelle and I. Paris Suarez. Detailed reliability assessment of MITC6 triangular shell elements. In preparation.

[6] P.S. Lee and K.J. Bathe. Development of MITC isotropic triangular shell finite elements. *Computers & Structures*, 82:945–962, 2004.

[7] I. Paris Suarez. *Robustesse des éléments finis triangulaires de coque.* PhD thesis, Université Pierre et Marie Curie - Paris VI, 2006.

Ocean Dynamics

Madeira Island wakes: barotropic instability study

Rui M. A. Caldeira [*]   Stefan Riha [†]   Aires dos Santos [‡]

Abstract

Madeira is a deep-sea island located in NE Atlantic ($33°N; 17°W$), its obstruction to the incoming oceanic and atmospheric flows induce the formation of leeward wakes. The ROMS - Regional Ocean Modeling System numerical toolkit was used, in a three-dimensional mode, to study the mesoscale ocean dynamics, leeward of Madeira Island. Results from numerical experiments showed strong vorticity generation, sensitive to the Reynolds number (Re) variability. Wake instabilities induce different behavior for cyclonic and anticyclonic eddies.

1 Introduction

Island wakes exhibit strong eddy activity with important consequences in the retention and transport of organic matter, and subsequent biological enrichment. Various physical processes such as filaments, small-scale upwelling cells or three-dimensional instabilities, enhance horizontal transport and vertical mixing of nutrient-rich deep water, when an upper surface current encounters oceanic islands [10] [11] [3] [6] [2]. The role of islands wakes in the biological enrichment of surface waters can not be neglected [5]. Motivated by this oceanographic context; by the fact that the most recent published discussion of Madeira Island geophysical wakes was [1]; and considering the recent advances in the numerical study of deep-sea island wakes [4], this talk explores the formation of oceanic island wakes, leeward of Madeira Island ($33°N; 17°W$). Recent numerical studies, the authors in [4] around an idealized cylinder suggest that an island wake in the deep ocean may not exhibit a strong cyclone-anticyclone asymmetry in barotropic instability regimes. However, the Madeira Island numerical study suggests the impossibility of asymmetric islands, such as Madeira, to generate solutions close to symmetry such as those often discussed in laboratory and idealized numerical experiments.

187

2 Numerical case study

2.1 ROMS configuration

The Regional Ocean Modeling (ROMS) is a free-surface, terrain-following, primitive equations ocean model. ROMS algorithms are intensively described in [8] [9]. The hydrostatic primitive equations for momentum are solved using a split-explicit time-stepping scheme. A cosine-shape time filter, centered at the new time level is used for the averaging of the barotropic fields [9]. Time-discretized uses a third-order accurate predictor (Leap-Frog) and corrector (Adams-Molton) time-stepping algorithm. A third-order upstream biased was used for advection in order to allow for the generation of steep gradients in the solution [7].

In order to study the island wake problem in a three-dimensional mode using ROMS, a methodology similar to [4] was followed. Nevertheless, unlike [4], the schematic bathymetry was not an idealized cylinder, it represented Madeira Island (NE Atlantic). The depth was assumed uniform around the island, in order to be able to isolate the effect of the island *per se*, from the effect of the surrounding seamounts. The island was centered in a channel like configuration with a prescribed inflow at the upstream boundary such that the zonal current depended only on the vertical shear. East(E) and West(W) channel boundaries were set to slippery-tangential are zero normal conditions, whereas boundaries around the island were set to zero-normal and no-slip flow. At the southern boundary, a clamped condition was used for density and radiation conditions were used for the outgoing baroclinic flows; a numerical sponge was applied upstream of the southern open boundary; viscosity and diffusivity incremented linearly $(50 - 600 m^2 s^{-1})$ over the last 5% of the model domain. At the bottom, the quadratic bottom friction law was applied. During the 150 days of model calculations, bottom velocities did not influence the upper layers solutions. Initial conditions were set equal to the inflow conditions in the entire domain, except at the E-W channel walls, and at the island points. To maintain geostrophic equilibrium in interior of the domain the following assumptions were made.

The velocity profile was taken as:

$$v(z) = \frac{c_1}{2} \left[1 + \tanh \left(\frac{z + h_s}{h_d} \right) \right] \tag{1}$$

Considering the thermal wind balance, density-field is constant in y-direction and can be written as a function of x and z:

$$\rho(x, z) = \rho_0 + h(z) - \frac{\rho_0 f}{g} \int_{x_m}^{x} \frac{\partial v}{\partial z} dx \tag{2}$$

Density anomaly $h(z)$ can be taken as:

$$h(z) = \delta_\rho \tanh\left(\frac{z + h_c}{h_t}\right) \tag{3}$$

Considering the linear equation of state, neglecting salinity:

$$\rho = \rho_0 - \rho_0 c_T (T - T_0) \tag{4}$$

Temperature can be written as:

$$T(x, z) = \frac{1}{c_T} - \frac{\rho}{\rho_0 c_T} + T_0 \tag{5}$$

A linear function was used for free surface:

$$\zeta(x) = \frac{c_1 f}{g}(x - 2x_m) \tag{6}$$

Where:

Symbol	Description	Units
h_s	central depth of shear layer	m
h_d	thickness of shear layer	m
c_1	near-surface v-velocity component	ms^{-1}
h_c	central depth of thermocline	m
h_t	thickness of thermocline	m
f	Coriolis-constant	$1s^{-1}$
c_T	coefficient in linear equation of state	$^oC^{-1}$
ρ_0	mean density	kgm^{-3}
T_0	mean temperature	^oC
δ_ρ	half of (approximate) density difference between bottom and surface	kgm^{-3}
x_m	middle of domain in east-west direction	m

2.2 Re sensitivity study

The effect of an island to incoming flow is a classic problem and can be studied by considering the balance between two forces: *inertial* and *frictional*. The ratio between inertial and frictional forces is known as the Reynolds number (Re), a dimensional number which is written as:

$$Re = \frac{UD}{\nu} \tag{7}$$

Where U is the unperturbed upstream velocity, D the horizontal scale of the object and ν the molecular kinematic viscosity. However, geostrophy i.e. rotation and stratification, differentiates oceanic and atmospheric wakes from wakes generated in homogeneous fluids, thus eddy viscosity(ν_e, viscosity hereafter) substitutes molecular viscosity(ν) [4]; allowing for comparison between laboratory and geophysical wakes. In the Madeira Island problem, D assumes an averaged value of 40 km; U, c_1 in (1), assumes a maximum value at the surface of 0.7 ms^{-1}. In order to vary Re, ν was changed. Viscosity is calculated explicitly in ROMS therefore it can be set to zero without excessive computational noise or instability (refer to discussion in [4], section 2.2 .

For Re=100 $\nu \sim 286$; for Re=200 $\nu \sim 143$; for Re=400 $\nu \sim 72$ and for Re=800 $\nu \sim 36$. The channel was represent in a regular grid 221×351×30 ($\delta x = \delta y \sim 1.5$ km). The velocity (1) and temperatures profiles (5) assumed the following parametrization: $\rho_0 = 1027 kgm^{-3}$; $T_0 = 10°C$; $c_T = 1.7^{-4}$; $g = 9.81 ms^{-1}$; $h_s = 600$; $h_d = 455$; $h_t = 1300$. The model was integrated for 150 days.

Rossby number (R_0) is another dimensional parameter which is often used to evaluate the effect of planetary rotation in the length scale of a geophysical process. R_0 is defined as:

$$R_0 = \frac{U}{Df} \tag{8}$$

When R_0 is small the effects of planetary rotation are large and the net acceleration comparably small. In this case study, $R_0 \sim 0.8$, thus implying a geostrophic constrain to flow evolution. Considering the Brunt-Väisälä buoyancy frequency ($N^2 = -\frac{g}{\rho_0}\frac{\partial \rho}{\partial z} \sim 10^{-3}$) and the baroclinic deformation radius $R_d = NH/f \sim 40km$; where $H \sim 10^2$, is the vertical scale set by the upstream flow and density profiles, h_d in (1); R_d is of the same order of magnitude as D in (8), (7), (9) therefore the Burger number ($Bu \approx 1$), considering:

$$Bu = \left(\frac{R_d}{D}\right)^2 \tag{9}$$

Figure 1 (upper panel) shows an horizontal slice of surface (z=-10m) vorticity($\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}$), with the flow field overlaid, for three different Re regimes. At Re ≈ 100, eddy shedding does not occur. Due to the asymmetric island shape vorticity is differently generated in the W side of Madeira than on the E. As Re increases the eddies develop transient meanders, detaching from the island at $Re \approx 200$. This detachment threshold is higher than that found for the symmetric island problem from [4], and certainly much higher than that found for non-rotating homogeneous flow problems i.e. $Re \approx 40$. For low Re, a weaker flow field maintains an intense re-circulation cell on the south coast. Increase Re fortifies cyclonic eddies and weakens anticyclones. As suggested by [4] the boundary shear layer around the island which is the main source of vorticity generation, is also related with changes in Re; its dependence can be fitted to a power law $Re^{-\alpha}$, where $1/4 < \alpha < 1/2$. The velocity shear near the island is the source
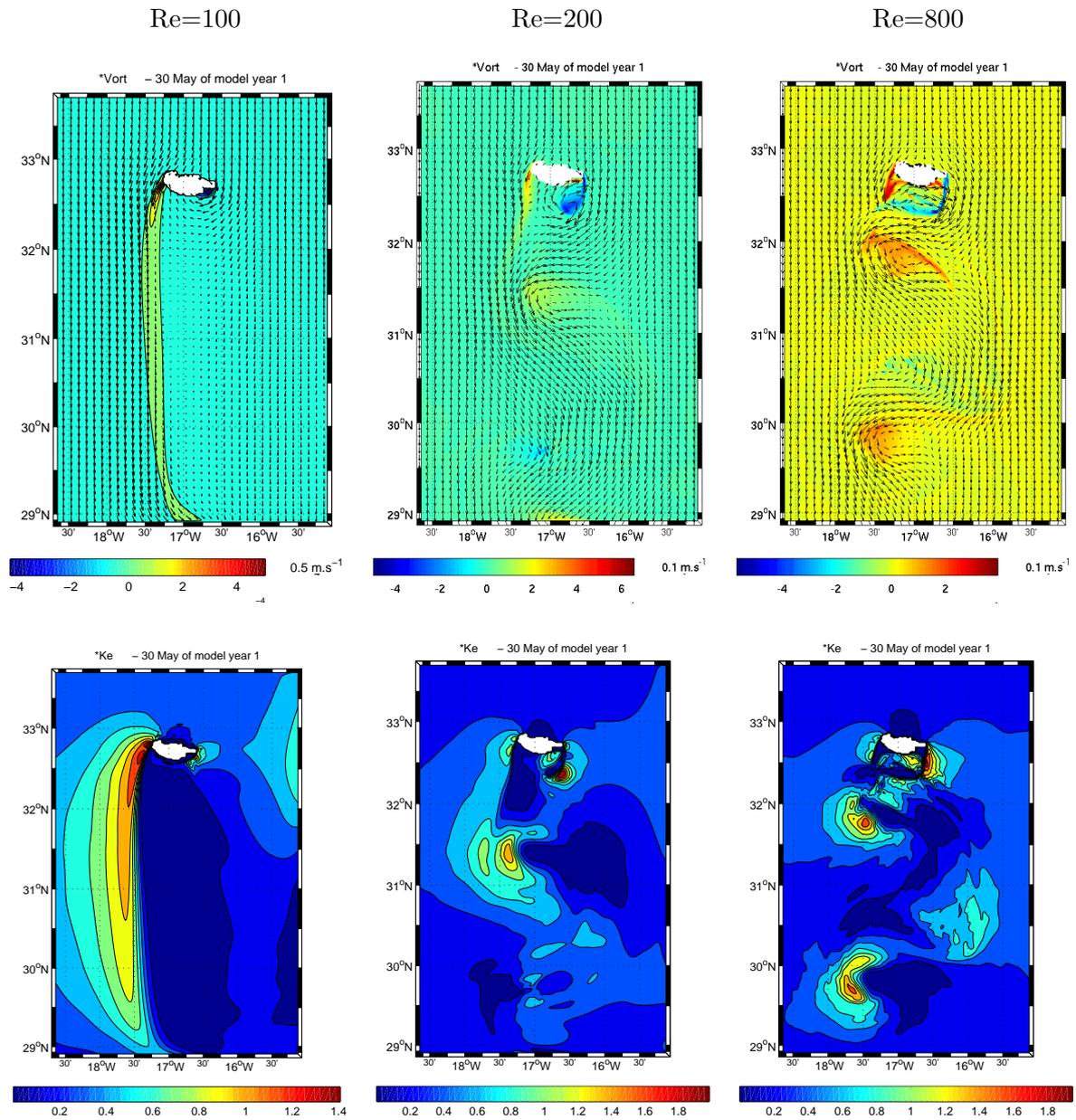
Re=100   Re=200   Re=800

Figure 1: Horizontal slices of: (i) vorticity (upper panel) and (ii) kinetic energy (lower panel), after 150 model days for: Re=100; Re=200; Re=800

of barotropic instability such that increasing shear enhances wake instability and increases eddy kinetic energy (Ke). In fact, in figure 1 (lower panel) $Ke = 1/2(u^2 + v^2)$ varies with Re. As suggested by [4], barotropic instability implies a positive volume-integrated conversion of mean kinetic energy to eddy kinetic energy (a.k.a.'barotropic conversion'). [4], suggested that barotropic conversion is largest at one island diameter, but it can occur up to two island diameters. This is also apparent from figure 1; assuming, in this case study, that most Ke originates from eddy activity, Ke is a good indicator of EKE (Eddy Kinetic Energy). Also apparent in figure 1 are the strips of intense Ke, at the island flanks, which suggest another location of intense barotropic conversion. Whereas [4], did not found dynamical asymmetry in the product vortices, the Madeira Island case, due to the asymmetric contouring of its coastline, showed strong asymmetries between cyclonic and anticyclonic eddies.

References

[1] R. M. A. Caldeira, S. Groom, P. Miller, and N. Nezlin. Sea-surface signatures of the island mass effect phenomena around Madeira Island, Northeast Atlantic. *Remote Sensing of the Environment*, 80:336–360, 2002.

[2] R. M. A. Caldeira, P. Marchesiello, N. P. Nezlin, P. M. DiGiacomo, and J. C. McWilliams. Island wakes in the Southern California bight. *J. Geophys. Res.*, 110(C11012), 2005. doi:10.1029/2004JC002675.

[3] D. E. Dietrich, M. J. Bowman, C. A. Lin, and A. Mestas-Nunez. Numerical studies of small island wakes. *Geophysics Astrophysics and Fluid Dynamics*, 83:195–231, 1996.

[4] C. Dong, J. C. McWilliams, and A. F. Shchepetkin. Island wakes in deep water. *Journal of Physical Oceanography*, 37(4):962, 2007.

[5] D. Hasegawa, H. Yamazaki, R. G. Lueck, and L. Seuront. How islands stir and fertilize the upper ocean. *Geophys. Res. Letters*, 31(16):L16303, 2004.

[6] K. J. Heywood, D. P. Stevens, and G. R. Bigg. The flow around the tropical island of Aldabra. *Deep Sea Res.*, 43(4):555–578, 1996.

[7] A. F. Shchepetkin and J. C. McWilliams. Quasi-monotone advection schemes based on explicit locally adaptive dissipation. *Monthly Weather Review*, 126:1541–1580, 1998.

[8] A. F. Shchepetkin and J. C. McWilliams. A method for computing horizontal pressure-gradient force in an oceanic model with a nonaligned vertical coordinate. *J. Geophys. Res.*, 108(C3):3090, 2003. doi:10.1029/2001JC001047.

[9] A. F. Shchepetkin and J. C. McWilliams. The regional ocean modeling system: A split-explicit, free-surface, topography following coordinates ocean model. *Ocean Modelling*, 9:347–404, 2005.

[10] M. Tomczak. Island wakes in deep and shallow water. *Journal of Geophysical Research*, 93:5153–5154, 1988.

[11] E. Wolanski, T. Asaeda, A. Tanaka, and E. Deleersnijder. Three-dimensional island wakes in the field, laboratory experiments and numerical models. *Continental Shelf Research*, 16:1437–1452, 1996.

Processes over submarine canyons: internal tide at Nazaré Canyon

Henrique Coelho[*]   Ana C. Garcia[†]   Ângela Canas[†]   Paulo C. Leitão[*]
Ramiro Neves[†]

Abstract

Submarine canyons might be a trap for sediments and organic matter. To study its influence on sediment transport we must understand how the ocean currents are affected by submarine canyons. In this work a few model results are compared with observed data to show the interaction of a large canyon (the Nazaré Canyon) with tidal currents.

1   Introduction

The Western Iberian Margin is characterized by a narrow shelf ranging from 25 to 50 km, with several canyons cutting the slope. The most noticeable of these structures is the narrow deep canyon of Nazaré, a 227 Km long *gouf* like structure [16] that cuts the continental shelf in a west-east direction almost all the way through the coast line (approximately 0.3 Km apart). The canyon floor decreases from a depth of 100 m at the head down to a depth of 4000 m. The average slope of the canyon axis is 2.2% with a maximum at 8.2% . Nazaré canyon is considered to have a strong structural origin [16]. The inexistence of a major river aligned with the head of the Canyon helps to support this idea. Furthermore, high values for short -term sediment deposition within the canyon were found [15], indicating high activity in this canyon. Since no major source of sediments is known in the area, it is possible that resuspension activity associated with internal waves play an important role in supplying and transporting fine sediments in the Canyon. The Western Iberian Margin sea level fluctuation is characterized predominantly by semidiurnal tide. The largest constituent, M2, has amplitude around 1m, and the second largest constituent, S2, has amplitude of 35 cm. The general circulation pattern in the West Iberian Margin has a strong seasonal regime (e.g. [9, 4, 13]). It is characterized by strong coastal upwelling in summertime, together with the generation of associated upwelling filaments that occurs in response to the strong

---

[*]Hidromod, Portugal. E-mail: `henrique.coelho@hidromod.com`

[†]Instituto Superior Técnico - MARETEC, Environment and Energy Section, Mechanics Department, Portugal.

north/northwesterly winds. As a result of this regime the general circulation pattern is dominated by an equatorward flow on the continental shelf and slope. During the winter, the northerly component of the wind weakens, or even reverses, reversing the surface flow that can be identified in satellite images as a relatively narrow and weak warm poleward current [8, 6, 11]. A poleward subsurface flow persists all year round extending from 200-300 m down to 1500 m (below the Mediterranean Water). Modeling results for circulation in the Western Iberian Margin [4] reproduced the major circulation patterns for this area and suggested a close relation between the topography and the decrease in transport further north. These results also showed that if the meridional component of windstress is not strong enough to reverse the flow, a permanent poleward flow will occur even during summer, if not on the shelf at least over the slope.

2 Data

Data was collected during five cruises of the NIOZ Research Vessel PELAGIA: The PE204, from 11 to 30 November 2002, the PE218 from 11October to 02 November 2003, the PE225, between 29 April and 24 Mai 2004, the PE236 from 20 April to 18 May 2005 and the PE252 between 30 August and 21 September 2006. During the five cruises some stations were repeated for a tidal cycle to observe the effect of the tide on the water column and also to search for the presence of internal tides. During the 2002 cruise two stations located in the canyon axis were chosen, one at around 800 m depth and another around 1600 m depth and repeated for the same tidal cycle. During the 2006 cruise four stations at the canyon axis were chosen, station 42 (300m), station 41 (720 m), station 37 (2700m) and station 38 (3500 m).

3 Model

The MOHID model (details available at `www.mohid.com`) was the adopted modeling tool in this study. A detailed description of the model can be found in [4]. The model is constituted by two domains: the broader domain - level 1 - encompassing the west coast of Iberia, extending from 37ºN to 40.2ºN and from 8.7ºW to 11.7ºW, and a nested domain - level 2 - for the Nazaré Canyon area from 38.8ºN to 40ºN and from 8.9ºW to 11.5º W. Horizontal grid spacing in the two domains is 1.5 km in both directions. For level 1 domain the model uses 27 vertical layers delimited by constant z-levels at surface and depths of 0, 15, 190, 304, 494, 544, 612, 727, 910, 1166, 1416, 1468, 1489, 1528, 1551, 1593, 1649, 1697, 1746, 1881, 1988, 2182, 2530, 2930, 3430, 4030 and 4800m. The bottom topography for this domain was derived

from ETOPO2[1] by means of an interpolation for the model grid followed by smoothing with a five-point Laplacian filter. The bottom cells are defined using a "partial step" approach instead of the traditional "full step" or "staircase" approach. The authors in [1] show that this approach minimizes the traditional problems associated with the staircase topography of z-level models. Level 2 domain vertical discretization is composed from bottom to surface by a set of fixed depth layers, having depths 1.7, 5.1 and 10.2m, chosen for better modelling the sediment dynamics, and an upper domain consisting in 15 sigma coordinate layers having relative thicknesses of 0.01, 0.03, 0.05, 0.08, 0.14, 0.20, 0.12, 0.1, 0.16, 0.05, 0.01, 0.01, 0.01, 0.01 and 0.02. This sigma coordinate discretization was chosen to reproduce adequately in the Nazare Canyon the annual mean density vertical gradient observed at the Iberian Margin. Biharmonic momentum diffusion coefficients are set to $1 \times 10^9 \mathrm{m}^4 \mathrm{s}^{-1}$. A constant turbulent diffusion of $10 \mathrm{m}^2/\mathrm{s}$ was considered for heat and salt transport. The model was initialised from a rest state with a null horizontal sea level gradient. The temperature and salinity fields are horizontally homogeneous. The vertical profile corresponds to the annual mean conditions observed along the Iberian Margin.

Lateral boundary conditions

In level 1 normal and tangential velocities were set to zero at the sidewalls and fresh water river input at coastal boundaries is not considered in this study. The western, southern and northern boundaries are open. A Blumberg and Kantha [3] radiation boundary condition was applied to the sea level. For salinity and temperature a null gradient was assumed.

In level 2 lateral boundary conditions are taken from the level 1 domain solution in the following way. A Flather [5] radiation boundary condition is applied for the sea level. For the horizontal velocity zonal and meridional components, temperature, salinity and cohesive sediment is considered for the 11 near lateral boundary cells a relaxation scheme [10]:

Sediment model

For bottom cohesive sediments, a flux term, F_b, (mass of sediment per unit bottom area per unit time) can be defined, corresponding to a source or sink for the suspended particulate matter in conditions of erosion or deposition, respectively. Consequently, at the bottom:

$$F_b = F_E - F_D$$

where F_E and F_D are the erosion and deposition fluxes, respectively. It is assumed that when bottom shear stress is smaller than a critical value for deposition, there is addition of matter to the bottom. Conversely, when the bottom shear is higher than a critical value, erosion occurs. The erosion algorithm used is based on the classical approach of Partheniades [12].

---

[1]U. S. Department of Commerce, Natural and Atmospheric Administration, National Geophysical Data Center, 2006, 2-minute Gridded Global Relief Data (Etopo2V2).

4 Results

Temperature, Salinity and Nephelometry were analyzed for the five cruises performed. During the November 2002 cruise CTD casts were performed along an axis transect and also some cross canyon transects. Temperature and salinity profiles revealed a well-defined surface mixed layer (SML) 100-150 m thick in the most seaward stations, separated from the deeper water masses by a very steep seasonal thermocline. Below this the North Atlantic Central Water (NACW) was encountered down to 500 m, the Mediterranean Water (MW) from 500-1500 m, and the Northern Atlantic Deep Water (NADW) below 1500 m. Two stations at the canyon axis (Station 38, at 1600 m and Station 41, at 800 m) were repeated for the same tidal cycle. By observing the salinity profiles from these stations for this tidal cycle (not shown) it is possible to notice the displacement of the salinity maximum at the MW level in the deeper station (station 38). In station 41 this effect can also be observed at the bottom of the surface mixed layer, but it is less evident. In station 38 an important bottom nepheloid layer is present. Although there is a considerable vertical displacement, the intensity of the nepheloid layer shows little variability. In station 41, on the other hand, both the vertical location and the intensity of the nepheloid layer show a large variability at around 100 m depth. Nevertheless, the time resolution in this experiment was insufficient to show that this variability has time scales similar to the semi-diurnal tide. For this reason during the 2006 cruise the vertical profiles were repeated for station 41 station every 30 minutes during 13 hours. The sinking of the salinity minimum during high tide was followed by a decrease in intensity of the BNL that progressively increases during the low tide. CDT data obtained at three cross-canyon sections identified as A, B and C are represented on Figure 1. In the deeper transect (A) a high turbid bottom nepheloid layer (BNL) at the level of the shelf-break could be observed in the southern rim of the canyon walls. Also in the southern wall, at 1500 m, below the MW, an intermediate nepheloid layer (INL) was observed spreading for nearly 15 km across the canyon. This INL extended to the northern wall, where it was centered at 2000 m. In the canyon axis an important BNL could also be observed. At cross-section B, a mid-canyon transect, we observed an important INL at nearly 200 m extending cross canyon. This INL might be an extension over the canyon of a BNL located over the surrounding continental shelf. A BNL seemed to be permanently present along the northern canyon wall down to 1000 m. Below this depth a progressive increase in nephelometry is observed all over the canyon axis. This data follows the observation in section A (although with a different amplitude - please note the different scales) indicating the presence of a permanent BNL along the axis. Transect (C) showed an important increase in turbidity generalized through the whole water column. However a maximum occurred at the shelf break level. This local maximum was followed by a local minimum and then by a monotonically increase with depth. Once more the BNL was present in the station at the canyon axis. In general, the turbidity maximum was found at mid-canyon, which may be connected with the channeling of the internal tide

198

energy along the canyon. The decrease near the canyon head could be due to the canyon meanderization that would dissipate the tidal energy. Figure 2 shows a transect along the canyon axis. Results indicate the presence of an important BNL at 1500 m, at the level of the bottom of the MW, this layer extends seaward for several kms. From this depth to almost the canyon head a continuous bottom nepheloid layer is formed along the canyon axis.

The model was forced exclusively with tides, using a realistic stratification profile and bathymetry, to verify if in such conditions it is possible to reproduce the observed patterns. The evolution of the water column sediment concentration with time for a station at 800 m deep on the canyon axis is comparable with data, showing the some kind of variability in the BNL at 800m deep. Figure 3 show the baroclinic component of the instantaneous zonal velocity in a transect along the canyon axis. It is evident the typical beams associated to maximums in velocity modulus resulting from the propagation of internal tide. In the next item the correlation between these beams, maximum suspended sediment concentration and generation spots for internal tide will be discussed. We will analyse now the model results for cross sections A, B and C referred in Figure 1 and the along axis transect (Figure 2) Although quantitatively there are important differences between model and data (see for example Figure 1 and Figure 4), qualitatively the model reproduces most of the observed patterns. One thing that we must keep in mind is that for model simulations bottom topography was interpolated and smoothed. This procedure is responsible for the differences seen in the shape of the canyon. For all the three sections we can state that salinity results are in close agreement with CTD data, since the same patterns can be found for all stations. For section A (see Figure 4) the model reproduces the maximum in sediment concentration observed at the bottom of the SML with the high turbid levels at the southernmost station following the same N-S gradient observed in field data. Another interesting aspect on model results is the capacity for reproducing the INL at the level of the bottom of MW. It not only reproduces this INL but also the observed S-N gradient with higher values decreasing from south to north. As observed in the field data the INL maximum is located in the southern flank around 1500 m deep. The model was able to reproduce also the BNL in section A. In section B, the model reproduces well the patterns observed with field data. The model reproduces correctly the INL below the SML and also the BNL at the canyon axis and along the entire northern wall. For cross-section C, results go along with those already obtained for sections A and B. Again the model is able to reproduce the nepheloid layers at the SML. Although field data does not show continuity between the INLs at the shelf break and the BNL at 300-400 m depth, this continuity is observed in model results. In the model there is a continuous high turbid layer starting on the shelf going along both flanks until the canyon axis. This results seem, nevertheless, very realistic. Suspended matter concentrations produced by the model along the canyon axis are shown in Figure 5. Maximum concentrations are observed near the bottom at the canyon head where strong velocities occur. Model also shows relative maximums at the

canyon floor around 1500 m depth extending seaward. These relative maximums are related with strong velocities generated by internal tide propagation along the canyon axis (see also Figure 3). Beside these local maximums, high sediment concentrations can be observed along the canyon axis, in agreement with field observations. The observed pattern in both situations (model and field data) is very similar, mainly the INL at the canyon axis is very realistic and in accordance with field data As mentioned earlier submarine canyons play a special role, as they may act as a trap for internal wave energy and by the capacity for canalizing this energy all way up to the canyon head. To better understand the role of Nazaré canyon in driving this energy we have calculated the specific perturbation potential energy (SPPE) according to [7] — not shown here. A large amount of the energy is concentrated near/over the canyon mouth and mid-canyon — but not over the canyon head, which might explain why the maximum sediment concentrations are found at mid-canyon and not at canyon-head. We note that SPPE is asymmetrically distributed in the cross-canyon direction, with a greater percentage lying along the southern wall. This is particularly evident in the canyon mouth. This kind of asymmetry in energy distribution is likely due to Coriolis rotational effects. Identical features have also been observed by the authors of [14]. In the surrounding areas of the canyon most of the SPPE concentrates in shelf break area. Some local maximums of SPPE over the shelf to the North of the canyon correspond probably to refraction of internal waves that were propagating along the canyon or to waves generated at the shelf break.

5   Discussion

Different stations were occupied in each cruise so it is not possible to compare results from one year to the other, for the same location. The 2002 and 2005 cruises were the only where several cross-canyon CTD transects were performed. For studying the role of internal waves in the canyon, cross-canyon transects are fundamental. For that reason and since the model run for the period of the 2002 cruise results from this particular cruise were addressed and discussed. Nevertheless, we would like to sustain that the performed analysis had the results from all the five cruises as background. During the 2006 cruise most attention was given to the continental shelf and slope involving the canyon area. For comparing results on the canyon and on a typical margin data from the 2006 cruise were also presented. On the canyon itself the main objective was to observe the effect of the semi-diurnal tide on the water column sediment distribution. Vertical displacements in the water masses, that reached 200 m, were observed on the stations where CTD profiles were repeated for a tidal cycle. These vertical displacements could be observed at the level of the superficial water and also the MW. Data from 2006 permit to confirm that these migrations have a time scale in the order of the semi-diurnal tide. The profile repeated in 2006 also reveals that the variability in the near bottom turbidity has the same time scale. Model results show the same pattern.

The variability in concentration near the bottom is very similar to the one observed. The model's only forcing condition was the tide, which leads to the conclusion that this variability is induced by tidal forcing. The similarity between model results and field observations is an indication that internal tides might have a dominant role in resuspension and concentration of bottom sediments. Data obtained during the several cruises performed show the existence of a permanent vertical stratification below the surface layer subjected to a strong seasonal variability [2]. This stratification is known to be favorable to the development of internal tides. Internal tides can be found nearly everywhere in the ocean. They are generated by the combined effect of barotropic tides and topography. Generation of the internal tides takes place where horizontal currents impinge on undersea mountains or ridges. The obliquely propagating internal waves are in turn subject to reflections and scattering. Reflection may lead to local concentration of energy.

To find a relation between internal tides and nepheloid layers one needs to consider the linear internal wave theory. Over sloping topography, bottom intensified flows and high mixing rates associated with internal waves can occur when the slope of the topography ($\alpha = \Delta h/\Delta x$) is similar to the angle that the group velocity vector (ϕ) makes with the horizontal. This angle is determined by $s = \pm\sqrt{(w^2 - f^2)/(N^2 - f^2)}$, where $s = \tan\phi$, w is the wave frequency, f the Coriolis parameter, and N the buoyancy frequency. Such slopes are termed critical slopes. These intensified flows may be the dominant process contributing to bottom shear stresses on continental slopes and hence be important in determining sediment resuspension.

The ratio α/s will determinate how the internal wave energy will propagate along the bottom slope. If the characteristic angle (CA) of the beam is steeper than the slope of the seafloor, internal tidal energy is transmitted landward as it bounces between the seafloor and the base of the mixed layer. If the CA is shallower than the slope, the energy is reflected back toward the deep ocean. When these two angles are equal, the energy is trapped near the seafloor, where it may be sufficient to stir up the sediments or at least to prevent suspended sediments from settling out. Considering this, a map for critical conditions was drawn (not shown) taking into account the ratio α/s for Nazaré Canyon and adjacent areas. To compute α/s, one needs the stratification and the bathymetry. In order to keep consistency the calculation was made using the same data used for the high resolution model calculations. Many near-critical areas are found along the continental slope to the north and to the south of Nazaré Canyon mainly between the 500 and 1500 m isobaths. Near the canyon mouth several critical areas are also found. Note in particular that the bottom slope is critical for M_2 tide at the southern end of section A previously referred. Along the Nazaré Canyon rims, all way up to the canyon head, near - critical areas are found. This might justify the large amount of suspended sediments found within the canyon. The authors of [14] explored the effect of canyon floor slope on internal tide energy generation and found that *"canyons that are near-critical along much of their entire length, ..., develop strong internal tides that*

propagate shoreward.". This seems to be approximately the case of Nazaré canyon and this explains the formation of INLs.

Results obtained with field data and model application do reveal the existence in Nazaré Canyon of important INL located at depths between 1000 and 1500 m and also a BNL that follows almost the entire canyon axis (as we saw is an area were much of the energy associated with internal tide is concentrated. The presence of an INL located at around 1000 m deep and a less evident one at 1600-1700 m deep, may be observed when looking to a section performed North of the canyon area during the 2006 cruise (not shown). Model results are concordant with field data for this section. Nevertheless they show an INL below 2000 m deep. As we saw depths between 1000 and 1600 m deep are exactly the near-critical areas for the semi-diurnal tide. So we may consider that these INL are caused by the near bottom mixing due to the semi-diurnal tide energy trapping in near-critical areas. Although we can infer this hypothesis from the field data, the fact that the model reproduces so well these patterns — INL in near critical areas for the internal tide — highly increases our confidence in this hypothesis. The fact that the only forcing mechanisms considered in the model configuration were tide and a density profile, leave no space to consider other cause then internal tide. Cross-canyon sections reveal that maximum concentrations for sediments are found at section B in the mid canyon. Concentrations are lower in the upper canyon (maximum at bottom is 0.5 mg/l) and much lower at the canyon mouth (section A). This should mean that, while at lower canyon there is no much energy available, at upper canyon a considerable portion of the energy was dissipated probably due to the canyon meandering. Another important observation is that sediment concentration is much higher in the canyon when compared with the slope north of the canyon. In fact in this region sediment concentrations at the INL do not exceed 0.1 mg/l (both in model and field data results). At the canyon sediment concentration at the INL located at 1000 m deep reaches 1mg/l in the field data and 0.5 mg/l in model results. This clearly shows that the canyon acts like a trap for the energy associated to the internal tide. This result is supported by the SPPE distribution that is clearly influenced by the canyon itself and by the "Promontório da Estremadura" south of the canyon were internal tide is likely to be generated. On the other hand Figure 3, were it can be observed the baroclinic component of velocity along all the meandering canyon axis, suggests a good correlation between near bottom velocity and suspended sediment concentration, mainly in the near-critical slope between the 2000 and 1000 m deep. Combining the observations performed during the different cruises between 2002 and 2006 and the simulations of the generation and propagation of the internal tide we can state that:

1. Most of the suspended sediment distribution observed in the canyon and at the continental slope north of the Canyon may easily be explained exclusively based on the propagation of the internal tide.

2. The INLs locations at the canyon do not diverge from the observed North of the canyon.
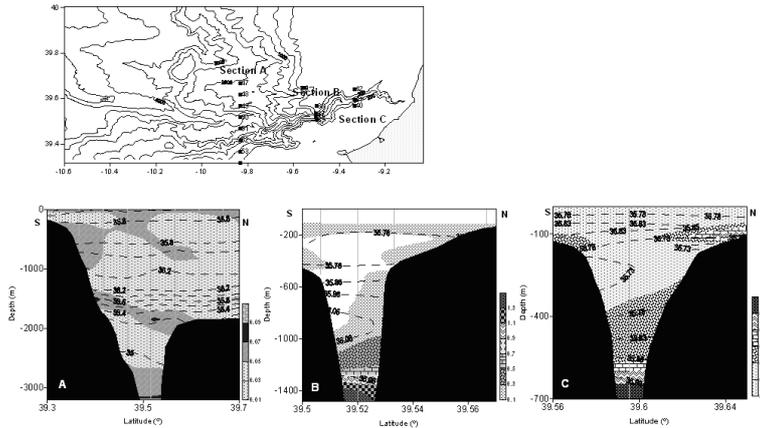
Figure 1: CTD data for Salinity (dashed lines) and Turbidity (contours) profiles at cross canyon sections. The sections location is shown in the upper panel. The left panel refers to section A (the most seaward section), the central panel to section B and the right to section C.

Nevertheless the concentration of sediments found at the canyon is much higher. This suggests that the energy associated to internal tides is mainly concentrated in the canyon area. This is confirmed by the model results. For one hand the canyon acts as a conduit for the energy to propagate shoreward, or at least to the area were the canyon meandering tends to dissipate some of this energy. On the other hand the ridges related with the canyon morphology and the "Promontório da Estremadura" (south of the canyon) tends to be generation spots for internal tides, increasing the available energy.

References

[1] A. Adcroft, C. Hill, and J. Marshall. Representation of topography by shaved cells in a height coordinate ocean model. *Monthly Weather Review*, 125(9):2293–2315, 1997.

[2] H. M. Van Aken. The hydrography of the mid-latitude northeast Atlantic Ocean I: The deep water masses. *Deep-sea Research I*, 47:757–788, 2000.

[3] A. Blumberg and L. H. Kantha. Open boundary conditions for circulation models. *J. Hydraulic Engineer.*, 111:237–255, 1985.

[4] H. S. Coelho, R. J. J. Neves, M. White, P. C. Leitão, and A. J. Santos. A model for ocean circulation on the Iberian coast. *Jour. Marine Systems*, 32(1–3):153–179, 2002.

[5] R. A. Flather. A tidal model of the northwest european continental shelf. *Mem. Soc. R. Sci. Liege*, 10(6):141–164, 1976.
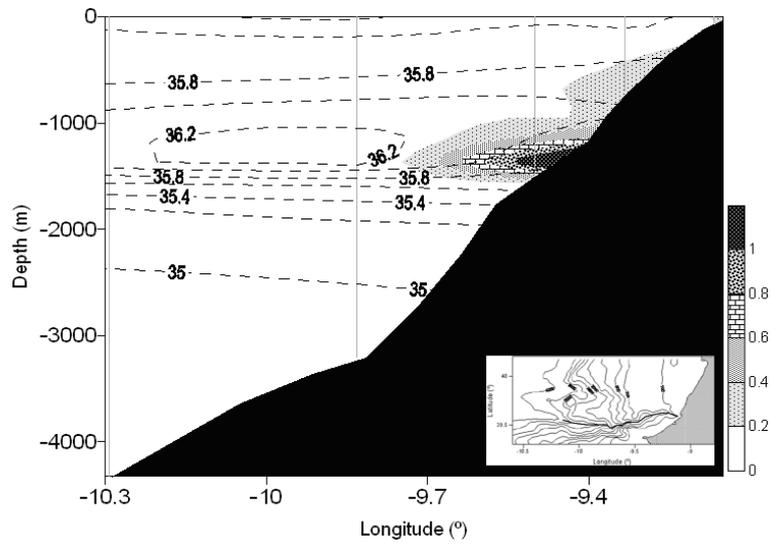
Figure 2: CTD data for Salinity (dashed lines) and Turbidity (contours) profiles in the along canyon axis section.
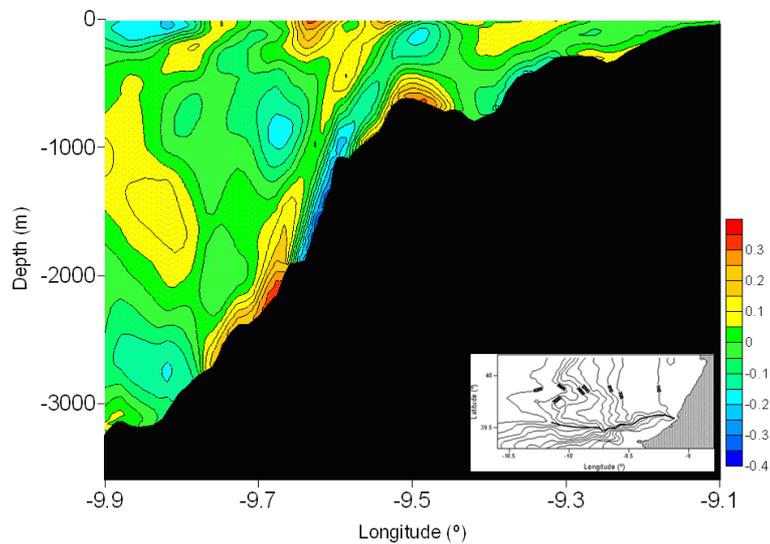
Figure 3: East-West baroclinic velocity contours in the along canyon section depicted at lower right corner of the panel during flood.
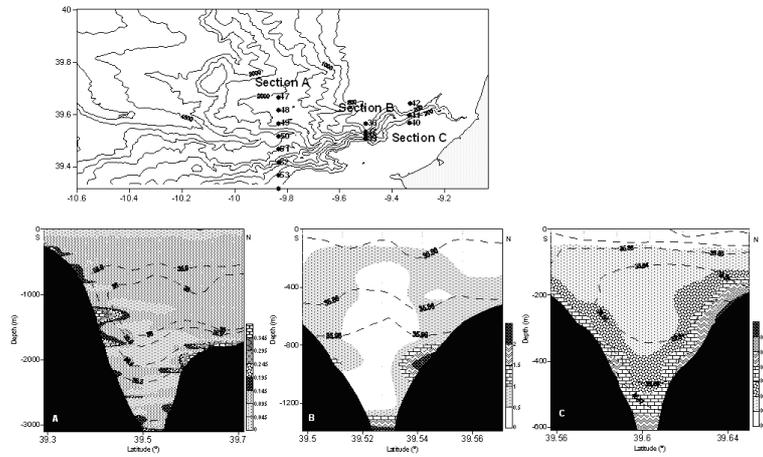
Figure 4: Model results for Salinity (dashed lines) and Turbidity (contours) profiles at cross canyon sections. The sections location is shown in the upper panel. The left panel refers to section A (the most seaward section), the central panel to section B and the right to section C.
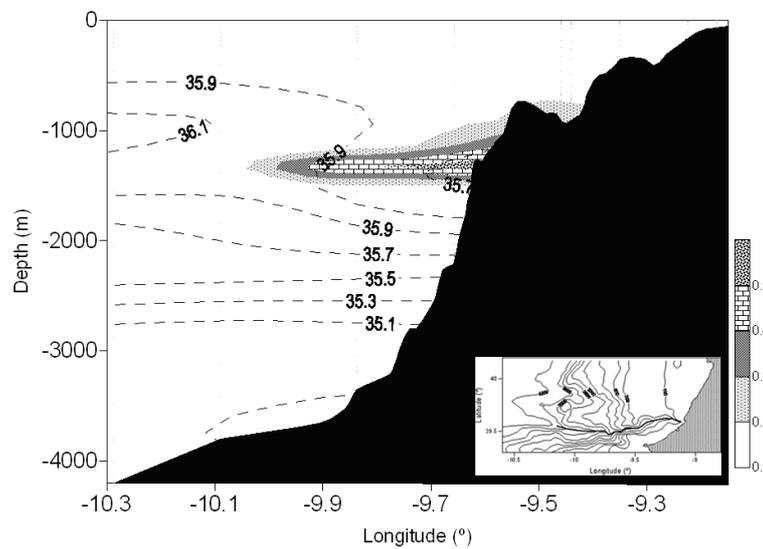
Figure 5: Model results for Salinity (dashed lines) and Turbidity (contours) profiles in the along canyon axis section.

[6] R. Frouin, A. Fiúza, I. Ambar, and T. J. Boyd. Observations of a poleward surface current off the coasts of Portugal and Spain during the winter. *J. Geophys. Res.*, 95:679–691, 1990.

[7] A. E. Gill. *Atmosphere-Ocean Dynamics*, volume 30 of *International Geophysics Series*. Academic Press, 1982.

[8] R. Haynes and E. D. Barton. A poleward flow along the Atlantic coast of the Iberian peninsula. *J. Geophys. Res.*, 95:11141–11425, 1990.

[9] R. Haynes, E. D. Barton, and I. Pilling. Development, persistence and variability of upwelling filaments off the Atlantic coast of the Iberian Peninsula. *J. Geophys. Res.*, 98:22681–22692, 1993.

[10] E. A. Martinsen and H. Engedahl. Implementation and testing of a lateral boundary scheme as an open boundary condition in a barotropic ocean model. *Coastal engineering*, 11(5–6):603–627, 1987.

[11] J. P. Mazé, M. Ahran, and H. Mercier. Volume budget of the eastern boundary layer off the Iberian Peninsula. *Deep-Sea Res.*, 44:1543–1574, 1997.

[12] E. Partheniades. Erosion and deposition of cohesive soils. *J. Hydr. Div., ASCE*, 91(1):105–139, 1965.

[13] A. Peliz, J. Dubert, A. M. P. Santos, P. B. Oliveira, and B. Le Cann. Winter upper ocean circulation in the Western Iberian basin — fronts, eddies and poleward flows: an overview. *Deep Sea Res. Part I: Oceanog. Res. Papers*, 52(4):621–646, 2005.

[14] E. T. Petruncio, J. D. Paduan, and L. K. Rosenfeld. Numerical simulations of the internal tide in a submarine canyon. *Ocean Modelling*, 4(3–4):221–248, 2002.

[15] S. Schmidt, H. C. Stigter, and T. C. E. van Weering. Enhanced short-term sediment deposition within the Nazaré canyon, North-East Atlantic. *Marine Geology*, 173(1-4):55–67, 2001.

[16] J.-R. Vanney and D. Mougenot. Un canyon sous-marin du type "gouf": le Canhão da Nazaré (Portugal). *Oceanologica Acta*, 13(1):1–14, 1990.

Operational oceanography. Zoom-in modeling for local applications

Emanuel Ferreira Coelho*

Abstract

Multiple operational forecasting schemes are usually available for the same scene, with multiple nesting and coupling options, covering large domains and including many regions of potential interest. However, local planning end-users usually ask for detailed and accurate one-scene, one-forecast products, and are highly sensitive to local data comparisons, whenever available. This work will show model training techniques that include stochastic modeling and assimilation tools, to improve the reliability and resolution of local end-products. This methodology is suited for end-to-end problems combining non-linearly related inter-disciplinary parameters, with different statistical properties, such as surface drift and ocean acoustic properties estimation.

1   Introduction

Coastal operational oceanography can be seen as a methodology to provide detailed and reliable multi-scale environmental information for littoral management, with a readiness and regular updates, depending on the final applications. Typically it includes modeling, observation networks, data fusion using Geographical Information System (GIS), and information delivery components with different levels of complexity, and combines unsupervised automated tools with subject matter expert efforts, into standard products tailored to identified end-users (e.g. [5], [11], [14], among others). Following an end-user perspective, multiple operational forecasting schemes are usually available for a given scene, with multiple nesting and coupling options, covering large domains and including many regions of potential interest. Different systems will likely use different local and remote sensing data. However, these local planning end-users usually are looking for detailed and accurate synthesized one-scene-one-forecast products are highly sensitive to local data comparisons, whenever available, and the required lead times for information processing and delivery may not be compatible with those of the standard products. The Rapid Environmental Assessment concept (REA) was

---

*Naval Research Laboratory (University of Southern Mississippi -visiting scientist) - NRL7320, Stennis Space Centre, MS39529-USA. E-mail: `coelho@nrlssc.navy.mil`

built within this framework as a process to tune environmental prediction systems to regional and local needs, within the required timeframes (e.g. [2]). Aiming to address this problem, the approach detailed below, combines observational (autonomous or supervised), model outputs and local data assimilation systems, through post-processing, in near real-time, to produce a recognized environmental product that subsequently can be used for operational and management purposes. Besides the availability of the required products, when addressing the predictability of stochastic environmental variables through numerical modeling, multiple sources of errors need to be considered. They are associated with the initialization and boundary conditions of models, numerical approximations, modeling strategies, impact of under-sampling in the assimilation process and unresolved scales. Several approaches have been used to address this problem that includes Monte-Carlo methods, producing ensemble based error estimates along with the predicted state variables (e.g. Lermusiaux, 2006). In order to guarantee error estimates are precise and accurate (i.e. the ensemble is calibrated), Kalman Filter based approaches have been successfully used to constraint the ensemble perturbations based on the observed errors (e.g. Bishop, 2001 and Evensen, 2003). However, ensemble calibration can also be achieved through post-processing using Bayesian statistics (e.g. [4], [2] and [12]), within the limits of the known cross-correlations among the observed and modeled variables. This work will show examples of these later post-processing techniques, applied to single variable problems using multi-model baysean averaging, to obtain short term local unbiased estimates and to generate an ensemble of possible values for the output variable, based on bootstrapping from probability distribution functions determined by the multi-model dispersion. Chapter II will introduce the Monte-Carlo experiments based on uniform distribution functions to reproduce the maximum entropy states, advected through the model domain. Chapter III will show two test cases during the MREA03 (off Elba-Italy) and MREA04 (west coast of Portugal) trials. Finally Chapter IV will introduce some final remarks and introduce on-going research using more advanced schemes based on the Ensemble Transform and Ensemble Transform Kalman Filter, as tested during the more recent MREA07 and LASIE trials (off Italy).

2 Ensemble generation

Numerical models estimates contain multi-scale physics with different statistical properties. Local data, though representative of certain scales and processes, will not be accurate to reproduce scales and time spans not well resolved by the observation network. These facts suggest treating assimilation and error prediction as a multi-scale problem such that several techniques have been proposed, like the multi-scale energy vorticity analysis (e.g. [9], [12]) and the mini-Hops downscaling (e.g. [3] and [8]). Under this limitation one can claim the operational model fields and observations could be used to span a domain containing most

of the true ocean-meteo-wave states, of a fully coupled system. Once several different model fields are available for a particular area and period of time, they will define the signal space of the possible states and reduce the orthogonal error space. The problem now will be how to use these two spaces, estimated during a training period of time preceding an analysis, to produce an optimal state-variable (or output combination) forecast estimate and the associated error span. One solution is to use averaging techniques (e.g. [2], [10]), where the problem is reduced to finding a set of weighting coefficients a_i such that

$$\sum_{x_0} \left\| S(x_0;t) - \sum_i a_i \Psi_j(x_0;t) \right\| \leq \varepsilon_{\mathrm{MIN}}$$

where $S(x_0;t)$ are the observations at locations x_0, Ψ_j are significant dynamic modes included in the multi-model suite, mapped and converted into the observed variables dimensions, and $\varepsilon_{\mathrm{MIN}}$ is a representation threshold.

Several methods can be used to find the set of coefficients a_i (e.g. using non-linear statistics as detailed in [12]). For this work, the problem is reduced to a single variable tracking and a sequential iterative method is used. Two iterative schemes are applied, one goes along the diagonal terms of the models fields covariance, for classes of modes with equivalent scales (MDS), and the other uses a hierarchical sequential method, fitting the residues from the previous fit into the next modes (LDS). This second method is more adequate for modes representing different scales. In the examples detailed below for surface drift estimation, the wind predictions are mapped into surface currents using a simple Ekman model, wave predictions are mapped into currents using a stoke drift model and together with the upper layer currents from the ocean model are used to generate the set of modes Ψ_j (see [2] for further details). The procedure of finding the coefficients combines both MDS and LDS techniques. It starts by computing an MDS estimate through all available ocean currents, then the residuals are fitted into a MDS of all available wind forced currents, and the final residuals are fitted into the Stokes drift estimates. The final sets of coefficients are then applied to the predicted amplitudes of each of the modes, in order to obtain local unbiased forecast estimates of the surface drift for the following cycle. The ensemble error field prediction is then carried using recursive non-linear Gaussian Bayesian tracking applied to surface drift problems such as:

$$\begin{aligned}
v(x;t) &= g_{t-dt}(x;t) + N(t-dt,0,G) \\
g_t(x,t) &= E_t(x,t) \times \{1 + \eta_c(x,0,c)\} + \eta_D(x,0,D)
\end{aligned}$$

where $E_t(x;t)$ corresponds to the drift velocity optimal combination of the multiple models using the set of coefficients $\{a_i\}$ estimated for the time t and time step dt, $g_t(x;t)$ is a velocity ensemble estimate, based on random perturbations of the optimal velocity along each trajectory and $v(x;t)$ is the true velocity of the true trajectory. This approach assumes the ensemble surface drift velocity errors to have two local random components. The first

component (η_D) represents the bias errors and is determined by a bias dispersion velocity (D), estimated from the analysis position error standard deviation, divided by the time step of the forecasts. The second component (η_C), represents the uncertainty in the velocity estimates, not resolved by the available models. It is represented here through a uniform error with zero mean and variance (C) determined by the local drifter relative velocity error. The error between the model ensemble estimate g, trained using the data from the previous forecast cycle ($t - dt$) and the real velocity field v can be assumed to have a Gaussian envelope distribution N with zero mean and variance G, such that $G_{t-dt} \approx C \cdot V + D$ (where V is a velocity magnitude scale) can be estimated from the learning period. These results can be used to test the efficiency and reliability of the method by tracking the magnitude of the predicted errors (smaller errors larger efficiency). Once new data become available one can compute the likelihood of an error equal or smaller than the observed, using the distribution N such that larger probabilities will mean the system was more reliable or accurate.

An example is displayed in Figure 1, for 14 June 2003, during the MREA03 trial. The "analysis RMS" (upper plot), relative velocity bias and correlation error (central plot) and "forecast RMS" (plot below), for drifters 1 (blue) and 2 (green) are displayed in the left panel, during the iteration phase. The jumps close to iteration 13 for drifter 1 and at iteration 16 for drifter 2 show random starts once an "analysis RMS" minimum was found. This method avoids trapping in local minima. At the end of the iteration process, Areas of Containment (AC) were computed using a "bootstrap" from maximum entropy uniform distributions, with 100 ensemble members, per drifter. The panel on the right shows the final estimated trajectories (solid lines) and observed drifter trajectories (dashed lines). The trajectories for the period -24 to 0 hour were used to estimate the ensemble weights and statistics, as explained above. Then, the tracers were re-initialized for the period 0 to 24 hours, starting from the observed drifters positions at the 0 hour. The purple areas displayed for each drifter are polygons containing all ensemble members, after a 24 hour period. The dots inside the AC's represent tracer's final positions and are indicative of the drifter spatial distribution functions, as could have been estimated using this technique.

3   Surface drift and velocity profiles estimation examples

Surface drift tests were carried during two trials in 2003. The MREA03 trial took place in the Eastern Ligurian sea. The ocean dynamics in this area are predominantly wind driven, though affected by irregular coastline and islands, superimposed to large scale compensating north-south flow (more intense along the Corsica channel), forced by the meridional temperature gradients. During this trial the Harvard Ocean Prediction System (HOPS), made available by Harvard University and the NCOM (using COAMPS atmospheric forcing and global NCOM boundary conditions) and SWAN, made available by the Naval Research Laboratory (NRL),
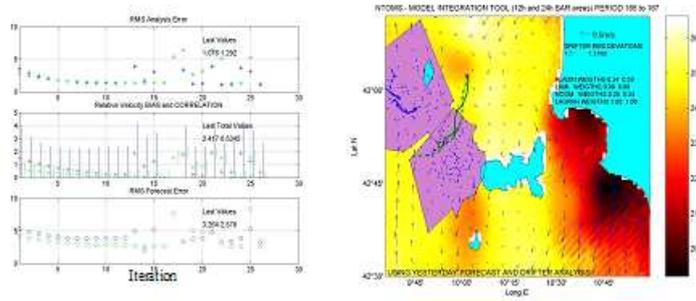
Figure 1: Surface drift training example for 14 June 2003, during the MREA03 trial. The background field in the right panel represents Navy Coastal Oceam Model (NCOM) sea surface temperature and the arrows are the NCOM forecast mean surface currents. This example combines two atmospheric models and two ocean models to produce optimal drift tracks.

were running in real-time, producing daily 48 hours forecasts of the ocean and wave fields. The HOPS system was also doing local temperature and salinity profile assimilation. Three meteorological models (COAMPS, ALADIN and LOKAL), running at different high resolutions and the wave model SWAN (forced by COAMPS) were also available. ALADIN outputs were made available by the Meteo France, through EPSHOM, and the LOKAL fields by the German Meteorological Service (Deutscher WetterDienst, DWD). Three ARGOS-CODE drifters were released in the area between Capraia-Elba-Italian Coast, in June and up to eight days of data were used for this analysis. The systems transmitted twice a day, through the ARGOS system, their GPS positions at a sampling rate of approximately 30 minutes. For the present study only positions between 13 and 20 June were considered. The second test area was the NL03 NATO exercise that took place in the Northern Irish Sea, close to Luce Bay (Scotland). Local dynamics were characterized by strong, spatially inhomogeneous, tidal currents, affected by islands and irregular coastlines, such that mean flow length scales were smaller then tidal excursions, imposing a nearly "chaotic" type of circulation. During this exercise the Proudman Ocean Laboratory Coastal-Ocean Monitoring System (POLCOMS) fields were made available through the UK Meteorological office. The wind field predictions were provided by the UK Meteorological office, together with LOKAL and COAMPS outputs. The SWAN wave fields, forced with COAMPS model, were also made available by the NRL. During this exercise, five ARGOS-CODE drifters were deployed by helicopter during September 2003 and up to eight days of data were used for model integration.

Figure 2 shows the results during the MREA03 (right panels) and NL03 (left panels). The horizontal axis represents efficiency taken as the standard deviations of ensemble velocity spread, computed from the analysis error of the training data set. The vertical axis show the reliability defined as the probability of having a forecast root-mean-square velocity error

(ensemble spread) larger than the observed error. Each dot represents one day forecast, and each symbol stands for a different drifter. The panels below are computed taking the ocean model upper layer velocity without any corrections, and are to represent the result one would get if using model outputs directly without further post-processing or model integration. Results show there is a significant increase in efficiency and reliability for both cases. This means if this training methodology was being used for Search and Rescue operations, the system would be providing search areas roughly 20% smaller and more accurate, up to 0.85 likelihoods, on average.
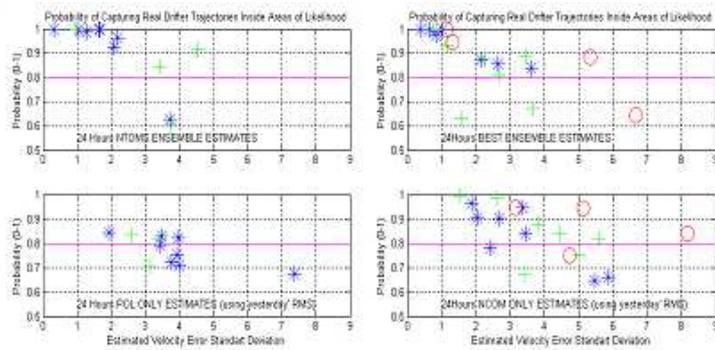
Figure 2: Model comparisons with CODE-ARGOS drifter data for the MREA03 (right) and NL03 (left) cases. Each dot represents one day average and each symbol a different drifter simulation.

These results were further tested during the MREA04 trial, off the coast of Portugal. Local dynamics were rather complex, combining tidal dynamics with wind driven currents, over the steep topography of a submarine canyon. During this trial, two HOPS runs made available by Harvard University and by Instituto Hidrográfico, two NCOM grids run by NRL (Ko, 2007), along with the atmospheric models ALADIN, made available by Meteo France and Instituto Meteorologia, LOKAL, made available by DWD, along with COAMPS and the wave model SWAN made available by NRL ([1]), were running in real-time. Historical data showed also non-linear internal wave trains to be an important local phenomenon; forcing strong shear and currents whenever present (e.g. [6]). This high frequency phenomenon was not being resolved by the available circulation models. Three ARGOS-CODE drifters were released during the month of April. The left panel of Figure 3 shows an example of a daily 48 hours forecast, with the predicted areas of containment, estimated for each 6 hours period, along with the observed and best predicted tracks. These estimates were using only NCOM and COAMPS integration, and we can see though the north and south estimates were accurate, the predicted areas of the drifter in the middle were not capturing the true trajectory.

The same training approach was used to assess the NCOM ocean model profile velocity
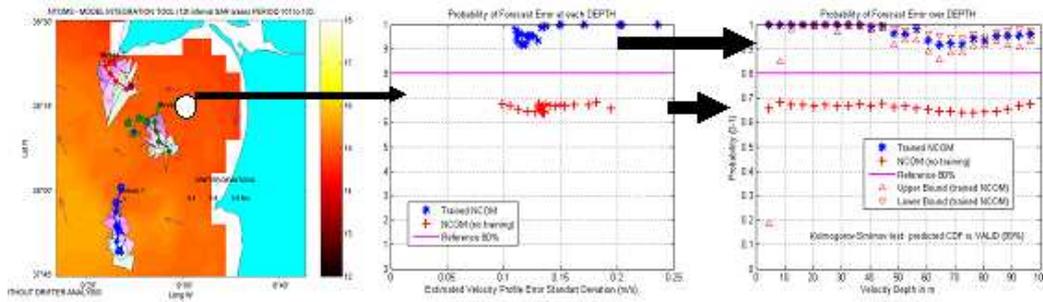
Figure 3: MREA04 example of April, 10 2004. The left panel shows drifter tracks along with the best estimate combining NCOM and COAMPS fields. The central and right panels show the training results for NCOM velocity estimation, at the location highlighted in the left panel.

estimates, using the BARNY ADCP mooring at the location shown as a white dot in the left panel (close to the central drifter). The scatter diagrams for the same period at each depth are shown in the central and right panels of Figure 3. The central panel shows training, although is not improving the magnitude of the prediction errors (precision), is improving the overall accuracy or reliability of the output currents. This is due to the fact NCOM was accurately estimating tidal currents that dominate the temporal variability ([7]), or possible variations around a mean value, such that training was only correcting the daily bias. The panel on the right in Figure 3 shows the probability of having the predicted errors larger than the observed, for each depth. Though, the upper layers show high likelihoods, further processing using a jackknife-bootstrapping technique, show these results in the upper 10m are highly uncertain or less accurate (error bounds shown as triangles in the plot). Since for the layers below the results seem accurate, this suggests the uncertainty of the surface currents to be due to phenomena not included in the model itself, namely to the uncertainty in the atmospheric forcing components or due to unresolved dynamics like non-linear internal waves. A simple scale analysis comparing possible mean advection in the upper layers and the actual drifter displacement, rules out this second hypothesis. To assess the atmospheric forcing uncertainty, we can compare the COAMPS, LOKAL and ALADIN 10m wind fields for a snapshot at April, 10th, as displayed in Figure 4. This example shows the models were not consistent with each other. Furthermore, ALADIN estimates show the on-set of a northeasterly wind event, likely due to a diurnal breeze effect and contouring the land steep topography north of the study area. This wind pattern is consistent with the drifter displacement further to the west, as observed, and with the wind measurements made on-board of the RV Alliance.

By considering the ALADIN fields in the multi-model training procedure, drifter trajectory prediction significantly improved, as we can see in Figure 5. The areas of containment shown
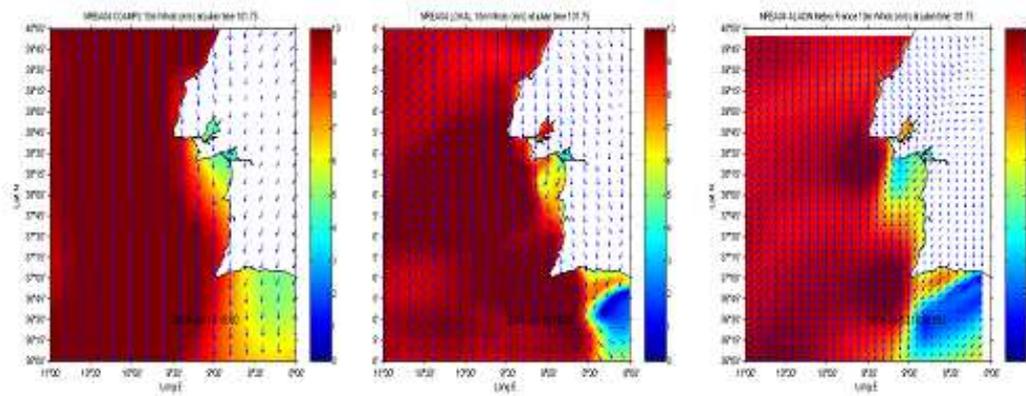
Figure 4: MREA04 atmospheric models 10m wind predictions for April, 10 at 1800. The left panel shows COAMPS estimates, the central the LOKAL and the right panel ALADIN fields. The COAMPS was running at 12km, the LOKAL and 6.5km and ALADIN at 4.5km resolutions.

in the left panel now include the actual trajectories for the central drifter and the overall statistics for the experiment, shown in the scatter diagram in the middle, and on the table on the right improve significantly. This result demonstrates this methodology can be only as good as the available models, such that by including more models we are in fact enlarging the vector-space of possible states, and hopefully improving local prediction skills.
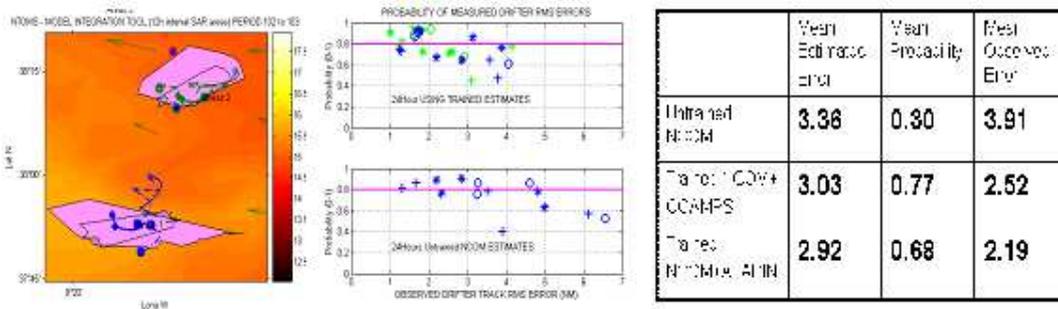
Figure 5: MREA04 results using NCOM, COAMPS and ALADIN combined. The panel in the left shows 12 hour interval areas of containment for the studied drifter tracks. The panel in the middle show the final statistics using untrained NCOM (below), NCOM and COAMPS (blue in upper plot) and NCOM combined with COAMPS and ALADIN (green dots). The table on the right shows the final statistics for the overall analysis period covering the month of April.

214

4 Final remarks

Ensemble training, combining operational models and data showed a strong potential for improving short term, local prediction. The examples discussed in this paper were focused in the surface drift problem, but the methodology has been successfully applied for sound speed and ocean state-variables estimation, with local and overall domain improvements (e.g. [13] and [8]). For analysis purposes, scatter plots were computed with the training error versus the probability of the forecast observed error being equal or smaller than the forecast error for sound speed profile estimation, using the several models and resolutions available during the MREA04 trial. The HOPS runs were performing local assimilation of profile data, hence starting with smaller bias error, while NCOM was only using sea surface temperature and large scale profile assimilation, through the outer nests. It is interesting to note that all models showed similar results after training with local temperature and salinity observations. Although, these results cannot be extended outside the observation locations and in time, since the training process did not include any information regarding state-variables covariances. To overcome this limitation, on-going research includes multiple run ensembles with perturbed physics and initial conditions through scaled analysis errors using techniques like the Ensemble Transform (ET) and Error Subspace Statistical Estimation (ESSE). Post-processing of the ensembles, combining data and model estimates, are also using Ensemble Transform Kalman Filter (Bishop, et.al, 2001) based approaches, similar to other ensemble data assimilation methods. These results will allow for future improvements and better identification of areas and periods impacted by local training.

Aknowledgement

References

[1] R. Allard, J. Dykes, Y. Hsu, J. Kaihatu, and D. Conley. A real-time nearshore wave and current prediction system. *Journal of Marine Systems*, 2007. doi:10.1016/j.jmarsys.2007.02.020.

[2] E. F. Coelho, M. Rixen, and R. Signell. NATO Tactical Ocean Modeling System: Concept applicability. Serial Report SR411, NATO Undersea Research Centre, 2005.

[3] E. F. Coelho and A. R. Robinson. *NATO Tactical Ocean Modelling: the mini-hops strategy in the mrea03 trial*, volume CP-47 of *NURC Conference Proceedings Series*. NATO, 2003.

[4] T. Gneiting, A. Westveld III, A. Raftery, and T. Goldman. Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation. Technical Report 449, Department of Statistics, University of Washington, 2004.

[5] J. Harding and J. Rigney. Operational oceanography in the U.S. Navy: A GODAE perspective. In *Ocean Weather Forecasting*, pages 467–482. Springer, NL, 2006. DOI 10.1007/1-4020-4028-8.

[6] J. Huthnance, H. Van-Aken, M. White, E. Barton, B. LeCann, E. F. Coelho, E. Fanjul, P. Miller, and J. Vitorino. Ocean margin exchange-water flux estimates. *Journal of Marine Systems*, 32:107–137, 2002.

[7] D. Ko, P. Martin, C. Rowley, and R. Preller. A real-time coastal ocean prediction experiment for MREA04. *Journal of Marine Systems*, 2007. doi:10.1016/j.jmarsys.2007.02.022.

[8] W. Leslie, A. R. Robinson, P. Haley Jr., O. Logutov, P. Moreno, P. Lermusiaux, and E. Coelho. Verification and training of real-time forecasting of multi-scale ocean dynamics for maritime rapid environmental assessment. *Journal of Marine Systems*, 2007. doi:10.1016/j.jmarsys.2007.02.001.

[9] X. S. Liang and A. R. Robinson. Localized multiscale energy and vorticity analysis I. Fundamentals. *Dynamics of Atmospheres and Oceans*, 2005. doi:10.1016/j.dynatmoce.2004.12.004.

[10] O. Logutov and A. R. Robinson. Multi-model fusion and error parameter estimation. *WMO DA symposium special issue, Q. J. R. Meteorol. Society.*, 131:3397–3408, 2006.

[11] N. Pinardi, I. Allen, E. Demirov, P. De Mey, G. Korres, A. Lascaratos, P.-Y. Le Traon, C Maillard, G. Manzella, and C. Tziavos. The Mediterranean ocean forecasting system: first phase of implementation (1998-2001). *Annales Geophysicae*, 21:3–20, 2003.

[12] M. Rixen and E. Coelho. The multi-scale multi-model super-ensemble cocktail: brake, shake and mix. serial report SR-444, NATO Undersea Research Centre, 2006.

[13] M. Rixen and E. Coelho. Operational prediction of acoustic properties in the ocean using multi-model statistics. *Ocean Modelling*, 11:428–440, 2006.

[14] A. R. Robinson and W. G. Leslie. dvanced systems for operational ocean forecasting of interdisciplinary fields and uncertainties. In *Maritime Rapid Environmental Assessment*, volume CP-47/CD-70-2003. NATO Undersea Research Centre, La Spezia, Italy, 2003.

Towards an hydrodynamic an biogeochemical operational model of the Portuguese coast

Paulo Leitão[*]   Guillaume Riflet[†]   Luís Fernandes[†]   Rodrigo Fernandes[†]

Angela Canas[†]   Pedro Pina[†]   Marcos Mateus[†]   Adélio Silva[*]

Ramiro Neves[†]

Abstract

This paper describes the on going efforts of IST (Technical University of Lisbon) and the company Hidromod in implementing a hydrodynamic and biogeochemical operational model of the Portuguese coast. It is presented the methodology used to downscale low-frequency pre-operational solutions (e.g Mercator) to the Portuguese coast scale. A special attention is given to the open boundaries conditions and to the model spin up. The methodology potential is illustrated describing the main applications done so far. There are several efforts underway to improve the presented model: (i) computational efficiency, (ii) spin up, (iii) data assimilation and (iv) validation.

1   Introduction

Operational oceanography has undergone rapid evolution in the last decade, especially in the framework of international projects like MFSTEP and Mersea. Until now, endeavours in the Atlantic Arc area have mainly focused on regional scales. However, increasing efforts have been made to implement operational systems also at coastal and local scales [17]. At these scales the universe of end users is larger, but Open Boundary Conditions (OBC) becomes a more complex problem. A way to overcome this difficulty is to downscale large scale operational systems. A consistent and clear definition of "operational model" is given by the American National Office for Integrated and Sustained Ocean Observations (Ocean.US - (`http://www.ocean.us/models`). In the words of Ocean.Us an operational model:

1. provides outputs at rates and in forms specified by decision makers and other user groups;

2. is well documented (detailed descriptions of underlying concepts, equations, simplifying assumptions, inputs and outputs are available);

---

[*]Hidromod SA, Portugal. First author e-mail: `paulo.chambel@hidromod.pt`.
[†]MARETEC, Secção de Ambiente e Energia - Dep. Mecânica, Instituto Superior Técnico, Portugal.

3. incorporates uncertainty into models for more useful, probabilistic predictions;

4. quantifies differences between model predictions and observations;

5. is operated by responsible organizations that ensure the quality and continuity of model predictions.

Having in mind a similar concept the company Hidromod and a research team of IST lead by Prof. Ramiro Neves decided to implement an operational hydrodynamic and biogeochemical model for the Portuguese coast. This is a long term project that involves a large team and a significant financial effort. A critical point is the interaction with partners with a complementary background. In this case the most relevant partnerships are with the Galician regional meteorological office (MeteoGalicia) with strong experience in weather forecast and with Mercator Ocean a French public company that provides global ocean forecasts. The operational model for the Portuguese coast aims to give support to fisheries, ecosystem health monitoring, and marine security. In this paper will be underline the efforts done to support ecosystem health monitoring. Hidromod and IST have managed and are managing several monitoring programs. The goal is to use operational model results as boundary conditions for several high resolution numerical applications being run to give support to the monitoring programs. In a first step only hydrodynamic forecast are being performed. In a second phase biogeochemical forecasts will be done for the entire Portuguese coast. A step in this direction is the INSEA EU project. This project is focused on the development of integrated management tools for coastal eutrophication assessment combining Models, Satellite Remote Sensing and in situ Measurements. INSEA aims to set-up and validate numerically robust ecological modelling systems in order to describe biogeochemical cycling of carbon and nutrients occurring under different hydrographical and trophic regimes, and to explore the system capabilities in a forecast mode to support coastal zone management issues. In the framework of this project Hidromod and IST are testing these concepts to the Tagus estuary and adjacent coastal area. In the future the tools being develop will be applied to the entire coast. In this paper we start by presenting the methodology being followed to implement an operational model. After, the potential of the methodology is illustrated with several results for different systems. Finally we discuss the work underway to improve the operational model being implemented for the Portuguese coast.

2   Materials and methods

2.1   MOHID system

Mohid Water Modelling System is a modular system created and developed by Instituto Superior Técnico and also used and improved by Hidromod. The development of MOHID started back in 1985 (for more details see `http://www.mohid.com`).

2.1.1 Hydrodynamic module

The MOHID hydrodynamic module solves the Navier-Stokes equations of a rotating fluid. The geophysical fluid is constrained to the hydrostatic and the Boussinesq approximations. The spatial discretization is done using a finite-volumes approach [14] similar to the one described by [6]. MOHID solves also a seawater density non-linear state equation depending on pressure, salinity and potential temperature. To calculate the turbulent vertical mixing, MOHID embeds GOTM [21]. By default the system uses a parametrization proposed in [5]. It is used a structure grid in the horizontal is an Arakawa C grid type. The vertical coordinate is a generic one, this way is possible choosing between several types of vertical discretizations (e.g. z-level, sigma and double-sigma coordinates). The baroclinic pressure gradient term is always calculated using a z-level approach, with a linear interpolation, in order to minimize spurious pressure-gradients [9]. A z-level vertical discretization was adopted for the 3D models. For the bottom layer a "partial step" approach was used instead of the traditional "full step" or "staircase" approach. The authors in [1] show that this approach minimizes the traditional problems associated with the staircase topography of z-level models ("staircase noise"). The temporal discretization is done using an alternate direction semi-implicit (ADI) method for the 2D mass balance equation (used to compute the SSH). For the 3D momentum (zonal and meridional velocities), heat and salt balance equations the vertical direction is computed implicitly while the horizontal directions are explicitly. The advection of momentum, heat and salt is computed using a total variation diminishing (TVD) scheme with a Superbee limiter. To dissipate high frequency noise is use a biharmonic filter for the velocities. The advantage of this methodology relatively to the Fickian diffusion is its ability to dissipate the high-frequency processes without significantly changing the lower frequency processes.

2.1.2 Biogeochemical modules

In the last years two basic water quality modules have been developed in the MOHID system for marine applications. Each model has its own level of detail and is best suited to specific systems:

1. WaterQuality module, initially developed using the US Environmental Protection Agency model. Despite successful improvements made in this code, the baseline philosophy has been rather untouched when it comes to nutrient cycles and biological/chemical processes. This model is best suited to applications in estuaries and coastal systems.

2. LIFE module [16], a detailed biogeochemistry pelagic model based on the ERSEM model. The model has a decoupled carbon-nutrients dynamics with an explicit parametrization of carbon, nitrogen, phosphorus and silica cycles. It considers a functional group approach with several groups of producers, consumers and decomposers. All living and

organic matter compartments of the model have variable stoichiometry, and the model also accounts for the synthesis of chlorophyll allowing a temporal and spatial variation of C:Chla ratios in producers populations.

2.2   Open boundary conditions

One of the principal conclusions of [19] and [13] was that to impose a consistent OBC it is necessary to have a good external solution. This importance is underlined in the review paper [4]. The methodology used here to define the OBC is similar to the one presented in [13]. A Flow Relaxation Scheme (FRS) applied to temperature (T), salinity (S) and velocities (U, V) [15] is combined with a radiation scheme from [8] for the barotropic mode. For the baroclinic mode a similar approach to the one presented in [13] was adopted in the MOHID system. In the MOHID case the user can choose to estimate the internal celerity based in the Orlanski approach or a constant celerity can be assumed. However both methodologies have problems, in the first case the celerity estimated based in the Orlanski method tends to be very noisy. In [4] the authors argue that the Orlanski-type conditions have considerable limitations due to the fact of Sommerfeld condition is justified only in the context of wave equations with a constant phase velocity. This is not the case of the internal waves. The authors in [10] have tested in schematic cases the baroclinic radiation estimating the internal celerity based in the Orlanski approach or assuming a constant celerity. The second approach present better results. However, estimating a constant celerity representative of the main internal wave energy propagating out of the domain with realistic density fields is a complex task. In the case of an application to the Algarve coast [11] it was assumed a internal celerity of 2 m/s. The authors in [18] estimated a value of approximately of 2 m/s for the celerity of internal waves associated with the first baroclinic Kelvin mode the celerity of internal waves for the California coast at the latitude of 39ºN, the same as this study. This approach give good results for simulations of 1 month time scale however for longer periods (1 year) there was the tendency to generation of noise near the North boundary. After testing the baroclinic radiation in different applications (schematic and realistic) estimating the celerity using the Orlanski approach or assuming a constant value, the IST and Hidromod modelling team concludes that is the best solution is not to radiate the baroclinic mode. This conclusion is similar to the one of [4]. These authors show similar experiences to ours and they justify them from the mathematical point of view why. Besides all the arguments presented an important aspect that must be highlighted is that when a Flow Relaxation Method is used to constrain density and the velocities a lot of the internal energy generated over the external data solution is damped by the nudging term. For the FRS boundary condition is usually assumed an exponential evolution of the relaxation period along 0.2^o was assumed, with 1 day in the boundary and 32 years out of the FRS area.

2.3 External solutions

The MOHID system allows the user to construct a tree of one-way nested models with no limitations on the number of nesting levels from the software point of view. By default, for each nesting level the reference solution for the OBC is the upper level in the MOHID nesting system. However, the user can add another solution linearly to the upper level model. The methodology usually followed consists in implementing two or more nesting levels. The first level (level 1) is 2D and tidal-driven. The second (level 2) and higher levels are 3D and baroclinic. The level 2 external solution results from adding linearly the level 1 solution plus a low frequency pre-operational solution like the ones provide by Mercator Ocean. For the higher levels the OBC is defined with a one-way nesting online with the immediately lower level. There is a great variety of global tidal solutions that can be use to define the open boundary of the level 1 (e.g. CSR4, FES2004, GOT00.2, NAO. 99b, TPXO6.2). These solutions became very common more or less 15 years ago. However, they only account for the effect of tide over SSH ad barotropic velocities. In the last years have also become available pre-operational models (Mercator, HYCOM-US, Topaz and FOAM) that can give a best estimate of the current state of the ocean low frequency processes.

2.4 Initial conditions and spin-up

The low frequency solution is adapted to the level 2 and higher levels grids with bilinear interpolation in the horizontal and linear interpolation in the vertical and in time. For the extrapolation procedure, the nearest neighbor approach was used. For the spin-up procedure, a methodology based on a slow connection of the forcing terms (baroclinic force, wind stress) was implemented. This methodology consists of defining an initial condition where the initial fields of salinity and temperature are interpolated from the low frequency solution, a null velocity field is assumed, and a (Sea Surface Height) SSH field with null gradient is also considered. A coefficient that varies linearly between 0 and 1 along the "connection" period of 5 days is multiplied by the baroclinic force and wind stress. Because the forces are slowly connected, the velocity reference solution of the OBC also needs to be slowly connected. The nudging term in the momentum equation is multiplied by a coefficient. In this way, the velocity field near the boundary also converges slowly to the reference solution. To minimize the perturbations suffered by the initial condition of salinity and temperature along the spin-up period, a relaxation time scale variable in time was also assumed for these properties. The idea is to assume a relaxation period that increases with time; in this way, in the beginning of the run the temperature and salinity fields have a stronger nudging when the external and internal activity is more intense due to the spin-up process. At the end of the spin-up period, the nudging in the model interior (out of the FRS area) is null. For the forces, a connection coefficient was assumed with a linear evolution over 5 days. For the reference

solution a quadratic evolution was imposed. This evolution was chosen based on a sensitivity analysis. This evolution allows, in the first instance, strong nudging across the entire domain. With time, the model tends to be free except in the flow relaxation scheme area. A 10-day connection period to the reference solution was assumed.

2.5   Atmospheric forcing

There is a problem related with the surface boundary condition. National (NOAA) and trans-national (ECMWF) Institutes provide atmospheric global solutions that can be used to force ocean models. These solutions do not have the resolution necessary to simulate high resolution atmospheric processes characteristic of the land/sea interface; an example is the wind curl associated with near shore topographic features. Along the Portuguese coast there are 3 areas where, in northerly winds, the wind curl is intense: Cape Raso, Sintra [10]; Cape Espichel, Arrábida; and the zonal ridge along the Algarve coast at Cape São Vicente [20]. One way is to downscale these global atmospheric solutions with mesocale numerical weather prediction systems. For the Portuguese applications the model is forced with high resolution atmospheric momentum and heat flux fields, results from an atmospheric modelling system based on the MM5 model [7]. IST runs a three domain operational system for forecast weather conditions for Portugal. The first domain runs with a resolution of 81 km, the second with 27 km and the third with 9 km. The OBC of the outer domain is defined using the global forecast provided by NCMRWF. The MM5 fields are interpolated into hourly fields for the MOHID coastal applications using triangulation interpolation in space and linear interpolation in time. The boundary conditions of vertical diffusion of momentum and turbulent kinetic energy are computed based on wind velocity fields stemming from the atmospheric model. Wind stress is computed function of the wind velocity. Atmospheric pressure forcing is also considered. Heat fluxes are computed using atmospheric model results for the following parameters: wind velocity, air temperature, solar radiation, relative humidity, cloud cover and downward long-wave radiation.

3   Main applications

The methodology presented earlier was and is being tested in several applications. The first application done aimed to simulate the Algarve coastal circulation using realistic forcing (viz., low frequency circulation, tide, high resolution atmospheric forcing). In the frame work of INSEA project the methodology tested for the first time for the Algarve coast is being extended to the entire Portuguese coast. IST is the leader of this project and Hidromod is one of the partners. One of the goals of this application is to generate realistic boundary conditions to very high resolution models run for in the framework scientific projects and monitoring projects. In the case o the INSEA project is run a high resolution model for

the Tagus area to study eutrophication tendencies in this area. MeteoGalicia is also running in a pre-operational mode the MOHID system for the Galician coast nested to the Spanish regional system ESEOAT develop in the framework of the Spanish project ESEOO (`http://www.eseoo.org`). The Azores University is also implementing a system very similar to the one being tested for the Portuguese coast but in this case in the framework of the Interreg project EASY also headed by IST with the consultancy support of Hidromod (`http://www.project-easy.info/`).

3.1 Algarve Coast

In the Algarve coastal application the low frequency open boundary conditions were defined via downscaling of the HYCOM-US operational solution for the Gulf of Cadiz. Atmospheric forcing was imposed using the MM5 high resolution mesoscale model (9 km resolution near the coast). A 3-level nesting system based on the MOHID numerical system was implemented. The higher nesting level had a horizontal resolution of $0.02^{\underline{o}}$ along the Algarve coast. The methodology was first validated qualitatively. A comparison between the numerical results and the conceptual model of the circulation described in the literature was made. A quantitative validation was also performed, based mainly on remote sensing data (Sea Surface Temperature - SST and altimetry) available for July 2004. The numerical system was able to reproduce many of the circulation features described in the literature (e.g.: Azores current recirculation in the Gulf of Cadiz, the upwelling jet, Mediterranean Water undercurrent, Mediterranean Outflow splitting, generation of meddies) and observed with remote sensing data (e.g., the signature in SST during a regime of upwelling relaxation). A detail description of this work was published in [11]. In areas like the Cape São Vicente the variability of primary production in summertime is strongly correlated to the intensity of upwelling. If the physical/biogeochemical model is able to reproduce the variability of surface chlorophyll, this means that the main biogeochemical processes as well as the horizontal and vertical transport and mixing are properly simulated. Having this joint validation in mind the hydrodynamic model was run coupled to a biogeochemical model (NPZ type). The model was able to capture the main chlorophyll patterns presented in the chlorophyll Modis data with 600 m ground resolution for July 2004.

3.2 Portuguese Coast

In the INSEA project application the nesting procedure is divided into two systems. In the first one the outer open boundary conditions are defined using has external data low-frequency and high-frequency large scale solutions. A three level nesting model is implemented. The first level (level 1) is a 2D barotropic tidal-driven model, and has variable horizontal resolution ($0.02^{\underline{o}}$-$0.04^{\underline{o}}$). This model domain covers most of the Atlantic coast of Iberia and Morocco.

The second (level 2) and third (level 3) levels are 3D baroclinic models. The second level model has 0.06^o horizontal resolution and the domain is similar to the level 1 domain. The third level has 0.02^o horizontal resolution and includes the Tagus Promontory area. The OBC for level 2 is defined by adding to the solution of level 1 (high frequency) the low frequency Mercator solution [3], which for this area has a resolution of 0.05^o. The surface boundary condition is imposed using high resolution results from the MM5 system run by IST. This system has a resolution of 9 km near the shore. This application is intended as a first step towards a future coastal operational model for the Portuguese coast, optimized to be of service to the tourism and fishing industries, which are two of the main economic activities in this region. This first nesting system is responsible for generating a realistic solution with a horizontal resolution of 0.02^o in the Tagus estuary area of influence. The second system aims to generate local solutions with horizontal resolution of 0.001^o. This system as a more flexible nature and aims to be used to generate very easily new nesting levels to answer to specific questions like determining the impact of the Guia submarine outfall (locate 2 km off-shore of the city of Cascais). The open boundary conditions are defined using an off-line nesting approach with the higher resolution solution produced by the first operational system described above. In the case of the biogeochemical component the solution adopted to nest the Tagus local application is to impose the climatologic atlas publish by NOAA for the nutrients using a nudging approach in the outer boundary. This application assumes that the only source of variability for the producers and decomposers is located inside the simulated area.

4   Future work

There are four areas where the methodology presented earlier is being improved: (i) computational efficiency, (ii) spin up, (iii) data assimilation and (iv) validation.

4.1   Computational efficiency

To improve computational efficiency the MOHID parallelization capabilities are being upgrade. The present version is able to run each nesting level in a different processor using MPI (message passing interface). The next version will be able to parallelize each nesting level using domain decomposition.

4.2   Spin up

The spin up method used consists in connecting slowly the forcing terms. Other spin up methods will be tested with the aim of reducing the spin up period. The alternative methods identified so far are the Dolph-Chebyshev digital filter [12]. An adaptation of this method consists in do a first free run (no slow connection of the forcing terms) during a short period

(e.g. 1 day). In a second step the noise of this first run is filter out. These spurious waves are associated with dynamic inconsistencies introduce in the solution by standard interpolation methods (e.g. triangulation) that do not conserve dynamic relations between properties. Finally the filter out solution is use as an initial condition. Another method is the Variational Initialization and FOrcing Platform (VIFOP) [2] this method consists in doing a first interpolation using a standard method (e.g. triangulation in horizontal and linear in vertical). In a second step is minimized a cost function that aims to preserve specific dynamic relations (e.g. geostrophic balance) and minimize the corrections to the interpolated field.

4.3   Data assimilation

A kalman filter assimilation module has been implemented in the MOHID system. To test this module the follow steps will be followed:

- Characterization of the error fields of the Tagus estuary model through comparison with available measurements;

- Assimilation of tide gauge water levels in the Tagus estuary model using the most appropriate assimilation configuration(s) from the SEEK method studied;

- Assimilation of salinity, temperature and chlorophyll in the Tagus estuary model using the most appropriate assimilation configuration(s) from the SEEK method studied;

- Assimilation of along track satellite altimetry in the Tagus costal area model using the most appropriate assimilation configuration(s) from the SEEK method studied.

In a second phase the kalman filter module will be tested in the Portuguese coast application.

4.4   Validation

The systematic validation of the Portuguese coastal operational model is being design considering remote sensing as the main source of field data (SST, SSH and chlorophyll). In the case of chlorophyll the level of uncertainty can be high. To reduce this uncertainty the MOHID developers aim to simulate explicitly the radiance distributions measure by the sensors from where is inferred the chlorophyll at the surface. The idea is to incorporate in the MOHID system the Hydrolight/Ecolight radiative transfer model. This numerical model computes radiance distributions and related quantities (irradiances, reflectances, diffuse attenuation functions, etc.) in the ocean. The model is based on the water absorption and scattering properties, the sky conditions, and the bottom boundary conditions computes the in-water light field and other quantities of interest to optical oceanographers, such as the water-leaving radiance and remote-sensing reflectance.

225

Acknowledgments

References

[1] A. Adcroft, C. Hill, and J. Marshall. Representation of topography by shaved cells in a height coordinate ocean model. *Monthly Weather Review*, 125(9):2293–2315, 1997.

[2] F. Auclair, P. Marsaleix, and C. Estournel. Truncation errors in coastal modelling : evaluation and reduction by an inverse method. *J. Atmos. Oceanic Technol.*, 17:1348–1367, 2000.

[3] P. Bahurel, P. De Mey, T. De Prada, E. Dombrowsky, P. Josse, C. Le Provost, P. Y. Le Traon, A. Piacentini, and L. Siefridt. Mercator, forecasting global ocean. *AVISO Altimetry Newsletter*, 8:14–16, 2001.

[4] E. Blayo and L. Debreu. Revisiting open boundary conditions from the point of view of characteristic variables. *Ocean Modelling*, 9:231–252, 2005.

[5] V. M. Canuto, A. Howard, Y. Cheng, and M. S. Dubovikov. Ocean turbulence Part I: one-point closure model. Momentum and heat vertical diffusivities. *Journal of Physical Oceanography*, 31:1413–1426, 2001.

[6] P. C. Chu and C. Fan. Finite volume ocean circulation model. a terrain-following crystal grid finite volume ocean circulation model (2002/4). *IEEE Journal of Oceanography*, 60:945–952, 2004.

[7] J. Dudhia, D. Gill, K. Manning, W. Wang, and C. Bruyere. *PSU/NCAR Mesoscale Modelling System, Tutorial Class Notes and Users Guide (MM5 Modelling System Version 3)*. Mesoscale and Microscale Meteorology Division, National Center for Atmospheric Research, Boulder, Colorado, USA, 2004. http://www.mmm.ucar.edu/mm5/documents/tutorial-v3-notes.html.

[8] R. A. Flather. A tidal model of the northwest european continental shelf. *Mem. Soc. R. Sci. Liege*, 10(6):141–164, 1976.

[9] N. Kliem and J. D. Pietrzak. On the pressure gradient error in sigma coordinate ocean models: A comparison with a laboratory experiment. *Journal of Geophysical Research*, 104(C12):29781–29800, 1999.

[10] P. Leitão. *Integration of Scales and Processes in the marine Environment Modelling*. PhD thesis, Technical Superior Institute, Lisbon, 2003.

[11] P. Leitão, H. Coelho, A. Santos, and R. Neves. Modelling the main features of the algarve coastal circulation during july 2004: A downscaling approach. *Journal of Atmospheric & Ocean Science*, 10(4):421–462, 2005.

[12] P. Lynch, D. Giard, and V. Ivanovici. Improving the efficiency of a digital filtering scheme. *Monthly Weather Review*, 125(8):1976–1982, 1997.

[13] P. Marchesiello, J. C. McWilliams, and A. Shchepetkin. Open boundary conditions for long-term integration of regional oceanic models. *Ocean Modelling*, 3:1–20, 2001.

[14] F. Martins, R. Neves, P. Leitão, and A. Silva. 3d modeling in the sado estuary using a new generic coordinate approach. *Oceanologica Acta*, 24:S51–S62, 2001.

[15] E. A. Martinsen and H. Engedahl. Implementation and testing of a lateral boundary scheme as an open boundary condition in a barotropic ocean model. *Coastal engineering*, 11(5–6):603–627, 1987.

[16] M. Mateus. *A process-oriented biogeochemical model for marine ecosystems: development, numerical study, and application.* PhD thesis, Instituto Superior Técnico, Lisbon, Portugal, 2006.

[17] D. Obaton, M. Garcia Sotillo, M. Ruiz Villareal, P. Montero, P. Chambel Leitao, R. Fernandes, Y. H. De Roeck, and J. White. Biscay Iberian and Irish shelves OOS physical modelling and assimilation team. State of art and plans. In *Proceedings of the 4th EUROGOOS Conference*, Brest, France, June 6–9 2005.

[18] L. Y. Oey. A forcing mechanism for the poleward flow off the southern California coast. *Journal of Geophysical Research*, 104(C6):13529–13539, June 15 1999.

[19] E. D. Palma and R. P. Matano. On the implementation of passive open boundary conditions for a general circulation model: The three-dimensional case. *Journal of Geophysical Research*, 105:8605–8627, 2000.

[20] P. Relvas and E. D. Barton. A separated jet and coastal counterflow during upwelling relaxation off Cape São Vicente (Iberian Peninsula). *Continental Shelf Research*, 25:29–49, 2005.

[21] L. Umlauf and H. Burchard. Second-order turbulence closure models for geophysical boundary layers. A review of recent work. *Continental Shelf Research*, 25:725–827, 2005.

Global ocean state estimation for climate research and applications

Rui M. Ponte* and the ECCO-GODAE group at MIT-AER†

Abstract

Understanding the role of the ocean in climate variability requires accurate knowledge of its 3-dimensional circulation and associated transports of heat, freshwater and other properties on a global scale. Here we describe on-going efforts in global ocean state estimation being pursued at MIT-AER as part of the ECCO-GODAE (Estimating Circulation and Climate of the Ocean-Global Ocean Data Assimilation Experiment) project.

The ECCO-GODAE state estimates ([8]) are produced by fitting, in a least-squares sense, the MIT general circulation model (MITgcm; [3, 2]) to most available in situ and satellite ocean observations (altimetric sea level records, space-based gravity measurements, sea surface temperature, subsurface temperature and salinity data from various instruments including modern Argo float profiles, scatterometer winds, etc.). First guess atmospheric forcing fields are those from the reanalysis products provided by the US National Center for Atmospheric Research-National Centers for Environmental Prediction and the analysis coincides with the modern instrumental period (1992–present).

The optimization procedure is based on the method of Lagrange multipliers ([8]) and makes use of the MITgcm adjoint codes to obtain estimates that are a "best" fit to all the data within known model and data errors (e.g., [5, 1]). Parameters currently adjusted in the optimization include surface atmospheric fluxes of momentum, heat and freshwater, and initial temperature and salinity fields. The optimized solutions are fully consistent with the model dynamics and its conservation principles of heat, momentum and other quantities, and can thus be readily probed for insight into the dynamics and forcing of the variable currents, density fields, sea level, and any other variables of interest (e.g., [7],[8],[4],[6]).

In this talk, we briefly review the methodology and various components (MITgcm, datasets, objective function, data weights) involved in the state estimation procedure, and explore the available solutions in the context of climate issues of major societal interest. Examples to be discussed include the determination of the variability in the Atlantic meridional overturning circulation at periods from days to years (e.g., [7]), and

---

*Atmospheric and Environmental Research, E-mail:rponte@aer.com.

†Principal Investigator: Carl Wunsch (Massachusetts Institute of Technology); Constantinos Evangelinos, Gael Forget, Patrick Heimbach, Charmaine King, Matthew Mazloff and Diana Spiegel (MIT); Sergey Vinogradov and Nadya Vinogradova (AER).

the study of decadal trends in regional and global mean sea level and how they relate to steric effects over the full water column and net heat and freshwater transfers at the air-sea interface (e.g., [9]).

We also preview how the ECCO-GODAE estimates will evolve in the next few years. Present plans involve increases in model grid resolution, expansion of the modeled domain to include the Arctic Ocean, implementation of a fully coupled ocean/sea-ice model and improved parameterizations of subgrid-scale processes, introduction of new datasets, and incorporation of new parameters in the control vector that is adjusted as part of the optimization.

References

[1] G. Forget and C. Wunsch. Estimated global hydrographic variability. *J. Phys. Oceanogr.*, 37(8):1997–2008, 2007. doi:10.1175/JPO3072.1.

[2] J. Marshall, A. Adcroft, C. Hill, L. Perelman, and C. Heisey. A finite-volume, incompressible Navier-Stokes model for studies of the ocean on parallel computers. *J. Geophys. Res.*, 102:5753–5766, 1997.

[3] J. Marshall, C. Hill, L. Perelman, and A. Adcroft. Hydrostatic, quasi-hydrostatic, and nonhydrostatic ocean modeling. *J. Geophys. Res*, 102(C3):5733–5752, 1997.

[4] R. M. Ponte, K. J. Quinn, C. Wunsch, and P. Heimbach. A comparison of model and GRACE estimates of the large-scale seasonal cycle in ocean bottom pressure. *Geophys. Res. Lett.*, 34(L09603), 2007. doi:10.1029/2007GL029599.

[5] R. M. Ponte, C. Wunsch, and D. Stammer. Spatial mapping of time-variable errors in Jason-1 and TOPEX/POSEIDON sea surface height measurements. *J. Atmos. Ocean. Technol.*, 24(6):1078–1085, 2007. doi:10.1175/JTECH2029.1.

[6] S. V. Vinogradov, R. M. Ponte, P. Heimbach, and C. Wunsch. The mean seasonal cycle in sea level estimated from a data-constrained general circulation model. J. Geophys. Res., in press, 2007.

[7] C. Wunsch and P. Heimbach. Decadal changes in the North Atlantic meridional overturning and heat flux. *J. Phys. Oceanogr.*, 36(11):2012–2024, 2006. doi:10.1175/JPO2957.1.

[8] C. Wunsch and P. Heimbach. Practical global ocean state estimation. *Physica D*, 230(1-2):197–208, 2007. doi:10.1016/j.physd.2006.09.040.

[9] C. Wunsch, R. M. Ponte, and P. Heimbach. Decadal trends in sea level patterns: 1993–2004. *J. Climate*, 20(24):5889–5911, 2007.

Miscellaneous Articles

Study of crack propagation behavior in the integrally stiffened panel and conventional stiffened panel

Muhammad Adeel*

Abstract

The crack growth characteristics of the integrally stiffened panel and conventional stiffened panel based on fracture mechanics and finite element analysis are investigated. A fourteen-stringer stiffened panel is analyzed for a central skin crack propagating towards the adjacent stringers. Stress intensity factors and fatigue crack propagation rates of both types of the stiffened panels are then compared. The analyses result show that integral stiffening cause higher stress intensity factor than conventional stiffened panel as the crack tip passes through the stringer and the integrally stiffened panel has less damage tolerance capability than the riveted stiffened panel.

Keywords: Integrally Stiffened Structure, Conventional Stiffened Structure, Stress Intensity Factor, Damage Tolerance, Finite Element Analysis.

1 Introduction

The continual need for low acquisition cost and the emergence of high speed machining and other technologies has brought about a renewed interest in large scale integral metallic structures for aircraft applications. The high performance levels in machines and equipment continue to place more exacting demands on the design of structural components. In aircraft, where weight is always a critical problem, integrally stiffened structures have proved particularly effective as a lightweight, high strength construction. Integral fuel tanks and pressurized shells e.g. wing fuel tank, cabin pressurized fuselage usually create sealing problems for a riveted structure. These problems are eliminated to a large extent by integral stiffeners. Other advantages of integral stiffened structures over riveted panels are improved performance through smoother exterior surfaces by reduction in number of attachments and nonbuckling characteristics of skin, increase in allowable stiffener compression loads by elimination of attached flanges and increased joint efficiencies under tension loads through the use of integral

---

*Master Student, Beihang University, Beijing, China. E-mail: `adeelned2006@gmail.com`.

doublers, etc. Components in which the integrally stiffened structures can be employed are jet engine components, aircraft and missile fuselages, jet engine pods, aircraft landing brake components, components with compound curvatures such as bulkheads and fairings, flat and curved panels, conical and other shell-like components. Nevertheless, applications of low cost, large -scale integral structures in damage tolerance critical areas such as the fuselage have been inhibited by a perceived lack of damage tolerance and by cost and manufacturing risks associated with size and complexity of the parts.

In the Integral Airframe Structures (IAS) Program, a feasible integrally stiffened fuselage concept was developed and analyses and tests were run to validate equal or better performance than conventional designs with regard to weight and structural integrity, while achieving a significant reduction in manufacturing cost [2]. In NASA report Linear elastic analysis of integrally stiffened panel with circumferential crack was performed using NASTRAN and STAGS. Stiffeners and the pad-up region were all modeled in detail using shell elements .Linear elastic analysis results in straight line plots of load versus K_1 at various crack lengths. It was found that as the crack approaches the thickness discontinuity at the stiffener base the stress intensity increases little with crack length. The analysis was also run with both NASTRAN and STAGS for a half crack length corresponding to a crack extending midway through the stiffener base on each side. Excellent agreement was found between two codes. While no further refinement of the model was done to more accurately evaluate the stress intensity in the vicinity of the thickness interface. The fracture analyses were conducted on the FAA/NASA stiffened panels using the STAGS code with the critical crack-tip-opening angle (CTOA) fracture criterion [4]. Comparisons were made between load-crack extension on stiffened and unstiffened panels with single cracks and multiple site damage (MSD). An assessment of the capability of the STAGS code with the critical CTOA failure criterion to predict residual strength was made. The results of residual strength pressure tests and nonlinear analyses of stringer-and frame-stiffened aluminum fuselage panels [5] with longitudinal cracks are presented. Two damage conditions are considered: a longitudinal crack located midway between stringers and a longitudinal crack adjacent to a stringer and along a row of fasteners in a lap joint that has multiple- site damage (MSD). At present it is readily possible to calculate the fatigue crack behavior, and particularly the residual strength of stiffened panels, to a reasonable degree of accuracy, provided the behavior of an unstiffened sheet of similar size is known [1]. The latter requirement does not set serious restrictions to the technical applicability of the method. Poe [3] predicted the crack growth behavior of a stiffened panel on the basis of unstiffened panel data, and compared his predictions with actual test data obtained from stiffened panels of different geometries. Poe's work contains many interesting results. It shows that a light stringer causes a smaller deceleration of crack growth, since it brings about less reduction of the stress intensity factor.

# 2   Description of stiffened panel geometry

This paper focuses on metallic integrally and conventional stiffened aircraft fuselage panels. Both types of the stiffened panels are constructed entirely of aluminum alloy. The structural configuration considered is shown in Figure 1. The overall dimensions of the panel include a 1.98 m length and a 2.642 m width. The skin is stiffened by 14-longitudinal Z-section stringers and two horizontal tear straps with a thickness of 4.32 mm. In the integrally stiffened panel, stringers are integrally machined with skin while a single line of rivets is used to attach stringer to the skin in conventional stiffened panel. The rivet spacing of one inch is used in the conventional stiffened panel.
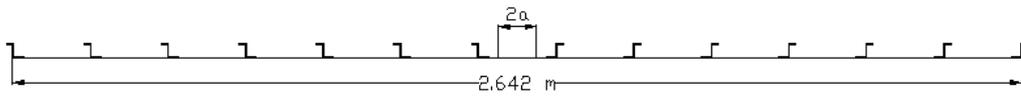
(a) Model for 14-stringer panel
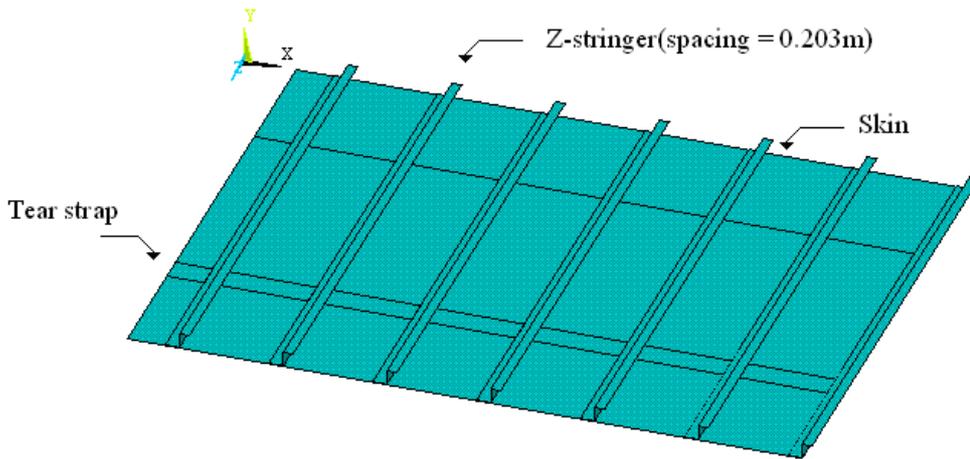
(b) $\frac{1}{4}$ Geometric model for 14-stringer panel
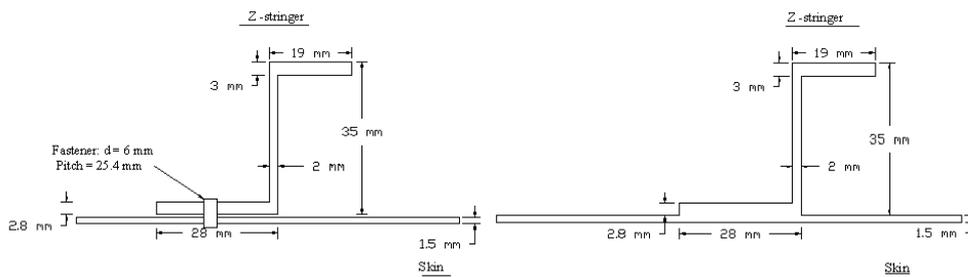
(c) Stringer riveted with skin

(d) Stringer integrally stiffened with skin

Figure 1: Geometry of integrally stiffened panel and riveted stiffened panel.

The physical properties of the model and loading are given as under.

Material properties:

$$E = 70\text{GPa}$$
$$\nu = 0.3$$
$$C = 5 \times 10^{-11}$$
$$m = 3$$

Uniform tensile load:

$$P = 222.411KN$$

3 Finite element model description

A quarter of both types of the stiffened panels with symmetric boundary conditions is analyzed using finite element analyses code ANSYS shown in Figure 2 and Figure 3. The SHELL181 element is used to model the skin and stringer in detail in both types of the stiffened panels. It is a 4-node element with six degrees of freedom at each node, translations in the x, y, and z directions and rotations about the x, y, and z-axes. It is finite strain element and is suitable for analyzing thin to moderately-thick shell structures. SHELL181 is well-suited for linear, large rotation, and/or large strain nonlinear applications. The rivets in the conventional stiffened panel are modeled using ANSYS spot weld feature based on the internal multipoint constraint (MPC) approach. This feature allows to model thin sheet components that are connected with spot welds, rivets, or fasteners. Singular elements are used around crack tip to pick up the singularity in the strain and stress at the crack tip. A circumferential crack is considered in the center of skin of both types of the stiffened panels and the same distributed tensile load is applied. The linear elastic static analysis is performed and stress intensity factor is calculated for both types of the stiffened panels. Similarly, stress intensity factor is obtained for various crack lengths while keeping same loading conditions. Stress intensity factor is also calculated for the unstiffened panel shown in Figure 4 for different crack length which has the same cross section area as the stiffened panel. A graph of stress intensity factor (K_1) vs. half crack length (a) is plotted for both types of the stiffened panels and the unstiffened panel shown in Figure 7. The rate of fatigue crack propagation (da/dn) is calculated using Paris law. Similarly, half crack length is plotted as function of rate of fatigue crack propagation shown in Figure 8.
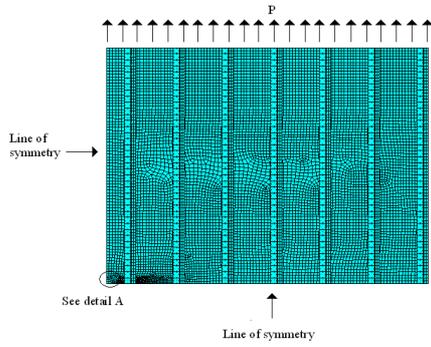
Figure 2: $\frac{1}{4}$ FE model of riveted stiffened panel
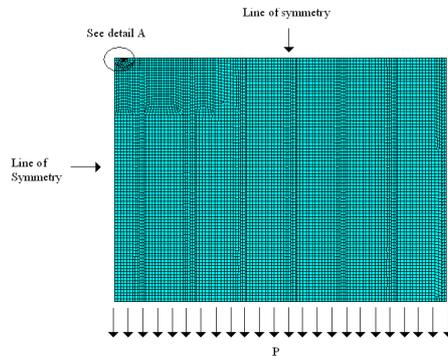
Figure 3: $\frac{1}{4}$ FE model of integrally stiffened panel
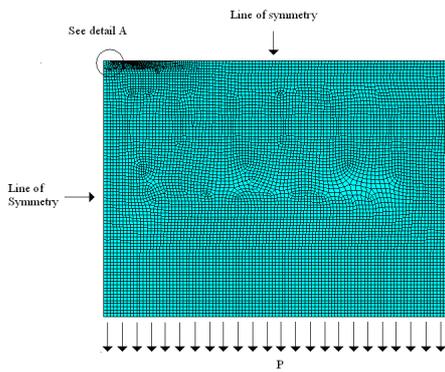
Figure 4: $\frac{1}{4}$ FE model of unstiffened panel


Figure 5: Detail A

237
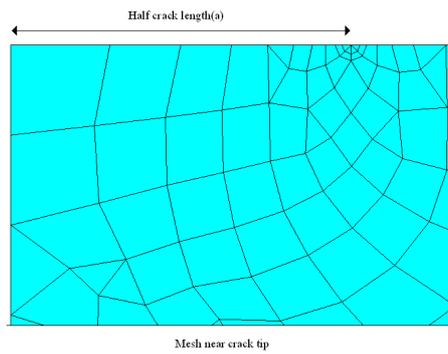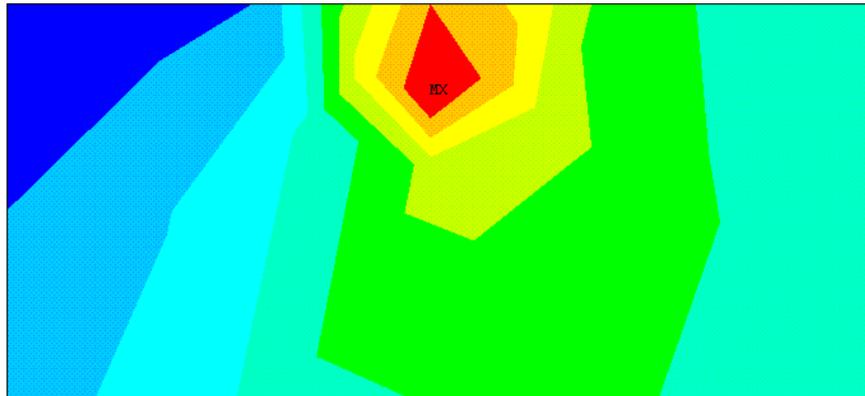
Figure 6: Von Mises stress around crack tip.
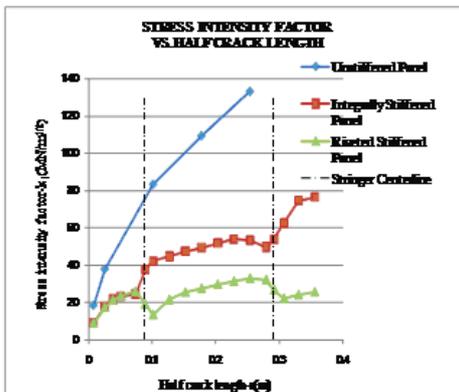
Figure 7: Graph of stress intensity factor versus half crack length for riveted, integral and unstiffened panels.
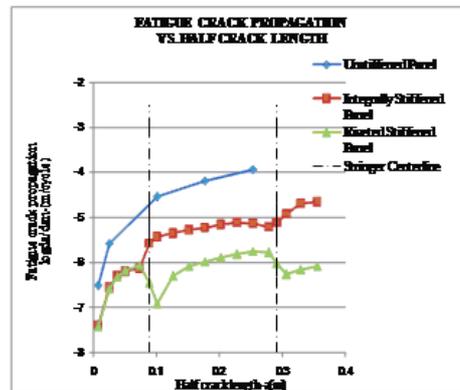

Figure 8: Graph of fatigue crack propagation versus half crack length for riveted, integral and unstiffened panels.

4 Results discussion and concluding remarks

The results depicted in Figure 7 indicate that as long as the crack tip is far from stringer the stress intensity factor increases with same rate in both types of the stiffened panels with increasing crack length. However, in the integrally stiffened panel the stress intensity factor increases while in the conventional stiffened panel stress intensity factor decreases as the crack passes through stringer. Similar trend is observed as the crack tip approaches the second stringer in both types of the stiffened panels. When crack has passed the stringer, the stiffening effect decreases, K_1 increases and so does da/dn. This trend is reflected by the plots in Figures 7 and 8. The finite element results of conventional stiffened panel show a

238

rise in von Mises stresses in the stringer as the crack tip passes through it. In the case of integral stiffeners the stiffening element cracks simultaneously with the skin. The results in Figure 8 show that integral stiffening causes less deceleration of crack growth compared to conventional stiffened panel in the vicinity of the stringer. It is also found that the value of stress intensity factor in unstiffened panel is higher than that in the integrally stiffened panel for the same crack length. The finite element analysis method used in this paper is verified by the results given in reference [2].

The results presented in this paper indicate that integral stiffening causes higher stress intensity factor than conventional stiffened panel as the crack tip passes through the stringer. Consequently we can conclude that in the case of integrally stiffened panel higher stresses will be developed around crack tip, it can be predicted from the above graphs that the integrally stiffened panel has less damage tolerance capability than the riveted stiffened panel.

References

[1] D. Broek. *Elementary Engineering Fracture Mechanics*. Martinus Nijhoff Publishers, 1982.

[2] R. G. Pettit, J. J. Wang, and C. Toh. Validated feasibility study of integrally stiffened metallic fuselage panels for reducing manufacturing costs. Technical Report CR-2000-209342, NASA, 2000.

[3] C. C. Poe. Crack propagation in stiffened panels. ASTM STP 486, 1971.

[4] B. R. Seshadri, J. C. Newman Jr., D. S. Dawicke, and R. D. Young. Fracture analysis of the FAA/NASA wide stiffened panels. Technical Report TM-1998-208976, NASA, 1998.

[5] R. D. Young, M. Rouse, D. R. Ambur, and J. H. Starnes Jr. Residual strength pressure tests and nonlinear analyses of stringer- and frame-stiffened aluminum fuselage panels with longitudinal cracks. Technical Report 19990021207, NASA, 1998.