16th Brazilian Congress of Mechanical Engineering

Engineering for the New Millennium

NONLINEAR AND CHAOTIC PHENOMENA IN ENGINEERING

ISBN 85-85769-06-6

Non Linear and Chaotic Phenomena in Engineering

TRB0104 5
ON LOCAL STABILITY ANALYSIS OF A NON-IDEAL PROBLEM
Dantas, Marcio Jose Horta
Balthazar, Jose Manoel

TRB0105 12
IDENTIFYING PERIODIC AND CHAOTIC EXPERIMENTAL TIME SERIES
Savi, Marcelo Amorim
Franca, Luiz Fernando Penna

TRB0157 21
NUMERICAL INVESTIGATION OF CHAOTIC MOTION IN A SHAPE MEMORY TWO-BAR TRUSS
Savi, Marcelo Amorim
Braga, Arthur Martins Barbosa
Pacheco, Pedro M.C.L.

TRB0160 29
BIFURCATIONS AND CRISES IN A SHAPE MEMORY OSCILLATOR
Savi, Marcelo Amorim
Machado, Luciano Guimarães
Pacheco, Pedro M. C. L.

TRB0165 39
TIME SERIES PREDICTION FROM A POLYNOMIAL MODELING: APPLICATION TO THE NONLINEAR PENDULUM
Savi, Marcelo Amorim
Pinto, Eduardo Gomes F.

TRB0247 49
LOWER BOUNDS OF DYNAMIC BUCKLING LOADS USING MELNIKOV CRITERION
Santee, Donald Mark
Gonçalves, Paulo Batista

TRB0301 59
NONLINEAR DYNAMICS OF A PSEUDOELASTIC OSCILLATOR USING A CONSTITUTIVE MODEL BASED ON PLASTICITY
Savi, Marcelo Amorim
Mamiya, Edgar N.
Viana, Dianne M.
Pacheco, Pedro M.C.L.

TRB0331 67
A NEW TREATMENT FOR RIGID BODIES COLLISION MODELS
Cataldo, Edson Luiz
Sampaio, Rubens

TRB0348 76
TRANSFER FROM THE EARTH TO THE HALO ORBITS WITH MINIMUM FUEL CONSUMPTION
Prado, Antonio Fernando Bertachini de Almeida
Correa, Annelisie Aiex
Stuchi, Terezinha

TRB0350 85
MINIMUM FUEL TRANSFERS BETWEEN HALO ORBITS
Prado, Antonio Fernando Bertachini de Almeida
Felipe, Gislaine
Beauge, Cristian

TRB0351 95
ORBITAL CONTROL OF A SATELLITE USING THE GRAVITY OF THE MOON
Prado, Antonio Fernando Bertachini de Almeida

TRB0353 104
AN ERROR ANALYSIS OF THE GRAVITATIONAL CAPTURE TRAJECTORIES
Prado, Antonio Fernando Bertachini de Almeida
Vieira-Neto, Ernesto

TRB0615 114
ON NON-IDEAL VIBRATIONS OF A NONLINEAR PORTAL FRAME: ANALYSIS OF THE PASSAGE THROUGH RESONANCE
Felix, Jorge
Balthazar, José Manoel
Brasil, Reyolando

TRB0981 123
DETECTING FIXED AND PERIODIC ORBITS IN CHAOTIC SYSTEMS
Roth, Bruno
Macau, Elbert Einstein Nerher
Terra, Maisa de Oliveira

TRB0982 131
NUMERICAL SIMULATION AND VISUALIZATION OF THE PROCESS OF EXTRACTION IN INDUSTRIAL INSTALLATIONS "CROWN-MODEL"
Cavalheiro, Marcos R. M.
Piva, Claudia

TRB0989 141
SHAPE AND STABILITY OF BALLOON ON THE TWO-FOR-ONE TWISTER
Marcicano, Joao Paulo Pereira
Tu, Carlos C. C.

TRB1003 150
SOME EXPERIMENTAL FINDINGS ON THE PLASTIC BUCKLING OF SHELLS UNDER AXIAL IMPACT LOADS
Alves, Marcílio
Barbosa, Giancarlo Micheli

TRB1209 156
IMPLEMENTATION OF THE LYAPUNOV FUNCTIONAL IN DIFFERENCE SCHEMES FOR THE SWIFT-HOHENBERG EQUATION
Pontes, Jose da Rocha Miranda
Christov, C. I.

TRB1527 165
GEO-LOCATION OF TRANSMITTERS USING IDEAL AND NON-IDEAL DATA, DOPPLER SHIFTS AND LEAST SQUARES
Sousa, Cristina Tobler
Kuga, Hélio Koiti
Setzer, Alberto Waingort

TRB1878 174
THE EARTH'S SHADOW AND THE SPACECRAFT'S ATTITUDE PROPAGATION
Cabette, Regina Elaine Santos
Zanardi, Maria Cecília
Moraes, Rodolpho Vilhena de

TRB1972 183
ON THE DYNAMICS OF AN ARMATURE CONTROLLED SERIES MOTOR MOUNTED ON AN ELASTICALLY SUPPORTED FOUNDATION
Mattos, Marcio Coelho de
Rocha, Roger

TRB2064 191
AN ANALYTICAL APPROACH FOR EVALUATION OF ORBITAL PERTURBATIONS DUE TO ATMOSPHERIC DRAG
de Moraes, Rodolpho Vilhena
Nicoletti, Gustavo Willrich

TRB2131 198
SEMI-ANALYTICAL DETERMINATION OF HALO ORBITS VIA THE LINDSTEDT-POINCARÉ METHOD
Prado, Antonio Fernando Bertachini De Almeida
Beauge, Cristian
Felipe, Gislaine

TRB2157 208
ANALYTICAL SOLUTION FOR SOME PROBLEMS WITH CHAOTIC RESPONSE
Rosales, Marta Beatriz
Filipich, Carlos Pedro

TRB2159 216
STRESS CONCENTRATION EVOLUTION IN SHOULDERED PLATES FOR NON-LINEAR ELASTIC MATERIALS
Gomes, Carlos J.
Troyani, Nando

TRB2340 222
FINITE VECTORIAL ROTATIONS
Bergweiler, Marcelo González
Pimenta, Paulo de Mattos

TRB2466 233
NONLINEAR DYNAMIC MODELING AND CONTROL OF AN UNDER-ACTUATED FLEXIBLE STRUCTURE MOUNTED
Góes, Luiz Carlos Sandoval
Costa, Alexsander
Soares, Álvaro M.S.
Grandinetti, Francisco José

TRB2484 243
TIKHONOV REGULARIZATION METHOD AND TRACKING
barbanti, Luciano
Balthazar, José M.
Dantas, M. J. H.

XVI CONGRESSO BRASILEIRO DE ENGENHARIA MECÂNICA
16th BRAZILIAN CONGRESS OF MECHANICAL ENGINEERING

On Local Stability Analysis of a non-Ideal Problem

Márcio José Horta Dantas
Faculdade de Matemática, UFU, 38408-100, Uberlândia, MG, Brazil
E-mail: marcio@ufu.br

José Manoel Balthazar
Instituto de Geociências e Ciências Exatas, CP 178, UNESP, 13500-230 Rio Claro, SP, Brazil
E-mail: **jmbaltha@rc.unesp.br**

Abstract: *In real engineering problems, it is of major importance, to consider non-ideal energy sources. They act on a vibrating system and at the same time they experience a reciprocal action from it. In this paper, we study a type of non-ideal problem where the interaction source energy -motion is accomplished through a special kind of friction (self-excited oscillations). We obtain results concerning both the stability and instability of the equilibrium points of this system. We also determine its bifurcation curves and found, in the set of parameters of the vibrating system, a Hopf Bifurcation.*

Keywords*: Non-Ideal Problems, Stability, Bifurcation, Central Manifolds*

1. Introduction

Recently, significant contributions have been made to the theory of vibrations of non-linear systems. Nevertheless, special kinds of vibrations, arising from the interaction of the system with the energy source, can not yet be completely explained by means of current theory. Usually the energy source is a DC motor, rotating with an angular velocity $\dot{\varphi}$. It is imperious to introduce two new concepts: an ideal energy source and a non-ideal energy source, as follows.

An ideal energy source is one that acts on the vibrating system, but doe not experience any reciprocal influence from the system.

A non-ideal source is one that acts on a vibrating system and at the same time experiences a reciprocal action from the system. Thus non-ideal systems have one more degree of freedom than their ideal counterparts.

We remark that changes in the parameters of the non-ideal vibrating system may be accomplished by changing the working conditions of the energy source.

The non-ideal problems have been investigated, in current literature, by a numbers of researchers. Prof. Kononenko (Kononenko, 1969) devoted an entire text to this subject and (Nayfeh and Mook, 1979) presents a complete review of this kind of problems up to 1974, completed by (Balthazar et al., 1999) and (Balthazar et al., 2001) from 1997 to 2001.

Particular cases of non-ideal problems with dry friction are discussed in (Alifov and Frolov, 1977), (Ibrahim, 1994) and recently by (Pontes et al., 2000).

The purpose of this paper is to deal with this kind of problem and to report some results of local analysis of a non-ideal vibrations, defined earlier in Kononenko´s book p.159 (Kononenko, 1969), concerning to self-excited systems.

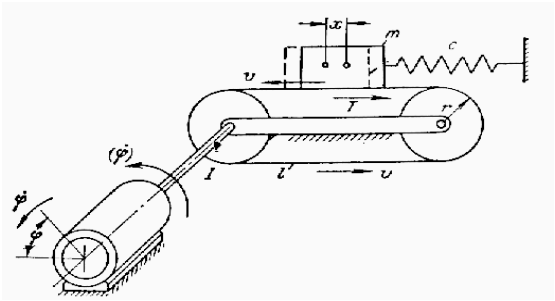


Figure 1- Self-excited oscillatory system model (Kononenko, 1969)

We remark that Prof. Kononenko doesn't study it, before, using the approach, which we will present in this paper.

2. The mathematical model

In order to investigate a kind of self-excited vibrating system with non-ideal energy source, we consider a mathematical model defined in Fig. (1), where we consider an inextensible belt l and stiffly jointed to a DC motor. The position of the mass is denoted by x, chosen so that, for x=0, the spring is unstressed.

In this mathematical model, it is assumed that it receives its energy from a DC motor with a specified characteristic $L(\dot{\varphi})$.

The endless belt moves with velocity $v = r\dot{\varphi}$, where φ is the angle turned through by the DC motor (rotation).

There are two forces acting on the mass $\mathcal{M}$:

- a linear spring whose stiffness constant is c and
- a kind of friction force T at the point of contact of the mass with the belt.

The force T is a function of the relative sliding velocity, between the belt and the block, denoted as, $v_{rel} = v - \dot{x}$. Then, the Lagrangian equations of motion for this mechanical model , is easily written in following form:

$$\begin{cases} m\ddot{x}+cx &= -H_2(\dot{x})+T(r\dot{\varphi}-\dot{x}) \\ I\ddot{\varphi} &= L(\dot{\varphi})-H_1(\dot{\varphi})-rT(r\dot{\varphi}-\dot{x}). \end{cases} \tag{1}$$

where

- I is the sum of the moments of inertia of the rotating parts,,
- $H_1(\dot{\varphi})$ the moment of the force resisting the rotation and
- $H_2(\dot{x})$ the force resisting the oscillatory motion.

Note that in the second equation the right side expresses the algebraic sum of the driving torque $L(\dot{\varphi})$ and the resisting torque.

We also assume that $L(\dot{\varphi})$ and $H_1(\dot{\varphi})$ are continuous and vary slowly and we notice that r =1, without loss of generality.

Note also that the first equation (1a) describes the motion of the oscillatory model system and the second equation the behavior of the energy source.

In this paper, we are interested in deal with the assumption when

$$H_2(\dot{x}) = H\dot{x}, \quad H = const. \tag{2}$$

$\dot{\varphi}(0) > 0$ (since the motor always rotates in an only direction)

We remark that if we assumed that the oscillations were approximately harmonic and the algebraic sum of the driving and resisting torque was small, then, consequently the acceleration $\ddot{\varphi}$ was small (it was, that the period of oscillation altered little in the period).

Then, using above property, rewritten the Lagrangian Eq. (1) in the suitable form

$$\begin{cases} \ddot{x}+\omega^2 x &= \dfrac{\varepsilon}{I}[-H_2(\dot{x})+T(r\dot{\varphi}-\dot{x})] \\ \ddot{\varphi} &= \dfrac{\varepsilon}{m}[L(\dot{\varphi})-H_1(\dot{\varphi})-rT(r\dot{\varphi}-\dot{x})]. \end{cases} \tag{1a}$$

Where

- $\omega^2 = \dfrac{c}{m}$ is the natural frequency of the system and
- ε i s the small parameter of the problem.

Using this above hypothesis, Prof. Kononenko (Kononenko, 1969) analyzed the stability of the problem, using a kind of perturbation theory. He also imposed that the function T was an algebraic polynomial of third degree, positive defined. He found that the stability conditions depended on the property of the energy source, that is, the gradient of the energy source characteristics, had a considerable influence on the stability of the problem.

Here, we will analyze the local stability this kind of problem, using different mathematical tools of those used by Prof. Kononenko in his book (Kononenko, 1969)(The physical model proposed by Prof. Kononenko was the motivation for this paper); they seem to be more adequate for this kind of problem. New hypothesis need to be introduced in this problem, as we will show latter.

2. Local Stability Analysis

Taking $y = \dot{x}$, $z = \dot{\varphi}$ and $M(\dot{\varphi}) = L(\dot{\varphi}) - H_1(\dot{\varphi})$ in Eq. (1) and by use of Eq. (2), we get the following equations

$$\begin{cases} \dot{x} &= y, \\ \dot{y} &= -\dfrac{c}{m}x - \dfrac{H}{m}y + \dfrac{T(z-y)}{m}, \\ \dot{z} &= \dfrac{M(z)}{I} - \dfrac{T(z-y)}{I}. \end{cases} \tag{3}$$

Now, let (x_0, y_0, z_0) be an equilibrium point of Eq. (3), then

$$\begin{cases} x_0 = \dfrac{T(z_0)}{c}, \\ y_0 = 0, \\ M(z_0) = T(z_0). \end{cases}$$

We may assume, without loss of generality that,

$$\begin{cases} z_0 = 0, \\ M(0) = T(0) = 0. \end{cases} \tag{4}$$

Then follow from Eq. (4) that *(0, 0, 0)* is an equilibrium point of Eq. (3).
Next, let us introduce the following notation:

$$\begin{cases} a = T'(0)/m, \\ k_1 = H/m, \\ k_2 = m/I \\ b = (T'(0) - M'(0))/I, \\ \omega = \sqrt{c/m}, \\ f(w) = T(w) - T'(0)w, \\ g(w) = M(w) - M'(0)w. \end{cases}$$

Clearly, k_1, k_2 and ω are positive constants and from Eq. (4) we deduce that $f(0) = g(0) = f'(0) = g'(0) = 0.$
We also have that Eq. (3) can be written as

$$\begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ -\omega^2 & -(a+k_1) & a \\ 0 & k_2 a & -b \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} 0 \\ f(z-y)/I \\ (g(z) - f(z-y))/I \end{pmatrix}. \tag{5}$$

Of course *a* and *b* are parameters of Eq. (5). Note that they will take in account, until the first-order approximation, the influence of the friction and of the characteristics of the motor of the vibrating system.

The characteristic polynomial of the linear part of Eq. (5) is given by

$$q(\lambda) = \lambda^3 + (a + b + k_1)\lambda^2 + ((a + k_1)b - k_2 a^2 + \omega^2)\lambda + \omega^2 b. \tag{6}$$

It is an immediate consequence of the classical Hurwitz Criterion that a polynomial $r(t) = t^3 + At^2 + Bt + C$, with real coefficients, has all roots with negative real part or zero real part if, and only if, $A \geq 0, B \geq 0, C \geq 0, AB \geq C.$

From this remark jointly with Eq. (6) and the Classical Center Manifold Theorem for Flows (see for example Guckenheimer et all (1983)), it is not difficult to prove following result:

Theorem 1:: *a) If* $a+b+k_1 < 0$ *or* $(a+k_1)b - k_2 a^2 + \omega^2 < 0$ *or* $b < 0$ *or* $(a+b+k_1)((a+k_1)b - k_2 a^2 + \omega^2) < \omega^2 b$ *Then (0, 0, 0) is an unstable equilibrium point of Eq. (5).*
b) If

$$\begin{cases} a+b+k_1 > 0, (\text{or } (a+k_1)b - k_2 a^2 + \omega^2 > 0), \\ b > 0, \\ (a+b+k_1)((a+k_1)b - k_2 a^2 + \omega^2) > \omega^2 b, \end{cases}$$

Then (0, 0, 0) is an asymptotically stable equilibrium point of Eq. (5).

3. Two-parameter Bifurcation

In this section, we will give some important results on the bifurcation's of Eq. (5) which depend on the parameters *a* and *b*.The proofs of the following results depend on a careful analysis of the behavior of the real part of the solutions of Eq. (6) and Stable Manifold Theorem (see Guckenheimer et all (1983)).

Theorem 2: *If (a, b) satisfies*

$$\begin{cases} a+b+k_1>0, (or \quad (a+k_1)b-k_2a^2+\omega^2>0\ ), \\ b>0, \\ (a+b+k_1)(\ (a+k_1)b-k_2a^2+\omega^2\ )-\omega^2 b=0, \end{cases}$$

or

$$\begin{cases} a+b+k_1<0, (or \quad (a+k_1)b-k_2a^2+\omega^2<0\ ), \\ b<0, \\ (a+b+k_1)(\ (a+k_1)b-k_2a^2+\omega^2\ )-\omega^2 b=0, \end{cases}$$

Then (0, 0, 0, (a, b)) is a point of bifurcation of Eq. (5).

Theorem 3: *For all* $a\in\mathbb{R}$, *(0, 0, 0, (a, 0)) is a bifurcation point of Eq. (5).*

4. One-parameter Bifurcation

In the Theorem 3, discussed before, the bifurcation points depend on two parameters. An interesting question is to know what happens when there is an only parameter.

We are going to study the existence of bifurcation points of the type *(0, 0, 0, a)* in the following system:

$$\begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ -\omega^2 & -(a+k_1) & a \\ 0 & k_2 a & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} 0 \\ f(z-y)/m \\ (g(z)-f(z-y))/I \end{pmatrix}. \tag{7}$$

The characteristic polynomial of the linear part of Eq. (7) is given by $q(\lambda)=\lambda^3+(a+k_1)\lambda^2+(-k_2a^2+\omega^2)\lambda$ and its roots are the following ones:

$$0, \frac{-(a+k_1)\pm\sqrt{(a+k_1)^2-4(\omega^2-k_2a^2)}}{2}. \tag{8}$$

In this section we are only interested in the case in which the roots, given by Eq. (8), have real part smaller or equal to zero. Therefore, by use of Theorem 1, it follows that $a+k_1\geq 0$ and $\omega^2-k_2a^2\geq 0$. Then we have four possibilities

$$\begin{cases} a+k_1>0 \quad and \quad \omega^2-k_2a^2>0, \\ a+k_1>0 \quad and \quad \omega^2-k_2a^2=0, \\ a+k_1=0 \quad and \quad \omega^2-k_2a^2>0, \\ a+k_1=0 \quad and \quad \omega^2-k_2a^2=0. \end{cases} \tag{9}$$

In this paper, we will obtain results for the case Eq. $(9)_1$. We remark that the cases Eq. $(9)_{2,3,4}$ can be studied using, for instance the Center Manifold Theorem and Normal Forms.

Theorem 4 *By setting* $a+k_1>0$ *and* $\omega^2-k_2a^2>0$ *, it follows that*

a) If $g''(0)<f''(0)$ *then (0, 0, 0) is an asymptotically stable equilibrium point of Eq. (7).*

b) If $f''(0)<g''(0)$ *then (0, 0, 0) is an unstable equilibrium point of Eq. (7).*

Sketch of the proof: From Eq. (8), we have three possibilities

$$\begin{cases}(a+k_1)^2 &= 4(\omega^2-k_2 a^2)^2,\\ (a+k_1)^2 &> 4(\omega^2-k_2 a^2)^2,\\ (a+k_1)^2 &< 4(\omega^2-k_2 a^2)^2.\end{cases}$$

By using some change of variables and the Central Manifold Theorem, see Guckenheimer et all (1983), we obtain, for all *three cases*, that Eq. (7) reduces to

$$\dot{u}=\frac{\omega^2(g''(0)-f''(0))}{2(\omega^2-k_2 a^2)I}u^2+O(u^3).$$

From this we get the desired result.
Next we will discuss the existence of Hopf bifurcations.

5. Hopf Bifurcation

Consider *(a, b)* such that

$$\begin{cases} b &> 0,\\ a+b+k_1 &> 0,\\ (a+b+k_1)((a+k_1)b-k_2 a^2+\omega^2) &= \omega^2 b.\end{cases} \tag{10}$$

Therefore, the roots of Eq. (6) are given by

$$\pm i\,\omega\sqrt{\frac{b}{a+b+k_1}} \text{ and } -(a+b+k_1).$$

Let a_0, b_0 be a solution of Eq. (10). Let us consider the following system

$$\begin{pmatrix}\dot{x}\\ \dot{y}\\ \dot{z}\end{pmatrix}=\begin{pmatrix}0 & 1 & 0\\ -\omega^2 & -(a_0+k_1) & a_0\\ 0 & k_2 a_0 & -b\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix}+\begin{pmatrix}0\\ f(z-y)/I\\ (g(z)-f(z-y))/I\end{pmatrix}. \tag{11}$$

Clearly, *b* is the parameter of Eq. (11).

Theorem 5: *In Eq. (11), assumes that $f''(0)=g''(0)=0$. Suppose that one of the following conditions holds:*

a) $g^{(3)}(0)\neq 0.$

b) $g^{(3)}(0)=0,\ f^{(3)}(0)\neq 0$ *and* $\omega^2\neq k_2 k_1^2\dfrac{(I-m)}{I}.$

c) $g^{(3)}(0)=0,\ f^{(3)}(0)\neq 0,\ \omega^2=k_2 k_1^2\dfrac{(I-m)}{I}$ *and*

$$-\left(\sqrt{I}\right)^7 k_1^2 k_2^2\sqrt{(I-m)}+3\left(\sqrt{I}\right)^5 m k_1^2 k_2^2\sqrt{(I-m)}-3\left(\sqrt{I}\right)^3 m^2 k_1^2 k_2^2\sqrt{(I-m)}+$$
$$\sqrt{I}\,m^3 k_1^2 k_2^2\sqrt{(I-m)}-I^3 m k_1 k_2+I^2 m^2 k_1 k_2-2I^4+I^4 m\neq 0.$$

Then there is finite subset S of $\mathbb{R}^2$ *satisfying Eq. (10), such that if* $(a_0,b_0)\in \mathbb{R}^2$*-S then* $(0,0,0,b_0)$ *is a Hopf bifurcation point.*

Sketch of the proof: By using Bézout Theorem, see Walker (1950), Theorem 5.4, pg.111, the Hopf Bifurcation Theorem (see Guckenheimer et all (1983)), the formulae (2.4), (2.5), pg.138 of Starzhinskii (1980) and a lot of algebraic computation, we got the desired result.

5. Summary of the results

In the Fig. 2, we summarized the results of the theorems 1, 2, 3 and 5.

We are representing in the plan, the parameters *a* and *b* that, until the first-order approximation, determine the qualitative behavior of the solutions of Eq. (5) in the neighborhood of the equilibrium point *(0,0,0).* We denote by *s* and *u* the dimension of stable and unstable manifold respectively.

For the parameters that are located in the limited areas by curves of the illustration, the dimension of the stable and unstable manifold is that indicated in the illustration. Besides, for the located parameters on the curves, bifurcation's can exist, as Hopf. Bifurcation.

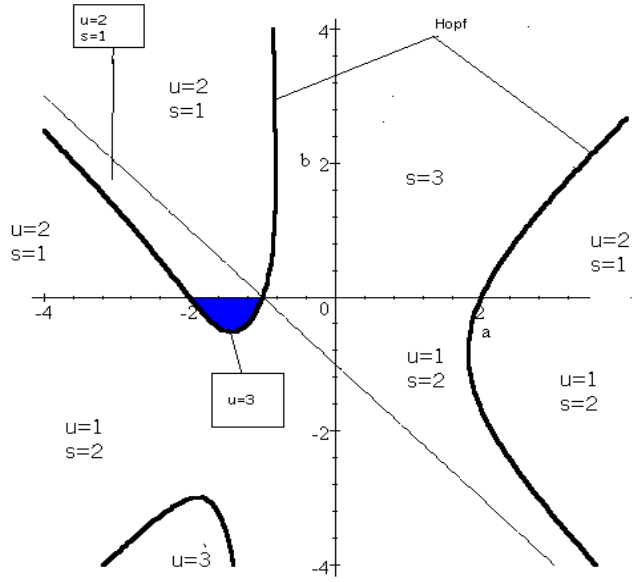


Figure 2- Sketch of the Solutions

6. Conclusions

We studied a type of non-ideal problem where the interaction source energy -motion is accomplished through a special kind of friction (self-excited oscillations).

We obtain results concerning both the stability and instability of the equilibrium points of this system and we determine its bifurcation curves (see Fig (2)). We also found, a Hopf Bifurcation, in the set of parameters of the vibrating system.

7. Acknowledgements

The authors acknowledge support by FAPESP, Fundação de Apoio à Pesquisa do Estado de São Paulo. The second author also thanks CNPq, Conselho Nacional de Pesquisas. Both are Brazilian Research Funding Agencies.

8. References

Alifov, A. A. and Frolov, K.V., 1997, " Investigation of Self-Excited Oscillations with Friction, Under Conditions of Parametric Excitation", Mekanika Tverdogo Tela, 12(4), 57-65.

Balthazar, J.M, Mook, D.T., Weber, H.I., A. Fenili, Belato, D., De Mattos, M.C. and Wieczorek, S., 1999," On vibrating Systems With A Limited Power Supply and Their applications to Engineering Sciences", In: Seminário Brasileiro de Análise, UNICAMP, Campinas, 2000, 137-227

Balthazar, J.M, Brasil, R.M.L.R.F., Weber, H.I., A. Fenili, Belato, D., Felix, J.L.P. Felix and Garzeri F.J., 2001," Review of New Vibration Issues Due to Non-Ideal Sources ", Dynamics and Their Controls", Edited by Udwadia, F. E. and Weber, H .I., Gordon and Breach Science Publishers,2001, submitted

Carr, J. and Naji Al-Amood, 1980, "Rate of Decay in Critical Cases I: Finite Dimensional Problems" Journal *of Mathematical Analysis and Applications* 75, 242-250.

Guckenheimer, J. and Holmes P., 1983, "Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields". New York Berlin Heidelberg Tokyo, Springer-Verlag, 473p.

Ibrahim, R. A., 1994," Friction-Induced Vibration, Chatter, Squeal, and Chaos Part II: Dynamics and Modeling", Applied Mechanics Reviews, 47(7), 227-253, 2

Kononenko, V.O, 1969, "Vibrating Systems with a Limited Power Supply" London, Illife Books LTD 236p.

Nayfeh,A.H. and Mook, D.T. ,1979, "Nonlinear Oscillations" New York: Wiley.

Pontes, B.R., Oliveira, V.A and Balthazar, J.M." On friction-Driven Vibrations in A Mass Block-Belt-Motor System with A Limited Power Supply", Journal of Sound and Vibration, 234(4), 713-723, 2000.

Starzhinskii V. M., 1980, “Applied Methods in the Theory of Nonlinear Oscillations”, Moscow, Mir Publishers,
Walker R. J., “Algebraic Curves”, Princeton, Princeton University Press, (1950).

XVI CONGRESSO BRASILEIRO DE ENGENHARIA MECÂNICA
16th BRAZILIAN CONGRESS OF MECHANICAL ENGINEERING

IDENTIFYING PERIODIC AND CHAOTIC EXPERIMENTAL TIME SERIES

Luiz Fernando P. Franca
Marcelo A. Savi
Instituto Militar de Engenharia
Departamento de Engenharia Mecânica e de Materiais
22.290.270 - Rio de Janeiro - RJ
E-Mail: savi@epq.ime.eb.br

Abstract. *The experimental analysis of nonlinear dynamical systems furnishes a scalar sequence of measurements, which may be analyzed using state space reconstruction and other techniques related to nonlinear analysis. The noise contamination is unavoidable in cases of data acquisition and, therefore, it is important to recognize techniques that can be employed for a correct identification of chaos. The present contribution discusses the experimental analysis of a nonlinear pendulum, considering state space reconstruction and the determination Lyapunov exponents. A procedure to construct Poincaré map of the signal is presented. The analyses of periodic and chaotic motions are carried out in order to establish a difference between them. Results show that it is possible to distinguish periodic and chaotic time series obtained from an experimental set up employing proper procedures even though noise suppression is not contemplated.*

Key-words: *Chaos, Time Series, Nonlinear Pendulum, Experimental Analysis.*

1. Introduction

The experimental analysis of nonlinear dynamical systems furnishes a scalar sequence of measurements. Therefore, a time series associated with system dynamics is available, being interesting to analyze it using state space reconstruction and other techniques related to nonlinear analysis. The noise contamination is unavoidable in cases of data acquisition and noise suppression is essential in signal processing, especially in chaos analysis. Many studies are devoted to evaluate noise suppression and its effects in the analysis of chaos. However, there are a small number of reports devoted to the effects of the system noise on chaos (Ogata *et al.*, 1997).

The analysis of nonlinear dynamical systems from time series involves state space reconstruction. Basically, there are two different methods for this aim: derivative coordinates and delay coordinates (Packard *et al.*, 1980; Takens, 1981; Broomhead & King, 1986). The method of delay coordinates has proven to be a powerful tool to analyze chaotic behavior of dynamical system. Ruelle (1979), Packard *et al.* (1980) and Takens (1981) have introduced the basic idea of this method and the main problem arising is the determination of the embedding parameters.

Nonlinear analysis also involves the determination of quantities, known as dynamical invariants, which are important to identify chaotic behavior. Lyapunov exponents are an example that evaluates sensitive dependence to initial conditions estimating the exponential divergence of nearby orbits. These exponents have been used as the most useful dynamical diagnostic tool for chaotic system analysis. The signs of the Lyapunov exponents provide a qualitative picture of the system's dynamics and any system containing at least one positive exponent presents chaotic behavior. The determination of Lyapunov exponents of dynamical system with an explicitly mathematical model, which can be linearized, is well established from the algorithm proposed by Wolf *et al.* (1985). On the other hand, the determination of these exponents from time series is quite more complex. Basically, there are two different classes of algorithms: trajectories, real space or direct method (Wolf *el al.*, 1985; Kantz, 1994; Rosenstein *et al.*, 1993); and perturbation, tangent space or Jacobian matrix method (Sano & Sawada, 1985; Eckmann *et al.*, 1986; Brown *et al.*, 1991; Briggs, 1990; Kruel *et al.*, 1993; Zeng *et al.*, 1991; Brown, 1993; Ellner *et al.*, 1992).

The main purpose of this contribution is to distinguish periodic and chaotic signals obtained from an experimental nonlinear pendulum. Noise suppression is not contemplated and the signal is analyzed without filtering. State space reconstruction and the determination of Lyapunov exponents are considered. Furthermore, it is necessary to present a procedure to construct a Poincaré map of the signal, which is also discussed. The choice of algorithms to be employed is based in the analysis of noise sensibility developed in Franca (2000). Algorithms proposed by Kantz (1994) and by Rosenstein *et al.* (1993) are conceived to estimate Lyapunov exponents. Results show that it is possible to distinguish periodic and chaotic signals employing proper procedures even though noise suppression is not contemplated.

2 - Experimental Apparatus

The experimental data related to the nonlinear pendulum response is obtained from the apparatus depicted in Figure 1. The pendulum is constructed by a disc with a lumped mass (1) and is connected to a rotary motion sensor (3). The dissipation is provided by a magnetic device (2), which is adjustable. A motor-string-spring device (4-5) provides the excitation for the pendulum. The motor (5), *PASCO ME-8750*, has the following characteristics: 12V DC, 0.3-3Hz and 0-0.3A. The signal measurement is done with the aid of two transducers. The rotary motion sensor (3), *PASCO encoder CI-6538*, has 1440 orifices and a precision of 0.25^0. The magnetic transducer (6) is employed in order to generate a frequency signal associated with the forcing frequency of the motor, which is used to construct the Poincaré map of the signal. The apparatus is connected with an A/D interface, *Science Workshop Interface 500* (*CI-6760*) where the

sampling frequency varies from 2Hz to 20kHz. The interface oversamples the signal 8 times for frequencies below 100Hz and a single time for higher sampling rates. Furthermore, this interface does not have any anti-aliasing filters and a 9V AC-DC adapter provides power supply.

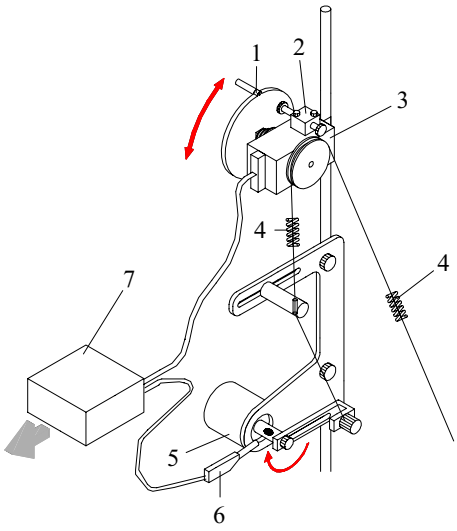


Figure 1. Experimental apparatus of the nonlinear pendulum: (1) Disc with concentrated mass; (2) Magnetic damping device; (3) Rotary motion sensor: PASCO CI-6538; (4) Spring; (5) DC Motor: PASCO ME-8750; (6) Magnetic transducer: TEKTRONIX; (7) Science workshop interface: PASCO CI-6760.

All signals are analyzed with the aid of the *Science Workshop Data Acquisition*, which allows one to evaluate angular velocity ($y = \dot{\theta}$) and angular position ($x = \theta$). Noise suppression is not contemplated and all signals are stored without filtering.

In order to perform the analysis of the nonlinear pendulum, one conceives that the time series is a sequence of angular position measured from the experiment, $s = x = \theta$. The apparatus also permits to measure the angular velocity $y = \dot{\theta}$, which is used to construct the real phase space (x versus y), employed to perform a visual validation of the reconstructed phase space.

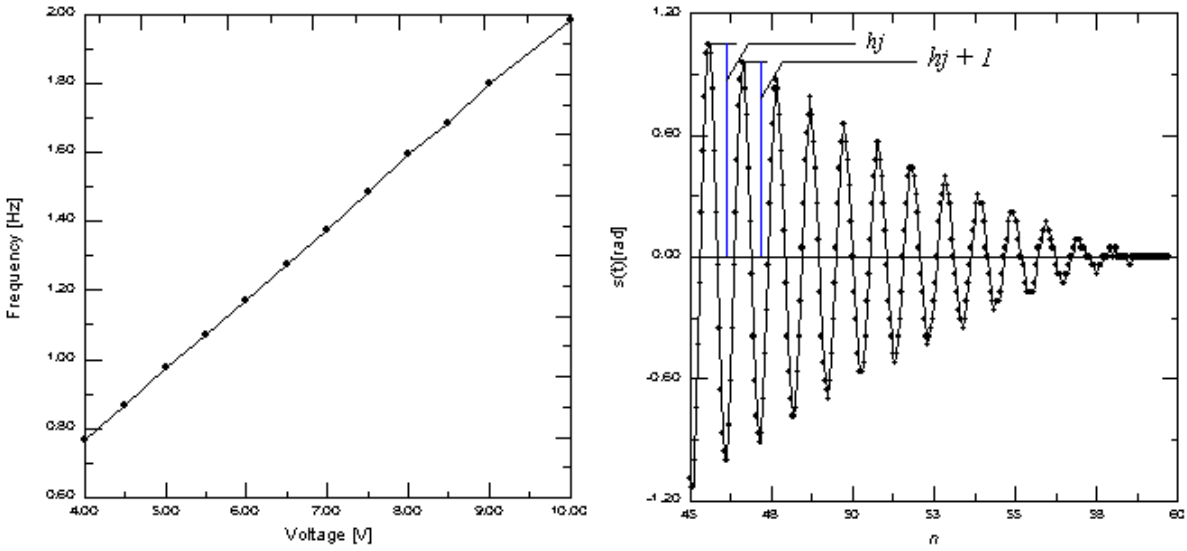
2.1 - Identification of System Parameters

Some characteristics of the apparatus are now discussed in order to identify some parameters of the pendulum. At first, the 12V DC motor is considered. Figure 2a shows the frequency (Ω_e) versus voltage curve which is important to identify the forcing frequency of the system. Other important parameter that needs to be quantified is the dissipation, related to a magnetic damping device. A convenient procedure is the logarithmic decrement (Meirovitch, 1986), which is defined verifying the ratio between any two consecutive displacement amplitudes. The analysis of this definition yields

$$\gamma = \frac{1}{j} \ln \frac{h_1}{h_{j+1}} \tag{1}$$

where h_1 and h_{j+1} are the amplitude of the displacements at time instants t_1 and $t_{j+1} = t_1 + jT_d$, respectively, with j being a integer number and $T_d = 2\pi/\omega_d$; ω_d is the free damping frequency. The time history of part of a free response signal is presented in Figure 2b, pointing the cited definitions. Under these assumptions, the non-dimensional viscous damping parameter, ς, is defined as follows

$$\varsigma = \frac{\gamma}{\sqrt{(2\pi) + \gamma^2}} \tag{2}$$



(a) (b)

Figure 2. Identification of system parameters: (a) Forcing frequency versus motor voltage; (b) Free response signal.

2.2 - Poincaré Maps

Poincaré map is an important tool to observe the response of a nonlinear system. Experimentally, this can be done in several ways. Moon (1992) presents a procedure employing a signal converter that stores the sampled data in a computer for display at a later time. Here, similar procedure is conceived in order to generate two signals: one associated with the motion and the other associated with the forcing frequency. The forcing frequency signal is generated with the aid of a magnetic transducer, which induces electric pulses when a reference bolt, connected to the motor, passes near it. These pulses are compared with the motion signal in order to generate a third signal representing the Poincaré Map where only measures in these time instants are contemplated.

3 - Periodic Signal

In order to analyze a periodic response of the experimental nonlinear pendulum, a period-2 signal is considered with N = 38,090 points, generated with a motor voltage V = 4.2V (Ω_e = 0.82Hz), a sample frequency Ω_s = 20Hz and a damping parameter ς = 0.0065. The time history evolution of part of the signal is presented in Figure 3.

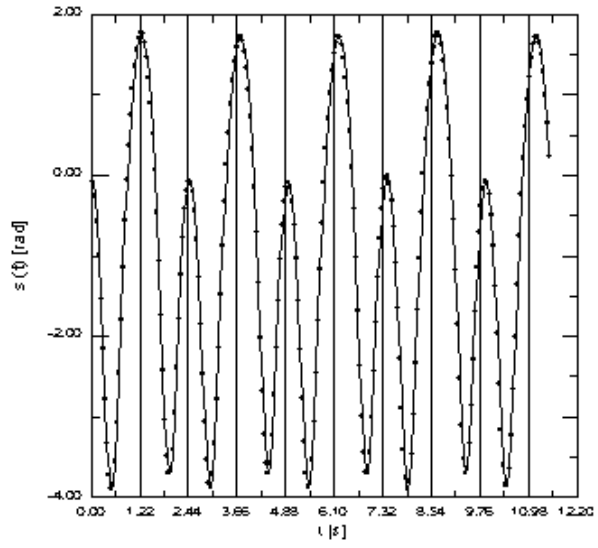
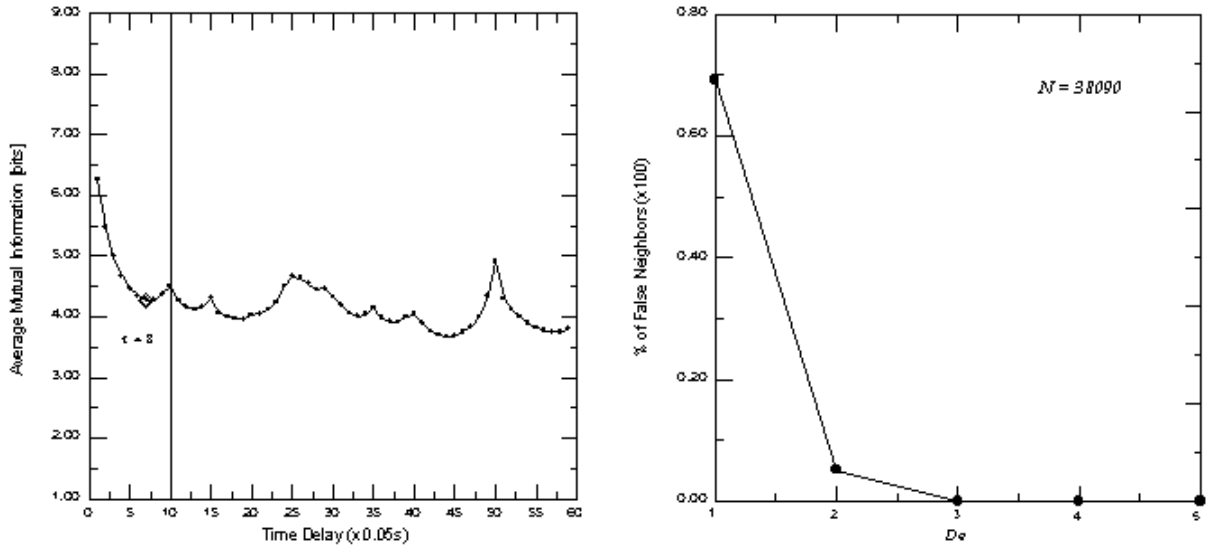


Figure 3. Periodic signal.

The state space reconstruction considers the signal derived from experiment to form a coordinate system that captures the structure of orbits in state space. This contribution employs the method of delay coordinates (Takens, 1981) and, therefore, it is necessary to determine delay parameters, τ and D_e. Here, the average mutual information method is employed to determine time delay (Fraser & Swinney, 1986) and the method of false nearest neighbors to estimate embedding dimension (Rhodes & Morari, 1997). Results of the analysis for the determination of these parameters is presented in Figure 4. Figure 4a shows the analysis of the mutual information versus time delay, and the first minimum

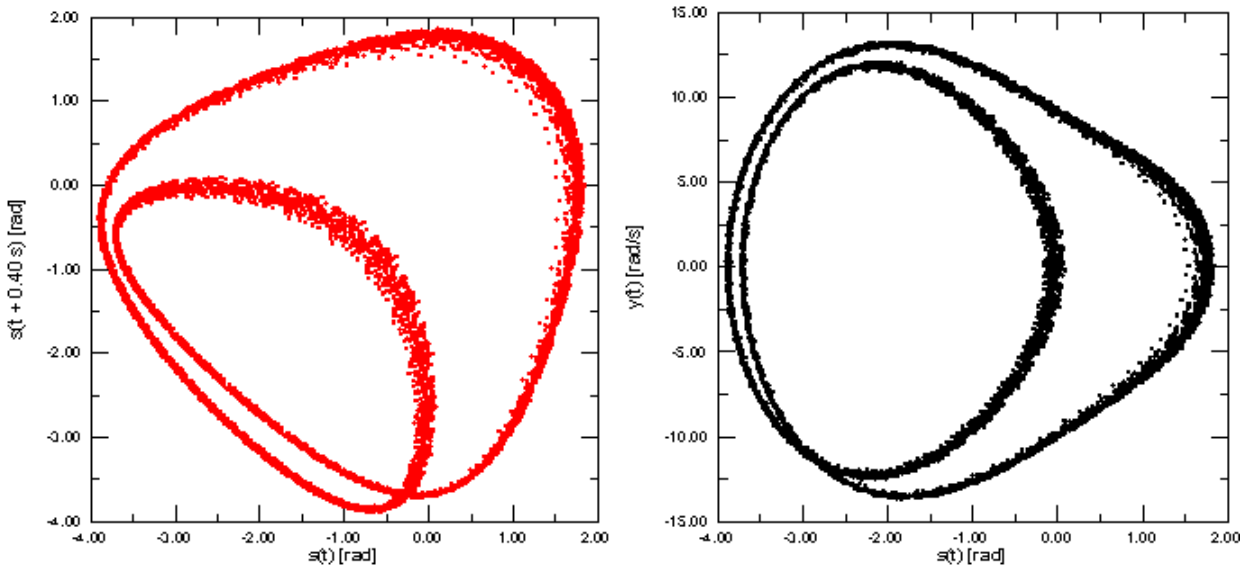
of this curve defines the time delay, τ = 8 x 0.05 = 0.40s. Figure 4b presents the curve of the percentage of false neighbor points versus embedding dimension, indicating that the embedding dimension needs to be between 3 and 4.



(a) (b)

Figure 4. Delay parameters associated with periodic signal. (a) Average mutual information versus τ; (b) Percentage of false neighbors versus D_e.

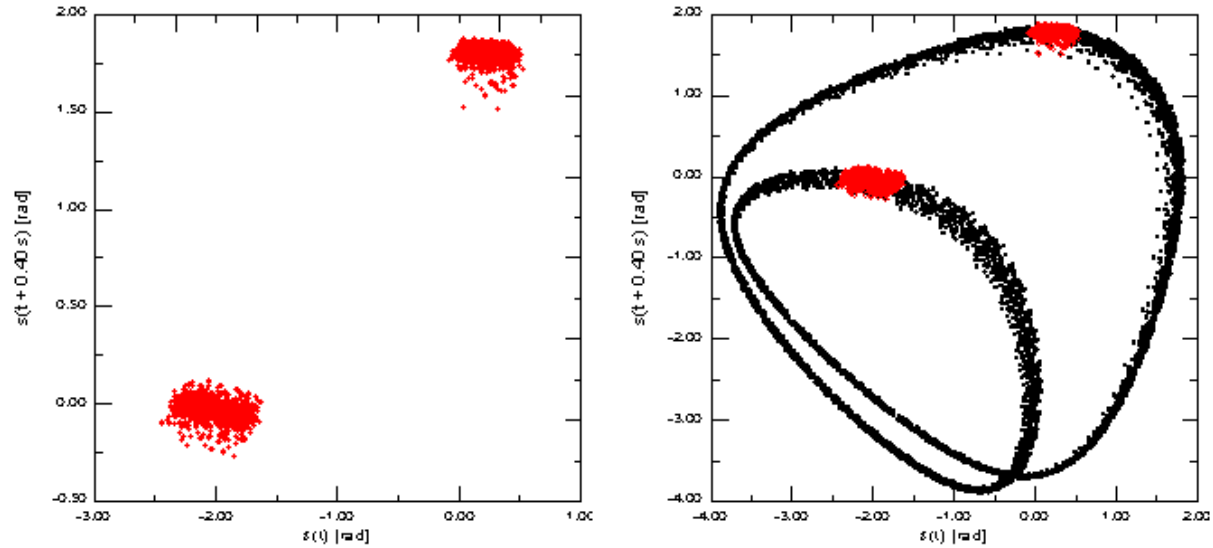
After the determination of delay parameters, state space can be reconstructed. Figure 5a presents the reconstructed phase space projected in 2 and 3 dimension while Figure 5b presents the real phase space measured in the experiment. Both spaces are similar from a topological point of view (Takens, 1981), presenting just a small coordinate change from one to another. Notice that the reconstructed phase space presents a closed curve that is typical of periodic motions.



(a) (b)

Figure 5. State space associated with periodic signal. (a) Reconstructed; (b) Real.

In order to construct a Poincaré map of the motion, two signals are considered: one associated with the motion and the other associated with the forcing frequency. The Poincaré map defined by this procedure is presented in Figure 6. Figure 6a presents the reconstructed Poincaré section while Figure 6b the phase space, also pointing out the Poincaré section. Notice that the Poincaré section shows two clouds of points, representing a period-2 motion. This result can be used to evaluate the noise level on data acquisition since only two points was expected. Therefore, the length of the range around each point is associated with noise.

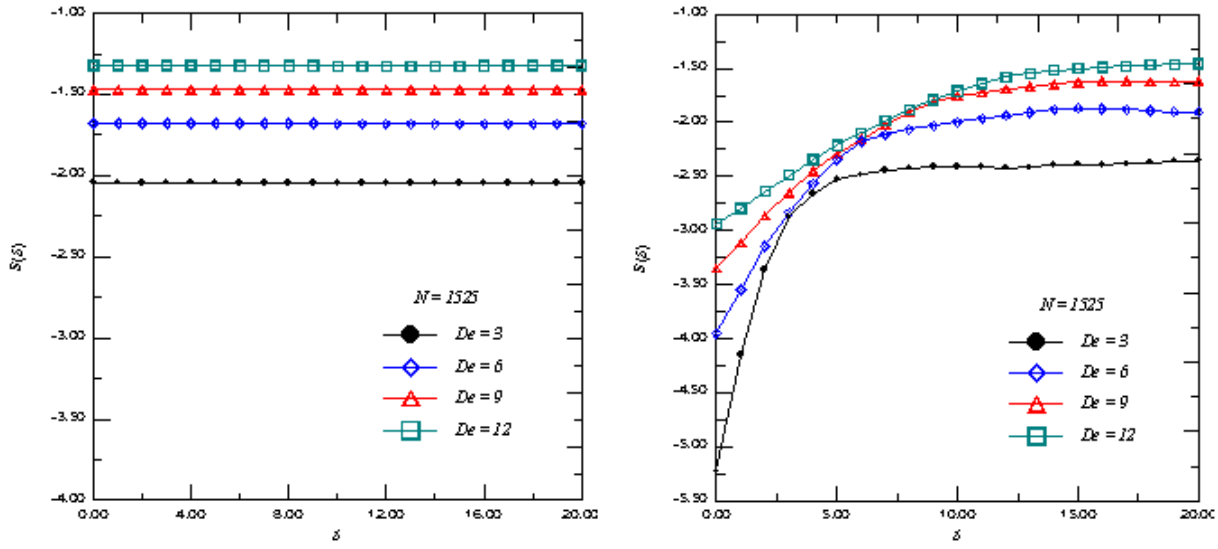


(a) (b)

Figure 6. Poincaré map associated with periodic signal: (a) Reconstructed; (b) Phase space.

Even though the system response presents a periodic-like characteristic, it is important to assure this conclusion with the determination of Lyapunov exponents. Since the greater exponent is the most important to diagnose chaotic motion, and taking into account the conclusions about noise sensibility presented in Franca (2000), algorithms proposed by Kantz (1994) and by Rosenstein *et al.* (1993) are conceived. The analysis is done regarding the Poincaré map of the signal. The use of these algorithms implies the determination of parameter ε before performing signal analysis. In order to "calibrate" the algorithm, a known signal is analyzed, for example a simple periodic motion, defining the correct value of this parameter. After this calibration, the parameter may be employed for all signals, allowing their correct identification.

Figure 7a presents the curve $S(\delta)$ versus δ predicted by the algorithm due to Kantz using $\varepsilon = 1.6$ and $D_e = 3, 6, 9, 12$. This curve has a null slope, meaning that $\lambda_{max} = 0$, and therefore, the signal is related to a periodic motion. On the other hand, Figure 7b presents the curve $S(\delta)$ versus δ predicted by the algorithm due to Rosenstein *et al.* for the same parameters. In this case, the curve presents a non-null slope, which indicates a positive exponent.



(a) (b)

Figure 7. Lyapunov exponents associated with periodic signal. (a) Kantz; (b)Rosenstein *et al.*

The difference between both algorithms is, perhaps, associated with the use of only one neighbor per time on the algorithm due to Rosenstein *et al.*, and not all neighbors within a certain neighborhood, which might induce larger statistical errors, especially in the presence of noise. This problem was pointed by Kantz (1994) and may explain different results obtained by both methods.

4 - Chaotic Signal

Chaos in the experimental nonlinear pendulum is analyzed considering a chaotic signal with N = 30589 points, generated with a motor voltage V = 4.2V (Ω_e = 0.82Hz), a sample frequency Ω_s = 20Hz and a damping parameter ς = 0.0125. The time history evolution of part of the signal is shown in Figure 8.

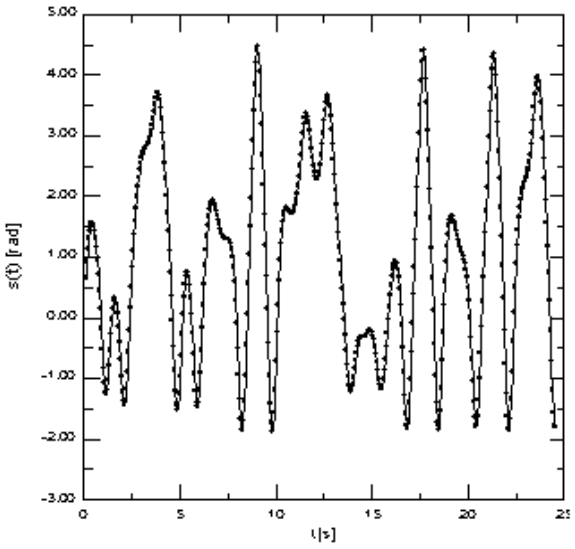
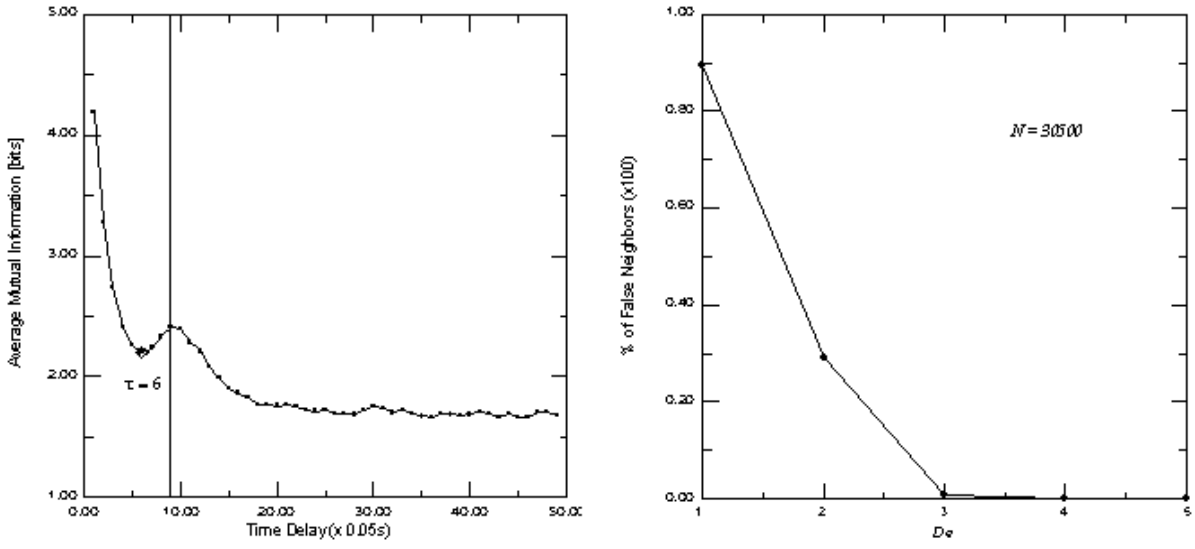


Figure 8. Chaotic signal.

The state space reconstruction from the experimental signal employing the method of delay coordinates is now in focus. Results of the analysis employed to determine the delay parameters are presented in Figure 9. Figure 9a shows the mutual information versus time delay, and the first minimum of the curve must be used as the time delay, furnishing τ = 6 x 0.05 = 0.30s. Figure 9b presents the curve of the percentage of false neighbor points versus embedding dimension, showing that the embedding dimension needs to be between 3 and 4. This result is in agreement with the one obtained for the periodic signal.

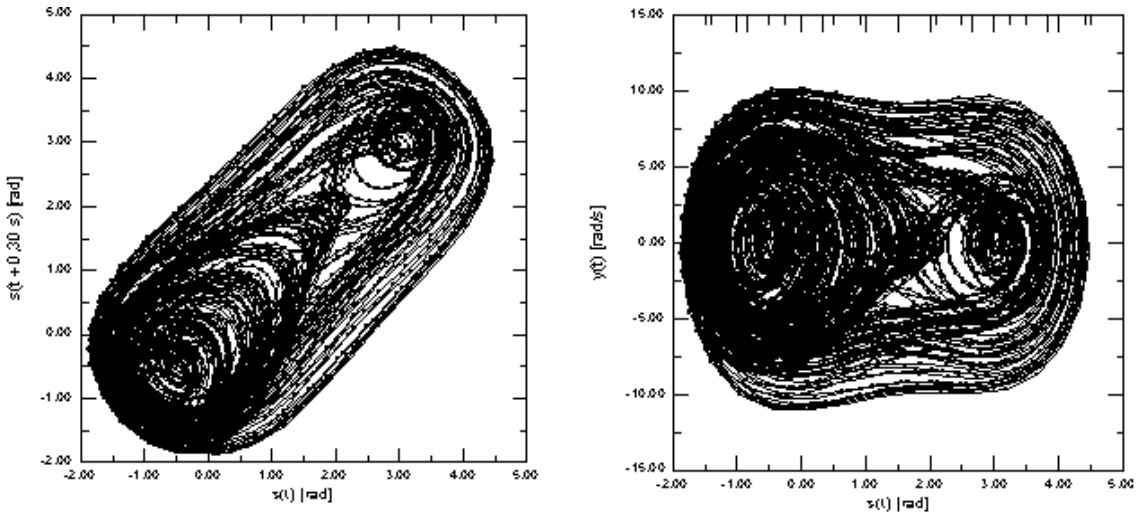
After the determination of delay parameters, it is possible to reconstruct state space. Figure 10a presents the reconstructed phase space while Figure 10b presents the real phase space measured in the experiment. Both spaces are similar from a topological point of view (Takens, 1981), presenting just a small coordinate change from one to another. Here, phase space presents a chaotic-like characteristic because the orbit is not a closed curve.

In order to construct the Poincaré map of the chaotic signal, the same procedure used in the preceding section is employed. The Poincaré map defined by this procedure is presented in Figure 11. Figure 11a shows the reconstructed Poincaré section while Figure 11b shows the real one. A strange attractor is clearly identified presenting a fractal-like structure. Nevertheless, it is useful to confirm this with the calculation of dynamical invariants.



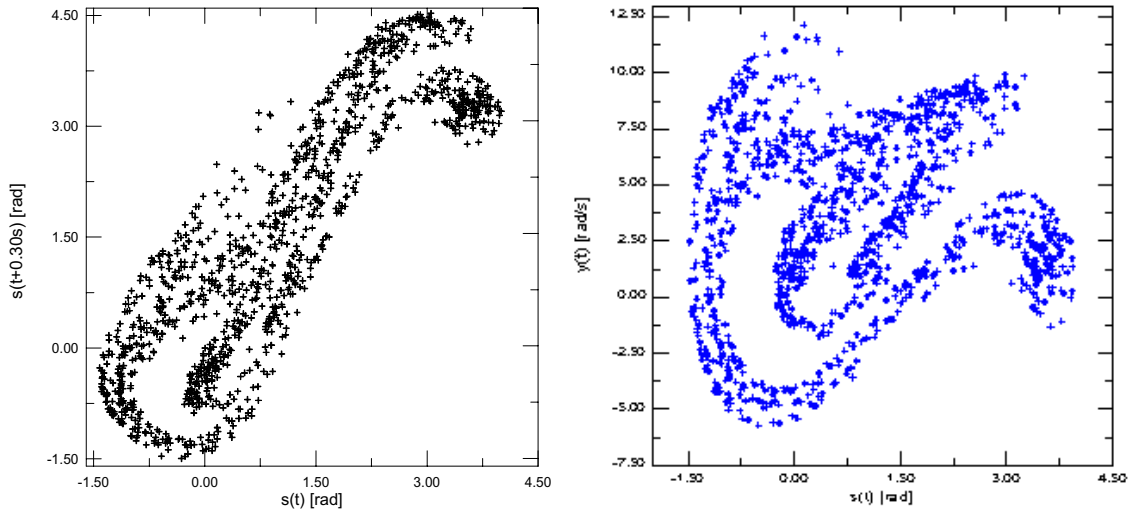
(a) (b)

Figure 9. Delay parameters associated with chaotic signal. (a) Average mutual information versus τ ; (b) Percentage of false neighbors versus D_e.



(a) (b)

Figure 10. State space associated with chaotic signal. (a) Reconstructed; (b) Real.



(a) (b)

Figure 11. Poincaré map of chaotic signal: (a) Reconstructed; (b) Real.

Even though the system response presents a chaotic-like characteristic, it is important to assure this conclusion with the determination of Lyapunov exponents. Employing algorithms due to Kantz (1994) and due to Rosenstein *et al.* (1993) and regarding the Poincaré map signal, it is possible to estimate the maximum Lyapunov exponent. With this aim, the following parameters need to be conceived: $\varepsilon = 1.6$ and $D_e = 6, 9, 12$. It should be emphasized that the value of ε is the same to the one used in the analysis of the periodic signal, discussed in the preceding section. This value is defined from the "calibration" of the algorithm and allows one to analyze different signals with the same parameters. Hence, it is possible to distinguish different kinds of motion.

The curve $S(\delta)$ versus δ, predicted by both algorithms, are presented in Figure 12. These curves present a linear range which tends to reach a stabilized value. The slope of the curve in this linear range estimates the maximum Lyapunov exponent and may be computed employing a linear regression. Hence, the algorithm due to Kantz (1994) furnishes $\lambda = 0.177 \pm 0.024$ while the algorithm due to Rosenstein *et al.* (1993) furnishes $\lambda = 0.153 \pm 0.010$. As expected, the system presents a positive exponent. Here, in contrast with the periodic signal analysis, both algorithms are capable to identify chaotic behavior.

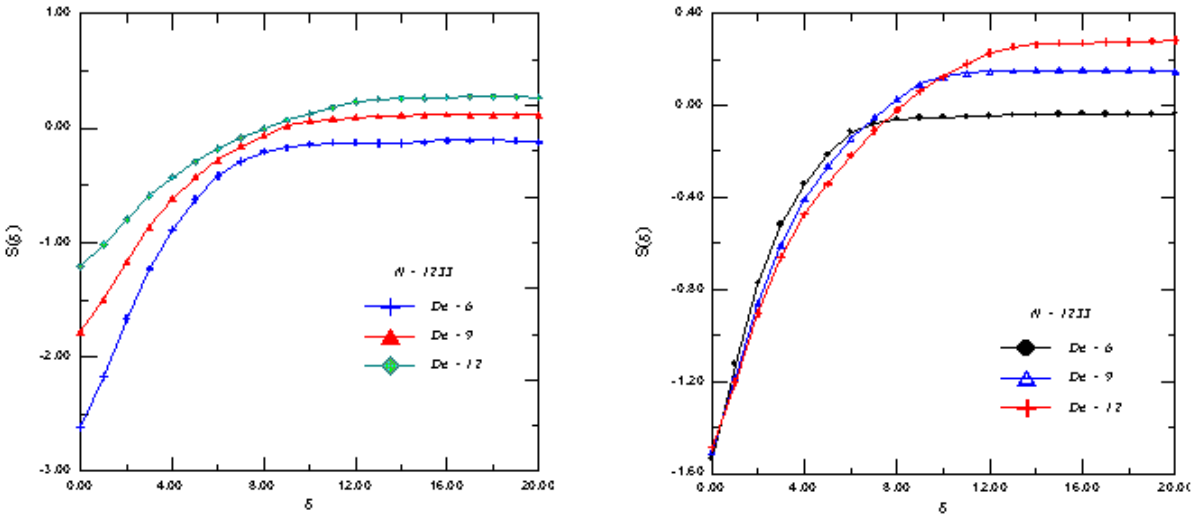


Figure 12. Lyapunov exponents associated with chaotic signal: $S(\delta)$ curves. (a) Kantz; (b) Rosenstein *et al.*

5. Conclusions

This contribution reports on the analysis of time series obtained from an experimental nonlinear pendulum. State space reconstruction is done employing the method of delay coordinates. Delay parameters are estimated with the average mutual information method to determine time delay and the false nearest neighbors method to estimate embedding dimension. Both methods present no noise sensibility. A procedure to construct the Poincaré Map is developed and presents good results. Lyapunov exponents are calculated employing the algorithms due to Kantz and due to Rosenstein *et al.* After performing the proposed calibration procedure, the Kantz's algorithm allows one to establish a difference between periodic and chaotic motion. The algorithm due to Rosenstein *et al.*, on the other hand, does not present good results for periodic motion. The authors agree that this contribution show that it is possible to distinguish periodic and chaotic time series obtained from an experimental set up without employing any kind of filters. Other dynamical systems must be analyzed in order to validate the present conclusions, however one believes that the procedures employed here can be applied to any system response.

6. Acknowledgments

The authors acknowledge the support of the Brazilian Research Council (CNPq).

7. References

Briggs, K., 1990, "An Improved Method for Estimating Lyapunov Exponents of Chaotic Time Series, *Physics Letters A*, v.151, n.1, 2, pp.27-32.

Broomhead, D.S. & King, G.P., 1986, "Extracting Qualitative Dynamics from Experimental Data", *Physica D*, v.20, pp.217-236.

Brown, R., 1993, "Calculating Lyapunov Exponents for Short and/or Noisy Data Sets", *Physical Review E*, v.47, n.6, pp.3962-3969.

Brown, R., Bryant, P. & Abarbanel, H.D.I., 1991, "Computing the Lyapunov Spectrum of a Dynamical System from an Observed Time Series", *Physical Review A*, v.43, n.6, pp.2787-2806.

Eckemann, J.-P., Kamphorst, S.O., Ruelle, D. & Ciliberto, S., 1986, "Lyapunov Exponents from Time Series", *Physical Review A*, v.34, n.6, pp.4971-4979.

Ellner, S., Nychka, D.W. & Gallant, A.R., 1992, "A Program to Estimate the Dominant Lyapunov Exponent of Noisy Nonlinear Systems from Time Series Data", *Institute of Statistics Mimeo Series*, #2235.

Franca, L.F.P., 2000, "Analysis of Chaotic Systems from Time Series: Application to Nonlinear Pendulum", M.Sc. Dissertation, Department of Mechanical and Materials Engineering - Military Institute of Engineering, in Portuguese.

Fraser, A.M. & Swinney, H.L., 1986, "Independent Coordinates for Strange Attractors from Mutual Information", *Physical Review A*, v.33, pp.1134-1140.

Kantz, H., 1994, "A Robust Method to Estimate the Maximal Lyapunov Exponent of a Time Series", *Physics Letters A*, v.185, pp.77-87.

Kruel, Th.–M., Eiswirth, M. & Schneider, F.W., 1993, "Computation of Lyapunov Spectra: Effect of Interactive Noise and Application to a Chemical Oscillator", *Physica D*, v.63, pp.117-137.

Meirovitch, L., 1986, "Elements of Vibration Analysis", McGraw-Hill.

Moon, F.C., 1992, "*Chaotic and Fractal Dynamics*", John Wiley & Sons.

Ogata, S., Iwayama, T. & Terachi, S., 1997, "Effect of System Noise on Chaotic Behavior in Rossler Type Nonlinear System", *International Journal of Bifurcation and Chaos*, v.7, n.12, pp.2872-2879.

Packard, N.J., Crutchfield, J.P., Fromer, J.D. & Shaw, R.S., 1980, "Geometry from a Time-Series", *Physical Review Letters*, v.45, n.9, pp.712-716.

Rosenstein, M.T., Collins, J.J. & De Luca, C.J., 1993, "A Practical Method for Calculating Largest Lyapunov Exponents from Small Data Sets", *Physica D*, v.65, pp.117-134.

Ruelle, D., 1979, "Ergotic Theory of Differentiable Dynamical Systems", *Mathématique of the Institut des Hautes Études Scientifiques*, n.5, pp.27.

Sano, M. & Sawada, Y., 1985, Measurement of the Lyapunov Spectrum from a Chaotic Time Series, *Physical Review Letters*, v.55, n.10, pp.1082-1085.

Takens, F., 1981, "Detecting Strange Attractors in Turbulence", Warwick, *Lecture notes in Mathematics*, v.898, Ed. D. Rand and L.-S Young, Springer, pp.366-381.

Wolf, A., Swift, J.B., Swinney, H.L. & Vastano, J.A., 1985, "Determining Lyapunov Exponents from a Time Series", *Physica D*, v.16, pp.285-317.

Zeng, X., Eykholt, R. & Pielke, R.A., 1991, "Estimating the Lyapunov Exponent Spectrum from Short Time Series of Low Precision", *Physical Review Letters*, v.66, pp.3229-3232.

XVI CONGRESSO BRASILEIRO DE ENGENHARIA MECÂNICA
16th BRAZILIAN CONGRESS OF MECHANICAL ENGINEERING

NUMERICAL INVESTIGATION OF CHAOTIC MOTION IN A SHAPE MEMORY TWO-BAR TRUSS

Marcelo A. Savi
Instituto Militar de Engenharia
Department of Mechanical and Materials Engineering
22.290.270 - Rio de Janeiro - RJ
E-Mail: savi@epq.ime.eb.br

Pedro M. C. L. Pacheco
CEFET/RJ
Department of Mechanical Engineering
20.271.110 - Rio de Janeiro - RJ
E-Mail: calas@cefet-rj.br

Arthur M. B. Braga
Pontifícia Universidade Católica do Rio de Janeiro
Department of Mechanical Engineering
22.453.900 - Rio de Janeiro - RJ
E-Mail: abraga@mec.puc-rio.br

Abstract. *The study of the structural response of two-bar trusses may be very helpful to understand some of the main stability characteristics of framed structures, as well as of flat arches and of many other physical phenomena associated with bifurcation buckling. This article is concerned with the dynamic response of a shape memory two-bar truss, which is an interesting example of a structural system that exhibit both kinematic and constitutive nonlinearities. A polynomial constitutive model is assumed to describe the behavior of the shape memory bars. Free and forced responses are investigated. Numerical simulations show that the system can easily reach a chaotic response.*

Key-words: : *Chaos, von Mises Truss, Shape Memory.*

1. Introduction

The study of the two-bar truss, also known as the von Mises truss, is important to define the main stability characteristics of framed structures as well as of flat arches and of many other physical phenomena associated with bifurcation buckling (Bazant & Cedolin, 1991). As depicted in Figure 1, this simple, plane, framed structure, is formed by two identical bars, both making an angle φ with an horizontal line, and free to rotate around their supports and at the joint. If the structure's mass is assumed to be lumped at the node, and only vertical, symmetrical motions of the truss are considered, the resulting discrete dynamical system is essentially one-dimensional.

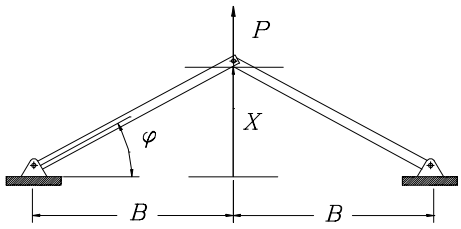


Figure 1. Two-bar truss (von Mises truss).

Despite the deceiving simplicity of the von Mises truss, its nonlinear dynamic response may exhibit a number of interesting, complex behaviors. For a given load level, two displacement configurations are possible. If the structure is loaded with a monotonically increasing force, the displacement path may jump from one configuration to another, presenting the snap-through behavior. The dynamic behavior is even richer when material nonlinearities are considered. In particular, the present contribution deals with two-bar trusses made from shape memory materials.

Pseudoelasticity and shape memory are both associated with thermoelastic martensitic transformations. The shape memory effect, present in various metallic alloys, is a phenomenon where plastically deformed objects may recover their original form after going through a proper heat treatment. The pseudoelastic behavior is characterized by complete strain recovery accompanied by large hysteresis in a loading-unloading cycle (Otsuka & Ren, 1999). Fibers of shape-memory alloys (SMAs) can be used to fabricate hybrid composites exhibiting these two different but related material behaviors. Detailed description of the shape memory effect and other phenomena associated with martensitic phase transformations, as well as examples of applications in the context of smart structures, may be found in van Humbeeck (1999), Duerig *et al.* (1999), Birman (1997), Rogers (1995), Shaw & Kyriakides (1995) and Schetky (1979).

Adaptive trusses, with shape memory actuators for low-frequency vibration control, are examples of dynamical systems that may behave as the structure considered in this paper. The term adaptive structure has been used to identify structural systems that are capable of changing their geometry or physical properties with the purpose of performing a specific task. An adaptive structure must be equipped with actuators that induce such controlled alterations. Among the many possible choices for actuators, those made of SMAs have shown a great potential in situations where high force, large strain, and low frequency structural control are needed. SMA actuators are easy to manufacture, relatively lightweight, and able of producing high forces or displacements with low power consumption.

This article is concerned with the dynamic response of a shape memory two-bar truss, which is an interesting example of a structural system that exhibit both kinematic and constitutive nonlinearities. A polynomial constitutive model is assumed to describe the behavior of the shape memory bars. Free and forced responses are investigated. It is shown that for a certain range of temperatures, the system may present up to eleven equilibrium configurations. Due to this rich structure, the system can easily reach a chaotic response even at moderate loads and frequencies.

2. Formulation

In the present investigation we consider a shape memory two-bar truss where each bar presents the shape memory and pseudoelastic effects. The two identical bars have length L and cross section area A. They form an angle φ with a horizontal line and are free to rotate around their supports and at the joint, but only on the plane formed by the two bars, (Figure 1). The critical Euler load of both bars is assumed to be sufficiently large so that buckling will not occur in the simulations reported here.

We further assume that the structure's mass is entirely concentrated at the junction between the two bars. Hence, the structure is divided into segments without mass, connected by nodes with lumped mass that is determined by static considerations. We consider only symmetric motions of the system, which implies that the concentrated mass, m, can only move vertically. The symmetric, vertical displacement is denoted by X. Under these assumptions, the balance of momentum is expressed through the following equation of motion

Consider a shape memory two-bar truss where each bar presents the shape memory and pseudoelastic effects. The two identical bars have length L and cross section area A. The bars form an angle φ with a horizontal line and are free to rotate, on its own plane, around the supports and at the joint (Figure 1). The critical Euler load of both bars is assumed to be sufficient large that buckling does not occurs for the presented simulations.

We further assume that the structure's mass is entirely concentrated at the junction between the two bras. Hence, the structure is divided into segments without mass, connected by nodes with lumped mass that is determined by static considerations. We consider only symmetric motions of the system, which implies that the concentrated mass, m, can only move vertically. The symmetric, vertical displacement is denoted by X. Under these assumptions, the balance of momentum is expressed through the following equation of motion,

$$-2F\sin\varphi + P = m\ddot{X} \tag{1}$$

where F is the force on each bar while P is an external force.

Shape memory and pseudoelastic behaviors of the bars are described by the polynomial constitutive model (Falk, 1980). This is a one-dimensional model that proposes a sixth degree polynomial free energy function in terms of the uniaxial strain, ε. The form of the free energy is chosen in such a way that its minima and maxima are respectively associated with the stability and instability of each phase of the SMA. As it is usual in one-dimensional models proposed for SMAs (Savi & Braga, 1993), three phases are considered: Austenite (A) and two variants of martensite ($M+$, $M-$). Hence, the free energy is chosen such that for high temperatures it has only one minimum at vanishing strain, representing the equilibrium of the austenitic phase. At low temperatures, martensite is stable, and the free energy must have two minima at non-vanishing strains. At intermediate temperatures, the free energy must have equilibrium points corresponding to both phases. Under these restrictions, the uniaxial stress, σ, is a fifth-degree polynomial of the strain (Savi & Braga, 1993), *i.e.*

$$\sigma = a_1(T - T_M)\varepsilon - a_2\varepsilon^3 + a_3\varepsilon^5 \tag{2}$$

where a_1, a_2 and a_3 are material constants, and T the temperature, while T_M is the temperature below which the martensitic phase is stable. If T_A is defined as the temperature above which austenite is stable, and the free energy has only one minimum at zero strain, it is possible to write the following condition,

$$T_A = T_M + \frac{1}{4}\frac{a_2^2}{a_1 a_3} \tag{3}$$

Therefore, the constant a_3 may be expressed in terms of other constants of the material. If the following strain definition is considered,

$$\varepsilon = \frac{L}{L_0} - 1 = \frac{\cos\varphi_0}{\cos\varphi} - 1 \tag{4}$$

the equation of motion may be rewritten as follows,

$$\begin{aligned} m\ddot{X} + \frac{2A}{L_0} X \Big\{ & [a_1(T - T_M) - 3a_2 + 5a_3] + [-a_1(T - T_M) + a_2 - a_3] L_0 (X^2 + B^2)^{-1/2} + \\ & + [3a_2 - 10a_3] \frac{1}{L_0} (X^2 + B^2)^{1/2} + [-a_2 + 10a_3] \frac{1}{L_0^2} (X^2 + B^2) - \frac{5a_3}{L_0^3} (X^2 + B^2)^{3/2} + \\ & + \frac{a_3}{L_0^4} (X^2 + B^2)^2 \Big\} = P(t) \end{aligned} \tag{5}$$

Considering a periodic excitation $P = P_0 \sin(\omega t)$, and introducing a linear viscous dissipation, the equation of motion may be written in non-dimensional form as,

$$\begin{aligned} x' &= y \\ y' &= \gamma \sin(\Omega\tau) - \xi y + x\{ -[(\theta - 1) - 3\alpha_2 + 5\alpha_3] + [(\theta - 1) - \alpha_2 + \alpha_3](x^2 + b^2)^{-1/2} \\ &\quad - [3\alpha_2 - 10\alpha_3](x^2 + b^2)^{1/2} + [\alpha_2 - 10\alpha_3](x^2 + b^2) + 5\alpha_3 (x^2 + b^2)^{3/2} - \alpha_3 (x^2 + b^2)^2 \} \end{aligned} \tag{6}$$

where $x = \frac{X}{L}$, $\omega_0^2 = \frac{2Aa_1T_M}{mL_0}$, $\gamma = \frac{P_0}{mL_0\omega_0^2}$, $\xi = \frac{c}{m\omega_0}$, $b = \frac{B}{L_0}$, $\tau = \omega_0 t$, $\Omega = \frac{\omega}{\omega_0}$,

$\alpha_2 = \frac{a_2}{a_1T_M}$, $\alpha_3 = \frac{a_3}{a_1T_M}$, $\theta = \frac{T}{T_M}$, $\theta_A = \frac{T_A}{T_M}$ and $()' = \frac{d()}{d\tau}$.

Numerical simulations are performed employing a fourth-order Runge-Kutta scheme with time steps chosen to be less than $\Delta\tau = 2\pi/200\Omega$. Nonlinear analysis involves the determination of quantities, known as dynamical invariants, which are important to identify chaotic behavior. Lyapunov exponents have been used as the most useful dynamical diagnostic tool for chaotic system analysis. These exponents evaluate the sensitive dependence to initial conditions estimating the exponential divergence of nearby orbits. The signs of the Lyapunov exponents provide a qualitative picture of the system's dynamics and any system containing at least one positive exponent presents chaotic behavior. Lyapunov exponents can also be used for the calculation of other invariant quantities as the attractor dimension, which may be determined by the Kaplan-Yorke conjecture (Kaplan & Yorke, 1983). The determination of Lyapunov exponents of dynamical system with an explicitly mathematical model, which can be linearized, is well established from the algorithm proposed by Wolf *et al.* (1985). Here, we employ this algorithm to calculate the Lyapunov spectrum, while the Lyapunov dimension is the parameter applied to define the attractor's dimension from the Lyapunov spectrum (Frederickson *et al.*, 1983; Farmer *et al.*, 1983; Wolf *et al.*, 1985).

In all simulations, we have used the material properties presented in Table 1. These values were chosen in order to, as shown in Figure 2, match experimental data obtained by Sittner *et al.* (1995) for a Cu-Zn-Al-Ni alloy at 373K. For the data in Table 1, the parameters defined in Equation (5) assume the values: $\alpha_2 = 4.34e-3$ and $\alpha_3 = 20.83e-6$. We further let $b = 0.866$, corresponding to a two-bar truss with an initial position $\varphi_0 = 30^0$.

Table 1. Material Properties.

a_1 (MPa/K)	a_2 (MPa)	a_3 (MPa)	T_M (K)	T_A (K)
523.29	1.868×10^7	2.186×10^9	288	364.3

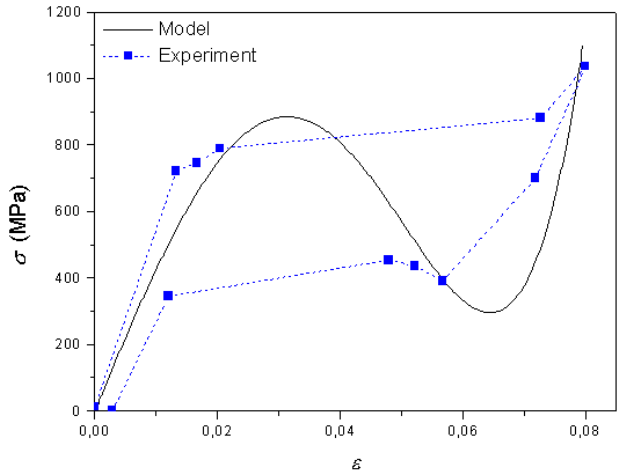


Figure 2. Stress-strain curve: experimental and predicted by polynomial model.

3. Free Vibration

In this Section, we discuss the free response of the shape memory two-bar truss. This is done by letting γ vanish in the equations of motion (5). It is well known that the von Mises truss presents three equilibrium points due to kinematics nonlinearity (Bazant & Cedolin, 1991). Of those, two are stable while the other one is unstable. In the case of a shape memory two-bar truss, constitutive nonlinearity introduces a different behavior. Denoting by $(\bar{x}, \bar{y})$ a point that makes the right-hand sides of equations of motion vanish, the following possibilities are found,

$$\bar{x} = 0 \text{ and } \bar{y} = 0 \tag{7}$$

$$\bar{x} = \pm\sqrt{1-b^2} \text{ and } \bar{y} = 0$$

$$\bar{x} = \pm\sqrt{\frac{1}{2\alpha_3}\left[\sqrt{\chi_1} \pm \sqrt{8\alpha_3\left(\alpha_2 + \sqrt{\chi_1}\right) + \chi_2}\right]} \text{ and } \bar{y} = 0$$

$$\bar{x} = \pm\sqrt{\frac{1}{2\alpha_3}\left[-\sqrt{\chi_1} \pm \sqrt{8\alpha_3\left(\alpha_2 - \sqrt{\chi_1}\right) + \chi_2}\right]} \text{ and } \bar{y} = 0$$

where $\chi_1 = \alpha_2^2 - 4\alpha_3(\theta - 1)$ and $\chi_2 = 2\alpha_3(1-b^2) + \alpha_2$. Of these eleven possibilities, only those that correspond to real numbers have physical meaning. Stability of these equilibrium configurations may be determined by the behavior of the system in their neighborhood. An analysis of the eigenvalues of the Jacobian matrix of the system reveals its local stability. Therefore,

(a) $0 < \theta < 1$ ($\alpha_2^2 < \chi_1 < 4\alpha_3 + \alpha_2^2$): the system has seven fixed points: The origin of the phase space and $\bar{x} = \pm\sqrt{1-b^2}$ are saddle points. The other four fixed points are centers. This is consistent with low temperature behavior of SMA, where two martensitic phases are stable.

(b) $1 \le \theta < \theta_A$ ($0 < \chi_1 \le \alpha_2^2$): the system has eleven fixed points, being five saddle points and the remaining are centers. The existence of six stable fixed points is explained by the stability of both martensitic phases and austenite in this range of temperature.

(c) $\theta \ge \theta_A$ ($\chi_1 \le 0$): the system has three fixed point. The origin and $\bar{x} = \pm\sqrt{1-b^2}$ are centers. Under this temperature range, austenite is the only stable phase in the stress-free SMA.

A map of equilibrium points as a function of temperature is presented in Figure 3 for Cu-Zn-Al-Ni alloy. At low temperatures, where the martensitic phase is stable ($0 \le \theta < 1$), there are seven equilibrium points, four of them stable while the others are unstable. By considering a higher temperature, where both martensitic and austenitic phases may coexist ($1 \le \theta < \theta_A$), the system exhibits five unstable and six stable equilibrium points. At an even higher temperature, where only the austenitic phase is present ($\theta \ge \theta_A$), the system has one unstable and two stable equilibrium points.

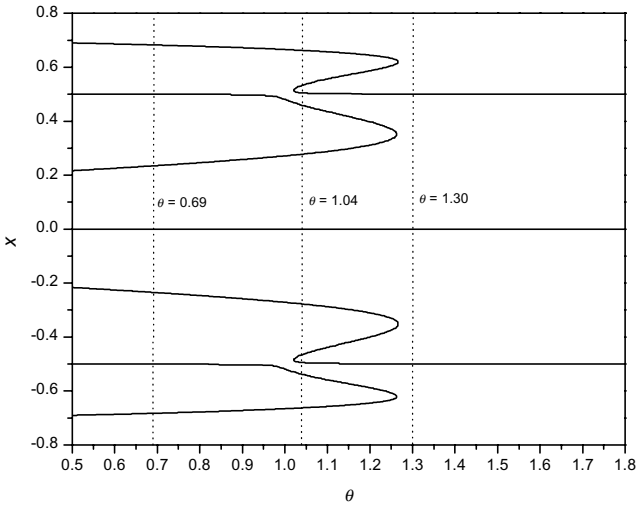


Figure 3. Map of equilibrium points as a function of temperature.

In order to illustrate the free response of the shape memory truss, we consider the non-dissipative system, that is, we let $\xi = 0$ in equation (5). Results from simulations are presented in the form of phase portraits. Figure 4 presents the free-response of the system at different temperatures. Figure 4a considers a temperature where the martensitic phase is stable ($\theta = 0.69$). There are, in this case, seven equilibrium points. From these, four are stable while the other three are unstable. At a higher temperature, when martensitic and austenitic phases are both present in the alloy ($\theta = 1.04$), the system, as depicted in Figure 4b, has five unstable and six stable equilibrium points. Figure 4c is representative of the free response at higher temperatures ($\theta = 1.30$), when the alloy is fully austenitic. As discussed above, in this case, the system has two centers and a single saddle point.

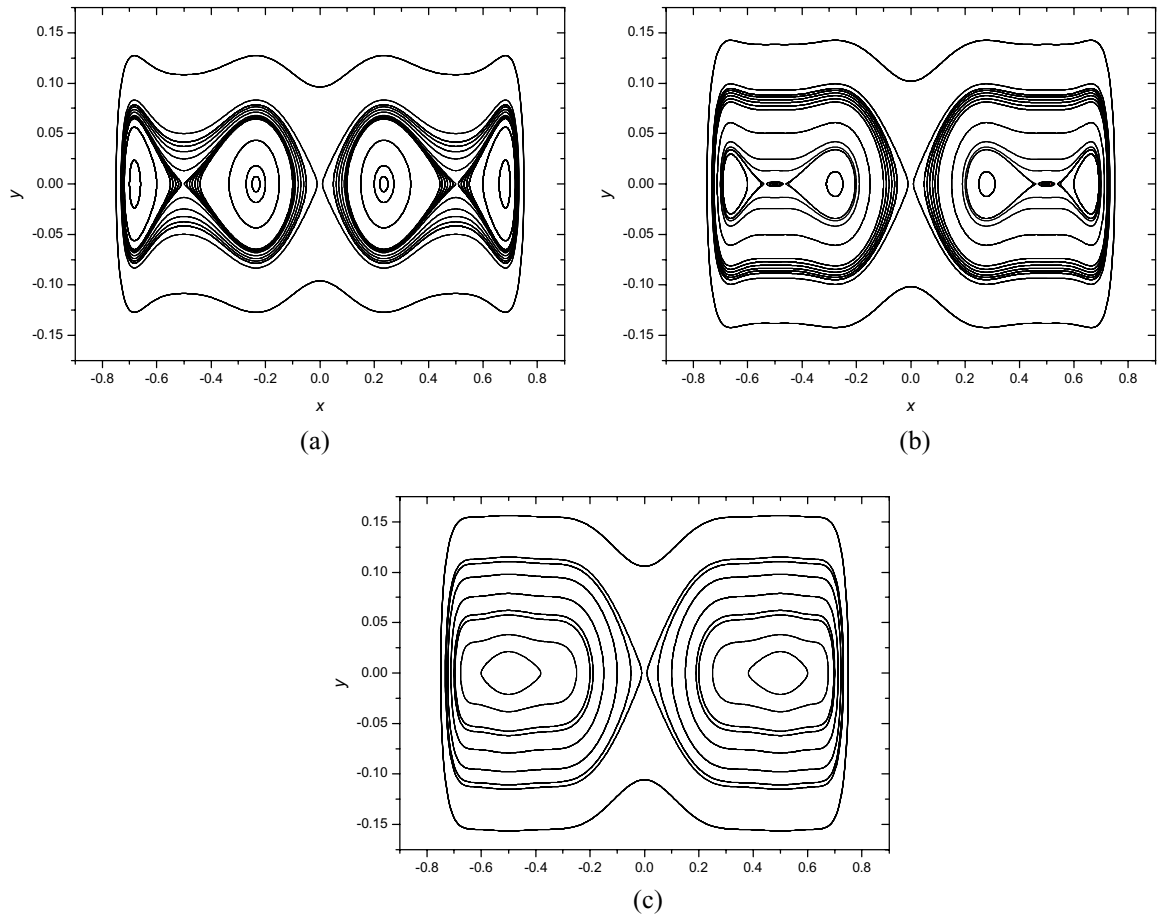


Figure 4. Phase portrait (a) $\theta = 0.69$; (b) $\theta = 1.04$; (c) $\theta = 1.30$.

4. Forced Vibration

In this section, we address the forced vibration response. The dynamic behavior is now much richer. Periodic, quasi-periodic and chaotic motions are all possible in the two-bar truss. In order to start the analysis, let us consider the bifurcation diagram which represents stroboscopically sampled displacement values, x, under the slow quasi-static increase of the driving force amplitude, γ. Different temperature, frequency and dissipation parameters are considered.

At first, we consider a temperature $\theta = 0.69$ where the martensitic phase is stable. The frequency parameter, Ω, is fixed at 0.5, while different values for dissipation are adopted. Figure 5 shows bifurcations and clouds of points related to chaotic motions.

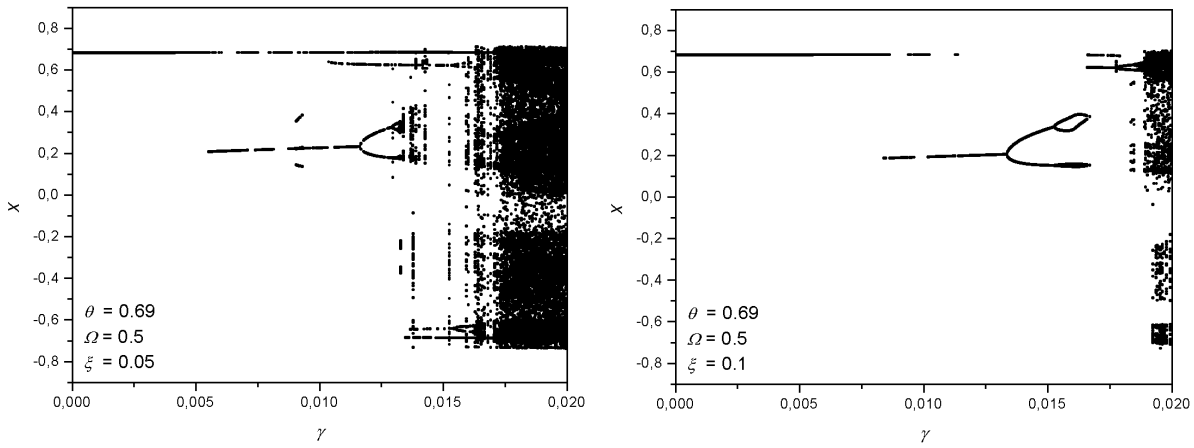


Figure 5. Bifurcation diagrams varying γ for $\theta = 0.69$ and $\Omega = 0.5$.

When $\Omega = 0.5$ and $\gamma = 0.020$, the two-bar truss exhibit chaotic motion. Poincaré mapping eliminates one dimension of the system by sampling the displacement and velocity stroboscopically at time intervals, summarizing the dynamical system behavior. Figure 6 shows Poincaré sections of these motions for two different dissipation parameters ($\xi = 0.05$ and $\xi = 0.1$). Both situations present strange attractors with fractal dimension, which may be evaluated from the Kaplan-Yorke conjecture (Wolf *et al.*, 1985). The Lyapunov spectrum associated with $\xi = 0.05$ is $\lambda_i = (+0.14, -0.21)$ and the Lyapunov dimension is $D = 1.66$. When $\xi = 0.1$, we obtain $\lambda_i = (+0.09, -0.23)$ and $D = 1.38$. As expected, higher dissipation tends to occupy a small region in phase space and the attractor dimension is smaller.

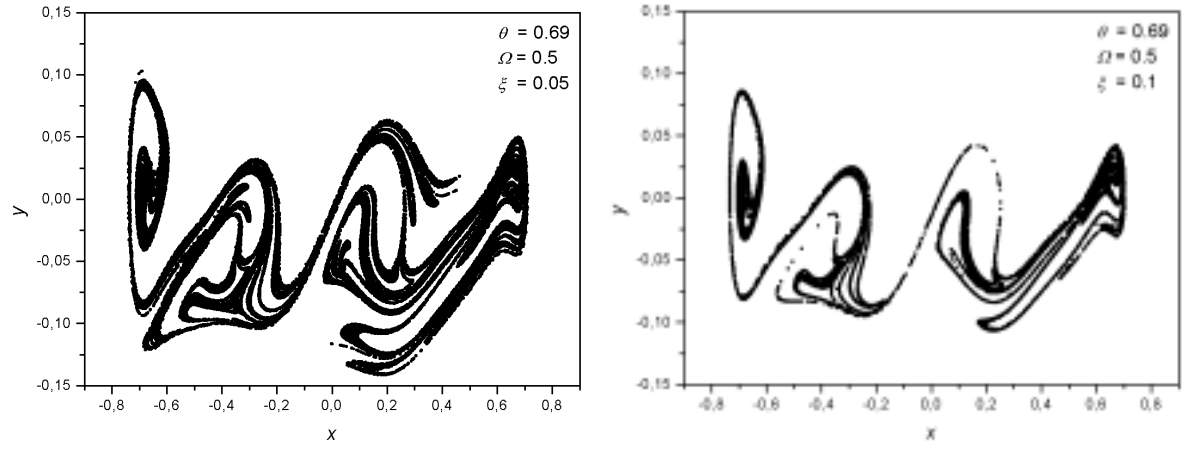


Figure 6. Strange attractors ($\theta = 0.69$, $\Omega = 0.5$, $\gamma = 0.020$).

Now, a higher temperature is considered and the austenitic phase is stable ($\theta = 1.30$). Bifurcation diagrams presented in Figure 7 shows the influence of the variation of the parameter γ for different dissipation values. Notice that for higher values of ξ, the behavior of the two-bar truss tends to be more regular.

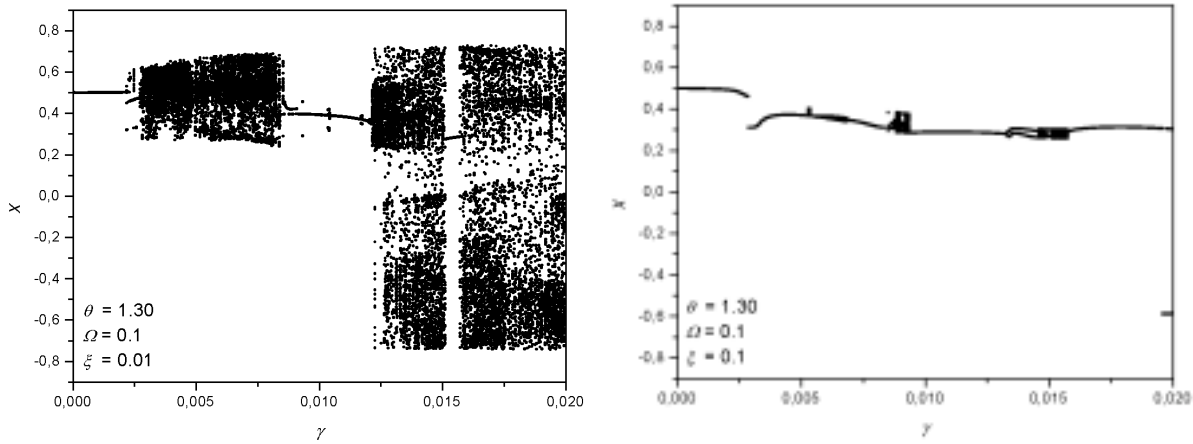


Figure 7. Bifurcation diagrams varying γ for $\theta = 1.30$ and $\Omega = 0.1$.

Let us consider an excitation with $\Omega = 0.1$ and $\gamma = 0.015$. With these parameters the two-bar truss present different behaviors for $\xi = 0.01$ and $\xi = 0.1$ (Figure 8). When $\xi = 0.01$, the system behaves chaotically and the Lyapunov spectrum is characterized by $\lambda_i = (+0.02, -0.04)$, while $D = 1.59$. When $\xi = 0.1$, a period-3 motion occurs.

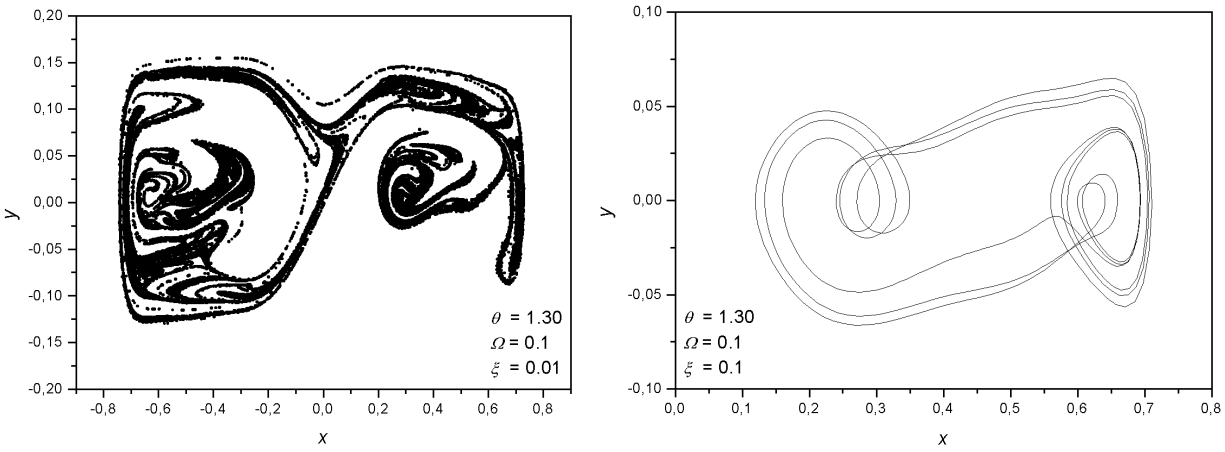


Figure 8. Response for $\theta = 1.30$, $\Omega = 0.1$, $\gamma = 0.015$. Strange attractor ($\xi = 0.01$) and period-3 motion ($\xi = 0.1$);

Considering a temperature where both martensite and austenite coexist, the strange attractor changes. Figure 9 shows the strange attractor related to the motion where $\theta = 1.04$, $\Omega = 0.1$, $\gamma = 0.015$ and $\xi = 0.01$. The Lyapunov spectrum of this motion is $\lambda_i = (+0.03, -0.05)$ while $D = 1.68$. As expected, this attractor has a dimension greater than that at higher temperature. Increasing the dissipation ($\xi = 0.1$) the system presents a period-3 motion.

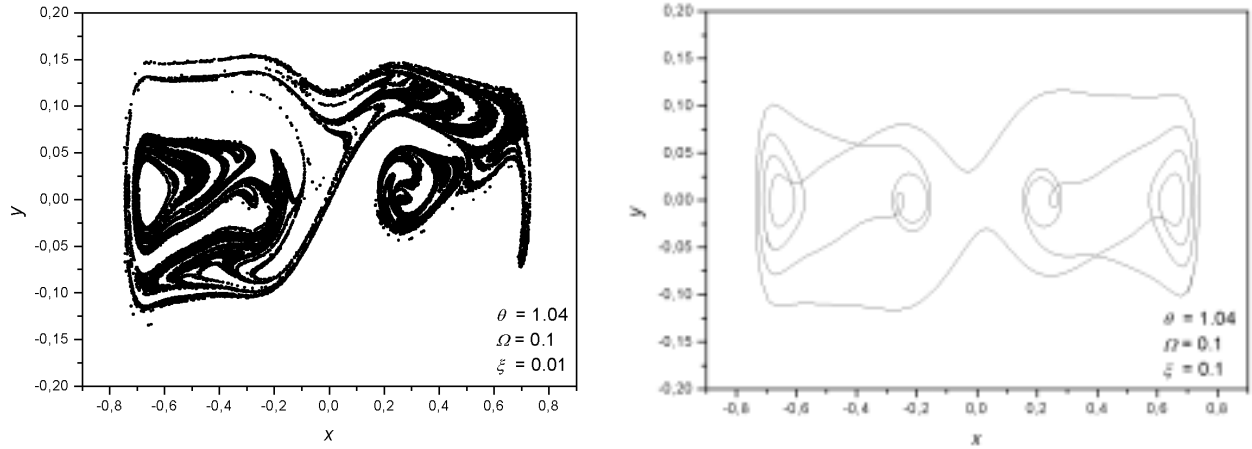


Figure 9. Response for $\theta = 1.04$, $\Omega = 0.1$, $\gamma = 0.015$. Strange attractor ($\xi = 0.01$) and period-3 motion ($\xi = 0.1$).

5. Conclusions

This article reports results from numerical simulations of the dynamical response of a shape memory two-bar truss. A polynomial constitutive model was assumed to describe the constitutive behavior of the bars. Free and forced responses were investigated. Numerical simulations were carried on employing a fourth order Runge-Kutta scheme for numerical integration, and the characterization of chaos was performed with Lyapunov exponents, evaluated with the aid of an algorithm proposed by Wolf *et al.* (1985). Results have shown that the system may present a number of interesting, complex behaviors. Basically, the system response is a combination of kinematics and constitutive nonlinearities. Due to this combination, it may present eleven equilibrium points, which are responsible for a very rich dynamics. Strange attractors of this system vary their pattern according with temperature.

6. References

Bazant, Z.P. & Cedolin, L., 1991, "*Stability of Structures*", Oxford Press.

Birman, V., 1997, "Review of Mechanics of Shape Memory Alloy Structures", *Applied Mechanics Review*, v. 50**,** pp. 629-645.

Duerig, T., Pelton, A. & Stöckel, D., 1999, "An overview of Nitinol Medical Applications", *Materials Science and Engineering A*, v.273-275, pp.149-160.

Falk, F., 1980, "Model Free Energy, Mechanics, and Thermodynamics of Shape Memory Alloys", *Acta Metallurgica*, v.28, pp.1773-1780.

Farmer, J.D., Ott, E. & Yorke, J.A., 1983, "The Dimension of Chaotic Attractors", *Physica 7D*, pp.153-180.

Frederickson, P., Kaplan, J., Yorke, E. & Yorke, J., 1983, "The Lyapunov Dimension of Strange Attractors", *Journal of Differential Equations*, 49, 185.

Kaplan J. L. & Yorke, J.A., 1983, "The Lyapunov Dimension of Strange Attractors", *Journal of Differential Equations*, v.49, pp.185.

Otsuka, K. & Ren, X., 1999, "Recent developments in the research of Shape Memory Alloys", *Intermetallics*, v.7, pp.511-528.

Rogers, C.A., 1995, "Intelligent Materials", *Scientific American*, September, pp.122-127.

Savi, M.A. & Braga, A.M.B., 1993, "Chaotic Vibration of an Oscillator with Shape Memory", *Journal of the Brazilian Society of Mechanical Sciences*, v.15, n.1, pp.1-20.

Schetky, L. M., 1979, "Shape Memory Alloys", *Scientific American*, v.241 pp.68-76.

Shaw, J. A. & Kyriakides, S., 1995, "Thermomechanical Aspects of NiTi", *Journal of Mechanics Phys. Solids*, v.43, pp. 1243-1281.

Sittner, P. Hara, Y. & Tokuda, M., 1995, "Experimental Study on the Thermoelastic Martensitic Transformation in Shape Memory Alloy Polycrystal Induced by Combined External Forces", *Metallurgical and Materials Transactions A*, v.16A, pp.2923-2935.

Thompson, J.M.T., 1993, "Chaos and Fractal Basin Boundaries in Engineering". In: *The Nature of Chaos* (ed. T. Mullin), Oxford Press.

van Humbeeck, J., 1999, "Non-medical Applications of Shape Memory Alloys", *Materials Science and Engineering A*, v. 273-275, pp.134-148.

Wolf, A., Swift, J.B., Swinney, H.L. & Vastano, J.A., 1985, "Determining Lyapunov Exponents from a Time Series", *Physica 16D*, pp.285-317.

XVI CONGRESSO BRASILEIRO DE ENGENHARIA MECÂNICA
16th BRAZILIAN CONGRESS OF MECHANICAL ENGINEERING

BIFURCAÇÕES E CRISES EM UM OSCILADOR COM MEMÓRIA DE FORMA

Luciano G. Machado
Marcelo A. Savi
Instituto Militar de Engenharia
Departamento de Engenharia Mecânica e de Materiais
22.290.270 – Rio de Janeiro - RJ
E-Mail: lmachado@epq.ime.eb.br, savi@epq.ime.eb.br

Pedro M. C. L. Pacheco
CEFET/RJ
Departamento de Engenharia Mecânica
20.271.110 - Rio de Janeiro – RJ
E-Mail: calas@cefet-rj.br

Resumo. *A dinâmica de sistemas com atuadores com memória de forma apresentam características intrinsecamente não-lineares e uma resposta muito rica. A resposta caótica é um exemplo deste tipo de comportamento. Este trabalho analisa alguns aspectos associados ao fenômeno da bifurcação em um oscilador com memória de forma onde a força de restituição é descrita a partir de um modelo constitutivo polinomial. O termo bifurcação está associado a uma mudança qualitativa na estrutura de uma solução, como consequência de uma variação dos parâmetros do sistema. O fenômeno da bifurcação está estreitamente relacionado com o caos. Os resultados da análise mostram que a resposta do oscilador com memória de forma apresenta cascatas de duplicação de período, direta e reversa, e crises na resposta do sistema.*

Palavras chave: *Memória de forma, caos, bifurcação, crise.*

1. Introdução

Os efeitos de memória de forma e pseudoelasticidade são fenômenos termoelásticos, associados à transformação de fase martensítica, presentes nas ligas com memória de forma (SMAs). As transformações de fase martensíticas são processos não-difusivos envolvendo fases sólidas que ocorrem a velocidades muito elevadas. Atribui-se a causa dessas transformações a diferença de energia livre entre as estruturas constituintes envolvidas no processo, o que induz modificações nas ligações químicas, tornando as transformações de fase de caráter essencialmente cristalográfico (Wasilevski, 1975). Essas transformações apresentam como principais características a não-dependência do tempo, forte dependência da temperatura e a propriedade de reversibilidade.

As ligas Ni-Ti, Cu-Zn, Cu-Zn-Al, Mg-Cu, Fe-Mn-Si, e Cr-Ni são algumas das SMAs que vêm motivando diversas aplicações em diferentes áreas do conhecimento, dentre as quais vale destacar as indústrias aeroespacial, automotiva e biomédica. Dentre estas, vale citar inúmeros dispositivos termo-ativados como válvulas pneumáticas para o controle de fluxo. No controle estático ou dinâmico de estruturas flexíveis, as SMAs podem ser utilizadas como sensores ou atuadores, sendo usualmente empregadas como fibras embebidas em uma matriz, formando o que se chama de compósito híbrido (Rogers *et al.*, 1991).

A dinâmica de sistemas com atuadores com memória de forma apresentam características intrinsecamente não-lineares e uma resposta muito rica. A resposta caótica é um exemplo deste comportamento. Savi & Braga (1993a,b) discutem a resposta caótica de osciladores onde a força de restituição é proporcionada por molas helicoidas com memória de forma. Savi & Pacheco (2002) estudam algumas características de osciladores com um e dois graus de liberdade, mostrando a presença de caos e hipercaos na resposta dos sistemas. Este trabalho analisa alguns aspectos associados às bifurcações em um oscilador com memória de forma onde a força de restituição é descrita a partir de um modelo constitutivo polinomial. O termo bifurcação está associado a uma mudança qualitativa na estrutura de uma solução, como consequência de uma variação dos parâmetros do sistema. Este termo foi usado pela primeira vez por Poincaré para expressar uma divisão das soluções de equilíbrio. O fenômeno da bifurcação está estreitamente relacionado com a existência do caos. Os resultados deste trabalho mostram cascatas de duplicação de período, direta e reversa, e crises na resposta do sistema.

2. Modelo Polinomial

Dentre todas as teorias fenomenológicas que se propõe a descrever os efeitos de memória de forma e pseudoelasticidade, o modelo polinomial é a mais simples. Este modelo foi proposto por Falk (1980) baseado na Teoria de Devonshire. Neste modelo, a Energia Livre de Helmholtz é uma função da temperatura, T, e da deformação unidimensional, E, isto é, $\psi = \psi(E, T)$ (Savi & Braga, 1993a).

A escolha desta energia é feita de tal forma que, para altas temperaturas, ela possua somente um ponto de mínimo em relação a deformação e, na ausência de forças externas, a única fase estável é a austenita. Para temperaturas intermediárias, ψ possui três mínimos. Além da austenita estável, para $E = 0$, encontra-se também a presença de duas variantes martensíticas, $M+$ e $M-$. Para baixas temperaturas, a energia livre possui dois pontos de mínimo que

representam as variantes de martensita *M+* e *M−* (Falk, 1980). Desta forma, a Energia Livre de Helmholtz assume a forma de um polinômio de sexto grau, enquanto a relação tensão-deformação assume a seguinte forma (Savi & Braga, 1993):

$$\sigma = \rho \frac{\partial \psi}{\partial E} = a(T - T_M)E - bE^3 + eE^5 \tag{1}$$

onde *a*, *b* e *e* são constantes positivas do material, T_M é a temperatura abaixo da qual só a fase martensítica é termodinâmica estável e ρ é a massa específica. Uma das dificuldades no uso do modelo polinomial é a falta de dados experimentais para avaliar as constantes do material.

3. Oscilador com Memória de Forma

Considere um oscilador, com um grau de liberdade, que consiste de uma massa m acoplada a um elemento de liga com memória de forma de comprimento *L* e um amortecedor, viscoso linear, cujo parâmetro é *c* (Figura 1). O sistema é excitado harmonicamente através de uma força $F = F_0 \operatorname{sen}(\Omega t)$.

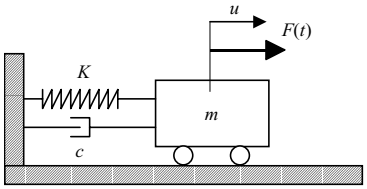


Figura 1. Oscilador com memória de forma.

A equação do movimento é dada por,

$$m\ddot{u} + c\dot{u} + \bar{a}(T - T_M)u - \bar{b}u^3 + \bar{e}u^5 = F_0 \operatorname{sen}(\omega t) \tag{2}$$

onde

$$\bar{a} = \frac{aA}{L} \;;\; \bar{b} = \frac{bA}{L^3} \;;\; \bar{e} = \frac{eA}{L^5} \tag{3}$$

Adimensionalizando a equação, chega-se a uma expressão do tipo:

$$U'' + \xi U' + (\theta - 1)U + \beta U^3 + \varepsilon U^5 = \delta \operatorname{sen}(\varpi \tau) \tag{4}$$

onde

$$U = \frac{u}{L} \;;\; \xi = \frac{c}{m\omega_0} \;;\; \theta = \frac{T}{T_M} \;;\; \beta = \frac{\bar{b}}{mL^3\omega_0^2} \;;\; \varepsilon = \frac{\bar{e}}{mL^5\omega_0^2} \;;$$
$$\delta = \frac{F_0}{mL\omega_0^2} \;;\; \omega_0^2 = \frac{\bar{a}T_M}{m} \;;\; \varpi = \frac{\omega}{\omega_0} \;;\; \tau = \omega_0 t \;;\; (\,)' = d(\,)/d\tau \tag{5}$$

A equação do movimento pode ser reescrita em termos de um sistema de equações de primeira ordem,

$$\begin{aligned} y_1' &= y_2 \\ y_2' &= \delta \operatorname{sen}(\varpi\tau) - \xi y_2 - (\theta - 1)y_1 + \beta y_1^3 - \varepsilon y_1^5 \end{aligned} \tag{6}$$

Neste trabalho, simulações numéricas são realizadas utilizando o método Runge Kutta de quarta ordem com passos menores que $\Delta t = 2\pi / 200$. Para todas as simulações estudadas consideram-se os parâmetros apresentados na Tabela 1. A caracterização do comportamento caótica é feita através dos expoentes de Lyapunov que avaliam a média da divergência exponencial entre trajetórias vizinhas. A determinação destes expoentes é feita a partir do algoritmo de Wolf *et al.* (1985).

Tabela 1. Parâmetros utilizados na simulação do oscilador de memória de forma.

ξ	0,2
ϖ	1
β	1,3e3
ε	4,7e5

4. Bifurcações

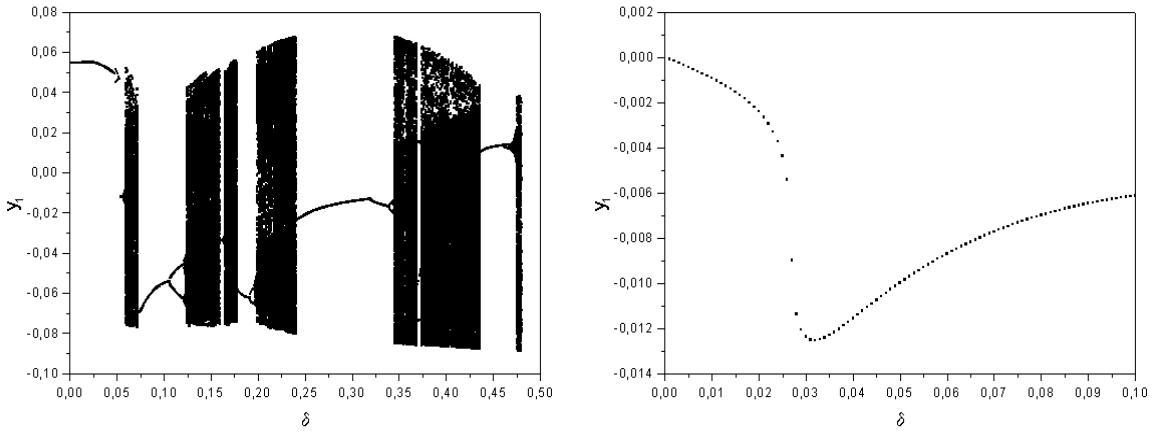
O termo bifurcação está associado a uma mudança qualitativa na estrutura de uma solução, como consequência de uma variação dos parâmetros do sistema. Este termo foi usado pela primeira vez por Poincaré para expressar uma divisão das soluções de equilíbrio. O fenômeno da bifurcação está estreitamente relacionado com a existência do caos no sentido de que um sistema dinâmico que não apresenta algum tipo de bifurcação não apresenta uma resposta caótica. Deve-se estar atento, no entanto, que a recíproca não é verdadeira, ou seja, um sistema que apresente bifurcações não necessariamente apresentará uma resposta caótica.

Os diagramas de bifurcação apresentam informações úteis para identificar a influência de um dado parâmetro na resposta de um sistema. O diagrama de bifurcação apresenta a distribuição estroboscópica da resposta do sistema a partir de uma variação lenta de um dado parâmetro (Thompsom & Stewart, 1986). Desta forma, é possível ter-se uma visão global sobre os efeitos da variação deste parâmetro na resposta.

Neste trabalho, tem-se interesse em considerar as bifurcações ocorridas em um oscilador com memória de forma. Analisam-se as bifurcações associadas ao parâmetro de forçamento, δ, assim como ao parâmetro relacionado com a temperatura, θ. A análise destas bifurcações permite identificar comportamentos interessantes tais como, cascatas de duplicação de período, direta e reversa, crises e *subductions*. Os diagramas de bifurcação apresentados neste trabalho avaliam a resposta do sistema em uma seção de Poincaré, desprezados os 300 primeiros períodos.

4.1. Bifurcações Associadas ao Parâmetro de Forçamento δ

Esta seção discute a influência das variações do parâmetro relacionado com a amplitude de forçamento, δ, na resposta do sistema. Inicialmente, considere os diagramas de bifurcação avaliados em duas temperaturas distintas. Na primeira θ = 0,7, tem-se que a fase martensítica é estável (Figura 2a). Nesta situação, observam-se regiões de bifurcações e regiões caóticas, caracterizadas por uma nuvem de pontos. Por outro lado, a temperatura θ = 3,5, Figura 2b, onde a fase austenítica é estável, apresenta um comportamento onde não surgem bifurcações ou caos. Desta forma, dedica-se especial interesse a condição onde $\theta < 1$, que representa temperaturas onde a fase martensítica é estável.



(a) (b)

Figure 2. Diagramas de bifurcação em δ. (a) θ = 0.7; (b) θ = 3.5.

Considere então uma temperatura θ = 0,7, onde a fase martensítica é estável. Observando o diagrama de bifurcação em um intervalo, $0 \le \delta \le 0{,}5$, vê-se o aparecimento de sucessivas regiões de comportamento caótico. Estas regiões estão separadas entre si por regiões de comportamento periódico (Figura 2a). A Figura 3a apresenta uma ampliação deste diagrama no intervalo $0 \le \delta \le 0{,}1$, onde observam-se os valores de δ para os quais têm-se comportamento periódico e caótico. Quando δ = 0,0575 o sistema perde a estabilidade e ocorre uma bifurcação do tipo duplicação de período. Em δ = 0,05875 acontece uma nova bifurcação. Estas bifurcações se repetem, sucessivamente, formando o que se chama de cascata de duplicação de períodos que aparece como uma sequência infinita de duplicação de períodos até o aparecimento do caos (Alligood *et al.*, 1997).

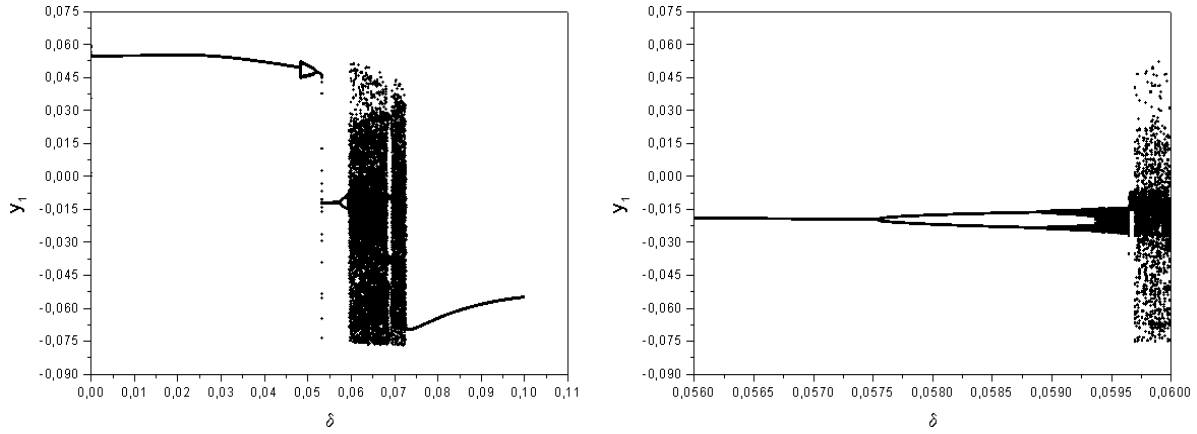


Figura 3. Diagrama de bifurcação em δ. (a) $0 \le \delta \le 0{,}5$; (b) $0{,}056 \le \delta \le 0{,}06$.

Uma nova ampliação destes diagramas no intervalo $0{,}056 \le \delta \le 0{,}060$ está mostrado na Figura 3b, permitindo que se observe a transição do comportamento periódico ao caótico. Quando $\delta = 0{,}0593$, ocorre um súbito aparecimento de um atrator caótico, associado a existência de um expoente de Lyapunov positivo. Este fenômeno é conhecido como *crise*. Grebogi *et al.* (1983) definem crise como sendo uma colisão entre um atrator caótico e uma órbita ou um ponto fixo instável. A crise que acontece no oscilador com memória de forma em $\delta = 0{,}0593$ é conhecida como *crise de contorno*, estando relacionada ao surgimento do atrator caótico e de sua bacia de atração.

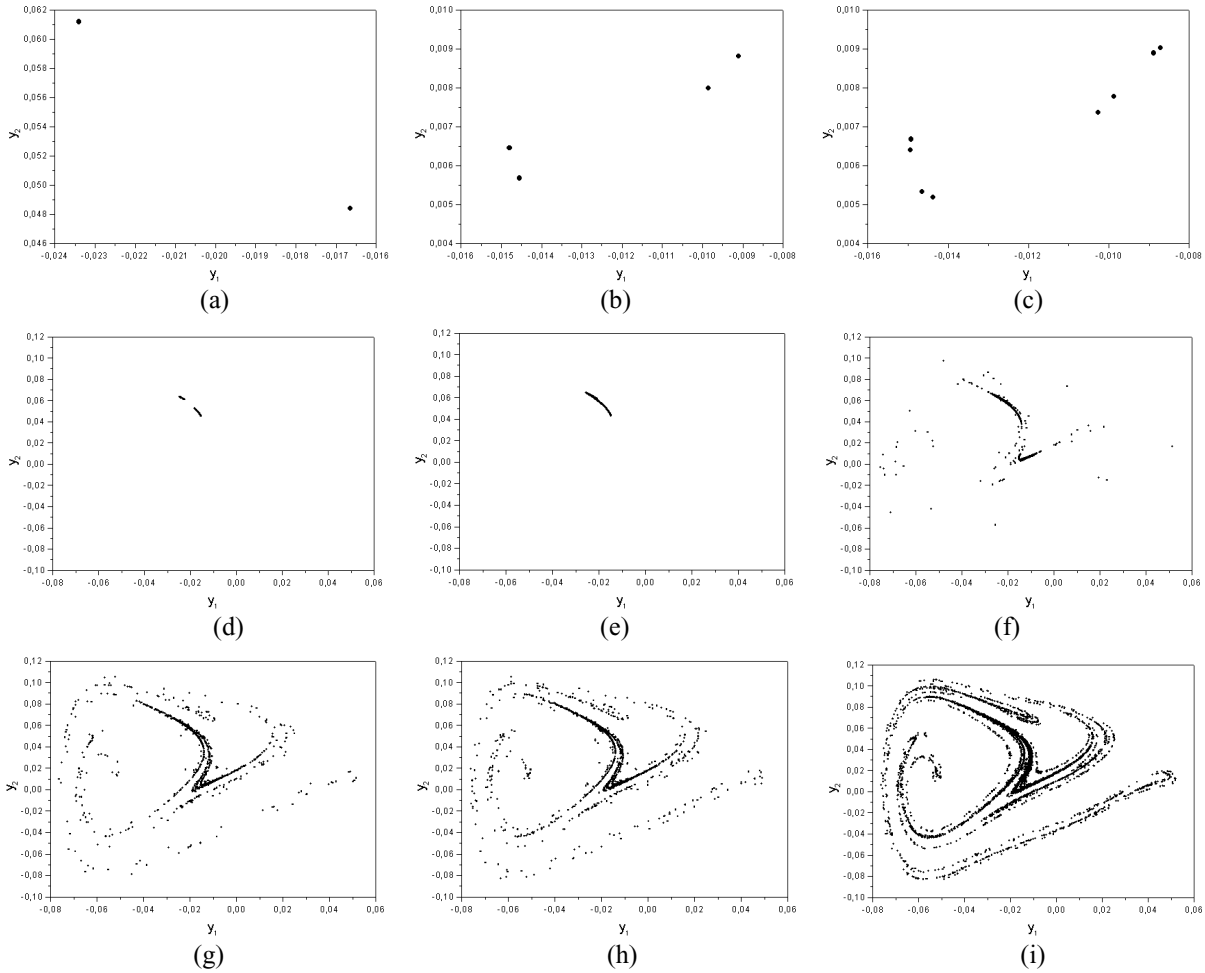


Figura 4. Surgimento e formação do atrator caótico. (a) $\delta = 0{,}05875$, período-2; (b) $\delta = 0{,}059$, período-4; (c) $\delta = 0{,}0592$, período-8 ; (d) $\delta = 0{,}0593$, caos ; (e) $\delta = 0{,}0595$, caos ; (f) $\delta = 0{,}05965$, caos ; (g) $\delta = 0{,}059675$, caos ; (h) $\delta = 0{,}05969$, caos ; (i) $\delta = 0{,}06$, caos.

A Figura 4 mostra seções de Poincaré relacionadas a diferentes valores do parâmetro δ. As Figuras 4a, 4b, 4c estão relacionadas ao comportamento periódico de período 2, 4 e 8, respectivamente. A partir da Figura 4d, quando $\delta = 0{,}0593$, tem-se comportamentos caóticos, identificados pela existência de um expoente de Lyapunov positivo. Na Figura 4d, observa-se o instante do surgimento do atrator, enquanto as Figuras 4e, 4f, 4g, 4h, 4i mostram a formação do atrator. Este processo de formação do atrator é definido como uma *crise interior* do sistema, estando relacionada com mudanças bruscas no tamanho ou forma do atrator. A Figura 5 mostra os espaços de fase referentes aos comportamentos apresentados na Figura 4.

Uma ampliação do atrator estranho da Figura 4d, momento do surgimento do caos, está mostrado na Figura 6a. Os expoentes de Lyapunov para os diferentes valores do parâmetro δ, são apresentados na Figura 6b, mostrando o instante em que o movimento passa a ser caótico

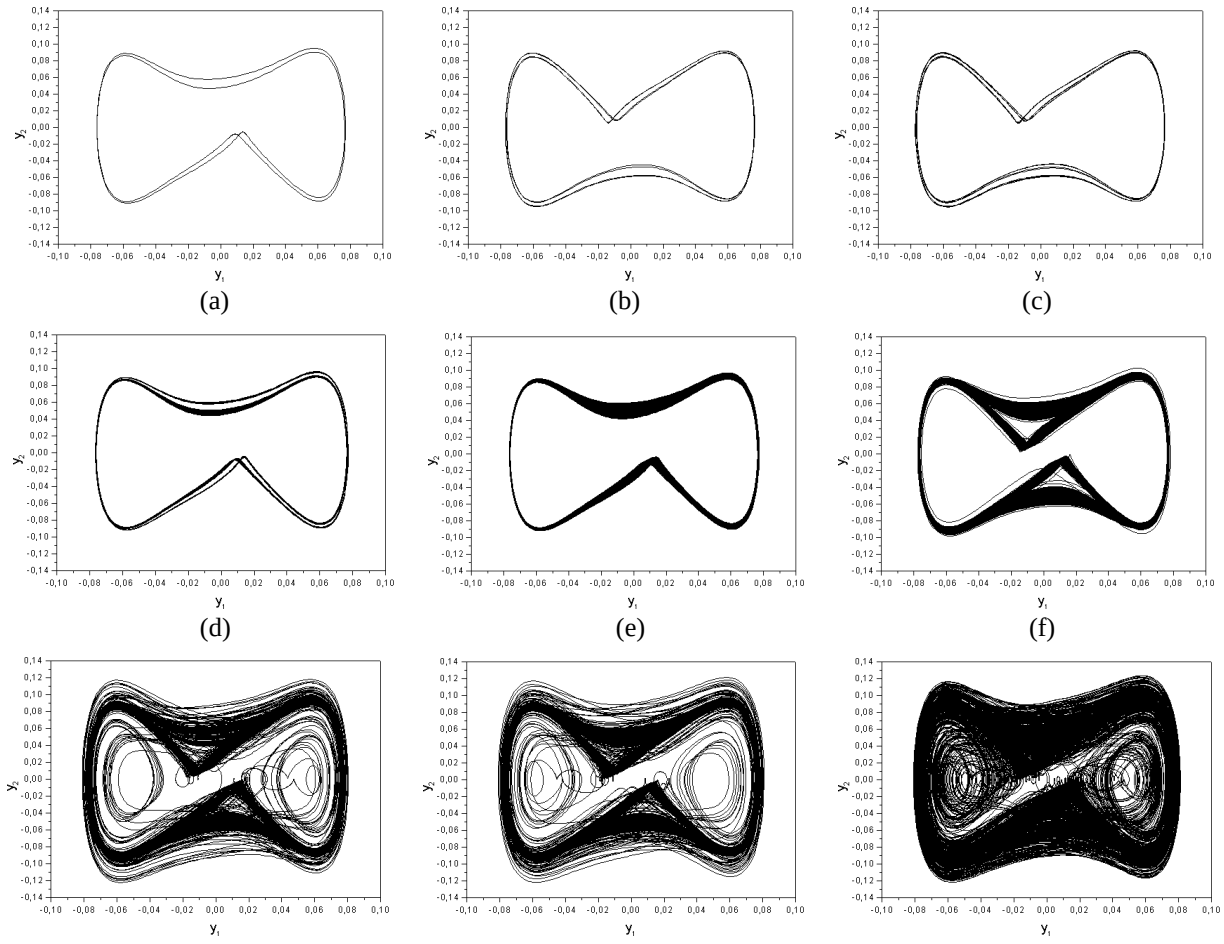


Figura 5. Espaços de fase da formação do atrator caótico. (a) δ = 0,05875, período-2; (b) δ = 0,059, período-4; (c) δ = 0,0592, período-8 ; (d) δ = 0,0593, caos ; (e) δ = 0,0595, caos ; (f) δ = 0,05965, caos ; (g) $\delta = 0{,}059675$, caos ; (h) $\delta = 0{,}05969$, caos ; (i) $\delta = 0{,}06$, caos.

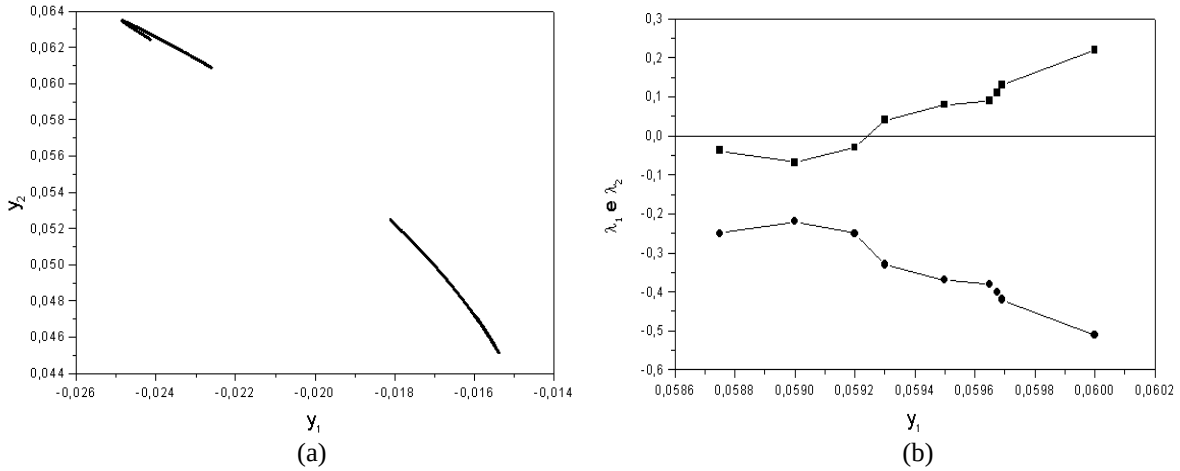


Figura 6. (a) Atrator caótico para $\delta = 0{,}0593$. (b) Expoentes de Lyapunov para diferentes valores de δ.

Observando o intervalo $0{,}068 \leq \delta \leq 0{,}070$, identifica-se a presença de uma janela periódica (Figura 7). A formação desta janela periódica está associada a um fenômeno conhecido como *subduction*. Grebogi *et al.* (1983) definem *subduction* como sendo o aparecimento de um atrator não-caótico dentro de um atrator caótico, sendo este último substituído pelo primeiro. A diferença entre *crise de contorno* e *subduction* é que a última não destrói a bacia de atração ao contrário da primeira. Nesta janela periódica, em particular, observa-se a presença de um comportamento periódico de período-5 na sua região central, se contrapondo ao caos nas suas bordas. Uma evolução das seções de Poincaré para diferentes valores do parâmetro δ contidos nesta janela estão mostrados na Figura 8, enquanto a Figura 9 mostra os espaços de fase para estes mesmos valores. A evolução dos expoentes de Lyapunov está mostrado na Figura 10.

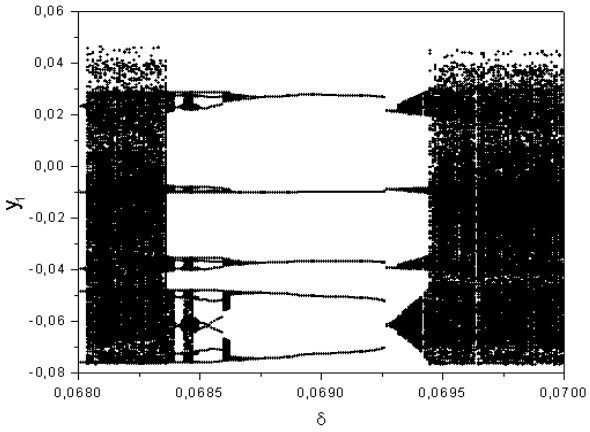


Figura 7. Janela periódica no intervalo $0{,}068 \leq \delta \leq 0{,}070$.

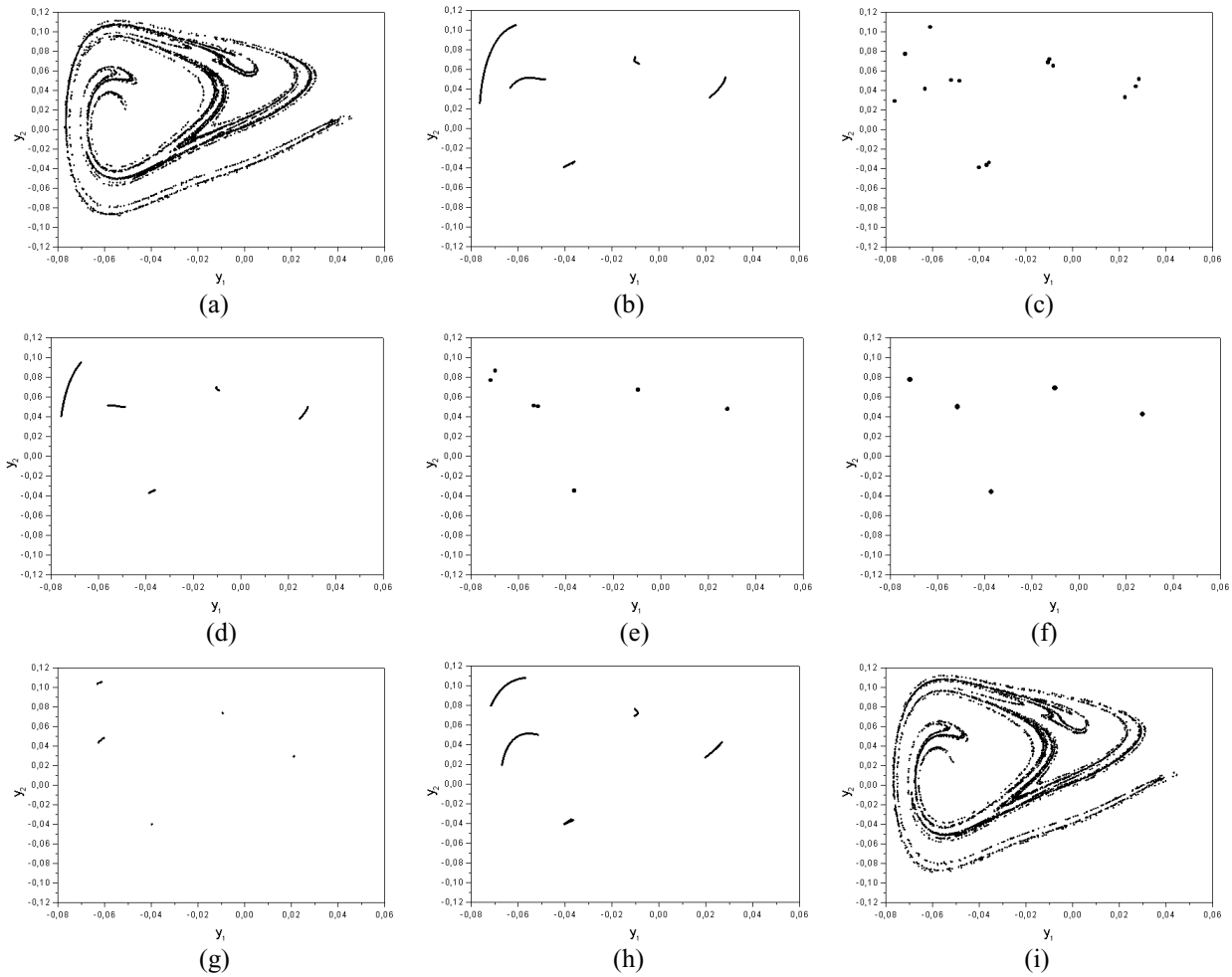


Figura 8. Surgimento e formação do atrator caótico. (a) $\delta = 0{,}0683$, caos; (b) $\delta = 0{,}06845$, caos ; (c) $\delta = 0{,}0685$, caos ; (d) $\delta = 0{,}0686$, caos ; (e) $\delta = 0{,}0687$, caos ; (f) $\delta = 0{,}0692$, período-5 ; (g) $\delta = 0{,}0693$, caos; (h) $\delta = 0{,}0694$, caos; (i) $\delta = 0{,}0695$, caos.

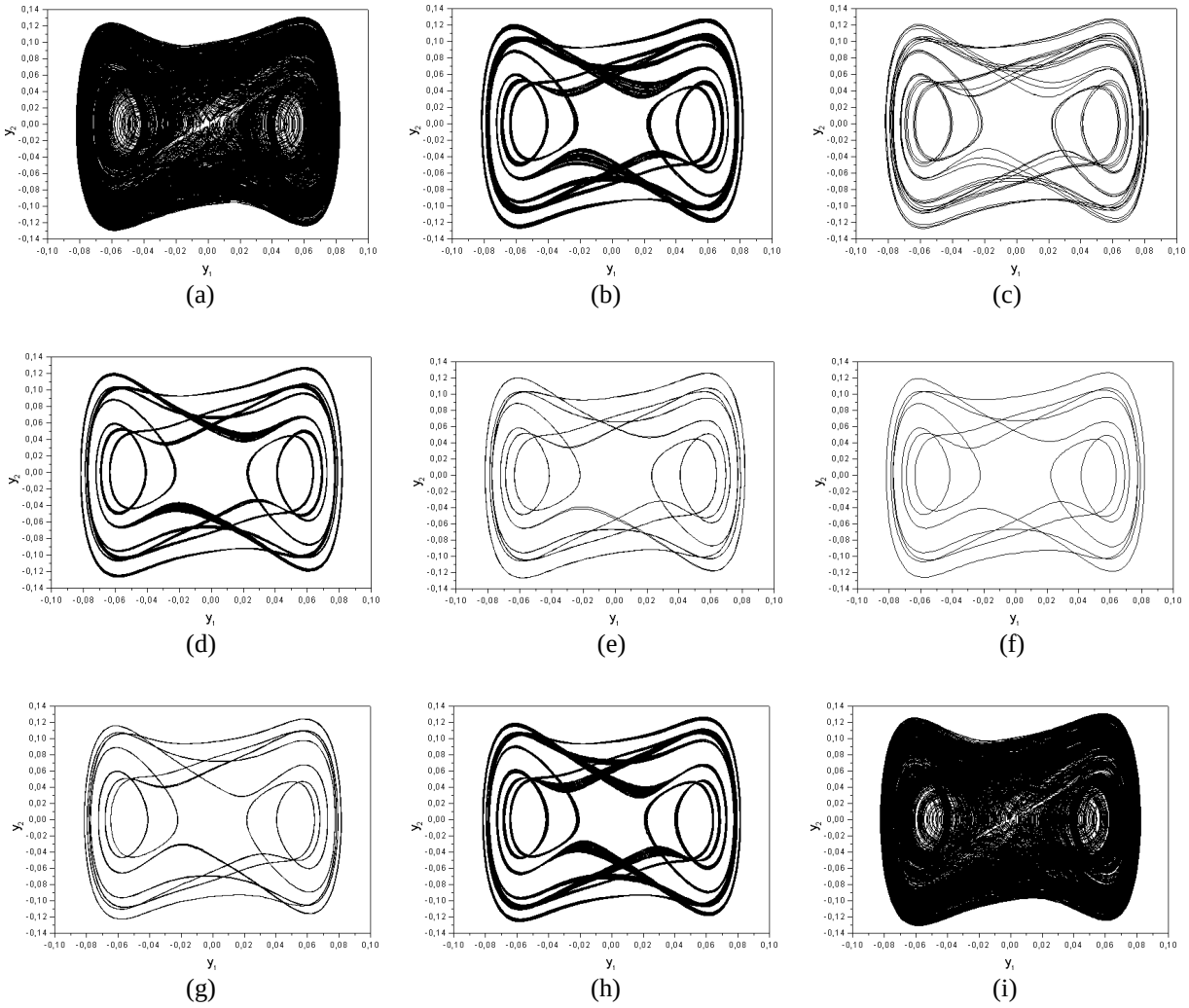


Figura 9. Espaços de fase na formação do atrator caótico. (a) δ = 0,0683, caos; (b) δ = 0,06845, caos ; (c) δ = 0,0685, caos ; (d) δ = 0,0686, caos ; (e) δ = 0,0687, caos ; (f) δ = 0,0692, período-5 ; (g) δ = 0,0693 , caos; (h) δ = 0,0694, caos, (i) δ = 0,0695, caos.

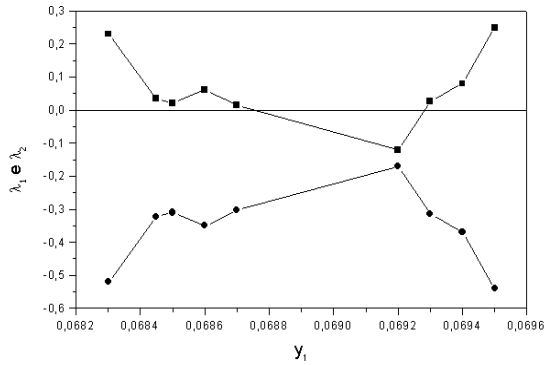
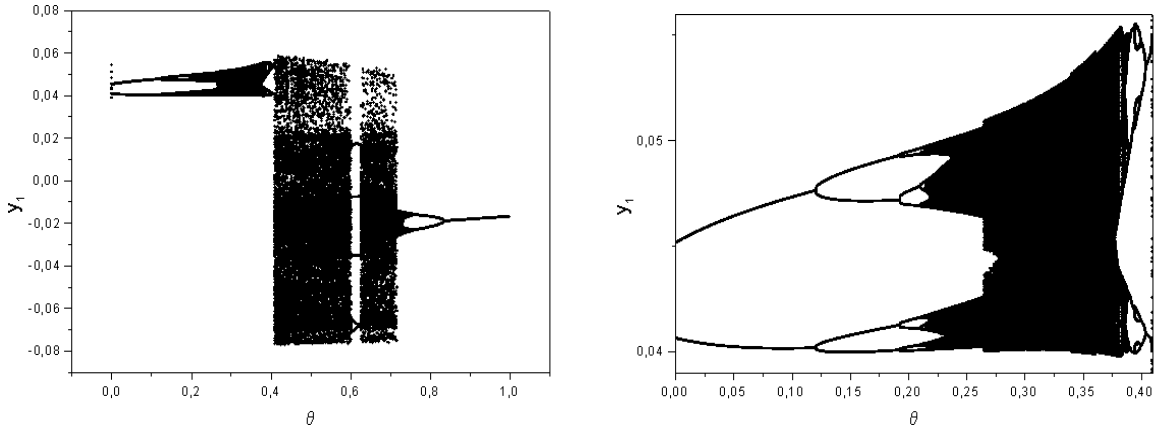


Figura 10. Expoentes de Lyapunov para diferentes valores de δ, das Fig.(10) e Fig.(11).

4.2. Bifurcações Associadas ao Parâmetro de Temperatura θ

Neste momento, são discutidas as bifurcações associadas às variações do parâmetro de temperatura θ. A Figura 11a apresenta o diagrama de bifurcação do parâmetro θ no intervalo $0 \leq \theta \leq 1$, para δ = 0,06. Uma grande variedade de respostas pode ser identificada. Em contraste com os diagramas relacionados com o parâmetro δ, as bifurcações não apresentam regiões de comportamento caótico separadas por regiões periódicas. Além disso, para $\theta > 1$, o sistema não mais apresenta comportamento caótico.

A Figura 11b mostra uma ampliação do diagrama no intervalo $0 \leq \theta \leq 0{,}405$. Em $\theta = 0$, surgem duas órbitas periódicas. Estas órbitas permanecem estáveis até $\theta = 0{,}12$. Neste ponto, o sistema perde a estabilidade, e cada uma das órbitas apresenta uma duplicação de período. Depois disso, as quatro órbitas periódicas sofrem outra perda de estabilidade e experimentam nova duplicação de período. Esta sequência se repete, sucessivamente, gerando uma cascata de duplicação de períodos.

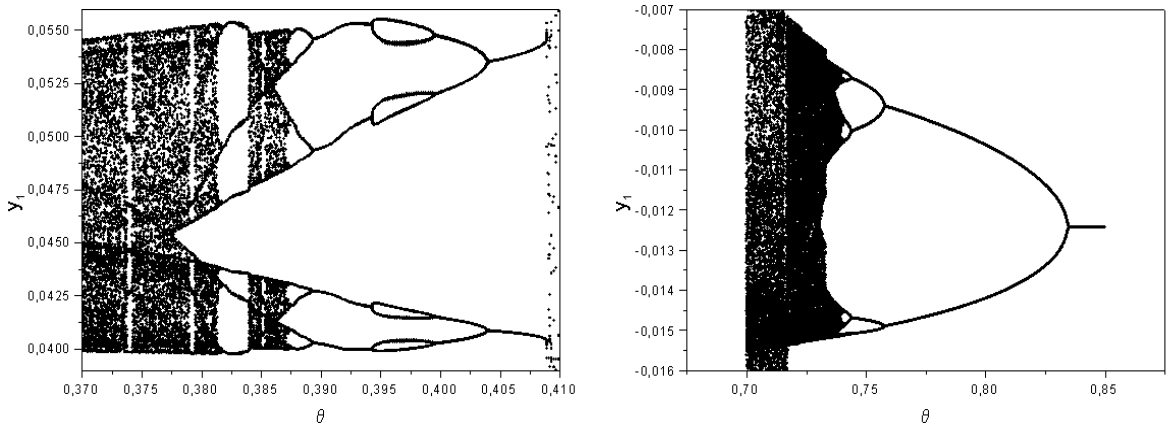


(a) (b)

Figura 11. Diagrama de bifurcação em θ. (a) $0 \leq \theta \leq 1$; (b) $0 \leq \theta \leq 0{,}405$.

Ampliando a Figura 11b no intervalo $0{,}37 \leq \theta \leq 0{,}41$ (Figura 12a), nota-se uma sucessão de janelas periódicas dentro de regiões com comportamento caótico até o valor de $\theta = 0{,}3875$. Depois, quando $0{,}3875 \leq \theta \leq 0{,}409$, vêem-se comportamentos periódicos, onde existem bifurcações do tipo duplicação de período, direta e reversa. Quando $\theta = 0{,}409$, surge uma crise no sistema e, novamente, tem-se a presença de caos. Considerando o intervalo $0{,}70 \leq \theta \leq 0{,}85$, mostrado na Figura 12b, visualiza-se o instante em que o caos dá lugar a uma cascata de duplicação de período reversa, em $\theta = 0{,}717$. Esta cascata reversa evolui até o comportamento periódico apresentar período-1 ($\theta = 0{,}85$). A partir daí, não se observa mais a presença de caos. Vale ressaltar que o aparecimento desta cascata reversa também está associada a uma crise do sistema. A evolução dos atratores associados a esta crise são semelhantes a apresentada no caso das bifurcações do parâmetro δ.

A Figura 13 mostra a evolução de seções de Poincaré para diferentes valores de θ, no intervalo $0{,}37 \leq \theta \leq 0{,}41$ (Figura 12a). A Figura 14, por sua vez, apresenta os espaços de fase para os mesmos valores de θ.



(a) (b)

Figura 12. Diagrama de bifurcação θ. (a) $0{,}37 \leq \theta \leq 0{,}41$; (b) $0{,}70 \leq \theta \leq 0{,}85$.

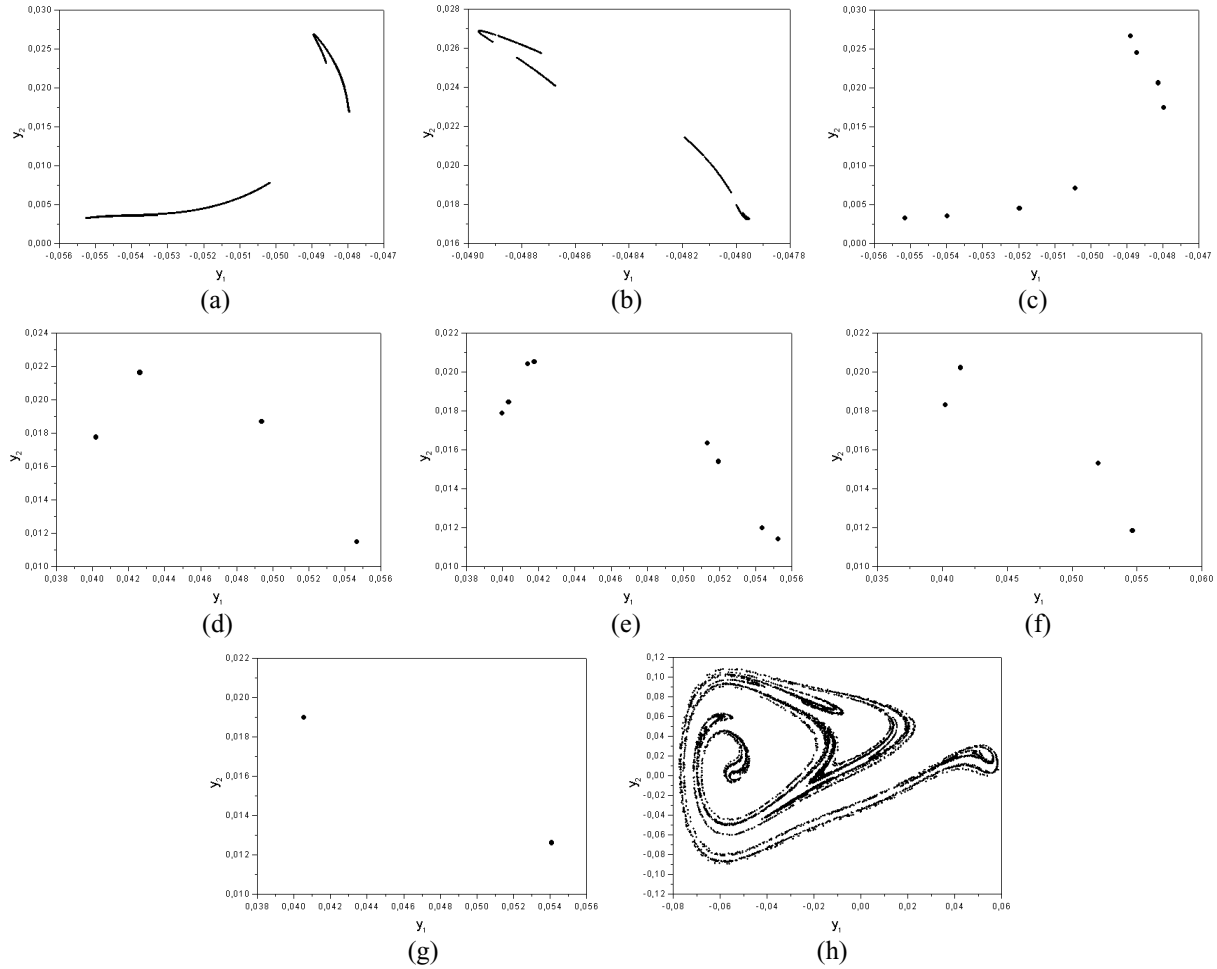


Figura 13. Surgimento e formação do atrator caótico. (a) $\theta = 0{,}386$, caos; (b) $\theta = 0{,}387$, caos; (c) $\theta = 0{,}3875$, período-8; (d) $\theta = 0{,}389$, período-4; (e) $\theta = 0{,}38975$, período-8; (f) $\theta = 0{,}4$, período-4; (g) $\theta = 0{,}408$, período-2; (h) $\theta = 0{,}409$, caos.

5. Conclusões

Este trabalho analisa alguns aspectos associados às bifurcações em um oscilador com memória de forma onde a força de restituição é descrita a partir de um modelo constitutivo polinomial. Neste trabalho, estudam-se variações de dois parâmetros: amplitude de forçamento e temperatura. Os resultados mostram cascatas de duplicação de período, direta e reversa, e crises na resposta do sistema. O sistema apresenta uma grande variedade de respostas em baixas temperaturas, onde a fase martensítica é estável, e em temperaturas intermediárias, onde tanto a martensita quanto a austenita são estáveis. Por outro lado, o sistema não apresenta comportamento caótico em altas temperaturas, onde somente a fase austenítica é estável.

6. Referências

Alligood, K.T., Sauer, T.D. & Yorke, J.A., 1997, "*Chaos: An Introduction to Dynamical Systems*", Springer-Verlag.

Falk, F., 1980, "Model Free Energy, Mechanics, and Thermodynamics of Shape Memory Alloys", *Acta Metalurica*, v.28, pp. 1773-1780.

Grebogi, C., Ott, E. & Yorke, J.A., 1983, "Crises, Sudden Changes in Chaotic Attractors, and Transient Chaos", *Physica D*, v.7, pp. 181-200.

Guckenheimer, J. & Holmes, P., 1983, "*Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*", Springer-Verlag.

Parker, T.S. & Chua, L.O., 1989, "*Practical Numerical Algorithm for Chaotic Systems*", Springer-Verlag.

Rogers, C.A., Liang, C. & Fuller, C.R., 1991, "Modeling of Shape Memory Alloy Hybrid Composites for Structural Acoustic Control", *Journal of Acoustic Society of America*, n.89, pp.210-220.

Savi, M.A. & Braga, A.M.B., 1993a, "Chaotic Vibration of an Oscillator with Shape Memory", *Journal of the Brazilian Society of Mechanical Science*, v.15, n.1, pp.1-20.

Savi, M.A. & Braga, A.M.B., 1993b, "Chaotic Response of a Shape Memory Oscillator with Internal Constraints", *COBEM 93 - XII Brazilian Congress of Mechanical Engineering*, Brasil, 7-10 December 1993, v.I, pp.33-36.

Savi, M.A. & Pacheco, P.M.C.L., 2002, "Chaos and Hyperchaos in Shape Memory Systems", *International Journal of Bifurcation and Chaos*, to appear.

Wasilevski, R.J., 1975, "On the Nature of the Martensitic Transformation", *Metalurgical Transactions*, v.6A, pp.1405-1418.

Wolf, A., Swift, J.B., Swinney, H.L. & Vastano, J.A., 1985, "Determining Lyapunov Expoents from a Times Series", *Physica D*, v.16, pp. 285-317.

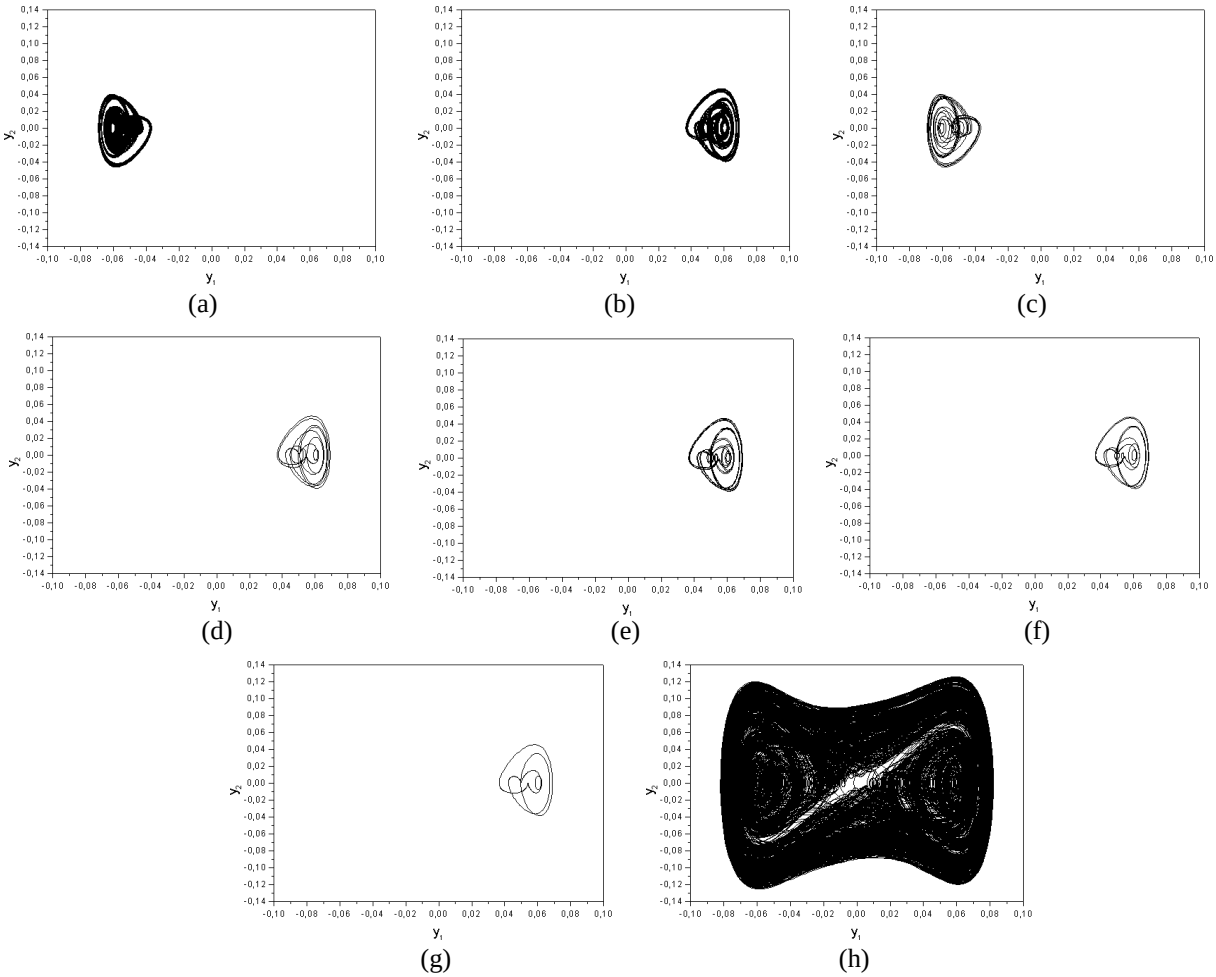


Figura 14. Espaços de fase na formação do atrator caótico. (a) $\theta = 0{,}386$, caos; (b) $\theta = 0{,}387$, caos; (c) $\theta = 0{,}3875$, período-8; (d) $\theta = 0{,}389$, período-4; (e) $\theta = 0{,}38975$, período-8; (f) $\theta = 0{,}4$, período-4; (g) $\theta = 0{,}408$, período-2; (h) $\theta = 0{,}409$, caos.

BIFURCATIONS AND CRISES IN A SHAPE MEMORY OSCILLATOR

Luciano G. Machado
Marcelo A. Savi
Instituto Militar de Engenharia - Department of Mechanical and Materials Engineering
22.290.270 – Rio de Janeiro - RJ

Pedro M. C. L. Pacheco
CEFET/RJ - Department of Mechanical Engineering
20.271.110 - Rio de Janeiro – RJ

Abstract. *The dynamical response of systems with shape memory actuators presents nonlinear characteristics and a very rich behavior. Chaotic response is an example of this kind of behavior. This contribution analyzes some aspects related to bifurcation phenomenon in a shape memory oscillator where the restitution force is described by a polynomial constitutive model. The term bifurcation is used to describe qualitative changes that occur in the orbit structure of a dynamical system as a consequence of parameter variations. Bifurcation phenomenon is closed related to chaos. Results of the analysis show that the response of the shape memory oscillator presents period doubling cascades, direct and reverse, and crises.*

Keywords: *Shape memory, chaos, bifurcation, crises.*

XVI CONGRESSO BRASILEIRO DE ENGENHARIA MECÂNICA
16th BRAZILIAN CONGRESS OF MECHANICAL ENGINEERING

PREDIÇÃO EM SÉRIES TEMPORAIS A PARTIR DE UMA MODELAGEM POLINOMIAL: APLICAÇÃO AO PÊNDULO NÃO-LINEAR

Eduardo Gomes Ferreira Pinto
Marcelo Amorim Savi
Instituto Militar de Engenharia
Departamento de Engenharia Mecânica e de Materiais
22290-270 – Rio de Janeiro – RJ – Brasil
E-Mail: savi@epq.ime.eb.br

Resumo. *A predição em séries temporais possui aplicações em várias áreas do conhecimento. Este trabalho tem como objetivo apresentar uma discussão sobre a predição em séries temporais não-lineares a partir de uma modelagem polinomial, avaliando a influência do tamanho da série e a sensibilidade à contaminação por ruído. Como aplicação, considera-se um pêndulo não-linear. Os sinais são obtidos através da integração numérica do modelo matemático, elegendo-se uma variável de estado como uma série temporal característica do sistema. A sensibilidade ao ruído é investigada a partir da contaminação do sinal por um ruído aleatório. Dois tipos de sinais são analisados: periódico e caótico. Os resultados mostram as dificuldades de se efetuar predições em séries temporais não-lineares.*

Palavras chave: *Predição, Séries temporais, Caos, Pêndulo não-linear.*

1. Introdução

A ligação mais direta entre a teoria do caos e o mundo real é a análise de séries temporais, nos termos da dinâmica não-linear (Schreiber, 1999). Uma série temporal é um conjunto discreto e numerável de valores de uma variável de estado de um sistema dinâmico. A predição em séries temporais tem como objetivo, a partir de uma série temporal escalar S_n (n = 1, …, N), denominada como *passado*, determinar uma estimativa para os seus valores seguintes, S_{N+1}, S_{N+2}, …, S_{N+p}, denominado *futuro*. Assim, os termos conhecidos da série são utilizados para estimar valores dos termos seguintes através de uma técnica de predição. Desta forma, estima-se uma série relacionada ao futuro: P_{N+1}, P_{N+2}, …, P_{N+p}. Este procedimento pode ser entendido como uma extrapolação.

Várias áreas do conhecimento têm interesse no estudo da predição em séries temporais dentre elas a engenharia (Miyano *et al.*, 2000; Cao *et al.*, 1997; Sakai & Aihara, 1994), a economia (Hsieh, 1991; Leung *et al.*, 2000), o marketing (Kuo, 2001), a meteorologia (Sfetsos, 2000; Toth *et al.*, 2000) e a medicina (Brandt *et al.*, 2000). Abarbanel *et al.* (1993), Lillekjendlie *et al.* (1994), Weigend & Gershenfeld (1994) Kantz & Schreiber (1997) e Schreiber (1999) fazem uma revisão sobre a análise de séries temporais, apresentando as principais técnicas de predição.

De uma maneira geral, a predição consiste em ajustar um modelo aos dados de uma série. Normalmente, este ajuste é feito através da minimização do erro na predição, através do método dos mínimos quadrados, determinando-se os coeficientes da técnica de predição. A Figura 1 mostra um esquema sobre o problema da predição. Considere uma série temporal conhecida (passado), S_n (cor azul). A partir dessa série, realiza-se uma predição dos termos seguintes (futuro), P_n (vermelho). Os termos da série prevista (vermelho) podem ser comparados com os da série original (verde), avaliando a precisão da técnica de predição.

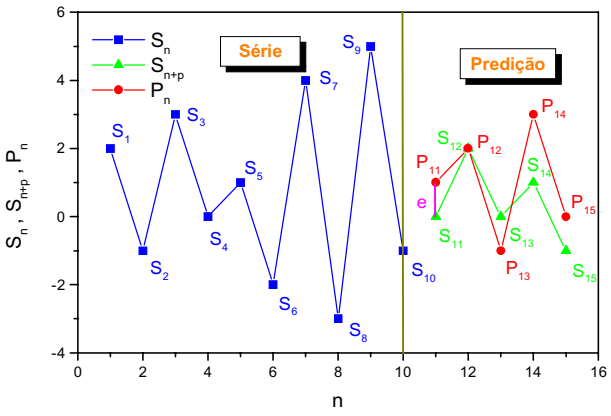


Figura 1. Predição em uma série temporal.

Salvino *et al.* (1995) definem a previsibilidade de uma série como uma grandeza que indica o quão extenso o passado deve ser utilizado para se determinar o futuro. Esta informação é importante para se estimar a confiabilidade e a viabilidade de uma predição, e é um parâmetro para comparar se um sinal é mais ou menos previsível do que outro. Os expoentes de Lyapunov fornecem uma indicação de quão distante de um determinado instante de tempo se pode realizar

predições confiáveis. A dimensão de informação, por outro lado, dá uma idéia de quão complexo o modelo deve ser (Casdagli, 1989).

As técnicas de predição podem ser classificadas em *lineares* ou *não-lineares*. As técnicas lineares são aquelas em que o modelo de predição satisfaz às condições de linearidade. As técnicas não-lineares, por sua vez, são as não incluídas na categoria das lineares. Outra classificação encontrada na literatura é a de *técnicas locais* ou *globais*. Por definição, técnicas locais são aquelas que variam ponto a ponto (ou vizinhança a vizinhança) no espaço de fase. As técnicas globais, por outro lado, são construídas em um passo e para todos os pontos em todo o espaço de fase (Abarbanel *et al.*, 1993). As técnicas denominadas *semi-locais* são técnicas globais com boas propriedades locais (Casdagli, 1989; Abarbanel *et al.*, 1993; Lillekjendlie *et al.*, 1994).

Dentre as técnicas de predição lineares mais utilizadas na literatura vale destacar: linear de ordem zero, média flutuante, auto-regressão e a média flutuante auto-regressiva (Weigend & Gershenfeld, 1994; Kantz & Schreiber, 1997; Schreiber, 1999). Com relação as técnicas não-lineares, vê-se que estas tem como primeiro passo realizar a reconstrução do espaço de estado (Farmer & Sidorowich, 1987; Casdagli, 1989). A partir daí pode-se, por exemplo, realizar uma predição não-linear de ordem zero (Farmer & Sidorowich, 1987; Abarbanel *et al.*, 1993; Kantz & Schreiber, 1997). Algumas alternativas mais sofisticadas incluem a predição polinomial e a predição por funções de base radial (Farmer & Sidorowich, 1987; Casdagli, 1989; Abarbanel *et al.*, 1993; Kantz & Schreiber, 1997; Hegger *et al.*, 1999; Schreiber, 1999). Para melhorar o desempenho das técnicas de predição não-lineares, surgem as estruturas capazes de decidir, que são as predições por redes neurais e por redes *wavelet* (Cao *et al.*, 1995), que combinam a decomposição *wavelet* e a estrutura das redes neurais.

O objetivo deste trabalho é apresentar uma discussão sobre predição em séries temporais não-lineares a partir de uma modelagem polinomial, avaliando a influência do tamanho da série e a sensibilidade à contaminação de ruído. Como aplicação, considera-se um pêndulo não-linear. As séries temporais são obtidas através da integração numérica do modelo matemático, elegendo-se uma variável de estado como um sinal característico do sistema. A sensibilidade ao ruído é investigada a partir da contaminação do sinal por um ruído aleatório. Dois tipos de sinais são analisados: periódico e caótico. Os resultados mostram as dificuldades de se efetuar predições em séries temporais não-lineares.

2. Predição a Partir de uma Modelagem por Funções Conhecidas

Considere uma série escalar S_n (n = 1, …, N), a partir da qual deseja-se estimar os seus valores seguintes, S_{N+1}, S_{N+2}, …, S_{N+p}. Assim, utilizam-se m termos conhecidos da série, passado, para estimar p valores, futuro, através de uma técnica de predição: P_{N+1}, P_{N+2}, …, P_{N+p}. Um procedimento geral é ajustar funções genéricas $\Phi_j\,(j=1,...,\alpha)$ aos m últimos pontos da série, o que permite avaliar a predição da seguinte forma,

$$P_{n+1} = \sum_{j=1}^{\alpha} a_j \Phi_j(S_n) = a_1 \Phi_1(S_n) + a_2 \Phi_2(S_n) + ... + a_\alpha \Phi_\alpha(S_n) \tag{1}$$

Os parâmetros $a_j\,(j=1,...,\alpha)$ devem ser obtidos de forma a minimizar um erro associado à predição. A literatura reporta uma série de possibilidades para se avaliar o erro envolvido na predição (Kantz &Schreiber, 1997). Uma das alternativas disponíveis é utilizar um erro *RMS* (*root mean squared*), e_{RMS}, definido a seguir:

$$e_{RMS} = \frac{1}{\sigma}\sqrt{\frac{1}{m}\sum_{n=N-m+1}^{N}(S_{n+1}-P_{n+1})^2} \tag{2}$$

onde σ é o desvio padrão da amostra. A minimização deste erro, implica que:

$$\frac{\partial e_{RMS}}{\partial a_i} = \frac{\partial}{\partial a_i}\left\{\frac{1}{\sigma}\sqrt{\frac{1}{m}\sum_{n=N-m+1}^{N}\left[S_{n+1}-\sum_{j=1}^{\alpha} a_j \Phi_j\right]^2}\right\} = 0 \quad (i=1,...,\alpha) \tag{3}$$

o que resulta no seguinte sistema,

$$\sum_{j=1}^{\alpha} C_{ij}\, a_j = \sum_{n=N-m+1}^{N} S_{n+1}\, \Phi_i \quad (i=1,...,\alpha) \tag{4}$$

onde C_{ij} é a *matriz de covariância* definida a seguir,

$$C_{ij} = \sum_{n=N-m+1}^{N} \Phi_i \, \Phi_j \quad (i, j = 1, ..., \alpha) \tag{5}$$

Desta forma, a função de aproximação é definida a partir da solução do sistema linear (4), o que torna possível efetuar a predição do termo P_{n+1}. As predições subsequentes podem ser feitas a partir de um procedimento recursivo.

Neste trabalho utilizam-se funções polinomiais para modelar o comportamento da série e fazer a predição. As predições são aferidas através do cálculo do *erro na amostra* ("*in-sample error*") ou pelo *erro na predição*, que são obtidos através da Eq. 2. No cálculo do erro na amostra, o valor de P_{n+1} (Eq. 1) é obtido nos termos da série que são utilizados para fazer o ajuste do preditor. No erro na predição, por sua vez, são utilizados os termos da série temporal que não foram inseridos no cálculo do primeiro erro.

As predições aqui tratadas podem ser classificadas como locais no sentido de que não consideram todo o espaço de estado, avaliando apenas a série escalar. As simulações são realizadas a partir do algoritmo de Hegger *et al.* (1999).

3. Pêndulo Não-linear

Considere um pêndulo não-linear onde θ define sua posição, ξ é um parâmetro relacionado ao coeficiente de amortecimento viscoso linear e ω_n está associado à freqüência natural do sistema. Admite-se um forçamento harmônico sendo μ a amplitude e Ω a freqüência de forçamento. A equação de movimento do sistema está apresentada a seguir.

$$\ddot{\theta} + \xi \dot{\theta} + \omega_n^2 \operatorname{sen}(\theta) = \mu \cos(\Omega t) \tag{3}$$

Esta equação pode ser reescrita como um sistema de equações de primeira ordem, $\dot{u} = f(u,t), u \in R^3$, onde $u_1 = x = \theta$, $u_2 = y = \dot{\theta}$, $u_3 = z = \Omega t$. Neste trabalho, os sinais são gerados a partir de simulações numéricas realizadas através do método Runge-Kutta de quarta ordem com passos inferiores a $\Delta t = 2\pi / 100$. Para isso, a variável de estado $x = \theta$ é escolhida como um sinal característico do pêndulo. Além disso, o sinal é contaminando com um ruído aleatório, $\eta = AR(-1, +1)$, onde A é a amplitude do ruído, e R (−1, +1) são números aleatórios que podem variar entre −1 e +1. Assim, o sinal a ser analisado é $S = x + \eta$. Note que, se η = 0, tem-se um sinal ideal, sem ruído. Desta maneira, é possível avaliar a influência do ruído na análise não-linear de séries temporais. Consideram-se, além do sinal ideal, A = 0, dois outros níveis de ruído: A = 0,314 e A = 0,628, representando respectivamente, 5% e 10% da máxima amplitude do sinal.

Investigam-se dois tipos de sinais: periódico de período-2 (μ = 2,56) e caótico (μ = 2,50). Os demais parâmetros são ξ = 0,2 e $\omega_n = \Omega$ = 1,0. A metodologia utilizada consiste em analisar diferentes graus do polinômio e diferentes tamanhos da série avaliando a influência da contaminação por ruído, comparando as predições aos valores da série no futuro.

3.1. Sinal Periódico

Considere um sinal de período-2 com N = 30000, obtido através de uma simulação numérica, a partir do qual deseja-se fazer a predição de valores futuros da série. A Figura 2 apresenta um trecho deste sinal.

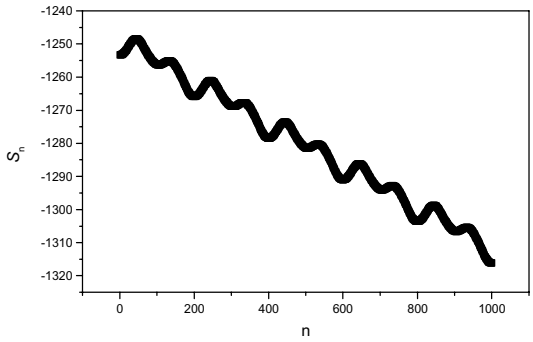
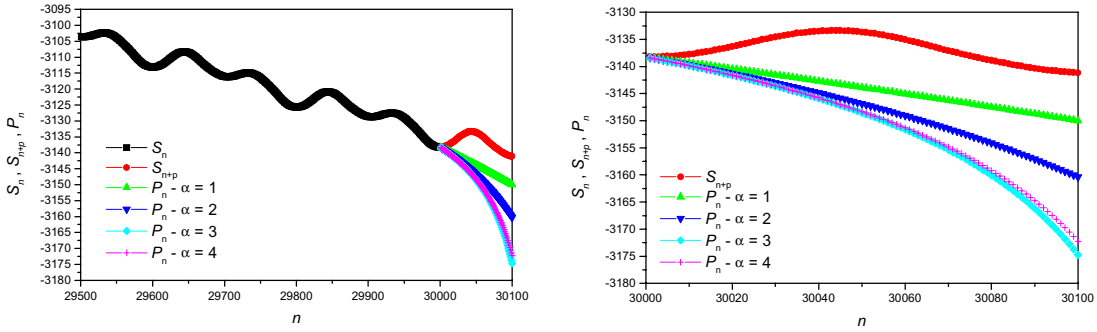


Figura 2. Série periódica.

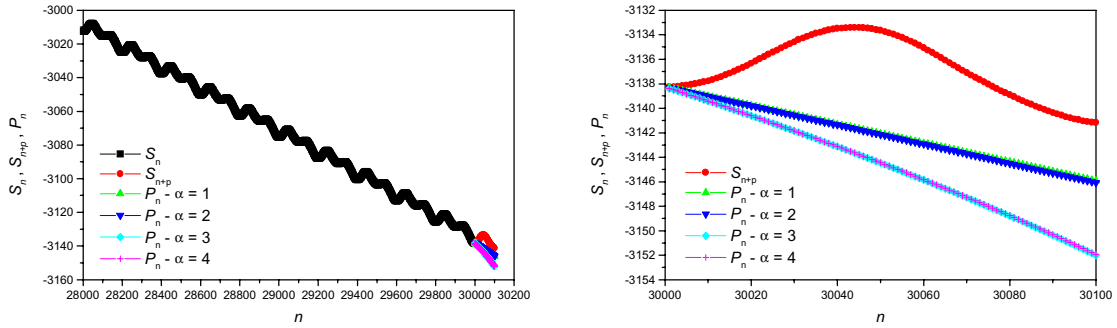
Inicialmente, considere a predição em séries temporais efetuada a partir de diferentes valores para o número de termos da série, m. Esta análise é realizada para diferentes ordens do polinômio de aproximação, α. A Figura 3a mostra a série temporal, o seu prolongamento e as diferentes predições efetuadas a partir de polinômios de ordem 1, 2, 3 e 4. A Figura 3b apresenta a região de predição mostrando a comparação entre o prolongamento da série e as predições. Estas

predições consideram m = 500, prevendo p = 100 pontos no futuro. A Figura 4, por sua vez considera m = 2000 e p = 100.



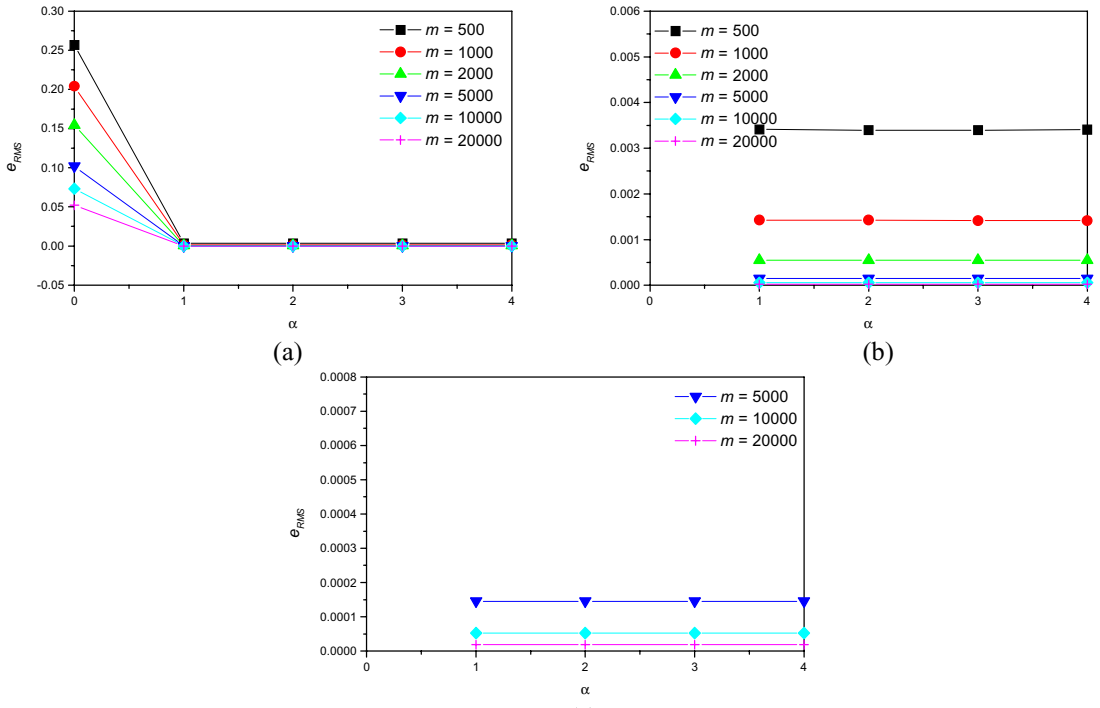
(a) (b)

Figura 3. Predição em série periódica, A = 0 (sem ruído), m = 500 e p = 100.



(a) (b)

Figura 4. Predição em série periódica, A = 0 (sem ruído), m = 2000 e p = 100.



(a) (b)

(c)

Figura 5. Erro na amostra em função do grau do polinômio e do número de termos da série na predição de um sinal periódico (A = 0, sem ruído, e p = 100): (a) α = 0, 1, 2, 3, 4; m = 500 a 20000; (b) α = 1, 2, 3, 4; m = 500 a 20000; (c) α = 1, 2, 3, 4; m = 5000 a 20000.

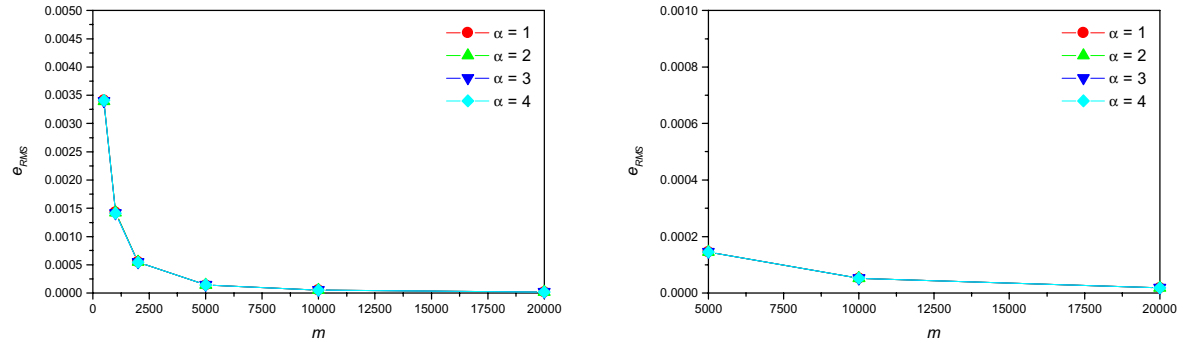


Figura 6. Erro na amostra em função do grau do polinômio e do número de termos da série na predição de um sinal periódico (A = 0, sem ruído, e p = 100). (a) m = 500 a 20000; (b) m = 5000 a 20000.

Uma análise do erro na amostra está em questão agora. A Figura 5 apresenta o erro na amostra para diferentes tamanhos da série temporal e ordens do polinômio. Como esperado, o aumento do tamanho da série tende a diminuir o erro na predição. Com relação a ordem do polinômio, vê-se que, embora o erro na amostra não se altere para diferentes valores de α, as predições são muito alteradas (Figuras 3-4). Ordens mais altas tendem a ser instáveis, causando uma divergência com relação à série original no futuro. Esta divergência, contudo, é atenuada na medida em que se aumenta o número de termos da série. Outra maneira de observar estes resultados está mostrada na Figura 6 que apresenta o número de termos da série na abscissa.

Neste ponto, avalia-se a influência do ruído na amostra de séries temporais periódicas. Para isso, consideram-se três níveis de ruído aleatório contaminando o sinal: A = 0 (sinal ideal), A = 0,314 e A = 0,628. Os resultados mostram que estes níveis de ruído apresentam uma pequena influência nas predições (Figura 7).

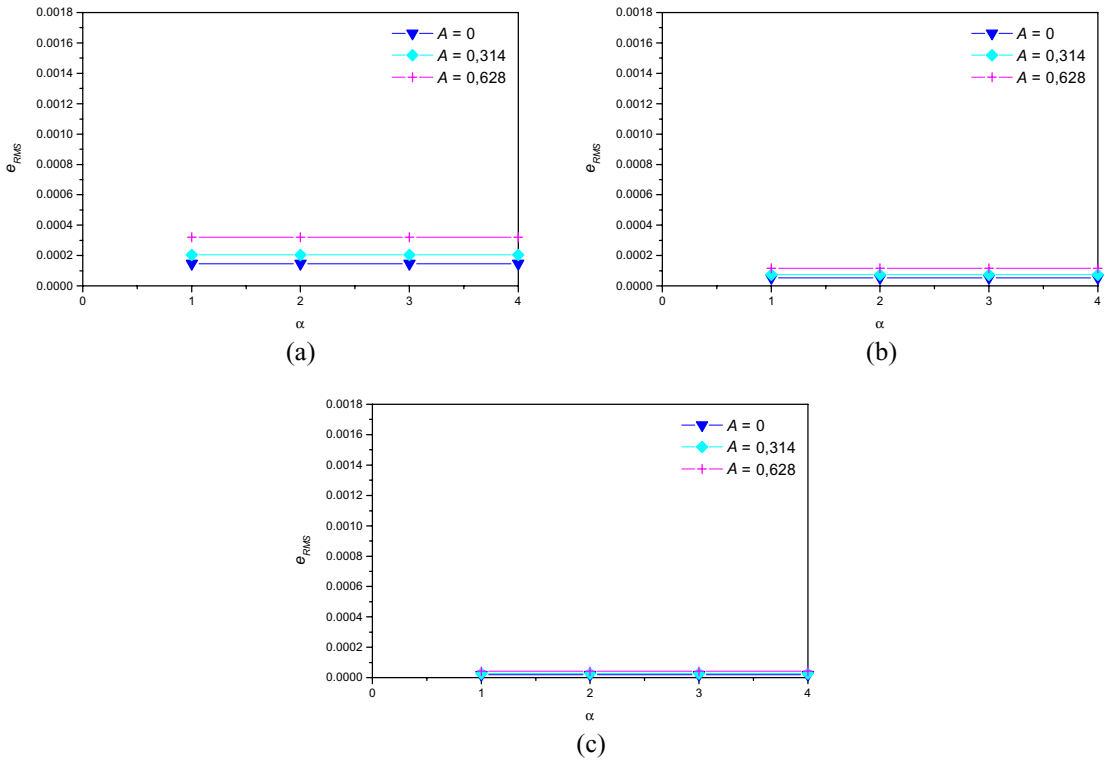


Figura 7. Erro na amostra mostrando a influência do ruído na predição em série periódica, p = 100. (a) m = 5000; (b) m = 10000; (c) m = 20000.

3.2. Sinal Caótico

Considere agora um sinal de caótico com $N = 30000$, obtido através de uma simulação numérica, a partir do qual deseja-se fazer a predição de valores futuros da série. A Figura 8 apresenta um trecho deste sinal.

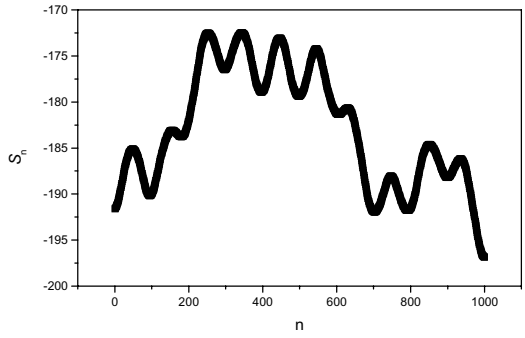


Figura 8. Série caótica.

A predição em séries temporais caóticas é efetuada a partir de diferentes valores para o número de termos da série, m, e diferentes ordens do polinômio de aproximação, α. A Figura 9a mostra a série temporal, o seu prolongamento e as diferentes predições efetuadas a partir de polinômios de ordem 1, 2, 3 e 4. A Figura 9b apresenta a região de predição mostrando a comparação entre o prolongamento da série e as predições. Estas predições consideram $m = 500$, prevendo $p = 100$ pontos no futuro. A Figura 10, por sua vez considera $m = 2000$ e $p = 100$.

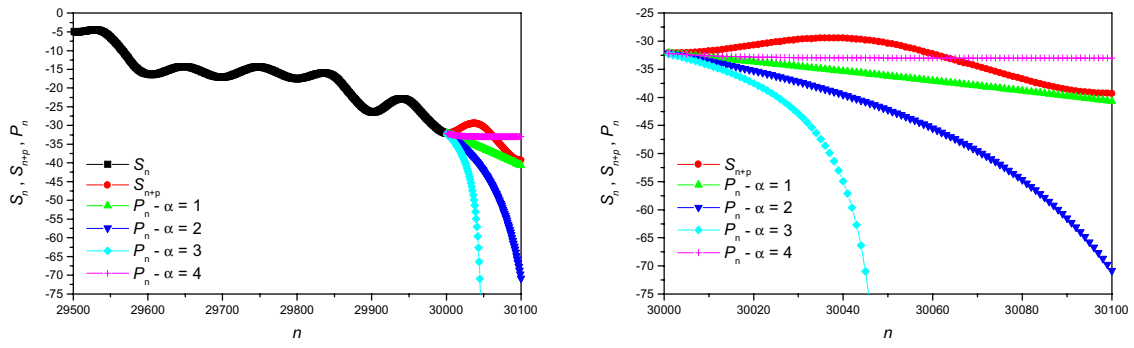


Figura 9. Predição em série caótica, $A = 0$ (sem ruído), $m = 500$ e $p = 100$.

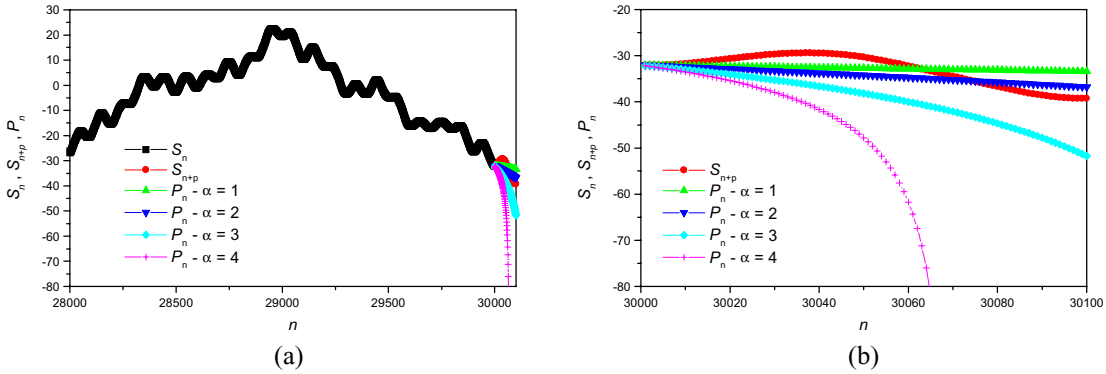


Figura 10. Predição em série caótica, $A = 0$ (sem ruído), $m = 2000$ e $p = 100$.

Uma análise do erro na amostra é tratada a seguir. A Figura 11 apresenta o erro na amostra para diferentes tamanhos da série temporal e ordens do polinômio. Os resultados para a série caótica são bem menos comportados do que o da série periódica, discutidos no item anterior. Neste caso, vê-se que o aumento do número de termos da série não necessariamente melhora a aproximação. Outra maneira de observar estes resultados está mostrada na Figura 12 que apresenta o número de termos da série na abscissa.

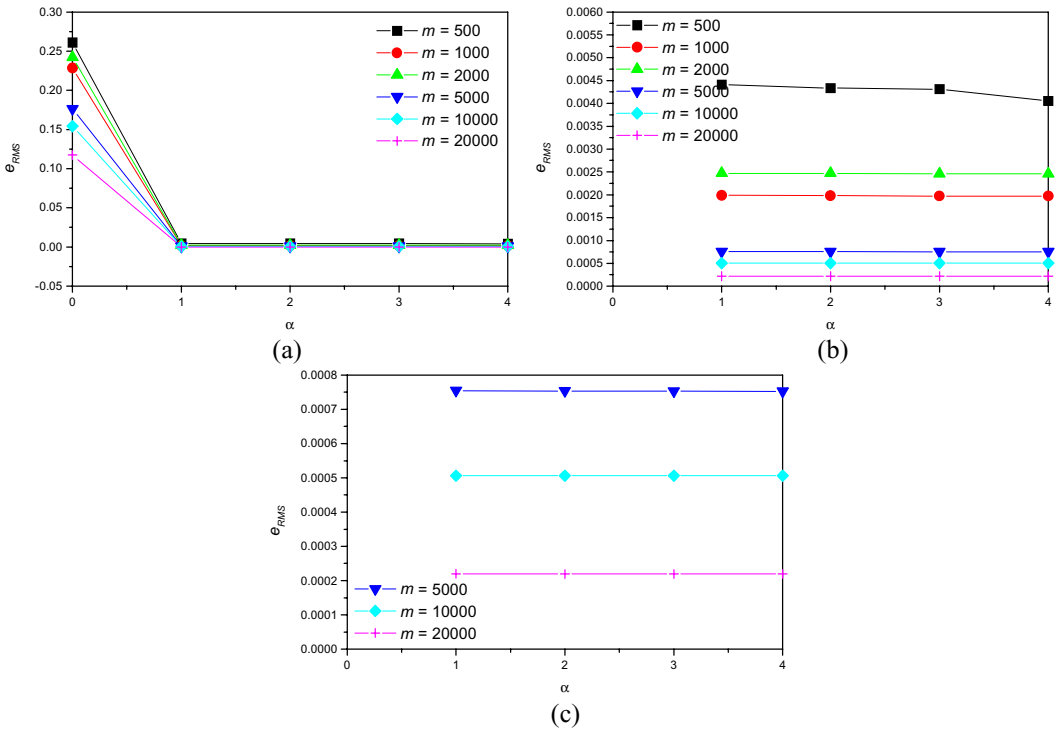


Figura 11. Erro na amostra em função do grau do polinômio e do número de termos da série na predição de um sinal caótico (A = 0, sem ruído, e p = 100): (a) α = 0, 1, 2, 3, 4; m = 500 a 20000; (b) α = 1, 2, 3, 4; m = 500 a 20000; (c) α = 1, 2, 3, 4; m = 5000 a 20000.

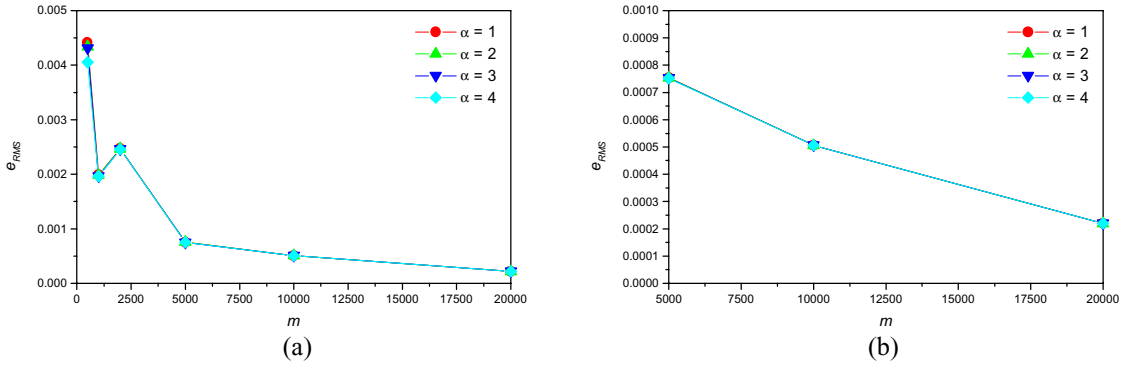


Figura 12. Erro na amostra em função do grau do polinômio e do número de termos da série na predição de um sinal caótico (A = 0, sem ruído, e p = 100). (a) m = 500 a 20000; (b) m = 5000 a 20000.

Considere agora uma análise da influência do ruído na predição de séries temporais caóticas. Para isso, três níveis de ruído aleatório são avaliados: A = 0 (sinal ideal), A = 0,314 e A = 0,628. Mais uma vez, pode-se verificar que estes níveis de ruído apresentam uma pequena influência nas predições (Figura 13).

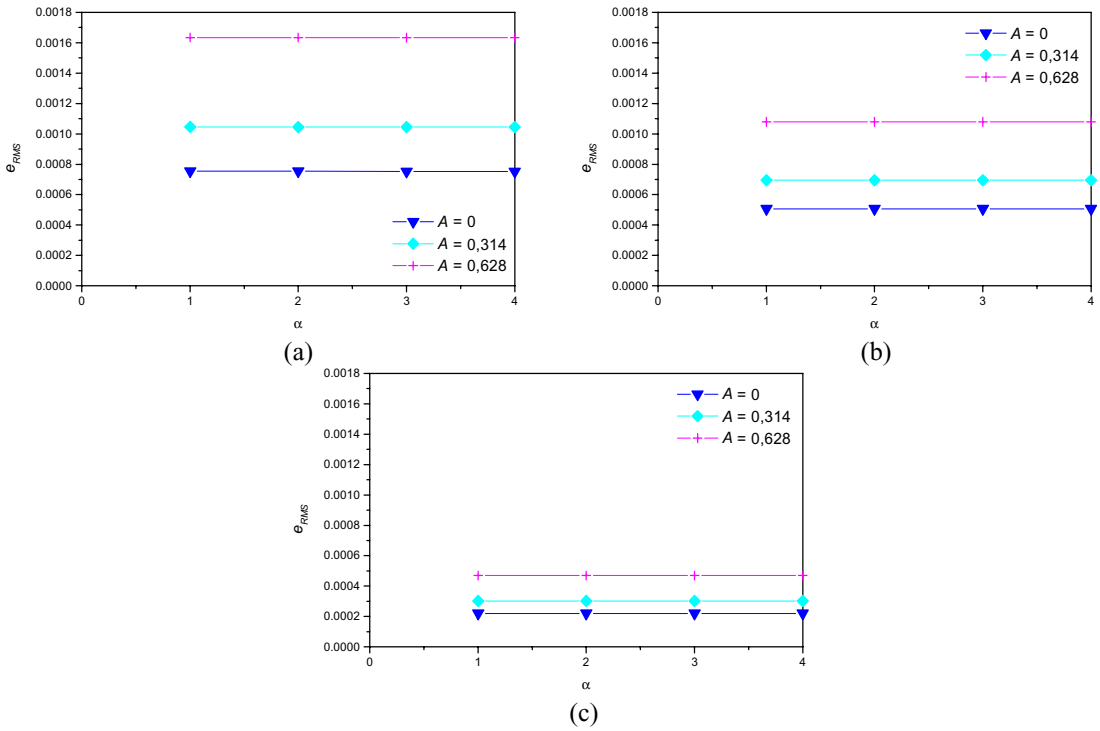


Figura 13. Erro na amostra mostrando a influência do ruído na predição em série caótica, p = 100. (a) m = 5000; (b) m = 10000; (c) m = 20000.

4. Considerações Sobre os Tipos de Séries

Esta seção tem como objetivo apresentar uma comparação entre a predição de sinais periódicos e caóticos. Para isso, apresentam-se comparações entre os erros nas predições dos dois tipos de sinais (Figura 15). Diferentes valores do número de termos para o ajuste (m = 20000 e m = 30000) são considerados para um sinal ideal (A = 0), com um polinômio de predição de grau um (α = 1). Conforme esperado, pode-se dizer que o sinal caótico é menos previsível que o periódico. Esta análise permite estipular uma região de previsibilidade para cada tipo de sinal. Por exemplo, para um erro aceitável e_{RMS} = 0,05 (linha horizontal), um sinal periódico admite uma previsão de longo prazo enquanto o sinal caótico admite uma previsão de, aproximadamente, 100 termos para m = 20000 e 200 termos para m = 30000.

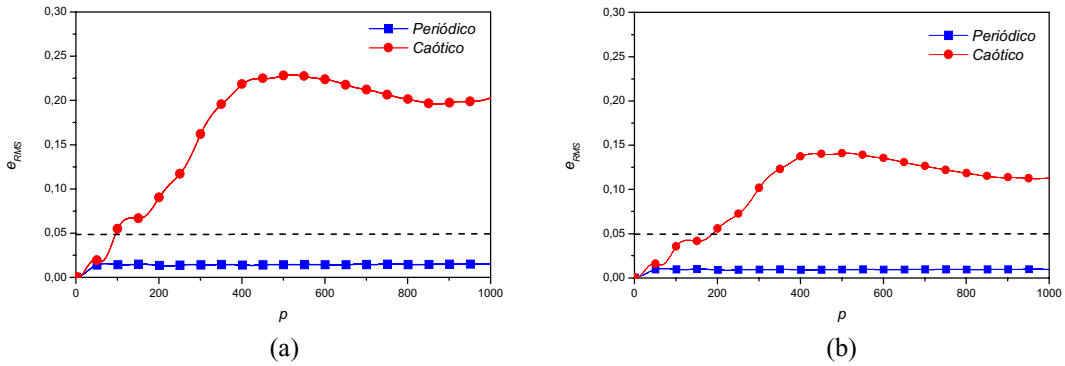


Figura 15. Erro na predição comparando os sinais periódico e caótico. (a) m = 20000; (b) m = 30000.

5. Conclusões

Este trabalho apresenta uma discussão sobre predição em séries temporais não-lineares a partir de uma modelagem polinomial, aplicado ao pêndulo não-linear. As séries temporais são obtidas através de integração numérica, elegendo-se uma variável de estado como um sinal característico do sistema. Dois tipos de sinais são analisados: periódico e caótico. Os resultados mostram que os sinais caóticos possuem uma previsibilidade menor que os sinais periódicos. No que diz respeito ao tamanho da série, o aumento do número de termos tende a melhorar a qualidade da predição nos sinais periódicos. O mesmo não pode ser dito sobre os sinais caóticos. Sobre a análise de séries contaminadas por ruído, pode-se dizer que a influência do ruído na predição não é significativa na técnica utilizada.

6. Referências

Abarbanel, H. D. I., 1995, "Tools for Analyzing Observed Chaotic Data", In Guran, A. & Inman, D. J., editors, *Smart Structures, Nonlinear Dynamics, and Control*, Prentice Hall, U.S.A., pp. 1 – 86.

Abarbanel, H. D. I.; Brown, R.; Sidorowich, J. J. & Tsimring, L. Sh., 1993, "The Analysis of Observed Chaotic Data in Physical Systems", *Reviews of Modern Physics*, Vol. 65, No. 4, pp. 1331 – 1392.

Brandt, M. E.; Ademoglu, A. & Pritchard, W. S., 2000, "Nonlinear Prediction and Complexity of Alpha EEG Activity", *International Journal of Bifurcation and Chaos*, Vol. 10, No. 1, pp. 123 – 133.

Cao, L.; Mees, A. & Judd, K., 1997, "Modeling and Predicting Non-Stationary Time Series", *International Journal of Bifurcation and Chaos*, Vol. 7, No. 8, pp. 1823 – 1831.

Cao, L.; Hong, Y.; Fang, H. & He, G., 1995, "Predicting Chaotic Time Series with Wavelet Networks", *Physica D*, Vol. 85, pp. 225 – 238.

Casdagli, M., 1989, "Nonlinear Prediction of Chaotic Time Series", *Physica D*, Vol. 35, pp. 335 – 356.

Farmer, J. D. & Sidorowich J. J., 1987, "Predicting Chaotic Time Series", *Physical Review Letters*, Vol. 59, No. 8, pp. 845 – 848.

Hegger, R.; Kantz, H. & Schreiber, T., 1999, "Practical Implementation of Nonlinear Time Series Methods: The TISEAN Package", *Chaos*, Vol. 9, No. 2, pp. 413 – 435.

Hsieh, D. A., 1991, "Chaos and Nonlinear Dynamics – Application to Financial Markets", *Journal of Finance*, Vol. 46, No. 5, pp. 1839 – 1877.

Kantz, H. & Schreiber, T., 1997, "*Nonlinear Time Series Analysis*", Cambridge University Press, United Kingdom, pp. 13 – 28, 42 – 57, 202 – 222.

Kugiumtzis, D.; Lillekjendlie, B. & Chistophersen, N., 1994, "Chaotic Time Series, Part I: Estimation of Some Invariant Properties in State-Space", *Modeling, Identification and Control*, Vol. 15, No. 4, pp. 205 – 224.

Kuo, R. J., 2001, "A Sales Forecasting System Based on Fuzzy Neural Network with Initial Weights Generated by Genetic Algorithm", *European Journal of Operational Research*, Vol. 129, pp. 496 – 517.

Leung, M. T.; Daouk, H. & Chen, A-S., 2000, "Forecasting Stock Indices: a Comparison of Classification and Level Estimation Models", *International Journal of Forecasting*, Vol. 16, pp. 173 – 190.

Lillekjendlie, B.; Kugiumtzis, D. & Chistophersen, N. (1994), "Chaotic Time Series, Part II: System Identification and Prediction", *Modeling, Identification and Control*, v.15, n.4, pp.225 – 243.

Masters, T., 1995, "*Neural, Novel & Hybrid Algorithms for Time Series Prediction*", John Wiley & Sons, U.S.A., pp. 181 – 237.

Myano, T.; Kimoto, S.; Shibuta, H.; Nakashima, K.; Ikenaga, Y. & Aihara K., 2000, "Time Series Analysis and Prediction on Complex Dynamical Behavior Observed in a Blast Furnace", *Physica D*, Vol. 135, pp. 305 – 330.

Press, W. H.; Teukolsky, S. A.; Vetterling, W. T. & Flannery, B. P., 1992, "*Numerical Recipes in C – The Art of Cientific Computing*", Cambridge University Press, U.S.A., pp. 43 – 50, 564 – 572, 656 – 714.

Sakai, K. & Aihara, K., 1994, "Nonlinear Vibrations in an Agricultural Implement System", *International Journal of Bifurcation and Chaos*, Vol. 4, No. 2, pp. 465 – 470.

Salvino, L. W.; Cawley, R.; Grebogi, C. & Yorke, J. A., 1995, "Predictability in Time Series", *Physics Letters A*, Vol. 209, pp. 327 – 332.

Schreiber, T., 1999, "Interdiciplinary Application of Nonlinear Time Series Methods", *Physics Reports*, Vol. 308, pp. 1 – 64.

Sfetsos, A., 2000, "A Comparison of Various Forecasting Techniques Applied to Mean Hourly Wind Speed Time Series", *Renewable Energy*, Vol. 21, pp. 23 – 35.

Toth, E.; Brath, A. & Montanari, A., 2000, "Comparison of Short-Term Rainfall Prediction Models for Real-Time Flood Forecasting", *Journal of Hydrology*, Vol. 239, pp. 132 – 147.

Weigend, A. S. & Gershenfeld, N. A., 1994, "*Time Series Prediction: Forecasting the Future and Understanding the Past*", Perseus Books.

TIME SERIES PREDICTION FROM A POLYNOMIAL MODELING: APPLICATION TO THE NONLINEAR PENDULUM

Eduardo Gomes Ferreira Pinto
Marcelo Amorim Savi
Instituto Militar de Engenharia
Department of Mechanical and Materials Engineering
22290-270 – Rio de Janeiro – RJ – Brazil
E-Mail: savi@epq.ime.eb.br

Abstract. *Time series prediction has application in many different areas. This contribution presents a discussion on the nonlinear time series prediction from a polynomial modeling, considering the influence of the number of data points and the noise sensibility. Nonlinear pendulum is considered as an application of the formulation. Signals are generated by numerical integration of the mathematical model, selecting a single variable of the system as a time series. Noise sensibility is evaluated introducing a random noise in the signal. Basically, the analyses of periodic and chaotic motions are carried out. Results show the main difficulties for the prediction, identifying some critical parameters.*

Keywords. *Prediction, Time series, Chaos, Nonlinear pendulum.*

XVI CONGRESSO BRASILEIRO DE ENGENHARIA MECÂNICA
16th BRAZILIAN CONGRESS OF MECHANICAL ENGINEERING

LOWER BOUNDS OF DYNAMIC BUCKLING LOADS USING MELNIKOV CRITERION

Donald Mark Santee
Departamento de Matemática
Campus de Catalão – Universidade Federal de Goiás
Caixa Postal 56 – CEP 75705-020 – Catalão, GO, Brazil
e-mail: santee@innet.psi.br

Paulo Batista Gonçalves
Departamento de Engenharia Civil
Pontifícia Universidade Católica do Rio de Janeiro
Rua Marques de São Vicente 225 – Gávea
CEP 22451-041 Rio de Janeiro, RJ, Brazil
e-mail: paulo@dec.puc-rio.br

Abstract *. Imperfection-sensitive structures usually become unstable for load levels below their linear buckling load associated to the perfect structure. In some cases as, for example, cylindrical shells under axial compression, there are experimental buckling loads 90% inferior to the theoretical load (Batista, 1979). For static loads, Koiter's Theory of Initial Postbuckling (Koiter, 1945) presents some expressions that allow the prediction of the buckling load taking into account the geometric non-linearity and the imperfection level. There are also some equivalent expressions for imperfection-sensitive structures under impact loads (Elishakoff, 1980). Because the expressions for the critical load were developed from a static equilibrium analysis, they actually calculate an upper bound for the load carrying capacity of the real structure (Santee, 1999). This work develops lower bound expressions for the buckling load of imperfection-sensitive structures under dynamic loads; special attention is given to periodic loads. Under this kind of load, the structure's dynamics becomes very complex, presenting various type of non-linear phenomena such as histeresis, saddle-node bifurcations, period doubling bifurcations and period doubling cascade bifurcations to chaos. The inspiration for the development of a lower bound to the buckling load is the behavior of the fundamental solution's basin of attraction boundary. This boundary is defined by the unstable manifold of the fundamental's solution corresponding saddle-point. Melnikov developed a procedure to determine an approximation for the first crossing of the stable and unstable manifolds of the saddle-point related to the fundamental stable solution. When the stable and unstable manifolds cross transversally at one point, they cross transversally at an infinite number of discrete points. Since the unstable manifold is the fundamental solutions basin of attraction boundary, this indicates that the basin of attractions becomes, at least partially, fractal. Thus the load level at which the tangling of the stable and unstable manifolds occurs, can be taken as the load that marks the beginning of the loss of stability of the structure, consequently a lower bound for the structures load capacity.*

Keywords: *Melnikov Method, Dynamic Buckling, Nonlinear Vibration, Imperfection-Sensitivity.*

1. Introduction

The main motivation for this work comes from a notorious and persistent discrepancy between theoretical and experimental results of the buckling loads of thin-walled elastic shell structures. For many years this discrepancy resisted all attempts to explain the large scatter in experimental results and why the magnitude of the critical load amounted sometimes to only a small percentage of the theoretical value. The general explanation for shell buckling was found when Koiter's general theory of buckling and postbuckling behavior of elastic structures became known. Koiter was the first to realize the immense importance of postbuckling behavior in connection with the presence of imperfection (El Nashie, 1992). Budiansky (1974) and other investigators provided further contributions in this field. However, for imperfection-sensitive structures, investigations are mainly concerned with static or step loads.

This paper addresses the problem of calculating the dynamic buckling load of imperfection-sensitive structures under a harmonic transversal load. Actually because of the resonance phenomenon this is one o the worst possible types of load. By itself the harmonic load is an important engineering load that can be caused, for example, by the vibration of an eccentric motor on or nearby the structure, or by the periodic excitation of the waves on an offshore structure.

The motion equation of an imperfection-sensitive structure when reduced to a single degree of freedom model can be expressed by

$$\ddot{x} + 2\eta\omega_0\dot{x} + \varepsilon + \omega_0^2 x + \beta x^2 - \gamma x^3 = F\cos(\Omega t) \tag{1}$$

In Eq. (1) η is the viscous damping parameter, ω_0 is the structure's natural frequency, ε is an imperfection parameter, β and γ are geometrical nonlinearity parameters, F is the magnitude of the externally applied load and Ω the loads frequency. The dots indicate derivation with respect to time t.

When $\gamma = 0$ with $\beta \neq 0$ the equation has only a quadratic nonlinearity parameter and the structure is said to be *liable to asymmetric bifurcation*. When $\beta = 0$ with $\gamma > 0$ the equation has only a cubic nonlinearity parameter and the structure is said to be *liable to unstable symmetric bifurcation*.

The challenge that is posted for engineering design is to find the load level F that the imperfection-sensitive structure can bear vibrating boundedly. The ideal solution to this problem would be an algebraic expression of the harmonic buckling load as a function of the other structural parameter.

2. The escape load

The solution set of Eq. (1) can be classified in bounded solutions and unbounded solutions. Bounded solutions indicate vibration of the structure. Unbounded solutions indicate ruin of the structure, as its displacements become increasingly large and incompatible with the structure's use and hypothesis embodied into the mathematical modeling. Unbounded solutions are also called *escape solutions*, or simply *escape*. In this work, as in the rest of the literature (Malasoma & Lamarque, 1994; Soliman, 1989; Szemplinska-Stupnika, 1995; Virgin, 1988), one is interested in the values of F and Ω that lead to escape solutions. The minimum value of F when all other parameters are maintained fixed is called an *escape load*. The underlying dynamics that ultimately leads to escape can be very complex, consequently the *escape boundary*, which is the set of escape loads in the parameter space, can be of fractal nature (Santee, 1999), thus by practical means, unpredictable.

Figure (1) shows the escape load as a function of the forcing frequency Ω, for the fixed set of parameters η=0.05, ε=0, ω_0=1, β=1 and γ=0 (a structure liable to asymmetric bifurcation). The escape boundary was obtained numerically following Algorithm 1. The algorithm considers two different kinds of applied load: A *suddenly applied load* (red line), and a *gradually increasing load* (black line). Figure (1) also shows the *escape load* for a structure under a step load of infinite duration. The gradually increasing case corresponds to the *static critical load*.

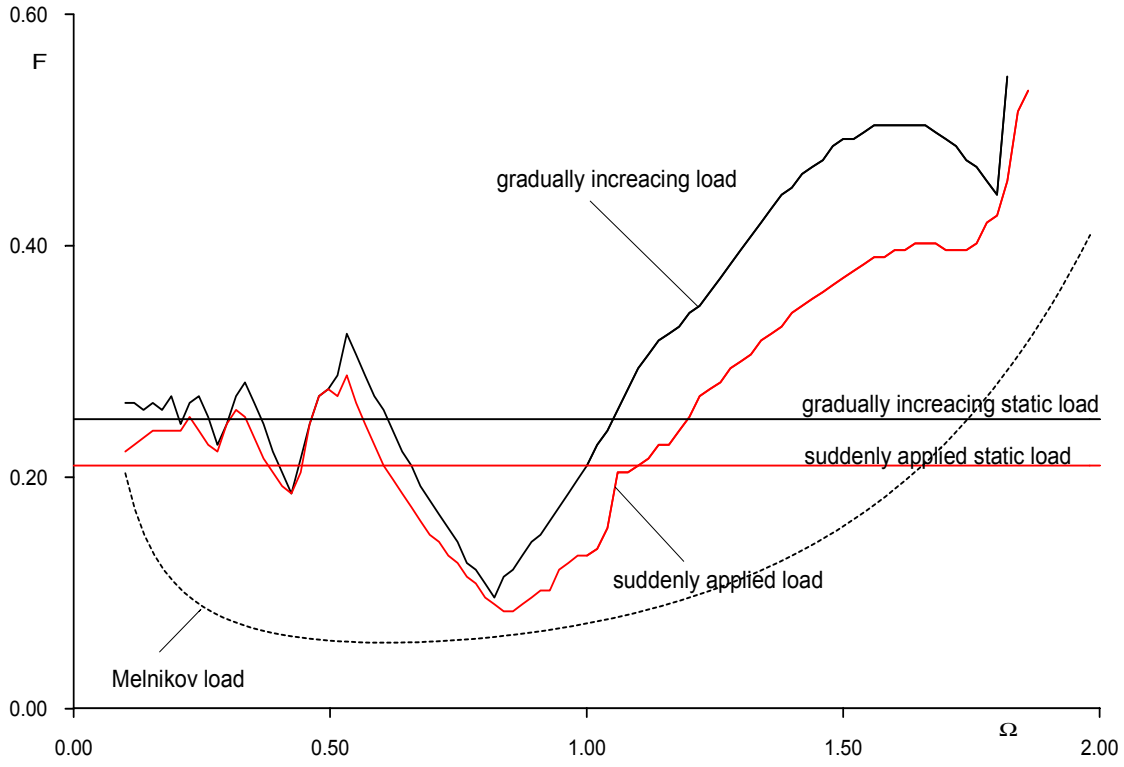


Figure 1. Escape boundary for a structure liable to asymmetric bifurcation with parameters η=0.05, ε=0, ω_0=1 e β=-1.

1. *Set an initial value for the excitation frequency* $\Omega = \Omega_{min}$.
2. *Zero the force parameter F and the initial conditions.*
3. *Integrate numerically during a time period* T_0 *, of r times the period of the exciting force*

$$T_0 = r\frac{2\pi}{\Omega}.$$

4. *If escape occurs in this time (with displacement and velocity superior to a pre-defined large number), save the value of F, zero F and the initial condition, and increment the frequency.*
5. *If escape does NOT occur, increment the value of F and update the initial conditions to a) Zero (for a suddenly applied load); or b) the previous final state (for a gradually increasing load);*
6. *Go to step 3 until the frequency* Ω *reaches a defined maximum value.*

Algorithm 1. Evaluation of the escape boundary.

In order to simulate the two kinds of applied load Alg. (1) can run using two different strategies regarding the initial conditions. For the suddenly applied load it uses a zero position and velocity initial condition at each increment of the applied load. For the gradually increasing load it uses the final position and velocity of the previous increment as the initial conditions of the current increment.

The escape boundary in Fig. (1) shows that as the value of the forcing frequency Ω varies, there is a series of valleys associated to the super harmonics of various orders culminating with a deep valley around the natural frequency. For high excitation frequencies the escape load is many times larger than the corresponding static escape load. One important aspect of the escape load is that it represents only an upper bound of the actual load bearing capacity of the structure. Because of dynamic perturbations, an imperfection-sensitive structure can escape at load levels much lower then the escape load, as will be shown later by the analysis of its basin of attraction.

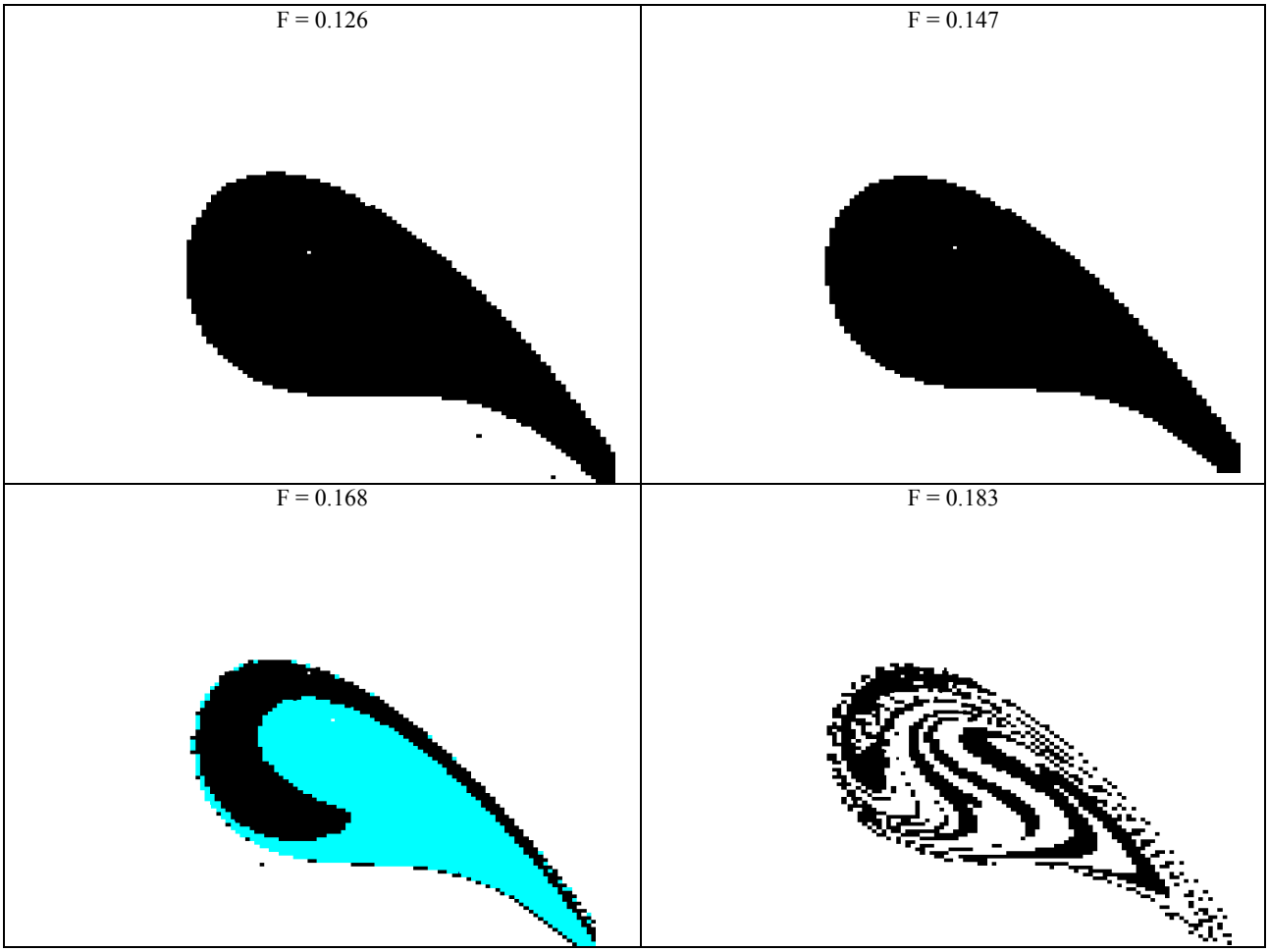


Figure 2. The evolution of the basin of attraction: sudden loss of integrity.

3. Basin of attraction and structural stability

Mathematically the *basin of attraction* of a periodic solution is the set of all initial conditions that lead to a solution (attractor) as time goes to infinity. This means that if a periodic solution has a large compact basin of attraction, it will be stable under finite perturbations. On the other hand, if it has a small, or fragmented, basin of attraction, small finite perturbations can lead the solution to escape even if the solution is stable. Thus a measure of the stability of the structure, in particular its safety has to be based on a global view of the behavior of the structure. This global view can be expressed mathematically by the characteristics of the basin of attraction and its boundary.

The concept of basin of attraction is based on the limit $t \to \infty$. Because of limitations in the numerical integration a practical concept for basin of attraction is used. This practical concept is the basin of *r*-attraction. A *basin of r-attraction* is the set of all initial conditions that lead to the neighborhood of the respective periodic solution in *r* times the forcing period T = 2 π/Ω. As the integration time is increased, the basin of *r*-attraction tends asymptotically to the basin of attraction. Our experience has shown that a basin of 32-attraction is a reasonable approximation.

Numerical explorations have shown that the way the basin of attraction changes as the load level *F* is increased, can be classified into one of two groups: One is that it gradually decreases until it vanishes completely; the other is that its shape remains basically the same as the load level increases, until is suddenly becomes fractal, as shown in Fig. (2). In these two types of basin of attraction evolution, the area tends to zero as the load is increased. One important thing to note is that even at the eminence of escape, when the basin of attraction is very small, the periodic solution represented by the fixed point in the Poincaré Map is a stable solution, making clear that when one uses the stability of the periodic solution as a measure of the structure's stability, this value furnishes only an upper bound to the true stable limit load.

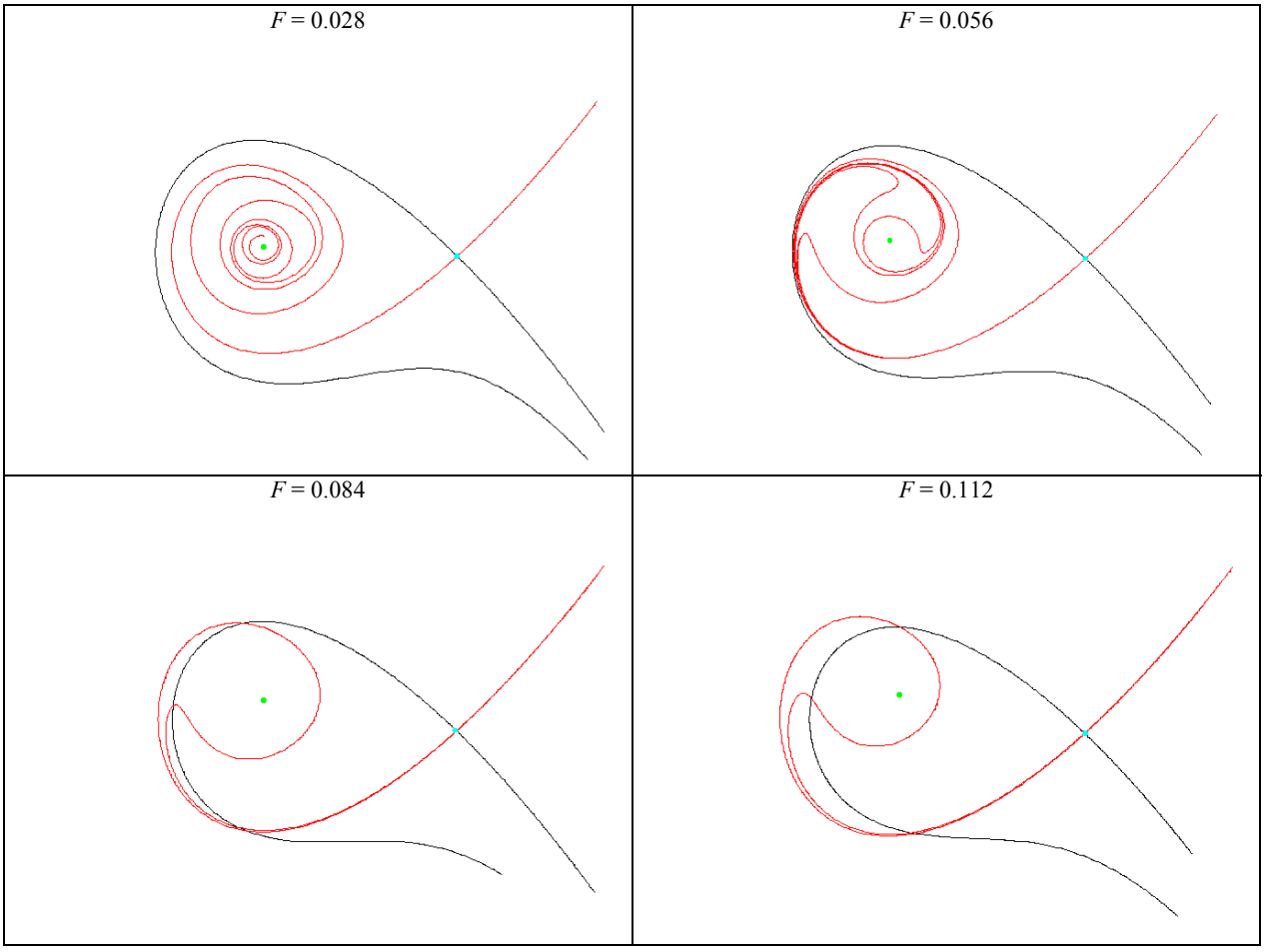


Figure 3. The change in the stable and unstable manifolds as the load *F* increases.

4. Stable and unstable manifolds: Melnikov's load

By definition a saddle point presents a stable behavior in the direction of some eigenvectors, and an unstable behavior in the direction of others. The set of initial conditions that lead the solution to the saddle point as $t \to \infty$ is the stable manifold, and the set of initial conditions that lead the solution to the saddle point as $t \to -\infty$ is the unstable manifold. The mere existence of a saddle point indicates the existence of more then one attractor, which can be another periodic, a quasi-periodic or chaotic solution, or infinity. The stable manifolds of the saddle point constitute the boundary between the two attractors.

When an imperfection-sensitive structure described by Eq. (1) is unloaded (F=0), the system has only one stable equilibrium point, and one saddle point. In this situation the saddle point's stable manifold goes smoothly around the stable equilibrium point marking the boundary of the basin of attraction of this point. One branch of the unstable manifold lies completely inside the basin of attraction leading to the stable equilibrium solution. The other branch lies outside the basin of attraction and tends to infinity.

Figure (3) shows how the stable (black line) and unstable (red line) manifolds change as the load level is increased. Note that they become closer together. From a certain load level F^{M} on, the stable and unstable manifolds cross transversally. There is a theorem (Guckenheimer & Holmes, 1983) that states that when the stable and unstable manifolds cross transversally at one point, they cross at an infinite number of points, thus this crossing indicates the beginning of the erosion of the basin of attraction.

The algebraic prediction of the first crossing of the stable and unstable manifolds can be used as a criterion for determining the stability of the structure. The mathematical Theory of Dynamical Systems presents o method, called Melnikov Method, to predict this crossing.

5. Melnikov's Method.

The starting point of the Melnikov Method is the *Melnikov Function*. This function gives a measure of the distance of the stable and unstable manifolds, when this distance is small (Moon, 1992; Wiggins, 1990). It is applied to problems where the damping is small, and when the algebraic expressions for the stable and unstable manifolds for zero damping are known. When both the damping and the externally applied force are small, the vector field of the system can be expressed generically by:

$$\begin{aligned} \dot{x} &= \frac{\partial E}{\partial v} + \xi g_1 \\ \dot{v} &= \frac{\partial E}{\partial x} + \xi g_2 \end{aligned} \tag{2}$$

where the vector $g = g(x,v,t) = (g_1, g_2)$, ξ is a small parameter and $E(x,v)$ is the total energy of the free and undamped system ($\xi = 0$). Also admit that $g(t)$ is periodic, that is, it satisfies the relation:

$$g(t+T) = g(t) \tag{3}$$

Melnikov's function is given by

$$M(t_0) = \int_{-\infty}^{\infty} g^* \cdot \nabla E(x^*, v^*) dt \tag{4}$$

where $g^* = g(x^*, v^*, t+t_0)$. $x^*(t)$ and $v^*(t)$ are the algebraic expressions of position and velocity of the stable and unstable manifolds for the conservative system.

The two manifolds cross when this distance is zero, such that:

$$M(t_0) = 0 \tag{5}$$

Equation (5) leads to an algebraic expression that can be used to calculate the load level F^{M} above which the tangling of the stable and unstable manifolds start. Next we apply the Melnikov method to imperfection-sensitive structures whose motion is described by Eq. (1).

6. Applying Melnikov's method to imperfection-sensitive structures.

Considering zero damping and zero external force, Eq. (1) can be integrated once which leads to the following total energy expression:

$$\frac{1}{2}\dot{x}^2 + \frac{1}{2}\omega_0^2 x^2 + \varepsilon x + \frac{1}{3}\beta x^3 - \frac{1}{4}\gamma x^4 = c \tag{6}$$

and c is the integration constant, which is determined by the initial conditions.

From the set of solutions of Eq. (6), the one of our concern is the solution that starts at the saddle point x_{saddle} and $v = 0$. This solution is of particular interest because it marks the limits of the potential well. Using this point as the initial conditions one can calculate the value of the integration constant c that expresses the desired solution:

$$\frac{1}{2}\omega_0^2 x_{saddle}^2 + \varepsilon x_{saddle} + \frac{1}{3}\beta x_{saddle}^3 - \frac{1}{4}\gamma x_{saddle}^4 = c = E_{saddle} \tag{7}$$

At this point Eq. (7) can be integrated once more, leading to:

$$\int \frac{dx}{\sqrt{2}\sqrt{E_{saddle} - \frac{1}{2}\omega_0^2 x^2 - \varepsilon x - \frac{1}{3}\beta x^3 + \frac{1}{4}\gamma x^4}} = \pm t + c \tag{8}$$

In Eq. (8) E_{saddle} is the energy value at the saddle point. And c_1 is a constant that can be adjusted in such a way that x has its maximum at $t = 0$. Generally it is not possible to obtain an analytical expression for the result of Eq.(8). But an algebraic expression of this solution is important for the derivation of the desired lower bound. For particular cases we have developed approximations to the solutions of this equation. These approximations are in fact one of the main contributions of this paper, and are presented next.

6.1. Structures liable to asymmetric bifurcation.

In the case of structures liable to asymmetric bifurcation ($\gamma = 0$) with no imperfection ($\varepsilon = 0$) it is possible to solve Eq. (8) algebraically. The solution is:

$$x_p(t) = -\frac{\omega_0^2}{\beta}\frac{e^{2\omega_0 t} - 4e^{\omega_0 t} + 1}{e^{2\omega_0 t} + 2e^{\omega_0 t} + 1} \tag{9}$$

The effect of the imperfection parameter ε can be introduced by observing that it does not change qualitatively the solution (the homoclinic orbit continues to be a homoclinic orbit), but changes only the position of the center and saddle point. Thus the approximate solution can be expressed by the following:

$$x(t) = A(\varepsilon) + L(\varepsilon)x_p[a(\varepsilon)t] \tag{10}$$

The coefficients A and L in Eq. (10) can be obtained by the restrictions:

$$\begin{aligned} &\lim_{t\to\infty} x(t) = x_{saddle} \\ &x(0) = x_{\max} \end{aligned} \tag{11}$$

This leads to:

$$A(\varepsilon) = -\frac{\varepsilon}{\omega_0^2}$$
$$L(\varepsilon) = 1 - \frac{2\beta\varepsilon}{\omega_0^4} \tag{12}$$

The time scale coefficient *a(ε)* in Eq. (10) can be obtained by applying Galerkin's method on the residue

$$R(t) = \frac{1}{2}\dot{x}^2 + \frac{1}{2}\omega_0^2 x^2 + \varepsilon x + \frac{1}{3}\beta x^3 - E_{saddle} \tag{13}$$

and by using the weight function $\delta x = x(t)$. This leads to:

$$a(\varepsilon) = 1 - \frac{\beta\varepsilon}{\omega_0^4} \tag{14}$$

Remembering that $v = \frac{dx}{dt}$, Melnikov's function becomes:

$$M(t_0) = \int_{-\infty}^{\infty} \overline{F} v \cos[\Omega(t + t_0)]dt - 2\overline{\eta}\omega_0 \int_{-\infty}^{\infty} v^2 dt \tag{15}$$

In Eq. (15) $F = \xi\overline{F}$, $\eta = \xi\overline{\eta}$ and ξ is the small perturbation parameter. By substituting $x(t)$ into Eq. (15) and calculating the integrals we obtain:

$$M(t_0) = \overline{F} sin(\Omega t_0)\left[-\frac{6L\pi\Omega^2}{a^2\beta sinh\left(\frac{\pi\Omega}{a\omega_0}\right)}\right] - 2\overline{\eta}\omega_0\left[\frac{6La\omega_0^5}{5\beta^2}\right] \tag{16}$$

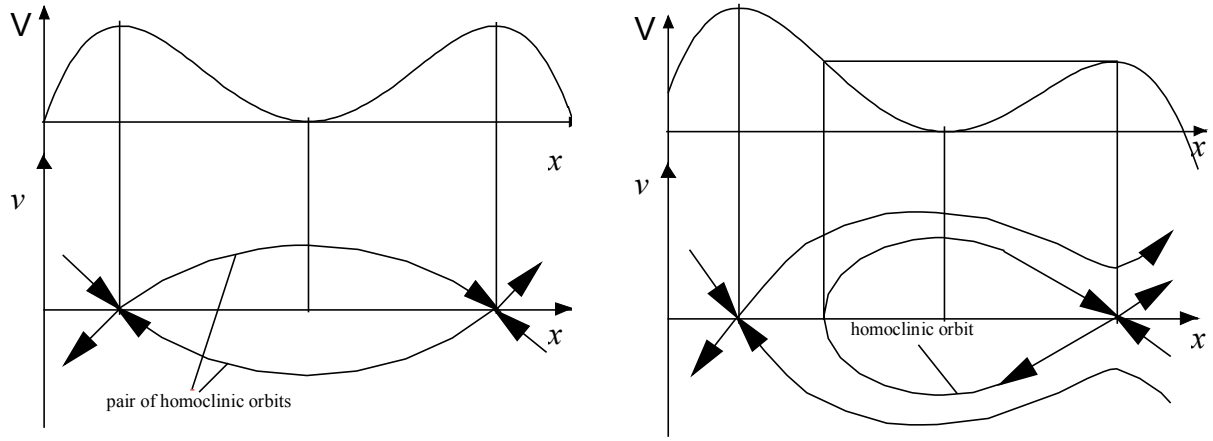
Thus $M(t_0)$ can be zero only when

$$F > F^M = \frac{\eta}{\pi}\left[\frac{2\omega_0^6}{5\beta\Omega^2} - \frac{\varepsilon\omega_0^2}{\Omega^2}\right] sinh\left[\frac{\pi\Omega}{\omega_0} + \frac{\pi\Omega\beta\varepsilon}{\omega_0^5}\right] \tag{17}$$

This is an expression that can be used to calculate the load level above which the structure's basin of attraction starts loosing its integrity.

6.2. Structures liable to unstable symmetric bifurcation.

For structures liable to unstable symmetric bifurcation without imperfection (ε = 0), the set of solutions of the equation contain a pair of heteroclinic orbits, and the potential energy function is symmetric in relation to the equilibrium point, as shown in Fig. (4). The presence of an imperfection changes the set of solutions qualitatively: the energy function is not symmetric, and the pair of heteroclinic orbits becomes a homoclinic orbit. For this reason we will consider the two cases: perfect and imperfect structures separately.



(a) – without imperfection: heteroclinic orbits (b) – with imperfection: homoclinic orbit

Figure 4. Breaking down of the pair of heteroclinic orbits into a homoclinic orbit due to the presence on imperfections.

6.2.1. Supposedly perfect structures.

For this case one can derive the exact solution of Eq. (8). This solution is:

$$x_p(t) = \frac{\omega_0}{\sqrt{\gamma}} \frac{e^{\sqrt{2}\omega_0 t} - 1}{e^{\sqrt{2}\omega_0 t} + 1} \tag{18}$$

Thus Melnikov's function is:

$$M(t_0) = \frac{\overline{F}}{\sqrt{\gamma}} \cos(\Omega t_0) \frac{\sqrt{2}\pi\Omega}{sinh\left(\frac{\pi\Omega\sqrt{2}}{2\omega_0}\right)} - \overline{\eta} \frac{4\sqrt{2}}{3} \frac{\omega_0^4}{\gamma} \tag{19}$$

which leads to the expression

$$F > F^M = \frac{4\eta\omega_0^4}{3\pi\Omega\sqrt{\gamma}} sinh\left(\frac{\sqrt{2}}{2} \frac{\pi\Omega}{\omega_0}\right) \tag{20}$$

for the crossing of the stable and unstable manifolds.

6.2.2. Imperfect structures.

The expression for the homoclinic orbit of the conservative system is inspired on the solution of the double well Duffing equation (Guckenheimer & Holmes, 1983). This solution was chosen because the underlying equation has a cubic nonlinearity and it has a known homoclinic orbit expression. From this we assume that the algebraic expression for the homoclinic orbit has the form:

$$x(t) = A(\varepsilon) + L(\varepsilon)\text{sech}[a(\varepsilon)t] \tag{21}$$

Applying the conditions expressed by Eq. (11) one obtains the coefficients A and L. The value of a is obtained from the condition

$$\dot{x}_{\max} = v_{\max} \tag{22}$$

In Eq. (22) $v_{\max}$ is the maximum velocity of the stable manifold of the conservative system. Thus we have

$$A(\varepsilon) = -\frac{\omega_0}{\sqrt{\gamma}} + \frac{|\varepsilon|}{2\omega_0^2} \tag{23a}$$

$$L(\varepsilon) = \frac{2\omega_0}{\sqrt{\gamma}} - \frac{\sqrt{2}\sqrt{|\varepsilon|}}{\sqrt{\omega_0}\,\gamma^{1/4}} - \frac{|\varepsilon|}{\omega_0^2} \tag{23b}$$

$$a(\varepsilon) = \frac{\sqrt{2}}{2}\omega_0 + \frac{\sqrt{|\varepsilon|}\,\gamma^{1/4}}{2\sqrt{\omega_0}} - \frac{\sqrt{2}|\varepsilon|\sqrt{\gamma}}{2\omega_0^2} \tag{23c}$$

With Eq. (21) and Eq. (23) Melnikov's function becomes:

$$M(t_0) = \overline{F}\,\frac{L\pi\Omega sin(\Omega t_0)}{a(\varepsilon)\cosh\left(\dfrac{\pi\Omega}{2a}\right)} - 2\omega_0\overline{\eta}\,\frac{2L^2(\varepsilon)a(\varepsilon)}{3} \tag{24}$$

which leads to:

$$F > F^M = \frac{2\eta\omega_0}{\pi\Omega}\left[\frac{2\omega_0^3}{3\sqrt{\gamma}} + \frac{\sqrt{2}\,\omega_0^{3/2}\sqrt{|\varepsilon|}}{3\gamma^{1/4}} - 2|\varepsilon|\right]\cosh\left[\frac{\sqrt{2}}{2}\frac{\pi\Omega}{\omega_0} + \frac{\pi\Omega\gamma^{1/4}\sqrt{|\varepsilon|}}{2\omega_0^{5/2}} + \frac{3\sqrt{2}\pi\sqrt{\gamma}\Omega|\varepsilon|}{4\omega_0^4}\right] \tag{25}$$

7. Melnikov Load as a lower bound

The algebraic expressions for the Melnikov Load given by Eq. (17), Eq. (20) and Eq. (25) allow a prediction of the load above which the stable solution's basin of attraction begins to loose its integrity by the tangling of the stable and unstable manifolds of the respective saddle point. Fig. (5) shows the Escape load and the Melnikov load for the case of a structure liable to asymmetric bifurcation. The dots are the escape load for the same structure with random perturbations of 10% on the parameters and a random perturbation of the amplitude and bandwidth of the applied load.

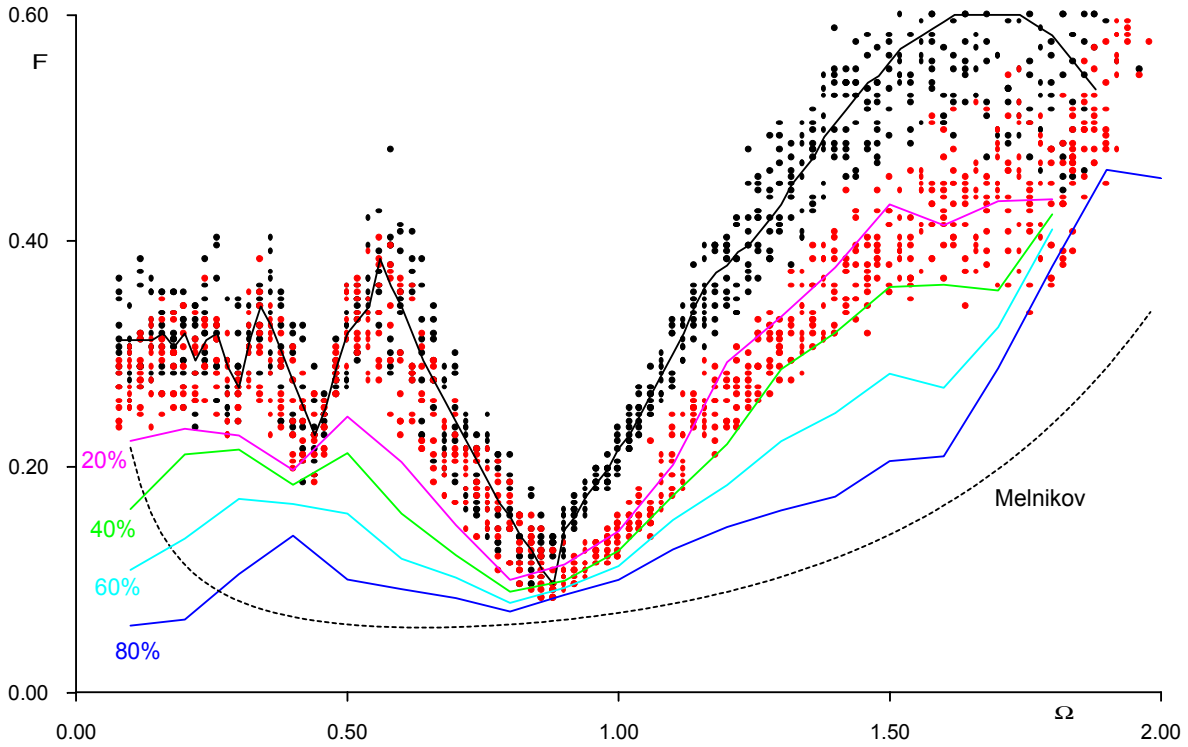


Figure 5. Comparing Melnikov Load with the buckling load of a realistic imperfection-sensitive structure.

The black dots indicate a gradually increasing load, and the red dots indicate a step load. The colored lines mark the size of the basin of attraction. This area is expressed as a percentage of the area of the unloaded structure's basin. The lines that are shown are 80% (dark blue), 60% (light blue), 40% (green) and 20% (pink). Finally Fig. (5) shows the Melnikov Load as a dashed line.

The presence of random imprecisions in the structure's parameters (stiffness, non-linearity, natural frequency, etc.) and in the applied load, make real structures become non-deterministic. In this sense the random perturbation of the parameters are a numerical simulation of a real structure. Note that the random perturbation of 10% of the actual value generates a large scatter in the values of the escape loads. Despite the large variation of the escape loads, the Melnikov load is always smaller indicating that it is a safe lower bound.

8. Conclusions

Just as the escape load can be considered an upper bound to the actual load capacity of the structure, the Melnikov load can be taken as a lower bound of this same capacity. Due to the simplicity of the expressions, Melnikov loads, given by Eq. (17), Eq. (20) and Eq. (25), can be used as a safe design recommendation for imperfection-sensitive structures under a periodic external load. These expressions are totally based on a mathematical reasoning that accounts for the effects of imperfection and dynamical perturbations on the structure.

9. References

Batista, R. C.,1979, "Lower Bound Estimates for Cylindrical Shell Buckling". Doctoral thesis - University College at London, England.

Budianky, B., 1974, "Theory of Buckling and Post Buckling Behavior of Elastic Structures", *Advances in Applied Mechanics*. Vol. 14, p. 1-65

El Naschie, M. S., 1992, "Stress, Stability and Chaos in Structural Engineering: An Energy Approach". McGraw-Hill, London.

Elishakoff, I., 1980, "Remarks on the Static and Dynamic Imperfection-Sensitivity of Nonsymmetric Structures" Journal of Applied Mechanics. Vol. 47, p. 111-115.

Guckenheimer, J. e Holmes, P., 1983, "Nonlinear Oscillations, Dynamical Systems and Bifurcation of Vector Fields". Springer-Verlag, New York.

Koiter, W. T., 1945, "On the Stability of Elastic Equilibrium". Doctoral thesis - Delft University, Holland.

Malasoma, J. -M and Lamarque, C. –H., 1994, "Chaotic Behavior of a Parametrically Excited Nonlinear Mechanical System". *Nonlinear Dynamics* 5: pp. 153-160.

Moon, F. C., 1992, "Chaotic and Fractal Dynamics: an Introduction for Applied Scientists and Engineers". Willey, New York.

Santee, D. M., 1999, "Non-linear vibrations and Instabilities of Imperfection-sensitive Structural Elements". Doctoral thesis - Catholic University of Rio de Janeiro, Brazil.

Soliman, M. S. e Thompson, J. M. T., 1989, "Integrity Measures Quantifying the Erosion of Smooth and Fractal Basins of Attraction". *J. of Sound and Vibration* 135(3), pp. 453-475.

Szemplinska-Stupnicka, W., 1995, "The Analytical Predictive Criteria for Chaos and Escape in Nonlinear Oscillators: A Survey". *Nonlinear Dynamics* 7, pp. 129-147.

Virgin, L. N., 1988, "On the Harmonic Response of an Oscillator with Unsymmetric Restoring Force". *J. of Sound and Vibration* 126. pp. 157-166.

Wiggins, S., 1990, "Introduction to Applied Nonlinear Dynamical Systems and Chaos". Springer-Verlag, New York.

XVI CONGRESSO BRASILEIRO DE ENGENHARIA MECÂNICA
16th BRAZILIAN CONGRESS OF MECHANICAL ENGINEERING

NONLINEAR DYNAMICS OF A PSEUDOELASTIC OSCILLATOR USING A CONSTITUTIVE MODEL BASED ON PLASTICITY

Edgar N. Mamiya
Dianne M. Viana
Universidade de Brasília
Departamento de Engenharia Mecânica
70.910.900 - Brasília - DF
E-Mail: mamiya@unb.br, dianne@enm.unb.br

Marcelo A. Savi
Instituto Militar de Engenharia
Departamento de Engenharia Mecânica e de Materiais
22.290.270 - Rio de Janeiro - RJ
E-Mail: savi@epq.ime.eb.br

Pedro M. C. L. Pacheco
CEFET/RJ
Departamento de Engenharia Mecânica
20.271.110 - Rio de Janeiro - RJ
E-Mail: calas@cefet-rj.br

Abstract. *The present contribution reports on the nonlinear dynamics of a single-degree of freedom pseudoelastic oscillator. Equations of motion are formulated using a constitutive model based on elastoplasticity to describe the restitution force of the oscillator. A numerical procedure based on the operator split technique associated with an iterative procedure is proposed. Results show some characteristics of the responses, showing the influence of internal dissipation.*

Key-words: *Nonlinear Dynamics, Pseudoelasticity, Shape Memory Alloys.*

1. Introduction

SMAs are a family of metals with the ability of changing shape depending on their temperature. SMAs undergo thermoelastic phase transformations, which may be induced either by temperature or stress. Shape memory and pseudoelasticity are effects associated with these transformations (Birman, 1997; Shaw & Kyriakides, 1995). Because of such remarkable properties, SMAs have found a number of applications in engineering sciences (van Humbeeck, 1999; Duerig *et al.*, 1999).

Vibration control, active buckling control, or more generally active structural modifications schemes can be imagine in these applications (Pietrzakowski, 2000; Rogers, 1995, 1990; Rogers *et al*., 1991). Vibration control has a great importance on many engineering problems and usually it is possible to adopt two strategies. Passive control, when no external energy fonts are considered, and active control where there are external fonts. The use of pseudoelastic materials on passive vibration control is associated with internal dissipation. On the other hand, the use on active vibration control is associated with the change of physical properties as a consequence of temperature variations, exploiting shape memory effect.

Since phenomena associated with martensitic transformation are intrinsically nonlinear, its dynamical response may present some characteristics not observed in linear systems. This contribution reports on the nonlinear dynamics of a single-degree of freedom pseudoelastic oscillator. Equations of motion are formulated using a constitutive model based on elastoplasticity to describe the restitution force of the oscillator (Viana & Mamiya, 2001). A numerical procedure based on the operator split technique (Ortiz *et al.*, 1983) associated with an iterative procedure to assure the convergence of the process is proposed (Savi & Pacheco, 1997). Results show some characteristics of the responses, showing the influence of internal dissipation.

2. Constitutive Model

The mechanical model is developed in the setting of one-dimensional media. As state variables, one considers the macroscopic strain ε, the transformation strain ε_T, which describes the inelastic strain associated with the volume fraction of martensite present in the material and the hardening variables α_{AM} and α_{MA} related to direct and reverse transformations, respectively. Dependence upon the absolute temperature θ is restricted, in this paper, to the material properties. The stress-strain relation is given by:

$$\sigma = E(\varepsilon - \varepsilon_T), \tag{1}$$

where E is the Young modulus of the material. Two inequality constraints are considered:

$$f_{AM}(\sigma, \varepsilon_T, \alpha_{AM}) = \sigma \frac{\varepsilon_T}{|\varepsilon_T|} - [\sigma_c(\varepsilon_T) + R(\alpha_{AM})] \leq 0 \tag{2}$$

$$f_{MA}(\sigma, \varepsilon_T, \alpha_{MA}) = [\sigma_c(\varepsilon_T) - R(\alpha_{MA})] - \sigma \frac{\varepsilon_T}{|\varepsilon_T|} \leq 0 \tag{3}$$

and define the elastic domain. The center σ_c and the radius R of the elastic domain are expressed, respectively, by:

$$\sigma_c(\varepsilon_T) = \overline{\sigma}_c + h|\varepsilon_T| \quad \text{and} \quad R(\alpha) = \overline{R} + a[1 - \exp(-b\alpha)] \tag{4}$$

where h, a and b are material properties.

The state variables ε_T, α_{AM} and α_{MA} evolve according to the flow rules:

$$\dot{\varepsilon}_T = (\dot{\gamma}_{AM} - \dot{\gamma}_{MA}) \frac{\varepsilon_T}{|\varepsilon_T|} \tag{5}$$

$$\dot{\alpha}_{AM} = \dot{\gamma}_{AM} - z\dot{\gamma}_{MA} \tag{6}$$

$$\dot{\alpha}_{MA} = \dot{\gamma}_{MA} - z\dot{\gamma}_{AM} \tag{7}$$

where $z >> 1$. The terms associated with the parameter z allows the hardening parameters α_{AM} and α_{MA} to recover their values back to zero whenever an inverse transformation takes place. This is a crucial feature for the description of the inner loops reported in the literature. The hardening variables α_{AM} and α_{MA}, together with the consistency parameters $\dot{\gamma}_{AM}$ and $\dot{\gamma}_{MA}$, are subjected to the following constraints:

$$\alpha_{AM} \geq 0, \quad \alpha_{MA} \geq 0, \quad \dot{\gamma}_{AM} \geq 0, \quad \dot{\gamma}_{MA} \geq 0, \quad \dot{\gamma}_{AM} f_{AM} = 0, \quad \dot{\gamma}_{MA} f_{MA} = 0 \tag{8}$$

A detailed description of this constitutive model can be found in Viana & Mamiya (2001).

3. Equations of Motion

Consider a single-degree of freedom pseudoelastic oscillator depicted in Figure 1. It consists of a mass, m, supported by a pseudoelastic element and a linear damper with coefficient c. The system is harmonically excited by a force $F = F_0 \sin(\omega t)$.

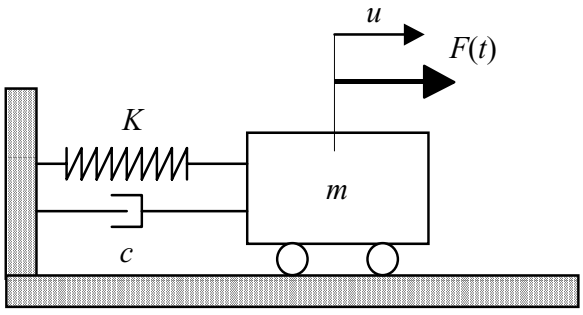


Figure 1. Single-degree of freedom oscillator.

Pseudoelastic behavior is described considering the constitutive model presented in the preceding section. Therefore, the restoring force is given by,

$$K = K(u,T) = \sigma A = EA(\varepsilon - \varepsilon_T) \tag{9}$$

where A is the cross-section area of the pseudoelastic element. By establishing the equilibrium of the system, equations of motion are written as follows

$$m\ddot{u} + c\dot{u} + K(u,T) = F_0 \operatorname{sen}(\omega t) \tag{10}$$

The non-dimensional form of equation (10) is given by:

$$U'' + \xi U' + \mu(U - \varepsilon_T) = \delta \operatorname{sen}(\Omega\tau) \tag{11}$$

where

$$U = \varepsilon = \frac{u}{L};\ \xi = \frac{c}{m\omega_0};\ \delta = \frac{F_0}{mL\omega_0^2};\ \Omega = \frac{\omega}{\omega_0};\ \omega_0^2 = \frac{E_R A}{mL};\ \mu = \frac{E}{E_R};\ \tau = \omega_0 t;\ (\,)' = d(\,)/d\tau \tag{12}$$

The constant E_R is an elastic modulus evaluated at a reference temperature. Hence, the parameter $\mu = \mu(\theta)$ defines a relation between elastic modulus in different temperatures.

The equation of motion may be rewritten in terms of a first order system, assuming $x = U$ and $y = U'$. Hence,

$$\begin{aligned} x' &= y \\ y' &= \delta \operatorname{sen}(\Omega\tau) - \xi y - \mu(x - \varepsilon_T) \end{aligned} \tag{13}$$

where the evolution of variable ε_T is defined by the flow rule (5).

4. Numerical Procedure

In order to solve the governing equations an algorithm based on the operator split technique is employed (Ortiz *et al.,* 1983), associated with an iterative procedure to assure the convergence of the process (Savi & Pacheco, 1997). In the first step of the procedure, equation (13) is integrated employing any classical scheme, like fourth order Runge-Kutta method, considering that variable ε_T is a known parameter which assumes the value from the previous time instant. The next step of solution procedure consists on a corrector step where the feasibility of this trial state is evaluated employing the algorithm to solve the constitutive equation (Viana & Mamiya, 2001). In order to assure the convergence of the process, these steps must be repeated until the values converge for two consecutive iterations.

Numerical simulations are performed considering a pseudolastic oscillator with a constant temperature θ = 373K. Therefore, μ = 1 for all simulations. The material properties were set as linear functions of the absolute temperature: $E = 5.075\theta - 1297$ (MPa), $\overline{\sigma}_c = 8\theta - 2440$ (MPa), $\overline{R} = 1.38\theta - 299$ (MPa), $h = 0.375\theta - 112.4$ (MPa), with exception of a and b, which were set to be zero.

5. Free Vibration

In this Section, the free response of the pseudoelastic oscillator is considered. This is done by letting δ vanish in the equations of motion (13). In order to illustrate the free response of the oscillator, a system without external dissipation is considered letting ξ = 0. Figure 2 presents the free-response of the system. Figure 2a presents the phase space showing an orbit related to an oscillatory motion around an equilibrium point. The response of the system presents a dissipative behavior, which tends to stabilize in a limit cycle. This dissipative behavior is associated with phase transformation, whose evolution can be observed in Figure 2b. Each cycle is related to a hysteresis loop in the stress-strain curve that promote energy dissipation (Figure 2c). This dissipation continues to exist until an elastic steady state is reached.

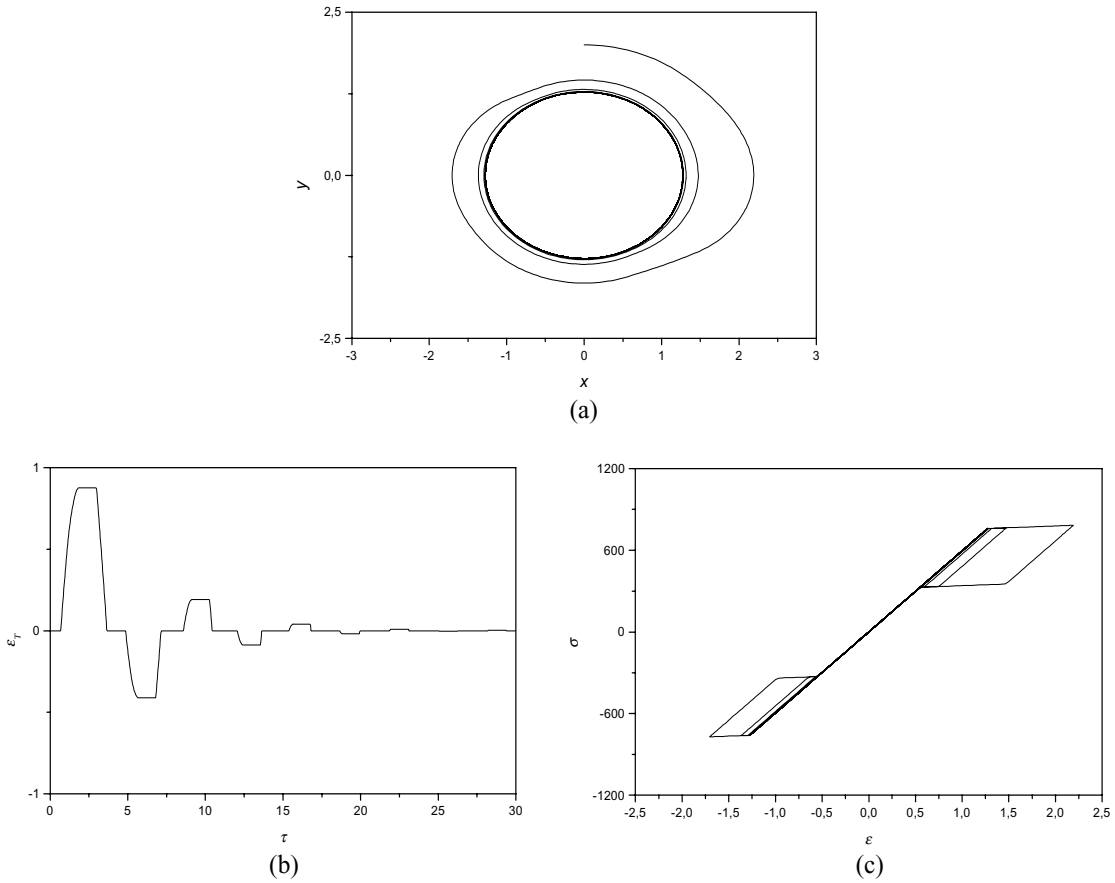


Figure 2. Free vibration response. (a) Phase space; (b) Evolution of ε_T variable; (c) Stress-strain curve.

6. Forced Vibrations

In this section, the forced vibration of the pseudoelastic oscillator without external dissipation is focused. At first, consider harmonic excitations defined by $\delta = 2$ and $\Omega = 1$. This means that the forcing frequency ω is equal to the natural parameter ω_0. Under this condition, the system presents a transient response, dissipating energy because of the hysteretic behavior (Figure 3). After this transient, a steady state behavior is reached presenting a period-1 response, also related to a hysteretic behavior (Figure 4).

At this point, consider harmonic excitations defined by $\delta = 2$ and $\Omega = 2$. This means that the forcing frequency ω is two times greater than the natural parameter ω_0. Once again, the system presents a transient response, dissipating energy because of the hysteretic behavior (Figure 5). After this transient, a steady state behavior is reached. In contrast with the preceding example, the system presents a period-2 response, related to an elastic oscillation (Figure 6).

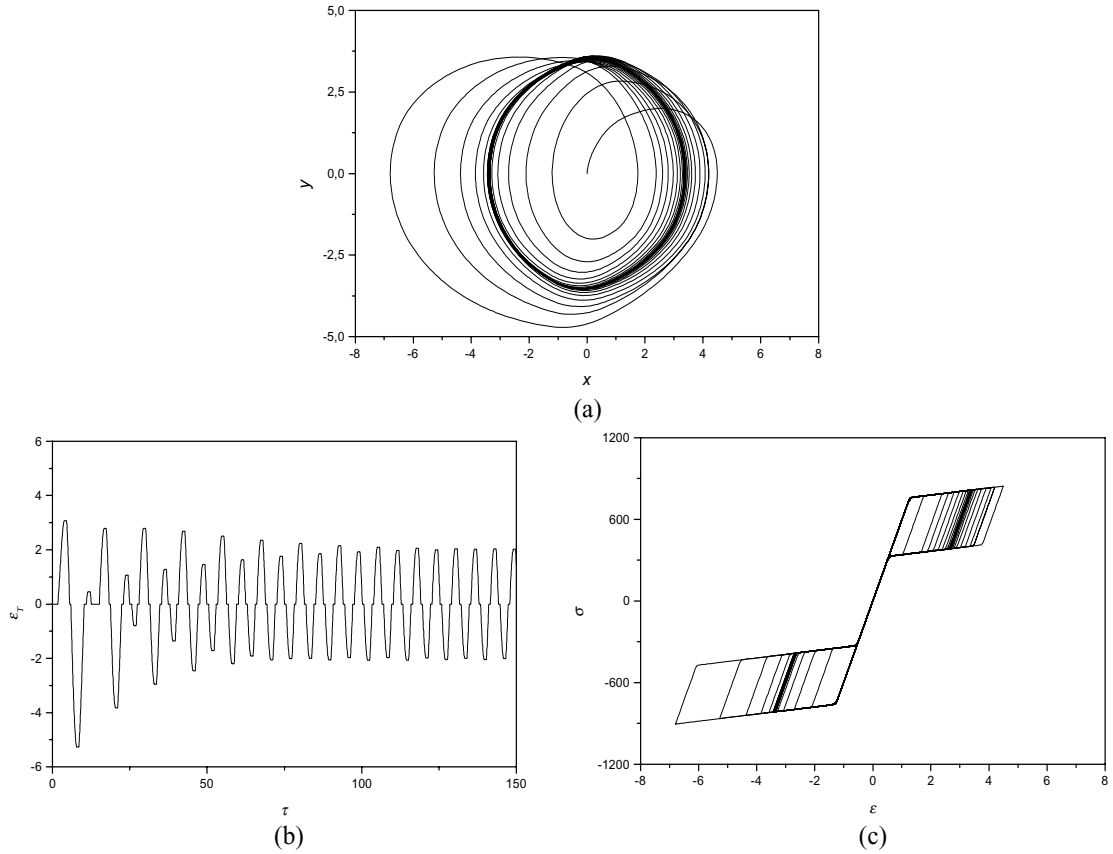


Figure 3. Forced vibration response, $\delta = 2$, $\Omega = 1$. (a) Phase space; (b) Evolution of ε_T variable; (c) Stress-strain curve.

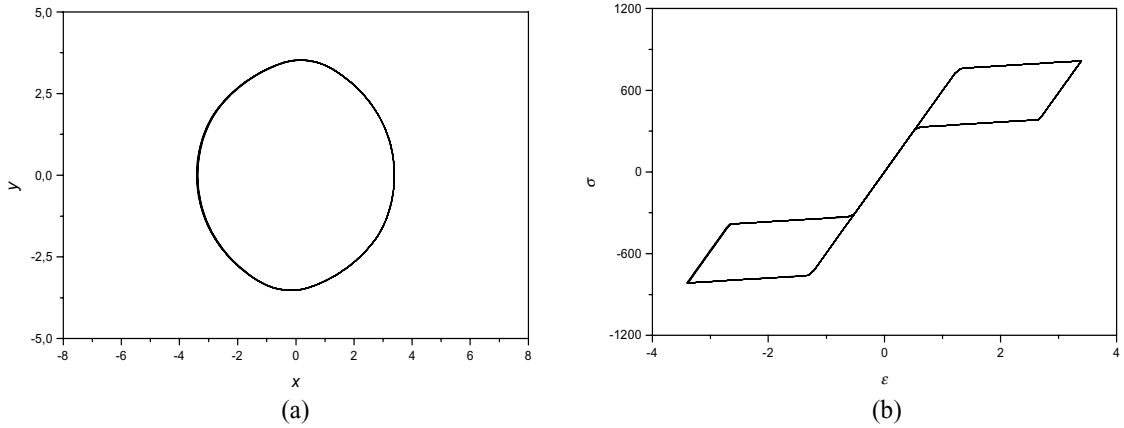


Figure 4. Steady state of forced vibration response, $\delta = 2$, $\Omega = 1$. (a) Phase space; (b) Evolution of ε_T variable; (c) Stress-strain curve.

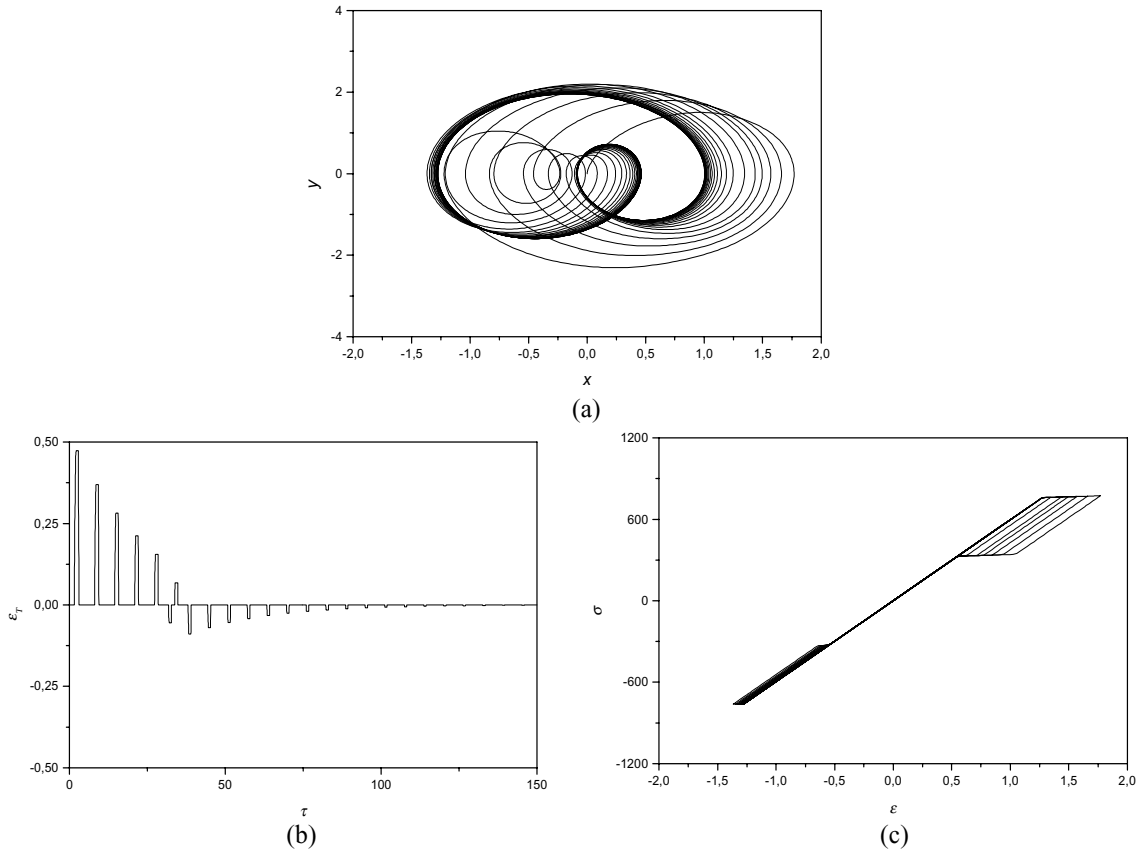


Figure 5. Forced vibration response, $\delta = 2$, $\Omega = 2$. (a) Phase space; (b) Evolution of ε_T variable; (c) Stress-strain curve.

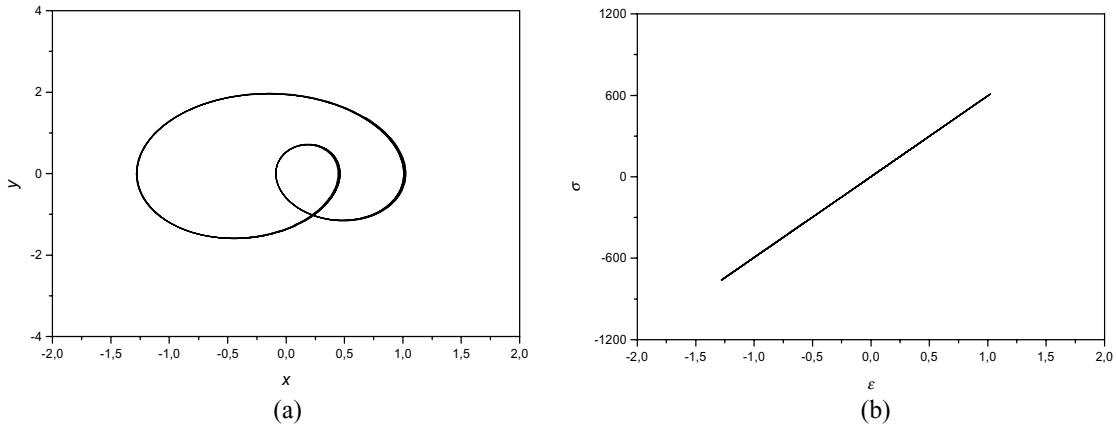


Figure 6. Steady state of forced vibration response, $\delta = 2$, $\Omega = 2$. (a) Phase space; (b) Evolution of ε_T variable; (c) Stress-strain curve.

A system with external dissipation ($\xi = 0.1$) and excitation parameters, $\delta = 2$ and $\Omega = 2$, is now focused. Under this new condition, the external dissipation tends to accelerate the transient response. The steady state is now related to a period-1 behavior. Figure 7a-b shows the phase space, presenting all the motion and the steady state response, respectively. Figure 7c shows the evolution of the variable ε_T, while Figure 7d presents the stress-strain curve.

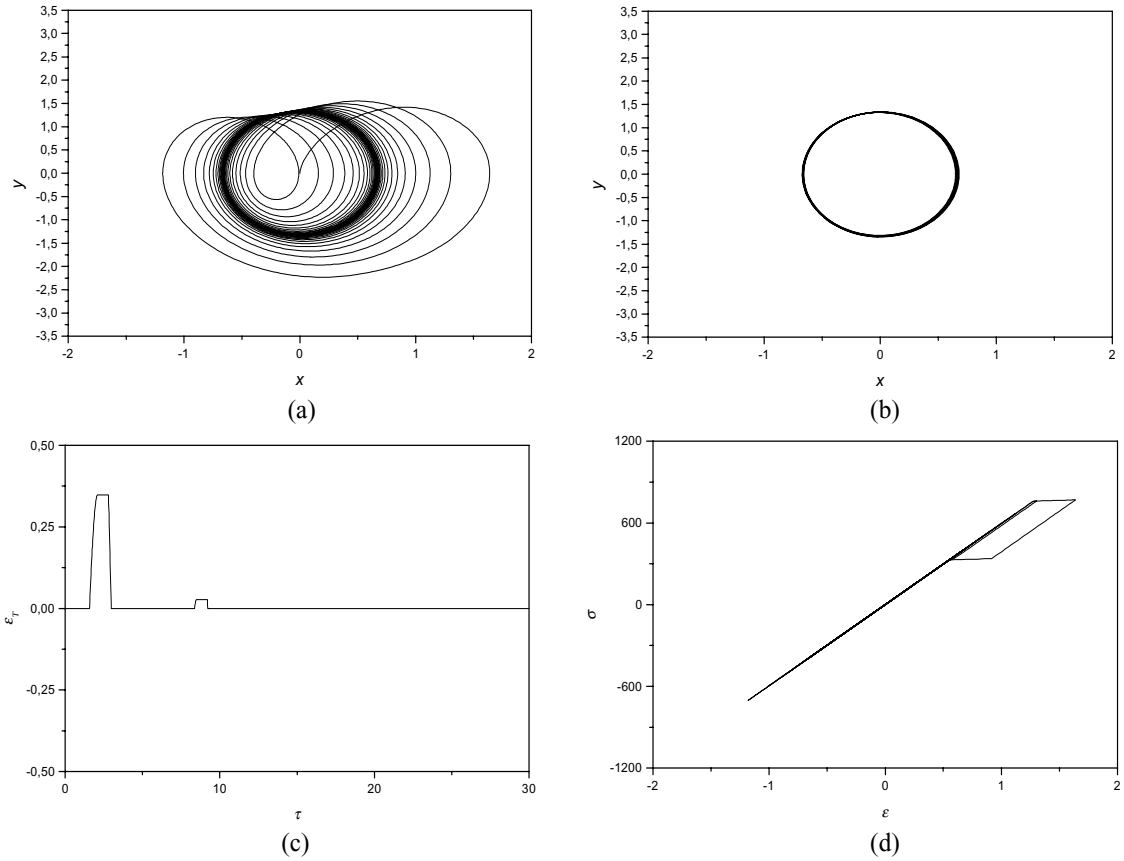


Figure 7. Forced vibration response, $\delta = 2$, $\Omega = 2$, $\xi = 0.1$. (a) Phase space; (b) Steady state; (c) Evolution of ε_T variable; (d) Stress-strain curve.

Notice that, even when no external dissipation is considered, the system dissipates energy because of hysteretic characteristic of the response. This internal dissipation may be useful on passive vibration control. Under this situation, the system may dissipate more energy for higher amplitudes of the response, avoiding undesirable situations. It should be pointed out the capacity of the proposed model to describe this behavior that is associated with internal loops observed during cyclic loads associated with incomplete phase transformations.

7. Conclusions

Numerical investigation of the nonlinear dynamics of a single-degree of freedom pseudoelastic oscillator is studied using a constitutive model based on elastoplasticity to describe the restitution force. A numerical procedure based on the operator split technique associated with an iterative procedure to assure the convergence of the process is proposed. Numerical simulations show some important dynamic characteristics of the system, showing the influence of internal dissipation. Furthermore, results show that the model is capable to capture complex behavior as internal loops observed during cyclic loads associated with incomplete phase transformations. This behavior may be exploited in passive vibration control.

8. Acknowledgments

The authors acknowledge the support of the CNPq and FAPERJ.

9. References

Birman, V., 1997, "Review of Mechanics of Shape Memory Alloy Structures", *Applied Mechanics Review*, v. 50**,** pp. 629-645.

Duerig, T., Pelton, A. & Stöckel, D., 1999, "An Overview of Nitinol Medical Applications", *Materials Science and Engineering A*, v.273-275, pp.149-160.

Ortiz, M., Pinsky, P.M. & Taylor, R.L., 1983, "Operator Split Methods for the Numerical Solution of the Elastoplastic Dynamic Problem", *Computer Methods in Applied Mechanics and Engineering*, v.39, pp.137-157.

Pietrzakowski, M., 2000, "Natural Frequency Modification of Thermally Activated Composite Plates", *Mécanique & Industries*, v.1, n.3, pp.313-320.

Rogers, C.A., 1990, "Active Vibration and Structural Acoustic Control of Shape Memory Alloy Hybrid Composites: Experimental Results", *Journal of Acoustical Society of America,* v.88, pp.2803-2811.

Rogers, C.A., Liang, C. & Fuller, C.R., 1991, "Modeling of Shape Memory Alloy Hybrid Composites for Structural Acoustic Control" , *Journal of Acoustical Society of America,* v.89, pp.210-220.

Rogers, C.A., 1995, "Intelligent Materials", *Scientific American*, September, pp.122-127.

Savi, M.A. & Pacheco, P.M.C.L., 1997, "Non-Linear Dynamics of an Elasto-Plastic Oscillator with Kinematic and Isotropic Hardening", *Journal of Sound and Vibration*, v.207(2), pp.207-226.

Shaw, J. A. & Kyriakides, S., 1995, "Thermomechanical Aspects of NiTi", *Journal of Mechanics Physics Solids*, v.43, pp.1243-1281.

van Humbeeck, J., 1999, "Non-medical Applications of Shape Memory Alloys", *Materials Science and Engineering A*, v.273-275, pp.134-148.

Viana, D. M. & Mamiya, E. N., 2001, "A Phenomenological Model for Shape Memory Alloys Inherited from Classical Plasticity", Submitted to *COBEM 2001 - XVI Congresso Brasileiro de Engenharia Mecânica*.

XVI CONGRESSO BRASILEIRO DE ENGENHARIA MECÂNICA
16th BRAZILIAN CONGRESS OF MECHANICAL ENGINEERING

A NEW TREATMENT FOR RIGID BODIES COLLISION MODELS

Edson Cataldo
Universidade Federal Fluminense
Departamento de Matemática Aplicada, Centro, Niterói, Brasil
Rua Mário Santos Braga, s/ no
24020-140 – Centro – Niterói – RJ – Brasil
ecataldo@mec.puc-rio.br

Rubens Sampaio
Pontifícia Universidade Católica do Rio de Janeiro
Departamento de Engenharia Mecânica
Rua Marquês de São Vicente, 225
22453-900 – Gávea – Rio de Janeiro – Brasil
rsampaio@mec.puc-rio.br

Abstract. *In general the motion of a body takes place in a confined environment and collisions of the body with the containing wall are possible. In order to predict the dynamics of a body in these conditions one must know what happens in a collision. Therefore, the problem is: if one knows the pre-collision dynamics of the body and the properties of the body and the wall one wants to predict the post-collision dynamics. This problem is quite old and it appeared in the literature in 1668. Up to 1984 it seemed that Newton's model was enough to solve the problem. But it was found that this was not the case and a renewed interest in the problem appeared. The aim of this paper is to treat the problem of plan collsions of rigid bodies, to classify the different models found in the literature and to present a new model, called C-S model, that is a generalization of most of these models.*

Keywords: *Colllisions, Dynamics, Modelling, Simulation, Rigid Bodies*

1. Introduction

Form the simplest observation, we can say that the dynamics of a body, or of a system with more than one particle, can be modeled properly if collisions are taken into account. In the works of Galileu and Descartes there are references to the collision between particles, but the first model of this problem seems to be due to John Wallis and Christopher Wren, independently, in 1668. Some great scientists such as Newton, Huygens, Coriolis, Darboux, Routh, Apple, Carnot and Poisson have also treated the problem. At the beginning of this century the problem generated some discussions, as we can see in the works of Painlevé (1905) and Klein (1910). But, up to 1984, all of these works used the theory developed by Newton or by Poisson and the dificulty was to include friction in the modelling, as was pointed out by Painlevé in his paper "Sur les lois de frottement de glissement".
In 1984, Kane (1984) published a work, in a journal with limited circulation, where he pointed out na apparent paradox: the application of Newton's theory with Coulomb's friction, universally accepted, in a problem of collisions of a double pendulum, conducted to generation of energy. What was wrong ?
In 1986, Keller (1986) presented a solution to Kane's paradox, but the solution was not easy to generalize. Keller's work was published in a journal with large circulation and arose widespread interest. In these thirteen years the interest has increased and there are some books totally dedicated to this topic, as the ones written by Glocker-Pfeiffer (1996), Brach(1991), Brogliato (1996) and Monteiro-Marques (1993).
Brach (1989) presented a model with linear equations containing some nondimensional parameters that characterize the collision and he defined "ratio between impulses" instead of coefficient of friction. However, his consideration did not give clear solutions to the problem when one considers *reverse sliding* during the collision. Stronge (1990) suggested a coefficient of restitution relating the energy during the compression phase to the energy during the expansion phase. Smith (1991,1992) presented a model with nonlinear equations. Wang-Mason (1992) applied the Routh's technique (1877) and compared the coefficients of restitution given by Newton and by Poisson. Sabine Durand (1996) studied the dynamics of systems with unilateral restirctions and included some ystems related to the collisions. Chatterjee (1997) presented new laws based in simple algorithms. He has not used many parameters and he obtained good results. Stoianovici and Hermuzlu (1996) have shown the limits of validity of some rigid bodies collision models. As their main interest was in Robotics, they focused in collisions of slender bodies at low velocities. Cathérine Cholet (1998) developed a new theory of rigid bodies collisions that satisfies the Principles of the Mechanics. Her work was based in he ideas introduced by Michel Frémond: a system formed by a set of rigid bodies is deformable because the relative positions between each pair of bodies vary. They discussed the theory and showed that it is coherent from the mathematical point of view and also experimentally validated.

2. Motion equations

The collision is modeled as instantaneous. we consider the generalized position of the system in the instnat t defined by $q = (q_1, q_2, \ldots, q_n)^t$. We consider the contact between two bodies C_1 and C_2 and let R be the force of reaction exerted by C_1 on C_2. Then we write $R = (R_N \; R_T)^t$.
The dynamics of the system is given by the Lagrangean equations:

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}}\right) - \frac{\partial T}{\partial q_i} = Q_i + r_i \tag{1}$$

with Q_i the contribution of the external generalized forces, r_i the generalized force due to the reaction in the contact and T the kinetic energy of the system. We should observe that r_i is only present when there is contact, otherwise it is null.
Considering only a planar situation, we have n parameters of position and two reactions in the contact (R_N and R_T) also unknown. Then, we need, not only the n equations obtained from Lagrange's equations but also two equations more, given by the collision laws that will be discussed later.

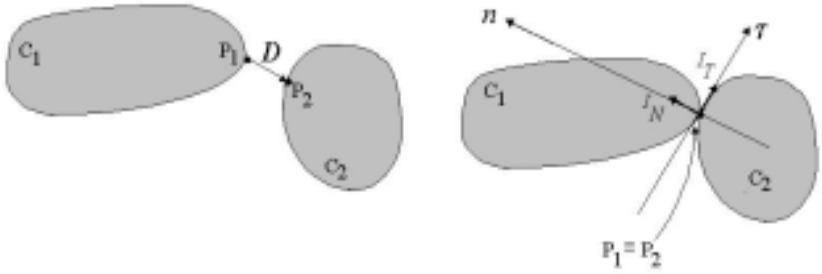


Figure 1: Collision between two bodies

We consider P_1 and P_2 the points of the bodies C_1 and C_2, respectively, that will be in contact in the collision. We denote by $\mathbf{D}$ the vector that represents the relative displacement between the two bodies and by $\dot{\mathbf{D}}$ the vector that represents the relative velocity between the bodies, as shown in the Fig. (1).

In the point of contact we represent the impulses in the normal and tangential directions by I_N and I_T. We use and $\mathbf{u}_n$ and $\mathbf{u}_\tau$, the unitary vectors of the normal direction (given by n) and tangential direction (given by τ), in a frame which we will call collision frame, shown in Fig. (1).

Evaluating the relative velocity between the contact points we have,

$$\dot{D} = \sum_{i=1}^{n} \frac{\partial P_2}{\partial q_i}\dot{q}_i + \frac{\partial P_2}{\partial t} - \sum_{i=1}^{n} \frac{\partial P_1}{\partial q_i}\dot{q}_i - \frac{\partial P_1}{\partial t} = \sum_{i=1}^{n}\left(\frac{\partial P_2}{\partial q_i} - \frac{\partial P_1}{\partial q_i}\right)\dot{q}_i + \frac{\partial P_2}{\partial t} - \frac{\partial P_1}{\partial t}$$

We use the notations,

$$W_T^i = \left(\frac{\partial P_2}{\partial q_i} - \frac{\partial P_1}{\partial q_i}\right).\mathbf{u}_\tau \quad , \quad W_N^i = \left(\frac{\partial P_2}{\partial q_i} - \frac{\partial P_1}{\partial q_i}\right).\mathbf{u}_n \; , \quad \tilde{\omega}_T = \left(\frac{\partial P_2}{\partial q_i} - \frac{\partial P_1}{\partial q_i}\right).\mathbf{u}_\tau \quad , \quad \tilde{\omega}_N = \left(\frac{\partial P_2}{\partial q_i} - \frac{\partial P_1}{\partial q_i}\right).\mathbf{u}_n$$

We consider, then, $\mathbf{W}_T$ the column vector in which the components are W_T^i and $\mathbf{W}_N$ the column vector in which the components are W_N^i.
We can write the normal ($\dot{D}_N$) and tangential ($\dot{D}_T$) components of $\dot{D}$ as $\dot{D}_N = \mathbf{W}_N^t\,\dot{q} + \tilde{\omega}_N$ and $\dot{D}_T = \mathbf{W}_N^t\,\dot{q} + \tilde{\omega}_T$.
Or we can write $\dot{\mathbf{D}} = [W]^t + \tilde{\omega}$, with

$\dot{\mathbf{D}} = \begin{pmatrix} \dot{D}_N \\ \dot{D}_T \end{pmatrix}$, $[W]^t = \begin{pmatrix} \mathbf{W}_N^t \\ \mathbf{W}_T^t \end{pmatrix}$ a matrix (2,n) and $\tilde{\boldsymbol{\omega}} = \begin{pmatrix} \tilde{\omega}_N \\ \tilde{\omega}_T \end{pmatrix}$.

The generalized force $\mathbf{r}$ can be written as

$$\mathbf{r} = \begin{pmatrix} \mathbf{W}_N & \mathbf{W}_T \end{pmatrix} \begin{pmatrix} R_N \\ R_T \end{pmatrix} \quad \text{or} \quad \mathbf{r} = [W]\mathbf{R}.$$

Integrating Eq. (1) in the interval $(t-\varepsilon, t+\varepsilon)$, with t the instant of collision and after some algebric manipulations we have the following equation

$$[M](\dot{\mathbf{q}}_E - \dot{\mathbf{q}}_A) = [W]\mathbf{I} = \begin{pmatrix} \mathbf{W}_N & \mathbf{W}_T \end{pmatrix} \begin{pmatrix} I_N \\ I_T \end{pmatrix}. \tag{2}$$

Our problem is to find $\dot{\mathbf{q}}_E$ and $\mathbf{I}$ given $[M]$, $[W]$ and $\dot{\mathbf{q}}_A$. Then, there are n equations and we want to find n+2 unknowns. Therefore, we need two more equations. These two equations are given by the restitution laws discussed later.

In some cases we can consider also na impulse of moment denoted by $\mathbf{I}_\theta$. In this case, the equation will be given by

$$[M](\dot{\mathbf{q}}_E - \dot{\mathbf{q}}_A) = [W]\mathbf{I} = \begin{pmatrix} \mathbf{W}_N & \mathbf{W}_T & \mathbf{W}_\theta \end{pmatrix} \begin{pmatrix} I_N \\ I_T \\ I_\theta \end{pmatrix}. \tag{3}$$

In this case, we have n equations to find n+3 unknowns. We need three more equations. These three equations will be given by the restitution laws.
We construct a collision model when we join the n equations that describe the motion of the system with the equations given by the restitution laws.
In order to solve the problem we use a strategy that consists in defining a process called *virtual process*. It is not related to time. We show a scheme in the Fig. (2) to ilustrate this idea.

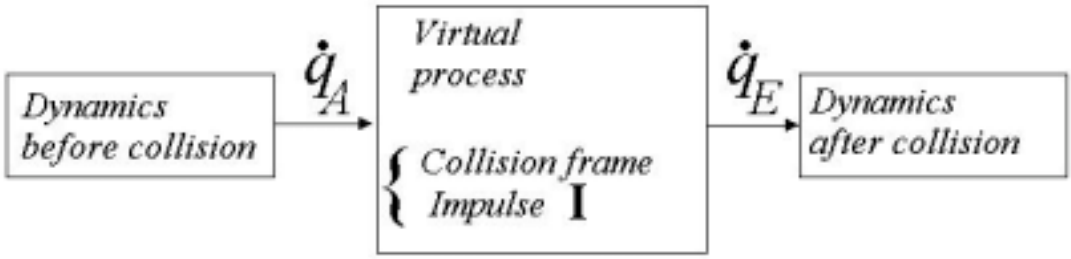


Figure 2: Virtual process sheme.

2.The local matrix mass

Instead of writing the equations in terms of $\dot{\mathbf{q}}$ we can use $\dot{\mathbf{D}}$. The vector $\mathbf{D}$ was shown in Fig. (1) and it is important because it monitors when the collision occurs.
We can write

$$\dot{\mathbf{D}} = [W]^t\dot{\mathbf{q}} + \tilde{\boldsymbol{\omega}} = \begin{pmatrix} W_N^t \\ W_T^t \end{pmatrix}\dot{\mathbf{q}} + \begin{pmatrix} \tilde{\omega}_N \\ \tilde{\omega}_T \end{pmatrix}. \tag{4}$$

Then,

$$\dot{\mathbf{D}}_E - \dot{\mathbf{D}}_A = [W]^t[M]^{-1}[W]\mathbf{I} = [\tilde{M}_L]\mathbf{I} \Rightarrow \mathbf{I} = [M_L](\dot{\mathbf{D}}_E - \dot{\mathbf{D}}_A).$$

when $[\widetilde{M}_L]$ is invertible and $[\widetilde{M}_L]^{-1} = [M_L]$. We call $[M_L]$ the local matrix mass.

3. Compression phase and expansion phase

In order to describe some of the collision models we will think, formally, that the change between the pre-collision velocity to the post-collision velocity occurs in two phases: the compression phase and the expansion phase. The virtual process will be composed by these two phases as it is shown schematically in Fig. (3).

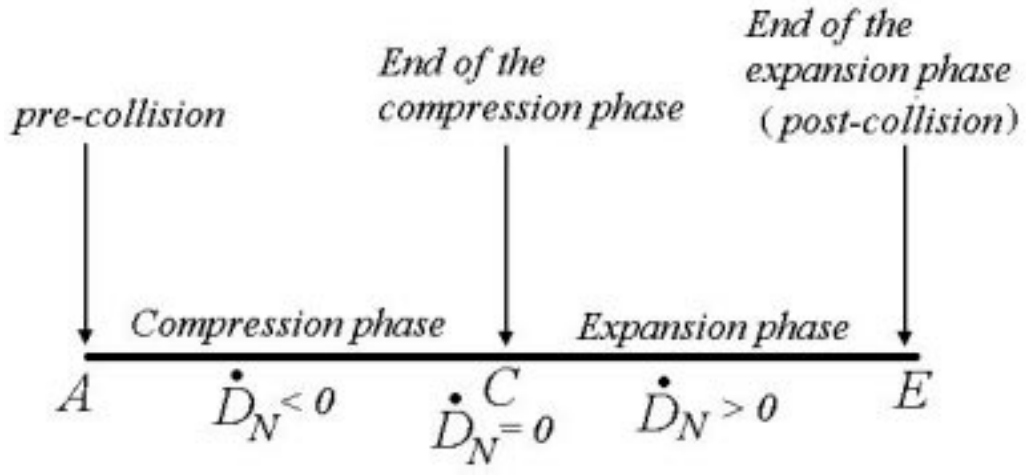


Figure 3: Compression phase and expansion phase.

4.A new collision model: the C-S model

We present a new collision model that tries to generalize some of the models from the literature and it also can predict some behavior that those models cannot.

The equations used are given in the following:

In the compression phase:

$$[M](\dot{\mathbf{q}}_C \quad \dot{\mathbf{q}}_A) = (\mathbf{W}_T \quad \mathbf{W}_N \quad \mathbf{W}_\theta)\begin{pmatrix} I_{NC} \\ I_{TC} \\ I_{\theta C} \end{pmatrix} \tag{5}$$

$$\begin{pmatrix} \dot{D}_{NC} \\ \dot{D}_{TC} \\ \dot{D}_{\theta C} \end{pmatrix} = \begin{pmatrix} \mathbf{W}_N^t \\ \mathbf{W}_T^t \\ \mathbf{W}_\theta^t \end{pmatrix}(\dot{\mathbf{q}}_C \quad \dot{\mathbf{q}}_A) + \begin{pmatrix} \dot{D}_{NA} \\ \dot{D}_{TA} \\ \dot{D}_{\theta A} \end{pmatrix} \tag{6}$$

To the restitution in the tangential direction we use the Coulomb's law in the form

$$\begin{cases} |I_{TC}| < \mu I_{NC} \Rightarrow \dot{D}_{TC} = 0 \\ I_{TC} = \mu I_{NC} \Rightarrow \dot{D}_{TC} \le 0 \\ I_{TC} = -\mu I_{NC} \Rightarrow \dot{D}_{TC} \ge 0 \end{cases} \tag{7}$$

We use the coefficient of moment in the compression phase given by

$$e_{mC} I_{\theta C} = -(1 + e_{mC}) \bar{J} \dot{D}_{\theta C}$$

(8)

$$\bar{J} = \frac{J_1 J_2}{J_1 + J_2}$$

(9)

with J_1 and J_2 the moment of inertia related to the center of mass of each body.

In the expansion phae:

$$[M](\dot{\mathbf{q}}_E \quad \dot{\mathbf{q}}_C) = \begin{pmatrix} \mathbf{W}_T & \mathbf{W}_N & \mathbf{W}_\theta \end{pmatrix} \begin{pmatrix} I_{NE} \\ I_{TE} \\ I_{\theta E} \end{pmatrix}$$

(10)

$$\begin{pmatrix} \dot{D}_{NE} \\ \dot{D}_{TE} \\ \dot{D}_{\theta E} \end{pmatrix} = \begin{pmatrix} \mathbf{W}_N^t \\ \mathbf{W}_T^t \\ \mathbf{W}_\theta^t \end{pmatrix} (\dot{\mathbf{q}}_E \quad \dot{\mathbf{q}}_C) + \begin{pmatrix} \dot{D}_{NC} \\ \dot{D}_{TC} \\ \dot{D}_{\theta C} \end{pmatrix}$$

(11)

We use the coefficient of moment in the compression phase given by

$$e_{mE} I_{\theta E} = -(1 + e_{mE}) \bar{J} \dot{D}_{\theta E} \quad ,$$

(12)

$$\bar{J} = \frac{J_1 J_2}{J_1 + J_2}$$

(13)

To the tangential restitution we consider

$$I_{TS} = \frac{1}{2}[\mu \nu I_{NE} \operatorname{sign}(I_{TC}) + e_n e_t I_{TC}] \ , \ 0 \le \nu \ , \ e_t \le 1$$

(14)

and then we use

If $I_{TC} \ge 0$, $I_{TS} \ge 0 \Rightarrow -\mu I_{NE} + 2I_{TS} \le I_{TE} \le \mu I_{NE}$

$$\begin{cases} -\mu I_{NE} + 2I_{TS} < I_{TE} < \mu I_{NE} \Rightarrow \dot{D}_{TE} = 0 \\ I_{TE} = +\mu I_{NE} \Rightarrow \dot{D}_{TE} \le 0 \\ I_{TE} = -\mu I_{NE} + 2I_{Ts} \Rightarrow \dot{D}_{TE} \ge 0 \end{cases}$$

If $I_{TC} \le 0$, $I_{TS} \le 0 \Rightarrow -\mu I_{NE} \le I_{TE} \le \mu I_{NE} + 2I_{TS}$

To the normal restitution we use the Poisson's coefficient given by $e_{np} = \frac{I_{NE}}{I_{NC}}$.

I_{NC} is the normal impulse in the of the compression phase and I_{NE} is the normal impulse in the end of the expansion phase.

4.Particular cases

As we had said, the C-S model generalizes some of the models from the literature. We will describe briefly three of these models and we will show what should be done, in the C-S model, to particularize the respective model.

4.1. Newton's model

This is the simplest model. It considers the coefficient of restitution given by Newton; that is, $e_n = -\frac{\dot{D}_{NE}}{\dot{D}_{NA}}$.

It does not consider friction (there will be notangential restitution) and it does not consider the impulse of moment. If in the C-S model we consider $e_{mC} = e_{mE} = -1$, $e_t = \nu = 0$ and $\mu = 0$ then we will obtain the Newton's model. It is important to observe that when we do not consider friction the Newton's coefficient of restitution and the Poisson's coefficient are equivalent. It can be seen in Cataldo and Sampaio(1999,2000).

4.2. Wang-Mason's model (considering the Poisson's coefficient of restitution)

This model uses the coefficient of restitution in the normal direction given by Poisson; that is, $e_{np} = \frac{I_{NE}}{I_{NC}}$. It considers friction given by Coulomb's law (modified, as we have presented) and it does not consider the impulse of moment.. If in the C-S model we consider $e_{mC} = e_{mE} = -1$ and $e_t = \nu = 0$ then we will obtain Wang-Mason's model.

4.3. Glocker-Pfeiffer's model

This model uses the coefficient of restitution in the normal direction given by Poisson. It considers friction and also the possible reversible portions of the tangential impulse. It does not consider the impulse of moment. If in the C-S model we consider $e_{mC} = e_{mE} = -1$ then we will obtain the Glocker-Pfeiffer's model.

4. New results using the C-S model

Using the C-S model we can observe some behaviors that couldn't be described using other models. As na example we consider the problem of a ball coliding with two barriers as shown in the Fig. (4).

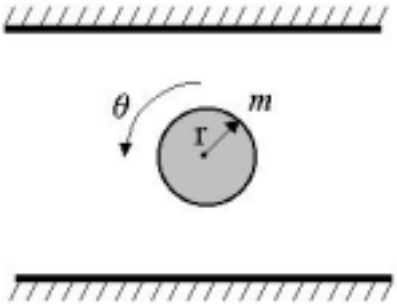


Figure 4: Collision of a ball with two barriers.

We consider the following values to the parameters and inital conditions: mass of the ball=1kg , $\theta_0 = 0$, $\dot{\theta}_0 = 0$, $x_0 = 0$, $\dot{x}_0 = 1$, $y_0 = 0.9\,\text{m}$, $\dot{y}_0 = -1\,\text{m/s}$, $e_{np} = 1$, distance between the barriers=1.01m, $\mu = 1$ and $r = 0.1$. If we consider $e_{mC} = -1$, $e_{mE} = -1$, $e_t = 0$ and $\nu = 0$ we obtain the same prediction obtained from Wang-Mason's model or Glocker-Pfeiffer's model. It is the behavior of a ball used, for example, in a table tennis game as shown in Fig. (5).

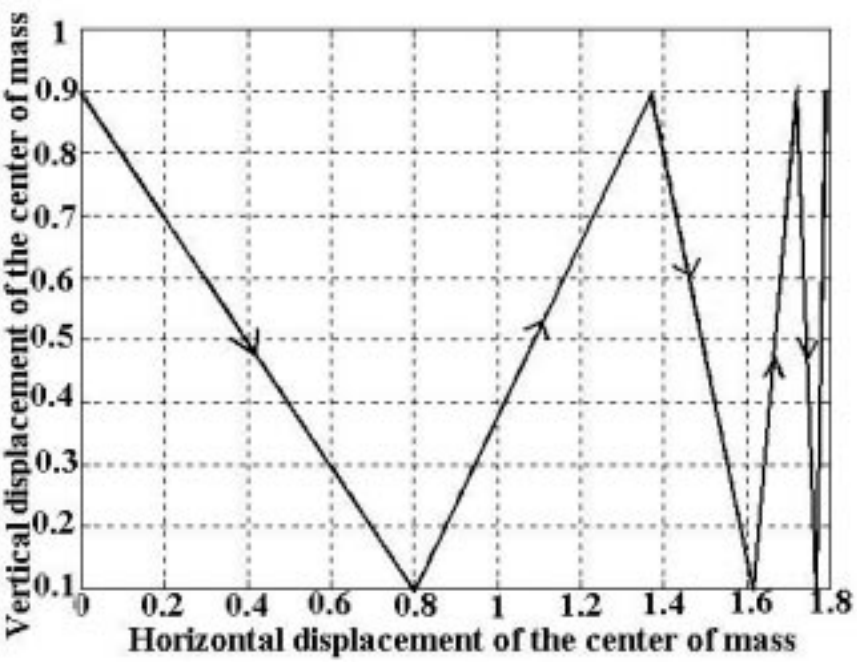


Figure 5: The C-S model using $e_{mC} = e_{mE} = -1$, $e_t = 0$ and $\nu = 0$. Units (m).

If we consider $e_{mC} = e_{mE} = -1$, $e_t = 1$ and $\nu = 1$ we obtain the same prediction obtained from Glocker-Pfeiffer's model in the case called superball-like behavior. It is the behavior of a ball made of steel and not hollow. Its trajectory is shown in the Fig. (6).

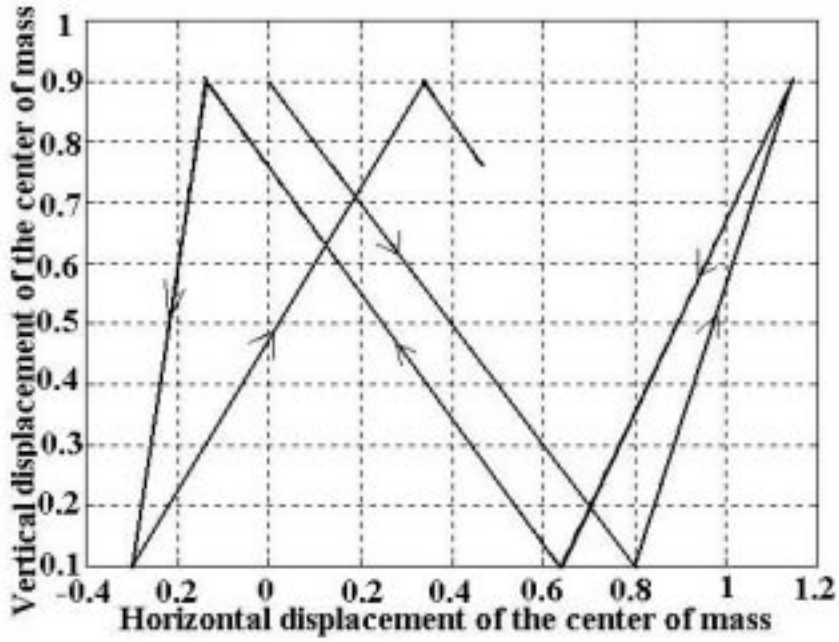


Figure 6: The C-S model using $e_{mC} = e_{mE} = -1$, $e_t = 1$ and $\nu = 1$. Units (m).

If we consider $e_{mC} = -1$, $e_{mE} = -0.5$, $e_t = 1$ and $\nu = 1$ we obtain a new behavior that couldn't be observed if we had used other models. We show the trajectory of the center of mass in the Fig. (7).

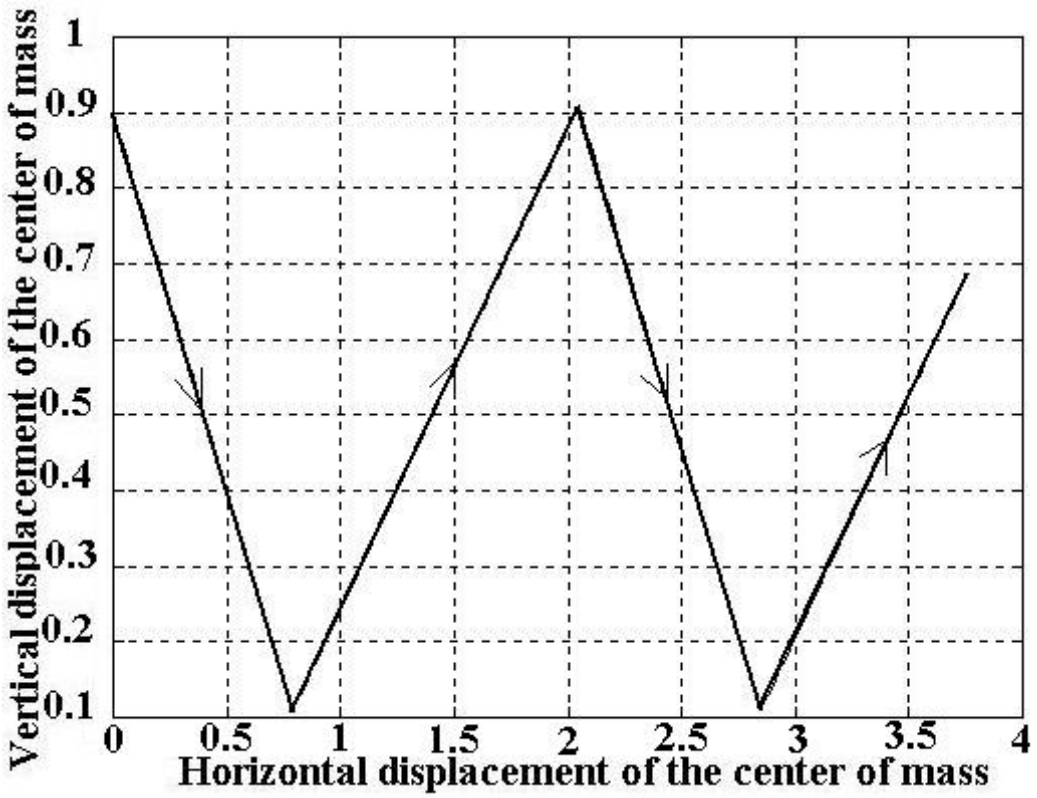


Figure 7: The C-S model using $e_{mC} = e_{mE} = -1$, $e_t = 1$ and $\nu = 1$. Units (m).

5. Conclusions

We studied rigid body collisions considering that these collisions are instantaneous. After making a systematic study of some rigid body collisions models we could formulate a new model: the C-S model. Using this model we could show some comparisons between models and we could present some behaviors that couldn't be obtained using other models. We showed simulations and animations using the C-S model in a way that would make us understand what was happening.

6. References

Brach, R. M., 1991, "Mechanical Impact Dynamics - Rigid Bodies Collisions", John Wiley \& Sons, New York, 260 p.
Brogliato, B., 1996, " Nonsmooth Impact Mechanics: Models, Dynamics and Control", Lectures Notes in Control and Information Sciences, Berlin, Springer.
Cataldo, E., 1999, "Simulation and modeling of Plan Rigid Bodies Collisions" (In Portuguese), Ph.D. Thesis, Pontificia Universidade Católica do Rio de Janeiro (PUC-Rio), RJ, Brazil, 295 p.
Cataldo, E. and Sampaio, R., 1999, " Comparing some Models of Collisions between Rigid Bodies ", Proceedings of PACAM VI/DINAME, Vol. 8, pp. 1301-1304.
Cataldo, E. and Sampaio, R., 1999, "Comparación entre Modelos de Colisión de Cuerpos", to appear in Revista Internacional de Métodos Numéricos y Diseño en Ingenieria.
Cataldo, E. and Sampaio, R., 2000, "Comparação entre previsões de alguns modelos de choque", Proceeding of CONEM2000, Natal, Rio Grande do Norte, Brazil. Work JC9812.
Cataldo, E. and Sampaio, R., 2000, "A new model of rigid bodies collision: the C-S model", to appear in Computational and Applied mathematics.
Chatterjee, A., 1997, "Rigid body collisions: some general considerations, new collision laws, and some experimental data", Ph.D. Thesis, Cornell University.
Cholet, C., 1997, "Chocs de solides rigides", Ph. D. Thesis in Maths (In French), Université Paris VI, Paris, France.
Kane, T. R., 1984, " A Dynamic Puzzle ", Stanford Mechanics Alumni Club Newsletter, pp. 6.
Keller, J. B., 1986, " Impact with Friction", ASME Journal of Applied Mechanics, Vol. 53, pp. 1-4.
Klein, F., 1910, " Zu Painlevés kritik des Coulombschenreibungsgestze", Zeitsch Math. Phys., Vol. 58, pp. 186-191.
Mason, M. T. and Wang, Y., 1992, "Two-Dimensional Rigid-Body Collisions with Friction", ASME Journal of Applied Mechanics, vol. 59, pp. 635-642.
Monteiro Marques, M. D. P., 1993, "Differential inclusions in nonsmooth mechanical problems. Shocks and dry friction" , Birkhauser, Boston.
Painlevé, P., 1905, "Sur Les Lois de Frottement de Glissement", C. R. Acad. Sci. Paris, Vol. 121, pp. 112-115, Vol. 141, pp. 401-405, Vol. 141, pp. 546-552.
Pfeiffer, F. and Glocker, C., 1996, "Multibody Dynamics with Unilateral Contacts", John Wiley \& Sons, New York, 317 p.
Routh, E. J., 1877, "An elementary treatise on the dynamics of a system of rigid bodies", 3rd edition, Mac Millan and Co, London, UK.

Smith, C. E., 1991, "Prediction Rebounds using Rigid-Body Dynamics", ASME Journal of Applied Mechanics, Vol. 58, pp. 754-758.
Smith, C. E. and Liu, P. P., 1992, "Coefficients of restitution", Journal of Applied Mechanics, Vol. 59, p. 963.
Stoianovici, D. and Hurmuzlu, Y., 1996, "A critical study of the applicability of rigid-Body collision theory", Journal of Applied Mechanics, Vol. 63, pp. 307-316.
Stronge, W. J., 1990, "Rigid Body Collisions with Friction", Proceedings of the Royal Society of London A, Vol. 431, pp. 169-181.

XVI CONGRESSO BRASILEIRO DE ENGENHARIA MECÂNICA
16th BRAZILIAN CONGRESS OF MECHANICAL ENGINEERING

TRANSFER FROM THE EARTH TO THE HALO ORBITS WITH MINIMUM FUEL CONSUMPTION

Annelisie Aiex Corrêa
Instituto Nacional de Pesquisas Espaciais –INPE
cp. 515 – 12227 010 - São José dos Campos –SP
Phone: (12) 345 6239 Fax: (12) 345 6228
anne@dem.inpe.br

Antônio Fernando Bertachini de Almeida Prado
Instituto Nacional de Pesquisas Espaciais –INPE
cp. 515 – 12227 010 - São José dos Campos –SP
Phone: (12) 345 6201 Fax: (12) 345 6228
prado@dem.inpe.br

Teresinha de Jesus Stuchi
Instituto de Física - Universidade Federal do Rio de Janeiro UFRJ
CT, bloco A, 68528 – 21945-970, Rio de Janeiro / RJ
tstuchi@if.ufrj.br

Abstract. *The purpose of this work is to study the transfer orbits from the Halo orbits around the collinear lagrangian point L_1 of the Earth-Moon system to a parking orbit close to the Earth. The family of the Halo orbits has been calculated using the numerical continuation method. The transfer orbits were studied in two different ways. The first is based on the two point boundary value problem and the second one uses the natural instability of the Halo orbits, where is enough to apply a small increment of velocity for the vehicle to escape from the orbit. The required ΔV to put this point of mass in an elliptical orbit around the largest primary is calculated assuming a fixed energy for the parking orbit.*

Keywords: *Halo orbits, transfer orbits, Lambert's problem*

1. Introduction

In the restricted three-body problem, one of the masses, m_3, is supposed to be infinitely small compared with the others two masses, called primary bodies, m_1 and m_2. The restriction of this problem lies in the fact that the neglected mass moves under the attraction forces of the two-body without disturbing their motion. In the problem studied here, the primaries are supposed to be moving in circular orbits around their centre of mass. For all the calculations we consider the Earth-Moon system, but it can be easily extended to others systems and the mass m_3 can be understood as a space vehicle.

In this dynamical system there are five equilibrium points called Euler-Lagrangian points. Three of them, L_1, L_2 and L_3, are lying in a straight line which passes by the centre of the primaries and the other two points, L_4 and L_5, form an equilateral triangle with the primaries as their vertices. A satellite placed at one of these points with zero velocity will remain in equilibrium because the gravitational and centripetal accelerations acting on the vehicle cancel each other. Although, the collinear points are not stable for all mass ratio it was shown by Farquhar (1970), it requires a control to maintain the satellite in the vicinity of the libration point.

It is well known that families of unstable periodic orbits exist around the libration points L_1, L_2 and L_3. The planar periodic orbits around these collinear points are called Lyapunov orbits and one of their bifurcation to the tridimensional family are called Halo orbits, see Fig. (1). Our motivation in studying the Halo orbits of the Earth-Moon system has two reasons. The orbits of the collinear equilibrium point between the primaries, L_1, could be used as a parking orbit for any lunar mission whereas the orbits of the lagrangian point on the left side of the Moon, L_2, could be used for observations of the Moon and the Earth simultaneously.

The Halo family orbits around the Lagrangian points L_1 and L_2 have been determinated by the numerical continuation method using the arc parameter which is well described in Gómez et al. (2001). This method requires the knowledge of one solution of a problem in a neighbourhood of the one considered.

We have performed some transfer orbits based on the two point boundary value problem. Although, such formulation does not take the time of transfer into account. When we introduce a time constraint, the two point boundary value problem becomes the well-known Lambert's problem (Prado, 1995) in the restricted three-body problem context. In our case, the set of boundary conditions considered was a point on the Halo orbit as the initial condition and a point close to the Earth as the final condition. The time of flight is fixed for all the transfers involved.

The instability of the Halo orbits provides an alternative of getting more economical transfer orbits than the ones provided by the Lambert's method. When we apply a small impulse to the vehicle it can be easily detached from the

orbit. We call such orbits as quasi-natural transfer orbits. In the last section of this paper we compare the ΔV required by both transfer methods applied.

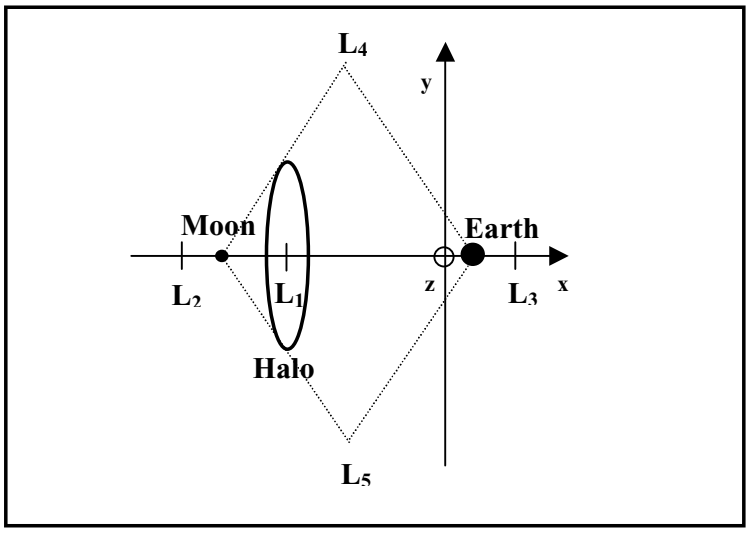


Figure 1. The set-up of the restricted three-body problem

2. Equations of Motion

The potential of the restricted three-body problem in a synodic coordinate system is given by

$$\Omega(x,y,z) = \frac{1}{2}(x^2+y^2) + \frac{(1-\mu)}{r_1} + \frac{\mu}{r_2} \tag{1}$$

where r_1 and r_2 are the distances from the Earth and the Moon to a space vehicle, respectively, and are given by

$$r_1^2 = (x-\mu)^2 + y^2 + z^2 \qquad r_2^2 = (x-(\mu-1))^2 + y^2 + z^2 \tag{2}$$

μ is the mass ratio equal to $\frac{m_2}{m_1+m_2} = 0.0121505$. In this normalized form, the distance between the primaries assumes an unitary value and the angular velocity is also equal to one, the equation of motion of the particle m_3 are:

$$\begin{aligned} \ddot{x} - 2\dot{y} &= \Omega_x \\ \ddot{y} + 2\dot{x} &= \Omega_y \\ \ddot{z} &= \Omega_z \end{aligned} \tag{3}$$

Those second order ordinary differential equations form a non-integrable system, but they present a first integral, known as the Jacobi integral

$$C = 2\Omega(x,y,z) - \dot{x}^2 - \dot{y}^2 - \dot{z}^2 \tag{4}$$

3. The Numerical Continuation Method

In this section we give a brief description of the numerical method used to compute the families of halo orbit. The essence of this method is the numerical integration of the characteristic curve along an arc parameter, it is detailed in references (Andreu; 1998; Gómez et al., 2001; Simó & Stuchi, 2000). We began the integration with a known initial conditions of the periodic orbit and the method is responsable to find the next set of approximated initial conditions in a vicinity of the one that we have started.

Firstly, we consider the initial point P_i on the Poincaré map y=0,

$$P_i = (x_{i,} y_i = 0, z_i, \dot{x}_i = 0, \dot{y}_i > 0, \dot{z}_i = 0) \tag{5}$$

the next point crosses y=0, with

$$P_f = (x_{f,} y_f = 0, z_f, \dot{x}_f = 0, \dot{y}_f < 0, \dot{z}_f = 0) \tag{6}$$

Due to the symmetry of this orbits respect to (x,z) plane, they pass through the surface y=0 with the components $\dot{x}$ and $\dot{z}$ equal to zero, so we can define two sets of periodic solutions as follow:

$$\begin{aligned} f_1(x, z, \dot{y})_f &= \dot{x}_f = 0 \\ f_2(x, z, \dot{y})_f &= \dot{z}_f = 0 \end{aligned} \tag{7}$$

Let $\overline{X} = (x, z, \dot{y})$ and $F = (f_1, f_2)$. The solutions of the nonlinear system (7) are the zeros of $F(\overline{X}_i) = 0$. If the rank of the Jacobian matrix of F is two, the solution of $F(\overline{X}_i) = 0$ describes a curve in the space $(x, z, \dot{y})$, where each point in that characteristic curve represents a set of initial conditions of the periodic orbits. To follow this curve we have used the numerical continuation method through an arc parameter *s* that is described in the works: Gómez et al., 1985; Andreu; 1998. To attain this task, the follow system should be integrated:

$$\frac{dx}{ds} = \frac{A_1}{A_0}, \quad \frac{dz}{ds} = \frac{A_2}{A_0}, \quad \frac{d\dot{y}}{ds} = \frac{A_3}{A_0} \tag{8}$$

where $A_1 = (f_z^1 f_{\dot{y}}^2 - f_z^2 f_{\dot{y}}^1)$, $A_2 = -(f_x^1 f_{\dot{y}}^2 - f_x^2 f_{\dot{y}}^1)$, $A_3 = (f_x^1 f_z^2 - f_x^2 f_z^1)$ and $A_0 = \sqrt{\sum_{i=1}^{3} A_i^2}$. To obtain the solution of the differential equations system, Eq. (8), a predictor-corrector method was implemented. The increment $d\overline{X} = (dx, dy, d\dot{z})$ when added to $\overline{X}_i = (x, z, \dot{y})$ give us the next point of a periodic orbit on the Poincaré section:

$$\overline{X}_{i+1} = \overline{X}_i + d\overline{X} \tag{9}$$

4. The Transfer Orbits

In this section we are have the goal of getting some optimal transfer orbits from the Halo orbit to an elliptical orbit around the Earth. To attain our objective we have investigated two manners of transferring a spacecraft: the first one, called quasi-natural transfer orbits and the second is based on the Lambert's problem. In this context, optimal transfer orbits are the ones that provide the lowest ΔV.

4.1. Quasi-natural Transfers

Due to the instability of the Halo orbits, the space vehicle could be detached from the orbit after a few revolutions on it. If a small displacement is given to the body on the Halo orbit it will escape easily, which means that one have a quasi-natural transfer orbit.

We have done some performance of this method from several points of the Halo orbit, but just a set of few points led to an approach to the Earth. It shows a disadvantage of using the quasi-natural transfers. The points that we have applied this method are shown in the Fig. (2). They allowed an approach of, approximately, 0.3 canonical units to the Earth.

We have noted that if we give the first impulse forward to x-positive direction, the vehicle can reach a point closer to the Earth. The second impulse is applied to capture the body, so the total ΔV is given by $\Delta V_i + \Delta V_f$, where:

$$\Delta V_i = k \, | v_x | \tag{10}$$

k represents a factor that should be increased on the velocity component, with a magnitude in the order of 10^{-2}, and

$$\Delta V_f = | \Delta V_{transfer} - V' | , \tag{11}$$

where V' is the velocity of a circular orbit around the Earth in a synodic frame and $V_{transfer}$ is the velocity of the spacecraft. To insert the vehicle on this circular orbit, the direction of $V_{transfer}$ should be parallel to the direction of the orbit upon arrival. Since, it does not happen, the vehicle enters in an elliptical orbit. The magnitude of V' is given by

$$V' = \sqrt{\frac{1-\mu}{r_1}} \,. \tag{12}$$

4.2. Lambert's Problem

When a system of differential equations is supposed to satisfy a set of initial and final conditions, the problem becomes the two point boundary value problem, where the time is a free variable. In our case, the set of initial conditions represents a point on the Halo orbit $P_A = (x_A, y_A, z_A)$ and the final condition is a point in an elliptical orbit near the Earth, $P_B = (x_B, y_B, z_B)$. In this way, there are infinite solutions, but, on the other hand, we can attribute the time of flight, adding a time constraint and, consequently, we are able to reduce the set of solutions.

Such a formulation, the "two point boundary value problem" plus a time constraint, in the restricted three-body problem is known as the Lambert's three-body problem, it is described in Prado, 1995. Our investigation begins on finding optimal trajectories to travel between the points A and B. The calculations of the total ΔV is determinated as in the case of quasi-natural transfer orbits.

$$\Delta V_A = | V_{orbital} - V_{transfer} | , \tag{13}$$

where $V_{orbital}$ is the Halo orbital velocity and $V_{transfer}$ is the velocity determinated by the Lambert's problem. These velocity should be enough to satisfy the time and the boundary constraints. Finally, the second impulse is:

$$\Delta V_B = | V_{transfer} - V' | . \tag{14}$$

and $\Delta V = \Delta V_A + \Delta V_B$.

5. Results

Contrary to the Lambert's method, which has got the well-defined boundary conditions, we had to adopt a criteria to stop the integration of the quasi-natural transfers: when a trajectory passes in a distance less or equal than 0.3 canonical units, this point is considered a good set of final conditions. The illustration of this method is shown in Figs. (2-3) in three different projections, on the planes (x,y), (x,z) and (y,z). A small impulse ΔV_i was applied on the marked point on the Halo orbit. This impulse is not strong enough to force the vehicle to leave the orbit in that marked point so, the vehicle escapes from the orbit in another region.

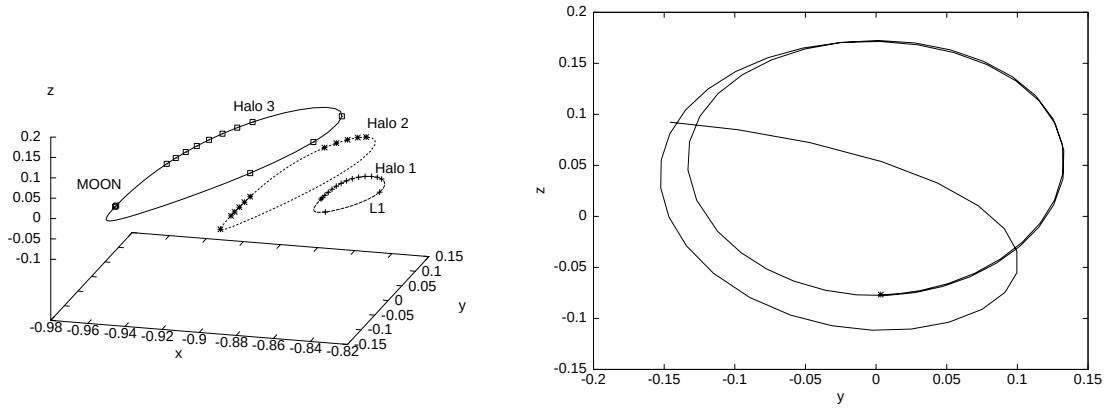


Figure 2. The Halo orbits studied with the points in which we have applied the models of transferring trajectories. The second graphic shows the escaping orbit in the quasi-natural problem, projected on the (xy) plane. A small increment of velocity was applied at the marked point.

This procedure was performed for the points marked in the three different halo orbits that belong to the same family, named as Halo 1, Halo 2 and Halo 3, as it is shown at the first graphic in the Fig. (2). For each point in those orbits we have also performed the Lambert's transfer approach covering a time interval from 0.5 to 10 canonical units of time, in steps of 0.4 units of time. The plots in the Figs. (4-6) show the comparison between the two methods applied on these points; the graphics at the left side of these figures show the total increment of velocity that was used in the quasi-natural transfer orbits and the graphics at the right side show the total increment of velocity required by the Lambert's method. It is seen that the first method applied is more economical than the Lambert's method, but it takes more time to reach the final conditions. The initial and final points are expressed in the Tables (2-4).

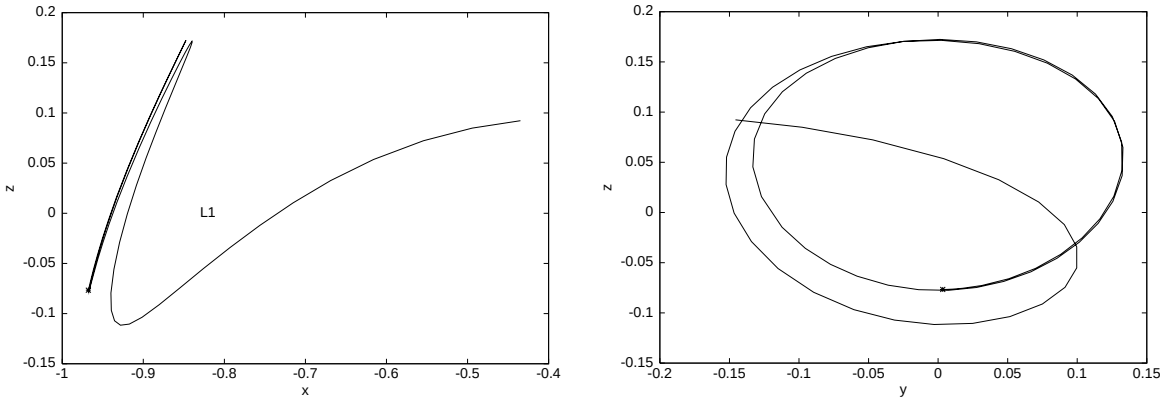


Figure 3. Escaping orbits in the quasi-natural problem, projection on the planes (xz) and (yz), respectively.

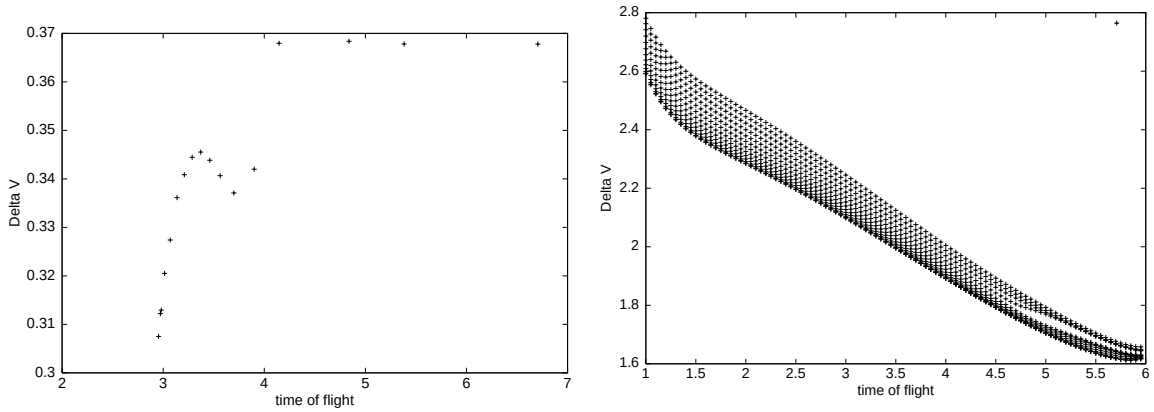


Figure 4. Comparison of consumption between the quasi-natural method and the Lambert method for the Halo 1

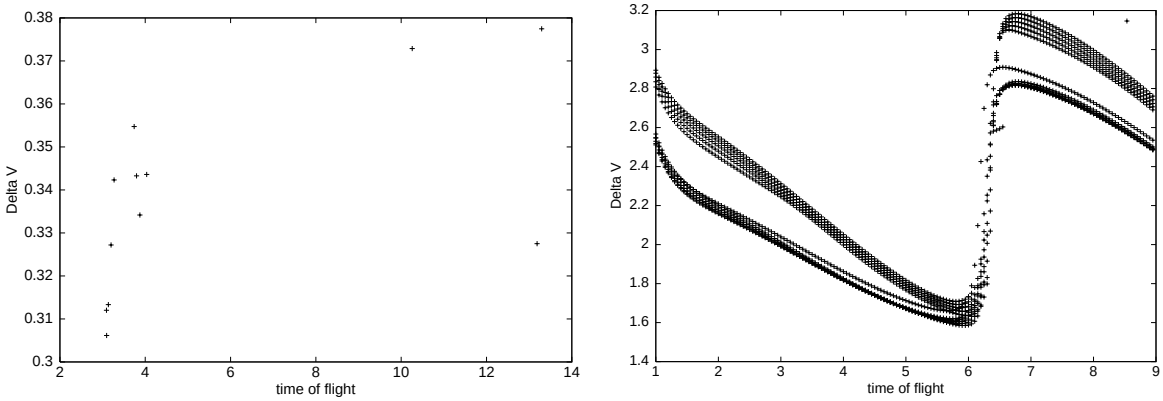


Figure 5. Comparison of consumption between the quasi-natural method and the Lambert method for the Halo 2

The Lambert's transfers simulations were performed as follow: at first, we discretized each Halo orbit in intervals of, approximately, 7^0. Then, for each point we have fixed different times constraints from 0.5 to 6 units of time in steps of 0.3. At the projection on the (y,z) plane, we attributed an origin to count the Halo angle, that is set on the z-positive direction. The result is shown in Figs. (7-8). The lowest ΔV is given at 180^0 from the origin. Also, we have seen that

this point has the highest velocity, $V_{orbital}$. Although, for the Halo 1, the minimum ΔV happens on the interval 100^0 to 240^0.

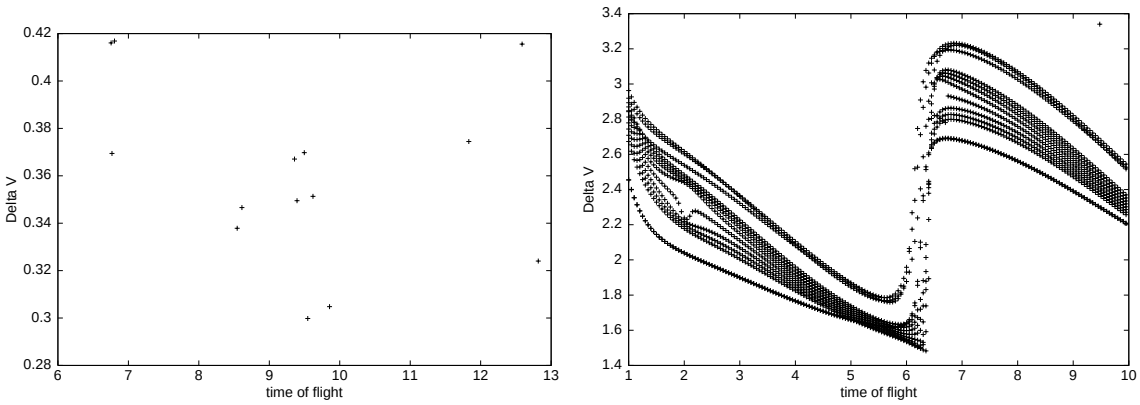


Figure 6. Comparison of consumption between the quasi-natural method and Lambert method for the Halo 3

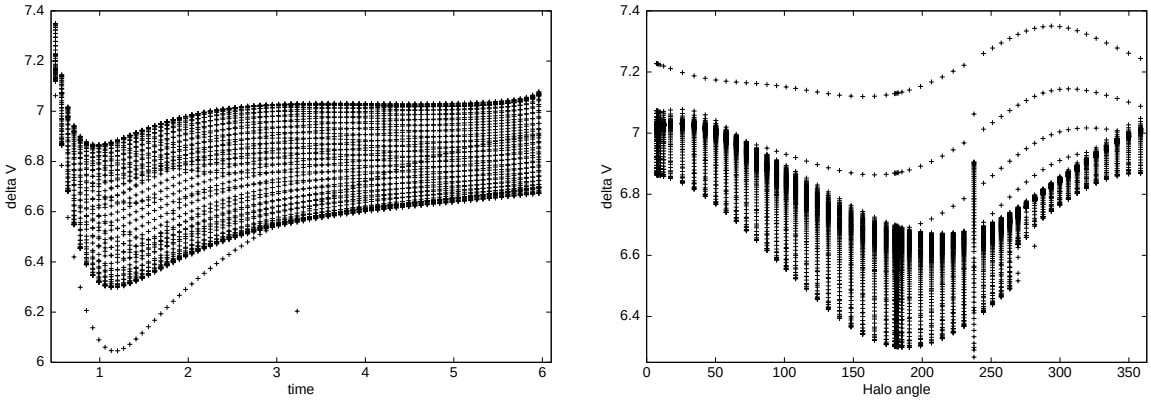


Figure 7. Lambert's method to transfer a spacecraft from the orbit Halo 1 to a distance of 0.0182 canonical units from the Earth

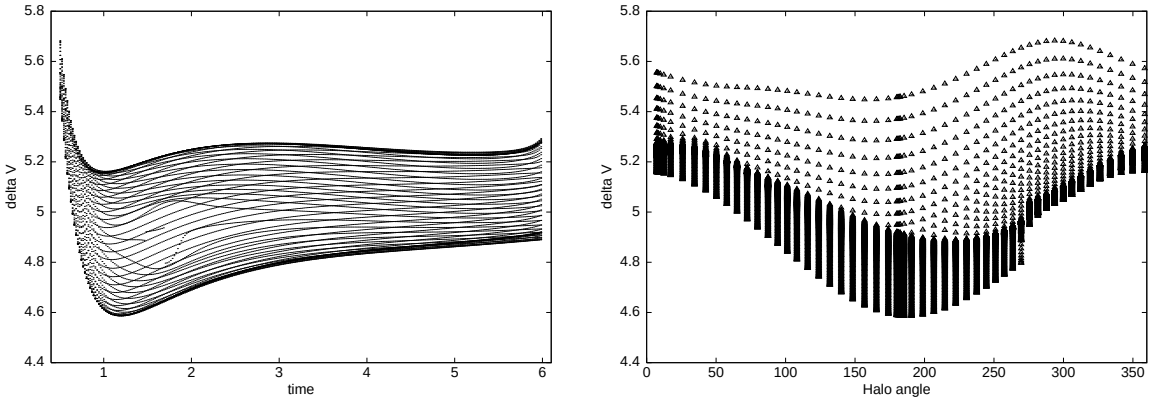


Figure 8. Lambert's method to transfer a spacecraft from the orbit Halo 1 to a distance 0.03 of canonical units from the Earth

6. Tables

Table 1. Initial Conditions of the Halo orbits used for all calculations in both methods

Initial Conditions	x	y	Z	$\dot{x}$	$\dot{y}$	$\dot{z}$
Halo 1	-0.85682660	0	-0.01745687	0	0.14233546	0
Halo 2	-0.90562287	0	-0.07786493	0	0.30253607	0
Halo 3	-0.96775851	0	-0.07673404	0	0.49163428	0

Table 2. Initial and final positions of the marked points on Halo 1 (see Fig. (2) above) which ones the quasi-natural method was applied using k = 0.05

x_0	y_0	z_0	$\dot{x}_0$	x_F	y_F	z_F	Time of Flight
-0.856806970	0.004155950	-0.017406470	0.001345150	-0.188073969	0.180757398	0.007569221	6.618000314
-0.856747990	0.008277400	-0.017256260	0.002708160	-0.017333753	-0.268147302	0.007823005	3.954000188
-0.856513680	0.016216790	-0.016676780	0.005522290	-0.014073873	-0.268486505	0.005124642	3.750000178
-0.856015150	0.025199880	-0.015522210	0.009231340	0.017100376	-0.269775805	0.004114668	3.621000172
-0.855224810	0.033673840	-0.013860250	0.013526010	0.050791439	-0.267175709	0.004216628	3.527000168
-0.854067170	0.041455480	-0.011690180	0.018445650	0.087216900	-0.259160437	0.004950526	3.450000164
-0.852452700	0.048282540	-0.009017610	0.023863420	0.120224445	-0.247318596	0.005880168	3.381000161
0.850356070	0.053685400	-0.005953470	0.029313730	0.109219759	-0.251818743	0.005540113	3.298000157
-0.847740830	0.057405050	-0.002557970	0.034348370	0.074422539	-0.262666171	0.004612468	3.211000153
-0.844601520	0.059140340	0.001096060	0.038419830	0.041222409	-0.268355418	0.004073614	3.134000149
-0.840993860	0.058552030	0.004901330	0.040894400	0.011972234	-0.269624975	0.004144981	3.070000146
-0.837081490	0.055343700	0.008686910	0.041125580	-0.013571196	-0.268594044	0.004941953	3.023000144
-0.833091820	0.049311200	0.012275010	0.038585920	-0.023431525	-0.267108025	0.006650255	3.006000143
-0.829478890	0.040797530	0.015339170	0.033221100	-0.004299305	-0.269218202	0.008355995	3.038000144
-0.823555490	-0.007304260	0.020064500	-0.006286670	0.012737533	-0.269660228	-0.000970623	5.443392849
-0.845308090	-0.058955290	0.000305790	-0.037669070	0.025052794	-0.269428603	0.002763643	4.904340007
-0.853496760	-0.044249780	-0.010706510	-0.020513290	-0.187553837	0.181349983	-0.007916565	13.315792433

Table 3. Initial and final positions of the marked points on Halo 2 (see Fig. (2) above) which ones the quasi-natural method was applied using k = 0.07

x_0	y_0	z_0	$\dot{x}_0$	x_F	y_F	z_F	Time of Flight
-0.905330040	0.017560540	-0.076594510	0.010045690	-0.008421343	-0.246287906	0.037293629	4.113000195
-0.904726650	0.030345160	-0.074027330	0.017625320	-0.039694178	-0.243025854	0.024272358	3.940000187
-0.903775730	0.042756290	-0.070107980	0.025383310	0.006089750	-0.248794179	0.020760886	3.872000184
-0.902428230	0.054822470	-0.064787010	0.033449230	-0.153817932	0.185539116	0.019014797	6.285000299
-0.870641720	0.111767370	0.026367950	0.087106630	-0.158058715	0.181755989	0.019657632	5.821000276
-0.863925160	0.109084040	0.041764320	0.086692320	0.018160090	-0.248666152	0.021208897	3.277000156
-0.856969830	0.103085610	0.057011650	0.083254490	-0.028475329	-0.245262009	0.021727287	3.204000152
-0.850120660	0.093649300	0.071469820	0.076604540	-0.046198123	-0.240760052	0.028888049	3.160000150
-0.843673190	0.080677070	0.084666740	0.066644650	-0.011189171	-0.245798145	0.037426121	3.169000151
-0.901615590	-0.060504550	-0.061694080	-0.037465750	-0.064300981	-0.236524648	0.026069871	13.248000629
-0.905592800	0.005657990	-0.077733710	0.003214170	-0.193780979	0.138921928	0.025515744	12.752000606

Table 4. Initial and final positions of the marked points on Halo 3 (see Fig. (2) above) which ones the quasi-natural method was applied using k = 0.07

x_0	y_0	z_0	$\dot{X}_0$	X_F	Y_F	Z_F	Time of Flight
-0.967745830	0.003350260	-0.076688800	0.003719030	0.205871230	-0.173310719	0.109536341	12.817000609
-0.967702110	0.007062790	-0.076532900	0.007840400	-0.088232505	-0.281793471	0.018269376	6.768000321
-0.953543550	0.096692410	-0.033212050	0.107212190	-0.190432397	-0.195221463	0.103453495	9.855000468
-0.949862110	0.104912280	-0.023519090	0.116120530	-0.102049077	-0.264751994	0.081276471	9.620000457
-0.945582570	0.112453010	-0.012781380	0.124171620	-0.052700410	-0.282445148	0.074724546	9.497000451
-0.940615470	0.119194210	-0.000908890	0.131214900	-0.035741297	0.290457815	0.057535675	6.803000323
-0.934864770	0.124942760	0.012186810	0.137023420	-0.031573648	0.290813615	0.058355395	6.755000321
-0.928235470	0.129417580	0.026578790	0.141281330	-0.047686777	-0.284079656	0.074309912	9.358000444
-0.920739830	0.132210130	0.042115800	0.143557090	-0.093413750	-0.268041937	0.080514887	9.394000446
-0.912540450	0.132902150	0.058389830	0.143443450	-0.186780159	-0.197992194	0.103348740	9.546000453
-0.84944869	0.030620960	0.169378120	0.031482940	-0.176206358	-0.232614442	-0.007202664	11.832000562
-0.852861890	-0.051027650	0.163763830	-0.052607210	-0.039511283	-0.294663900	-0.022259998	8.612000409
-0.878199810	-0.110945300	0.121037720	-0.116671800	-0.108656564	-0.270421686	0.045964974	8.544000406

7. Conclusions

It was seen that the Lambert's transfer method appears to be more accurate than the quasi-natural transfers, but it spends more impulse to attain the same conditions. To perform a rendezvous manoeuvres the Lambert's method is more indicated because we have control over the time and the boundary condition.

8. Acknowledgements

We are thankful to FAPESP for supporting this work, under the contract number 99/11457-2.

9. References

Andreu M. A., 1998, "The Quasi-Bicircular Problem", Doctored Thesis, Universitat de Barcelona, Barcelona, Spain.

Breakwell J. V., Brown J. V., 1979, "The Halo Family of 3-Dimensional Periodic Orbits in the Earth-Moon Restricted 3-Body Problem", Celestial Mechanics, Vol. 20, pp 389-404.

Brouke R. A., 1968, "Periodics Orbits in the Restrited Three Body Problem with Earth-Moon Masses", JPL, Technical Report 32-1168.

Damario L. D. D., Edelbaum T. N.:, 1974, "Minimum Impulsive Three Body Trajectories", AIAA Journal, Vol. 12, pp. 455-462.

Farquhar R. W., 1970, "The Control and use of Libration Point Satellites", Nasa, TR R-346.

Farquhar R. W., 1980, "Trajectories and orbital Maneuvers for the First Libration-Point Satellite", Journal of Guidance, Control and Dynamics, Vol. 3 , No. 6, pp. 549-554.

de Felipe G., Beaugé C., Prado A. F. B. A., 2000, "Determinação Analítica de Órbitas do Tipo Halo", Boletim da Sociedade Astronômica Brasileira, Vol. 20, No 1.

Gómez G., Simó C., 1993, "Study of the Transfer from the Earth to a Halo Orbit around the Equilibrium Point L_1", Celestial Mechanics and Dynamical Astronomy, Vol. 56, pp. 541-562.

Gómez G., Llibre J., Martínez R., Simó C., 2001, "Dynamics and Mission Design Near Libration Points", Vol. 1, Fundamentals: The Case of Collinear Libration Points, World Scientific.

Gómez G., Llibre J., Martínez R., Simó C., 2001, "Dynamics and Mission Design Near Libration Points", Vol. 2, Fundamentals: The Case of Triangular Libration Points, World Scientific.

Gómez G, Simó C., Jorba À., Masdemont J., 2001, "Dynamics and Mission Design Near Libration Points", Vol. 3, Fundamentals: Advanced Methods for Collinear Points, World Scientific.

Gómez G, Simó C., Jorba À., Masdemont J., 2001, "Dynamics and Mission Design Near Libration Points", Vol. 4, Fundamentals: Advanced Methods for Triangular Points, World Scientific.

Howell K. C., Pernicka H. J., 1993, "Station Keeping Method for libration Point Trajectories", Journal of Guidance, Control and Dynamics, Vol. 16, No. 1, pp. 151-159.

Howell K. C., Mains D. L., Barden B. T., 1994, "Transfer Trajectories form Earth Parking Orbits to Sun-Earth Halo Orbits", AAS/AIAA Spaceflight Mechanics Meeting, Cocoa Beach, FL.

Popescu M., Cardos V., 1995, "The Domain of Initial Conditions for the Class of Three-Dimensional Halo Periodical Orbits", Acta Astronautica, Vol. 36, No. 4, pp. 193-196.

Prado A. F. B. A., Broucke R.., 1995, "Transfer Orbits in Restricted Problem", Journal of Guidance, Control and Dynamics, Vol. 18, No. 3, pp. 593-598.

Simó C., Stuchi T. J., Central Stable/Unstable Manifolds and the Destruction of KAM Tori in the Planar Hill Problem, Phisica D, 2000, Vol. 140, pp. 1-32.

Szebehely V., 1967, "Theory of Orbits", Academic Press, New York.

XVI CONGRESSO BRASILEIRO DE ENGENHARIA MECÂNICA
16th BRAZILIAN CONGRESS OF MECHANICAL ENGINEERING

MINIMUM FUEL TRANSFERS BETWEEN HALO ORBITS

Gislaine de Felipe
Instituto Nacional de Pesquisas Espaciais - INPE
São José dos Campos - SP - 12227-010 - Brasil
Phone (123)256197 - Fax (123)25-6226
E-mail: GISLAINE@DEM.INPE.BR

Antonio Fernando Bertachini de Almeida Prado
Instituto Nacional de Pesquisas Espaciais - INPE
São José dos Campos - SP - 12227-010 - Brasil
Phone (123)256197 - Fax (123)25-6226
E-mail: PRADO@DEM.INPE.BR

Cristian Beauge
Universidade Nacional de Córdoba
Laprida 854, (5000), Córdoba, Argentina
E-mail: BEAUGE@DEM.INPE.BR

Abstract. *This paper has the goal of studying the problem of orbital transfers between two Halo orbits in the Sun-Earth system. Halo orbits are special three-dimensional trajectories that exist around the Lagrangian points of the restricted three-body problem. These orbits are studied in several papers, since they have important applications in astronautics. The first step involved in this research is to perform the determination of the Halo orbits. To do that, an analytic calculation is performed using the Linstedt-Poincaré method. The present paper considers that a maneuver will be performed to transfer the spacecraft from an initial to a final Halo orbit. The control that will be used to achieve that goal is constituted by a pair of instantaneous change in the velocity of the spacecraft at the beginning and at the end of the maneuver. A numerical algorithm based in the Lambert Problem is built to calculate the transfer orbits. The two orbits are divided in several points and the algorithm is applied to each pair of points. Finally, the solution that gives the minimum fuel consumption is found. This maneuver can be used to: i) change the orbit of the spacecraft to allow a second application in a different Halo orbit; ii) to perform station keeping in a Halo orbit that is escaping from its nominal orbital parameters.*

Key Words*: Astrodynamics, Halo Orbits, Orbital Maneuvers.*

1. Introduction

To study the problem of transfer orbits between Halo orbits, the restricted three-body problem in three dimensions is used as the mathematical model. It is assumed that the total system (Earth + Sun + spacecraft) satisfies the hypotheses: i) all the bodies are point masses; ii) the Earth and the Sun are in circular orbits around their mutual center of mass. Then, the goal is to study the motion of the spacecraft governed by these two masses. The Halo orbits are trajectories that exist around the Lagrangian points, that are the well-known equilibrium points that appear in the rotating frame of the circular restricted three-body problem (e. g. Szebehely, 1967). They are very important for astronautical applications. Since they are five points of equilibrium in the equations of motion, it means that a particle located at one of those points with zero relative velocity will remain there indefinitely. The collinear points (L_1, L_2 and L_3) are always unstable and the triangular points (L_4 and L_5) are stable in the present case (Earth-Sun system). They are all very good points to locate a space station, since they require a small amount of ΔV (and fuel) for station keeping. The triangular points are especially good for this purpose, since they are stable equilibrium points. In the nomenclature used in this paper, L_1 is the collinear Lagrangian point that exists between the Sun and the Earth. It is located about 1,496,867 km from the Earth. L_2 is the collinear Lagrangian point that exists behind the Earth. It is located about 1,506,915 km from the Earth.

There are many papers in the literature that refers to Halo orbits. In a general form, it is possible to divide them in three groups:

1) Papers that concentrate on the determination of the orbits, like: Farquhar *et al*. (1977), Howell and Breakwell (1984), Breakwell and Brown (1979), Farquhar (1972), Richardson (1980), Farquhar (1970), Farquhar (1991), Popescu and Cardos (1995), Popescu (1986), Farquhar and Kamel (1973), Richardson (1980a e 1980b). These papers describe the Halo orbits and show approximate analytical solutions, that can later be used as a starting point to find accurate numerical orbits;

2) Papers that consider the problem of transferring the spacecraft between a parking orbit around the Earth and a Halo orbit. Among these, we can mention Gómez *et al*. (1993), Howell *et al*. (1994), Stalos *et al*. (1993), Starchville and Melton (1997), D' Amario and Edelbaum (1974);

3) A third line of research studies maneuvers between Halo orbits. Some good samples are: Simó (1987), Howell and Pernicka (1993), Howell and Gordon (1992), Farquhar (1970), Farquhar *et al.* (1980), Hiday and Howell

4) (1992), Gingis (1992), Gordon and Howell (1992), Farquhar (1980), Popescu (1985). In this category it is also possible to include some papers that consider the Rendezvous between two spacecrafts, like: Jones (1993), Jones and Bishop (1993, 1994).

2. The restricted three-dimensional three-body problem

The model used in this paper is the well-known circular restricted three-body problem. This model assumes that two main bodies (M_1 and M_2) are orbiting their common center of mass in circular Keplerian orbits and a third body (M_3), with negligible mass, is orbiting these two primaries. The motion of M_3 is supposed to be affected by both primaries, but it does not affect their motion (Szebehely, 1967). The canonical system of units is used, and it implies that: i) The unit of distance (l) is the distance between M_1 and M_2; ii) The angular velocity (ω) of the motion of M_1 and M_2 is assumed to be one; iii) The mass of the smaller primary (M_2) is given by $\mu = m_2/(m_1+m_2)$ (where m_1 and m_2 are the real masses of M_1 and M_2, respectively) and the mass of M_1 is (1-μ), so the total mass of the system is one; iv) The unit of time is defined such that the period of the motion of the primaries is 2π; v) The gravitational constant is one. Table 1 shows the values for these parameters for the Sun(M_1)-Earth(M_2) system, which is the case considered in this paper.

Table 1. Canonical system of units.

Unit of distance	149,596,000 km
Unit of time	58.13 days
Unit of velocity	29.8 km/s

There are several systems of reference that can be used to describe the three-dimensional restricted three-body problem (see Szebehely, 1967). In this paper the rotating system is used. In the rotating system of reference, the origin is the center of mass of the two massive primaries. The horizontal axis (x) is the line that connects the two primaries at any time. It rotates with an angular velocity ω in such way that the two massive primaries are always on this axis. The vertical axis (y) is perpendicular to the (x) axis. In this system, the positions of the primaries are: $x_1 = -\mu$, $x_2 = 1-\mu$, $y_1 = y_2 = 0$. In this system, the equations of motion for the massless particle are (Szebehely, 1967):

$$\ddot{x} - 2\dot{y} = x - (1-\mu)\frac{x+\mu}{r_1^3} - \mu\frac{x-1+\mu}{r_2^3} \tag{1}$$

$$\ddot{y} + 2\dot{x} = y - (1-\mu)\frac{y}{r_1^3} - \mu\frac{y}{r_2^3} \tag{2}$$

$$\ddot{z} = -(1-\mu)\frac{z}{r_1^3} - \mu\frac{z}{r_2^3} \tag{3}$$

where r_1 and r_2 are the distances from M_1 and M_2, given by

$$r_1^2 = (x+\mu)^2 + y^2 + z^2; \quad r_2^2 = (x-1+\mu)^2 + y^2 + z^2 \tag{4-5}$$

3. The Halo Orbits

The goal of this paper is to find transfer orbits between Halo orbits, which are periodic orbits in the three-dimensional restricted three-body problem in the neighborhood of the collinear Lagrangian points (e. g. Breakwell and Brown, 1979). Fig. (1) shows a schematic view of those orbits. Those transfers are required to make possible orbital maneuvers for space missions. To perform this task, it is necessary to look at the characteristics of the Halo orbits and explore the non-linear system. To determine those orbits, a possible approach is to find an analytical approximation for the family of periodic orbits by using the Linstedt-Poincare method, which allows us to obtain one solution, in a series form, until a very high order. As a reference orbit it is possible to use the linear solution.

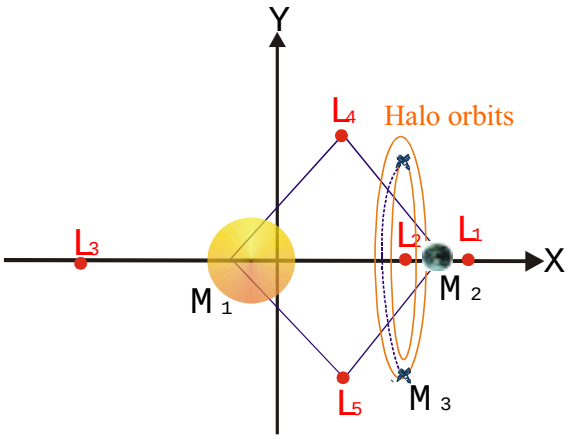


Figure 1. Geometry of the problem.

The basic idea of this method is to use the non-linear part of the differential equations written as a Legendre polynomial (Richardson, 1980). The solution of the linear approximation can be written as:

$$x = A\exp\{i\omega_o t\}; \quad y = KA\exp\{i\omega_o t\}; \quad z = B\exp\{i\nu_o t\} \tag{6}$$

where ω_0, ν_0 are the fundamental frequencies in the plane and perpendicular to the plane; A, B are the amplitudes (in complex notation) in the x and z axis, respectively, and K is the relation between the amplitudes in the x and y axis. After that, we know that when the non-linear terms of the differential equations are included in the variational equations, we have that the general solution will depend on the linear solution in the plane and in the z axis. Remember that in this case the motions are no longer separable. So, it is possible to write the general solution as a Fourier series of the type

$$x = \sum_{i,j,k,l} X_{i,j,k,l} A^i B^j \exp\{k\omega + l\nu\}t \tag{7}$$

where $X_{i,j,k,l}$ are complex coefficients. The same is true for the variables y and z. Their determination is made by replacing the proposed solutions in the differential equations, identifying the terms of the same harmonic and power and solving the algebraic equations that are left. The new frequencies can be determined as a series of power in the amplitudes

$$\omega = \sum_{i,j} \omega_{i,j} A^i B^j; \quad \nu = \sum_{i,j} \nu_{i,j} A^i B^j \tag{8}$$

where $\omega_{0,0}$; $\nu_{0,0}$ are the frequencies of the linear system. In this way, it is possible to determine periodic orbits if we request that ω and ν be both equal (i.e. $\omega = \nu$).

Following this determination, it is possible to perform the validation of the solution found by this method, using numerical integration of the restricted three-body problem with the initial condition found by the analytical method.

After finding a set of initial conditions (x, y, z, Vx, Vy, Vz) that allows the spacecraft to enter in an orbit around the Lagrangian point, it is possible to use the Lambert Problem approach in the restricted three-body problem, (which is described below) to find the orbital transfers. It is necessary to choose two points that belong to two different Halo orbits and vary the time of flight involved.

4. The Lambert's Three-Body Problem

The problem that is considered in the present paper is that of finding trajectories to travel between two points fixed in the rotating frame, that belong to two different Halo orbits. Since those points are in known positions, this problem can be formulated as:

"Find an orbit (in the three-body problem context) that makes a spacecraft leave a given point A and go to another given point B". That is the famous TPBVP (two point boundary value problem). There are many orbits that satisfy this requirement, and the way used in this paper to find families of solutions is to specify a time of flight for the transfer. Then, the problem becomes the Lambert's three-body problem, that can be formulated as:

"Find an orbit (in the three-body problem context) that makes a spacecraft leave a given point A and go to another given point B, arriving there in a specified time of flight". Then, by varying the specified time of flight it is possible to find a whole family of transfer orbits and study them in terms of the required ΔV, energy, initial flight path angle, etc.

To solve this problem, it is necessary to remember that the restricted three-body problem has no general analytical solution, so numerical integration is the only possible approach to solve it. Then, it is necessary to follow the steps:

i) Guess a initial velocity $\vec{V}_i$, so together with the initial prescribed position $\vec{r}_i$, the complete initial state is known;

ii) Integrate numerically the equations of motion for a specified transfer time t;

iii) Check the final position $\vec{r}_f$ obtained from the numerical integration with the prescribed final position. If there is an agreement (difference less than a specified error allowed) the solution is found and the process can stop here. If there is no agreement, an increment in the initial guessed velocity $\vec{V}_i$ is made and the process goes back to step ii).

The method used to find the increment in the guessed variables ($\vec{V}_i$) is the standard gradient method, as described in Press et. alli, 1989. The routines available in this reference are also used in this research with minor modifications.

5. Results

To simulate orbital transfers, the approach used in this paper is to discretize the initial and the final Halo orbits into several points. For a preliminary study this number of points was chosen equal to eight. The points were chosen in a such way that the time to travel between them is constant. The fact that the velocity of the spacecraft is not constant, makes those points to be closer to each other in the regions where the velocity is lower. This number can be increased to get a more accurate minimum cost maneuver. Then, each point of the initial orbit is combined with each point of the final orbit. For each pair of points, the Lambert problem is solved for a series of times of flight. Then, the maneuver that has the minimal fuel consumption for each pair is listed in a table. After all the simulations are performed, we have a table that gives you the minimum fuel consumption maneuver for the whole transfer. Fig. (2) shows several Halo orbits discretezed in points. From this collective view it is possible to see the three-dimensional character of the orbits.

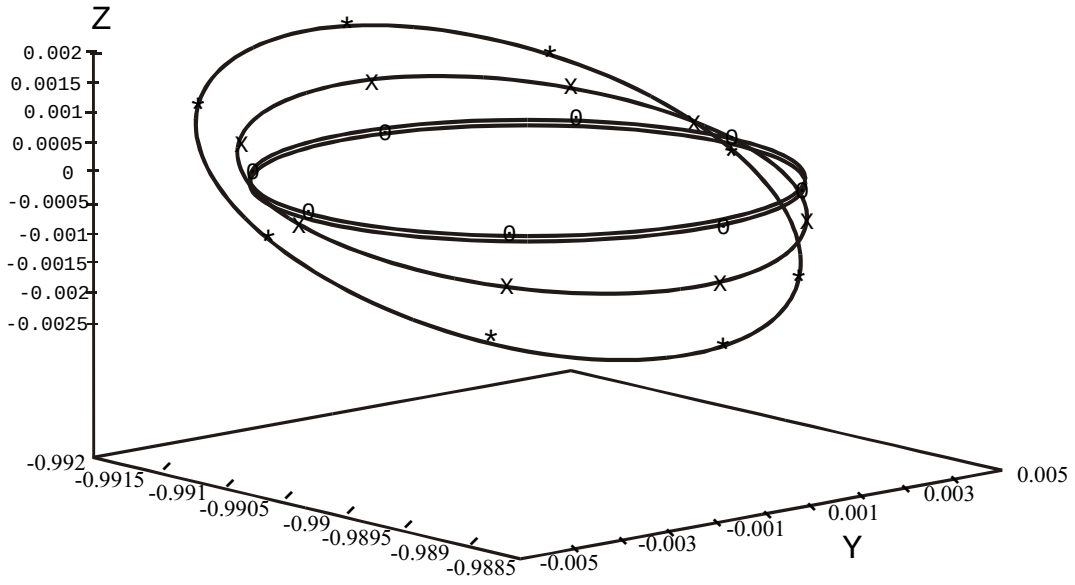


Figure 2. Halo orbits showing the points for the transfers.

For this first study, five different Halo orbits were generated with the Linstedt-Poincare method. They all belong to the same family around the Lagrangian point L_1. The difference between them can be expressed by a single parameter, usually the parameter B, that represents the linear amplitude in the z axis. In terms of this parameter, the orbits used in this research are: 0.001, 0.002, 0.008, 0.100, 0.200. Table 2 shows the position and velocity, in the rotating frame, for the eight points in which the orbits were divided and for all five orbits. This table also defines the notation used for the points. Table 3 shows the distance from the Lagrangian point L_1 ($|\vec{r}|$) and the magnitude of the velocity ($|\vec{V}|$) for all the points, to improve our physical view of those orbits. From this table, it is clear that the velocity is higher when the spacecraft is closer to the Lagrangian point and that this velocity does not change very much with the amplitude.

Table 2. Position and Velocity, in the rotating frame, for the points at the Halo Orbit

β = 0.001	x	y	z	Vx	Vy	Vz
a	0.9915764	0	9.0146638E-5	0	-9.2802658E-3	0
b	0.9905927	-4.4068124E-3	-1.7595656E-5	-2.9308972E-3	3.0436429 E-5	-2.0207181E-4
c	0.9888314	0	-1.1173384E-4	0	8.8361299E-3	0
d	0.9905927	4.4068124E-3	-1.7595656E-5	2.9308972E-3	3.0436429E-5	2.0207181E-4
f	0.9139200	-3.1786984E-3	5.5462655E-5	-1.2066819 E-3	-6.4206765 E-3	-1.6725628 E-4
g	0.9894499	3.2139562E-3	-8.3025874E-5	-2.7426946E-3	6.0320211E-3	-1.3484908E-4
h	0.9894499	3.2139562E-3	-8.3025874E-5	2.7426946E-3	6.0320211E-3	-1.3484908E-4
i	0.9139200	3.1786984E-3	5.5462655E-5	1.2066819E-3	-6.4206765E-3	1.6725628E-4

β = 0.002	x	y	z	Vx	Vy	Vz
1	0.9915791	0	1.8027433E-4	0	-9.6372878E-3	0
2	0.9913730	-3.2886017E-3	1.0512785E-4	-1.2940653E-3	-6.1670399E-3	-3.4361399E-4
3	0.9905184	-4.4036820E-3	-4.5490277E-5	-3.0013907E-3	4.8843334E-4	-4.0031494E-4
4	0.9893819	-3.0581790E-3	-1.7277878E-4	-2.6358010E-3	6.3576697E-3	-2.5543628E-4
5	0.9888315	0	-2.2349013E-4	0	8.8422068E-3	0
6	0.9893819	3.0581790E-3	-1.7277878E-4	2.6358010E-3	6.3576697E-3	2.5543628E-4
7	0.9905184	4.4036820E-3	-4.5490277E-5	3.0013907E-3	4.8843334E-4	4.0031494E-4
8	0.9913730	3.2886017E-3	1.0512785E-4	1.2940653E-3	-6.1670399E-3	3.4361399E-4

β = 0.008	x	y	z	Vx	Vy	Vz
9	0.9915761	0	7.2118220E-5	0	-9.6269149E-3	0
10	0.9913704	-3.2857156E-3	4.2065759E-5	-1.2918301E-3	-6.1617643E-3	-1.3742438E-4
11	0.9905169	-4.3997934E-3	-1.8178376E-5	-2.9985447E-3	4.8672559E-4	-1.6012900E-4
12	0.9893812	-3.0557504E-3	-6.9098170E-5	-2.6342045E-3	6.3520242E-3	-1.0218788E-4
13	0.9888312	0	-8.9385994E-5	0	8.8354003E-3	0
14	0.9893812	3.0557504E-3	-6.9098170E-5	2.6342045E-3	6.3520242E-3	1.0218788E-4
15	0.9905169	4.3997934E-3	-1.8178376E-5	2.9985447E-3	4.8672559E-4	1.6012900E-4
16	0.9913704	3.2857156E-3	4.2065759E-5	1.2918301E-3	-6.1617643E-3	1.3742438E-4

β = 0.1	x	y	z	Vx	Vy	Vz
17	0.9916646	0	8.9835988E-4	0	-9.9311651E-3	0
18	0.9914460	-3.3814689E-3	5.2044320E-4	-1.3575077E-3	-6.3154853E-3	-1.7256193E-3
19	0.9905619	-4.5132517E-3	-2.3374731E-4	-3.0818872E-3	-5.3731726E-4	-2.0004922E-3
20	0.9894013	-3.1264760E-3	-8.6863238E-4	-2.6808429E-3	6.5167892E-3	-1.2724627E-3
21	0.9888426	0	1.1210284E-3	0	9.0336156E-3	0
22	0.9894013	3.1264760E-3	-8.6863238E-4	2.6808429E-3	6.5167892E-3	1.2724627E-3
23	0.9905619	4.5132517E-3	-2.3374731E-4	3.0818872E-3	5.3731726E-4	2.0004922E-3
24	0.9914460	3.3814689E-3	5.2044320E-4	1.3575077E-3	6.3154853E-3	1.7256193E-3

β = 0.2	x	y	z	Vx	Vy	Vz
25	0.9919299	0	1.7783072E-3	0	-1.0822374E-2	0
26	0.9916728	-3.6602402E-3	1.0088883E-3	-1.5513607E-3	-6.7533095E-3	-3.4979149E-3
27	0.9906992	-4.8383083E-3	-5.0606701E-4	3.3244333E-3	6.9133561E-4	-3.9942181E-3
28	0.9894669	-3.3275079E-3	-1.7663177E-3	-2.8152400E-3	6.9894204E-3	-2.5166626E-3
29	0.9888835	0	-2.2641394E-3	0	9.5971956E-3	0
30	0.9894669	3.3275079E-3	-1.7663177E-3	2.8152400E-3	6.9894204E-3	2.5166626E-3
31	0.9906992	4.8383083E-3	-5.0606701E-4	3.3244333E-3	6.9133561E-4	3.9942181E-3
32	0.9916728	3.6602402E-3	1.0088883E-3	1.5513607E-3	-6.7533095E-3	3.4979149E-3

Table 3. Distance from the Lagrangian point L_1 and the magnitude of the velocity.

Point	$\lvert \vec{r} \rvert$	$\lvert \vec{v} \rvert$
a	1.5880606E-3	9.2802658E-3
b = d	4.4477486E-3	2.9380226E-3
c	1.6806183E-3	8.8361299E-3
f = i	7.6137303E-2	6.5352232E-3
g = h	3.2602283E-3	6.6276569E-3
1	7.4199902E-2	9.6372878E-3
2 =8	3.5687747E-3	6.3107096E-3
3 =7	4.4353963E-3	3.0671102E-3
4 = 6	3.1230101E-3	6.8871371E-3
5	1.1807439E-3	8.8422068E-3
9	1.5868396E-3	9.6269149E-3
10 = 16	3.5638065E-3	6.2972255E-3

Point	$\lvert \vec{r} \rvert$	$\lvert \vec{v} \rvert$
11 = 15	4.4311610E-3	3.0420081E-3
12 = 14	3.1167481E-3	6.8773313E-3
13	1.6812778E-3	8.8354003E-3
17	1.8995584E-3	9.9311651E-3
18 = 24	3.7184239E-3	6.6862503E-3
19 = 23	4.5552299E-3	3.7133149E-3
20 = 22	3.2980301E-3	7.1606299E-3
21	1.6047733E-3	9.0336156E-3
25	3.1623802E-3	1.0822374E-2
26 = 32	4.1525897E-3	7.7620434E-3
27 = 31	4.9159963E-3	5.2424784E-3
28 = 30	3.8035198E-3	7.9442535E-3
29	2.5204487E-3	9.5971956E-3

After having all those points calculated, the approach used to find the minimum transfer maneuvers is the following: i) Choose one point in the initial orbit and one point in the final orbit; ii) Specify a time for the transfer; iii) Solve the Lambert' s Problem for this pair of points and transfer time. Step ii) and iii) are repeated for a series of different transfer times to show the minimum fuel consumption maneuver between the two points considered. After that the points are varied in both initial and final orbits to search for the minimum transfer between the two orbits. Of course that this procedure gives an approximation of the minimum, since it uses a discretization of the two orbits involved. This accuracy increases with the number of points considered in the orbits and this process of division has no limit.

As an example, transfers starting at the point " a" in the initial orbit are considered. All the other four Halo orbits that were determined are studied as the final orbit of the maneuver. Table 4 and 5 show the results. The step of time used to solve the Lambert' s problem was 0.1. The values shown are the following: T is the transfer time in canonical units, ΔV is the total increment of velocity (the two impulses) required for the transfer, α_1 is the angle between the velocity of the spacecraft when it is in the initial halo orbit and its velocity at the same point when it is in the transfer orbit at the point when the first impulse is applied, α_2 is the angle between the velocity of the spacecraft when it is in the final halo orbit and its velocity at the same point when it is in the transfer orbit at the point when the second impulse is applied and ϕ is the transfer angle, that is the angle that the spacecraft travels during the transfer. All the angles are expressed in degrees. Fig. (3) shows all the pair of orbits involved in the transfer, as seen in the y-z plane of the rotating frame. The scales are different from plot to plot to make a better view of the transfers. The variation in the x axis is small, and the transfers are close to planar.

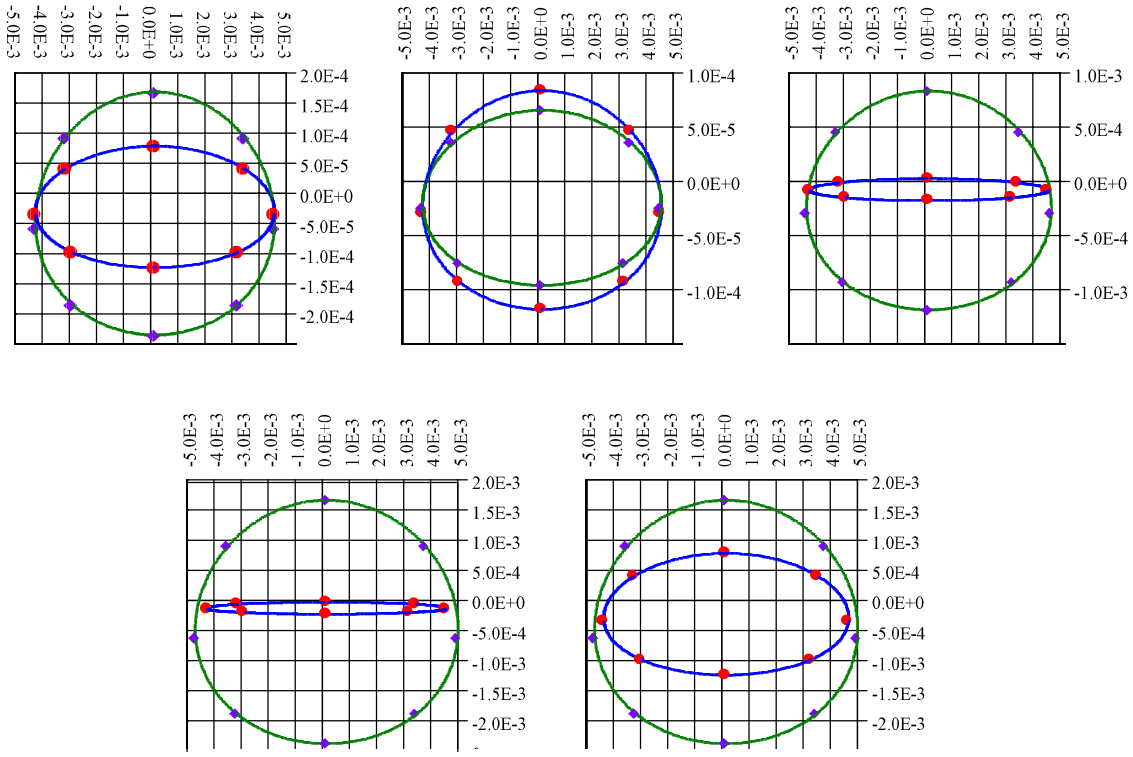


Figure. 3. Orbits involved in the transfer, as seen in the y (horizontal)-z (vertical) plane of the rotating frame.

The results show that the minimum transfer is located in positions that have an angle of transfer close to 90 degrees, with the time of transfer always equal to 0.4 or 0.8. Those minimums correspond to trajectories that have smaller values for the angles α_1 and α_2. This rule is valid for all the transfers. Simulations for transfers starting at points different than "a" in the initial orbit are not shown here to save space, but they have similar results.

Next, Table 5 shows all the minimum transfers that were found as a function of the final orbit. In this Table the step of time used to solve the Lambert's Problem was reduced to 0.01, to search for the minimum with more accuracy. This reduction showed to be very effective in searching transfers with less fuel consumption, because it finds transfers that has smaller angles α_1 and α_2. It shows that when the amplitude of the final orbit increases by a large factor (transfers from 0.001 to 0.100 or higher) the cost of the maneuver increases with the amplitude, as expected. But, when the differences are not too large (transfers from 0.001 to 0.002) other factors can interfere, like the possibility of finding transfers with smaller values for the angles α_1 and α_2.

Table 4. Minimum fuel consumption transfers.

a to:	T	Δv	α_1	α_2	ϕ
1	1.3	0.0182788	68.67	68.44	0
2	0.4	0.0005760	0.97	2.38	88
3	0.8	0.0005426	0.39	8.94	91
4	1.1	0.0013703	2.77	4.89	93
5	1.5	0.0034253	10.84	9.19	180
6	0.8	0.0185753	104.98	41.38	267
7	0.9	0.0183725	122.34	18.92	269
8	1.2	0.0176982	123.22	8.51	272
9	1.3	0.0180844	67.78	67.80	0
10	0.4	0.0004091	0.28	0.57	89
11	0.8	0.0004887	0.31	8.21	90
12	1.1	0.0012671	2.24	4.44	91
13	1.5	0.0019600	4.57	2.83	180
14	0.8	0.0185510	104.96	41.23	269
15	0.9	0.0183067	122.31	16.06	270
16	1.2	0.0174681	123.21	5.21	271
17	0.3	0.0210007	87.69	95.52	0
18	0.4	0.0037769	8.17	20.91	81
19	0.8	0.0024066	2.94	30.29	93
20	1.1	0.0047056	13.59	19.75	106
21	1.3	0.0120948	43.31	36.22	180
22	0.8	0.0191311	105.51	42.99	255
23	0.9	0.0194998	123.19	37.94	267
24	1.1	0.0197938	122.11	31.11	279
25	0.8	0.0211989	83.256	72.19	0
26	0.5	0.0077145	16.03	38.37	75
27	0.8	0.0052961	6.98	47.90	96
28	0.6	0.0185206	112.40	9.33	118
29	1.1	0.0157302	58.76	46.06	180
30	0.8	0.0206828	106.92	45.98	242
31	0.9	0.0218282	125.58	50.62	264
32	0.5	0.0240766	139.26	43.33	285
17 to:	**T**	**Δv**	**α_1**	**α_2**	**ϕ**
25	0.5	0.0215656	86.46	78.06	0
26	0.4	0.0043818	9.08	19.98	75
27	0.9	0.0073551	4.08	87.63	96
28	1.1	0.0054996	15.67	19.15	118
29	1.3	0.0136204	48.50	37.27	180
30	0.9	0.0215394	104.17	64.18	242
31	0.9	0.0235611	124.55	97.02	264
32	1.0	0.0246829	122.34	67.23	285

Table 5. Minimum fuel consumption transfers for different amplitudes.

From a to:	T	Δv	α_1	α_2	ϕ
3	0.77	0.000264388	0.31	3.94	91
10	0.39	0.000060374	0.03	0.28	89
19	0.77	0.002348159	2.90	29.34	93
27	0.76	0.004047831	0.45	45.70	96
From 17 to:	**T**	**Δv**	**α_1**	**α_2**	**ϕ**
26	0.41	0.004341998	9.26	19.41	75

6. Conclusions

An analytical algorithm was implemented to find Halo orbits around the Lagrangian point L_1 of the Earth-Moon system. The method worked very well and several orbits were found. After that, an algorithm based in the Lambert´s Problem was adapted and implemented to solve the problem of transfer orbits between two Halo orbits of the same family using a bi-impulsive control. The results showed several transfers that can be identified by the transfer time. From this family, it is possible to find the transfer with minimum cost. The general behavior of the minimum transfers is that the impulses are applied as close as possible to the direction of motion of the spacecraft in the Halo orbits and that the angle the spacecraft travels during the transfer is close to 90 degrees. The cost increases with the difference in amplitude of the Halo orbits, with the exception of orbits with small difference, where other factors can dominate the cost. Those results can be used by mission designers to plan missions that include orbital transfers like that.

7. Acknowledgment

The authors are grateful to the Foundation to Support Research in the São Paulo State (FAPESP) for the research grants received under Contracts 2000/14769-4, 2000/7074-0 and 1999/08740-4 and to CNPq (National Council for Scientific and Technological Development) - Brazil for the contract 300221/95-9.

8. References

Breakwell, J. V., and Brown, J. V., 1979, The Halo Family of 3-Dimensional Periodic Orbits in the Earth-Moon Restricted 3-Body Problem, Celestial Mechanics, vol. 20, pp. 389-404.

D' Amario, L. D. D., and Edelbaum, T. N., 1974, Minimum impulse three body trajectories, AIAA, Journal 12, 455-462.

Farquhar, R. W. and Kamel, A. A., 1973, Quasi-periodic orbits about the transfer libration point, Celestial Mechanics, vol. 7, pp. 458-474.

Farquhar, R. W., 1970, The Control and use of Libration Point Satellites, NASA, TR R-346, September.

Farquhar, R. W., 1970, The Moon' s Influence on the Location of the Sun-Earth Exterior Libration Point, Celestial Mechanics, vol.2.

Farquhar, R. W., 1972, A Halo-Orbit Lunar Station, Astronautics and Aeronautics, pp. 59-63.

Farquhar, R. W., 1980, Trajectories and orbital maneuvers for the first libration-point satellite, Journal of Guidance, Control and Dynamics, vol. 3, pp. 549-554.

Farquhar, R. W., 1991, Halo-Orbit and Lunar-Swingby Missions of the 1990' s, Acta Astronautica, vol. 24, 1991, pp. 227-234.

Farquhar, R. W., Muhonen, D. P., and Richardson, D. L., 1977, Mission Design for a Halo Orbiter of the Earth, Journal of Spacecraft and Rockets, vol. 14, n. 13, pp. 170-177, March.

Farquhar, R. W., Muhonen, D. P., Newman, C. and Heuberger, H., 1980, Trajectories and orbital Maneuvers for the first Libration Point Satellite, Journal of Guidance, Control and Dynamics, vol. 3, n. 6, pp. 549-554.

Gingiss, A. J., Navigation Analysis of Earth-Moon Libration Point Missions, M.S. Thesis, Massachusetts Inst. of Techology, The Charles Stark Draper Laboratory Rep. CSDL-T-1133, Cambrige, MA, June 1992.

Gómez, G., Jorba, A., Masdemont, J. and Simó, C., 1993, Study of transfer from the Earth to a Halo Orbit Around the Equilibrium Point L_1, Celestial Mechanics and Dynamical Astronomy, vol. 56, pp. 541-562.

Gordon, S. C., and Howell, K. C., 1992, Orbit Determination Error Analysis and Comparison of Station-Keeping Costs for Lissajous and Halo-Type Libration Point orbits, Advances in Astronautical Sciences, vol. 79, Part. I, 1992, pp. 117-136.

Hiday, L. A., Howell, K. C., 1992, Transfers between libration-point orbits in the elliptic restricted problem. AAS/AIAA Spaceflight Mechanics Meeting, paper AAS 92-126, Colorado.

Howell, K. C, and Gordon, S. C., 1992, Orbit Determination Error Analysis and Station Keeping Strategy for Sun-Earth L_1 Libration Point Orbits, School of Aeronautics and Astronautics, Purdue University, West Lafayette, Indiana.

Howell, K. C, and Pernicka, H. J., 1993, Station Keeping Method for Libration Point Trajectories, Journal of Guidance, Control, and Dynamics, vol. 16, n. 1, pp. 151-159, January-February.

Howell, K. C. and Breakwell, J. V., 1984, Almost Rectilinear Halo Orbits, Celestial Mechanics, vol. 32, pp. 29-52.

Howell, K. C., Mains, D. L., and Barden B. T., 1994, Transfer Trajectories from Earth Parking Orbits to Sun-Earth Halo Orbits, AAS paper 94-160, AAS/AIAA Spacefligtht Mechanics Meeting, Cocoa Beach, FL, February 14-16.

Jones, B. L., and Bishop, R. H., 1993, H_2 Optimal Halo Orbit Guidance, Journal of Guidance, Control, and Dynamics, vol. 16, n.6, pp. 1118-1124.

Jones, B. L., and Bishop, R. H., 1993, Stable Orbit Rendezvous for a Small Radius Translunar Halo Orbit, Advances in Astronautical Sciences, vol. 82, Part. I, pp. 585-604.

Jones, B. L., and Bishop, R. H., 1994, Rendezvous Targeting and Navigation for a Translunar Halo Orbit, Journal of Guidance, Control and Dynamics, vol. 17, n. 5, pp. 1109-1114, September-October.

Popescu, M. and Cardos, V., 1995, The Domain of initial Conditions for the Class of three-dimensional Halo Periodical Orbits, Acta Astronautical, vol. 36, n.4, pp. 193-196.

Popescu, M., 1985, Optimal transfer from Lagrangian points, Acta Astronautica, vol. 12, pp. 225-228.

Popescu, M., 1986, Auxiliary problem concerning optimal pursuit on Lagrangian orbits, Journal of Guidance, Control and Dynamics, vol. 9, pp. 717-719.

Press, W. H.; Flannery, B. P.; Teukolsky S. A. and Vetterling, W. T., 1989, Numerical Recipes, Cambridge University Press, NewYork.

Richardson, D. L., 1980, Halo Orbit Formulation for the ISEE-3 Mission, Journal Guidance and Control, vol. 3, n. 6, pp. 543-548.

Richardson, D. L., 1980a, Analytic Construction of periodic orbits about the collinear points, Celestial Mechanics, vol. 22, pp. 241-253.

Richardson, D. L., 1980b, A note on a Lagrangian formulation for motion about the collinear points, Celestial Mechanics, vol. 22, pp. 231-236.

Simó, C., Gómez, G., Llibre J., Martínez R. and Rodríguez J., 1987, On the Optimal Station Keeping Control of Halo Orbits, Acta Astronautica, vol. 15, n. 6/7, pp. 193-197.

Stalos, S., Folta, D., Short, B., Jen, J., and Seacord, A., 1993, Optimum Transfer to a Large-Amplitude Halo Orbit for the Solar and Heliospheric Observatory (SOHO) Spacecraft, paper n. 93-553, AAS/GSFC International Symposium on Space Flight Dynamics, April.

Starchville, Jr. T. F., Melton, G. R., 1997, Optimal Low-Thrust Trajectories to Earth-Moon L_2 Halo Orbits (Circular Problem), AAS/AIAA Astrodynamics Specialist Conference, Sun Valley, Idaho, August 4-7, paper 97-714.

Szebehely, V., 1967, Theory of Orbits, Academic Press, New York.

XVI CONGRESSO BRASILEIRO DE ENGENHARIA MECÂNICA
16th BRAZILIAN CONGRESS OF MECHANICAL ENGINEERING

ORBITAL CONTROL OF A SATELLITE USING THE GRAVITY OF THE MOON

Antonio Fernando Bertachini de Almeida Prado
Instituto Nacional de Pesquisas Espaciais - INPE - Brazil
São José dos Campos - SP - 12227-010 - Brazil
E-mail: PRADO@DEM.INPE.BR

Abstract*. In this paper, a study is made in the problem of the orbital control of an Earth´s satellite using the gravity of the Moon. The main objective is to study a technique to decrease the fuel consumption of a plane change maneuver to be performed in a satellite that is in orbit around the Earth. The main idea of this approach is to send the spacecraft to the Moon using a single impulsive maneuver, use the gravity field of the Moon to make the desired plane change of the trajectory, and then return the spacecraft to its nominal semi-major axis and eccentricity using a bi-impulsive Hohmann type maneuver. The spacecraft is assumed to start in a circular orbit in the plane of the lunar orbit around the Earth and the goal is to put it in a similar orbit that differs from the initial orbit only by the inclination. A description of the close approach maneuver is made in the three-dimensional space. Analytical equations based in the patched conics approximation are used to calculate the variation in velocity, angular momentum, energy and inclination of the spacecraft that realizes this maneuver. Then, several simulations are made to evaluate the savings involved.*

Key Words: *Astrodynamics, Orbital Maneuver, Swing-by.*

1. Introduction

The literature shows several applications of the swing-by technique. Some of them can be found in Swenson (1992), that studied a mission to Neptune using swing-bys to gain energy to accomplish the mission; Weinstein (1992), that made a similar study for a mission to Pluto; Farquhar and Dunham (1981), that formulated a mission to study the Earth's geomagnetic tail; Farquhar, Muhonen and Church (1985), Efron, Yeomans, and Schanzle (1985) and Muhonen, Davis, and Dunham (1985), that planned the mission ISEE-3/ICE; Flandro (1966), that made the first studies for the Voyager mission; Byrnes and D'Amario (1982), that design a mission to flyby the comet Halley; D'Amario, Byrnes and Stanford (1981 and 1982) that studied multiple flyby for interplanetary missions; Marsh and Howell (1988) and Dunham and Davis (1985) that design missions with multiple lunar swing-bys; Prado and Broucke (1994), that studied the effects of the atmosphere in a swing-by trajectory; Striepe and Braun (1991), that used a swing-by in Venus to reach Mars; Felipe and Prado (1999), that studied numerically a swing-by in three dimensions, including the effects in the inclination; Prado (1996), that considered the possibility of applying an impulse during the passage by the periapsis; Prado and Broucke (1995), that classified trajectories making a swing-by with the Moon. The most usual approach to study this problem is to divide the problem in three phases dominated by the "two-body" celestial mechanics. Other models used to study this problem are the circular restricted three-body problem (like in Broucke (1988), Broucke and Prado (1993), and Prado (1993)) and the elliptic restricted three-body problem (Prado (1997)).

The goal of this paper is to use analytical equations for the variations of velocity, energy, angular momentum and inclination to find an economical strategy to change the inclination of the orbit of a spacecraft by using a close approach with the Moon. The idea is to replace the expensive impulse required to make the plane change by a series of three impulses: the first one used to send the spacecraft in a planar trajectory to the Moon, to use the swing-by to make a free plane change and then to use the two remaining impulses to make a Hohmann type planar transfer to change the semi-major axis and eccentricity that resulted from the swing-by to the desired nominal values. This maneuver was used by Hughes to save the AsiaSat 3 satellite after a failure with the rocket Proton, that delivered the spacecraft in a wrong orbit. Another advantage of this approach is that it allows the possibility of studying the Moon, making it a double objective mission. This maneuver will be compared with a direct single impulsive maneuver. There are other maneuvers that consider more impulses, like the bi-elliptic transfer with plane change, but they are not considered in this paper. Among the several sets of initial conditions that can be used to identify uniquely one swing-by trajectory, the following five variables are used (see Fig. (1)): V_p, the velocity of the spacecraft at periapsis of the orbit around the secondary body; two angles (α and β), that specify the direction of the periapsis of the trajectory of the spacecraft around the secondary body in a three-dimensional space; r_p the distance from the spacecraft to the center of the secondary body in the moment of the closest approach to this secondary body (periapsis distance); γ, the angle between the velocity vector at periapsis and the intersection between the horizontal plane that passes by the periapsis and the plane perpendicular to the periapsis that holds $\vec{V}_p$.

Figure (1) also shows the sequence for this maneuver. It is assumed that the system has three bodies: a primary (M_1, the Earth) and a secondary (M_2, the Moon) bodies with finite masses that are assumed to be in circular orbits around their common center of mass and a third body with negligible mass (the spacecraft) that has its motion governed by the two other bodies. The spacecraft leaves the point A, passes by the point P (the periapsis of the trajectory of the spacecraft in its orbit around M_2) and goes to the point B. The points A and B are chosen in a such way that the influence of the Moon at those two points can be neglected and, consequently, the energy can be assumed to remain constant after B and before A (the system follows the two-body celestial mechanics). The initial conditions are clearly identified in Fig. (1). The distance r_p is not to scale, to make the figure easier to understand. The result of this maneuver is a change in velocity, energy, angular momentum and inclination in the Keplerian orbit of the spacecraft around the central body.

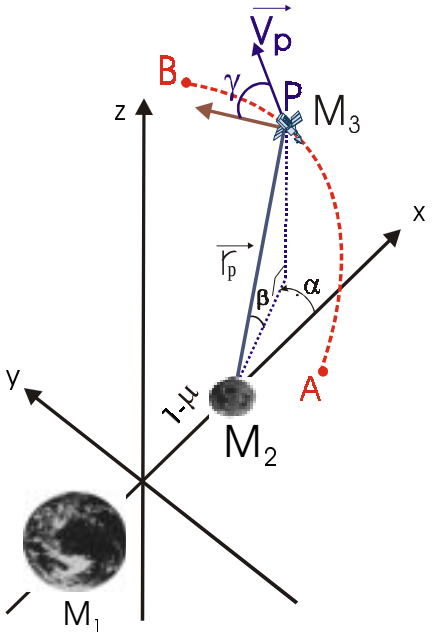


Figure 1. The Swing-By in Three Dimensions.

2. Mathematical model

To start the description of the mathematical model used in this research, we briefly describe the orbital change of a single particle subjected to a close approach with the Moon under the "patched-conics" model. It is assumed that the particle is in an initial orbit around the Earth with given semi-major axis (a) and eccentricity (e) such that it crosses the orbit of the Moon and that this orbit is coplanar with the orbit of the Moon. Then, the swing-by is assumed to be planar for the equations developed below. The periapsis distance (r_p) for the swing-by is assumed to be known. The canonical system of units is used, and it implies that: i) The unit of distance is the distance between the Earth and the Moon; ii) The angular velocity of the motion of the Earth and the Moon is assumed to be one; iii) The mass of the Moon is given by $\mu = \dfrac{m_2}{m_1 + m_2}$ (where m_1 and m_2 are the real masses of the Earth and the Moon, respectively) and the mass of the Earth is $(1-\mu)$, so the total mass of the system is one; iv) The unit of time is defined such that the period of the motion of the Earth and the Moon is 2π; v) The gravitational constant is one.

The first step is to obtain the magnitude of the first impulse, the one used to send the spacecraft from its initial parking orbit around the Earth to the Moon. It is assumed that the initial orbit around the Earth is circular with semi-major axis a_0 and that the transfer orbit is chosen to have semi-major axis a and eccentricity e. The minimum value for the semi-major axis that allows a transfer to the Moon is given by $(1+a_0)/2$. In this way, the magnitude of the impulse required for this transfer and the eccentricity of the transfer orbit are given by:

$$\Delta V_1 = \sqrt{\frac{2(1-\mu)}{a_0} - \frac{(1-\mu)}{a}} - \sqrt{\frac{(1-\mu)}{a_0}} \qquad (1)$$

$$e = 1 - \frac{a_0}{a} \tag{2}$$

Then, it is possible to calculate the magnitude of the velocity of the particle with respect to the Earth in the moment of the crossing with the Moon's orbit (V_i), as well as the true anomaly of that point (θ). They come from

$$V_i = \sqrt{(1-\mu)\left(2 - \frac{1}{a}\right)} \tag{3}$$

$$\theta = \cos^{-1}\left[\frac{1}{e}\left(a(1-e^2) - 1\right)\right] \tag{4}$$

In those equations the fact that the distance between the Earth and the Moon is one was used and we took only the positive value of the true anomaly. Next, it is calculated the angle between the inertial velocity of the particle and the velocity of the Moon (the flight path angle γ), as well as the magnitude of the velocity of the particle with respect to the Moon in the moment that the approach starts (V_∞). They are given by (assuming a counter-clock-wise orbit for the particle)

$$\gamma = \tan^{-1}\left[\frac{e\sin\theta}{1 + e\cos\theta}\right] \tag{5}$$

$$V_\infty = \sqrt{V_i^2 + V_2^2 - 2V_i\, V_2\, \cos\gamma} \tag{6}$$

Later we can use the fact that the velocity of the Moon around the Earth (V_2) is one. Figure (2) shows the vector addition used to derive the equations.

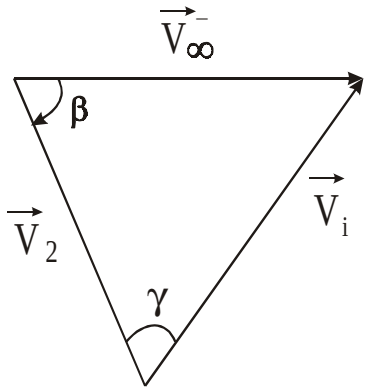


Figure 2. Vector addition during the close-approach.

The angle β shown in Fig. (2) is given by

$$\beta = \cos^{-1}\left[-\frac{V_i^2 - V_2^2 - V_\infty^{-2}}{2V_2\, V_\infty^-}\right] \tag{7}$$

Those information allow us to obtain the turning angle (2δ) of the particle around the Moon, from

$$\delta = \sin^{-1}\left(1 \Big/ 1 + \frac{r_p V_\infty^2}{\mu}\right) \tag{8}$$

The angle of approach (α) has two values, depending if the particle is passing in front or behind Jupiter. These two values will be called α_1 and α_2. They are obtained from $\alpha_1 = \pi + \beta + \delta$ and $\alpha_2 = 2\pi + \beta - \delta$. In this paper only the first value is used, because the second one does not improve the efficiency of the swing-by maneuver.

Assuming that the spacecraft is coming from the same plane that the Moon orbits around the Earth and that the angle β and the periapsis distance r_p are used as control variables of the maneuver (they can be reached with maneuvers that has negligible consumption), we can obtain the following relation for the angle γ

$$\gamma = \sin^{-1}\left[-\mathrm{tg}(\delta)\mathrm{tg}(\beta)\right] \tag{9}$$

Now it is necessary to have equations for the inertial velocities of the spacecraft before ($\vec{V}_i$) and after ($\vec{V}_0$) the close approach. They are (Prado, 2000)

$$\vec{V}_i = \vec{V}_\infty^- + \vec{V}_2 = V_\infty \sin\delta(\cos\beta\cos\alpha, \cos\beta\sin\alpha, \sin\beta) + V_\infty \cos\delta(-\sin\gamma\sin\beta\cos\alpha - \cos\gamma\sin\alpha, -\sin\gamma\sin\beta\sin\alpha + \\ + \cos\gamma\cos\alpha, \cos\beta\sin\gamma) + (0, V_2, 0) \tag{10}$$

$$\vec{V}_0 = \vec{V}_\infty^+ + \vec{V}_2 = -V_\infty \sin\delta(\cos\beta\cos\alpha, \cos\beta\sin\alpha, \sin\beta) + V_\infty \cos\delta(-\sin\gamma\sin\beta\cos\alpha - \cos\gamma\sin\alpha, -\sin\gamma\sin\beta\sin\alpha + \\ + \cos\gamma\cos\alpha, \cos\beta\sin\gamma) + (0, V_2, 0) \tag{11}$$

From those equations, it is possible to obtain expressions for the variations in velocity, energy and angular momentum of the spacecraft. They are

$$\Delta\vec{V} = \vec{V}_0 - \vec{V}_i = -2V_\infty \sin\delta(\cos\alpha\cos\beta, \cos\beta\sin\alpha, \sin\beta) \tag{12}$$

which implies that

$$\Delta V = \left|\Delta\vec{V}\right| = 2V_\infty \sin\delta \tag{13}$$

$$\Delta E = \frac{1}{2}\left(V_0^2 - V_i^2\right) = -2V_2 V_\infty \cos\beta\sin\alpha\sin\delta \tag{14}$$

For the angular momentum ($\vec{C}$) the results are:

$$\vec{C}_i = \vec{R}\times\vec{V}_i = d\,V_\infty(0, -\sin\beta\sin\delta + \cos\beta\cos\delta\sin\gamma, \frac{V_2}{V_\infty} + \cos\alpha\cos\delta\cos\gamma + \cos\beta\sin\alpha\sin\delta - \cos\delta\sin\alpha\sin\beta\sin\gamma) \tag{15}$$

$$\vec{C}_0 = \vec{R}\times\vec{V}_0 = d\,V_\infty(0, \sin\beta\sin\delta - \cos\beta\cos\delta\sin\gamma, \frac{V_2}{V_\infty} + \cos\alpha\cos\delta\cos\gamma - \cos\beta\sin\alpha\sin\delta - \cos\delta\sin\alpha\sin\beta\sin\gamma) \tag{16}$$

Where $\vec{R} = (d, 0, 0)$ is the position vector of the Moon. Then

$$\Delta\vec{C} = \vec{C}_0 - \vec{C}_i = 2d\,V_\infty \sin\delta(0, \sin\beta, -\cos\beta\sin\alpha) \tag{17}$$

and

$$\left|\Delta\vec{C}\right| = 2dV_\infty \sin\delta\left(\cos^2\beta\sin^2\alpha + \sin^2\beta\right)^{1/2} \tag{18}$$

Using the definition of angular velocity $\omega = \frac{V_2}{d}$ it is possible to get

$$\omega\Delta C_Z = -2V_2 V_\infty \cos\beta\sin\alpha\sin\delta = \Delta E \tag{19}$$

The inclination before the close approach is assumed to be zero and the equation for the inclination after the passage is obtained by the following set of equations

$$\left|\vec{C}_o\right| = dV_\infty\left((\sin\beta\sin\delta - \cos\beta\cos\delta\sin\gamma)^2 + \left(\frac{V_2}{V_\infty} + \cos\alpha\cos\delta\cos\gamma - \cos\beta\sin\alpha\sin\delta - \cos\delta\sin\alpha\sin\beta\sin\gamma\right)^2\right)^{1/2} \tag{20}$$

$$C_{oZ} = dV_\infty \left(\frac{V_2}{V_\infty} + \cos\alpha\cos\delta\cos\gamma - \cos\beta\sin\alpha\sin\delta - \cos\delta\sin\alpha\sin\beta\sin\gamma \right) \tag{21}$$

$$\mathrm{Cos}(i_o) = \frac{C_{oZ}}{|\vec{C}_o|} = \frac{1}{\sqrt{1 + \left(\dfrac{\sin\beta\sin\delta - \cos\beta\cos\delta\sin\gamma}{\dfrac{V_2}{V_\infty} + \cos\alpha\cos\delta\cos\gamma - \cos\beta\sin\alpha\sin\delta - \cos\delta\sin\alpha\sin\beta\sin\gamma} \right)^2}} \tag{22}$$

Where $\vec{C}_o$ is the final angular momentum and i_o is the final inclination. The subscript Z stands for the z-component of the angular momentum. Equation (22) is a function of the two variables used as the control for the maneuver: the angle β and the periapsis distance r_p. The values for the other variables are known: $V_2 = 1$, V_∞ is given by Eq. (6), δ is given by Eq. (8), γ is given by Eq. (9) and α is obtained from $\alpha = \pi + \beta + \delta$. After all those considerations and also using the values $a = 0.51$; $a_0 = 0.017$, the final equation for the inclination is given by

$$\begin{aligned}
i_o = \mathrm{Cos}^{-1}\Bigg[\Bigg(& 1 - \frac{0.822\mathrm{Cos}[\beta]\mathrm{Sin}\left[3.234 + \mathrm{Sin}^{-1}\left[\frac{1}{1+55.807rp}\right]\right]}{1+55.807r_p} + \frac{0.822\mathrm{Sin}[\beta]\mathrm{Sin}\left[3.234 + \mathrm{Sin}^{-1}\left[\frac{1}{1+55.807rp}\right]\right]\mathrm{Tan}[\beta]}{1+55.807r_p} \\
& + 0.822\sqrt{1 - \frac{1}{(1+55.807r_p)^2}}\,\mathrm{Cos}\left[3.234 + \mathrm{Sin}^{-1}\left[\frac{1}{1+55.807r_p}\right]\right]\sqrt{1 - \frac{\mathrm{Tan}^2[\beta]}{(1+55.807r_p)^2\left(1 - \frac{1}{(1+55.807r_p)^2}\right)}}\Bigg) \\
& \Bigg(\frac{2.701\mathrm{Sin}^2[\beta]}{(1+55.807r_p)^2} + \Bigg(1 - \frac{0.822\mathrm{Cos}[\beta]\mathrm{Sin}\left[3.234 + \mathrm{Sin}^{-1}\left[\frac{1}{1+55.807r_p}\right]\right]}{1+55.807r_p} + \frac{0.822\mathrm{Sin}[\beta]\mathrm{Sin}\left[3.234 + \mathrm{Sin}^{-1}\left[\frac{1}{1+55.807r_p}\right]\right]\mathrm{Tan}[\beta]}{1+55.807r_p} \\
& + 0.822\sqrt{1 - \frac{1}{(1+55.807r_p)^2}}\,\mathrm{Cos}\left[3.234 + \mathrm{Sin}^{-1}\left[\frac{1}{1+55.807r_p}\right]\right]\sqrt{1 - \frac{\mathrm{Tan}^2[\beta]}{(1+55.807r_p)^2\left(1 - \frac{1}{(1+55.807r_p)^2}\right)}}\Bigg)^2\Bigg)^{-1/2}\Bigg]
\end{aligned} \tag{23}$$

Figure (3) shows the three-dimensional view of this function. The function is not defined for all values of the angle β and the periapsis distance r_p, because the equation has several square roots and inverse trigonometric functions that have limits of in their definitions.

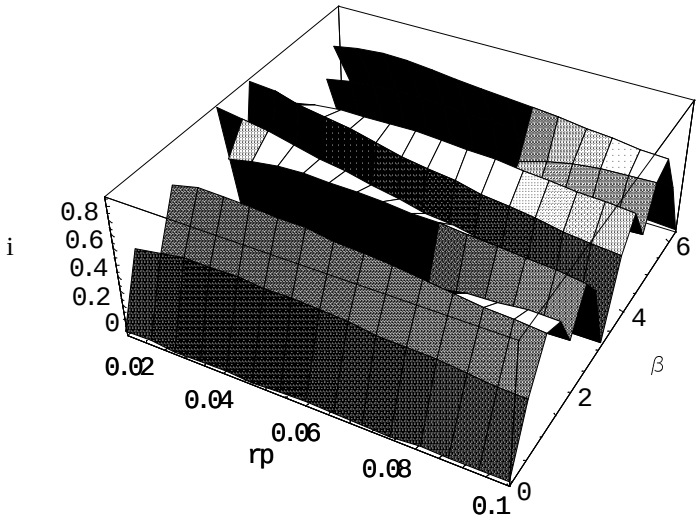


Figure 3. Inclination obtained by the swing-by (in radians) as a function of the angle β (in radians, at the right side of the figure) and the periapsis distance r_p (in canonical units, at the left of the figure).

After that it is necessary to obtain the semi-major axis (a_f) and the eccentricity (e_f) of the orbit after the close approach. They are given by

$$a_f = \frac{(1-\mu)}{2(1-\mu) - V_0^2} \tag{24}$$

$$e_f = \sqrt{1 - \frac{C_0^2}{(1-\mu)a_f}} \tag{25}$$

With those quantities available, it is possible to obtain the impulses required to perform the maneuver in inclination of the spacecraft under two approaches:

i) The standard maneuver of applying one impulse in the spacecraft when it passes by the intersection of the two orbits. For this maneuver the impulse required is given by (Chobotov, 1996)

$$\Delta V_H = 2\sqrt{\frac{(1-\mu)}{a_0}}\sin\left(\frac{\Delta i}{2}\right) \tag{26}$$

where Δi is the inclination desired for the maneuver.

ii) The proposed maneuver of sending the spacecraft to the Moon in a planar maneuver, to use the close approach with the Moon to change the inclination of the orbit with no cost, and then performing a two-impulsive maneuver to change the semi-major axis (a_f) and the eccentricity (e_f) of the orbit after the close approach back to their original values a and zero. This bi-impulsive maneuver has a cost giving by

$$\Delta V_2 = \left| \sqrt{\frac{2(1-\mu)}{a_f(1+e_f)} - \frac{(1-\mu)}{a_f}} - \sqrt{\frac{2(1-\mu)}{a_f(1+e_f)} - \frac{(1-\mu)}{a_t}} \right| \tag{27}$$

$$\Delta V_3 = \left| \sqrt{\frac{2(1-\mu)}{a_0} - \frac{(1-\mu)}{a_t}} - \sqrt{\frac{(1-\mu)}{a_0}} \right| \tag{28}$$

where $a_t = \dfrac{a_f(1+e_f)+a_0}{2}$.

So, the total velocity increment is given by

$$\Delta V_t = \Delta V_1 + \Delta V_2 + \Delta V_3 \tag{29}$$

3. Results

Several simulations were performed with the equations shown above. Figure (4) shows the difference in fuel consumption ($\Delta V_t - \Delta V_H$), measured in canonical units of increment of velocity to be applied in the spacecraft (1 canonical unit is equal to 1.023 km/s) as a function of the angle β, in radians, for three values of the periapsis distance 0.0048 (1845 km), 0.0524 (20142 km) and 0.1000 (38440 km). It is visible that the difference in fuel consumption decreases with this variable, what is expected because the effects of the swing-by increases when the spacecraft is passing closer to the Moon and the inclination change is higher, increasing the savings of the maneuver that uses the swing-by. This is confirmed by Fig. (5), which shows the inclination change, in radians, obtained by the swing-by maneuver. Another property shown by this figure is that the interval of definition of the function increases when the value of the periapsis distance increases. In particular, it is interesting to note that when the periapsis distance goes from 0.0048 to 0.0524 two new regions of savings for the maneuver proposed appears, close to β = 1 and β = 5. It means that a very careful analysis of those parameters has to be made before making a final decision of a practical maneuver.

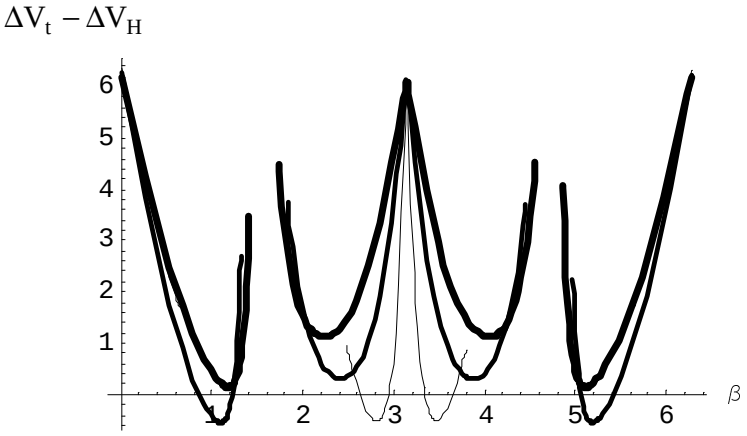


Figure 4. $\Delta V_t - \Delta V_H$ (canonical units) as a function of β (in radians) for a_0 = 0.017 and a = 0.51. The darkest line represents the results for r_p = 0.1, the intermediate line represents the results for r_p = 0.0524 and the thinnest line represents the results for r_p = 0.0048.

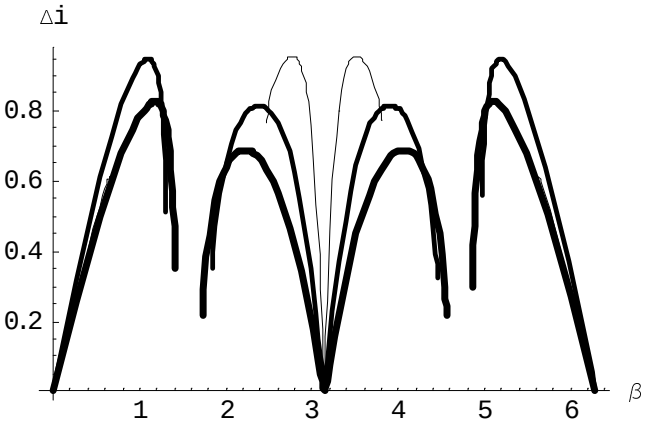


Figure 5. Inclination change (in radians) as a function of β (in radians) for a_0 = 0.017 and a = 0.51. The darkest line represents the results for r_p = 0.1, the intermediate line represents the results for r_p = 0.0524 and the thinnest line represents the results for r_p = 0.0048.

Next, the effect of the semi-major axis of the transfer orbit from the Earth to the Moon is studied. Figure (6) shows the same difference in fuel consumption ($\Delta V_t - \Delta V_H$) as a function of β for a_0 = 0.017 and r_p = 0.0048 and considering

three values for a: 0.51 (the minimum to achieve the orbit of the Moon), 0.56 and 0.61. It is clear that the savings of the swing-by maneuver increase when the value of this variable increases. Simulations with a larger number of values for this variable show that this behavior stays the same until the infinity. The reason is that the increase in fuel consumption in the first impulse required to obtain a larger value of the semi-major axis is small when compared to the gains obtained with a swing-by performed with a higher approaching velocity.

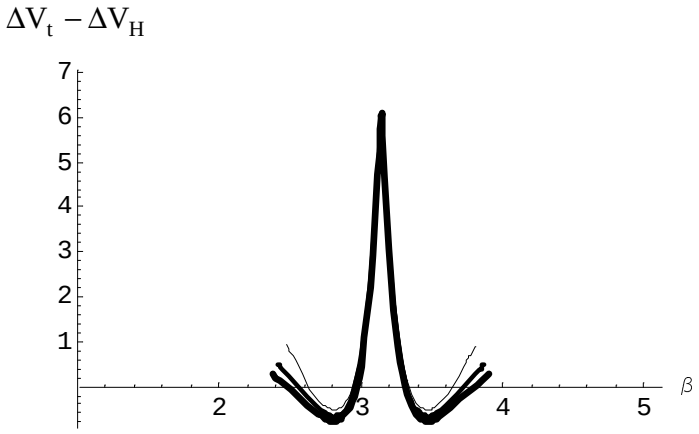


Figure 6. $\Delta V_t - \Delta V_H$ (in canonical units) as a function of β (in radians) for $a_0 = 0.017$ and $r_p = 0.0048$. The darkest line represents the results for a = 0.61, the intermediate line represents the results for a = 0.56 and the thinnest line represents the results for a = 0.51.

Then, the effects of varying the initial orbit of the spacecraft around the Earth are considered. Figure (7) shows the difference in fuel consumption ($\Delta V_t - \Delta V_H$) as a function of β for $r_p = 0.0048$ (1845 km), a = 0.51 (196044 km) and three values for the semi-major axis of the initial orbit: 0.017 (6535 km), 0.024 (9226 km) and 0.030 (11532 km). The results show that the advantages of the swing-by maneuver increases when the value of the initial semi-major axis decreases. This result is expected because the costs of the inclination change increase very much when the initial orbit gets closer to the Earth.

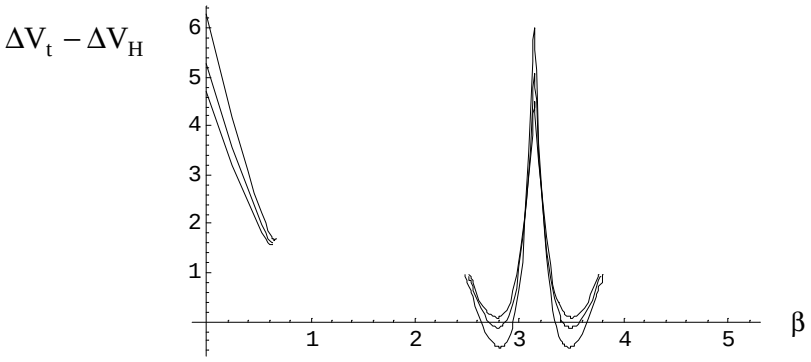


Figure 7. $\Delta V_t - \Delta V_H$ (in canonical units) as a function of β (in radians) for $r_p = 0.0048$ and a = 0.51. The bottom line represents the results for $a_0 = 0.017$, the intermediate line represents the results for $a_0 = 0.024$ and the superior line represents the results for $a_0 = 0.030$.

4. Conclusions

A set of analytical equations is used to describe the swing-by in three dimensions and to evaluate the variation in the inclination of the orbit of a spacecraft that is passing by the Moon. Then, it is possible to compare analytically two schemes to make an orbital maneuver that has only an inclination change in the orbit as its goal: i) the standard single impulse maneuver that applies the impulse in the crossing point between the two orbits and, ii) a three-impulsive transfer that applies the first impulse to send the spacecraft to the Moon, uses the gravity field of the Moon to perform a zero cost inclination change maneuver, and then makes two more impulses to accomplish a planar Hohmann transfer from the orbit that is obtained after the swing-by and the desired final orbit. The results showed that in several circumstances this second maneuver can be more economical than the standard maneuver, if a proper choice of the periapsis position is

made. It is also showed that this maneuver is more economical if: i) the initial parking orbit is close to the Earth, ii) the semi-major axis of the transfer orbit from the Earth to the Moon tends to infinity, iii) the periapsis of the close approach to the Moon has the smallest possible value. Then, this research can be used to design maneuvers for low Earth orbit satellites that requires larger inclination changes.

5. Acknowledgment

The author is grateful to the Foundation to Support Research in the São Paulo State (FAPESP) for the research grant received under Contract 2000/14769-4 and to CNPq (National Council for Scientific and Technological Development) - Brazil for the contract 300221/95-9.

6. References

Broucke, R.A. & A.F.B.A. Prado (1993). "Jupiter Swing-By Trajectories Passing Near the Earth", *Advances in the Astronautical Sciences*, Vol. 82, No 2, pp. 1159-1176.

Broucke, R.A. (1988). "The Celestial Mechanics of Gravity Assist", AIAA Paper 88-4220.

Byrnes, D.V. & L.A. D'Amario (1982) "A Combined Halley Flyby Galileo Mission," AIAA paper 82-1462. In: AIAA/AAS Astrodynamics Conference, San Diego, CA, Aug.

CARVELL, R. (1985), "Ulysses-The Sun From Above and Below," *Space*, Vol. 1, pp. 18-55, Dec. 85-Feb. 86.

Chobotov, V. A. (1996) *Orbital motion – Second Edition*. Reston: American Institute of Aeronautics and Astronautics, 1996.

D'Amario, L.A., D.V. Byrnes & R.H Stanford. (1981). "A New Method for Optimizing Multiple-Flyby Trajectories, " *Journal of Guidance, Control, and Dynamics*, Vol. 4, No 6, pp. 591-596.

D'Amario, L.A., D.V. Byrnes & R.H. Stanford (1982). "Interplanetary Trajectory Optimization with Application to Galileo," *Journal of Guidance, Control, and Dynamics*, Vol. 5, No. 5, pp. 465-471.

Dunham, D. & S. Davis (1985). "Optimization of a Multiple Lunar-Swing by Trajectory Sequence," *Journal of Astronautical Sciences*, Vol. 33, No. 3, pp. 275-288.

Efron, L., D.K. Yeomans & A.F. Schanzle (1985). "ISEE-3/ICE Navigation Analysis," *Journal of Astronautical Sciences*, Vol. 33, No. 3, pp. 301-323.

Farquhar, R., D. Muhonen & L.C Church (1985). "Trajectories and Orbital Maneuvers for the ISEE-3/ICE Comet Mission", *Journal of Astronautical Sciences*, Vol. 33, No. 3, pp. 235-254.

Farquhar, R.W. & D.W. Dunham (1981). "A New Trajectory Concept for Exploring the Earth's Geomagnetic Tail", *Journal of Guidance, Control and Dynamics*, Vol. 4, No. 2, pp. 192-196.

Felipe, G. & A.F.B.A. Prado (1999). "Classification of Out of Plane Swing-by Trajectories". *Journal of Guidance, Control and Dynamics*, Vol. 22, No. 5, pp. 643-649.

Felipe, G. and Prado, A.F.B.A., "Study of the Inclination Change in Three-Dimensional Swing-By Trajectories". 22nd International Symposium on Space Technology and Science, pp. 60 (00-j-02). Congress realized in 28-May to 04-June, 2000, Morioka, Japan.

Flandro, G. (1966). "Fast Reconnaissance Missions to the Outer Solar System Utilizing Energy Derived from the Gravitational Field of Jupiter," *Astronautical Acta*, Vol. 12, No. 4.

Marsh, S.M. & K.C. Howell (1988). "Double Lunar Swing by Trajectory Design," AIAA paper 88-4289.

Muhonen, D., S. Davis, & D. Dunham (1985). "Alternative Gravity-Assist Sequences for the ISEE-3 Escape Trajectory," *Journal of Astronautical Sciences*, Vol. 33, No. 3, pp. 255-273.

Prado, A.F.B.A. & R.A. Broucke (1994). "A Study of the Effects of the Atmospheric Drag in Swing-By Trajectories," *Journal of the Brazilian Society of Mechanical Sciences*, Vol. XVI, pp. 537-544.

Prado, A.F.B.A. & R.A. Broucke (1995). "A Classification of Swing-By Trajectories using The Moon". *Applied Mechanics Reviews*, Vol. 48, No. 11, Part 2, November, pp. 138-142.

Prado, A.F.B.A. (1993). "Optimal Transfer and Swing-By Orbits in the Two- and Three-Body Problems", Ph.D. Dissertation, Dept. of Aerospace Engineering and Engineering Mechanics, Univ. of Texas, Austin, TX.

Prado, A.F.B.A. (1996). "Powered Swing-By". *Journal of Guidance, Control and Dynamics*, Vol. 19, No. 5, pp. 1142-1147.

Prado, A.F.B.A. (1997). "Close-approach Trajectories in the Elliptic Restricted Problem", *Journal of Guidance, Control, and Dynamics*, Vol. 20, No. 4, pp. 797-802.

PRADO, A.F.B.A. (2000), "An Analytical Description of the Close Approach Maneuver in Three Dimensions". 51th International Astronautical Congress (IAF-00-A.5.05). Rio de Janeiro, Brazil, 02-06 October, 2000.

Striepe, S.A. & R.D. Braun (1991). "Effects of a Venus Swing by Periapsis Burn During an Earth-Mars Trajectory," *The Journal of the Astronautical Sciences*, Vol. 39, No. 3, pp. 299-312..

Swenson, B.L. (1992). "Neptune Atmospheric Probe Mission", AIAA Paper 92-4371.

Szebehely, V. (1967). *Theory of Orbits*, Academic Press, New York, Chap. 10.

Weinstein, S.S. (1992). "Pluto Flyby Mission Design Concepts for Very Small and Moderate Spacecraft", AIAA Paper 92-4372.

XVI CONGRESSO BRASILEIRO DE ENGENHARIA MECÂNICA
16th BRAZILIAN CONGRESS OF MECHANICAL ENGINEERING

SENSIBILIDADE DAS TRAJETÓRIAS DE CAPTURA GRAVITACIONAL TEMPORÁRIA

Ernesto Vieira Neto
FEG-UNESP - CP 205 Guaratinguetá CEP 12.500-000 SP – BRAZIL
ernesto@feg.unesp.br

Antonio Fernando Bertachini de Almeida Prado
Instituto Nacional de Pesquisas Espaciais - INPE
São José dos Campos - SP - 12227-010 - Brasil
Phone (123)256197 - Fax (123)25-6226
E-mail: PRADO@DEM.INPE.BR

Resumo. *Este trabalho estuda o problema de captura gravitacional utilizando o problema restrito circular de três corpos. Na captura gravitacional a energia do problema de dois corpos de uma partícula, muda de positiva para negativa. Essa mudança é temporária, depois de algum tempo ela volta a ser positiva e o veículo espacial escapa das proximidades do corpo celeste. A importância desta captura temporária é que ela pode ser usada para diminuir a quantidade de combustível gasto em missões que transferem um veículo espacial das proximidades de um planeta para as proximidades de um outro planeta, como o sistema Terra-Lua. O objetivo é aplicar um impulso durante a captura temporária e transforma-la numa captura permanente. O objetivo desse impulso é decrescer ainda mais a energia de dois corpos do veículo espacial. Neste trabalho é estudado o problema da determinação de trajetórias com mínima energia, bem como o comportamento das trajetórias de captura gravitacional quando ocorre um pequeno desvio na velocidade antes de ocorrer a captura gravitacional. Para realizar o estudo, um grande número de trajetórias são integradas numericamente sendo sistematicamente adicionado erros graduais no vetor velocidade inicial.*

Palavras chave:. *Astrodinâmica, Captura Gravitacional, Problema de Três Corpos, Manobras Orbitais.*

1. Introdução

A passagem de um veículo espacial entre duas órbitas é conhecida como transferência orbital. O método de transferência mais conhecido é o de Hohmann (1925). Porém, existem diversos outros métodos que podem ser aplicados, dependendo basicamente do tipo de problema a ser resolvido. Neste trabalho serão abordadas algumas das características que envolvem um tipo de transferência que foi desenvolvida no início da década de 90 por Belbruno e Miller (1990a,b), Yamakawa (1992) e outros. Esta transferência utiliza uma manobra de *swing-by* em torno da Lua, que impulsiona o veículo espacial para longe da influência gravitacional do sistema Terra-Lua. Nesta região é possível a aquisição de uma trajetória de captura gravitacional temporária em torno da Lua utilizando uma quantidade muito pequena de combustível. Dessa forma, pode-se economizar energia tanto no início da manobra quanto no final.

Nessa transferência, a captura gravitacional é uma das principais etapas dessa manobra e, no entanto, a mais desconhecida. É então feito uma análise numérica, seguindo as linhas de pesquisa dos trabalhos existentes na literatura até o momento. As análises paramétricas que foram feitas levam a conclusões importantes que poderão servir de indicação para um trabalho analítico futuro. Ainda assim, é feito um esforço no sentido de encontrar justificativas para os resultados numéricos obtidos no presente trabalho. Ficou demonstrado nos trabalhos de Yamakawa *et al.* (1993) e Belbruno (1987; 1990) que a utilização da captura gravitacional na manobra de transferência produz uma economia de combustível em relação à transferência de Hohmann.

O trabalho de Yamakawa (1992) foi o primeiro a sistematizar a forma de estudo do problema de captura gravitacional aplicado à astrodinâmica. A aplicação do conceito de energia de dois corpos leva a resultados importantes, facilitando o entendimento da captura gravitacional, que é estudado usando a dinâmica do problema restrito de três corpos. Seguindo estes conceitos, este trabalho faz uma busca de trajetórias de mínima energia, bem como uma análise da sensibilidade que o veículo espacial possui em relação às condições iniciais quando este entra na esfera de influência gravitacional da Lua.

Neste trabalho será analisado o que ocorre às trajetórias de captura gravitacional quando erros são introduzidos na velocidade inicial. A sensibilidade é uma forma de entender o quanto o comportamento das trajetórias é alterado quando pequenos erros nas condições iniciais ocorrem. Para isso as trajetórias que iniciam num ponto distante da Lua serão avaliadas numericamente. Como ainda não existem na literatura critérios para estabelecer a sensibilidade das trajetórias de captura gravitacional, resolveu-se aplicar inicialmente uma abordagem numérica.

2. Conceitos Básicos Relativos ao Problema de Captura Gravitacional

O fenômeno de captura é estudado utilizando o problema de três corpos. Muitos sistemas importantes podem ser modelados dessa forma, como no caso da Terra, Lua e um veículo espacial; ou ainda o Sol, um planeta e um veículo espacial; etc.. O modelo de três corpos pode ser dividido em duas categorias:

a) no problema geral de três corpos, onde cada corpo afeta mutuamente o movimento dos outros dois corpos;

b) no problema restrito de três corpos, onde um dos corpos tem massa desprezível em relação aos outros dois e não influencia o movimento desses.

No problema restrito ainda existem quatro possibilidades de estudo, correspondente ao movimento relativo entre os dois primários: a) hiperbólico; b) parabólico; c) elíptico; d) circular.

Neste trabalho será utilizada a dinâmica do problema restrito circular de três corpos, cuja dinâmica, no sistema baricêntrico girante, é apresentada a seguir (Roy, 1988):

$$\ddot{x} - 2\dot{y} - x = -(1-\mu)\frac{x+\mu}{r_1^3} - \mu\frac{x-1+\mu}{r_2^3} \qquad \ddot{y} + 2\dot{x} - y = -y\left(\frac{1-\mu}{r_1^3} + \frac{\mu}{r_2^3}\right) \tag{1}$$

onde $r_1^2 = (-\mu - x)^2 + y^2, r_2^2 = (1-\mu-x)^2 + y^2$ e μ é o parâmetro gravitacional do sistema.

A captura gravitacional referida neste trabalho é de caráter temporário, já que é impossível a captura gravitacional permanente neste modelo (ver Tanikawa (1983)). Mesmo assim, a captura gravitacional temporária no problema de transferência de um veículo espacial da Terra para a Lua é de grande importância. No momento da captura gravitacional o veículo espacial está com uma energia de dois corpos em relação à Lua menor, o que implica que o consumo de combustível para a estabilização da órbita (tornar a órbita circular em torno da Lua) é menor.

Para se ter uma idéia da captura gravitacional, a Fig. (1) mostra a mesma trajetória de captura no sistema baricêntrico girante e no sistema fixo, com origem no planeta Terra. Alguns dos termos que aparecerão com freqüência neste texto serão apresentados a seguir.

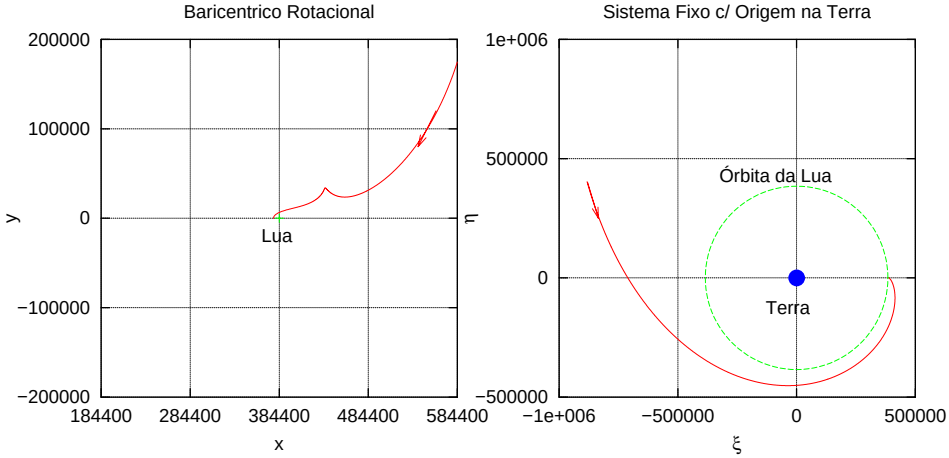


Figura 1. Trajetória de captura gravitacional

2.1. C3 (Energia de Dois Corpos)

Expressa pelo dobro da soma da energia cinética e da energia potencial do veículo espacial em relação a um dos primários (no caso deste trabalho, a Lua):

$$C3 = V^2 - \frac{2\mu}{r} \tag{2}$$

onde r e V são a distância e a velocidade do veículo espacial em relação à Lua, e μ é a massa adimensional (ou específica) da Lua. Para os sistemas de dois corpos C3 é constante, desde que nenhuma outra força, além da atração gravitacional do corpo central, exista. Nota-se que C3 = 2E, onde E é a energia total por unidade de massa do corpo M_3 dada pelo problema de dois corpos. Pode-se estabelecer a regra: se C3 > 0 a órbita será uma hipérbole; se C3 = 0 a órbita será uma parábola e se C3 < 0 a órbita será uma elipse.

2.2. Esfera de Influência

A esfera de influência da Lua será definida como uma esfera cujo raio é de 100.000 km. Esta esfera de influência está além da esfera de influência clássica (Roy, 1988), cujo raio é da ordem de 66.000 km para a Lua. Este número está próximo da distância crítica para a Lua, que é definida por $r = (2\mu)^{1/3}$ (Yamakawa, 1992), onde μ, neste caso, é a massa da Lua (o Apêndice A de Yamakawa , 1992, apresenta a demonstração da equação da distância crítica). A distância crítica marca a fronteira onde a energia de dois corpos C3 ainda é negativa, o que é uma das condições para a captura gravitacional.

2.3 Captura Gravitacional

Baseado no que foi colocado até agora, pode-se definir captura gravitacional de duas formas distintas (Yamakawa, 1992). Definição 1: "Embora o valor de C3 de um veículo espacial com respeito a Lua seja positivo fora da esfera de influência (ou seja, o veículo espacial está numa órbita hiperbólica), ele se torna negativo na periluna (a órbita é transformada em elíptica) sem o uso de forças não gravitacionais". Essa definição é conceitual. Uma segunda definição, mais adequada para análise numérica, é a definição 2: "Considera-se que ocorreu uma captura gravitacional quando ocorrem os seguintes fatos: a) C3 com respeito a Lua é negativo na periluna; b) A distância entre a Lua e o veículo espacial excede 100000 km antes de completar 50 dias, sem que haja uma colisão com a Lua". Nos cálculos computacionais que serão utilizados nesse trabalho para o estudo da captura gravitacional, esta segunda definição será aplicada, devido a sua objetividade e por ser praticável do ponto de vista computacional.

3. Descrição do procedimento

Foi utilizado o modelo do problema restrito circular de três corpos. Os cálculos foram feitos no sistema de coordenadas girantes. O objetivo deste trabalho é o de estudar trajetórias que terminem em captura gravitacional com a Lua. Para fazer isto, foram geradas várias trajetórias cujas características são explicadas abaixo.

a) Foi utilizado o sistema Terra-Lua como sistema de binários para o modelo dinâmico;

b) O movimento é estudado no plano do sistema Terra-Lua. Isso se deve ao fato de que foi demonstrado por Yamakawa (1992) que a captura fora desse plano não melhora as eventuais economias de combustível;

c) O ponto de chegada de cada trajetória fica a uma distância de 100 km da superfície da Lua (r_p = 1838 km a partir do centro da Lua) e na integração numérica o passo de tempo é negativo (portanto as condições ditas iniciais, são na realidade as condições finais da órbita, após a captura). Para especificar completamente a posição final, é necessário conhecer o valor do ângulo α, que é o ângulo da posição da periluna. Este ângulo é medido a partir da linha Terra-Lua, no sentido anti-horário, a partir do lado da Lua oposto a Terra (ver Fig. (2));

d) A magnitude da velocidade inicial V na periluna é calculada a partir do valor de C3 (Equação 3). A direção inicial da velocidade é escolhida como sendo perpendicular à linha que une o veículo espacial ao centro da Lua, apontando na direção anti-horária para as órbitas diretas e na direção horária para as órbitas retrógradas.;

e) É considerada como sendo uma órbita de escape (logo, de captura no sentido normal do tempo), uma órbita que satisfaça a definição 2 mostrada na seção anterior. A Fig. (2) mostra o ponto P onde o veículo espacial escapa da esfera de influência. O ângulo que define este ponto é chamado de ângulo da posição de entrada (Yamakawa, 1992) e é designado pela letra grega β.

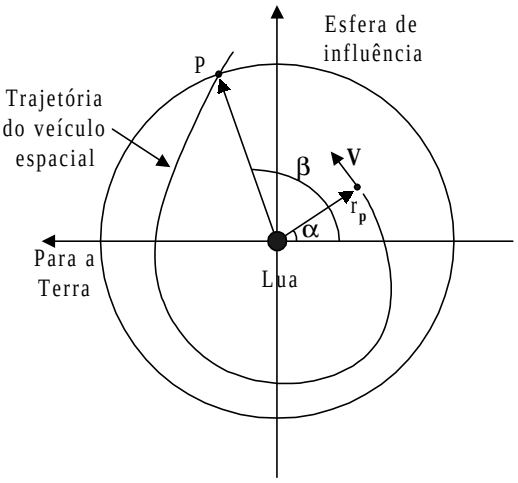


Figura 2. Descrição dos parâmetros da captura.

4. Estudo de Capturas com C3 Mínimo

Os resultados obtidos por Yamakawa (1992) levam a uma fronteira de mínima energia C3. Essa fronteira representa o menor valor possível de C3 para o qual ainda existem trajetórias com captura gravitacional. A fronteira para o sistema Terra-Lua é apresentada na Fig. (3) para as capturas diretas e retrógradas. Nestas Figuras a variável radial representa o valor absoluto de C3 em unidades canônicas. A variável angular é o ângulo α, que representa o ângulo da posição da periluna. Para alcançar estas energias mínimas foram feitas integrações das trajetórias com o ângulo α fixo e variando C3 de 0 até –0.3 unidades canônicas em passos de –0.01. Somente os menores valores de C3 que proporcionam captura gravitacional estão apresentadas nas figuras abaixo. Com a observação dessas figuras pode-se concluir que existem regiões mais favoráveis à captura gravitacional, no sentido em que a energia mínima C3 é menor do que o valor dessa mesma grandeza em outras regiões. É interessante notar que a captura retrógrada, de forma geral, proporciona valores

mínimos de C3 maiores do que os valores obtidos na captura direta, embora as diferenças não sejam grandes. A explicação para esse efeito é que na captura direta existe uma variação de velocidade maior do que no caso da captura retrógrada. Ou seja, as forças perturbativas atuam com um impulso contínuo menor no caso da captura retrógrada do que no caso da captura direta fazendo com que esse tipo de captura não alcance os valores de C3 tão baixo quanto no caso da captura direta.

Existe uma interessante descontinuidade entre 57º e 58º. Para explicar esse fenômeno a Fig. (4) apresenta duas trajetórias, ambas com valor de C3 = -0.16. Uma trajetória tem α = 57.5º e a outra tem α = 57.6º. Nota-se que para a mesma energia, mudando apenas 0.1º em α nas condições iniciais, uma trajetória consegue sair da esfera de influência gravitacional da Lua, enquanto que a outra colide com a Lua. Essa particularidade pode ser facilmente observada na Fig. (4), pois a trajetória do veículo espacial inicialmente se afasta da Lua, depois faz um retorno para uma passagem próxima da Lua. Neste ponto a primeira trajetória continua seu caminho se afastando definitivamente da Lua, enquanto que a segunda trajetória colide com a Lua. Quando α é variado, mantendo-se a mesma energia de dois corpos C3, o vetor velocidade mantém sua intensidade fixa, mas varia sua direção. Isto provoca as descontinuidades apresentadas na Fig. (3).

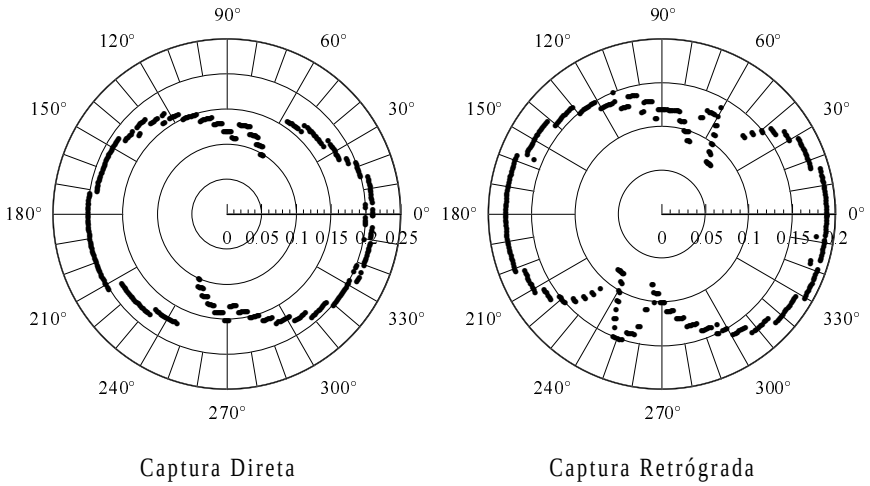


Figura 3 – C3 mínimo do sistema Terra-Lua para capturas diretas e retrógradas.

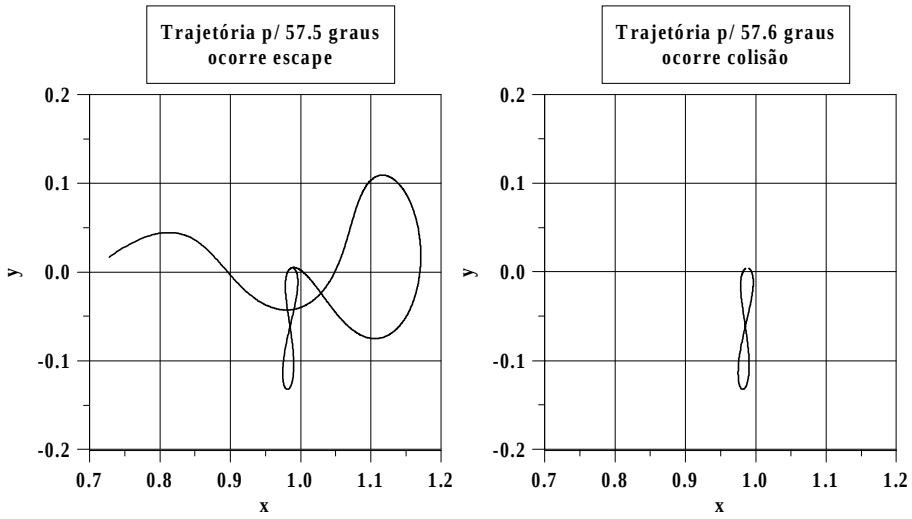


Figura 4 – Trajetórias com C3 = –0.16.

5. Estudo da Sensibilidade dessas Trajetórias

Para saber o quanto trajetórias de captura gravitacional são sensíveis às condições iniciais foram escolhidas algumas trajetórias que são capturadas pela Lua em seu lado oposto a Terra. A essas trajetórias foram adicionadas pequenas perturbações na direção e na magnitude da velocidade, mantendo os valores da posição (x, y) sem modificações. Em seguida foi verificado o efeito que essa perturbação inicial provoca na trajetória. Para medir esse efeito foi tomado o valor x, no momento em que a trajetória cruza pela primeira vez o eixo x (y = 0). Neste mesmo ponto também a distância do ponto de captura foi medido.

5.1. Erros na Magnitude da Velocidade

Os erros na magnitude da velocidade foram produzidos a partir do vetor velocidade inicial (obtidos conforme explicado na Seção anterior). Com o valor da magnitude desse vetor uma determinada porcentagem que foi adicionada ao vetor velocidade inicial foi especificado. Agora um novo vetor de estado, com a mesma posição do vetor anterior mas com a velocidade acrescida de um erro, é então propagado com um passo de tempo positivo até o tempo zero, que é o momento da passagem da trajetória pela periluna desejada. Neste trabalho são estudados dois casos com duas trajetórias diferentes.

Caso 1: O comportamento da trajetória depois de adicionado erro na velocidade inicial está mostrado na Fig. (5). As condições iniciais dessa trajetória levam a uma periluna com C3 = –0.1 e $\alpha = 150^o$.

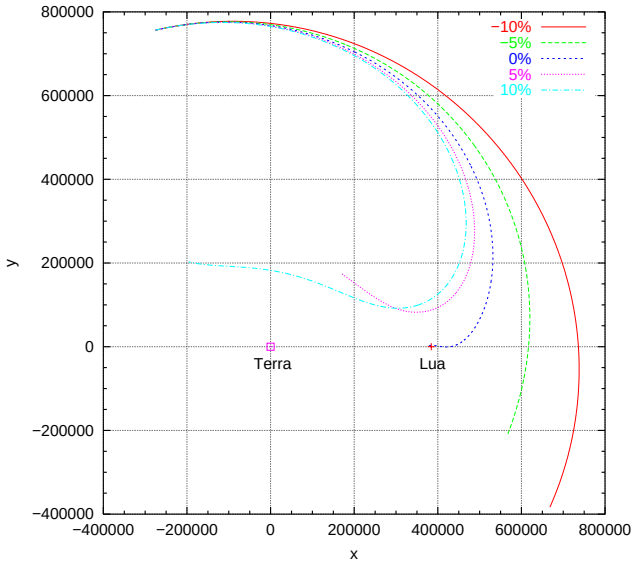


Figura 5. Trajetórias com erros na magnitude da velocidade para condições de periluna C3 = -0.1 e $\alpha = 150^o$.

Nota-se claramente que em todos os casos as trajetórias estão alteradas em relação a nominal (0%). Erros da ordem de ±5% são suficientes para fazer com que o veículo espacial não se aproxime da Lua.

Com as mesmas condições de periluna da Figura anterior, foram gerados os gráficos da Fig. (6) que apresentam o primeiro valor positivo de x quando y = 0 (ou seja, o primeiro ponto de cruzamento do eixo horizontal) e sua distância em relação à posição de captura inicial. A distância em relação à posição de captura é definida como:

$$d = \sqrt{\left(x - x_p\right)^2 + \left(y - y_p\right)^2} \tag{3}$$

onde (x, y) é a posição do veículo espacial e (x_p, y_p) é a coordenada da periluna da trajetória nominal (0%). Cada ponto da Figura representa uma trajetória. Pequenos valores foram somados a velocidade inicial. Esses valores foram de -10% a +10%, ou seja, as primeiras trajetórias têm valores inferiores à velocidade inicial da trajetória nominal e as últimas trajetórias têm valores superiores à velocidade inicial da trajetória nominal.

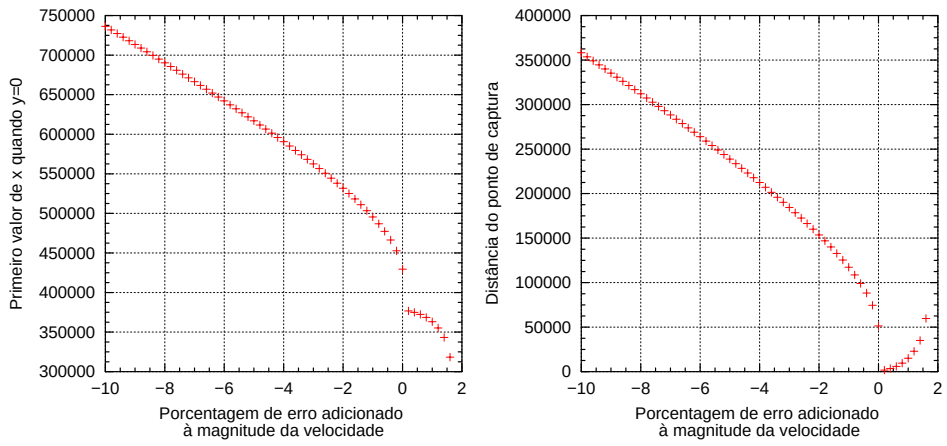


Figura 6. Resultados para erros de até 10% na magnitude da velocidade.

A Figura (6) mostra o comportamento das trajetórias desde o momento em que elas são 10% menores do que a velocidade da trajetória nominal e que termina em captura gravitacional até o momento em que elas superam em 10% a velocidade da trajetória nominal. Como mostrado na Fig. (5), a maioria das trajetórias que não cortam o eixo positivo de x, têm velocidade maior do que a velocidade inicial, e por isso os gráficos da Fig. (6) terminam antes de 2%. A trajetória nominal (0%) corta o eixo x pela primeira vez em ≈429454 km, a trajetória com −10% da velocidade inicial em ≈736340 km, e a trajetória com 1.6% da velocidade inicial (a última com x > 0) em ≈318377 km. Ou seja, existe uma diferença de aproximadamente 306886 km entre a trajetória com erro negativo e trajetória inicial e de 111077 km entre a trajetória com erro positivo e a trajetória inicial (lembrando que a Lua está a 384400 km de distância da Terra, então a trajetória perturbada com um acréscimo na velocidade passa pelo eixo x na região entre a Terra e a Lua). Para um melhor entendimento do fenômeno, optou-se por investigar a região de erro em torno de 1%. A Figura (7) apresenta esses resultados.

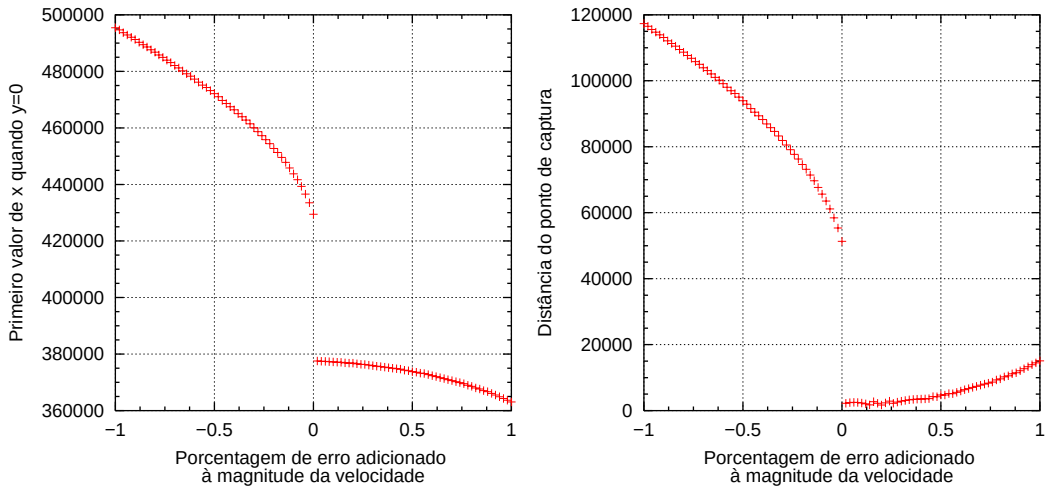


Figura 7. Resultados para erros de até 1% na magnitude da velocidade.

Na Figura (7) os primeiros valores de x diminuem para erros negativos e positivos de forma diferente: de forma mais rápida para os erros negativos e de forma mais lenta para os erros positivos. Ou seja, enquanto a velocidade é aumentada a trajetória do veículo espacial vai se aproximando da Lua, ocorre uma descontinuidade devido à presença da Lua, em seguida as trajetórias se afastam da Lua pelo lado oposto. Analisando a distância do ponto de captura, ela diminui quando os erros são negativos e aumenta quando os erros são positivos. E também ocorre uma descontinuidade em 0% (a trajetória nominal). A Figura (5) ajuda a explicar a descontinuidade. As trajetórias com erros negativos na velocidade cruzam o eixo x depois da Lua se aproximando dela. Por isso o valor de x diminui, assim como a distância do ponto de captura. Depois, as trajetórias com erros positivos na velocidade cruzam o eixo x antes da Lua e se afastam em direção à origem do sistema. Ou seja, o valor de x continua diminuindo, enquanto que a distância do ponto de captura aumenta.

Caso 2: Existe uma fronteira que separa as trajetórias que são capturadas pelo lado oposta da Lua (que vem de fora do sistema Terra-Lua) das trajetórias que são capturadas pelo lado da Terra (que entram na esfera de influência gravitacional da Lua vindo das proximidades da Terra) (Vieira Neto, 1999). Esse caso explora uma trajetória nessa fronteira. A Figura (8) tem como condição final para a periluna $\alpha = 172^{o}$ e C3 = –0.19. As trajetórias de captura gravitacional são mais sensíveis para este nível de energia. É possível ver que a trajetória nominal (0%) faz várias revoluções em torno da Lua antes de alcançar a periluna desejada (veja a ampliação na Fig. (9)). Um fato diferente neste caso, em relação aos estudados anteriormente, é que as trajetórias com erros se afastam mais da trajetória inicial do que nos casos anteriores. As trajetórias com velocidades iniciais menores se afastam do sistema Terra-Lua (a espiral crescente ocorre porque o sistema de referências está girando). Enquanto que as trajetórias com velocidades iniciais maiores ficam presas dentro do sistema.

A Figura (10) apresenta o comportamento da sensibilidade das trajetórias em relação à adição de erros na magnitude da velocidade na escala de –1% a 1%. O comportamento é semelhante ao da Fig. (6) do primeiro caso. O fato das trajetórias com menor velocidade inicial estarem se afastando já foi explicado pela Fig. (8). As trajetórias com velocidades iniciais maiores passam mais próximo do ponto de captura e também revolucionam em torno da Lua como a trajetória inicial. O interessante é que quanto maior a velocidade inicial, menos revoluções em torno da Lua, e essas trajetórias depois escapam na direção da Terra.

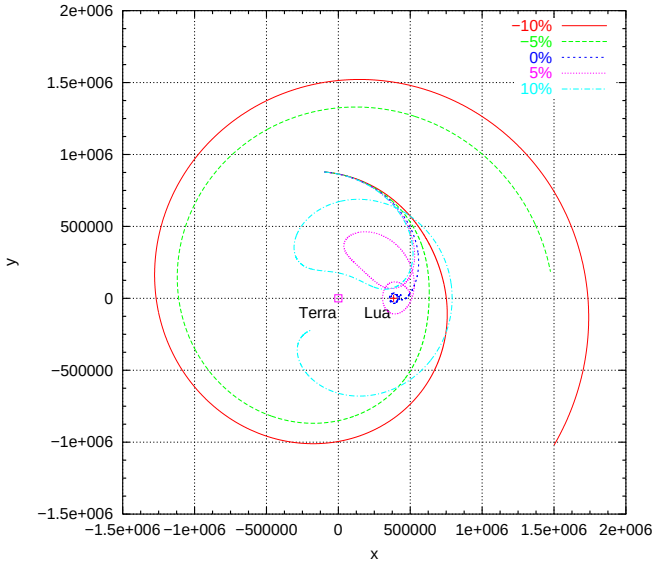


Figura 8. Trajetórias com erros na magnitude da velocidade para condições de periluna C3 = -0.19 e $\alpha = 172^{o}$.

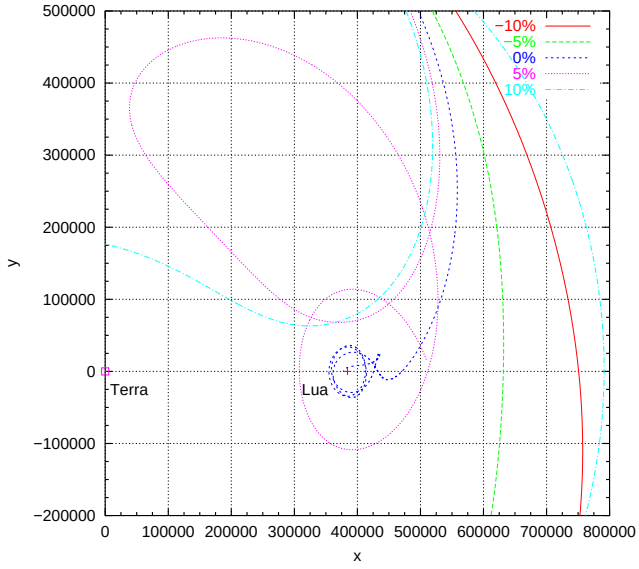


Figura 9. Ampliação da Fig. (8)

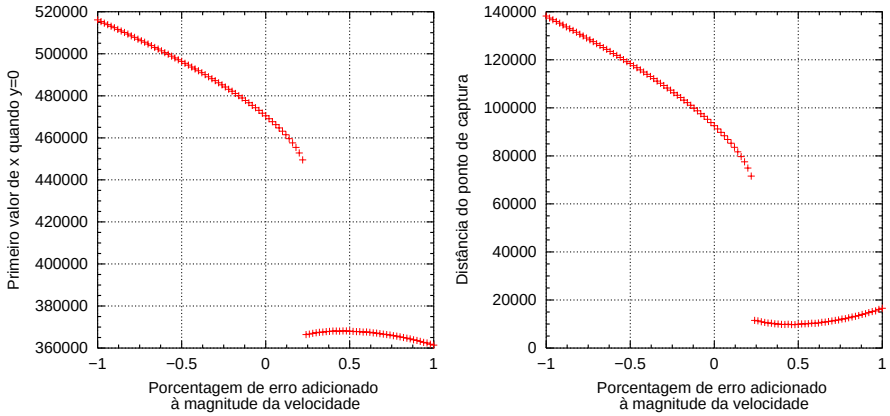


Figura 10. Resultados para erros de até 1% na magnitude da velocidade.

Do estudo destes dois casos é possível concluir que as trajetórias de captura gravitacional apresentam uma grande sensibilidade aos erros. Principalmente a erros em velocidade maiores que 10%, como mostra o primeiro caso. Além

disto, estes três casos apresentados demonstram que é difícil obter resultados gerais nesse tipo de estudo. Os dois casos demonstram comportamentos bem diferentes a respeito da posição em que cortam o eixo x.

5.2. Erros na Direção da Velocidade

Dando continuidade a esse estudo de sensibilidade, agora serão adicionados erros à direção inicial da velocidade. O objetivo é compreender o comportamento das trajetórias quando elas são perturbadas quanto à direção da velocidade. Aqui também são estudados os mesmos dois casos.

Caso 1: A Figura (11) apresenta as trajetórias perturbadas com erros na direção da velocidade. As condições para a periluna é $\alpha = 150^o$ e C3 = -0.1.

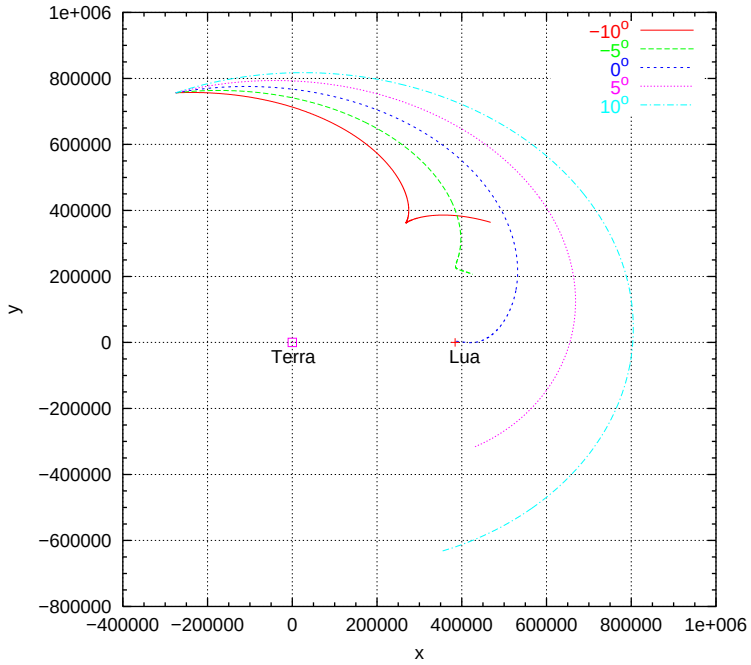


Figura 11. Trajetórias com erros na direção da velocidade para condições de periluna C3 = -0.1 e $\alpha = 150^o$.

Como se pode ver, existe uma diferença significativa entre as trajetórias com erro na magnitude e as trajetórias com erro na direção. Estas últimas apresentam uma separação da trajetória inicial maior do que as do caso anterior. A soma de um ângulo positivo na direção da velocidade faz com que a trajetória se desvie para a esquerda, e com a soma de um ângulo negativo, a trajetória se desvia para a direita. As figuras subseqüentes mostrarão outras diferenças.

A Figura (12) tem correspondência com a Fig. (6), apesar de serem invertidas uma em relação à outra. A inversão ocorre devido à definição da soma do ângulo à velocidade inicial, que faz com que um aumento do ângulo a direção da velocidade inicial gire numa direção anti-horária e vice-versa.

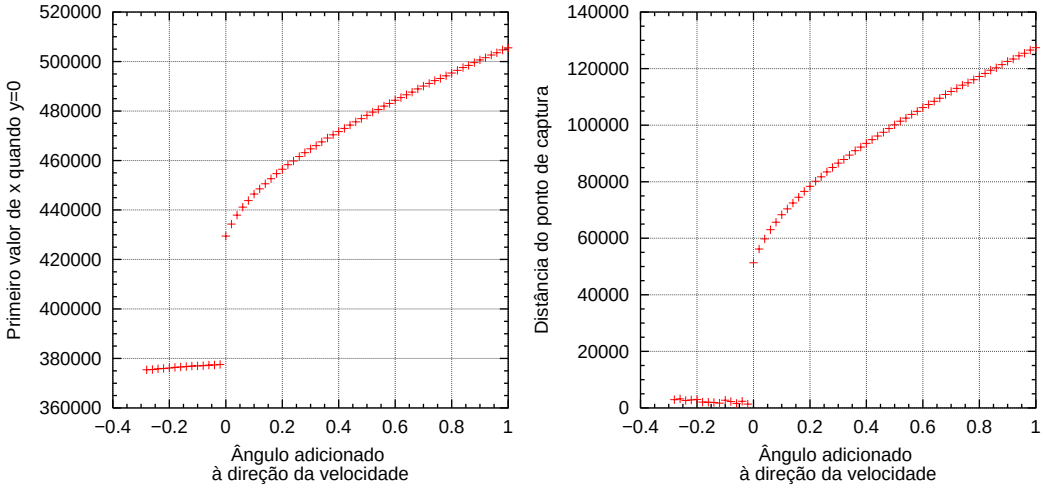


Figura 12. Resultados para erros de até 1^o na direção da velocidade para condições de periluna C3 = -0.19 e $\alpha = 172^o$.

A Figura (12) apresenta o resultado para o intervalo de -1° a 1°. Nessa magnitude, apesar das Figuras serem inversas, o comportamento fica muito similar ao do caso dos erros somados à magnitude.

Caso 2: As trajetórias para o condição de periluna com $\alpha = 172^o$ e C3 = -0.19 estão apresentados na Fig. (13).

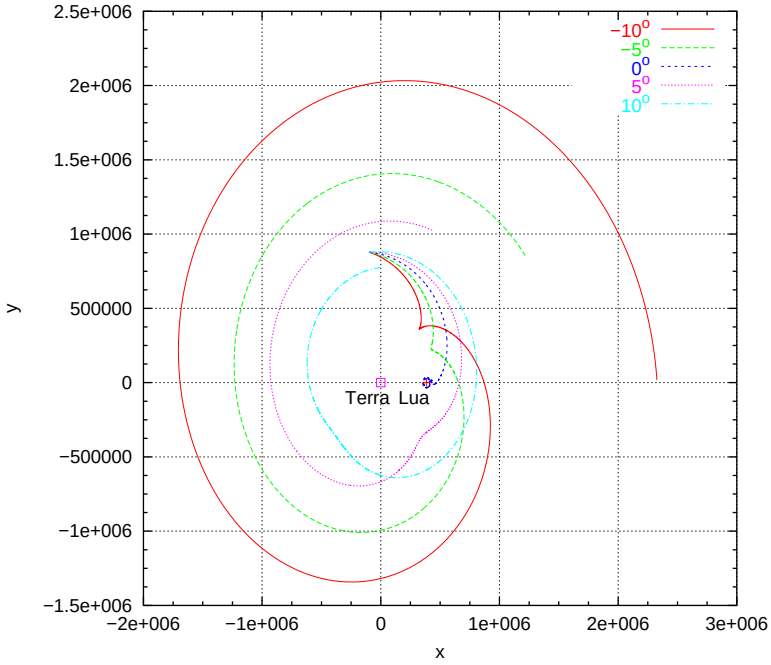


Figura 13. Trajetórias com erros na direção da velocidade.

A Figura (14) apresenta o comportamento destas trajetórias. O caso apresentado na Fig. (14) mostra uma descontinuidade diferente das anteriores, onde o ponto de cruzamento do eixo x dá um salto de quase 60000 km e se estabiliza dentro de uma falha. Isto ocorre porque as trajetórias que sofrem a quantidade de desvio mostrada na escala, ao passarem próximo da Lua, elas ficam presas dando várias revoluções. Isto acontecia também no caso em que o erro era adicionado na magnitude da velocidade, mas não com a mesma intensidade que neste caso.

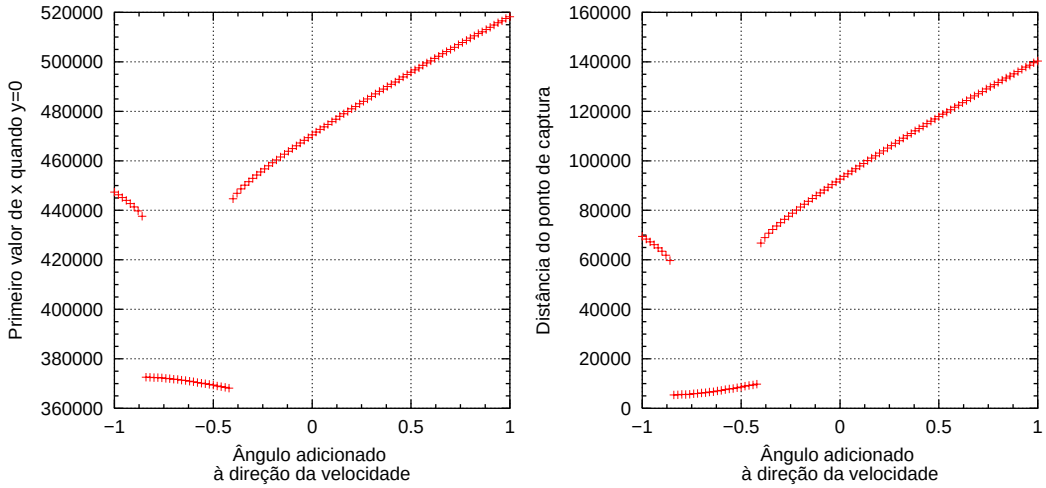


Figura 14. Resultados para erros de até 1° na direção da velocidade.

6. Conclusões

Um resultado interessante mostrado nesse estudo é o fato de terem sido encontradas regiões em que a captura gravitacional ocorre com um valor de C3 mais negativo do que em outras. Alem disso, o presente trabalho pode explicar as descontinuidades existentes na curva de C3 mínimo, que não haviam sido explicadas anteriormente. Na seqüência foram efetuados testes de sensibilidade que demonstram que as trajetórias de captura gravitacional são sensíveis quando expostas a pequenos erros nas condições iniciais. Esta sensibilidade dificulta uma tentativa de otimização global da transferência Terra-Lua.

Erros na magnitude da velocidade são mais problemáticos do que os erros em direção. Mas nos dois casos ocorrem faixas de erro em que é possível fazer a correção com um baixo custo de combustível. Além disso, apesar de estarem a uma certa distância da periluna desejada, a energia de dois corpos C3 apresenta um valor menor em relação a manobra que não considera a captura gravitacional, o que possibilita fazer uma manobra com um consumo menor de combustível.

7. Referências

Belbruno, E.A. Examples of the nonlinear dynamics o ballistic capture and escape in the Earth-Moon system. AIAA **Astrodynamics Conference**, Portland, 1990. (AIAA-90-2896)

Belbruno, E.A. Lunar capture orbits, a method of constructing Earth-Moon trajectories and the lunar gas mission. AIAA/DGLR/JSASS **International Eletric Propulsion Conference**, 19., Colorado Springs, 1987. (AIAA-87-1054) 10 p.

Belbruno, E.A.; Miller, J.K. **A ballistic lunar capture trajectory for the japonese spacecraft Hiten**. Pasadena, JPL, 1990a, (JPL IOM-90.4-1731).

Belbruno, E.A.; Miller, J.K. **A ballistic lunar capture trajectory for the Lunar Observer mission**. 1990b, (JPL IOM-90.4-1752).

Hohmann, W. **Die erreichbarkeit der himmelskorper**. Munique: s.e., 1925. 1 v.

Roy, A.E. **Orbital motion**. Bristol: Adam Hilger, 1988.

Tanikawa, K. Impossibility of the capture of retrograde satellites in the restricted three-body problem. **Celestial Mechanics**, v. 29, n.4, p. 367-402, 1983.

Vieira Neto, E. Estudo numérico da captura gravitacional temporária utilizando o problema restrito de três corpos. Tese de doutorado, INPE, 1999. (INPE-7033-TDI/663)

Yamakawa, H. **On Earth-Moon transfer trajectory with gravitational capture**. Dr. Dissertation, (University of Tokyo). December, 1992.

Yamakawa; H.; Kawaguchi; J.; Ishii; N.; Matsuo, H. On Earth-Moon transfer trajectory with gravitational capture. In: AAS/AIAA Astrodynamics Specialist Conference, Vitoria, August 1993. **Proceedings**. Hampton, NASA, 1993.

An Error Analysis of the Gravitational Capture Trajectories

Ernesto Vieira Neto
FEG-UNESP - CP 205 Guaratinguetá CEP 12.500-000 SP – BRAZIL
ernesto@feg.unesp.br

Antônio Fernando Bertachini de Almeida Prado
DEM-INPE – São José dos Campos CEP SP – BRAZIL
prado@dem.inpe.br

Abstract. *This paper studies the problem of gravitational capture in the regularized restricted three-body problem. For gravitational capture it is understood a phenomenon where a massless particle changes its two-body energy around one of the primaries from positive to negative. This capture is always temporary and, after some time, the two-body energy changes back to positive and the massless spacecraft leaves the neighborhood of the primary. The importance of this temporary capture is that it can be used to decrease the fuel expenditure for a mission going from one of the primaries to the other, like an Earth-Moon mission. The goal is to apply an impulse to the spacecraft during this temporary capture to accomplish a permanent capture. Since the goal of this impulse is to decrease the two-body energy of the spacecraft, its magnitude will be smaller if applied during this temporary capture. This paper presents results of the search of minimum energy trajectories. It is particularly verified the behavior of those trajectories if a small deviation of the velocity before the capture occurs. To make this study, a large number of trajectories starting close to the secondary body are numerically integrated backward in time. The initial position and velocity of the spacecraft are changed and it is verified if an escape occurs for every trajectory.*

Keywords. *Astrodynamics, Gravitational Capture, Problem of Three Bodies, Orbital Maneuvers.*

XVI CONGRESSO BRASILEIRO DE ENGENHARIA MECÂNICA
16th BRAZILIAN CONGRESS OF MECHANICAL ENGINEERING

On Non-Ideal Vibrations of Nonlinear Portal Frame: Analysis of the Passage Through Resonance

Jorge L. Palacios
Faculdade de Engenharia Mecânica da Universidade Estadual de Campinas, C.P. 6122, CEP 13083-970, Campinas, SP, Brazil
e-mail: jfelix@fem.unicamp.br

José M. Balthazar
Instituto de Geocências e Ciências Exatas da Universidade Estadual Paulista, C.P. 178, CEP 13500-230, Rio Claro, SP, Brazil
e-mail: jmbaltha@rc.unesp.br

Reyolando M.L.R.F. Brasil
Escola Politécnica da Universidade de São Paulo, C P 61546, CEP 9524-970, SP, Brazil
e-mail: rmlrdfbr@usp.br

Abstract: *The vibrations of a simple portal frame structure induced by a non-ideal energy source are studied computationally. The interaction of the structure with the excitation source may lead to the occurrence of interesting phenomena during the forward passage through the several resonance states of the systems. The equations of motion of a simplified three-degree-of-freedom model are obtained via a Lagrangian approach having quadratic non-linearities and modulation equations that govern the amplitudes and phases of the interacting modes obtained from the method of averaging are used to investigate the non-ideal vibrations. Special attention is focused in passage through resonance when the excitation frequency is near the second modal frequency and under internal resonance conditions presented of the form 1:2. The solutions are obtained via Runge-Kutta-Fehlberg method for a set of control parameter. The results of approach are plotted in time evolution curve. The saturation and jump phenomenon are observed at certain levels of excitation.This saturation appear in the energy transference from a higher frequency mode to a lower frequency mode due to internal resonance and jump appear on the frequency responce curve as a discontinuity which indicates a region where steady-state conditions do not exist.*

Keywords: *Internal Resonance, Nonlinear Dynamics, Non-Ideal Power Supply, Averaging Method.*

Nomenclature

h = Height of column
EI = Beam and column bending rigity
k = Stiffness
g = Gravity
m = Mass
m_0 = unbalanced mass
M = Mass of motor
L = Length of beam
r = Eccentricity of the DC motor
I_m = Moment of inertia of rotor
u = Horizontal displacement
v = Vertical displacement
q_j = Modal generalized displacement
$\dot{q}_3$ =Rotor speed
α_j, μ_j =Constants
a_i =Modal amplitudes
ξ_j =Modal phases
σ_j =Detuning parameters
A_j, β_j, Δ =Parameters
ε = Small parameter of the problem
ω_j = Natural frequencys of modes
c = Modal linear viscous damping
$\hat{a}$ = Related to voltage of DC Motor
$\hat{b}$ = Constant related to of DC motor
Ω = Frequency of non-ideal energy source
$\Theta(\dot{q})$ = Characteristic driving torque of the motor
$\Psi(\dot{q})$ =Resisting torque of the motor

1. Introduction

In the present work, the case of simple portal frame structure induced by a non-ideal energy source is studied. Recently, significant contributions have been made to the theory of vibrations of non-linear systems. Nevertheless, special kinds of vibrations, arising from the interaction of the system with the energy source, can not yet be completely explained by means of current theory.

It is convenient to introduce two new concepts: an ideal energy source and a non-ideal energy source, as follows.

An ideal energy source is one that acts on the vibrating system, but does not experience any reciprocal influence from the system. A non-ideal source is one that acts on a vibrating system and at the same time experiences a reciprocal action from the system. Changes in the parameters of the system may be accomplished by changing the working conditions of the energy source.

These interactions may become especially active when the energy source has very limited power and they will be more visible in the resonance regions. That is, we assume that the difference between the natural frequency ω of the system and the frequency of the exciting force (for example, a DC motor) Ω is small, i.e., a detuning parameter Δ=ω-Ω is small.

In an ideal system, a motor mounted on a structure, we assume that the motor requires a certain input (Power) to produce a certain output (RPM) regardless of the motion of the structure. If we consider the same system as non-ideal, this may be not the case. Hence, we are interested in what happens to the motor, input, and output, as the response of the system changes.

Vibrating problems with a limited power supply have been investigated by a number of researchers. (Kononenko, 1969) devoted an entire text to this subject and (Balthazar et al., 1999) and (Balthazar et al., 2001) presents a complete review of these kinds of problems up to 2001.

The purpose of this paper is to deal with this kind of problem and to report some results for a simple non-ideal Portal Frame Structure of geometrically non linear behavior excited by a Limited Power Supply (non-ideal problem).

As early as 1970 (Barr and Macwanell, 1971) studied a simple portal frame under support motion, but nonlinear elastic forces were not taken into account. A study of nonlinear oscillations of portal frames under a single ideal harmonic excitation can be found in (Mazzilli and Brasil, 1995).

Recently a study of nonlinear oscillations of portal frames under several ideal loads can be found in (Brasil, 1999). The non-ideal case appears in (Balthazar and Brasil, 2000), (Brasil, Palacios and Balthazar, 2000) and (Brasil, Palacios and Balthazar, 2001).

The present paper reports results of numerical simulations based in a mathematical model of non-ideal simple portal frame vibrating system reduced to a three-degrees of freedom dynamical system deduced by Lagrangian approach and results of numerical simulations of the modulation equations deduced by a method of averaging to investigate the behavior of the system near the resonance region (the frequency of the second mode is in resonance with the frequency mean of the energy source) and the physical and geometric properties of the frame are chosen to tune the natural frequencies of these two modes into a 1:2 internal resonance. In this case of modal interactions, if the high-frequency mode is directly excited, the response of the low-frequency mode could be very large due to an energy transfer to this mode from the high-frequency mode (Nayfeh, 1979). The torque resisting the rotation of the rotor has been taken as the characteristic of the DC motor (energy source).

2. Equations of Motion

The structure will be modeled as a three-degree-of-freedom system. The adopted model is shown in Figure 1 (portal Frame and motor).

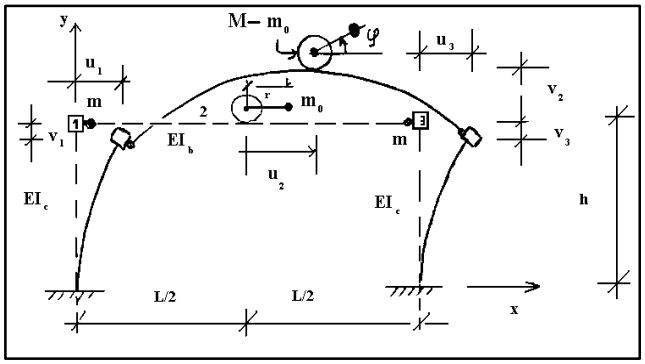


Figure 1. Model of the portal frame with a non-ideal motor.

The generalized coordinate q_1 is related to the horizontal displacement of the central section of the beam in the sway mode (with natural frequency w_1) and q_2 to the mid-span vertical displacement of the same section of beam in the first symmetrical mode (with natural frequency w_2). An unbalanced non-ideal motor is placed at the mid-span of the beam. The angular displacement of its rotor is adopted as generalized coordinate q_3.

The original equations of motion (Brasil and Balthazar, 2000) are:

$$\ddot{q}_1 + w_1^2 q_1 = e[-a_5 q_1 q_2 + a_1(\ddot{q}_3 sin(q_3) + \dot{q}_3^2 \cos(q_3)) - m_1 \dot{q}_1]$$
$$\ddot{q}_2 + w_2^2 q_2 = e[-a_6 q_1^2 + a_2(-\ddot{q}_3 \cos(q_3) + \dot{q}_3^2 sin(q_3)) - m_2 \dot{q}_2 - a_8]$$
$$\ddot{q}_3 = e[a_3 \ddot{q}_1 sin(q_3) - a_4 \ddot{q}_2 \cos(q_3) - a_7 \cos(q_3) + F(\dot{q}_3)] \quad (1)$$

Here, the constants a_j, $j = 1,\ldots,8$ are defined in the earlier works in (Balthazar and Brasil, 2000), (Brasil, Palacios and Balthazar, 2000) and (Brasil, Palacios and Balthazar, 2001), the parameter e is a small dimensionless parameter.

Here, we consider real, limited power supply, motors. For simplicity, their characteristic curves of the energy source (DC motor) are assumed to be straight lines of form $\Theta(\dot{q}_3)-\Psi(\dot{q}_3)=\hat{a}-\hat{b}\dot{q}_3$. Note that the parameter $\hat{a}$ is related to the voltage and $\hat{b}$ is a constant for each model of motor considered. The voltage will be the control parameter of the problem.

3. The Averaged Equations

The motions described by Eq. (1) near resonance region will be studied by the method of averaging (Kononenko, 1969) and (Nayfeh and Mook, 1979). To apply the method of averaging, we first use the method of variation of parameters.

When ε=0, the solutions of Eq. (1) can be expressed as

$$q_1 = A_1\cos(\frac{1}{2}q_3-\beta_1),\ q_2 = A_2\cos(q_3-\beta_2),\ q_3=\Delta t \tag{2}$$

Subject to the constraints

$$\dot{q}_1 = -\frac{1}{2}A_1\omega_2 sin(\frac{1}{2}q_3-\beta_1),\ \dot{q}_2 = -A_2\omega_2 sin(q_3-\beta_3),\ \dot{q}_3=\Delta \tag{3}$$

Where $A_1, A_2, \beta_1, \beta_2$ and Δ are constants, which are sometimes referred to as parameters.

When $\varepsilon\neq 0$, we assume that the solution of Eq. (1) is still given by Eq. (2) but with time varying $A_1, A_2, \beta_1, \beta_2$ and Δ ; that is, $A_1=A_1(t), A_2=A_2(t), \beta_1=\beta_1(t), \beta_2=\beta_2(t)$ and $\Delta=\Delta(t)$. This is why this approach is called the method of variation of parameters. We restrict our attention to a narrow band of frequencies around the natural frequency ω_2 introducing the detuning parameter σ_2 and detuning parameter σ_1 in the presence of internal resonance

$$\Delta=\omega_2+\varepsilon\sigma_2,\ \omega_2=2\omega_1+\varepsilon\sigma_1 \tag{4}$$

and we average the transformed equations over the interval $[0\ \ 2\pi]$ (Palacios, Balthazar and Brasil, 2001) for obtain the following set of five autonomous first order equations:

$$\begin{Bmatrix} \frac{da_1}{dt} \\ \frac{d\chi_1}{dt} \\ \frac{da_2}{dt} \\ \frac{d\chi_2}{dt} \\ \frac{d\Omega}{dt} \end{Bmatrix} = \begin{Bmatrix} -\frac{\varepsilon}{\omega_2}[-\frac{\alpha_5}{2}a_1a_2 sin(\chi_2-2\chi_1)+\frac{\mu_1}{2}a_1\omega_2] \\ \varepsilon\frac{1}{2}[\sigma_2+\sigma_1]-\frac{\varepsilon}{2a_1\omega_2}\alpha_5 a_1a_2\cos(\chi_2-2\chi_1) \\ -\frac{\varepsilon}{\omega_2}[\frac{\alpha_6}{4}a_1^2 sin(\chi_2-2\chi_1)+\frac{\alpha_2}{2}\Omega^2\cos(\chi_2)+\frac{1}{2}\mu_2 a_2\omega_2] \\ \varepsilon\sigma_2+\frac{\varepsilon}{a_2\omega_2}[-\frac{\alpha_6}{4}a_1^2\cos(\chi_2-2\chi_1)+\frac{1}{2}\alpha_2\Omega^2 sin(\chi_2)] \\ \varepsilon[\frac{1}{2}\alpha_4 a_2\omega_2\Omega\cos(\chi_2)+F(\Omega)] \end{Bmatrix} \tag{5}$$

The above autonomous differential equations determine both the amplitude and phases of the $O(\varepsilon)$ approximations of the modal response $q_1(t)$, $q_2(t)$ and $q_3(t)$ governed by Eq. (1). These equations, (5), will henceforth be referred to as the averaged equations.

The first approximation is,

$$q_1 = a_1\cos(\frac{1}{2}\Omega t-\chi_1),\ q_2 = a_2\cos(\Omega t-\chi_2),\ \ q_3=\Omega t \tag{6}$$

where the a_n, χ_n and Ω are given by the Eq. (5). Constant solutions of Eq. (5) represent steady-state periodic/quasi-periodic motions of system Portal Frame/Non-ideal motor. Dynamic solutions of Eq. (5) imply amplitude and phase modulated motions of the system.

4. Results and Observations

To verify the theoretical results of steady state responses, the Eq. (1) and Eq. (5) are numerically integrated using the 4th and 5th order Runge-Kutta-Fehlberg algorithm with variable time steps (Burden and Faires, 1993). The control parameters used are: $\hat{a}$=0.35, $\hat{a}$=0.38, $\hat{a}$=0.41 and $\hat{a}$=0.43 in the equation of the torque characteristic $\Theta(\dot{q}_3) - \Psi(\dot{q}_3) = \hat{a} - \hat{b}\dot{q}_3$.

4.1 Dynamic Solutions of Averaged Equations

In Figure 2, we show time history of the amplitudes a_1 (associated with the horizontal sway of the frame), a_2 (associated to the vertical motion of beam where the non-ideal motor is mounted) and the speed of non-ideal motor Ω that are determined from Eq. (5), observed in the passage through resonance with a slowly increasing value of control parameter $\hat{a} \in$ [0.35 0.50].

In Figure 2(a) the speed of non-ideal motor is below the natural frequency $\mathtt{W}_2$ and we observe that the amplitude a_2 increases in the relation to the amplitude a_1 for the control parameter $\hat{a}$=0.35.

Figures 2(b), 2(c), shows that the speed of non-ideal motor is near the natural frequency $\mathtt{W}_2$ and we observe that the amplitude a_2 stops increasing and the amplitude a_1 jumps to considerable values, potentially dangerous.

This above property is the well-known phenomenon of "saturation" and energy transference from a high frequency low amplitude mode to a low frequency high amplitude mode due to internal resonance for the control parameter $\hat{a}$=0.38 and $\hat{a}$=0.41.

The Figure 2(d), show that the speed of non-ideal motor is above the natural frequency $\mathtt{W}_2$ and we observe that the amplitude a_2 increases in the relation to the amplitude a_1 for the control parameter $\hat{a}$=0.43.

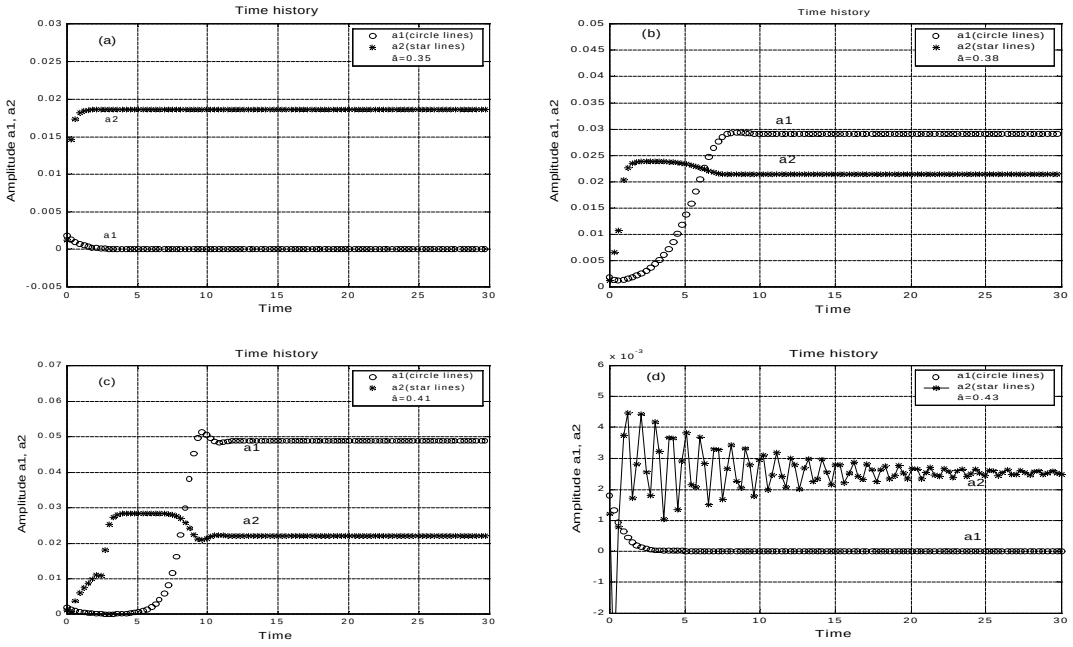


Figure 2. Time history of the amplitudes a_1 (circle lines), a_2 (star lines) with control parameters $\hat{a}$=0.35 for (a), $\hat{a}$=0.38 for (b), $\hat{a}$=0.41 for (c) and $\hat{a}$=0.43 for (d).

4.2 Constant Solutions of Averaged Equations

In this section, we investigate the constant solutions of Eq. (5) and are determined by the following system of equations:

$$\begin{Bmatrix} -\frac{\alpha_5}{2}a_1a_2 sin(\chi_2-2\chi_1)+\frac{\mu_1}{2}a_1\omega_2 \\ \frac{1}{2}[\sigma_2+\sigma_1]a_1-\frac{1}{2\omega_2}\alpha_5 a_1a_2\cos(\chi_2-2\chi_1) \\ \frac{\alpha_6}{4}a_1^2 sin(\chi_2-2\chi_1)+\frac{\alpha_2}{2}\Omega^2\cos(\chi_2)+\frac{1}{2}\mu_2 a_2\omega_2 \\ \sigma_2 a_2+\frac{1}{\omega_2}[-\frac{\alpha_6}{4}a_1^2\cos(\chi_2-2\chi_1)+\frac{1}{2}\alpha_2\Omega^2 sin(\chi_2)] \\ \frac{1}{2}\alpha_4 a_2\omega_2\Omega\cos(\chi_2)+F(\Omega) \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{Bmatrix} \tag{7}$$

Equation (7) that is solved by analytical method or numerically applying the Newton-Raphson method.

There are two types of constant solutions:

I. We analyze constant solutions of the averaged equations (5) when $a_1 = 0$ and $a_2 \neq 0$.

From Eq. (7) we find, an expression for the amplitude

$$a_2 = \frac{\alpha_2\Omega^2}{\omega_2\sqrt{4(\Omega-\omega_2)^2+\mu_2^2}} \tag{8}$$

an expression for the phase of the oscillation

$$\cot\chi_2 = -\frac{\mu_2}{2(\Omega-\omega_2)} \tag{9}$$

and an equation for determining the frequency of excitation of the motor

$$M(\Omega) = S(\Omega) \tag{10}$$

where

$$M(\Omega) = \frac{1}{I_m + m_0 r^2}(\hat{a}-\hat{b}\Omega), \quad S(\Omega) = \frac{\alpha_4\mu_2\omega_2^2}{2\alpha_2\Omega}a_2^2 \tag{11}$$

Thus, to first approximations are

$$q_1 = 0,\ q_2 = a_2\cos(\Omega t-\xi_2),\ q_3 = \Omega t. \tag{12}$$

In Figure 3, the functions S and the characteristic of the motor M are plotted as functions of Ω. These curves are obtained directly from (11).

In Figure 4, shows the steady-state solutions of q_2, obtained by the Eq. (12). For values of the excitation frequency Ω, obtained by the Eq. (10) or by the intersection of the curves M and S indicated on the Figure 3.

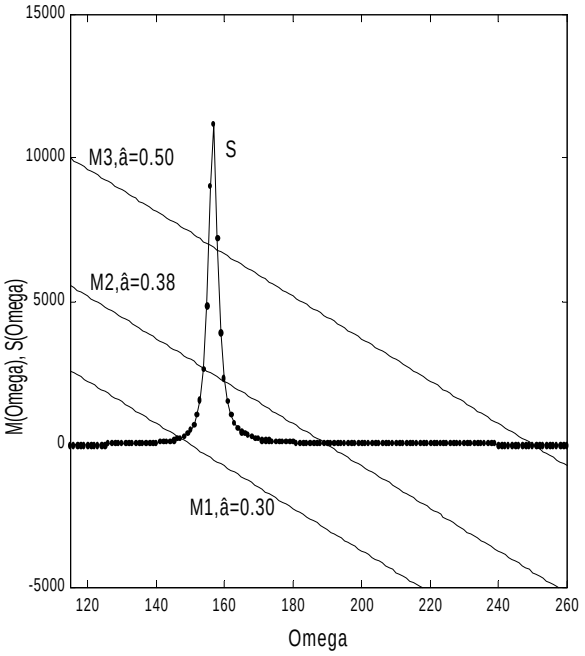


Figure 3. Characteristics of the motor M and S as function of the excitation frequency Ω. See equation (11).

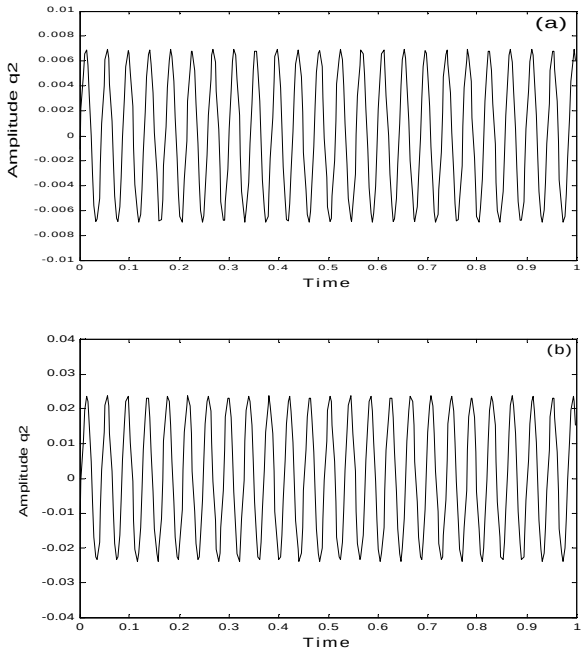


Figure 4. Modal solutions q_2, see Eq. (12), for (a) M_1 and (b) M_2.

II. We analyze constant solutions of the averaged equations (5) when $a_1 \neq 0$ and $a_2 \neq 0$.

Equation (7) that is solved numerically applying the Newton-Raphson method. In Figure 5, shows the constants solutions a_1 and a_2 as function of excitation frequency Ω. There is the jump phenomenon and a saturation phenomenon associated with varying the control parameter $\hat{a}$.

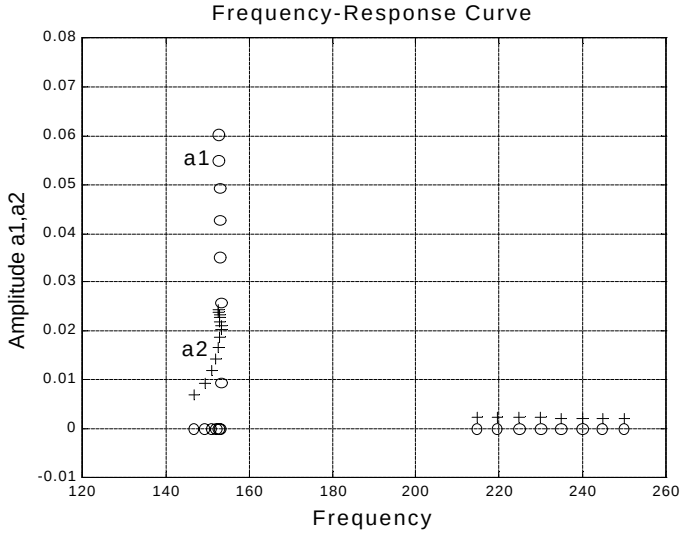
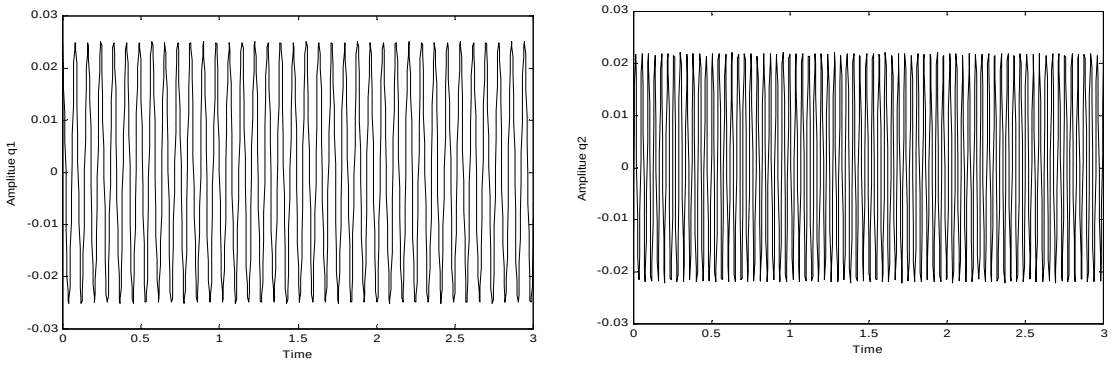


Figure 5. Frequency-response curves for portal frame-motor structure for the case $\Omega \approx w_2$ and $w_2 \approx 2w_1$

Thus, to first approximation for Ω=153.58, a_1=0.025, a_2=0.022, x_1=6.11 and x_2=8.43 corresponding the control parameter $\hat{a}$=0.38, we obtain

$$q_1 = a_1\, cos\left(\frac{1}{2}\Omega t - \xi_1\right),\ q_2 = a_2\, cos(\Omega t - \xi_2),\ q_3 = \Omega t\,. \tag{13}$$

In Figure 6, we show the steady-state solutions of Eq. (13) and a exchange of energy between the two modes q_1 and q_2 corresponding the control parameter $\hat{a}$=0.38.



(a) (b)

Figure 6. Steady-state solutions of q_1 in (a), q_2 in (b) corresponding the constant solutions of averaged equations with the control parameter $\hat{a}$=0.38.

4.3 Dynamic Solutions of Equations of Motion

In this section, we investigate the solutions of Eq. (1), with $\varepsilon = 1$, for the control parameter $\hat{a} \in [0.30, 0.50]$ and we show typical frequency-response curve in Figure 7. The response curve exhibits the saturation and jump phenomenon. The idea of the numerical simulations we are about to present is to obtain, for each power setting of the motor, the steady-state amplitude of the two generalized coordinates:

$$aq_1 = \frac{(max(q_1) - min(q_1))}{2},$$
$$aq_2 = \frac{(max(q_2) - min(q_2))}{2}, \tag{14}$$

and the corresponding resulting speed of the shaft of the motor:

$$\text{Freq} = mean\left(\frac{dq_3}{dt}\right). \tag{15}$$

where

$q_1, q_2, and\ \frac{dq_3}{dt}$ are steady-state solutions of the Eq. (1) and aq_1, aq_2 are the amplitudes de q_1, q_2 respectively and "Freq" describe the mean excitation frequency of the motor that is near to the frequency of the second mode.

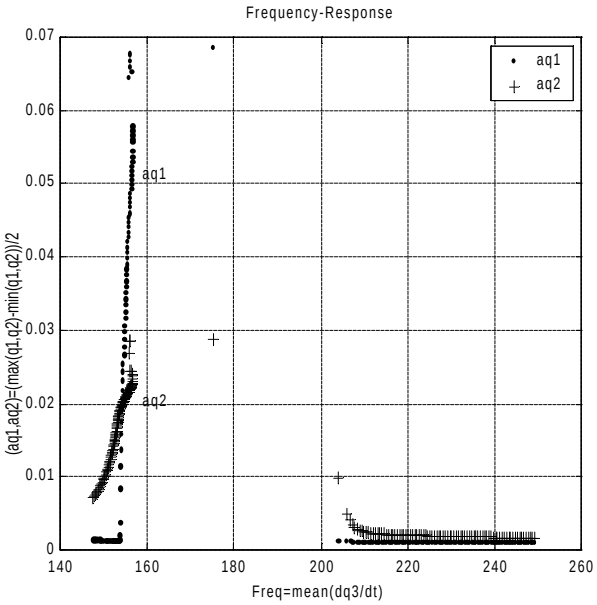


Figure 7. Frequency-Response for the first and second mode when the control parameter $\hat{a} \in [0.30, 0.50]$. Line point corresponding the amplitude of q_1 and line plus corresponding the amplitude of q_2 as function of motor speed $\frac{dq_3}{dt}$.

In Figure 7, we show that for certain power settings (increasing the values of control parameter $\hat{a}$ in equation $\Theta - \Psi = \hat{a} - \hat{b}\dot{q}_3$), the amplitude of the coordinate q_2, associated to the vertical movement of beam where the motor is mounted, stops increasing and the amplitude of the coordinate the q_1, associated with the horizontal sway of the frame, jumps to considerable values, potentially dangerous. This is the well-known phenomenon of saturation and energy transference from a higher frequency mode to a lower frequency mode due to internal resonance.

Concluding remarks

Oscillations of a portal frame of non-linear behavior excited by a non-ideal motor (with limited power supply) were studied via the method of averaging.

We have compared numerical results of the averaged equation (5) and the original governing equation (1).

We consider this averaging method to be an excellent tool to study the characteristics of motion of a non-ideal system.

We Analyzed the possibility of occurrence of Sommerfeld effect (getting stuck in resonance, that is, not enough power to reach higher speed regimes with lower energy consumption) and the possibilities of occurrence of saturation of a high frequency low amplitude mode and transference of energy to a low frequency high amplitude mode.

Acknowledgements

The authors acknowledge support by FAPESP, Fundação de Apoio à pesquisa do Estado de São Paulo. The second and third authors also thank CNPq, Conselho Nacional de Pesquisas. Both are Brazilian Research Funding Agencies.

References

Balthazar, J.M, Mook, D.T., Weber, H.I., A. Fenili, Belato, D., De Mattos M.C. and Wieczorek, S., 1999," On Vibrating Systems with a Limited Power Supply and Their Applications ", Brazilian Seminar of Analysis, State University of Campinas, Short Course, pp.: 137-227.

Balthazar, J.M, Brasil, R.M.L.R.F., Weber, H.I., A. Fenili, Belato, D., Felix, J.L.P. Felix and Garzeri F.J., 2001," Review of New Vibration Issues Due to Non-Ideal Sources ", submitted

Barr, DC. And DC. Mcwannell. Parametric Instability in Structures under Support Motion, Journal of Sound of Vibration V. 14, N4, p. 491-509, 1971.

Brasil, R.M.F.L, and Balthazar, J.M., 2000, " Nonlinear Oscillations of a Portal Frame Structure Excited by a Non-ideal Motor", Proceedings of 2000 2nd International Conference- Control of Oscillations and Chaos, Edited by F.L. Chernousko, A. L. Fradlov, Volume 2 of 3, St. Petersburg, Russia, pp. 275-278.

Brasil, R.M.F.L, Palacios, J.L and Balthazar, J.M., 2000," Some comments on Numerical analysis of Nonlinear Vibrations of a Civil Structure Induced by a Non-Ideal Source", Proceedings of Computational Methods in Engineering' 2000, CD ROM, 21st Iberian Latin- American Congress on Computational Methods in Engineering, December 6- 8, Rio de Janeiro, Brazil.

Brasil, R.M.F.L, Palacios, J.L and Balthazar, J.M., 2001," On the Nonlinear Dynamic Behavior of a Non-Ideal Machine Foundation: Numerical Simulations" Proceedings of Third International conference on Nonlinear Dynamics, Chaos, Control, and Their Applications in Engineering Sciences, Campos do Jordão, Brazil, July 31-august 4, pp. 326-354.

Brasil, R.M.F.L, 1999, "Multiple Scales Analysis of Nonlinear Oscillations of A Portal Frame Foundations for Several Machines. Journal of the Brazilian Society of Mechanical Sciences", Vol. 21,N0.4, pp.641-654.

Burden, , R.L. and Faires, J.D.,1993, "Numerical Analysis", PWS Publishing Company, Boston.

Mazzilli, C.E.N. and Brazil, R.M.F.L., 1995, "Effect of static Loading on the Nonlinear Vibrations of a Three-time Redundant Portal Frame: Analytical and Numerical Studies", Nonlinear Dynamics, Vol. 8, pp. 347-366.

Nayfeh, A. H., 1981, "Introduction to Perturbation Techniques", Wiley-Interscience, New York.

Nayfeh, A. H. and Mook, D. T., 1979, "Nonlinear Oscillations" , Wiley-Interscience, New York.

Kononenko, V., 1969, "Vibrating Systems with Limited Power Supply", Illife.

Palacios, J.L, Balthazar, J.M. and Brasil, R.M.F.L., 2001, "On Non-ideal Dinamics of Nonlinear Portal Frame analysis using Averaging Method ", Proceding of the Nineth International Symposium on Dynamic Problems of Mechanics, IX DINAME, Florianopolis, Brazil, pp. 143-147.

XVI CONGRESSO BRASILEIRO DE ENGENHARIA MECÂNICA
16th BRAZILIAN CONGRESS OF MECHANICAL ENGINEERING

DETERMINAÇÃO DE ÓRBITAS FIXAS E PERIÓDICAS EM SISTEMAS CAÓTICOS

Roth, Bruno Aloís Forlin
Instituto Nacional de Pesquisas Espaciais - INPE,
Laboratório Associado de Computação e Matemática Aplicada - LAC,
Av. dos Astronautas, 1758. S. J. Campos, SP, C.P. 515, 12227-010, Brasil.
roth@pgrad.inpe.br

Macau, Elbert Einstein Neher
Instituto Nacional de Pesquisas Espaciais - INPE,
Laboratório de Integração e Testes - LIT,
Av. dos Astronautas, 1758. S. J. Campos, SP, C.P. 515, 12227-010, Brasil.
elbert@lit.inpe.br

Terra, Maisa de Oliveira
Instituto Nacional de Pesquisas Espaciais - INPE,
Laboratório de Integração e Testes - LIT,
Av. dos Astronautas, 1758. S. J. Campos, SP, C.P. 515, 12227-010, Brasil.
maisa@lit.inpe.br

Resumo. *A determinação de órbitas periódicas de sistemas caóticos é de fundamental importância para a caracterização de sistemas dinâmicos. Órbitas periódicas intáveis revelam o esqueleto do conjunto caótico. Quantidades que caracterizam o movimento caótico, tais como, dimensões, expoentes de Lyapunov e entropia topológica, podem ser calculados apartir do conjunto de órbitas periódicas que esta densamente presente no invariante caótico. Neste artigo, alguns métodos de determinação de pontos fixos e periódicos intáveis em sistemas dinâmicos são explorados. A eficiência e aplicabilidade desses métodos são verificados. Nosso objetivo é desenvolver um ambiente de análise integrado e uma biblioteca numérica para determinação de pontos fixos e órbitas periódicas em sistemas caóticos.*

Palavras chave:. *caos, órbitas periódicas instáveis, sistema dinâmico não-linear.*

1 Introdução

Em sistemas dinâmicos em regime de evolução caótica, o movimento é organizado em torno de órbitas periódicas instáveis. Estas órbitas capturam o esqueleto topológico do movimento, no sentido de que toda órbita pode ser aproximadamente reconstituída através das órbitas periódicas instáveis de curto período. Podemos calcular quantidades que caracterizam o movimento caótico, tais como dimensões, expoentes de Lyapunov, entropia topológica, seguindo as órbitas periódicas instáveis do sistema (Auerbach, 1987; Cvitanovic, 1971). Em (Cvitanovic, 1971) uma expansão em série em torno das órbitas periódicas instáveis de curto período foi introduzida e aplicada para avaliação de grandezas representativas da evolução caótica. As expansões de curvatura (Cvitanovic, 1971) do somatório das órbita periódicas (Gutzwiller, 1971; Balian, 1974; Miller, 1975) são uma ferramenta igualmente poderosa para a avaliação das ressonâncias nos sistemas caóticos Hamiltonianos (Cvitanovic-II, 1988). O presente trabalho tem por objetivo explorar comparativamente os dois métodos de detecção de órditas periódicas que são considerados os mais eficientes quando se conhece a descrição matemática do sistema dinâmico. São estes métodos Schmelcher e Diakonos (SD) (Schmelcher, 1998) e Davidchack e Lai (DL) (Davidchack, 2000). De fato, o segundo método é, de certa forma, uma evolução do primeiro. Seus autores afirmam, inclusive, que este método é poderoso o suficiente para determinar todos as órbitas periódicas existentes no invariante caótico. Entretando, questões fundamentais merecem ser consideradas, em especial relacionadas a eficiência, facilidade de implementação e custo computacional. Em adição, faz-se ainda necessário entender exatamente o mecanismo que intermedia a evolução de um método ao outro.

A determinação das órbitas periódicas instáveis de um invariante caótico se justifica por vários motivos. No estudo de sistemas, freqüentemente há interesse em se levantar grandezas estatísticas de longo prazo, tais como, médias, expoentes de Lyapunov, dimensões e outros invariantes, como a *medida natural* e densidade de *probabilidade*. Estas quantidades estatísticas são fisicamente significativas somente quando a medida que está sendo considerada é gerada por uma trajetória típica no espaço da fase. Esta *medida* é chamada *medida natural* (Bowen, 1975) se é invariante em relação a evolução da dinâmica. Consequentemente, é de importância fundamental compreender e caracterizar a medida natural em termos de quantidades dinamicamente fundamentais. Não há nada mais fundamental do que expressar a medida natural em termos das órbitas periódicas embutidas no conjunto caótico invariante (Lai, 1997). Em (Grebogi, 1988) a medida natural invariante é calculada em termos do valor dos autovalores das órbitas periódicas instáveis embutidas no atrator caótico, sendo expressa e testada somente para o exemplo especial de uma dinâmica hiperbólica.

Existem casos onde a densidade da probabilidade ou a medida natural gerada por trajetórias típicas de sistemas dinâmicos

caóticos é altamente singular. Uma trajetória originada de uma condição inicial aleatória na bacia da atração visita diferentes partes do atrator caótico com probabilidades totalmente diferentes. Chame-se de *quente* as regiões cuja probabilidade de visita das trajetórias é alta, de *frio* as regiões cuja probabilidade de visita das trajetórias é baixa (Grebogi, 1988). Tais regiões *quente* e *frio* podem se entrelaçar em escalas arbitrariamente finas. Neste sentido, os atratores caóticos possuem uma estrutura multifractal, uma propriedade que se reflete na medida natural. Em (Farmer, 1983; Grassberger, 1983; Hentschel, 1983; Grassberger, 1985) há uma descrição do mecanismo que permite quantificar como as órbitas caóticas se distribuem nos atratores. Os espectros de dimensão fractal (Grassberger, 1983; Hentschel, 1983; Grassberger, 1985) são sensíveis às características da estrutura de pontos *quentes* e *frios* no atrator. Para uma classe grande de atratores caóticos, o número infinito das órbitas periódicas instáveis embutidas no atrator fornece a chave para a compreensão de tais estruturas (Grebogi, 1988).

O conjunto de órbitas periódicas instáveis de um sistema dinâmico caótico é também útil para caracterizar a estrutura de magnificação dos atratores estranhos. As órbitas periódicas instáveis são embutidas densamente no atrator (Ekmann, 1985) e seu comprimento fornece uma hierarquia natural das escalas mais finas do atrator. O número de órbitas periódicas cresce exponencialmente com o tamanho do período, permitindo assim uma série de aproximações melhoradas da estrutura total do atrator. Desde que as órbitas periódicas e seus autovalores são invariante sob transformações de coordenadas, a função de magnificação fornece uma codificação invariante dos detalhes microscópicos do sistema. A estrutura multifractal do atrator expressa em termos do espectro (Halsey, 1986) pode ser diretamente extraída da função de magnificação convergida (Auerbach, 1990).

Muitos sistemas caóticos admitem uma adequada dinâmica simbólica (Hao, 1989). Seja o mapa logístico de uma dimensão (May, 1976)

$$x_{n+1} = r x_n (1 - x_n). \tag{1}$$

Para uma trajetória típica, a dinâmica simbólica correspondente pode ser definida associando-se dois símbolos 0 e 1 aos pontos da trajetórias que se situam entre $0 \leq x_n < x_c = \frac{1}{2}$ e $x_c < x_n \leq 1$, respectivamente. Neste exemplo, há uma correspondência de um-a-um entre os trajetórias no espaço de fase e as seqüências do itinerário representado pela seqüência semi-infinita dois símbolos. Para mapas de duas dimensões não existe nenhuma receita para identificar uma partição adequada. Existe uma conjectura de que a partição adequada passa através dos pontos preliminares da tangência entre as variedades estáveis e instáveis (Graassberger-II, 1985; Giovannini, 1991; Chritiansen, 1991), que ocorre estritamente apenas para sistemas específicos tais como o mapa de Hénon (Graassberger-II, 1985; Giovannini, 1991; Henon, 1976). Para outros sistemas, algumas considerações adicionais devem ser empregadas, como o atrator que se dobra no exemplo do atrator de Duffing (Giovannini, 1991), ou nas considerações da simetria no exemplo do mapa padrão (Chritiansen, 1991). É também possível construir uma partição baseando-se em uma análise topológica (Lefranc, 1994). Entretanto, esta aproximação pode somente ser aplicada aos mapas de duas dimensões obtidos da superfície de Poincaré da seção de fluxo de três dimensões. Até o momento não existe nenhuma aproximação eficiente para identificar a partição na dinâmica caótica com dimensão alta. É de importância fundamental para o estudo de dinâmica caótica, encontrar a partição adequada, sobretudo em relação a aplicações tecnológicas importantes tais, como comunicação com caos (Hayes, 1993).

Dado um sistema caótico, sabe-se que a partição pode ser especificada usando o conjunto das infinitas órbitas periódicas instáveis embutidas no conjunto invariante dinâmico subjacente (Baddi, 1994). Cada órbita periódica instável precisa ser representada por uma seqüência simbólica única, se a partição estiver adequada. Assim, é possível atribuir um símbolo a cada ponto de órbita periódica instável de tal maneira que a exigência acima seja satisfeita até algum período grande. O número N_p de pontos da órbita cresce exponencialmente em função do período $p : N_p \sim e^{hTp}$, onde $h_T > 0$ é a entropia topológica do conjunto caótico (Davidchack, 2000).

Através da detecção de órbitas periódicas instáveis é possível testar a presença de determinismo nos dados experimentais, assim como o comportamento de um sistema dinâmico pode ser predito por elas (Aguirre, 1998).

Uma aplicação importante da determinação de órbitas periódicas instáveis está no controle de caos (Ott, 1990), onde este é o passo essencial. Por estas razões muita atenção vem sendo dada a procura de métodos de detecção das órbitas periódicas instáveis nos modelos teóricos e dados experimentais.

Este trabalho se divide como se segue. Na seção seguinte, faremos s decrição do método Schmelcher e Diakonos (SD) (Schmelcher, 1998). Na seção 3, enfocaremos o método Davidchack e Lai (DL) (Davidchack, 2000). Na seção 4 apresentaremos os resultados que fundamentam a comparação entre os métodos. Finalmente, apresentaremos uma conclusão com uma análise de resultados.

2 Método de Schmelcher e Diakonos (Schemelcher, 1998)

A idéia básica deste método consiste em partir do sistema dinâmico discreto N-dimensional em regime caótico e através de transformações lineares apropriadas, obter um sistema dinâmico transformado, tal que todos pontos fixos deste novo sistema sejam os mesmos do sistema original e com as mesmas localizações no espaço de fase. Entretanto as transformações lineares, apropriadamente escolhidas, mudam a estabilidade dos pontos fixos e órbitas periódicas, de tal forma que ao invés de instáveis, como no sistema original, passam a ser estáveis na nova dinâmica, sendo diferentes transformações lineares responsáveis pela estabilidade de diferentes conjuntos de pontos.

Vamos considerar um sistema dinâmico discreto completamente caótico N-dimensional dado por

$$U : \vec{r}_i + 1 = \vec{f}(\vec{r}_i). \tag{2}$$

Por U ser completamente caótico, ele apresenta apenas pontos fixos instáveis. Note que órbitas periódicas de período p correspondem a termos $f^p(\vec{r}_i)$ em lugar de $f(\vec{r}_i)$ na Eq. (2) acima.

O objetivo proposto é o de construir a partir da Eq. (2) um sistema dinâmico S_k diferente contendo os mesmos pontos fixos (PF) de U. Utilizando a transformação L_k abaixo, os PF que eram instáveis passam a estáveis.

$$L_k : U \rightarrow S_k \tag{3}$$

Devido a estabilidade dos pontos fixos no sistema dissipativo S_k, toda trajetória de S_k depois de algumas iterações dirige-se para um ponto fixo $\vec{r}_F$, ou sai dos limites do invariante caótico do sistema original U. Por construção $\vec{r}_F$ também é um ponto fixo do sistema U. Para cumprir a correspondência um a um entre os pontos fixos de U e S_k, a transformação L_k deve em geral ser linear. Portanto:

$$S_k : \vec{r}_{i+1} = \vec{r}_i + \Lambda_k(\vec{f}(\vec{r}_i) - \vec{r}_i), \tag{4}$$

onde Λ_k é uma matriz NxN constante inversível. Reescrevemos $\Lambda_k = \lambda C_k$, tal que λ seja um parâmetro real, $1 \gg \lambda > 0$, e C_k matrizes ortogonais cujos elementos são apenas $\{-1, 0, 1\}$. O número destas matrizes é dado por $2^n n!$ e são listadas para $n = 1$ e $n = 2$ na Tab. (1). A definição na Eq. (4) satisfaz a correspondência um para um entre os pontos fixos de U e S_k: Se $\vec{r}_i = \vec{r}_F$ é um ponto fixo de U, o termo a direita da Eq. (4) desaparece e portanto $\vec{r}_F$ também é ponto fixo de S_k; por sua vez, se $\vec{r}_F$ é ponto fixo de S_k, e Λ_k não é singular, o termo a direita da Eq. (4) deve ser igual a 0 para $\vec{r}_i = \vec{r}_F$, então $\vec{r}_F$ é também ponto fixo de U. Assim, as leis dinâmicas U e S_k possuem pontos fixos em posições idênticas no espaço.

Tabela 1: Matrizes C_k para uma e duas dimensões.

Dimensão	***Matrizes*** C_k
1D	$C_1 = 1$ $C_2 = -1$
2D	$C_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ $C_2 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$ $C_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ $C_4 = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$ $C_5 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ $C_6 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$ $C_7 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ $C_8 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

O método acima de detecção de pontos periódicos instáveis para um sistema dinâmico caótico envolve o parâmetro λ que deve ser suficientemente pequeno a fim transformar pontos fixos instáveis, do sistema caótico original, para estáveis através da transformações L_k. Quanto maior for o período a ser encontrado menor deverá ser o parâmetro λ. Entretanto, o parâmetro não pode ser muito pequeno, pois a convergência na iteração das leis dinâmicas transformadas será muito lenta.

Há uma maneira simples de evitar o ajuste do parâmetro λ, tomando-se o limite $\lambda \rightarrow 0$ o que nos leva a

$$\lim_{\lambda \rightarrow 0} \frac{(\vec{r}_{i+1} - \vec{r}_i)}{\lambda} = \dot{\vec{r}} = C_k(\vec{f}(\vec{r}_i) - \vec{r}_i). \tag{5}$$

Esta equação representa a formulação contínua da transformação do sistema dinâmico discreto Eq. (4).

As soluções da Eq. (5) não dependem do parâmetro λ, podendo ser facilmente resolvido usando um integrador de equações diferenciais ordinárias (ex: Runge-Kutta (Press, 1992)).

Na preposição de (SD) as sementes (condições iniciais) para a Eq. (5) devem ser obtidas ou pela determinação de uma grade sobre o atrator caótico ou por pontos aleatótios pertencentes a um certo limite do espaço de fase contendo o invariante caótico de interesse. A evolução temporal destas sementes para cada uma das matrizes C_k pode *ou* levar as proximidades da solução do ponto fixo *ou* pode divergir, saindo da região do espaço de fase que contém o atrator caótico.

As órbitas periódicas estáveis dos sistemas transformados podem então ser encontradas simplesmente iterando estes sistemas. Para demonstrar a eficiência deste método, este foi aplicado ao atrator do mapa bidimensional de Ikeda (Ikeda, 1979). Assim, as órbitas periódicas instáveis até o período 13 do atrator estranho de Ikeda são encontradas. Subseqüentemente, deriva-se seus coeficientes da estabilidade, usa-se as expansões de acordo com a teoria da órbita periódica e determina-se os expoentes médios de Lyapunov, a dimensão fractal, e a entropia topológica do atrator de Ikeda.

3 Método de Davidchack e Lai (Davidchack, 2000).

Este método usa um esquema iterativo baseado no método semi-implícito de Euler (Press, 1992), e tem as seguintes propriedades favoráveis:

- Perto de um ponto da órbita periódica exibe uma convergência rápida e similar a do método tradicional de Newton-Raphson NR (Press, 1992);
- Quando afastado dos pontos da órbita é similar ao método do SD (Schemelcher, 1998) e, conseqüentemente, apresenta convergência.

Este método DL (Davidchack, 2000) combina os métodos SD (Schemelcher, 1998) e NR (Press, 1992) como podemos observar na Tabela (2):

Tabela 2: Representação esquemática comparativa das prescrições dos métodos SD, NR, DL.

	SD	***NR***
Original:	$\vec{r}_{i+1} = \vec{r}_i + \Lambda_k(\vec{f}(\vec{r}_i) - \vec{r}_i)$	$\vec{f}(\vec{r}_i) - \vec{r}_i = -J(\vec{r}_i)\delta(\vec{r}_i)$
Equivalências:	$g(\vec{r}_i) \equiv \vec{f}(\vec{r}_i) - \vec{r}_i$ $\delta(\vec{r}_i) \equiv \vec{r}_{i+1} - \vec{r}_i$ $\frac{1}{\lambda} \equiv \beta\|\|g(\vec{r}_i)\|\|$	$g(\vec{r}_i) \equiv \vec{f}(\vec{r}_i) - \vec{r}_i$
SD & NR:	$\beta\|\|g(\vec{r}_i)\|\|\delta(\vec{r}_i) = C_k g(\vec{r}_i)$	$-C_k J(\vec{r}_i)\delta(\vec{r}_i) = C_k g(\vec{r}_i)$
DL:	$[\mathbf{1}\beta\|\|g(\vec{r}_i)\|\| - C_k J(\vec{r}_i)]\delta(\vec{r}_i) = C_k g(\vec{r}_i)$	

- Nas proximidades do ponto fixo, $||g(\vec{r}_i)|| = \sqrt{\sum_k (f_k^p(\vec{r}_i) - r_k)^2} \rightarrow 0$, e a convergência quadrática do NR[48] é preservada pelo método;
- Quando afastado do ponto fixo e valores grandes de β, o método preserva quase que total a propriedade de convergência global característica do SD.

Resolvendo-se o sistema de equações lineares obtemos $\delta(\vec{r})$ e iteramos nossas variáveis por

$$\vec{r}_{n+1} = \vec{r}_n + \delta(\vec{r}_n). \tag{6}$$

Para encontrar pontos fixos ou pontos de período 2, tanto uma grade fina de pontos definida sobre o atrator ou uma grade de pontos aleatórios contidos na região do atrator caótico podem ser utilizadas como semente para iteração do algoritmo de DL.

Para períodos $p > 2$, os autores deste método afirmam que os pontos periódicos de períodos inferiores a p são suficientes para iniciar a busca de pontos de órbitas de período p, o que é muito mais efetivo do que usar pontos aleatórios selecionados no espaço da fase ou no atrator. Na maioria dos casos é suficiente usar apenas pontos das órbita de período $p-1$ a fim de detectar todas as órbitas periódicas instáveis de período p.

O calculo das órbitas periódicas instáveis, com esta estratégia, do mapa de Ikeda-Hammel-Jones-Moloney, de período 22 com 106 pontos necessita o mesmo tempo computacional que o método SD (Schemelcher, 1998) requer para período 13 com 6000 pontos. Devido a eficiência do método, o cálculo de órbitas periódicas instáveis nos sistemas de dimensões elevadas é possível.

4 Resultados

Nesta seção apresentamos resultados de nossa análise comparativa entre os dois métodos descritos. Procuramos nesta análise verificar a eficácia do método em localizar todos os pontos periódicos existentes para cada período, a precisão das respostas fornecidas, tempo de processamento, eficiência para buscas de órbitas de períodos longos e possíveis alterações e otimizações aplicáveis a cada um dos métodos.

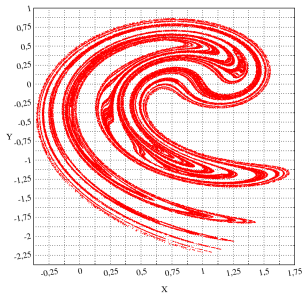


Figura 1: Atrator de Ikeda

Os testes foram realizados a partir da integração numérica da descrição contínua de SD dada por

$$\vec{r}_{i+1} = C_k(\vec{f}(\vec{r}_i) - \vec{r}_i) \tag{7}$$

e pela iteração do algoritmo de DL que baseia-se na solução do sistema linear dado por

$$[\mathbf{1}\beta||g(\vec{r}_i)|| - C_k J(\vec{r}_i)]\delta(\vec{r}_i) = C_k g(\vec{r}_i) \tag{8}$$

Ambos foram aplicados à busca de PF e soluções periódicas do Mapa de Ikeda definido por

$$f(x,y) = (A + B(x\cos(\omega) - y\sin(\omega)), B(x\sin(\omega) + y\cos(\omega))), \tag{9}$$

onde $\omega = K - \frac{E}{1+x^2+y^2}$, (x, y) são as duas variáveis dinâmicas e os valores adotados para os quatro parâmetros são $A = 1.0$, $B = 0.9$, $K = 0.4$ e $E = 6.0$, que correspondem ao atrator caótico apresentado na Fig.(1).

Os parâmetros numéricos do método de SD são passo inicial do algoritmo de Runge-Kutta de quarta ordem de passo variável, número máximo de iterações por condição inicial, tolerância exigida e tolerância que diferencia soluções distintas. Todos fixos conforme prescrição dos autores. Já para o segundo método, utilizamos dois valores para o parâmetro β, valores estes definidos de acordo com o período das solucões procuradas. Por sua vez, o numéro máximo de iterações permitidas foi definido como função linear de β dada por $100 + 2\beta$.

Para período 1 e 2, são utilizadas 10000 condições iniciais sobre os limites do espaço de fase definidos por $\{-1.0, 2.0; -3.0, 1.5\}$, que contém o atrator caótico como pode ser verificado pela Fig.(1). Obtivemos como resposta os pontos fixos da Tab. (3). Como se pode observar, com ambos os métodos encontramos os mesmos pontos fixos, porém com precisões bem distintas. Para facilitar a inspeção visual, na Tab. (3) representamos em negrito os algarismos decimais iguais entre soluções dos dois métodos.

Tabela 3: Pontos fixos de período 1 e 2 utilizado $C_1 - C_5$ e $N = 10000$ condições iniciais aleatórias sobre o atrator.

PF	***SD***	***LD***
1a	(**5.3275462294078**834e-01, **2.4689677271101**148e-01)	(**5.3275462294079**023e-01, **2.4689677271101**318e-01)
1b	(**1.1142696145813**442e+00, **-2.2856944609869**831e+00)	(**1.1142696145814**104e+00, **-2.2856944609861**602e+00)
2a	(**6.2160432324656**689e-01, **6.0593364761280**199e-01)	(**6.2160432324652**715e-01, **6.0593364761281**554e-01)
2b	(**5.0983725038941**063e-01, **-6.0836992864915**973e-01)	(**5.0983725038942**507e-01, **-6.0836992864917**161e-01)

A fim de testar a estacionaridade destas soluções, aplicamos o Mapa de Ikeda a estes pontos. Os valores obtidos de distâncias entre sucessivas iterações e pontos iniciais como função do número de iterações são apresentadas na Fig. (2). Verifica-se que as soluções fornecidas pelo método de DL foram muito mais precisas que os SD. Apenas para o ponto $p2b$, a distâncias não se mantem muito próxima a zero. Já para as soluções de SD, todos as trajetórias se afastam do ponto candidato a ponto fixo após algumas dezenas de iterações.

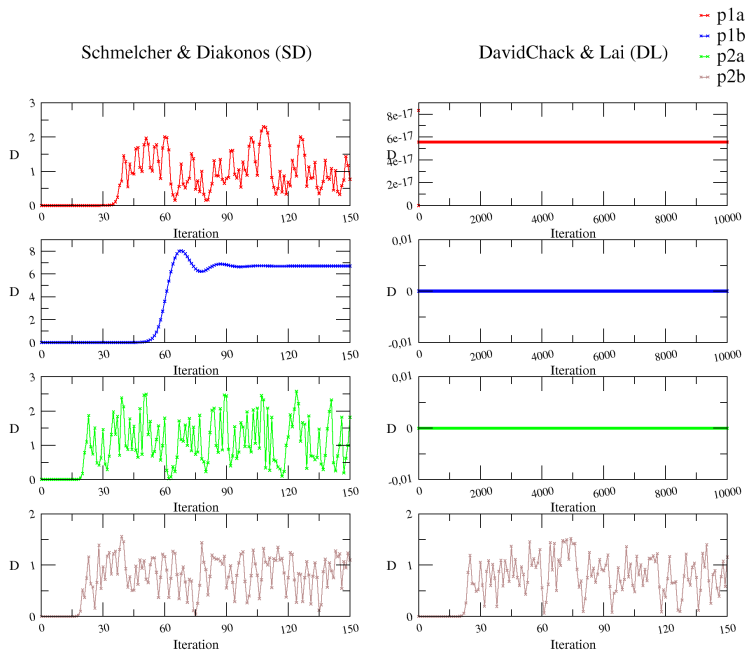


Figura 2: Resultado do teste de estacionaridade das soluções fornecidas pelos métodos de SD e DL segundo aplicações do Mapa de Ikeda. Apresentamos a distância euclideana entre as soluções fornecidas por SD e DL e suas sucessivas iterações pelo Mapa do Ikeda como função do número de iterações.

Para busca de pontos fixos de período 3, utilizamos como semente de ambos métodos apenas as respectivas soluções de períodos 1 e 2 que cada um dos métodos forneceu. A Tab. (3) apresenta um sumário das soluções finais obtidas a partir de cada um dos dois pontos fixos de período $1 (p1a \ e \ p1b)$ e dos dois pontos de período $2 (p2a \ e \ p2b)$ mediante a aplicação dos métodos usando cada uma das oito matrizes C_k. Os estados finais atingidos foram: alguma solução de período 1 ou 3, divergência para fora da região do atrator caótico do Mapa do Ikeda e mesmo ponto inicial (estacionaridade observada para alguns casos de ponto fixo de período 1). Já na Fig. (3), apresentamos as trajetórias obtidas no espaço de fase (x, y). A primeira e terceira coluna referem-se ao método de SD, enquanto a segunda e quarta ao método de DL. Diferentes matrizes C_k correspondem aos diferentes gráficos. Em cada um deles, apresentamos as trajetórias iteradas para cada uma das quatro condições inicias utilizando-se uma das matrizes C_k conforme indicado.

Tabela 4: Soluções finais obtidas pelas iterações dos métodos de SD e DL utilizando-se todas matrizes C_k e partindo de PF de $1P$ e $2P$ como condições iniciais.

PF	***SD***								***DL***							
	Matrizes C_k *d* = divergência para fora do invariante caótico *e* = comportamentos estáticos *l* = laços (ciclos limites) símbolos restantes referem-se aos rótulos adotados para cada ponto fixo (vide Fig. (3))															
	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8
p1a	e	d	d	l	e	d	e	e	e	e	e	e	e	e	e	e
p1b	d	e	d	d	e	e	d	e	e	e	e	e	e	e	e	e
p2a	1a	3ba	d	3bc	d	d	l	3ac	1a	3ba	d	3bc	d	d	1a	3ac
p2b	3aa	d	d	l	3ba	d	l	3aa	3aa	d	d	3bb	3ba	d	1a	3aa

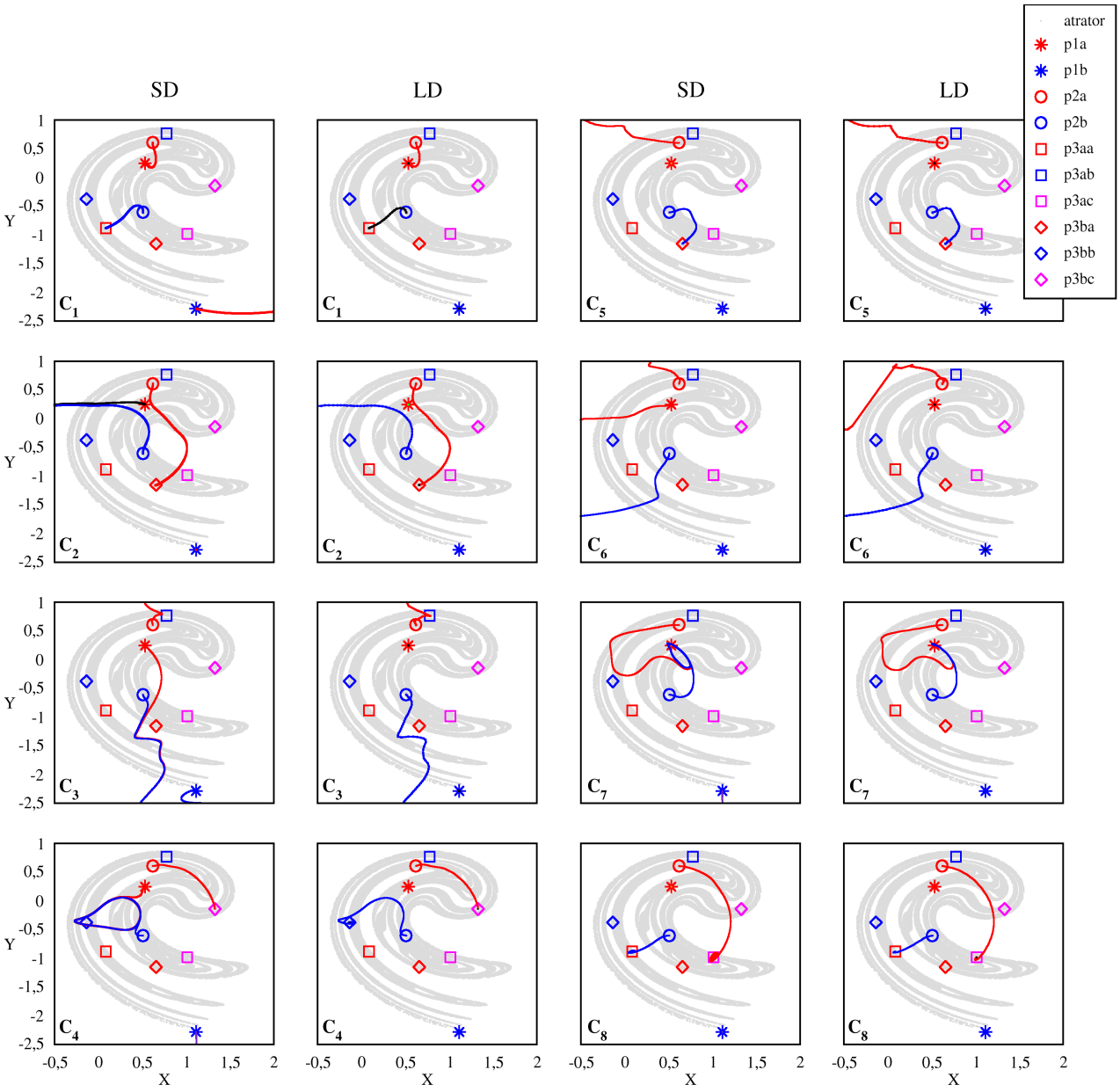


Figura 3: Trajetórias obtidas pela busca de PF de período 3 com a iteração dos dois métodos utilizando as 8 matrizes C_k e partindo de condições iniciais de período 1 e 2.

5 Conclusão

Analisamos neste trabalho dois métodos de detecção de órbitas periódicas instáveis em sistemas dinâmicos em regime caótico recentemente propostos na literatura: os métodos de Schmelcher & Diakonos (SD) e de Davichack & Lai (DL). Podemos verificar de nossos estudos que o método SD apresenta algumas características indesejáveis, porém de fácil solução. São elas:

1. Presisão das soluções muito pior que as obtidas por DL;
2. Elevado tempo de processamento quando soluções com maiores precisões são exigidas;
3. Perda de tempo computacional em longos laços quando ciclos limites da dinâmica de SD são atingidos, próximos a região contendo um PF de interesse (não encontrado por esta iteração da dinâmica de SD).

Algumas estratégias simples, adotadas pelo método de DL, podem ser aplicadas a SD de modo a eliminar estas características que degradam fortemente a eficiência deste método.

Em relação as precisões das soluções finais e longos tempos de processamento para refiná-las, uma estratégia natural é utilizar o algoritmo de NR quando o método de SD leva a proximidades de um condidato a PF. Esta condição pode ser verificada através do teste de $||g(\vec{r})||$. Para valores de $||g(\vec{r})||$ menores que um certo parâmetro fixo, o método de NR torna-se responsável por melhorar a precisão da solução.

Os laços longos referidos no item 3 podem ser eliminados pela simples introdução do NR ou através de interrupções de evoluções temporais longas que não estejam convergindo para solução.

Em nossos testes verificamos que apenas utilizando soluções de períodos inferiores nem todos PF de período 3 foram encontrados por *ambos métodos* (vide Fig. (3)e Tab. (4)). Entretanto a aplicação do mapa do Ikeda às soluções encontradas leva a estes pontos que faltam.

Para períodos superiores (até 13) verificamos que a utilização apenas de pontos de períodos $(p-1)$ como sementes para iniciar a busca de pontos de períodp p seguidas pela iteração pelo mapa de Ikeda das soluções fornecidas sã suficientes para encontrar todos os PF de um dado período *pelo método de SD*. Esta investigação não havia ainda sido feita nos trabalhos que citamos, uma vez que apenas conjuntos de muitos pontos aleatórios sobre o invariante caótico haviam sido utilizados para realizar a mesmo tarefa.

É nosso objetiva investigar a eficiência do método de SD após a inclusão de todos esses aprimoramentos e então comparar novamente ambos métodos.

6 Agradecimentos

Este trabalho foi financiado parcialmente pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq.

7 Referências

Aguirre, L. A. Souza, A. V. An Algorithm for estimaing fixed points of dynamical systems from time series. International Journal of Bifurcation and Chaos, v. 8, n. 11, p. 2203-2213, Abr. 1998.

Auerbach, D.; Cvitanovic, P.; Eckmann, J.; Gunaratne G.; Procaccia I. Exploring Chaotic Motion Through Periodic Orbits. Physical Review Letters, v.58, n. 23, p. 2387-2389, Jun. 1987.

Auerbach, D. Scaling of periodic orbits in two-dimensional chaotic system. Physical Review A, v. 41, n. 12, p. 6692-6701, Jun. 1990.

Baddi, R.; Brun, E.; Finardi; Flepp, L.; Holzner, Parisi, J; Reyl, C; Simonet, J. . Physical Review Letters, v. , n. 66, p. 1389 1994.

Balian, R.; Bloch, C. Ann. Physical (NY), v. 85 p. 514, 1974.

Bowen, R.; Ruelle, D. Invent. Math, v. 79 p. 181, 1975.

Christiansen, F.; Politi, A. . Physical Review A, v. , n. 24 p. 1837 1991.

Cvitanovic, P. Invariant measurement of strange sets in terms of cycles. Physical Review Letters, v. 61, n. 23, p. 2729, Jun. 1988.

Cvitanovic-II, P.; Eckhardt B. Periodic-Orbit Quantization of Chaotic Systems. Physical Review Letters, v. 63, n. 8, p. 823-826 Jun. 1988.

Davidchack, R. L.; Lai, Y.; Bollt, E. M.; Dhamala, M. Estimating generating partitions of chaotic systems by unstable periodic orbits. Physical Review E, v.61, n. 2, p. 1353-1356, Fev. 2000.

Eckmann, J. P.; Ruelle, D. . Rev. Mod, Phys, v. 57, n., p. 617, 1985.

Farmer, J. D; Ott, E.; Yorke, J.A. Physica (Amsterdam), v. 7D p. 153, 1983.

Giovannini, F; Politi, A. . Physical Review A, v. , n. 24, p. 1837 1991.

Grassberger, P. . Physical Review Letters, v. 97a, n., p. 227, 1983.

Grassberger, P. . Physical Review Letters, v. 107a, n., p. 101, 1985.

Grassberger-II, P.; Kantz, H. . Physical Review A, v. , n. 113, p. 235, 1985.

Grebogi, C.; Ott, E.; Yorke J. A. Unstable periodic orbits and the dimensions of multifractal chaotic attractors. Physical Review A, v. 37, n. 5, p. 1711-1724, Mar. 1988.

Gutzwiller, M. C. Jornal Math Physical, v. 8, 10, 11, 12 p. 1979, 1004, 1791, 343, 1967, 1969, 1970, 1971.

Halsey, T. C.; Jensen, M. H.; Kadanoff, L. P.; Procaccia, I.; Sharaiman, B. I. . Physical Review A, v. , n. 33, p. 1141, 1986.
Hao, B.-L. Elementary Symbolic Dynamics and Chaos in Dissipative Systems. World Scientific, Singapore, v. , n. , p. , 1989.
Hayes, S.; Grebogi, C.; Ott, E. . Physical Review Letters, v. , n. 70, p. 3031 1993.
Hénon, M. . Comm. Math. Phys., v. , n. 50 p. 69 1976.
Hentschel, H. G. E.; Procaccia, I. . Physical Review D, v. 8, n., p. 435, 1983.
Ikeda, K. . Opt. Commun, v. , n. 30, p. 257, 1979.
Lai, Y.; Nagai, Y.; Grebogi C. Characterization of the Natural Measure by Unstable Periodic Orbits in Chaotic Attractors. Physical Review Letters, v. 79, n. 4, p. 649-652, Jul. 1997.
Lefranc, M.; Glorieux, P.; Papoff, F.; Molesti F.; Arimondo E. . Physical Review Letters, v. , n. 73, p. 1364 1994.
May, R. . Nature, London, v. , n. 261, p. 459, 1976.
Miller, W. M.; Bloch, C. Ann. Physical (NY), v. 30 p. 77, 1975.
Ott, E.;Grebogi, C.; Yorke J. A. Controling Caos. Physical Review Letters, v. , n. 64 p. 1196, Out. 1990.
Press, W. H.; Teukolsky, S. A. Numerical Recipes in C: The Art of Scientific Computing. 2.ed. Cambridge: Press Syndicate, 1992. 994 p.
Schmelcher, P.; Diakonos, F. K. General apprach to the localization of unstable periodic orbits in chaotic dynamical system. Physical Review E, v. 57, n. 3, p. 2739-2746, Mar. 1998.

8 Direitos Autorais

DETECTING FIXED AND PERIODIC ORBITS IN CHAOTIC SYSTEMS

Roth, Bruno Aloís Forlin
Instituto Nacional de Pesquisas Espaciais - INPE,
Laboratório Associado de Computação e Matemática Aplicada - LAC,
Av. dos Astronautas, 1758. S. J. Campos, SP, C.P. 515, 12227-010, Brasil.
roth@pgrad.inpe.br

Macau, Elbert Einstein Nehrer
Instituto Nacional de Pesquisas Espaciais - INPE,
Laboratório de Integração e Testes - LIT,
Av. dos Astronautas, 1758. S. J. Campos, SP, C.P. 515, 12227-010, Brasil.
elbert@lit.inpe.br

Terra, Maisa de Oliveira
Instituto Nacional de Pesquisas Espaciais - INPE,
Laboratório de Integração e Testes - LIT,
Av. dos Astronautas, 1758. S. J. Campos, SP, C.P. 515, 12227-010, Brasil.
maisa@lit.inpe.br

Abstract. *The detection of periodic orbits of chaotic systems is of fundamental importance for characterization of dynamical systems. Unstable periodic orbits reveal the skeleton of the chaotic set, and even quantities which characterize the behavior of the chaotic system, such as dimensions, Lyapunov expoents, and topological entropy, can be calculated with infinite set of these orbits. In this paper some methods of determination of unstable fixed point an periodic points in dynamical systems are explored. The efficiency and the aplicability of those methods are verified. Our final goal is the development of an analysis enviroment and a library for fixed and periodic point determination in chaotic systems.*

Keywords:. *chaos, unstable periodic orbits, nolinear dynamical systems.*

XVI CONGRESSO BRASILEIRO DE ENGENHARIA MECÂNICA
16th BRAZILIAN CONGRESS OF MECHANICAL ENGINEERING

SIMULAÇÃO NUMÉRICA E VISUALIZAÇÃO DOS PROCESSOS DE EXTRAÇÃO EM INSTALAÇÕES INDUSTRIAIS "CROWN-MODEL"

Marcos Ronaldo Melo Cavalheiro
UNIJUI – Universidade Regional do Noroeste do Estado do Rio Grande do Sul, Departamento de Tecnologia, Caixa Postal 560, CEP: 98700-000, Ijuí, RS, Brasil
E-mail: mrmc@detec.unijui.tche.br

Claudia Piva
Universidade Regional do Noroeste do Estado do Rio Grande do Sul, Programa de Pós Graduação em Modelagem Matemática, Caixa Postal 560, CEP 98700-000 Ijuí, RS, Brasil
E-mail: claudiap@main.unijui.tche.br

Resumo. *O trabalho refere-se à modelagem matemática dos processos no trajeto conexivo do extrator "Crown Model", que é um fragmento importante do modelo matemático deste extrator. O esquema físico do trajeto considera os seguintes fenômenos: movimento da matéria-prima e miscela pelo trajeto; transferência de óleo entre fases sólida e poro, entre fases poro e bulk; difusão de todo campo extrativo; etc. O modelo matemático apresenta-se por três equações em derivadas parciais, não lineares bidimensionais, evolutivas, com condições de contorno da largura entre os trechos horizontais com trajeto, e considerando a vazão da recirculação da miscela. O algoritmo baseia-se no método de linhas. Foi criado um aplicativo para simular os processo no extrator (linguagem C++) e código para visualizar estes processos (linguagem Delphi). Com estas ferramentas foram realizadas as simulações para todo o extrato que mostraram a concordância dos resultados com as leis físico-químicas, e com as propriedades das características operacionais do extrator deste tipo.*

Palavras chave: *Extração de Óleo, Instalação Industrial, Modelagem Matemática, Visualização Gráfica.*

1. Introdução

Um dos setores importantes da indústria de produção de alimentos é o setor de extração de óleo vegetal de grãos (soja, milho, girassol, arroz...) por solvente. Entre os extratores industriais contemporâneos destaca-se o extrator "Crown Model", com grande potência na fabricação de óleo (Bockisch, 1998; Erickson,1995). A modelagem matemática dos processos ocorridos neste extrator pode contribuir essencialmente para sua projeção com maior confiança e para determinar os regimes operacionais ótimos. No trabalho (Benetti, 2001), foi apresentado o modelo matemático dos processos em trechos horizontais, sessões de drenagem e carregamento do extrator "Crown Model". O presente trabalho complementa definitivamente o seu modelo matemático, desenvolvendo o modelo dos processos no trajeto conexivo. O esquema principal de todo o extrator é apresentado na Fig 1. A descrição geral dos processos nos trechos horizontais é dada em (Benetti, 2001). Observa-se somente que os componentes principais que passam através do extrator são a matéria-prima (um meio poroso especialmente preparado de grãos que inclui óleo vegetal), o solvente (hexano) que extrai o óleo e a miscela (a mistura liquida do hexano com óleo).

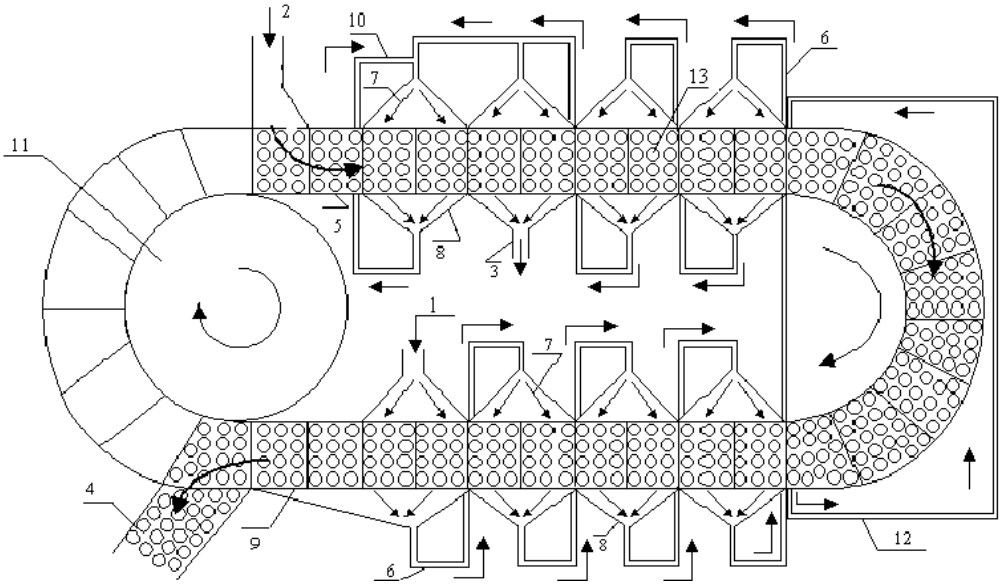


Figura 1. Esquema principal do extrator "CROWN MODEL": 1- Entrada do hexano; 2- Entrada da matéria-prima; 3- Saída da miscela concentrada; 4- Saída do farelo branco; 5- Fita do transportador; 6- Tubos de miscela; 7- Distribuidores; 8- Bandejas; 9- Zona de drenagem; 10- Recirculação da miscela na seção de carregamento; 11- Engrenagem; 12- Tubo de recirculação.

Os espaços externos entre as partículas da matéria-prima (Fig. 2-b) formam a fase bulk, com porosidade ε_b, que durante a extração está preenchida por miscela (a mistura do hexano com óleo), com concentração *(C)* de óleo. A partícula inclui a fase poro, com porosidade ε_p, que ocupa os espaços internos que também é preenchida por miscela, mas com concentração de óleo $C^p > C$. A fase sólida com concentração C^N no início da extração, contém somente o óleo, mas à medida que ocorre a lavagem, há saída de óleo da fase sólida e entrada do hexano nesta fase.

Em geral, a matéria-prima e a miscela movem-se em direções contrárias. A matéria-prima sob a ação da tremonha desloca-se pelo trecho superior através do trajeto conexivo, e o trecho inferior no sentido horário. A miscela move-se na direção contrária de modo cruzado, iniciando do trecho inferior para o trecho superior. Alguma vazão da miscela passa de cima para baixo através do trajeto conexivo, que obriga a montagem de um tubo (Fig. 2-a) para recirculação da miscela. Além dessa propriedade, o trajeto conexivo possui outras particularidades, a saber:

- a matéria-prima move-se pelo trajeto com velocidade que se altera pelo raio (as camadas mais distantes do centro movem-se mais rapidamente), por isso, perto da superfície exterior na parte acima do trajeto, surgem espaços vazios; mas na parte abaixo do trajeto a matéria-prima desloca-se sob força da gravidade à superfície exterior, e espaços vazios aparecem já perto da superfície inferior (observa-se que na entrada e na saída do trajeto conexivo a espessura da camada da matéria-prima é igual a espessura nos trechos lineares);
- no trajeto conexivo entra também a miscela envolvida pela matéria-prima. De acordo com o modo da ligação do trecho e trajeto, a miscela pode preencher toda espessura da camada ou alguma parte dela. No trajeto conexivo que não possui fundo perfurado, a miscela começa a mover-se com a velocidade da matéria-prima mas, com o aumento da coordenada angular (θ), aumenta-se a influência da força da gravidade e a miscela começa a mover-se mais rápido do que a matéria-prima e a diferença entre elas atinge o valor máximo quando (θ=90º). É evidente que, de acordo com a lei da conservação da massa, a área do fluxo da miscela neste caso diminui-se e alguma parte do fluxo da matéria-prima não vai ter temporariamente o contato com a miscela; quando a miscela atinge a parte abaixo (θ >90º), a influência da força gravitacional vai diminuir e a área de contato com a matéria-prima aumenta-se, mas pode acontecer que a miscela vai passar pelos espaços 8 (Fig 2-a) livres da matéria-prima.

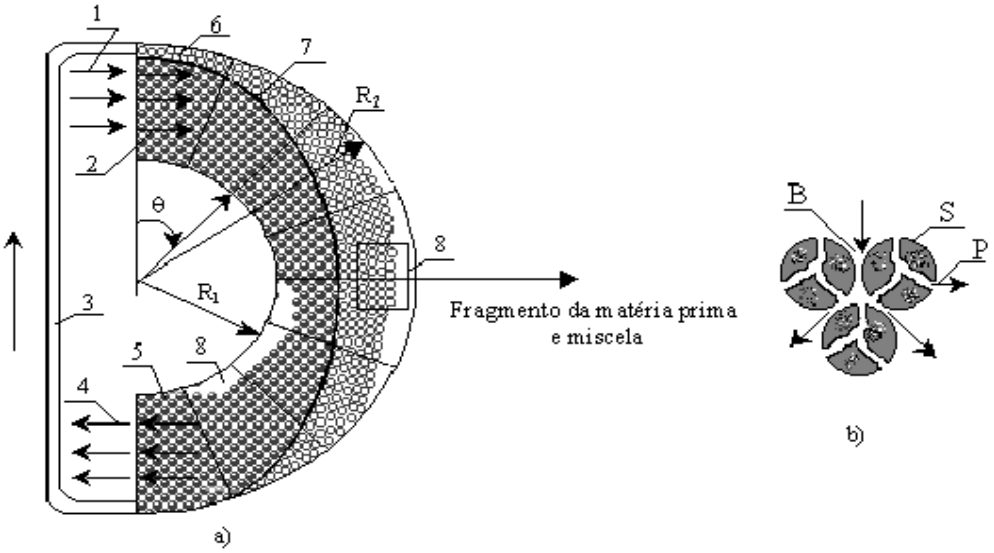


Figura 2. Esquema dos fluxos e das fases no trajeto conexivo do Extrator "Crown Model": 1- fluxo da miscela no fim do trecho superior; 2- fluxo da miscela no início do trajeto conexivo; 3-tubo com fluxo recirculante; 4- fluxo da miscela no inicio do trecho inferior; 5-fluxo da miscela no fim do trajeto conexivo; 6- limite do fluxo da miscela; 7- zona sem miscela; 8- zona sem matéria-prima; B- fase bulk; S- fase sólida; P- fase poro.

Então, pode-se observar que no conexivo trajeto também como nos trechos horizontais, existe a "lavagem" (ou seja, o movimento da miscela relativamente da matéria-prima) mas com a velocidade variável ao longo do trajeto. No resultado ocorrem a transferência do óleo da fase sólida até a fase bulk, e transferência equivalente do hexano em direções contrárias (da fase bulk até a fase sólida). Observa-se que devido a velocidade notável do transportador, as concentrações C^N e C^p podem estar em desequilíbrio.

Logo, observa-se que o quadro dos processos da extração em trajeto conexivo é mais complicado do que em trechos lineares. Mas os dados experimentais mostraram que no trajeto conexivo, a alteração das concentrações C, C^p e C^N não são grandes. Por isso o esquema dos processos em trajeto conexivo pode ser simplificado, usando as hipóteses seguintes:

- A miscela entra no trajeto conexivo com a velocidade do transportador (u).
- À medida que ela se desloca pelo trajeto, surge a velocidade relativa da miscela V_r, que altera-se conforme a projeção da força de gravidade na direção do movimento.
- A velocidade máxima da miscela situa-se na posição (θ=90º), é igual a velocidade vertical da miscela nos trechos horizontais.

– Para determinar a transferência da massa entre as fases poro e bulk, utiliza-se a velocidade efetiva V_t^{ef} , introduzida pela Eq. (1). Sua dedução é apresentada no trabalho (Piva, 2001),:

$$V_t^{ef} = V_m \operatorname{sen}\theta / \left(1 + \left(V_m / u\right)\operatorname{sen}\theta\right) \tag{1}$$

onde V_m – Velocidade vertical da miscela nos trechos.
– A matéria-prima move-se pelo trajeto com velocidade *u*.
– O fluxo de recirculação possui a concentração média na saída do trajeto e distribui-se uniformemente na miscela, da fase bulk, que entra no trajeto.
– A temperatura é uniforme por todo o trajeto.
– As paredes celulares das partículas de matéria-prima estão totalmente rompidas.
– A quantidade de óleo que sai das partículas substitui-se por uma mesma quantidade volumétrica de miscela.
– A distribuição das concentrações C^p e C^N, na entrada do trajeto, é igual a sua distribuição na saída do trecho superior.
– A distribuição das concentrações *C*, na entrada do trajeto, determina-se pela distribuição na saída do trecho superior, considerando a injeção do fluxo de recirculação.

O esquema físico considera os seguintes fenômenos em sua interação:
– Distribuições bidimensionais das concentrações *C* ,C^p e C^N pelo campo do trajeto.
– Difusão do óleo pela miscela ao longo de todo o trajeto, nas direções angular e radial.
– A influência do movimento da miscela relativamente da matéria-prima, na transferência do óleo.
– Transferência do óleo entre as fases poro e bulk.
– Transferência não equilibrada entre fases sólida e poro.

2. Modelo Matemático do Trajeto Conexivo

2.1. Deduções das Equações Modelo Matemático do Trajeto Conexivo

O modelo matemático é realizado de acordo com o esquema físico e apresenta-se por três equações em derivadas parciais, em coordenadas eulerianas, cilíndricas e inclui três variáveis independentes (t - tempo, *r* – radial *e* θ - angular). Na Fig 3 apresenta-se o esquema de divisão do campo de extração do trajeto conexivo, por células com tamanhos, $\left(r + \Delta r/2\right)\Delta\theta$, Δr

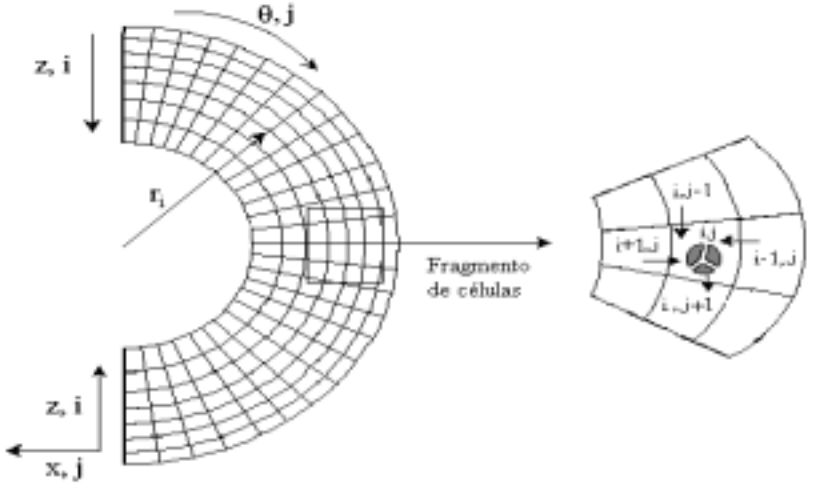


Figura 3. Esquema de divisão do campo de extração.

Deduzindo a equação da alteração do óleo na fase bulk e considerando que o volume ocupado por esta fase na célula *(i, j)* é $\varepsilon_b \Delta r\left(r_i + \Delta r/2\right)\Delta\theta$, e a concentração do óleo nesta fase, durante o tempo Δt, altera-se de $C_{i,j}^{t}$ até $C_{i,j}^{t+\Delta t}$, podemos escrever o termo para a alteração do conteúdo de óleo na célula *(i, j)*:

$$\varepsilon_b \Delta r\left(r_i + \Delta r/2\right)\Delta\theta\left(C_{i,j}^{t+\Delta t} - C_{i,j}^{t}\right) \tag{2}$$

Durante o tempo Δt, entre as células *(i, j-1)*; *(i, j)* e *(i, j+1)* , na direção radial, por causa da difusão, passa a quantidade de óleo:

$$E_S(r_i + \Delta r)\Delta\theta\varepsilon_b\Delta t\left(C_{i-1,j} - C_{i,j}\right)/\Delta r \; + \; E_S r_i\Delta\theta\varepsilon_b\Delta t\left(C_{i+1,j} - C_{i,j}\right)/\Delta r \tag{3}$$

onde, E_s é o coeficiente da dispersão (Cussler, 1997).

Analogamente, na direção angular, temos entre as células *(i, j-1)*; *(i, j)* e *(i, j+1)*:

$$E_S\Delta r\varepsilon_b\Delta t\left(C_{i,j-1} - C_{i,j}\right)/\left((r_i + \Delta r/2)\Delta\theta\right) + E_S\Delta r\varepsilon_b\Delta t\left(C_{i,j+1} - C_{i,j}\right)/\left((r_i + \Delta r/2)\Delta\theta\right) \tag{4}$$

Na célula *(i, j)*, a transferência de óleo da fase poro para a fase bulk é:

$$k_f(1-\varepsilon_b)a_p\Delta r\Delta\theta(r_i + \Delta r/2)\left(C^p_{i,j} - C_{i,j}\right)\Delta t \tag{5}$$

onde k_f é o coeficiente da transferência do óleo entre fases poro e bulk; E_S é o coeficiente de dispersão *(m^2/s); a_p* é a superfície de contato entre fases poro e bulk *(1/m)* e τ é o tempo *(s)*.

O termo que representa o fluxo conectivo de óleo na fase bulk da célula *(i, j-1)* para célula *(i, j)*, é dado por:

$$r_i\omega'\Delta r\Delta\theta\left(C_{i,j-1} - C_{i,j}\right)\Delta t \tag{6}$$

onde $\omega' = u/R_1$, que é a velocidade angular da tremonha no trajeto conexivo, R_1 – é o raio interno do trajeto.

Então, de acordo com a lei de conservação da espécie, temos a seguinte igualdade:

$$\begin{aligned}
&\varepsilon_b\Delta r(r_i + \Delta r/2)\Delta\theta\left(C^{t+\Delta t}_{i,j} - C^{t}_{i,j}\right) = E_S(r_i + \Delta r)\Delta\theta\varepsilon_b\Delta t\left(C_{i-1,j} - C_{i,j}\right)/\Delta r + \\
&E_S\left(\left(C_{i+1,j} - C_{i,j}\right)/\Delta r\right)r_i\Delta\theta\varepsilon_b\Delta t + k_f(1-\varepsilon_b)a_p\Delta r(r_i + \Delta r/2)\Delta\theta\left(C^p_{i,j} - C_{i,j}\right)\Delta t + \\
&E_S\left(\left(C_{i,j-1} - C_{i,j}\right)/(r_i + \Delta r/2)\Delta\theta\right)\Delta r\varepsilon_b\Delta t + E_S\left(\left(C_{i,j+1} - C_{i,j}\right)/\left((r_i + \Delta r/2)\Delta\theta\right)\right)\Delta r\varepsilon_b\Delta t + \\
&r_i\omega'\left(C_{i,j-1} - C_{i,j}\right)\Delta r\varepsilon_b\Delta t
\end{aligned} \tag{7}$$

Dividindo a Eq. (7) por $\varepsilon_b\Delta r(r_i + \Delta r/2)\Delta\theta\Delta t$ e simplificando, e realizando a passagem do limite Δr , $\Delta\theta$, $\Delta t \to 0$ obtemos a equação para alteração de óleo na fase bulk:

$$\partial C/\partial t = E_S\left(\partial^2 C/\partial r^2 + 1/r\,\partial C/\partial r\right) + E_S\,1/r^2\,\partial^2 C/\partial\theta^2 + k_f a_p\,(1-\varepsilon_b)/\varepsilon_b\left(C^p - C\right)\omega'\partial C/\partial\theta \tag{8}$$

Analogamente, obtemos as equações das fases poro e sólida, respectivamente:

$$\partial C^p/\partial t = -\left((1-\varepsilon_b)/\varepsilon_p\right)\left[k_N\left(C^N - E_d C^p\right)\right] - \left(k_f a_p/\varepsilon_p\right)\left(C^p - C\right) - \omega'\left(\partial C^p/\partial\theta\right) \tag{9}$$

$$\partial C^N/\partial t = -k_N\left(C^N - E_d C^p\right) - \omega'\left(\partial C^N/\partial\theta\right) \tag{10}$$

onde, k_N - é o coeficiente da transferência do óleo da fase sólida para fase poro e E_d é o coeficiente de equilíbrio entre fases sólida e poro.

2.2. Condições Iniciais e de Contorno

Para resolver o sistema de Eq. (8), (9), (10) no trajeto conexivo, as condições iniciais para as fases bulk, sólida e poro são:

$$C(\theta, r, 0) = C_0(\theta, r)\,;\; C^N(\theta, r, 0) = C^N_0(\theta, r)\,;\; C^p(\theta, r, 0) = C^p_0(\theta, r) \tag{11}$$

para $\theta = 0...\pi$; $r = R_1...R_2$, onde R_2 – é o raio externo do trajeto.

As condições de contorno para fase bulk são as seguintes:

- na entrada do trajeto, (considerando o fluxo de recirculação):

$$C(\theta, r, t) = (u_m/u)C(z, x, t) + (1 - (u_m/u))\int_{R_1}^{R_2} (C(\theta_f, r, t)dr)/L_r \tag{12}$$

para $\theta = 0; r = R_1...R_2; t > 0;\ z = 0....L_r;\ x = m_1.X_S; \theta_f = \pi$, onde L_r é a altura do leito; m_1 é número de seções do trecho superior e X_S é a largura de lavagem de uma seção.

- na superfície inferior

$$\partial C/\partial r(\theta, r, t) = 0 \text{ ; para } \theta = 0...\pi;\ r = R_1;\ t > 0 \tag{13}$$

- na superfície superior

$$\partial C/\partial r(\theta, r, t) = 0 \text{ ; para } \theta = 0...\pi;\ r = R_2;\ t > 0 \tag{14}$$

na saída do trajeto:

$$C(\theta, r, t) = C(z, x_2, t) \tag{15}$$

para $\theta = \pi; r = R_1...R_2; t > 0;\ z = 0....L_r;\ x_2 = 0;$

As condições de contorno para a fase sólida e poro na entrada do trajeto:

$$C^N(\theta, r, t) = C^N(z, x, t);\quad C^p\ (\theta, r, t) = C^p(z, x, t) \tag{16}$$

para $\theta = \pi; r = R_1...R_2; t > 0;\ z = 0....L_r;\ x = m_1 X_s;$

O coeficiente de transferência de massa (k_f) entre as fases poro e bulk, determina-se pela fórmula (Treybal, 1963).

$$k_f = (Sh.D_{AB})/d_p\ ,\ Sh = a\ \mathrm{Re}^b\ .Sc^{0,42} \quad \text{e} \quad \mathrm{Re} = (V_{ef}\, d_p)/V \tag{17}$$

onde:

a) $a = 2,4$ e $b = 0,34$ para $0,08 < Re < 125$

b) $a = 0,442$ e $b = 0,69$ para $125 < Re < 5000$

D_{AB} – é o coeficiente de difusão; d_p – diâmetro condicional das partículas da matéria-prima; Re – número de Reynolds; *S*c – número de Schmidt.

A velocidade efetiva V_{ef} necessita-se para determinar o número *Re* é dada pela Eq. (1).

3 Algoritmo de Cálculo e a sua Verificação

Para resolver o sistema de Eq. (8), (9), (10) junto com as equações dos trechos horizontais apresentados em (Benetti, 2001), optou-se aplicar o método de linhas (Lambert,1978), onde as equações em derivadas parciais (EDP) transformam-se em equações diferenciais ordinárias (EDO's), deixando as derivadas temporais e discretizando as derivadas espaciais. Criando uma malha através da divisão do domínio do trajeto em *n* camadas e N_{pc} setores (Fig. 3), obtemos o seguinte sistema de equações diferenciais ordinárias (EDO's):

$$\begin{aligned} dC_{i,j}/d\tau = {} & T_1\,(r_i + \Delta r)\left(C_{i-1,j} - 2C_{i,j} + C_{i+1,j}\right)/(r_i + \Delta r/2) - T_2\,\left(C_{i+1,j} - C_{i,j}\right)/(r_i + \Delta r/2) + \\ & + R_t\left(C^p_{i,j} - C_{i,j}\right) + T_3\,\left(C_{i,j-1} - 2C_{i,j} + C_{i,j+1}\right)/(r_i + \Delta r/2)^2 + T_4\, r_i\,\left(C_{i,j-1} - C_{i,j}\right)/(r_i + \Delta r/2) \end{aligned} \tag{18}$$

$$dC_{i,j}^{N}/d\mathrm{t} = -k_N\left(C_{i,j}^{N} - E_d C_{i,j}^{p}\right) - T_4\left(C_{i,j}^{N} - C_{i,j-1}^{N}\right) \tag{19}$$

$$dC_{i,j}^{p}/d\mathrm{t} = -R_{ep}\left(C_{i,j}^{N^{\mathrm{t}}} - E_d C_{i,j}^{p}\right)/\Delta\mathrm{t} - T_5\left(C_{i,j}^{p} - C_{i,j}\right) + T_4\left(C_{i,j-1}^{p} - C_{i,j}^{p}\right) \tag{20}$$

onde: $T_1 = E_S/\Delta r^2$; $T_2 = E_S/\Delta r$; $T_3 = E_S/\Delta\mathrm{q}^2$; $T_4 = \mathrm{q}'/\Delta\mathrm{q}$; $R_t = k_f a_p\left(1-\mathrm{e}_b\right)/\mathrm{e}_b$;

$T_5 = k_f a_p/\mathrm{e}_p$; $R_{ep} = \left(1-\mathrm{e}_p\right)/\mathrm{e}_p$;

com $i = 1...n$ e $j = 1...N_{pc}$.

Este sistema resolve-se pelo método de Runge-Kutta de 4 ordem. O aplicativo (CROWN) para integração do sistema das equações Eq. (18 - 20) junto com as equações dos trechos horizontais (Benetti, 2000) foi criado em linguagem C++.

Para avaliar a precisão de cálculos e determinar os intervalos seguros de alteração dos parâmetros da malha nos domínios de integração, foram realizadas duas séries computacionais:

a) alterar os parâmetros (n, N_{h1}, N_{pc} e N_{h2}) quando o passo h é constante;

b) alterar o passo h quando os parâmetros da malha são constantes.

Para caso a) os resultados comparativos apresentam-se na tabela 1. Evidencia-se que aumentando os parâmetros: n, N_{h1}, N_{pc} e N_{h2} até limite admissível por potência das redes computacionais (ao redor de 90.000 equações diferenciais ordinárias) que:

- não se observam casos de divergência dos cálculos;
- as características principais: concentração da miscela C_u na saída 3 (Fig. 1), perdas do óleo junto com fases poro ($Q_{ol}^{p,D}$) e sólido ($Q_{ol}^{N,D}$), concentração na fase bulk na zona de drenagem (C_D) não alteram-se significativamente.

Tabela 1 – Influência dos parâmetros da malha (n, N_{h1}, N_{pc}, N_{h2}) com passo de integração h = const.

N	N_{h1}	N_{pc}	N_{h2}	k_{bc}	C_u (%)	$Q_{ol}^{p,D}$ (l/s)	$Q_{ol}^{N,D}$ (l/s)	C_D (%)
3	8	12	8	972	24,87	$0{,}1744.10^{-3}$	$0{,}1968.10^{-3}$	0,4222
5	10	15	10	2025	24,92	$0{,}1618.10^{-3}$	$0{,}1850.10^{-3}$	0,4032
7	15	20	15	4200	24,97	$0{,}1542.10^{-3}$	$0{,}1775.10^{-3}$	0,3922
10	20	30	20	8100	25,00	$0{,}1496.10^{-3}$	$0{,}1730.10^{-3}$	0,3853
15	30	40	30	18000	25,05	$0{,}1463\cdot10^{-3}$	$0{,}1698\cdot10^{-3}$	0,3805
20	40	50	40	31800	25,09	$0{,}1449\cdot10^{-3}$	$0{,}1686\cdot10^{-3}$	0,3789
30	50	60	50	59400	25,12	$0{,}1440\cdot10^{-3}$	$0{,}1678\cdot10^{-3}$	0,3777
35	55	65	55	76125	25,13	$0{,}1438\cdot10^{-3}$	$0{,}1677\cdot10^{-3}$	0,3777
36	55	65	55	78300	25,13	$0{,}1437\cdot10^{-3}$	$0{,}1676\cdot10^{-3}$	0,3776
38	60	70	60	90060	Não calcula por falta de memória computacional			

Obs: a cor cinza indica o regime básico escolhido, k_{bc} – quantidade das equações diferenciais.

Para explicar a estabilidade dos cálculos foi aplicado o critério de Courant (Oran & Boris, 1987) nas direções radiais (S_R) e angulares (S_q) dos trechos horizontais. Os valores S_R e S_q tem que ser maiores que 1 para assegurar a estabilidade dos cálculos. Da Tabela 1 temos para malha do regime básico S_R = 4 e S_q = 3,6 e para malha mais fina S_R = 1,1 e S_q = 1,3, ou seja, os critérios S_R e S_q são satisfeitos para todas as variantes apresentadas na Tabela 1.

Resultados com alteração do passo de integração apresentam-se na Tabela 2 (para dois tipos de tamanhos de malha).

Na tabela 2 observam-se os casos de divergência, onde para a malha mais fina a divergência realizou-se com passo de integração menor do que para a malha mais grossa. A divergência acontecia sempre que S_R e S_q eram menor que 1.

Tabela 2 – Alteração do passo de integração para dois tipos de tamanho de malha.

h	n =20 N_{h1} = 40 N_{h2} =40 N_{pc}=50 k_{bc}=31800		n =15 N_{h1} =30 N_{h2} =30 N_{pc} =40 k_{bc} =18000	
	C_u	P_{ol} (%)	C_u	P_{ol} (%)
1,25	0,2509	0,6185	0,2505	0,6235
2	0,2509	0,6185	0,2505	0,62346
2,5	div	Div	0,2505	0,62341
2,666			0,2500	0,6188
2,777			0,2500	0,61817
3			0,2501	0,619346
3,333			div	Div

Obs.: div – caso de divergência.

4. Aplicativos para Simulações e Visualização dos Processos de Extração.

A apresentação dos resultados por dados numéricos dificulta a análise e compreensão dos fenômenos que ocorrem durante o processo de extração. Portanto foi elaborado mais um aplicativo (V_CROWN) para possibilitar a visualização gráfica dos mesmos em diversos momentos do tempo.

Este aplicativo foi desenvolvido em linguagem DELPHI (Swan, 1999), por esta possuir uma boa biblioteca gráfica. O V_CROWN possui interface gráfica moderna e amigável, além de possibilitar a leitura da simulação numérica, permitindo a visualização da distribuição das concentrações de óleo na fase bulk, poro e sólida pelo campo de extração (trechos, bandejas, trajeto conexivo). Esquema simplificado do aplicativo V_CROWN é apresentado na Fig 4.

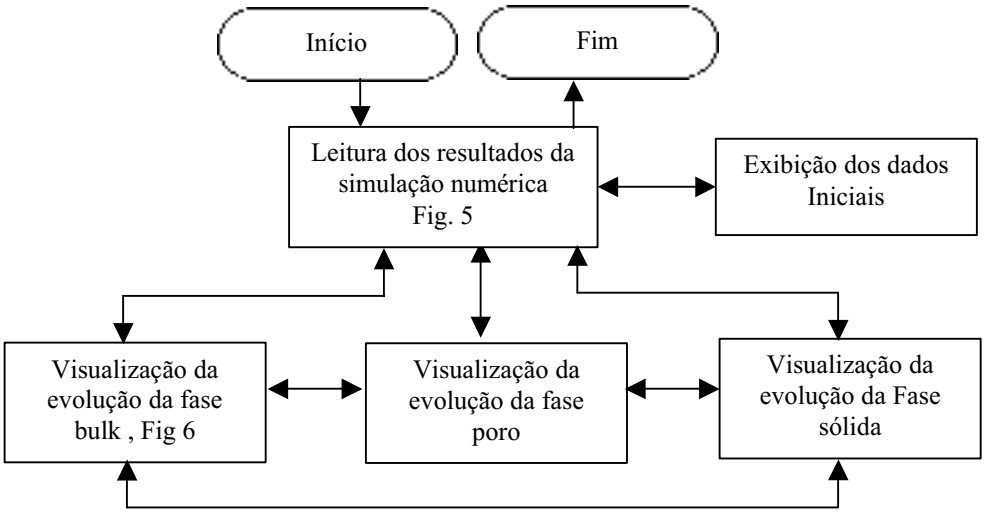


Figura 4. Esquema de ligação dos módulos no aplicativo V_CROWN.

Os Módulos de fase bulk, poro e sólida, possuem o mesmo sistema de visualização das concentrações de óleo no extrator no decorrer do tempo. Foi inserida uma legenda por cores que auxilia a identificação destas concentrações. Tem-se a que representa a variação do tempo, durante o processo de extração. Com essa variação, tem-se uma animação gráfica na figura que representa o extrator. Através dos "botões" pode-se passar de um módulo para outro com facilidade e rapidez.

Através desta animação pode-se realizar uma análise mais detalhada e com maior facilidade dos fenômenos físico-químicos que ocorrerem durante o processo de extração de óleo.

Figura 5. Módulo da visualização dos dados parâmetros geométricos e de regime.

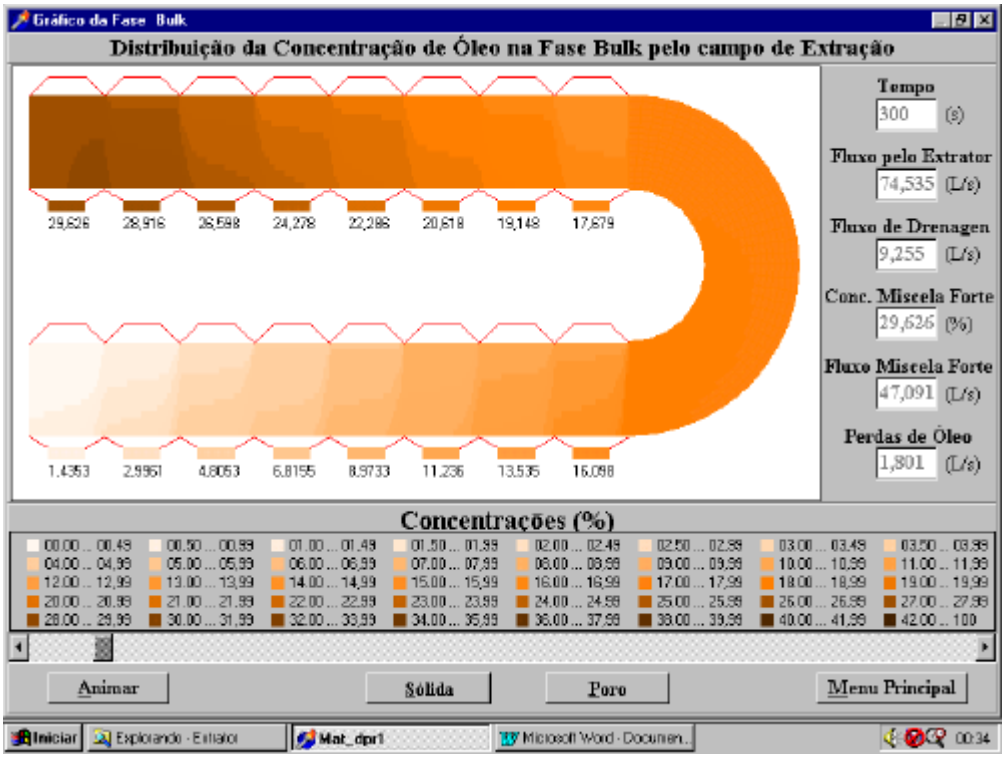


Figura 6. Módulo de visualização da distribuição das concentrações de óleo na fase bulk no campo de extração.

5. Pesquisas Numéricas dos Processos em todo o Extrator

Baseando-se nas características do extrator “Crown Model” (Empresa INTECNIAL - Erechim-RS) e dos dados bibliográficos, foram escolhidos para as simulações numéricas, os seguintes dados (Tab 3):

Tabela 3 – Parâmetros do extrator, propriedades de óleo, solvente e da matéria-prima

$n=10$	$L_r=0,8$ m	$H_l=3,66$ m	$V_b=0,5$ m^3
$a_p=65$ (1/m)	$d_p=0,005$ m	$\epsilon_b=0,42$	$\epsilon_p=0,28$
$D_{AB}=1,3.10^{-9}$ m^2/s	$E_s=1.10^{-4}$ m^2/s	$\rho_{ol}=910$ kg/m^3	$E_d=0,2$
$X_A=16,2$ m	$m_1=6$	$N_{h1}=20$	$N_t=0,19$
$R_1=2,6$ m	$R_2=3,4$ m	$N_{pc}=30$	$k_N=0,01$
$X_C=16,2$ m	$m_2=6$	$N_{h2}=20$	

onde n é o número de camadas horizontais (parâmetro de malha); H_L a comprimento do leito; V_b o volume da bandeja; N_t a concentração mássica do óleo na matéria-prima; X_A o comprimento do trecho superior; N_{h1} número de camadas verticais no trecho superior (parâmetro de malha); N_{pc} o número de camadas verticais no trajeto conexivo; X_C o comprimento do trecho inferior; m_2 o número de seções do trecho inferior; N_{h2} número de camadas verticais no trecho inferior; para parâmetro de malha.

Observa-se que o modelo acoplado do extrator foi identificado no trabalho (Benetti, 2001), onde também foi mostrada a concordância com os dados experimentais. Na Fig. 7 apresenta-se a distribuição das concentrações médias $\overline{C}, \overline{C}^N, \overline{C}^p$ pela altura, ao longo do extrator (trecho superior + trajeto conexivo + trecho inferior), no regime estacionário. Analisando este regime foram constatados vários fenômenos que ocorrem durante o processo de extração, dos quais destacamos:

- o desequilíbrio das concentrações, entre as fases sólida e poro, na zona da entrada da matéria-prima, este efeito explica-se pela pequena diferença inicial entre concentrações $\overline{C}^N$ e $\overline{C}^p$ na etapa de carregamento ($\overline{C}^N \approx 0,25$; $\overline{C}^p \approx 0,20$, quando em equilíbrio a relação entre eles tem que ser $\overline{C}^N = 0,2\overline{C}^p$), por esta causa também observa-se uma grande transferência de óleo da fase sólida para poro, explicada pela forte tendência em estabelecer o equilíbrio;
- na entrada de cada seção há uma diferença brusca na concentração $\overline{C}$ este efeito é conseqüência da propriedade do fluxo contracorrente cruzado, pois a miscela da bandeja da seção posterior cai na superfície da miscela da seção anterior;
- na entrada do trajeto conexivo observa-se um aumento na concentração $\overline{C}$, que ocorre, devido ao fluxo de recirculação do trajeto quando a miscela mais concentrada da saída do trajeto ($\overline{C}_{st}$) transfere-se na sua entrada, aumentando bruscamente a concentração da miscela da fase bulk ;
- no trajeto conexivo, observa-se uma pequena alteração na concentração do óleo das fases, devido a insignificante velocidade de lavagem da matéria-prima (V_f), neste trajeto;
- no limite entre o trajeto conexivo e o trecho inferior, ocorre uma queda forte nas concentrações $\overline{C}$ e $\overline{C}^p$ pois a matéria-prima esta sendo lavada com miscela de menor concentração.

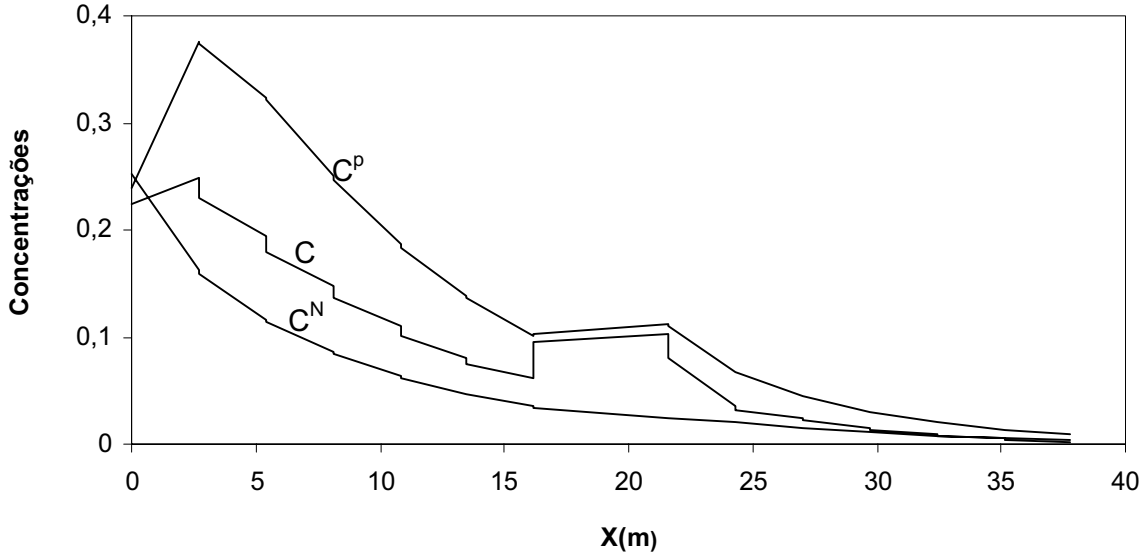


Figura 7. Alteração das concentrações médias $\overline{C}, \overline{C}^N, \overline{C}^p$ ao longo do extrator no regime estacionário. X(m) – coordenada corrente ao longo do extrator.

6. Conclusão

Foi elaborado o esquema físico e o modelo matemático do trajeto conexivo do extrator industrial "CROWN MODEL", que consideram: difusão pelo campo da extração; transferência de massa entre fases sólida e poro, e entre fases poro e bulk; movimento da matéria-prima e miscela; existência do fluxo de recirculação; influência da gravidade na velocidade da miscela pelo trajeto.

Foi elaborado o algoritmo de cálculo que é baseado no método de linhas, e o fragmento do aplicativo correspondente para o trajeto. Juntamente com o modelo de extração (Benetti, 2001), nos trechos horizontais (Fig.1), este fragmento constitui-se um modelo acoplado dos processos no extrator "Crown Model".

As simulações numéricas realizadas para todo o extrator, mostraram a concordância dos resultados com as leis físico-químicas e com as propriedades das características operacionais do extrator deste tipo: existência do desequilíbrio entre fases poro e sólido; os valores máximos das concentrações $\overline{C}^{p}$ e $\overline{C}$ na zona da entrada da matéria-prima; queda abrupta das concentrações $\overline{C}$, entre seções, transferências insignificantes dos óleos no trajeto conexivo.

7. Agradecimentos

O presente trabalho foi realizado com o apoio de CAPES. Os autores agradecem a empresa INTECNIAL pelo fornecimento dos parâmetros geométricos e dados experimentais do extrator CROWN-MODEL.

8. Referências

Benetti, R. C. 2001, "M*odelagem matemática dos processos nos trechos horizontais do extrator "Crown-Model" e sua identificação"* Ijuí, 132 p. Dissertação (Mestrado em Modelagem Matemática) – UNIJUI.

Bockisch, M., 1998, *" Fats and Oils Handbook"*. AOCS Press, Illinois.

Cussler, E. L.,1997, *"Diffusion: Mass transfer in Fluid Systems"*. 2. ed. University of Minnesota: Cambridge, 580 p.

Erickson, David R.,1995, "*Practical handbook of soybean processing and utilization"*. Illinois: AOCS Press, p. 56-116.

Lambert, J. D. ,1978, "*Numerical methods for ordinary differential systens – the initial value problem"*, New York: John Wiley, 283 p..

Oran, E.S., Boris, J.P., 1987, *"Numerical Simulation of Reactive Flow"*, Elsevier, New York, 660 p.

Piva, C. ,2001, "*Modelagem matemática dos processos no trajeto conexivo do extrator "Crown-Model" e simulações numéricas"*. Ijuí, 117 p. Dissertação (Mestrado em Modelagem Matemática) – UNIJUI.

Ramos, S. et all., 1999, "*Dados de medição do Extrator".* INTECNIAL – RS

Swan, T, 1999, "*Delphi 4: Bibilia do Programador"*, Berkeley, São Paulo - SP

Treybal, R. E., 1963, *"Liquid Extraction"* 2. ed. New York, McGraw-Hill. Series in Chemical Engineering.

Yao, P., 1995, *"Borland C^{++} 4.0: programação for Windows"*, Makron Books, São Paulo - SP.

NUMERICAL SIMULATION AND VISUALIZATION OF THE PROCESS OF EXTRACTION IN INDUSTRIAL INSTALLATIONS "CROWN-MODEL".

Marcos Ronaldo Melo Cavalheiro
UNIJUI – Regional University of state of Rio Grande do Sul, Technology of Department, P.O. Box 560, CEP: 98700-000, Ijuí, RS, Brazil
E-mail: mrmc@detec.unijui.tche.br

Claudia Piva
UNIJUI - Regional University of state of Rio Grande do Sul, Program of Graduation Mathematical Modeling, P.O. Box 560, CEP 98700-000 Ijuí, RS, Brazil

Abstract. *The present refers to the mathematical modeling of processes in the connective passage of the extractor know as "Crown Model" like an important fragment of the mathematical model of this extractor. The physical outline of the connective r considers the following features: the raw material and miscela is moved along of the connective passage; oil transfer between solid phases and pore, between phase pore and bulk flow; diffusion for the whole extractive field; etc. The mathematical model is composed by three partial differential equation, non-linear, bi dimensional, transient, with boundary conditions in the laterals and between domains, considering miscelas´s recirculation. The algorithm base on the method of lines. A code was developed to simulate the processes in the extractor (C++ language). Other code was developed simulation sly to visualize these processes (Delphi language). With theses tools the simulation were accomplished for the whole extractor, which showed good agreement between the results and the physical chemistries laws and, with the properties of the operational characteristics of the extractor of this type.*

Keywords. *Extraction of oil, Installations Industrial, Mathematical Modeling, Graphic Visualization*

XVI CONGRESSO BRASILEIRO DE ENGENHARIA MECÂNICA
16th BRAZILIAN CONGRESS OF MECHANICAL ENGINEERING

Formato e Estabilidade do Balão Têxtil no Processo de Torção "Dois por Um" Descendente

J.P.P. Marcicano
Departamento de Engenharia Mecatrônica e Sistemas Mecânicos da Escola Politécnica da Universidade de São Paulo, Av. Prof. Mello de Moraes 2231, 05508-900, Sao Paulo, SP, Brasil.

C.C.C. Tu
Departamento de Engenharia Mecatrônica e Sistemas Mecânicos da Escola Politécnica da Universidade de São Paulo, Av. Prof. Mello de Moraes 2231, 05508-900, Sao Paulo, SP, Brasil.

Resumo. *Neste trabalho, apresentam-se modelos matemáticos e programas de computador para calcular a forma e a estabilidade de balões têxteis com um loop restritos por limitadores cilíndricos. O formato do balão foi calculado usando o "shooting method" para resolver o problema a valores de contorno (TPBVP). O princípio da mínima energia potencial (Lagrange) foi usado para analisar a estabilidade dos balões. Um novo método foi desenvolvido para determinar o coeficiente de atrito entre o fio e o limitador cilíndrico.*
Os resultados dos modelos foram comparados com valores experimentais. Os experimentos foram conduzidos em um dispositivo simples, com um fuso construído a partir de um motor de furadeira manual. O desvio médio entre os resultados, simulação e experimental, foi de 0,8% para o formato e 2,0 % para a estabilidade.
Foi verificado que a aceleração de Coriolis pode ser desprezada para a maioria das situações práticas.

Palavras chave: *balão têxtil, estabilidade, retorcedeira dois por um.*

1. Introdução

Para fazer tecido são utilizados fios longos e contínuos. Porém os materiais originalmente disponíveis para tecelagem, como o algodão e lã, têm fibras curtas. Há muito tempo o homem descobriu que agregando as fibras curtas em paralelo e depois torcendo-as pode-se produzir um fio longo, contínuo e resistente. Este processo, entretanto, requer um número muito grande de torções. De Barr (Barr, 1965) calculou a quantia de 3,2 x 10^{12} torções para produzir uma única camisa. Para inserir estas torções foram desenvolvidas muitas máquinas de alta rotação no decorrer da revolução industrial.

Existem dois métodos conhecidos para torcer os fios: o concêntrico e o rotativo .Em ambos os métodos, a cada volta do dispositivo, uma volta é introduzida no fio. No concêntrico, a aceleração centrípeta do fio é menor do que no rotativo. Esta aceleração é a principal causa da tensão no fio durante o processamento. Apesar de trabalhar com tensões maiores, o método rotativo é empregado na maioria das máquinas modernas de alta velocidade por permitir formas contrutivas mais adequadas. Neste método, devido a rotação, o fio assume visualmente um formato característico parecido com um balão. A estabilidade deste balão depende do ajuste da rotação e tensão no fio. Matendo-se a tensão e aumentando-se a rotação, o balão aumenta de tamanho e pode colapsar, o mesmo ocorre mantendo-se a rotação e diminuindo-se a tensão. Para conter o diâmetro do balão, costuma-se utilizar limitadores na forma de anéis ou cilindros.

O processo de torção "dois por um", ou de dupla torção, baseado na soma dos princípios rotativo e concêntrico foi patenteado em 1755 na Inglaterra por C.F. Kirkman.

O balão livre tem sido estudado desde o século 18, O problema de corda submetida a rotação no vácuo foi inicialmente estudado por D.Bernoulli (1700-1782) e por L. Euler (1707-1783). Ludicke em 1881 publicou um dos primeiros trabalhos teóricos aplicados as máquinas de produção de fios têxteis. Desde então muitos outros trabalhos teóricos e experimentais foram realizados.

Soluções aproximadas para as equações do balão foram dadas inicialmente por Grishin em 1965, Hornegger e Fehr, Capello (Capello,1958) , Bracewell e Greenhalgh. Estas soluções de forma parcial satisfaziam as necessidades de projeto.

As soluções numéricas, surgiram com a disponibilidade de computadores, obtidas por: Mack (Mack,1953), Hannah (Hannah,1955), Crank (Crank,1953) e Padfield (Padfield,1958).

Em 1965, o livro de Barr e Catling (Barr,1965), revisou grande parte dos trabalhos sobre o balão livre, feitos por pesquisadores do ocidente. Este livro revisa também os estudos sobre estabilidade do balão, porém o número de referências sobre estabilidade é muito pequeno com relação ao estudo do formato. As únicas referências dadas sobre estabilidade são de Grishin e Greenwood.

A estabilidade de balões livres também foi estudada na Rússia por Yakubouskii (Yakubouskii,1961) que utilizou o método de Bifurcação e considerou os efeitos da força de Coriolis; Zak (Zak,1967) e (Zak,1968) utilizou o método de Pertubação e concluiu que a estabilidade é parcialmente controlada pelo arrasto aerodinâmico. Porém, Turteltaub e Bejar (Turteltaub,1976), no ocidente, com o método de Pertubação concluiram que o efeito do arrasto aerodinâmico é desprezível em relação à estabilidade.

Kothari e Leaf. (Kothari,1979) apresentaram um modelos numéricos para o processo “dois por um” sem limitadores. Lisini (Lisini,1981) determinou a geometria de balões livres e restritos utilizando o método de elementos finitos. Tu (Tu,1982) determinou o formato do balão restrito utilizando o “shooting method” e analisou experimentalmente a estabilidade. Batra (Batra,1989) resolveu o formato do balão no processo de ring-spinning. Lisini (Lisini,1992) obteve soluções dinâmicas de balões livres no processo de ring-spinning utilizando um método numérico do tipo diferenças finitas para resolver o sistema de equações diferenciais em função da posição e tempo.

Fraser (Fraser,1993a,1993c) analisou o processo dois por um com cilindro limitador através do método de bifurcação, demonstrou a natureza não linear do problema e a multiplicidade de soluções . O mesmo autor em 1993 (Fraser,1993b) através do método de Bifurcação analisou a estabilidade no espaço de soluções para o processo de “ring spinning” livre, e em 1996, Fraser, Clark, Ghosh e Zeng (Fraser,1996) o processo de ring-spinning restrito.

A partir do experimento realizado por Hall (Clark,1998), onde mediu-se a tração em função do comprimento do fio para um balão livre em um dispositivo que controla o comprimento do fio, foram desenvolvidos três trabalhos para análise da estabilidade do balão. O primeiro foi apresentado por Zhu, Hall e Rahn (Zhu,1997) , a estabilidade das soluções foi verificada pela análise dos sinais dos autovalores da equação de movimento obtida pelo método de discretização e linearização de Galerkin. O segundo, apresentado por Clark, Fraser, Sharma e Rahn em 1998 (Clark,1998) a estabilidade foi analisada por um modelo baseado no método de Bifurcação para localizar as regiões instáveis e o método de pertubação para verificar a estabilidade das soluções através da análise do sinal dos autovalores. E no terceiro, Zhu e Rahn (Zhu,2000), utilizam o Teorema de Bifurcação de Hopf para analisar a estabilidade e amplitude de vibração do fio. Na aplicação dos resultados desses trabalhos deve-se considerar o fato de que em balões têxteis o comprimento do fio não é fixo, normalmente tem-se o controle da tensão.

A revisão da literatura mostra que as modelagens atualmente disponíveis, mesmo modelando numericamente e/ou experimentalmente, não resolveram o problema de estabilidade para o balão com limitador cilíndrico.

Este artigo apresenta a modelagem teórica e experimental do balão têxtil restrito, são calculadas as soluções em regime permanente e a estabilidade dessas soluções. A equação de movimento do fio é obtida pela equação de Newton e resolvida numericamente como um problema a valores de contorno, as equaçoes diferenciais são integradas com o método de Runge-Kutta de quarta ordem e o problema a valores de contorno é resolvido com o método de Newton-Raphson. A estabilidade das soluções é analisada pelo critério de estabilidade de Lagrange ou da mínima energia potencial total. Os resultados dos modelos são verificados experimentalmente.

2. Modelos Matemáticos – Formato e Estabilidade

Para deduzir as equações de movimento do fio para o balão com cilindro limitador considere a Fig.(1) . O fio de masssa por unidade de comprimento m, gira com velocidade angular constante w saindo por um guia com velocidade constante V e sob força trativa T_0 no ponto O, atinge o cilindro limitador de raio Ro em z igual a z_1, mantém-se em contato com o cilindro até z igual a z_2 onde sai e termina em z igual a z_f e raio R_f .

No caso de retorcedeiras dois por um com cilindro limitador tem-se duas regiões distintas, uma livre onde a força de arraste aerodinâmico atua sobre o fio, e a outra onde o fio mantém-se em contato com o limitador onde considera-se a reação normal N a parede e a força de atrito μN na direção tangencial. A força peso mg é considerada nas duas regiões.

Na região onde o fio toca o limitador, o raio permanece com valor constante então dr/ds = 0 e $d^2r/ds^2=0$.

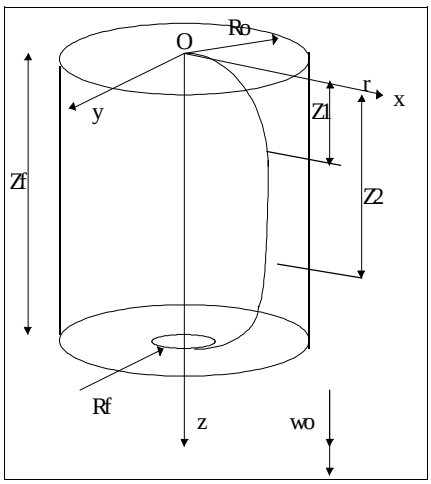


Figura 1.Balão com limitador cilíndrico

As equações de movimento do fio no trecho onde o fio toca o limitador são dadas por:

$$-r(V\frac{d\Theta}{ds}-w)^2=\frac{-Tr}{m}(\frac{d\Theta}{ds})^2-\frac{N}{m} \qquad (1)$$

$$rV^2\frac{d^2\Theta}{ds^2}=\frac{1}{m}\left(\frac{d}{ds}\left(T\,r\frac{d\Theta}{ds}\right)+\mu N\right) \tag{2}$$

$$V^2\frac{d^2z}{ds^2}=\frac{1}{m}\left(\frac{d}{ds}\left(T\frac{dz}{ds}\right)+mg\right) \tag{3}$$

Onde T é a força trativa e a s a coordenada medida ao longo do fio.
Considerando-se que o fio é inextensível vale a relação:

$$\frac{dr}{ds}^2+\left(r\frac{d\Theta}{ds}\right)^2+\frac{dz}{ds}^2=1 \tag{4}$$

Adimensionalizando-se as equações de movimento com o seguinte conjunto de adimensionais:

r' = r / R_o
z'= z / R_o
V' = V / w R_o
T' = T / m w^2 R_o^2
s' = s / R_o
g' = g / R_o w^2
p_o= 8 R_o C_n ρ d / m

Doravante os adimensionais serão expressos sem o apóstofo.
Substituindo-se dθ /ds por L e dr/ds por K, pode-se obter as expressões das derivadas de s,T,θ e L com relação a z, para a região do balão onde o fio toca o limitador.

$$\frac{ds}{dz}=\left(1-K^2-(rL)^2\right)^{-1/2} \tag{5}$$

$$\frac{d\Theta}{dz}=L\frac{ds}{dz} \tag{6}$$

$$(V^2-T)\left[1+\left(r\frac{ds}{dz}L\right)^2\right]\frac{dL}{dz}=\frac{ds}{dz}\left[-g\frac{ds}{dz}L+\mu(VL-1)^2-\mu T L^2\right] \tag{7}$$

$$\frac{dT}{dz}=\frac{ds}{dz}^2\left[-g-(V^2-T)L\,r^2\frac{dL}{dz}\right] \tag{8}$$

Na região livre, as equações de movimento do fio são:

$$V^2\frac{d^2r}{ds^2}-r\left(V\frac{d\Theta}{ds}-w\right)^2=\frac{1}{m}\left[\frac{d}{ds}\left(T\frac{dr}{ds}\right)-T\,r\frac{d\Theta}{ds}^2\right] \tag{9}$$

$$rV^2\frac{d^2\Theta}{ds^2}+2V\frac{dr}{ds}\left(V\frac{d\Theta}{ds}-w\right)=\frac{1}{m}\left[\frac{d}{ds}\left(T\,r\frac{d\Theta}{ds}\right)+T\frac{dr}{ds}\frac{d\Theta}{ds}+\frac{Cn\,\rho\,d}{2}w^2r^2\right] \tag{10}$$

$$V^2\frac{d^2z}{ds^2}=\frac{1}{m}\left[\frac{d}{ds}\left(T\frac{dz}{ds}\right)+mg\right] \tag{11}$$

Onde C_n é o coeficiente de arraste aerodinâmico, ρ a massa específica do ar e d o diâmetro aparente do fio.
Considerando-se a condição de inextensibilidade do fio e adimensionalizando-se obtém-se as derivadas de s,r,θ,T,K e L com relação a z.

$$\frac{ds}{dz}=\left(1-K^2-(rL)^2\right)^{-1/2} \tag{12}$$

$$\frac{dr}{dz}=K\frac{ds}{dz} \tag{13}$$

$$\frac{d\Theta}{dz} = L\frac{ds}{dz} \tag{14}$$

$$\frac{dT}{dz} = -r\frac{dr}{dz} - \frac{p_o r^3}{16}\frac{d\Theta}{dz} - g \tag{15}$$

$$(V^2 - T)\frac{dK}{dz} = \frac{ds}{dz}[r(V^2 - T)L^2 - 2rVL + r] + \frac{dT}{dz}K \tag{16}$$

$$r(V^2 - T)\frac{dL}{dz} = \frac{ds}{dz}[2VK + \frac{p_o r^2}{16}] + L[r\frac{dT}{dz} - 2\frac{ds}{dz}K(V^2 - T)] \tag{17}$$

As equações de movimento do fio formam um sistema de equações diferenciais não lineares, neste caso ao que se saiba, não é possível calcular uma solução analítica. Por método numérico é possível integrar as equações das derivadas das variáveis e com as condições de contorno conhecidas nos extremos calcular a posição e a força trativa de cada ponto do fio. Neste trabalho as equações diferenciais foram integradas pelo método de Runge-Kutta de quarta ordem e a solução do problema a valores de contorno foi feita pelo método de Newton-Raphson , onde os valores iniciais foram ajustados para satisfazer as condições finais.

As condições de contorno foram estabelecidas para os quatro pontos da Fig. 1. Nos pontos O e Z_f a condição é que o fio deve passar por esses pontos, em O, z=0 e r=0, e em z_f, $z=z_f$ e $r=R_f$. Em z_1 e z_2 foi estabelecido que as coordenadas, a tração e as derivadas das coordenadas em z devem ser contínuas.

O teorema de estabilidade de Lagrange estabelece que se em alguma posição de um sistema conservativo a energia potencial tem um mínimo absoluto, então esta posição é a posição de equilíbrio estável do sistema.

De acordo com Gantmacher (Gantmacher,1970) , o teorema permanece válido para um sistema não conservativo que é obtido de um sistema conservativo pela adição de forças giroscópicas e dissipativas. Entretanto nota-se que a posição de equilíbrio é preservada se forças giroscópicas ou dissipativas são adicionadas ao sistema. E o teorema também é válido se a posição de equilíbrio não for um mínimo absoluto.

No caso de estruturas , Tauchert (Tauchert,1974) define a estabilidade para estruturas em equilíbrio submetidas a forças estáticas, uma posição é dita de de equilíbrio estável se quando submetida a um deslocamento virtual compátivel com as restrições geométricas, retorna a posição inicial quando o deslocamento é retirado, caso não retorne a posição é considerada de equilíbrio instável.

A estabilidade pode também ser definida em termos da energia potencial total Π da estrutura, a energia potencial total é composta por duas parcelas, o potencial das forças externas V_e e a energia interna U. Se Π aumenta durante um deslocamento virtual então a posição de equilíbrio é estável, caso Π diminua ou permaneça constante a posição é instável.

A equivalência entre as definições pode ser demonstrada considerando-se um sistema em equilíbrio estável. Pela segunda definição a variação da energia interna U é maior que a variação do trabalho $-V_e$ feito pelas forças externas durante o deslocamento.Então quando o deslocamento virtual é retirado, existe uma ação de restauração, e o sistema retorna para a configuração inicial.

No caso do fio em equilíbrio sob uma força trativa T_i aplicada na extremidades e comprimento s_i , a energia potencial total com relação a posição de equilíbrio para um deslocamento que mude o comprimento para $S=s_i+\delta$ mantendo-se T_i constante, pode ser calculado por:

$$\Pi = T_i(S - s_i) + (E_p(S) - E_p(s_i)) \tag{18}$$

No caso dos balões com atrito e arraste aerodinâmico, a variação da energia potencial pode ser calculada pelo trabalho externo dado pela integral de $-T_0\,dS$, isto é válido se as forças atuantes sobre o fio forem potenciais. Na dedução das equações de movimento do fio , as derivadas temporais foram substituidas por derivadas parciais em s e em t, as derivadas em t foram anuladas para obter a solução em regime permanente, desse modo as forças atuantes no fio ficaram independentes das velocidades dr/dt e $r\,d\theta/dt$, segundo Gantmacher (Gantmacher,1970), se as forças generalizadas Q não dependerem das velocidades generalizadas e existir uma função Π tal que:

$$Q_i = \frac{-\partial \Pi}{\partial q_i} \tag{19}$$

então as forças Qi são potenciais. Desse modo, a variação da energia potencial pode ser calculada pelo trabalho externo. A expressão utilizada para calcular a variação da energia potencial total de um balão com comprimento s_i e força trativa na extremidade de T_i torna-se:

$$\Pi = T_i(S - s_i) - \int_{s_i}^{S} T_0(S)\,dS \tag{20}$$

para o balão ser estável a energia potencial total deve ser mínima então $d\Pi/dS=0$ e $d^2\Pi/dS^2 > 0$ para $s=s_i$ derivando-se a energia potencial total em S mantendo-se T_i fixo obtém-se:

$$\frac{d\Pi}{dS} = T_i - T_0(S) \tag{21}$$

em $S=s_i$,$T_0(s_i)=T_i$ então $d\pi/dS =0$, satisfazendo a primeira condição e,

$$\frac{d^2\Pi}{dS^2} = \frac{-dT_0(S)}{dS} \tag{22}$$

Dessa forma, quando dT_0/dS é menor que zero as soluções são estáveis, a transição ocorre quando T_0 é mínimo e as soluções são instáveis quando dT_0/dS é maior que zero.

A equação 22 pode ser utilizada para verificar a estabilidade do fio quando submetido a uma força trativa constante na extremidade e as forças atuantes no sistema forem potenciais.

3. Resultados das Simulações Numérica e Experimental - Formato e Estabilidade

O formato de um balão restrito por limitador cilíndrico foi medido por método fotográfico. Os valores adimensionalizados utilizados das variáveis no ensaio para determinação do formato são: tração inicial T_0=1,24, aceleração da gravidade g=0,460 10^{-3} ,velocidade de alimentação do fio V=3,14 10^{-3} ,altura final Z_f=2,50, raio final R_f=0,636, arraste aerodinâmico p_0=5,9 e coeficiente de atrito entre o fio e o limitador μ=0,242.

O formato predito pelo modelo foi obtido pela solução do problema a valores de contorno, através da integração das equações de movimento do fio utilizando-se valores iniciais e condições finais compatíveis com o balão simulado no ensaio.

Os resultados experimentais e do modelo teórico são apresentados no gráfico da Fig 2 , onde apresenta-se em função do ângulo θ, os valores do raio projetado medidos experimentalmente R_{ne} e os valores preditos pelo modelo R_{nc}

Comparando-se os resultados pode-se dizer que o modelo foi capaz de calcular com boa precisão o formato do balão. O erro porcentual entre o valor médio medido e predito pode ser explicado pela dispersão das medidas experimentais representadas pela incerteza expandida.

O valor do coeficiente de atrito utilizado no modelo teórico foi obtido através de ajuste do ângulo final do balão. O formato do balão calculado pelo modelo teórico depende do valor do coeficiente de atrito, como obteve-se bons resultados na determinação do formato, pode-se dizer que o método experimental utilizado para medir o formato do balão através da fotografia em plano superior demonstrou ser adequado para determinação do coeficiente de atrito entre o fio e parede do limitador. Este método viabiliza a determinação do coeficiente de atrito em retorcedeiras ou fiadeiras industriais onde o limitador não é transparente.

Na determinação da estabilidade das soluções, no ensaio de estabilidade mediu-se a mínima força trativa que torna o balão instável. Neste ensaio a mínima força trativa foi medida para diferentes alturas de balão, os valores das outras variáveis são os mesmos do ensaio para determinação do formato.

A mínima força trativa predita foi calculada através da minimização de T_0 com relação ao comprimento total do balão S, com o método de minimização quadrática, T_0 e S foram calculados através da integração das equações de movimento do fio e solução do problema a valores de contorno.

A comparação entre os resultados experimentais e preditos são apresentados na Fig. 3.

O coeficiente de correlação entre T_{0exp} e Z_f vale 0,998 calculado com os pontos exibidos na Fig. 3, o que demonstra que T_{0exp} é influenciado por Z_f , e a relação é próxima a linear .

A relação entre a tração inicial e comprimento total do fio calculada para balões com altura de 2,96 pode ser observada na Fig. 4. Os gráficos de energia potencial total para as soluções com comprimento 4,136 e 4,186 são apresentados em Fig. 5 e Fig. 6 respectivamente. Verifica-se que o balão com comprimento 4,136 que é menor que o comprimento correspondente a mínima força trativa, a energia potencial total apresenta um mínimo, sendo por isso estável, enquanto que o balão com comprimento de 4,186 apresenta um máximo o que corresponde a uma posição de equilíbrio instável.

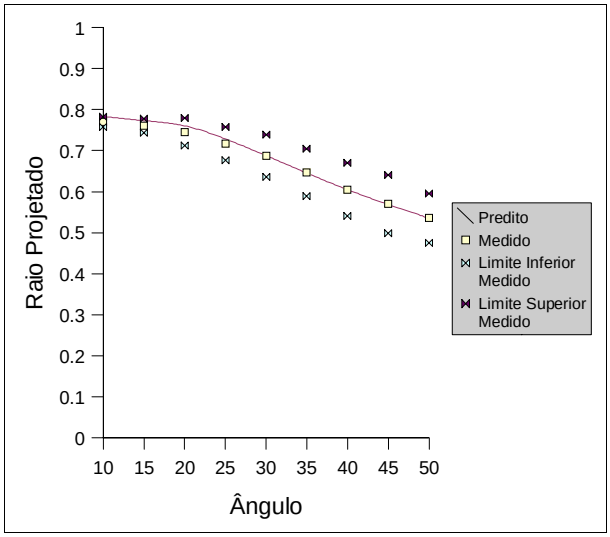


Figura 2. Raio projetado em função do ângulo, valores medidos e preditos

Os resultados obtidos com o modelo de Estabilidade de Lagrange utilizado para analisar a estabilidade através da variação da energia potencial total das soluções de balões com atrito e arrastre aerodinâmico concordam com os resultados experimentais. Nos casos analizados neste trabalho, verificou-se a dependência da estabilidade com a derivada da tração inicial com relação ao comprimento total. As soluções em equilíbrio estável quando dT_0/dS é menor que zero , instável quando dT_0/dS é maior que zero.e a transição ocorrendo na força trativa mínima onde dT_0/dS é nulo.

A influência da força de Coriolis foi verificada calculando-se a mínima força trativa com velocidade de alimentação do fio igual zero, outras componentes de força foram afetadas, mas como elas dependem de V^2 e a ordem de grandeza de V é 10^{-3} então a influência dessas componentes de força tem ordem de grandeza 10^{-6} podendo ser desprezadas. Como resultado verifica-se que a força de Coriolis exerce uma influência da ordem de 10^{-3} , no valor da mínima força trativa no fio e de 10^{-2} no comprimento do balão. Este resultado concorda com o obtido por outros pesquisadores, Fraser (Fraser,1993c) considerou dentre as forças inerciais somente a força centrípeta, Tu (Tu,1982) também adotou o mesmo procedimento, mas ponderou que com o desenvolvimento das máquinas têxteis para aumentar a produtividade a velocidade de alimentação do fio deverá aumentar.

A constatação experimental da existência de soluções com dois loops em retorcimento de fio com cilindro limitador, discorda da conclusão do artigo de Fraser (Fraser,1993a), que o cilindro limitador elimina a multiplicidade de soluções para uma dada configuração geométrica. A eliminação da multiplicidade de soluções ocorre de acordo com Tu (Tu,1982) para balões no vácuo restritos por cilindro limitador sem atrito.

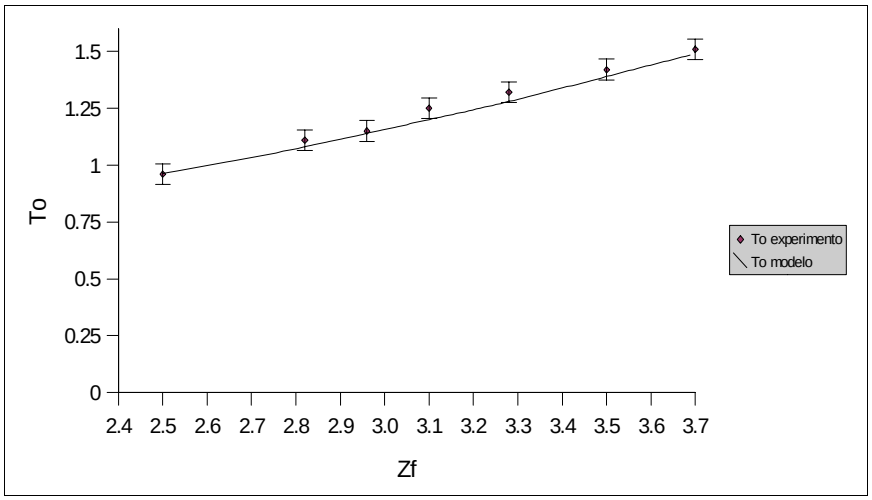


Figura 3. Mínima força trativa medida e calculada em função da altura do balão

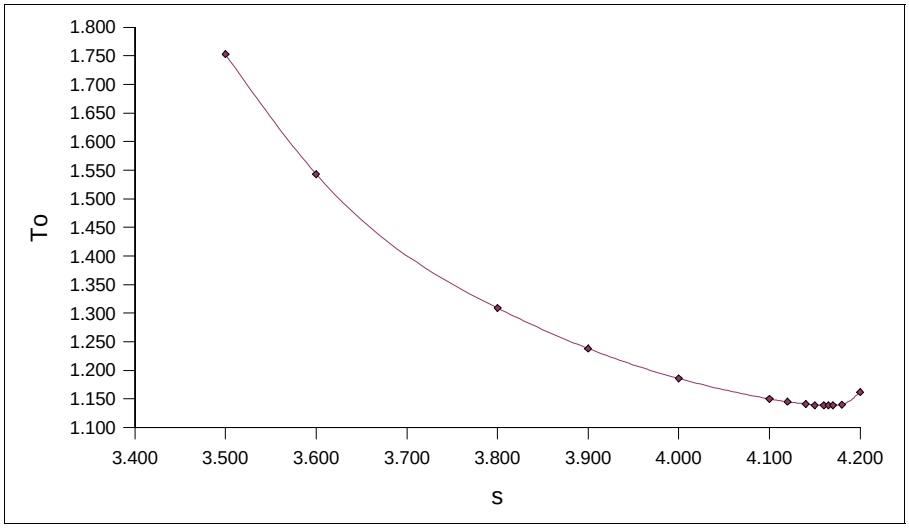


Figura 4. Tração inicial em função do comprimento para os balões de altura Z_f=2,96

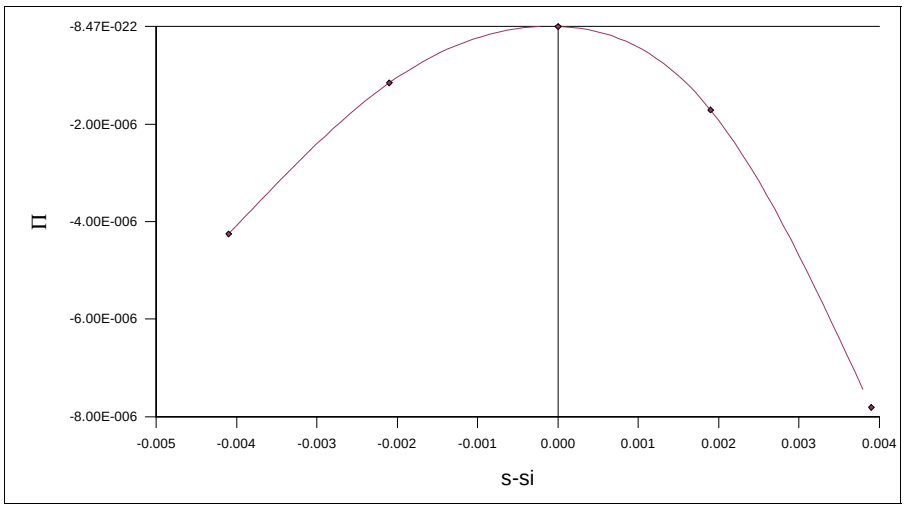


Figura 5. Variação da Energia Potencial Total para o Balão de comprimento 4,186.

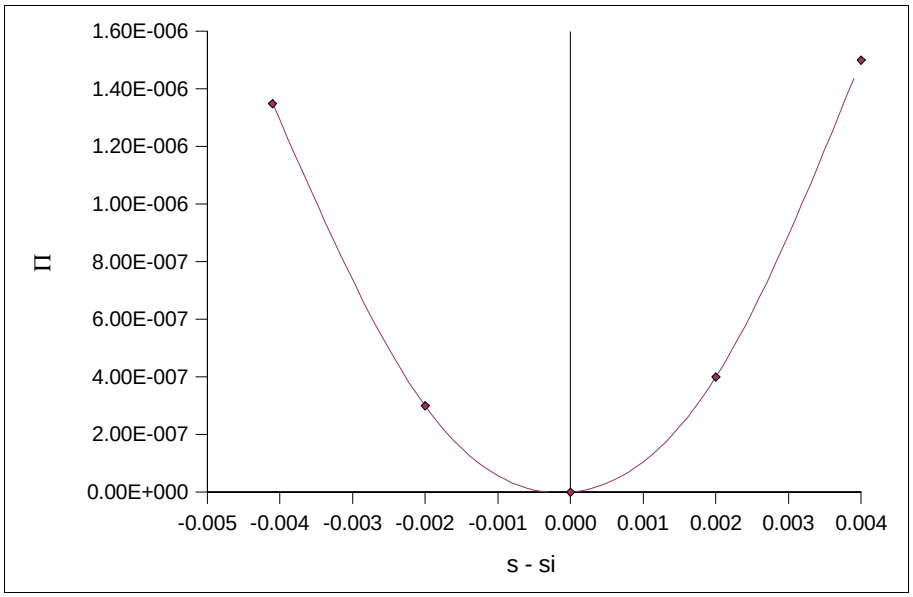


Figura 6. Variação da Energia Potencial Total para o Balão de comprimento 4,136

4.Conclusões

O conjunto de equações diferenciais apresentado para calcular o formato e a força trativa do fio em regime permanente considera todos os termos de aceleração inercial pode simular com razoável precisão o processo têxtil dois por um descendente.

Quanto a estabilidade demonstrou-se que o critério de Estabilidade de Lagrange pode ser utilizado para verificar a estabilidade de soluções e determinar a força trativa que torna o balão instável.

As influências da força de Coriolis e de outros termos dependentes da velocidade de alimentação do fio na estabilidade de balões com baixa velocidade de alimentação do fio são pequenas.

O método experimental para medir o formato do balão, por fotografia em plano superior mostrou-se adequado para determinar o coeficiente de atrito entre o fio e a parede do limitador, podendo ser utilizado na prática onde os limitadores são feitos de metal.

O conjunto de programas desenvolvido pode ser utilizado para calcular o formato do balão, coeficiente de atrito e tração mínima de modo viável e com precisão para fins práticos em microcomputadores.

5.Referências

Barr,A.E. e Catling,H.,1965, .The Principles and Theory of Ring Spinning. The Textile Institute and Butterworth & Co Publishers Ltd, volume V.

Batra,S.K.; Ghosh,T.K.; Zeidman,M.I.,1989, An Integrated Approach to Dynamic Analysis of the Ring Spinning Process Part II: with Air Drag. Text.Res.Inst, 59,pp. 416-424.

Capello,A.,1958,. The Spinning Balloon, Part II, J.Text Inst., v. 49, pp T579.

Clark,J.D. Et alli.,1998, The Dynamic Response of a ballooning Yarn: Theory and Experiment. Proc.R.Soc.Lond, A 454, pp. 2767-2789.

Crank,J. A ,1953,Theoretical Investigation of Cap and Ring Spinning Systems, Text. Research, v23, pp 266-276.

Fraser,W.B,1993a,. On the dynamics of the two-for one twister. Proc.Royal Soc. London, A447, pp. 409-425.

Fraser,W.B.,1993c, Air Drag and Friction in the Two-for-One Twister: Results from the Theory, J.Text.Inst, 84, no 3.

Fraser,W.B.; Clark,J.D. ; Ghosh, T.K.; Zeng,Q.,1996, The Effect of a Control Ring on the Stability of the Ring-Spinning Balloon, Proc.R.Soc.Lond, A 452, pp. 47-62.

Fraser,W.B.,1993b, On the theory of ring spinning. Phil.Trans. Proc. Royal Soc. London, A, pp. 439-468.

Gantmacher, F.,1970, Lectures in Analytical Mechanics.Mir Publishers.Moscou.pp 169-173.

Hannah,M.,1955, Applications of a theory of Spinning Balloon, Part I , J.Text.Inst.,v46, pp. T1-T16.

Kothari,V.K. e Leaf,G.A.V.,1979, The Unwinding of Yarns from Packages Part IV: TWO-FOR-ONE Twisting: The Storage Disc and Balloon. J.Text.Inst, No 5, pp. 184-192.

Lisini,G.G. et alli.,1981, Determination of Balloon Surface in Textile Machines - A Finite Segment Approach. ASME Journal of Engineering for Industry, v 103, pp.424-30.

Lisini,G.G. et alli.,1992, A Comparison of Stationary and Non-Stationary Mathematical Models for the Ring-Spinning Process; J.Text.Inst,83, no 4.

Mack,C.,1953, Theoretical Study of Ring and Cap Spinning Balloon Curves (with and without air drag). J.Text.Inst., vol 44, no 11, pp T483.

Padfield, D.G.,1958, The motion and tension of an unwinding thread I. Proc.Royal Soc. London , A245, p.382-407.

Tauchert, T.R.,1974, Energy Principles in Structural Mechanics . McGraw-Hill, p.199-210.

Tu,C.C.,1982, Shape and Stability of the cilindrically constrained Yarn Balloon. Texas. Tese, University of Texas at Austin.

Turteltaub,M.J.; Bejar,M.A.,1976, Balloon Collapse in Ring-Spinning. Transactions of the ASME, pp 658-663.

Yakubouskii,Y.V.,1961, Contribution to the Multiple Balloon Problem. Tech of Text Industry U.S.S.R., no 3, pp 42-54.

Zak,M.A.,1967, The Criterion of the Configuration Stability of the Ballooning Yarn, Tech. of. Textile Ind. U.S.S.R., no 5, p 148-153.

Zak,M.A.,1968, On the Loss of Stability of the Shape of an Ideally Flexible String. Journal of Applied Math. And Mech., pp 1110.

Zhu,F. ; HALL,K. ; RAHN,C.D.,1997, Steady State Response and Stability of Ballooning Strings in Air. Int.J. Non-Linear Mechanics, vol 33, no 1, pp 33-46.

Zhu,F. , RAHN,C.D.,2000, Limit Cycle Prediction for Ballooning Strings. Int. J. of Non-Linear Mechanics, 35, pp. 373-383.

Shape and Stability of Balloon on the two-for-one twister

J.P.P. Marcicano

Departamento de Engenharia Mecatrônica e Sistemas Mecânicos da Escola Politécnica da Universidade de São Paulo, Av. Prof. Mello de Moraes 2231, 05508-900, Sao Paulo, SP, Brasil.

C.C.C. Tu
Departamento de Engenharia Mecatrônica e Sistemas Mecânicos da Escola Politécnica da Universidade de São Paulo, Av. Prof. Mello de Moraes 2231, 05508-900, Sao Paulo, SP, Brasil.

Abstract.*Theoretical derivations and computer programs were written that predict both the shape and the stability point of single loop, cylindrically constrained, yarn balloons. The shape was calculated using the shooting method for solving TPBVP. The principle of minimum potential energy (Lagrange) was used to predict the stability of the balloons. A new method was also developed for measuring the crosswise sliding friction coefficient between the yarn and the cylindrical balloon limiter.*
The theoretical calculations were compared with experimentally measured values. These measurements were performed on a simple, yet adequate, ballooning yarn apparatus whose spindle was a converted hand grinder. The average standard deviation of the balloon shape measurements was 0,8 % and 2,0% for stability. The predicted shape and stability points agreed with the measured values to within these errors.
It was proven that the Coriolis acceleration can be neglected for most practical situations.

Keywords: *textile balloon, stability, two-for-one spindle.*

XVI CONGRESSO BRASILEIRO DE ENGENHARIA MECÂNICA
16th BRAZILIAN CONGRESS OF MECHANICAL ENGINEERING

SOME EXPERIMENTAL FINDINGS ON THE PLASTIC BUCKLING OF SHELLS UNDER AXIAL IMPACT LOADS

Marcílio Alves
Group of Solid Mechanics and Structural Impact
Department of Mechatronics and Mechanical Systems Engineering
University of São Paulo
maralves@usp.br

Giancarlo Barbosa Micheli
Group of Solid Mechanics and Structural Impact
Department of Mechatronics and Mechanical Systems Engineering
University of São Paulo
giancarlo.micheli@poli.usp.br

Abstract. Many devices used to absorb impact energy rely on the progressive folding of tube-like structures. However, for sufficiently slender shells, a less efficient energy absorption mechanism might occur, i.e. global buckling. In this article, the transition length marking the threshold between progressive and global buckling is explored experimentally by a series of static and dynamic tests. In particular, the dynamic testes were performed in a specially designed rig called drop hammer, capable to yield high impact energies. By fixing the impact mass and some tube lengths, the impact velocity was changed until transition took place. Also, the influence of a wide range of temperatures on the buckling transition was obtained. Lastly, it was explored the validity of the laws of similarity in order to provide some insights on whether scaled models can be used in the design and analysis of structures under impact loads. The results indicate that the critical length for buckling transition is strongly dependent of the impact speed and temperature. The tests also showed, though not conclusively, that the phenomenon seems not to obey the laws of scaling.

Key words:. buckling, temperature, impact, shells, scaling.

1. Introduction

In order to mitigate the damage in passengers, cargo and structures caused by collision, drop of objects and other impact loads, it is important to employ energy absorption devices. They are conceived in many different ways, using materials as diverse as wood, oil, metals and composites (Jones, 1997). Many efficient devices for impact energy absorption are based on the progressive crushing of tube-like structures. Indeed, the axial collapse of tubes is one of the basic mechanisms used for the absorption of impact energy (Reid, 1993). The current studies range from analytic (Alexander, 1960 and Jones, 1989), experimental, Andrews et al. (1983), and/or numeric, (Karagiozova and Jones, 2000 and Karagiozova et al., 2000) approaches, aimed at tackling several aspects in the collapse of tubes, such as maximum load, medium values of deceleration, deformation patterns, energy dissipation efficiency, etc.. Clearly, industries like the automotive one is much concerned with the structural response of their vehicles in impact situations (McGregor et al., 1993 and Jones, 1989), and by using shell-like structures a vehicle can be designed to support static, dynamic and impact loads.

The ideal mechanism of impact energy absorption using tubes consists in its folding along the length. This phenomenon, called progressive buckling (Jones, 1997), can take place either statically or in the range of speeds where automobile accidents occur. Due to the influence of the velocity on the material properties, e.g. flow stress, and to the mass of the tubes which need to be displaced during the impact, it is established that inertia effects are most important in determining the device deformation pattern and, hence, its efficiency (Karagiozova et al., 2000 and Su et al., 1995). Acordingly, the denomination progressive dynamic buckling is commonly used to designate the axial impact of shell-like structures.

Oppose to progressive dynamic buckling, under certain conditions, slender tubes can collapse in a global way where just a few hinges are formed. Of course that, in this collapse mechanism, the absorption of impact energy is very little efficient. Since there is a tendency in the automobile industry to use slender structures, it is important to explore geometric parameters, boundary and initial conditions which determine one or other mode of collapse of the structure. This allows a designer to conceive a structure which could collapse progressively, if the aim were to absorb larger impact energy. Figure (1) illustrates these two basic collapse mechanisms of a tube under axial impact load.

Some mechanisms of absorption of impact energy in tubes are reasonably well understood (Reid, 1993; Jones and Abramowicz, 1985 and Abramowicz and Jones, 1986). In the several available methods of analysis of the progressive dynamic buckling phenomenon (Alexander, 1960 and Lindberg and Florence, 1987), the solutions are usually dependent on the tube geometry (Singace et al., 1995) and on the initial and boundary conditions (Murase and Jones, 1993).

(a) (b)

Figure 1. Different collapse mechanisms of a shell under axial impact. (a) progressive and (b) global.

It is all reasonable to suppose that, according to the various parameters quoted above, there should be a transition between global and progressive buckling, such that tubes of same transverse section may exhibit global or progressive collapse when varying the impact speed, the length of the tube and the boundary conditions. It is interesting to note that most of the results available in the open literature refer to tubes loaded mechanically at room temperature. Little work has been done on buckling and post-buckling of tubes under thermal loads, despite its application on the design of airplanes and rockets (Thornton, 1993). For the thermal loads considered in Feldman (1996), for instance, the critical buckling load is significantly affected by the temperature.

In spite of the development of sophisticated computational tools, the area of structural impact still requires a great number of experiments, many of them prohibitive of being accomplished in real scale prototypes. Accordingly, tests in small scale structural models allows the development of potentially cheaper prototypes (Jones, 1989 and Singer et al., 1997). This is particularly true for big structures like ships, capsules transporting nuclear fuel, sea platforms, helicopters, etc. This leads to the third point explored in this article where the validity of the scaling laws is investigated.

2. Experimental Procedure

The basic equipment used during this experimental program is a free fall hammer, designed and built at the University of São Paulo, Fig. (2). It consists in two ten meters long I beams positioned along the walls of the building. These beams are fastened obliquely at the top by two U beams and at the bottom by a four-leg frame, which allows a wide space for movement of test specimens and operator. Lying on the I beams face there are two guides, strictly standing in the vertical position. These guides allow the impact mass, ranging from 10 to 200kg, to slide. Hence, energies of up to 20kJ can be made available by varying the impact mass and height. All the impact position and release of the mass is made by a computer controlled system, comprising in house software and hardware.

Figure 2. The drop hammer rig.

The structure in test rests on an anvil that needed to be designed to have a large mass in order to avoid losses of energy to the foundations. A total mass of around 3700kg was the result of a mounted anvil in concrete and steel plates.

Aluminum tubes of external diameter 50.8mm and thickness 2mm were submitted to axial impact as well as quasi-static loads. It is important to observe that, typically, mechanical properties of room temperature aluminum are little sensitive to the strain rate and, hence, to the impact velocity. Different tubes lengths of 360mm, 500mm and 650mm were tested at room temperature, while a tube of 360mm was tested at four temperatures of -70°C, 22°C, 300°C and 407°C, as measured by thermocouples fixed on the tube and covered by an insulating film. The tubes were heat up in an industrial electric oven and cooled down with dry ice in a specially built thermal box.

Finally, in relation to the above commented laws of similarity, two series of tubes were used: one as the room temperature tubes described in this section, ie lengths of 360mm, 500mm, 650mm x 50.8mm x 2mm, to which a scaling factor of 1 was attributed, and tubes of length 180mm, 250mm, 325mm x 25.4mm x 1mm, resulting in a scaling factor of ½, as listed in Tab. (1).

Table 1. Table of dimensions of the tubes and their respective experiments.

Tube	Scaling factor	External diameter (m)	Thickness (m)	Lengths (m)	Experiments
A	1	0.0508	0.002	0.360-0.500-0.650	Velocity Temperature Scaling
B	1/2	0.0254	0.001	0.180-0.250-0.325	Scaling

3. Results

3.1. Static results

As already indicated, a major objective of this work is to quantify at what extent the impact velocity selects one between two of the preferable modes of buckling, i.e. how the velocity triggers progressive or global buckling. Experimentally, one valid approach would be to fix the impact velocity, hence the drop height, and to vary the shell length. In fact, a similar procedure was done for the quasi-static case where tubes of different lengths were compressed in a press, with the resultant deformation pattern being analyzed. By so doing, it was found a critical quasi-static length of

$$L_{cr} = 0{,}350\text{m} \tag{1}$$

for the aluminum tubes, with Fig. (3) summarizing the tests results. It is interesting to note that a change in just 1mm in the tube length is capable to trigger one or other mode of collapse.

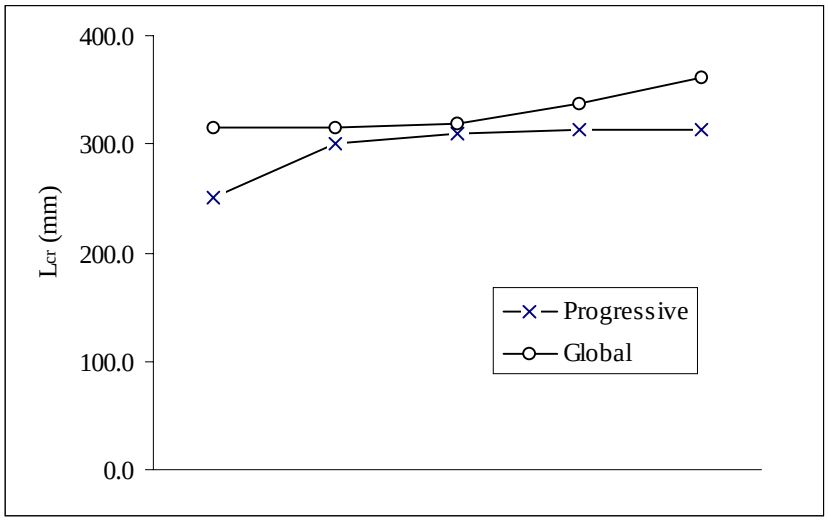


Figure 3. Experimental data describing the influence of the tube length on the mode of collapse.

3.2. Impact tests of room temperature

Additional tests were then accomplished for the impact case. Here, instead of varying the length of the tubes for a fixed speed, it was decide to fix the tube length and to vary the velocity. The reason is that the speed is a parameter easily changed by varying the impact height, since it was assumed that equation

$$V = \sqrt{2gH} \tag{2}$$

where H is the impact height, gives a good approximation for the velocity of the impact mass sliding down the guides. The electronic position control of the impact mass allowed an easy adjustment of the impact velocity, so the tests were accomplished in a simpler way. The impact mass was fixed in 120kg and the results are depicted in Fig. (4). It is evident that the impact velocity plays an important role in the transition between one or other mode of collapse. The higher the impact speed, the more stable the tube, i.e. progressive buckling occurs.

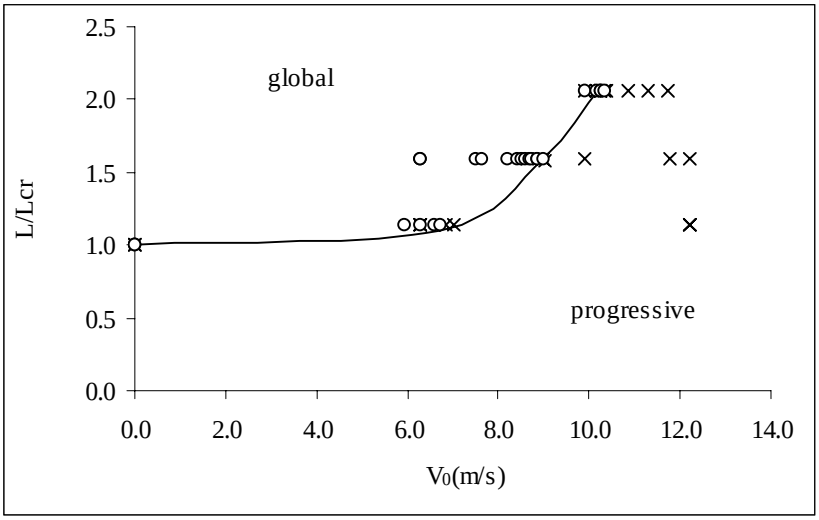


Figure 4. Influence of the impact velocity on the mode of collapse of aluminum tubes.

3.3. Impact tests of various temperatures

A series of tests was performed aimed to study the influence of temperature on the buckling transition, Fig(5). The tubes were heat up in an oven or cooled down in a thermal box. The temperature profile along the tube length, as obtained with thermocouples, indicated that it is uniform, with a small variation at the extreme of the tube which lies on the platform. To avoid such a variation, the tubes were rested on thin insulating pads, which assured a nearly constant temperature along the whole tube length before impact. The temperature at the exact moment of the test was obtained without the need of thermocouples; if they were fixed on the tubes at the actual tests they would be severely damaged.

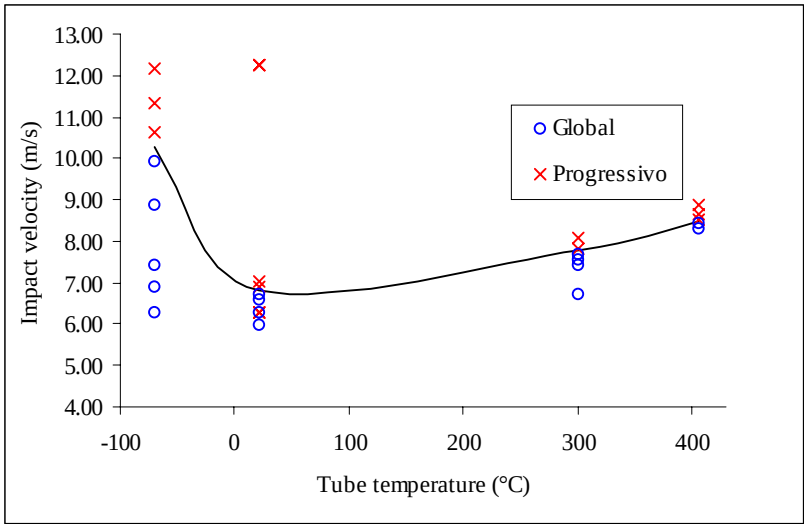


Figure 5. Influence of temperature and velocity on buckling transition of aluminum tubes of L=360mm.

To determine the temperature at the actual impact test, the temperature profile at different length positions along time was measured. Since it was verified that the heat exchange coefficient was approximately constant, it was just necessary to wait a fixed time with the tube already on the impact platform, until the desired temperature was reached. By so proceeding, the transition speed triggering global or progressive buckling was measured for tubes with L=360mm, as indicated in Fig. (5).

The last series of tests to be described are related to scaling. By performing quasi-static tests in the same way as described before, a transition length of 143mm was obtained for the tubes scaled by a factor of ½. Interesting to observe that the scale factor used, i.e. ½, would suggest a transition length of Lcr=315 x ½ = 157.5mm, where 315mm is the critical length obtained for the scaling factor one. Hence, these quasi-static tests indicate that the phenomenon does not obey the laws of scaling. It could be argued that the material of the tubes for the A and B series, see Tab. (1), have different mechanical properties, which remains to be verified. On the other hand, Abramowicz and Jones (1997) suggest that the transition into one or other collapse mode does not seem to be dependent on the material. A first conclusion is then that the transition length marking global or progressive buckling does not obey the laws of scaling in the case of quasi-static loads.

As for the dynamic tests, the same methodology as for the scaling factor 1 was employed for the tubes scaled by the factor ½. The results presented in Fig. (6) are plot dimensionless, normalized using the quasi-static critical length. It can be appreciate that the two curve sets do not coincide, as it would be expected if the scaling laws were valid.

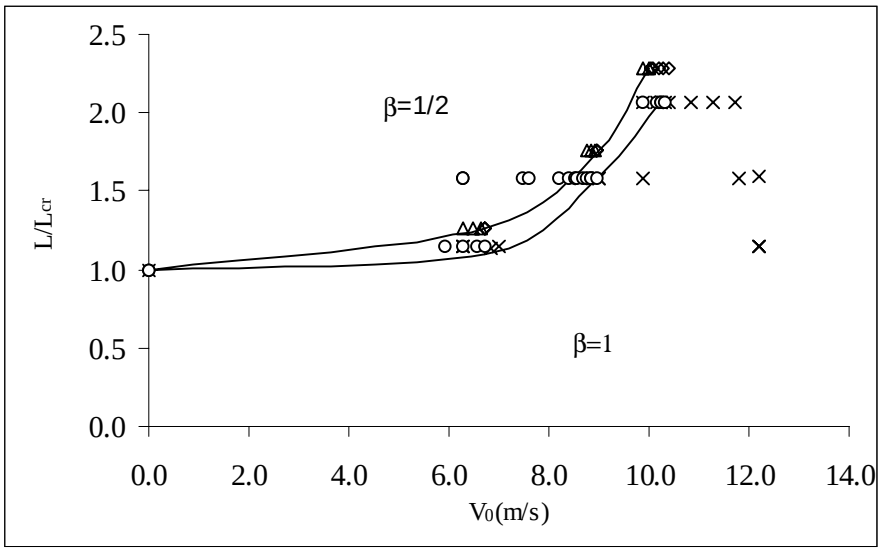


Figure 6. Dimensionless buckling transition length for different impact velocities and two scaled sets of tubes.

4. Discussion

Perhaps the most important fact obtained by this experimental program involving axial impact of tubes is the strong influence of temperature and impact velocity on the mode of collapse of tubes.

It is interesting to notice that few are the works exploring the transition among global and progressive buckling of tubes. Abramowicz and Jones (1997) explored this phenomenon recently, suggesting a simple model to predict the transition length marking the major global and progressive collapse modes. However, their model is unable to capture the influence of material and impact velocity, which certainly points to the need of more theoretical and experimental programs.

The present work was conceived to enhance the knowledge on buckling transition from a purely experimental point of view. This is a first step in better understand the collapse of impacted structures and yields the necessary motivation for additional theoretical studies on the field, as in fact is being now pursued.

The three sets of results shown here are new in the literature. Certainly, they need to be enhanced in many ways. Nevertheless, they do show interesting features not revealed before.

Firstly, it was found a significant increase of the critical length for the transition between global and progressive buckling, as the impact velocity increases. Hence, the larger the impact velocity the longer a tube can be in order to still collapse progressively. This is an important finding which should be taking into consideration in the design of impact energy absorbers.

As expected, temperature plays an important role in determining the transition length between progressive and global buckling. It can be concluded from Fig. (5) that there exists a critical temperature for which the transition length decreases for temperatures higher than this critical value. For smaller temperatures, the transition length increases with

the temperature increase. The determination of this critical point using the current experimental data would be not accurate since more tests would be necessary around 0°C.

For the region above the critical point, larger temperatures would make the material more sensitive to the strain rate (ASM, 1985), leading to global buckling. Accordingly, for temperature higher than the critical one, the transition length reduces. A decrease of the temperature bellows the critical value also leads to a reduction of the critical length as higher impact velocities are required for a progressive buckling response of the tube.

Finally, Fig. (6) indicates that the laws of scaling were not fully satisfied for the dynamic case, although care should be exercised since the discrepancy is not large. Surely, it would be necessary to explore further this non-scaling behavior by, for instance, exploring the phenomenon for other scale factors and by measuring the mechanical properties of the tubes.

5. Conclusion

This work presented experimental results on the transition phenomenon between global and progressive buckling of tubes subject to impact loads and to different temperatures. It was find that both temperature and impact velocity strongly influence the mode of collapse and, hence, the capability of the tubes to absorb more or less impact energy. The experimental program was accomplished thanks to an impact rig called drop hammer, specially designed and built at our structural impact laboratory.

6. Acknowledgements

Part of this project was accomplished with budget from the Fundação de Amparo à Pesquisa do estado de São Paulo, grant n. 99/03559-0.

7. References

Abramowicz, W. and Jones, N., 1986, “Dynamic progressive buckling of circular and square tubes”, International Journal of Impact Engineering, Vol.4, pp. 243-270.

Abramowicz, W. and Jones, N., 1997, “Transition from initial global bending to progressive buckling of tubes loaded statically and dynamically”, International Journal of Impact Engineering, Vol.19, pp. 415-437.

Alexander, J.M., 1960, “An approximate analysis of the collapse of cylindrical shells under axial loading”, Quart. J. Mech. Appl. Math., Vol.13, pp. 10-15.

Andrews, K.R.F., England, G.L. and Ghani, E., 1983, “Classification of the axial collapse of circular tubes under quasi-static load”, International Journal of Mechanical Sciences, Vol.35, No. 7, pp. 597-613.

ASM,1985, “Metals Handbook Desk Edition”.

Feldman, E., 1996, “The effect of temperature-dependent material properties on elasto-viscoplastic buckling behaviour of non-uniformly heated MMC plates”, Composite Structures, Vol.35, pp. 65-74.

Jones, N. and Abramowicz, W., 1985, “Static and dynamic axial crushing of circular and square tubes”, Metal Forming and Impact Mechanics, S.R.Reid, editor. Pergamon, New York, pp. 225-247.

Jones, N., 1989, “Recent studies on the dynamic plastic behavior of structures”, Applied Mechanics Review, Vol.42, pp. 95-115.

Jones, N., 1997, “Structural Impact”, Cambridge University Press.

Karagiozova, D. and Jones, N., 2000, “Dynamic elastic-plastic buckling of circular cylindrical shells under axial impact”, International Journal of Solids and Structures, Vol.37, No. 14, pp. 2005-2034.

Karagiozova, D., Alves, M. and Jones, N., 2000. “Inertia effects in axisymmetrically deformed cylindrical shells under axial impact, International Journal of Impact Engineering, Vol.24, No. 10, pp. 1083-1115.

Lindberg, H.E. and Florence, A.L., 1987, “Dynamic Pulse Buckling : Theory and Experiment”, Martinus Nijhoff, Norwell MA.

McGregor, I.J., Meadows, D.J., Scott, C.E. and Seeds, A.D., 1993, “Impact performance of aluminum structures”, In: Structural Crashworthiness and failure (eds. N. Jones & T. Wierzbicki), Elsevier Applied Science, Barking, Essex, pp. 385-421.

Murase, K., and Jones, N., 1993, “The variation of modes in the dynamic axial plastic buckling of circular tubes”, In: NK Gupta, editor. Plasticity and Impact Mechanics. Wiley Eastern Limited, pp. 222-237.

Reid, S.R., 1993, “Plastic deformation mechanisms in axially compressed metal tubes used as impact energy absorbers”, International Journal of Mechanical Sciences, Vol.35, No. 12, pp. 1035-1052.

Singace, A.A., Elsobky, H., and Reddy, T.Y., 1995, “On the eccentricity factor in the progressive crushing of tubes” International Journal of Solids and Structures, Vol.32, No. 24, pp. 3589-3602.

Singer, J., Arbocz, J., Weller, T., 1997, “Buckling Experiments”, Wiley, Vol.1.

Su, X.Y., Yu, T.H. and Reid, S.R., 1995, “Inertia-sensitive impact energy absorbing systems. Part I – effect of inertia and elasticity”, International Journal of Impact Engineering, Vol.16, pp. 561-672.

Thornton, E.A., 1993, “ThermalBuckling of Plates and Shells”, Applied Mechanics Reviews, ASME, Vol.46, n. 10, p. 485-506.

XVI CONGRESSO BRASILEIRO DE ENGENHARIA MECÂNICA
16th BRAZILIAN CONGRESS OF MECHANICAL ENGINEERING

IMPLEMENTATION OF THE LYAPUNOV FUNCTIONAL IN DIFFERENCE SCHEMES FOR THE SWIFT-HOHENBERG EQUATION

C. I. Christov
Dept. of Mathematics, University of Louisiana at Lafayette,
Lafayette, LA 70504-1010
christov@louisiana.edu

J. Pontes
Metallurgy and Materials Engineering Department – EE/COPPE/UFRJ PO Box 68505
21945-970 Rio de Janeiro RJ, Brazil
jopontes@.ufrj.br

Abstract. Using the operator-splitting method Christov & Pontes (2000, 2001) developed a second-order in time implicit difference scheme for solving the Swift-Hohenberg equation (S-H) which describes pattern formation in Rayleigh-Benard cells. For each time step the scheme involves internal iterations which improve the stability and increase the accuracy with which the Lyapunov functional for S-H is approximated. Different cases of pattern formation were treated and it was shown that the new scheme reaches the stationary pattern several times faster than the previously used first-order in time schemes. In this work we review the main steps concerning the derivation of the second-order in time scheme and prove that the scheme strictly satisfies a discrete approximation of the Lyapunov functional. Results of numerical simulations conducted in four large boxes with two levels of forcings and two different random initial conditions are presented, illustrating that this scheme preserves the non-increasing time-dependent behaviour of the functional. The rate of change of the functional is generaly slow save the pecipituous downfalls during the time intervals in which the pattern changes qualitatively.

Keywords: Nonlinear Systems, Bénard Convection, Implicit Methods, Finite Difference Methods

1 Introduction

A widely accepted model for the thermal convection in a thin layer of fluid heated from below is the so-called Swift-Hohenberg equation (SH, for breivity) (Swift & Hohenberg, 1977) which is a nonlinear parabolic equation containing fourth-order space-derivatives (generalized diffusion equation). It describes the pattern formation in fluid layers confined between horizontal well-conducting boundaries.

Unlike the models that describe wave propagation on the surface of convective layer, the SH equation possesses a Lyapunov functional which ensures the potential behaviour of the solution. One of the properties of the potential evolution is that only spatial chaos is allowed whereas spatio-temporal chaos is not. However, if one uses a difference scheme which does not faithfully represent the Lyapunov functional one might encounter numerically some non-physical effects. For instance, when the Lyapunov functional is only approximately enforced, then initially the solution approaches one of the attractors in the functional space but eventually leaves the domain of attraction of that particular steady-state solution after being "kicked" from one of the small disturbances the latter arising from the inadequate approximation. Then the solution may keep wandering between several of the attractors exhibiting spatio-temporal chaotic behavior which is not possible for the original model.

Solving SH numerically is a challenge because of the interplay between the higher-order diffusion (higher-order spatial derivatives) and nonlinearity on the one hand, and the presence of a Lyapunov functional – on the other. A computationally efficient difference scheme for SH was developed by Christov et al (1997) where a first-order in time implicit time-stepping is used in the framework of the operator-splitting methods.

There are very few papers in the literature which deal with the problems connected to the numerical implementation of an additional integral constraint on the solution, such as the Lyapunov functional. For dissipative systems (Swift-Hohenberg or complex Landau-Ginzburg equation) an elucidating discussion can be found in the paper by Vinals et al (1991), and the one by Montagne et al (1996). For the case of Lyapunov constraint in a conservative system (such as Nonlinear Schrödinger equation) one is referred to the work of Ivonin (1999). Though some fine approximations of the additional integral constraint are presented in the above mentioned papers, they are not in fact strictly preserving approximations. We believe, the strict

implementation is crucial if one is to deal with more complex models.

In this work we review the main steps concerning the derivation of the second-order in time scheme and prove that the scheme strictly satisfies a discrete approximation of the Lyapunov functional (see also Christov & Pontes 2000, 2001). The scheme employs internal iterations to secure adequate approximation of the nonlinear term. Results of numerical simulations conducted in four large boxes are presented, illustrating that this scheme preserves the non-increasing time-dependent behaviour of the functional.

2 Posing the Problem

Consider a rectangular region $D : \{x \in [0, L_x,],\ y \in [0, L_y]\}$ with boundary ∂D at which different types of b.c. can be imposed. Consider the following Generalized Diffusion Equation (GDE):

$$\frac{\partial u}{\partial t} = -D(\Delta + \kappa^2)^2 + F(u) \equiv -D\Delta\Delta u - 2D\kappa^2\Delta u - D\kappa^4 + F(u)\ . \tag{2.1}$$

For the cubic nonlinearity one has the following potential function

$$F(u) = -\frac{dU(u)}{du} = \varepsilon(x)u - gu^3\ , \quad U(u) = -\frac{\varepsilon(x)}{2}u^2 + \frac{g}{4}u^4\ . \tag{2.2}$$

Eq.(2.1) with (2.2) acknowledged is the Swift-Hohenberg equation (SH, for brevity) derived for the Rayleigh-Bénard convection to account for the formation of convective rolls in high Prandtl number fluid layers. The variable $u(x, y, t)$ describes the horizontal planform of the temperature deviation from the conductive profile.

The correct set of lateral b.c. is the one which secures that the evolution of the "energy" $\int v^2$ depends only on its production or dissipation in the bulk, but not on the surface. In other words, in the balance equation for the evolution of energy

$$\frac{d}{dt}\frac{1}{2}\int_D v^2 dxdy = -\oint_{\partial D} v\frac{\partial a_4(x,y,t)\Delta v}{\partial n}\mathrm{d}l - \oint_{\partial D} a_4(x,y,t)\Delta v\frac{\partial v}{\partial n}\mathrm{d}l - \oint_{\partial D} a_2(x,y,t)v\frac{\partial v}{\partial n}\mathrm{d}l -$$
$$\int_D a_0(x,y,t)(\Delta v)^2\mathrm{d}x\mathrm{d}y - \int_D a_2(x,y,t)(\nabla v)^2\mathrm{d}x\mathrm{d}y - \int_D a_0(x,y,t)v^2\mathrm{d}x\mathrm{d}y$$

one has to make the surface integrals vanish. This can happen if one of the following admissible b.c. conditions are imposed

$$v = \frac{\partial v}{\partial n} = 0\ , \quad v = \Delta v = 0\ , \quad \frac{\partial v}{\partial n} = \frac{\partial \Delta v}{\partial n} = 0\ , \qquad (x, y) \in \partial D, \tag{2.3}$$

where n stands for the outward normal direction to the boundary ∂D. We call $(2.3)_1$ and $(2.3)_2$ "generalized Dirichlet conditions" of the first and second kind, respectively. The condition $(2.3)_3$ involves only derivatives at the boundary, hence the coinage – "generalized Neumann condition".

The main feature of (2.1) is that the damping of perturbations occurs via the fourth-order spatial derivatives, while the term with second-order spatial derivatives enhances the perturbations. This allows the occurrence of a linear bifurcation of the solution for certain values of the parameters and/or the size of domain. The nontrivial solutions branch out from a (generally motionless) reference state. Their shapes are the result of the interplay between the complicated linear operator and the nonlinearity. From the perspective of GDE it is clear that a bifurcation can take place only for sufficiently large domains whose size is commensurate with the length scales of the patterns.

SH admits a non-increasing Lyapunov functional

$$\Psi = \int_D \left\{ -\frac{\varepsilon}{2}u^2 + \frac{g}{4}u^4 + \frac{D}{2}\left[(\Delta u)^2 - 2\kappa^2(\nabla u)^2 + \kappa^4 u^2\right] \right\} \mathrm{d}x\,\mathrm{d}y, \tag{2.4}$$

$$\frac{\partial \Psi}{\partial t} = -\int_D \left(\frac{\partial u}{\partial t}\right)^2 \mathrm{d}x\,\mathrm{d}y < 0\ , \tag{2.5}$$

which rules out the complex temporal or spatio-temporal behavior in the long time (turbulent, oscillatory, or chaotic) and allows formation of steady convective patterns. In their turn these steady patterns can be quite complicated in shape, e.g., spatially chaotic.

Consider the rectangle $x \in [0, L_x]$, $y \in [0, L_y]$ and the Dirichlet b.c. of the first and second kind:

$$u = \frac{\partial u}{\partial x} = 0 \text{ for } x = 0,\ L_x; \qquad u = \frac{\partial u}{\partial y} = 0 \text{ for } y = 0,\ L_y, \tag{2.6}$$

$$u = \frac{\partial^2 u}{\partial x^2} = 0 \text{ for } x = 0,\ L_x; \qquad u = \frac{\partial^2 u}{\partial y^2} = 0 \text{ for } y = 0,\ L_y, \tag{2.7}$$

Respectively, the Neumann condition is

$$\frac{\partial u}{\partial x} = \frac{\partial^2 u}{\partial x^2} = 0 \quad \text{for} \quad x = 0,\ L_x, \qquad \frac{\partial u}{\partial y} = \frac{\partial^2 u}{\partial y^2} = 0 \quad \text{for} \quad y = 0,\ L_y, \tag{2.8}$$

There is no restriction to use mixed types of b.c. which are combinations of Dirichlet and Neumann conditions. If the scheme and algorithm perform properly for the "pure" cases (including the Neumann one), then they will do the same for the mixed cases, since any admissible (in the sense of (2.3)) mixture of boundary conditions yields a well posed boundary value problem. For the sake of simplicity we restrict ourselves in the present work to Dirichlet conditions of the first kind.

3 Difference Scheme

3.1 Implicit Time-Stepping

It is not possible to achieve a strict satisfaction of a discrete version of the Lyapunov functional with an explicit scheme which has only two time stages (levels). On the other hand, it is not obvious how the problem can be solved through using multi-level scheme. It is *a-priori* clear that a scheme which is both implicit in time and nonlinear will possess the necessary symmetry to accommodate for the existence of a non-increasing functional. The key to the scheme which satisfies the additional integral constraint (the Lyapunov functional) is the approximation of the nonlinear potential term. A scheme for (2.1) which is implicit and nonlinear reads

$$\begin{aligned}\frac{u^{n+1}-u^n}{\tau} =& D\left[-\frac{\partial^4}{\partial x^4} - \frac{\partial^4}{\partial y^4} - 2k^2\frac{\partial^2}{\partial x^2} - 2k^2\frac{\partial^2}{\partial y^2} - 2\frac{\partial^4}{\partial x^2 \partial y^2} - k^4\right]\frac{u^{n+1}+u^n}{2} \\ & - \frac{U(u^{n+1})-U(u^n)}{u^{n+1}-u^n},\end{aligned} \tag{3.1}$$

where $U(u)$ stands for the potential of the nonlinear force acting upon the system. For the particular case of quartic potential the nonlinear term adopts the form:

$$-\frac{U(u^{n+1})-U(u^n)}{u^{n+1}-u^n} = \frac{\varepsilon(x)}{2}[u^{n+1}+u^n] - \frac{g}{4}\left[(u^{n+1})^3 + (u^{n+1})^2 u^n + u^{n+1}(u^n)^2 + (u^n)^3\right].$$

3.2 Internal Iterations

The scheme (3.1) is nonlinear and can be solved only by means of iterating the solution within a given time step. Additional benefit from the iterations is that they allow us to alleviate a possible problem connected with the inversion of the linear operators when they are not negative definite. This kind of complications can be expected only when $max\{L_x, L_y\}k_0 \geq \alpha$, where α can be estimated from the inclusion theorem. As already mentioned, a simple consequence of the non-definiteness of the linear operator is the occurrence of a linear bifurcation of the stationary problem. We tackle this complication by means of an explicit approximation of the second-order terms. Thus the iterative scheme reads

$$\begin{aligned}\frac{u^{n,k+1}-u^n}{\tau} =& \left[-D\frac{\partial^4}{\partial x^4} - D\frac{\partial^4}{\partial y^4} - Dk^4\right]\frac{u^{n,k+1}+u^n}{2} + \left[-2Dk^2\frac{\partial^2}{\partial x^2} - 2Dk^2\frac{\partial^2}{\partial y^2} - \right. \\ & \left. 2D\frac{\partial^4}{\partial x^2 \partial y^2} + \varepsilon(x)\right]\frac{u^{n,k}+u^n}{2} - \frac{g}{4}\left[(u^{n,k})^2 + u^{n,k}u^n + (u^n)^2\right]u^{n,k+1} - \frac{g}{4}(u^n)^3.\end{aligned} \tag{3.2}$$

Here the superscript $(n, k+1)$ designates the current ("new") iteration of the unknown set function while the indices (n, k) and (n) distinguish the quantities known from the previous iteration and the previous time step, respectively. The scheme with internal iterations is linear for $u^{n,k+1}$. The internal iterations are conducted until the following criterion is satisfied

$$\frac{\max\|u^{n,K+1} - u^{n,K}\|}{\max\|u^{n,K+1}\|} < \delta, \tag{3.3}$$

for certain $k = K$. Then the last iteration gives the sought function on the new time stage, $u^{n+1} \stackrel{\text{def}}{=} u^{n,K+1}$.

The gist of the concept of internal iterations is that the same time step is repeated until convergence. Since the iterative process begins from an initial condition which is the value of the sought function from the previous time step, the number of internal iterations needed for convergence depends heavily on the magnitude

of the time increment τ. For smaller τ the initial condition for the iterations is closer to the sought function and the number of iterations is expected to be small. The trade-off is that a very small τ requires a larger number of time steps which increases the overall number of arithmetic operations per nodal point. Conversely, an inappropriately large τ will bring about a larger number of internal iterations per time step increasing significantly the computational time needed to achieve a single time step dispelling thus the advantage of the larger "strides" (the faster time-stepping). The dependence of the number of internal iterations on τ is nonlinear and leaves a room for optimization. Our numerical experiments show that the calculations are cost effective if the number of internal iterations is in the interval $4 \leq K \leq 16$. This estimate calls for a reduction of the time step when faster processes are treated for which the evolution from a given time stage to the next time stage involves a significant deformation of the field. This means that when faster temporal transients are involved, the usage of larger time steps τ leading to $K \gg 20$ is not justified regardless to the fact that formally speaking the implicit scheme is still stable.

3.3 The Splitting

The inversion of the matrix of Eq. (3.2) is a rather costly procedure even though it is sparse. The 3D case is drastically more expensive. Moreover that the internal iterations require the process to be repeated several times during each time step. Then it is only natural to introduce operator splitting in order to minimize the operations per unit iteration and hence per one time step.

We settle here for the *second Douglas scheme* (Douglas, 1956) (also called "scheme of stabilizing correction" (Yanenko, 1971) which gives the full-time-step approximation for non-commutative operators and is more robust for nonlinear problems than ADI (see Yanenko, 1971) for a review of the splitting schemes and strategies). Another advantage of the stabilizing correction is that for linear problems in 3D it is absolutely stable, while ADI is not. We generalize the Douglas scheme for fourth-order operators and modify it to be second order accurate in time (a Crank-Nicolson type of scheme) as follows

$$\begin{aligned}
\frac{\tilde{u}-u^n}{\tau} &= L_{11}^{n,k}\tilde{u} + L_{22}^{n,k}u^n + \frac{1}{2}\left[-D\frac{\partial^4}{\partial x^4} - D\frac{\partial^4}{\partial y^4} - Dk^4 + \frac{g}{2}(u^n)^2\right]u^n \\
&+ \frac{1}{2}\left[-L_{12} - L_1 - L_2\right]\left[u^{n,k} + u^n\right] \\
\frac{u^{n,k+1}-\tilde{u}}{\tau} &= L_2^{n,k}(u^{n,k+1} - u^n)
\end{aligned}$$

where

$$L_{11}^{n,k} \stackrel{\text{def}}{=} -\frac{D}{2}\frac{\partial^4}{\partial x^4} - \frac{D}{4}k^4 + \frac{g}{8}\left[\left(u^{n,k}\right)^2 + u^{n,k}u^n + (u^n)^2\right] \tag{3.4}$$

$$L_{22}^{n,k} \stackrel{\text{def}}{=} -\frac{D}{2}\frac{\partial^4}{\partial y^4} - \frac{D}{4}k^4 + \frac{g}{8}\left[\left(u^{n,k}\right)^2 + u^{n,k}u^n + (u^n)^2\right] + \frac{\varepsilon(x)}{2} \tag{3.5}$$

$$L_{12} \stackrel{\text{def}}{=} 2D\frac{\partial^4}{\partial x^2\partial y^2}, \qquad L_1 \stackrel{\text{def}}{=} 2Dk^2\frac{\partial^2}{\partial x^2} + \frac{\varepsilon(x)}{2}, \qquad L_2 \stackrel{\text{def}}{=} 2Dk^2\frac{\partial^2}{\partial y^2} + \frac{\varepsilon(x)}{2} \tag{3.6}$$

In order to show that the splitting approximates the original implicit scheme we rewrite (3.4) as follows

$$\begin{aligned}
[E - \tau L_{11}^{n,k}]\tilde{u} &= [E + \tau L_{22}^{n,k}]u^n + \frac{\tau}{2}\left[-L_{12} - L_1 - L_2\right]\left[u^{n,k} + u^n\right] \qquad (3.7)\\
&+ \frac{\tau}{2}\left[-\frac{\partial^4}{\partial x^4} - \frac{\partial^4}{\partial y^4} - k^4 + \frac{g}{2}(u^n)^2\right]u^n, \\
[E - \tau L_{22}^{n,k}]u^{n,k+1} &= \tilde{u} - \tau L_{22}^{n,k}u^n
\end{aligned}$$

Now we are prepared to eliminate the intermediate variable $\tilde{u}$. This is done after applying the operator $[E - \tau L_{11}^{n,k}]$ to the second of equations (3.7) and adding the result to the first one, namely

$$\begin{aligned}
&[E - \tau L_{11}^{n,k}][E - \tau L_{22}^{n,k}]u^{n,k+1} = [E + \tau L_{22}^{n,k}]u^n - \tau[E - \tau L_{11}^{n,k}]L_{22}^{n,k}u^n \\
&+\frac{\tau}{2}\left[-L_{12} - L_1 - L_2\right]\left[u^{n,k} + u^n\right] + \frac{\tau}{2}\left[-\frac{\partial^4}{\partial x^4} - \frac{\partial^4}{\partial y^4} - k^4 + \frac{g}{2}(u^n)^2\right]u^n,
\end{aligned}$$

or else,

$$\begin{aligned}
[E + \tau^2 L_{11}^{n,k}L_{22}^{n,k}]\frac{u^{n,k+1}-u^n}{\tau} &= (L_{11}^{n,k} + L_{22}^{n,k})u^{n,k+1} + \frac{1}{2}\left[-L_{12} - L_1 - L_2\right]\left[u^{n,k} + u^n\right] \\
&+ \frac{1}{2}\left[-D\frac{\partial^4}{\partial x^4} - D\frac{\partial^4}{\partial y^4} - Dk^4 + \frac{g}{2}(u^n)^2\right]u^n. \qquad (3.8)
\end{aligned}$$

Upon acknowledging (3.4), (3.5), (3.6), it is readily shown that (3.8) is in fact (3.2) save the *positive definite* operator of norm larger than unity

$$B \equiv E + \tau^2 L_{11}^{n,k} L_{22}^{n,k} = E + O(\tau^2), \tag{3.9}$$

acting upon the time difference $(u^{n+1} - u^n)/\tau$. Acting upon the time difference means that the operator B has no influence on the steady-state result. The fact that $||B|| > 1$ means that the splitting scheme is more stable than the original implicit scheme. The splitting scheme approximates the desired scheme in full-time steps within the adopted order of approximation $O(\tau^2)$. Thus, employing a splitting does not degrade the temporal approximation of the scheme. In other words, the splitting scheme coincides with the original scheme within the order of approximation of the latter.

4 Implementation of the Lyapunov Functional

After the iterations converge one has $u^{n+1} = u^{n,k+1} = u^{n,k}$ and hence one arrives at a nonlinear scheme in full-time steps which is exactly the difference approximation of scheme (3.1) when in the latter the operators L are replaced by their standard second-order representations in finite differences, Λ (see Christov et al, 1997), namely

$$\frac{u^{n+1} - u^n}{\tau} = (\Lambda_{11} + \Lambda_{22} - \Lambda_1 - \Lambda_2 + 2\Lambda_{12})\frac{u^{n+1} + u^n}{2} + \frac{U(u^{n+1}) - U(u^n)}{u^{n+1} - u^n}\,, \tag{4.1}$$

where

$$\frac{U(u^{n+1}) - U(u^n)}{u^{n+1} - u^n} = -\frac{\varepsilon}{2}\left[u^{n+1} + u^n\right] + \frac{g}{4}\left[(u^{n+1})^3 + (u^{n+1})^2 u^n + u^{n+1}(u^n)^2 + (u^n)^3\right].$$

Note that in the left-hand side we neglected the $O(\tau^2)$ contribution to the operator B from (3.9) as asymptotically vanishing with respect to unity.

Upon multiplying eq.(4.1) by $(u_{ij}^{n+1} - u_{ij}^n)/\tau$ and taking the sum over the spatial indices one obtains

$$\begin{aligned}\sum_{i=2}^{M-1}\sum_{j=2}^{N-1}\left[\frac{u_{i,j}^{n+1} - u_{i,j}^n}{\tau}\right]^2 &= \sum_{i=2}^{M-1}\sum_{j=2}^{N-1}\left\{\frac{\varepsilon}{2\tau}\left[\left(u_{i,j}^{n+1}\right)^2 - \left(u_{i,j}^n\right)^2\right] - \frac{g}{4\tau}\left[\left(u_{i,j}^{n+1}\right)^4 - \left(u_{i,j}^n\right)^4\right]\right\} \\ &+ \sum_{i=2}^{M-1}\sum_{j=2}^{N-1}\frac{u_{i,j}^{n+1} - u_{i,j}^n}{\tau}[\Lambda_{11} + \Lambda_{22} - \Lambda_1 - \Lambda_2 + 2\Lambda_{12}]\frac{u_{i,j}^{n+1} + u_{i,j}^n}{2}\end{aligned} \tag{4.2}$$

The last term is manipulated using integration (summation) by parts and the discrete form of boundary conditions for u (2.6), (2.7) are acknowledged. We demonstrate the procedure on the difference operators in x-direction and for Λ_{11} we get

$$\begin{aligned}\frac{1}{2\tau}\sum_{i=2}^{M-1} u_{ij}^{n+1}\Lambda_{11}u_{i,j}^{n+1} &\stackrel{\text{def}}{=} -\frac{D}{2h_x^4\tau}\sum_{i=2}^{M-1} u_{ij}^{n+1}\left[u_{i-2,j}^{n+1} - 4u_{i-1,j}^{n+1} + 6u_{i,j}^{n+1} - 4u_{i+1,j}^{n+1} + u_{i+2,j}^{n+1}\right] \\ &\equiv -\frac{D}{2h_x^4\tau}\sum_{i=3}^{M-2} u_{ij}^{n+1}\left[u_{i-2,j}^{n+1} - 4u_{i-1,j}^{n+1} + 6u_{i,j}^{n+1} - 4u_{i+1,j}^{n+1} + u_{i+2,j}^{n+1}\right] \\ &= -\frac{D}{2h_x^4\tau}\sum_{i=2}^{M-1}\left[2u_{i+1,j}^{n+1}u_{i-1,j}^{n+1} - 4u_{i-1,j}^{n+1}u_{i,j}^{n+1} + 6u_{i,j}^{n+1}u_{i,j}^{n+1} - 4u_{i+1,j}^{n+1}u_{i,j}^{n+1}\right] \\ &- \frac{D}{2h_x^4\tau}\left[-u_{Mj}^{n+1}u_{M-2j}^{n+1} - u_{M-1j}^{n+1}u_{M-3j}^{n+1} - u_{4j}^{n+1}u_{2j}^{n+1} - u_{3j}^{n+1}u_{1j}^{n+1}\right] \\ &- \frac{D}{2h_x^4\tau}\left[4u_{Mj}^{n+1}u_{M-1j}^{n+1} + 4u_{M-1j}^{n+1}u_{M-2j}^{n+1} + 4u_{3j}^{n+1}u_{2j}^{n+1} + 4u_{2j}^{n+1}u_{1j}^{n+1}\right] \\ &- \frac{D}{2h_x^4\tau}\left[-6\left(u_{M-1j}^{n+1}\right)^2 - 6\left(u_{2j}^{n+1}\right)^2\right] = -\frac{D}{2\tau}\sum_{i=2}^{M-1}\left[\frac{u_{i+1,j}^{n+1} - 2u_{i,j}^{n+1} + u_{i-1,j}^{n+1}}{h_x^2}\right]^2\end{aligned} \tag{4.3}$$

Similarly for Λ_1

$$-\frac{1}{2\tau}\sum_{i=2}^{M-1} u_{ij}^{n+1}\Lambda_1 u_{i,j}^{n+1} \stackrel{\text{def}}{=} -\frac{2Dk^2}{2h_x^2\tau}\sum_{i=2}^{M-1} u_{ij}^{n+1}\left[u_{i-1,j}^{n+1} - 2u_{i,j}^{n+1} + u_{i+1,j}^{n+1}\right]$$

$$= -\frac{2Dk^2}{2h_x^2\tau}\left\{\sum_{i=2}^{M-1}\left[u_{i,j}^{n+1}u_{i+1,j}^{n+1} - \frac{1}{2}\left(u_{i+1,j}^{n+1}\right)^2 - \frac{1}{2}\left(u_{i,j}^{n+1}\right)^2\right] + \frac{1}{2}\left[\left(u_2^{n+1}j\right)^2 - \left(u_{Mj}^{n+1}\right)^2\right]\right\}$$

$$-\frac{2Dk^2}{2h_x^2\tau}\left\{\sum_{i=2}^{M-1}\left[u_{ij}^{n+1}u_{i-1,j}^{n+1} - \frac{1}{2}\left(u_{i,j}^{n+1}\right)^2 - \frac{1}{2}\left(u_{i-1,j}^{n+1}\right)^2\right] - \frac{1}{2}\left[\left(u_1^{n+1}j\right)^2 + \left(u_{M-1j}^{n+1}\right)^2\right]\right\}$$

$$= \frac{2Dk^2}{4\tau}\sum_{i=2}^{M-1}\left\{\left[\frac{u_{i+1,j}^{n+1} - u_{i,j}^{n+1}}{h_x}\right]^2 + \left[\frac{u_{i,j}^{n+1} - u_{i-1,j}^{n+1}}{h_x}\right]^2\right\} \tag{4.4}$$

Making use of the same technique on can show that

$$\frac{1}{2\tau}\sum_{i=2}^{M-1} u_{ij}^n\Lambda_{11}u_{i,j}^{n+1} - u_{ij}^{n+1}\Lambda_{11}u_{i,j}^n = 0, \qquad \frac{1}{2\tau}\sum_{j=2}^{N-1} u_{ij}^n\Lambda_{22}u_{i,j}^{n+1} - u_{ij}^{n+1}\Lambda_{22}u_{i,j}^n = 0,$$

$$\frac{1}{2\tau}\sum_{i=2}^{M-1} u_{ij}^n\Lambda_1 u_{i,j}^{n+1} - u_{ij}^{n+1}\Lambda_1 u_{i,j}^n = 0, \qquad \frac{1}{2\tau}\sum_{j=2}^{N-1} u_{ij}^n\Lambda_2 u_{i,j}^{n+1} - u_{ij}^{n+1}\Lambda_2 u_{i,j}^n = 0.$$

In the same vein are treated the terms connected with the y spatial derivative. The approximation of the mixed fourth derivative make use of the above derivations in both x- and y-direction. For the sake of convenience we denote

$$\Phi_{ij} = \frac{u_{i,j-1} - 2u_{ij} + u_{ij+1}}{h_y^2}.$$

Then

$$\sum_{i=2}^{M-1} u_{ij}\frac{\Phi_{i-1,j} - 2\Phi_{ij} + \Phi_{i+1j}}{h_x^2} = \sum_{i=2}^{M-1}\Phi_{ij}\frac{u_{i-1,j} - 2u_{ij} + u_{i+1j}}{h_x^2}$$

$$+\frac{u_{M-1j}\Phi_{Mj} - u_{1j}\Phi_{2j}}{h_x^2} + \frac{u_{Mj}\Phi_{M-1j} - u_{2j}\Phi_{1j}}{h_x^2} \equiv \sum_{i=2}^{M-1}\Phi_{ij}\frac{u_{i-1,j} - 2u_{ij} + u_{i+1j}}{h_x^2},$$

and hence

$$2\sum_{i=2}^{M-1}\sum_{j=2}^{N-1} u_{ij}\Lambda_{12}u_{ij} = 2\sum_{i=2}^{M-1}\sum_{j=2}^{N-1}\frac{u_{i-1,j} - 2u_{ij} + u_{i+1j}}{h_x^2}\cdot\frac{u_{i,j-1} - 2u_{ij} + u_{ij+1}}{h_y^2}$$

Now it is readily shown that the right-hand side of (4.2) is the time difference of the Lyapunov functional Ψ

$$\frac{\Psi^{n+1} - \Psi^n}{\tau} = -\sum_{i=2}^{M-1}\sum_{j=2}^{N-1}\left(\frac{u_{i,j}^{n+1} - u_{i,j}^n}{\tau}\right)^2 \tag{4.5}$$

$$\Psi^n = \sum_{i=2}^{M-1}\sum_{j=2}^{N-1}\left[-\frac{\varepsilon}{2}\left(u_{i,j}^n\right)^2 + \frac{g}{4}\left(u_{i,j}^n\right)^4 + \frac{Dk^4}{2}\left(u_{i,j}^n\right)^2\right]$$

$$-\frac{2Dk^2}{4}\sum_{i=2}^{M-1}\sum_{j=2}^{N-1}\left[\frac{u_{i+1,j}^n - u_{i,j}^n}{h_x}\right]^2 + \left[\frac{u_{i,j}^n - u_{i-1,j}^n}{h_x}\right]^2 + \left[\frac{u_{i,j+1}^n - u_{i,j}^n}{h_y}\right]^2 + \left[\frac{u_{i,j}^n - u_{i,j-1}^n}{h_y}\right]^2$$

$$+\frac{D}{2}\sum_{i=2}^{M-1}\sum_{j=2}^{N-1}\left[\frac{u_{i+1,j}^n - 2u_{i,j}^n + u_{i-1,j}^n}{h_x^2} + \frac{u_{i,j+1}^n - 2u_{i,j}^n + u_{i,j-1}^n}{h_y^2}\right]^2$$

The last formula presents a $O(\tau^2 + h_x^2 + h_y^2)$ approximation to the original Lyapunov functional (2.5) for the differential equation. The important point about difference version of the Lyapunov functional (4.5) is that it is strictly enforced provided that the internal iterations converge. Its satisfaction does not depend on the truncation error.

5 Numerical Results

The main subject of the present work is to demonstrate the impact of the implementation of the Lyapunov functional on the numerical dynamics. As a featuring example we consider a uniformly loaded system. From the point of view of dynamical systems the behavior of the solution is essentially the same as of the ramped system (Christov et al, 1997) and the only differences are connected with the fact that larger number of solutions with different symmetries can take place for the uniformly loaded system.

Before all we should mention that we checked the spatial and temporal discretizations on a simple case when the whole domain of the flow is occupied by a single convective cell. All mandatory tests involving doubling the temporal and spatial resolutions confirmed the second-order approximation of the scheme and the discrete implementation of the Lyapunov functional. The convergence of the internal iterations is tested for different values of the tolerance δ defined in eq. (3.3). The quantitative differences between the patterns are negligible for $\delta \leq 10^{-5}$. This allows us to choose $\delta = 10^{-6}$.

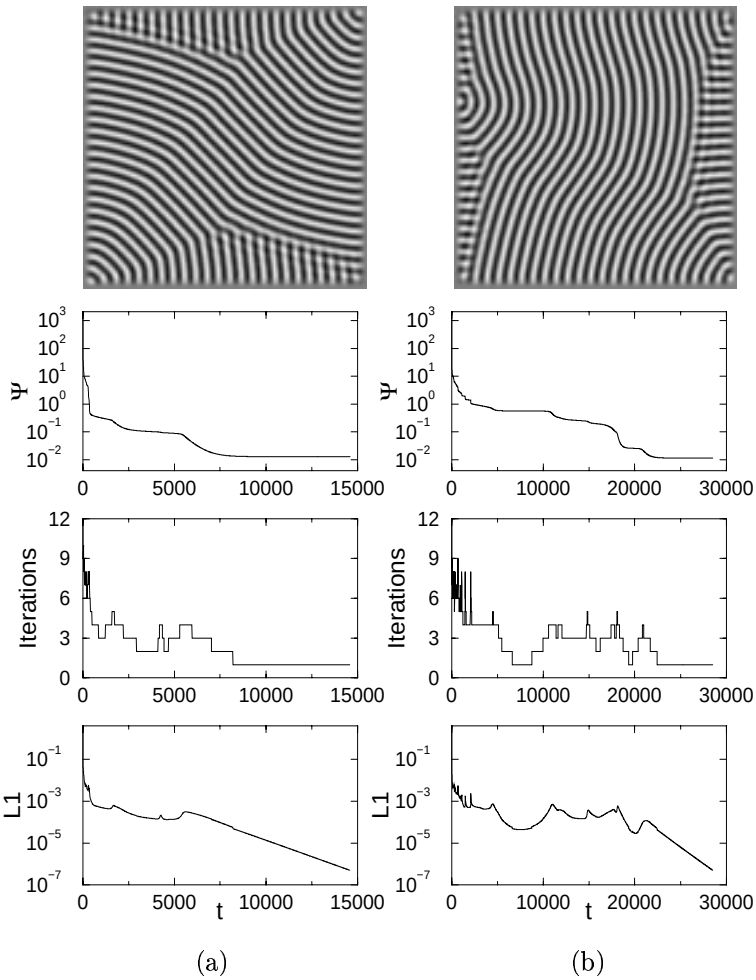


Figure 1: Results obtained in two simulations run in a 50×50 box with a forcing $\varepsilon = 0.25$ and two different random initail conditions. Steady-state patterns, the time evolution of the associated Lyapunov finctional, number of internal iterations accoplished in each time step and time evolution of L_1 are shown from top to bottom. Random initial condition seeds were generated in an IBM SP2 RISC 6000 using the *random* function and seeds 280449 (a) and 123456 (b).

In this work we present the results of four numerical simulations conducted in a large box with size $L_x = 50$ and $L_y = 50$ (referred to in what follows as a 50×50 box). The first two simulations were conducted with a forcing $\varepsilon = 0.25$ and two different random initial conditions whereas we adopted $\varepsilon = 0.5$ in the other two and same initial conditions used in the former ones. The results of these four simulations are summarized in Figs. (1) and (2).

For the sake of definiteness and backward compatibility with the works from the literature we select $\kappa_0 = 3.1172$, $g = 12.9$, and $D = 0.015$ which values correspond to a typical Rayleigh-Bénard convection with pattern formation (see, e.g. Swift & Hohenberg, 1977). When the size of the box is $L_x = L_y = 50$, the selected value for k_0 allows roughly 50 convective cells in each direction. A staggered mesh with 302×302 points is used which gives a spatial resolution of approximately 6 points per convective cell (12 points per wavelength). The time step used in all simulations was $\tau = 0.02$.

The random initial condition is constructed by means of a random generator. The value of u in the first point (lower-left corner) is calculated with the random generator for a given initial seed and then the rest if the grid points are filled row-wise using the previous point as a seed for the next point. Finally, the initial field is renormalized to [-1,1]. Thus, the pattern can be referred by the number of the initial seed. We assess the rate of evolution of the pattern during the simulation by monitoring the relative L_1 norm defined as

$$L_1 = \frac{1}{\tau} \frac{\sum_{i,j} |u_{i,j}^{n+1} - u_{i,j}^{n}|}{\sum_{i,j} |u^{n+1}|}, \tag{5.1}$$

which roughly corresponds to the ratio between the spatial average of the modulus of time derivative u_t and

the spatial average of the modulus of the function itself.

The calculations begin from the random initial condition and proceeded until $L_1 \leq 5 \times 10^{-7}$, when it can be assumed that the motion is virtually steady. In all simulations the Lyapunov functional $\Psi(t)$ and L_1 rapidly decrease at the very beginning of the evolution. In this stage most of the spectral content of the random initial condition is filtered and only wavelengths close to κ_0 survive. This is the stage in which the structure is formed and the amplitude of the pattern saturates. The pattern is characterized by a large nuber of defects and short correlation lenghts. The second stage is characterized by elimination of defects, resulting in longer correlation lenghts. The *phase* of the patterns evolve with minor changes in the average amplitude. The Lyapunov functional decreases at a much lower rate and L_1 norm evolves slower. However, the evolution of L_1 is not monotonical and accelerates when defects collapse. This behaviour is precisely captured by the sudden increases of the norm $L_1(t)$ and abrupt decreases in the Lyapunov functional.

We monitored also the number of internal iterations. In the time intervals of slow evolution, the number of internal iterations is around 1-3. As expected, this number increases in the time intervals of rapid evolution (the same intervals in which the L_1 norm increases).

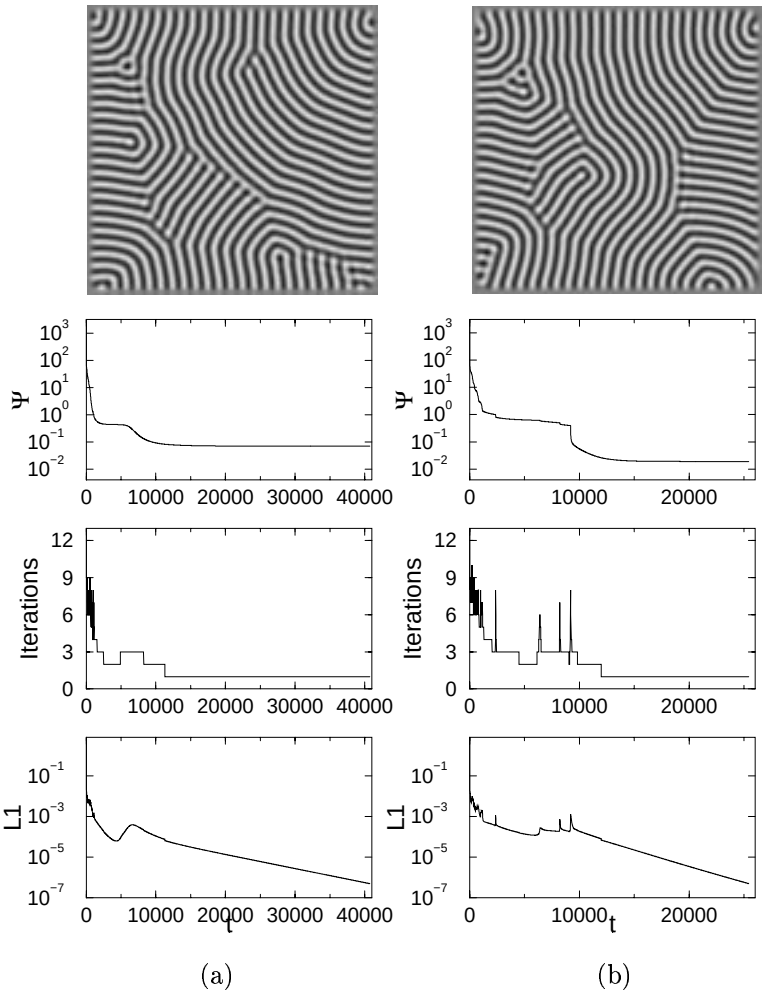


Figure 2: Results obtained in two simulations run in a 50×50 box with a forcing $\varepsilon = 0.50$ and two different random initail conditions. Steady-state patterns, the time evolution of the associated Lyapunov finctional, number of internal iterations accoplished in each time step and time evolution of L_1 are shown from top to bottom. Random initial condition seeds were generated in an IBM SP2 RISC 6000 using the *random* function and seeds 280449 (a) and 123456 (b).

Manipulating the initial condition we succeeded to obtain a plethora of solutions as illustrated in Figs. (1) and (2). The effect of increasing the forcing ε applied to the system is also shown in the figures: higher forcings lead to a higher density of defects and shorter correlation lengths due to the fact that a wider band of wavelengths becames linealy unstable.

6 Conclusions

In the present work we review and applied the operator-splitting difference scheme proposed by Christov and Pontes (2000, 2001), for the numerical solution of nonlinear diffusion equations containing fourth-order space-derivatives. The scheme is of second-order approximation both in time and space and does not contain artificial dispersion, hence the disturbances are quickly attenuated. It is fully implicit owing to the use of internal iterations. The main characteristic of the scheme is that a discrete version of the Lyapunov functional which holds for the original differential equation is strictly implemented.

The performance of the scheme is demonstrated for the evolution of the solution of the Swift-Hohenberg (SH) equation in four simulations conducted in large boxes with two forcing levels and two different random initial conditions. The SH equation models the Rayleigh-Bénard convection in a horizontal layer. The strict implementation of the Lyapunov functional with its non-increasing behaviour is clearly demonstrated. The numerically obtained Lyapunov functional decreases more rapidly when the pattern undergoes qualitative

changes (when the solution enters the immediate vicinity of the respective attractor). In all cases the evolution ended up in a spatially chaotic pattern. No temporal chaos was possible due to the strict satisfaction of the Lyapunov functional.

The proposed difference scheme can serve as a model for constructing approximations to nonlinear physical system when additional constraints, such as Lyapunov functional, are imposed on the solution.

7 Acknowledgements

The work of C.I.C. is supported in part by Grant LEQSF(1999-2002)-RD-A-49 from the Louisiana Board of Regents. J. P. acknowledges financial support from CNPq (Brazil) through grant No. 452399/00-9(NV). The simullations were performed on the IBM SP2 RS-6000 of the NACAD-UFRJ Computer Center at the Federal University of Rio de Janeiro, Brazil.

References

CHRISTOV, C. I., AND PONTES, J. A second-order scheme for evolution equations modelling Bénard convection. In *Proceedings of the* 8^{th} *Brazilian Congress of Thermal Engineering and Sciences* (Brazil, 2000). Paper S06P01 (CD).

CHRISTOV, C. I., AND PONTES, J. Numerical scheme for Swift-Hohenberg equation with strict implementation of the Lyapunov functional. *Mathematical and Computer Modelling* ((To appear)).

CHRISTOV, C. I., PONTES, J., WALGRAEF, D., AND VELARDE, M. G. Implicit time splitting for fourth-order parabolic equations. *Comput. Methods Apll. Mech. Engrg 148* (1997), 209–224.

DOUGLAS, J., AND RACHFORD, H. H. On the numerical solution of heat conduction problems in two and three space variables. *Trans. Amer. Math. Soc. 82* (1956), 421–439.

IVONIN, I. A., PAVLENKO, V. P., AND PERSSON, H. Self-consistent turbulence in the two-dimensional nonlinear Schodinger equation with a repulsive potential. *Phys. Rev. E 60* (1999), 492–499.

MONTAGNE, R., HERNADEZ-GARCIA, E., AND SAN-MIGUEL, M. Numerical study of a Lyapunov functional for the complex Ginzburg-Landau equation. *Physica D 96* (1996), 44–65.

SWIFT, J., AND HOHENBERG, P. C. Hydrodynamic fluctuations at the convective instability. *Phys. Rev. A 15* (1977), 319–328.

VINALS, J., HERNADES-GARCIA, E., SAN MIGUEL, M., AND TORAL, R. Numerical study of the dynamic aspects of pattern selection in the stochastic Swift-Hohenberg equation, in one dimension. *Phys. Rev. A 44* (1991), 1123–1133.

YANENKO, N. N. *Method of Fractional Steps.* Gordon and Breach, NY, 1971.

XVI CONGRESSO BRASILEIRO DE ENGENHARIA MECÂNICA
16th BRAZILIAN CONGRESS OF MECHANICAL ENGINEERING

LOCALIZAÇÃO GEOGRÁFICA DE TRANSMISSORES USANDO DADOS IDEAIS E NÃO IDEAIS, DESVIO DOPPLER E MÍNIMOS QUADRADOS

Cristina Tobler de Sousa
INPE/DMC, Av. dos Astronautas, 1758, C.P. 515, CEP 12227-010, São José dos Campos, Brasil
cristina@dem.inpe.br

Hélio Koiti Kuga
INPE/DMC, Av. dos Astronautas, 1758, C.P. 515, CEP 12227-010, São José dos Campos, Brasil
hkk@dem.inpe.br

Alberto W. Setzer
INPE/LMO, Av. dos Astronautas, 1758, C.P. 515, CEP 12227-010, São José dos Campos, Brasil
asetzer@met.inpe.br

Resumo. *Neste trabalho descreve-se um novo procedimento implementado no Instituto Nacional de Pesquisas Espaciais (INPE) para se obter a localização geográfica de transmissores usando desvio Doppler e estimação por lotes baseada em técnicas estatísticas de mínimos quadrados. A localização geográfica de transmissores através de satélites e seu monitoramento atendem necessidades várias, como busca e resgate de pessoas em locais remotos, monitoramento de bóias oceanográficas, deslocamento de animais, aeronaves, pessoas, equipamentos, tanto para pesquisas científicas como sistemas de segurança. Resultados com dados simulados sob condições ideais, degradadas, e com erros realísticamente inseridos nas medidas de desvio Doppler, mostram, através de análise de erros, que o localizador geográfico funciona apropriadamente. Foram introduzidos também erros nas efemérides do Satélite de Coleta de Dados (SCD-2) e no Satélite Sino Brasileiro de Recursos Terrestres (CBERS-1) utilizados neste trabalho, através de desvios nos seus movimentos médios. Este trabalho modela as medidas de desvio Doppler e a técnica de mínimos quadrados não linear aplicada ao problema. Resultados e análises obtidos através do procedimento de localização geográfica de transmissores, se encontram neste trabalho, demonstrando a precisão na localização obtida.*

Palavras chave: *desvio Doppler, transmissores, localização, mínimos quadrados*

1. Introdução

No Brasil, as atividades de localização geográfica são aplicadas principalmente na Biologia e na Oceanografia. Na Biologia, são fixados mini-transmissores em animais selvagens para monitoramento dos seus deslocamentos e hábitos (Mantovani; Muelbert et al., 2000, por exemplo). Na oceanografia, são lançadas bóias de deriva nos oceanos (Kampel, 1997; Stevenson, 1998) para acompanhamento do seu deslocamento. Estas atividades têm aplicações ainda na localização e resgate de aeronaves e embarcações em situações de emergência, (SARSAT, 2000) e na localização de pesquisadores do Programa Antártico (PROANTAR) na Antártica (Setzer, 1997) em missões de campo.

No presente, o monitoramento geográfico de transmissores é realizado no Brasil através da compra dos dados de localização (latitude e longitude) do sistema francês Argos (CLS, 1989). O sistema processa os sinais recebidos pelos satélites meteorológicos americanos da série "National Oceanic and Atmospheric Administration" (NOAA) (Kidwell, 1991), e os disponibiliza com atraso de várias horas. O mais recente procedimento de localização geográfica de transmissores apresentado neste trabalho (Sousa, 2000), possibilita a obtenção das medidas de localização geográfica, de forma independente, ustilizando-se satélites e estações de recepção brasileiros.

Na seção a segui, encontra-se resumido o modelo do problema de localização geográfica, utilizando medidas de desvio Doppler e técnicas de mínimos quadrados não lineares. A técnica de geolocalização de transmissores determina entre outros parâmetros, a latitude, longitude e altitude de transmissores situados na superfície terrestre, em tempo-quase-real, ou seja, imediatamente após a coleta das medidas.

Os resultados de dados simulados encontram-se descritos em tabelas e gráficos, com as respectivas análises.

2. Modelo básico: localização de transmissores

A localização de um transmissor (ou Plataforma de Coleta de Dados - PCD) pode ser determinada medindo-se o desvio Doppler dos sinais de freqüência devido à velocidade do satélite relativa ao transmissor.

A localização geográfica de PCDs considera que durante a passagem dos satélites SCD-2 e CBERS-1, sinais de freqüência UHF transmitidas pelas PCDs são recebidos a bordo do satélite, que são imediatamente (em tempo real) retransmitidos para estações de recepção terrestres, onde são processados. Isto acontece quando o círculo visibilidade do satélite inclui a estação e o transmissor.

A velocidade do satélite relativa ao transmissor ($v\cos(\alpha)$) será denotada por $\dot{\rho}$, é dada pela Equação do efeito Doppler (Resnick, 1968) como segue:

$$\tilde{n} = \frac{(f_r - f_t)}{f_t} c, \qquad (1)$$

onde:

a) *fr* é a freqüência recebida pelo satélite;
b) *ft* é a freqüência de referência (ou nominal) enviada pelo transmissor;
c) (*fr* - *ft*) é o desvio Doppler devido à velocidade relativa transmissor-satélite;
d) *c* é a velocidade da luz;
e) α é o ângulo entre o vetor velocidade *v* do satélite e o de posição do satélite relativa ao transmissor.

O gerador de efemérides de satélites utiliza o modelo NORAD SPG8 (Hoots, 1980) para obter a órbita do satélite no instante da medida de desvio Doppler. As efemérides atualizadas são utilizadas na Eq. (1) que modela as observações. Portanto, sejam as observações modeladas por:

$$y=h(x)+v \tag{2}$$

onde:

a) *y* é um vetor de *m* observações;
b) *h(x)* é uma função não linear vetorial de dimensão *m* dos parâmetros *x* a serem estimados e das efemérides do satélite;
c) *v* é um vetor de dimensão *m* dos erros de observação.

a solução dos mínimos quadrados não linear é dada por (Bierman, 1977):

$$H_1 \ddot{a}\hat{x} = \ddot{a}y_1 \tag{3}$$

onde $\delta\hat{x} = (\hat{x} - \bar{x})$, H_1 é uma matriz triangular superior, e portanto a solução $\ddot{a}\hat{x}$ é obtida diretamente. O método torna-se iterativo ao tomar a estimativa $\hat{x}$, a partir da Equação (3.25), como sendo o novo valor de referência $\bar{x}$. Assim, calcula-se um novo valor para $\delta\hat{x}$ sucessivamente, até quando essa correção diferencial tender a zero, indicando a convergência do método. A matriz H_1 é o resultado da transformação ortogonal de Householder *T* (Lawson, 1972), tal que:

$$\begin{bmatrix} H_1 \\ \hline 0 \end{bmatrix} = T \begin{bmatrix} P_0^{-1/2} \\ \hline W^{1/2}H \end{bmatrix} \tag{4}$$

onde:

a) *H* é a matriz de derivadas parciais das observações em relação ao estado calculada em torno da referência;
b) *W* é a matriz de peso das componentes do vetor das medidas *y* (ou ponderação dos resíduos);
c) $P_0^{-1/2}$ é a matriz raiz quadrada de informação inicial das variáveis de estado.

O δy_1, é tal que:

$$\begin{bmatrix} \ddot{a}y_1 \\ \hline \ddot{a}y_2 \end{bmatrix} = T \begin{bmatrix} P_0^{-1/2}\ddot{a}\hat{x}_0 \\ \hline W^{1/2}\ddot{a}y \end{bmatrix} \tag{5}$$

onde δy é o vetor dos resíduos. A função custo final pode ser escrita como:

$$J = \left\| \ddot{a}y_1 - H_1 \ddot{a}\hat{x} \right\|^2 + \left\| \ddot{a}y_2 \right\|^2 \tag{6}$$

com $J_{min} = \left\| \ddot{a}y_2 \right\|^2$, onde J_{min} é o custo mínimo.

3. Análises, simulações e qualificação do geolocalizador

Nesta seção descrevem-se simulações computacionais do geolocalizador desenvolvido, com o objetivo de qualificá-lo, verificar seu desempenho sob condições ideais e não ideais mas controladas, e analisá-lo sob aspectos de precisão nas mais diversas situações.

Inicialmente são feitas simulações onde se geram diversos arquivos de dados a partir de passagens do satélite SCD-2 sobre uma estação de recepção conhecida, onde os sinais de freqüência recebidos foram simulados usando a equação de desvio Doppler, as efemérides do satélite e as localizações geográficas nominais dos transmissores. Esse satélite é representativo do satélite quase equatorial (SCD-2, com 25° de inclinação).

A partir desses conjunto de arquivos, vários testes foram realizados com a finalidade de se observar o efeito de erros não inseridos (condição ideal - subseção 3.1), e inseridos propositadamente para simular condições realísticas (condição não ideal – subseção 3.2) tanto nas efemérides como nas medidas Doppler. As simulações e os resultados foram obtidos através de algoritmos desenvolvidos na linguagem "Fortran 77" e compilado em PC (computador pessoal) para fins de portabilidade, onde os "bias" não são estimados.

Os resultados e análises se encontram em tabelas e gráficos, com um total de 20 arquivos simulados, onde cada arquivo eqüivale a uma passagem do satélite e abrange toda a curva Doppler. Dos 20 arquivos, 10 são simulados com alta taxa de amostragem, ou seja, cerca de 600 medidas de desvio Doppler por passagem, o que eqüivale a uma medida a cada segundo; e os outros 10 estão com taxa de amostragem menor, ou seja, uma medida a cada 90 segundos. O tempo decorrido por passagem é de cerca de 10 minutos (para satélites baixos), que eqüivale ao intervalo em que o círculo de visibilidade do satélite inclui a estação receptora e o transmissor.

Foram considerados três transmissores situados em diferentes localizações fixas, com os seguintes valores das respectivas longitudes, latitudes e freqüências nominais:

a) PCD: 32692 (próximo a Cuiabá)
   Longitude = 304.02° (leste)
   Latitude = 17.17° (sul)
   Freqüência nominal = 401.65 MHz
b) PCD: 32694 (Amazonas)
   Longitude = 300.97° (leste)
   Latitude = 6.78° (sul)
   Freqüência nominal = 401.65 MHz
c) PCD: 113 (Peru)
   Longitude = 282.84999° (leste)
   Latitude = 11.7761° (sul)
   Freqüência nominal = 401.62 MHz

Para esses transmissores, utilizou-se as efemérides do satélite quase equatorial SCD-2 e a estação receptora que se encontrava em Cuiabá, onde 4 arquivos (passagens de satélite) foram simulados usando o 32692, 2 usando o 32694, e 14 usando o 113.

Conhecendo-se a posição geográfica (longitude, latitude) de referência dos transmissores, e a calculada pelo software, a expressão que descreve o erro (tri-dimensional) em quilômetros é dada como:

$$Erro(km) = \left\| r_{ref} - r_{cal} \right\| \tag{7}$$

onde:

a) *Erro (km)* é o erro em quilômetros da norma da diferença entre os vetores r_{ref} e r_{cal};
b) r_{ref} é o vetor posição de referência do transmissor;
c) r_{cal} é o vetor posição calculado pelo "software" (geolocalizador);
d) ll · ll é a norma do vetor.

A representação geométrica está indicada na Fig. (1), como segue:

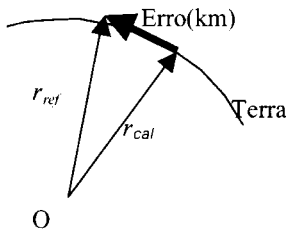


Figura 1 - Representação geométrica do erro em quilômetros.

3.1. Simulações sob condições ideais: alta e baixa amostragens, sem erro nas efemérides e sem erro nas medidas de desvio Doppler

Duas taxas de amostragem das medidas de desvio Doppler foram consideradas:

i) alta amostragem com 600 pontos por passagem, correspondendo a uma taxa de aproximadamente 1 Hz;
ii) baixa amostragem com 7 pontos por passagem correspondendo a uma medida a cada 90 segundos (~0,01 Hz).

Os resultados considerando os 20 arquivos s imulados sob condições ideais, sem erro nas efemérides e sem erro nas medidas Doppler, estão indicados na Tab. (1).

A coluna do erro médio em quilômetros (km), que aparece na tabela seguinte, representa a média aritmética dos erros das localizações geográficas em relação à posição verdadeira do transmissor (conforme Eq. (1)), onde considerou-se 5 passagens ascendentes e 5 descendentes do satélite, por amostragem. As colunas seguintes: máximo e mínimo, representam os valores máximos e mínimos absolutos das 5 passagens; as colunas denominadas por "desc." e "asc." informam se a órbita do satélite foi descendente ou ascendente; e σ é o desvio padrão em relação à amostra de 5 passagens.

Tabela 1. Condições ideais

No de Pass.	Amos-tragem	Erro médio $\pm\sigma$ (km)		Máximo (km)		Mínimo (km)	
		desc.	asc.	desc.	Asc.	desc.	asc.
10	600	5.E-05 ± 4.E-05	9.E-05 ± 4.E-05	1.E-04	2.E-04	1.E-05	5.E-05
10	7	4.E-04 ± 4.E-04	4.E-04 ± 2.E-04	1.E-03	7.E-04	1.E-04	2.E-04

Observando a Tab. (1), o erro médio e a dispersão da localização se encontram entre 10^{-4} km e 10^{-5} km, mostrando que o geolocalizador desenvolvido, usando dados que simulam condições ideais, fornece resultados precisos. Os arquivos com alta amostragem produziram resultados uma ordem de grandeza mais precisos do que os de baixa amostragem, ou seja, uma quantidade maior de medidas implica em resultado estatístico melhor.

3.2. Impacto de erros (simulados)

Foram considerados, nesta seção, os mesmos transmissores, taxas de amostragem, satélites e estações de recepção da seção anterior. Os resultados e as análises estão indicados em tabelas e gráficos.

Considerou-se desta feita os arquivos simulados sob condições não ideais, ou seja, erros inseridos propositadamente. Os erros considerados foram:

- erros aleatórios nas medidas Doppler;
- erros tendenciosos nas medidas Doppler;
- erros nas efemérides dos satélites.

Com a finalidade de se aproximar de condições realísticas, o estimador de "bias" no algoritmo desenvolvido foi desligado, de sorte que o geolocalizador irá assumir que não existe "bias" nas medidas Doppler.

3.2.1. Erros aleatórios nas medidas doppler

Erros aleatórios de 1 Hz, 10 Hz e 100 Hz foram inseridos usando uma distribuição normal padrão (média nula e variância unitária) nas medidas Doppler. Utilizou-se o satélite SCD-2, e os resultados da precisão da localização estão representados na Tab. (2).

Tabela 2. Caso com erros aleatórios nas medidas doppler

No de Pass.	Erros Aleat.(Hz)	Amostra-gem	Erro médio $\pm\sigma$ (km)		Máximo (km)		Mínimo (km)	
			desc.	asc.	desc.	asc.	desc.	asc.
10	1	600	0.02 ± 0.00	0.02 ± 0.00	0.02	0.02	0.01	0.01
10	1	7	0.28 ± 0.21	0.15 ± 0.04	0.59	0.22	0.07	0.12
10	10	600	0.18 ± 0.03	0.19 ± 0.04	0.21	0.22	0.12	0.11
10	10	7	2.84 ± 2.09	1.50 ± 0.36	6.00	2.21	0.69	1.22
10	100	600	1.79 ± 0.32	1.93 ± 0.43	2.14	2.19	1.21	1.07
10	100	7	31.38 ± 24.67	14.99 ± 3.61	69.00	22.09	6.86	12.14

Observando a Tab. (2), os valores da coluna do erro médio aumentaram na medida que o erro aleatório inserido nas medidas Doppler aumentou. O mesmo aconteceu na quantidade de amostras por passagem. Para baixa amostragem (7),

os erros em geral foram maiores do que para alta amostragem (600), cerca de uma ordem de grandeza. As colunas referentes ao máximo e mínimo, indicam que os máximos e mínimos erros absolutos aumentam consideravelmente quando a taxa de amostragem é menor. Estes resultados permitem avaliar quais os valores máximos e mínimos que este tipo de erro aleatório nas medidas Doppler pode causar. Em baixas amostragens um erro aleatório de 1 Hz causa erros de algumas centenas de metros, de 10 Hz causa erros de alguns quilômetros, e 100 Hz algumas dezenas de quilômetros. Os resultados de alta amostragem (600 por passagem) são para referência e nunca serão obtidas na prática.

A representação gráfica do resultado individual das 60 passagens implícitas na Tab. (2) estão indicados na Fig. (2).

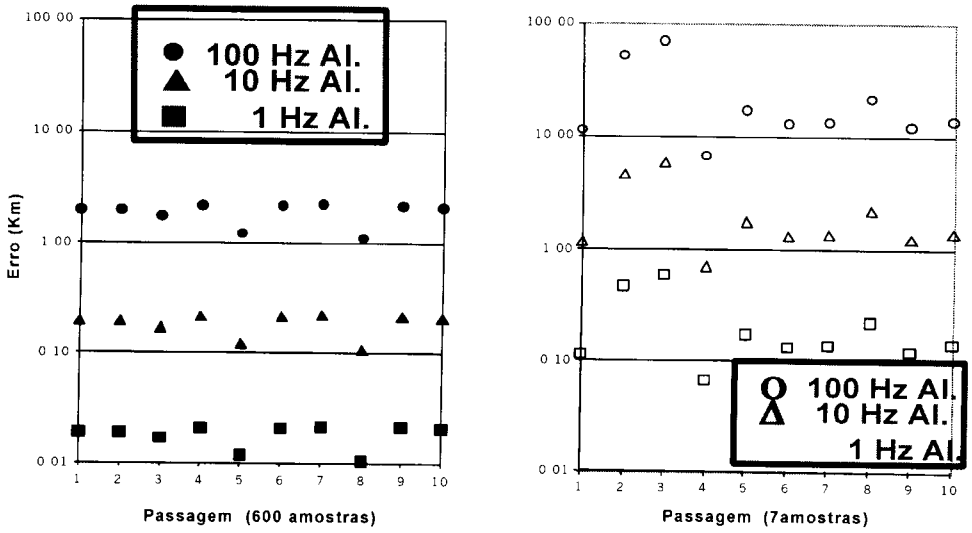


Figura 2. Erro aleatório para passagens contendo 600 e 7 amostras.

Nesta análise, 60 passagens foram consideradas com taxas de 600 e de 7 amostras por passagem. O eixo x das abcissas indica o número da passagem, o eixo y das ordenadas, em escala logarítmica de base 10, se refere à variação do erro em quilômetros na localização, onde o símbolo quadrado representa o resultado ao se inserir erros aleatórios de 1 Hz na medida Doppler; o triângulo, 10 Hz; e o círculo, 100 Hz. Os gráficos do lado esquerdo das figuras consideram alta amostragem, e os do lado direito, baixa amostragem. Comparando os dois gráficos da Fig. (2), verifica-se que os resultados dos erros da localização obtidos com a alta amostragem foram menores, cerca de uma ordem de grandeza logarítmica, do que aqueles com os de baixa amostragem.

3.2.2. Erros tendenciosos nas medidas Doppler

Neste caso, para analisar o efeito de erros sistemáticos, foram inseridos erros tendenciosos de 1 Hz, 10 Hz e 100 Hz nas medidas Doppler, usando os satélites SCD-2 e o NOAA-12. Os resultados estão representados na Tab. (3).

Tabela 3. Caso com erros tendenciosos nas medidas doppler

No de Pass.	Erro Tend.(Hz)	Amos-tragem	Erro médio $\pm\sigma$ (km)	
			desc.	asc.
10	1	600	0.19 ± 0.04	0.26 ± 0.02
10	1	7	0.19 ± 0.03	0.25 ± 0.01
10	10	600	1.88 ± 0.39	2.58 ± 0.17
10	10	7	1.95 ± 0.26	2.51 ± 0.09
10	100	600	18.81 ± 3.84	25.77 ± 1.72
10	100	7	19.44 ± 2.56	25.10 ± 0.90

Observando a Tab. (3), os valores da coluna do erro médio aumentam quando os 3 valores de erros tendenciosos inseridos nas medidas Doppler também aumentam. O erro médio para satélites com órbita ascendente aparenta ser maior do que os de órbita descendente, porém não de maneira significativa a ponto de ser conclusiva. Entretanto, pode-se notar, que devido ao erro inserido nas medidas Doppler ser tendencioso e não de natureza aleatória, os resultados permaneceram na mesma ordem de grandeza, tanto para alta como baixa amostragem. Desta forma conclui-se que erros tendenciosos introduzem um nível de erro que dependem em menor grau da taxa de amostragem. Portanto, justifica-se a introdução do estimador de "bias" no algoritmo de geolocalização.

3.2.3. Erros de efemérides com e sem erros de observação

Nesta seção são considerados erros de 10 e 100 km nas efemérides das órbitas do satélite, para verificar como a precisão nas efemérides, impactam na precisão da localização. Conjuntamente, foram incluídos erros de observação (ou medidas Doppler) de 10 e 100 Hz aleatórios e/ou tendenciosos.

Os resultados considerando apenas erros nas efemérides, ou também adicionando erros de observação estão prescritos na Tab. (4).

Tabela 4. Caso com erros de efemérides e de observação

No de Pass.	Erros Efem.(km)	Erros Doppler (Hz)	Amos-tragem	Erro médio ±σ (km)	
				desc.	Asc.
10	10	Nenhum	600	8.8 ± 0.1	8.6 ± 0.03
10	10		7	8.8 ± 0.1	8.6 ± 0.03
10	10	10 aleat.	600	8.9 ± 0.1	8.7 ± 0.04
10	10		7	9.3 ± 1.1	8.3 ± 0.3
10	10	10 tend.	600	10.5 ± 0.3	10.9 ± 0.1
10	10		7	10.5 ± 0.3	10.9 ± 0.1
10	10	10 aleat.	600	10.5 ± 0.3	11.0 ± 0.1
10	10	+ 10 tend.	7	10.6 ± 0.4	10.7 ± 0.1
10	100	Nenhum	600	88 ± 0.7	86 ± 0.3
10	100		7	88 ± 0.8	86 ± 0.3
10	100	100 aleat.	600	89 ± 0.9	87 ± 0.5
10	100		7	93 ± 12.0	83 ± 2.9
10	100	100 tend.	600	105 ± 2.7	109 ± 0.6
10	100		7	105 ± 2.6	109 ± 0.7
10	100	100 aleat.	600	105 ± 2.7	110 ± 0.6
10	100	+ 100 tend.	7	107 ± 4.2	107 ± 0.8

Observando a Tab. (4), as localizações obtidas tiveram, naturalmente, os maiores erros ao se inserirem todos os tipos de erros simultaneamente. Pode ser notado que a contribuição maior foi devida aos erros inseridos nas efemérides. Os erros aleatórios são filtrados pelo algoritmo de mínimos quadrados e produzem imprecisão marginal. Já os erros sistemáticos ("bias") se traduzem em maior imprecisão final, e o nível de erros adicionado é semelhante ao obtido na Tab. (2). Pode-se concluir que, a precisão nos elementos "two-lines" (efemérides) produz impacto direto na precisão da localização, ou seja, 10 km de erro nas efemérides resulta erros na localização da mesma ordem de grandeza. Logo para os satélites analisados pode-se inferir que precisões da ordem de 1-2 quilômetros só poderão ser obtidos se os erros nas efemérides forem inferiores a 2 km, os erros aleatórios menores que 10 Hz, e os erros tendenciosos menores que 10 Hz, em condições de baixa taxa de amostragem.

A Fig. (3) com 600 e com 7 amostras respectivamente, sintetizam graficamente os diversos tipos de erros simulados versos o erro médio da localização, em escala logarítmica de base 10, conforme valores indicados nas tabelas anteriores, considerando órbita ascendente.

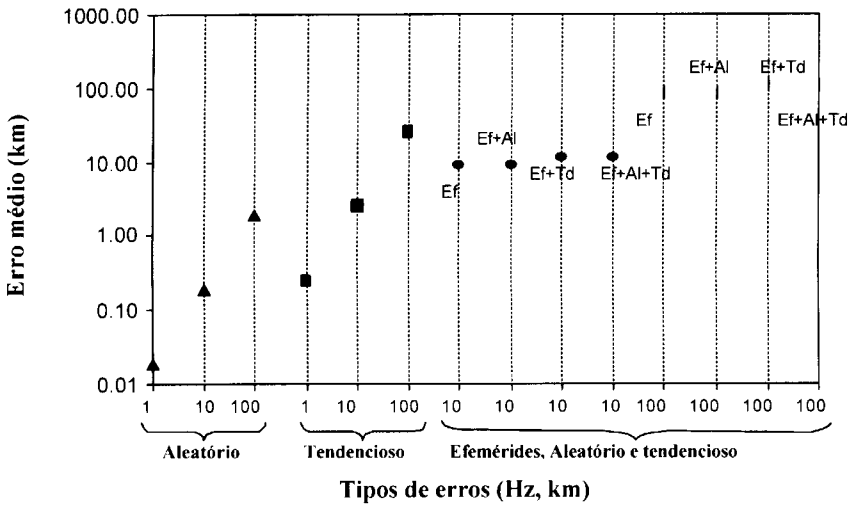


Figura 3. Variação do erro médio de localização versos o tipo de erro considerado, com 600 amostras.

Observando a Fig. (3), os triângulos e os quadrados representam a variação do erro médio em quilômetros, conforme valores da Tab. (2) e da Tab. (3), onde o eixo das abcissas indica a quantidade de erros aleatórios e tendenciosos inseridos nas medidas Doppler e efemérides. Pode-se notar, que a contribuição tendenciosa prevalece sobre a aleatória. Os círculos e losangos indicam o erro em quilômetros obtido, considerando todos os tipos de erros inseridos, tanto nas efemérides como nas medidas Doppler, conforme indicados na área do gráfico. Pode-se perceber que não houve acréscimo significativo ao se inserir, além de erros nas efemérides, erros aleatórios e tendenciosos. A mesma análise pode ser feita considerando 7 amostras por passagem do satélite como mostra a Fig. (4).

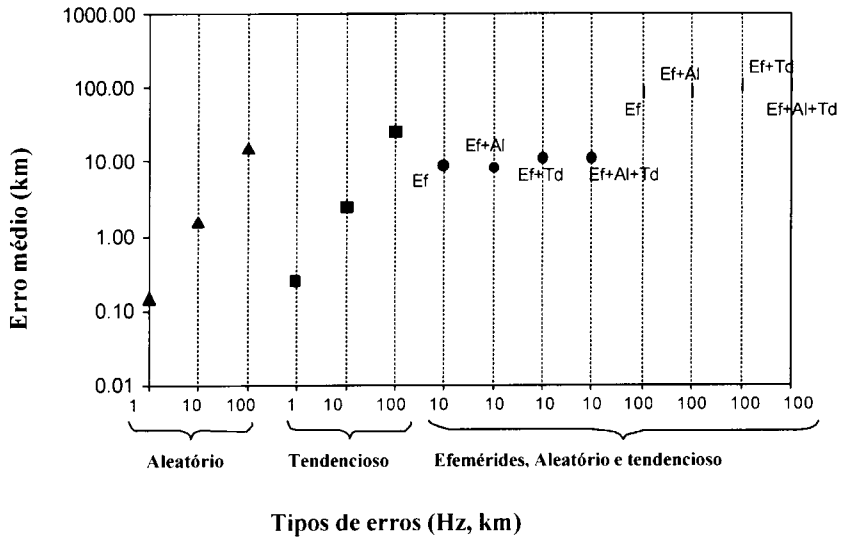


Figura 4. Variação do erro médio de localização versos o tipo de erro considerado, com 7 amostras.

Pode-se notar na Fig. (4), uma diferença em relação à 2, na qual os erros aleatórios, considerando 7 amostras por passagem, tiveram maior influência do que no caso com 600 amostras.

4. Conclusões

A Tab. (5) apresenta resumidamente os resultados. Nelas consideram-se as duas amostragens (600 e 7), as duas órbitas (descendente e ascendente), os diversos tipos de erros inseridos nas medidas Doppler e nas efemérides. Baseando-se nestas tabelas, algumas conclusões podem ser obtidas.

Tabela 5. Síntese de resultados para alta amostragem (600 amostras)

No de pass	Tipo de erro inserido	Amostra-gem	SCD-2 desc.	SCD-2 asc.	Amostra-gem	SCD-2 desc.	SCD-2 asc.
			Erro	Médio (Km)		Erro	Médio (Km)
4.E-04	4.E-04	600	5.E-05	9.E-05	7	4.E-04	4.E-04
0.28	0.15		0.02	0.02		0.28	0.15
2.8	1.5		0.2	0.2		2.8	1.5
31.4	15.0		1.8	1.9		31.4	15.0
0.2	0.3		0.2	0.3		0.2	0.3
2.0	2.5		1.9	2.6		2.0	2.5
19.4	25.1		18.8	25.8		19.4	25.1
8.8	8.6		8.8	8.6		8.8	8.6
9.3	8.3		8.9	8.7		9.3	8.3
10.5	10.9		10.5	10.9		10.5	10.9
10.6	10.7		10.5	11.0		10.6	10.7
88.0	86.0		88.0	86.0		88.0	86.0
93.0	83.0		89.0	87.0		93.0	83.0
105.0	109.0		105.0	109.0		105.0	109.0
107.0	107.0		105.0	110.0		107.0	107.0

Observando o caso onde as freqüências recebidas pelos satélites são simuladas sob condições ideais, ou seja, sem inserção proposital de erros, pode-se concluir que o localizador está operando adequadamente.

Quanto às diferenças entre os resultados da simulação encontrados para os erros aleatórios e os tendenciosos, pode-se concluir, que os erros aleatórios são filtrados. Eles produzem localizações cuja média do erro, tende a zero, principalmente para alta taxa de amostragem (600 medidas). A média dos erros mantém certa coerência e proporcionalidade; assim, para alta amostragem, erros aleatórios de 1, 10 e 100 Hz produzem erros de 0.02 e 0.2 e 2 km, respectivamente. Para 7 amostras obteve-se resultados igualmente consistentes: erros aleatórios de 1, 10 e 100 Hz produzem erros de 0.2 e 2 e 20 km, respectivamente. Os resultados da simulação com erros tendenciosos, por sua vez, produzem erros que degradam igualmente a localização em ambas as amostragens (600 e 7).

Estes resultados eram esperados de acordo com as características intrínsecas dos erros considerados. Pode-se concluir, que quanto menor o erro apresentado nas medidas Doppler, menor o erro de localização em relação à posição verdadeira, principalmente para alta taxa de amostragem. A inclusão de erros tendenciosos prejudica em maior grau os resultados de localização do que os erros aleatórios.

A localização, com tendências nas medidas Doppler, foi calculada sem estimar a deriva nas medidas. Entretanto, estas tendências podem ser retiradas, se suas estimativas forem realizadas no processo de localização.

Observando os resultados da Tab. (5), ao inserir erros de efemérides, conclui-se que é fundamental minimizá-los, pois o localizador não consegue compensá-los. Os erros na localização causados pelas efemérides dos satélites são aproximadamente iguais em magnitude aos erros na localização dos transmissores.

Por fim, nota-se que as precisões da localização independem da passagem do satélite sobre o transmissor ser ascendente ou descendente.

5. Referências

CLS - Service Argos: guide to the Argos System., 1989, Toulouse, September.

Bierman, G.J., 1977, Factorization methods for discret sequential estimation. New York., Academic.

Kampel, M.; Stevenson, M.R., 1997, "Heat transport estimates in the surface layer of the Antarctic polar front using a satellite tracked drifter - first results." In: Proceedings of V International Congress of the Brazilian Geophysical Society, São Paulo.

Lawson, C.L.; Hanson, R.J., 1972, "Solving least Squares Problems." Englewood Cliffs, Prentice.

Muelbert, C.M.M.; et al., 2000, "Movimentos sazonais de elefantes marinhos do sul da ilha elefante, shetland do sul, antártica, observações através de telemetria de satélites." (in Portuguese) VII Seminars on Antartica research. São Paulo 2000. Program and abstracts. São Paulo, USP. IG. p.38.

Resnick, R., 1968, "Introdução à relatividade especial." New York, Wiley.

Sousa, T.S. 2000, "Geolocalização de transmissores com satélites usando desvio Doppler em tempo quase real." (in Portuguese) PhD dissertation, São José dos Campos, INPE.

Setzer, A.W., 1997, "Manual do sistema de dados remotos." (in Portuguese) São José dos Campos, INPE, ProAntar project.

Stevenson, M. R.; Dias-Brito, D.; Stech, J. L.; Kampel, M., 1998, How do cold biota arrive in a tropical bay near Rio de Janeiro Bay, Brazil. Continental Shelf Research, v. 18, n. 13, p. 1595-1612.
Techno - Sciences., 2000, COSPAR/SARSAT. [online] www.technosci.com/ 2000, Oct. 2000.

GEO-LOCATION OF TRANSMITTERS USING IDEAL AND NON-IDEAL DATA, DOPPLER SHIFTS AND LEAST SQUARES

Cristina Tobler de Sousa
INPE/DMC, Av. dos Astronautas, 1758, C.P. 515, CEP 12227-010, São José dos Campos, Brasil
cristina@dem.inpe.br

Hélio Koiti Kuga
INPE/DMC, Av. dos Astronautas, 1758, C.P. 515, CEP 12227-010, São José dos Campos, Brasil
hkk@dem.inpe.br

Alberto W. Setzer
INPE/LMO, Av. dos Astronautas, 1758, C.P. 515, CEP 12227-010, São José dos Campos, Brasil
asetzer@met.inpe.br

Abstract. In this paper we describe method implemented at INPE (Brazilian Space Research Institute) for obtaining the geographic localization (geo-location) of transmitters via satellites, using Doppler shifts and batch estimator based on least-squares technique. Near real-time geographical location of transmitters through satellites is used to monitor and rescue people in remote areas, or to track animals and oceanographic buoys for scientific research. The geo-location results were obtained through simulated sets of data. First, simulations for ideal cases were carried out. Afterwards, there were introduced random and bias noise in the nominal frequency, to characterize the different levels of errors encountered in practice. Finally, satellite ephemeris errors were introduced through deviations in the mean motion of satellites SCD-2 and CBERS-1 utilized in this work. This paper addresses the modeling of the Doppler shift measurements and the non-linear least squares technique applied to the problem. The results and analysis of the geographic location method are shown and depicted in the paper, demonstrating the location accuracy achievable by such a near real-time system technique.

Keywords. *Doppler shifts, transmitter, geo-location, least squares*

XVI CONGRESSO BRASILEIRO DE ENGENHARIA MECÂNICA
16th BRAZILIAN CONGRESS OF MECHANICAL ENGINEERING

THE EARTH'S SHADOW AND THE SPACECRAFT'S ATTITUDE PROPAGATION

Regina Elaine Santos Cabette
UNESP - Faculdade de Engenharia de Guaratinguetá-Departamento de Matemática
Av. Ariberto Pereira da Cunha, 333 - Pedregulho
cabette@uol.com.br
Maria Cecília Zanardi
UNESP - Faculdade de Engenharia de Guaratinguetá-Departamento de Matemática
Av. Ariberto Pereira da Cunha, 333 - Pedregulho
cecília@feg.uneps.br
Rodolpho Vilhena de Moraes
UNESP - Faculdade de Engenharia de Guaratinguetá-Departamento de Matemática
Av. Ariberto Pereira da Cunha, 333 - Pedregulho
rodolpho@feg.unesp.br

Abstract: *A semi-analytical approach is proposed to study the rotational motion of an artificial satellites under the influence of torque due to the solar radiation pressure and taking into account the influence of Earth's shadow. The equations of motion are given in the Hamiltonian formalism and the Andoyer variables are used to describe the spacecraft's rotational motion. The model for the shadow function (ψ) considers three specific regions: shadow ($\psi = 0$), penumbra ($0 < \psi < 1$) and full light ($\psi = 1$), in order to get the variations of the state variables close to the real. The considered model was developed by Kabelac (1988) and takes into account geometrical and physical factors. The semi-analytical process is applied to get the solution for the equations of motion, using the mapping of the shadow function. By mapping and computing the shadow function, we can get the periods in which the satellite is in the penumbra and when it is not illuminate (the torque due to the solar radiation pressure is zero). When the satellite is totally illuminated or it is in the penumbra, a known analytical solution is used to predict the satellite's attitude. The numerical simulation shows that the transition from the shadow region to the illuminate region was attenuated because the penumbra region is included in the shadow function.*

Key Words*: Rotational Motion, Artificial Satellites, Solar Radiation Pressure, Earth's Shadow Function.*

1. Introduction

The study of solar radiation pressure torque and its influence in the motion of artificial satellites has been developed and improved for several authors. Efficient models for describing the influence of Earth's shadow in the torque due to the solar radiation pressure have been presented. Thus an analytical or numerical solution for the system of equations that describe the rotational motion of this artificial satellites can be determined with more precision.

Ferraz Mello (1972) proposed a mathematic model for the study of the effects of the direct solar radiation pressure in the orbit of an artificial satellite. The equations of the orbital motion are integrated using the Von Zeipel method to eliminate the periodic terms. The shadow function was assigned one when the satellite was iluminated by the Sun and zero when the satellite was in the Earth's shadow. This function was defined for angle λ, formed between the geocentric directions of the satellite and the dark pole of the Earth.

The Ferraz-Mello model for the shadow function was used by Vilhena de Moraes and Zanardi (1997) to analyze the equations of rotational motion when the Earth's shadow is considered in modeling the solar radiation torque. A semi-analytical procedure was proposed to solve the equations of motion.

Kabelac (1988) provided a new theory for the shadow function. In this theory were considered the geometrical factors: - the mutual position of Sun, Earth and satellite, - the shape of the conical surface shadows and - the effect of refraction. The following physical factors were also taken into account: - light diffusion due to refraction, - atmospheric absorption, - light diffusion and - the effect of ozone. The results were compared to photometric measurements of the sunset, as observed from the satellite. The study was made including the passage of the satellite through the Sun light, shadow and penumbra. This will be the model used in this paper.

Here Andoyer's variables (l_1,l_2,l_3) and (L_1,L_2,L_3) are used to describe the motion of the satellite around its center of mass. Andoyer's metric variables L_i, i=1,2,3, are defined as follows: L_2 is the modulus of the rotational angular momentum, L_1 and L_2 are respectively, the projection of L_2, on the z-axis of the body frame and on the inertial Z-axis. The Andoyer's angular variables, l_i, i=1,2,3, are related with the orientation of the body frame with respect to the inertial frame.

The objective of this paper is to include the effects of the Earth's penumbra in the rotational motion of an artificial satellite under the influence of solar radiation torque using the Kabelac model for the shadow function. A semi-analytical approach is proposed to solve the equation for the rotational motion when the Earth's shadow is taken into account in the modeling of the solar radiation torque.

2. Equations for the Rotational Motion

Using Andoyer's variables L_i and l_i (i = 1, 2, 3) considering the torque produced by solar radiation pressure, the equations for the rotational motion can be expressed as follows:

$$\frac{dL_i}{dt} = -\frac{\partial H}{\partial l_i} + \psi S_i \ , \quad \frac{dl_i}{dt} = \frac{\partial H}{\partial L_i} + \psi P_i \tag{1}$$

here H is the Hamiltonian for the torque-free problem and P_i and S_i are the functions of the solar radiation torque components expressed in the system of the principal axes of inertia. These functions depend on the Andoyer's variables, on the Sun satellite distance, on the right ascension and declination of the Sun, on the orbital elements of the satellite and on the physical and geometrical properties of the satellites (Zanardi & Vilhena de Moraes, 1999). ψ is a function of the orbital elements of the satellite, of the longitude of the Sun and ecliptic's obliquity. In this paper ψ is described by Kabelac's model, which is presented in the section 3.

When the Earth's shadow is not considered (ψ = 1), an analytical solution for the system (1) was obtained for a circular cylindrical satellite and using the Lagrange's method of variation parameter (Zanardi & Vilhena de Moraes, 1994). This solution shows that the solar radiation torque produces periodic variations in L_2 and L_3, and secular and periodic variations in all of the angular variables l_1, l_2, l_3. In this case, this solution shows that there is no influence of the solar radiation torque on the variable L_1.

3. Shadow Function

In the version of the Earth's shadow function theory presented by Kabelac (1988), are considered the geometrical factors (the mutual position of Sun, Earth and satellite, the shape of the conical surface shadows and the effect of refraction) and the physical factors (light diffusion due to refraction, atmospheric absorption, light diffusion and the effect of ozone). Thus the shadow function is composed by a physical shadow function ψ_f and a geometrical shadow function ψ_g:

$$\psi = \psi_g . \psi_f \tag{2}$$

3.1. Geometrical Shadow Function

3.1.1. Trajectory of the Light Ray and Its Geometry

We shall consider a ray p, propagating from an arbitrary point of the Sun, which passes through the Earth's atmosphere and reaches the surface of the satellite.

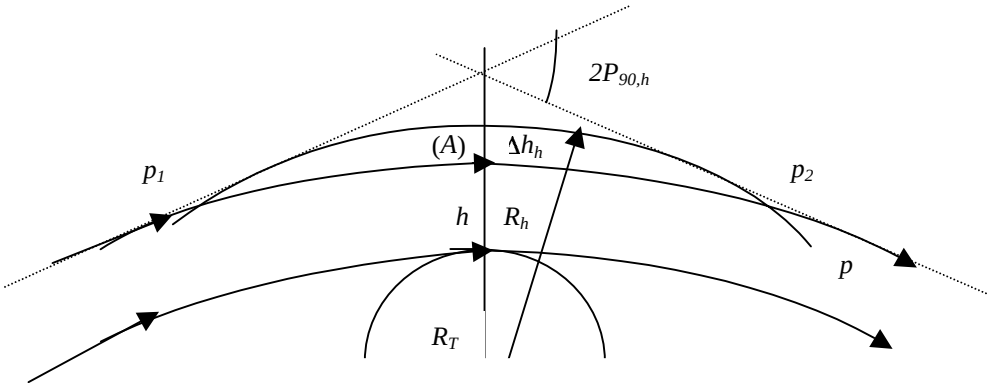


Figure 1 - Trajectory of the light rays where: h is the height of the Earth's atmosphere, $2P_{90,h}$ is the atmospheric refraction angle in the height h, R_h is the radius of the sphere in the height h, R_T is the radius of the Earth.

In Figure 1, the trajectory of the ray is represented and the principal parameters associated with them are shown.

3.1.2. Shadow Cones

The regions that present conic shapes at the moment that the Earth hides the Sun light, when observed from the satellite are denominated shadow cones. Then it's necessary to study the geometry of the problem. In order to do this, consider three shadow cones: cone k_i, umbra region, cone k_e, penumbra region and cone k_E, full light region, which are shown in Fig. 2. In Fig. 3, $\Delta\chi$ represents the angle under which the center of the Sun is seen from the satellite placed in the point $P(x,y,z)$ above the contour of the Earth and it is given for the following expression (Kabelac, 1988).

$$\Delta\chi = 2P_{90} - \mathrm{R}_0/\mathrm{r}_\mathrm{O} - \alpha \tag{3}$$

where: $\Delta\chi$ is the angle related to the ray p and depends on z, α and R_0;
r_O is the distance Earth-Sun;
$R_0 = (R_T + \Delta h_0)\ cos\ P_{90}$;

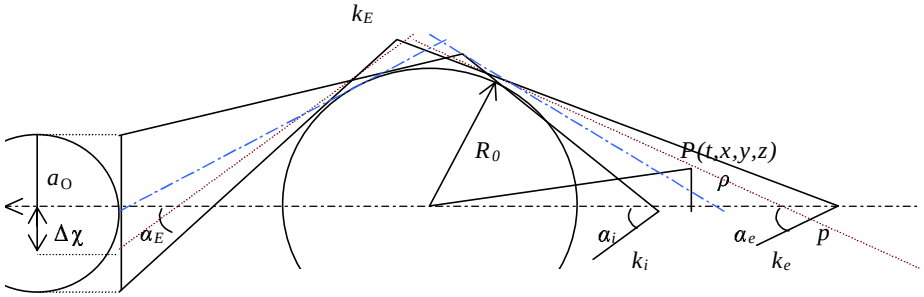


Figure 2 - Shadow Cones, where: α_i is the angle related to the umbra region, α_e is the angle related to the penumbra region, α_E is the angle related to the totally illuminated region, ρ is given by: $\rho = (x^2 + y^2)^{1/2}$ and x, y, z are the components of the Earth-satellite vector position in the system Earth-Sun (the system Earth-Sun (Oxyz) is a geocentric system where the z axis points towards the Sun).

In the shadow cones, the geometrical shadow function assumes the values: inside the cone k_i (umbra) $\psi_g = 0$, outside k_i and inside the cone k_e (penumbra) $0 < \psi_g < 1$ and outside k_e (full illumination) $\psi_g = 1$.

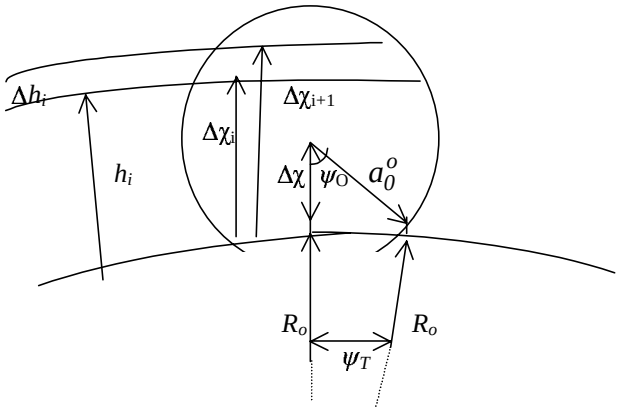


Figure 3 - View angle observed from the satellite (Angle $\Delta\chi$).

The orbital elements of the satellite trajectory around the Earth are known and the vector position **r** is expressed in the orbital system (*Oxoyozo*). The components of **r** in the Earth-Sun system can be obtained using the rotation matrix that relates these systems. This rotation matrix will depend on the satellite orbital elements, the Ecliptic obliquity and the longitude of the Sun (Vilhena de Moraes e Zanardi, 1997).

The geometrical shadow function (Kabelac, 1988) is expressed by:

$$\psi_g = 1 - \frac{1}{\pi}[\psi_\mathrm{O} - sen\psi_\mathrm{O} cos\psi_\mathrm{O} + \frac{2}{3}(\frac{R_0^o}{a_0^o})\psi_T^3] \tag{4}$$

where ψ_O and ψ_T are shown in Fig. 3. The numerical implementation of ψ_g is presented in Cabette (2001).

3.2. Physical Shadow Function

In the Kabelac's development of the Physical Shadow Function is analyzed the illumination effect on the entry and exit of the solar rays in the Earth's atmosphere. Based on the studies of lunar eclipses developed by Link (1956), it can be assumed that:

$$\psi_f = \psi_{f1}\,\psi_{f2}\,\psi_{f3}\,\psi_{f4} \tag{5}$$

where $\psi_{f1} = 10^{-d_1}$, $\psi_{f2} = 10^{-d_2}$, $\psi_{f3} = 10^{-d_3}$, $\psi_{f4} = 10^{-d_4}$.

The symbol d_1 takes into account the influence of the refraction, the symbol d_2 takes into account the atmospheric absorption and the symbols d_3 and d_4 consider the influence of the light diffusion and the effect of ozone respectively. These symbols can be taken from Kabelac (1988), to heights from zero to 100 km above the Earth surface. It's important to note that the solar rays that pass through the Earth's atmosphere and are hitting on the surface of the satellite are gradually attenuated by the Earth's atmosphere, when the satellite is in the penumbra region, which means: - the lower the height h, the denser is the Earth's atmosphere and more attenuated will be the solar rays that will reach the satellite surface. Therefore, the physical shadow function in this case will have a smaller value. - while the height h increases, the Earth atmosphere decreases, and its influence under the solar rays that reach the satellite surface in the orbit will be smaller. Thus, the physical shadow function will have larger value. The numerical implementation of ψ_f is also presented in Cabette (2001).

4. Semi-Analytical Solution Approach

It will be presented here a semi-analytical process to solve the system (1). In this process a mapping for the function ψ is performed and a known analytical solution is used in order to compute the variables describing the rotational motion when $\psi = 1$.

This analytical solution (Zanardi & Vilhena de Moraes, 1994) is given by:

$$L_1 = \bar{L}_{10} \tag{6}$$

$$L_k = \bar{L}_{k0} + \delta L_k \,, \; k = 2,3 \tag{7}$$

$$l_j = \bar{l}_{j0} + (\bar{n}_{j0} + \Delta \bar{n}_j)\, t + \delta l_j \,, \; j = 1,2 \tag{8}$$

$$l_3 = \bar{l}_{30} + \Delta \bar{n}_3 \, t + \delta l_3 \tag{9}$$

where: $\bar{L}_{i0}$ e $\bar{l}_{i0}$, (i = 1, 2, 3) are constants depending on the initial conditions;

δL_k e δl_i are periodic functions;

$\bar{n}_{j0}$ (j = 1,2) e $\Delta \bar{n}_i$ are linear coefficients associated with the torque-free problem and with solar radiation pressure torque, respectively.

This solution shows that:

1) the projection of the rotational angular momentum on the Oz axis principal inertial system is not influenced by the torque derived from the solar radiation pressure;

2) the solar radiation torque produces secular and periodic variations in the angular variables l_i , i = *1, 2, 3*, and only periodic variations in the rotational angular momentum L_2 and in this projection on the inertial OZ axis.

An accumulative error can be introduced in the solution (6-9) if the shadow is not considered, because when the satellite is not illuminated, the variables l_1, l_2, L_2 and L_3 are constants. Then the semi-analytical process is developed, and in this procedure the shadow function is obtained numerically by Eq. (4) and analytical solution (6-9) is used when the satellite isn't in the Earth's shadow. This approach is similar to the semi-analytical one presented in Vilhena de Moraes & Zanardi (1997), when the shadow function is described by Ferraz -Mello model (Ferraz-Mello, 1972).

The semi-analytical solution is given by the following algorithm:

1) from t_0 to t_1, the satellite is in the shadow. The solar radiation torque is zero;

2) from t_1 to t_2 the satellite is in penumbra region and the influence of the solar radiation torque is partial. The shadow function gets the value given by Eq. (2), ($0 < \psi < 1$). The analytical solution (6-9) can be used, with the terms associated with the solar radiation torque multiplied by ψ , t_1 being the initial time of integration and t_2 the final time of the integration;

3) from t_2 to t_3 the satellite is illuminated and the analytical solution is valid, t_2 being the initial time of integration and t_3 the final time of integration for the computation of analytical solution in the interval t_2 to t_3;

4) from t_3 to t_4 the satellite is in penumbra and the procedure is similar to the item 2;

5) from t_4 to t_5 the satellite is not affected by the solar radiation torque. In this interval, the variables L_2, L_3, l_1 and l_3 remain constants, with value fixed at t_4;

6) after t_5 the procedure will be repeated.

4.1. Numerical Simulation

In order to analyze orders for the magnitude of the solar radiation torque, and to validate our semi-analytical approach, let us consider an hypothetical satellite with its orbital and physical characteristics given in Tab. 1.

Table 1: Characteristics of the satellite, where A, B and C are the principal moments of inertia

Orbital Elements	Andoyer Variables (at initial time t_0)	Physical/Geometrical Elements
e = 0,01617	$l_1 = l_2 = \pi/2$ rad	A = B = 3.9499$\times 10^5$ kg m^2
a = 6.95964 $\times 10^3$ km	$l_3 =$ 5.87 rad	C = 1.0307$\times 10^5$ kg m^2
I = 0.5533 rad	$L_1 = 0$	M = 11.550 kg
$\Omega = \pi/6$ rad	L_2 = 9.7307$\times 10^3$ kgm^{2}s^{-1}	orbital period = 5778 s
w = 0 rad	L_3 = -2.9956$\times 10^3$ kgm^{2}s^{-1}	rotational period = 255 s

The analytical solution (valid when $\psi = 1$) and the proposed semi-analytical solution ($0 \leq \psi \leq 1$) are shown in Fig. 4 at 0 - 10000 s. In this simulation, it is assumed that $\psi_f = 1$. The magnitude of the effects due to the solar radiation torque in the variations of Andoyer's variables, at 10000s, are 10^{-8}rad for l_1, 10^{-7} rad for l_2 and l_3, 10^{-8} kgkm^2/s for L_3 and 10^{-10} for L_2. These results agree with the results obtained by Vilhena de Moraes & Zanardi (1997). The passage by the penumbra region was fast (about 8 s) for the considered satellite (Cabette, 2001), because of that, it is not possible to note this transition in Fig. 4 e 5. The results for the passage by two initial penumbra regions, t_1 to t_2 and t_3 to t_4, are shown in Figs. 6-10. Then, it is possible to observe that the transition between the shadow and the illuminated region is attenuated, due to the penumbra.

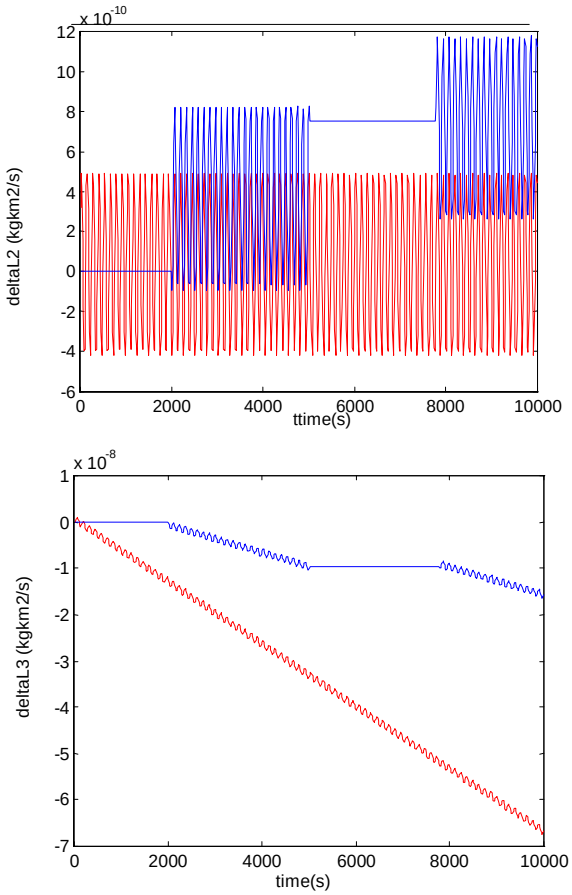


Figure 4: Influence of the Earth's shadow on the metric variables deltaL_2 = L_2 - L_{20} and deltaL_3 = L_3 - L_{30} (--- satellite passes through shadow , --- satellite always illuminated)

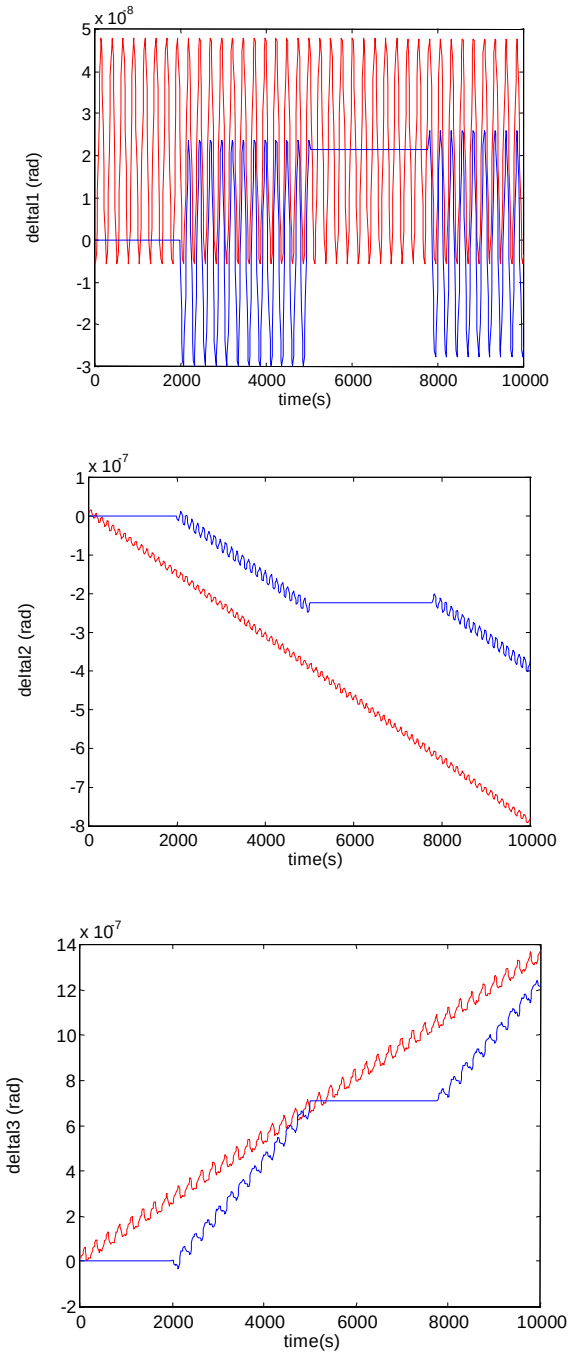


Figure 5: Influence of the Earth's shadow on the angular variables deltal_1 = l_1 - l_{10} , deltal_2 = l_2 - l_{20} - (L_2/A) t and deltal_3 = l_3 - l_{30} (--- satellite passes through shadow , --- satellite always illuminated)

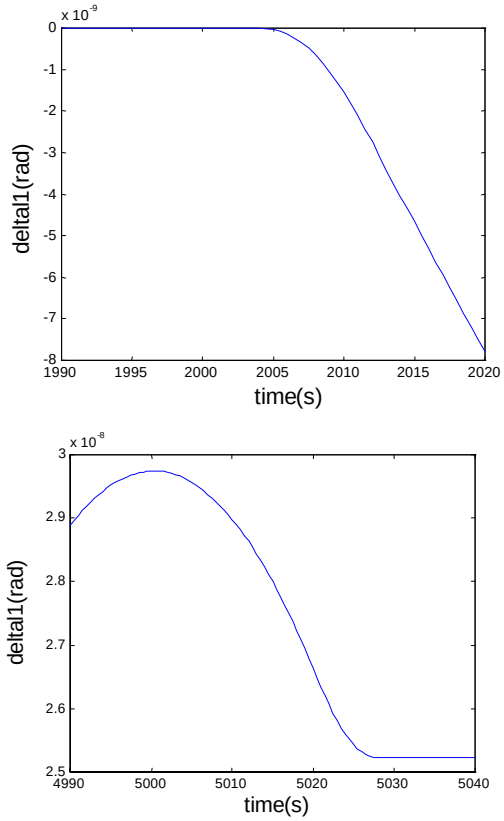


Figura 6: Influence of the Earth's shadow on the angular variable l_1, during the passage by two initial penumbra regions (delta $l_1 = l_1 - l_{10}$).

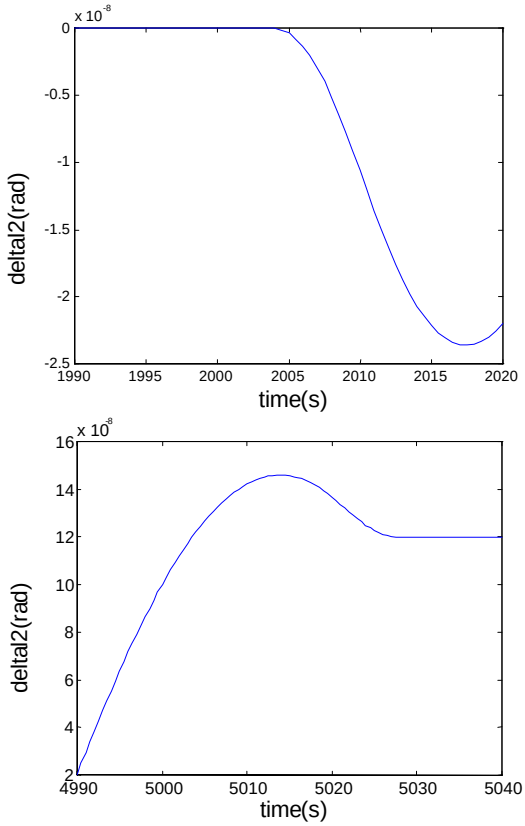


Figura 7: Influence of the Earth's shadow on the angular variable l_2, during the passage by two initial penumbra regions (delta $l_2 = l_2 - l_{20} - (L_{20}/A)\, t$).

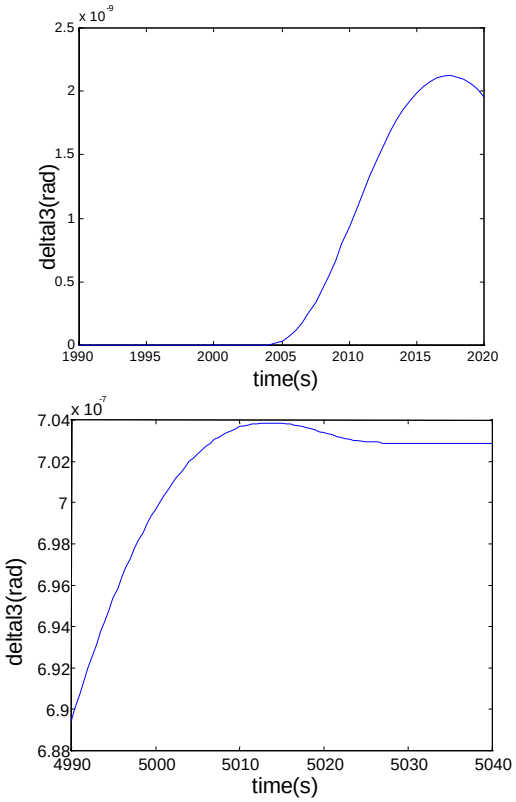


Figura 8: Influence of the Earth's shadow on the angular variable l_3, during the passage by two initial penumbra regions (delta l_3 = l_3 - l_{30}).

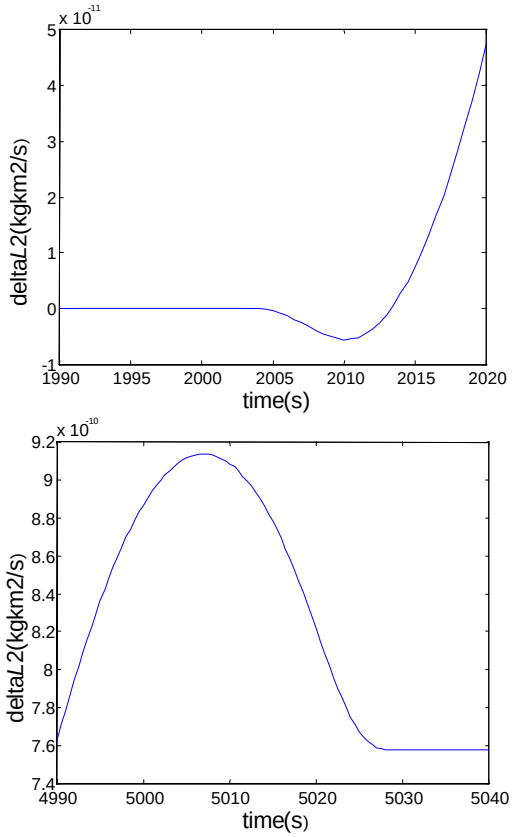


Figura 9: Influence of the Earth's shadow on the metric variable L_2, during the passage by two initial penumbra regions (delta L_2 = L_2 - L_{20}).

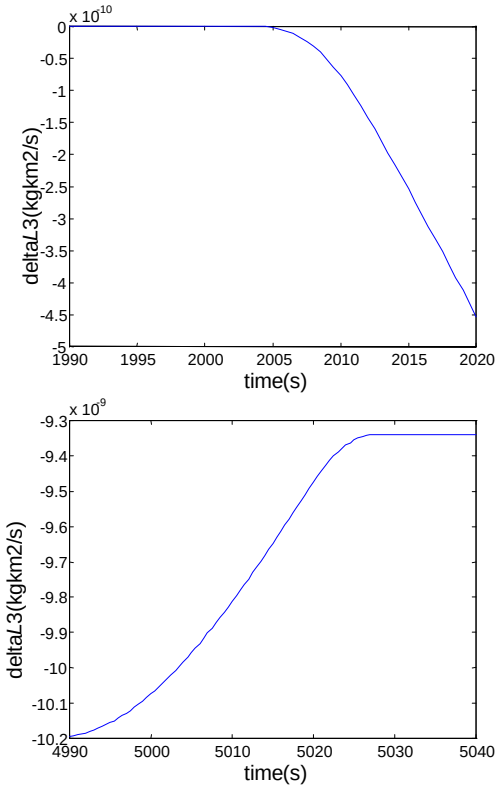


Figura 10: Influence of the Earth's shadow on the metric variable L_3, during the passage by two initial penumbra regions (delta L_3 = L_3 - L_{30}).

5. Conclusion

A semi-analytical procedure was proposed to predict the satellite's attitude under the influence of Earth's shadow over the solar radiation torque, including the penumbra region. In this procedure the mapping of the shadow's function is constructed, using the Kabelac's model (Kabelac, 1988), and a known analytical solution, valid when the satellite is illuminated, is used in order to propagate the attitude. The results shows that the transition from the shadow region to the illuminated region was attenuated because the penumbra region is including in the shadow function. By the results presented, the influence of the solar radiation torque is about 10^{-8} rad for the variable l_1, 10^{-6} rad for the l_2 and l_3, 10^{-10} kgkm^2/s for the L_2 and 10^{-7} kgkm^2/s for the L_3. Here the shadow function was described by the Kabelac's model and it was important in order to get a more real situation for the rotational motion of the satellite.

References

Cabette, R. E. S., 2001, "A Sombra da Terra e a Propagação de Atitude de Satélites Artificiais", Master Thesis, UNESP-FEG, Guaratinguetá, SP, Brazil.

Ferraz Mello, S., 1972, "Analytical Study of the Earth's Shadowing Effects on Satellite Orbits", Celestial Mechanics, v. 5, p. 80-101.

Kabelac, J., 1988, "Shadow Function - Contribution to the Theory of the Motion of Artificial Satellites", Bull. Astron. Inst. Czechosl., p. 213.

Lum, K. Y., Bloch, A. M., 1999, "Generalized Serret-Andoyer Transformation and Applications for the Controlled Rigid Body", Dynamics and Control, v. 9, p. 39-66.

Vilhena de Moraes, R., Zanardi, M.C., 1997, "Influence of Earth's Shadow on the Rotational Motion of Artificial Satellite Perturbed by Solar Radiation Torque", *Cospar*, v. 19, n.11, p. 1715-1718.

Zanardi, M. C.; Vilhena de Moraes, R., 1999, "Analytical and Semi-Analytical Analysis of an Artificial Satellite's Rotational Motion", Celestial Mechanics and Dynamical Astronomy, v. 75, p.227-250.

Zanardi, M. C., Vilhena, R. M., 1994, "Effects of Solar Radiation Torque on Satellite Spin and Attitude", *J. of the Braz. Soc. Mechanical Sciences*, v. XVI, p. 532.

XVI CONGRESSO BRASILEIRO DE ENGENHARIA MECÂNICA
16th BRAZILIAN CONGRESS OF MECHANICAL ENGINEERING

ON THE DYNAMICS OF AN ARMATURE CONTROLLED SERIES MOTOR MOUNTED ON AN ELASTICALLY SUPPORTED FOUNDATION

M. C. Mattos
Departamento de Engenharia Mecânica, Universidade Federal do Espírito Santo - Vitória – ES
mcoelho@npd.ufes.br

R. Rocha
Departamento de Engenharia Mecânica, Universidade Federal do Espírito Santo - Vitória – ES

Abstract. *This paper presents some analysis on the dynamics of an unbalanced series direct current motor mounted on an elastically supported table through numerical simulations. The governing equations of motion are obtained from the Lagrange's equations. We consider that a non-ideal energy source supplies the energy to the system, and analyze the interactions between dynamics of the foundation and dynamics of the motor. Several authors, as one may see in the paper, worked with non-ideal problems and, particularly, with unbalanced motors mounted on elastic foundations. This permits to compare the overall dynamics of the series motor with that of the independently excited motor. One of the important characteristics of the series motor is its high torque at low rotation speed. This characteristic can be important to determine if the rotational motion of the motor will be captured, or not, by resonance of the foundation. Comments on this subject, as well as on the performance of the motor, are presented.*

Keywords: *Non-ideal systems, Nonlinear systems, Sommerfeld effect, Nonlinear vibrations*

1. Introduction

The study of the interactions between the several parts that compose a given nonlinear system is important to understand several phenomena observed in practical situations. In this work we are concerned with the interactions between the dynamics of a series DC motor and the dynamics of its foundation. This problem can be included in the study of non-ideal systems because the primary source of energy to excite the structure (foundation) is the kinetic energy of the motor. On the other hand, the dynamics of the motor can be altered by the dynamics of the foundation. Furthermore, the most of the electrical DC motors don't have a non-limited power supplier. Hence, we have the case in which the excitation is influenced by the response of the structure, that is, this system is non-ideal. The first kind of non-ideal problem to arise in the literature is the so-called Sommerfeld effect (Kononenko, 1969). Evan-Iwanowski (1976) and Dimentberg *et al* (1994) gave further contributions to this non-ideal problem. Nayfeh & Mook (1979) present a comprehensive and complete review of different approaches until 1979. Recently, Balthazar *et al* (1996-a) analyzed, with interesting details, the governing equations of a non-ideal problem, containing quadratic and cubic nonlinearities, for a cantilevered beam supporting a non-ideal energy source at its free end. Balthazar *et al* (1996-b), Balthazar *et al* (1996-c) and Wieczorek & Mook (1997) present some experimental results on this problem. The numerical simulation of a simplified dynamics of this problem, using an unbalanced rotor attached to an elastic nonlinear support with internal and external damping and driven by a non-ideal energy source were analyzed by Balthazar *et al* (1996-d). The referred works use experimental data to estimate the characteristic curves of the motor. Hence the torque generated by the motor is not known explicitly, but only implicitly, through curves torque versus rotation speed of the motor. However, this curve is not generally obtained when the interactions between the motor and the foundation take place. In this work we take into account the equations that govern the response of the motor and, then, make a coupling between the mechanical equations, that is, the equations that govern the response of the foundation, and the electrical ones. Taking explicitly in account the equations of the motor presents some advantages:

(a) The characteristic curve of the motor is easily included in the model of the problem.

(b) It's possible to preview the mechanical response of the system and the electrical one.

(c) If the motor have any kind of control, it's possible to consider it in the model easily.

Furthermore, we have considered another category of motor, a DC series motor, to analyze the interactions. Mattos (Mattos-1999) have studied the foundation-motor coupled dynamics for independently excited motor, so that it will be possible to make some comparisons between them. First, it is necessary to note that series motor presents very high torque if its rotation speed reaches to low values. Then, the response of the system closed to resonance frequencies can be quite different that for independently excited motors.

One of the reasons to understand the dynamics of this kind of system is related to stiffness of the machines and structures. In general, a reduction in weight of a mechanical system implies in a reduction of its stiffness, at least if there is no change in its material. It expected that mechanical systems will be more flexible as possible. To compensate this, active control will be applied more and more. However, as flexible is the system more interactions will become important to describe it completely.

2. Physical system

Figure 01 shows a simplified sketch of the system, which consists of elastically supported table (foundation) and an unbalanced motor. The physical parameters indicated in that figure represent:

- k_1 and k_2 : linear and nonlinear stiffness elements, respectively.
- z and θ : translational and rotational coordinates of the displacements of the motor.
- M, J and m : mass of the table + motor set, moment of inertia of the rotor and unbalancing mass, respectively.
- c and c_θ : damping acting at the translational and rotational coordinates, respectively.
- r : distance between the unbalancing mass and the shaft of the motor

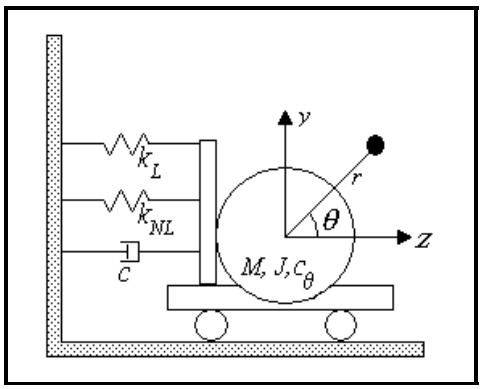


Figure 01 – Physical System

A conventional analysis of the system is presented in several books that cover mechanical vibration problems, for example, Thompson (1986). According to the conventional model, there is no coupling between the vibration of the foundation and the rotational motion of the rotate machine on that. In this work we take into account this coupling. Furthermore, we have considered that the machine is an electrically driven series motor, which electrical characteristics are also taken into account. Figure 02 shows the scheme of an armature controlled series DC motor. In this case the current in the armature and field circuits are the same. This kind of motor presents high torque at low rotation speeds. Since the first natural frequencies of several structures are very low, the characteristics of this motor may show something different that the independently excited motor.

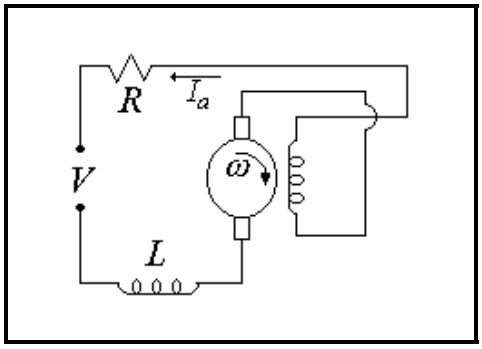


Figure 01 – Series DC motor electrical scheme

3. Obtaining the governing equations of the system

In order to obtain the governing equations of the motion of the motion of the foundation we have applied Lagrange's equations. The first step is obtain the kinetic (T) and potential (W) energies of the system, which can be written as:

$$W = \frac{K_L z^2}{2} + \frac{K_{NL} z^4}{4} + mgr\sin\theta \tag{1}$$

$$T = \frac{M}{2}\left(\frac{dz}{dt}\right)^2 + \frac{J}{2}\left(\frac{d\theta}{dt}\right)^2 + \frac{mv^2}{2} \tag{2}$$

where *v* is the velocity of the mass *m*. But it is easy to show that

$$v^2 = \left(\frac{dz}{dt} - r\frac{d}{dt}\sin \quad \right)^2 + \left(r\frac{d}{dt}\cos \quad \right)^2 \tag{3}$$

Now we can apply Lagrange's equation with respect to the coordinates z and . Then we have

$$\frac{d}{dt}\left(\frac{\partial T}{\partial z'}\right) - \frac{\partial T}{\partial z} + \frac{\partial W}{\partial z} = -cz' \text{ and } \frac{d}{dt}\left(\frac{\partial T}{\partial \ '}\right) - \frac{\partial T}{\partial} + \frac{\partial W}{\partial} = -c \quad ' + T_I - T(t) \tag{4}$$

where $z' = \frac{dz}{dt}$, T_I is the torque induced by the motor over its shaft and $T(t)$ is the load torque.

Applying eq. 4 to our system we obtain

$$(M+m)\frac{d^2z}{dt^2} + c\frac{dz}{dt} + k_L z + k_{NL} z^3 - mr\frac{d^2}{dt^2}\sin \quad - mr\left(\frac{d}{dt}\right)^2 cos \quad = 0 \tag{5}$$

$$\left(J + mr^2\right)\frac{d^2}{dt^2} + c \quad \frac{d}{dt} - mr\frac{d^2z}{dt^2}\sin \quad + mgr\cos \quad = T_I - T(t) \tag{6}$$

According to the theory for DC series motor (Chapman-1994), the governing equation for the electric circuit of the motor is

$$L\frac{dI_a}{dt} + R_a I_a + e_b = V(t) \tag{7}$$

where e_b is the fem between the terminals of the armature of the motor. This is called counter electromotive force and can be written as $e_b = k_b I_a \omega$ **.** The torque generated by the motor is written as (Chapman-1994),

$$T_I = k_t I_a^2 \tag{8}$$

Equations 7 and 8 show first difference between the series and independently excited motor, since the torque generated by the field circuit over the armature cannot be considered as a linear function of the current. Another nonlinearity is related to the counter electromotive force, which is not a linear function, but a product between the current and the rotation speed of the motor.

To obtain the governing equations of the system in their dimensionless form we define the non-dimensional time as $= {}_0 t$ and the non-dimensional displacement as $x = z/r$. In order to set a standard state from which the system is described we have taken a system without interactions between the rotational and translational motion, and excited by a voltage that leads the motor to rotate at the resonance frequency of the foundation. Then we define the new voltage, current and rotation speed as a ratio between themselves and that obtained for the standard state. Furthermore, we have defined several ratios between the parameters of the system in order to describe it overall in terms of non-dimensional parameters.

Equation 9 presents the definition of a set of relations between the parameters of the system.

$$\left\{ \begin{array}{lll} \omega_0 = \sqrt{\frac{k_1}{M+m}} & \frac{c}{M+m} = 2\xi\omega_0 & \frac{c}{J+mr^2} = 2\xi \ \omega_0 \\ v = \sqrt{\frac{k_t}{c \ \omega_0}} \cdot \frac{1}{R + k_b\omega_0} V & i = \sqrt{\frac{k_t}{c \ \omega_0}} \ I_a & \hat{T} = \frac{T(t)}{\left(J+mr^2\right)\omega_0^2} \\ \frac{m}{M+m} = \mu & \frac{mr^2}{J+mr^2} = \varepsilon & \frac{R_a}{L_a\omega_0} = \rho \\ \frac{g}{\omega_0^2 r} = \kappa & \frac{k_{NL} r^2}{k_L} = \sigma & \frac{k_b k_t}{R_a c} = \lambda \end{array} \right\} \tag{9}$$

Hence we obtain

$$\frac{d^2x}{d\tau^2} + 2\varsigma\frac{dx}{d\tau} + x + \sigma x^3 - \mu\left[\frac{d^2\theta}{d\tau^2}\sin\theta + \left(\frac{d\theta}{d\tau}\right)^2\cos\theta\right] = 0 \tag{10}$$

$$\frac{d^2\theta}{d\tau^2} + 2\varsigma_\theta\frac{d\theta}{d\tau} - \varepsilon\left[\frac{d^2x}{d\tau^2}\sin\theta - \kappa\cos\theta\right] = 2\varsigma_\theta i^2 - \hat{T} \tag{11}$$

$$\frac{di_a}{d\tau} + \rho i_a + \rho\cdot\lambda\frac{d\theta}{d\tau} = \rho(1+\lambda)v \tag{12}$$

To integrate the equations (10) to (12) in a safer way we should manipulate them. Note that both equations (06) and (07) have explicitly the second derivative of x and φ. This generates an algebraic loop during the integration of the equations and the results can be expurious. After some manipulations we have

$$\begin{aligned}&(1-\mu\varepsilon\sin^2\theta)\frac{d^2x}{d\tau^2} + 2\varsigma\frac{dx}{d\tau} + x + \sigma x^3 + 2\varsigma_\theta\mu\frac{d\theta}{d\tau}\sin\theta + \frac{\mu\varepsilon\kappa}{2}\sin 2\theta \\ &- 2\mu\varsigma_\theta i\sin\theta + \mu m\sin\theta + \mu\left(\frac{d\theta}{d\tau}\right)^2\cos\theta = 0\end{aligned} \tag{13}$$

$$\begin{aligned}&(1-\mu\varepsilon\sin^2\theta)\frac{d^2\theta}{d\tau^2} + 2\varsigma_\theta\frac{d\theta}{d\tau} + 2\varsigma\varepsilon\frac{dx}{d\tau}\sin\theta + \varepsilon x\sin\theta - \varepsilon\sigma x^3\sin\theta - \\ &-\frac{\varepsilon\mu}{2}\left(\frac{d\theta}{d\tau}\right)^2\sin 2\theta + \varepsilon\kappa\cos\theta - 2\varsigma_\theta i^2 + m = 0\end{aligned} \tag{14}$$

$$\frac{di}{d\tau} + \rho i + \rho\cdot\lambda\frac{d\theta}{d\tau} = \rho(1+\lambda)v \tag{15}$$

Now it's easy to see that the coefficients of the highest derivative terms are not constant. Note that these coefficients could reach to zero if the product $\mu\varepsilon$ reaches to the unity. However, it's not possible in practice because both μ and ε are less than the unity.

Equations (13) to (15) were analyzed through numerical simulations. Note that the governing equations take in account the gravity field and the mechanical and electrical coupling, which can be important to explain some phenomena observed in practical situations.

4. Simulation results for series motor

To integrate equations (09)-(11) we applied a 5th order varying step Runge-Kutta integrator. The parameters we have set to our analysis are shown in table 01. Note that we have, for instance, neglected the influence of the gravity. However, we have considered three values for the nonlinear stiffness (null, negative and positive). Hence we can infer whether the response observed in the laboratory tests may be due to presence of nonlinear terms like this.

Table 1. Parameter values used for simulation.

Parameter	Value	Parameter	Value
ξ_1	8.0000e-2	λ	2.0000e+2
ξ_2	1.4958e-3	μ	2.8643e-3
σ	-1.0884e-3 | 0 | 1.0884e-3	ρ	1.5915e+0
κ	0	ε	3.235e-1

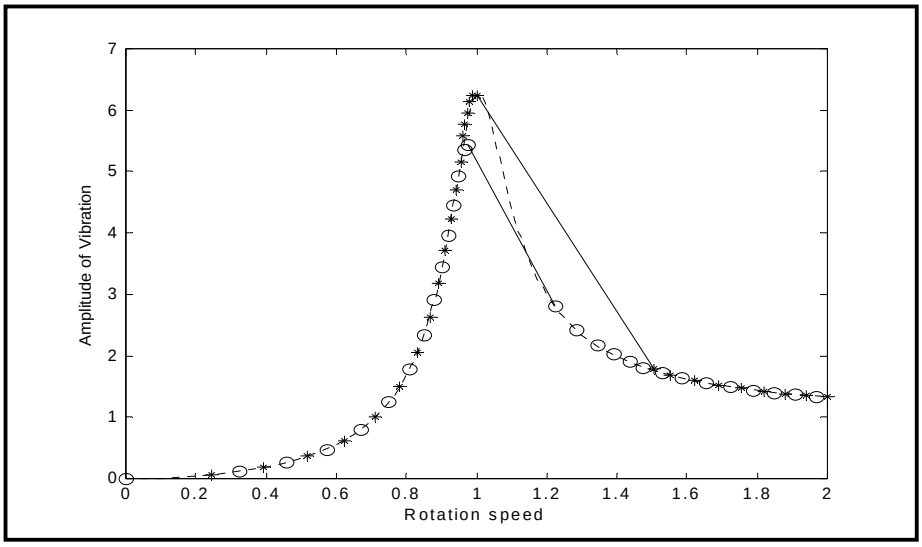


Figure 03 - Rotation speed vs. applied voltage. ∗: Increasing voltage o: Decreasing voltage

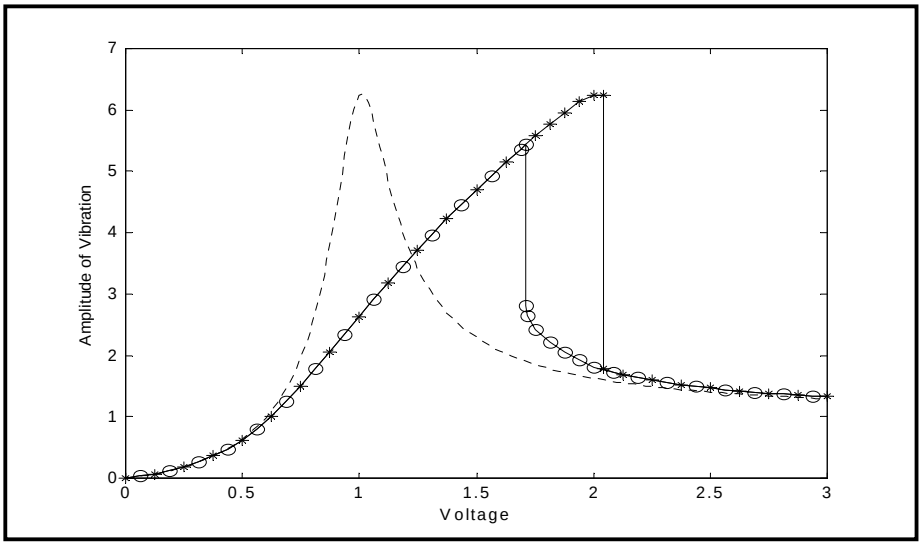


Figure 04 - Vibration Amplitude vs. applied voltage ∗: Increasing voltage o: Decreasing voltage

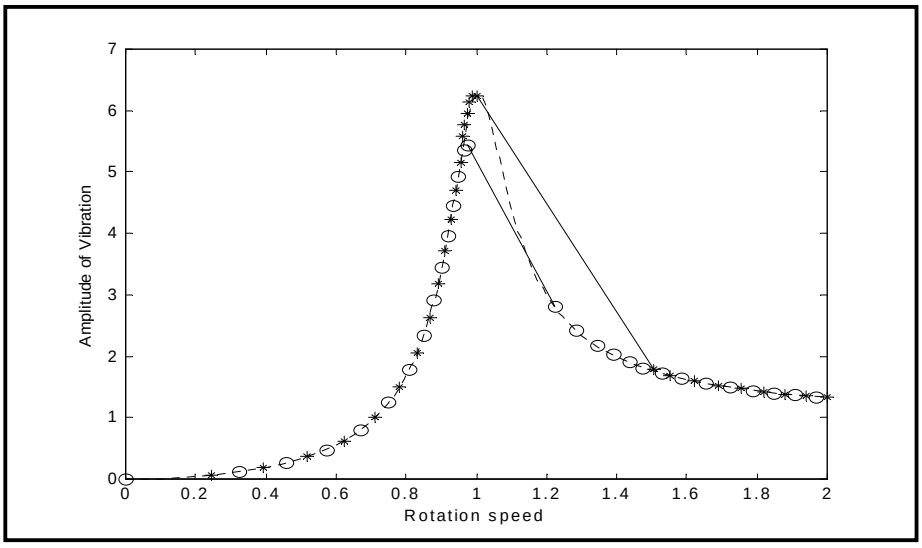


Figure 05 - Vibration amplitude vs. rotation speed. ∗: Increasing voltage o: Decreasing voltage

Figure 03 shows the rotation speed of the motor as a function of the applied voltage. The dashed line indicates the behavior of the motor without any interaction with the foundation. A deviation between the conventional behavior and that observed when foundation-motor interactions take place can be observed, manly for voltages between 1 and 2 volts. At approximately 2 volts, the rotation speed jumps from 0.95 to 1.4 (approximate values) for increasing voltage. For decreasing voltage the jump occurs at 1.2 volts, from 1.2 to 0.95 (approximate values). It can be also observed that, after the jump for increasing voltage, the rotation speed of the motor does not reach to the value expected for the system without interactions. Figure 04 shows jumps that occur in the amplitude of vibration of the foundation. We can note that, when the rotation speed jumps to a higher value, the amplitude of vibration jumps to a lower one, and vice versa. The dashed line in figure 04 is related to the response of the foundation if no interactions with the motor take place. Figure 05 shows the two jumps simultaneously in the frequency response of the system, where we can see a rotation speed range, between 1.0 and 1.2 approximately, in which the motion of the system is unstable. Note that there are not any point (* or o) in this range. Here again the dashed line is related to the frequency response of the foundation without interactions with the motor.

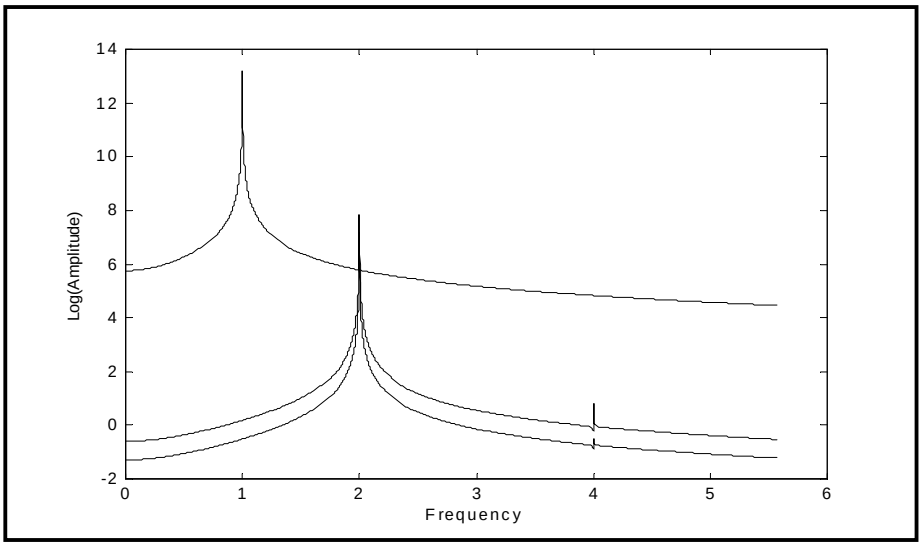


Figure 06 - Spectra for hardening cubic stiffness and null gravity when the motor rotates closed to the resonance vibration of the table, rotation of the shaft and demanded current (up to down)

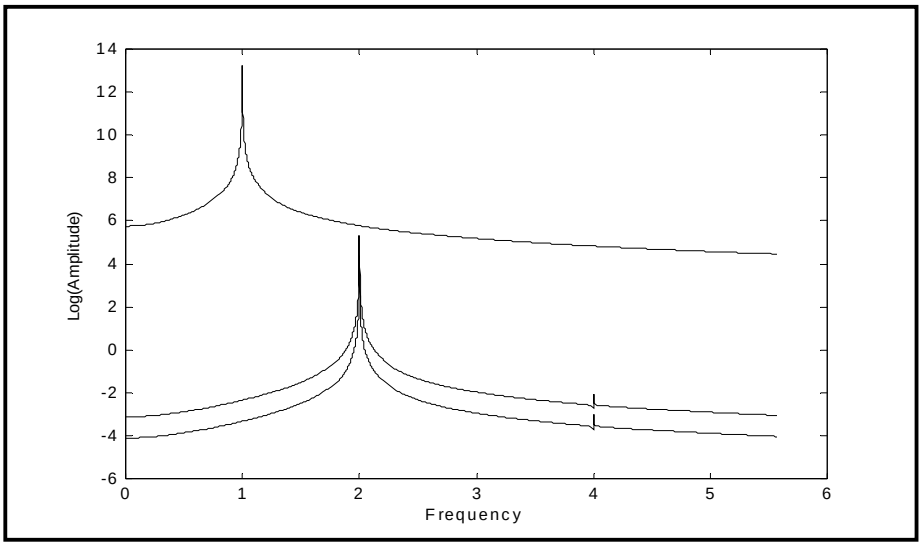


Figure 07 - Spectra for linear stiffness and null gravity when the motor rotates closed to the resonance vibration of the table, rotation of the shaft and demanded current (up to down)

Figure 06 shows spectra of the response of the system for vibration of the foundation, rotation of the shaft of the motor and current demanded by the motor. According to the conventional model for an unbalanced rotate machine on an elastically supported foundation the foundation oscillates, in the stationary state, in the same frequency of the machine. This model neglects the coupling between *"rotational dynamics"* and *"translational dynamics"*.

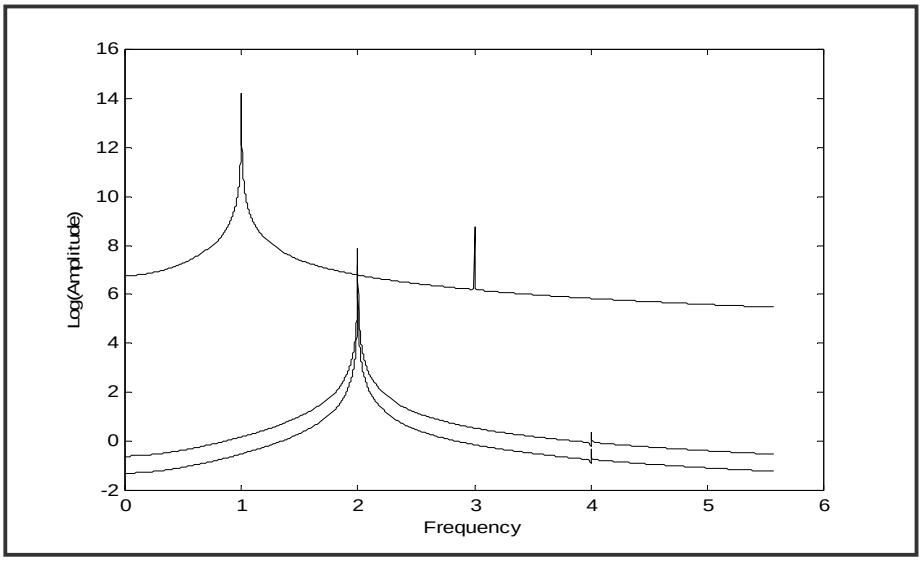


Figure 08 - Spectra for softening cubic stiffness and null gravity when the motor rotates closed to the resonance vibration of the table, rotation of the shaft and demanded current (up to down)

We can see that, as previewed by the conventional model, the response of the foundation is dominated by the excitation frequency, that is, the rotation of the motor. However, the spectra for the rotation speed and current indicates that there is no stationary state in this case, at least in the same way as that expected according to the conventional model, which indicates constant rotation speed and current in the stationary state. Note that both the rotation speed and current oscillate at a frequency that is twice the rotation speed of the motor (equal to the oscillation frequency of the foundation). In other words, if the mean value of the rotation speed of the motor is closed to natural frequency of the foundation, the rotation is not constant, but oscillates closed to the double of its mean value. The demanded current oscillates at the same frequency as the rotation speed of the motor. The results for linear stiffness and softening cubic stiffness are qualitative the same as that in fig. 06, as we can see in figs. 07 and 08.

4. A brief note on performance of the motor

Figure 09 presents the power consumption of the motor against the amplitude of vibration of foundation. Since the amplitude of the motion of the foundations increases much closed to the resonance, the ratio between the power available to accelerate the shaft and the load and the power demanded by the motor from the electrical source will decrease. This means that the performance of the motor will decrease during the passage by the resonance. This can be concluded also from fig. 03, since the rotation speed of the motor is less than the expected rotation speed if the motor was mounted on a very stiff structure.

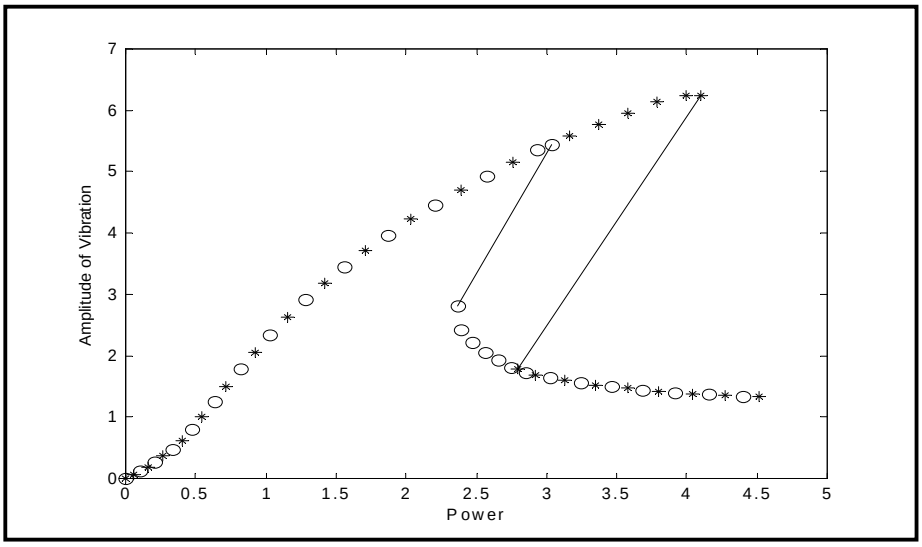


Figure 09 - Amplitude of vibration of the foundation against the power consumption of the motor.

Another aspect to be observed from the point of view of the performance is the variation of the rotation speed of the motor. Since the rotation speed is constant, an extra torque will be required to accelerate the motor. Since this extra torque is not the goal of the system, the performance will decrease. According to figs. 06-08, we can approximate the rotation speed of the motor and the current as

$$\eta = \eta_0 + \varepsilon_1 \sin(2t) + \varepsilon_2 \cos(2t) \tag{16}$$

$$i = i_0 + \alpha_1 \sin(2t) + \alpha_2 \cos(2t) \tag{17}$$

where $\eta = \frac{d\omega}{d\tau}$.

Since the load torque *T(t)* is equal to zero, the effective torque to be maintained by the motor is $2\varsigma_\theta \frac{d\theta}{d\tau}$. However, all the other terms in equation 11 should be considered. A complete analysis of the performance was carried out yet, but we have find that it is possible to manipulate the governing equations to obtain a closed form solution for a performance coefficient.

5. Concluding remarks

This works presents some advances in study of non-ideal systems. Particularly we cite:

- The interactions between the dynamics of the motor and the dynamics of the base induce a non-constant rotation speed of the motor. The oscillation of the rotation speed is equal to twice its mean value.
- The cubic term for nonlinear stiffness is responsible, as we could expect, for odd harmonics in the frequency response of the system.
- The current demanded by the motor varies with the same frequencies of the vibration of the base. We have not studied yet if there is a problem when this oscillation frequency is equal to the resonant frequency of the electrical circuit of the motor.
- The new presentation form of the governing equations of the systems permits us to analyze easily the dependence of the inertia terms with the rotation of the motor's shaft.
- The consideration of the motor's governing equations permits us to analyze what happens with the current in the motor's circuit. Consequently, we can describe the torque curve better in comparison with the conventional procedure, which imposes a particular characteristic for the torque-rotation curve. Note that we can also take easily in account the characteristic of load (external) curve.
- The variation of the performance of the system is commented. The performance of the motor decreases around the resonance frequency of the foundation.

6. References

Balthazar, J. M. et al - On Non-Ideal Flexible Resonance Crossing: Modeling and General Characteristics; Sixth Conference On Nonlinear Stability and Dynamics. 1996 (a).

Balthazar, J. M. et al - On Non-Ideal Euler-Bernouilli Cantilever Beam Supporting an Unbalanced Motor at its Tip End; Preprint. 1996 (b).

Balthazar, J. M. et al - Note on Non-Ideal Dynamical Systems; In Nonlinear Dynamics, Chaos, Control and Their Applications to Engineering Sciences. 1996 (c).

Balthazar, J. M. - Numerical Simulations on a Non-Ideal Dynamic System; Preprint. 1996 (d).

Chapman, R; Electrical Machines; J. Wiley and Sons; New York; 1994.

Dimentberg, M. et al - Passage through Critical Speed With Limited Power by Switching System Stiffness; Nonlinear and Stochastic Dynamics 192, DE- Vol 78, 57-67. 1994.

Evan-Iwanowski, R. M. - Resonance Oscillators in Mechanical Systems; Elsevier. 1976.

Kononenko, V. O - Vibrating Systems With a Limited Power Supply ; llife. 1969.

Mattos, M. C.; On the Dynamics of an Armature Controlled Series Motor Mounted on an Elastically Supported Foundation; XX COBEM; Águas de Lindóia – SP; 1999,

Nayfeh, A. H & D. T. Mook - Nonlinear Oscillations; John Wiley & Sons. 1979.

Thompson, W; Vibrações Mecânicas; J. Wiley and Sons; New York; 1986.

Wieczorek, S & D.T. Mook - Experimental Investigation of Transverse Vibration of a Cantilevered Beam Excited by a Non-Ideal Energy Source; V PACAM, Mayaguez; Puerto Rico. 1997.

XVI CONGRESSO BRASILEIRO DE ENGENHARIA MECÂNICA
16th BRAZILIAN CONGRESS OF MECHANICAL ENGINEERING

AN ANALYTICAL APPROACH FOR EVALUATION OF ORBITAL PERTURBATIONS DUE TO ATMOSPHERIC DRAG

Rodolpho Vilhena de Moraes
DMA-FEG-UNESP, Av. Ariberto P. da Cunha, 333, Guaratinguetá, SP, CEP: 12516-410
(rvm@feg.unesp.br)

Gustavo Willrich Nicoletti
FEG-UNESP (bolsista PIBIC/ CNPq), Av. Ariberto P. da Cunha, 333, Guaratinguetá, SP, CEP: 12516-410
(mec98039@feg.unesp.br)

Abstract: The equations describing the motion of artificial satellites under the influence of the atmospheric drag is non-linear and an analytical formal solution is proposed to solve them. Some models to describe the atmospheric density are presented and discussed comparatively. The presented models are compared with the numerical method CIRA, commonly used in space centers. It is shown that for some regions the analytical termospheric model TD88 agrees quite well with the CIRA model and it is an useful model for analytical solutions. The equations are expressed in canonical form. Secular perturbations are obtained and thus long and short period perturbations. Resonance conditions are analyzed.

Key words: orbital perturbations, atmospheric drag, atmospheric density

1. INTRODUCTION

The atmosphere exerts on an artificial satellite a force, called atmospheric drag, which is of the dissipative type. At high altitudes the ambient air is extremely rarefied, however, although of small magnitude, atmospheric drag is one of the dominant orbital perturbations for satellites orbiting below 700 km. The decreasing of the acceleration, due to the drag, is larger when the satellite is close to the perigee. The reduction of the speed of the satellite in the perigee results in a reduction of the height of the subsequent apogee. Nevertheless, the perigee height stays practically unaffected and the eccentricity of the orbit increases. Thus the orbit of the satellite contracts and become more nearly circular, and later spiral, causing the fall of the satellite. Mainly after the fundamental work done by Cook., King-Hele and Walker, (Cook et al.,1961) several authors, using different techniques, have been detailing the effects of the atmosphere in the motion of artificial satellites (Vilhena de Moraes, 1994).

2. ATMOSPHERIC DRAG

We call atmospheric drag force the component of the aerodynamic force parallel to the flow. This force is oriented along the velocity vector but in opposite direction. It is commonly adopted to represent the modulus of the atmospheric drag $\vec{D}$ force in the form (King-Hele,1964):

$$D = \frac{1}{2}\rho V^2 S C_D$$

Here ρ is the atmospheric density, V represents the magnitude of the satellite's velocity relative to the atmosphere itself. The combination $\frac{1}{2}\rho V^2$ is essentially a kinetic energy (per unit of volume) and is called the Dynamic Pressure. S is a reference area, that can be taken as the area perpendicular to the direction of motion. The dimensionless drag coefficient C_D is the drag coefficient which depends on the shape and structure of the satellite.

3. THE TERMOSFERIC MODEL TD-88

The atmospheric density depends essentially on the altitude above the Earth's sea level but also on several other factors such as temperature, time of day, season, latitude, and solar activity. In order to study the influence of the atmospheric drag in the motion of artificial satellites several analytic models have been proposed to describe the atmospheric density (Vilhena de Moraes, 1994).

The classical exponential density model, as treated by Brouwer and Hori in their theory about the motion of an artificial satellite (Brouwer and Hori, 1961), in spite of been extensively used in analytical theories, it do not take into account the above mentioned factors.

Due to dependence on several parameters, some of them with behavior apparently chaotic, it is almost impossible to obtain an actual model to describe completely the time variation of the atmospheric density. A comprehensive and didactic study about the properties of the upper atmosphere, including details of effects due to the solar influence, can be found in the papers published by Jacchia in the 60's (Jacchia 1963, a, b , c)

Generally, when realistic models are used for the atmospheric density, they hinder the analytic solution of the equations of motion. However, the TD 88 analytical model proposed by Sehnal (Sehnal, 1988) is extremely convenient for analytic developments. Such model, in the version of Pospíšilová (Sehnal and Pospíšilová, 1988), is characterized by a multiplication of functions depending upon several parameters such as solar flux, geomagnetic index Kp, latitude of satellite, day of the year and local time.

Thus, following the referred authors, the density can be expressed in the form:

$$\rho = f_x f_0 k_0 \sum_{n=1}^{7} h_n g_n \quad ,$$

where:

$$f_x = 1 + a(F_x - F_b)$$

$$k_0 = 1 + a_3 (K_p - 3)$$

$$f_n = \frac{(F_b - 60)}{160}$$

$$f_0 = a_2 + f_n$$

The dependence with the altitude are described by the following functions:

$$h_n = K_{n,j} + \sum_{j=1}^{3} K_{n,j} \exp\left(\frac{(120 - h)}{29j}\right).$$

The functions g are given by:

$$g_1 = 1$$

$$g_2 = \frac{f_n}{2} + a_4$$

$$g_3 = \sin(d - p)\sin\phi$$

$$g_4 = (a_5 f_n + 1)\sin(d - p_4)$$

$$g_5 = (a_6 f_n + 1)\sin 2(d - p_5)$$

$$g_6 = (a_7 f_n + 1)\sin(t - p_6)\cos\varphi$$

$$g_7 = (a_8 f_n + 1)\sin 2(t - p_7)\cos^2(\varphi),$$

were: $K_{n,j}$ and a_i are numerical constants (calculated), p_n are phases (calculated), F_x stands for the solar flow measured on 10,7cm per day, Fb is the solar flux averaged over three solar rotations, Kp is the local geomagnetic index three hours before the local time, h is the altitude of the satellite in km, d is the day count of the year, . φ is the latitude and t is the local time.

The TD 88 analytical model doesn't differ much from good current atmospheric models. As an example, Fig. (1), Fig. (2), Fig. (3), and Fig. (4). show, for different conditions, a comparison between values of density given by CIRA 86 and those obtained by the thermospheric model TD-88, and those obtained using the model proposed by Jacchia (Jacchia, 1977). The variation of density with the local time, for a particular condition, is shown in Fig. (6) using the density computed by the TD-88 model.

Particularly, the use of the TD-88 model in analytical developments leads to the same level of complexity as the Brower-Hori model and, as it was already pointed out, is more realistic.

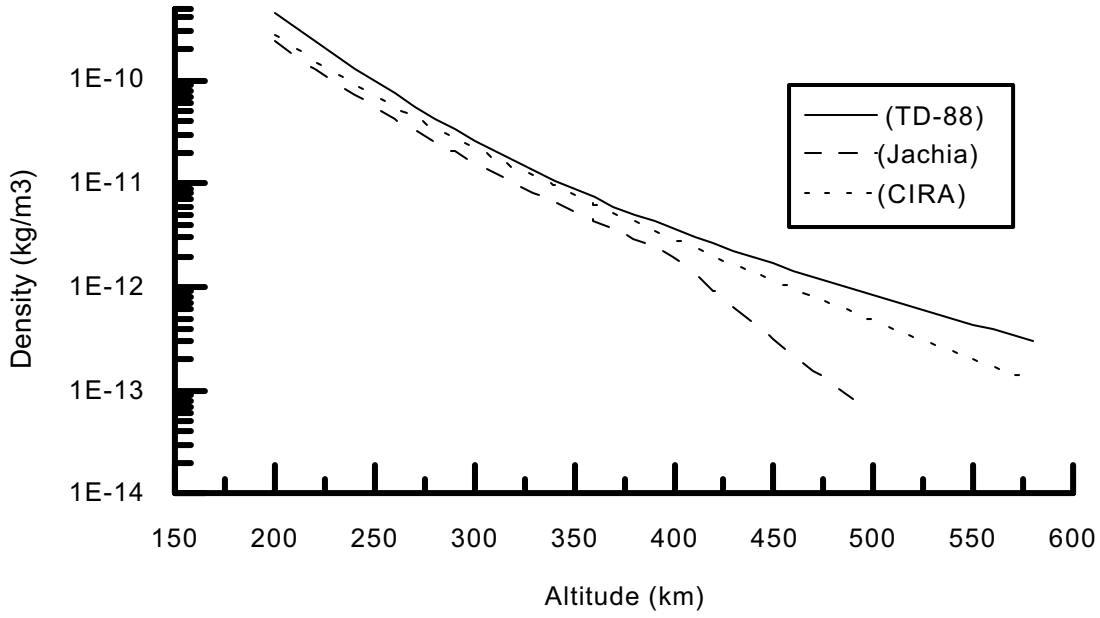


Figure 1. Density variations with altitude under the conditions: d=264, F_b= Fx=150, Kp=4, t=3, φ =0, T=900°

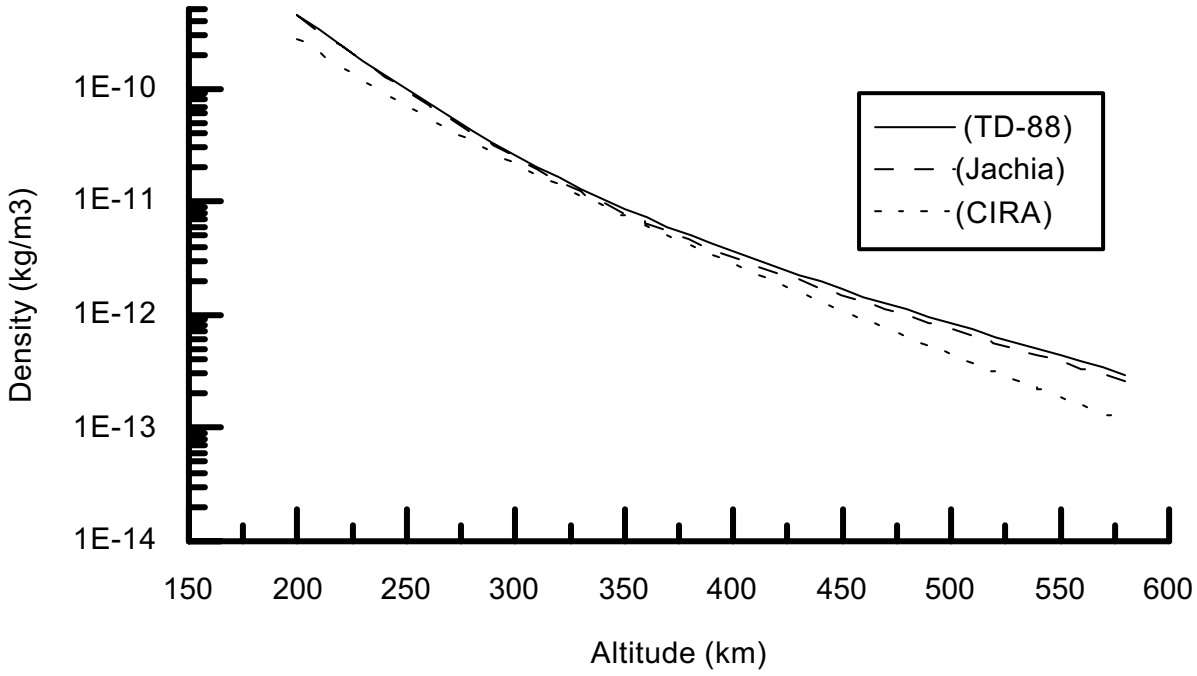


Figure 2. Density variation with altitude under the conditions: d=80, h=3, F_b= F_x=150, Kp=4, t=3, T = 902°

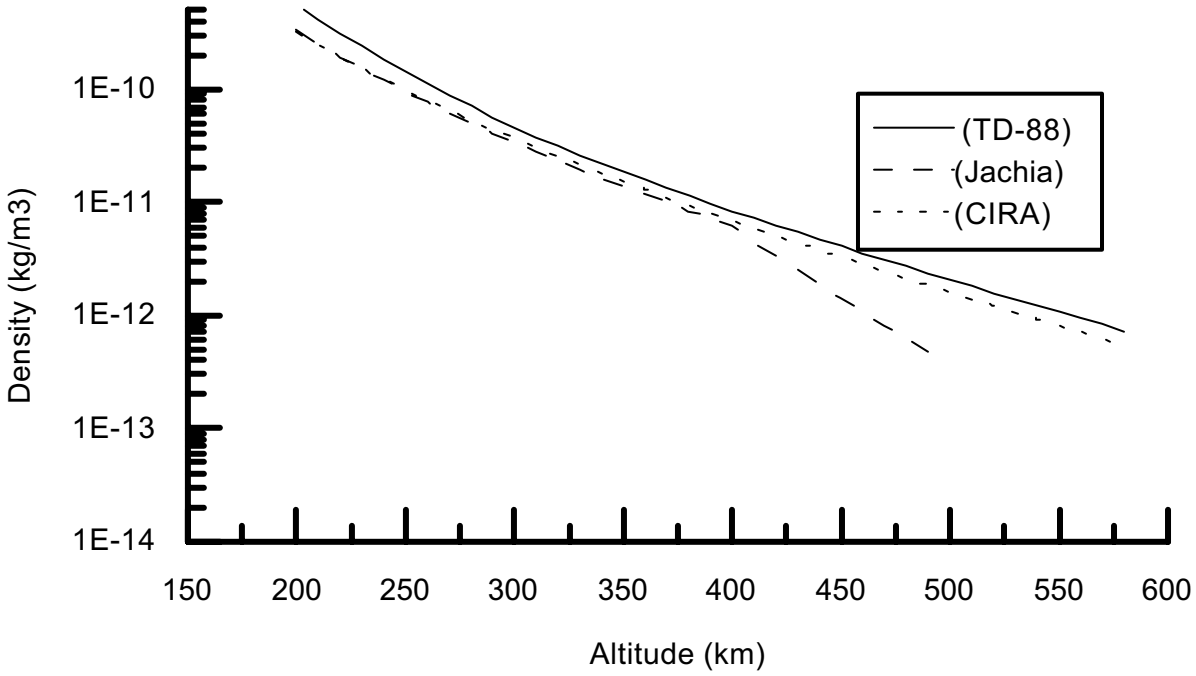


Figure 3. Density variations with altitude under the conditions: d=264, F_b= Fx=150, Kp=4, t=15, φ =0º, T= 1190°

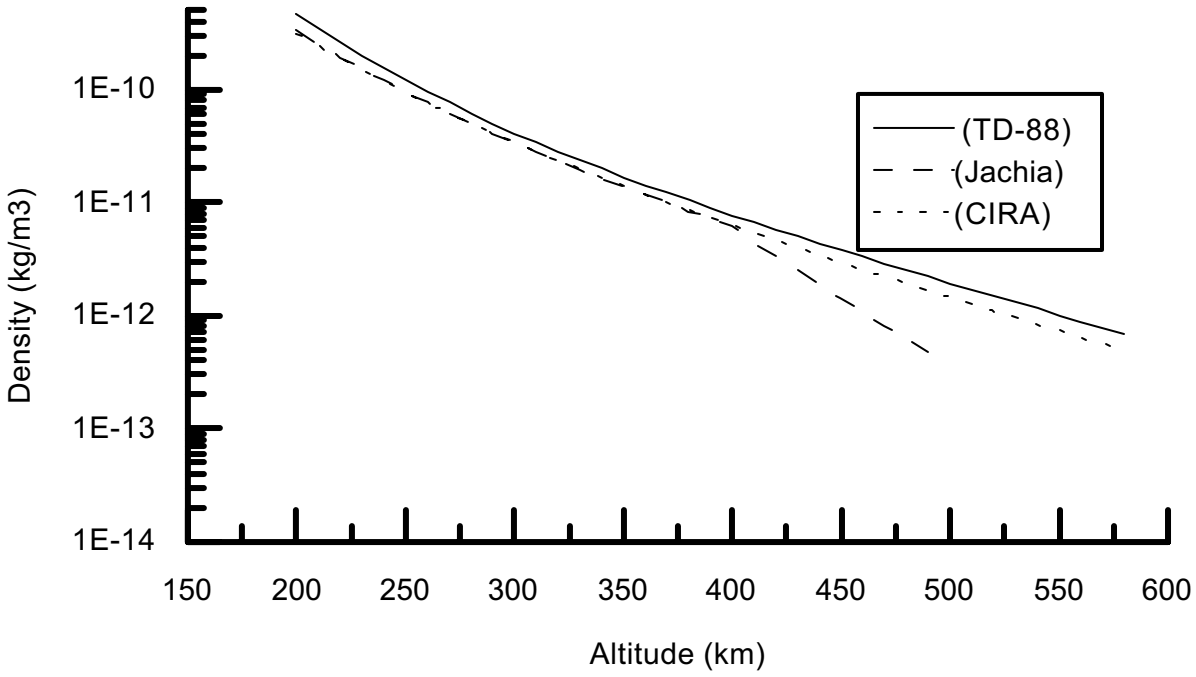


Figure 4. Density variations with altitude under the conditions : d =356, t=15, F_b= F_x=150, Kp=4, T= 1200

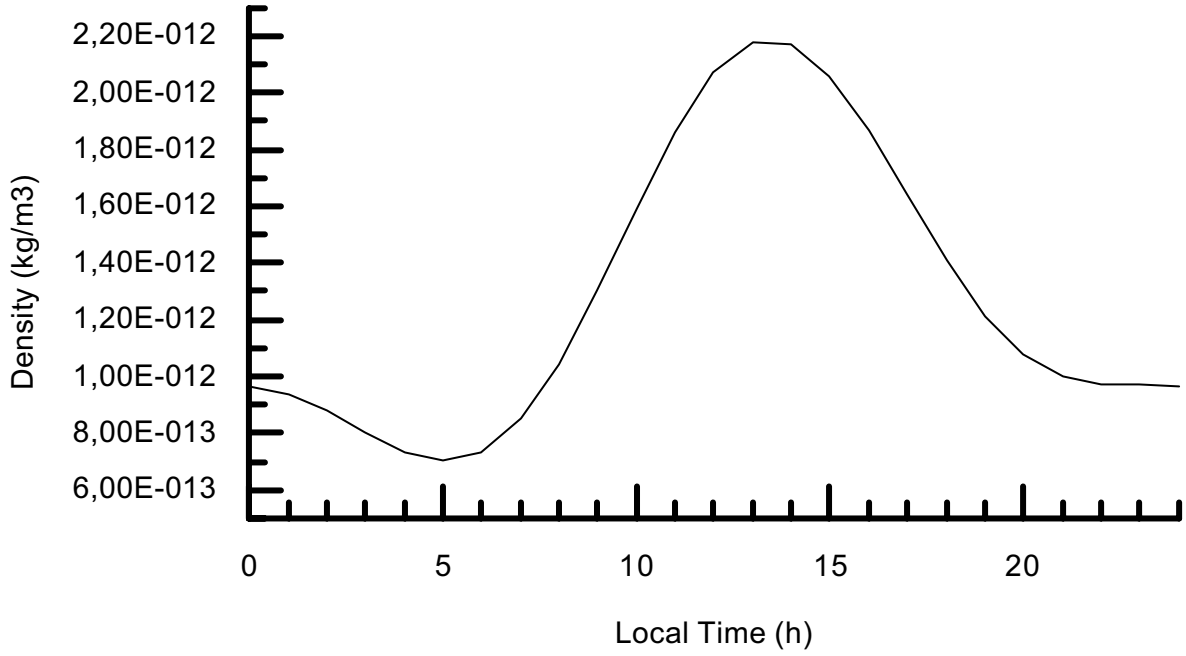


Figure 5. Variation of the density with the local time at 500 km of altitude, d=80, h=3, F_b= F_x=150, Kp=4.

4. EQUATIONS OF MOTION

Using canonical variables L_i and l_i (i =1,2,3) the equations describing the orbital motion of an artificial Earth satellite submitted to monogenic and polygenic forces can be expressed in the extended canonical form as

$$\frac{dL_i}{dt} = \frac{\partial F}{\partial l_i} + P_i$$

$$\frac{dl_i}{dt} = -\frac{\partial F}{\partial L_i} + Q_i$$

Here F is the Hamiltonian and P_i and Q_i, due to non-conservative forces, are given by (Brouwer and Hori, 1961):

$$P_i = \sum_k X_k \frac{\partial \dot{x}_i}{\partial l_k} \quad , \quad Q_i = \sum_k X_k \frac{\partial x_i}{\partial L_k} ,$$

where X_k represents the components of the non-conservative acceleration expressed in an equatorial Cartesian system with coordinates x_i, i= 1, 2, 3. Considering drag perturbations we have

$$X_k = -\frac{SC_D}{2m}\rho V(\dot{x}_k + \delta_k)$$

where δ_k is a parcel due to the contribution of the rotation velocity of the atmosphere

Neglecting the rotation of the atmosphere, P_i and Q_i can be expressed in terms of the canonical variables as follows

$$P_1 = -\frac{SC_D}{2m}\rho V L_1[(2a/r)-1] \quad ; \quad Q_1 = -\frac{SC_D}{2m}\rho V[2e(1-\frac{1}{8}e^2+...)\sin M+...) + \\ +\frac{2}{e}\frac{L_2}{L_1}(1-\frac{7}{8}e^2+...)\sin M+...]$$

$$P_2 = -\frac{SC_D}{2m}\rho V L_2 \quad ; \quad Q_2 = \frac{SC_D}{2m}\rho V\frac{2}{e}(1-\frac{7}{8}e^2+...)\sin M+...$$

$$P_3 = -\frac{SC_D}{2m}\rho V L_3 \quad ; \quad Q_3 = 0$$

Initially it will be considered here F containing only the contribution of the main secular terms due to the second zonal harmonic. So,

$$F = -\frac{\mu^2}{2L_1^2}+\frac{1}{4}\frac{\mu^4 J_2}{L_1^3 L_2^2}\left(3\frac{L_3^2}{L_2^2}-1\right)$$

where μ is the terrestrial gravitational constant and J_2 is a constant associated with the Earth flattening.

Using the TD88 model, and expressing the local time and the latitude φ in terms of the orbital elements and taking into account only secular terms we have (Vilhena de Moraes, 1989):

$$\rho_s = f_x f_0 k_0 \sum_{n=1}^{7} G_{n,0}\left\{K_{n,o}+\sum_{j=1}^{3}K_{n,j}A_j[(1+\frac{c_j^2}{4})F+A_{0j}]\right\}$$

where G_{no} , A_j , F and A_{nj} are functions of the metric variables.

The components P_i and Q_i can be expressed in terms of the orbital elements. If only secular terms are considered we have (Vilhena de Moraes, 1989, 1983, 1981):

$$P_i = -\frac{SC_D}{2m}\rho_s p_0^i \quad , \quad Q_i = 0,$$

and then,

$$\frac{dL_i}{dt} = -\frac{SC_D}{2m}f_x f_0 k_0 \sum_{n=1}^{7} G_{n,0}\left\{K_{n,o}+\sum_{j=1}^{3}K_{n,j}A_j[(1+\frac{c_j^2}{4})A_{0j}]\right\}p_0^i$$

$$\frac{dl_i}{dt} = N_i$$

where $N_i = \dfrac{\partial F}{\partial L_i}$

Considering J_2 as a small parameter and the drag perturbations as a second order effect, the secular perturbations due to the drag can be easily found solving this system by the method of the successive approximations. In fact, since up to zero order $L_i = L_{i0}$ = constant and $l_i = n_i t + l_{io}$, where l_{io}, are constants, up to to the first order we have:

$$L_i = A_i t + C_i$$

where A_i and C_i are constants. Up to first order, there exists no drag influence on the angular variables l_i.

In order to compute short and long periodic terms, firstly we must develop the P_i and Q_i in the neighborhood of the drag free problem, that is

$$P_i = \sum_{j=1}^{3} \left(\frac{\partial P_i}{\partial L_j} \right)^* \delta L_j + \sum_{j=1}^{3} \left(\frac{\partial P_i}{\partial l_j} \right)^* \delta l_j$$

$$Q_i = \sum_{j=1}^{3} \left(\frac{\partial Q_i}{\partial L_j} \right)^* \delta L_j + \sum_{j=1}^{3} \left(\frac{\partial Q_i}{\partial l_j} \right)^* \delta l_j$$

and the resulting expressions can be put in the following form

$$P_i = -\frac{SC_d}{2m} f_x f_0 k_0 \sum_j p_j^i (L_{10}, L_{20}, L_{30}) \cos(\alpha_j l_1 + \beta_j l_2 + \gamma_j l_3)$$

$$Q_i = -\frac{SC_d}{2m} f_x f_0 k_0 \sum_j q_j^i (L_{10}, L_{20}, L_{30}) \sin(\alpha_j l_1 + \beta_j l_2 + \gamma_j l_3)$$

Since the solution for the drag free problem is known, introducing these expressions in the equations of motion, the resulting system can be easily solved by any perturbation method.

5. CONCLUSIONS

Looking at Fig. (1) to Fig.(5) it is easy to recognize that there is no substantial difference between the values of the density found by the models TD-88, CIRA 86 and Jacchia 77.

The TD-88 atmospheric model proposed to describe analytically the density atmospheric profile enables an analytical solution for the equations of motion of an artificial Earth satellites under the influence of the atmospheric drag.

6. ACKNOWLEDGEMENTS

This work has been partially supported by the Brazilian National Council of Research (CNPq).

7. REFERENCES

Brouwer, D. and Hori, G. I. , 1961. “Theoretical Evaluation of Atmospheric Drag Effects in the Motion of an Artificial Satellite”, Astron. J. 66, 193.

Cook, G. E., King-Hele, D., Walker, D. M., 1961 “The Contraction of Satellite Orbits under Influence of Air Drag”, Proc. of Royal Soc., 264, 88-121.

Jacchia, Luigi G. , 1975, "The Earth's Upper Atmosphere - I", Sky and Telescope, March, 155-159.

Jacchia, Luigi G. , 1975, "The Earth's Upper Atmosphere - II", Sky and Telescope, April, 229, 232

Jacchia, Luigi G. , 1975, "The Earth's Upper Atmosphere - III", Sky and Telescope, May, 294, 299.

Jacchia, Luigi G., 1977, "Thermospheric Temperature, Density, and Composition: New Models", SpeciaL Report.

King-Hele, D., 1964, “Theory of Satellite Orbits in na Atmosphere”, Butterwords, London.

Sehnal, L., 1988, “Termospheric Model TD 88”, Bull. Astronon-Inst. Czechosl., 39,120-127.

Sehnal, L. and Pospísilová, 1988 “Termospheric Model TD 88”, Pre-print. Nº. 67, Astronon-Inst. Czechosl. Acad. Sci., Ondrojov..

Vilhena de Moraes, R. ,1981, "Combined Solar Radiation Pressure and Drag Effects on the Orbits of Artificial Satellite", Cel. Mech., 25, 281-292.

Vilhena de Moraes, R. and Fitzgibbon, M. T. R., 1983, "Comment on Coupled Perturbations on Orbits of Artificial Satellites", The Motion of Planets and Natural and Artificial Satellites, S. F. Mello and P. Nacozy (eds), Universidade de São Paulo, 175-182.

Vilhena de Moraes, R., 1989, "Analytical Computation of Drag Perturbations Using Thermospheric Model TD-88", Orbital Dynamics of Natural and Artificial Objects, R. Vieira Martins, D. Lazaro and W. Sessin ((eds.)), Obs. Naciona, Rio de Janeiro, 135-144.

Vilhena de Moraes, R, 1994, “Non Gravitational Disturbing Forces”, Adv. Space. Res., Vol. 14, nº 5, (5)45-(5)68.

XVI CONGRESSO BRASILEIRO DE ENGENHARIA MECÂNICA
16th BRAZILIAN CONGRESS OF MECHANICAL ENGINEERING

SEMI-ANALYTICAL DETERMINATION OF HALO ORBITS VIA THE LINDSTEDT-POINCARÉ METHOD

Cristian Beauge
Universidade Nacional de Córdoba
Laprida 854, (5000), Córdoba, Argentina
E-mail: BEAUGE@DEM.INPE.BR

Gislaine de Felipe
Instituto Nacional de Pesquisas Espaciais - INPE
São José dos Campos - SP - 12227-010 - Brasil
Phone (123)256197 - Fax (123)25-6226
E-mail: GISLAINE@DEM.INPE.BR

Antonio Fernando Bertachini de Almeida Prado
Instituto Nacional de Pesquisas Espaciais - INPE
São José dos Campos - SP - 12227-010 - Brasil
Phone (123)256197 - Fax (123)25-6226
E-mail: PRADO@DEM.INPE.BR

Abstract. *This paper presents results of the application of the Lindstedt-Poincaré method for the determination of three-dimensional periodic orbits in the vicinity of the collinear Lagrange points in the restricted circular three-body problem. These solutions, known in the literature as "Halo" orbits, are of great practical importance and have been used as reference trajectories for several artificial satellites and probes launched in the last 20 years. The principal aim of this communication is to introduce a simple and straightforward analytical method for the computation of small-to-moderate amplitude solutions, together with an innovative approach for extending this approach to any amplitude, no matter how large. Examples of Halo orbits are shown in the case of the Earth-Moon system, and the results compared with numerical integrations of the exact differential equations.*

Key Words: Astrodynamics, Halo Orbits, Orbit Determination.

1. Introduction

Consider the circular restricted three-body problem in a rotating barycentric coordinate system *(X,Y,Z)*, such that both massive bodies (i.e. $m_1 > m_2 > 0$) are always located in the *X* axis, and the *XY* plane coincides with their plane of motion. In this scenario we find five equilibrium solutions, the so-called Lagrange points, which are denoted by L_i ($i=1,...,5$). Two of them (L_4 and L_5) are located in the *XY* plane forming equilateral triangles with the primaries. The remaining points all lie along the *X* axis: L_1 between both massive bodies, L_2 beyond and L_3 between m_1 and minus infinity. For values of m_2/m_1 smaller than a critical value, the equilateral points are stable. All three "collinear" points, on the other hand, are unstable for all values of the masses. They are usually referred to as fixed points of type *Center* x *Saddle* x *Center*; unstable (i.e. *Center* x *Saddle*) for motions in the plane, and stable (i.e. *Center*) in the *Z* axis. Three-dimensional periodic orbits around these unstable points, also unstable by nature, are called *Halo orbits*.

It seems at first hand that such trajectories are among the least desirable places in which to place an artificial satellite. Nevertheless, since the beginning of the 1970s serious studies began to be undertaken on this regard. Possibly the first was due to Farquhar. When NASA was still planning possible scenarios for the exploration of the Moon, serious consideration began to be placed on the problem of communication between the Earth and a hypothetical exploratory base located in the far side of the Moon. It is known that, due to the spin-orbit resonance between the Moon's rotational motion and its translation around the center of mass, the far side is always invisible from our planet, thus making impossible any contact with ground control. This question was analyzed, among others, by Farquhar (1970), Farquhar and Kamel (1973) and Richardson and Kary (1975). As a solution, they proposed the localization of a communication satellite in a Halo orbit around the L_2 Lagrange solution. Due to its periodicity and particular shape, the trajectory would avoid all periodic occultations by the Moon, assuring permanent contact with the astronauts. The question of the intrinsic instability of the orbit was solved proposing periodic station keeping maneuvers whenever the true trajectory departed sufficiently from the nominal solution.

Although the end of the *Apollo* project appeared as a serious blow to these projects, soon other applications of halo orbits were proposed and made reality. Examples are the satellites developed for the study of Solar wind in the vicinity of our planet. Solar wind consists of atoms and ions that are driven by the Sun. Studying their composition, astronomers hope to obtain important data to estimate the chemical composition of the Solar System at the time of its formation and, thus, are of great importance for models of planetary formation. However, there exist serious difficulties in the detection and analysis of Solar wind. Since it is made up of electrically charged particles, it is

deflected by the terrestrial magnetosphere and, consequently, cannot be studied from the vicinity of our planet. The only solution is to place a satellite outside the "bow shock".

Of all the possible orbits for such a satellite, the best suited is precisely a Halo orbit around the L_1 point of the Sun-Earth system. The main advantage over normal heliocentric orbits is the fact that it is always between the Sun and our planet, thus allowing a permanent observation of both bodies without eclipses of any sort. This type of orbit was used in 1978 by the probe ISEE-3 (*International Sun-Earth Explorer-3*), where it remained for several years. Almost twenty years later, the Solar wind in the L_1 point was once again monitored by the SOHO (*Solar & Heliospheric Observatory*), launched in 1995 and, more recently, by the ACE (*Advanced Composition Explorer*) launched in 1997, and still in operation. Probably the most ambitious projects are yet to come. The GENESIS probe (scheduled for 2002) will not only continue the observation of the Sun but will also collect samples of Solar wind and, after a three-year mission, bring them back to Earth. Thus it represents the first program, since Apollo, to bring extraterrestrial material to Earth.

In the present paper we review the determination of Halo orbits in the circular restricted three-body problem via a classical analytical method known as Lindstedt's Device or the Lindstedt-Poincaré method. Its main advantage over other analytical approaches lies in its simplicity and high precision. An important drawback, however, is the fact that it is only convergent for low to moderate amplitudes of the periodic orbits. We will therefore briefly introduce a modification of the equations of motion that allows us to overcome this limitation. The manuscript is organized as follows: Section 2 presents the equations of motion in the vicinity of a Lagrange fixed point, together with Richardson's expansion of the pseudo-potential terms. These equations form the basis for the application of the analytical method. Lindstedt's device is reviewed in Section 3, and explicit formula for Halo orbits are presented as well. Section 4 discusses the extension of the method for large-amplitude orbits. Section 5 shows examples of Halo orbits in the Earth-Moon and Sun-Earth systems, while comparisons with exact numerical simulations follow in Section 6. Finally, conclusions close the paper in Section 7.

2. Orbits in the Vicinity of the Collinear Lagrange Points

Since we are basically interested in orbits in the vicinity of one of the collinear equilibrium points, let us begin by introducing a transformation of coordinates (X,Y,Z)→ (x,y,z) to new variables centered around L_i:

$$\begin{aligned} X &= -\gamma_i x + \mu + A \\ Y &= -\gamma_i y \\ Z &= \gamma_i z, \end{aligned} \tag{1}$$

where γ_i is the distance from L_i to the nearest primary, and $A = -1 + \gamma_1$ (for L_1), $A = -1 - \gamma_2$ (for L_2). In terms of these new variables, the equations of motion can be written as (see Szebehely, 1967):

$$\begin{aligned} \ddot{x} - 2\dot{y} &= x - \frac{(1-\mu)}{r_1^3}(x-\mu) - \frac{\mu}{r_2^3}(x+1-\mu) \\ \ddot{y} + 2\dot{x} &= y - \frac{(1-\mu)}{r_1^3} y - \frac{\mu}{r_2^3} y \\ \ddot{z} &= -\frac{(1-\mu)}{r_1^3} z - \frac{\mu}{r_2^3} z \end{aligned} \tag{2}$$

Here $\mu = m_2/(m_1+m_2)$, $r_1^2 = (x-\mu)^2 + y^2 + z^2$ and $r_2^2 = (x+1-\mu)^2 + y^2 + z^2$. In order to develop the nonlinear terms in the right-hand side, we expand them as function of Legendre polynomials P_n, defined by means of the expression:

$$\frac{1}{\sqrt{(x-a)^2 + (y-b)^2 + (z-c)^2}} = \frac{1}{D} \sum_{n=0}^{\infty} \left(\frac{\rho}{D}\right)^n P_n\left(\frac{ax+by+cz}{D\rho}\right) \tag{3}$$

with $D^2 = a^2 + b^2 + c^2$ and $\rho^2 = x^2 + y^2 + z^2$. After some straightforward calculations (see Richardson, 1980) the equations of motion for the massless particle can be written in power series of the coordinates as:

$$\begin{aligned}
\ddot{x} - 2\dot{y} - (1+2c_2)x &= \sum_{n\geq 2}(n+1)c_{n+1}T_n \\
\ddot{y} + 2\dot{x} + (c_2 - 1)y &= y\sum_{n\geq 2} c_{n+1}R_{n-1} \\
\ddot{z} + c_2 z &= z\sum_{n\geq 2} c_{n+1}R_{n-1}
\end{aligned} \tag{4}$$

The c_n coefficients are given by:

$$c_n(\mu) = \frac{1}{\gamma_i^3}\left((\pm 1)^n \mu + (-1)^n \frac{(1-\mu)\gamma_i^{n+1}}{(1 \mp \gamma_i)^{n+1}}\right) \tag{5}$$

where the upper sign is used for L_1 and the lower for L_2. Functions T_n, R_n can be obtained by the following recurrence relations:

$$R_n(x,y,z) = -\frac{2n+2}{n+2}T_n + \frac{2n+3}{n+2}xR_{n-1} - \frac{2n+1}{n+2}\rho^2 R_{n-2} \tag{6.1}$$

$$T_n(x,y,z) = \frac{2n-1}{n}xT_{n-1} - \frac{n-1}{n}\rho^2 T_{n-2} \tag{6.2}$$

with initial values $T_0 = 1$, $T_1 = x$ and $R_0 = 1$; $R_1 = -3x$.

Equations (4) constitute our final expression for the motion of a massless particle in the vicinity of the collinear libration point L_i. Note that the right-hand side contains only nonlinear terms and that the expansion (3) in Legendre polynomials is convergent only for values of $\rho < 1$; this then constitutes a limit for the possible values of the amplitudes of the Halo orbits determined via this approach.

3. The Lindstedt-Poincaré Method

Amongst the various perturbation techniques applied to the problem of halo orbits, probably the simplest and most straightforward is the Lindstedt's device. Contrary to other approaches, it does not require extensive knowledge of the Hamiltonian perturbation methods or the development/use of complex algebraic computer software. Although the sheer size of the formulas may seem uncomfortable at first, they are very easily understood and, once the basis of the method is grasped, can be reproduced and extended with very little trouble. The basic idea of the method is as follows: Starting from the linearized system (i.e. zeroing all the right-hand members of Eq. (4)), we can determine the linear solution as:

$$\begin{aligned}
x &= \alpha e^{i\omega_0 t} \\
y &= \kappa\alpha e^{i\omega_0 t} \\
z &= \beta e^{i\nu_0 t}
\end{aligned} \tag{7}$$

where ω_0, ν_0 are the linear planar and off-plane frequencies of motion and α, β are the amplitudes (in complex notation) in the *x* and *z* axis respectively. Finally, κ is a constant that relates the amplitude in the *y*-axis with that in the other planar dimension. Now, we know that the inclusion of nonlinear terms (i.e. right-hand side of (4)) will yield a solution that will contain contributions from both the linear planar motion as well as the linear motion in the *z* axis. Thus, we can write the general solution as a Fourier series of the type:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \sum_{i,j,k,l} \begin{pmatrix} X_{i,j,k,l} \\ Y_{i,j,k,l} \\ Z_{i,j,k,l} \end{pmatrix} \alpha^i \beta^j e^{\sqrt{-1}(k\omega + l\nu)t} \tag{8}$$

in terms of new complex coefficients $X_{i,j,k,l}$, $Y_{i,j,k,l}$, $Z_{i,j,k,l}$. We suppose similar expansions for the coefficients T_n, R_n in the right-hand sides. The values of the coefficients of Eq. (8) are obtained introducing these series into the original differential equations, identifying terms of equal harmonic and degree (in α and β) and solving the resulting algebraic equations. The new frequencies (also affected by the nonlinear terms) can be obtained as power series of the linear amplitudes α, β in the form:

$$\omega = \sum_{i,j} \omega_{i,j} \alpha^i \beta^j \tag{9.1}$$

$$\nu = \sum_{i,j} \nu_{i,j} \alpha^i \beta^j \tag{9.2}$$

where $\omega_{0,0} = \omega_0, \nu_{0,0} = \nu_0$.

The general quasi-periodic solutions, in which $\omega \neq \nu$, are usually referred to as *Lissajous* orbits and are of limited practical importance since they periodically cross the line of sight between both primaries. *Halo* orbits can be considered as a particular case of Lissajous trajectories, in the case of $\omega = \nu$. Unfortunately, periodic orbits, although present in the nonlinear system, have no linear counterpart. Thus, in order to use the Lindstedt-Poincaré method, we must reformulate our original system Eq. (4). This can be accomplished introducing a new function Δ (which will ultimately be set to zero) and rewriting the differential equations as:

$$\begin{aligned} \ddot{x} - 2\dot{y} - (1+2c_2)x &= \sum_{n\geq 2} (n+1) c_{n+1} T_n \\ \ddot{y} + 2\dot{x} + (c_2 - 1)y &= y \sum_{n\geq 2} c_{n+1} R_{n-1} \\ \ddot{z} + (c_2 - \Delta)z &= z \sum_{n\geq 2} c_{n+1} R_{n-1} \end{aligned} \tag{10}$$

Once again, this function can be expanded in power series of the amplitudes: $\Delta = \sum_{i,j} \Delta_{i,j} \alpha^i \beta^j$. The linear terms $\Delta_{1,0}$, $\Delta_{0,1}$ are now chosen so as to generate a periodic orbit in the linear system. The high-order terms $\Delta_{i,j}$ are finally chosen so as to guarantee $\Delta = 0$ (see Gomez *et al.* 1998, Andreu and Simo 1999, Gomez *et al.* 2000).

4. Extension to Large-Amplitude Orbits

The formula deduced in the previous section is convergent for all Halo orbits with amplitudes smaller than a certain critical value associated with $\rho = 1$ (see equation 3). This limit is not critical for most practical applications since the trajectories of most Halo-type artificial satellites do not extend beyond $\rho = 0.8$. Nevertheless, in some cases it proves important to (at least) analyze the structure of Halo orbits beyond this limit. In such a scenario, the formula deduced from Richardson' s expansion proves useless, and a different methodology must be employed. In this section we introduce such an approach. Although it implies a certain complication in the formula (and thus is not recommended for $\rho < 1$) it is convergent for all values of the amplitude.

We begin with equations (2). However, instead of expanding the right-hand sides in Legendre polynomials, we introduce two additional variables to our system:

$$\begin{aligned} s_1 &= \frac{(1-\mu)}{r_1^3} \\ s_2 &= \frac{\mu}{r_2^3} \end{aligned} \tag{11}$$

Differentiating these expressions, we can rewrite our new (extended) system as:

$$\begin{aligned} \ddot{x} - 2\dot{y} &= x - s_1(x-\mu) - s_2(x+1-\mu) \\ \ddot{y} + 2\dot{x} &= y - s_1 y - s_2 y \\ \ddot{z} &= -s_1 z - s_2 z \\ \dot{s}_1 &= -3 s_1 \frac{\dot{r}_1}{r_1} \\ \dot{s}_2 &= -3 s_2 \frac{\dot{r}_2}{r_2} \end{aligned} \tag{12}$$

The procedure to solve this system is similar to the one detailed in Section 3. We use the Lindstedt-Poincaré method in exactly the same manner. The difference with Richardson' s technique, however, lies in the expansion of the right-hand side of the last two equations. Instead of using Legendre polynomials, we work with operations on Fourier series. First, we use Eq. (8) to obtain a Fourier series of r_1, r_2. Since these quantities are always greater than zero, we can invert these series and obtain expressions for the coefficients of the series $(r_1)^{-1}$, $(r_2)^{-1}$. This can be done

analytically with little trouble, although a certain care must be taken with the manipulation of the coefficients. Introducing these into Eq. (12), we finally obtain formula analogous to those used in classical studies.

5. Results

This method was implemented and applied to both the Earth-Moon and the Sun-Earth sub-systems. In each we generated several Halo orbits around the Lagrangian point L_1 for different values of the linear amplitude β. Figure 1 shows the three-dimensional view and the x-y, y-z and x-z projections for the Earth-Moon case. Figure 2 shows the same views in the Sun-Earth system. The values used for β varied from 0.1 to 0.9 in steps of 0.1. Values of β higher than 0.9 gives very poor results and are not shown here. Figure 3 shows the variation of the non-linear frequency ω as a function of β. We can see that the period of the Halo orbit is very sensitive with respect to this amplitude; its value being significantly higher for the Earth-Moon system than for the Sun-Earth case. Figure 4 shows the variation of the amplitude in the x-direction α as a function of β. It is visible that the relation is close to linear for both systems and that the amplitudes are higher for the Sun-Earth system.

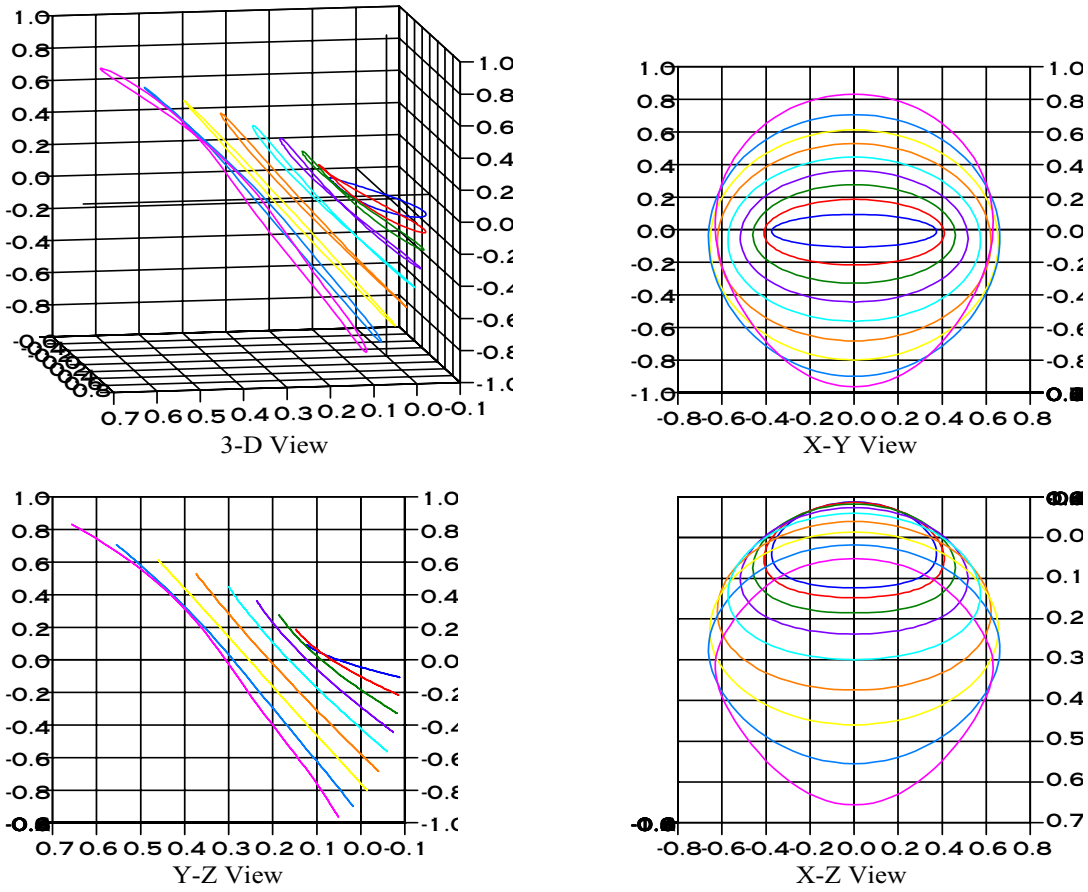


Fig. 1. Halo Orbits in the Earth-Moon System (3-D, X-Y, Y-Z,X-Z).

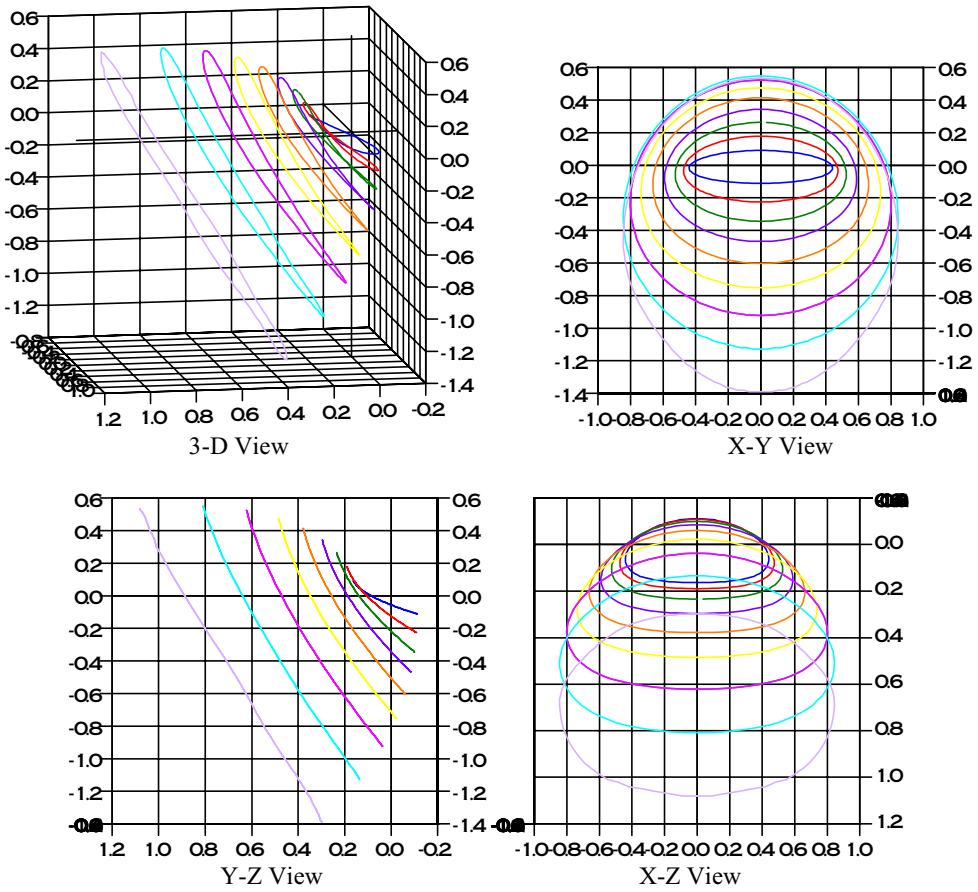


Fig. 2. Examples of Halo Orbits in the Sun-Earth System (3-D, X-Y, Y-Z, X-Z).

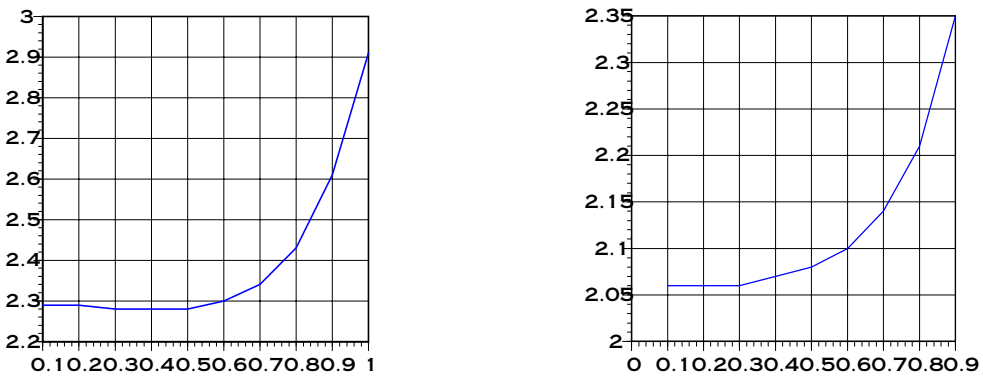


Fig. 3. Frequency of the Halo orbit as a function of β for the Earth-Moon system (left) and Sun-Earth system (right).

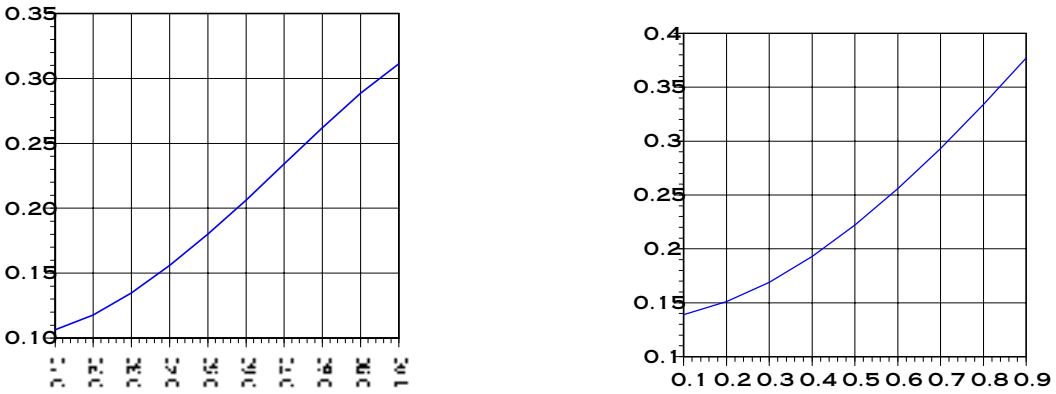


Fig. 4. Amplitude in the x-direction (α) of the Halo orbit as a function of β for the Earth-Moon system (left) and Sun-Earth system (right).

6. Comparison with Numerical Simulations

Having determined a number of solutions with the analytical approach, the next step was to check these results against the exact numerical integration of the restricted three-body problem. The results of this comparison are shown in Figs. 5 and 6 for different values of β for the Sun-Earth system. Initial conditions for the orbits generated by numerical integration were taken equal to those deduced from the analytical estimation. Figure 5 shows the difference between the two orbits, also as function of time. This difference is calculated at every instant of time by

$$\delta = \sqrt{\delta x^2 + \delta y^2 + \delta z^2 + \delta v_x^2 + \delta v_y^2 + \delta v_z^2} \tag{13}$$

Where δx, δy, δz are the differences in the components of the position vector between the trajectory given by the analytical approximation and the trajectory given by the numerical integration and δv_x, δv_y, δv_z are the differences in the components of the velocity vector between the same trajectories. Figure 6 shows both sets of orbits superposed as function of time. Each plot has a collection of orbits generated by the same initial conditions, but integrated over different periods of time. The numbers written at the right of the plots represents the integration time of each trajectory in canonical units (one canonical units is defined such that the period of the orbit is 2π). This procedure was used to determine the time that the orbit is destroyed. Figure 7 shows the orbits individually for the case β = 0.32. It is visible that the destruction of the orbit occurs for a time between 7 and 8.

We can clearly see that the accuracy of the analytical method decreases with larger values of β; this is expected since the size of the non-linear terms is a direct function of the amplitude. Nevertheless, even for moderate-to-large values of the amplitude, the approximation obtained by the Lindsted-Poincaré method is very good.

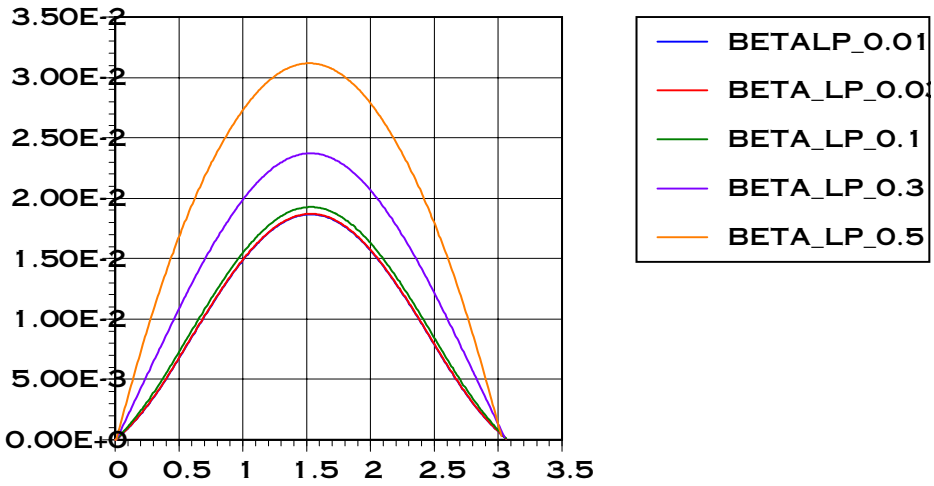


Fig. 5. Difference Between the Two Orbits (Analytical and Numerical) as Function of Time.

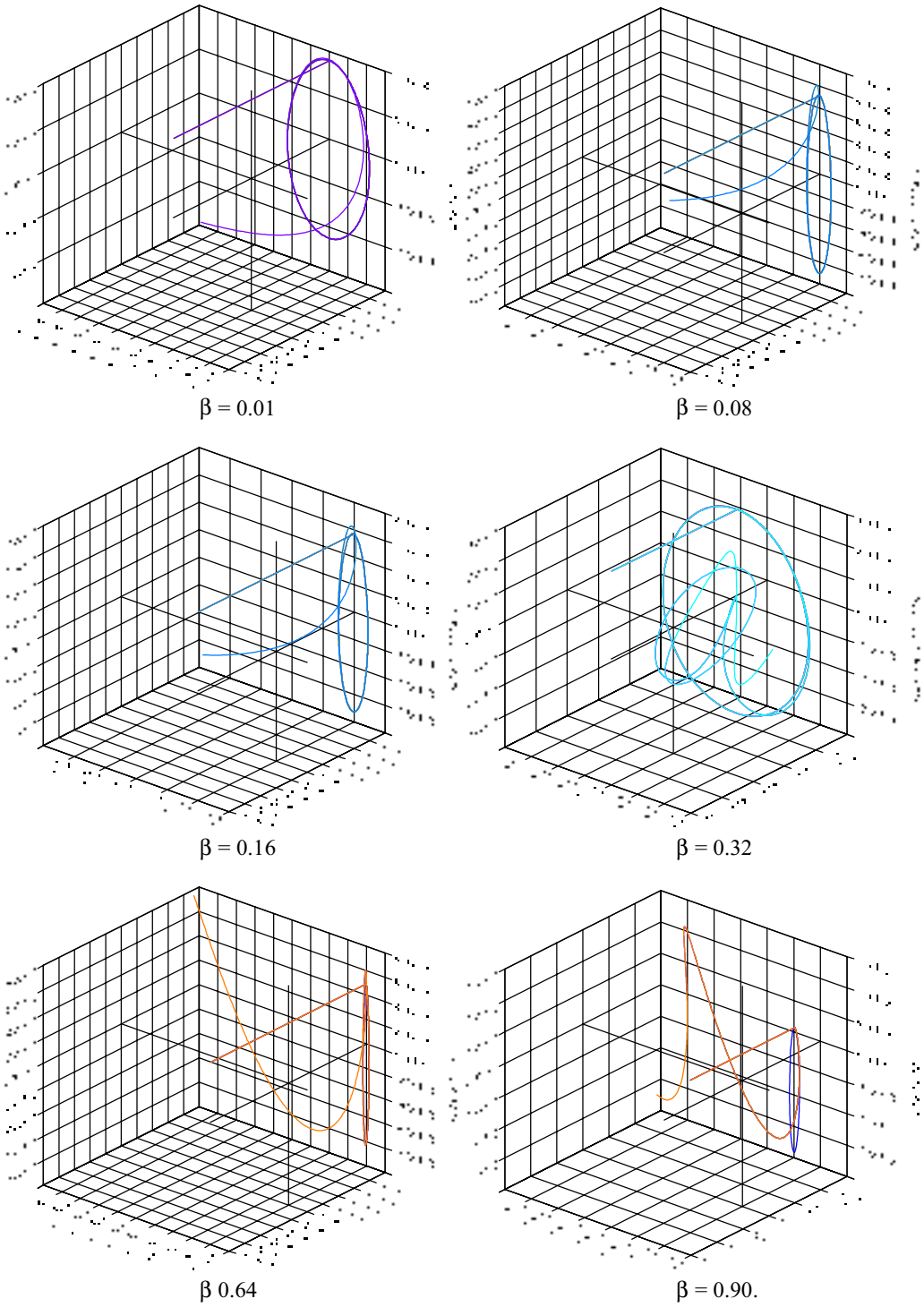


Fig. 6. Analytical and Numerical Trajectories for β = 0.01, 0.08, 0.16, 0.32, 0.64, 0.90.

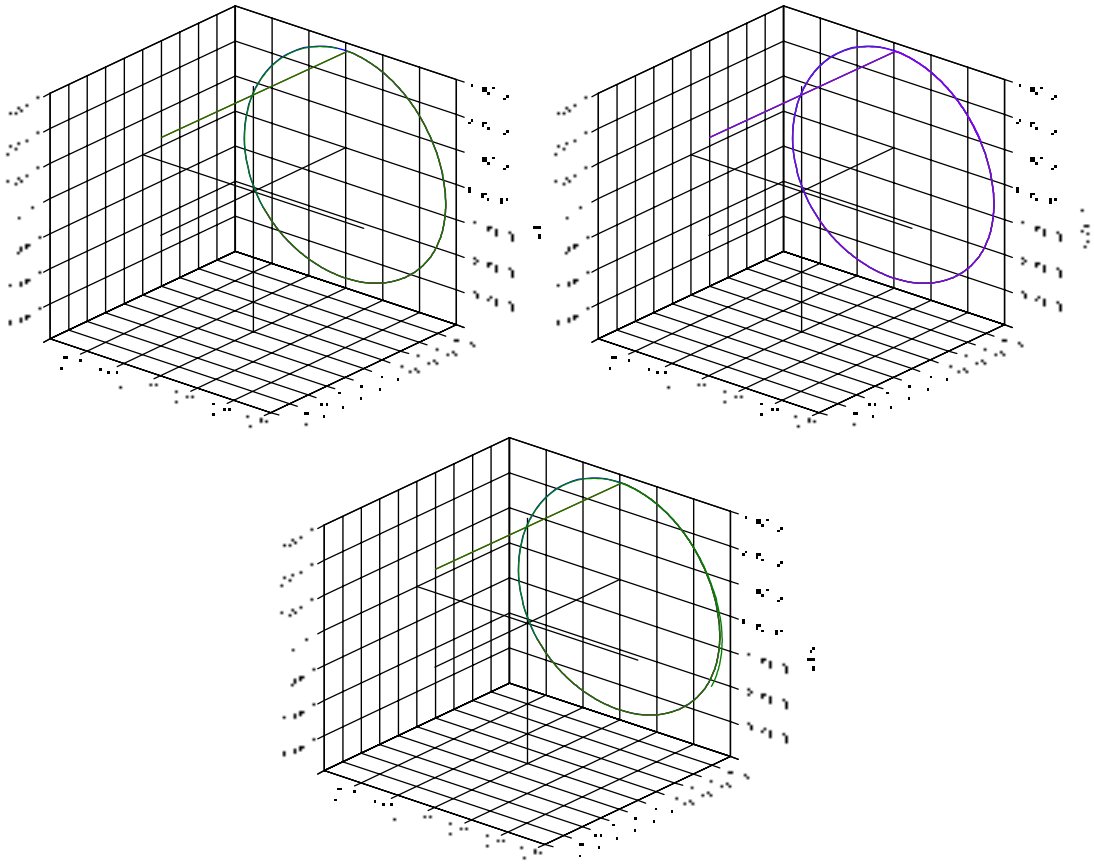


Fig. 7. Time Evolution of the Trajectories for $\beta = 0.32$.

7. Conclusions

In this communication we presented an analytical algorithm to calculate Halo orbits in the circular restricted three-body problem. Several periodic orbits were generated for both the Earth-Moon and the Sun-Earth systems. Comparisons with numerical simulations of the exact equations show that the orbits remain close to each other for about three revolutions, although the exact time is a function of the amplitude. After this interval, the differences between the full and the truncated solution become significant. It is important to bear in mind that these periodic orbits are unstable; thus, any discrepancy between the initial conditions and the center manifold will grow exponentially with time. For this reason, the proximity between the analytical and numerical solutions for as long as three complete revolutions is more than adequate. Another characteristic visible in the results shown in this paper is that in the Sun-Earth system the trajectories are closer to each other than in the Earth-Moon system. This fact occurs because the mass of the Moon (in the Earth-Moon system) relative to the total mass of the system is higher than the mass of the Earth (in the Sun-Earth system) and as a consequence its gravity field acts stronger in the trajectories.

8. Acknowledgments

The authors are grateful to the São Paulo State Science Foundation (FAPESP) for the research grants received under Contracts 2000/14769-4, 2000/7074-0 and 1999/08740-4 and to CNPq (Brazilian National Council for Scientific and Technological Development) for contract 300221/95-9.

9. References

Andreu, M.A. e Simó, C. (1999). Translunar Halo Orbits in the Quasi-Bicircular Problem, em "The Dynamics of Small Bodies in the Solar System", NATO, ASI (Maratea, Italia).

Farquhar, R. W. (1970). The Control and Use of Libration-Point Satellites, NASA TR R-346.

Farquhar, R. W. e Kamel, A.A. (1973). Quasi-Periodic Orbits Around the Translunar Libration Points, *Celestial Mechanics*, **7**, 458.

Gomez, G., Howell, K., Masdemont, J. e Simó, C. (1998). Station-Keeping Strategies for Translunar Libration Point Orbits. *AAS 98-168.*

Gomez, G., Masdemont, J. e Simó, C. (2000). Quasi-Halo Orbits associated with Libration Points, preprint.
Richardson, D.L. (1980). Analytic Costruccion of Periodic Orbits Around the Collinear Points. *Celestial Mechanics*, 22, 241-253.
Szebehely, V. (1967). "Theory of Orbits: The Restricted Problem of Three Bodies", Academic Press, New York.

XVI CONGRESSO BRASILEIRO DE ENGENHARIA MECÂNICA
16th BRAZILIAN CONGRESS OF MECHANICAL ENGINEERING

ANALYTICAL SOLUTION FOR SOME PROBLEMS WITH CHAOTIC RESPONSE

Carlos P. Filipich
Department of Engineering, Universidad Nacional del Sur, 8000 Bahía Blanca, Argentina
Mechanical Systems Analysis Group, FRBB, Universidad Tecnológica Nacional, 8000 Bahía Blanca, Argentina

Marta B. Rosales
Department of Engineering, Universidad Nacional del Sur, 8000 Bahía Blanca, Argentina
CONICET, Argentina
mrosales@criba.edu.ar

Abstract. *In this paper the analytical solution of nonlinear ordinary differential systems is addressed. The problems herein solved are classical in the related literature and some exhibit chaotic behavior in certain ranges of the involved parameters despite being simple-looking deterministic systems. The solutions are approached by means of algebraic series in the time variable that lead to elementary recurrence algorithm. This methodology is an alternative to the standard numerical techniques (e.g. Runge-Kutta, Newmark, central difference) and ensures the theoretical exactness of the response. The desired numerical precision is attained using time steps several times larger than the usual ones in the numerical integration schemes. Three examples illustrate the methodology: a) Duffing equation: b) mass-spring system with spring of response proportional to a hyperbolic tangent; c) Lorenz equations. The diagrams of the trajectories vs. time and phase plots are shown. The simplicity of the technique is apparent. Neither divergence nor numerical damping have been detected in the reported results. An available analytical solution may be an additional tool within the standard qualitative analysis. At present the solution of higher order problems and others governed by partial differential equations are under study.*

Keywords. *non-linear, differential equations, algebraic series*

1. Introduction

Non-linear differential problems governed by ordinary differential equations (ODE´s) are herein addressed by an analytical solution using algebraic series. The previous manipulation of the equations leads to very convenient recurrence algorithms that ensure the exactness of the solution as well as the computational efficiency of the method. The approach is straightforward and is illustrated with three examples, a) Duffing equation: b) mass-spring system with spring of response proportional to a hyperbolic tangent; c) Lorenz equations. In all the cases the results are given in plots (state variables vs. time, phase plots). Neither divergence nor numerical damping was found in any case. The availability of an analytical solution may be also a helpful tool in the qualitative analysis of non-linear equations.

Usually this type of differential systems is tackled by means of numerical tools (time integration schemes, e.g. Runge-Kutta, Newmark method, central difference, see for instance Bathe, 1995). The authors have addressed similar problems with a variational method named WEM (Rosales and Filipich, 2000 and Rosales and Filipich, 2001).

2. Analytical solution

In this section the general algebra of the approach is stated. In the following sections the three examples are presented.

Let us consider a continuous function $x = x(\tau)$ with $0 \le \tau \le 1$. We will denote the expansion in algebraic series as

$$[x] \equiv \sum_{k=0}^{N} a_{1k} \tau^k \tag{1}$$

and for powers *m*

$$[x^m] \equiv \sum_{k=0}^{N} a_{mk} \tau^k \tag{2}$$

In order to fulfill an *algebraic consistence condition* (A.C.) the following relationships have to be satisfied

$$[x^m] = [x^{m-1}][x] \tag{3}$$

After replacing the series expressions in each factor of this equation, one obtains the following recurrence expression

$$a_{mk} = \sum_{p=0}^{k} a_{(m-1)p} a_{1(k-p)} \qquad \text{or} \qquad a_{mk} = \sum_{p=0}^{k} a_{(m-1)(k-p)} a_{1p} \tag{4}$$

Now let us expand an analytical function $f = \hat{f}(x) = \hat{f}(x(\tau)) = f(\tau)$ in Taylor series

$$\hat{f}(x) = \sum_{m=0}^{M} \alpha_m x^m \tag{5}$$

where α_m are known and in particular, we denote

$$[1] = \sum_{k=0}^{N} \delta_{0k} \tau^k \tag{6}$$

where $a_{0k} = \delta_{0k}$ and δ_{0k} are the Kronecker delta´s. If we substitute Eq. (2) in Eq. (5) we may write

$$[f(\tau)] = \sum_{k=0}^{N} \varphi_k \tau^k \qquad \varphi_k = \sum_{m=0}^{M} \alpha_m a_{mk} \tag{7}$$

This expression will be used for any analytical function. Now is we have to deal with a rational function $\hat{F}(x)$

$$\hat{F}(x) = \frac{\hat{g}(x)}{\hat{f}(x)} = \frac{g(\tau)}{f(\tau)} = F(\tau) \tag{8}$$

being $\hat{g}(x)$ and $\hat{f}(x)$ analytical functions and $\hat{f}(0) \neq 0$ and $\hat{g}(x) = \sum_{m=0}^{M} \beta_m x^m$ and β_m are known. Then it is possible to write

$$[g(\tau)] = \sum_{k=0}^{N} \varepsilon_k \tau^k \text{ (a)} \qquad \varepsilon_k = \sum_{m=0}^{M} \beta_m a_{mk} \text{ (b)} \tag{9}$$

If we denote

$$[F(\tau)] = \sum_{k=0}^{N} \lambda_k \tau^k \tag{10}$$

Now the A.C. must be applied

$$[F(\tau)][f(\tau)] = [g(\tau)] \tag{11}$$

$$\left(\sum_{k=0}^{N} \lambda_k \tau^k \right) \left(\sum_{k=0}^{N} \varphi_k \tau^k \right) = \sum_{k=0}^{N} \varepsilon_k \tau^k \quad \text{where} \quad \varepsilon_k = \sum_{p=0}^{k} \varphi_p \lambda_{(k-p)} \tag{12}$$

The λ_k are unknowns and the sets φ_k and ε_k are known. It is apparent that $\lambda_0 = \varepsilon_0 / \varphi_0$. Now the recurrence relationship for λ_k is

$$\lambda_k = \frac{\varepsilon_k - \sum_{p=1}^{k} \varphi_p \lambda_{(k-p)}}{\varphi_0} \quad \text{where} \quad \varphi_0 \neq 0 \quad \text{and} \quad k = 1,2,...,N \tag{13}$$

It should be noted that $\varphi_0 \neq 0$ in order for $F(0)$ to exist. Also all the expansions can be done with respect to any τ_0. Here $\tau_0 = 0$.

3. Illustrative examples.

Three examples will be addressed with this technique: a Duffing equation (SDOF second order ODE), a non-linear second order ODE (nonlinear spring of hyperbolic tangent reaction) and the Lorenz equations (system of three ODE's). In this section the analytical statement for each case is introduced. The numerical results are shown in Section 3.

3. 1. Duffing equation

In this subsection the well-known differential problem introduced by Duffing is tackled. It may represent several physical problems such as a beam at buckling load, an electrical circuit with a nonlinear inductor, a non-linear spring with mass system. The differential problem is stated as follows

$$\begin{aligned} &\ddot{u} + p^* \dot{u} + q^* u^3 = B^* \hat{f}(t) \\ &u(0) = U_0 \\ &\dot{u}(0) = V_0^* \end{aligned} \tag{14}$$

where $u = \hat{u}(t)$. Now let us assume T an interval of interest and a non-dimensionalized variable $\tau = t/T$. Accordingly $u = \hat{u}(t) = \hat{u}(T\tau) = v(\tau)$; $\hat{f}(t) = \hat{f}(T\tau) = f(\tau)$. Thus the problem (1) is now written as

$$\begin{aligned} &v'' + pv' + qv^3 = Bf(\tau) \\ &v(0) = U_0 \\ &v'(0) = V_0 \end{aligned} \tag{15}$$

in which $p = Tp^*$; $q = T^2 q^*$; $B = T^2 B^*$; $V_0 = TV_0^*$ and $v' = dv/d\tau$, etc. Let us propose the following algebraic series to represent powers of v

$$\begin{aligned} &\left[v^m\right] = \sum_{i=0}^{N} a_{mi} \tau^i \\ &\left[v'\right] = \sum_{i=0}^{N-1} \varphi_{1i} a_{1(i+1)} \tau^i; \quad \left[v''\right] = \sum_{i=0}^{N-2} \varphi_{2i} a_{1(i+2)} \tau^i; \quad \varphi_{1i} = (i+1); \quad \varphi_{2i} = (i+1)(i+2) \end{aligned} \tag{16}$$

The load may be expanded in Taylor series as

$$\left[f(\tau)\right] = \sum_{i=0}^{N} \alpha_i \tau^i \tag{17}$$

The non-linear differential Eq. (15) is now written

$$\sum_{i=0}^{N-2} \tau^i \left[\varphi_{2i} a_{1(i+2)} + p\varphi_{1i} a_{1(i+1)} + q a_{3i} - B\alpha_i\right] = 0 \tag{18}$$

A recurrence relationship is obtained and together with the A.C. it is possible to find all the values of the coefficients a_{mi} starting from the initial conditions. That is, from Eq. (4)

$$a_{2i} = \sum_{p=0}^{i} a_{1p} a_{1(i-p)} \qquad \text{and} \qquad a_{3i} = \sum_{p=0}^{i} a_{2p} a_{1(i-p)} \tag{19}$$

Consequently from Eq. (18) and since $a_{10} = U_0$ and $a_{11} = V_0$ the following recurrence expression yields

$$a_{1(i+2)} = \frac{B\alpha_i - p\varphi_i a_{1(i+1)} - q a_{3i}}{\varphi_{2i}} \qquad (i = 0,1,2,\cdots,N-2) \tag{20}$$

with which the problem may be solved. Figure 1 shows the trajectory as a function of the dimensionless time (τ=t/T) with *T*=0.35 for the case $U_0 = 3$, $V_0^* = 4$, $p^* = 0.05$, $q^* = 1$, $B^* = 7.5$, $\hat{f}(t) = \cos t$ with *N*=20.

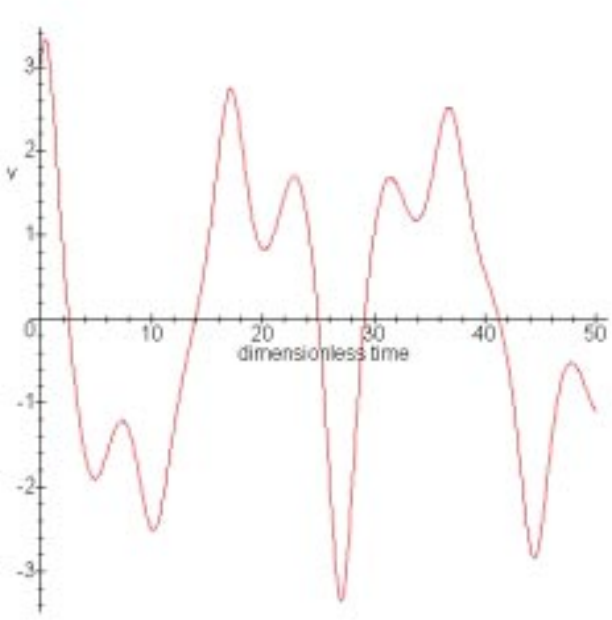


Figure 1. Solution of the forced Duffing equation. $U_0 = 3$, $V_0^* = 4$, $p^* = 0.05$, $q^* = 1$, $B^* = 7.5$.

Figure 2 depicts the corresponding phase plot. These results are identical to the ones reported in Thompson and Stewart, 1986. It should be mentioned that this problem is extremely sensitive to the initial conditions. That is, is for instance one takes $U_0 = 3.1$, $V_0^* = 4.1$ the solution varies quite significantly from the one reported here, as Thompson and Stewart comment. The consequence of this feature is chaos. Consequently the numerical behavior of the methodology is relevant to the reliability of the results.

3. 2. Hyperbolic tangent nonlinearity

Let us analyze a mass-spring system in which the spring reacts with a hyperbolic tangent law. Here the following non-linear homogeneous equation is going to be solved (after non-dimensionalization)

$$\begin{aligned} &v'' + q\tanh(v) = 0 \\ &v(0) = U_0 \\ &v'(0) = V_0 \end{aligned} \tag{21}$$

with $v = v(\tau)$ and $v' = dv/d\tau, \ldots$. It should be noted that this autonomous system with a two-dimensional phase space will not exhibit physical chaos. Notwithstanding its solution using numerical integration schemes may not conduce to adequate responses, i.e. numerical chaos or damping is sometimes introduced by the method itself.

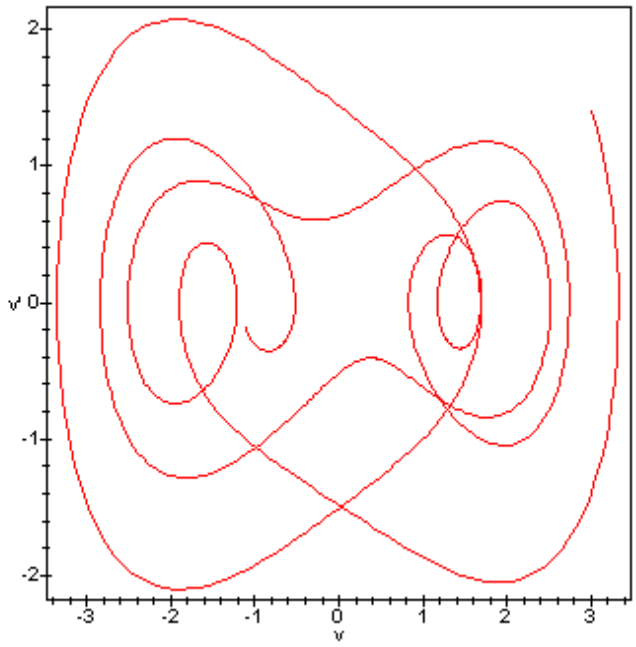


Figure 2. Phase plot of the Duffing problem. Idem Fig. 1.

The Taylor series for a hyperbolic tangent is

$$\tanh(v) = \sum_{m=0}^{M} \alpha_m v^m \tag{22}$$

The next power series is proposed for v

$$[v] = \sum_{i=0}^{N} a_{1i} \tau^i \tag{23}$$

and the second derivative and powers may be written as

$$[v''] = \sum_{i=0}^{N-2} \varphi_{2i} a_{1(i+2)} \tau^i; \quad [v^m] = \sum_{i=0}^{N} a_{mi} \tau^i; \tag{24}$$

Using this notation the non-linear differential Eq. (21) is reduced to

$$\sum_{i=0}^{N} \tau^i \left[\varphi_{2i} a_{1(i+2)} + q \sum_{m=0}^{M} \alpha_m a_{mi} \right] = 0$$

It is possible to obtain a recurrence relationship for a first set of coefficients

$$a_{1(i+2)} = -\frac{q \sum_{m=0}^{M} \alpha_m a_{mi}}{\varphi_{2i}} \qquad (i = 0,1,2,\cdots, N-2) \tag{25}$$

The algorithm should start by setting the initial conditions with which $a_{10} = U_0$ and $a_{11} = V_0$. The A.C. Eq. (4) permits to find a_{m0} and a_{m1}. Now, making use of the recurrence Eq. (25) a new pair of coefficients is found: a_{12} and a_{13}. The loop starts again using the A.C. and the recurrence alternatively until the desired order is attained. Recall that *M* is the number of terms in the Taylor series (it will be shown that *M*=15 is more than sufficient) and *N* is the number of terms in the power series.

The equations stated in this section were solved with the recurrence Eq. (25). In this case *q*=1, *N*=20, *T*=1 and $U_0 = 0.4$, $V_0 = 0$. For this particular selection of the parameters the solution is periodic, as is observed in Fig. 3. The phase plot (not reported here) is a perfect circle. As mentioned before this particular example does not have the conditions to physical chaos.

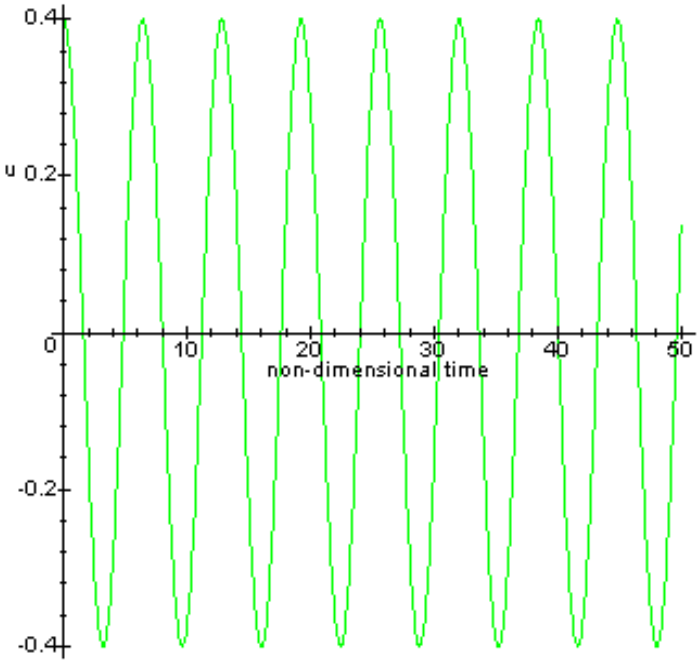


Figure 3. Mass-spring system with hyperbolic tangent non-linearity. *q*=1, *N*=20, *T*=1 and $U_0 = 0.4$, $V_0 = 0$. Number of steps 5.

It should be noted that in certain cases, when handled with standard numerical techniques (for instance Newmark, Houbolt, alpha methods), numerical damping or chaos may arise for certain time steps (Xie, 1996). Even in certain "hard" tests no convergent solutions were found with these schemes. In Rosales and Filipich (2000) this problems were overcome using a variational methodology.

3. 3. Lorenz equations

The well-known problem of Lorenz (see for instance Strogatz, 1994), who discovered in 1963 the chaotic motions of a model of convection of rolls in the atmosphere, is herein solved with algebraic series. The behavior of this problems is such that very small changes in the initial conditions lead to different solutions but, as Lorenz showed, bounded in a three dimensional butterfly-shaped domain. The three-dimensional system derived by Lorenz is

$$\begin{aligned} \dot{u} &= \sigma^* (v - u) \\ \dot{v} &= r^* u - v - uw \\ \dot{w} &= uv - b^* w \end{aligned} \tag{26}$$

where $\sigma^*, r^*, b^* > 0$ are parameters; $u = \hat{u}(t)$; $v = \hat{v}(t)$; $w = \hat{w}(t)$. Here the variable $\tau = t/T$ is introduced. The following algebraic series are proposed

$$[u] = \sum_{i=0}^{N} A_i \tau^i; \quad [v] = \sum_{i=0}^{N} B_i \tau^i; \quad [w] = \sum_{i=0}^{N} C_i \tau^i; \tag{27}$$

and the derivatives

$$[u'] = \sum_{i=0}^{N-1} \varphi_{1i} A_{i+1} \tau^i; \quad [v'] = \sum_{i=0}^{N-1} \varphi_{1i} B_{i+1} \tau^i; \quad [w'] = \sum_{i=0}^{N-1} \varphi_{1i} C_{i+1} \tau^i; \tag{28}$$

After replacement of these series in the differential system a recurrence system of equations is obtained as follows

$$A_{i+1} = \frac{\sigma}{\varphi_{1i}}(A_i - B_i); \quad B_{i+1} = \frac{1}{\varphi_{1i}}(rA_i - TB_i - TS_i); \quad C_{i+1} = \frac{1}{\varphi_{1i}}(TZ_i - bC_i) \tag{29}$$

with

$$S_i = \sum_{p=0}^{i} A_p C_{i-p}; \quad Z_i = \sum_{p=0}^{i} A_p B_{i-p} \tag{30}$$

The Lorenz problem was solved using the recurrence Eq. (29) for the data reported in the caption of Fig. 4. The response is chaotic, as found by Lorenz in 1963.

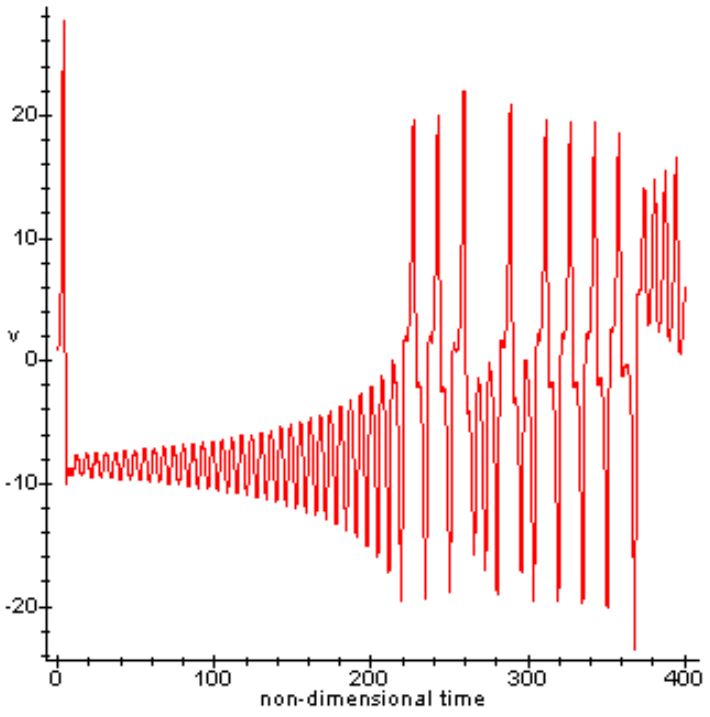


Figure 4. Lorenz attractor. Trajectory *v.* $u_0 = 0$, $v_0 = 1$, $w_0 = 0$. $T = 0.1$. $\sigma^* = 10, r^* = 28, b^* = 8/3$. *N*=20.

The phase plot of Fig. 5 shows a bounded domain of the trajectory. Again this problem is very sensitive to the initial conditions and poses a challenge to any numerical tool.

4. Conclusions.

In this paper the analytical solutions of three non-linear differential problems governed by ODE´s have been addressed with algebraic series. Some of these problems are known to be sensitive dependent on initial conditions.

The aim is to provide of a theoretically exact solution as an alternative to standard numerical techniques. The resulting recurrence algorithms are very efficient and, in general, the time steps are larger than the usual ones in other integration schemes.

The availability of an analytical solution may be useful to perform qualitative analysis, which lead to the study of bifurcations and chaos.

5. Acknowledgements

This work has been partially supported by grants for PGI from SGCyT of the Universidad Nacional del Sur and for PIP form CONICET, both institutions from Argentina. The second author is a research scientist of CONICET.

6. References.

Bathe, K-J.,. 1995, "Finite Element Procedures", Ed. Prentice-Hall.

Rosales, M.B. and Filipich, C.P., 2000, "An Alternative Technique for Time Integration of Dynamic Equations", Proceedings of the XXIXX South American Congress on Structural Engineering (Jubileo Ing. Ricaldoni), Vol. 1, CDROM, Punta del Este, Uruguay.

Rosales, M.B. and Filipich, C.P., 2001, "Time Integration of Non-Linear Dynamic Equations by means of a Direct Variational Method", submitted to the Journal of Sound and Vibration (under review).

Strogatz, S.H., 1994, "Nonlinear Dynamic and Chaos", Ed. Addison-Wesley Pub.Co., Reading, Massachussets, U.S.A., 498 p.

Thompson, J.M.T. and Stewart, H.B., 1986, " Nonlinear Dynamic and Chaos", Ed. John Wiley & Sons Ltd. 376 p.

Xie, Y.M., 1996, " An Assessment of Time Integration Schemes for Non-Linear Dynamic Equations", Journal of Sound and Vibration, Vol. 192, pp. 321-331.

7. Copyright Notice

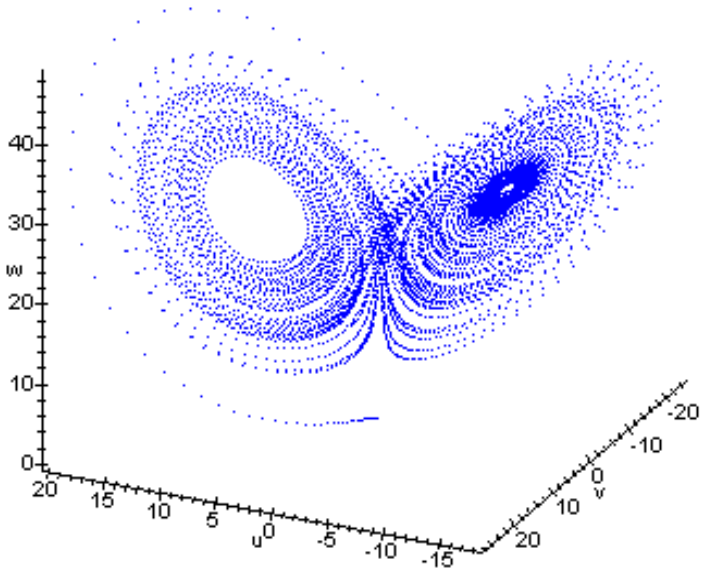


Figure 5 . Lorenz strange attractor. Idem Fig. 4.

XVI CONGRESSO BRASILEIRO DE ENGENHARIA MECÂNICA
16th BRAZILIAN CONGRESS OF MECHANICAL ENGINEERING

STRESS CONCENTRATION EVOLUTION IN SHOULDERED PLATES FOR NON-LINEAR ELASTIC MATERIALS

Carlos J. Gomes A.
Mechanical Engineering Department, Carnegie Mellon University, Pittsburgh, USA
cgomes@andrew.cmu.edu

Nando Troyani
Departamento de Ingeniería Mecánica, Núcleo de Anzoátegui, Universidad de Oriente, Venezuela
ntroyani@cantv.net

Abstract This work determines de evolution of the theoretical stress concentration factor for shouldered plates of nonlinear elastic material subjected to tension in the context of linear kinematics mechanical response, using the finite element method. The non-linearity is handled using an incremental procedure that also takes care of both the non-uniformity and the anisotropy induced by the loading process, by locally using the concept of equivalent elastic modulus. It is determined that for materials which exhibit softening the concentration diminishes in a way that depends on the variation of the tangent modulus of the material, a development considered favorable. The employed numerical procedure satisfies the essential equilibrium requirement at each computational iteration.

Keywords. *Stress concentration factor, nonlinear materials, finite element, predictor corrector.*

1. Introduction

It has been established that for linear elastic materials the theoretical stress concentration factor, K_t, is only a function of the geometry of the member and a function of the type of loading as well, but independent of the load magnitude. Recently, Troyani et al (2001) show that their magnitude depends also on the length of the member. In addition, for nonlinear materials K_t depends on the local magnitude of stress as well, as is indicated in Savin (1968) and more recently, using the so-called simple incremental procedure (SIP), in Gomes and Troyani (1998).

In this work K_t for shouldered plates of non linear elastic material subjected to tension, is determined for the case when $D/d =$ 1.5, see Fig. (1), for various values of the r/d ratio, for semi-infinite (sufficiently long) plates in the direction of the applied load, the semi infinite concept is precisely defined in Troyani et al (2001.)

The basic definition for K_t, Pilkey (1997), is used here in the form given in Eq. (1) below,

$$K_t = \frac{\sigma_{max}}{\sigma_{nominal}} \tag{1}$$

The nominal stress, $\sigma_{nominal}$, is computed using the reduced section corresponding to dimension d, Fig. (1), with S representing uniform end traction and considering a unit thickness plate, according to,

$$\sigma_{nominal} = S \tag{2}$$

The maximum stress, σ_{max}, is computed using triangular finite elements in the context of plane stress elasticity, in a numerical strategy based on an incremental predictor corrector procedure (IPCP). Linear kinematics is used through this work, that is, the standard linearized relationship between displacements and strain are used. For the work reported here the aluminum alloy S-6063-T0 (experimentally determined stress-strain diagram shown in Fig. (2).) was used.

2. Incremental predictor corrector calculations

There are a variety of numerical strategies to treat the kind of problem examined herein, see Desai and Abel (1972) and Zienckiewics and Taylor (1994), for instance. Out of these, a particular form of IPCP was developed for the work presented here and selected to compare this particular strategy with existing ones (reported elsewhere) and to be able to compute strains and corresponding stresses at each incremental step in the procedure in order to compute the evolution of K_t.

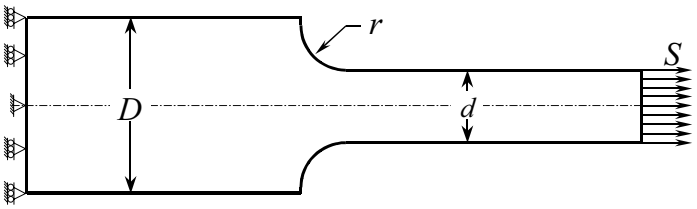


Figure 1. Geometry, boundary conditions and significant parameters.

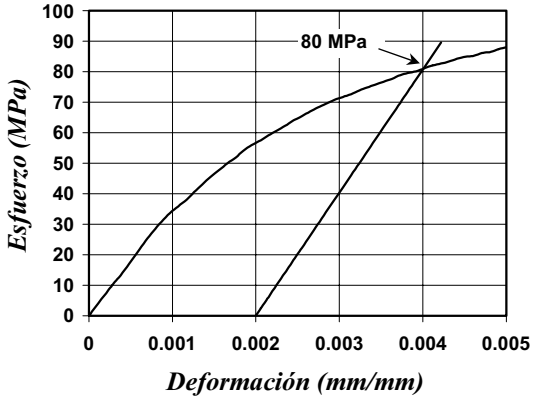


Figure 2. Stress-strain diagram for aluminum alloy S-6063-T0 up to yielding (0.2 %).

The IPCP consists of dividing the total load in load increments, not necessarily of the same magnitude as long as they satisfy Eq. (3) below.

$$\mathbf{f_T} = \mathbf{f_o} + \sum_{\mathbf{i}=1}^{\mathbf{M}} \Delta \mathbf{f_i} \tag{3}$$

$\mathbf{f_T}$, $\mathbf{f_o}$ and $\Delta\mathbf{f_i}$ represent total load, initial load and i^{th} load increment respectively.

The load increments are applied sequentially and within each increment the material response throughout the domain is considered linear but non-uniform in the following fashion.

1. The material properties (tangent modulus), first predictor, are determined for each element at the beginning of the load increment, based on the local stress.
2. A first estimate for the displacements, strains and stresses is computed for the load increment at each element.
3. The material properties, second predictor, are then estimated at the end of the load increment based on the local stress corresponding to each element.
4. The displacements and strains and stresses for the load increment are computed on the basis of the average material properties (corrector), estimates 1 and 3 above.
5. Repeat procedure 1-4, for each successive load increment until the total load is attained.

Formally, the process proceeds according to tangent modulus evaluations based on

$$E^{c}_{i-1} = \frac{E^{p}_{i-1} + E'_{i-1}}{2} \tag{4}$$

E^{p}_{i-1} : Local tangent modulus at the beginning of the i^{th} load increment.

E'_{i-1} : Local tangent modulus predicted at the end of the i^{th} load increment.

E^{c}_{i-1} : Local tangent modulus corrected at the end of the i^{th} load increment.

So the nonlinear problem is solved as a succession of piecewise linearized problems, this is represented in Fig. (3) where a graphical comparison is shown between the SIP and IPCP, for a typical increment.

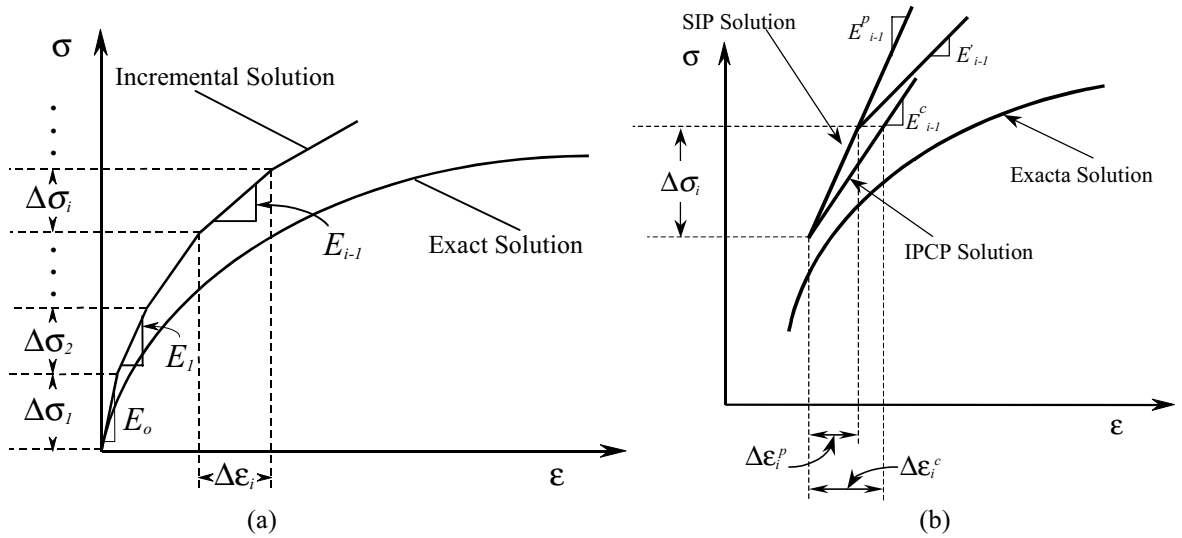


Figure 3. (a) Incremental scheme, and (b) predictor corrector scheme.

Clearly for the IPCP strategy, Fig. (3b) two preliminary calculations are necessary to obtain the results for each load increment, Consequently the method is computational costlier than the SIP method, but appreciably more accurate.

The current values of the strain tensor and stress tensor are obtained with the following expressions

$$\boldsymbol{\varepsilon}_{\mathbf{i}} = \boldsymbol{\varepsilon}_{\mathbf{o}} + \sum_{\mathbf{j}=1}^{\mathbf{i}} \Delta\boldsymbol{\varepsilon}_{\mathbf{j}} \tag{5}$$

$$\boldsymbol{\sigma}_{\mathbf{i}} = \boldsymbol{\sigma}_{\mathbf{o}} + \sum_{\mathbf{j}=1}^{\mathbf{i}} \Delta\boldsymbol{\sigma}_{\mathbf{j}} \tag{6}$$

$\boldsymbol{\sigma}_i$, $\boldsymbol{\varepsilon}_i$: Stress and strain tensors after the j^{th} load increment respectively.
$\boldsymbol{\sigma}_o$, $\boldsymbol{\varepsilon}_o$: Initial stress and strain tensors.
$\Delta\boldsymbol{\sigma}_j$, $\Delta\boldsymbol{\varepsilon}_j$: Increments of the stress and the strain tensors at the j^{th} load increment respectively.

The two methods are, strictly speaking, applicable only to uniaxial stresses, since the tangent modulus depends on the local stresses (strains) intensity and these are different in general for the two principal directions in 2-dimensional problems, transforming what was initially an isotropic material into a non-uniform anisotropic one, that is, stress induced anisotropy. This difficulty was solved by assuming that a representative local (element) tangent modulus could be estimated on the basis of the von Misses effective stress since the material treated here is ductile. Same treatment could be given to the Poisson ratio. However, it was considered constant here only as a result of lack of available experimental values.

3. Computational procedure

The procedure described was coded by modifying an existing linear elastic finite element FORTRAN program based on the constant strain triangle (CST) element, and by adding the necessary additional routines.

Validation of the procedure was achieved by attempting to numerically simulate the experimentally determined stress-strain diagram of Fig. (2) (a one dimensional problem that can be considered a particular case of plane stresses) using both the IPCP and the SIP strategies.

Figure (4) shows the deviations, from the curve in Fig. (2), as a function of the number of load increments of the calculated results using both the SIP and the IPCP. Additionally the results from the finite element based software; ANSYS® using the Newton-Raphson method is included.

As can be seen in Fig. (4) convergence of the IPCP scheme is notably faster than the SIP scheme, based on error evaluation performed using Eq. (7) below. However, its convergence speed is not as faster as the Newton-Raphson speed when the number of load increment is higher.

Figure (4) shows that the IPCP strategy converges faster than the SIP, in such a way that it is computationaly less costly than the SIP. That is, in spite of the fact that the IPCP scheme requires two predicting computations for each load increment (vs one computation for the SIP for each load increment), the overall improvement in convergence clearly favors the use of this strategy in terms of computational cost.

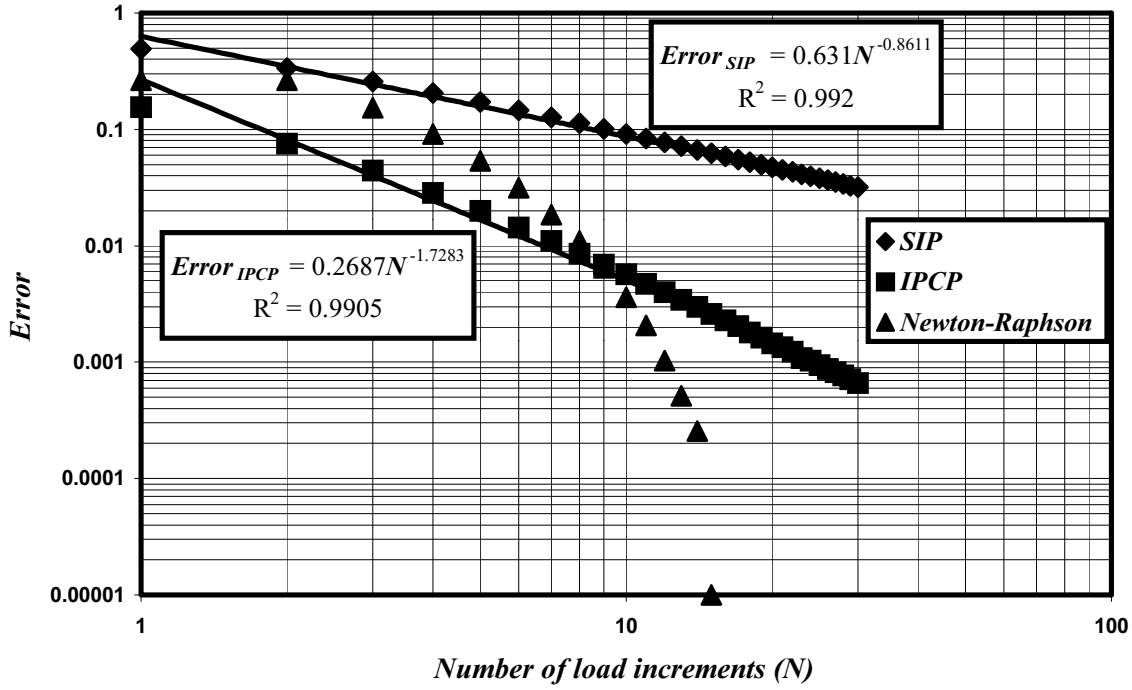


Figure 4. Percentage of error as a function of the number of load increments for the SIP and IPCP scheme.

$$Error = \frac{(\varepsilon_{real} - \varepsilon_N)}{\varepsilon_{real}}.100\% \tag{7}$$

ε_{real}: Experimental strain.
ε_N: Numerically determined strain for N number of load increments.

In order to increase the computational efficiency of both numerical procedures, the Gibbs, Poole and Stockmayer (1976) algorithm for band minimization was used.

4. Results and Analysis

Figure (5) shows a typical finite element mesh utilized in the course of the calculations reported here.

Figure (6) shows the evolution of the stress concentration factor for the case studied here using the numerical procedure described in the previous two sections. In the numerical tests the load was increased until the yield stress was reached at any local part of the domain. Guaranteeing in this fashion that the material remained in the elastic regime throughout the calculations.

This figure exhibits a decrease in K_t values as the stress level is increased (load increase). A phenomenon attributable to the fact that as the stresses are increased the material becomes progressively softer, less rigid, a fact that is reflected in smaller tangent moduli, see Fig. (2).

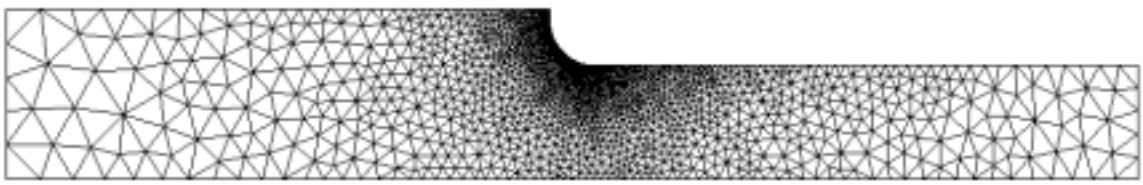


Figure 5. Typical finite element mesh, r/d = 0.2, D/d = 1.5.

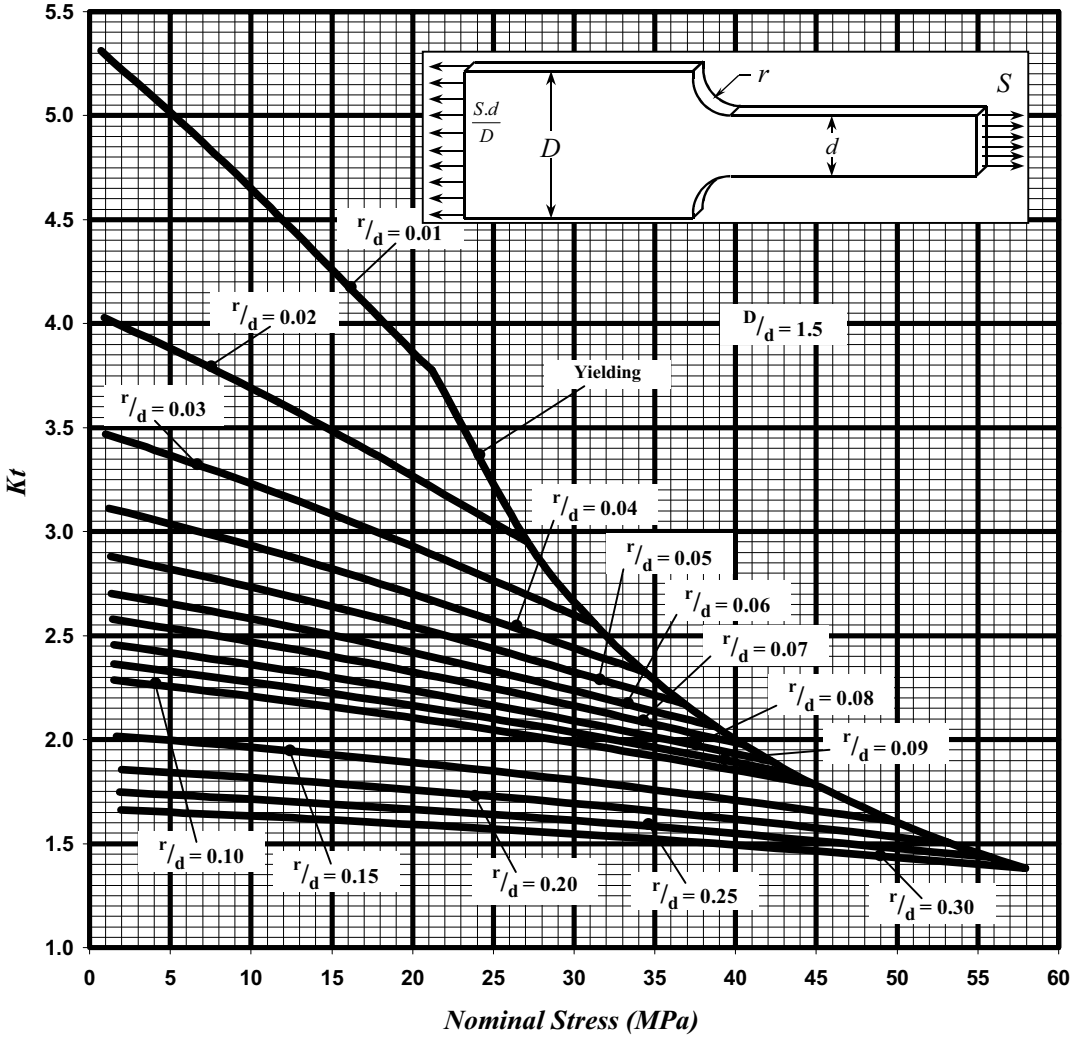


Figure 6. Stress concentration factor evolution as a function of the nominal stress level.

5. Concluding remarks

This work presents the numerically obtained in-plane stress concentration factors for shouldered plates subjected to tension, made of a non-linear elastic material, in particular, for the S-6063 T0 aluminum alloy.

It was determined that these factors decrease with increasing stress levels, due to material softening. One of the resulting effects of this decrease is a redistribution of stresses over the critical cross section to compensate for the stress decrease at the maximum stress position, in order to satisfy equilibrium. This observation is believed to be universal for this kind of materials.

The fact that the stress concentration in the critical point in the studied part decreases with increase in load could be considered as a beneficial feature of the kind of material treated here in terms of the overall strength exhibited by loaded mechanical elements.

6. Acknowledgement

This work was partially founded by Consejo de Investigación at Universidad de Oriente, Venezuela.

7. References

Desai, C. & Abel, J., 1972, “Introduction to the Finite Element Method,” Van Nostrand Reinhold Company, The United States, 477 p.

Frocht, M. M., 1935, "Factors of stress concentration photoelastically determined,” Trans. ASME, Applied Mechanics Section, Vol. 57, pp A67-A68.

Gibbs N., Poole, W. and Stockmayer, P., 1976, "An Algorithm for Reducing the Bandwidth and Profile of a Sparse Matrix," SIAM Journal on Numerical Analysis, Vol. 13, pp. 236-250.

Gomes, C. y Troyani, N., 1998, "Evolución del Factor de Concentración de Esfuerzos en Materiales Elásticos no Lineales", Procceding of IV Congreso Internacional de Métodos Numéricos en Ingeniería y Ciencias Aplicadas (CIMENICS98), Ciudad Guayana, Venezuela, pp. AE13- AE20.

Peterson, R., 1974, "Stress Concentration Factor," John Waley & Sons, Canadá, 317 p.

Pilkey, W. D., 1997, "Peterson's stress concentration factors," John Wiley & Sons, 2nd edition.

Pollonais, Y., Gomes, C. y Troyani, N., 1999, "Factor de concentración de esfuerzos para placas cortas con entallas en U simétricas sometidas a tensión," Procceding of IV Congreso Iberoamericano de Ingeniería Mecánica, CIDIM'99, Santiago de Chile, Chile.

Savin, G. N., 1968, "Stress distribution around holes," traslation of "Raspredeleniye Naprayazheniy Okolo Olverstiy." "Naukova Dumka" Press, Kiev.

Shigley J. and Mischke C., 1990, "Diseño en Ingeniería Mecánica," Mc Graw Hill, México, 883 p.

Sterlacci, G., Troyani, N. and Gomes, C., 2000, "Stress concentration factor of a U-shaped circumferential groove in cylinders of finite length under uniform axial load," Procceding of V Congreso Internacional de Métodos Numéricos en Ingeniería y Ciencias Aplicadas, CIMENICS'2000, Puerto la Cruz, Venezuela, pp. AE1- AE6.

Troyani, N., Gomes, C. and Sterlacci, G., 2001, "Theoretical Stress Concentration Factor for Short Rectangular Plates With Centered Circular Holes," Journal of Mechanical Design. Accepted for publication.

Zienkiewicz, O. and Taylor, R., 1994, "El Método de los Elementos Finitos," Mc Graw Hill, España, 650 p.

XVI CONGRESSO BRASILEIRO DE ENGENHARIA MECÂNICA
16th BRAZILIAN CONGRESS OF MECHANICAL ENGINEERING

FINITE VECTORIAL ROTATIONS

Marcelo González Bergweiler
Voith Paper Máquinas e Equipamentos Ltda.
Marcelo.Bergweiler@ig.com.br

Paulo M. Pimenta
Departamento de Estruturas e Fundações da Universidade de São Paulo
PPimenta@usp.br

Abstract. *Starting with the definition of a new product between vectors, the Geometric Product, and its use in simple geometric constructions, generic rotations as a composition of two binary rotations are obtained. This leads naturally to the definition of quaternions. After presenting some classical algebraic properties and an important relation between additive and multiplicative rotations, the legitimate use of vectors for representing finite rotations is stated, showing that the lack of commutativity of finite rotations as an argument against the use of vectorial parameters is nothing but prejudice. The elegance of quaternions in obtaining the link between the angular velocity vector and the time derivative of the different vectorial parameters is used. By introducing the concept of objectivity very generic algebraic relations between these tensors, their derivatives and the rotation tensor are shown. An additive integration algorithm is presented as a natural alternative to the original multiplicative one.*

Keywords. *geometric product, quaternions, finite rotations, objectivity, rigid body dynamics.*

1. Introduction

Rigid body dynamics uses almost exclusively Euler parameters. A much more natural approach, the use of vectorial parameters, is banned from textbooks as an example of " vector candidates that do not obey the law of vector addition" (e.g. Goldstein, 1982; Crisfield, 1997; Thompson, 1963).

As engineers are rather practical oriented minds, this advice is being more and more solemnly ignored, as may be seen from many advanced papers dealing with vectorial parameters (e.g. Cardona and Geradin, 1988). After all there is no need to add these " objects" , as it is well known that we only multiply rotations.

In this paper we show that the name vector parameter is fully justified, and more, we will shamelessly add them. In fact we will state that the segregation of vectorial rotation parameters is nothing but prejudice. A very effective rigid body integration algorithm (Simo and Wong, 1991), already using vector parameters but in multiplicative flavor, is here naturally presented in an additive taste.

For this sake we start from the very beginning and introduce a new product between vectors, the Geometric Product, showing its use in simple geometric constructions. With the help of this new product, generic rotations as a composition of two binary rotations (half turn rotations) are obtained. By inverting the historical process, the geometric product and its simple algebraic properties lead naturally to the definition of quaternions. We then define the tangent vector of rotation and the rotation vector.

After presenting some classical algebraic properties (composition of rotations, the Euler-Rodrigues formulas), and an important relation between additive and multiplicative rotations, the legitimate use of vectors for representing finite rotations is stated, showing that the lack of commutativity of finite rotations as an argument against the use of vectorial parameters is nonsense.

The next step is the link between the angular velocity vector and the time derivative of the different vectorial parameters. Again the elegance of quaternions in obtaining these tensors is evident. Very generic algebraic relations between these tensors and the rotation tensor are shown, and also between their derivatives. Here the concept of objectivity, borrowed from continuum mechanics, is crucial to prove these properties in a very generic way, the specific vectorial parameter used being irrelevant.

Having built the mathematical basis and the difficulty of finite rotations demystified, obvious integration algorithms present themselves. We restrict our presentation to the cited algorithms, referring to Bergweiler (2000) for a systematic development.

2. An Introduction to Quaternions and Rotations

In literature pertinent to rotations quaternions are frequently mentioned. However, it is difficult to find modern texts on them, old references describing quaternions as a generalization of complex numbers, what can be repulsive to some readers. Besides that, when dealing with rotations, it is not explained satisfactorily why they represent a rotation.

Our presentation will be very simple and objective, inverting the historic process and obtaining them from vectors, geometrically much more appellative than complex numbers. Quaternions will be the result of a new product between vectors, which we will call the *geometric product*, a concept borrowed from Hestenes (1987). This new product will be associative, permitting a division between vectors, thus enabling an effective algebraic manipulation.

2.1. The Geometric Product

We know at least seven different types of products involving vectors. Besides the product between a scalar and a vector we know the dot and vector products between two vectors. We also know the dyadic product (tensor product) and the external (Grassmann product) between two vectors and the (internal) product between a tensor and a vector. Finally in a two-dimensional space there also exists the complex product between two vectors.

The dot and vector products give only an incomplete measure of the relative position of its two factors. It is cumbersome to manipulate them algebraically because, although both products satisfy distributivity, only the scalar product is commutative and none of them is associative, so that a division is not possible.

When multiplying a scalar by another object we usually do not to use any multiplicative symbol, perhaps due to the fundamental character of this operation. For this same reason (soon we will learn that to be coherent we must use the same symbology) we will use none in our *geometric product* between two vectors **u** and **v**, which we define as

$$\mathbf{uv} \equiv -\mathbf{u}\cdot\mathbf{v} + \mathbf{u}\times\mathbf{v} \tag{1}$$

The denomination geometric product is suggested by the fact that this product contains the sine and the cosine of the angle between the two vectors, thus retaining all the geometric information of their relative position. That we are adding a vector to a scalar strikes the eye. In fact, formally perhaps we should have first introduced an " ordered" pair consisting of a scalar and a vector and then define the algebraic properties which justify the notation above. We should remember that when writing a complex number in the form $z = a + bi$ we also " add" objects of different gender, the ordered pair of real numbers remaining implicit. In practice we revoke the old saying that we can not add apples to oranges. In fact their sum results in a fruit salad, Hestenes (1987). The fruit salad in (1) is called a quaternion.

The negative signal in the geometric product is so defined in order to turn the product associative. To prove the associativity, which means that $(\mathbf{uv})\mathbf{w} = \mathbf{u}(\mathbf{vw})$, we should remember the formula for the double vector product and the relations for the mixed product. We must point out that when developing $(\mathbf{uv})\mathbf{w}$ we obtain $(-\mathbf{u}\cdot\mathbf{v} + \mathbf{u}\times\mathbf{v})\mathbf{w}$ so that **w** multiplies scalarly $\mathbf{u}\cdot\mathbf{v}$ and geometrically $\mathbf{u}\times\mathbf{v}$, two different multiplications but represented intentionally by the same " symbol" . The definition of the full quaternionic product overcomes this technical detail.

We have the following relations, which result from the scalar product being commutative and the vector product anti-commutative

$$\begin{aligned} \mathbf{u}\cdot\mathbf{v} &= -\frac{\mathbf{uv} + \mathbf{vu}}{2} \\ \mathbf{u}\times\mathbf{v} &= \frac{\mathbf{uv} - \mathbf{vu}}{2} \end{aligned} \tag{2}$$

Hestenes remarks that these relations show the fundamental character of the geometric product, since it is formally possible to recover both the scalar and the vector product, that is, the relative position of the vectors.

Noticing that $\mathbf{u}^2 \equiv \mathbf{uu} = -\mathbf{u}\cdot\mathbf{u}$ is a scalar, the multiplication of **u**, provided it is not zero, by

$$\mathbf{u}^{-1} \equiv \frac{\mathbf{u}}{\mathbf{u}^2} \tag{3}$$

at left or right results in unity (a scalar), so that we now have an inverse for our vectors. The associativity allows us then to make use of the important algebraic property of cancellation.

As distributivity is valid for both the scalar and the vector product, it is also valid for the geometric product, which is only a combination of them. We have seen that the geometric product is associative, but it is certainly not commutative (we find the same situation in matrix algebra). Technically, the set of quaternions with the addition and multiplication operations is called a division ring, where only commutativity is missing to form a field, like that of the real or complex numbers.

2.2. Example

The geometric product may be used with advantage in simple problems. A straight line is given by one of its points and a vector defining its orientation, A and **b**. For any given point O we want to find a point H on the line such that **h**=HO is perpendicular to **b**, Fig.(1).

As $\mathbf{h}\cdot\mathbf{b} = 0$, the geometric product is given by

$$\begin{aligned} \mathbf{hb} &= \mathbf{h}\times\mathbf{b} \\ &= \mathbf{a}\times\mathbf{b} \end{aligned} \tag{4}$$

where **a**=AO. After multiplying both sides by the inverse of **b,** and remembering that the geometric product is associative, the desired result is simply

$$\mathbf{h} = (\mathbf{a}\times\mathbf{b})\mathbf{b}^{-1} \tag{5}$$

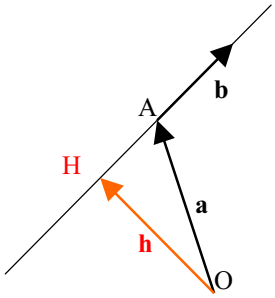


Figure 1.

If we want to use traditional notation we use (3), discard the dot product and obtain

$$\begin{aligned}\mathbf{h} &= \frac{(\mathbf{b}\times\mathbf{a})\times\mathbf{b}}{\mathbf{b}\cdot\mathbf{b}} \\ &= \mathbf{a} - \mathbf{b}\frac{\mathbf{a}\cdot\mathbf{b}}{\mathbf{b}\cdot\mathbf{b}}\end{aligned} \tag{6}$$

In more complex problems the lack of commutativity is a big problem and the geometric product looses its elegance. It was Gibbs in 1880 who claimed that the scalar and vector products should be the building blocks for vectorial systems. Hamilton, who invented quaternions in 1843, and his followers fought a bitter decade long battle and had to surrender for the simplicity of vectors and their products (Crowe, 1967).

2.3. Binary Rotations

Let us now focus on the today main use of quaternions. Initially we will interpret the expression

$$\mathbf{u}' \equiv -\mathbf{nun} \tag{7}$$

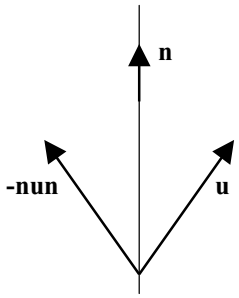


Figure 2.

where **n** is a versor and **u** any vector. We have that

$$\begin{aligned}-\mathbf{nun} &= -\mathbf{n}(-\mathbf{u}\cdot\mathbf{n} + \mathbf{u}\times\mathbf{n}) \\ &= (\mathbf{u}\cdot\mathbf{n})\mathbf{n} + \mathbf{n}\cdot(\mathbf{u}\times\mathbf{n}) - \mathbf{n}\times(\mathbf{u}\times\mathbf{n}) \\ &= (\mathbf{u}\cdot\mathbf{n})\mathbf{n} - \mathbf{n}\times(\mathbf{u}\times\mathbf{n}) \\ &= \mathbf{u}_{parallel} - \mathbf{u}_{perpendicular}\end{aligned} \tag{8}$$

so that in a plane Eq. (7) is a reflection of **u** on **n** and in space a rotation of 180 degrees around **n**, Fig.(2), called by Altmann (1986) a binary rotation.

2.4. Generic Rotations

We will now show that any rotation may be reached by two successive binary rotations. We want to turn **u** around an axis defined by a versor **e**, perpendicular to the plane defined by versors **n** and **m**, which have a relative angle of $\theta/2$ between each other. The vector **u** does in general not lie in this plane, but $\mathbf{u}_{nm}$ is the projection on it and $\mathbf{u}_e$ is what Hestenes (1987) calls the *rejection*, the component perpendicular to the plane, thus parallel to **e**. When we apply the binary rotation **n**, this component will result inverted. If we now apply a second binary rotation, this time around **m**, the effect will be that of reverting the inversion, obtaining the rotated vector **u**" , Fig. (3). The final result is that of a rotation of the vector **u** around **e** by an angle θ, given by

$$\mathbf{u}'' = \mathbf{mnunm} \tag{9}$$

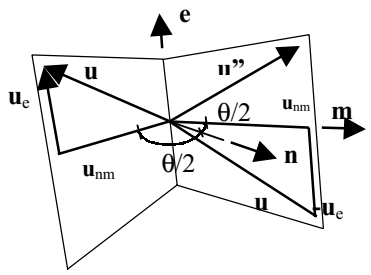


Figure 3.

If we observe that

$$\begin{aligned}\mathbf{nm} &= \mathbf{n}\times\mathbf{m} - \mathbf{n}\cdot\mathbf{m} \\ &= \sin\theta/2\,\mathbf{e} + \cos\theta/2\end{aligned} \tag{10}$$

we come to the conclusion that, provided **n** and **m** remain on the plane perpendicular to **e** and with a relative angle of $\theta/2$, their position relative to **u** does not matter, so that **mnunm** rotates any vector **u** by double the angle between **n** and **m**, from **n** to **m**.

2.5. The Quaternion

The geometric product of two vectors gives a scalar plus a vector whose " sum" is called a *quaternion*. Any sum of a scalar r and a vector **w** may be interpreted as the result of the geometric product of two vectors (not unique). In other words, for any r and **w** there exist **u** and **v** such that

$$Q \equiv r + \mathbf{w} = \mathbf{uv} \tag{11}$$

Thus we can find the inverse of Q provided that neither **u** nor **v** be zero, so neither r nor **w**

$$Q^{-1} = \mathbf{v}^{-1}\mathbf{u}^{-1} = \frac{\mathbf{vu}}{\mathbf{v}^2\mathbf{u}^2} = \frac{r-\mathbf{w}}{v^2u^2} \tag{12}$$

The conjugate of a quaternion is defined analogously to that of a complex number

$$Q^T \equiv r - \mathbf{w} = \mathbf{vu} \tag{13}$$

and it is easy to show that the inverse is simply given by

$$Q^{-1} = \frac{Q^T}{QQ^T} \tag{14}$$

also in analogy to complex numbers. A quaternion with the property $QQ^T=1$ is said to be *unitary*. A quaternion with a null scalar part is called a pure quaternion, a denomination originated when vectors did not exist formally.

2.6. Quaternions are Rotations

We described a rotation with the help of two geometric products. The corresponding quaternion expression is easily shown to be

$$\mathbf{u}'' = q\mathbf{u}q^T \tag{15}$$

where we defined

$$\begin{aligned} q &\equiv c + \mathbf{s} \\ \mathbf{s} &\equiv s\mathbf{e} \\ c &\equiv \cos\theta/2 \\ s &\equiv \sin\theta/2 \end{aligned} \tag{16}$$

Now, being λ a non-zero scalar and noticing that q is unitary, we can write

$$\mathbf{u}'' = \lambda q\mathbf{u}\frac{\lambda q^T}{\lambda^2 q q^T} \tag{17}$$

and as $\lambda q = Q$, a generic quaternion, we have

$$\mathbf{u}'' = Q\mathbf{u}Q^{-1} \tag{18}$$

That is, any quaternion, even not unitary, is a rotation. It suffices to use the inverse instead of the conjugate in the transformation formula. For example

$$\begin{aligned} T &\equiv q/c = 1 + \mathbf{t} \\ \mathbf{t} &\equiv \mathbf{s}/c \end{aligned} \tag{19}$$

so that a rotation may also be written as

$$\mathbf{u}'' = T\mathbf{u}T^{-1} \tag{20}$$

Cardona and Geradin (1988) call **t** the *rotational pseudo-vector*; we call it the *tangent rotation vector*. It will be very useful in what follows.

2.7. An Algebraic Justification

We sum up saying that a generic rotation may be represented by two binary rotations in sequence, what corresponds to the geometric product of two versors, which in turn is a unit quaternion.

We will now indicate a purely algebraic way to show that a quaternion represents a rotation. We need the almost trivial statement that if a linear transformation **R** preserves length, then $\mathbf{R}\mathbf{R}^T$=1, so that **R** must be a rotation. We point out, however, that **R** may be *proper* or *improper* depending on the sign of its determinant, that is, depending if the relative orientation of any vector triple is preserved or not.

The justification found in literature to show that a quaternion represents a rotation is to show that $q\mathbf{u}q^T$ preserves the length of **u**. This demonstration is trivial if we know the geometric properties of quaternions, what we already did with the help of the geometric product, but more lengthy if we had to use only brute force and the definition of the quaternion product. Besides, it would be necessary to show that the orthogonal tensor is proper, what is still more tedious.

There is no doubt that the simple geometric properties of a binary rotation associated to the geometric product give a much more efficient way to justify that a quaternion represents a rotation.

2.8. The Rodrigues Formulae

After a little algebra to eliminate the geometric products in (15) we obtain

$$\mathbf{u}'' = \mathbf{u} + 2c\mathbf{s}\times\mathbf{u} + 2\mathbf{s}\times(\mathbf{s}\times\mathbf{u}) \tag{21}$$

what is called by Altmann (1986) the conical transformation, and may be deduced at once by direct geometric reasoning. It was deduced in 1840 by the Frenchman of possible Portuguese origin (Altmann, 1986) Benjamin M. Olinde Rodrigues.

We can also use traditional vector notation for the composition of rotations. A body is submitted to a rotation given by $\mathbf{t}_1$, followed by a second rotation $\mathbf{t}_2$. In analogy to tensor composition of rotations, the quaternion of the compound rotation is given multiplying T_1 and T_2 in the proper order

$$T_2 T_1 = 1 - \mathbf{t}_2 \cdot \mathbf{t}_1 + \mathbf{t}_2 + \mathbf{t}_1 + \mathbf{t}_2 \times \mathbf{t}_1 \tag{22}$$

The resulting quaternion is not normalized in the sense of Eq. (19), were we saw that multiplying a quaternion by a scalar does not affect the rotation it represents. The normalized quaternion is

$$T_{21} = 1 + \frac{\mathbf{t}_2 + \mathbf{t}_1 + \mathbf{t}_2 \times \mathbf{t}_1}{1 - \mathbf{t}_2 \cdot \mathbf{t}_1} \tag{23}$$

which means that the tangent rotation vector of the compound rotation is given by

$$\mathbf{t}_{21} = \frac{\mathbf{t}_2 + \mathbf{t}_1 + \mathbf{t}_2 \times \mathbf{t}_1}{1 - \mathbf{t}_2 \cdot \mathbf{t}_1} \tag{24}$$

We will use this relation in a moment. It is evident from this formula that the composition of rotations is not commutative. It has been observed that, although the axis of the compound rotation is different if we reverse the order, the modulus, that is, the rotation angle, is the same. This is evident since $\mathbf{t}_2 \times \mathbf{t}_1$ is orthogonal both to $\mathbf{t}_1$ and $\mathbf{t}_2$.

2.9. Change of Reference

In the numeric treatment of rotation dynamics the most convenient approach is to specify a reference configuration of the rigid body under study and to describe the current position by means of a rotation (and a translation) relative to this reference position. But we may want to define a second reference position. What we will do now is to relate these two references. For the original reference $\mathbf{u}_1$ we have

$$\mathbf{u} = T(\mathbf{t})\mathbf{u}_1 T^{-1}(\mathbf{t}) \tag{25}$$

If we use the second reference $\mathbf{u}_2$, for the same vector $\mathbf{u}$ we may write

$$\mathbf{u} = T(\Delta\mathbf{t}^*)\mathbf{u}_2 T^{-1}(\Delta\mathbf{t}^*) \tag{26}$$

where $\Delta\mathbf{t}^*$ is the rotation relative to the second reference. Soon it will be clear why we choose this writing. The second reference is rotated from the original one by $\mathbf{t}_2$

$$\mathbf{u}_2 = T(\mathbf{t}_2)\mathbf{u}_1 T^{-1}(\mathbf{t}_2) \tag{27}$$

So extracting $\mathbf{u}_1$ and inserting in (25) we have

$$\mathbf{u} = T(\mathbf{t})T^{-1}(\mathbf{t}_2)\mathbf{u}_2 T(\mathbf{t}_2)T^{-1}(\mathbf{t}) \tag{28}$$

which compared to (26) results in

$$\begin{aligned} T(\Delta\mathbf{t}^*) &= T(\mathbf{t})T^{-1}(\mathbf{t}_2) \\ &= T(\mathbf{t})T(-\mathbf{t}_2) \end{aligned} \tag{29}$$

Defining $\Delta\mathbf{t} \equiv \mathbf{t} - \mathbf{t}_2$ and using the composition formula (24) we have finally

$$\Delta\mathbf{t}^* = \frac{\Delta\mathbf{t} + \mathbf{t}_2 \times \Delta\mathbf{t}}{1 + \mathbf{t}_2 \cdot \mathbf{t}_2 + \mathbf{t}_2 \cdot \Delta\mathbf{t}} \tag{30}$$

This is the key to transform an additive relation to a multiplicative one. In fact we may rewrite (29) as

$$T(\Delta\mathbf{t}^*)T(\mathbf{t}_2) = T(\mathbf{t}_2 + \Delta\mathbf{t}) \tag{31}$$

and we hope it is now clear why we have written $\Delta\mathbf{t}^*$, as it represents a multiplicative increment of the rotation (a tensor or quaternion), whereas $\Delta\mathbf{t}$ represents an additive increment, but this time of the tangent rotation vector. We remark that there is no interest in an additive increment for the rotation tensor, because it would not be orthogonal and difficult to obtain directly. If $\Delta\mathbf{t}$ is small there is a strong connection between $\Delta\mathbf{t}^*$ and the angular velocity vector, see (38).

2.10. The Rotation Vector

After having introduced the tangent vector of rotation we can now define the *rotation vector* itself, which is simply

$$\boldsymbol{\theta} \equiv \theta\, \mathbf{e} \tag{32}$$

Let us now introduce the notation that a bar under a vector represents the tensor equivalent to the vector product, that is, if $\mathbf{w} = \mathbf{u} \times \mathbf{v}$ then $\underline{\mathbf{u}}$ is a tensor such that $\mathbf{w} = \underline{\mathbf{u}}\mathbf{v}$. Using the rotation vector, the rotation (21) is given by

$$\mathbf{R} = 1 + \frac{sin\theta}{\theta}\underline{\boldsymbol{\theta}} + \frac{1}{2}\left(\frac{sin\theta/2}{\theta/2}\underline{\boldsymbol{\theta}}\right)^2 \tag{33}$$

which is the Euler-Rodrigues formula. Inspired by the quaternion notation we also introduced a notation where a unit tensor is implicit when adding a scalar to a tensor. The corresponding expression using the tangent vector of rotation is simply

$$\mathbf{R} = 1 + 2\frac{\underline{\mathbf{t}} + \underline{\mathbf{t}}^2}{1 + \mathbf{t}\cdot\mathbf{t}} = (1 + \underline{\mathbf{t}})(1 + \underline{\mathbf{t}})^{-T} \tag{34}$$

The expansion of (33) in Taylor series is given by

$$\mathbf{R} = 1 + \underline{\boldsymbol{\theta}} + \frac{1}{2}\underline{\boldsymbol{\theta}}^2 + \frac{1}{3!}\underline{\boldsymbol{\theta}}^3 + \ldots \equiv \exp(\underline{\boldsymbol{\theta}}) \tag{35}$$

which has a very simple geometric interpretation (Argyris, 1982), and

$$\mathbf{R} = 1 + 2\left(\underline{\mathbf{t}} + \underline{\mathbf{t}}^2 + \underline{\mathbf{t}}^3 + \ldots\right) \tag{36}$$

2.11. The Prejudice

The classical Euler angles are called *rotation parameters*. The rotation vector and the tangent rotation vector are also rotation parameters. Although vectorial rotation parameters have a very simple geometric interpretation, only now is their use spreading out. Probably this late recognition is due to the fact that the lack of commutativity when composing rotations is in apparent conflict to the parallelogram rule for vector addition and vector commutativity (e.g. Thomson, 1963). Rotation vectors are legitimate rotation parameters and can be vectorially added without restrictions, but their sum does not compare to the rotation vector of the composition of rotations. There is no dilemma whatsoever if we recognize that the composition of rotations is a product of tensors, in general different to the vector sum of rotation vectors. We restate this thoughts in the following way:

It is a prejudice to demand that the vector sum of rotation parameters be equivalent to the parameter for the multiplication of rotation tensors. The vector sum of rotation parameters is commutative, does represent a rotation, but its interpretation depends on the specific parameter. The multiplication of tensors, which is not commutative, is a different geometric operation.

2.12. Closure

We believe that the definition of the geometric product between vectors and its use to obtain a composition of binary rotations is the most natural way to introduce quaternions because it is based on strong geometric motivation.

The representation of a rotation by two successive reflections is known, (Hestenes, 1987). Apparently it is easier to use reflections than rotations in spaces with dimension greater than 3, but in three-dimensional space a reflection is a binary rotation with reversed sign, so that a generic rotation is a composition of two binaries (Pars, 1965), a trivial consequence, physically much more appelative, but scarcely mentioned in literature.

We presented an interesting complement to traditional vector algebra, inspirated by Hestenes (1987), who, however, used the much more powerful n-dimensional Clifford Algebra with their external products. The quaternion is only a by-product and his geometric product contains a bivector, an oriented area, in place of our vector product.

3. Angular Velocity

Let us now develop an expression that relates the angular speed to the derivative of the rotation vector. The tensor so obtained is fundamental for the numeric treatment of rotations. We will change notation and put a bar over a vector if it is in the reference configuration. Taking the derivative of (15) we can easily show that the angular speed is given by

$$\boldsymbol{\omega} = 2\dot{q}q^T \tag{37}$$

Noticing that $q = cT$ and also that the scalar component of the above equation is zero we obtain finally

$$\begin{aligned} \boldsymbol{\omega} &= \mathbf{Y}\dot{\mathbf{t}} \\ \mathbf{Y} &\equiv 2\frac{1+\underline{\mathbf{t}}}{1+\mathbf{t}\cdot\mathbf{t}} \end{aligned} \tag{38}$$

Also, developing (37) directly results in

$$\begin{aligned} \boldsymbol{\omega} &= \boldsymbol{\Gamma}\dot{\boldsymbol{\theta}} \\ \boldsymbol{\Gamma} &\equiv 1 + \frac{2s^2}{\boldsymbol{\theta}\cdot\boldsymbol{\theta}}\underline{\boldsymbol{\theta}} + \frac{1-\frac{2sc}{\theta}}{\boldsymbol{\theta}\cdot\boldsymbol{\theta}}\underline{\boldsymbol{\theta}}^2 \end{aligned} \tag{39}$$

It is common to find the transpose of these tensors in literature. In Bergweiler (2000) we present a straightforward method to calculate their inverse. We remark that the inverses have a singularity so that updated Lagrangian numerical methods are preferred.

We will also need the first and second derivatives. As the derivative is a linear operator, we can define second order tensors **V** and **W** so that for any vector **a**,

$$\begin{aligned} D\boldsymbol{\Gamma}(\boldsymbol{\theta})[d\boldsymbol{\theta}]\mathbf{a} &= \mathbf{V}(\boldsymbol{\theta},\mathbf{a})d\boldsymbol{\theta} \\ D^2\boldsymbol{\Gamma}(\boldsymbol{\theta})[d\boldsymbol{\theta},d\boldsymbol{\theta}]\mathbf{a} &= \mathbf{W}(\boldsymbol{\theta},d\boldsymbol{\theta},\mathbf{a})d\boldsymbol{\theta} \end{aligned} \tag{40}$$

4. Vectorial Rotation Parameters and Objectivity

Schulz and Filippou (2001) remark that the Euler-Rodrigues formula (33) satisfies the geometrically obvious fact that for every rotation **Q**,

$$\mathbf{R}(\mathbf{Q}\boldsymbol{\theta}) = \mathbf{Q}\mathbf{R}(\boldsymbol{\theta})\mathbf{Q}^T \tag{41}$$

We recognize that this is the requirement of *objectivity*, a concept borrowed from continuum mechanics. By taking the derivative of (41) we obtain

$$\boldsymbol{\Gamma}(\mathbf{Q}\boldsymbol{\theta}) = \mathbf{Q}\boldsymbol{\Gamma}(\boldsymbol{\theta})\mathbf{Q}^T \tag{42}$$

so that the tensor $\boldsymbol{\Gamma}$ is also objective, which should not be a surprise. As an example, the rotation function given by $\hat{\mathbf{R}}(\boldsymbol{\theta}) = \exp(\mathbf{Q}\boldsymbol{\theta})$, where $\mathbf{Q}$ is any constant rotation different from the identity tensor, is not objective, and the properties derived below do not apply. But it is easy to show that $\hat{\mathbf{R}}(\boldsymbol{\theta}) = \exp(f(\boldsymbol{\theta}\cdot\boldsymbol{\theta})\boldsymbol{\theta})$, where f is any scalar function, is objective, see Eq. (34), although $\hat{\mathbf{R}}(\boldsymbol{\theta}) = \exp(f(\boldsymbol{\theta})\boldsymbol{\theta})$ is not. In this section we will obtain useful expressions based only on the concept of objectivity.

Before continuing with the rotation function we will first investigate the class of generic tensor functions of a vectorial parameter $\boldsymbol{\rho}$, $\boldsymbol{\Gamma} = \hat{\boldsymbol{\Gamma}}(\boldsymbol{\rho})$. The only requirement for this functions is that they must be objective

$$\hat{\boldsymbol{\Gamma}}(\exp(\underline{\boldsymbol{\theta}})\boldsymbol{\rho}) = \exp(\underline{\boldsymbol{\theta}})\hat{\boldsymbol{\Gamma}}(\boldsymbol{\rho})\exp^T(\underline{\boldsymbol{\theta}}) \tag{43}$$

for any rotation vector $\boldsymbol{\theta}$.

We know that $\exp(\lambda\underline{\boldsymbol{\theta}})\boldsymbol{\theta} = \boldsymbol{\theta}$ for any λ. Let us prove that $\boldsymbol{\rho}$ is also an axis of $\hat{\boldsymbol{\Gamma}}$. Setting $\boldsymbol{\theta} = \lambda\boldsymbol{\rho}$ in (43) we obtain

$$\hat{\boldsymbol{\Gamma}}(\boldsymbol{\rho}) = \exp(\lambda\underline{\boldsymbol{\rho}})\hat{\boldsymbol{\Gamma}}(\boldsymbol{\rho})\exp^T(\lambda\underline{\boldsymbol{\rho}}) \tag{44}$$

which means that $\boldsymbol{\rho}$ is a symmetry axis of $\boldsymbol{\Gamma}$, an eigenvector. For an objective rotation we then have

$$\hat{\mathbf{R}}(\boldsymbol{\rho})\boldsymbol{\rho} = \boldsymbol{\rho} \tag{45}$$

Now we will prove that $\hat{\boldsymbol{\Gamma}}(-\boldsymbol{\rho}) = \hat{\boldsymbol{\Gamma}}^{\mathrm{T}}(\boldsymbol{\rho})$. Multiplying each side of (44) by $\mathbf{r}$ and $\mathbf{e}$, where $\mathbf{r}$ is any vector perpendicular to $\mathbf{e}$, which is the versor of $\boldsymbol{\rho}$, and letting $\lambda\boldsymbol{\rho} = \pi\mathbf{e}$, a binary rotation, we have

$$\mathbf{r}\cdot\hat{\boldsymbol{\Gamma}}(\boldsymbol{\rho})\mathbf{e} = -\mathbf{r}\cdot\hat{\boldsymbol{\Gamma}}(\boldsymbol{\rho})\mathbf{e} \tag{46}$$

so that

$$\mathbf{r}\cdot\hat{\boldsymbol{\Gamma}}(\boldsymbol{\rho})\mathbf{e} = 0 \tag{47}$$

which is a restatement that $\mathbf{e}$ is an eigenvector of $\boldsymbol{\Gamma}$ and a justification for (45). Using a similar argument we can prove that

$$\mathbf{e}\cdot\hat{\boldsymbol{\Gamma}}(\boldsymbol{\rho})\mathbf{r} = 0 \tag{48}$$

what is not so obvious geometrically. Now the transpose is such that $\mathbf{a}\cdot\hat{\boldsymbol{\Gamma}}(\boldsymbol{\rho})\mathbf{b} = \mathbf{b}\cdot\hat{\boldsymbol{\Gamma}}^T(\boldsymbol{\rho})\mathbf{a}$ for any vectors $\mathbf{a}$ and $\mathbf{b}$. Let $\boldsymbol{\Pi}$ be a binary rotation perpendicular to $\boldsymbol{\rho}$ so that $\boldsymbol{\Pi}\boldsymbol{\rho} = -\boldsymbol{\rho}$ and $\boldsymbol{\Pi} = \boldsymbol{\Pi}^T$. Thus

$$\begin{aligned}\mathbf{a}\cdot\hat{\boldsymbol{\Gamma}}(-\boldsymbol{\rho})\mathbf{b} &= \mathbf{a}\cdot\boldsymbol{\Pi}\hat{\boldsymbol{\Gamma}}(\boldsymbol{\rho})\boldsymbol{\Pi}^T\mathbf{b}\\ &= \boldsymbol{\Pi}\mathbf{a}\cdot\hat{\boldsymbol{\Gamma}}(\boldsymbol{\rho})\boldsymbol{\Pi}\mathbf{b}\\ &= \mathbf{b}\cdot\hat{\boldsymbol{\Gamma}}(\boldsymbol{\rho})\mathbf{a}\\ &= \mathbf{a}\cdot\hat{\boldsymbol{\Gamma}}^T(\boldsymbol{\rho})\mathbf{b}\end{aligned} \tag{49}$$

The last but one passage is not obvious but may be proved decomposing $\mathbf{a}$ and $\mathbf{b}$ in components parallel and perpendicular to $\mathbf{e}$ and using (47) and (48). Finally, as $\mathbf{a}$ and $\mathbf{b}$ are arbitrary, we have

$$\hat{\boldsymbol{\Gamma}}(-\boldsymbol{\rho}) = \hat{\boldsymbol{\Gamma}}^T(\boldsymbol{\rho}) \tag{50}$$

5. Additional Consequences of Objectivity

In what follows we will drop the caret to simplify notation, and keep in mind that the properties do not depend on the specific set used. The symbols $\{\boldsymbol{\theta}, \mathbf{R}, \boldsymbol{\Gamma}, \mathbf{V}, \mathbf{W}\}$ will represent any specific set, so that $\boldsymbol{\theta}$ represents any vectorial parameter, not necessarily the rotation vector. We remark that the requirement of objectivity lies only on $\{\mathbf{R}, \boldsymbol{\theta}\}$, as in consequence it is automatically fulfilled by the derived tensors.

The properties from last section are enough to prove the following

$$\mathbf{R} = 1 + \underline{\boldsymbol{\theta}}\boldsymbol{\Gamma} \tag{51}$$

$$\mathbf{R}\boldsymbol{\Gamma} = \boldsymbol{\Gamma}^T \tag{52}$$

The last one may be found in Pimenta (1987) and is useful in numerical implementations because it saves a tensor multiplication, and is used on the algorithms below. Also we can prove that for any vectors $\mathbf{a}$ and $\mathbf{b}$,

$$\begin{aligned}&\mathbf{V}(\boldsymbol{\theta},\mathbf{a})\mathbf{b} = \mathbf{V}(\boldsymbol{\theta},\mathbf{b})\mathbf{a} + (\boldsymbol{\Gamma}\mathbf{b})\times\boldsymbol{\Gamma}\mathbf{a}\\ &\mathbf{V}^T(\boldsymbol{\theta},\mathbf{a})\mathbf{b} = \mathbf{V}^T(-\boldsymbol{\theta},-\mathbf{b})\mathbf{a}\\ &\mathbf{V}^T(\boldsymbol{\theta},\mathbf{a}) = \mathbf{V}(\boldsymbol{\theta},\mathbf{a}) + \boldsymbol{\Gamma}\underline{\mathbf{a}}\boldsymbol{\Gamma}^T\\ &\mathbf{R}\mathbf{V}(\boldsymbol{\theta},\mathbf{a})\mathbf{b} = \mathbf{V}(-\boldsymbol{\theta},-\mathbf{b})\mathbf{a}\end{aligned} \tag{53}$$

and similarly for any $\mathbf{a}$,$\mathbf{b}$ and $\mathbf{c}$,

$$\begin{aligned}
&\mathbf{W}(\boldsymbol{\theta},\mathbf{a},\mathbf{b})\mathbf{c} = \mathbf{W}(\boldsymbol{\theta},\mathbf{c},\mathbf{b})\mathbf{a} \\
&\mathbf{W}^T(\boldsymbol{\theta},\mathbf{a},\mathbf{b})\mathbf{c} = \mathbf{W}^T(-\boldsymbol{\theta},\mathbf{a},\mathbf{c})\mathbf{b} \\
&\mathbf{W}\left(\boldsymbol{\theta},\dot{\boldsymbol{\theta}},\mathbf{b}\right) = \mathbf{W}^T\left(\boldsymbol{\theta},\dot{\boldsymbol{\theta}},\mathbf{b}\right) - \dot{\boldsymbol{\Gamma}}^T\,\underline{\mathbf{b}}\boldsymbol{\Gamma} - \boldsymbol{\Gamma}^T\underline{\mathbf{b}}\,\dot{\boldsymbol{\Gamma}}
\end{aligned} \tag{54}$$

These expressions are very helpful for developing the equations of non-linear rods and to prove algebraically that the stiffness matrix is symmetric (Bergweiler, 2000).

6. The Additive Algorithm

In Simo and Wong (1991) we find a multiplicative algorithm for the integration of the movement of a rigid body such that momentum and kinetic energy are conserved. Let $\bar{\mathbf{I}}$ be the inertia tensor at the reference configuration, $\mathbf{L}$ the angular momentum and $\mathbf{m}_{1/2}$ the external moment calculated at half the time step (mid point rule). Given $\mathbf{R}_1, \bar{\boldsymbol{\omega}}_1, \mathbf{L}_1$, the state at stage 1, find the rotation increment $\Delta\bar{\boldsymbol{\theta}}$ such that, after a time step Δt, momentum in stage 2 is conserved:

Multiplicative Algorithm

$$\begin{aligned}
&\Delta\bar{\boldsymbol{\theta}} := \Delta t\,\bar{\boldsymbol{\omega}}_1 \\
&\texttt{repeat} \\
&\quad \bar{\boldsymbol{\omega}}_2 := \frac{2\Delta\bar{\boldsymbol{\theta}}}{\Delta t} - \bar{\boldsymbol{\omega}}_1 \\
&\quad \Delta\bar{\mathbf{R}} := \exp\left(\Delta\underline{\bar{\boldsymbol{\theta}}}\right) \\
&\quad \mathbf{R}_2 := \mathbf{R}_1\Delta\bar{\mathbf{R}} \\
&\quad \Delta\bar{\boldsymbol{\Gamma}}^{-T} := \boldsymbol{\Gamma}^{-1}\left(-\Delta\bar{\boldsymbol{\theta}}\right) \\
&\quad \bar{\mathbf{L}}_2 := \bar{\mathbf{I}}\,\bar{\boldsymbol{\omega}}_2 \\
&\quad \mathbf{L}_2 := \mathbf{R}_2\bar{\mathbf{L}}_2 \\
&\quad \mathbf{g} := \mathbf{L}_2 - \mathbf{L}_1 - \Delta t\mathbf{m}_{1/2} \\
&\quad \mathbf{K} := \mathbf{R}_2\left(\frac{2}{\Delta t}\bar{\mathbf{I}}\Delta\bar{\boldsymbol{\Gamma}}^{-T} - \underline{\bar{\mathbf{L}}}_2\right) \\
&\quad \Delta\Delta\bar{\mathbf{o}} := -\mathbf{K}^{-1}\mathbf{g} \\
&\quad \Delta\bar{\boldsymbol{\theta}} := \log\left(\Delta\bar{\mathbf{R}}\exp\left(\Delta\Delta\bar{\mathbf{o}}\right)\right) \\
&\texttt{until } \left|\Delta\Delta\bar{\mathbf{o}}\right| < tolerance\left(\left|\Delta\bar{\boldsymbol{\theta}}\right|\right)
\end{aligned}$$

Additive Algorithm

$$\begin{aligned}
&\Delta\bar{\boldsymbol{\theta}} := \Delta t\,\bar{\boldsymbol{\omega}}_1 + \frac{\Delta t^2}{2}\dot{\boldsymbol{\omega}}_1 \\
&\texttt{repeat} \\
&\quad \bar{\boldsymbol{\omega}}_2 := \frac{2\Delta\bar{\boldsymbol{\theta}}}{\Delta t} - \bar{\boldsymbol{\omega}}_1 \\
&\quad \Delta\bar{\mathbf{R}} := \exp\left(\Delta\underline{\bar{\boldsymbol{\theta}}}\right) \\
&\quad \mathbf{R}_2 := \mathbf{R}_1\Delta\bar{\mathbf{R}} \\
&\quad \Delta\bar{\boldsymbol{\Gamma}}^{T} := \boldsymbol{\Gamma}\left(-\Delta\bar{\boldsymbol{\theta}}\right) \\
&\quad \bar{\mathbf{L}}_2 := \bar{\mathbf{I}}\,\bar{\boldsymbol{\omega}}_2 \\
&\quad \mathbf{L}_2 := \mathbf{R}_2\bar{\mathbf{L}}_2 \\
&\quad \mathbf{g} := \mathbf{L}_2 - \mathbf{L}_1 - \Delta t\mathbf{m}_{1/2} \\
&\quad \mathbf{K} := \mathbf{R}_2\left(\frac{2}{\Delta t}\bar{\mathbf{I}} - \underline{\bar{\mathbf{L}}}_2\Delta\bar{\boldsymbol{\Gamma}}^{T}\right) - \Delta t D\mathbf{m}_{1/2}\left(\Delta\bar{\boldsymbol{\theta}}\right) \\
&\quad \Delta\Delta\bar{\boldsymbol{\theta}} := -\mathbf{K}^{-1}\mathbf{g} \\
&\quad \Delta\bar{\boldsymbol{\theta}} := \Delta\bar{\boldsymbol{\theta}} + \Delta\Delta\bar{\boldsymbol{\theta}} \\
&\texttt{until } \left|\Delta\Delta\bar{\boldsymbol{\theta}}\right| < tolerance\left(\left|\Delta\bar{\boldsymbol{\theta}}\right|\right)
\end{aligned} \tag{55}$$

If the rotations are stored as quaternions the log extraction is easily done in the multiplicative algorithm. We observe that Simo and Wong do not take the derivative of the moment, which is technically required for a faster convergence but in simulations revealed of no influence. The additive version updates the rotation vector, whereas the original algorithm technically updates the " integral of the angular velocity" , see Bergweiler (2000) for a thorough discussion.

7. Final Remark

Our additive linearization is a mathematical consequence of the Newton-Raphson root finding algorithm and so needs no physical justification. This very fact was the key to realize that the multiplication of tensors and sum of vectors are two different ways to compose rotations, as explained earlier in section 2.11. We also think that our version of the linearization is " cleaner" and hope that the comparison of both algorithms helps to dissipate the already mentioned prejudice.

8. References

Altmann, S.L., 1986, **Rotations, quaternions and double groups**, Clarendon Press, Oxfford.

Argyris, J.A., 1982, **An excursion into large rotations**, Computer Methods in Applied Mechanics and Engineering, Vol.32, pp.85-155.

Bergweiler, M.González, 2000, **Tratamento numérico de rotações finitas e sua aplicação em corpos rígidos e barras**, Thesis (Ph.D) - Escola Politécnica da Universidade de São Paulo, Departamento de Engenharia Civil.

Cardona, A. and Geradin, M., 1988, **A beam finite element non-linear theory with finite rotations**, International journal for numerical methods in engineering, Vol.26, pp.2403-2438.

Crisfield, M.A., 1997, **Non-linear finite element analysis of solids and structures. Volume 2: advanced topics**, John Wiley & Sons, Chichester.

Crowe, M.J., 1967, **A history of vector analysis**, Dover, New York. [Reprint,1994]

Goldstein, H.,1980, **Classical mechanics**, Addison Wesley, Reading.

Hestenes, D., 1987, **New foundations for classical mechanics**, D.Reidel, Dordrecht.

Pars, L.A., 1965, **A treatise on analytical dynamics**, Heinemann, London.

Pimenta, P.M., 1987, **Rotações finitas,** Boletim técnico PEF/8714, Escola Politécnica da Universidade de São Paulo.

Schulz, M., Filippou, F.C, 2001, **Non-linear spatial Timoshenko beam element with curvature interpolation**, International journal for numerical methods in engineering, Vol.50, pp.761-785.

Simo, J.C., Wong, K., 1991, **Unconditionally stable algorithms for rigid body dynamics that exactly preserve energy and angular momentum**, International journal for numerical methods in engineering, Vol.31, pp.19-52.

Thomson, W.T., 1963, **Introduction to space dynamics**, Dover, New York. [Reprint,1986]

Wilson, E.B., 1901, **Vector analysis (of Gibbs)**. 8th pr. 1943, Yale University Press, New Haven. [Reprint by Gordon and Breach, NewYork, 1966]

XVI CONGRESSO BRASILEIRO DE ENGENHARIA MECÂNICA
16th BRAZILIAN CONGRESS OF MECHANICAL ENGINEERING

NONLINEAR DYNAMIC MODELING AND CONTROL OF AN UNDER-ACTUATED FLEXIBLE STRUCTURE MOUNTED MICROMANIPULATOR SYSTEM

Luiz Carlos S. Góes
Instituto Tecnológico de Aeronáutica – CTA, 12.228-900 São José dos Campos, SP, Brazil
goes@mec.ita.cta.br

Alexsander Costa
Instituto Tecnológico de Aeronáutica – CTA, 12.228-900 São José dos Campos, SP, Brazil
alexsta@mec.ita.cta.br

Álvaro M. S. Soares
Departamento de Engenharia Mecânica – UNITAU, 12.080-000 Taubaté, SP, Brazil
alvaro@inf.unitau.br

Francisco J. Grandinetti
Departamento de Engenharia Mecânica – UNITAU, 12.080-000 Taubaté, SP, Brazil

Abstract. *The objective of this paper is to model and control an underactuated Flexible Structure Mounted Micromanipulator System (FSMS). The system consists of one degree-of-freedom rigid micromanipulator mounted on the tip of a flexible beam. The beam deflection and joint motion occur in a plane perpendicular to the gravity direction. The Lagrangian formulation for mixed rigid and flexible coordinates is employed to derive the nonlinear dynamic model of the FSMS. The assumed mode method with two flexible modes is used to approximate the dynamics of the flexible supporting beam. It is analytically shown that the model-based control law leads to a stable closed-loop response. The results are illustrated by simulation of an FSMS assembled at the Mechatronics Laboratory of the University of Taubaté.*

Keywords *Modeling, Nonlinear Control, FSMS.*

1. Introduction

Investigations on nonlinear dynamic modeling and control of Flexible Structure Mounted Manipulator System (FSMS) have attracted a great deal of interest due to special applications related to free-floating space robots, and operation of long-reach flexible robotic manipulator arms with large operational work space and simultaneous high-accuracy of its end-effector (Xu and Kanade, 1993). Typical practical applications of this kind of system include the on-orbit teleoperated control experiments with the ETS VII free-flying space robot (Yoshida, Nenchev, Inaba and Oda, 2000), and the experimental control of the so-called micro-macro manipulator system, where a small dexterous (micro) manipulator is mounted on a long-reach flexible (macro) manipulator structure (Lew, and Trudnowski, 1996).

The robotic systems mounted on a flexible link is a dynamical system characterized by infinite number of degree of freedom, governed by nonlinear, coupled, ordinary and partial differential equations. The exact solution of such a mathematical model is normally not feasible, and the infinite dimensional model imposes severe constrains on the design of active controllers as well (Shiffman, 1993). Hence it's necessary to discretize the continuous system by using assumed modes, finite elements (FEM), or lumped-parameters. The link deflection with respect to its rigid configuration is assumed to be small and a linear theory of elasticity is selected. The dynamic system is then modeled as an underactuated robotic system (Jain and Rodriguez, 1993; Oriolo and Nakamura, 1991) with active or controlled degrees of freedom represented by the joint coordinates of the micro-macro manipulator, and passive degrees of freedom represented by the modal coordinates of the flexible links (Nenchev, 2000).

Previous experimental work on FSMS and LRMS (Long Reach Manipulator System) include among others: the MIT LRMS named "Shaky" (Mavroidis, Dubowsky, Raju, 1995), the VES-II: Vehicle Emulation System Model II (Dubowsky *et al.*, 1994), the Micro/Macro Flexible-Link Manipulator test bed (Lew, and Trudnowski, 1996), and the Dual-Arm TREP-II LRMS Experiment (Nenchev, Yoshida, and Uchiyama, 1999). The later experiment consists of a planar two-joint rigid link manipulator attached to the free end of a double beam representing an elastic base. The latter has inspired the design and the assembly of a similar FSMS experiment at the University of Taubaté (UNITAU).

The objective of this work is to describe the nonlinear dynamic modeling, the positional and active vibration control of a FSMS system assembled at the UNITAU Mechatronics Laboratory. The flexible structure is a cantilever beam mounted on the tip of a three degree of freedom cylindrical (r, θ, z), rigid Macro-Manipulator (Ma-M). The flexible Euler-Bernoulli beam has a low stiffness in the horizontal direction, in the same plane of motion of the polar Micro-Manipulator (Mi-M). In this study it is assumed that the cylindrical Ma-M is used only to position the FSMS in its workspace, and after this is kept stationary. Figure 1 shows a view of the Mi-M mounted on the flexible beam (Góes, Grandinetti, Soares and Nenchev, 2000).

In this work, the analytical model of the system is derived from Lagrange's equation and the assumed mode method. The Mi-M has only one rotational degree of freedom and the motion of the cylindrical Ma-M is used only to locate the FSMS on its workspace, and excite the vibracional degrees of freedom of the flexible beam. A generalized flexible homogeneous transformation and symbolic manipulation are applied to derive the equations of motion of the complete Ma-Mi-Manipulator, (Góes, Grandinetti, and Soares, 2000). In this work a simpler model is considered, where the generalized coordinates are the angular displacement of the Mi-M joint, $q_r(t)$, and the modal coordinates, $\mathbf{q}_f(t)$, associated with the elastic displacement of the beam, w(x, t). The generalized force is the joint torque of the Mi-M system, u(t).

Figure 1. View of the micromanipulator mounted on the tip of flexible structure of the macro-manipulator.

2. Dynamic Equations of the FSMS System.

The elastic displacement of the beam, w(x,t), and the rigid joint angle, $q_r(t)$, are shown schematically in Figure 2. Here, L_f is the undeformed length of the beam, m_L is the mass of the motor armature, J_m is the motor inertia, m_r is the link mass, and J_r is the link moment of inertia with respect to the CG.

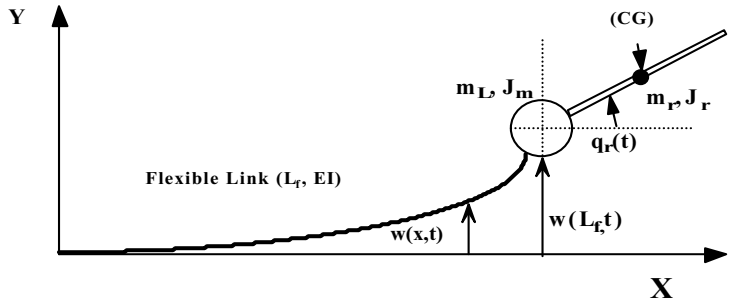


Figure 2. Schematic view of the micromanipulator mounted on a flexible structure.

The general form of Lagrange's equation is given bellow, where the index *(k=1,2,3)* includes both the active, $q_1=q_r(t)$, and passive, $q_2=q_{f1}(t)$, $q_3=q_{f2}(t)$, degrees of freedom of the FSMS. The generalized forces, τ_k, for this system is obtained through the principle of virtual work, where u(t) is the applied torque at the micromanipulator joint,

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_k}\right)-\frac{\partial T}{\partial q_k}+\frac{\partial V}{\partial q_k}=\tau_k \quad (k=1,2,3) \tag{1}$$

$$\delta W = u\,\delta q_r$$

In the assumed modes method, the flexibility is approximated by a finite series of space-dependent functions multiplied by time-dependent amplitude functions (flexible coordinates), as given by Eq. (2). In this work, we use as admissible function a special class of shape functions (Junkins, 1989), given by Eq (3) that can be considered to be comparison functions of a clamped-free beam with a tip mass, since it satisfies both geometric and natural boundary conditions. It is important to note that , here, the chosen shape function only satisfy the geometric boundary conditions of FSMS at the clamped end, and hence belong to the class of admissible functions only, without any loss of generality of the methodology (Theodore and Ghosal, 1995).

In this study only the first two modes are retained since the actuators and sensors cannot operate in the high frequency range associated with higher order modes. Thus, the elastic deformation, w(x,t) is written as

$$w(x,t) = \sum_{i=1}^{2} \phi_i(x) q_i(t) \tag{2}$$

$$\phi_i(x) = 1 - \cos\left(i\pi x / L_f\right) + 0.5(-1)^{i+1}\left(i\pi x / L_f\right)^2 \tag{3}$$

where, $\phi_i(x)$ (i=1,2) are the admissible shape function, $q_i(t)$ are generalized modal coordinates, and L_f is length of the flexible link.

The total kinetic energy of the FSMS system includes both the kinetic energy of the flexible beam with a tip mass, m_L, due to the motor armature, given by Eq. (4), and the kinetic energy of the micromanipulator system due to the motion of its joint and rigid link. Here, we will consider and effective rotary inertia (J_e), combining the motor and link inertia.

$$T_f = (1/2)\left[m_L \left(\dot{w}(L_f,t)\right)^2 + \int_0^{L_f} \rho(x) \left(\dot{w}(x,t)\right)^2 dx \right] \tag{4}$$

$$T_r = \frac{1}{2} m_r |\dot{r}|^2 + \frac{1}{2} J_e \dot{q}_r^2 = \frac{1}{2} m_r \left[0.25 L_r^2 \dot{q}_r^2 + L_r \cos(q_r)\dot{q}_r \dot{w}(L_f,t) + \left(\dot{w}(L_f,t)\right)^2\right] + \frac{1}{2} J_e \dot{q}_r^2 \tag{5}$$

where, m_L is mass of the DC servomotor (kg), $\rho(x)$ is mass per unit length of the beam (kg/m), E is modulus of the elasticity of the beam (N/m^2), and I(x) is area moment of inertia of the beam (m^4).

In the above expression, Eq.(5), use was made of the position of the center of gravity (CG) of the rigid link of the M i-M, as shown in Figure 2, and given by Eq.(6

$$r(t) = \left(L_f + 0.5 L_r \cos(q_r)\right)\hat{i} + \left(w(L_f,t) + 0.5 L_r sin(q_r)\right)\hat{j} \tag{6}$$

Since the motion of flexible beam and of M i-M joint is in the horizontal plane no gravity related terms appear in the FSMS potential energy, V_r,

$$V_r = 0 \tag{7}$$

remaining only the terms related to beam elasticity:

$$V_f = (1/2)\int_0^{L_f} EI(x)\left(w''(x,t)\right)^2 dx \tag{8}$$

Substituting Eqs. (4, 5 and 8) into Lagrange's equation Eq.(1) and performing some algebraic manipulation, one arrives at the nonlinear dynamic equations of the FSMS system. The resulting equations can be written in the matrix form below,

$$\begin{bmatrix} M_{rr} & M_{rf1} & M_{rf2} \\ M_{f1r} & M_{f1f1} & M_{f1f1} \\ M_{f2r} & M_{f2f1} & M_{f2f1} \end{bmatrix} \begin{Bmatrix} \ddot{q}_r \\ \ddot{q}_{f1} \\ \ddot{q}_{f2} \end{Bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ N_{f1r} & 0 & 0 \\ N_{f2r} & 0 & 0 \end{bmatrix} \begin{Bmatrix} \dot{q}_r \\ \dot{q}_{f1} \\ \dot{q}_{f2} \end{Bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & K_{f1f1} & K_{f1f2} \\ 0 & K_{f2f1} & K_{f2f2} \end{bmatrix} \begin{Bmatrix} q_r \\ q_{f1} \\ q_{f2} \end{Bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u \tag{9}$$

where the elements of this matrix are given by, i, j = 1, 2,

$$[M_{rr}] = 0.25 m_r L_r^2 + 0.5 J_e \tag{10}$$

$$[M_{rf}]_i = [M_{fr}]_i^T = 0.5 m_r L_r \mathbf{f}_i(L_f)\cos(q_r) \tag{11}$$

$$[M_{ff}]_{ij} = (m_r + m_r)\mathbf{f}_i(L_f)\mathbf{f}_j(L_f) + \int_0^{L_f} \mathbf{f}_i(x)\mathbf{f}_j(x)dx \tag{12}$$

$$[N_{fr}]_i = -0.5 m_r L_r \mathbf{f}_i(L_f) sin(q_r)\dot{q}_r \tag{13}$$

$$[K_{ff}]_{ij} = \int_0^{L_f} EI\mathbf{f}_i''(x)\mathbf{f}_j''(x)dx \tag{14}$$

$$\{q\}_i = \{q_r \quad q_{fi}\}^T \tag{15}$$

3. The Experimental Assembly

The experimental setup is controlled by a computer platform with AD/DA converters, driven by a real time control program written in C language. The geometrical and physical parameters of the assembled FSMS system are presented in the table 1, below.

Modulus of elasticity of the beam (N/m^2)	E	7.1e10
Area moment of inertia of the beam (m^4)	I	2.97e-10
Mass of the rigid link (kg)	m_r	0.45
Mass of the DC servomotor (kg)	m_L	0.68
Length of the rigid link (m)	L_r	0.30
Length of the flexible link (m)	L_f	0.81
Mass per unit length of the beam (kg/m)	ρ	0.32
Thickness of the beam (m)	B	5.2e-3
Height of the beam (m)	H	2.54e-2

Table 1. Physical parameters of the flexible link of the FSMS system.

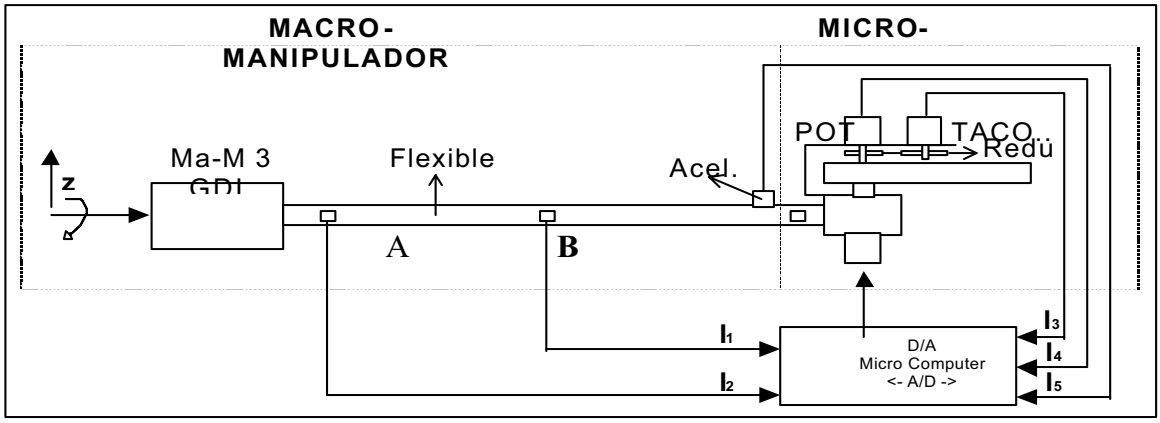


Figure 3. Experimental scheme of the control of the FSMS

Figure (3) shows the experimental assembly used in the frequency response function identification, and modal structure determination of the stationary base FSMS. The experimental assembly is composed of the following elements:

- Flexible link: beam clamped by the base in the Ma-M;
- Mi-M with 1 dof;
- Strain-gage sensor disposed in half bridge set-up;

- Piezoelectric accelerometer conditioned by charge amplifier, with vibration pre-amplifier gain V_e = 9.0 mV/g and the accelerometer sensibility S_s, = 10.10 pC/g;
- One turn potentiometer with gain k_T = 6.02 V/rd;
- DC tacogenerator with gain k_t = 0.62 V/(rd/s);
- DC servomotor, with permanent magnet, with velocity gain k_v = 13.0 V/(rd/s);
- HP dynamic analyzer, model 35660-A.

4. Nonlinear Control of the FSMS

The model of the FSMS is used to derive the control strategy and simulate the active damp of the mechanical vibrations, induced in the flexible beam (Lew and Trudnowski, 1996). This strategy consists in the cancellation of the vibracional modes using the inertial coupling matrix between the flexible link and the M-M.

According to Nenchev et al. (1999), the positional control of the M-M can be defined in terms of acceleration or velocity control variables. One possible control strategy attempts to determine a trajectory of the M-M that results in a constant coupling moment, with this the M-M moves in its workspace without exciting vibrations in the flexible link, since the reaction force in the flexible link will be null.

Equation (9) represents the nonlinear dynamics of the FSMS with one generalized coordinate (q_r) actively controlled by the motor torque, u(t), and two passive flexible coordinates (q_{f1}, q_{f2}). The active coordinate is controlled through the equation,

$$M_{rr}\ddot{q}_r + M_{rf1}\ddot{q}_{f1} + M_{rf2}\ddot{q}_{f2} = \tau \tag{16}$$

The passive generalized coordinates are related by the matrix equation,

$$\begin{bmatrix} M_{f1r} \\ M_{f2r} \end{bmatrix}\ddot{q}_r + \begin{bmatrix} M_{f1f1} & M_{f1f2} \\ M_{f2f1} & M_{f2f2} \end{bmatrix}\begin{Bmatrix} \ddot{q}_{f1} \\ \ddot{q}_{f2} \end{Bmatrix} + \begin{bmatrix} N_{f1r} \\ N_{f2r} \end{bmatrix}\dot{q}_r + \begin{bmatrix} K_{f1f1} & K_{f1f2} \\ K_{f2f1} & K_{f2f2} \end{bmatrix}\begin{Bmatrix} q_{f1} \\ q_{f2} \end{Bmatrix} = 0 \tag{17}$$

or in a equivalent form,

$$M_{fr}\ddot{q}_r + M_{ff}\ddot{q}_f + N_{fr}\dot{q}_r + K_{ff}q_f = 0 \tag{18}$$

Equation (18) expresses the dynamic constraint between the active and passive coordinates, and provided the flexible inertia matrix is non-singular, one obtains the acceleration of the passive (flexible) coordinates, as

$$\ddot{q}_f = -M_{ff}^{-1}M_{fr}\ddot{q}_r - M_{ff}^{-1}N_{fr}\dot{q}_r - M_{ff}^{-1}K_{ff}q_f \tag{19}$$

Substituting the above expression into Eq.(18), a reduced order equation is obtained for the active joint acceleration,

$$\tilde{M}_{rr}\ddot{q}_r + \tilde{C}_r\dot{q}_r + \tilde{K}_{ff}q_f = \tau \tag{20}$$

where, the tilde components are obtained by projecting the passive components on the active coordinate space by means of a common dynamic coupling expression, given by, $M_{rf} \cdot M_{ff}^{-1}$ (Nenchev, 2000),

$$\tilde{M}_{rr} = M_{rr} - M_{rf}M_{ff}^{-1}M_{fr} \tag{21}$$

$$\tilde{C}_r = -M_{rf}M_{ff}^{-1}N_{fr} \tag{22}$$

$$\tilde{K}_{ff} = -M_{rf}M_{ff}^{-1}K_{ff} \tag{23}$$

The FSMS control can then be divided in two distinct phases as proposed by (Theodore and Ghosal, 1995). In the first control phase a joint space inverse dynamics control is implemented, as illustrated in Figure (4), bellow,

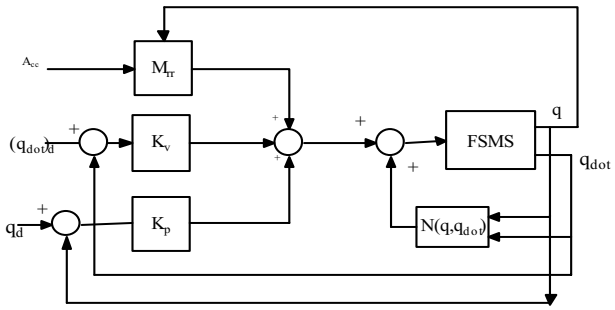


Figure 4. Schematic diagram for active joint control of the FSMS using inverse dynamics control.

The function of the nonlinear inner loop indicated in this figure, $n(q,\dot{q})$, is to provide a linear and decoupled dynamics for the active joint acceleration, whereas the outer loop is to provide zero-error tracking control and stabilization of the joint coordinate.

The nonlinear inverse dynamics control law for the FSMS is then given by,

$$t = \tilde{M}_{rr}\ddot{q}_r^d + K_v\dot{e} + K_p e + \tilde{C}_r\dot{q}_r + \tilde{K}_{ff}q_f \tag{24}$$

which upon substitution on Eq.(20) gives the following asymptotically stable equation for the positional error of the active joint, $e = q_r^d - q_r$, provided that M_{rr} is non-singular,

$$\ddot{e} + M_{rr}^{-1}K_v\dot{e} + M_{rr}^{-1}K_p e = 0 \tag{25}$$

The second control phase, the active joint is commanded to actively damp the vibrations induced in the elastic structure of the FSMS. The flexible equation of motion, Eq.(is excited by the force and torque reactions, Γ, induced by the motion of the M+M. The integral of the reaction force is denoted the coupling moment of the FSMS, Π, (Nenchev, Yoshida, and Uchiyama, 1999)

$$\Gamma = \begin{bmatrix} f^T & t^T \end{bmatrix} = M_{fr}\ddot{q}_r + \dot{M}_{rf}\dot{q}_r = \frac{d}{dt}\Pi \tag{26}$$

$$\Pi = M_{fr}\dot{q}_r \tag{27}$$

The objective of the proposed vibration control strategy is to find a joint acceleration control that upon substitution on the flexible equation of motion, Eq. (19), gives a stable closed loop equation for the passive coordinates. The closed loop control law with velocity and position feedback, automatically cancel the coupling terms of the reaction matrix, and introduces an active damping term in the vibration of the flexible element, thus altering the oscillation frequency and damping of the closed loop system. The active vibration control is given by,

$$\ddot{q}_r = M_{fr}^{+}(G_f^v\dot{q}_f + G_f^p q_f - \dot{M}_{fr}\dot{q}_r) \tag{28}$$

where, M_{fr}^{+} is the Moore-Penrose pseudoinverse of the coupling matrix M_{rf}. Substituting Eq. (28) into (18), gives the closed loop behavior of the passive coordinate dynamics,

$$M_{ff}\ddot{q}_f + G_f^v\dot{q}_f + (K_{ff} + G_f^p)q_f = 0 \tag{29}$$

Using the modified dynamics of the active joint coordinate, Eq.(20), one can find the torque required to give the acceleration expressed by Eq.(28), which is required to damp the vibration of the flexible link. By this procedure one arrives at an expression of a unified torque control, composed of two phases, one for positioning the active joint and the other to damp the vibrations of the elastic supporting structure.

5. Control Simulation of the FSMS

The proposed control law for the FSMS was tested by simulation in the Matlab/Simulink environment. The Simulink block diagram is composed of three main dynamic blocks representing the active coordinate dynamics, given

by Eq.(16), and two passive dynamic blocks, given by Eq.(17). The outputs of these blocks are the active and passive accelerations, which are recursively integrated to find velocities and position coordinates.

The dynamics of the passive coordinates are calculated using the actual active joint acceleration. In a similar manner the passive dynamics depends on the actual values of the active coordinate through the respective coupling matrices. The control torque has two parts, one is represented by the inverse dynamics control, as shown schematically in Figure (4). This torque component is calculated from the position and velocity errors of the active joint. The position and velocity errors are feedback with gains, K_p and K_v, respectively.

The second part of the torque control corresponds to the vibracional damping of the flexible structure. This excitation signal gives the required acceleration, Eq. (28), necessary to introduce the active damping in the FSMS. The damping is controlled by the velocity gain, G_f^v, while the frequency of oscillations is controlled by the position gain, G_f^p. These gains are chosen to give the required attenuation of the FSMS oscillations. Without any original structural damping, the indicated gains ($G_f^p = 0$, and $G_f^v = 800$) can attenuate the vibrations in a time scale of a few seconds.

The reference trajectory is a function of the specified active joint acceleration, velocity and position. This trajectory corresponds to a slew motion of approximately 60 degrees in one second. This trajectory takes in account the actual torque limitation of the experiment, which is on the order of 0.1 Nm.

The two-phase control strategy is governed by the switching control signal derived from the main simulation time base. In the first 1.5 second, the inverse dynamics control is operating. This time is necessary to reach the stable endpoint commanded in the active joint control. This control signal excites the vibracional degrees of freedom of the FSMS, which are then damped by switching the control to give a torque corresponding to a previously calculated joint acceleration, as given by Eq.(28). The Matlab/Simulink simulation scheme is as shown in Figure (5), bellow.

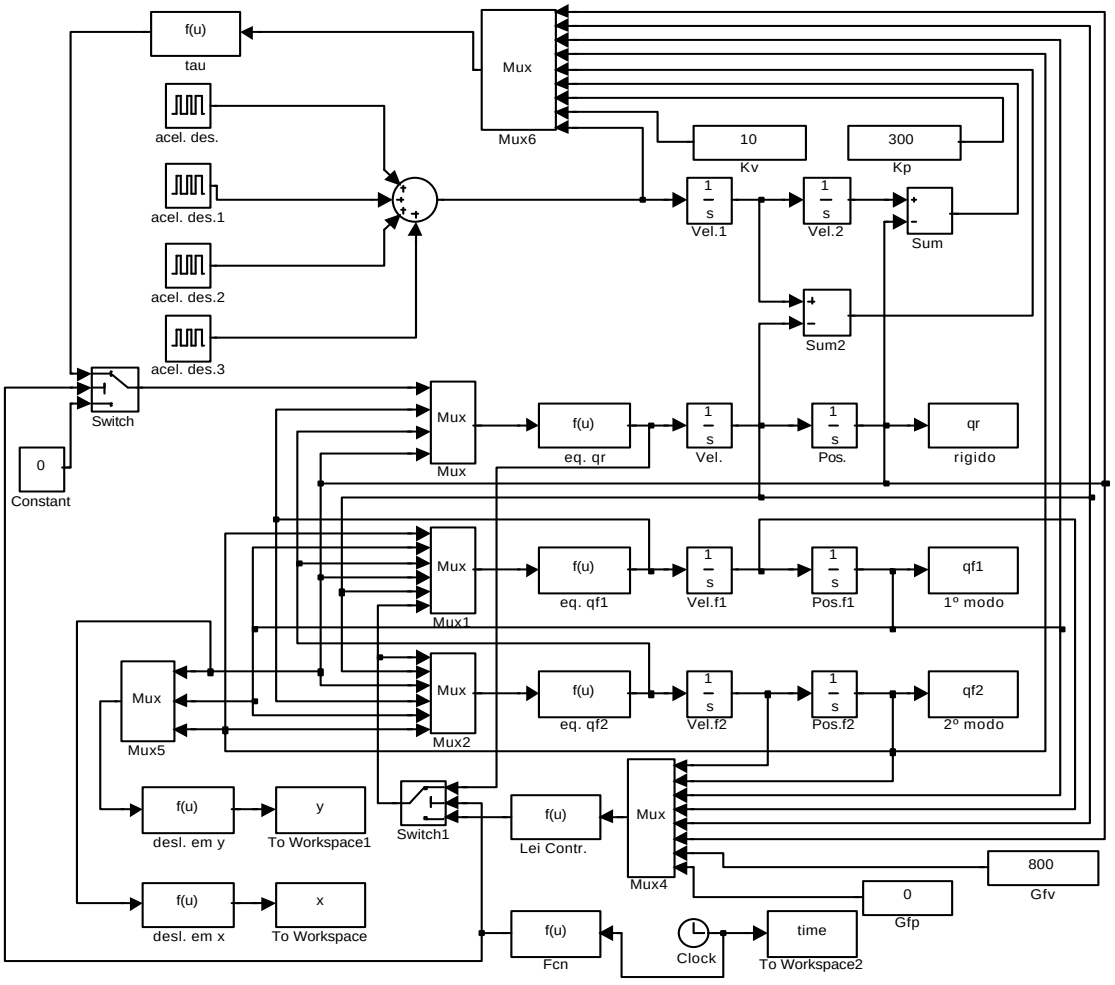


Figure 5. Simulink Block Diagram of the FSMS Control Law.

Figure (6) shows the time variation of the active joint coordinate, $q_r(t)$. The simulated movement corresponds to a slew motion of the rigid link by an angle of 1 rd (approximately 60 angular degrees) in one second. This figure shows that the prescribed final angle is achieved with zero error. After the position control, the joint angle is moved in such way to damp the vibrations of the flexible link

The control torque is shown in Figure (7). In this figure one can see the switching time between the two-mode controller at the time 1.5 seconds. The switching is characterized by an

The damped motion of the flexible link is shown in Figure (8). This graph shows the time variation of the flexible modal coordinate, $q_{f1}(t)$, associated with the first modal coordinate in the assumed mode expansion . The damped oscillations decay with an exponential rate taking approximately 6 seconds to damp out the first mode.

The time evolution of the second modal coordinate, $q_{f2}(t)$, is shown in Figure (9). Compared with the firs mode, the cancellation effect is not so effective, since a residual oscillation persists even after 10 seconds. This is in part because only the first mode was optimized through the choice of the velocity feedback gain, G_f^v.

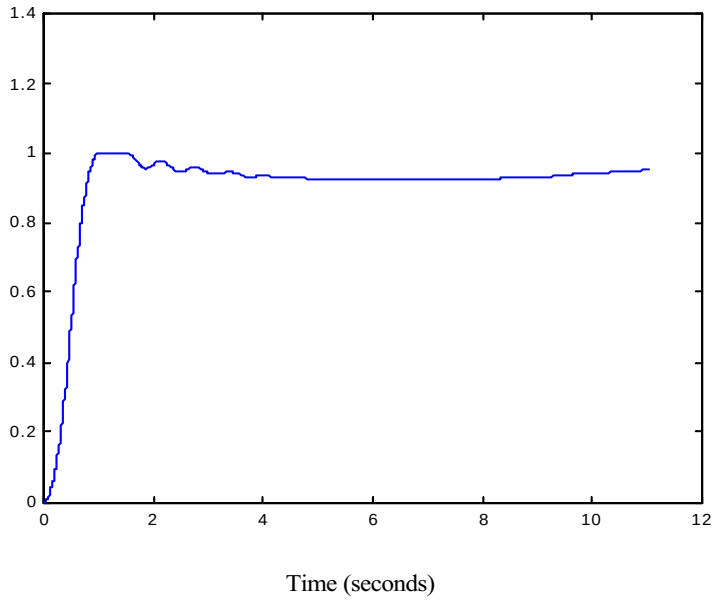


Figure 6. Time variation of the active joint coordinate measured in radias, qr(t).

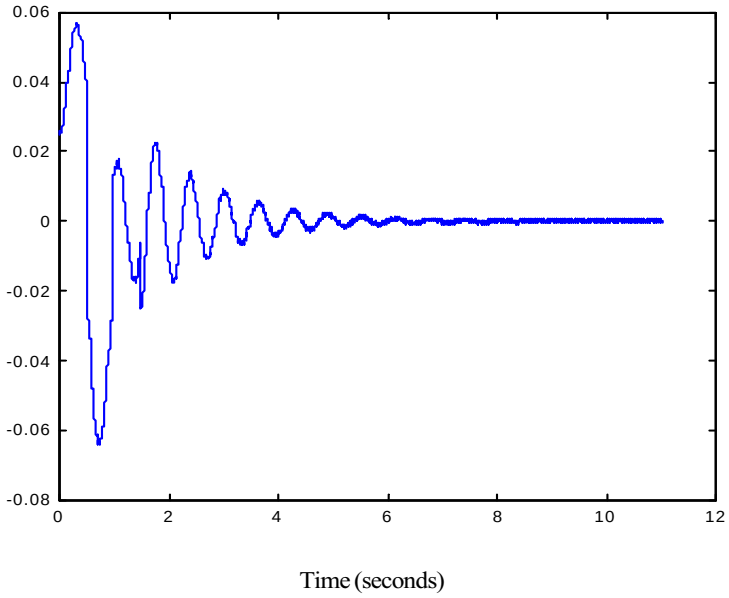


Figure 7. Control Torque for the FSMS System

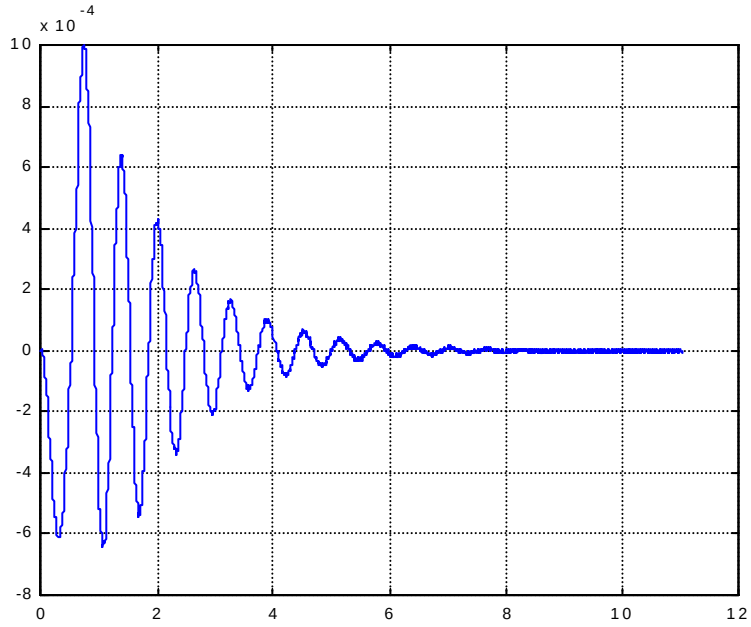


Figure 8. Simulation of the time evolution of the modal coordinate, $q_{f1}(t)$.

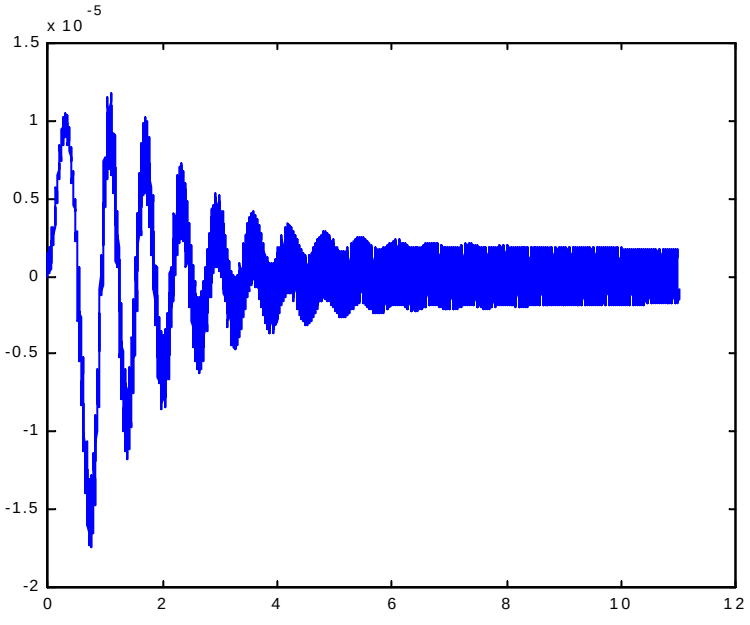


Figure 9. Simulation of the time evolution of the second modal coordinate, $q_{f2}(t)$.

7. Conclusions

In this work, a discrete model based on the assumed modes technique was derived for an existing FSMS system. The Lagrange's equations were used to derive a set of closed equations of motion of the FSMS, with up to two flexible modes. This model was used to simulate the dynamics of the FSMS under a nonlinear model-based control law. The control law was used successfully, since it resulted in a stable closed loop system, where the elastic oscillations are effectively damped by the active control strategy, as shown in Fig. (8) and (9). The simulation of this control strategy was also successful, since the control torque is within the capacity of the real actuator of the FSMS (bellow 0.1 Nm). A small departure from the equilibrium position is observed for the active joint angle. This is due to the drift in the control law associated with the joint acceleration law. In steady state regime, the joint acceleration is zero corresponding to a constant joint velocity, If this value is not zero a small drift can be observed in the joint angle position after the positional control is switched off. A new control strategy is been derived, based on joint velocity control, such as to avoid this small position drift in the final active vibration control phase of the FSMS

8. References

Dubowsky, S., Durfee, W., Corrigan, T., Kulinski, A., and Muller, U., 1994, "A Laboratory Test Bed for Space Robotics: The VES-II", IEEE/RSJ Int. Conf. on Intelligent Robot System, Munich, Germany.

Góes, L.C.S., Grandinetti, F.J., Soares, A.M., and Nechev, D., 2000, "Controller of Micro/Macro Hybrid Manipulator System", Congresso Nacional de Engenharia Mecânica, CONEM 2000, Natal, RN, Brazil.

Góes, L.C., Grandinetti, F.J., and Soares, A.M., 2000, "Dynamic Modeling and Experimental Identification of Flexible Structure Mounted Manipulator System", Proc. Robotics 2000, ASCE International Conference, Albuquerque, NM, Vol.1, USA, pp. 283-292.

Junkins, J.L., 1989, "Application of Modern State Space Analysis in Spacecraft Dynamics, Estimation and Control, Tutorial Keynote Lecture, AAS Guidance and Control Conference, Keystone, CO., Advances in the Astronautical Sciences", Vol. 68, Univelt, San Diego, CA, pp. 555-678.

Jain, A., and Rodriguez, G., 1993, "An Analysis of the Kinematic and Dynamics of Underactuated Manipulators", IEEE Trans. On Robotics and Automation, Vol. 9(4):411-422.

Lew, J.Y., and Trudnowski, D.J., 1996, "Vibration Control of a Micro/Macro-Manipulator System, IEEE Control System, Vol. 1 (2).

Mavroidis, C., Dubowski, S., and Raju, V., 1995, "End-Point Control of Long Reach Manipulator System", Proc. Ninth World Congress of IFToMM, Milano, Italy, pp.1740-1744.

Nenchev, D., 2000, "Underactuated Multibody System: Reaction Null Space Based Decomposition", International Conference on Nonlinear Dynamics, Chaos, Control and Their Applications in Engineering Sciences, Campos de Jordão, Brazil (in press).

Nenchev, D., Yoshida, K., and Uchiyama, M., 1999, "Reaction Null Space Control of Free-Floating and Elastic Base Robots", "RBCM – J. of the Brazilian Soc. Mechanical Sciences, Vol. 21(2):175-184.

Oriolo, G., and Nakamura, Y., 1991, "Free-Joint Manipulators: Motion Control under Nonholonomic Constraint", Proc. IEEE/RSJ Int. Workshop on Intelligent Robots and Systems'91, pp.1248-1253.

Nenchev, D., Yoshida, K., Vichitkulsawat, P., and Uchiyama, M., 1999, "Reaction null space control of flexible structures mounted manipulating systems, IEEE Trans. Robotics and Automation, Vol. 15(6).

Shiffman, J.J., 1993. "Lyapunov Functions and the Control of the Euler-Bernoulli Beam, Int. J. Control, 57(4):971-990.

Theodore, R.J., and Ghosal, A., 1995, "Comparison of the Assumed Modes and Finite Element Models for Flexible Multi-Link Manipulators", The International Journal of Robotics Research, Vol 14(2):91-111.

Xu, Y., and Kanade, T., 1993, "Space Robotics: Dynamics and Control", Kluwer Academic Publishers.

Yoshida, K., Nenchev, D., Inaba, N., and Oda, M., 2000, "Extended ETS-VII Experiments for Space Robot Dynamics and Attitude Disturbance Control", Proc. of the 22nd ISTS Conference, Morioka, Japan (in press).

XVI CONGRESSO BRASILEIRO DE ENGENHARIA MECÂNICA
16th BRAZILIAN CONGRESS OF MECHANICAL ENGINEERING

TIKHONOV REGULARIZATION METHOD AND TRACKING

L. Barbanti
IME-USP-P.O.Box 66281-São Paulo-SP-Brasil- 05315-970
barbanti@ime.usp.br

J. M. Balthazar
IGCE-UNESP-P.O.Box178-Rio Claro-SP-Brasil-13500-230
jmbaltha@rc.unesp.br

M. J. H. Dantas
Mat-UFU-Uberlandia-MG-Brasil-38400-000
marcio@ufu.br

Abstract. *When tracking a trajectory in a nonideal system the unstabilities in the zero dynamics lead to ill-posed problems. By means of the Tikhonov regularization method, a class of equations converging to the original one can be defined allowing in this way an approximating scheme for a stable solution in the tracking procedure.*

Keywords. *Tracking solutions, Tikhonov regularization, inverse problem, approximating scheme for ODE, nonideal systems.*

1. Introduction

The arguments in this paper comes from the need to make stable the control choices for synthesis in nonideal coupled systems mainly in the problem of tracking a trajectory by means of the input-output linearized process associated to the original one. The correspondence between the trajectory to be tracked and the corresponding tracking control is stable only in a very few cases. So, in general, small inaccuracies in the performing controls lead to big errors in the corresponding trajectories showing that the problem in general is ill-posed. Using the Tikhonov regularization method for ill-posed systems we install a scheme of equations approximating in a stable mode the input-output linearized one. Throughout this work we will be using freely the notations and concepts as in Isidori(1989) and Tikhonov and Arsenin(1977).

2. Tracking trajectories

Let be the controlled ODE with an observable variable y,

$$\begin{aligned} \frac{dx}{dt} &= A(x) + g(x)u \\ y &= h(x) \end{aligned} \tag{1}$$

where $x \in R$ and $A(\cdot) : R^n \to R^n$ and $g, h : R^n \to R$ and $u \in L_1^{loc}$.

Let us consider the input-output linearization for Eq. (1),

$$y^{(r)} = v \tag{2}$$

where r $(r \leq n)$ is the relative degree for Eq. (1) and

$$v = L_g(L_A^{r-1}h)u + L_A^r h \tag{3}$$

with $L_m p$ representing the usual Lie derivative of the mapping m in the direction p.

Consider now the problem of tracking a fixed trajectory $y_\#$ in Eq. (2). Defining

$$y^* = \left(y, \frac{dy}{dt}, \ldots, y^{(r-1)}\right)$$

and

$$y_\#^* = (y_\#, \ldots, y_\#^{(r-1)})$$

and the tracking error:

$$e = y^* - y_\#^*$$

we get- following Tikhonov and Arsenin (1977) and by choosing $k_1, k_2, \ldots, k_r$ in such a way that the polynomial (in the variable p) $k_1 + k_2 p + \cdots + k_r p^{(r-1)} + p^r$ has all roots strictly in the left half plane- that the tracking control $u_\#$ is defined by:

$$u_\# = \frac{1}{L_g(L_A^{r-1}h)}(-L_f^r h + y_\#^{(r)} - k_r e^{(r)} - \cdots - k_1 e). \tag{4}$$

Following Isidori (1989) if the zero-dynamics for Eq. (1) is stable we have the problem of tracking a trajectory solved according Eq. (2) by using the control done by the Eq. (4). However, if stability is not allowed this is not true even in very simple two-dimensional processes - with $x = (z, w)$ - like the following:

$$\begin{aligned} \frac{dz}{dt} &= w + u \\ \frac{dw}{dt} &= -u \\ y &= z. \end{aligned}$$

In general if the zero-dynamics in the system in Eq. (1) is unstable we use the Tikhonov regularization method for get an approximating scheme of equations to the Eq. (1) in such a way the synthesis of $u_\#$ it turns in a stable procedure. In fact: let us consider the performance (or realization) operator

$$\Lambda(y_\#) = u_\#.$$

It is easily seen that Λ is a linear affine operator but in general with no continuous inverse showing in this way that when in real situations we could not expect an effective reliability in the system. In other words: small perturbations in $u_\#$ do not lead to the proximity of $\Lambda^{-1}(u_\#) = y_\#$. By other hand, sometimes it is necessary to consider during an experiment, an executable small perturbations of $u_\#$, namely $u_\#^\delta$, with $\|u_\# - u_\#^\delta\| \leq \delta$ and/or the fact that $u_\#^\delta \notin R(\Lambda)$ (with $R(\Lambda)$ being the range of Λ). In both the cases, by using then the Tikhonov method it is possible to synthesize a regularization scheme

$$\Lambda^* = (\alpha I + \Lambda^* \Lambda)\Lambda_\alpha$$

(where Λ^* denotes the adjoint operator of Λ, $\alpha = \alpha(\delta) > 0$, and I is the identity operator) in such a way

$$\Lambda_\alpha(u_\#^\delta) = u_\#^\delta \to \Lambda^{-1}(u^\delta)$$

in a stable manner.

In this way we had defined a scheme of equations $(1)_\alpha$ converging in a convenient topology to the Eq. (1). Namely $(1)_\alpha$ is the equation

$$\begin{aligned} \frac{dx}{dt} &= A_\alpha(x) + g(x)u \\ y &= h(x) \end{aligned}$$

having Λ_α as its performance operator.

Further, as shown in Barbanti and Balthazar (2001), there is an $\alpha_0 > 0$ such that for all $\alpha > \alpha_0$ we have $(1)_\alpha$ being a stable process. Notice that elsewhere we are making use of the above $(1)_\alpha \to (1)$ scheme to get chaos controllability in Eq. (1).

3. References

Barbanti, L., Balthazar, J. M., 2001, "Nonideal coupled systems and the Tikhonov approximating sequence for inverse solutions", Preprint.

Isidori, A., 1989, "Non linear control systems:an introduction", Springer Verlag.

Tikonov, A. N., Arsenin, V. Y.,1977, "Solution of ill-posed problems", Winston and Sons.