

REVISTA BRASILEIRA DE CIÊNCIAS MECÂNICAS

PUBLICAÇÃO DA ABCM
ASSOCIAÇÃO BRASILEIRA DE CIÊNCIAS MECÂNICAS



MEMBROS DA DIRETORIA DA ABCM

Luiz Bevilacqua (Presidente)

Hans Ingo Weber (Vice-Presidente)

Augusto Cesar Galeão (Secretário Geral)

Mauro Sergio Pinto de Sampaio (Secretário)

Maurício Nogueira Frota (Diretor de Patrimônio)

Vol. VIII nº 4 - Nov. 1986

**EDITOR
RESPONSÁVEL**
Rubens Sampaio

**EDITOR
EXECUTIVO**
J. M. Freire

**CONSELHO
EDITORIAL**

Abimael F. D. Loula
Arthur J. V. Porto
Berend Snoeijer
Bernardo Horowitz
C. S. Barcellos
D. E. Zampieri
Duraíd Mahrus
E.O. Taroco Aliano
F. Venâncio Filho
F. E. Mourão Saboya
Giulio Massarani
Guillermo Creuss
Hans Ingo Weber
Henner A. Gomide
Jan Leon Scieszko
Jerzy T. Sielawa
J. J. Espindola
Liu Hsu
Maurício N. Frota
Miguel H. Hirata
Nelson Back
Néstor Zouain
Nivaldo L. Cupini
O. Maizza Neto
Pedro Carajilesco
Sergio Colle

Editorial

Dynamic response of a rigid disc mounted on a flexible shaft under non-linear excitations 261

V. Steffen Jr. - Membro da ABCM e SEM

F.P. Lépre - Membro da ABCM

E.B. Teodoro

Universidade Federal de Uberlândia

Departamento de Engenharia Mecânica

Uberlândia - MG

Randon vibrations and noise generation in railway wheels 273

K. Popp

Institute of Mechanics

University of Hannover

FRG

State estimation of multivariable systems using the duality concept 285

Agenor de Toledo Fleury

Atair Rios Neto

Instituto de Pesquisas Espaciais - INPE/MCT

São José dos Campos/SP

Uma formulação consistente do método dos elementos finitos para resolução de problemas de transporte difusivo-convectivo 309

Eduardo Gomes do Carmo - Membro da ABCM

COPPE/UFRJ

Augusto Cesar Galeão - Membro da ABCM

INCC/CNPq



**A REVISTA BRASILEIRA DE CIÊNCIAS MECÂNICAS
É PUBLICADA COM O APOIO**

**DO CNPq E FINEP
COMPANHIA VALE DO RIO DOCE
IBM DO BRASIL**

Conselho da ABCM eleito para o Biênio 86/87

Eng^o Carlos Alberto Couto, FINEP
Dr. Edgardo Taroco, LNCC/CNPq
Dr. Hazim Ali Al-Qureshi, ITA
Dr. Henner Alberto Gomide, UFU
Suplente - Dr. Jaime Tupiassú Pinto de Castro, PUC/RJ
Eng^o José Augusto R. do Amaral, NUCLEN
Suplente - Eng^o José Carlos Balthazar, UnB
Prof. José de Mendonça Freire, PUC/RJ
Suplente - Dr. Kazuo Hatakeyama, CEPED/BA
Dr. Raúl A. Feijóo, LNCC/CNPq
Dr. Rubens Sampaio Filho, PUC/RJ
Dr. Tito Luiz da Silveira, Fund. Souza Marques
Dr. Valder Steffen Junior, UFU

AGRADECIMENTO

A Direção da Revista Brasileira de Ciências Mecânicas agradece a contribuição de todos os revisores que se empenharam para a boa apresentação dos trabalhos publicados durante este ano de 1986:

Alcir de Faro Orlando, PUC/RJ
Angela Ourivio, Nieckele, PUC/RJ
Antonio Santos Vargas, PUC/RJ
Arno Blass, UFSC
Carlos Alberto de Almeida, PUC/RJ
Clovis Raimundo Maliska, UFSC
Giulio Massarani, COPPE/UFRJ
Hans Ingo Heber, UNICAMP
Jaime Tupiassú Pinto de Castro, PUC/RJ
Jan Leon Scieszko, COPPE/UFRJ
João Sampaio D'Ávila, UFSE
Jorge Guilherme S. Patiño, PUC/RJ
Jorge Teixeira Freire, UFSC
José Alberto dos Reis Parise, PUC/RJ
José Luiz de França Freire, PUC/RJ
Krishnaswama Rajagopal, COPPE/UFRJ
Luís Fernando Alzuguir Azevedo, PUC/RJ
Maurício Nogueira Frota, PUC/RJ
Miguel Hiroo Hirata, COPPE/UFRJ
Paulo Jorge Serpa Paes Leme, PUC/RJ
Paulo Roberto de Souza Mendes, PUC/RJ
Pedro Carajilescov, ITA
Sidney Stuckenbruck, PUC/RJ
Washington Braga Filho, PUC/RJ

DYNAMIC RESPONSE OF A RIGID DISC MOUNTED ON A FLEXIBLE SHAFT UNDER NON-LINEAR EXCITATIONS

V. Steffen Jr. - Membro da ABCM e SEM

F. P. Léopore - Membro da ABCM

E. B. Teodoro

Universidade Federal de Uberlândia

Departamento de Engenharia Mecânica

Uberlândia - MG

ABSTRACT

This paper presents a mathematical model for a rigid disc mounted on a flexible shaft supported by ball bearings. The disc is excited by a contactless magnetic force. The equations of motion are integrated using a Runge-Kutta technique and the frequency response is obtained by a FFT technique. The response of the system is discussed for different situations.

INTRODUCTION

The authors have been studying the possibilities of amplitude of vibrations reductions by introducing an extra magnetic force when the rotor system is magnetically excited [1],[2].

In [3], it is presented a one-degree-of-freedom system moving transversely and axially on a magnetic field. For the first case it can be observed an augmentation of the natural frequency value when increasing the magnetic flux. In the second case, a reduction is observed for natural frequency value when increasing the magnetic flux. These results show that the magnetic attraction can be used to modify the system stiffness.

Different works have been conducted to characterize a magnetical bearing [4],[5] and using a control feedback system to keep the amplitudes of vibrations at acceptable levels [6],[7].

In all above applications magnetic non-linear forces excite the rotor systems. This paper analyses the system response when magnetic forces are applied to a flexible rotor.

MATHEMATICAL MODEL

The physical model for the rotor-system is a vertical simple supported flexible shaft carrying a symmetrical rigid disc. Figure 1 shows the system geometric configuration and the fixed and rotating coordinates axes.

The equations of motion can be obtained using the Lagrange Energy Method and they can be written in the matrix form:

$$\begin{bmatrix} m & 0 & 0 & 0 \\ 0 & J & 0 & 0 \\ 0 & 0 & m & 0 \\ 0 & 0 & 0 & J \end{bmatrix} \begin{Bmatrix} \ddot{x} \\ \ddot{\theta} \\ \ddot{y} \\ \ddot{\phi} \end{Bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I\Omega \\ 0 & 0 & 0 & 0 \\ 0 & -I\Omega & 0 & 0 \end{bmatrix} \begin{Bmatrix} \dot{x} \\ \dot{\theta} \\ \dot{y} \\ \dot{\phi} \end{Bmatrix} + \begin{bmatrix} k_d & k_{ad} & 0 & 0 \\ k_{ad} & k_a & 0 & 0 \\ 0 & 0 & k_d & k_{ad} \\ 0 & 0 & k_{ad} & k_a \end{bmatrix} \begin{Bmatrix} x \\ \theta \\ y \\ \phi \end{Bmatrix} = \begin{Bmatrix} F \\ 0 \\ 0 \\ 0 \end{Bmatrix} \quad (1)$$

where:

x, θ, y, ϕ are the generalized coordinates;

m, J, I are the inertia coefficients;

Ω is the rotation of the shaft

k_a, k_d, k_{ad} are the shaft stiffness coefficients;

F is the excitation force.

The non-linear, magnetic contactless force is represented by $F(x) = C/d^2$, where d is the instantaneous gap and C is a magnetic field constant ($N \cdot m^2/A^2$). For an electromagnet $C = (NI)^2 \mu_s / 2$, where NI is the Ampère-turns, μ is the magnetic permeability and S is the pole-area. For the case where a permanent magnet is describing a sinusoidal motion in the radial direction beginning from a starting gap d_0 , F can be expressed by:

$$F = F_x = C / (d_0 - x_0 \sin \omega_e t - x) \quad (2)$$

where x_0 is the value of the amplitude of the permanent magnet motion and ω_e is the exciting frequency.

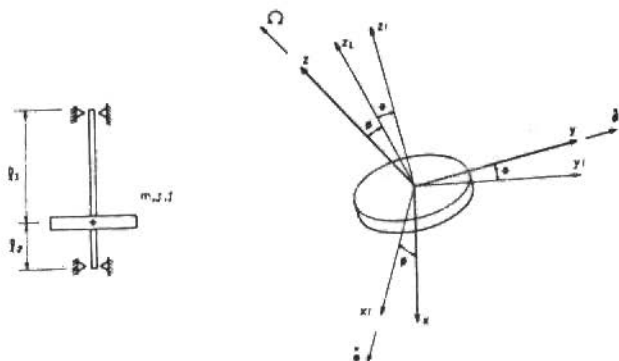


Figure 1. Flexible rotor

HOMOGENEOUS CASE

The equation (1) is integrated (exact solution) when $F = 0$ and the natural frequencies and critical speeds are determined and shown by Figure 2.

For these calculations the values of the parameters are taken from a real laboratory model, which is a vertical flexible steel shaft ($\phi 9\text{mm}$, 0.461m length). The shaft supports a rigid disc ($\phi 150\text{mm}$, $m = 1.833\text{kg}$, $I = 0.0052\text{kg}\cdot\text{m}^2$), which is located at $l_2 = 0.137\text{m}$.

The stiffness parameters k_d and k_a are dynamically identified as $k_d = 50129\text{N/m}$ and $k_a = 2427\text{N}\cdot\text{m}/\text{rd}$. The coupling stiffness parameter k_{ad} is calculated: $k_{ad} = 8878\text{N}\cdot\text{m}/\text{m}\cdot\text{rd}$. Tests made with the laboratory model show that the damping factors associated to all considered vibrating modes are very small ($\xi \approx 5.0 \times 10^{-4}$).

Figure 2 shows the influence on the values of the natural frequencies when considering $k_{ad} = 0$.

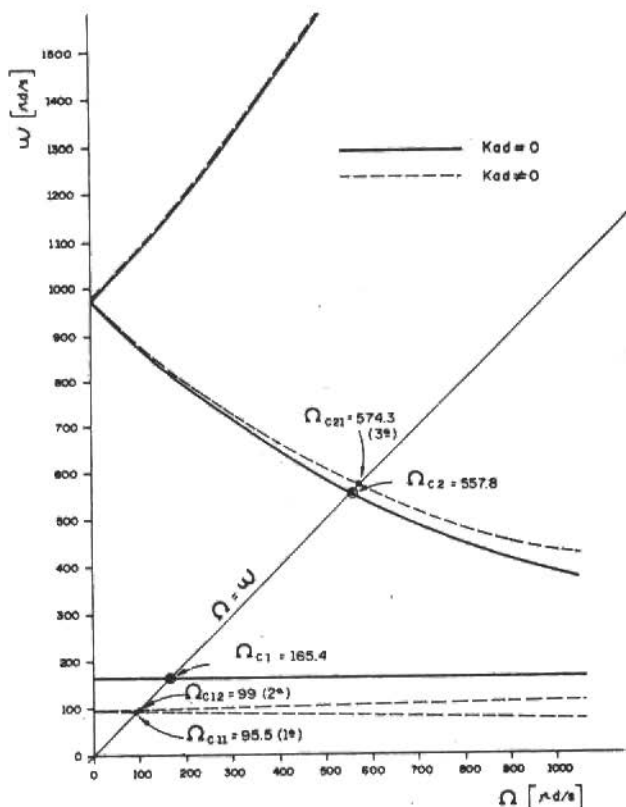


Figure 2. Natural frequencies and critical speeds

NON-HOMOGENEOUS CASE

The time-response for the non-homogeneous system is obtained by the integration of equation (1) using a Matrix Runge-Kutta technique. The value for the magnetic constant C is taken from experimental verification. In the laboratory, the excitation is provided by a permanent magnet driven by an electrodynamic shaker which is placed close to the disc at an initial gap $d_0 = 3.0\text{mm}$, in such way that the amplitude of the imposed sinusoidal motion is 1.0mm . This way, for the region of interest ($3.0 \pm 1.0\text{mm}$) the mean value of C is $1.048 \times 10^{-5} \text{ N}\cdot\text{m}/\text{A}^2$.

The frequency response is obtained from the time-response using a FFT software. This way it is possible to determine the natural frequencies for different speeds of rotation Ω of the rotor when it is excited by non-linear magnetic forces.

Figure 3 shows $x(t)$ and $\theta(t)$ for $\Omega = \omega_e = 400$ rd/s.

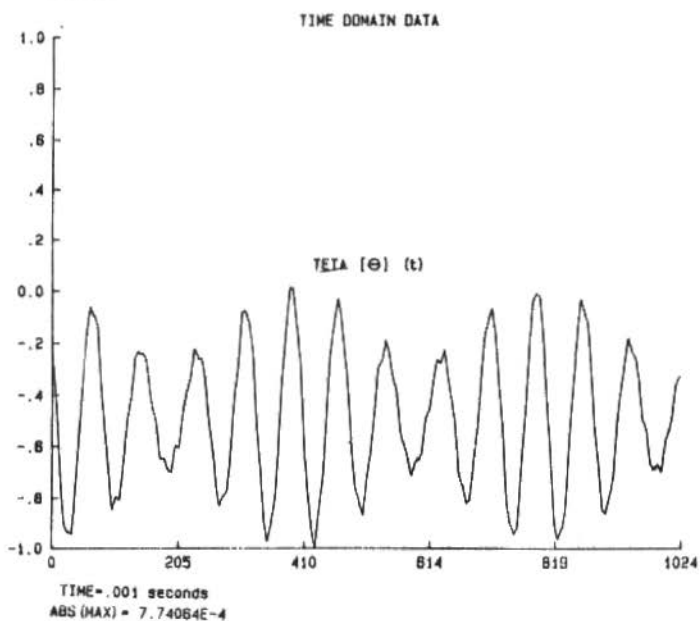
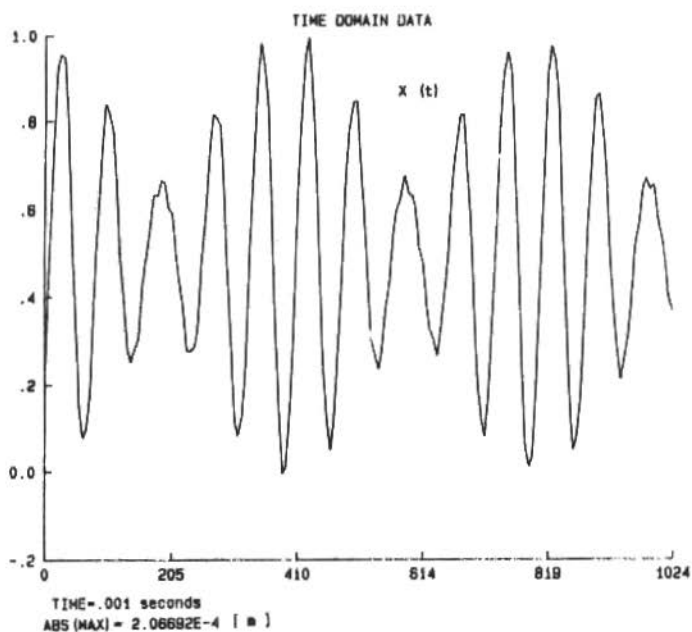


Figure 3. $x(t)$ and $\theta(t)$, $\Omega = \omega = 400$ rd/s, $C = 1.048 \times 10^{-5}$ Nm²/A²

Figure 4 shows the magnitude and phase spectra calculated for $x(t)$ and Figure 5 shows the magnitude and phase spectra for $\theta(t)$.

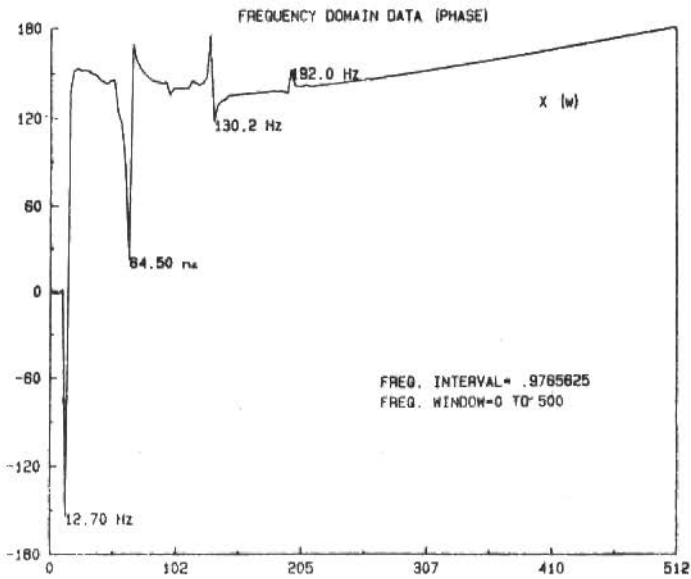
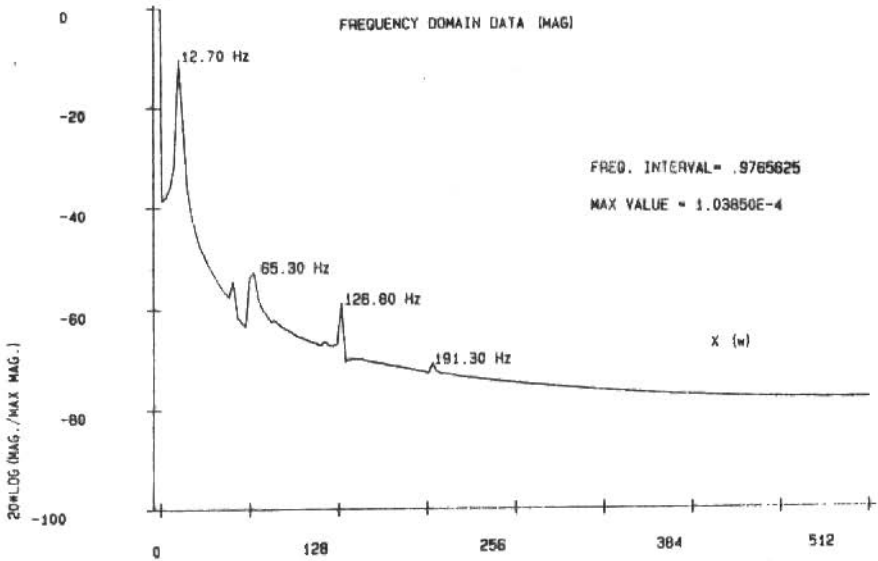
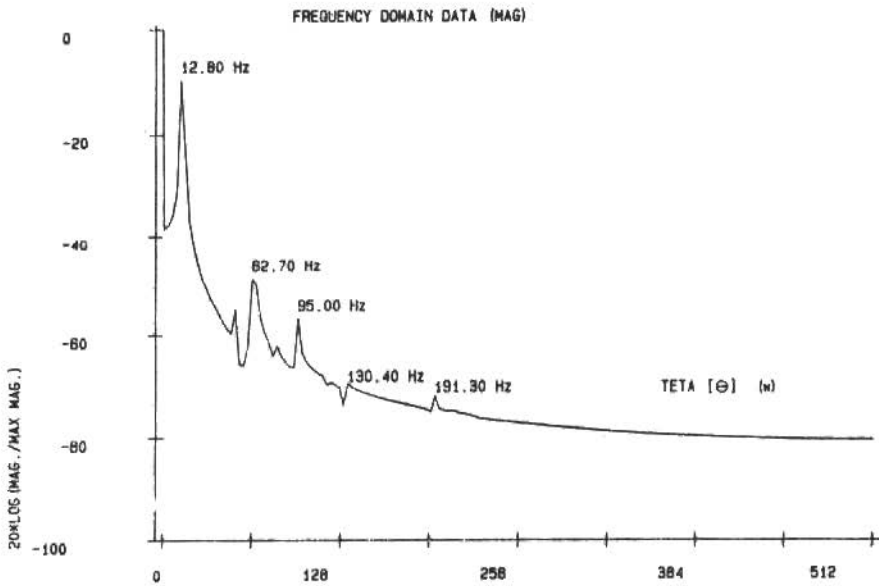


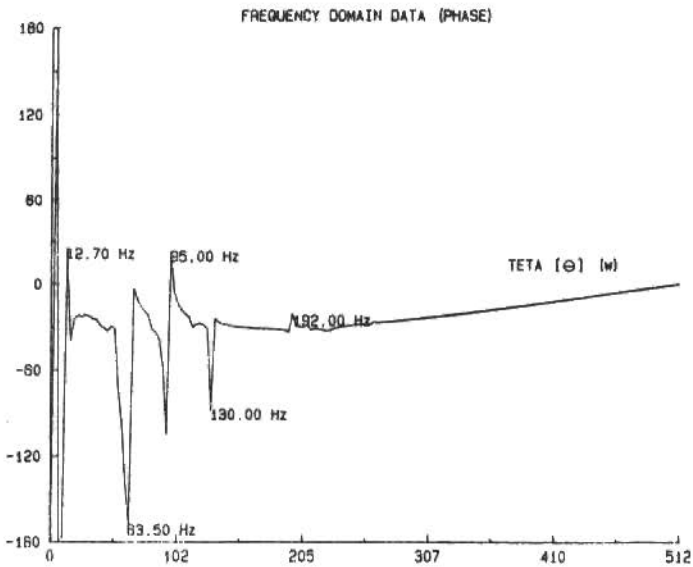
Figure 4. Magnitude and phase spectra calculated for $x(t)$

$$\Omega = \omega_e = 400 \text{ rd/s}; C = 1.048 \times 10^{-5} \text{ Nm}^2/\text{A}^2$$



FREQ. INTERVAL = .9765625

MAX VALUE = 3.80085E-4



FREQ. INTERVAL = .9765625

FREQ. WINDOW = 0 TO 500

Figure 5. Magnitude and phase spectra calculated for $\theta(t)$

$$\Omega = \omega_e = 400 \text{ rd/s}; C = 1.048 \times 10^{-5} \text{ Nm}^2/\text{A}^2$$

The lower frequencies are presented on Table 1, for different situations.

Table 1. Lower frequencies

DIRECT METHOD (Figure 2)		RUNGE KUTTA INTEGRATION AND FFT ANALYSIS		
$\Omega = 0$	$\Omega = 400$ rd/s	$\Omega = 400$ rd/s $\omega\Omega = 0$ Initial Conditions $X(0) = 1\text{mm}$ $\dot{X}(0) = 5.24 \times 10^{-2}$ rd $Y(0) = \dot{Y}(0) = 0$	$\Omega = 400$ rd/s $\omega\Omega = 0$ $C = 1.048 \times 10^{-3}$ Nn ² /A ² Initial Conditions $X(0) = 1\text{mm}$ $\dot{X}(0) = 5.24 \times 10^{-2}$ rd $Y(0) = \dot{Y}(0) = 0$	$\Omega = \omega\Omega = 400$ rd/s $C = 1.048 \times 10^{-3}$ Nn ² /A ²
—	14.40	—	—	—
15.28	15.90	15.50	12.70	12.70
—	105.0	94.70	96.0	62.4 (1x0) 94.8
155.2	—	—	—	126.8 (2x0)
—	234.80	219.0	223.8	191.3 (3x0)
[Hz]	[Hz]	[Hz]	[Hz]	[Hz]

It can be noticed that the values of the frequencies have a tendency to change in the presence of the magnetic excitation, suggesting that the stiffness of the system has been modified. This is an interesting behavior of the system considering that this effect could be used to cross critical speeds safely.

Figure 6 shows the locus of the disc. It can be observed that the vibratory motion occurs in region displaced from the origin. This displacement is due to the magnetic exciting force applied at the x - direction.

Another magnetic force represented by equation (3) can be introduced for reducing the vibrations level of the disc [1]:

$$F_C = C / [d_{oc} - x_{oc} \text{sen}(\omega_c t + \Delta) + x]^2 \quad (3)$$

where Δ is a phase angle which is taken into account when $\omega_e = \omega_c$.

Figure 7 shows the locus of the disc when F_C is applied at $t = 0.40$ sec, when $\Omega = \omega_e = \omega_c = 400$ rd/s and $\Delta = \pi$.

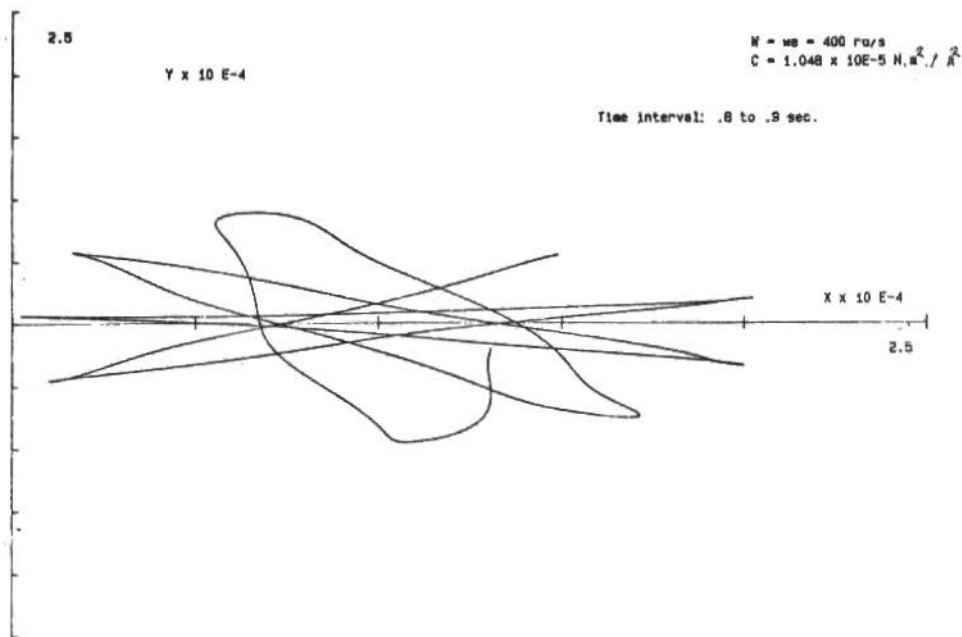


Figure 6. Locus of the disc

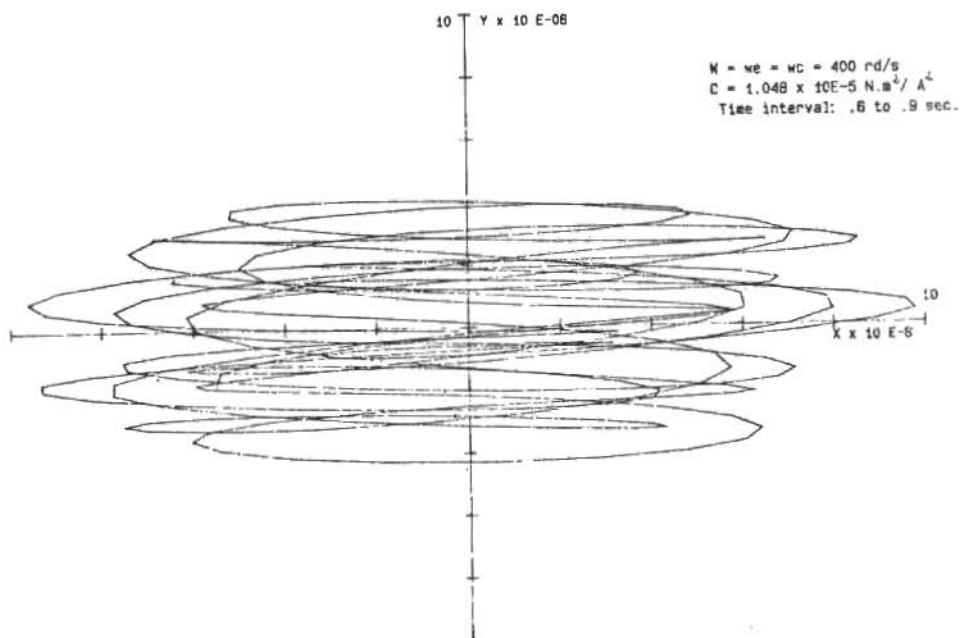


Figure 7. Locus of the disc-amplitude controlled

It can be seen that when F_c is applied, the motion of the disc occurs again in a region close to the origin and this motion is significantly more "organised". To show the possibilities of vibrations reductions when F_c is introduced, it is now supposed that the shaft operating speed of rotation is 400 rd/s (between two criticals), and $\omega = \Omega$. The frequency $\omega_c = 1.2 \Omega$, and F_c are applied when $t = 0.40$ s. Figure 8 and 9 show $x(t)$ and $\theta(t)$ respectively, and the reductions in the RMS values of x, θ - amplitudes are 8.3 and 7.4 times respectively.

Finally it is now presented the case where force F_c is applied when $t = 0.40$ s and $\Omega = \omega = \omega_c = 400$ rd/s. In this case the phase angle of F_c (related to F) is π . The RMS of x - amplitude after the application of F_c is 24.6 times smaller. A reduction of 15.9 times is found in the RMS value of θ - amplitude. Figure 10 shows $x(t)$.

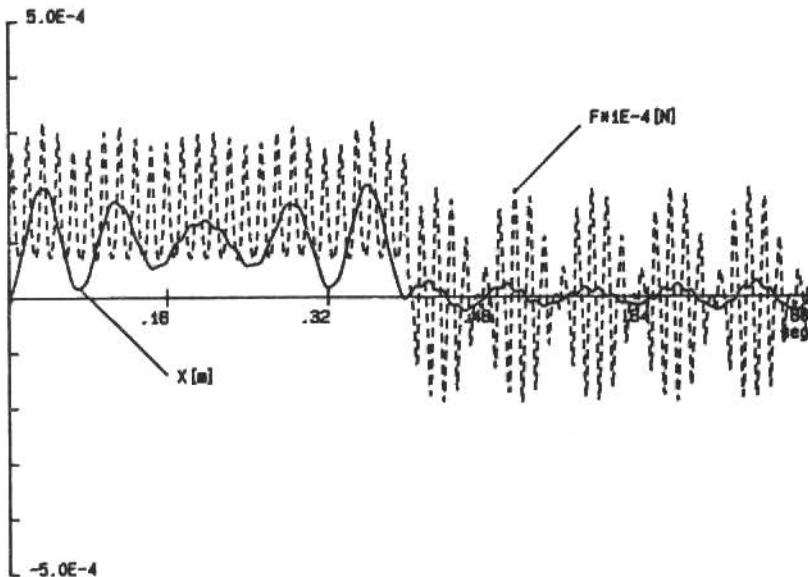


Figure 8. $x(t)$ $\Omega = \omega = 400$ rd/s and $\omega_c = 480$ rd/s
Control force applied when $t = 0.40$ sec.

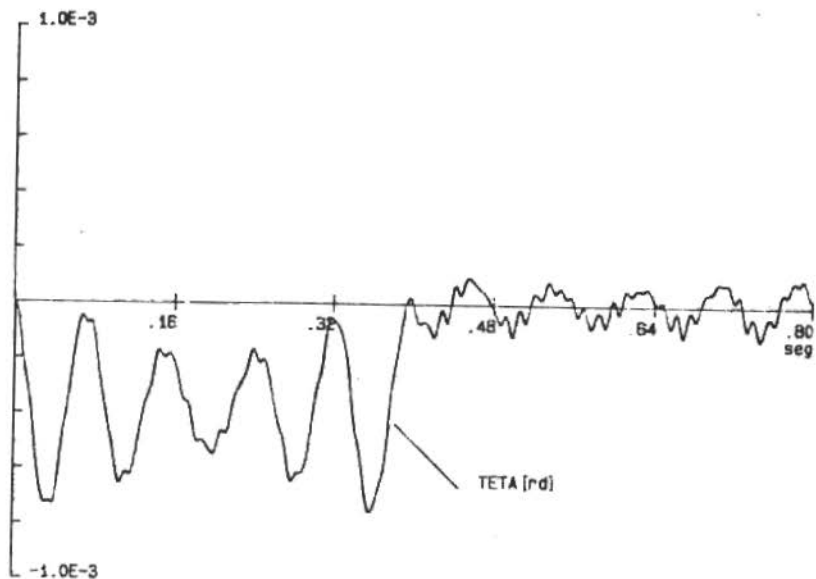


Figure 9. $\theta(t)$ for $\Omega = \omega = 400$ rd/s and $\omega_c = 480$ rd/s
Control force applied when $t = 0.40$ sec.

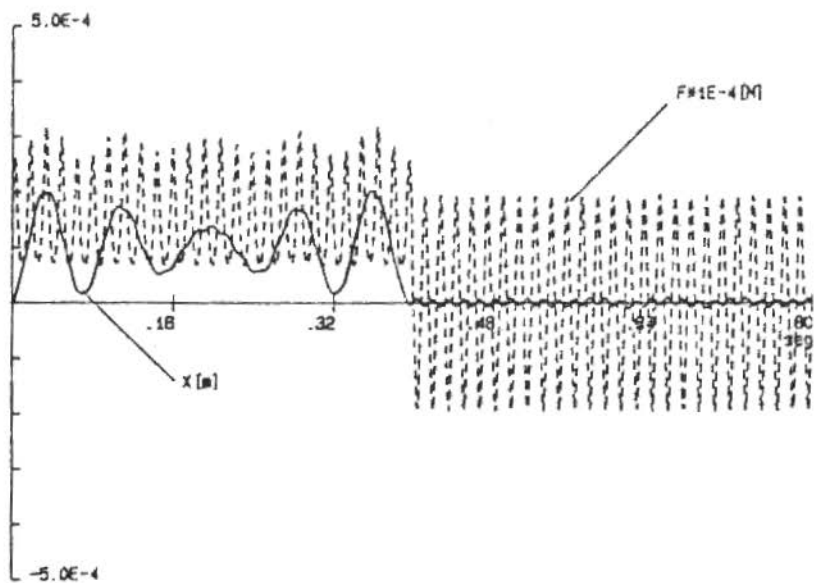


Figure 10. $x(t)$ for $\Omega = \omega = \omega_c = 400$ rd/s
Control force applied when $t = 0.40$ sec.

CONCLUSIONS

A method to analyse the response of a rotor-system in the presence of non-linear magnetic excitations is presented.

The system stiffness is changed in the presence of magnetic excitations.

Significant reductions of the amplitudes of vibrations can be obtained when introducing an extra magnetic force.

REFERENCES

- [1] Steffen Jr., V. and Lépore N., F.P. - A control force to reduce magnetically excited rotor vibration. Mécanique, Matériaux Electricité, n.404, France (1984).
- [2] Steffen Jr., V. and Lépore N., F.P. - Dynamical behavior of a flexible rotor under non-linear excitation. Brazilian Congress of Mechanical Engineering - COBEM 85, São José dos Campos, ITA CTA, Dec. 1985.
- [3] Lépore N., F.P. and Steffen Jr., V. - Mancal magnético. Technical Report to CNPq Proc. Nb.402510/80, Brazil (1983).
- [4] Schweitzer, G. and Ulbrich, H. - Magnetic bearings - A Novel type of suspension. Second Int. Conf. on Vibrations in Rotating Machinery, Inst. of Mech. Eng., Cambridge, C273/80 (1980).
- [5] Weber, H.I. and Zampieri, D.E. - Betrachtung zur anwendung aktiver lager bei schnellaufenden rotoren. VDI, Berichte Nr. 269 (1976).
- [6] Schweitzer, G. and Lange, R. - Characteristics of a magnetic rotor bearing for active vibration control. Conf. on Vibration Rotaring Machinery, Inst. Mech. Eng., Cambridge (1976) C 239/76.
- [7] Zampieri, D.E. - Einfluss einer aktiven magnetischen lagerung auf das dynamische verhalten eines rotors. Dr. Ing. Thesis, Institut B für Mechanik der Universität Stuttgart (Germany), 1981.

RANDOM VIBRATIONS AND NOISE GENERATION IN RAILWAY WHEELS

K. Popp

Institute of Mechanics
University of Hannover
FRG

INTRODUCTION

High frequency bending vibrations of railway wheels lead to sound radiation and noise generation. One can distinguish between squealing noise due to stick-slip excitation during curving motion and rolling noise due to random excitation which also occurs during tangent motion. Squealing noise has been extensively investigated in [1], [2], where also literature reviews can be found. In the present paper the rolling-noise problem is considered in detail. The aim is an integrated analysis comprising the dynamic response of a wheel-disc to broad band random excitation, the sound radiation, and the sound power calculation taking account of the frequency-dependent sensitivity of the human ear (A-evaluation). The resulting A-evaluated sound power represents a scalar measure of the acoustic quality of a wheel and can serve as a performance index to classify various types of wheel-disc design. The analysis of wheel models with different cross-sections shows a significant influence of the geometric parameters on the radiated sound power. Thus, an acoustical optimal wheel design is possible.

The most important assumptions in the present paper are rigid rails and point contact between wheel and rail. In the contact point the excitation force is applied. The time dependence of the excitation force is approximated by a stationary white noise

process with zero mean. However, a colored noise excitation process can also be applied in the methods described subsequently.

NATURAL VIBRATIONS

For the analysis of natural wheel vibrations the finite element model applied in [1],[2] has proven itself well. Ring elements are used based on Mindlin's theory of thick plates which takes into account the effects of shear deformation and rotatory inertia. The finite element model results in the natural circular frequencies $\omega_{k\ell}$ and the corresponding bending mode shapes $W_{k\ell}(r,\phi)$ for the ℓ -th natural vibration with k modal diameters, where r,ϕ denote the polar coordinates of the wheel-disc mid plane. A comparison of calculated results with holographic measurements for a 1:5.5 scale wheel model shows excellent agreement even at higher frequencies, cf. Figure 1. In the worst case the frequency error amounts to only 3%. In this paper the mode shapes are utilized for a modal expansion of the wheel's vibration response to random excitation.

RANDOM VIBRATIONS AND SOUND POWER ANALYSIS

System Description. The system under consideration is shown in the block diagram, Figure 2a. The system behavior can be analyzed in the frequency domain as well as in the time domain. In the following both approaches are applied. First the description of the wheel disc dynamics is given.

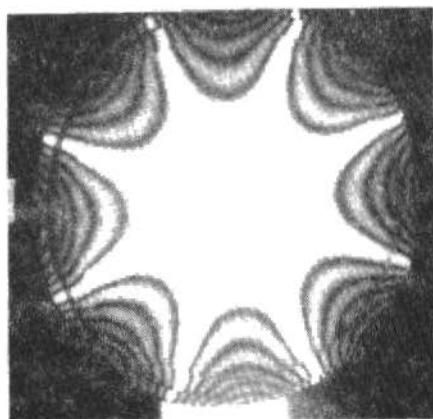
The modal expansion of the wheel deflection $w(r,\phi,t)$ reads

$$w(r,\phi,t) = \underline{W}^T(r,\phi) \underline{c}(t) \quad , \quad (1)$$

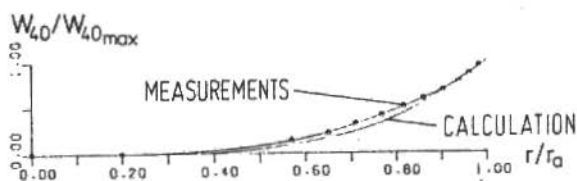
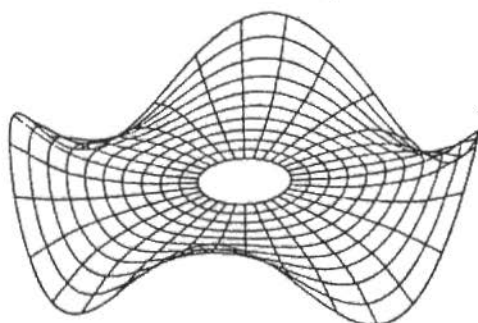
where $\underline{W}(r,\phi)$ comprises the bending modes $W_{k\ell}(r,\phi)$. The corresponding $n \times 1$ -vector $\underline{c}(t)$ of modal coordinates follows from the differential eq., cf. [1],[2],

$$\ddot{\underline{c}}(t) + \underline{\eta} \dot{\underline{c}}(t) + \underline{\omega}^2 \underline{c}(t) = \underline{\mu}^{-1} \underline{W}(r_c, \phi_c) f(t) \quad , \quad f(t) \sim (0,q) \quad . \quad (2)$$

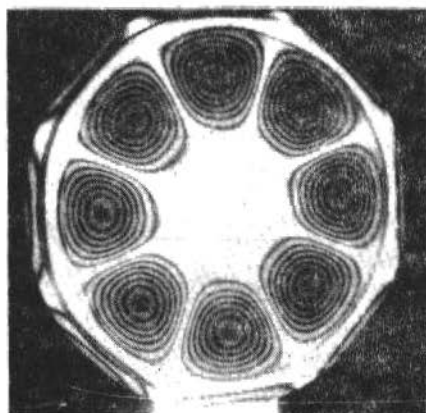
Here, $\underline{\omega}^2 = \underline{\text{diag}}(\omega_{k\ell}^2)$ is the matrix of natural frequencies and $\underline{\mu}$ denotes the matrix of generalized masses. The stochastic excitation



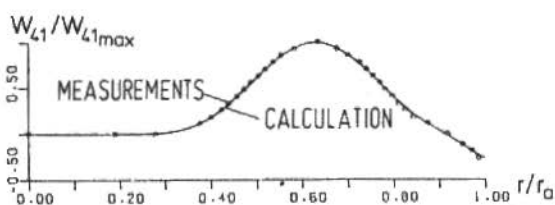
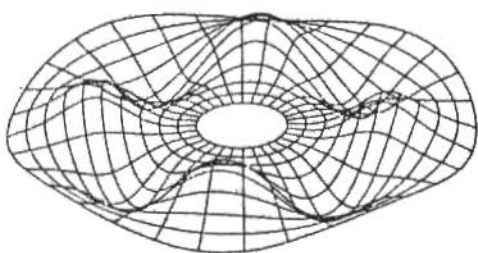
$f_{40M} = 8543 \text{ Hz}$, $f_{40C} = 8488 \text{ Hz}$
 $\Delta f = 0.6\%$



MODE 4/0



$f_{41M} = 14121 \text{ Hz}$, $f_{41C} = 13855 \text{ Hz}$
 $\Delta f = 1.9\%$



MODE 4/1

Figure 1. Comparison of calculated and measured mode shapes

force $f(t)$ is characterized by zero mean and correlation function $R_f(\tau) = q \delta(\tau)$, where q is the noise intensity, τ the correlation time, and δ denotes the Dirac-function. The excitation acts at the contact point with coordinates r_c, ϕ_c . In (2) modal damping is added, where η denotes the damping coefficient. Corresponding to (1), (2), the complex transfer function $F(r, \phi, \omega)$ of the wheel reads

$$F(r, \phi, \Omega) = \underline{W}(r, \phi) [\omega^2 (1 + i\eta\Omega) - \Omega^2 \underline{E}]^{-1} \underline{u}^{-1} \underline{W}(r_c, \phi_c) \quad (3)$$

where Ω denotes the frequency variable and \underline{E} is the unity matrix. The excitation process can be described by the power spectral density $S_f(\Omega) = q/2\pi = \text{const}$ which follows from the correlation function by means of a Fourier transform. Now, the power spectral density $S_w(r, \phi, \Omega)$ of the wheel deflections follows from the relation

$$S_w(r, \phi, \Omega) = |F(r, \phi, \Omega)|^2 S_f(\Omega) \quad (4)$$

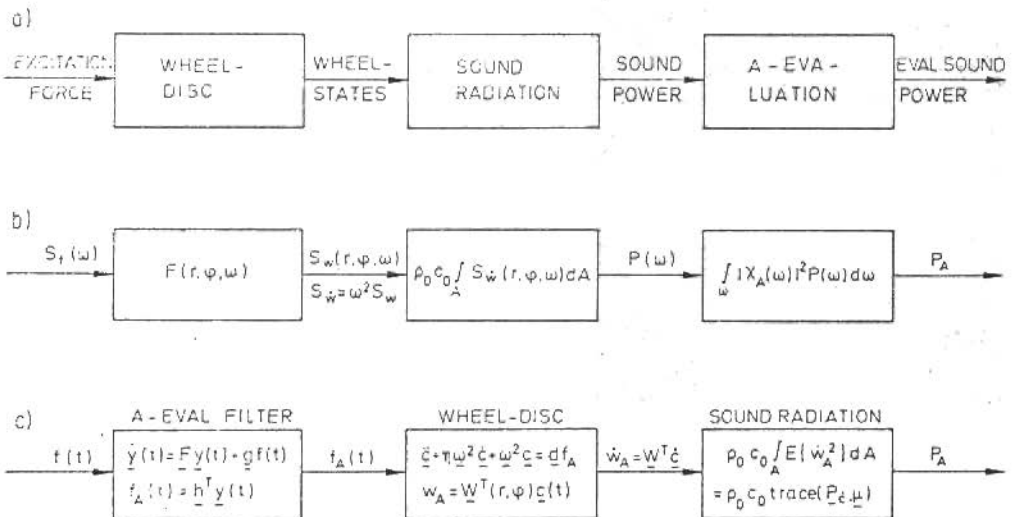


Figure 2. Block diagrams:
 a) system description,
 b) analysis in the frequency domain,
 c) analysis in the time domain.

Sound Power. The sound power P radiated into the half space can be calculated using the radiation efficiency σ , cf. [3],

$$P = \sigma \rho_0 c_0 \int_A \bar{w}^2(r, \phi) dA, \quad dA = r d\phi dr, \quad (5)$$

where ρ_0, c_0 are the density and the sound velocity of the surrounding air, respectively, and \bar{w}^2 is the mean square velocity of the wheel bending vibrations. In the interesting frequency range of railway wheels it yields $\sigma \approx 1$, cf. [1]. Thus, from (5) the sound power follows using the frequency domain and the time domain approach, respectively,

$$P = \rho_0 c_0 \int_A \int_{\Omega} S_w^*(r, \phi, \Omega) d\Omega dA = \rho_0 c_0 \int_A E\{\dot{w}^2(r, \phi, t)\} dA. \quad (6)$$

Here, $S_w^*(r, \phi, \Omega)$ is the power spectral density of the wheel velocity due to bending vibration and $E\{\cdot\}$ denotes the expectation operator. Utilizing (4) the frequency domain approach results in

$$S_w^*(r, \phi, \Omega) = \Omega^2 S_f(r, \phi, \Omega), \quad (7)$$

$$P(\Omega) = \rho_0 c_0 \int_A \Omega^2 |F(r, \phi, \Omega)|^2 S_f(\Omega) dA, \quad (8)$$

$$P = \int_{\Omega} P(\Omega) d\Omega = \rho_0 c_0 \int_A \int_{\Omega} \Omega^2 |F(r, \phi, \Omega)|^2 S_f(\Omega) d\Omega dA, \quad (9)$$

where the sound power spectral density $P(\Omega)$ has been introduced.

Regarding (1), the time domain approach leads to

$$E\{\dot{w}(r, \phi, t)\dot{w}(r, \phi, t)\} = \underline{W}^T(r, \phi) E\{\dot{\underline{c}}(t)\dot{\underline{c}}^T(t)\} \underline{W}(r, \phi), \quad (10)$$

$$\underline{P}_c = E\{\dot{\underline{c}}(t)\dot{\underline{c}}^T(t)\}, \quad (11)$$

$$P = \rho_0 c_0 \int_A \underline{W}^T(r, \phi) \underline{P}_c \underline{W}(r, \phi) dA, \quad (12)$$

Here, \underline{P}_c denotes the covariance matrix of the modal velocities which follows from a covariance analysis. For this, eq.(2) is represented in the state space,

$$\dot{\underline{x}}(t) = \underline{A} \underline{x}(t) + \underline{b} f(t) \quad , \quad f(t) \sim (0, q) \quad , \quad (13)$$

where

$$\underline{x}(t) = \begin{bmatrix} \underline{c}(t) \\ \dot{\underline{c}}(t) \end{bmatrix} \quad , \quad \underline{A} = \begin{bmatrix} 0 & \underline{E} \\ -\omega^2 & -\eta \omega^2 \end{bmatrix} \quad , \quad \underline{b} = \begin{bmatrix} 0 \\ \underline{u}^{-1} \underline{W}(r_C, \phi_C) \end{bmatrix} \quad . \quad (14)$$

Now, the covariance matrix $\underline{P}_x = E\{\underline{x}(t) \underline{x}^T(t)\}$ in steady state is given by the solution of the Ljapunov matrix eq., cf. [4],

$$\underline{A} \underline{P}_x + \underline{P}_x \underline{A}^T + q \underline{b} \underline{b}^T = \underline{0} \quad . \quad (15)$$

The solution matrix \underline{P}_x can be partitioned,

$$\underline{P}_x = \begin{bmatrix} \underline{P}_C & * \\ * & \underline{P}_C^* \end{bmatrix} \quad , \quad (16)$$

where the covariance matrix \underline{P}_C required in (12) appears as a submatrix.

A-Evaluation. The frequency depending sensitivity of the human ear is given in [5] as transfer function $\chi_A(\Omega)$,

$$\chi_A(\Omega) = \frac{A\Omega^4}{\left(1 + i \frac{\Omega}{\Omega_1}\right)^2 \left(1 + i \frac{\Omega}{\Omega_2}\right)^2 \left(1 + i \frac{\Omega}{\Omega_3}\right) \left(1 + i \frac{\Omega}{\Omega_4}\right)} \quad , \quad (17)$$

where

$$\left. \begin{aligned} A &= (41 \cdot 75 \cdot 10^9)^{-1} \text{ s}^7 \quad , \\ \Omega_1 &= 2\pi \cdot 20 \cdot 6 \text{ s}^{-1} \quad , \\ \Omega_2 &= 2\pi \cdot 737 \cdot 9 \text{ s}^{-1} \quad , \\ \Omega_3 &= 2\pi \cdot 207 \cdot 7 \text{ s}^{-1} \quad , \\ \Omega_4 &= 2\pi \cdot 12200 \text{ s}^{-1} \quad . \end{aligned} \right\} \quad (18)$$

The corresponding time domain representation results in an A-evaluation filter of order six,

$$\dot{\underline{y}}(t) = \underline{F} \underline{y}(t) + \underline{g} u(t) \quad , \quad \underline{u}_A(t) = \underline{h}^T \underline{y}(t) \quad , \quad (19)$$

where $u(t)$, $u_A(t)$ are input and output variables, respectively, the 6 1-vector $\underline{y}(t)$ comprises the filter state variables, the matrix \underline{F} and the vectors \underline{g} , \underline{h} of appropriate dimensions characterize the filter dynamics,

$$\underline{F} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -a_6 & -a_5 & -a_4 & -a_3 & -a_2 & -a_1 \end{bmatrix}, \quad \underline{g} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \quad \underline{h} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ a \\ 0 \end{bmatrix}. \quad (20)$$

The entries have been calculated as

$$\begin{aligned} a &= 7.397586 \cdot 10^9, \\ a_1 &= 158.8817 \cdot 10^3, \\ a_2 &= 6.734726 \cdot 10^9, \\ a_3 &= 33.4358 \cdot 10^{12}, \\ a_4 &= 27.75376 \cdot 10^{15}, \\ a_5 &= 5.3029 \cdot 10^{18}, \\ a_6 &= 308.8492 \cdot 10^{18}. \end{aligned} \quad (21)$$

The frequency response $|\chi_A(\Omega)|$ of the filter is shown in Figure 3. It lies within the tolerance bandwidth specified in [5] for any frequency. The maximum deviation from the nominal values is smaller than 0.4 dB.

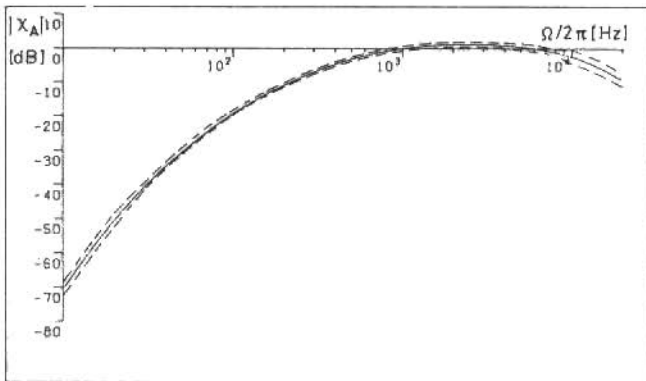


Figure 3. Frequency response of the A-evaluation filter ———, tolerance bandwidth specified in [5] - - - - -,

A-Evaluated Sound Power. The A-evaluated sound power P_A can be calculated in the frequency domain as well as in the time domain. In frequency domain, cf. Figure 2b, follows from (8), (9), (17),

$$P_A = \int_{\Omega} |\chi_A(\Omega)|^2 P(\Omega) d\Omega = \rho_0 c_0 \int_A \int_{\Omega} \Omega^2 |\chi_A(\Omega)|^2 |F(r, \phi, \Omega)|^2 S_f(\Omega) d\Omega dA \quad (22)$$

The time domain approach starts from the A-evaluated wheel velocity $\dot{w}_A(r, \phi, t)$,

$$\dot{w}_A(r, \phi, t) = \underline{W}^T(r, \phi) \dot{\underline{c}}_A(t) \quad (23)$$

$$E\{\dot{w}_A^2(r, \phi, t)\} = \underline{W}^T(r, \phi) E\{\dot{\underline{c}}_A(t) \dot{\underline{c}}_A^T(t)\} \underline{W}(r, \phi) \quad (24)$$

$$\hat{\underline{P}}_C = E\{\dot{\underline{c}}_A(t) \dot{\underline{c}}_A^T(t)\} \quad (25)$$

$$P_A = \rho_0 c_0 \int_A E\{\dot{w}_A^2(r, \phi, t)\} dA = \rho_0 c_0 \int_A \underline{W}^T(r, \phi) \hat{\underline{P}}_C \underline{W}(r, \phi) dA \quad (26)$$

Thus, filtering of the vector $\dot{\underline{c}}(t) = [\dot{c}_1(t), \dots, \dot{c}_n(t)]^T$ of modal velocities is required. However, since the transfer function $\chi_A(\Omega)$ is scalar, cf. (17), the filter can be shifted to the system input resulting in the same input-output relation in steady state. This is obvious from the system representation in the frequency domain, Figure 2b. After the filter shift, the corresponding block diagram is shown in Figure 2c. This is the basis for the analysis in the time domain. Instead of n modal velocities $\dot{c}_i(t)$, $i = 1(1)n$, only one variable, $f(t)$, has to be filtered, which reduces the computational amount tremendously. Using the filter eq. (19), where now $u(t) \equiv f(t)$, $u_A(t) \equiv f_A(t)$ hold, and rewriting the wheel eq. (13) with $f(t)$ as input variable, leads to

$$\dot{\underline{y}}(t) = \underline{F} \underline{y}(t) + \underline{g} f(t), \quad f(t) \sim (0, q) \quad (27)$$

$$f_A(t) = \underline{h}^T \underline{y}(t) \quad (28)$$

$$\dot{\underline{x}}(t) = \underline{A} \underline{x}(t) + \underline{b} f_A(t), \quad \underline{x}(t) = \left[\dot{\underline{c}}_A^T(t) \dot{\underline{c}}_A^T(t) \right]^T \quad (29)$$

Here, \underline{F} , \underline{g} , \underline{h}^T are given in (20) and \underline{A} , \underline{b} in (14). The eq. (27) - (29) can be combined to one state equation describing the overall system dynamics,

$$\dot{\underline{\hat{x}}}(t) = \underline{\hat{A}} \underline{\hat{x}}(t) + \underline{\hat{b}} f(t), \quad f(t) \sim (0, q), \quad (30)$$

where

$$\underline{\hat{x}}(t) = \begin{bmatrix} \underline{y}(t) \\ \underline{x}(t) \end{bmatrix}, \quad \underline{\hat{A}} = \begin{bmatrix} \underline{F} & \underline{0} \\ \underline{b} \underline{h}^T & \underline{A} \end{bmatrix}, \quad \underline{\hat{b}} = \begin{bmatrix} \underline{g} \\ \underline{0} \end{bmatrix}. \quad (31)$$

Analogous to (15), the covariance matrix $\underline{\hat{P}} = E\{\underline{\hat{x}}(t) \underline{\hat{x}}^T(t)\}$ in steady state follows from the Ljapunov matrix equation

$$\underline{\hat{A}} \underline{\hat{P}} + \underline{\hat{P}} \underline{\hat{A}}^T + q \underline{\hat{b}} \underline{\hat{b}}^T = 0, \quad (32)$$

which can efficiently be solved using e.g. the algorithms given in [6]. The solution matrix $\underline{\hat{P}}$ can be partitioned,

$$\underline{\hat{P}} = \begin{bmatrix} \underline{\hat{P}}_y & * & * \\ * & \underline{\hat{P}}_c & * \\ * & * & \underline{\hat{P}}_c^* \end{bmatrix}, \quad (33)$$

where the covariance matrix $\underline{\hat{P}}_c^*$ required in (26) appears as a submatrix. The computation of the A-evaluated sound power P_A , cf. (26), can further be simplified. Utilizing the modified generalized mass matrix $\underline{\hat{\mu}}^*$,

$$\underline{\hat{\mu}}^* = \text{diag}(\mu_i^*) = \int_A \underline{W}(r, \phi) \underline{W}^T(r, \phi) dA, \quad (34)$$

which is diagonal since the eigenfunctions are orthogonal, it follows from (26)

$$P_A = \rho_0 c_0 \text{tr}(\underline{\hat{P}}_c^* \underline{\hat{\mu}}^*) = \rho_0 c_0 \sum_{i=1}^n \hat{\sigma}_{c_i}^2 \mu_i^*. \quad (35)$$

Here, $\hat{\sigma}_{c_i}^2$, $i = 1(1)n$, denote the variances of the A-evaluated modal velocities. Thus, a purely algebraic method for the computation of the A-evaluated sound power P_A has been developed.

A comparison of the two approaches resulting in the A-evaluated sound power P_A , cf. (22) and (35), shows clearly the advantages of the time domain analysis with respect to the numerical effort required for the calculation. Applying the algebraic method, cf. (32), (35), leads to computing times of at least factor ten faster than in case of the frequency domain analysis. This aspect is important if the evaluated sound power P_A is used as performance index in an optimization procedure for an acoustical optimal wheel-disc design.

Instead of the white noise excitation process also a colored noise process can be used. Then, in the time domain approach the system input has to pass an appropriate shape filter in addition to the A-evaluation filter.

EXAMPLE

The developed algebraic method has been applied to the analysis of four different railway wheels. Figure 4 shows the

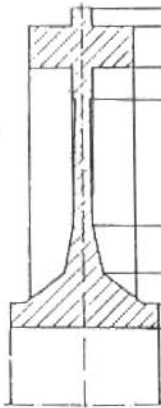
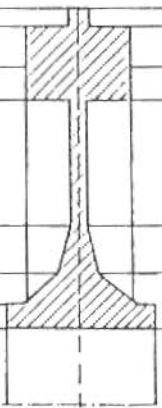
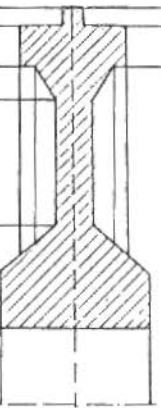

WHEEL TYPE	N	A	B	C
WHEEL GEO - METRY				
P_A / P_{AN}	1	0,627	1,087	0,271

Figure 4. Analysis of different types of wheel-disc design
($\sigma = 0.0001$)

corresponding cross-sections as used in the finite element model. The geometry of wheel type N is taken from a normal wheel. Compared to N, wheel type A shows a thicker rim while wheel type B shows a thicker web and thicker hub. Wheel type C consists only of a disc rim. The A-evaluated sound power has been calculated using the structural damping coefficient $\mu = 0.0001$. The results are related to the sound power P_{AN} of wheel type N. It turns out that wheel type A radiates only 63% of the sound power of P_{AN} , because the high input impedance of the thick rim has a beneficial influence. Wheel type B is worst than wheel type N. However, best results are gained for the ring, wheel type C, due to the much smaller surface. These results show that the sound power radiation depends strongly on the wheel-disc geometry. Thus, an acoustical optimal wheel design within the given geometry limits is possible.

CONCLUSIONS

An integrated analysis of the vibration behavior and sound radiation of railway wheels under broad band random excitation has been performed. The calculation of the corresponding sound power is given in frequency domain and time domain as well. The frequency dependent sensitivity of the human ear (A-evaluation) has been taken into account. Here, the A-evaluation which is given in the frequency domain has been transformed into time domain resulting in an A-evaluation filter. Based on the covariance analysis a very efficient algebraic method for the calculation of the A-evaluation sound power is given and applied to classify various types of wheel-disc design. It turns out that the cross-section geometry of a wheel-disc has a strong influence on the radiated sound power. Thus, an acoustical optimal wheel design is possible, where the A-evaluated sound power can serve as performance index.

REFERENCES

- [1] Schneider, E. - Schwingungsverhalten und Schallabstrahlung von Schienenrädern. , Dr.-Ing. Diss. University of Hannover, Fortschr.-Ber. VDI Reihe 11 Nr. 74, Düsseldorf 1985.
- [2] Popp, K.; Schneider, E; and Irretier, H. - Noise generation in railway wheels due to rail-wheel contact forces. Proc. 9th IAVSD-Symp-. held in Linköping, Sweden, June 24-28, 1985.

- [3] Morfey, C.L. - Theory of acoustics II. In: Noise and Vibration (Eds.: White, R.G., Walker, J.G.) Chichester 1982, p.64.
- [4] Müller, P.C. and Schiehlen, W. - Lineare schwingungen. Wiesbaden 1976.
- [5] DIN IEC 651: Schallpegelmesser, 12. 1981.
- [6] Bartels, R.H. and Stewart, G.W. - Solution of the matrix equation $AX+XB=C$. Communications of the ACM, 15 (9) : 820-826, 1972.

STATE ESTIMATION OF MULTIVARIABLE SYSTEMS USING THE DUALITY CONCEPT

Agenor de Toledo Fleury

Atair Rios Neto

Instituto de Pesquisas Espaciais - INPE/MCT

São José dos Campos/SP

ABSTRACT

This work presents the development of a new estimator for dynamic systems based upon the duality concept between state estimator and optimal control problems. This leads to an estimator algorithm similar to the Extended Kalman Filter for use in real time nonlinear system, which can recursively be approximated by linear systems. The duality concept is used to transform the original estimation problem into an equivalent one of virtual control. This control problem is then used to generate an adaptive, locally convergent algorithm where, instead of full state estimation, one has to estimate just a control vector with smaller dimension than the state vector. The control formulation also allows the development of convergence acceleration criteria. Three convergence criteria, as well as adaptive noise methods to compensate for model errors, are then developed and used with the proposed estimator in the test problem created to validate the estimator. The algorithm is implemented in a digital computer to estimate the orbit of a low Earth orbit satellite under simulated conditions. Numerical results include test cases considering different initial values for the estimator and different standard deviations for the observations provided by three Earth stations, allowing a preliminary evaluation of the estimator characteristics.

INTRODUCTION

A very common situation arising when one has to estimate the state of a multivariable system is the problem of lack of knowledge about the system dynamics. In this case, one has to adopt a dynamic model for the estimator which is only a crude approximation of reality. Unmodelled dynamic effects, specially in the case of non linear systems, can cause divergence of the estimates.

In order to avoid divergence, error compensation techniques are usually employed with different versions of the Kalman Filter (e.g. Maybeck, 1979, Jazwinski, 1970; Gelb et al., 1974). The techniques usually explore the information given by the observation residues to either directly estimate the unmodelled effects (e.g.: Tapley and Ingram, 1973; Cruz and Rios Neto, 1980; Rios Neto and Cruz; 1985) or to condition the state error covariance matrix to keep the capability of the estimator to extract information from the new observations (e.g. Jazwinski, 1969; Rios Neto and Kuga, 1981, 1982, 1985). In the resulting procedures there is always an increase in the number of variables to be estimated.

This work presents a new alternative scheme for the multi-variable state estimation problem, specially for nonlinear systems. The duality concept between the estimation and optimal control problems is explored to transform the original estimation problem into one of tracking the observations with a virtual control. The key idea is that, independently of the dynamic model for the estimator, if the system is completely controllable and observable, one can choose a desirable control action to drive the system towards a region defined by the observations, in a finite time interval. Without the need of increasing the number of estimated variables, the observation residues are used to estimate the virtual control necessary to update the estimate of the state. Besides that, the virtual nature of the tracking control allows the possibility of choosing the control actions adequate to better extract the information contained in the observations (Rios Neto and Fleury, 1984; Fleury, 1985). Therefore, the proposed estimator is indicated for situations where there is a great lack of knowledge about the system dynamics but there is a high local level of information in the observations.

The virtual formulation also allows to get some advantages

from the system structure. As will be explained in item 3.2, one can define a control matrix correction in order to improve convergence.

Preliminary tests of the proposed estimator were done under digitally simulated conditions for the problem of real time orbit estimation of a low altitude artificial satellite. Results are shown to be satisfactory for this exploratory phase of the procedure.

PROPOSED PROCEDURE

The problem to be solved is the state estimation of a multi-variable dynamic system of the type:

$$\dot{x} = f(x,t) + f^n(x,t) + \bar{G}(t) w(t), \quad (1)$$

$$y(t_k) = h_k(x(t_k), t_k) + v(t_k), \quad k = 1, 2, \dots \quad (2)$$

where x is the $n \times 1$ state vector; $w(t)$ and $v(t_k)$ are $m \times 1$ and $r \times 1$ independent Gaussian white noises with the usual hypothesis of $w(t)$ being independent of the past state and $v(t_k)$ being independent of the state, with distributions defined by zero means and covariances:

$$E[w(t) w^T(\tau)] = Q(t) \delta(t-\tau), \quad (3)$$

$$E[v(t_k) v^T(t_j)] = R(k) \delta_{kj}, \quad (4)$$

where $\delta(t-\tau)$ is the Dirac delta function and δ_{kj} is the Kronecker symbol.

In this problem, the term $f^n(x,t)$ in Equation (1) represents the unknown part of the dynamic model, which usually cannot be included in the estimator model because of lack of knowledge about the system dynamics. Consider now a typical discretization interval, (t_k, t_{k+1}) . In the prediction phase of an extended Kalman filter (e.g. Jazwinski, 1970), a nominal trajectory is generated by:

$$\dot{\bar{x}} = f(\bar{x}, t); \quad \bar{x}(t_k) = \hat{x}(t_k | t_k) \quad (5)$$

where $\hat{x}(t_k | t_k)$ is the estimate in t_k .

Linearizing the system given in Equation (1) around the nominal trajectory and defining:

$$\Delta x(t) = x(t) - \bar{x}(t), \quad (6)$$

one obtains a first order approximation of the propagated error as given by:

$$\Delta x(t_{k+1}) = \phi(t_{k+1}, t_k) \Delta x(t_k) + \int_{t_k}^{t_{k+1}} \phi(t_{k+1}, s) \bar{G}(s) w(s) ds \quad (7)$$

In Equation (7), $\phi(\dots)$ is the state transition matrix associated to the linear system:

$$\dot{\Delta x}(t) = f_x(\bar{x}, t) \Delta x(t) + \bar{G}(t) w(t), \quad (8)$$

that is:

$$\dot{\phi}(t, t_k) = f_x(\bar{x}, t) \phi(t, t_k) ; \phi(t_k, t_k) = I \quad (9)$$

where the subindex x indicates a partial derivation with respect to the state.

The propagated error can be regarded as the a priori information in t_{k+1} which is the information based on previously processed observations:

$$\Delta x(t_{k+1}) = \Delta \bar{x}(t_{k+1} | t_k) + \eta(k+1 | k). \quad (10)$$

Since the nominal trajectory was taken with the initial value in t_k equal to the estimated value at that time, then the propagated estimate in t_{k+1} is necessarily zero. Therefore, if the approximations of Equations (5) and (7) are assumed, there results:

$$\Delta x(t_{k+1}) = 0 + \eta(k+1 | k), \quad (11)$$

where $\eta(k+1 | k)$ is zero mean conditioned on the observations already processed with covariances given by:

$$E[\eta(k+1|k) \eta^T(k+1|k)] \triangleq P(k+1|k) =$$

$$\phi(t_{k+1}, t_k) P(k|k) \phi^T(t_{k+1}, t_k) + \int_{t_k}^{t_{k+1}} \phi(t_{k+1}, s) \bar{G}(s) Q(s) G^T(s) \phi^T(t_{k+1}, s) ds. \quad (12)$$

In the proposed procedure the state estimation of the system in Equation (1) is approximated by the estimation of a "virtual control". Defining:

$$\dot{x}_c = f(x_c, t) + G(t) u(t), \quad (13)$$

the control vector $u(t)$ is to be estimated in order to force the controlled state x_c to be a good approximation of the true state x by tracking the observations given by Equation (2). From an heuristic point of view, the idea is quite simple.

Consider a typical interval (t_k, t_{k+1}) . Since the model of the estimator is not a good approximation to the real system, within the interval t_k to t_{k+1} , the propagated trajectory will deviate relative to the true one. A control action, based on the observations residues, is then calculated in t_{k+1} to update the estimate of the state vector. This action changes the initial condition for the nominal trajectory in t_{k+1} up to a point closer to the true trajectory than the original one. In some sense, this procedure resembles the extended Kalman filter but it must be pointed out that the virtual control scheme requires a number of estimated control components just equal to the number of degrees of freedom of the system (controllability) and also requires a number of observations with enough level of information which guarantees the calculation of the control action in t_{k+1} (observability). There results an adaptive procedure that forces the dynamic model in the estimator towards the true trajectory.

To obtain the control $u(t)$ in the interval t_k to t_{k+1} , one assumes u as a first order perturbation and takes the controlled trajectory as:

$$\dot{\bar{x}}_C = f(\bar{x}_C, t), \quad \bar{x}_C(t_k) = \hat{x}(t_k) = \hat{x}(t_k | t_k). \quad (14)$$

Using the same steps given by Equations (6) and (7):

$$\Delta \dot{x}_C = f_x(\bar{x}_C, t) \Delta x_C + G(t) u(t), \quad \Delta x_C(t_k) = 0, \quad (15)$$

where $u(t)$ is modelled as a step process and calculated to satisfy:

$$\Delta x_C(t_{k+1}) = \Delta x(t_{k+1}), \quad (16)$$

$$y(t_{k+1}) = h_{k+1}(\bar{x}_C(t_{k+1}) + \Delta x_C(t_{k+1}), t_{k+1}) + v(t_{k+1}). \quad (17)$$

From Equations (11) and (16) there results:

$$0 = \Delta x_C(t_{k+1}) + \eta(k+1|k), \quad (18)$$

But from Equation (15) one obtains:

$$\Delta x_C(t_{k+1}) = \left(\int_{t_k}^{t_{k+1}} \phi_C(t_{k+1}, s) G(s) ds \right) u(t_k) \triangleq \gamma(k+1, k) u(t_k) \quad (19)$$

where $\phi_C(t, t_k)$ is the transition matrix associated to Equation(15). From the linearization of Equation (17) one gets:

$$\Delta y(t_{k+1}) = \frac{\partial}{\partial x_C} h_{k+1}(\bar{x}_C(t_{k+1}), t_{k+1}) \Delta x_C(t_{k+1}) + v(t_{k+1}) \quad (20)$$

where high order terms have been disregarded.

Finally, combining the results of Equations (18), (19) and (20), the following problem of parameter estimation results:

$$0 = \gamma(k+1, k) u(t_k) + \eta(k+1|k), \quad (21)$$

$$\Delta y(t_{k+1}) = H(k+1) \gamma(k+1, k) u(t_k) + v(t_{k+1}). \quad (22)$$

where

$$H(k+1) \triangleq \frac{\partial}{\partial x_C} h_{k+1}(\bar{x}_C(t_{k+1}), t_{k+1}). \quad (23)$$

Using a Gauss-Markov minimum variance estimator (e.g. Liebelt, 1967 and Maybeck, 1979), estimates of $u(t_k)$ are obtained:

$$\hat{u}(t_k) = P_u(k+1|k) \gamma(k+1, k) H^T(k+1) R^{-1}(k+1) \Delta y(t_{k+1}) \quad (24)$$

$$P_u(k+1|k) = [\gamma^T(k+1, k) P^{-1}(k+1|k) \gamma(k+1, k) + \gamma^T(k+1, k) \cdot H^T(k+1) R^{-1}(k+1) H(k+1) \gamma(k+1, k)], \quad (25)$$

where $R(k+1)$ and $P(k+1|k)$ are as defined in Equations (4) and (12).

To recover the estimate of the state, one shall combine Equations (19) and (24) to get:

$$\Delta \bar{x}_c(t_{k+1}) = \gamma(k+1, k) \hat{u}(t_k) \quad (26)$$

and take the approximation:

$$\bar{x}(t_{k+1}|t_{k+1}) = \bar{x}(t_{k+1}|t_k) + \Delta \bar{x}_c(t_{k+1}). \quad (27)$$

To recover the covariance matrix of the error in the estimate, it is only necessary to consider Equations (16), (19) and (27) to get:

$$e(t_{k+1}|t_{k+1}) \triangleq x(t_{k+1}) - \bar{x}(t_{k+1}|t_{k+1}) = \Delta x - \Delta \bar{x}_c \quad (28)$$

and

$$P(k+1|k+1) \triangleq E[e(t_{k+1}|t_{k+1}) e^T(t_{k+1}|t_{k+1})] = \gamma(k+1, k) P_u(k+1|k) \cdot \gamma^T(k+1, k). \quad (29)$$

APPLICATION

The procedure was tested under digital simulation (Burroughs B6800) for the case of real time orbit determination of a low altitude satellite with:

zero excentricity, 42° inclination, 250 km altitude,
 $A/m = 0,00076 \text{ m}^2/\text{kg}$, area over mass ratio,
 $C_D = 2.0$, drag coefficient.

To simulate the observations, a true orbit was generated by numerical integration, using a dynamic model including the influence of gravity (up to J_6 zonal and C_{44} , S_{44} tesseral coefficients), atmospheric drag and perturbations of Sun and Moon. The model adopted for $f(x_c, t)$, in Equation (13), only included the gravitational effects up to J_2 , thus characterizing a situation of lack of knowledge quite serious in the dynamics of the system. It is shown in Kuga (1982) that the application of the Extended Kalman Filter to the same problem, without any error compensation technique, cause divergence of the estimates.

Bearing in mind the approximations taken in the proposed procedure, one must expect some difficulties in dealing with this type of problem. Therefore, error compensation techniques are employed.

Adaptive State Noise Estimation

To overcome the ill-conditioning on the state error covariance matrix due to the approximation of Equation (7) for the propagation of the error and due to the nonlinearities in the observations, within the applications both $P(k+1|k)$ and $P_u(k+1|k)$ are adapted by using an Adaptive State Noise Estimation technique (Rios Neto and Kuga, 1981, 1982, 1985). This technique is the generalization of a procedure by Jazwinski (1969) and consists of adding noise to the system through the state noise covariance matrix Q .

Consider the typical interval (t_k, t_{k+1}) and assume that $Q(k)$ for this interval is diagonal or can be diagonalized:

$$Q(k) = \begin{bmatrix} q_{1k} & & Q \\ & \ddots & \\ 0 & & q_{mk} \end{bmatrix}. \quad (30)$$

Defining the observation residue, $r(k+1|k)$, and the true residue $r_v(k+1|k)$, as:

$$r(k+1|k) = y(t_{k+1}) - \bar{y}(t_{k+1}) \quad , \quad (31)$$

$$r(k+1|k) = r_v(k+1|k) + v_{k+1} \quad , \quad (32)$$

where $\bar{y}(t_{k+1})$ is the observation vector calculated with the propagated state; $r_v(k+1|k)$ is the observation residue which would be obtained if the observations were not contaminated by noise.

Now, the parameters $q_{i,k}$, $i = 1, \dots, m$ must be chosen so as to maximize the probability of occurrence of the true residue $r_v(k+1|k)$, that is, the criterion to calculate $q_{i,k}$, $i = 1, \dots, m$ is stated by:

$$\begin{aligned} \max_p [r_v(k+1|k)]. \\ q \geq 0 \end{aligned} \quad (33)$$

It can be shown (Jazwinski, 1969; Rios Neto and Kuga, 1985) that this criterion leads to:

$$(r_v(k+1|k))_0^2 = E[r_v^2(k+1|k)], \quad (34)$$

where the subindex 0 is to mean the occurred value of $r_v(k+1|k)$.

Developing both sides of Equation (34), accounting for Equation (32) and manipulating them properly, there results an expression called Pseudoobservations Equation, which allows to calculate the parameters $q_{i,k}$ in the form:

$$y'(t_{k+1}) = H'(k+1)q_k + v'_{k+1}, \quad (35)$$

where $y'(t_{k+1})$ is the $r \times 1$ pseudoobservations vector, q_k is the $m \times 1$ vector which contains the parameters $q_{i,k}$ to be estimated and v'_{k+1} is a $r \times 1$ Gaussian white noise (see Appendix for the necessary steps to obtain this equation).

The solution for q_k can now be obtained through the use of a linear Kalman Filter.

Employing the technique described above, $P_u(k+1|k)$ is substituted for:

$$P_u^C(k+1|k) = P_u(k+1|k) + Q_u(k+1|k) \quad (36)$$

in Equation (29), while $P(k+1|k)$ is generated by:

$$P(k+1|k) = \Phi(k+1, k) P(k|k) \Phi^T(k+1, k) + \bar{\gamma}(k) Q(k) \bar{\gamma}^T(k). \quad (37)$$

where

$$\bar{\gamma}(k) \triangleq \int_{t_k}^{t_{k+1}} \phi(t_{k+1}) \bar{G}(s) ds$$

Matrices $Q_u(k+1|k)$ and $Q(k)$ are adapted in each step.

Control Matrix Correction

Generally, when one is faced with the problem of state estimation of mechanical systems, it is usual to have the matrix \bar{G} in:

$$\dot{x} = f(x, t) + \bar{G}w(t) \quad (38)$$

as:

$$\bar{G} = \begin{bmatrix} 0 & (n-m) \times m \\ I & m \times m \end{bmatrix}, \quad (39)$$

where 0 and I are the null and the identify matrices, respectively.

This means that the coefficients in \bar{G} just make the coupling between the state variables and noises.

In the "virtual control" context it is possible to define the control matrix G in Equation (13) in order to augment the coupling between estimated controls and state variables and thus accelerate the procedure convergence. In this paper, three types of corrections for the control matrix G are proposed. In what follows, a forced variable is defined as a state variable which is directly excited by noise in Equation (38). The other state variables are called nonforced variables.

Direct Correction. The simplest form to define a control matrix with the characteristic described before is to consider:

$$G = \begin{bmatrix} C_p & I_{n-m} \\ C_v & I_m \end{bmatrix}, \quad (40)$$

where C_p and C_v are positive coefficients, held fixed during the time propagation.

Despite the advantage given by the simple form, this correction technique has the drawbacks of using a trial and error method to choose C_p , C_v and of adding the corrections to the state variables with the same sign.

Restoration. The ideal of restoration was first used in numerical gradient-like methods to solve optimal control problems, where one has to deal with different requirements of searching a gradient step towards the extremal and satisfying the terminal constraints (Miele, 1975).

In the present case, restoration is used to modify C_p and C_v in Expression (40), thus giving priority of convergence to either the forced state variables or the nonforced state variables, depending on the system behavior. In order to make a decision to change C_p and C_v a convergence measure must be available. Estimated errors, e_p for the nonforced variables and e_v for the forced ones, and defined by:

$$e_p(k) = \frac{1}{(n-m)\sigma_r} \left[\sum_{i=1}^{n-m} P_{ii}(k|k) \right]^{1/2}, \quad (41)$$

$$e_v(k) = \frac{1}{m\sigma_v} \left[\sum_{i=n-m+1}^m P_{ii}(k|k) \right]^{1/2}, \quad (42)$$

can be used to indicate convergence. In Equations (41) and (42), P_{ii} are the elements on the diagonal of the $P(k|k)$ matrix and σ_r , σ_v represent the expected standard deviation of the errors between estimated and true variables after convergence of the problem. With these definitions, e_p , $e_v = 1$ mean convergence to the true trajectory. The technique can now be implemented through the following may: the procedure is initialized with the Direct Correction; after \bar{n} steps, a test is done. If $e_p > 2$ and $e_v > 2$, C_p and C_v are not changed, if $0 < e_p < 2$ and $e_v > 2$, only C_v is changed; if $0 < e_v < 2$ and $e_p > 2$, only C_p is changed and finally, if $0 < e_p < 2$ and $0 < e_v < 2$, both coefficients are changed. Propagation is started again with the new coefficients and after \bar{n} steps a new test is done.

With this technique, besides the simple form, coefficients

C_p and C_v are changed during the time interval of interest. However, the same disadvantages described for the Direct Correction are still present.

Automatic Correction. In the automatic correction technique, matrix G is generated at each step inside the algorithm, reinforcing the adaptive characteristic of the proposed procedure. Determination of the elements in G is done by solving, step by step, a deterministic linear optimal control problem. This control problem arises if one considers the estimation procedure as a controller that must drive the system from a given point in t_k to satisfy the observations in t_{k+1} , as mentioned before. Consider, for (t_k, t_{k+1}) , the propagation of the first order perturbation Δx :

$$\Delta x(t_{k+1}) = \Phi(k+1, k) \Delta x(t_k) + \gamma(k+1, k) u(t_k) \quad (43)$$

Assuming that the G matrix is constant in this interval:

$$\gamma(k+1, k) = \int_{t_k}^{t_{k+1}} \Phi(t_k, s) G(s) ds = \int_{t_k}^{t_{k+1}} \Phi(t_k, s) ds G_k \triangleq \bar{B}_k G_k \quad (44)$$

and using Equation (44) in (43):

$$\Delta x(t_{k+1}) = \Phi(k+1, k) \Delta x(t_k) + \bar{B}_k G_k u(t_k). \quad (45)$$

Consider now that all terms in the right-hand side of Equation (45) are known except the G_k matrix. In t_{k+1} , it is expected that $\Delta y(t_{k+1})$ and this corresponds to minimize a quadratic criterion given by:

$$J = [\Delta y(t_{k+1}) - H_{k+1} \Delta x(t_{k+1})]^T [\Delta y(t_{k+1}) - H_{k+1} \Delta x(t_{k+1})] \quad (46)$$

subject to the constraint given by Equation (45).

Defining G_k as:

$$G_k = \begin{bmatrix} G_1, k \\ G_2, k \end{bmatrix} = \begin{bmatrix} g_n & 0 \\ 0 & g_{n-m, m} \\ g_{1, n-m+1} & 0 \\ 0 & g_{n, m} \end{bmatrix} \quad (47)$$

and making:

$$\frac{\partial J}{\partial G_k} = 0 \quad (48)$$

there results a linear system to calculate the nonnull elements of the G_k matrix in the form:

$$D_{k+1} g_k = \Delta y(t_{k+1}) - H_{k+1} \phi(k+1, k) \Delta x(t_k), \quad (49)$$

where g_k is the $n \times 1$ vector of elements g_{ij} , as in Equation (47).

It is important to point out that the D_{k+1} matrix in Equation (49) is $(r \times n)$ dimensional and, thus, the linear system shall be solved by a deterministic least squares method or by an equivalent recursive method, if $r \neq n$.

The automatic correction technique has the distinguished feature of providing coefficients for G_k adjusted at each step, with variable signs and modules. However, an one-step lag has to be imposed to G_k since it was admitted that terms of Equation (45) were already known. This means that matrix G_k , which will be used in (t_k, t_{k+1}) is calculated with the values obtained in (t_{k+1}, t_k) . Another problem to mention is the fact that one can only use estimator propagated states in the technique and this represents only an approximation to the true states.

All the techniques to correct matrix G were applied to the orbit estimation problem. Results, as will be seen, were very satisfactory. As a final remark, it must be said that these techniques were developed for the general case and can be used with the proposed procedure in different problems (Fleury, 1985).

OBSERVATIONS

In the test problem of orbit estimation of a low altitude satellite, observations were generated combining the true orbit data with the location of 3 fictitious symmetric topocentric tracking stations, to get at each 1 second range and range-rate data, contaminated by white Gaussian noise with standard deviations:

$$\sigma_\rho = 10.0 \text{ m}; \quad \sigma_{\dot{\rho}} = 0,1 \text{ m/s.}$$

RESULTS

Results for the two cases, where different initial conditions are taken for the proposed estimator, are shown in this section. The virtual control $u(t)$ in Equation (13) is a 3×1 vector since it is possible to "control an orbit" using three independent forces, one for each axis. Parameters for analysis are true and estimated position errors, true and estimated velocity errors and range and range-rate normalized residues given, respectively by:

$$\Delta r(k) = \left\{ \sum_{i=1}^3 [x_i(k) - \hat{x}_i(k)]^2 \right\}^{1/2} \quad k = 1, 2, \dots \quad (50)$$

$$\Delta \hat{r}(k) = \left\{ \sum_{i=1}^3 P_{ii}(k|k) \right\}^{1/2} \quad k = 1, 2, \dots \quad (51)$$

$$\Delta v(k) = \left\{ \sum_{r=4}^6 [x_r(k) - \hat{x}_r(k)]^2 \right\}^{1/2} \quad k = 1, 2, \dots \quad (52)$$

$$\Delta \hat{v}(k) = \left\{ \sum_{i=4}^6 P_{ii}(k|k) \right\}^{1/2} \quad k = 1, 2, \dots \quad (53)$$

$$r_{\rho}(k) = \frac{1}{n_e \sigma_{\rho}} \left\{ \sum_{i=1}^{n_e} [y_{\rho}(k) - \hat{y}_{\rho}(k)]_i \right\} \quad k = 1, 2, \dots \quad (54)$$

$$\dot{r}_{\rho}(k) = \frac{1}{n_e \sigma_{\dot{\rho}}} \left\{ \sum_{i=1}^{n_e} [\dot{y}_{\rho}(k) - \hat{\dot{y}}_{\rho}(k)]_i \right\} \quad k = 1, 2, \dots \quad (55)$$

where y_{ρ} , \dot{y}_{ρ} are the actual (simulated) observations of range and range-rate respectively and \hat{y}_{ρ} , $\hat{\dot{y}}_{\rho}$ are the estimated observations of range and range-rate respectively. In the definitions (54) and (55) n_e means number of fictitious stations observing the satellite.

To guarantee convergence the true and estimated errors must converge to the same small values and the range and range-rate normalized residues must lie between the limits $\pm 3\sigma$.

NORMAL CASE

In this case, the estimator is initialized with errors of 80m in position and 0,6m/s in velocity. This corresponds to a situation where a good a priori estimate of the orbit is available as a result of preliminary orbit determination. Figures 1 and 2 show the results obtained with the use of the Direct Correction with coefficients $C_p = 4.0$, $C_v = 10.0$

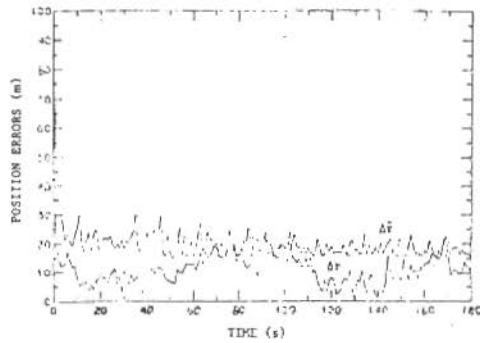


Figure 1. True and estimated position errors:
Normal case, Direct correction

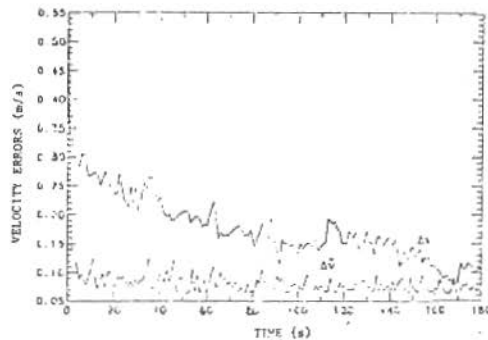


Figure 2. True and estimated velocity errors:
Normal case, Direct correction

Figures 3 and 4 show the application of the Restoration Correction to the same problem. Coefficients were chosen as one among the values 30.0, 1.0 and 0. for C_p ; and as one among the values 3.0, 4.0 and 1.0 for C_v .

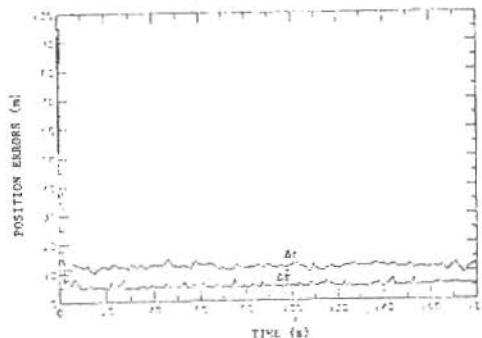


Figure 3. True and estimated position errors:
Normal case, Restoration

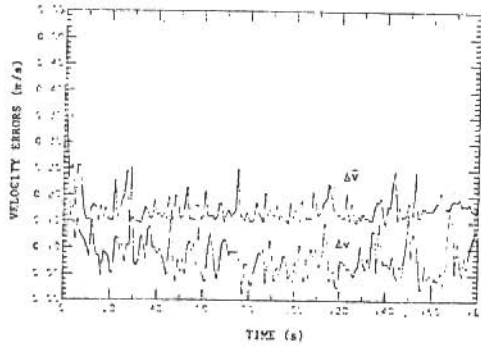


Figure 4. True and estimated velocity errors:
Normal case, restoration

Finally, Figures 5 and 6 present the results obtained with the use of the Automatic Correction.

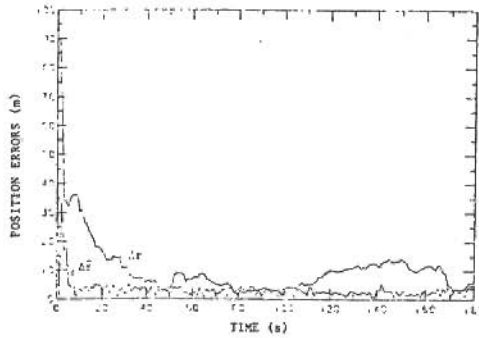


Figure 5. True and estimation position errors:
Normal case, automatic correction

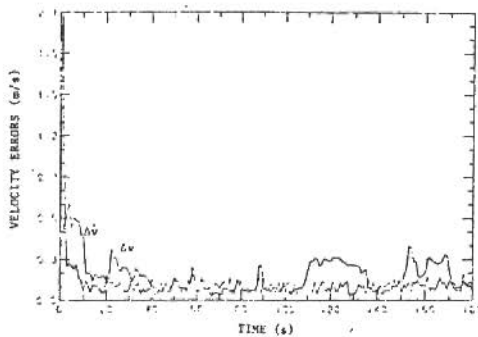


Figure 6. True and estimated velocity errors:
Normal case, restoration

The analysis of Figures 2 to 7 shows that one can attain good results using any of the correction techniques, since convergence to values close to the true orbit occurs for the three methods. Best performance is achieved using Automatic Correction, as expected, although this is the most time consuming of the proposed techniques (Fleury, 1985). Another important feature which can be extracted comparing Figures 1 and 2 with Figures 3 and 4 is that an improvement in position convergence is obtained when Restoration is used instead of Direct Correction.

CRITICAL CASE

In this case, the estimator is initialized with errors of 1000m in position and 10 m/s in velocity. This corresponds to a situation where the initial error are larger than the usual initial errors obtained in the preliminary orbit determination.

The results obtained with the Direct Correction technique are shown in Figures 7, 8, 9 and 10.

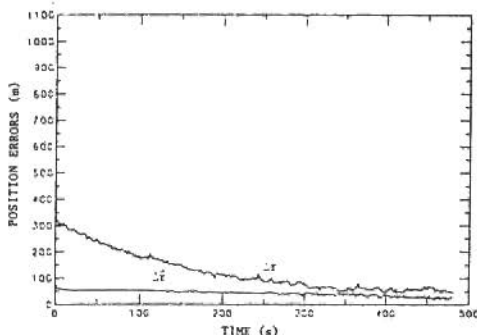


Figure 7. True and estimation position errors: Critical case, direct correction

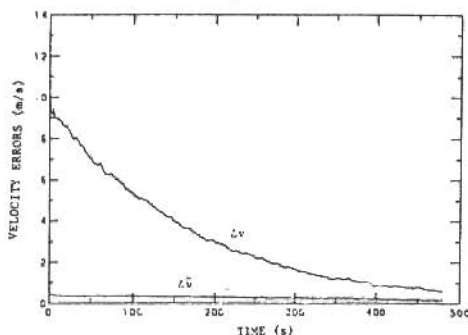


Figure 8. True and estimated velocity errors: Critical case, direct correction

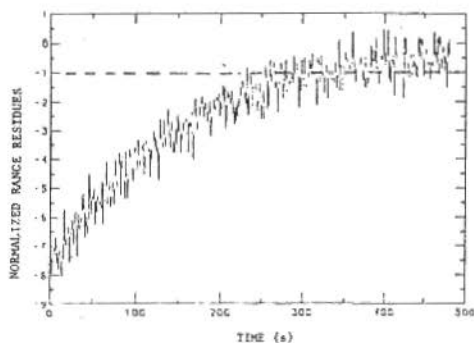


Figure 9. Range normalized residues:
Critical case, direct correction

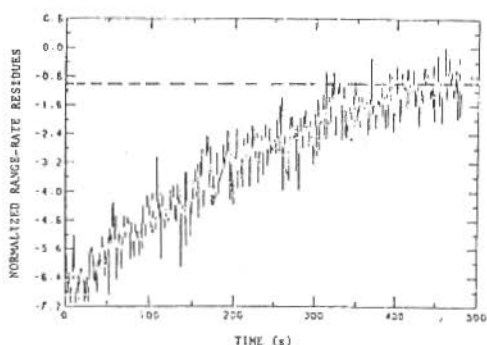


Figure 10. Range-rate normalized residues:
Critical case, direct correction

Figures 11, 12, 13 and 14 presents the results accomplished with the proposed procedure aided by the Automatic Correction technique.

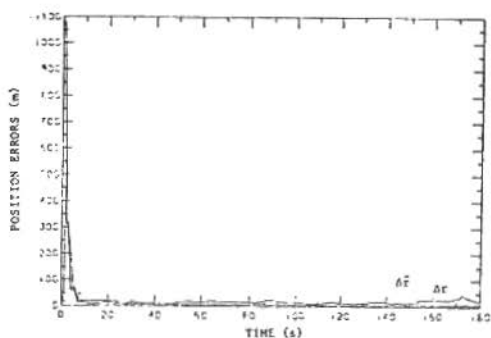


Figure 11. True and estimated position errors:
Critical case, automatic correction

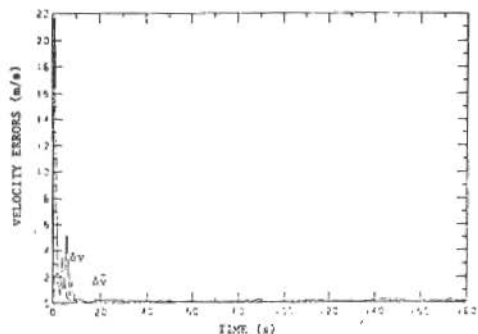


Figure 12. True and estimated velocity errors:
Critical case, automatic correction

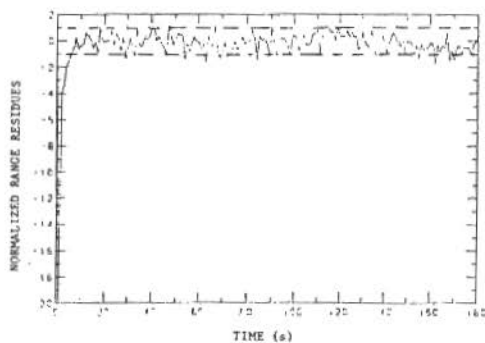


Figure 13. Range normalized residues:
Critical case, automatic correction

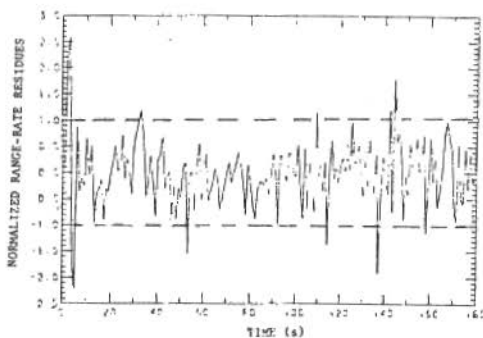


Figure 14. Range-rate normalized residues:
Critical case, automatic correction

The analysis of Figures 7 to 14 shows very good results for the Automatic Correction and a satisfactory convergence with the use of the Direct Correction. In this last case, the 1000m and 10 m/s initial errors are reduced to approximately 30 m and 0,4m/s. Although the final errors are reasonably small, the time interval for convergence is quite long. Notice also that the estimated errors in the transient phase of convergence are very optimistic when compared to the real ones. These characteristics show that there are still further efforts to be done in this Direct Correction technique to reach a better performance.

On the other hand, results obtained with the Automatic Correction scheme are comparable to results presented by other techniques usually employed in nonlinear system estimation such as the Extended Kalman Filter aided by Model Compensation Technique (e.g. Kuga, 1982). Convergence is very fast and normalized residues are maintained between the limites of $\pm 1\sigma$ during all time propagation.

CONCLUSIONS

A new approach to state estimation of nonlinear systems has been presented, where the concept of duality was explored to transform the estimation problem into one of determining a virtual control. As indicated by the test done in this exploratory phase, it is expected to be a valid alternative for the case where a great lack of knowledge in the dynamics exists, but a good level of information is locally provided by the observations.

In this work, one only started to explore the possibilities opened by the virtual nature of control in Equation (13). Presently, efforts are being done to further explore these possibilities, mostly in terms of the structure of the virtual control $u(t)$ and of the structure of $G(t)$, seeking for other forms of automatic and adaptive generation of this matrix. Besides this, one is also paying efforts in trying to tune the filter with constant levels of noise to avoid extra efforts imposed by adaptive estimation of these noises.

REFERENCES

- [1] Cruz, J.J.; Rios Neto, A. Estimating the state of ships in the presence of unmodelled accelerations. American Towing Tank Conference, 19, Ann Arbor, Michigan, USA, 1980.
- [2] Fleury, A.T. Estimadores de estado de sistemas dinâmicos baseados no conceito de dualidade (Dynamical systems state estimators based upon the duality concept). Doctoral thesis, Escola Politécnica da Universidade de São Paulo, 1985.
- [3] Gelb, A.; Kasper Jr. J.F.; Nash Jr., R.A.; Price, C.F., Sutherland Jr., A.A. Applied Optimal Estimation. Cambridge, MA, MIT, 1974.
- [4] Jazwinski, A.H. Adaptive filtering. Automatica, 5: 475-485, 1969.
- [5] Jazwinski, A.H. Stochastic processes and filtering theory. New York, Academic, 1970 (Mathematics in Science and Engineering, 64).
- [6] Kuga, H.K. Estimaco adaptativa de rbita aplicada a satlites de baixa altitude (Adaptive orbit estimation applied to low altitude satellites) MSc Thesis (Space Sciences). So Jos dos Campos, INPE, Feb. 1982 (INPE-2316-TDL/079).
- [7] Liebelt, P.B. An introduction to optimal estimation. Menlo Park, CA, Addison Wesley, 1967.
- [8] Miele, A. Recent advances in gradient algorithms for optimal control problems. J. Optimization Theory and Applications, 17(5/6): 361-430, 1975.
- [9] Maybeck, P.S. Stochastic models, estimation and control, v.1, New York, Academic, 1979 (Mathematics in Science and Engineering, 141-1).
- [10] Rios Neto, A.; Cruz, J.J. A Stochastic rudder control law for ship path-following autopilots. Automatica, 21(4): 371-384, 1985.
- [11] Rios Neto, A.; Fleury, A.T. Utilizao das nooes de dualidade no desenvolvimento de um estimador de estado (Use of the duality concept in the development of a state estimator). So Jos dos Campos, INPE, Sept. 1984 (INPE-3288-RPI/110).
- [12] Rios Neto, A.; Kuga, H.K. Estimaco adaptativa e em tempo real de rbita de satlites artificiais a baixa altitude (Real-time adaptive orbit estimator of artificial satellites at

- low altitude). In: Sixth Brazilian Congress of Mechanical Engineering, Rio de Janeiro, 1981. Proceedings. ABCM, 1981, v. B., p. 405-414.
- [13] Rios Neto, A.; Kuga, H.K. Estimaco adaptativa de rudo no estado para o Filtro de Kalman (State noise adaptive estimation for Kalman Filter). In: Fourth Brazilian Congress of Automatica, Campinas, 1982. Proceedings. SBA, 1982, v.1, p. 101-105.
- [14] Rios Neto, A.; Kuga, H.K. Kalman filtering state noise adaptive estimation. Presented at the 2nd. IASTED International Conference in Telecommunication and Control, Rio de Janeiro, 1985.
- [15] Tapley, B.D.; Ingram, D.S. Orbit determination in the presence of unmodelled accelerations. IEEE Transactions Automatic Control, AC-18(4): 369-373, Aug. 1973.

APPENDIX

Adaptive State Noise Estimation: Pseudoobservation Equation

As seen in the text, to estimate the diagonal elements of $Q(k)$, one imposes consistency between the observation residues and their statistics (Jazwinski, 1969). This corresponds to impose that the diagonal elements of $Q(k)$ must assume values that maximize the probability of occurrence of the true residue of each observation (Rios Neto and Kuga, 1982). To understand the resulting procedure, particularly the pseudoobservation equation, the key steps are reproduced in this Appendix.

The observation residue $r(k+1|k)$ is defined as in Equation (31) by:

$$r(k+1|k) = y(t_{k+1}) - \bar{y}(t_{k+1}) = y(t_{k+1}) - H_{(k+1)} \bar{x}(t_{k+1}) \quad (A.1)$$

and the true residue, (Equation 32), by:

$$r_v(k+1|k) = r(k+1|k) - v_{k+1}. \quad (A.2)$$

Under the hypothesis of having normal probability distributions the criterion of statistical consistence is realized by imposing:

$$(r_v(k+1|k))_0^2 = E [r_v^2(k+1|k)]. \quad (\text{A.3})$$

where the subindex 0 is to mean to occurred value of $r_v(k+1|k)$. Equation (A.3) is the condition to maximize the probability of occurrence of the value $(r_v(k+1|k))_0$. Using Equation (A.2) and Equation (8) of Section 2 in both sides of Equation (A.3), there results after some algebraic manipulations:

$$\begin{aligned} & (r^2(k+1|k) - 2 r(k+1|k) v(t_{k+1}) + v^2(t_{k+1}))_0 = \\ & = H_{k+1} \phi(t_{k+1}, t_k) P(k|k) \phi^T(t_{k+1}, t_k) H_{k+1}^T + H_{k+1} \bar{Y} R \bar{Y}^T H_{k+1}^T \end{aligned} \quad (\text{A.4})$$

Define a noise v'_{k+1} as:

$$v'_{k+1} = -2(r(k+1|k))_C v(t_{k+1}) + v^2(t_{k+1}) - R_{k+1} \quad (\text{A.5})$$

Then

$$E[v'_{k+1}] = 0, \quad (\text{A.6})$$

$$E[(v'_{k+1})^2] = 4(r^2(k+1|k))_C R_{k+1} + 2R_{k+1}^2, \quad (\text{A.7})$$

where $(r(k+1|k))_C$ is the calculated value, corresponding to the occurred value of the random variable $r(k+1|k)$.

With this definition and using the fact that $Q(k)$ is a diagonal matrix, it is reasonable to adopt, from Equation (A4), the Pseudoobservation equation in the form:

$$y'(t_{k+1}) = H'(k+1) q_k + v'_{k+1}, \quad (\text{A.8})$$

where

$$\begin{aligned} y'(t_{k+1}) &= (r^2(k+1|k))_C + R(k+1) - \\ &- H_{k+1} \phi(t_{k+1}, t_k) P(k|k) \cdot \phi^T(t_{k+1}, t_k) H_{k+1}^T, \end{aligned} \quad (\text{A.9})$$

$$H'(k+1) = [(\sum_{j=1}^n H_{ij} \cdot \bar{y}_{j1})^2; (\sum_{j=1}^n H_{ij} \bar{y}_{j2})^2; \dots; (\sum_{j=1}^n H_{ij} \bar{y}_{jm})^2], \quad (\text{A.10})$$

and $q_k = \{q_{k1}, \dots, q_{km}\}$ is the $m \times 1$ vector which contains the parameters to be estimated.

To estimate the vector q_k , the following algorithm is implemented:

- Prediction or propagation phase

$$\hat{q}(k+1|k) = \hat{q}(k|k), \quad (\text{A.11})$$

$$P^q(k+1|k) = P^q(k|k) + W^q(k|k), \quad (\text{A.12})$$

where $W^q(k|k)$ is a diagonal matrix of small constants which is added to $P^q(k|k)$ to guarantee that $P^q(k+1|k)$ is not smaller than the computer numerical zero.

- Filtering phase: follows the same steps of the linear Kalman Filter (e.g. Kuga, 1982; Maybeck, 1979).

UMA FORMULAÇÃO CONSISTENTE DO MÉTODO DOS ELEMENTOS FINITOS PARA RESOLUÇÃO DE PROBLEMAS DE TRANSPORTE DIFUSIVO-CONVECTIVO

Eduardo Gomes do Carmo - Membro da ABCM
COPPE/UF RJ

Augusto Cesar Galeão - Membro da ABCM
INCC/CNPq

RESUMO

Um novo método, particularmente apropriado para a resolução de problemas de transporte difusivo-convectivo, com forte preponderância dos efeitos de convecção, é desenvolvido. Mostra-se que esse método enquadra-se na classe de aproximações do tipo Petrov-Galerkin. Através de alguns exemplos comprova-se a sua potencialidade na resolução de camadas limites internas e/ou externas.

ABSTRACT

A new method specially designed to solve highly convective transport problems is proposed. Using a variational approach it is shown that this weighted residual method belongs to a class of Petrov-Galerkin's approximation. Some examples are presented in order to demonstrate the adequacy of this method in predicting internal or external boundary layers.

INTRODUÇÃO

A resolução numérica das equações de transporte, e das equações de Navier-Stokes, tem sido objeto, nos últimos anos, de um grande número de trabalhos publicados. Tal interesse decorre, em grande parte, das dificuldades numéricas associadas à solução de escoamentos com altos Reynolds, ou fenômenos de transporte governados por elevados números de Peclet. Em tais situações, caracterizadas pela preponderância dos termos convectivos (acentuados gradientes) sobre os difusivos (operador de Laplace), ocorrem frequentemente problemas de choque e/ou problemas típicos de camada limite.

Em termos de soluções numéricas baseadas na técnica de diferenças finitas, as características de estabilidade e maior ordem de aproximação do operador de diferença centrada, quando utilizado para a discretização de operadores elípticos, explicam o seu uso na aproximação dos termos difusivos. Por outro lado, como ressaltado em [1], o uso de diferença centrada para o tratamento da parte convectiva, deve-se possivelmente, a uma errônea noção de consistência, baseada no erro de discretização numa expansão em série de Taylor. Assim, se em termos do erro de discretização, o operador de diferença central é de 2ª ordem, em termos do erro de truncamento global, que obviamente é o que importa do ponto de vista de soluções aproximadas de equações diferenciais, esse operador conduz a uma aproximação de "3ª ordem" para as derivadas segundas (termos difusivos), e de 2ª ordem para as derivadas primeiras (termos convectivos). Logo, esse não é um procedimento consistente. Pior do que isso, a diferença centrada não apresenta as características de estabilidade numérica, quando aplicada a derivadas de ordem ímpar. Decorre daí, que em problemas predominantemente convectivos, o uso desse procedimento conduz a uma resposta oscilatória, que se propaga ao longo dos pontos nodais do domínio. Por isso, nem mesmo o recurso da utilização de malhas extremamente refinadas, na região de elevados gradientes, consegue, normalmente, fazer desaparecer essas oscilações espúrias.

Os primeiros intentos, no sentido de contornar essa dificuldade, foram feitos em problemas uni-dimensionais, empregando-se diferenças não-centradas (1ª ordem) para o termo convectivo. Do ponto de vista físico, essa aproximação era mais condizente, visto que as informações carregadas pelo escoamento eram mais fortemente ponderadas à montante. Do ponto de vista matemático, a maior estabilidade

desse operador, na discretização de derivadas primeiras, era responsável pela eliminação das oscilações [1]. No entanto, do ponto de vista numérico, esse procedimento era equivalente à adição de uma difusividade artificial [1,2].

Se se interpreta que o esquema de diferença central é que é, na realidade, o responsável pela introdução de uma difusividade artificial negativa, justifica-se entender que o emprego de diferença não-centrada corrige esse problema pela adição de uma difusividade numérica. Surgem então os chamados esquemas híbridos, (combinação ponderada do operador de diferença centrada para os termos difusivos e "one-sided upstream difference" para os termos convectivos), e os esquemas ditos de "upwind" ótimo, em que o coeficiente de difusividade artificial era calculado de forma a fornecer, (para um problema relativamente simples: modelo uni-dimensional de transporte em regime e sem termo de fonte), aproximações nodalmente exatas [3]. No entanto, como mostrado em [1], o uso deste procedimento para a abordagem de problemas de transporte predominantemente convectivos, transientes e/ou contendo termos de fonte, conduzia a resultados excessivamente difusivos; ou ainda, a sua generalização a domínios bi-dimensionais produziam muitas vezes piores resultados que aqueles obtidos empregando-se diferenças centradas, devido à presença de uma marcante difusividade artificial, perpendicular à direção do escoamento.

Esse aspecto, exaustivamente discutido em [1] e [4], decorre pura e simplesmente do fato de que se por um lado o operador de diferença não-centrada de 1ª ordem aumenta a estabilidade do esquema numérico, por outro lado, ele pode comprometer sensivelmente a precisão da solução.

De novo, analisando-se do ponto de vista do erro de truncamento global esse algoritmo é inconsistente, e neste sentido aquele autor propõe a utilização de um esquema de diferença não-centrada de 3ª ordem, para a aproximação dos termos convectivos, combinado com diferença centrada, para os termos difusivos, com o que retém-se globalmente as características de consistência, estabilidade e precisão, e uma relativa simplicidade para a sua generalização a domínios bi ou tri-dimensionais [4]. No entanto, a despeito dessas características, as soluções obtidas com o emprego desse algoritmo apresentam oscilações localizadas na região próxima à formação de camada limite.

Também a construção de modelos de elementos finitos, para a resolução numérica de problemas de escoamento e/ou transporte, teve (ou está tendo) uma evolução semelhante àquela verificada com o método de diferenças finitas. Assim, as primeiras aplicações tiveram por base o método de Galerkin, possivelmente pelo uso generalizado dessa formulação na resolução de problemas estruturais, para os quais, via de regra, estimativas de convergência ótima [5], decorrentes da V-elipticidade dos operadores normalmente envolvidos nesse tipo de problemas já eram então bem conhecidas.

Ocorre que, nos problemas de escoamento e/ou transporte, a presença do termo convectivo destrói a simetria do operador, e o método de Galerkin não apresenta mais a propriedade de melhor aproximação. Além disso, as equações discretas, resultantes da utilização desse método, apresentavam a mesma estrutura daquelas obtidas via diferenças centradas. Em outras palavras, em problemas predominantemente convectivos, as soluções de Galerkin apresentavam as mesmas oscilações nodais, àquela altura já bem conhecidas por aqueles que empregavam diferenças finitas.

Ora, se operadores de diferenças não-centradas tinham características estabilizadoras quando usados na discretização do termo convectivo, algo semelhante deveria ser procurado no contexto de elementos finitos. Era preciso, portanto, adaptar à formulação variacional empregada, um adequado método de resíduos ponderados, de forma que as equações discretas, assim geradas, reproduzissem o modelo físico do problema. Modificar o espaço de funções peso era a solução natural [6].

A primeira aplicação deste procedimento, o chamado método de Petrov-Galerkin, a uma equação de transporte em regime, num domínio unidimensional, é feita em [7]. Aí são utilizadas funções peso contínuas, construídas adicionando-se às funções de interpolação clássicas, funções que ponderavam mais fortemente as informações à montante do escoamento. Assim como nos procedimentos de diferenças finitas, aqui também o parâmetro de "upsind" é escolhido de forma a eliminar as oscilações ou, em certos casos, fornecer soluções nodalmente exatas.

Generalizações deste procedimento para problemas bi-dimensionais são posteriormente apresentadas em [8] e [9]. Ocorre que naqueles trabalhos a matriz de difusividade era também afetada pela pon

deração, o que não acontecia nos problemas unidimensionais, e o que de certa forma contrariava a experiência adquirida com diferenças finitas, onde apenas as derivadas convectivas tinham a sua aproximação modificada.

Volta-se de novo ao método de Galerkin. Alguns autores propõem um esquema de integração modificada para a obtenção da matriz convectiva [10,11]. Outros adotam a idéia da difusividade artificial segundo a direção do escoamento, criando a terminologia "dissipação de balanceamento anisotrópico" [12].

Assim como já havia acontecido com alguns modelos de diferenças finitas do tipo "upwind", também com os modelos clássicos de elementos finitos, do tipo "upwind", uma excessiva difusão ortogonal às linhas de corrente estava presente. Se esse inconveniente podia ser remediado pela consideração da difusividade artificial, agindo tão somente segundo a direção do escoamento, o mesmo não acontecia quando termos transientes e de fonte ocorriam na equação do problema. Alguns exemplos numéricos ressaltando esses aspectos são apresentados em [2], onde é mostrada a superioridade do modelo "SUPG-Streamline Upwind Petrov-Galerkin" [13] na resolução desses problemas.

Do ponto de vista matemático esse último método consiste basicamente numa formulação do tipo Petrov-Galerkin, onde as funções peso, que são descontínuas, são construídas adicionando-se às funções clássicas do método de Galerkin, uma perturbação, definida pelos autores como "streamline upwind", que atua somente na direção das linhas de corrente. Essas funções peso modificadas incidem sobre todos os termos da equação, de forma que a condição de ortogonalidade do resíduo em relação às mesmas é satisfeita, resultando, portanto, numa formulação consistente de resíduos ponderados. Em última instância, esse modelo é uma generalização de [7] a domínios multi-dimensionais.

Conforme mostrado em [14] o método SUPG apresenta elevada ordem de precisão e boas propriedades de estabilidade. Assim, se as soluções exatas são suficientemente regulares, estimativas globais de erro "ótimas" ou "quase-ótimas" são obtidas. Para soluções exatas típicas de camada limite, estimativas locais de erro ótimas ou quase-ótimas são previstas fora de uma pequena vizinhança, envolvendo a camada limite. Em outras palavras, para esses problemas as so

luções do método SUPG apresentam oscilações localizadas numa região próxima à de desenvolvimento da camada limite. Claro, portanto, que se a formação da camada limite decorre de uma descontinuidade causada pela condição de contorno prescrita, uma possível solução é refinar a malha nessa região. Se a descontinuidade é interna, uma estratégia é combinar esse modelo com um esquema auto-adaptativo de refinamento de malha.

Uma outra estratégia é refinar o método, de forma que o modelo tenha a capacidade de absorver possíveis descontinuidades. Um grande avanço neste sentido é apresentado em [15], onde seus autores mostram que nem sempre a linha de corrente é a direção apropriada de "upwind". Como, na construção do método "Streamline upwind", parte-se da idéia da difusividade artificial negativa causada pelo método de Galerkin, a perturbação adicionada à função peso procura gerar uma difusividade positiva, na direção do escoamento. Isso é seguido via um fluxo convectivo segundo as linhas de corrente. Claro que qualquer campo ortogonal ao gradiente da quantidade transportada, gera um fluxo convectivo nulo, o qual, obviamente, não altera a difusividade artificial buscada no método "Streamline upwind". Isso em essência traduz a idéia anterior desenvolvida em [15]. Partindo dessa idéia, aqueles autores utilizam um princípio de máximo, satisfazendo o método de Petrov-Galerkin, e desenvolvem um elemento triangular apropriado à resolução desses problemas.

Esse modelo gera um esquema iterativo, de forma, que pode ser visto como um modelo internamente auto-adaptativo. Os resultados obtidos com o mesmo são bastante superiores àqueles fornecidos com o SUPG, principalmente no que diz respeito à grande precisão na determinação de camadas limite. O seu grande inconveniente é a extrema dificuldade a generalizações.

Mais recentemente um novo modelo também do tipo Petrov-Galerkin consistente, que adiciona à função peso do método SUPG um termo visando a captura de descontinuidades, foi proposto em [16]. De certa forma ele combina as características do "Streamline upwind" com a observação básica ressaltada em [15]. Os resultados apresentados em [16] mostram uma melhor performance deste modelo, em relação ao SUPG original, na representação de camadas limite internas e/ou externas, muito embora, em problemas regulares, note-se um pequeno decréscimo na precisão global.

O modelo que será apresentado neste trabalho baseia-se em dois aspectos básicos:

- i) o primeiro, já exaustivamente discutido nesta introdução, é a inadequação do método de Galerkin na obtenção de soluções aproximadas de problemas de transporte predominantemente convectivos. Isso nos leva à adoção do método de Petrov-Galerkin pelas vantagens também já discutidas anteriormente;
- ii) o segundo, relaciona-se com a construção do espaço de funções teste. Neste sentido duas observações ainda não explicitadas são fundamentais. Se o problema é puramente convectivo o fluxo é carregado segundo as linhas de corrente. Essa é a solução exata do problema. Porém, em termos de solução aproximada, a direção de transporte é também aproximada. Logo as funções teste devem ser capazes de representar esta condição. Ou seja, quanto mais próxima a solução aproximada se encontrar da solução exata, tanto mais próxima deve estar a direção aproximada de fluxo, da direção física de transporte, isto é: das linhas de corrente.

Com isso o modelo procura se auto-adaptar, no sentido de fornecer a melhor aproximação.

A primeira parte desse trabalho é dedicada à formulação variacional do problema e à sua forma discretizada via Petrov-Galerkin. A seguir discutem-se rapidamente os métodos SUPG e SUPG + operador de captura.

Na segunda parte mostra-se a construção das funções peso especialmente apropriadas para a solução do problema e apresenta-se a montagem do modelo discreto. Posteriormente ressaltam-se algumas características do modelo proposto e mostra-se que o modelo SUPG + operador de captura é um caso particular do mesmo.

Finalmente alguns exemplos são analisados, e os resultados numéricos obtidos com o modelo desenvolvido neste trabalho são comparados com outros existentes.

FORMULAÇÃO DO PROBLEMA

Neste trabalho estamos interessados na resolução do problema de transporte de uma grandeza escalar num meio fluido, cujo modelo matemático é descrito pelo problema de valor de contorno com condi-

ção inicial definido em termos:

i) da equação de difusão-convecção

$$\dot{\phi} + u \cdot \nabla \phi + \operatorname{div}(-K\nabla\phi) = f(x,t) \quad \text{em } \Omega \quad (1a)$$

ii) das condições de contorno

$$\begin{aligned} \phi(x,t) = g(x,t) & ; \quad x \in \Gamma_g & \Gamma_g \cup \Gamma_q = \Gamma \\ -K\nabla\phi \cdot n = q(x,t) & ; \quad x \in \Gamma_q & \Gamma_g \cap \Gamma_q = \text{vazio} \end{aligned} \quad (1b-c)$$

iii) da condição inicial

$$\phi(x,0) = \phi_0(x) \quad ; \quad x \in \Omega \quad (1d)$$

Nas expressões anteriores, $\Omega \subset \mathbb{R}^n$ ($1 \leq n \leq 3$) é uma região finita de fronteira suave Γ e normal exterior n . O campo escalar $\phi(x,t)$ fornece, para cada $x \in \Omega$ e para cada tempo $t \in [0,t)$, o valor da grandeza imersa no meio fluido, cujo escoamento é definido pelo campo de velocidade $u(x,t)$, suposto conhecido. As funções f , g , q e ϕ_0 são dados do problema; $K(x)$ é o tensor difusividade; $\nabla\phi$ denota o gradiente de ϕ ; $\operatorname{div}(\cdot)$ é o divergente; (\cdot) produto escalar e $(\cdot) = \partial/\partial t$.

Se designamos por:

a) $L_2(\Omega)$, ao espaço das funções quadrado-integráveis, com produto interno

$$\langle \psi, \theta \rangle = \int_{\Omega} \psi \theta \, d\Omega \quad ,$$

e norma

$$\| \psi \|_0 = \langle \psi, \psi \rangle^{1/2} < + \infty \quad ;$$

b) $H_1(\Omega) \equiv \{ \psi ; \psi \in L_2(\Omega) ; (\nabla\psi_i) \in L_2(\Omega) ; i = 1, 2, 3 \}$ ao espaço de Sobolev cujos elementos têm norma

$$\| \psi \|_1 = [\langle \psi, \psi \rangle + \langle \nabla\psi, \nabla\psi \rangle]^{1/2} \quad ;$$

e se introduzimos:

c) o conjunto S de funções teste (cinematicamente admissíveis)

$$S \equiv \{\psi(x,t); p/ \text{ cada } t \in [0,T]; \psi \in H_1(\Omega); \psi|_{\Gamma} = g\}$$

d) o espaço V das variações admissíveis

$$V \equiv \{\hat{\psi}; \hat{\psi} \in H_1(\Omega); \hat{\psi}|_{\Gamma} = 0\} ;$$

então:

Problema (P1): A solução do problema proposto é a função $\phi \in S$, tal que, para cada tempo $t \in [0,T]$ e para toda $\hat{\phi} \in V$, satisfaz a relação

$$\langle \dot{\hat{\phi}} + u \cdot \nabla \phi, \hat{\phi} \rangle + \langle K \nabla \phi, \nabla \hat{\phi} \rangle - \langle f, \hat{\phi} \rangle - \langle \langle q, \hat{\phi} \rangle \rangle = 0 ; \quad (2a)$$

onde:

$$\langle \langle q, \hat{\phi} \rangle \rangle = \int_{\Gamma_q} q \hat{\phi} \, d\Gamma , \quad (2b)$$

e no instante inicial:

$$\langle \phi - \phi_0, \hat{\phi} \rangle = 0 \quad (2c)$$

SOLUÇÕES APROXIMADAS

Seja agora τ_h uma partição de elementos finitos definida em Ω , construída através da união de elementos e , que não se sobrepõem.

Diremos que ϕ^h é uma aproximação de Galerkin do problema anterior, se para cada tempo $t \in [0,T]$, e para toda $\hat{\phi}^h \in V^h$, essa aproximação satisfaz às relações:

$$\begin{aligned} \text{i)} & \quad \langle \dot{\hat{\phi}}^h + u \cdot \nabla \phi^h, \hat{\phi}^h \rangle + \langle K \nabla \phi^h, \nabla \hat{\phi}^h \rangle - \langle f, \hat{\phi}^h \rangle - \langle \langle q, \hat{\phi}^h \rangle \rangle = 0 \\ \text{ii)} & \quad \phi^h - \tilde{\phi}^h \in V^h ; \quad \forall \tilde{\phi}^h \in S^h \end{aligned} \quad (3a-b)$$

onde:

$$\begin{aligned} V^h & \equiv \left\{ \hat{\phi}^h \in C^0(\Omega); \phi^h|_e \in P^k, \forall e \in \tau_h; \hat{\phi}^h|_{\Gamma_g} = 0 \right\} \\ S^h & \equiv \left\{ \phi^h \in C^0(\Omega); \phi^h|_e \in P^k, \forall e \in \tau_h; \phi^h|_{\Gamma_g} = g \right\} , \end{aligned} \quad (3c-d)$$

sendo P^k o espaço dos polinômios de grau menor ou no máximo igual a k . Além disso, no instante inicial:

$$\langle \phi^h - \phi_0, \hat{\phi}^h \rangle = 0 \quad (3e)$$

Em resumo: no método de Galerkin, as mesmas funções que interpolam localmente, em cada elemento finito, a solução aproximada ϕ^h , são usadas, também, para a construção do espaço de funções peso V^h . Em outras palavras, se o problema envolve condições de contorno homogêneas, (do tipo Dirichlet), os espaços V^h e S^h coincidem. Já no método de Petrov-Galerkin isso não ocorre, ou seja o espaço de funções peso não coincide com o espaço de funções teste.

Na próxima seção apresentamos, resumidamente, o método desenvolvido na referência [16], conhecido como SUPG + operador de captura; uma variante do método "SUPG-Streamline Upwind Petrov-Galerkin" [13].

O MÉTODO SUPG + OPERADOR DE CAPTURA

Neste método [13,16] o espaço de funções peso é constituído por funções da forma

$$w^h = \hat{\phi}^h + p, \quad (4a)$$

onde $\hat{\phi}^h \in V^h$, e p é uma função descontínua definida como:

$$p = \alpha u \cdot \nabla \hat{\phi}^h + \beta \frac{u \cdot \nabla \phi^h}{|\nabla \phi^h|^2} \nabla \phi^h \cdot \nabla \hat{\phi}^h. \quad (4b)$$

α e β são conhecidas como funções de "upwind" (ver [7,9,16]), e para $\beta \equiv 0$ tem-se o método SUPG.

A solução aproximada $\phi^h \in S^h$ é então determinada de forma a satisfazer a condição:

$$\begin{aligned} & \langle \hat{\phi}^h + u \cdot \nabla \phi^h, \hat{\phi}^h \rangle + \langle K \nabla \phi^h, \nabla \hat{\phi}^h \rangle - \langle f, \hat{\phi}^h \rangle - \langle \langle q, \hat{\phi}^h \rangle \rangle + \\ & + \sum_e \int_{\Omega_e} \left[\hat{\phi}^h + u \cdot \nabla \phi^h + \text{div}(-K \nabla \phi^h) - f \right] p \, d\Omega = 0 \quad \forall \hat{\phi}^h \in V^h. \end{aligned} \quad (5)$$

Como já foi mencionado na introdução, e de acordo com os resultados apresentados nas referências [13,16], este método quando

aplicado a problemas predominantemente convectivos apresenta um desempenho bastante superior ao do método de Galerkin. No entanto, como será mostrado em alguns exemplos a serem apresentados no final desse trabalho, na presença de termos transientes ou termos de fonte, e desde que o problema envolva a formação de camada limite, mesmo com o uso desse método podem ser observadas oscilações de grande intensidade, numa região próxima à da formação da camada limite.

Na próxima seção desenvolveremos um novo método que elimina substancialmente a ocorrência dessas oscilações.

O MÉTODO PROPOSTO

Suponhamos que de alguma forma uma solução aproximada ϕ^h , do problema variacional

$$\langle \dot{\phi} + u \cdot \nabla \phi, \hat{\phi} \rangle + \langle K \nabla \phi, \nabla \hat{\phi} \rangle - \langle f, \hat{\phi} \rangle - \langle \langle h, \hat{\phi} \rangle \rangle = 0 \quad \forall \hat{\phi} \in V \quad (6)$$

tenha sido obtida. O seguinte problema é então proposto:

Problema (P2): determinar o campo v^h , tal que:

$$i) \quad v^h \cdot \nabla \phi^h + F(\phi^h) = 0 \quad \text{em } \Omega_e; \quad F(\phi^h) = \dot{\phi}^h + \text{div}(-K \nabla \phi^h) - f \quad (7a)$$

$$ii) \quad \|v^h - u\|_0^2 = \int_{\Omega_e} (v^h - u) \cdot (v^h - u) d\Omega \quad \text{seja mínimo.} \quad (7b)$$

Minimizar (7b) com a restrição (7a) é equivalente a determinar o par (v^h, λ) que torna estacionário o funcional:

$$I[v^h, \lambda] = \int_{\Omega_e} \left[\frac{1}{2} (v^h - u) \cdot (v^h - u) + \lambda (v^h \cdot \nabla \phi^h + F(\phi^h)) \right] d\Omega \quad (8)$$

Estamos assumindo que no sub-domínio $\Omega_e \subset \Omega$ as funções $v^h, \nabla \phi^h \in F(\phi^h)$ são suficientemente regulares de forma que $I[v^h, \lambda]$ tenha sentido. Da condição $\delta I = 0$ obtemos:

$$iii) \quad \text{o multiplicador } \lambda = - \frac{(v^h - u) \cdot \nabla \phi^h}{|\nabla \phi^h|^2}; \quad |\nabla \phi^h| = (\nabla \phi^h \cdot \nabla \phi^h)^{1/2} \neq 0 \quad (9a)$$

$$iv) \quad \text{o campo } v^h = u - \frac{u \cdot \nabla \phi^h + F(\phi^h)}{|\nabla \phi^h|^2} \nabla \phi^h \quad (9b)$$

A solução anterior nos diz que para cada ϕ^h o valor mínimo é

alcançado para:

$$\|v^h - u\|_0^2 = \sum_e \int_{\Omega_e} \frac{\varepsilon^2}{|\nabla\phi^h|^2} d\Omega \quad (10a)$$

onde,

$$\varepsilon = u \cdot \nabla\phi^h + F(\phi^h) \quad (10b)$$

é o resíduo correspondente à solução aproximada ϕ^h , em Ω_e . Em outras palavras: em cada sub-região Ω_e , o campo de velocidades "real", da solução aproximada ϕ^h , e que se encontra mais próximo, (proximidade no sentido da norma em $L_2(\Omega_e)$), do campo de velocidade real da solução exata ϕ , é aquele que verifica (9b).

Como ligar esses resultados com a busca de uma melhor aproximação do problema (P1)? Suponhamos, por exemplo, que houvésemos adotado o método de Galerkin. Neste caso, como já vimos, S^h e V^h coincidem, e quando o problema é predominantemente convectivo, soluções aproximadas de caráter oscilatório, e fisicamente não-realísticas, são comumente observadas. Se refinamos a malha, e portanto, aumentamos a dimensão do sub-espaço V^h , obviamente o resíduo diminui, e $\|v^h - u\|_0$ também decresce. A questão é, e se fixamos a discretização?

Neste caso, uma possível solução seria: uma vez determinada uma certa aproximação $\phi_k^h \in S^h$, calcular o correspondente campo v_k^h ; com este determinar nova aproximação $\phi_{k+1}^h \in S^h$, e novo v_{k+1}^h , e assim sucessivamente. Tal procedimento nos levaria ao seguinte esquema iterativo:

$$\begin{aligned} & \langle \phi_{k+1}^h + u \cdot \nabla\phi_{k+1}^h, \hat{\phi}^h \rangle + \langle K\nabla\phi_{k+1}^h, \nabla\hat{\phi}^h \rangle - \langle f, \hat{\phi}^h \rangle - \langle \langle h, \hat{\phi}^h \rangle \rangle = \\ & = \sum_e \int_e \left[\frac{u \cdot \nabla\phi_k^h + F(\phi_k^h)}{|\nabla\phi_k^h|^2} \nabla\phi_k^h \cdot \nabla\phi_{k+1}^h \right] \hat{\phi}^h d\Omega \quad (11) \end{aligned}$$

Desta maneira a solução aproximada e o campo de velocidades aproximado devem se auto-adaptar de forma a fornecer o menor erro quadrático na solução. O inconveniente deste esquema é que o processo é desacoplado, e portanto, deve-se esperar uma "reduzida estabilidade".

Um procedimento mais consistente pode ser obtido se exigirmos

que a solução aproximada verifique, simultaneamente, (6) e (7a-b). Antes porém algumas conclusões devem ser ressaltadas:

a) qualquer que seja ϕ^h , a equação (7a) é satisfeita com v^h definido em (9b)

b) a solução $v^h = u$ existe e ocorre sempre que

$$\sum_e \int_{\Omega_e} \frac{[u \cdot \nabla \phi^h + F(\phi^h)]^2}{|\nabla \phi^h|^2} d\Omega = 0 \quad ; \quad |\nabla \phi^h| \neq 0 \quad . \quad (12)$$

Portanto: dentre todos os ϕ^h , aquele para o qual a expressão acima é satisfeita, verifica, em cada sub-domínio Ω_e , a equação de transporte. Segue-se daí que:

Problema (P3): A função $\phi^h \in S^h$ que satisfaz, simultaneamente, a expressão (3a-e) e a condição

$$\sum_e \int_{\Omega_e} \frac{[u \cdot \nabla \phi^h + F(\phi^h)]^2}{|\nabla \phi^h|^2} \nabla \phi^h \cdot \nabla \hat{\phi}^h d\Omega = 0 \quad , \quad \forall \nabla \hat{\phi}^h \quad , \quad (13)$$

é uma solução aproximada do problema de difusão-convecção.

Realmente, se a expressão acima deve ser satisfeita para todo $\nabla \hat{\phi}^h$, em particular ela é válida para $\nabla \hat{\phi}^h \equiv \nabla \phi^h$, e neste caso

$$\sum_e \int_{\Omega_e} [u \cdot \nabla \phi^h + F(\phi^h)]^2 d\Omega = 0 \quad . \quad (14)$$

Note-se ainda que em (13) nenhuma menção é feita com relação às condições de contorno do problema. Essas são verificadas pela satisfação simultânea de (3a-e).

Antes de apresentarmos a expressão final do método proposto, um último aspecto deve ser ressaltado. As soluções ϕ^h , caracterizadas no problema (P3) devem necessariamente satisfazer a condição $\nabla \phi^h \neq 0$ pois somente assim o problema (P2) tem sentido.

Isso nos remete à consideração de duas situações distintas. A primeira, caracterizada por fenômenos puramente convectivos; a outra, oposta a essa, constituída exclusivamente por fenômenos difusivos. Vejamos cada uma de per si:

Problema (P4): Transporte puramente convectivo.

A aproximação ϕ^h , da solução do problema puramente convectivo

$$u \cdot \nabla \phi = 0 \quad \text{em } \Omega, \quad (15)$$

pode ser obtida da minimização do funcional

$$I(\phi) = \frac{1}{2} \langle u \cdot \nabla \phi^h, u \cdot \nabla \phi^h \rangle \quad (16)$$

o que variacionalmente equivale à exigência da satisfação de

$$\langle u \cdot \nabla \phi^h, u \cdot \nabla \hat{\phi}^h \rangle = 0 \quad \forall \hat{\phi}^h \in V^h, \quad (17)$$

Do ponto de vista da determinação de aproximações numéricas, a minimização da forma quadrática (16) é mais conveniente do que a aproximação de Galerkin:

$$\langle u \cdot \nabla \phi^h, \hat{\phi}^h \rangle = 0 \quad \forall \hat{\phi}^h \in V^h, \quad (18)$$

empregada na formulação clássica (ver eq.(3a)).

Observemos que a expressão (17) acrescenta uma informação adicional no cálculo da solução aproximada, que é, justamente buscar fluxos transportados por convecção segundo as linhas de corrente.

Realmente, se $u \in U$ é um campo vetorial, é possível associar, ao mesmo, o campo tensorial

$$T \equiv (u \otimes u), \quad (19a)$$

tal que

$$Tv = (u \otimes v) = (u \cdot v)u; \quad \forall v \in U \quad (19b)$$

Portanto, se a expressão (17) é escrita como:

$$\langle u \cdot \nabla \phi^h, u \cdot \nabla \hat{\phi}^h \rangle = \langle (u \cdot \nabla \phi^h)u, \nabla \hat{\phi}^h \rangle, \quad (20)$$

e identificamos $v \equiv \nabla\phi^h$, vemos que o mínimo está sendo buscado dentre todos os fluxos na direção do escoamento u .

Problema (P5): Transporte puramente difusivo

Neste caso a aproximação construída no problema (P3) não tem sentido, já que agora a solução buscada deve satisfazer:

$$F(\phi) = 0 \quad \text{em } \Omega . \quad (21)$$

É bem sabido que nesta situação a aproximação de Galerkin, obtida da condição

$$\langle F(\phi^h), \hat{\phi}^h \rangle = 0 \quad \forall \hat{\phi}^h \in V^h , \quad (22)$$

é ótima. Ainda assim podemos tentar uma formulação semelhante àque la proposta no problema (P4), identificando o campo ϕ , que satisfaz à equação de difusão (21), pela exigência de que

$$\langle F(\phi), u \cdot \nabla \hat{\phi} \rangle = 0 \quad \forall \hat{\phi} \in V . \quad (23)$$

Não é difícil de ver que a equação de Euler-Lagrange associada a (23) é:

$$\text{div}(F(\phi)u) = 0 \quad \text{em } \Omega . \quad (24)$$

Portanto, ainda que $F(\phi) = 0$ seja uma solução de (24), qualquer campo ϕ , tal que

$$\nabla F \cdot u = -(\text{div } u)F , \quad (25)$$

verifica também essa equação.

Em resumo, a forma (23) não é adequada para caracterizar a solução de um problema exclusivamente difusivo. Ora, em que medida en tão justifica-se a apresentação desses dois últimos problemas? Jus tifica-se juntamente se o problema a ser abordado envolve termos con vectivos e difusivos.

Consideremos então a seguinte expressão

$$\sum_e \int_{\Omega_e} \left[u \cdot \nabla \phi^h + F(\phi^h) \right] v^e u \cdot \nabla \hat{\phi}^h d\Omega = 0, \quad \forall \hat{\phi}^h \in V^h; \quad (26)$$

em que v^e é uma função de $\Omega_e \times [1, \infty)$ em \mathbb{R} , e que satisfaz as seguintes condições:

$$i) \quad 0 \leq v^e(x, \lambda) < 1; \quad \forall x \in \Omega_e, \quad \forall \lambda \in [1, \infty) \quad (27)$$

$$ii) \quad v^e(x, \lambda) = 0 \quad \text{em} \quad \Gamma_e \quad \text{contorno de} \quad \Omega_e, \quad \forall \lambda \quad (28)$$

$$iii) \quad \lim_{\lambda \rightarrow \infty} \left\| v^e(x, \lambda) - 1 \right\|_{0, \Omega_e} = 0; \quad (29)$$

A determinação de v^e será dada no Apêndice B.

Assim sendo, a igualdade (26) mais as restrições (27) a (29) implicam em que

$$\sum_e \int_{\Omega_e} \text{div} \left\{ \left[u \cdot \nabla \phi^h + F(\phi^h) \right] v^e u \right\} \hat{\phi}^h d\Omega = 0, \quad \forall \hat{\phi}^h \in V^h. \quad (30)$$

Portanto, para " λ " suficientemente grande, se exigirmos que a expressão (26) seja satisfeita simultaneamente com a forma (3a-e), estamos melhorando a aproximação, visto que, neste caso, também o resíduo local da equação de transporte, convectado pelo escoamento u , está sendo ortogonalizado em relação as mesmas funções base $\hat{\phi}^h$, o que deve contribuir para reduzir substancialmente as oscilações, não-físicas, que eram obtidas quando se utilizava exclusivamente a forma (3a-e) em problemas predominantemente convectivos.

Concluindo-se: se à expressão (3a-e) acrescentam-se termos da forma (13) e (26), tem-se uma melhor representação do campo aproximado ϕ^h .

Finalmente, das discussões anteriores, tem-se que a solução aproximada $\phi^h \in S^h$ deve satisfazer, para cada tempo $t \in [0, T)$, à expressão:

$$\begin{aligned} 0 = & \langle \hat{\phi}^h + u \cdot \nabla \phi^h, \hat{\phi}^h \rangle + \langle K \nabla \phi^h, \nabla \hat{\phi}^h \rangle - \langle f, \hat{\phi}^h \rangle - \langle \langle q, \hat{\phi}^h \rangle \rangle + \sum_e \int_{\Omega_e} v^e \beta \left[\hat{\phi}^h + u \cdot \nabla \phi^h + \right. \\ & \left. + \text{div}(-K \nabla \phi^h) - f \right]^2 \frac{\nabla \phi^h \cdot \nabla \hat{\phi}^h}{|\nabla \phi^h|^2} d\Omega + \sum_e \int_{\Omega_e} v^e \alpha \left[\hat{\phi}^h + u \cdot \nabla \phi^h + \right. \\ & \left. + \text{div}(-K \nabla \phi^h) - f \right] u \cdot \nabla \hat{\phi}^h d\Omega, \quad \forall \hat{\phi}^h \in V^h. \quad (31) \end{aligned}$$

Se analisarmos a equação (31) verificamos que a solução aproximada ϕ^h S^h é determinada usando-se o método de Petrov-Galerkin, com funções peso da forma:

$$w^h = \hat{\phi}^h + v^e \left\{ \alpha u + \beta \frac{[u \cdot \nabla \phi^h + F(\phi^h)] \nabla \phi^h}{|\nabla \phi^h|^2} \cdot \nabla \hat{\phi}^h \right\} \quad (32)$$

Para " λ " suficientemente grande, e se adotarmos na definição das funções livres α e β a mesma metodologia apresentada em [16] (ver Apêndice A) vemos que o método aqui proposto coincide com o método "SUPG + operador de captura", desde que:

- i) $\dot{\phi}^h \equiv 0$ (ausência de termos transientes)
- ii) $f \equiv 0$ (ausência de termos de fonte)
- iii) $\text{div}(-K\nabla \phi^h) \equiv 0$ (interpolação linear)

Quando isso não ocorre o método aqui desenvolvido apresenta uma performance superior ao SUPG e ao SPUG + operador de captura, como demonstram os exemplos numéricos que serão apresentados a seguir.

RESULTADOS NUMÉRICOS

Nos três exemplos teste que serão apresentados nesta seção, os resultados numéricos obtidos com o método aqui proposto serão comparados com a solução exata desses problemas e com os resultados obtidos quando são utilizados os métodos SUPG e o SUPG + operador de captura.

Em todos os problemas estaremos interessados na determinação de soluções estacionárias em um domínio bi-dimensional, em cujo contorno é conhecido o valor da função (ver Figura 1). Admitiremos que o meio é homogêneo e isotrópico (coeficiente de difusividade $k=10^{-8}$), e que o escoamento é unidimensional e constante.

No cálculo das soluções numéricas foram empregadas malhas uniformes, (10×10) ou (20×20) , constituídas por elementos bi-lineares com integração numérica de Gauss (2×2) . Dado o caráter iterativo do procedimento desenvolvido na seção anterior, bem como do procedimento apresentado em [16] (ver seção "Método SUPG+Operador de Captura"), adotou-se em ambos os casos, o limite $\max |\phi_k^i - \phi_{k-1}^i| < 10^{-3}$ ($i = \text{n}^\circ$; $k = \text{ordem de iteração}$) como critério de convergência.

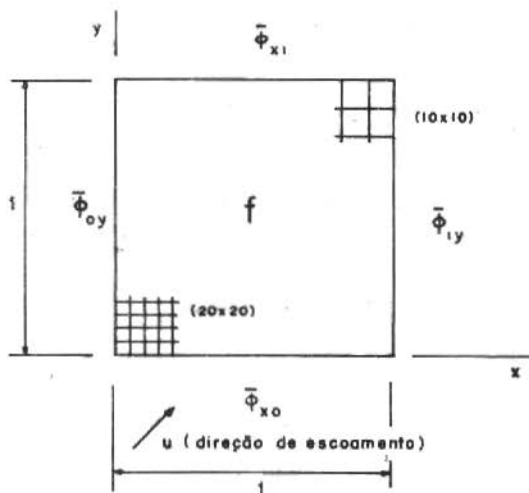


Figura 1. Definição do problema

1º Exemplo: Neste problema o termo de fonte é nulo ($f \equiv 0$), e o escoamento se processa obliquamente em relação aos eixos globais x - y , com uma velocidade de fluxo cujas componentes em relação a esses eixos valem $u_x = -1$ e $u_y = 2$. Para as condições de contorno assumiu-se que:

$$\bar{\phi}_{x0} = \begin{cases} 0; & \text{se } 0 \leq x < 0.7 \\ 1; & \text{se } 0.7 \leq x \leq 1 \end{cases}$$

$$\bar{\phi}_{1y} = \begin{cases} 1; & \text{se } 0 \leq y \leq 0.9 \\ 10(1-y); & \text{se } 0.9 < y \leq 1 \end{cases}$$

$$\bar{\phi}_{x1} = \bar{\phi}_{oy} = 0$$

Nas Figuras 3 e 4 estão mostrados, em elevação, os resultados obtidos empregando-se uma malha de (10×10) , correspondendo a um Peclet de malha $P_e^u = 2.2 \times 10^7$ (ver Apêndice A). Na Figura 2 apresenta-se a solução exata, calculada nos mesmos pontos nodais.

Apesar do Peclet resultante ser extremamente elevado, os resultados mostrados na Figura 3 concordam razoavelmente bem com a solução exata do problema. São observadas pequenas oscilações na vi-

zinhança da camada limite interna, e um certo "arredondamento" na formação dessa camada limite, indicando uma solução mais difusiva que a real. Quando comparados com os resultados plotados na Figura 4, (SUPG), facilmente observa-se o efeito favorável dos termos não-lineares, (associados à determinação da direção v^h), na redução das oscilações. Essa mesma conclusão já havia sido previamente observada em [16]. Note que, para esse problema específico, o método desenvolvido na seção anterior fica idêntico ao SUPG + operador de captura.

Para caracterizar o efeito do refinamento da discretização, na Figura 6 mostra-se a solução obtida com uma malha de (20×20) . O ganho de precisão é evidente; e a proximidade com a solução exata (Figura 5) é bastante acentuada, muito embora o Peclet de malha ($Pe^u = 1.1 \times 10^7$) seja da mesma ordem de grandeza daquele anteriormente reportado. Com relação à solução via SUPG (Figura 7), mesmo com o refinamento da malha, os resultados são, ainda assim, piores que aqueles mostrados na Figura 3.

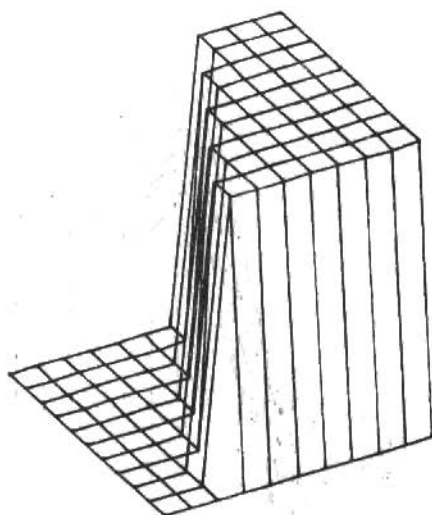


Figura 2. Solução exata

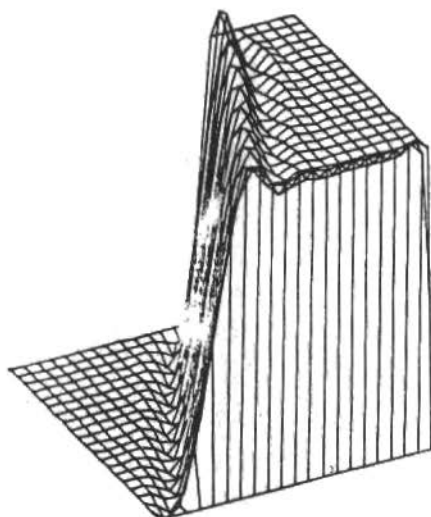


Figura 7. Solução pelo método SUPG (malha 20x20)

2º Exemplo: Neste problema o termo de fonte é constante em todo o domínio, e vale $f = 1$. O escoamento se processa na direção x (ver Figura 1) com uma velocidade $u_x = 1$. As condições de contorno são homogêneas, isto é,

$$\bar{\phi}_{x0} = \bar{\phi}_{x1} = \bar{\phi}_{0y} = \bar{\phi}_{1y} = 0 .$$

A solução exata desse problema (um plano inclinado a 45°) está mostrada na Figura 8.

Na Figura 9 tem-se, para uma malha de (10×10) ($\text{Peclet} = 1.0 \times 10^7$), a solução numérica gerada pelo modelo desenvolvido neste trabalho. É marcante a capacidade do mesmo na resolução das camadas limites externas. Oscilações são praticamente inexistentes. Situação distinta ocorre se observamos as Figuras 10 e 11. Vê-se que, para esse problema o operador de captura (Figura 10) não consegue reduzir as oscilações previstas na solução calculada via SUPG (Figura 11), chegando mesmo a apresentar maiores picos nas camadas limites laterais.

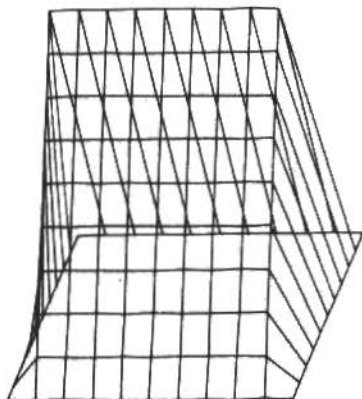


Figura 8. Solução exata

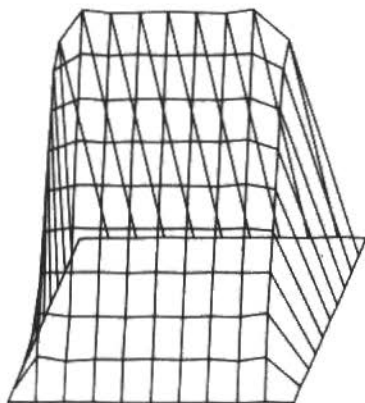


Figura 9. Solução pelo método atual (malha 10x10)

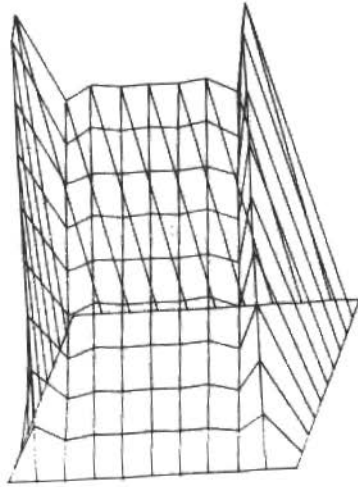


Figura 10. Solução pelo método SUPG + operador de captura (malha 10x10)

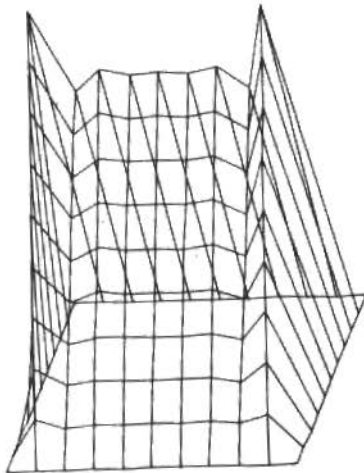


Figura 11. Solução pelo método SUPG (malha 10x10)

3º Exemplo: Este caso difere do anterior exclusivamente pelo termo de fonte empregado, que neste caso é definido pela função:

$$f = \begin{cases} 1; & \text{se } 0 < x \leq 0.5 \\ -1; & \text{se } 0.5 < x < 1 \end{cases}$$

A solução exata; um "telhado de abas" simétricas em relação à "cumieira" $y = 0.5$, com inclinação 1:1, está mostrada na Figura 12.

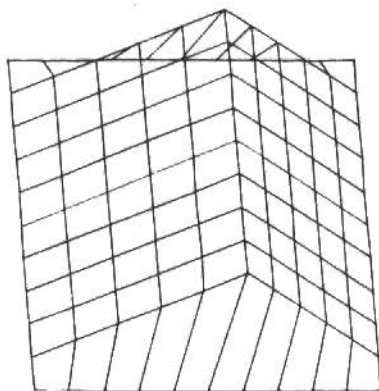


Figura 12. Solução exata

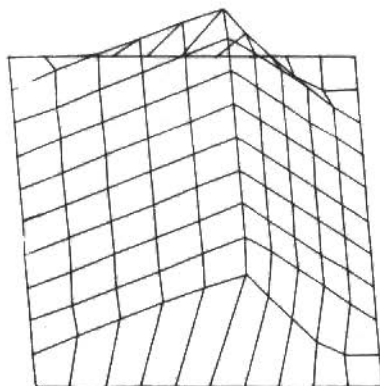


Figura 13. Solução pelo método atual (malha 10x10)

As soluções calculadas pelos três métodos, usando-se uma malha de (10×10) , estão apresentadas nas Figuras 13 a 15. De novo o método proposto apresenta uma performance superior aos outros dois. E mais uma vez, as oscilações do método SUPG são de menor intensidade que aquelas verificadas, introduzindo-se, nas funções peso, o termo associado ao operador de captura.

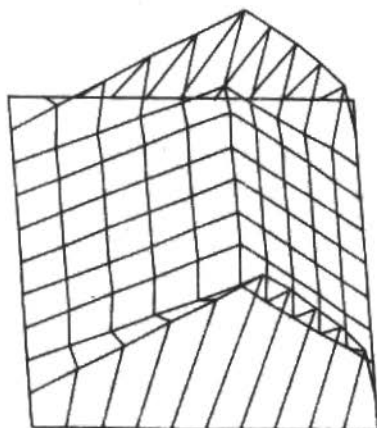


Figura 14. Solução pelo método SUPG + operador de captura
(malha 10×10)

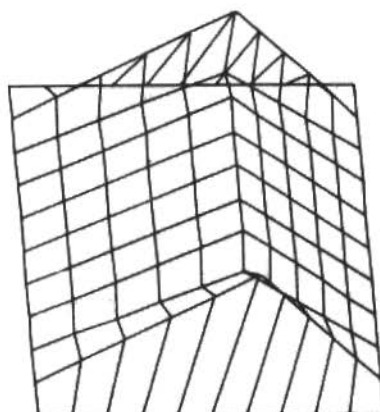


Figura 15. Solução pelo método SUPG (malha 10×10)

CONCLUSÕES

Neste trabalho apresentamos uma formulação consistente do método dos elementos finitos, para resolução de problemas difusivo-convectivos. A partir dessa formulação foi proposto um novo método que mostrou-se bastante eficaz na obtenção de soluções numéricas desses problemas, mesmo nas situações em que o transporte se dá quase que exclusivamente por convecção.

A fácil generalização desse método, permitindo a utilização de diferentes tipos de elementos finitos, evidencia a sua potencialidade. No entanto, a sua aplicação exige a resolução de um sistema de equações não-lineares, mesmo que o problema seja linear.

Os resultados numéricos apresentados demonstraram a capacidade de desse método na resolução de camadas limites internas e/ou externas, e a sua superioridade em relação aos métodos SUPG e SUPG + operador de captura.

Generalização do novo método a sistema de equações está em andamento. Além disso, desenvolvimentos estão sendo feitos no sentido da determinação das funções de ponderação, α e β , ótimas.

REFERÊNCIAS

- [1] Leonard, B.P., A survey of finite differences of opinion on numerical muddling of the incomprehensible defective convection equation. Finite Element Methods for Convection Dominated Flows, AMD vol.34, ed.T.J. Hughes (ASME, New York, 1979).
- [2] Brooks, A.N. and Hughes, T.J.J., Streamline upwind Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations. Comput. Meths. Appl. Mech. Engrg., 32 :199-259 (1982).
- [3] Spalding, D.B., A novel finite difference formulation for differential expressions involving both first and second derivatives. Int. J. Numer. Meths. Engrg., 4 :551-559 (1972).
- [4] Leonard, B.P., A stable and accurate convective modelling procedure based on quadratic upstream interpolation. Comput. Meths. Appl. Mech. Engrg., 19 :59-98 (1979).
- [5] Strang, G. and Fix, G.J., An analysis of the finite element method. Prentice-Hall (1973).
- [6] Zienkiewicz, O.C.; Gallagher, R.H. and Hood, P., Newtonian and non-Newtonian viscous incompressible flow. 2nd Conf. Mathe-

- matics of finite elements and applications, ed. J.R.Whiteman (Academic Press, New York, 1975).
- [7] Christie, I.; Griffiths, D.F. and Mitchell, A.R., Finite element methods for second order differential equations with significant first derivatives. Int. J. Numer. Meths. Engrg., 10 : 1389-1396 (1976).
- [8] Heinrich, J.C.; Huyakorn, P.S.; Zienkiewicz, O.C. and Mitchell, A.R., An upwind finite element scheme for two-dimensional convective transport equation. Int. J. Numer. Meths. Engrg., 11 : 134-143 (1977).
- [9] Heinrich, J.C. and Zienkiewicz, O.C., Quadratic finite element schemes for two-dimensional convective - transport problems. Int. J. Numer. Meths. Engrg., 11 : 1831-1844 (1977).
- [10] Hughes, T.J., A simple scheme for developing "upwind" finite elements. Int. J. Numer. Meths. Engrg., 12 : 1359-1365 (1978).
- [11] Hughes, T.J.; Liu, W.K. and Brooks, A., Finite element analysis of incompressible viscous flows by the penalty function formulation. I. Comput. Phys., 30 : 1-60 (1979).
- [12] Kelly, D.W.; Nakazawa, S.; Zienkiewicz, O.C. and Heinrich, J.C., A note on upwinding and anisotropic balancing dissipation in finite element approximations to convective diffusion problems. Int. J. Numer. Meths. Engrg., 15 : 1705-1711 (1980).
- [13] Hughes, T.J. and Brooks, A., A theoretical framework for Petrov-Galerkin methods with discontinuous weighting functions: Application to the streamline upwind procedure. Finite Element in Fluids, vol.IV, eds. R.H. Gallagher et al. (Wiley London, 1982).
- [14] Johnson, C.; Nävert, U. and Pitkäranta, J., Finite element methods for linear hyperbolic problems. Comput. Methds.Appl. Mech. Engrg., 45 : 285-312 (1984).
- [15] Mizukami, A. and Hughes, T.J., A Petrov-Galerkin finite element method for convection-dominated flows: An accurate upwind technique for satisfying the maximum principle. Comput.Meths. Appl. Mech. Engrg., 50 : 181-193 (1985).
- [16] Hughes, T.J.; Mallet, M. and Mizukami, A., A new finite element formulation for computational fluid dynamics: II Beyond SUPG. Comput. Meths. Appl. Mech. Engrg., 341-355 (1986).

APÊNDICE A

Seja Δ_e^h dado por:

$$\Delta_e^h = - \frac{[u \cdot \nabla \phi^h + F(\phi^h)]}{|\nabla \phi^h|^2} \nabla \phi^h \quad (\text{A.1})$$

Desta forma, o campo v^h em Ω_e é dado por:

$$v^h = u + \Delta_e^h \quad (\text{A.2})$$

Compatibilizando v^h e u , escolheremos ω_e^h da seguinte forma:

$$\omega_e^h = \tau_1 \frac{v^h}{|u|} - \tau_2 \frac{\Delta_e^h}{|\Delta_e^h|} \quad (\text{A.3})$$

Seja $x^e = x^e(\varepsilon)$ o mapeamento geométrico do domínio gerador para Ω_e . O mapeamento inverso é escrito como $\varepsilon^e = \varepsilon^e(x)$.

Sejam os vetores a^e e b^e ; dados por:

$$a^e = (\nabla \varepsilon^e) u \quad (\text{A.4})$$

$$b^e = (\nabla \varepsilon^e) \Delta_e^h \quad (\text{A.5})$$

Fazendo uso da metodologia utilizada em [7] e [16], definimos h_e^u e h_e^Δ de modo que:

$$h_e^u = 2|u| / |a^e| \quad (\text{A.6})$$

$$h_e^\Delta = 2|\Delta_e^h| / |b^e| \quad (\text{A.7})$$

O número de Peclet de malha é definido como:

$$P_e^u = \frac{1}{2} \frac{h_e^u |u|}{k} \quad (\text{A.8})$$

$$P_e^\Delta = \frac{1}{2} \frac{h_e^\Delta |\Delta_e^h|}{k} \quad (\text{A.9})$$

Sejam $\delta^u(P_e^u)$ e $\delta^\Delta(P_e^\Delta)$ dados por:

$$\delta^u(P_e^u) = \max \left\{ 0, c_2 - \frac{c_2}{c_1 P_e^u} \right\} \quad (\text{A.10})$$

$$\delta^\Delta(P_e^\Delta) = \max \left\{ 0, c_2 - \frac{c_2}{c_1 P_e^\Delta} \right\} \quad (\text{A.11})$$

onde as constantes c_1 e c_2 , dependem do tipo de elemento utilizado.

Por exemplo, para elementos lineares em uma dimensão, bilineares em duas dimensões, e trilineares em três dimensões os valores c_1 e c_2 são:

$$c_1 = 1 \quad (\text{A.12})$$

$$c_2 = 1 \quad (\text{A.13})$$

Com essas definições τ_1 e τ_2 são expressos pelas relações:

$$\tau_1 = \frac{h_e^u \delta^u(P_e^u)}{2} \quad (\text{A.14})$$

$$\tau_2 = \begin{cases} 0, & \text{se } \tau_1 = 0 \\ \frac{h_e^\Delta \delta^\Delta(P_e^\Delta)}{2} \end{cases} \quad (\text{A.15})$$

Fazendo uso de (A.1) a (A.3) obtemos:

$$w_e^h = \frac{\tau_1 u}{|u|} + \left[\frac{\tau_2}{|\Delta_e^h|} - \frac{\tau_1}{|u|} \right] \left[\frac{u \cdot \nabla \phi^h + F(\phi^h)}{|\nabla \phi^h|^2} \right] \nabla \phi^h \quad (\text{A.16})$$

Seja agora w_e^h a função peso dada por:

$$w_e^h = \begin{cases} \hat{\phi}^h + v^e \{ w_e^h \cdot \nabla \hat{\phi}^h \}; & \text{se } \frac{\tau_2}{|\Delta_e^h|} \geq \frac{\tau_1}{|u|} \\ \hat{\phi}_e^h + v^e \left\{ \frac{\tau_1}{|u|} u \cdot \nabla \hat{\phi}^h \right\}; & \text{se } \frac{\tau_2}{|\Delta_e^h|} < \frac{\tau_1}{|u|} \end{cases} \quad (\text{A.17})$$

Assim, por comparação com a equação 32 da seção "Método Proposto", as funções α , β , são determinadas e valem:

$$\alpha = \frac{\tau_1}{|u|} \quad (\text{A.18})$$

$$\beta = \max \left\{ 0, \frac{\tau_2}{|\Delta_e^h|} - \frac{\tau_1}{|u|} \right\} \quad (\text{A.19})$$

APÊNDICE B

A função "sino", $v^e(x, \lambda)$, pode ser definida para os elementos tradicionais, em termos das coordenadas locais, da seguinte forma:

I. UNI-DIMENSIONAL

$$v^e(x, \lambda) = (-1 / (\lambda(1-\varepsilon^2)))$$

$$-1 \leq \varepsilon \leq 1 \quad ; \quad x = x(\varepsilon) \quad (\text{B.1})$$

II. BI-DIMENSIONAL

Elemento Triangular

$$v^e(x, \lambda) = \exp(-1 / (\lambda(27 \cdot L_1 \cdot L_2 \cdot L_3))) \quad (\text{B.2})$$

$x = x(L_1, L_2, L_3)$, onde L_1, L_2 e L_3 são as coordenadas de área.

Quadrilátero

$$v^e(x, \lambda) = \exp(-1 / (\lambda(1-\varepsilon^2)(1-\eta^2)))$$

$$-1 \leq \eta \leq 1 \quad ; \quad x = x(\varepsilon, \eta) \quad (\text{B.3})$$

$$-1 \leq \varepsilon \leq 1$$

III. TRI-DIMENSIONAL

Tetraedro

$$v^e(x, \lambda) = \exp(-1 / (\lambda(256 \cdot L_1 \cdot L_2 \cdot L_3 \cdot L_4))) \quad (\text{B.4})$$

$x = x(L_1, L_2, L_3, L_4)$, onde L_1, L_2, L_3, L_4 são as coordenadas de volume.

Paralelepípedo

$$v^e(x, \lambda) = \exp(-1 / \lambda \cdot (1 - \varepsilon^2) \cdot (1 - \delta^2) \cdot (1 - \eta^2))$$

$$-1 \leq \delta \leq 1$$

$$-1 \leq \eta \leq 1 \quad ; \quad x = x(\varepsilon, \eta, \delta) \quad (B.5)$$

$$-1 \leq \varepsilon \leq 1$$

Das equações (b.1) a (B.5), vê-se que a função "sino" $v^e(x, \lambda)$ satisfaz as condições (27) a (29) da seção "Método Proposto".

Na prática trabalhamos com λ suficientemente grande, o que em termos numéricos é equivalente a assumir $v^e(x, \lambda) = 1$.

ÍNDICE GERAL

VOL. VIII - Nº 1

- Vibrating Subsoilers: A way reduce power and draft forces in soil preparation 5
- Léopore N., F.P.**
Steffen Jr., V.
Mech. Eng. Department
Federal University of Uberlândia
38400 Uberlândia - MG
Brasil
- Flow past an oscillating infinite vertical plate with oscillating plate temperature 15
- V. M. Soundalgekar**
7/10, Vivekanand housing Soc.
Saraswat Colony
Dombivali (E) 421 201
India
- S. T. Revankar**
Dept. of Nuclear Engng.
University of California
Berkeley - CA
USA
- Análise da solidificação em lingotamento contínuo nas várias etapas do processo 21
- Carlos Lineu de Farias e Alves**
CTA-ITA-IEM-IEM
- Roberto Vanuchi Fernandes**
EEI
- Análise da troca de calor por radiação difusa em uma cavidade cinzenta trapezoidal 31
- Rogério Martins Saldanha da Gama, Membro da ABCM**
LNCC/CNPq
- Análise experimental de tensões térmicas em trilhos longos soldados 45
- Jaime Tupiassú Pinho de Castro, Membro da ABCM**
José Luiz de França Freire, Membro da ABCM
Deprtº de Engenharia Mecânica - PUC/RJ

VOL. VIII- Nº 2

- Stability of centrifuges partially filled with liquid 73
Dr. Eberhard Brommundt
Institut für Technische Mechanik (Mechanikzentrum)
Technische Universität Braunschweig
D-3300 Braunschweig
- Dynamics of the relative motion of a solid particle and a viscous fluid confined in cylindrical ducts of different geometries 87
Roberto Guimarães Pereira, Membro da ABCM
Maurício Nogueira Frota, Membro da ABCM
Departamento de Engenharia Mecânica
Pontifícia Universidade Católica do Rio de Janeiro
- Determinação teórica e experimental da velocidade de queima de propelente sólido 105
Aparecida Minhoko Kawamoto
IAE - Divisão de Química
CTA - São José dos Campos, SP
Gordiano Faria de Alvim Filho
ITA - Departamento de Aeronáutica
CTA - São José dos Campos, SP
- Sobre as soluções de problemas de troca de energia radiante térmica 121
Rogério Martins Saldanha da Gama – Membro da ABCM
LNCC/CNPq
- Análise não-linear do equilíbrio de estruturas unidimensionais elásticas planas 131
Heraldo Silva da Costa Mattos
Rubens Sampaio Filho, Membro da ABCM
Depto de Engenharia Mecânica, PUC/RJ
- Modeling of flexible multibeam systems by rigid-elastic superelements 151
Werner O. Schiehlen
Jochen Rauh
University of Stuttgart, Stuttgart, F. R. G.

VOL. VIII- Nº 3

- Machine-foundation-soil Interaction; combined
continuum and boundary element approach 169
L. Gaul
Institut für Mechanik
University of the German Armed Forces Hamburg
- Toward the utilization of computer intelligence in
metal machining 199
Dr. Hejat Olgac
University of Connecticut – USA
- Sobre a interpretação do tensor parcial de tensão e da
força difusiva em misturas sólido-fluido 219
Rogério Martins Saldanha da Gama – Membro da ABCM
LNCC/CNPq
Rubens Sampaio – Membro da ABCM
DEM-PUC/RJ
- Stability analysis of dissipative nonlinear dynamical
systems 232
Edwin J. Kreuzer
Institute B of Mechanics
University of Stuttgart
Pfaffenwaldring 9
D-7000 Stuttgart 80, F. R. G.

Vol. VIII, nº 4

Editorial

Dynamic response of a rigid disc mounted on a flexible shaft under non-linear excitations 261

V. Steffen Jr. - Membro da ABCM e SEM

F.P. Lépore - Membro da ABCM

E.B. Teodoro

Universidade Federal de Uberlândia

Departamento de Engenharia Mecânica

Uberlândia - MG

Randon vibrations and noise generation in railway wheels 273

K. Popp

Institute of Mechanics

University of Hannover

FRG

State estimation of multivariable systems using the duality concept 285

Agenor de Toledo Fleury

Atair Rios Neto

Instituto de Pesquisas Espaciais - INPE/MCT

São José dos Campos/SP

Uma formulação consistente do método dos elementos finitos para resolução de problemas de transporte difusivo-convectivo 309

Eduardo Gomes do Carmo - Membro da ABCM

COPPE/UFRJ

Augusto Cesar Galeão - Membro da ABCM

INCC/CNPq

