Ministério da Aeronáutica
Departamento de Pesquisa e Desenvolvimento

Instituto de Atividades Espaciais
Centro Técnico Aeroespacial
12.200 - São José dos Campos - SP
Brasil

EIGENVALUE TRANSLATION METHOD FOR MODE CALCULATIONS

by

Edgardo Gerck

and

Carlos H. Brito Cruz

10 de Novembro de 1978

Relatório EAV-10/78

ABSTRACT

A new method is described for the first few modes calculations in a interferometer that has several advantages over the ALLMAT subroutine, the Prony Method and the Fox and Li Method.

In the illustrative results shown for same cases it can be seen that the eigenvalue translation method is typically 100 fold times faster than the usual Fox and Li Method and 10 times faster than ALLMAT.

I. Introduction

-. The purpose of this paper is to present a physically interesting and very efficient method for the calculation of the first few dominant modes of a general interferometer.

To calculate the resonant optical modes the Kirchhoff integral equation is usually transformed in a matrix equation:

$$Mx = \lambda x$$

If the matrix equation is to accurately represent the integral equation, the order of the matrix M must be large.

The current criterium is that the eigenvector x must have from 6 N to 10 N components [1] where N is the resonator Fresnel number.

As pointed out by Siegman [1] , it is useful to have a simple algorithm to calculate the dominant eigenmodes of a, for example, 100 x 100 matrix, which means an interferometer with N = 10. With these large matrices it is not recommended to use the "optimal" procedure as suggested by Sanderson and Streifer [2]. While we are in general interested only on the first few dominant modes the recommended subroutine IBM-ALLMAT will iterate needlessly to calculate 100 modes of the 100 x 100 matrix.

This means that computer time will be used to calculate typically less than 10 actual modes and about 90 random collections of points due to the non-convergence of the process. Besides, we also have to store large amounts of data in rapid access memory.

Siegman [1] has proposed the Prony method to achieve a higher efficiency in calculating only the dominant modes of the eigenequation. The Prony method also does not need to store the matrix M. We present here the eigenvalue translation method which has the following advantages:

(1) The iterative process of wave boucing requires a number typically 100 fold less of trips than the Fox and Li method and is faster than ALLMAT for the dominant modes calculation regardless of the order of the matrix M;

(2) ease implementation on a computer, less than 50 lines of a FORTRAN program;

(3) no auxiliary subroutines (like polynomial root finder) are needed;

(4) Due to the fast convergence, the process can be visualized on a display to analyse mode beating;

(5) yields eigenvalues and eigenvectors simultaneously;

(6) can be used in multiple mirror systems where there could not be a definite integral equation;

(7) since it is a propagation method there is no need to store the matrix M.

The proposed method seems to be optimal for the first four eigenmode calculation (that is the (0,0), (0,1), (1,0) and (1,1) modes) in the sense of simplicity and efficiency. It can be applied to lower order modes at a slight cost on total simplicity.

II. The Integral Equation

The integral equation for the optical resonator has the following general form

$$\alpha_{p\ell} f_{p\ell}(r) = \int K_\ell(r,s) A(s) f_{p\ell}(s) ds \qquad (1)$$

where the kernel K may depend on ℓ as for a cavity in cilyndrical coordinates. The A(s) factor is usually A(s) = s or A(s) = 1. Note that as $f_{p\ell}$ appears on both sides of eq. (1), the equation represents either single-trip propagation for symmetric modes or round-trip for the unsymmetric modes. These two cases exaust the possibilities for all interferometers.

Defining the scalar product

$$(f,g) = \int f(s)g(s)A(s)ds \tag{2}$$

we may write (1) as the scalar product

$$\alpha_{p\ell} f_{p\ell} = (K_\ell, f_{p\ell}) \tag{3}$$

The kernel $K_\ell(r,s)$ is always symmetric [10] for any cavity, including misalignments, misfiguring of mirrors, etc. This implies in $(f_{p\ell}, f_{qn}) = \delta_{pq}^{\ell n}$ through normalization.

In transforming the integral equation (1) in matrix form by the Gaussian quadrature formula for example, the Gaussian weights and the function A(s) will render M unsymmetric in the equation

$$\lambda x = Mx \tag{5}$$

As the scalar product must be conserved [11] in this transformation we may rewrite (5) as

$$\lambda x = (M', x) \tag{6}$$

where $M'_{ij} = K_\ell(r_i, s_j)$. Hence M' will always be symmetric in this formulation of (1). This will have a direct influence on the convergence rate of the method.

Let us describe the proposed method in the general context of (1).

III. Eigenvalue translation

i) Description

Starting with an arbitrary function h^0, decomposable through the scalar product (2) on the $f_{p\ell}$ basis:

$$h^0(s) = \sum_{i=0}^{\infty} a_i f_{i\ell}(s) \qquad \text{(for a given } \ell\text{)}$$

we construct the sequence:

$$h^n = (K_\ell, h^{n-1}) = (K_\ell^n, h^0) = \sum_{i=0}^{\infty} a_i \alpha_{i\ell}^n f_{i\ell} \tag{7}$$

In the usual Fox and Li method the sequence is continued until convergence to the dominant mode occurs. But because the dominant eigenvalue $\alpha_{0\ell}$ is near to the next $\alpha_{1\ell}$ in absolute value, convergence is slow. Is is clear that if we can somehow eliminate the $(p = \text{even}, \ell)$ or the $(p = \text{odd}, \ell)$ modes, the convergence rate of the sequence to the then dominant mode shall increase.

Another idea might be to selectively amplify the p = odd or the p = even modes to increase convergence to the p-set chosen.

In the proposed method we use both effects, i.e. selective filtering and amplification by summing a constant μ to all eigenvalues of the equation (1). This is the technique of eigenvalue translation [7] .

Mathematically, if the equation

$$\alpha_{p\ell} f_{p\ell}(r) = \int K_\ell(r,s) A(s) f_{p\ell}(s) ds$$

has eigenvalues $\alpha_{p\ell}$ then the equation

$$(\alpha_{p\ell} + \mu) f_{p\ell}(r) = \int \{K_\ell(r,s) A(s) + \mu\delta(r-s)\} f_{p\ell}(s) ds$$

where $\delta(r-s)$ is the Dirac delta, has eigenvalues $\alpha'_{p\ell} = \alpha_{p\ell} + \mu$, translated by μ, and with the same eigenvectors. If we work with the matrix equation

$$Mx = \lambda x$$

we need only to form the matrix $M' = M + I\mu$ to obtain $\lambda' = \lambda + \mu$ with x unchanged. However, in our case as we have a delta distribution in the kernel, there will be a slight difference, as we shall see in item 2.

If a $\alpha'_{p\ell}$ is translated outside the unit circle, this implies in gain for this mode.

Conversely if it is translated inwards this results in an increased loss. Our idea would be to use this feature to accelerate convergence.

Since all $\alpha_{p\ell}$ are in the unit circle ($|\alpha_{p\ell}|<1$), this has two implications on μ:

(i) $|\mu|=1$ is a sufficient condition to translate any desired eigenvalue outside the unit circle.

(ii) μ can always be a real parameter, as can be seen by simple geometric considerations on the complex plane.

A positive μ would determine the dominant eigenvalue on the right half-plane and a negative μ could be used for the left half-plane.

However, the second condition (ii) is very sensitive to the value of $|\mu|$. To see this, first rewrite (7) as:

$$h^n = a_0 \alpha_{0\ell}^n \left(f_{0\ell} + \left(\frac{a_1}{a_0}\right)\left(\frac{\alpha_{1\ell}}{\alpha_{0\ell}}\right)^n f_{1\ell} + \dots \right)$$

where $\left|\frac{\alpha_{1\ell}}{\alpha_{0\ell}}\right|^2 < 1$ by hypothesis. Our first aim could be to determine the dominant eigenvalue on the right half-plane, thus $\mu > 0$.

As a worst case of this analysis on μ, suppose that the dominant eigenvalue in the RHP is $\alpha_0 = 0 + i\varepsilon$ and the second dominant is $\alpha_1 = \delta + i0$, with $|\varepsilon| > \delta > 0$.

After translation by $\mu + i0$ we will have:

$$\left|\frac{\alpha_1{}^1}{\alpha_0{}^1}\right|^2 = \frac{(\delta + \mu)^2}{\varepsilon^2 + \mu^2} \tag{8}$$

and if (8) is to be still smaller than one the condition on μ is

$$0 < \mu < \frac{\varepsilon^2 - \delta^2}{2\delta} \tag{9}$$

As for the dominant modes $|\delta| \doteq |\varepsilon|$, the condition (9) for convergence to $\alpha_{0\ell}$ is very sensitive on μ. The argument can be easily understood from Fig. 1.

The same reasoning applies to the LHP case.

To overcome this difficulty we used the following procedure: with $\mu = 0$, N iterations are made and with the calculated value of the $\alpha_{0\ell}$ we make $\mu = s\alpha_{p\ell}$, the constant s chosen between 0 and 1 (s = 0.5 for example). Other N iterations are made and μ readjusted as before, continuing the process until convergence occurs, within the desired error limit. Typically N = 10 yields good results.

Although the dominant modes of an interferometer have nearly the same diffraction losses, by selectively filtering and amplifying a set of modes we were able to decrease by a factor of 100 the number of bounces inside the interferometer to achieve a stationary state pattern.

2) Method:

Using (6) we calculate the translated equation:

$$\alpha'_{p\ell} f_{p\ell} = (M', f_{p\ell}) + I\mu f_{p\ell} \tag{10}$$

Eq. (10) is not identical to

$$\alpha'_{p\ell} f_{p\ell} = (M' + I\mu, f_{p\ell})$$

because I comes from a delta function and must not be affected by A(s) and the quadrature.

Specifying an arbitrary vector h^0 we form the sequence

$$h^n = (M', h^{n-1}) + Ih^{n-1} \quad ; \; n = 1,\ 2,\ \ldots.$$

which converges to the dominant mode $f_{0\ell}$. The eigenvalue $\alpha'_{0\ell}$ is calculated by the Rayleigh quotient

$$\alpha'_{0\ell} \simeq \frac{(h^{n+1}, h^n)}{(h^n, h^n)} \tag{11}$$

using the scalar product (2). The $\alpha_{0\ell}$ is calculated by $\alpha_{0\ell} = \alpha'_{0\ell} - \mu$.

According to Ralston [8], because M' is always symmetric under (2), the error in (11) is given by

$$\alpha'_{0\ell} = \frac{(h^{n+1}, h^n)}{(h^n, h^n)} + 0\left[\left(\frac{\alpha_{1\ell}}{\alpha_{0\ell}}\right)^{2n}\right]$$

If we had not used (2) we would have for the error the expression

$$0\left[\left(\frac{\alpha_{1\ell}}{\alpha_{0\ell}}\right)^n\right]$$

implying a much slower convergence rate because of the matrix non-symmetry.

Thus we see one more advantage of formulating the algebraic eigenproblem as (6).

To calculate the higher order modes we can use one of two techniques:

(i) matrix deflation;

(ii) by taking the scalar product we can extract the $(0,\ell)$ mode from the original distribution and with the modified h achieve convergence to the then dominant mode.

The second process is to be preferable because cumulative errors do not set in. However we feel that the eigenvalue translation method is actually superior for the first four modes calculation and that for higher modes the question is more a matter of taste.

IV. Integral Equations Used

To make clear the concepts and definitions we shall be using in this article let us state the integral equations we will be working with:

IV.1 Circular Geometry - we shall assume identical circular mirrors of diameter 2a, radius of curvature R, set apart with a centerline distance d. The usual integral equation for the resonant one-trip mode $f_{p\ell}$ is:

$$\gamma_{p\ell} f_{p\ell}(r) = \int_0^1 \frac{F}{2\pi}\, i^{\ell+1} J_\ell(Frs) e^{-igF(r^2+s^2)/2} \cdot s f_{p\ell}(s) ds$$

where $g = 1 - d/R$ and $F = ka^2/d$.

As $f_{p\ell}$ is a mode in resonance the eigenvalues $\gamma_{p\ell}$ are real for any cavity, due to the resonance condition. The factor e^{-ikd} is a modal factor of the form $\exp(-i(kd)_{p\ell})$.

This constant factor, which has its origin in the kernel, has the characteristics of a (p,ℓ) eigenvalue and cannot remain in the kernel if we want to solve for the resonant $f_{p\ell}$ modes, because it is an unknown factor before obtainning the solution.

The eigenequation to be solved is then:

$$\left[\gamma_{p\ell}e^{i(kd)_{p\ell}}\right] f_{p\ell}(r) = \frac{F}{2\pi}\, i^{\ell+1} \int_0^1 J_\ell(Frs)e^{-igF(r^2+s^2)/2} . s f_{p\ell}(s)ds \tag{1}$$

which has complex eigenvalues $\alpha_{p\ell} = \gamma_{p\ell}e^{i(kd)_{p\ell}}$.

To change the problem to that of a symmetrical kernel, let us specify the scalar product in the vectorial spaces involved as

$$(f,g)_c = \int_0^1 f(s)g(s)sds \tag{2}$$

Denoting the kernel by $K_\ell(r,s)$:

$$K_\ell(r,s) = \frac{F}{2\pi}\; i^{\ell+1} e^{-igF(r^2+s^2)/2}$$

which is symmetrical, i.e., $K_\ell(r,s) = K_\ell(s,r)$. We see that $\alpha_{p\ell} f_{p\ell}(r)$ is the projection by $K_\ell(r,s)$ of $f_{p\ell}(s)$ with the scalar product (2):

$$\alpha_{p\ell} f_{p\ell} = (K_\ell, f_{p\ell})$$

with $(f_{p\ell}, f_{qn}) = \delta^{pq}_{\ell n}$ through normalization.

It is interesting to note that for infinite mirrors, this projection induces a transform $H_{g,\ell}$ such that $H_{0,\ell}$ is the usual generalized Hankel Transform.

II.2 Rectangular Geometry - we shall assume infinite strip mirrors, of width 2a (from -a to a), with centerline spacing d, for which the integral equation is:

$$\alpha_j f_j(y) = \left(\frac{iF}{2\pi}\right)^{1/2} \int_{-1}^{1} e^{iFxy} e^{-igF(x^2+y^2)/2} f_j(x)dx$$

In this case the scalar product is

$$(f,g)_R = \int_{-1}^{1} f(x)g(x)dx \qquad (3)$$

In the limit for infinite mirrors, this equation will yeld a transform, which for the case $g = 0$ is the usual Fourier Trasnform.

Let us describe now the results obtained.

V. Results:

The confocal stable resonator is the symmetrical cavity with smallest diffraction loss for the dominant modes and due to this, was used as a test for efficiency, as a worst case analysis. However its eigenvalues are always real [9] , so that it is a particular case that cannot be a conclusive test for the eigenvalue translation method.

Due to this, other configurations were calculated, so that besides the confocal case we will analyse an unstable and stable strip configurations.

V.1 Confocal case - a resonator Fresnel number N = 1.5 was used in two test cases, A and B. The graphs will show the radial amplitude distribution in a half mirror. We have used a fixed μ because the $\alpha_{p\ell}$ are real.

Case A

A uniform field distribution is launched from one of the mirrors; the Gauss-Legendre quadrature method was used to obtain a 32 x 32 matrix.

For $\mu = 0$ the distribution are shown in figure 3; convergence is attained after 2500 bounces.

For $\mu = 1$ we can see in figure 3 that after only 34 bounces the steady-state field is reached.

For $\mu = .5$ the steady-state Gaussian field is obtained in only 20 bounces.

We can see that with $\mu = 0$ (the Fox and Li method) the amplitude distribution was different in each mirror before the steady-state due to odd-even mode beating. With $\mu = 1$ under the same conditions, there was no oscillatory mode beating (due to p = odd mode filtering) and convergence was attained in 1/60th the number of bounces for $\mu = 0$. With $\mu = .5$ the settling time (Fig. 2) is less and the dominant mode is obtained in only 20 bounces, that is more than 100 times faster.

The effect of μ on the convergence rate is illustrated in figure 2.

A similar plot is obtained for $\alpha_{1\ell}$ with $\mu<0$.

Since the asymptotic behaviour with $\mu = \pm 1$ implies a long settling time and with $\mu = 0$ the convergence is too oscillatory due to mode beating [3], μ_{opt} is in the range $0<|\mu|<1$.

Next we used a different profile for h^0, as follows.

Case B.

We specified for h^0 a random distribution of points with null average amplitude, ranging from -.5 to +.5 in amplitude. It was employed a 100 x 100 matrix, with points equally spaced.

The high matrix order was necessary due to the high spatial frequency content of the launched wave.

The results are depicted in Figure 4, for $\mu = 0$, .5 and -.5. We see the markedly different behaviour between the $\mu = 0$ and the $\mu \neq 0$ cases. For $\mu = .5$ the steady-state was reached in 2465 bounces less than with $\mu = 0$ and the (1,0) mode was reached in 2470 less bounces than in the usual Fox and Li method.

The random distribution was used to excite a large number of modes of the cavity.

V.2 Unstable and Stable Strip Case

For a variety of configurations the dominant eigenvalue was calculated by the Fox and Li method, ALLMAT and eigenvalue translation using a 96 x 96 matrix. The results scale easily to other matrix dimensions N because diagonalization time in ALLMAT is proportional to N^3 and for vector propagation each iteration time is proportional to N^3 also.

The results are sumarized in Table 1. We see that the proposed method is of the order of 100 times faster than Fox and Li method ($\mu = 0$ case) and 10 times faster than ALLMAT.

In this case we have used the adaptive μ technique with N = 10 and s = 0.5 because $\alpha_{p\ell}$ is complex. This is the general method.

VI. Conclusions

To calculate the first four dominant modes of an interferometer the process of eigenvalue translation is recommended for large or small matrices alike. The process is simple, does not need auxiliary subroutines and is faster than those currently used. It can be readily extended to higher order modes as already explained, although loosing in simplicity.

The reason for the increase in efficiency over other methods is the effective use of the information available on the eigenvalues, v.g. $|\alpha_{p\ell}|<1$. This technique is equivalent to a selective mode filtering and amplification for $|\alpha_{0\ell}+\mu|>1$ implies power gain as the wave bounces to and fro.

The effect of μ on convergence is seen in Figure 2.

Since the assimptotic behaviour with $|\mu|=1$ implies a long settling time, and with $\mu=0$ the convergence is too oscillatory due to mode beating [3], μ_{opt} is in the range $0<|\mu|<1$.

We have used $|\mu| \simeq 0.5$ by specifying $\mu = s\alpha_{p\ell}$ for $s = .5$ with good results. The optimum value depends of course on the detailed mode structure which is by assumption unknown.

This method can also be applied even when there is not a single definite integral equation because then the eigenvalues of each kernel could be translated and therefore improve the convergence.

For unstable cavities, the steady-state patterns are attained in as little as 30 bounces because of the higher loss for the non-dominant modes.

VII. Acknowledgements

We are indebted to Prof. O. Andrade for his many fruitful suggestions and to the Universidade de São Carlos-USP-Instituto de Física e Química.

This study was supported in part by CNPq.

References

1 - A.E. Siegman and H.Y. Miller, Appl. Opt. 9, 2729)1970).

2 - R.L. Sanderson and W. Streifer, Appl. Opt. 8, 131 (1969).

3 - A.G. Fox and T. Li, Bell Syst. Tech., J., 453 (1961).

4 - G.D. Boyd and H. Kogelnik, Bell Syst. Tech., J., 1347 (1962).

5 - W.S.C. Chang, Principles of Quantum Electronics (Addison-Wesley, 1969), Chap. 7, p. 277.

6 - H. Blok, Diffraction Theory of Open Resonators, PhD Thesis, Rotterdam (1970).

7 - A.R. Gourlay and G.A. Watson, Computational Methods for Matrix Eigenproblems (John Wiley & Sons, 1973) Chap. 4, ps. 42 and 45.

8 - A. Ralston - A First Course in Numerical Analysis, McGraw-Hill, New York (1965).

9 - G.D. Boyd and J.P. Gordon - BSTJ, 40, 489 (1961).

10 - $K_\ell(r,s)$ is the Green function of the problem and as the light path is always reversible in an interferometer or laser, $K_\ell(r,s) = K_\ell(s,r)$.

11 - The discret limit of the matrix equation must be equal to the integral equation. Thus the scalar product must have the same definition in both cases.

FIGURE CAPTIONS

Figure 1: Vectorial composition depicting complex eigenvalue translation in a worst case analysis for the Right-Hand Plane.

Figure 2: Eigenmode evolution to steady-state for each value of the translation μ parameter, $\mu = 0$; $\mu = 0.5$; $\mu = 1$. The compromise between fast rising time and small settling time can be observed. The unshifted case ($\mu = 0$) corresponds to the Fox and Li method.

Figure 3: Field distribution at the mirrors M1 and M2 after a given number of bounces. h^0 is a constant initial field distribution launched from M2. A set is shown for each value $\mu = 0$; $\mu = 0.5$; $\mu = 1$. The convergence rate can be checked against Fig. 2.

Figure 4: As in Fig. 3 but now for a random initial field distribution of null average and a negative value of the translation parameter, $\mu = -0.5$.

TABLE CAPTIONS

Table I: Sumary of results for unstable and stable resonators showing direct comparison of computing time between Fox and Li method, ALLMAT and eigenvalue translation.

g	N	Neq	α_{00}	Time [s] Fox-Li	Time [s] ALLMAT	Time [s] Eig.Trans.
0.2	2.0	–	+0.99494 -i0.10050	>2400	500	100
0.8	2.0	–	+0.89430 -i0.44716	820	500	60
1.4	3.0	3,0	+0.55979 -i0.61541	400	500	80

Table 1.

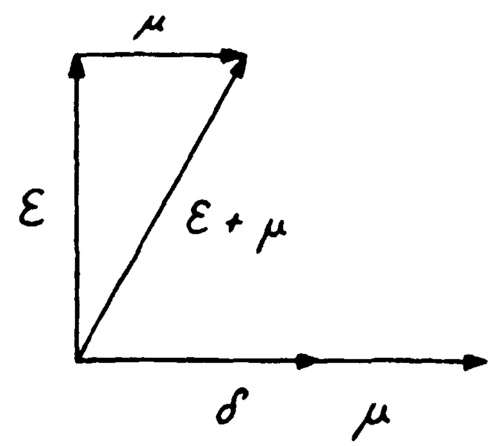

Figure 1

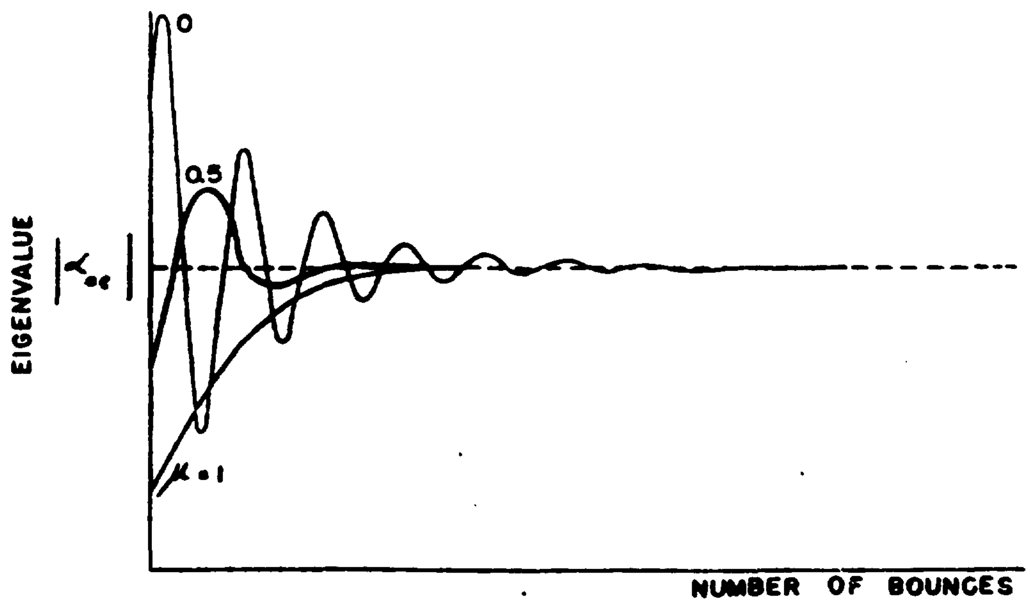

FIGURE-2

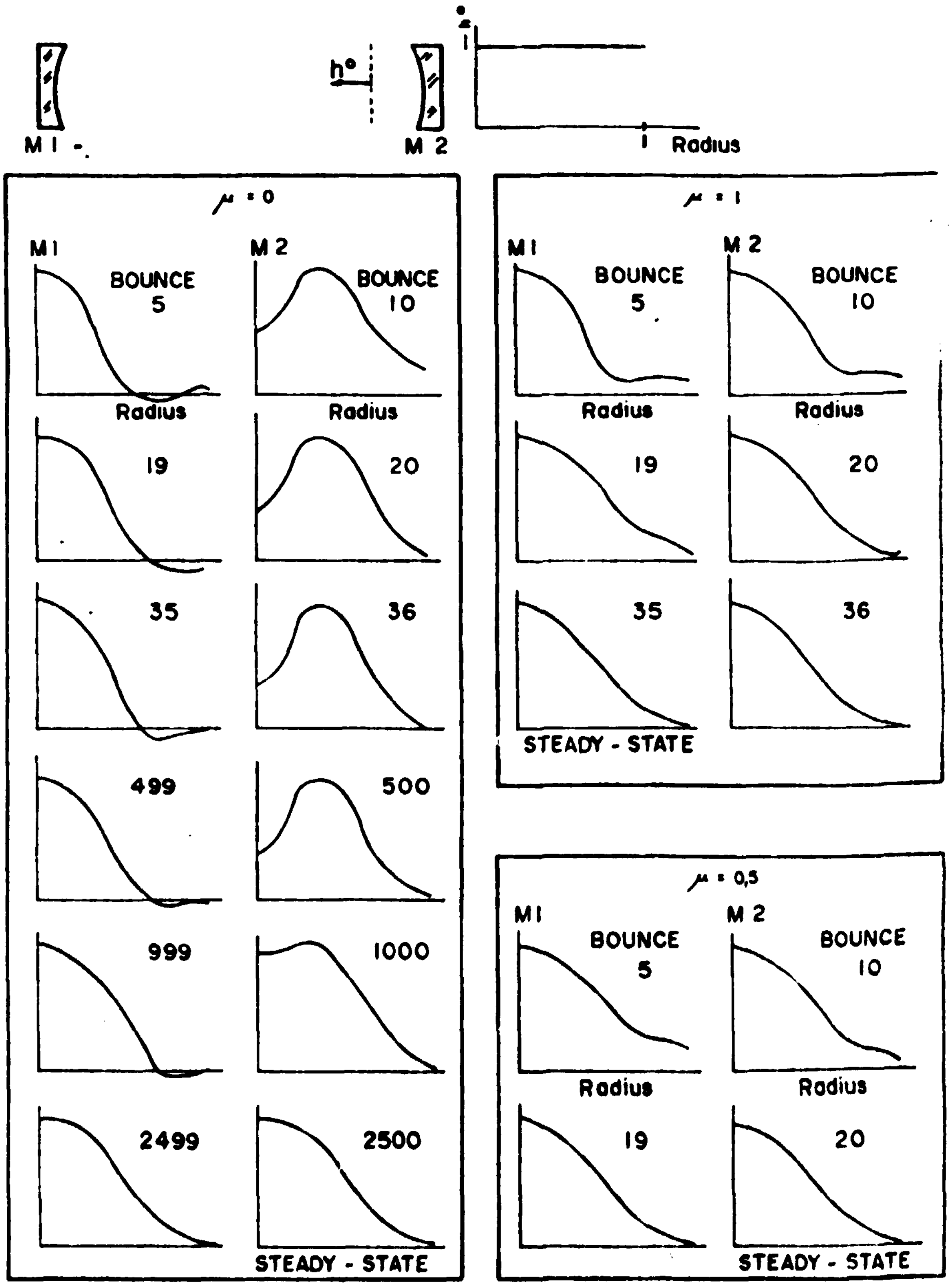
M 1
h°
M 2
1
1
Radius
μ = 0
M 1
BOUNCE
5
M 2
BOUNCE
10
Radius
Radius
19
20
35
36
499
500
999
1000
2499
2500
STEADY - STATE
μ = 1
M 1
BOUNCE
5
M 2
BOUNCE
10
Radius
Radius
19
20
35
36
STEADY - STATE
μ = 0,5
M 1
BOUNCE
5
M 2
BOUNCE
10
Radius
Radius
19
20
STEADY - STATE

FIGURE - 3

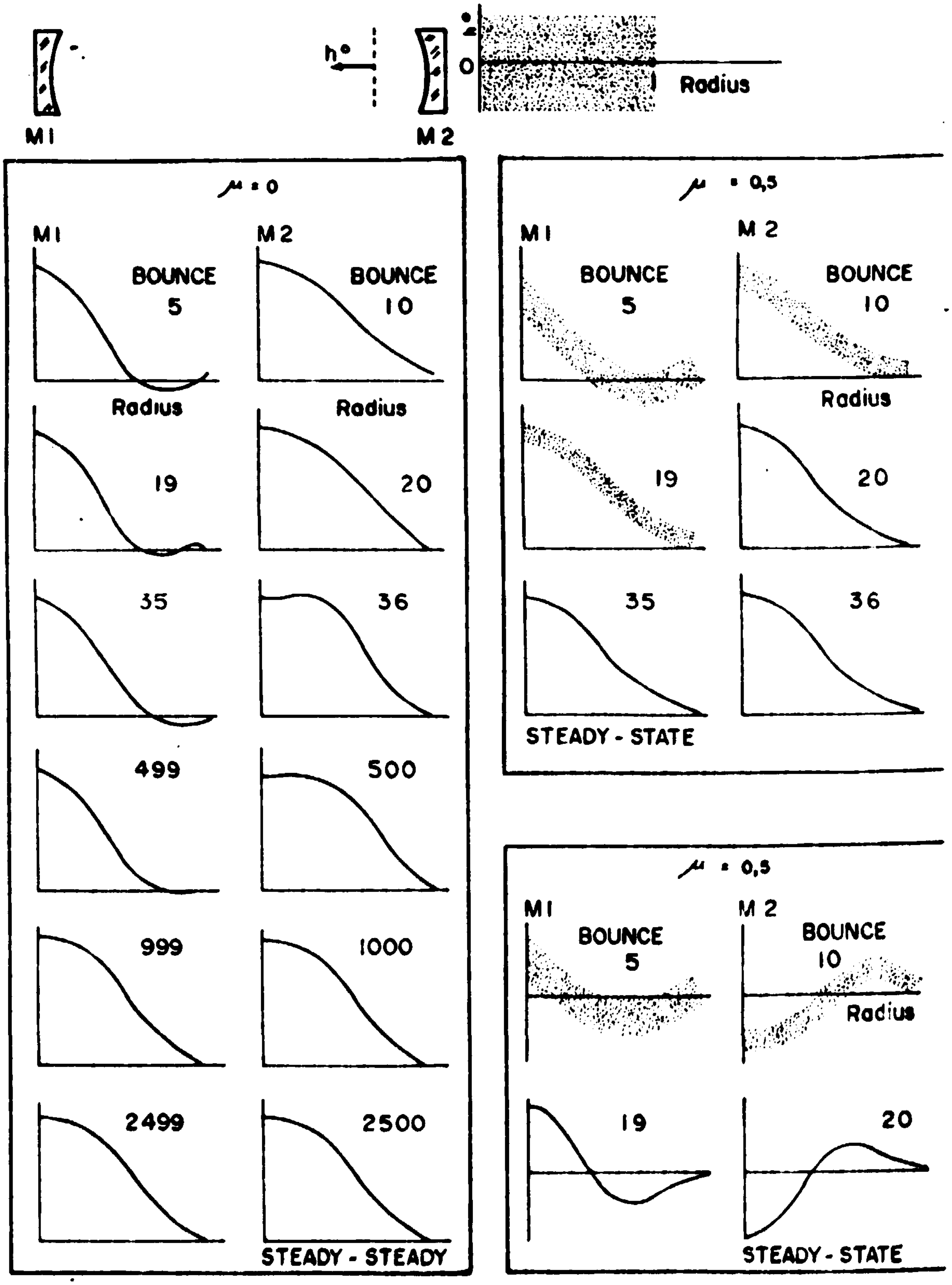
M1
M2
h°
0
Radius
μ = 0
M1
M2
BOUNCE 5
BOUNCE 10
Radius
Radius
19
20
35
36
499
500
999
1000
2499
2500
STEADY - STEADY
μ = 0,5
M1
M2
BOUNCE 5
BOUNCE 10
Radius
19
20
35
36
STEADY - STATE
μ = 0,5
M1
M2
BOUNCE 5
BOUNCE 10
Radius
19
20
STEADY - STATE

FIGURE - 4