

## Reduction of model in linear and non-linear dynamics: Method of RITZ

---

### Summary:

This document presents the principle of reduction of model by projection on reduced basis (method of Ritz). The base most usually used is the modal base.

The problems of truncation due to the use of a reduced base are mentioned. Corrections of truncation are proposed.

The description and the properties of the algorithms of resolution of the system of differential equations of the second order obtained in transitory analysis are presented in the document [R5.06.04].

## Contents

---

1 Introduction.....	3
2 Methods of reduction of Ritz into linear.....	3
2.1 General description.....	3
2.1.1 Continuous formulation.....	3
2.1.2 Approximation finite elements.....	3
2.2 Projection on reduced basis.....	4
2.3 Projection on modal basis.....	4
2.4 Modal truncation error.....	6
2.5 Corrections of modal truncation.....	8
2.5.1 Static correction a posteriori.....	8
2.5.2 Addition of static modes to the base.....	9
3 Extension of the methods of linear reduction of Ritz in non-.....	10
3.1 Problem general.....	10
3.2 Indication of the error of projection.....	10
4 Use in Code_Aster.....	11
5 Bibliography.....	12
6 Description of the versions of the document.....	12

## 1 Introduction

Starting from a description of the geometry and materials of the structures, the finite element method makes it possible to create a precise and reliable but large-sized model. In the case of a problem of dynamics, one wishes to calculate the answer of a system for various moments (analyzes transitory) where for various frequencies (analyzes harmonic). The size of the model finite elements obtained is often irreconcilable with the number of calculations necessary to get all the desired results.

For a restricted whole of dynamic stresses, there exists generally a subspace of low dimension making it possible to describe the dynamic behavior of the structure under specific requests.

The projection of the model on a restricted basis is called method of Ritz or Rayleigh-Ritz.

This document comprises the following points:

- a presentation of the methods of Ritz, their use into linear,
- a detail of the possible corrections of truncation,
- generalization into nonlinear of the methods of Ritz,
- two simple examples of illustration.

## 2 Methods of reduction of Ritz into linear

### 2.1 General description

#### 2.1.1 Continuous formulation

The method of Ritz consists in projecting displacement on a restricted basis of functions checking the conditions kinematics of the problem:

$$\tilde{u}(M, t) = \sum_{i=1}^n \eta_i(t) \psi_i(M) \quad \text{éq 2.1.1-1}$$

Displacement is described by a series of independent forms  $\{\psi_i(M); i=1 \dots n\}$  multiplied by amplitudes functions of time  $\{\eta_i(t); i=1 \dots n\}$ .

The difficulty consists in defining this family of form  $\{\psi_i(M); i=1 \dots n\}$  who contrary to the functions of form of the finite element method are nonworthless on most of the structure.

The quality of the approximation is related to the fact that displacements obtained have a good approximation in the subspace generated by  $\text{Vect}\{\psi_i(M), i=1 \dots n\}$ .

#### Projection on modal basis

It is known that the clean modes  $\{\phi_i(M); i=1 \dots \infty\}$  generate the space of the fields kinematically acceptable. Displacement breaks up according to:

$$u(M, t) = \sum_{i=1}^{\infty} \eta_i(t) \phi_i(M) \quad \text{éq 2.1.1-2}$$

The option most usually used for the method of Ritz then consists in taking as bases projection them  $n$  first modes:

$$\tilde{u}(M, t) = \sum_{i=1}^n \tilde{\eta}_i(t) \phi_i(M) \quad \text{éq 2.1.1-3}$$

Displacement obtained is an approximation of real displacement.

It can be interesting to add to  $n$  first modes, of other forms (see [§2.6.2]).

#### 2.1.2 Approximation finite elements

In the case of an approximation of displacement by finite elements displacement is already approximate within the space of functions of forms:

$$u^h(M, t) = \sum_{i=1}^{Nh} q_i(t) N_i(M) \quad \text{éq 2.1.2-1}$$

One notes  $U$  the vector of the degrees of freedom of displacement:  $U(t) = [q_1(t), q_2(t), \dots, q_{Nh}(t)]$  ;

### Method of Ritz in finished dimension

If  $n < Nh$ , method of Ritz applied to the field  $u(M, t)$  comes then like a second approximation:

$$\tilde{U}(t) = \sum_{i=1}^n \eta_i(t) \psi_i \quad \text{éq 2.1.2-2}$$

with  $\{\psi_{i, i=1 \dots n}\}$  the base of  $n$  vectors independent and kinematically acceptable. One poses

$$\Psi = [\psi_1, \psi_2, \psi_3, \dots, \psi_n]. \text{ From where the matrix writing: } U = \Psi \eta \quad \text{éq 2.1.2-3}$$

## 2.2 Projection on reduced basis

Let us consider the following differential connection obtained by a method finite elements:

$$\mathbf{M} \ddot{U} + \mathbf{C} \dot{U} + \mathbf{K} U = \mathbf{F} \quad U \in R^{Nh} \quad \text{éq 2.2-1}$$

The solution sought in the form [éq 2.1.2-3]. While considering the same form for virtual displacement, it comes:

$$\Psi^T \mathbf{M} \psi \ddot{\eta} + \Psi^T \mathbf{C} \Psi \dot{\eta} + \Psi^T \mathbf{K} \Psi \eta = \Psi^T \mathbf{F} \quad \eta \in R^n \quad \text{éq 2.2-2}$$

where:  $\eta$  is the vector of generalized displacements,  $\bar{\mathbf{K}} = \Psi^T \mathbf{K} \Psi$  et  $\bar{\mathbf{M}} = \Psi^T \mathbf{M} \Psi$  matrices of generalized stiffness and mass are called respectively.

The system [éq 2.2-2] is generally a coupled differential connection, the generalized matrices which compose it are in the full case general even if at the beginning the matrices  $\mathbf{M}$  and  $\mathbf{K}$  were hollow. One thus loses the structure particular to the profit of a size of problem much more reduced  $n \times n$ .

In the case general, the system [éq 2.2-2] provides only one approximate solution of the system [éq 2.2-1]. The mistake which one makes is called truncation error.

One has no information on the value of this error. It can be very large if the subspace of projection is badly selected. It is known only that this error decreases when the size of the base of projection increases.

If one has information a priori on the form of the solution, one can choose in an effective way the base of projection in order to minimize this error. For example, if it is known that the solution is made up only of movements of solid body, it can restrict to 6 the dimension of space.

Thereafter, one chooses the base of the clean modes as bases projection.

## 2.3 Projection on modal basis

### Clean modes

The modes are defined like the couples  $\{(\omega^i, \Phi^i)_{i=1 \dots Nh}\}$  solutions of the equation:

$$(K - \omega^2 M) \Phi = 0 \quad \text{éq 2.3-1}$$

**Note:**

*It is advisable to check that the modes calculated by approximation finite elements are sufficiently representative:  $(\omega^{h_i}, \Phi^{h_i}) \approx (\omega_i, \Phi_i)$ . One can consider that the approximation finite elements is correct when the modal deformations present a higher wavelength in keeping with meshes of the grid (the concept wavelength is a generalization of the notion definite on the equation of the waves, one can twice define it as the length between two nodes of the modal deformation).  
Thereafter one omits, the index  $h_i$  corresponding to the approximation finite elements.*

## Quotient of Raleigh : energy interpretation

The own pulsations and forms can be defined as the solutions of the problem of minimization according to :

$\forall i \in [1, Nh] : \Phi_i$  minimize under space  $R^{Nh} - Vect \{ \Phi_j, j \in [0, i-1] \}$  the functional calculus:

$$R(X) = \frac{X^T K X}{X^T M X} \text{ on pose: } \omega_i^2 = \frac{\Phi_i^T K \Phi_i}{\Phi_i^T M \Phi_i} = R(\Phi_i) \quad \text{éq 2.3-2}$$

## Method of reduction

A method of reduction very largely employed for the linear problems is the method of modal recombination. It consists in choosing as bases projection them  $n$  first clean modes of the structure  $\{ \Phi_{i, i=1 \dots n} \}$ .

$$\tilde{U}(t) = \sum_{i=1}^n \eta_i(t) \Phi_i \quad \text{éq 2.3-3}$$

Always let us consider the following differential connection:

$$\mathbf{M} \ddot{U} + \mathbf{C} \dot{U} + \mathbf{K} U = \mathbf{F} \quad U \in R^{Nh} \quad \text{éq 2.3-4}$$

Clean modes  $\{ \Phi_{i, i=1 \dots Nh} \}$  have the property to be  $\mathbf{M}$  and  $\mathbf{K}$  orthogonal, i.e. one has the following relations:

$$\begin{aligned} \Phi_i^T \mathbf{M} \Phi_j &= m_i \delta_{ij} \\ \Phi_i^T \mathbf{K} \Phi_j &= k_i \delta_{ij} \end{aligned}$$

$\delta$  is the symbol of KRONECKER

$m_i$  modal mass or generalized mass of the mode is called  $i$

$k_i$  modal rigidity or generalized rigidity of the mode is called  $i$

Matrices projected of  $\mathbf{M}$  and  $\mathbf{K}$  on the basis of clean mode are thus diagonal; it is one of the advantages which justified the use of the modal base as bases projection. The system [éq 2.3 - 4] project on the basis of clean mode first of the system is written:

$$\begin{pmatrix} \backslash & 0 & 0 \\ 0 & m_i & 0 \\ 0 & 0 & \backslash \end{pmatrix} \ddot{\eta} + \Phi^T \mathbf{C} \Phi \dot{\eta} + \begin{pmatrix} \backslash & 0 & 0 \\ 0 & k_i & 0 \\ 0 & 0 & \backslash \end{pmatrix} \eta = \Phi^T \mathbf{F}_{ext} \quad \text{éq 2.3-5}$$

The projection of the matrix  $\mathbf{C}$  has no reason in any general information to be also diagonal. If the system is strongly deadened (presence of shock absorbers on the structure), this matrix will not be diagonal.

### Note:

*As opposed to what do many software, Code\_Aster allows in this case to integrate the system of modal equations coupled without diagonalisation of the matrix of generalized damping. The method of integration is in this case an implicit method of NEWMARK or clarifies EULER.*

On the other hand, if only damping entering concerned is a structural damping (internal dissipation of material for a homogeneous structure) it is then licit to make the assumption of a damping proportional, still called assumption of BASILE, in this case  $\mathbf{C}$  express yourself like linear combination of  $\mathbf{M}$  and  $\mathbf{K}$  (damping of RAYLEIGH), and its projection on the clean modes is diagonal (cf Doc. [R5.05.04] on the modeling of damping).

In this case, the system [éq 2.3-4] is divided into p linear differential equations of the second order uncoupled. The answer of the system is then the recombination of the response of p oscillators simple associates to the clean modes, from where the expression of "modal superposition" used usually.

Each differential equation is written  $m_i$  :

$$m_i \ddot{\eta}_i + c_i \dot{\eta}_i + k_i \eta_i = f_i \quad \text{éq 2.3-6}$$

or while dividing by the modal mass:

$$\ddot{\eta}_i + 2 \xi_i \omega_i \dot{\eta}_i + \omega_i^2 \eta_i = \frac{f_i}{m_i} \quad \text{éq 2.3-7}$$

with:

$$\xi_i \text{ amortissement modal réduit} = \frac{c_i}{c_{\text{critique}}} = \frac{c_i}{2 m_i \cdot \omega_i}$$

This equation can be solved very simply in the frequential field:

$$\hat{\eta}_i = \frac{\hat{f}_i(\omega)}{m_i \cdot (-\omega^2 + 2 \cdot \xi_i \omega_i \omega + \omega_i^2)} \quad \text{éq 2.3-8}$$

where  $\hat{\quad}$  represent the transform of FOURIER and  $\omega$  the frequency of excitation.

Particular digital methods the such integral of DUHAMEL make it possible to pass this expression of the frequential field to the temporal field. (see for example Doc. [R5.05.01] on a method of temporal integration).

## 2.4 Modal truncation error

In the case of the modal recombination with damping proportional, one can highlight the truncation error which one makes while projecting on the basis of clean mode first of the system. Indeed, if one considers the complete base of N clean modes of the discretized problem, there is equivalence between the initial problem and the problem project. Thus the exact solution of the problem discretized by finite elements is written:

$$U = \sum_i^{Nh} \eta_i \Phi_i$$

where the generalized coordinates are solution of:

$$\ddot{\eta}_i + 2 \xi_i \omega_i \dot{\eta}_i + \omega_i^2 \eta_i = \frac{f_i}{m_i}$$

summation extending on **all clean modes** system (of finished size).

By solving the problem with a reduced number of clean modes,  $n < Nh$  . The solution obtained is the following one:

$$\tilde{U} = \sum_{i=1}^n \eta_i \Phi_i$$

The mistake made by truncating the base of representation of the solution is thus:

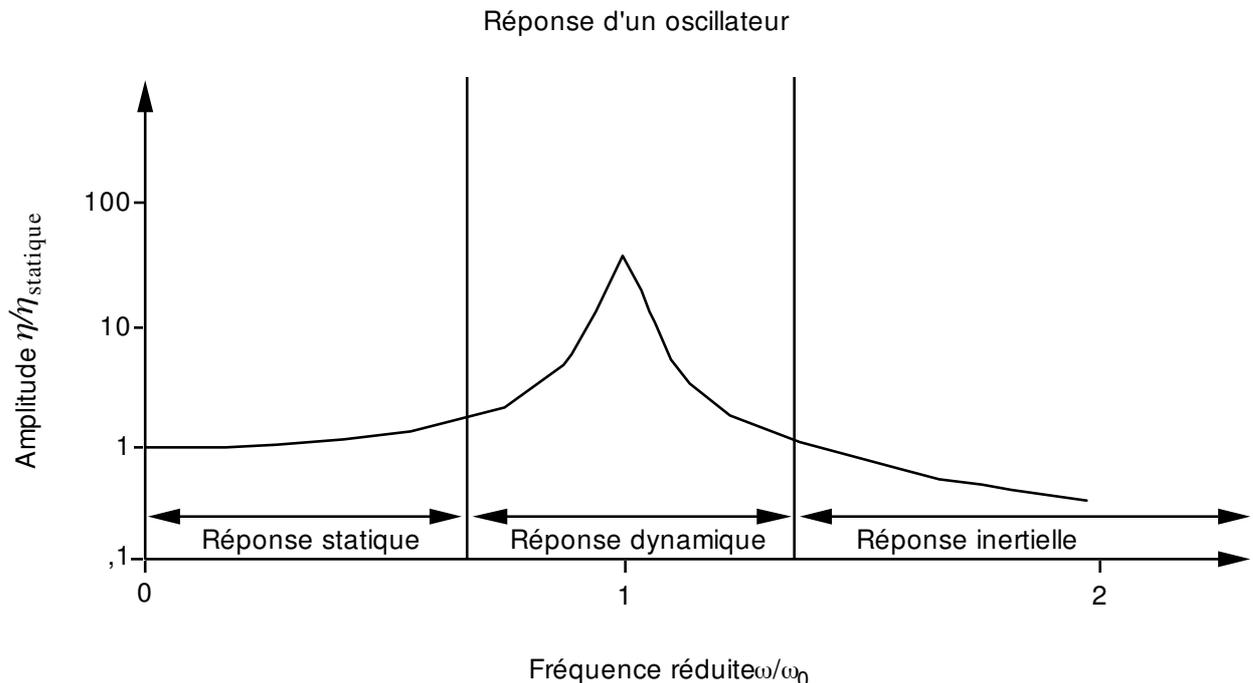
$$\mathbf{E} = U - \tilde{U} = \sum_{i=n+1}^{Nh} \eta_i \Phi_i \quad \text{éq 2.4-1}$$

In the frequential field the expression of the error is:

$$\hat{\mathbf{E}}_{(\omega)} = \hat{U} - \hat{\tilde{U}} = \sum_{i=n+1}^{Nh} \frac{\Phi_i^t \hat{\mathbf{F}}_{(\omega)}}{m_i} \cdot \frac{1}{\omega_i^2 - \omega^2 + 2j\xi_i \omega \omega_i} \cdot \Phi_i \quad \text{éq 2.4-2}$$

the summation is carried out on all the neglected modes of the system.

Let us study the relative answer  $\eta/\eta_{statique}$  of an oscillator to a purely sinusoidal excitation of variable frequency (diagram below), with  $\eta_{statique}$  coefficients of the static answer corresponding to a static force. One can distinguish three intervals in the spectrum where the oscillator has a different behavior. In low frequency ( $\omega \ll \omega_0$ ) the oscillator has a static answer. Around  $\omega_0$  the oscillator has a dynamic response (amplification of the mode), and high frequency the oscillator answers in an inertial way ( $\frac{1}{\omega^2}$  dominating term).



Let us suppose that the excitation of the system, defined by the vector  $\mathbf{F}(\omega)$ , is with narrow band, in particular that it is worthless for frequencies higher than  $\omega_{max}$  given.

In this case, to represent the answer of the linear system correctly, it is necessary undoubtedly to take into account all the modes having a pulsation lower than  $\omega_{max}$ , because the latter will answer in a dynamic way the excitation.

On the other hand, modes such as  $\omega_i \gg \omega_{max}$  nevertheless have a static contribution to the answer of the system. These are often these modes that one does not take into account.

By making a development limited in  $\omega$  in the vicinity of 0 . One obtains the principal part of the error which is:

$$\hat{\mathbf{E}}_{(\omega)} = \hat{U} - \tilde{U} = \sum_{i=p+1}^n \frac{\Phi_i^T \cdot \hat{\mathbf{F}}_{(\omega)}}{k_i} \cdot \left( 1 - 2j\xi_i \frac{\omega}{\omega_i} + 0 \left( \frac{\omega}{\omega_i} \right) \right) \cdot \Phi_i \quad \text{éq 2.4-3}$$

The error is all the more small as generalized rigidities of the neglected modes are large. In theory thus, it will be necessary to take all the most flexible modes until the residual flexibility of an additional mode is in negligible relative value compared to the sum of the flexibilities already taken into account:

$$\frac{1}{k_{n+1}} \ll \sum_{i=1}^n \frac{1}{k_i}$$

However, it is observed that by neglecting the modes of high frequency one makes a systematic mistake on the answer of the system (even in low frequency). There exist various possibilities that we will detail now to correct the answer in the beach  $[0, \omega_{\max}]$  where the modes were chosen.

## 2.5 Corrections of modal truncation

To mitigate the problem of truncation due to the neglected modes, it is necessary to try to estimate their effect in the field of frequency  $[0, \omega_{\max}]$  who interests us. We saw that the neglected modes having an own pulsation such as  $\omega_i \gg \omega_{\max}$  have a contribution known as static to the answer of the system in the field  $[0, \omega_{\max}]$ . The techniques of correction consist in calculating this static contribution.

### 2.5.1 Static correction a posteriori

The truncation error, by considering only the static answer of the neglected modes (transformed opposite of the principal part of the error) is:

$$\mathbf{E}_{(t)} = U - \tilde{U} \approx \sum_{i=n+1}^{Nh} \frac{\Phi_i^T \cdot \mathbf{F}_{(t)}}{k_i} \cdot \Phi_i \quad \text{éq 2.5.1-1}$$

But a priori the neglected modes as their generalized rigidities are unknown. On the other hand, one can determine the complete static response of the system to a loading  $\mathbf{F}$ , the latter is worth:

$$U = \mathbf{K}^{-1} \cdot \mathbf{F}_{(t)} \approx \sum_{i=1}^{Nh} \frac{\Phi_i^T \cdot \mathbf{F}_{(t)}}{k_i} \cdot \Phi_i$$

The correction which should be made is thus:

$$\sum_{i=n+1}^{Nh} \frac{\Phi_i^T \cdot \mathbf{F}_{(t)}}{k_i} \cdot \Phi_i \approx \mathbf{K}^{-1} \cdot \mathbf{F}_{(t)} - \sum_{i=1}^n \frac{\Phi_i^T \cdot \mathbf{F}_{(t)}}{k_i} \cdot \Phi_i$$

The corrected solution of the answer of the system is thus worth:

$$U = \tilde{U} + \mathbf{E} \approx \tilde{U} + \mathbf{K}^{-1} \cdot \mathbf{F}_{(t)} - \sum_{i=1}^n \frac{\Phi_i^T \cdot \mathbf{F}_{(t)}}{k_i} \cdot \Phi_i \quad \text{éq 2.5.1-2}$$

This correction is called a posteriori, because it does not intervene in the dynamic resolution of the linear system and can be calculated only well later on. If  $\mathbf{F}_{(t)}$  breaks up into  $k$  produced functions of time by functions of the coordinates of space, this correction requires a factorization of  $\mathbf{K}$  and  $k$  resolutions.

This method has the advantage of not increasing the number of vectors taken into account in the base. This method is applicable in the case of an excitation to narrow band, or at least having a known cut-off frequency. The correction is exact in the low frequency field but can distort the answer of the system in high frequency [§Annexe1].

## 2.5.2 Addition of static modes to the base

Let us suppose that the loading  $\mathbf{F}(t)$  is written:

$$\mathbf{F}(t) = \sum_j \alpha_j(t) \cdot \mathbf{F}_j$$

The second way of correcting the truncation error consists in adding to the base of the initial clean modes of **static modes**  $\Psi_j$  defined as the deformation in each effort  $\mathbf{F}_j$  given:

$$\Psi_j = \mathbf{K}^{-1} \cdot \mathbf{F}_j \quad \text{éq 2.5.2-1}$$

The new base of projection to be considered is the following one:

$$\hat{\Phi} = [\Phi_1, \Phi_2, \dots, \Phi_p, \Psi_1, \Psi_2, \dots, \Psi_m] = [\Phi \Psi] \quad \text{éq 2.5.2-2}$$

The components generalized to use are the following ones:

$$\hat{\eta} = [\eta_1, \eta_2, \dots, \eta_p, \mu_1, \mu_2, \dots, \mu_m] = [\eta, \mu] \quad \text{éq 2.5.2-3}$$

The problem project on the supplemented basis is:

$$\begin{bmatrix} \begin{pmatrix} \backslash & 0 & 0 \\ 0 & m_i & 0 \\ 0 & 0 & \backslash \end{pmatrix} & \Phi^T \cdot \mathbf{M} \cdot \Psi \\ \Psi^T \cdot \mathbf{M} \cdot \Phi & \Psi^T \cdot \mathbf{M} \cdot \Psi \end{bmatrix} \begin{Bmatrix} \ddot{\eta} \\ \ddot{\mu} \end{Bmatrix} + \begin{bmatrix} \begin{pmatrix} \backslash & 0 & 0 \\ 0 & k_i & 0 \\ 0 & 0 & \backslash \end{pmatrix} & \Phi^T \cdot \mathbf{K} \cdot \Psi \\ \Psi^T \cdot \mathbf{K} \cdot \Phi & \Psi^T \cdot \mathbf{K} \cdot \Psi \end{bmatrix} \begin{Bmatrix} \eta \\ \mu \end{Bmatrix} = \begin{Bmatrix} \Phi^T \cdot \mathbf{F} \cdot \alpha \\ \Psi^T \cdot \mathbf{F} \cdot \alpha \end{Bmatrix} \quad \text{éq 2.5.2-4}$$

It is noted that one lost the diagonal character of the generalized matrices, but the advantage obtained is that the base supplemented with static modes makes it possible to correctly represent the low frequency behavior of the initial system.

For example it is simple to show that at worthless frequency the solution of this system is:

$$\eta = 0 \quad \text{and} \quad \mu = \alpha \quad \text{who is the exact solution of the initial static problem.}$$

One presents in appendix 1, the comparison on a discrete system with 3 degrees of freedom between the exact solution, the solution projected on 1 mode, that projected on a mode with a static correction and the solution consisted 1 clean mode and 1 static mode.

One realizes that the addition of static modes makes it possible to extend beyond the interval  $[0, \omega_{\max} = \max(\omega_j)]$  the good dynamic representation of the system. This technique thus seems very interesting, it has the virtue to carry out the correction immediately what will be interesting for the nonlinear methods where one needs the knowledge of physical displacements to each step of time.

Implementation the practical of this technique is specified in [U2.06.04].

## 3 Extension of the methods of linear reduction of Ritz in non-

### 3.1 Problem general

The non-linear problem of dynamics discretized without damping can be generally put in the following form:

$$\mathbf{M} \cdot \ddot{\mathbf{X}} + \mathbf{G}(\mathbf{X}) = \mathbf{F}(t) \quad \mathbf{X} \in R^n \quad \text{éq 3.1-1}$$

$\mathbf{G}(\mathbf{X})$  is a non-linear function of  $\mathbf{X}$  who represents the internal forces of the system as all the other forces which are dependent on displacement,  $\mathbf{F}$  the vector of the external forces and  $\mathbf{M}$  the matrix of mass of the system.

The matrix of tangent rigidity system is by definition:

$$\mathbf{K}_{(\mathbf{x})}^{tg} = \frac{\partial \mathbf{G}}{\partial \mathbf{X}}(\mathbf{x}) \quad \text{éq 3.1-2}$$

It makes it possible to define a modal base at every moment by:

$$\left( -\omega_{i(\mathbf{x})}^{tg^2} \mathbf{M} + \mathbf{K}_{i(\mathbf{x})}^{tg} \right) \cdot \Phi_{i(\mathbf{x})}^{tg} = 0 \quad \text{éq 3.1-3}$$

The modes thus defined depend on  $\mathbf{X}$ , therefore moment  $t$ .

Knowing that the calculation of the modal base is very expensive in computing times, the idea to want to project with each step of time the model on a modal basis, then to solve, is irrelevant compared to a direct resolution.

The method most usually used consists in defining a base of projection while adding to the modes calculated on an initial configuration of the forms making it possible to project to it not linearity.

Example: if it not linearity comes from a specific shock, one proposes to enrich the modal base with static modes allowing to project the effort undergone by the structure lasting the shock [R5.06.04].

The method of Ritz remains always relevant in nonlinear calculations, if the selected base makes it possible to correctly project displacements and the efforts.

The nonlinear problem project on a basis  $\Psi$  unspecified is written:

$$\Psi^T \mathbf{M} \Psi \cdot \ddot{\eta} + \Psi^T \mathbf{G}(\mathbf{X}) = \Psi^T \mathbf{F} \quad \eta \in R^n \quad \text{éq 3.1-4}$$

Two possibilities are then possible:

- nonthe linearities are located and one can evaluate to it not linearity on the basis of projection:  $\mathbf{G}(\mathbf{X}) = \mathbf{G}(\Psi \eta)$ .
- The problem to be solved is a nonlinear differential connection in  $\eta$  of smaller size. Various strategies are possible to solve this problem, depending primarily on the technique of integration which one wishes to use.
- nonthe linearities are total, and it is necessary to pass by again within the space of physical degrees of freedom to calculate the internal forces:  $\mathbf{G}(\mathbf{X})$ .

This second method is more expensive it is much less current.

One presents in appendix 2, the comparison on a system with 3 degrees of freedom with nona linearity in  $x^3$  between the exact solution and the solution obtained by the method above with 1 then 2 modes. One realizes that it is necessary to take more modes counts some than for the linear problem. On the other hand, on this example 2 modes are enough very well to describe the system.

### 3.2 Indication of the error of projection

*Warning : The translation process used on this website is a "Machine Translation". It may be imprecise and inaccurate in whole or in part and is provided as a convenience.*

Copyright 2019 EDF R&D - Licensed under the terms of the GNU FDL (<http://www.gnu.org/copyleft/fdl.html>)

For the nonlinear problems the physical direction amongst modes to be taken into account is completely lost, and if the methods of reduction always give a solution it is necessary to know the degree of confidence which one can grant to them. A way of proceeding, which is a little expensive but essential is to calculate the residue of the initial system to each step of time. It is defined by:

$$\mathbf{R} = \mathbf{M} \cdot \ddot{\mathbf{X}} + \mathbf{G}(\mathbf{X}) - \mathbf{F}(t)$$

This vector residue is unfortunately not null, it is only its projection on the basis used which is it.

A standard can then be calculated for this residue; more the standard of the residue will be small more one will be able to grant confidence to the solution.

To use a relative value, one may find it beneficial to calculate the following fraction:

$$r = \frac{\|\mathbf{R}\|}{\max(\|\mathbf{F}(t)\|, \|\mathbf{G}(\mathbf{X})\|, \|\mathbf{M} \cdot \ddot{\mathbf{X}}\|)} \quad \text{éq 3.2-1}$$

**Note:**

*This indicator is not currently established in Code\_Aster.*

## 4 Use in Code\_Aster

In *Code\_Aster*, the methods of Ritz are usable in transient primarily by the operator `DYNA_TRAN_MODAL` [U4.53.21].

A phase of projection of the matrices of rigidity and mass on a basis of vectors is carried out by the operators `PROJ_MATR_BASE` [U4.63.12] and `PROJ_VECT_BASE` [U4.63.13].

The generalized dynamic problem is then solved in the operator `DYNA_TRAN_MODAL` by a diagram of explicit integration (EULER or DEVOGELEARE) or implicit (NEWMARK). The characteristics and properties of the diagrams of integration are presented in the note [R5.06.04]. For the structures for which the assumption of BASILE does not apply (damping nonproportional) one will project also the matrix of damping which does not become diagonal. The integration of the coupled system is done then obligatorily with the implicit scheme (NEWMARK) or clarifies (EULER).

Localised non-linearities are specified directly in the operator `DYNA_TRAN_MODAL`. One can introduce localised non-linearities of the standard shock and friction (see [R5.06.03] Modeling of the shocks and frictions), modal forces function of displacement or speed.

The static corrections of truncation a posteriori are available in the case of a single excitation (see [R4.05.01] seismic Answer).

The addition of static modes can be done by using the operators as a preliminary `MODE_STATIQUE` [U4.52.14] and `DEFI_BASE_MODAL` [U4.64.02]. When the problem comprises non-linearities only the explicit diagrams can be used.

For total non-linearities [éq 3.1-4], it is possible to use the order `DYNA_NON_LINE` [U4.53.01] with the operand `PROJ_MODAL`, who allows to calculate with each step of time the internal forces according to the physical degrees of freedom, then to project the problem on a modal basis.

An operation of return to the physical base is then necessary to obtain the physical sizes such as displacement, speed or acceleration on the structure. This operation is carried out by the operator `REST_GENE_PHYS` [U4.63.31] for the whole of the structure or `RECU_FONCTION`, (key word factor `RESU_GENE`) [U4.32.01] for the follow-up of a size in a node. In the case of them total non-linearities (calculation with `DYNA_NON_LINE`) the operator is used `REST_COND_TRAN` [U4.63.33].

More generally, the approach of Ritz can be used in harmonic calculation by the order `DYNA_LINE_HARM` [U4.53.22] and of spectral concentration of power by the order `DYNA_ALEA_MODAL` [U4.53.23].

Finally the dynamic under-structuring can be regarded as a method of Ritz specific [R4.06.02].

## 5 Bibliography

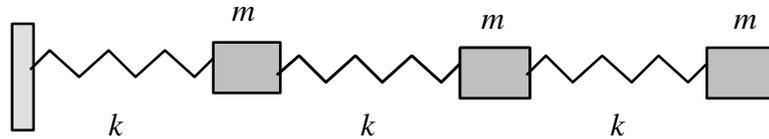
- [1] BATHE - WILSON: Finite Element procedures in Analysis Engineering
- [2] MR. GERADIN, D. RIXEN: Theory of the vibrations: application to the dynamics of the structures – Masson 1993
- [3] J.F.IMBERT: Analyses of the structures by finite elements - cepadues editions 1979
- [4] BELYTSCHKO - LIU - PARK: Innovative methods for not linear problems
- [5] EWINS D.J. Modal Testing: Theory and practice Reserch Studies Close LTD
- [6] R.E. RICKELL: Non-linear dynamics by mode superposition Methods Computer in Applied Mechanics and Engineering (1976) flight 7
- [7] P. LUKKUNAPRASIT: Dynamic answer of year elastic, viscoplastic system in modal coordinates. Earthquake Engineering & Structural Dynamics (1980)
- [8] G. JACQUART: Methods of RITZ in non-linear dynamics - Application to systems with shock and friction localised - Report EDF-DER HP-61/91.105

## 6 Description of the versions of the document

Version Aster	Author (S) Organization (S)	Description of the modifications
6	G. JACQUART (EDF R & D /AMV)	Initial text
7,4	L.RATIER, G. JACQUART (EDF R & D /AMA, EDF/CNPE of Tricastin)	

## Annexe 1

Let us consider the discrete system with three masses according to:



The matrices of rigidity and mass are:

$$\mathbf{M} = \begin{pmatrix} m & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & m \end{pmatrix} \quad \mathbf{K} = \begin{pmatrix} k & -k & 0 \\ -k & 2k & -k \\ 0 & -k & 2k \end{pmatrix}$$

That is to say:  $\omega_0^2 = \frac{k}{m}$

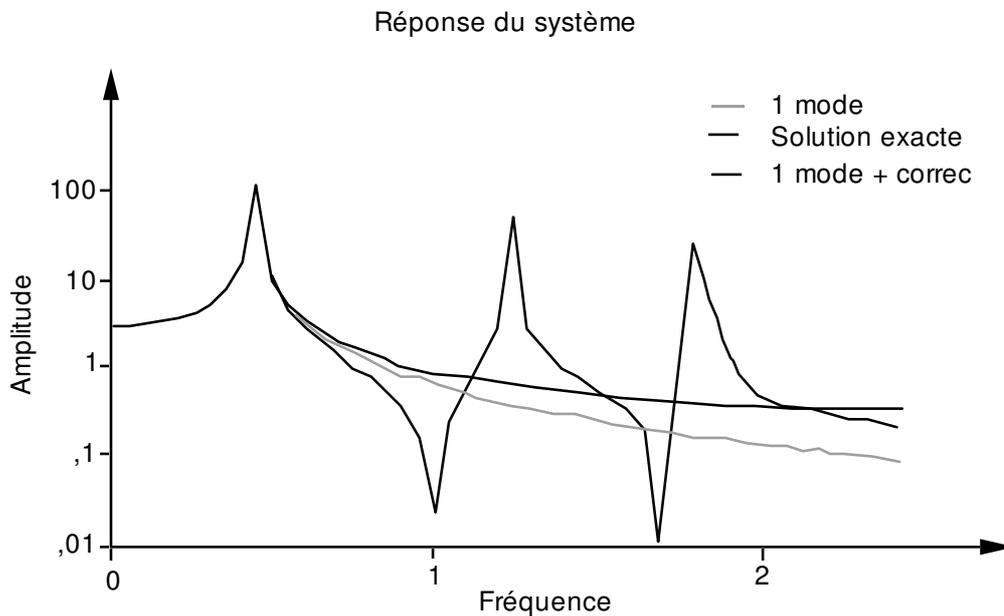
The clean modes and their pulsation are worth:

$$\omega_1^2 = 0,198 \omega_0^2, \quad m_1 = 1,841 \quad m, \quad \Phi_1 = \begin{pmatrix} 1 \\ 0,802 \\ 0,445 \end{pmatrix}$$

$$\omega_2^2 = 1,555 \omega_0^2, \quad m_2 = 2,863 \quad m, \quad \Phi_2 = \begin{pmatrix} 1 \\ -0,555 \\ -1,247 \end{pmatrix}$$

$$\omega_3^2 = 3,247 \omega_0^2, \quad m_3 = 9,296 \quad m, \quad \Phi_3 = \begin{pmatrix} 1 \\ -2,247 \\ 1,802 \end{pmatrix}$$

Let us compare the answers of the system modelled by only one clean mode with or without static correction:



It is noted that the static correction makes it possible to correct the low frequency answer, the model with 1 mode plus correction sticks perfectly to the exact solution in low frequency. On the other hand, in high frequency (beyond the first mode), this correction results in over-estimating the answer enormously. The use of the static correction will have to thus be used with prudence and within the framework of an excitation with narrow band.

Let us look at what the method of addition of a static mode gives.

If one applies a unit force to item 1, the static deformation is worth:

$$\Psi_s = \frac{1}{k} \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$$

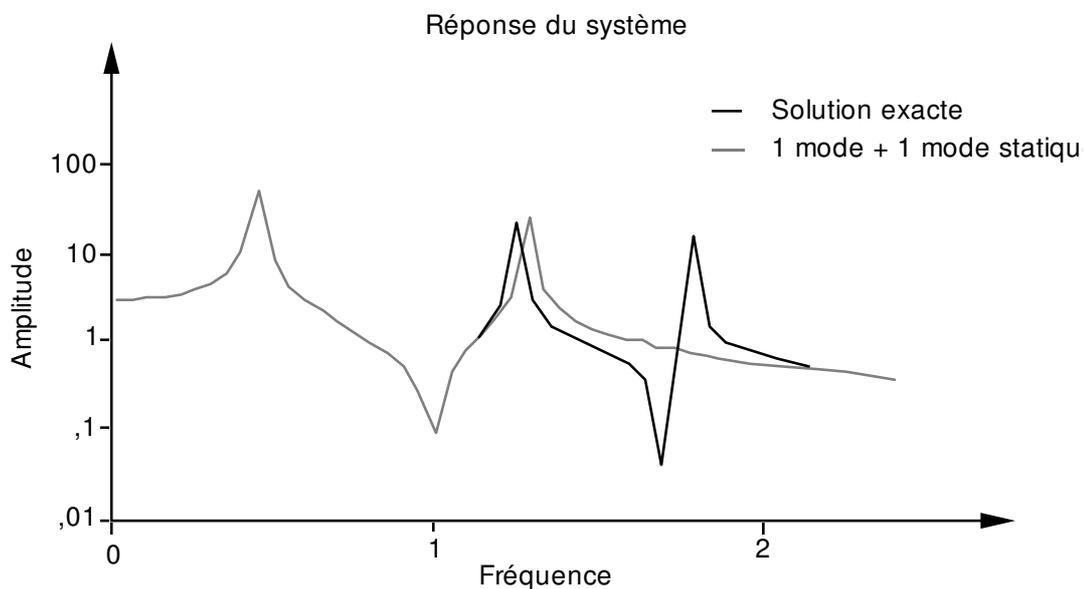
The projected matrices of mass and rigidity which one obtains are the following ones:

$$\hat{\mathbf{M}} = \begin{pmatrix} 1,841 m \frac{5,049}{\omega_0^2} & \\ & 14 \frac{m}{k^2} \end{pmatrix} \quad \text{et} \quad \hat{\mathbf{K}} = \begin{pmatrix} 0,365 k & 1 \\ 1 & \frac{3}{k} \end{pmatrix}$$

this system has as Eigen frequencies:

$$\omega_1^2 = 0,198 \omega_0^2 \quad \text{and} \quad \omega_2^2 = 1,667 \omega_0^2$$

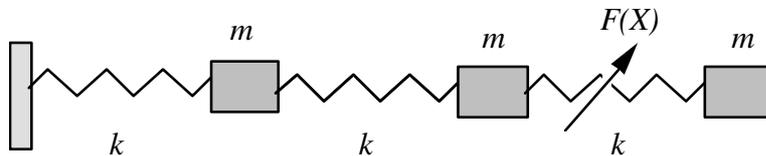
The answer of the system modelled with a clean mode and a static mode is the following one:



One realizes that one corrects very well in low frequency, (effect of correction static), one models well the dynamics of the system beyond the first mode taken into account. On the other hand the effect of the second mode is badly represented (shift on the frequency).

## Annexe 2

Let us consider the discrete system with three masses according to:



The matrices of stiffness and mass are:

$$\mathbf{M} = \begin{pmatrix} m & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & m \end{pmatrix} \quad \mathbf{K} = \begin{pmatrix} k & -k & 0 \\ -k & 2k & -k \\ 0 & -k & 2k \end{pmatrix}$$

Let us make this system non-linear by adding a term of force interns between  $x_1$  and  $x_2$  cubic:

$$F = k \cdot (x_1 - x_2)^3$$

Let us seek to evaluate the response of this system to a forced excitation of frequency close to the first Eigen frequency of the linear system (one chose  $\omega = 0,18\omega_0$ ), with an important amplitude  $F_m = 3 \cdot k$ .

In this configuration, the answer of the system cannot be evaluated by the answer of the linear system (the cubic term is well too important), it is necessary to implement a non-linear calculation with pseudo-forces as one showed in [§3.2].

One can see in the report [bib8] the curves of the transitory results of this method, by taking into account one or 2 clean modes of the initial linear system.

With only one clean mode, one realizes that one makes a relatively important mistake (reaching 50% sometimes), on the other hand it is satisfactory to note that the extrema of the vibrations are rather well envisaged. One could have hoped that while exciting in on this side first Eigen frequency it would have is enough to only one clean mode to model the answer of the system, one sees here that it is not the case. As one often notes it for non-linear systems, the system also answers with the surharmoniques ones of the frequency of excitation.

On the other hand, by taking 2 clean modes to model the response of this structure to 3 ddl, one gets a very satisfactory result (a few % of error on the amplitude), with the eye one has difficulty distinguishing the difference. This shows that by choosing a sufficiently rich base of projection one can thanks to a method of pseudo - forces model a dynamic system very well complexes with non-linearities.