

## **About the POD Model Reduction in Computational Mechanics for Nonlinear Continuous Dynamical Systems**

R. Sampaio<sup>1</sup> and C. Soize<sup>2</sup>

### **Summary**

An analysis of the efficiency of the reduced models constructed using the POD-basis and the LIN-basis is presented in nonlinear dynamics for continuous elastic systems discretized by the finite element method. The POD-basis is the basis constructed with the POD method while the LIN-basis is the basis derived from the generalized eigenvalue problem associated with the underlying linear conservative part of the system and usually called the eigenmodes of vibration. The efficiency of the POD-basis or the LIN-basis is related to the speed of convergence in the frequency domain of the solution constructed with the reduced model with respect to its dimension. A basis will be more efficient than another if it converges more rapidly than the other. An example is presented in order to analyze the efficiency of the POD-and LIN-bases. It is concluded that the POD-basis is not more efficient than the LIN-basis for the example treated in nonlinear elastodynamics.

### **Introduction**

The number of papers dealing with Proper Orthogonal Decomposition (POD) [1], [2] (also known as Karhunen-Loève basis (KL)) to construct reduced models has increased a lot in diverse fields. The objective of this paper is to compare the efficiency of the reduced model constructed with the POD-basis with the one constructed with the LIN-basis for nonlinear dynamics of continuous elastic systems discretized by the finite element method. We mean by POD-basis the basis constructed with the POD method. LIN-basis means the basis derived from the generalized eigenvalue problem associated with the underlying linear part of the nonlinear system and is usually called the eigenmodes of vibration. We mean by efficiency of the POD-basis or the LIN-basis the speed of convergence in the frequency domain of the solution constructed with the reduced model with respect to its dimension. In this paper we are not interested in constructing a reduced model adapted to a given excitation, which is the case if one uses the POD-basis. We are interested in constructing a reduced model as a predictive model for any excitation, which is the case if one uses the LIN-basis. In linear and nonlinear elastodynamics it is usual to use the LIN-basis to construct the reduced model to predict the response to any excitation. It should be noted that the POD-basis strongly depends on the excitation of the system while the LIN-basis does not depend and gives a reduced model valid for all excitations. Nevertheless in this paper the comparison

---

<sup>1</sup>PUC-Rio, Mechanical Engineering Department, Rio de Janeiro, RJ, Brazil

<sup>2</sup>University of Marne la Valle, Laboratoire de Mecanique, 5 bd Descartes, 77454 Marne la Vallee, France, christian.soize@univ-mlv.fr

of the efficiency of the two bases will be limited to the response to a given excitation. This paper is limited to the presentation of the finite element formulation and the computational results obtained. All the continuous and theoretical aspects are developed in [3].

### Finite Element Model

We consider a nonlinear continuous dynamical system whose finite element model is given by the following matrix equation in  $R^p$ ,

$$[M]\ddot{y}(t) + [D]\dot{y}(t) + [K]y(t) + f_{NL}(y(t), \dot{y}(t)) = f(t) \quad , \quad \forall t \in R \quad (1)$$

in which  $y(t)$  is the  $R^p$ -vector of the degrees of freedom,  $\dot{y}(t)$  and  $\ddot{y}(t)$  are the  $R^p$ -vectors of the velocities and accelerations, and where  $[M]$ ,  $[D]$ ,  $[K]$  and  $f_{NL}(y(t), \dot{y}(t))$  are the mass, damping, stiffness matrices and the nonlinear vector forces. The vector load is written as  $f(t) = a g(t) f_0$ , in which  $a$  is the amplitude and  $f_0$  is a normalized vector describing the spatial distribution of the load. The impulse  $t \mapsto g(t)$  is a square integrable real-valued function on  $R$  whose Fourier Transform  $\omega \mapsto \hat{g}(\omega) = \int_R e^{-i\omega t} g(t) dt$  has a bounded support  $\underline{B}_e \cup B_e$  with  $B_e = [\Omega_c - \Delta\Omega/2, \Omega_c + \Delta\Omega/2]$  and  $\underline{B}_e = [-\Omega_c - \Delta\Omega/2, -\Omega_c + \Delta\Omega/2]$ , in which  $\Omega_c$  is the central frequency and  $\Delta\Omega$  is the bandwidth. In addition it is assumed that  $\max_{\omega \in B} |\hat{g}(\omega)| = 1$ .

### The LIN-basis as the Normal Mode Basis

We introduce the *normal modes generalized eigenvalue problem*  $[K]z = \hat{\mu}[M]z$  for symmetric matrices related to the underlying linear part of the nonlinear dynamical system. The eigenvalues  $0 < \hat{\mu}_1 \leq \hat{\mu}_2 \leq \dots \leq \hat{\mu}_p$  are the square of the eigenfrequencies and the associated eigenvectors  $z^1, z^2, \dots, z^p$  are the eigenmodes. Let  $\langle x, y \rangle = \sum_{j=1}^p x_j y_j$  be the Euclidean inner product of  $x$  and  $y$  belonging to  $R^p$ . We then have the usual orthogonality properties  $\langle [K]z^\alpha, z^\beta \rangle = \hat{\mu}_\alpha \delta_{\alpha\beta}$  and  $\langle [M]z^\alpha, z^\beta \rangle = \delta_{\alpha\beta}$ .

### The POD-basis Constructed with the POD Method

The construction of the POD-basis is presented in the deterministic case. The stochastic case is similar and corresponds to the Karhunen-Loève decomposition, although the two terminologies are sometimes used in the literature without discrimination between the deterministic and the stochastic cases. It should be noted that for the linear case, since the mass, damping and stiffness are symmetric positive-definite matrices, an optimal basis independent of vector load  $f$  can be constructed and allows an optimal reduced model to be constructed for any vector load  $f$  (see [4], [5]). As proved in these references, this optimal basis does not coincide with the LIN-basis. For the present nonlinear case, it is important to emphasize that the POD-basis is well adapted to represent a given response due to a given vector load. The finite element mesh and the finite elements used for the discretization of the POD linear operator are the same as the finite element discretization of the

weak formulation of the nonlinear boundary value problem. The finite element discretization of the POD eigenvalue problem for the nonlinear boundary value problem yields [3] the *POD generalized eigenvalue problem* for matrices which is written as  $[A]z = \widehat{\lambda} [H]z$  in which  $[A] = \int_R [H]y(t) ([H]y(t))^T dt$  where  $[A]$  is a positive symmetric  $p \times p$  real matrix,  $y(t)$  is the unique solution of Eq. (1) and where  $[H]$  is the positive-definite symmetric  $p \times p$  real matrix corresponding to the finite element discretization of the inner product in  $L^2$  (this matrix can be constructed as the mass matrix for which the mass density is equal to 1 in all the domain). The eigenvalues are positive numbers such that  $\widehat{\lambda}_1 \geq \widehat{\lambda}_2 \geq \dots \geq \widehat{\lambda}_p \geq 0$ . The associated eigenvectors  $z^1, z^2, \dots, z^p$  constitute the finite element approximation of the POD-basis and satisfy the orthogonality properties  $\langle [A]z^\alpha, z^\beta \rangle = \widehat{\lambda}_\alpha \delta_{\alpha\beta}$  and  $\langle [H]z^\alpha, z^\beta \rangle = \delta_{\alpha\beta}$ .

### Reduced Model and Observation of the Nonlinear Dynamical System

Let  $\{z^1, \dots, z^p\}$  be an algebraic basis of  $R^p$ . Such a basis can be either the LIN-basis or the POD-basis. The reduced model is obtained by projection of Eq. (1) on the subspace  $V_N$  of  $R^p$  spanned by  $\{z^1, \dots, z^N\}$  with  $N \ll p$ . Let  $[Z_N]$  be the  $(p \times N)$  real matrix whose columns are the vectors  $\{z^1, \dots, z^N\}$ . The generalized load vector  $F^N(t) = [Z_N]^T f(t)$  belongs to  $R^N$ . The generalized mass, damping and stiffness matrices  $[M_N] = [Z_N]^T [M] [Z_N]$ ,  $[D_N] = [Z_N]^T [D] [Z_N]$  and  $[K_N] = [Z_N]^T [K] [Z_N]$  are positive-definite symmetric  $(N \times N)$  real matrices. The reduced model of the finite element approximation defined by Eq. (1) is written as  $y^N(t) = [Z_N] q^N(t)$  in which the vector  $q^N(t) \in R^N$  of the generalized coordinates verifies the nonlinear differential equation,

$$[M_N] \ddot{q}^N(t) + [D_N] \dot{q}^N(t) + [K_N] q^N(t) + F_{NL}^N(q^N(t), \dot{q}^N(t)) = F^N(t) \quad , \quad \forall t \in R \quad , \quad (2)$$

where, for all  $q$  and  $p$  in  $R^N$ ,  $F_{NL}(q, p) = [Z_N]^T f_{NL}([Z_N]q, [Z_N]p)$ . It is assumed that the nonlinearities are such that Eq. (1) has a unique solution such that  $y$  and  $\dot{y}$  are square integrable vector-valued functions on  $R$ . Let  $\widehat{y}(\omega) = \int_R e^{-i\omega t} y(t) dt$  be the Fourier Transform of  $y$ . Let  $h(\omega) = (4\pi)^{-1} \{ \langle \omega^2 [M] \widehat{y}(\omega), \overline{\widehat{y}(\omega)} \rangle + \langle [K] \widehat{y}(\omega), \overline{\widehat{y}(\omega)} \rangle \}$  be the energy density in the frequency domain where the overline denotes the complex conjugate. Using the reduced model and denoting the Fourier Transform of  $q^N$  as  $\widehat{q}^N$ , the approximation  $h^N(\omega)$  of  $h(\omega)$  is written as  $h^N(\omega) = (4\pi)^{-1} \{ \langle \omega^2 [M_N] \widehat{q}^N(\omega), \overline{\widehat{q}^N(\omega)} \rangle + \langle [K_N] \widehat{q}^N(\omega), \overline{\widehat{q}^N(\omega)} \rangle \}$ .

### Numerical Solver

Let  $B = [-\omega_{\max}, \omega_{\max}]$  be the frequency band of analysis such that  $B_e \cup B_e \subset B$  in which  $\omega_{\max} = 2\pi f_{\max}$  is such that  $|\int_R \|\widehat{y}(\omega)\|^2 d\omega - \int_{-\omega_{\max}}^{\omega_{\max}} \|\widehat{y}(\omega)\|^2 d\omega| \leq \varepsilon$  with  $\varepsilon$  an *a priori* given precision. The time step is taken as  $\Delta t = 1/(2f_{\max})$  and the time integration is  $\widehat{T} = n_{\text{time}} \Delta t$  with  $n_{\text{time}}$  a positive integer. The integration in  $R$  is

approximated by an integration over the finite interval  $[t_0, t_1]$  in which  $t_0 = -\widehat{T}/2$  and  $t_1 = \widehat{T}/2 - \Delta t$ . The sampling time points are  $t_k = t_0 + k\Delta t$ . The frequency step is  $\Delta\omega = 2\omega_{\max}/n_{\text{freq}}$  with  $n_{\text{freq}} = n_{\text{time}}$ . The sampling frequency points are  $\omega_k = -\omega_{\max} + k\Delta\omega$ . Equations (1) and (2) are integrated over  $[t_0, t_1]$  using an implicit step by step time-integration method (Newmark scheme) with zero initial conditions at  $t_0$ . At each sampling time point  $t_k$  the nonlinear algebraic equation deduced from Eq. (1) or Eq. (2) is solved using an iteration method (fixed point). The matrix  $[A]$  is computed by  $[A] \simeq \Delta t \sum_{k=0}^{n_{\text{time}}-1} [H]y(t_k) ([H]y(t_k))^T$ . The eigenvectors  $z^1, \dots, z^N$  associated with the  $N$  largest eigenvalues  $\widehat{\lambda}_1 \geq \dots \geq \widehat{\lambda}_N$  of the POD generalized eigenvalue problem (or associated with the  $N$  smallest eigenvalues  $\widehat{\mu}_1 \leq \dots \leq \widehat{\mu}_N$  of the normal modes generalized eigenvalue problem) can nowadays be computed for very large generalized eigenvalue problems using an iterative method based on the subspace iteration method or the Lanczos method [6],[7]. Using such iteration method, a very efficient algorithm can be constructed (see [3]). It should be noted that the POD generalized eigenvalue problem can also be solved using the snapshot method (see [8]).

### Numerical studies of the efficiency of the POD-basis with respect to the LIN-basis

In order to define the error functions allowing the efficiency of the two reduced models constructed with the LIN-basis and with the POD-basis to be evaluated, we denote by  $h_{\text{LIN}}^N(\omega)$  the quantity  $h^N(\omega)$  when the LIN-basis is used and by  $h_{\text{POD}}^N(\omega)$  when the POD-basis is used. We then defined the following error functions, depending on the dimension  $N$  of the reduced model, such that  $e_{\text{REF-LIN}}(N) = \widetilde{\omega}_{\max}^{-1} \int_0^{\widetilde{\omega}_{\max}} (\log_{10} h(\omega) - \log_{10} h_{\text{LIN}}^N(\omega)) d\omega$  measures the error between the reference solution and the LIN-basis solution,  $e_{\text{REF-POD}}(N) = \widetilde{\omega}_{\max}^{-1} \int_0^{\widetilde{\omega}_{\max}} (\log_{10} h(\omega) - \log_{10} h_{\text{POD}}^N(\omega)) d\omega$  measures the error between the reference solution and the POD-basis solution,  $e_{\text{LIN-POD}}(N) = \widetilde{\omega}_{\max}^{-1} \int_0^{\widetilde{\omega}_{\max}} (\log_{10} h_{\text{LIN}}^N(\omega) - \log_{10} h_{\text{POD}}^N(\omega)) d\omega$  measures the error between the LIN-basis solution and the POD-basis solution.

The continuous elastic system is constituted of two coupled subsystems. The first subsystem is a linear continuous elastic system constituted of a Timoshenko beam with added dissipation. The second nonlinear subsystem is constituted of a nonsymmetric distributed nonlinearities. The geometrical properties of the beam are: length  $1m$ , width  $0.1m$ , height  $0.1m$ . The boundary conditions are of a cantilever beam. The beam is homogeneous, isotropic, whose material properties are: density  $7500kg/m^3$ , Young's modulus  $2.1 \times 10^{10}N/m^2$ , Poisson's coefficient  $0.3$ , shearing correction factor  $5/6$ . The damping model is introduced by the model damping rate which is  $0.02$  for the first three modes of the uncoupled subsystem,  $0.01$  for the fourth mode and  $0.005$  for the others. The finite element model of the cantilever beam is constituted of  $100$  2-nodes Timoshenko

beam elements. The first six computed eigenfrequencies of the uncoupled subsystem are 26.9, 162.7, 432.9, 794.1, 1219.2 and 1685.3 Hz. The second subsystem is constituted of a distributed density of nonsymmetric nonlinear stiffness producing forces transversally to the beam. At each finite element node of the mesh of the beam the function  $f_{NL}$  is then independent of the velocity and is constructed using  $f_{NL}(y) = 0$  if  $y \leq 0$  and  $f_{NL}(y) = k_0 y^3$  if  $y > 0$ , with  $k_0 = 2.9301 \times 10^{18}$  N/m. For the vector load, the amplitude  $a$  is equal to 1 and the force is a point force applied at the free end of the beam. The impulse function is such that  $g(t) = (\pi t)^{-1} \{ \sin(t(\Omega_c + \Delta\Omega/2)) - \sin(t(\Omega_c - \Delta\Omega/2)) \}$  whose Fourier Transform is  $\hat{g}(\omega) = 1_{B_e \cup B_e^c}(\omega)$ , for which the bandwidth  $\Delta f = 1400$  Hertz and the central frequency  $f_c = 701$  Hertz such that  $\Delta\Omega = 2\pi\Delta f$  and  $\Omega_c = 2\pi f_c$ . Then the frequency band of excitation contains the first five eigenfrequencies of the first linear subsystem. The value of  $f_{max}$  has been calculated in order to obtain a good accuracy for the time integration scheme and is 12000 Hz. On the other hand a convergence analysis has been performed with respect to the time integration  $\hat{T}$  in order that the coupled system be at rest for  $t = \hat{T}/2$  with a good accuracy. This time integration is defined by the value of  $n_{time}$  whose necessary value is 32768. Figure 1(left) displays the graph of the function  $f \mapsto \log_{10} h(2\pi f)$  over  $[0, 8000]$  Hz for the reference solution. Figure 1(right) displays the graph of the function  $j \mapsto \log_{10}(\hat{\lambda}_j/\hat{\lambda}_1)$ . This figure shows the decreasing speed of the eigenvalues which is related to the convergence speed of the POD-solution with respect to  $N$ . For illustrating the

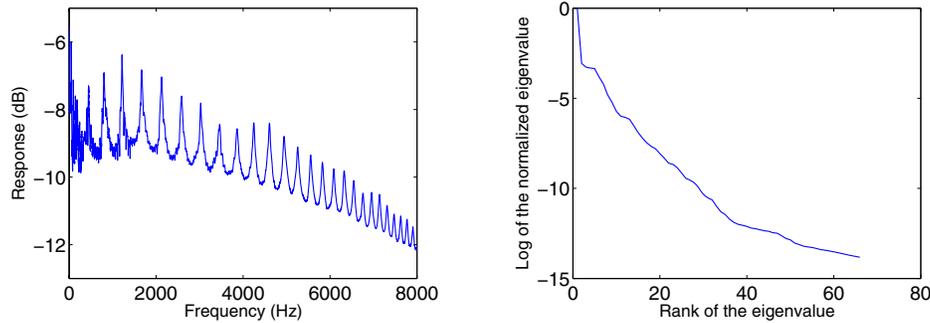


Figure 1: *Left figure:* Graph of  $f \mapsto \log_{10} h(2\pi f)$  for the reference solution; horizontal axis is frequency in Hertz; vertical axis is  $f \mapsto \log_{10} h(2\pi f)$ . *Right figure:* Graph of  $j \mapsto \log_{10}(\hat{\lambda}_j/\hat{\lambda}_1)$ ; horizontal axis is rank  $j$  of the eigenvalue; vertical axis is  $\log_{10}(\hat{\lambda}_j/\hat{\lambda}_1)$ .

convergence in the frequency domain we show the results for a given  $N$ . Figure 2(left) corresponds to  $N = 5$  and shows three curves related to the reference solution (thick solid line), LIN-solution (red thin solid line) and POD-solution (blue thin solid line). It should be noted that the thin solid lines may be superposed and hence some of them are not visible. The error functions are computed for

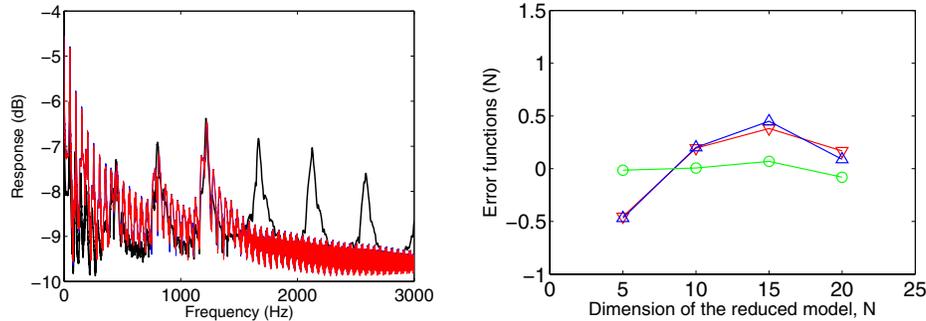


Figure 2: *Left figure:* For  $N=5$ , graph of  $f \mapsto \log_{10} h(2\pi f)$  for the reference solution (black line), RML-solution (red line) and RMNL-solution (blue line); horizontal axis is frequency in Hertz; vertical axis is  $f \mapsto \log_{10} h(2\pi f)$ . *Right figure:* Efficiency of the LIN-solution and POD-solution in function of dimension  $N$  of the reduced model; graphs of the error functions versus  $N$ :  $e_{\text{REF-LIN}}(N)$  (triangle down),  $e_{\text{REF-POD}}(N)$  (triangle up),  $e_{\text{LIN-POD}}(N)$  (circle).

$\tilde{\omega}_{\max} = 2\pi \times 8000$  rad/s. Figure 2(right) displays the three error functions allowing the efficiency of the LIN-solution and the POD-solution in function of dimension  $N$  of the reduced model to be carried out. The analysis of this figure shows that the LIN-basis and the POD-basis have the same efficiency with respect to the dimension  $N$  of the reduction in the frequency band  $[0, 8000]$  Hz, the convergence being reached for  $N = 20$ . We mean by the same efficiency the fact that the two bases need the same value of  $N$  to get convergence.

### Conclusion

This paper has been devoted to the analysis of the efficiency of the reduced models constructed using the POD-basis and the LIN-basis in nonlinear dynamics for continuous elastic systems. The efficiency of the POD-basis or the LIN-basis is related to the speed of convergence in the frequency domain of the solution constructed with the reduced model with respect to its dimension. A basis will be more efficient than another if it converges more rapidly than the other. It can be concluded that the POD-basis is not more efficient than the LIN-basis for the example treated and also for the six other examples presented in [3].

### References

1. Loève, M. (1963): *Probability Theory*, 3rd ed., Van Nostrand, New York.
2. Holmes, P., Lumley, J.L. and Berkooz, G. (1997): *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*, Cambridge University Press, Cambridge.
3. Sampaio, R. and Soize, C. (2006): "Remarks on the Efficiency of POD and KL Methods for Model Reduction in Nonlinear Dynamics of Continuous

- Systems.", *International Journal for Numerical Methods in Engineering*, unpublished, submitted in February 2006.
4. Soize, C. (1998): "Reduced Models in the Medium Frequency Range for General Dissipative Structural-Dynamics Systems", *European Journal of Mechanics A/Solids*, Vol. 17, pp. 657-685.
  5. Ohayon, R. and Soize, C. (1998), *Structural Acoustics and Vibration*, Academic Press, San Diego.
  6. Chatelin, F. (1993): *Eigenvalues of Matrices*, Wiley, New York.
  7. Bathe, K.J. and Wilson, E.L. (1976): *Numerical Methods in Finite Element Analysis*, Prentice Hall, New York.
  8. Sirovich, L. (1987): "Turbulence and the dynamics of coherent structures, Part I: coherent structures", *Quarterly of Applied Mathematics*, Vol. 45, pp. 561-571.

