Nonlinear normal modes, Part I: A useful framework for the structural dynamicist

G. Kerschen, M. Peeters, J.C. Golinval, A.F. Vakakis

1. Nonlinear normal modes: a brief historical perspective and current state-of-the-art

The concept of a normal mode is central in the theory of linear vibrating systems. Besides their obvious physical interpretation, the linear normal modes (LNMs) have interesting mathematical properties. They can be used to decouple the governing equations of motion; i.e., a linear system vibrates as if it were made of independent oscillators governed by the eigensolutions. Two important properties that directly result from this decoupling are:

1. **Invariance**: if the motion is initiated on one specific LNM, the remaining LNMs remain quiescent for all time.
2. **Modal superposition**: free and forced oscillations can conveniently be expressed as linear combinations of individual LNM motions.

In addition, LNMs are relevant dynamical features that can be exploited for various purposes including model reduction (e.g., substructuring techniques [1]), experimental modal analysis [2], finite element model updating [3] and structural health monitoring [4].

Clearly, though, linearity is an idealization, an exception to the rule; nonlinearity is a frequent occurrence in real-life applications [5]. For instance, in an aircraft, besides nonlinear fluid–structure interaction, typical nonlinearities include
The years to come: nonlinearities, this paper is an attempt to highlight several aspects that might drive their development and exploitation in nonlinear dynamics. There are several reasons supporting this statement:

1. Nonlinear systems can exhibit extremely complex behaviors which linear systems cannot. These phenomena include jumps, bifurcations, saturation, subharmonic, superharmonic and internal resonances, resonance captures, limit cycles, modal interactions, and chaos.
2. NNMs have two important limitations compared to their linear counterpart. First and foremost, the principle of superposition, which is the cornerstone of linear theory, does not apply to nonlinear systems. Second, the lack of orthogonality relations satisfied by the NNMs complicates their exploitation as bases for order reduction of the nonlinear dynamics.
3. The vast majority of the NNM literature deals with low-order lumped-mass models (systems with typically a few degrees of freedom—DOFs) or focuses on specific resonances of distributed parameter systems.
4. Most existing constructive techniques for computing NNMs are based on asymptotic approaches and rely on fairly involved mathematical developments.

Any attempt to apply traditional linear analysis to nonlinear systems results, at best, in a suboptimal design. Thus, there is a need for efficient, analytically rigorous, broadly applicable analysis techniques for nonlinear structural dynamics. In this context, nonlinear normal modes (NNMs) offer a solid theoretical and mathematical tool for interpreting a wide class of nonlinear dynamical phenomena, yet they have a clear and simple conceptual relation to the LNMs, with which practicing structural engineers are familiar. Other appealing features of the NNMs are that they are capable of handling strong structural nonlinearity and that they have the potential to address the individualistic nature of nonlinear systems.

The most straightforward definition of an NNM is a vibration in unison of the system (i.e., a synchronous oscillation). NNMs were pioneered in the 1960s thanks to the seminal work of Rosenberg [6–8]. They were further studied in the 1970s by Rand [9–11] and Manevitch and Mikhlin [12]. They were regarded as a theoretical curiosity until the beginning of the 1990s when they were given a new impetus through the efforts of Vakakis et al. [13–18] and Shaw and Pierre [19–22]. Since then, a large body of literature has addressed, with notable success, the qualitative and quantitative analysis of nonlinear phenomena using NNMs (see, e.g., [23–44]). For a few typical applications and comprehensive reviews, the reader can refer to [17,18,45,46].

However, most structural engineers still view NNMs as a concept that is foreign to them, and they do not yet consider NNMs as a useful concept for structural dynamics. There are several reasons supporting this statement:

1. There have been very few attempts to compute NNMs using numerical methods [33,47–52]. Interestingly, algorithms for the continuation of periodic solutions are really quite sophisticated and advanced (see, e.g., [53,54]), yet they have not been fully exploited for the computation of nonlinear modes. In Part II, we support that these numerical algorithms pave the way for an effective and practical computation of NNMs.
2. Time–frequency analysis is a versatile tool for analyzing nonstationary signals; i.e., signals whose spectral contents vary with time. It has been successfully exploited in structural dynamics, e.g., for linear and nonlinear system identification [55–57]. Unlike the Fourier transform, a method such as the wavelet transform models the time evolution of the dominant frequency components of a signal. Because this method can deal adequately with the frequency–energy dependence inherent to nonlinear oscillations, we believe that it is the ideal companion to the NNMs.
3. An appropriate graphical depiction of the NNMs is key to their exploitation. In this study, we show that the representation of NNMs in a frequency–energy plot is particularly convenient. It facilitates the interpretation of the dynamics and clarifies the bifurcations that connect, generate or eliminate the different NNM branches.

Another objective of the present paper is to describe and to illustrate in a simple manner the fundamental properties of NNMs. This is achieved to convince the structural dynamicist not necessarily acquainted with them that they are a useful framework for the analysis of nonlinear vibrating structures.

This paper is organized as follows. In the next section, the two main definitions of NNMs are provided. In addition, their fundamental properties are described, and their representation in a frequency–energy plot is introduced. In Section 3, the different means of computing the NNMs are briefly reviewed and assessed. The potential applications of NNMs to ‘linear’
and nonlinear modal analysis, nonlinear model reduction and the study of localization phenomena are then discussed in Section 4. In Part II, the NNM computation is carried out using numerical continuation algorithms.

2. NNMs: What are they?

In this study, the free response of discrete mechanical systems is considered, assuming that continuous systems have been spatially discretized using, e.g., the finite element method. The equations of motion are

\[
M \ddot{x}(t) + Kx(t) + f_{nl}[x(t), \dot{x}(t)] = 0
\]

where \(M\) is the mass matrix; \(K\) is the stiffness matrix; \(x\), \(\dot{x}\) and \(\ddot{x}\) are the displacement, velocity and acceleration vectors, respectively; \(f_{nl}\) is the nonlinear restoring force vector, assumed to be regular.

To illustrate the different concepts, a two-degree-of-freedom (2DOF) system with a cubic stiffness is chosen. The system is depicted in Fig. 1, and its motion is governed by the equations

\[
\begin{align*}
\ddot{x}_1 + (2x_1 - x_2) + 0.5x_1^3 &= 0 \\
\ddot{x}_2 + (2x_2 - x_1) &= 0
\end{align*}
\]

For comparison purposes, the underlying linear system

\[
\begin{align*}
\ddot{x}_1 + (2x_1 - x_2) &= 0 \\
\ddot{x}_2 + (2x_2 - x_1) &= 0
\end{align*}
\]

is depicted in Fig. 1, and its time series is shown in Fig. 2. The LNM motions of system (3) in the configuration space are depicted in Fig. 3.
is also considered in this study. The time series corresponding to in-phase and out-of-phase normal mode motions of the linear system (3) are depicted in Fig. 2. Motion in the configuration space (i.e., in the plane of the displacements $x_1(t)$ and $x_2(t)$) is given in Fig. 3. Obviously, LNM motions correspond to straight lines in this plane.

2.1. Definition of an NNM

There exist two main definitions of the NNMs in the literature, due to Rosenberg [6–8] and Shaw and Pierre [19–22]. There have been additional definitions, which include a complex-valued invariant manifold formulation [58,59] and group theoretic definitions [17], but they are not described herein.

Historically, Lyapunov and Poincaré contributions served as the cornerstone of the NNM development. For $n$-DOF conservative systems with no internal resonances, Lyapunov showed that there exist at least $n$ different families of periodic solutions around the stable equilibrium point of the system. At low-energy, the periodic solutions of each family are in the neighborhood of a LNM of the linearized system. These $n$ families define $n$ NNMs that can be regarded as nonlinear extensions of the $n$ LNMs of the underlying linear system. Poincaré and Poincaré-Dulac theorems are fundamental for establishing the theory of normal form, which was later exploited for NNM computation [23,39].

2.1.1. Rosenberg’s definition

During the normal mode motion of a linear conservative system, each system component moves with the same frequency and with a fixed ratio amongst the displacements of the components. Targeting a straightforward nonlinear extension of the LNM concept, Rosenberg defined an NNM as a vibration in unison of the system (i.e., a synchronous oscillation). This definition requires that all material points of the system reach their extreme values and pass through zero simultaneously and allows all displacements to be expressed in terms of a single reference displacement.

For illustration, the time series and the configuration space of in-phase and out-of-phase NNM motions during the free response of system (2) are depicted in Figs. 4 and 5, respectively. The modal lines of this nonlinear system are curves, resulting from the nonlinear relationship between the coordinates during the periodic motion. These curved NNMs, termed nonsimilar NNMs by Rosenberg, are generic in nonlinear systems, and their existence certainly complicates the concept of orthogonality between modes [60,61]. As discussed in Section 2.2.1, a particular feature of these NNMs is that their shape depends on the total energy present in the system. When special spatial symmetries exist, the NNMs may degenerate into (energy-invariant) straight modal lines, as in the linear case (see [14] and Section 2.2.3).

![Fig. 4. Time series of NNM motions of system (2) (——$x_1(t)$; ---$x_2(t)$). Left plot: in-phase NNM ($[x_1(0)\ x_2(0)\ \dot{x}_1(0)\ \dot{x}_2(0)]=\begin{bmatrix} 3.319 & 11.134 & 0 & 0 \end{bmatrix}$); right plot: out-of-phase NNM ($[x_1(0)\ x_2(0)\ \dot{x}_1(0)\ \dot{x}_2(0)]=\begin{bmatrix} -10.188 & 0.262 & 0 & 0 \end{bmatrix}$).](image)

![Fig. 5. NNM motions of system (2) in the configuration space. Left plot: in-phase NNM; right plot: out-of-phase NNM.](image)
2.1.2. The invariant manifold approach

Shaw and Pierre proposed a generalization of Rosenberg’s definition that provides a direct and elegant extension of the NNM concept to damped systems. Based on geometric arguments and inspired by the center manifold technique [62], they defined an NNM as a two-dimensional invariant manifold in phase space. Such a manifold is invariant under the flow (i.e., orbits that start out in the manifold remain in it for all time), which extends the invariance property of LNMs to nonlinear systems. In order to parameterize the manifold, a single pair of state variables (i.e., both the displacement and the velocity) are chosen as master coordinates, the remaining variables being functionally related to the chosen pair. Therefore, the system behaves like a nonlinear single-DOF system on the manifold.

Geometrically, LNMs are represented by planes in phase space, and NNMs are two-dimensional surfaces that are tangent to them at the equilibrium point. For illustration, the manifolds corresponding to in-phase and out-of-phase NNMs of system (2) are given in Fig. 6.

2.1.3. The approach followed in this study

At first glance, Rosenberg’s definition may appear restrictive in two cases:

1. This definition, as such, cannot be easily extended to nonconservative systems.
2. In the presence of internal resonances (i.e., when two or more NNMs interact), some coordinates may have a dominant frequency component different than that of the other coordinates (e.g., some coordinates may vibrate faster than others). In this case, the system no longer vibrates in unison. This is illustrated in Fig. 7 for an internally resonant NNM (3:1 internal resonance) of system (2).

However, these two limitations can be circumvented. Firstly, as shown in Section 4.2 and in [33,40], the damped dynamics can often be interpreted based on the topological structure and bifurcations of the NNMs of the underlying undamped system. We also note that, due to the lack of knowledge of damping mechanisms, engineering design in industry is often based on the conservative system, and this even for linear vibrating structures. Secondly, realizing that the motion is still periodic in the presence of internal resonances, Rosenberg’s definition of an NNM can be extended to a (nonnecessarily

Fig. 6. Two-dimensional invariant manifolds of system (2) with the corresponding LNMs. Left plot: in-phase LNM and NNM; right plot: out-of-phase LNM and NNM.

Fig. 7. Internally resonant NNM (3:1 internal resonance; \( [x_1(0) x_2(0) \dot{x}_1(0) \dot{x}_2(0)] = [8.476 54.263 0 0] \)). Left plot: time series (-----, \( x_1(t) \); ---, \( x_2(t) \)); right plot: configuration space.
synchronous) periodic motion of the conservative system. This extended definition is particularly attractive when targeting a numerical computation of the NNMs. It enables the nonlinear modes to be effectively computed using algorithms for the continuation of periodic solutions, which are really quite sophisticated and advanced. This NNM definition is considered throughout the present study.

2.2. Fundamental properties

NNMs have intrinsic properties that are fundamentally different from those of LNMs. They are reviewed and illustrated in what follows.

2.2.1. Frequency–energy dependence

One typical dynamical feature of nonlinear systems is the frequency–energy dependence of their oscillations. One important consequence is that the frequency response functions (FRFs) of nonlinear systems are no longer invariant. For illustration, the FRFs of system

\[
\ddot{x}_1 + (0.02x_1 - 0.01\dot{x}_2) + (2x_1 - x_2) + 0.5x_1^3 = F \cos \omega t \\
\ddot{x}_2 + (0.02x_2 - 0.01\dot{x}_1) + (2x_2 - x_1) = 0
\]

have been computed using a modified version of the algorithm described in Part II and are depicted in Figs. 8 and 9 for \(F\) varying between 0.002 and 0.2 N.

The modal curves and frequencies of oscillation of NNMs also depend on the total energy in the system. In contrast to linear theory, this energy dependence prevents the direct separation of space and time in the governing equations of motion, which complicates the analytical calculation of the NNMs.

Returning to the undamped system (2), Fig. 10 shows the time series, the configuration space, the power spectral density (PSD) and two-dimensional projections of the phase space of three in-phase NNM motions of increasing energies. The NNM motion at low-energy resembles that of the in-phase LNM of the underlying linear system (3). The modal curve is a straight line, there is one main harmonic component in the system response, and the motion in phase space is a circle. For the motion at moderate energy, the NNM motion is now a curve, and the presence of two harmonic components can be detected. A clear departure from the LNM (harmonic) motion is observed. At high-energy, this is even more enhanced.

Fig. 8. Nonlinear frequency response functions close to the first resonant frequency (five different forcing amplitudes: 0.002, 0.01, 0.05, 0.1, 0.2 N). Left plot: \(x_1\); right plot: \(x_2\).

Fig. 9. Nonlinear frequency response functions close to the second resonant frequency (five different forcing amplitudes: 0.002, 0.01, 0.05, 0.1, 0.2 N). Left plot: \(x_1\); right plot: \(x_2\).
For instance, the motion in phase space is a strongly deformed ellipse. When moving from the low- to the high-energy NNM, the period of the motion decreases from 6.28 to 4.755 s. This is due to the hardening characteristic of the cubic spring. For illustration, Fig. 11 also superposes the three-in-phase NNMs in the configuration space.

To further illustrate the frequency–energy dependence of the NNMs, the harmonic balance method can be applied to system (2). This method expresses the periodic motion of a system by means of a finite Fourier series [63]. For simplicity, a series with a single harmonic component is considered

\[
\begin{align*}
x_1(t) &= A \cos \omega t, \\
x_2(t) &= B \cos 2\omega t
\end{align*}
\]

This ansatz is plugged into the equations of motion (2). Expanding \( \cos^3 \omega t \) in terms of \( \cos \omega t \) and \( \cos 3\omega t \), and balancing all the coefficients of the \( \cos \omega t \) terms yields

\[
\begin{align*}
-\omega^2 A + (2A - B) + 0.5 \frac{3A^3}{4} &= 0 \\
-\omega^2 B + (2B - A) &= 0
\end{align*}
\]
Analytic expressions for coefficients $A$ and $B$ are then readily obtained

$$A = \pm \sqrt{\frac{8(\omega^2 - 3)(\omega^2 - 1)}{3(\omega^2 - 2)}}$$

$$B = \frac{A}{2 - \omega^2}$$

The square root exists in the two frequency intervals

$$\omega_1 \in [1, \sqrt{2}] \quad \text{and} \quad \omega_2 \in [\sqrt{3}, +\infty[$$

noting that $\omega = 1 \text{ rad/s}$ and $\omega = \sqrt{3} \text{ rad/s}$ are the two natural frequencies of the underlying linear system (3). In the first (second) frequency interval, $B$ has the same (opposite) sign as $A$; an in-phase (out-of-phase) NNM motion is observed for initial conditions $[x_1(0) \ x_2(0) \dot{x}_1(0) \dot{x}_2(0)] = [A \ B \ 0 \ 0]$.

The (conserved) total energy during the free response of system (2) is

$$\text{Total Energy} = \text{Kinetic Energy} + \text{Potential Energy} = \frac{A^2}{2} + \frac{(B - A)^2}{2} + \frac{B^2}{2} + \frac{0.5 A^4}{4}$$

which, according to Eqs. (7) and (8), demonstrates the frequency–energy dependence of NNM motions.

An appropriate graphical depiction of the NNMs is key to their exploitation. The usual representation in the literature is to plot the motion amplitude at a particular DOF as a function of frequency. Due to the frequency–energy dependence, we believe that the representation of NNMs in a frequency–energy plot (FEP) is particularly convenient $[33, 40]$. An NNM motion is represented by a point in the FEP, which is drawn at a frequency corresponding to the minimal period of the periodic motion and at an energy equal to the conserved total energy during the motion. A branch, represented by a solid line, is a family of NNM motions possessing the same qualitative features (e.g., the in-phase NNM motions of a 2DOF system).

![Frequency–energy plot of system (3). LNM motions depicted in the configuration space are inset.](image1)

![Frequency–energy plot of system (2). NNM motions depicted in the configuration space are inset. The horizontal and vertical axes in these plots are the displacements of the first and second DOFs, respectively; the aspect ratio is set so that increments on the horizontal and vertical axes are equal in size to indicate whether or not the motion is localized to a particular DOF.](image2)
As a point of comparison, the FEP of the linear system (3) is shown in Fig. 12. Because the natural frequencies do not depend on energy, this FEP comprises two horizontal lines at the two resonant frequencies of the system. The LNM motions represented in the configuration space are inset and are obviously unaltered by the energy level.

The FEP of the nonlinear system (2) was computed using the method proposed in Part II and is shown in Fig. 13. The backbone of the plot is formed by two branches, which represent in-phase \((S_{11}^+ )\) and out-of-phase \((S_{11}^-)\) synchronous NNMs. The letter \(S\) refers to symmetric periodic solutions for which the displacements and velocities of the system at half period are equal but with an opposite sign to those at time \(t = 0\). As shown in the next section, unsymmetric periodic solutions may also be encountered and are denoted by a letter \(U\). The indices in the notations are used to mention that the two masses vibrate with the same dominant frequency. The FEP clearly shows that the nonlinear modal parameters have a strong dependence on the total energy in the system:

1. The frequency of both the in-phase and out-of-phase NNMs increases with the energy level, which reveals the hardening characteristic of the system.
2. The modal curves change for increasing energies. The in-phase NNM tends to localize to the second DOF (i.e., it resembles a vertical curve), whereas the out-of-phase NNM localizes to the first DOF (i.e., it resembles an horizontal curve). This localization property is a key feature of nonlinear systems that is briefly discussed in Section 4.4. It is discussed extensively in [17] and exploited for vibration mitigation in [33,40,64,65].

The comparison between Figs. 12 and 13 also reveals that NNMs have a clear and simple conceptual relation to the LNMs.

2.2.2. Modal interactions—internally resonant NNMs

Another salient feature of nonlinear systems is that NNMs may interact during a general motion of the system. Nonlinear modal interactions have been studied extensively in the literature (see, e.g., the monograph [45]). A case of particular interest is when the linear natural frequencies are commensurate or nearly commensurate \([60,66–68]\). An energy exchange between the different modes involved may therefore be observed during the internal resonance. For instance, exciting a high-frequency mode may produce a large-amplitude response in a low-frequency mode. Vibration absorbers exploiting these energy transfers have been studied in [69].

Internally resonant NNMs have no counterpart in linear systems. Considering system (2) and according to the discussion in the previous section, the FEP in Fig. 13 does not seem to feature internally resonant NNMs. However, when carrying out the NNM computation at higher energy levels, Fig. 14 shows that another branch of periodic solutions, termed a tongue, emanates from the backbone branch \(S_{11}^+\). On this tongue, denoted \(S_{31}\), there is a 3:1 internal resonance between the in-phase and out-of-phase NNMs.

Surprisingly, the ratio of the linear natural frequencies of system (2) is \(\sqrt{3}\). Due to energy dependence, a 3:1 ratio between the two frequencies can still be realized, because the frequency of the in-phase NNM increases less rapidly than that of the out-of-phase NNM. This clearly shows that NNMs can be internally resonant without necessarily having commensurate linear natural frequencies, a feature that is rarely discussed in the literature [70,71]. This also underlines that important nonlinear phenomena can be missed when resorting to perturbation techniques, which are limited to small-amplitude motions.

![Fig. 14. Frequency–energy plot of system (2) featuring a 3:1 internal resonance between the in-phase and out-of-phase NNMs.](image)
To better illustrate the resonance mechanism, the branch $S_{11}/C_0$ is represented in the FEP of Fig. 15 at the third of its frequency. This is relevant, because a periodic solution of period $T$ is also periodic with period $3T$, and the resulting branch is therefore denoted $S_{33}/C_0$. It turns out that a smooth transition from $S_{11}$ to $S_{33}/C_0$ occurs on tongue $S_{31}$. This transition is also depicted in Fig. 16 where the evolution of the configuration space and of the Fourier coefficients is shown for several points on $S_{31}$ or in its vicinity. Starting from NNM (a), an in-phase motion characterized by two perceptible harmonic components is observed. From (a) to (d), the relative importance of the third harmonics grows, as clearly confirmed by the motion in the configuration space. Moving from (d) to (e) corresponds to a drastic qualitative change in the dynamics. Firstly, the first harmonics has completely disappeared for both oscillators. Secondly, the signs of the coefficients of the third harmonics are opposite. Overall, this means that an out-of-phase NNM motion is realized. Eventually, through a 3:1 internal resonance, the motion ends up on $S_{33}/C_0$ or, equivalently, on $S_{11}/C_0$. From (f) to (h), the relative importance of the third harmonics diminishes, and a motion qualitatively similar to that at (a) is observed. However, the configuration space of NNM (h) reveals the presence of a fifth harmonics, which is a precursor to the gradual development of tongue $S_{51}$.

This indicates that other resonance scenarios exist in this seemingly simple system. The frequency of the out-of-phase NNM motions on $S_{11}$ steadily increases for increasing energies, whereas the NNM motions on $S_{11}$ have their frequency asymptotically approaching a value close to $\sqrt{3}$ rad/s. Following this reasoning, we expect the existence of a countable infinity of internal resonance cases (e.g., 2:1, 4:1, 5:1, etc.). To confirm this conjecture, additional tongues have been computed numerically and are represented in Fig. 17. These tongues emanate from $S_{11}$ and coalesce into $S_{11}$ following a mechanism similar to that described above (Fig. 16). To illustrate the rich dynamics, a few representative NNMs of system (2) are depicted in Fig. 18. Such a complex dynamics was first observed in [33] for a system with an essential nonlinearity. It is interesting that this can also be reproduced for more generic nonlinear systems.

2.2.3. Mode bifurcations and stability

A third fundamental property of NNMs is that their number may exceed the number of DOFs of the system. Due to mode bifurcations, not all NNMs can be regarded as nonlinear continuation of normal modes of linear systems [17,72,73]. Internally resonant NNMs are one example. Another possible example corresponds to the NNM bifurcations of the system

$$\ddot{x}_1 + x_1 + x_1^3 + K(x_1 - x_2)^3 = 0$$

$$\ddot{x}_2 + x_2 + x_2^3 + K(x_2 - x_1)^3 = 0$$  \hspace{1cm} (11)

for variations of the coupling spring $K$ [14]. This system possesses similar NNMs that obey to the relation $x_2(t) = c x_1(t)$. Eliminating $x_2$ from Eqs. (11) yields

$$\ddot{x}_1 + x_1 + (1 + K(1 - c^3)x_1^3 = 0$$

$$\ddot{x}_1 + x_1 - \frac{1}{c^3}(K(1 - c^3) + c^3)x_1^3 = 0$$  \hspace{1cm} (12)

Because both equations must lead to the same solution, it follows

$$K(1 + c)(c - 1)^3 = c(1 - c^2), \quad c \neq 0$$  \hspace{1cm} (13)
Fig. 16. Internally resonant NNMs (3:1 resonance). Top plot: close-up of the tongue S31 in the frequency–energy plot. Bottom plots: configuration space (horizontal axis: $x_1$; vertical axis: $x_2$) and Fourier coefficients of a series containing cosine terms only (gray, $x_1$; black, $x_2$).
Eq. (13) means that system (11) always possesses two modes characterized by \( c = \pm 1 \) that are direct extension of the LNMs. However, this system can possess two additional similar NNMs that cannot be captured using linearization procedures. At \( K = 0.25 \), these NNMs bifurcate from the out-of-phase mode, as shown in Fig. 19.

Another important characteristic of NNMs is that they can be stable or unstable, which is in contrast to linear theory where all modes are neutrally stable. In this context, instability means that small perturbations of the initial conditions that generate the NNM motion lead to the elimination of the mode oscillation. Therefore, unstable NNMs are not physically realizable. The NNM stability analysis can be performed numerically or analytically. In Fig. 20, stability is computed numerically through the eigenvalues of the monodromy matrix. In other studies, analytical results are obtained through Floquet theory after adding small perturbations to the periodic solutions. For a detailed stability analysis of the NNMs, the reader can refer to [14, 16, 17, 74, 75].
Bifurcations and stability are interrelated concepts, because a change in stability occurs through a bifurcation. For instance, the bifurcation in system (11) generates a pair of stable/unstable NNMs (Fig. 19). Returning to system (2), another illustration of NNM stability is shown in the FEP of Fig. 20. When the tongue $U_21$ bifurcates from $S_11+$, the NNMs on this latter branch lose stability. A detailed description of this tongue and the related dynamical mechanisms (e.g., symmetry-breaking bifurcation) is beyond the scope of this paper. This figure also shows that stability can be lost when a turning point is encountered.

3. NNMs: How to compute them?

Different methods for computing NNMs of discrete and continuous systems are briefly described in this section. They are classified in two categories, according to whether the computation relies on analytical or numerical methods. This discussion is by no means a comprehensive overview of the past and current approaches and focuses on the free response of nonlinear vibrating systems. There is no attempt to summarize the methods dealing with the forced response case.

3.1. Analytical techniques

Rosenberg was the first to develop constructive techniques for computing NNMs of discrete conservative oscillators. Rand obtained explicit approximate expressions for modal curves in 2DOF systems by applying a perturbation method to a modal equation [10]. Manevitch and Mikhlin reduced the problem of computing nonsimilar NNMs to a set of singular boundary value problems, which were then solved using power series expansions [12] (see Section 3.1.1). The book by Vakakis et al. [17] summarizes the developments until the 1990s.

The early 1990s witnessed a resurgence of interest in the NNMs with the works of Vakakis [13–16] and Shaw and Pierre [19–22]. Simple discrete systems were first studied [15,19,21], but the generalization to continuous systems [16,22] soon followed. For continuous systems, two main approaches exist in the literature. The first approach is to study directly the original partial differential [16,22]. An alternative method is to discretize the governing nonlinear partial differential equations into an infinite set of ordinary differential equations that is truncated to a finite number of equations [20]. The two alternatives are compared in [76,80] using the invariant manifold approach.

3.1.1. An energy-based formulation

This formulation relies on Rosenberg’s work [8] and expresses an NNM as a modal curve in the configuration space. It was further developed by Manevitch and Mikhlin for discrete conservative oscillators [12] and exploited in a few other studies [15,17]. To illustrate the method, it is applied to system (2)

$$\ddot{x}_1 + (2x_1 - x_2) + 0.5x_1^3 = 0$$

$$\ddot{x}_2 + (2x_2 - x_1) = 0$$

(14)

When the system vibrates along an NNM, the displacement $x_2$ is linked to $x_1$ through the expression of the modal curve $\hat{x}_2$

$$x_2 = \hat{x}_2(x_1)$$

(15)

The objective of the method is to eliminate the time derivatives from the equations of motion (14). To compute the second time derivative of $x_2$, relation (15) is differentiated twice using the chain rule

$$\ddot{x}_2 = \dddot{x}_2(x_1^2) + \ddot{x}_2 \dot{x}_1$$

(16)
where prime denotes differentiation with respect to \( x_1 \). This expression involves the second time derivative of \( x_1 \), which is readily obtained from the equations of motion

\[
\dot{x}_1 = -2x_1 + \ddot{x}_2 - 0.5x_1^3
\]  

(17)

It then remains to compute the first time derivative of \( x_1 \) appearing in Eq. (16). To this end, a first integral of motion expressing explicitly the conservation of energy during the motion is written by multiplying Eq. (17) by \( \dot{x}_1 \) and integrating

\[
\dot{x}_1^2 = 2 \int_0^{x_1} \dot{x}_1 \, dx_1 = -2 \int_{x_1}^{X_1} [2u - \dot{x}_2(u) + 0.5u^2] \, du
\]  

(18)

where \( X_1 \) is the maximum amplitude attained by \( x_1 \); i.e., when \( \dot{x}_1 = 0 \). The derivatives are substituted into the second of Eqs. (14), which yields the equation governing the modal curve:

\[
\dot{x}_2^2 \{ -2 \int_{x_1}^{X_1} [2u - \dot{x}_2(u) + 0.5u^2] \, du \} + \ddot{x}_2^2 [ -2x_1 + \dot{x}_2 - 0.5x_1^3 ] + (2\dot{x}_2 - x_1) = 0
\]  

(19)

Because the coefficient of the highest derivative vanishes when \( x_1 = X_1 \), this functional equation is singular at the maximum equipotential surface. It must therefore be supplemented by a boundary condition

\[
[\dot{x}_2^2 [ -2x_1 + \dot{x}_2 - 0.5x_1^3 ] + (2\dot{x}_2 - x_1) \} \big|_{x_1 = X_1} = 0
\]  

(20)

which expresses that the nonlinear mode intersects orthogonally the maximum equipotential surface in the configuration space. Eq. (19) does not depend on the time variable, and its solution is amenable to a power series expansion:

\[
\dot{x}_2(x_1) = \dot{x}_2(0)(x_1) + \alpha \dot{x}_2(1)(x_1) + \epsilon^2 \dot{x}_2(2)(x_1) + O(\epsilon^3)
\]  

(21)

This formulation was extended to undamped continuous systems in [16]. The displacement of any point of the system is expressed in terms of a single reference displacement \( x_0(t) = x(s_0, t) \) by the functional relation

\[
x(s, t) = X(s, x_0(t))
\]  

(22)

where \( s \) is the spatial coordinate, and \( X \) is a modal function characterizing the considered NNM. Then, an integral equation expressing the conservation of energy during the motion is used in conjunction with Eq. (22) to eliminate the time derivatives from the equations of motion. Eventually, the equation governing the modal function \( X \) is obtained and is solved using power series.

In the presence of internal resonances, the folding of the NNMs in the configuration space may result in multivalued relationship among the various coordinates (see Fig. 7). This has been nicely addressed in [68] by considering NNMs in an appropriately defined modale space.

### 3.1.2. The invariant manifold approach

The invariant manifold approach [19–22] is similar in spirit to the energy-based formulation. The difference with the previous approach is that a pair of state variables (i.e., both the displacement and the velocity) are chosen as master coordinates, the remaining variables being functionally related to the chosen pair:

\[
x(s, t) = X_1(s, x_0(t), x_0(t)) \quad \text{and} \quad \dot{x}(s, t) = X_2(s, x_0(t), x_0(t))
\]  

(23)

These relations define a two-dimensional invariant manifold in phase space. By taking the time derivative of these constraint equations and using the chain rule differentiation, the explicit time dependence from the equations of motion can be eliminated. Eventually, this yields a set of partial differential equations governing the modal functions \( X_1 \) and \( X_2 \). These equations are as difficult to solve as the original problem, but the solution can be approximated using power series. In summary, this is a six-step procedure:

1. choose master coordinates \( x_0(t), \dot{x}_0(t) \);
2. express slaved coordinates as \( X_1[s, x_0(t), \dot{x}_0(t)], \dot{X}_2[s, x_0(t), \dot{x}_0(t)] \);
3. use the invariant manifold technique to eliminate time dependence;
4. approximate a local solution using polynomial expansion of \( X_1 \) and \( X_2 \) in terms of \( x_0(t) \) and \( \dot{x}_0(t) \);
5. substitute expansions into time-independent partial differential equations governing the geometry of the manifold, and solve polynomial expansion of \( X_1 \) and \( X_2 \);
6. replace the slaved coordinates with their expansions, thus eliminating them from the system.

For systems with internal resonances, a multi-mode invariant manifold is considered in [66] to account for the influence of several modes. For instance, when two modes are resonant, the master coordinates comprise two pairs of state variables, and the resulting invariant manifold is four-dimensional. The invariant manifold approach was also reformulated using a complex framework [58], which was then extended to systems with internal resonances [59].
3.1.3. The multiple scales method

One perturbation method that has received considerable attention for the NNM computation is the method of multiple scales [31, 51, 58, 60, 61, 77, 78]. Governing partial differential equations can be attacked directly (i.e., without spatial discretization) with this method. The first step is to introduce a small nondimensional parameter $\epsilon$ to indicate the smallness of the nonlinear terms. The solution is then sought in the form of an asymptotic expansion [63].

The underlying idea of the multiple scales method is to consider expansions in terms of multiple independent time scales, instead of a single independent variable $t$

$$x(s, t) = \alpha_1(s, T_0, T_1, T_2, \ldots) + \epsilon^2 \alpha_2(s, T_0, T_1, T_2, \ldots) + \cdots \quad \text{with} \quad T_i = \epsilon^i t$$

where $T_0$ is a time scale characterizing the fast motion of the system (i.e., the motion occurring at the dominant frequency of the NNM). Because the generic motion of a nonlinear system is not harmonic, other time scales are necessary to describe the motion; these are the slow time scales $T_1, T_2, \ldots$. An increasingly accurate approximation is therefore obtained as additional time scales enter in the analysis. The approximating functions $x_i(s, T_0, T_1, T_2, \ldots)$ are then determined after integration of linear differential equations for each order of $\epsilon$, and imposition of solvability conditions, which correspond to the elimination of secular terms.

3.1.4. Other approaches

The method of normal forms was first employed by Lamarque and Jézéquel [23] and Nayfeh [79] using a complex formulation. A real normal theory for the NNM computation was then proposed in [80, 81] for conservative systems and in [39] for nonconservative systems. The philosophy of the method is to seek a nonlinear change of coordinates$^1$ that brings the equations of motion into the simplest possible form, termed the normal form [62]. In other words, the objective of the method is to eliminate as many as possible of the nonlinear terms from the governing equations, which is similar in spirit to the decoupling of the equations of motion provided by the LNMs of a linear system. However, a complete decoupling of the equations is generally not possible for nonlinear systems, and only the invariance property can be enforced. Eventually, the computed normal form dictates the dynamics of the system on the invariant manifold.

Another technique that assumes that the NNM motion is periodic in time and approximates it by means of a finite Fourier series is the harmonic balance method; e.g.,

$$x(s, t) = \sum_{n=0}^{N} a_n^1(s) \cos n\omega t + \sum_{n=0}^{N} a_n^2(s) \sin n\omega t$$

By substituting this relation into the governing equations of motion and ‘balancing the harmonics’, the nonlinear modes can be computed by solving nonlinear boundary value problems for the $a_n$ [82–89]. Because analytical solutions are available in a limited number of cases (mostly when a single harmonic component is considered; see Section 2.2.1), numerical methods are often used to solve the resulting equations. The harmonic balance method can therefore be viewed as a semi-analytical technique.

A method similar in spirit to the harmonic balance method and to the Galerkin-based approach in [48] was introduced in [41, 90]. The most distinctive feature of this formulation is that the modal vector and the corresponding frequency depend on the amplitude but also on the total phase variable. The dynamics is defined by a differential equation, governing the total phase motion, from which the period of the oscillations is deduced.

3.2. Numerical techniques

Most existing constructive techniques for computing NNMs are based on asymptotic approaches. Despite that robust algorithms for the computation of isolated periodic orbits [91] and for the continuation of a family of orbits [53, 54] have been developed, it is somewhat surprising that there have been very few attempts to compute NNMs using numerical methods [33, 47–52].

One of the first approaches was proposed by Slater in [47]. Based on Rosenberg’s definition, the procedure integrates directly the governing equations of motion over one period using numerical algorithms (e.g., Runge–Kutta and Newmark). It comprises two steps:

1. An isolated periodic solution corresponding to a specific energy level is computed by modifying iteratively the initial conditions governing the free response of the system. This is carried out using optimization algorithms that minimize a periodicity condition (i.e., a cost function representing the lack of periodicity of the current iterate).
2. Low-energy modal curves and the corresponding periods of oscillation are first computed, taking the normal modes and natural frequencies of the underlying linear system as initial guesses. The energy is then gradually increased with the previously computed NNM as an initial guess for the next NNM.

$^1$ To recover the linear results at small-amplitude motions, the applied coordinate transformations must be near-identity. As a result, the method can only be used in the neighborhood of an equilibrium point.
This step-wise type of continuation of periodic solutions is called sequential continuation [92]. Similarly, shooting algorithms coupled with sequential continuation were considered in Lee et al. [33,93] and Bajaj et al. [51,94] to numerically solve the nonlinear boundary value problem that defines a family of NNM motions.

A more sophisticated continuation method is the so-called asymptotic-numerical method [95]. It is a semi-analytical technique that is based on a power series expansion of the unknowns parameterized by a control parameter. It is utilized to follow the NNM branches in conjunction with the harmonic balance method in [88] or with finite difference methods in [50,96]. Another well-established method implemented in the AUTO software is the pseudo-arclength continuation. It is the approach used for the NNM calculation in Part II.

Based on the invariant manifold approach, Pesheck et al. [48,97] developed a meaningful numerical extension of it. In the original formulation, the master variables are the position and velocity in Cartesian coordinates, and the solution is sought using a polynomial expansion. In the proposed Galerkin-based approach, an alternative set of coordinates is defined (i.e., the amplitude and phase of the considered mode), and the polynomial approach is replaced by a Galerkin method. Eventually, a set of nonlinear algebraic equations is obtained and solved using local optimization algorithms.

Finally, we note that computer implementation of both the multiple scales and the invariant manifold approach have been carried out in [77,98] and applied to finite element models of planar frames and beams.

3.3. Assessment of the different methodologies

Analytical methodologies have the advantage that NNMs can be constructed symbolically, which is certainly useful for gaining insight into the dynamics and for performing parametric studies. Among other things, they clearly highlight the frequency–energy dependence of the NNMs. The fundamental drawbacks of these techniques are that (i) they are quite analytically involved and require a careful treatment in the presence of internal resonances; (ii) the resultant dynamics are only accurate for small-amplitude motions; and (iii) the upper bound for these motions is not known a priori.

The energy-based formulation is an elegant approach, but, because it is based on symmetry arguments, it requires that the nonlinearities be of odd order. It is also a priori limited to undamped systems. The invariant manifold approach does not present these limitations. However, though its basic definition allows for large-amplitude motions, the constructive technique, which relies on power series expansions, is limited to small motions. Regarding perturbation analysis (e.g., the multiple scales method), it can now be performed using symbolic manipulation programs, but their application to nonlinear systems with more than a few DOFs remains involved. Different analytical methods have been compared in [79,94], and the results obtained were consistent. In [79], the author reports that the method of multiple scales is the simplest and involves the least algebra. Finally, we note that the harmonic balance method yields solutions which are generally valid over a much larger domain compared to the aforementioned methodologies. However, because analytic expressions of the resulting equations are available only in a limited number of cases, it should be regarded more as a numerical technique.

Analytical approaches may become inaccurate in the moderate to strongly nonlinear range of motion and are limited to systems with low dimensionality. Their utility for real-world structures is therefore questionable. In this context, numerical methods have certainly the potential to ‘push the envelope’ and to make nonlinear modal analysis more accessible to the practicing structural engineer. The key advantage of these methods is that they lend themselves fairly easily to strongly nonlinear systems with large-amplitude motions, which is nicely evidenced in [97,103]. In addition, most of them provide an exact solution to the NNM calculation. Their fundamental drawback is that they rely on extensive numerical simulations and are still computationally intensive.

The Galerkin-based invariant manifold approach is one of the most effective techniques for building reduced-order models of nonlinear vibrating systems. It is truly versatile and can be applied to a large variety of nonlinear dynamic systems, including nonconservative, gyroscopic and piecewise-linear systems, with an accuracy controlled over the chosen amplitude range. One possible limitation is that the interpretation of the NNMs is complicated when multi-mode invariant manifolds, which are higher-dimensional surfaces, are computed.

On the contrary, a particularly appealing feature of the continuation of periodic solutions is that the resulting NNMs have a clear conceptual relation to the LNM, with which practicing structural engineers are familiar (see Section 2.2.1). As discussed in Section 4.2, this makes it a promising technique for developing a practical nonlinear analog of experimental modal analysis, which is well-established for linear systems. In this framework, the implementation of sequential continuation techniques is truly straightforward, and the calculations can be performed with limited user interaction. They represent the ideal starting point for the dynamicist not necessarily acquainted with the numerical calculation of the NNMs. However, their computational efficiency is limited, and they are likely to fail when a turning point or a bifurcation is encountered. Effective alternatives are those based on more sophisticated continuation techniques (e.g., the asymptotic-numerical method and the pseudo-arclength continuation). One limitation of the continuation of periodic solutions is that it is not clear how they can be extended to nonconservative systems. Nevertheless, as shown in Section 4.2, the damped dynamics can be interpreted based on the topological structure and bifurcations of the NNMs of the underlying undamped system.
4. NNMs: Why are they useful?

The objective of this section is to describe several applications where NNMs represent a useful framework for the structural dynamicist. Specifically, we highlight how useful the NNMs are for modal analysis and system identification and how they may be exploited in conjunction with time–frequency analysis in order to extend the existing linear methodologies [2]. Nonlinear model reduction and the study of localization phenomena are also discussed.

4.1. ‘Linear’ modal analysis

Modal analysis and testing of linear mechanical structures has been developed over the past 40–50 years, and the techniques available today are mature and advanced [2]. Clearly, though, linearity is an idealization, an exception to the rule; nonlinearity is a frequent occurrence in real-life applications. In the presence of nonlinear phenomena, the structural dynamicist should therefore ask the question: Can I still use the linear modes? Obviously, the answer depends on the type of the nonlinearity and on the excitation level.

In this context, we believe that the computation of the NNMs and their representation in a FEP is a robust and accurate tool to decide whether or not the linear framework is still applicable. It can be used to determine which modes (and to what extent) are sensitive to the nonlinearity. Going back to Fig. 13, it is clear that, until an energy of $10^{-1}C_0$, the mode shapes and natural frequencies are unaffected by the nonlinearity and can safely be used. Beyond this critical energy level, both the in-phase and out-of-phase modes show a significant departure from the LNMs and become dependent on the total energy in the system.

As another example, the FEP of system
\[
\ddot{x}_1 + (2x_1 - x_2) = 0 \\
\ddot{x}_2 + (2x_2 - x_1 - x_3) + 0.5x_2^3 = 0 \\
\ddot{x}_3 + (2x_3 - x_2) = 0
\]

is depicted in Fig. 21. The linear modal parameters remain unchanged until approximately an energy of $10^{-1}$. Another interesting finding is that the nonlinearity has no influence whatsoever either on the frequency or on the mode shape of the second mode.

4.2. Nonlinear modal analysis

When it is certain that the system is excited in the nonlinear range, the linear framework should be abandoned in favor of a nonlinear modal analysis. Any attempt to apply traditional linear analysis in this context results, at best, in a suboptimal design.

Considering again system (2) as a first example, its FEP in Fig. 13 greatly helps to understand how the modal curves deform under the action of the nonlinearity. The in-phase NNM tends to localize to the second DOF, whereas the out-of-phase NNM localizes to the first DOF. Regarding the corresponding frequency of oscillation, both modes are characterized by a hardening behavior due to the presence of the cubic spring.

As a second example, a planar cantilever beam discretized by 20 finite elements and with a cubic spring at the free end is now considered (see Table 1 for the geometrical and mechanical properties). This models a real nonlinear beam that was used as a benchmark for nonlinear system identification during the European action COST F3 [99]. The first two modes are

![Figure 21](image-url)
plotted in the FEPs of Figs. 22 and 23, respectively. Considering the same energy level, the first modal curve seems somewhat more affected by the nonlinearity compared to the second modal curve. Their frequencies of oscillation undergo a strong increase with increasing energy levels. The FEPs also highlight the presence of two tongues, revealing the existence of internal resonances. The tongue in Fig. 22 corresponds to a 5:1 internal resonance between the first and second modes of the beam. When the energy gradually increases along the tongue, a smooth transition from the first mode to the second mode occurs following a dynamical mechanism similar to that described in Section 2.2.2. Similarly, a 5:1 internal resonance between the second and fourth modes is observed in Fig. 23. These internal resonances occur despite that the linear natural frequencies are not commensurate, as also discussed in Section 2.2.2.

These two examples demonstrate that such a nonlinear modal analysis is an important tool for thoroughly understanding the system’s vibratory response in the nonlinear regime. Clearly, this cannot be achieved using linearization procedures. However, because the general motion of a nonlinear system cannot be expressed as a superposition of individual NNM motions and because the modes in all these figures are computed based on the underlying undamped system, the practical utility of the nonlinear modal analysis might appear, at first, questionable.
A first motivation to compute and exploit the NNMs is that forced resonances in nonlinear systems occur in their neighborhoods. The knowledge of the NNMs can therefore provide valuable insight into the structure of the resonances, a feature of considerable engineering importance [17]. For illustration, system (4) is considered. In Figs. 24 and 25, the backbone \( S_{11}^+ \) of the frequency–energy plot is superposed to the nonlinear FRFs of Figs. 8 and 9. It can be observed that the backbone of the FEP traces the locus of the frequency response peaks for both the in-phase and out-of-phase modes. Furthermore, Fig. 26 compares the forced response of the system close to the first resonance (for \( F = 0.1 \), see the square in Fig. 24) to the free response of the corresponding point of the NNM backbone. An excellent agreement is obtained between the two types of motion.

A second motivation is that the damped dynamics closely follows the NNMs of the underlying undamped system. To demonstrate this, a time–frequency analysis method, the continuous wavelet transform (CWT) [100], is used. In contrast to the Fourier transform, which assumes signal stationarity, the CWT involves a windowing technique with variable-sized

---

**Fig. 24.** Nonlinear frequency response functions close to the first resonant frequency (five different forcing amplitudes: 0.002, 0.01, 0.05, 0.1, 0.2 N). The dashed line is the backbone \( S_{11}^+ \) of the frequency–energy plot. Left plot: \( x_1 \); right plot: \( x_2 \).

**Fig. 25.** Nonlinear frequency response functions close to the second resonant frequency (five different forcing amplitudes: 0.002, 0.01, 0.05, 0.1, 0.2 N). The dashed line is the backbone \( S_{11}/C_0 \) of the frequency–energy plot. Left plot: \( x_1 \); right plot: \( x_2 \).

**Fig. 26.** Free \((F = 0)\) and forced responses \((F = 0.1)\) of system (4) in the configuration space. ---, forced response; ----, free response.
regions. Small time intervals are considered for high-frequency components, whereas the size of the interval is increased for lower-frequency components. The CWT can therefore track the temporal evolution of the instantaneous frequencies, which makes it an effective tool for analyzing nonlinear signals. The usual representation of the transform is to plot its modulus as a function of time and frequency in a three-dimensional or contour plot. To use the CWT in conjunction with the FEP, a different representation is proposed herein. The CWT is represented in a frequency–energy plot by substituting the instantaneous energy in the system for time.

The free response of system

\[
\ddot{x}_1 + 0.03\dot{x}_1 + (2x_1 - x_2) + 0.5x_1^3 = 0 \\
\ddot{x}_2 + 0.01\dot{x}_2 + (2x_2 - x_1) = 0
\]

is depicted in Figs. 27 and 28 for an excitation of the in-phase and out-of-phase NNMs, respectively. The left plot is the theoretical FEP, that is the FEP computed from the equations of motion. The right plot is the ‘experimental’ FEP, calculated directly from the time series: (i) the backbone is provided by the CWT, and (ii) the modal curves are obtained by representing the time series in the configuration space for one oscillation around a specific energy level. For comparison, the theoretical backbone is represented by a solid line in the experimental FEP. A perfect agreement is obtained between the two FEPs, which shows that the undamped NNMs are attractors for the damped trajectories. In the present case, the modal damping ratios are 1% and 0.6%, but we note that this result holds for higher damping ratios.

Fig. 29 displays the free response of the planar cantilever beam excited at its first mode (with a damping matrix equal to the mass matrix, \(C = M\)). It shows that similar conclusions can also be reached for more complex systems.

Even if a possible criticism of the proposed approach is that it defines an NNM as a periodic solution of the underlying undamped system, these two examples support that they still give a very accurate picture of the damped dynamics. These results also show that the CWT is the ideal companion to the NNMs. We believe that the combined use of the FEP and the CWT represents a suitable framework for developing a new nonlinear system identification method, which could be viewed as a practical nonlinear analog of experimental modal analysis.
For instance, one specific application that could ultimately benefit from the proposed advancements is aircraft ground vibration testing (GVT)\cite{101}. GVTs are performed on aircraft prototypes before their first flight and provide critical data for flutter analysis. Identification of an accurate model during GVTs could allow the effects of nonlinearity on flutter to be explored by calculation prior to the flight test programme. Such an improvement would increase the aeroelastic prediction capabilities.

### 4.3. Reduced-order modeling

In a recent series of works\cite{39,48,67,81,97,102}, it was shown that NNMs can provide effective bases for constructing reduced-order models of the dynamics of discrete and continuous nonlinear oscillators.

Specifically, Touzé et al. performed a comparative study of reduced-order models of large-amplitude vibrations of shell structures of different configurations using either LNMs or NNMs\cite{102}. They showed that one or two NNMs were sufficient for accurately capturing the shell dynamics, and even the bifurcation structure of the dynamics that resulted from the nonlinear interaction of two shell modes in internal resonance. By contrast, multiple linear modes were necessary to achieve the same accuracy. For illustration, a specific application taken from\cite{102} is shown in Fig. 30. It depicts the frequency response curve of the nondimensionalized amplitude of the transverse displacement of a hyperbolic paraboloid panel under harmonic excitation. The harmonic excitation is applied at the center of the panel, and its frequency is in the vicinity of the first eigenfrequency. Comparing the reference (exact) computational solution to reduced-order models based on the leading LNM and NNM, respectively, the accuracy of the NNM-based model and its superiority over the LNM-based model are established. In this example, 15 LNMs were required to obtain results of similar accuracy.

These results demonstrate that NNMs hold promise for low-order reduction of structural models with many DOFs (e.g., finite element computational models). It is the application which has received the most attention so far in the literature. The reader can refer for instance to\cite{34,39,42–44,46,102–106} and references therein for further detail. Even though NNMs do not possess orthogonality properties (as do the LNMs), the resulting models are still expected to be much
more accurate compared to their linear counterpart (especially for systems with strong or even nonlinearizable nonlinearities). The reason for the enhanced accuracy of NNM-based reduced models lies in their invariance property and in the fact that they represent exact solutions of the free and forced nonlinear dynamics of the oscillators considered (i.e., oscillations in the neighborhoods of structural resonances can be captured by either a single NNM or a small set of NNMs when internal resonances occur).

4.4. Study of localization phenomena

Localization and motion confinement are observed when vibrational energy tends to be confined to one particular area of a structure. They have first been observed for periodic linear structures presenting a structural irregularity (e.g., mistuned bladed disks [107]). Nonlinear localization has been studied extensively by Vakakis and co-workers (see, e.g. [17,108–110]). One of its distinctive features is that it can occur in periodic structures even in the absence of structural disorder.

Although the energy is shared between the two oscillators at low-energy, the NNMs of system (2) localize to either DOF for increasing energies. As evidenced in Fig. 13, the in-phase NNM tends to localize to the second DOF, whereas the out-of-phase NNM localizes to the first DOF. Clearly, this localization property is to be attributed to the frequency–energy dependence of nonlinear oscillations.

The fact that some NNMs spatially confine vibrational energy can find applications in vibration mitigation of mechanical systems. For instance, the nonlinear energy pumping phenomenon directly exploits this property in order to transfer irreversibly vibrational energy from a primary structure to a nonlinear vibration absorber (see, e.g. [33,40,64,65,111]). Some authors are also exploiting mode localization for the design of microelectromechanical systems (MEMS) [37].

5. Conclusion

To robustly and accurately model nonlinearity in realistic vibrating structures is one of the greatest challenges in structural engineering. In this context, NNMs certainly represent a useful framework for the dynamicist. They have a clear conceptual relation to the linear normal modes, yet they can highlight nonlinear phenomena that are unexpected (and unexplainable) from a linear viewpoint.

The two main definitions, the fundamental properties and different analytical and numerical methods for computing NNMs were reviewed and illustrated with numerical examples. We have also highlighted that even seemingly simple nonlinear systems can exhibit very complicated dynamics. The 2DOF system investigated herein is characterized by an intricate NNM structure with (presumably) a countable infinity of internal resonances and strong motion localization in either oscillators. One interesting finding is that the internal resonances occur without necessarily having commensurate linear natural frequencies. This is rarely discussed in the literature and is a consequence of the frequency–energy dependence of the NNMs.

Because there is very little work that addresses the application of NNMs to real-word structures, we have identified several aspects that might drive their development in the future:

- Algorithms for the continuation of periodic solutions provide a very accurate computation of the NNMs of strongly nonlinear systems. Despite their computational burden, they certainly pave the way for an effective and practical computation of the nonlinear modes. They are described and discussed in more details in Part II of this study.
- The wavelet transform is a versatile time–frequency analysis method that can track the temporal evolution of the frequency of oscillation of NNMs.
- A frequency–energy plot is a suitable tool to represent the NNMs and to interpret the dynamics of nonlinear systems.

Using the combination of these three tools, we can relate the damped dynamics to the different branches of periodic solutions in the frequency–energy plot. These tools should help to extend experimental modal analysis, which is well-established for linear systems, to a practical nonlinear analog based on force appropriation.

One of the most limiting features of NNMs is that the general motion of a nonlinear system cannot be expressed as a superposition of individual NNMs. Even if there is no reason to believe that this limitation will be resolved soon, this paper has shown that the NNMs still provide a valuable tool for understanding (and possibly exploiting) the effects of structural nonlinearities on the dynamics.

Acknowledgments

The authors would like to thank Professors L.A. Bergman and D.M. McFarland, and Dr. Lee, Nucera and Panagopoulos for all the constructive discussions on the NNM concept.
References


