Available online at www.sciencedirect.com
ScienceDirect
Probabilistic Engineering Mechanics 23 (2008) 289–306
PROBABILISTIC ENGINEERING MECHANICS
www.elsevier.com/locate/probengmech

New partial differential equations governing the joint, response–excitation, probability distributions of nonlinear systems, under general stochastic excitation☆

Themistoklis P. Sapsis[a], Gerassimos A. Athanassoulis[b,*]

[a] *Massachusetts Institute of Technology, Department of Mechanical Engineering, United States*
[b] *National Technical University of Athens, Department of Naval Architecture and Marine Engineering, Heroon Polytechniou 9, 15773 Athens, Greece*

Received 19 December 2007; accepted 19 December 2007
Available online 10 January 2008

Abstract

In the present work the problem of determining the probabilistic structure of the dynamical response of nonlinear systems subjected to general, external, stochastic excitation is considered. The starting point of our approach is a Hopf-type equation, governing the evolution of the joint, response–excitation, characteristic functional. Exploiting this equation, we derive new linear partial differential equations governing the joint, response–excitation, characteristic (or probability density) function, which can be considered as an extension of the well-known Fokker–Planck–Kolmogorov equation to the case of a general, correlated excitation and, thus, non-Markovian response character. These new equations are supplemented by initial conditions and a marginal compatibility condition (with respect to the known probability distribution of the excitation), which is of non-local character. The validity of this new equation is also checked by showing its equivalence with the infinite system of moment equations. The method is applicable to any differential system, in state-space form, exhibiting polynomial nonlinearities. In this paper the method is illustrated through a detailed analysis of a simple, first-order, scalar equation, with a cubic nonlinearity. It is also shown that various versions of Fokker–Planck–Kolmogorov equation, corresponding to the case of independent-increment excitations, can be derived by using the same approach.

A numerical method for the solution of these new equations is introduced and illustrated through its application to the simple model problem. It is based on the representation of the joint probability density (or characteristic) function by means of a convex superposition of kernel functions, which permits us to satisfy *a priori* the non-local marginal compatibility condition. On the basis of this representation, the partial differential equation is eventually transformed to a system of ordinary differential equations for the kernel parameters. Extension to general, multidimensional, dynamical systems exhibiting any polynomial nonlinearity will be presented in a forthcoming paper.

Keywords: Stochastic dynamics; Stochastic differential equations; Functional Differential equations; Correlated stochastic excitation; Generalized Fokker–Planck–Kolmogorov equation; Non-Markovian responses; Characteristic functional; Kernel density functions

1. Introduction

Many problems occurring in applied sciences and engineering are successfully modelled as stochastic differential equations. A very important class of such problems are those modelled as stochastically excited, nonlinear, dynamical systems. Well-known examples include the dynamic responses of ships and other man-made structures and systems under the influence of wind-generated waves in the sea [46,58,5,1], the dynamic responses of buildings and bridges under the influence of earthquakes [36,16,31], as well as the dynamic responses of structures and vehicles under the influence of wind forces [48,33,52,27]. In all these cases the excitation loads are assumed to be known stochastic processes, either Gaussian or non-Gaussian, as in the case of wind loads. Their probabilistic and correlation structure can be (and, usually, have been)

☆ This work was initiated in 2004 in the context of the Diploma Thesis of TPS, under the supervision of GAA at the National Technical University of Athens, Dept. of Naval Architecture and Marine Engineering.

* Corresponding author. Tel.: +30 210 7721136; fax: +30 210 7721397.
E-mail addresses: sapsis@mit.edu (Th.P. Sapsis), mathan@central.ntua.gr (G.A. Athanassoulis).

0266-8920/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.probengmech.2007.12.028

inferred by means of statistical data analysis and, in most cases, have been conveniently parameterized for easy reference and use in calculations. Most of the foundational facts and aspects concerning the stochastic modelling philosophy in engineering and applied science, and the corresponding mathematical background can be found nowadays in book form; see, e.g., [33,49, 52,44].

The ultimate objective in the analysis of such problems is to obtain a complete probabilistic description of the response process, permitting us to answer any important questions about the response dynamics. Examples of such questions concerns the distributions of local extrema, of upcrossing rates at certain levels, of the first passage time associated with a critical level value, etc. To make this possible we need, in principle, to know the whole Kolmogorov hierarchy of the n-fold, joint, probability distributions $F_{x(t_1)x(t_2)\ldots x(t_n)}(a_1, a_2, \ldots, a_n)$ of the n-variate response random variables $(x(t_1), x(t_2), \ldots, x(t_n))$ at any collection of time instances $(t_1, t_2, \ldots, t_n)$ or, equivalently and more concisely, the Characteristic Functional (Ch.Fl) of the response process. Because of the obvious difficulties of this general concept of solution of the probabilistic dynamics problems, there is a constant tendency – at least in the applied and engineering literature – to avoid such an approach, resorting to simpler (partial) solution concepts.

An important and extensively studied context, permitting a relatively easy characterization of the probabilistic responses of a dynamical system, occurs if we assume that the excitation is a process with independent increments (see, e.g., [43,17, 51,24,25,30]). The key feature in this context is that the response vector, in the state-space formalism, is a Markovian process and, thus, its probability density function is governed by the Fokker–Planck–Kolmogorov (FPK) equation (in the case of a Gaussian excitation) or by reasonable extensions of the FPK equation (in the case of non-Gaussian excitation). Approximate techniques of solution of these equations have been thoroughly surveyed in [18]. Interestingly enough, there have been identified broad classes of problems in which analytic solutions of the classical FPK equation are available (see, e.g., [12,51,56,41]), making this approach even more attractive.

An approximate method dealing with nonlinear systems under general stochastic excitation is the Statistical Linearization Method (see, e.g., [44]), which is based on the approximation of the full system by a 'statistically equivalent' linear one. Some variations of the method, concerning local linearization in the phase space, have been recently presented [40], giving promising results. It is also possible to develop approximate solution schemes by replacing the given dynamical system by a "statistically equivalent" nonlinear system provided that the latter belongs to the class of problems which can be solved exactly. This method has been applied to various particular problems in the last three decades; see e.g. [37,13,14,44,42,59].

Another well-known and extensively used method that can be applied to any type of stochastic excitation and to any type of nonlinearity, is the method of moments, which reduces the initial stochastic dynamics problem to an infinite system of deterministic differential equations for the moment functions [7,43]. This infinite system is truncated and becomes closed (in the case of nonlinear problems) by means of appropriate closure schemes. Then, it is solved numerically, providing us with some (incomplete) information about the probabilistic characterization of the response process.

Another method, in principle well-known but in very little use for solving practical problems in stochastic dynamics, is the one based on the characteristic functional (Ch.Fl) of the full probability measure associated with the dynamic response process. The first step in this direction was made by Hopf [29] who derived a Functional Differential Equation (FDE) for the Ch.Fl associated with the probabilistic solution of the Navier–Stokes equations. This approach, known as *the statistical approach to turbulence*, has been developed further by many authors (see, e.g., [35,38,21,20]), and, eventually, led to the derivation and exploitation of various transport-diffusion equations for pdfs of the velocities and composition in turbulent reactive flows [32,39]. In parallel, a simpler version of the same approach has been developed and applied to finite-dimensional dynamical systems, governed by Stochastic Ordinary Differential Equations (SODEs). See, e.g., [7]. Such Hopf-type FDEs are always linear, and govern the Ch.Fl of the sought-for probability measure or – depending on the specific formulation – the Ch.Fl of the joint, response–excitation, probability measure. In recent years successful attempts have been reported towards the analytic determination of the response Ch.Fl for some classes of linear problems, even avoiding the explicit use of Hopf's FDE [11,8,9]. For some non-linear problems, the Ch.Fl can be expressed as a formal infinite-dimensional (functional) integral [38], which is of little (or no) practical use.

In this paper, Hopf's FDE is taken as the starting point of the probabilistic analysis of the considered stochastic dynamics problem. Because of the generality of Hopf's approach, the method is applicable to any (at least) polynomially non-linear system and any kind of stochastic excitation. Nevertheless, for reasons of simplicity and clarity, our study will be carried out on a specific, first-order, dynamical system, with cubic nonlinearity. The excitation process will be assumed, in principle, completely known, with a given (arbitrary) correlation structure and continuous (or smoother) sample functions. This implies a non-Markovian character of the response, making the approach based on the FPK equation inapplicable. Exploiting the Hopf FDE, new Partial Differential Equations (PDEs) governing the joint, response–excitation, characteristic functions (ch.f.), are derived. The corresponding equations for the joint pdfs are also obtained, by applying a Fourier transformation. These new PDEs, which are always linear, can be considered as a systematic and rigorous generalization of the FPK-type equations to the case of correlated excitation and non-Markovian responses. As an additional test of validity of these new PDEs, we show that they produce the correct infinite system of the moment equations. The same approach, i.e. starting from the Hopf FDE, is also applied to derive FPK equations, for the case of independent-increment excitation. A lack of rigor occurs here, when the sample functions of the excitation process are not continuous,

since the validity of the basic properties of the Ch.Fl exploited in our treatment are questionable in this case. Nevertheless, it is shown, in Section 6, that the equation obtained using our approach, by means of formal calculations, coincides with the extended FPK equation (for the same system) obtained by other, well-established methods [25]. This allows us to conjecture that a rigorous extension of the present theory to the case of independent-increment excitation might be possible, probably under some appropriate conditions.

Abbreviations

The following abbreviations – some of which have already been introduced above – will be consistently used in the following:

B-space	Banach space
ch.f(s)	characteristic function(s)
Ch.Fl(s)	characteristic functional(s)
F-derivative	Frechet derivative
FDE(s)	functional differential equation(s)
FPK	Fokker–Planck–Kolmogorov
ODE(s)	ordinary differential equation(s)
PDE(s)	partial differential equation(s)
pdf(s)	probability density function(s)
SODE(s)	stochastic ODE(s)

2. Preliminaries and notation

In this work we consider ODEs (systems) of the form (in state space formulation):

$$\dot{x}(t) = G(x(t)) + y(t), \qquad x(t_0) = x_0, \tag{2.1}$$

where x and y are scalar-valued or N-vector-valued, continuous (or smoother) functions, defined at least on an interval $I \equiv [t_0, T]$ (that is, $x, y : [t_0, T] \equiv I \to \mathbb{R}^N$), and $G : \mathbb{R}^N \to \mathbb{R}^N$, $N = 1$ or $N > 1$, is also a continuous (or smoother) function. Both the excitation $y(\cdot)$ and the initial conditions x_0 will be assumed known stochastic elements (function and variable, respectively). In contrast with the standard approach, followed in the case of an Ito SODE, the excitation $y(\cdot)$ is allowed to be smooth (e.g., k-times continuously differentiable), exhibiting any type of correlation structure in time. Thus, the sample functions $x(t)$ and $y(t)$ are considered as elements of smooth-function B-spaces, denoted by $\mathscr{X}$ and $\mathscr{Y}$, respectively. Our main results will refer to the case $N = 1$, $\mathscr{Y} = C^k(I)$, $I \subseteq \mathbb{R}$, $k = 0$ or $k > 0$, and $\mathscr{X}$ a similar space with smoother elements. The whole methodology can be extended to the vector case $N > 1$ with the usual trouble (see [4] for a detailed analysis of a second-order system).

The topological dual spaces of $\mathscr{X}$ and $\mathscr{Y}$ are also B-spaces and will be denoted by $\mathscr{X}' = \mathscr{U}$ and $\mathscr{Y}' = \mathscr{V}$. The symbols $\langle u, x\rangle$ and $\langle v, y\rangle$ denote the standard duality pairings between $\mathscr{X}$ and $\mathscr{U}$, and $\mathscr{Y}$ and $\mathscr{V}$, respectively.

The underlying probability space is denoted by $(\Omega, \mathscr{B}(\Omega), \mathscr{P}_\Omega)$, where Ω is an abstract version of the sample (trial) space, $\mathscr{B}(\Omega)$ is the family of Borel sets of Ω, and $\mathscr{P}_\Omega$ is the corresponding probability measure over Ω. The stochastic processes x and y are measurable maps $x, y : \Omega \to \mathscr{X}, \mathscr{Y}$, which define the induced probability spaces $(\mathscr{X}, \mathscr{B}(\mathscr{X}), \mathscr{P}_\mathscr{X})$ and $(\mathscr{Y}, \mathscr{B}(\mathscr{Y}), \mathscr{P}_\mathscr{Y})$, respectively. We shall also need and consider the joint process $x \times y : \Omega \to \mathscr{X} \times \mathscr{Y}$ with induced probability space $(\mathscr{X} \times \mathscr{Y}, \mathscr{B}(\mathscr{X} \times \mathscr{Y}), \mathscr{P}_{\mathscr{X}\times\mathscr{Y}})$. In the following we shall use the notation x or $x(\cdot)$ or $x(\cdot; \omega)$, and similarly for y, for the random element, and $x(t; \omega)$, $t \in [t_0, T] \equiv I \subseteq \mathbb{R}$, $\omega \in \Omega$, and similarly for y, for the sample functions, in accordance with the needs of the discussion.

The finite-dimensional distributions, densities and characteristic functions of the random element $x(\cdot; \omega)$ will be denoted by $F_{x(t_1)\ldots x(t_M)}(\alpha_1, \ldots, \alpha_M)$, $f_{x(t_1)\ldots x(t_M)}(\alpha_1, \ldots, \alpha_M)$ and $\phi_{x(t_1)\ldots x(t_M)}(\upsilon_1, \ldots, \upsilon_M)$, respectively. This implies a convenient notation for the joint random element $(x(\cdot; \omega), y(\cdot; \omega))$; for example $f_{x(t_1)x(t_2)y(t_1)y(t_2)y(t_3)}(\alpha_1, \alpha_2, \beta_1, \beta_2, \beta_3)$ is the $(2 + 3)$-order joint probability density function, and $\phi_{x(t_1)x(t_2)y(t_1)y(t_2)y(t_3)}(\upsilon_1, \upsilon_2, \nu_1, \nu_2, \nu_3)$ is the corresponding characteristic function. The usual (finite-dimensional) mean value operator (ensample average) will be denoted by $\mathbf{E}^\omega[\cdot]$. For example, the mean value function of the random element $x(\cdot; \omega)$ will be written as $m_{x(t)} = \mathbf{E}^\omega[x(t; \omega)]$. Slight variations (simplifications) of this notation will be introduced later, in accordance with the needs of the presentation.

Infinite-dimensional (global) moments, are defined by integrating over the whole sample space $\mathscr{X}$ with respect to the probability measure $\mathscr{P}_\mathscr{X}$ (see, e.g., [33,55,19]). For example, the mean (first moment) $m_\mathscr{X}$ is defined to be this element of $\mathscr{X}$, for which the following scalar equation holds true:

$$\langle u, m_\mathscr{X}\rangle = \int_\mathscr{X} \langle u, x\rangle \, \mathscr{P}(\mathrm{d}x), \quad \forall u \in \mathscr{U}, \tag{2.2a}$$

where $\mathscr{U} \equiv \mathscr{X}'$. Furthermore, the correlation operator (second moment) is defined to be this linear operator $R_{\mathscr{X}\mathscr{X}} : \mathscr{U} \to \mathscr{X}$, for which the following scalar equation is valid $\forall u, w \in \mathscr{U}$:

$$\langle w, R_{\mathscr{X}\mathscr{X}} u\rangle = \int_\mathscr{X} \langle w, x\rangle \langle u, x\rangle \, \mathscr{P}(\mathrm{d}x). \tag{2.2b}$$

The integrals appearing in the right-hand side of Eq. (2.2a) are infinite-dimensional (functional) integrals over B-spaces. (For detailed definitions and conditions ensuring existence of these integrals see references stated above or [15].) In general, the functional integral of any bounded, measurable, continuous functional $\mathscr{G} : \mathscr{X} \to \mathbb{C}$, with respect to a probability measure $\mathscr{P}$, is well defined, and will be denoted by $\int_\mathscr{X} \mathscr{G}(x) \, \mathscr{P}(\mathrm{d}x)$.

Measures and integrals over infinite-dimensional vector spaces are related with the corresponding finite-dimensional ones through the concepts of cylinder sets, cylinder measures and cylinder functionals. Let $\mathscr{X}$ be a separable B-space, $\mathscr{U}$ be the dual of $\mathscr{X}$, and $u_1, \ldots, u_Q$, be Q linearly independent elements of $\mathscr{U}$. Then, to any element $x \in \mathscr{X}$ we associate the Q-dimensional projection $\Pi_{u_1,\ldots,u_Q} : \mathscr{X} \to \mathbb{R}^Q$, defined by

$$\Pi_{u_1,\ldots,u_Q}[x] = \left(\langle u_1, x\rangle, \ldots, \langle u_Q, x\rangle\right). \tag{2.3}$$

The inverse of $\Pi_{u_1,\ldots,u_Q}[\cdot]$, applied to the Borel sets $B\left(\mathbb{R}^Q\right)$, defines the cylinder sets of $\mathscr{X}$. The existence of a probability measure $\mathscr{P}_\mathscr{X}$ on $\mathscr{X}$ implies the existence of Q-dimensional

(marginal) measures $P_{u_1,\ldots,u_Q}$ on $\mathbb{R}^Q$, associated with the random vectors $\left(\langle u_1, x(\cdot;\omega)\rangle, \ldots, \langle u_Q, x(\cdot;\omega)\rangle\right)$ by means of the relation

$$P_{u_1,\ldots,u_Q}(E_Q) = \mathscr{P}_{\mathscr{X}}\left(\Pi^{-1}_{u_1,\ldots,u_Q}[E_Q]\right) \tag{2.4}$$

for any $E_Q \in \mathscr{B}(\mathbb{R}^Q)$.

Consider now an arbitrary cylinder functional $\mathscr{G}: \mathscr{X} \to \mathbb{C}$, that is a functional of the form

$$\mathscr{G}(x) = g\left(\langle u_1, x\rangle, \ldots, \langle u_Q, x\rangle\right), \quad x \in \mathscr{X}, \tag{2.5}$$

where $g: \mathbb{R}^Q \to \mathbb{C}$ is an arbitrary, measurable, integrable function. In this case, the infinite-dimensional integral of $\mathscr{G}(x)$ with respect to the probability measure $\mathscr{P}$ over the space $\mathscr{X}$, can be expressed as a Q-dimensional integral by means of the formula:

$$\int_{\mathscr{X}} \mathscr{G}(x)\,\mathscr{P}(\mathrm{d}x) = \int_{\mathbb{R}^Q} g(\boldsymbol{a})\,P_{u_1,\ldots,u_Q}(\mathrm{d}\boldsymbol{a}). \tag{2.6}$$

Eqs. (2.5) and (2.6) provide us with a powerful method for evaluating integrals over infinite-dimensional (function) spaces. They will be referred to as the (Q-dimensional) *Projection Theorem*.

3. A brief review on the characteristic functional and its basic properties

For stochastic processes with values in a linear space (as e.g. the space of continuous functions $C(I)$, $I = [t_0, T]$), the full probability measure is completely characterised by means of the corresponding characteristic functional. See, e.g., [23, Chapter V, Sec.3], [55, Chapter IV]. In this section we recall the definition and some basic properties of the Ch.Fl for probability measures defined on separable B-spaces.

3.1. Definition of the characteristic functional

Definition 3.1. Let $\mathscr{X}$ be a separable B-space and $\mathscr{P} = \mathscr{P}_{\mathscr{X}}$ be a probability measure defined on it. The Ch.Fl $\mathscr{F}$ of $\mathscr{P}$ is a cylinder functional defined on the dual space $\mathscr{X}' = \mathscr{U}$ by the formula

$$\mathscr{F}(u) = \int_{\mathscr{X}} \mathrm{e}^{\mathrm{i}\langle u,x\rangle}\,\mathscr{P}(\mathrm{d}x), \quad u \in \mathscr{U}. \tag{3.1}$$

This integral always exists provided that the corresponding probability measure is well defined.

3.2. Infinite-dimensional (global) moments

Let the Ch.Fl be differentiable in the sense of Frechet. In order to calculate the F-derivative $D\mathscr{F}(u)$, we make use of the Gateaux derivative (which always exists for a F-differentiable map). Thus, we have

$$D\mathscr{F}(u)[z] = \left.\frac{\mathrm{d}\mathscr{F}(u+\varepsilon z)}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} = \mathrm{i}\cdot\int_{\mathscr{X}} \langle z,x\rangle\,\mathrm{e}^{\mathrm{i}\langle u,x\rangle}\,\mathscr{P}(\mathrm{d}x), \quad u,z \in \mathscr{U}. \tag{3.2}$$

Setting $u = 0$, we obtain

$$D\mathscr{F}(0)[z] = \mathrm{i}\cdot\int_{\mathscr{X}} \langle z,x\rangle\,\mathscr{P}(\mathrm{d}x), \quad z \in \mathscr{U}. \tag{3.3}$$

Since $D\mathscr{F}(0)[z]$ is a continuous, linear functional with respect to z, there should exist an element $m \in \mathscr{X}$, such that

$$\langle z, m\rangle = -\mathrm{i}\cdot D\mathscr{F}(0)[z] = \int_{\mathscr{X}} \langle z,x\rangle\,\mathscr{P}(\mathrm{d}x). \tag{3.3$'$}$$

Comparing the above equation with Eq. (2.2a), it is easily seen that the element $m \in \mathscr{X}$ of Eq. (3.3)$'$ coincides with the mean value $m_{\mathscr{X}}$ of the probability measure $\mathscr{P}$. The correlation operator $R_{\mathscr{X}\mathscr{X}}$ can be associated in a similar way with the second F-derivative of the Ch.Fl. In this case we have

$$\langle w, R_{\mathscr{X}\mathscr{X}} z\rangle = -D^2\mathscr{F}(0)[z,w] = \int_{\mathscr{X}} \langle w,x\rangle\langle z,x\rangle\,\mathscr{P}(\mathrm{d}x), \quad z,w \in \mathscr{U}.$$

3.3. Finite-dimensional (point) moments

In the case where the space $\mathscr{X}$ is a function space, apart from infinite-dimensional (global) moments, we are also interested in finite-dimensional moments associated with finite-dimensional projections $(x(t_1;\omega), x(t_2;\omega), \ldots, x(t_n;\omega))$, for any set of time instances $(t_1, t_2, \ldots, t_n)$. This kind of moments can be obtained also by differentiating the Ch.Fl, this time using Volterra functional derivatives. (See, e.g., [57] or [7].) Volterra derivatives, e.g. the first-order one $\delta\mathscr{F}(u)/\delta u(t)$, can be calculated either by applying the original definition to the functional, or by applying the Frechet derivative $D\mathscr{F}(u)[z]$ at $z(\cdot) = \delta(\cdot - t)$. Following the second approach, and using Eqs. (3.3) and (2.6), we obtain

$$\frac{\delta\mathscr{F}(0)}{\delta u(t)} \overset{\text{def}}{=} D\mathscr{F}(0)[\delta(\cdot - t)] = \mathrm{i}\int_{\mathscr{X}} x(t)\,\mathscr{P}(\mathrm{d}x) = \mathrm{i}\int_{\mathbb{R}} a\,\mathrm{d}F_{x(t)}(a) = \mathrm{i}\mathbf{E}^{\omega}[x(t;\omega)]$$

and thus

$$\mathbf{E}^{\omega}[x(t;\omega)] = \frac{1}{\mathrm{i}}\frac{\delta\mathscr{F}(0)}{\delta x(t)}. \tag{3.4a}$$

Similarly we obtain

$$\mathbf{E}^{\omega}[x(t_1;\omega)\,x(t_2;\omega)] = \frac{1}{\mathrm{i}^2}\cdot D^2\mathscr{F}(0)[\delta(\cdot - t_1), \delta(\cdot - t_2)] = \frac{1}{\mathrm{i}^2}\frac{\delta^2\mathscr{F}(0)}{\delta u(t_1)\,\delta u(t_2)}, \tag{3.4b}$$

as well as analogous expressions for higher-order moments. Working similarly, and using appropriate generalized functions, we can derive equations for higher-order moments involving both the values of the random element at some time instances, and the values of its derivatives either at the same or at different

time instances. As an example we give the formula:

$$\mathbf{E}^{\omega}\left[x'\left(t_1;\omega\right)x\left(t_2;\omega\right)\right]=\frac{1}{(-\mathrm{i})\,\mathrm{i}}\cdot D^2F\left(0\right)\times\left[\delta'\left(\cdot-t_1\right),\delta\left(\cdot-t_2\right)\right].$$

4. Hopf-type equation for the characteristic functional

In order to illustrate the derivation of Hopf-type FDEs for nonlinear dynamical systems, and pave the way to the next section, where these equations will be exploited to produce new PDEs for finite-dimensional ch.fs, we shall restrict ourselves here to a specific case of a simple (scalar, first-order) dynamical system having a cubic nonlinearity, which is described by the following SODE:

$$x'\left(t;\omega\right)+\mu x\left(t;\omega\right)+kx^3\left(t;\omega\right)=y\left(t;\omega\right), \tag{4.1a}$$

$$x\left(t_0;\omega\right)=x_0\left(\omega\right), \tag{4.1b}$$

where μ,k are deterministic constants, $x_0\left(\omega\right)$ is a random variable with known ch.f $\phi_0\left(\upsilon\right),\upsilon\in\mathbb{R}$, and the excitation $y\left(\cdot,\omega\right)$ is a real-valued random function, with sample space $\mathscr{Y}$, probability measure $\mathscr{P}_y$, and Ch.Fl $\mathscr{F}_y\left(v\right),v\in\mathscr{Y}'=\mathscr{V}$. The sample space $\mathscr{Y}$ can be taken to be a quite general, separable, B-space. In the present work, it will be taken as a space $Y=C^k\left(I\right),I\subseteq\mathbb{R}$, for some $k\in\mathbb{N}\cup\{0\}$.

Standard existence and uniqueness theory (see, e.g., [10], or [49]) ensure that there is a stochastic process $x\left(\cdot;\omega\right)$, with sample space $\mathscr{X}=C^{k+1}\left(I\right)$ and probability measure $\mathscr{P}_x$, and a joint probability space $\left(\mathscr{X}\times\mathscr{Y},\mathscr{B}\left(\mathscr{X}\times\mathscr{Y}\right),\mathscr{P}_{xy}\right)$, such that the joint process $\left(x\left(\cdot;\omega\right),y\left(\cdot;\omega\right)\right)$ verifies the SODE (4.1).

The joint, response–excitation, probability measure $\mathscr{P}_{xy}$ is equivalently described by the joint Ch.Fl

$$\mathscr{F}_{xy}\left(u,v\right)=\int_{\mathscr{Y}}\int_{\mathscr{X}}\mathrm{e}^{\mathrm{i}\left(\langle u,x\rangle+\langle v,y\rangle\right)}\mathscr{P}_{xy}\left(\mathrm{d}x,\mathrm{d}y\right). \tag{4.2}$$

We shall now use the SODE (4.1) in order to obtain an FDE for $\mathscr{F}_{xy}\left(u,v\right)$. Let us consider the Volterra u-partial derivative of $\mathscr{F}_{xy}$ at time t:

$$\frac{\delta\mathscr{F}_{xy}\left(u,v\right)}{\delta u\left(t\right)}=\int_{\mathscr{Y}}\int_{\mathscr{X}}\mathrm{i}x\left(t\right)\mathrm{e}^{\mathrm{i}\left(\langle u,x\rangle+\langle v,y\rangle\right)}\mathscr{P}_{xy}\left(\mathrm{d}x,\mathrm{d}y\right). \tag{4.3}$$

Since the sample space $\mathscr{X}$ consists of smooth functions, we can differentiate (4.3) with respect to t, obtaining:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\delta\mathscr{F}_{xy}(u,v)}{\delta u(t)}=\int_{\mathscr{Y}}\int_{\mathscr{X}}\mathrm{i}x'(t)\mathrm{e}^{\mathrm{i}\left(\langle u,x\rangle+\langle v,y\rangle\right)}\mathscr{P}_{xy}(\mathrm{d}x,\mathrm{d}y). \tag{4.4}$$

Further, we compute the three-fold u-partial Volterra derivative of $\mathscr{F}_{xy}\left(u,v\right)$ at time instants $t_1,t_2,t_3\in I$:

$$\frac{\delta^3\mathscr{F}_{xy}\left(u,v\right)}{\delta u\left(t_1\right)\delta u\left(t_2\right)\delta u\left(t_3\right)}=\int_{\mathscr{Y}}\int_{\mathscr{X}}\mathrm{i}x\left(t_1\right)\mathrm{i}x\left(t_2\right)\mathrm{i}x\left(t_3\right)\times\mathrm{e}^{\mathrm{i}\left(\langle u,x\rangle+\langle v,y\rangle\right)}\mathscr{P}_{xy}\left(\mathrm{d}x,\mathrm{d}y\right). \tag{4.5}$$

Setting $t_1=t_2=t_3=t$ in the latter, and combining with Eqs. (4.3), (4.5) and (4.1a), we get

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}t}\frac{\delta\mathscr{F}_{xy}\left(u,v\right)}{\delta u\left(t\right)}+\mu\frac{\delta\mathscr{F}_{xy}\left(u,v\right)}{\delta u\left(t\right)}-k\frac{\delta^3\mathscr{F}_{xy}\left(u,v\right)}{\delta u\left(t\right)^3}\\&=\mathrm{i}\int_{\mathscr{Y}}\int_{\mathscr{X}}\left[x'\left(t\right)+\mu x\left(t\right)+kx^3\left(t\right)\right]\times\mathrm{e}^{\mathrm{i}\left(\langle u,x\rangle+\langle v,y\rangle\right)}\mathscr{P}_{xy}\left(\mathrm{d}x,\mathrm{d}y\right)\\&\overset{(4.1a)}{=}\mathrm{i}\int_{\mathscr{Y}}\int_{\mathscr{X}}y\left(t\right)\mathrm{e}^{\mathrm{i}\left(\langle u,x\rangle+\langle v,y\rangle\right)}\mathscr{P}_{xy}\left(\mathrm{d}x,\mathrm{d}y\right).\end{aligned} \tag{4.6}$$

Clearly, the last double functional integral can be expressed as a v-partial Volterra derivative:

$$\int_{\mathscr{Y}}\int_{\mathscr{X}}\mathrm{i}y\left(t\right)\mathrm{e}^{\mathrm{i}\left(\langle u,x\rangle+\langle v,y\rangle\right)}\mathscr{P}_{xy}\left(\mathrm{d}x,\mathrm{d}y\right)=\frac{\delta\mathscr{F}_{xy}\left(u,v\right)}{\delta v\left(t\right)}. \tag{4.7}$$

Combining (4.6) and (4.7) we derive the sought-for, Hopf-type, FDE that governs the joint Ch.Fl $\mathscr{F}_{xy}\left(u,v\right)$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\delta\mathscr{F}_{xy}\left(u,v\right)}{\delta u\left(t\right)}+\mu\frac{\delta\mathscr{F}_{xy}\left(u,v\right)}{\delta u\left(t\right)}-k\frac{\delta^3\mathscr{F}_{xy}\left(u,v\right)}{\delta u\left(t\right)^3}=\frac{\delta\mathscr{F}_{xy}\left(u,v\right)}{\delta v\left(t\right)}. \tag{4.8a}$$

Eq. (4.8a) is a linear FDE involving Volterra functional derivatives, as well as ordinary time derivatives. The cubic nonlinearity of the initial SODE corresponds to the 3-fold Volterra derivative $\delta^3\mathscr{F}_{xy}/\delta u\left(t\right)^3$. From the above derivation it is clear that any nth-order polynomial nonlinearity of the initial differential equation is transformed to an n-fold Volterra derivative in the corresponding Hopf-type FDE. Another important feature of Eq. (4.8a) is that it holds true for any continuous functionals $u\in\mathscr{U},v\in\mathscr{V}$.

Eq. (4.8a) has to be supplemented by an appropriate initial condition, expressing that the probability measure associated with the initial state $x\left(t_0,\omega\right)$ is given. This condition can be implemented by means of the joint Ch.Fl $\mathscr{F}_{xy}\left(u,v\right)$ as follows. Setting $v=0$ (to restrict ourselves to the response process only) and $u=\upsilon\cdot\delta\left(\cdot-t_0\right),\upsilon\in\mathbb{R}$, (to concentrate only at the initial time instant), will result in

$$\begin{aligned}&\mathscr{F}_{xy}\left(\upsilon\delta\left(\cdot-t_0\right),0\right)\\&=\int_{\mathscr{Y}}\int_{\mathscr{X}}\mathrm{e}^{\mathrm{i}\left(\langle\upsilon\delta(\cdot-t_0),x\rangle+\langle 0,y\rangle\right)}\mathscr{P}_{xy}\left(\mathrm{d}x,\mathrm{d}y\right)\\&=\int_{\mathscr{X}}\mathrm{e}^{\mathrm{i}\langle\upsilon\delta(\cdot-t_0),x\rangle}\mathscr{P}_x\left(\mathrm{d}x\right)=\phi_0(\upsilon),\end{aligned}$$

where $\phi_0\left(\upsilon\right)$ is the ch.f of $x\left(t_0,\omega\right)=x_0\left(\omega\right)$. Hence, the initial condition can be expressed as

$$\mathscr{F}_{xy}\left(\upsilon\delta\left(\cdot-t_0\right),0\right)=\phi_0(\upsilon),\quad\upsilon\in\mathbb{R}. \tag{4.8b}$$

5. Derivation of new PDEs for joint response–excitation characteristic functions

In this section we shall exploit the Hopf-type FDE (4.8), obtained above, to derive new PDEs for the joint,

response–excitation, ch.f when the excitation is a known stochastic process either with a.e. continuous sample functions or smoother. In contrast with the case of an independent-increment excitation process, where the randomness of the excitation "regenerates" every time instant and allows us to write explicitly an equation involving only the response density (the well-known FPK equation), in the case of a stochastic excitation with smooth sample functions, the randomness evolves, in general, in a smoother way, as a result of the finite correlation time, making necessary to consider response and excitation jointly.

The causality principle dictates that the current value $x(t;\omega)$ of the response, depends only on the history of the excitation $y(t_0 \le s < t;\omega)$. However, this does not prevent the stochastic dependence between $x(t;\omega)$ and $y(t+\varepsilon;\omega)$, $\varepsilon > 0$, which is a natural result of the smoothness and the finite correlation time of the excitation, $C_{yy}(t+\varepsilon, t) \neq 0$.

We shall proceed to derive a PDE for the joint ch.f $\phi_{x(t)y(t)}(\upsilon,\nu)$ corresponding to the pair of random variables $(x(t;\omega), y(t;\omega))$, $t =$ fixed. To this end we apply Eq. (4.8a), above, to the pair

$$u = \upsilon\cdot\delta(\cdot - t), \qquad v = \nu\cdot\delta(\cdot - s), \tag{5.1}$$

$(\upsilon, \nu \in \mathbb{R})$ and take the limit $s \to t$, after some manipulations. For the first term of Eq. (4.8a) (see also Eq. (4.4)), we obtain

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\delta\mathscr{F}_{xy}(u,v)}{\delta u(t)}\right)\Bigg|_{\substack{u=\upsilon\cdot\delta(\cdot-t)\\ v=\nu\cdot\delta(\cdot-s)}}\\
&= \iint_{\mathscr{X}\times\mathscr{Y}} \mathrm{i}\frac{\mathrm{d}x(t)}{\mathrm{d}t}\exp\{\mathrm{i}\upsilon x(t) + \mathrm{i}\nu y(s)\}\mathscr{P}_{xy}(\mathrm{d}x,\mathrm{d}y)\\
&= \frac{1}{\upsilon}\frac{\partial}{\partial t}\left[\iint_{\mathscr{X}\times\mathscr{Y}} \exp\{\mathrm{i}\upsilon x(t) + \mathrm{i}\nu y(s)\}\mathscr{P}_{xy}(\mathrm{d}x,\mathrm{d}y)\right]\\
&= \left[\text{Projection Theorem}\right]\\
&= \frac{1}{\upsilon}\frac{\partial}{\partial t}\left[\iint_{\mathbb{R}\times\mathbb{R}} \exp\{\mathrm{i}\upsilon x + \mathrm{i}\nu y\} f_{x(t)y(s)}(x,y)\,\mathrm{d}x\mathrm{d}y\right]\\
&= \frac{1}{\upsilon}\frac{\partial\phi_{x(t)y(s)}(\upsilon,\nu)}{\partial t}.
\end{aligned}$$

Taking now the limit $s \to t$, we get

$$\begin{aligned}
&\lim_{s\to t}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\delta\mathscr{F}_{xy}(u,v)}{\delta u(t)}\right)\Bigg|_{\substack{u=\upsilon\cdot\delta(\cdot-t)\\ v=\nu\cdot\delta(\cdot-s)}}\\
&= \frac{1}{\upsilon}\frac{\partial\phi_{x(t)y(s)}(\upsilon,\nu)}{\partial t}\Bigg|_{s=t}.
\end{aligned} \tag{5.2}$$

Working similarly, we readily obtain the following results concerning the remaining terms appearing in Eq. (4.8a):

$$\frac{\delta\mathscr{F}_{xy}(\upsilon\cdot\delta(\cdot-t), \nu\cdot\delta(\cdot-t))}{\delta u(t)} = \frac{\partial\phi_{x(t)y(t)}(\upsilon,\nu)}{\partial\upsilon}, \tag{5.3}$$

$$\frac{\delta^3\mathscr{F}_{xy}(\upsilon\cdot\delta(\cdot-t), \nu\cdot\delta(\cdot-t))}{\delta u(t)^3} = \frac{\partial^3\phi_{x(t)y(t)}(\upsilon,\nu)}{\partial\upsilon^3}, \tag{5.4}$$

$$\frac{\delta\mathscr{F}_{xy}(\upsilon\cdot\delta(\cdot-t), \nu\cdot\delta(\cdot-t))}{\delta v(t)} = \frac{\partial\phi_{x(t)y(t)}(\upsilon,\nu)}{\partial\nu}. \tag{5.5}$$

Combining Eqs. (5.2)–(5.5) with the FDE (4.8a), we obtain the following PDE for the joint ch.f $\phi_{x(t)y(t)}(\upsilon,\nu)$, of the pair of random variables $(x(t;\omega), y(t;\omega))$, for every $t > t_0$:

$$\begin{aligned}
&\frac{1}{\upsilon}\frac{\partial\phi_{x(t)y(s)}(\upsilon,\nu)}{\partial t}\Bigg|_{s=t} + \mu\frac{\partial\phi_{x(t)y(t)}(\upsilon,\nu)}{\partial\upsilon}\\
&\quad - k\frac{\partial^3\phi_{x(t)y(t)}(\upsilon,\nu)}{\partial\upsilon^3} = \frac{\partial\phi_{x(t)y(t)}(\upsilon,\nu)}{\partial\nu}.
\end{aligned} \tag{5.6a}$$

Now, since the stochastic process $y(\cdot,\omega)$ is given, its ch.f $\phi_{y(t)}(\nu)$ is known. Hence, the y-marginal of the joint ch.f $\phi_{x(t)y(t)}(\upsilon,\nu)$ has to coincide with $\phi_{y(t)}(\nu)$, resulting in the following *marginal compatibility condition:*

$$\phi_{x(t)y(t)}(0,\nu) = \phi_{y(t)}(\nu), \quad \nu\in\mathbb{R}, t \ge t_0. \tag{5.6b}$$

In addition, the initial condition (4.8b) implies the following *initial condition* to $\phi_{x(t)y(t)}(\upsilon,\nu)$:

$$\phi_{x(t_0)y(t_0)}(\upsilon,0) = \phi_{x(t_0)}(\upsilon) = \phi_0(\upsilon), \quad \upsilon\in\mathbb{R}. \tag{5.6c}$$

Finally, two obvious, yet essential, conditions that the sought-for function $\phi_{x(t)y(s)}(\upsilon,\nu)$ should obey are the following:

$$\phi_{x(t)y(s)}(0,0) = 1, \quad t, s \ge t_0, \tag{5.6d}$$

$$\phi_{x(t)y(s)}(\upsilon,\nu) \quad \text{is non-negative definite w.r.t. } \upsilon,\nu, \text{ for any } t, s \ge t_0 \tag{5.6e}$$

which come directly from the fact that it is a characteristic function. The last two conditions will be referred to as *constitutive conditions*.

The above problem (5.6a)–(5.6e) can be equivalently reformulated in terms of the corresponding joint, response–excitation, pdf $f_{x(t)y(s)}(a,\beta)$. Recalling that $f_{x(t)y(s)}(a,\beta)$ and $\phi_{x(t)y(s)}(\upsilon,\nu)$ constitute a Fourier transform pair, i.e.

$$\phi_{x(t)y(s)}(\upsilon,\nu) = \mathscr{F}^{a\to\upsilon}_{\beta\to\nu}\left\{f_{x(t)y(s)}(a,\beta)\right\},$$

and applying the inverse Fourier transformation to (5.6a)–(5.6c) and (5.6e), we readily obtain the partial differential equation

$$\begin{aligned}
&\frac{\partial f_{x(t)y(s)}(a,\beta)}{\partial t}\Bigg|_{s=t} + \frac{\partial}{\partial a}\left[\left(\mu a + k a^3\right) f_{x(t)y(t)}(a,\beta)\right]\\
&\quad + \frac{\partial}{\partial a}\left[\beta f_{x(t)y(t)}(a,\beta)\right] = 0,
\end{aligned} \tag{5.7a}$$

the following alternative forms of the *marginal compatibility condition* and the *initial condition*

$$\int_{\mathbb{R}} f_{x(t)y(t)}(a,\beta)\,\mathrm{d}a = f_{y(t)}(\beta), \quad \beta\in\mathbb{R}, \ t \ge t_0, \tag{5.7b}$$

$$\int_{\mathbb{R}} f_{x(t_0)y(t_0)}(a,\beta)\,\mathrm{d}\beta = f_{x(t_0)}(a) = f_0(a), \quad a\in\mathbb{R}, \tag{5.7c}$$

as well as the corresponding new forms of the *constitutive conditions*

$$\int_{\mathbb{R}\times\mathbb{R}} f_{x(t)y(s)}(a,\beta)\,\mathrm{d}a\mathrm{d}\beta = 1, \quad t, s \ge t_0 \tag{5.7d}$$

$$f_{x(t)y(s)}(a,\beta) \ge 0, \quad \text{for any } a,\beta\in\mathbb{R} \text{ any } t, s \ge t_0. \tag{5.7e}$$

To the best of our knowledge, Eqs. (5.6a)–(5.6e) and (5.7a)–(5.7e), governing the evolution of the joint, response–excitation, ch.f $\phi_{x(t)y(t)}(\upsilon, \nu)$ and pdf $f_{x(t)y(t)}(a, \beta)$, first appeared in [3]. They can be considered as a new kind of mathematical model, providing us with the probabilistic characterization of the response $x(t, \omega), \omega \in \Omega$, for each $t \in I$, obtained by taking the y-marginal of the joint ch.f or the joint pdf. This mathematical model is valid for any kind of stochastic excitation with a.e. continuous (or smoother) sample functions, having any (known) probabilistic structure.

Although the mathematical analysis (solvability theory) of problem (5.6a)–(5.6e) or (5.7a)–(5.7e) is an open problem, existing numerical evidence, presented in Section 8 (see also [45]), suggests that it might be well-posed under reasonable assumptions.

In concluding this section we should emphasize that the above approach can be generalized in order to obtain similar, linear, PDEs for the joint, $N - x$ and $M - y$, ch.f

$$\phi_{x(t_1)\ldots x(t_N)y(s_1)\ldots y(s_M)}(\upsilon_1, \ldots \upsilon_N, \nu_1, \ldots, \nu_M)$$

(or the corresponding joint pdf), along with appropriate (marginal compatibility and initial) conditions. This point will be further discussed in forthcoming work [4]. It seems that in this way it is possible to construct a closed (finitely-solvable) hierarchy of linear problems providing us with the full hierarchy of the finite-dimensional probabilities of the stochastic response $x(\cdot; \omega)$.

6. Derivation of the FPK equation for the case of independent increment excitation

Eqs. (5.6) and (5.7) hold true for any kind of stochastic excitation process, provided that the latter has at least a.e. continuous sample functions. We shall now turn to the most studied case, those of an Ito SODE, where $y(t; \omega)$ represents the generalized derivative of an independent-increment process. In this case the response $x(t; \omega)$ is continuous but not differentiable. Thus, the treatment based on the Hopf equation, as developed in Section 5, is not valid, since the duality pairings (5.1) are not applicable. The question arises if it is possible to treat this case also by a similar method, starting from the Hopf equation and obtaining the usual FPK equation — which involves only the response ch.f (or pdf). In the present section we shall show how this is possible, by resorting back to the FDE for a finite-difference version of the SODE (4.1). The crucial property, to be exploited in this case, is the independence of the current response value $x(t; \omega)$ from the future increment $\Delta_\tau z(t; \omega) = z(t + \tau; \omega) - z(t; \omega), \tau > 0$, of the excitation. Everything presented in this section can be generalized to multidimensional nonlinear dynamical systems.

Let us rewrite the SODE (4.1a) and (4.1b) in a finite-difference form:

$$\frac{x(t + \tau; \omega) - x(t; \omega)}{\tau} + \mu x(t; \omega) + k x^3(t; \omega) = \frac{\Delta_\tau z(t; \omega)}{\tau}, \tag{6.1a}$$

$$x(t_0; \omega) = x_0(\omega), \tag{6.1b}$$

where $z(\cdot; \omega)$ is a known, real-valued process with independent increments, and $x_0(\omega)$ is a known random variable. The time increment τ is assumed to be positive, $\tau > 0$, and this is essential in what follows.

The sample functions of the stochastic process $z(\cdot; \omega)$ may be either continuous functions (as in the case of a Wiener process) or piecewise-continuous functions (as in the case of a general Levy process containing Poisson components). In the first case (continuous sample functions), it is clear that the previously developed approach can be applied to Eq. (6.1). In the second case (*cadlag* sample functions) the applicability of the same arguments is not yet clear. Since, however, the space of cadlag functions is given the structure of a Banach space [28] with a well-defined topological dual [53,54], the rigorous extension of the present theory for this case is expected to be possible, probably under some appropriate conditions. Indeed, it is remarkable that the obtained PDE for the ch.f of the response $x(\cdot; \omega)$, by applying our method at a formal level, coincides with the corresponding one (for the same system) obtained by other methods even in a case where the excitation process has not continuous sample functions.

Consider the joint Ch.Fl of the two random functions $x(t; \omega), \tau^{-1}\Delta_\tau z(t; \omega)$,

$$\mathscr{F}_{x(\tau^{-1}\Delta_\tau z)}(u, v) = \int_{\mathscr{X}}\int_{\mathscr{X}} \mathrm{e}^{\mathrm{i}(\langle u, x\rangle + \langle v, \tau^{-1}\Delta_\tau z\rangle)} \times \mathscr{P}_{xy}\left(\mathrm{d}x, \mathrm{d}\left(\tau^{-1}\Delta_\tau z\right)\right), \tag{6.2}$$

where $\mathscr{X}$ is an appropriate function space. The functional $\mathscr{F}_{x(\tau^{-1}\Delta_\tau z)}(u, v)$ is parametrically dependent on the time increment $\tau > 0$. Working similarly as in Section 4, we find that Eqs. (6.1a) and (6.1b) for the processes $x(t; \omega)$ and $\Delta_\tau z(t; \omega)$ are equivalent to the following Hopf-type FDE and initial condition for the Ch.Fl $\mathscr{F}_{x(\tau^{-1}\Delta_\tau z)}(u, v)$:

$$\tau^{-1}\left(\frac{\delta \mathscr{F}_{x(\tau^{-1}\Delta_\tau z)}(u, v)}{\delta u(t + \tau)} - \frac{\delta \mathscr{F}_{x(\tau^{-1}\Delta_\tau z)}(u, v)}{\delta u(t)}\right) + \mu \frac{\delta \mathscr{F}_{x(\tau^{-1}\Delta_\tau z)}(u, v)}{\delta u(t)} - k\frac{\delta^3 \mathscr{F}_{x(\tau^{-1}\Delta_\tau z)}(u, v)}{\delta u(t)^3} = \frac{\delta \mathscr{F}_{x(\tau^{-1}\Delta_\tau z)}(u, v)}{\delta v(t)}, \tag{6.3a}$$

$$\mathscr{F}_{x(\tau^{-1}\Delta_\tau z)}(\upsilon\delta(\cdot - t_0), 0) = \phi_0(\upsilon), \quad \upsilon \in \mathbb{R}. \tag{6.3b}$$

Taking the limit of both sides of (6.3a) as $v \to 0$, and noting that

$$\lim_{v\to 0}\frac{\delta \mathscr{F}_{x(\tau^{-1}\Delta_\tau z)}(u, v)}{\delta u(t)} = \frac{\delta \mathscr{F}_{x(\tau^{-1}\Delta_\tau z)}(u, 0)}{\delta u(t)} = \frac{\delta \mathscr{F}_x(u)}{\delta u(t)},$$

we obtain

$$\tau^{-1}\left(\frac{\delta \mathscr{F}_x(u)}{\delta u(t + \tau)} - \frac{\delta \mathscr{F}_x(u)}{\delta u(t)}\right) + \mu\frac{\delta \mathscr{F}_x(u)}{\delta u(t)} - k\frac{\delta^3 \mathscr{F}_x(u)}{\delta u(t)^3} = \lim_{v\to 0}\frac{\delta \mathscr{F}_{x(\tau^{-1}\Delta_\tau z)}(u, v)}{\delta v(t)}. \tag{6.4}$$

In accordance with Eq. (4.7), the functional derivative appearing in the right-hand side of (6.4) is expressed as:

$$\begin{aligned}\frac{\delta \mathscr{F}_{x(\tau^{-1}\Delta_\tau z)}(u,v)}{\delta v(t)} &\equiv I_\tau(u,v)\\ &= \mathrm{i}\int_{\mathscr{Z}}\int_{\mathscr{X}}\left(\tau^{-1}\Delta_\tau z(t)\right)\mathrm{e}^{\mathrm{i}\left(\langle u,x\rangle+\langle v,\tau^{-1}\Delta_\tau z\rangle\right)}\\ &\quad\times\mathscr{P}_{x(\tau^{-1}\Delta_\tau z(t))}\left(\mathrm{d}x,\mathrm{d}\left(\tau^{-1}\Delta_\tau z\right)\right).\end{aligned}\tag{6.5}$$

In the present case, because of the specific form of the excitation (independent-increment process), the response $x(t;\omega)$ is stochastically independent from the future increment of the excitation $\tau^{-1}\Delta_\tau z(t;\omega)=\tau^{-1}(z(t+\tau;\omega)-z(t;\omega))$ (at this point we make essential use of the assumption $\tau>0$). As a consequence, the joint probability measure $\mathscr{P}_{x(\tau^{-1}\Delta_\tau z)}\left(\mathrm{d}x,\mathrm{d}\left(\tau^{-1}\Delta_\tau z\right)\right)$ can be written in multiplicative form as follows

$$\mathscr{P}_{x(\tau^{-1}\Delta_\tau z)}\left(\mathrm{d}x,\mathrm{d}\left(\tau^{-1}\Delta_\tau z\right)\right)=\mathscr{P}_x(\mathrm{d}x)\cdot\mathscr{P}_{\tau^{-1}\Delta_\tau z}\times\left(\mathrm{d}\left(\tau^{-1}\Delta_\tau z\right)\right),$$

and the double functional integral appearing in the right-hand side of Eq. (6.5), can be factorised and simplified as follows:

$$\begin{aligned}&I_\tau(u,v)\\ &= \mathrm{i}\int_{\mathscr{Z}}\int_{\mathscr{X}}\left(\tau^{-1}\Delta_\tau z(t)\right)\mathrm{e}^{\mathrm{i}\left(\langle u,x\rangle+\langle v,\tau^{-1}\Delta_\tau z\rangle\right)}\mathscr{P}_x(\mathrm{d}x)\\ &\quad\times\mathscr{P}_{\tau^{-1}\Delta_\tau z(t)}\left(\mathrm{d}\left(\tau^{-1}\Delta_\tau z\right)\right)\\ &= \mathrm{i}\int_{\mathscr{Z}}\left(\tau^{-1}\Delta_\tau z(t)\right)\mathrm{e}^{\mathrm{i}\langle v,\tau^{-1}\Delta_\tau z\rangle}\mathscr{P}_{\tau^{-1}\Delta_\tau z}\left(\mathrm{d}\left(\tau^{-1}\Delta_\tau z\right)\right)\\ &\quad\cdot\int_{\mathscr{X}}\mathrm{e}^{\mathrm{i}\langle u,x\rangle}\mathscr{P}_x(\mathrm{d}x)\\ &= \mathrm{i}\int_{\mathscr{Z}}\left(\tau^{-1}\Delta_\tau z(t)\right)\mathrm{e}^{\mathrm{i}\langle v,\tau^{-1}\Delta_\tau z\rangle}\mathscr{P}_{\Delta_\tau z}(\mathrm{d}(\Delta_\tau z))\cdot\mathscr{F}_x(u).\end{aligned}$$

The identity $\mathscr{P}_{\tau^{-1}\Delta_\tau z}\left(\mathrm{d}\left(\tau^{-1}\Delta_\tau z\right)\right)=\mathscr{P}_{\Delta_\tau z}(\mathrm{d}(\Delta_\tau z))$ has been used in deriving the last member of the above equation. Substituting the above result in Eq. (6.5) and getting back to (6.4) we find

$$\begin{aligned}&\tau^{-1}\left(\frac{\delta\mathscr{F}_x(u)}{\delta u(t+\tau)}-\frac{\delta\mathscr{F}_x(u)}{\delta u(t)}\right)+\mu\frac{\delta\mathscr{F}_x(u)}{\delta u(t)}-k\frac{\delta^3\mathscr{F}_x(u)}{\delta u(t)^3}\\ &= \mathrm{i}\mathscr{F}_x(u)\cdot\lim_{v\to0}\int_{\mathscr{Z}}\left(\tau^{-1}\Delta_\tau z(t)\right)\\ &\quad\times\mathrm{e}^{\mathrm{i}\langle v,\tau^{-1}\Delta_\tau z\rangle}\mathscr{P}_{\Delta_\tau z}(\mathrm{d}(\Delta_\tau z)).\end{aligned}\tag{6.6}$$

We are interested in the limiting case $\tau\to0^+$. Thus, taking the limit in both sides of (6.6) we obtain

$$\begin{aligned}&\lim_{\tau\to0^+}\tau^{-1}\left(\frac{\delta\mathscr{F}_x(u)}{\delta u(t+\tau)}-\frac{\delta\mathscr{F}_x(u)}{\delta u(t)}\right)+\mu\frac{\delta\mathscr{F}_x(u)}{\delta u(t)}\\ &\quad-k\frac{\delta^3\mathscr{F}_x(u)}{\delta u(t)^3}\\ &= \mathrm{i}\mathscr{F}_x(u)\cdot\lim_{\tau\to0^+}\lim_{v\to0}\int_{\mathscr{Z}}\left(\tau^{-1}\Delta_\tau z(t)\right)\mathrm{e}^{\mathrm{i}\langle v,\tau^{-1}\Delta_\tau z\rangle}\\ &\quad\times\mathscr{P}_{\Delta_\tau z}(\mathrm{d}(\Delta_\tau z)),\end{aligned}\tag{6.7}$$

which holds true for any $u\in\mathscr{U}$, and $v\to0$ in $\mathscr{V}$. In order to find an equation for the one-dimensional ch.f, Eq. (6.7) is applied for $u=v\cdot\delta(\cdot-t)$, $v=v\tau\cdot\delta(\cdot-t)$. This specific choice for the amplitude of v ensures that $v\to0$ whenever $\tau\to0^+$, thus transforming the double limit in the right-hand side of Eq. (6.7) into a single one:

$$\begin{aligned}&\frac{1}{v}\frac{\partial\phi_{x(t)}(v)}{\partial t}+\mu\frac{\partial\phi_{x(t)}(v)}{\partial v}-k\frac{\partial^3\phi_{x(t)}(v)}{\partial v^3}\\ &= \mathrm{i}\phi_{x(t)}(v)\cdot\lim_{\tau\to0^+}\int_{\mathscr{Z}}\left(\tau^{-1}\Delta_\tau z(t)\right)\\ &\quad\times\exp(\mathrm{i}v\Delta_\tau z(t))\,\mathscr{P}_{\Delta_\tau z}(\mathrm{d}(\Delta_\tau z)).\end{aligned}\tag{6.8}$$

In order to calculate the limit of the right-hand side of Eq. (6.8) we consider the marginal ch.f

$$\phi_{\Delta_\tau z(t)}(v)=\int_{\mathscr{Z}}\exp\{\mathrm{i}v(\Delta_\tau z(t))\}\,\mathscr{P}_{\Delta_\tau z}(\mathrm{d}(\Delta_\tau z)),\tag{6.9}$$

and take the limiting value of its τ-derivative as $\tau\to0^+$:

$$\begin{aligned}&\lim_{\tau\to0^+}\frac{\partial\phi_{\Delta_\tau z(t)}(v)}{\partial\tau}\\ &= \lim_{\tau\to0^+}\frac{\partial}{\partial\tau}\int_{\mathscr{Z}}\exp\{\mathrm{i}v(\Delta_\tau z(t))\}\,\mathscr{P}_{\Delta_\tau z}(\mathrm{d}(\Delta_\tau z))\\ &= \lim_{\tau\to0^+}\int_{\mathscr{Z}}\mathrm{i}v\frac{\partial\Delta_\tau z(t)}{\partial\tau}\mathrm{e}^{\mathrm{i}v\Delta_\tau z(t)}\mathscr{P}_{\Delta_\tau z}(\mathrm{d}(\Delta_\tau z))\\ &= \mathrm{i}v\lim_{\tau\to0^+}\int_{\mathscr{Z}}\frac{\Delta_\tau z(t)}{\tau}\mathrm{e}^{\mathrm{i}v\Delta_\tau z(t)}\mathscr{P}_{\Delta_\tau z}(\mathrm{d}(\Delta_\tau z)).\end{aligned}\tag{6.10}$$

The last term in the above equation coincides – apart from the factor $\mathrm{i}v$ – with the term appearing in the right-hand side of Eq. (6.8). Thus, combining Eqs. (6.8) and (6.10) we obtain

$$\begin{aligned}&\frac{\partial\phi_{x(t)}(v)}{\partial t}+\mu v\frac{\partial\phi_{x(t)}(v)}{\partial v}-kv\frac{\partial^3\phi_{x(t)}(v)}{\partial v^3}\\ &= \phi_{x(t)}(v)\lim_{\tau\to0^+}\frac{\partial\phi_{\Delta_\tau z(t)}(v)}{\partial\tau},\quad v\in\mathbb{R}.\end{aligned}\tag{6.11}$$

The corresponding equation for the pdf is easily derived from Eq. (6.11) by applying the inverse Fourier transformation

$$\begin{aligned}&\frac{\partial f_{x(t)}(a)}{\partial t}+\frac{\partial}{\partial a}\left[\left(\mu a+ka^3\right)f_{x(t)}(a)\right]\\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}v(y-a)}f_{x(t)}(y)\\ &\quad\times\lim_{\tau\to0^+}\frac{\partial\phi_{\Delta_\tau z(t)}(v)}{\partial\tau}\mathrm{d}y\mathrm{d}v.\end{aligned}\tag{6.12}$$

The term $\lim_{\tau\to0^+}\frac{\partial\phi_{\Delta_\tau z(t)}(v)}{\partial\tau}$ expresses the specific probabilistic characteristics of the independent-increment excitation

process which affects the response probabilities. We will now proceed to examine some special cases (specific excitation processes) in order to compare our equations with analogous results obtained by other authors with different methods.

Case 1: Gaussian, independent-increment excitation process

Let us assume that $z(\cdot;\cdot) : I \times \Omega \to \mathbb{R}$ is the Wiener process, i.e. for arbitrary t and $\tau > 0$ the increment $z(t+\tau;\omega) - z(t;\omega)$ has a Gaussian distribution with

$$E^{\omega}[z(t+\tau;\omega) - z(t;\omega)] = 0 \quad \text{and}$$
$$E^{\omega}\left[(z(t+\tau;\omega) - z(t;\omega))^2\right] = b^2\tau. \tag{6.13}$$

Then the characteristic function of the increment $z(t+\tau;\omega) - z(t;\omega)$,

$$\phi_{\Delta_\tau z(t)}(\upsilon) = \exp\left\{-\frac{1}{2}b^2\upsilon^2\tau\right\}, \tag{6.14}$$

hence,

$$\lim_{\tau\to 0^+}\frac{\partial\phi_{\Delta_\tau z(t)}(\upsilon)}{\partial\tau} = -\frac{1}{2}b^2\upsilon^2. \tag{6.15}$$

Substituting Eq. (6.15) to Eq. (6.11) will result to the following FPK equation

$$\frac{\partial\phi_{x(t)}(\upsilon)}{\partial t} + \mu\upsilon\frac{\partial\phi_{x(t)}(\upsilon)}{\partial\upsilon} - k\upsilon\frac{\partial^3\phi_{x(t)}(\upsilon)}{\partial\upsilon^3} = -\frac{1}{2}b^2\upsilon^2\phi_{x(t)}(\upsilon), \quad \upsilon\in\mathbb{R}. \tag{6.16}$$

This equation coincides with the one given by Grigoriu [25, Example2] for the same case.

Case 2: Compound Poisson excitation process (piece-wise continues sample functions)

Assume now that $z(\cdot;\cdot) : I \times \Omega \to \mathbb{R}$ is a compound Poisson process, i.e.

$$z(t;\omega) = \sum_{n=1}^{N(t;\omega)}\hat{z}_n(\omega),$$

where $N(t;\omega)$ is a Poisson process, and $\hat{z}_n(\omega), n = 1,2,\ldots$ are real-valued, independent identically distributed random variables, independent of $N(t;\omega)$. The Poisson process $N(t;\omega)$ follows the distribution

$$F(k) = \frac{[\lambda(t)]^k}{k!}e^{-\lambda(t)}, \quad k = 0,1,\ldots, \tag{6.17}$$

where $\lambda(t)$ is an increasing positive function characterizing the process. The characteristic function of the increment $z(t+\tau;\omega) - z(t;\omega)$ is given by

$$\phi_{\Delta_\tau z(t)}(\upsilon) = \exp\left\{(\lambda(t+\tau) - \lambda(t))\left(\phi_{\hat{z}}(\upsilon) - 1\right)\right\}. \tag{6.18}$$

[49, Ch. I]. Thus, we have

$$\lim_{\tau\to 0^+}\frac{\partial\phi_{\Delta_\tau z(t)}(\upsilon)}{\partial\tau} = \lambda'(t)\left(\phi_{\hat{z}}(\upsilon) - 1\right) = \lambda'(t)\,\mathbf{E}^{\omega}\left[e^{i\upsilon\hat{z}} - 1\right]. \tag{6.19}$$

Substituting Eq. (6.19) into Eq. (6.11) will result in the equation

$$\frac{\partial\phi_{x(t)}(\upsilon)}{\partial t} + \mu\upsilon\frac{\partial\phi_{x(t)}(\upsilon)}{\partial\upsilon} - k\upsilon\frac{\partial^3\phi_{x(t)}(\upsilon)}{\partial\upsilon^3} = \phi_{x(t)}(\upsilon)\cdot\lambda'(t)\left(\phi_{\hat{z}}(\upsilon) - 1\right), \quad \upsilon\in\mathbb{R}. \tag{6.20}$$

Assuming now that $N(t;\omega)$ is a homogeneous Poisson process with $\lambda(t) = \rho t$, and the system is linear ($k = 0$), we obtain the specific FPK equation

$$\frac{\partial\phi_{x(t)}(\upsilon)}{\partial t} + \mu\upsilon\frac{\partial\phi_{x(t)}(\upsilon)}{\partial\upsilon} = \rho\mathbf{E}^{\omega}\left[e^{i\upsilon\hat{z}} - 1\right]\phi_{x(t)}(\upsilon), \quad \upsilon\in\mathbb{R}, \tag{6.21}$$

which coincides with the one given by Grigoriu [25, Example 4], obtained by using different, well-established techniques. That is, a formal application of our method leads to the correct equation even in this case of an excitation with non-continuous sample functions.

7. Moment equations from the new PDE (5.6)

It is worth noticing that the PDE (5.6a), derived at Section 5, can reproduce the infinite set of moment equations corresponding to the dynamical system equation (4.1a). This is a very important consistency result that can be interpreted twofold. From the point of view of the new PDE (5.6a), it provides an independent check of validity. From the point of view of the infinite system of moment equations, it provides an "integrating scheme" permitting the replacement of the infinite system of ODEs by a single linear PDE. The remaining of this section is devoted to the proof of the above mentioned consistency result.

Let us denote by $M_{nm}(t,s) = \mathbf{E}^{\omega}\left[x^n(t;\omega)\cdot y^m(s;\omega)\right]$, $n,m = 0,1,\ldots$, the joint (n,m)th- order moment of $x(t;\omega)$ and $y(s;\omega)$. Then, by direct integration of Eq. (4.1a), it is easily seen that infinite system of moment equations has the form

$$\frac{1}{n+1}\cdot\left.\frac{\mathrm{d}M_{n+1,m}(t,s)}{\mathrm{d}t}\right|_{s=t} + \mu M_{n+1,m}(t,t) = -kM_{n+3,m}(t,s) + M_{n,m+1}(t,s). \tag{7.1}$$

We shall now derive the same Eq. (7.1) using the PDE (5.6a). Recall first that

$$\left.\frac{\partial^{n+m}\phi_{x(t)y(s)}(\upsilon,\nu)}{\partial\upsilon^n\partial\nu^m}\right|_{\substack{\upsilon=0\\ \nu=0}} = i^{n+m}\mathbf{E}^{\omega}\left[x^n(t;\omega)\cdot y^m(s;\omega)\right] = i^{n+m}M_{n,m}(t,s), \quad n,m = 0,1,\ldots. \tag{7.2}$$

By direct differentiation of Eq. (5.6a) we obtain

$$\left.\frac{\partial^{n+m+2}\phi_{x(t)y(s)}(\upsilon,\nu)}{\partial\upsilon^{n+1}\partial\nu^m\partial t}\right|_{s=t} + \mu\frac{\partial^{n+m+1}}{\partial\upsilon^{n+1}\partial\nu^m}\left[\upsilon\frac{\partial\phi_{x(t)y(t)}(\upsilon,\nu)}{\partial\upsilon}\right] + k\frac{\partial^{n+m+1}}{\partial\upsilon^{n+1}\partial\nu^m}\left[\upsilon\frac{\partial^3\phi_{x(t)y(t)}(\upsilon,\nu)}{\partial^3\upsilon}\right] = \frac{\partial^{n+m+1}}{\partial\upsilon^{n+1}\partial\nu^m}\left[\upsilon\frac{\partial\phi_{x(t)y(t)}(\upsilon,\nu)}{\partial\nu}\right], \quad \upsilon,\nu\in\mathbb{R}. \tag{7.3}$$

In accordance with (7.2), the first term in the left-hand side of Eq. (7.3) can be written as

$$\frac{\partial}{\partial t}\left.\frac{\partial^{n+m+1}\phi_{x(t)y(s)}(\upsilon,\nu)}{\partial\upsilon^{n+1}\partial\nu^{m}}\right|_{\substack{\upsilon=0\\ \nu=0}} = \mathrm{i}^{n+m+1}\frac{\partial M_{n+1,m}(t,s)}{\partial t}. \tag{7.4a}$$

To proceed with the remaining three terms in Eq. (7.3), use will be made of the following lemma:

Lemma 7.1. *For every C^n-differentiable function $f:\mathbb{R}\to\mathbb{R}$, we have*

$$\left.\frac{d^n}{\mathrm{d}x^n}\left[xf(x)\right]\right|_{x=0} = n\left.\frac{\mathrm{d}^{n-1}f(x)}{\mathrm{d}x^{n-1}}\right|_{x=0}. \quad \blacksquare$$

Hence,

$$\left.\frac{\partial^{n+m+1}}{\partial\upsilon^{n+1}\partial\nu^{m}}\left[\upsilon\frac{\partial\phi_{x(t)y(t)}(\upsilon,\nu)}{\partial\upsilon}\right]\right|_{\substack{\upsilon=0\\ \nu=0}} = (n+1)\left.\frac{\partial^{n+m+1}\phi_{x(t)y(t)}(\upsilon,\nu)}{\partial\upsilon^{n+1}\partial\nu^{m}}\right|_{\substack{\upsilon=0\\ \nu=0}}, \tag{7.4b}$$

$$\left.\frac{\partial^{n+m+1}}{\partial\upsilon^{n+1}\partial\nu^{m}}\left[\upsilon\frac{\partial^{3}\phi_{x(t)y(t)}(\upsilon,\nu)}{\partial^{3}\upsilon}\right]\right|_{\substack{\upsilon=0\\ \nu=0}} = (n+1)\left.\frac{\partial^{n+m+3}\phi_{x(t)y(t)}(\upsilon,\nu)}{\partial\upsilon^{n+3}\partial\nu^{m}}\right|_{\substack{\upsilon=0\\ \nu=0}}, \tag{7.4c}$$

and

$$\left.\frac{\partial^{n+m+1}}{\partial\upsilon^{n+1}\partial\nu^{m}}\left[\upsilon\frac{\partial\phi_{x(t)y(t)}(\upsilon,\nu)}{\partial\nu}\right]\right|_{\substack{\upsilon=0\\ \nu=0}} = (n+1)\left.\frac{\partial^{n+m+1}\phi_{x(t)y(t)}(\upsilon,\nu)}{\partial\upsilon^{n}\partial\nu^{m+1}}\right|_{\substack{\upsilon=0\\ \nu=0}}. \tag{7.4d}$$

Substituting Eqs. (7.4b)–(7.4d) in Eq. (7.3) we obtain

$$\left(\frac{1}{(n+1)}\frac{\partial}{\partial t}\frac{\partial^{n+m+1}\phi_{x(t)y(s)}(\upsilon,\nu)}{\partial\upsilon^{n+1}\partial\nu^{m}} + \mu\frac{\partial^{n+m+1}\phi_{x(t)y(t)}(\upsilon,\nu)}{\partial\upsilon^{n+1}\partial\nu^{m}} + k\frac{\partial^{n+m+3}\phi_{x(t)y(t)}(\upsilon,\nu)}{\partial\upsilon^{n+3}\partial\nu^{m}}\right)_{\substack{\upsilon=0\\ \nu=0}} = \left.\frac{\partial^{n+m+1}\phi_{x(t)y(t)}(\upsilon,\nu)}{\partial\upsilon^{n}\partial\nu^{m+1}}\right|_{\substack{\upsilon=0\\ \nu=0}}. \tag{7.5}$$

By applying Eqs. (7.2) and (7.4a), the above equation reduces to the infinite system of moment equation (7.1). This completes the proof of the consistency result announced at the beginning of this section.

8. Kernel density representation of joint pdfs

Clearly, problem (5.6) – either in the form (5.6a)–(5.6e) or in the form (5.7a)–(5.7e) – exhibits some peculiarities making it distinctly different from the usual initial-boundary value problems for PDEs, coming from problems of Mathematical Physics. These peculiarities reflect the probabilistic origin of the present problem.

In the remaining part of this paper, an original (particle-type) method for the numerical solution of problem (5.6) (or (5.7)) is developed, and some first, illustrative, numerical results are presented. The main tool, on which the formulation of the numerical scheme relies, is the representation of the sought-for pdf and ch.f by means of convex superpositions of kernel density function s (kdfs) and their Fourier transformation, the kernel characteristic function s (kch.fs), respectively. A short presentation of the basic facts about kdfs is given below.

Kernel density functions constitute a key notion/tool within the framework of nonparametric statistical estimation. See, e.g., [47]. In our approach, a kdf $K(x;x_*,h)$ is mainly thought of as a generalized (non-symmetric) summability kernel, appropriate to represent pdfs [22]. The defining properties of an M-variate kdf are the following:

(Pr.1) $K(x;x_*,h)$ is a continuous, real-valued function defined on a domain of the form $\mathsf{D}_K = A\times A\times \mathsf{M}^{\mathrm{NonNegDef}}_{M\times M}$, where $A\subseteq\mathbb{R}^M$ is taken to be contained in (or to be equal to) the support of the target pdf, say $f(x)$, which is to be represented (see Lemma 8.1 and Theorem 8.1), and $\mathsf{M}^{\mathrm{NonNegDef}}_{M\times M}$ is the set of non-negative definite, $M\times M$-matrices, which can serve as covariance matrices.

(Pr.2) $K(x;x_*,h)\geq 0$, for $(x;x_*,h)\in\mathsf{D}_K$.

(Pr.3) $\int_A K(x;x_*,h)\,\mathrm{d}x_* = 1$, for $(x,h)\in A\times\mathsf{M}^{\mathrm{NonNegDef}}_{M\times M}$.

(Pr.4) $\lim_{\|h\|\to 0}\int_{\|x-x_*\|>\delta} K(x;x_*,h)\,\mathrm{d}x_* = 0$, for any $x_*\in A$ and $\delta>0$.

A kernel characteristic function is defined as the Fourier transformation of a kdf. Clearly, properties (Pr.2), (Pr.3) ensure that each kdf is a pdf on its own. The ***shape*** of the kernel function $K(x;x_*,h)$ is controlled by its covariance matrix h, also called *bandwidth* (or *shape*) *parameter*. h quantifies the spreading of the kernel probability mass around its "***center***" x_*. Another – simpler and in many cases adequate – choice of the shape parameter is the M-variate vector of the eigenvalues of the covariance matrix. In this sense, the domain $\mathsf{D}_K = A\times A\times \mathsf{M}^{\mathrm{NonNegDef}}_{M\times M}$ can be (and will be) simplified as $A\times A\times[0,\infty)^M$.

Using the defining properties (Pr.1)–(Pr.4), and only these, it is not difficult to prove the following

Lemma 8.1. *If $f(x)$ is a continuous pdf and $K(\cdot;\cdot,\cdot)$ is any kernel function satisfying* (Pr.1)–(Pr.4)*, then, for any x,*

$$\lim_{\|h\|\to 0}\int_A K(x;x_*,h)f(x_*)\,\mathrm{d}x_* = f(x). \quad \blacksquare \tag{8.1}$$

That is, as the bandwidth decreases, the kernel function shrinks around its "center" x_*, having the weak asymptotic limit

$$K(x;x_*,h)\xrightarrow{\|h\|\to 0}\delta(x-x_*). \tag{8.2}$$

On the other hand, as the bandwidth increases the kernel function spreads out.

Theorem 8.1. *The set of all convex finite superpositions of the form* $\sum_{n=1}^{N} p_n K(x; x_n, h_n)$*, where* $p_1 + p_2 + \cdots + p_N = 1$*,* $p_n \geq 0$ *for all n, and* $K(\cdot;\cdot,\cdot)$ *is any kernel function satisfying* (Pr.1)–(Pr.4)*, is dense within the set of all continuous pdfs supported in A. That is, given any continuous pdf* $f(x)$*, a specific kernel function* $K(x; x_*, h)$*, and an arbitrary (small) number* $\varepsilon > 0$*, there exist a bandwidth parameter* h_**, a finite set of centers* $\{x_n\}_{n=1}^{N}$ *in A, and a vector* $\mathbf{p} = (p_1, p_2, \ldots, p_N)$ *lying in the positive cone of* $\mathbb{R}^N$*, such that*

$$\max_{x\in A} \left| f(x) - f^N(x) \right| < \varepsilon, \tag{8.3a}$$

where

$$f^N(x) = \sum_{n=1}^{N} p_n K(x; x_n, h_n). \quad \blacksquare \tag{8.3b}$$

The clue of the proof of this theorem is Lemma 8.1, in conjunction with the properties of the Riemann sum approximation of the integral $\int_A K(x; x_*, h) f(x_*)\, \mathrm{d}x_*$ [2]. The technical details are omitted. The above theorem makes clear that any (continuous) pdf can be approximated, as closely as it is required, by a representation of the form (8.3b).

9. Reformulation of the problem by using kernel density representations

We shall now apply the pdf representation (8.3b) (or the corresponding ch.f representation, obtained by means of a Fourier transformation) in order to reformulate problem (5.7) (or (5.6a)) in a way facilitating its numerical solution. Again here and subsequently, as in the introduction, $f_{\mathbf{xy}} = f_{x(t)y(s)}(a,\beta)$, $\phi_{xy} = \phi_{x(t)y(s)}(\upsilon,\nu)$ are four-argument, two-variate, joint, response–excitation pdf and ch.f, respectively. For clarity, in the present and the subsequent sections, vector or matrix quantities will be explicitly denoted by using bold letters.

Applying the representation (8.3b) for the pdf, and the corresponding one for the ch.f we define the approximants

$$f^N_{x(t)y(s)}(a,\beta) = \sum_{k=1}^{N} p_k(t,s)\, K\left(a,\beta; \mathbf{m}^k(t,s), \mathbf{h}^k(t,s)\right), \tag{9.1}$$

$$\phi^N_{x(t)y(s)}(\upsilon,\nu) = \mathscr{F}^{a\to\upsilon}_{\beta\to\nu}\left\{ f^N_{x(t)y(s)}(a,\beta)\right\} = \sum_{k=1}^{N} p_k(t,s)\, \tilde{K}\left(\upsilon,\nu; \mathbf{m}^k(t,s), \mathbf{h}^k(t,s)\right). \tag{9.2}$$

Here $\mathbf{m}^k = (m^k_x, m^k_y)$ is the location parameter, namely the position of the most probable (highest) value of the kdf, and $\mathbf{h}^k$ is the shape parameter, represented either by the 2×2-covariance matrix of the kdf or by the two eigenvalues of the latter (both pictures will be applied to the numerical treatment). For the numerical computations, $K\left(a,\beta;\mathbf{m}^k,\mathbf{h}^k\right)$ is taken to be a Gaussian pdf. (See, e.g., [26, Sec. 2.9].)

Our main goal now is to exploit the representations (9.1) and (9.2) in order to solve the system (5.6) or the equivalent (5.7). Conditions (5.6d) and (5.6e), or the equivalent (5.7d) and (5.7e), are automatically satisfied since the approximants (9.1) and (9.2) are by construction pdfs and ch.fs, respectively.

To facilitate the discussion, we define the linear differential operators

$$\mathcal{L}\bullet = \left.\frac{\partial \bullet}{\partial t}\right|_{s=t} + \mu\frac{\partial[a\bullet]}{\partial a} + k\frac{\partial\left[a^3\bullet\right]}{\partial a} - \frac{\partial[\beta\bullet]}{\partial a}. \tag{9.3a}$$

$$\tilde{\mathcal{L}}\bullet = \left.\frac{\partial \bullet}{\partial t}\right|_{s=t} + \mu\upsilon\frac{\partial\bullet}{\partial\upsilon} - k\upsilon\frac{\partial^3\bullet}{\partial\upsilon^3} - \upsilon\frac{\partial\bullet}{\partial\nu}. \tag{9.4a}$$

And rewrite Eqs. (5.6a) and (5.7a) in the following concise form:

$$\mathcal{L}\left[f_{xy}\right](a,\beta,t) = 0, \quad (a,\beta)\in\mathbb{R}^2, t\geq t_0, \tag{9.3b}$$

$$\tilde{\mathcal{L}}\left[\phi_{xy}\right](\upsilon,\nu,t) = 0, \quad (\upsilon,\nu)\in\mathbb{R}^2, t\geq t_0. \tag{9.4b}$$

It is interesting to note here that the two equivalent formulations –(5.7a) or (9.3) in terms of the pdf, and (5.6a) or (9.4a),(9.4b) in terms of the ch.f – are both useful and they will be considered in parallel, since the conceptual arguments are better stated using the pdf formulation, while the numerical analysis is better developed using the ch.f formulation.

Substituting the approximation (9.1) into (9.3), we obtain

$$\sum_{k=1}^{N}\mathcal{L}\left[p_k K\left(a,\beta;\mathbf{m}^k,\mathbf{h}^k\right)\right] = 0, \quad (a,\beta)\in\mathbb{R}^2. \tag{9.5}$$

Let us denote by $\varepsilon\left(\mathbf{h}^j\right)$ the radius of the effective support of $K\left(a,\beta;\mathbf{m}^j,\mathbf{h}^j\right)$. $\varepsilon\left(\mathbf{h}^j\right)$ will be taken and always kept to be small. Since each kernel function $K\left(a,\beta;\mathbf{m}^j,\mathbf{h}^j\right)$ is taken to be concentrated around its center $\mathbf{m}^j = \left(m^j_x, m^j_y\right)$ and it is positive there, Eq. (9.5), restricted in a neighborhood $\mathscr{N}\left(\mathbf{m}^j, \varepsilon\left(\mathbf{h}^j\right)\right)$, is locally equivalent with the equation

$$\sum_{k=1}^{N}\mathcal{L}\left[p_k K\left(a,\beta;\mathbf{m}^k,\mathbf{h}^k\right)\right]\cdot K\left(a,\beta;\mathbf{m}^j,\mathbf{h}^j\right) = 0, \quad (a,\beta)\in\mathscr{N}\left(\mathbf{m}^j,\varepsilon\left(\mathbf{h}^j\right)\right). \tag{9.6}$$

Assuming that the system of neighborhoods $\left\{\mathscr{N}\left(\mathbf{m}^j,\varepsilon\left(\mathbf{h}^j\right)\right), j=1,\ldots,N\right\}$ covers the essential support of the sought-for density function f_{xy}, we can assert that the global equation (9.5) is equivalent to the system of local equations

$$\sum_{k=1}^{N}\mathcal{L}\left[p_k K\left(a,\beta;\mathbf{m}^k,\mathbf{h}^k\right)\right]\cdot K\left(a,\beta;\mathbf{m}^j,\mathbf{h}^j\right) = 0, \quad \forall j\in\{1,\ldots,N\}, \text{ and } \forall (a,\beta)\in\bigcup_{j=1,\ldots,N}\mathscr{N}\left(\mathbf{m}^j,\varepsilon\left(\mathbf{h}^j\right)\right). \tag{9.7}$$

By taking a Fourier transformation, Eq. (9.7) is equivalently rewritten as

$$\sum_{k=1}^{N} \tilde{\mathcal{L}}\left[p_k \tilde{K}\left(\upsilon, \nu; m^k, h^k\right)\right] * \tilde{K}\left(\upsilon, \nu; m^j, h^j\right) = 0,$$
$$\forall j \in \{1, \ldots, N\}, \text{ and } \forall (\upsilon, \nu) \in \mathbb{R}^2, \tag{9.8}$$

where $*$ denotes the convolution operator. Although the latter equation could be considered as being more complicated than Eq. (9.7), an efficient numerical solution scheme will be based on it.

10. A two-level numerical solution scheme for the set of Eq. (9.8)

To proceed to the numerical solution, use will be made of a specific choice of the kdf. Assuming a Gaussian density as the kdf, we have

$$K\left(a, \beta; \mathbf{m}^k, \mathbf{C}^k\right) = \frac{1}{2\pi\sqrt{|\det[\mathbf{C}]|}} \times \exp\left[-\frac{1}{2}\begin{pmatrix} a - m_x^k \\ \beta - m_y^k \end{pmatrix}^{\mathrm{T}} \left[\mathbf{C}^k\right]^{-1} \begin{pmatrix} a - m_x^k \\ \beta - m_y^k \end{pmatrix}\right], \tag{10.1a}$$

with corresponding kernel characteristic function

$$\tilde{K}\left(\upsilon, \nu; \mathbf{m}^k, \mathbf{C}^k\right) = \exp\left(\mathrm{i}\begin{bmatrix} m_x^k \\ m_y^k \end{bmatrix}^{\mathrm{T}} \cdot \begin{bmatrix} \upsilon \\ \nu \end{bmatrix} - \frac{1}{2}\begin{bmatrix} \upsilon \\ \nu \end{bmatrix}^{\mathrm{T}} \cdot \left[\mathbf{C}^k\right] \cdot \begin{bmatrix} \upsilon \\ \nu \end{bmatrix}\right), \tag{10.1b}$$

where

$$\mathbf{m}^k = \mathbf{m}^k(t, s) = \left(m_x^k(t), m_y^k(s)\right) \tag{10.2a}$$

is the mean vector, and

$$\mathbf{C}^k = \mathbf{C}^k(t, s) = \begin{pmatrix} C_{xx}^k(t, t) & C_{xy}^k(t, s) \\ C_{yx}^k(s, t) & C_{yy}^k(s, s) \end{pmatrix} \tag{10.2b}$$

is the covariance matrix of our Gaussian kdf. As we have already mentioned above, a dual realization of the shape parameter will be considered herewith. Apart from the covariance matrix $\mathbf{C}^k$, the vector $\mathbf{h}^k = \left(h_x^k, h_y^k\right)$ having as elements the two eigenvalues of the matrix $\mathbf{C}^k$, will also be used in this case.

Our numerical solution scheme will be implemented by restricting the kdf to be highly concentrated, so that the effective supports of any pair of two different kernels to be practically non-overlapping. This permits us to neglect the interaction between any pair of Gaussian kernels, i.e. to disregard the summation in the left-hand side of Eq. (9.7) and its equivalent Eq. (9.8). Thus, under the above assumption, which is equivalent with the condition $\left\|\mathbf{h}^k\right\| < \varepsilon_1$, for all $k \in \{1, \ldots, N\}$, where ε_1 is an appropriate (small) constant, Eq. (9.8) simplifies to

$$\tilde{L}\left[p_j \tilde{K}\left(\upsilon, \nu; \mathbf{m}^j, \mathbf{C}^j\right)\right] = 0, \quad t \geq s \geq t_0,$$
$$\text{and } \forall j \in \{1, \ldots, N\}, \text{ and } \forall (\upsilon, \nu) \in \mathbb{R}^2. \tag{10.3a}$$

Furthermore, assuming the amplitudes p_j are positive and piecewise constant, the above equation is further simplified to

$$\tilde{\mathcal{L}}\left[\tilde{K}\left(\upsilon, \nu; \mathbf{m}^j, \mathbf{C}^j\right)\right] = 0, \quad \text{within each time interval}$$
$$\tau^{(\ell)} \leq s \leq t \leq \tau^{(\ell+1)}, \forall j \in \{1, \ldots, N\},$$
$$\text{and } \forall (\upsilon, \nu) \in \mathbb{R}^2. \tag{10.3b}$$

On the basis of the above discussion, a two-level (two-time scale) approach comes into the scene:

a. Solve the set of independent equation (10.3b) within each interval $\tau^{(\ell)} \leq s \leq t \leq \tau^{(\ell+1)}$ (this is the *short-time phase* or *inner-cycle phase*), and then
b. Come back to the complete representation and update the values of the amplitudes p_j, passing from the interval $\left[\tau^{(\ell)}, \tau^{(\ell+1)}\right]$ to the interval $\left[\tau^{(\ell+1)}, \tau^{(\ell+2)}\right]$ (this is the *coarse-time phase* or *the outer-cycle phase*).

The *criterion* for defining the sequence of coarse updating times $\tau^{(\ell)}$, $\ell = 1, 2, 3, \ldots$, is formulated as a sufficient condition for the validity of the assumptions underlying the derivation of the set of independent equation (10.3b). It turns out that the most critical assumption is the restriction of each kdf to be highly concentrated around its center. As expected, because of the diffusive character of the problem, it has been found that, during the short-time phase solution, kernel parameters evolve in a way leading to a continuous increase of the variance parameter $\left\|\mathbf{h}^k\right\|$. (See, for example, Figs. 2(c) and 3(c), in Section 11, and the discussion therein.) The growth of the quantity $\left\|\mathbf{h}^k\right\|$ leads to the spreading of the mass of the corresponding kdf, which results in the violation of the assumption of negligible interaction between the kernels.

Thus, the set of kernel parameters $\mathbf{m}^k(t, s)$ and $\mathbf{C}^k(t, s)$ evolve in accordance with the simplified dynamical equation (10.3b) from time $\tau^{(\ell)}$, until the spreading index $\left\|\mathbf{h}^k(t)\right\|$, of some kernel, exceeds a certain critical value, say $\varepsilon_1 > 0$. This value of t is taken to be the next updating time $\tau^{(\ell+1)}$. At that time instant, the inner-cycle (short-time) solution phase is interrupted, and an approximation of the total joint pdf $f^N_{x(\tau^{(\ell+1)})y(\tau^{(\ell+1)})}(a, \beta)$ is calculated by means of Eq. (9.1), in the specific form:

$$f^N_{x(\tau^{(\ell+1)})y(\tau^{(\ell+1)})}(a, \beta) = \sum_{k=1}^{N} p_k\left(\tau^{(\ell)}\right) K\left(a, \beta; \mathbf{m}^k\left(\tau^{(\ell+1)}, \tau^{(\ell+1)}\right), \mathbf{h}^k\left(\tau^{(\ell+1)}, \tau^{(\ell+1)}\right)\right). \tag{10.4}$$

Then, the calculated pdf (10.4) is re-approximated, by using a new set of kdfs, satisfying the concentration condition $\left\|\mathbf{h}^k\right\| = \varepsilon_2 < \varepsilon_1$, with different amplitudes $p_k\left(\tau^{(\ell+1)}\right)$. The latter are calculated by means of an optimization algorithm (used also for the set up of the initial conditions), which is described in

the Appendix. After the updating of the amplitudes, the next inner-cycle begins, and the procedure continues as described above.

During each time interval $\left[\tau^{(\ell)}, \tau^{(\ell+1)}\right]$ the amplitudes are considered constant and, thus, globally, p_j are piecewise constant functions of time. In fact, the evolution of the amplitudes p_j is much slower than the evolution of the kernel parameters $\mathbf{m}^k$ and $\mathbf{h}^k$, and this is what justifies the piece-wise constant assumption for p_j in our numerical scheme. C.f. [50]. An improved numerical solution, taking also into account the evolution of p_j in a continuous fashion, can be constructed and will be published elsewhere.

It should be stressed that the accuracy of the method proposed and developed herewith is critically dependent on the threshold value ε_1 for the variance parameter (spreading index) $\left\|\mathbf{h}^k(t)\right\|$.

10.1. A local-moment method for the numerical solution of Eq. (10.3b)

We are now focusing on the numerical treatment of Eq. (10.3b). For each value of $j \in \{1, \ldots, N\}$, Eq. (10.3b) contains three unknown functions, namely the response mean value $m_x^j(t)$, and covariances $C_{xx}^j(t,t)$ and $C_{xy}^j(t,s)$, which should be determined, and two known functions, namely the excitation mean value $m_y^j(s)$ and the autocovariance $C_{yy}^j(s,s)$, introducing the appropriate, inner-cycle, excitation. Thus, any solution scheme of Eq. (10.3b) should provide us with a number of equations (hopefully three) governing the evolution of the three unknown functions, along with the evidence that introducing the obtained solution in the operator $\tilde{\mathcal{L}}[\tilde{K}(\upsilon, \nu;\ \mathbf{m}^j, \mathbf{C}^j)]$ will result in 0 (at least approximately) for all values of $(\upsilon, \nu) \in \mathbb{R}^2$.

Since the (Gaussian) kernel $\tilde{K}\left(\upsilon, \nu;\ \mathbf{m}^j, \mathbf{C}^j\right)$ is $C^\infty\left(\mathbb{R}^2\right)$ in (υ, ν) and dies out as $\|(\upsilon, \nu)\| \to \infty$, Eq. (10.3b) is equivalent to the following system of localized moment equations:

$$\frac{\partial^{p+q}}{\partial^p \upsilon \partial^q \nu} \tilde{\mathcal{L}}\left[\tilde{K}\left(\upsilon, \nu; \mathbf{m}^j, \mathbf{C}^j\right)\right]\Big|_{\substack{\upsilon=0 \\ \nu=0}} = 0,$$

$$\forall\,(p,q) \in \mathbb{N}_0 \times \mathbb{N}_0, \quad \mathbb{N}_0 = \{0, 1, 2, 3, \ldots\}. \tag{10.5}$$

Exploiting the specific (Gaussian) form of the kernel, and considering the cases $(p,q) = (1,0)$, $(2,0)$ and $(1,1)$, the following three (nonlinear) ODEs are obtained from (10.5):

$$m_{x,t}^j(t) + \mu m_x^j(t) + 3k m_x^j(t)\, C_{xx}^j(t,t) + k\left[m_x^j(t)\right]^3 = m_y^j(t), \tag{10.6a}$$

$$C_{xx,t}^j(t,s) + \mu C_{xx}^j(t,s) + 3k m_x^j(t)\, C_{xx}^j(t,s)\, m_x^j(t) + 3k C_{xx}^j(t,t)\, C_{xx}^j(t,s) = C_{xy}^j(t,s), \tag{10.6b}$$

$$C_{xy,t}^j(t,s) + \mu C_{xy}^j(t,s) + 3k C_{xx}^j(t,t)\, C_{xy}^j(t,s) + 3k m_x^j(t)\, C_{xy}^j(t,s)\, m_x^j(t) = C_{yy}^j(t,s). \tag{10.6c}$$

These equations involve the three unknown functions $m_x^j(t)$, $C_{xx}^j(t,t)$ and $C_{xy}^j(t,s)$, and they are differential equations with respect to t, parametrically dependent on s. (No derivatives with respect to s appear.) They should be satisfied for all values of (t,s) such that $\tau^{(\ell)} \le s \le t \le \tau^{(\ell+1)}$. We are especially interesting in the solution of system (10.6) on the diagonal $s = t$.

By direct calculations it has been found that, if the three moment equations (10.6a)–(10.6c) holds true, then various other – but not all – (p,q)-moment equations are also satisfied. In any case, the system (10.6a)–(10.6c) is closed and can be efficiently solved, providing us with a reasonable approximation of the evolution of the kernel parameters $m_x^j(t)$, $C_{xx}^j(t, s=t)$ and $C_{xy}^j(t, s=t)$. When the value of $\left\|\mathbf{h}^j(t)\right\|$ exceeds the threshold value ε_1, the current inner-cycle phase is finished and the procedure switches to outer-cycle phase.

The numerical solution of the set of nonlinear ODEs (10.6) is implemented by using the method of the *quasilinearization* [6,34]. Taking advantage of the symmetry properties of the correlation matrix, the equations can be solved on the 'diagonal', that is around $s = t$. The sequence of time instants for the numerical scheme has the form

$$(t,s) : (t_i, t_i) \to (t_{i+1}, t_i) \to (t_{i+1}, t_{i+1}).$$

An important aspect of the present method is its suitability for parallel computation. Parallelization techniques can be applied both to the dynamical evolution of the kernels and to the optimization algorithm. In the first case the algorithm can take advantage of the independent evolution of each kernel. For the parallelization of the optimization algorithm we can split the group of kdfs into subgroups and then independently approximate each subgroup by new kernels with small variance. Hence, we can probably succeed fast computations for systems of higher dimensions, subjected to general (smooth) excitation.

Table 1
System parameters

System parameters	Case I	Case II
μ	1	1
k	1	−1

11. Numerical examples

We shall now apply the above described numerical scheme to the numerical determination of the response pdf of a dynamical system (4.1), excited by a known stochastic process (see below), with system parameters μ and k having the values given in Table 1, under Cases I and II.

By performing a stability analysis to problem (4.1) we found that for $\mu/k > 0$ (Case I in Table 1), the nonlinear system has one stable fixed point located at zero. A pitchfork bifurcation occurs at $\mu/k = 0$, and the fixed point at zero becomes unstable in the semi-axis $\mu/k < 0$ (Case II). In the same region $(\mu/k < 0)$ two symmetric stable points appear at $\pm\sqrt{|\mu/k|}$. Hence, we have the bifurcation diagram shown in Fig. 1.

On the basis of the above described dynamical features of the studied problem, it is natural to expect that, in Case I, the evolved pdf will become eventually a unimodal distribution

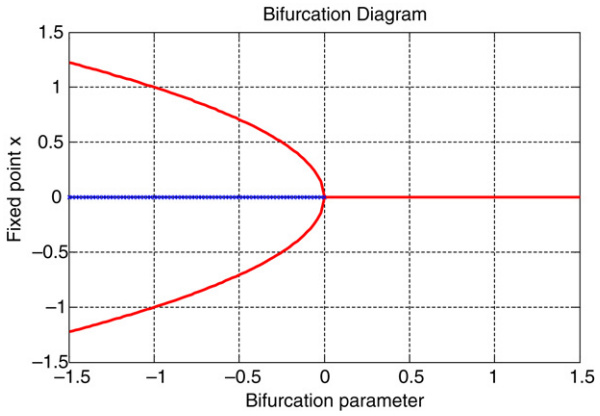

Fig. 1. Bifurcation diagram for system (4.1) with respect to the bif. parameter μ/k.

centered at zero, while in Case II, the probability will concentrate around the pair of the two symmetric stable fixed points $\pm\sqrt{\mu/|k|}$, hence ultimately a bimodal distribution will appear. Since the stable fixed points are global attractors we expect to attain these results after some time, independently of the initial density. The numerical results to be presented and discussed below clearly comply with this behavior, dictated by the qualitative analysis of the studied system.

Consider first Case I, with a bimodal initial pdf, defined as a convex superposition of two Gaussians with parameter values $m_1 = 0, m_2 = 0.6$, $\sigma_1 = 0.1$, $\sigma_2 = 0.6$, and amplitudes $p_1 = 0.4$ and $p_2 = 0.6$, respectively. This initial pdf is shown in Fig. 2(b), at the section $t = 0$. The excitation process is taken to be, in this case, a Gaussian stationary random function with zero mean and covariance function given by

$$C_{YY}(\tau) = \frac{1}{2}\cos^2(2\tau)\,. \tag{11.1}$$

Numerical results are presented in Fig. 2. More specifically, in the two upper plots of this figure (Fig. 2(a) and (b)), the evolution of the probability density $f_{x(t)}(a)$ is shown, for the time interval $0 < t < 1.4$ s, large enough to get the steady state response pdf. Also, in the same figure (Fig. 2(a)) the orbits of $m_x^k(t)$ are plotted by using thick black lines. The apparent discontinuities every 0.2 s are due to the reapproximation of the calculated density by means of a new convex superposition of kdfs with smaller variance every time the concentration parameters h^k exceeds the critical value ε_1 (which in this example was taken to be $\varepsilon_1 = 0.3$). In Fig. 2(c) the evolution of the variance for some kdfs of the response density is shown. The diffusive character of the evolution (strictly increasing variances with respect to time) is clearly seen in the numerical results. Again, the apparent discontinuities are due to the

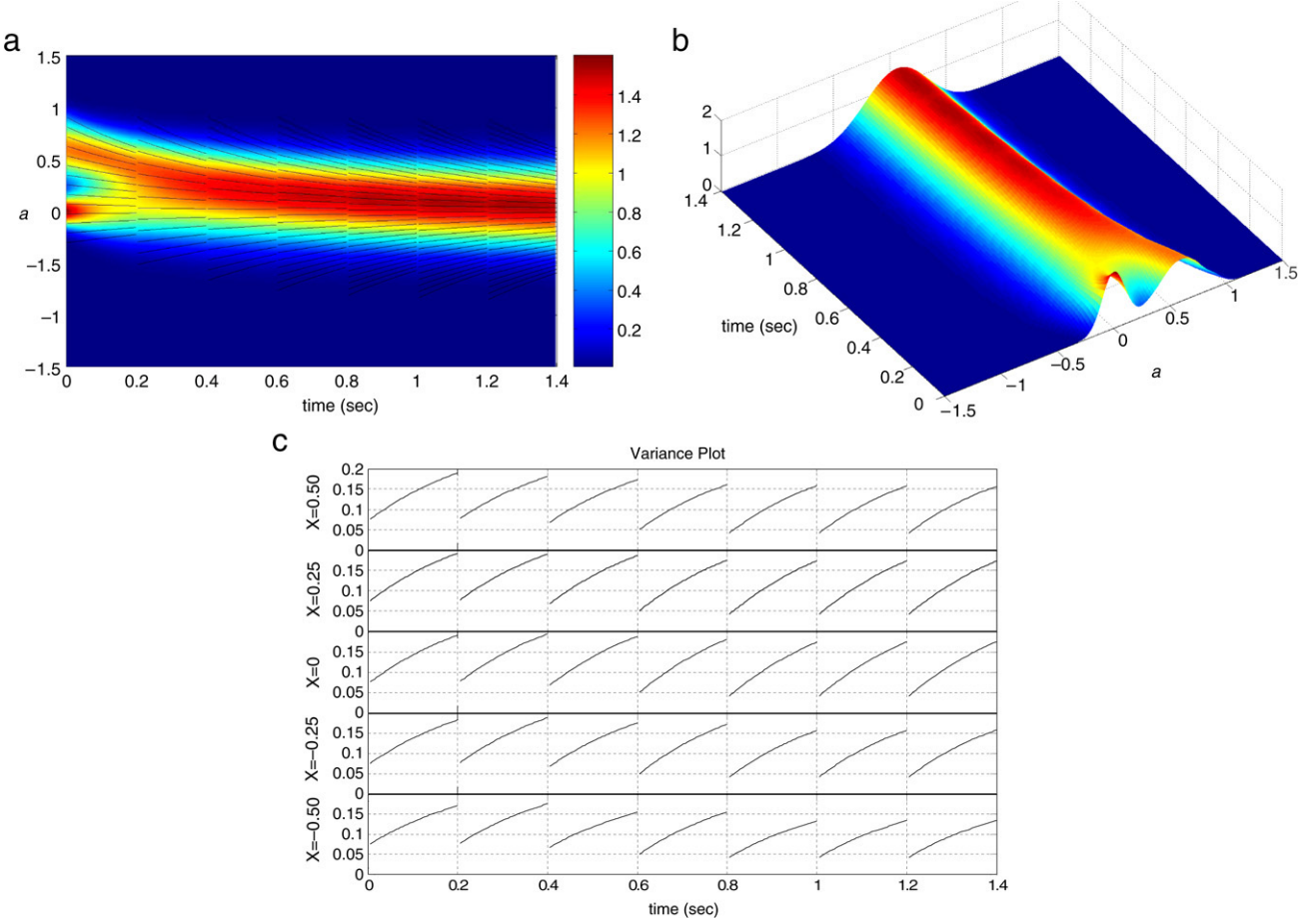

Fig. 2. (a) Response pdf $f_{x(t)}(a)$ and $m_x^j(t)$ curves for Case I with stationary excitation. (b) 3D plot of the response pdf. (c) Variance plots for some kdfs.

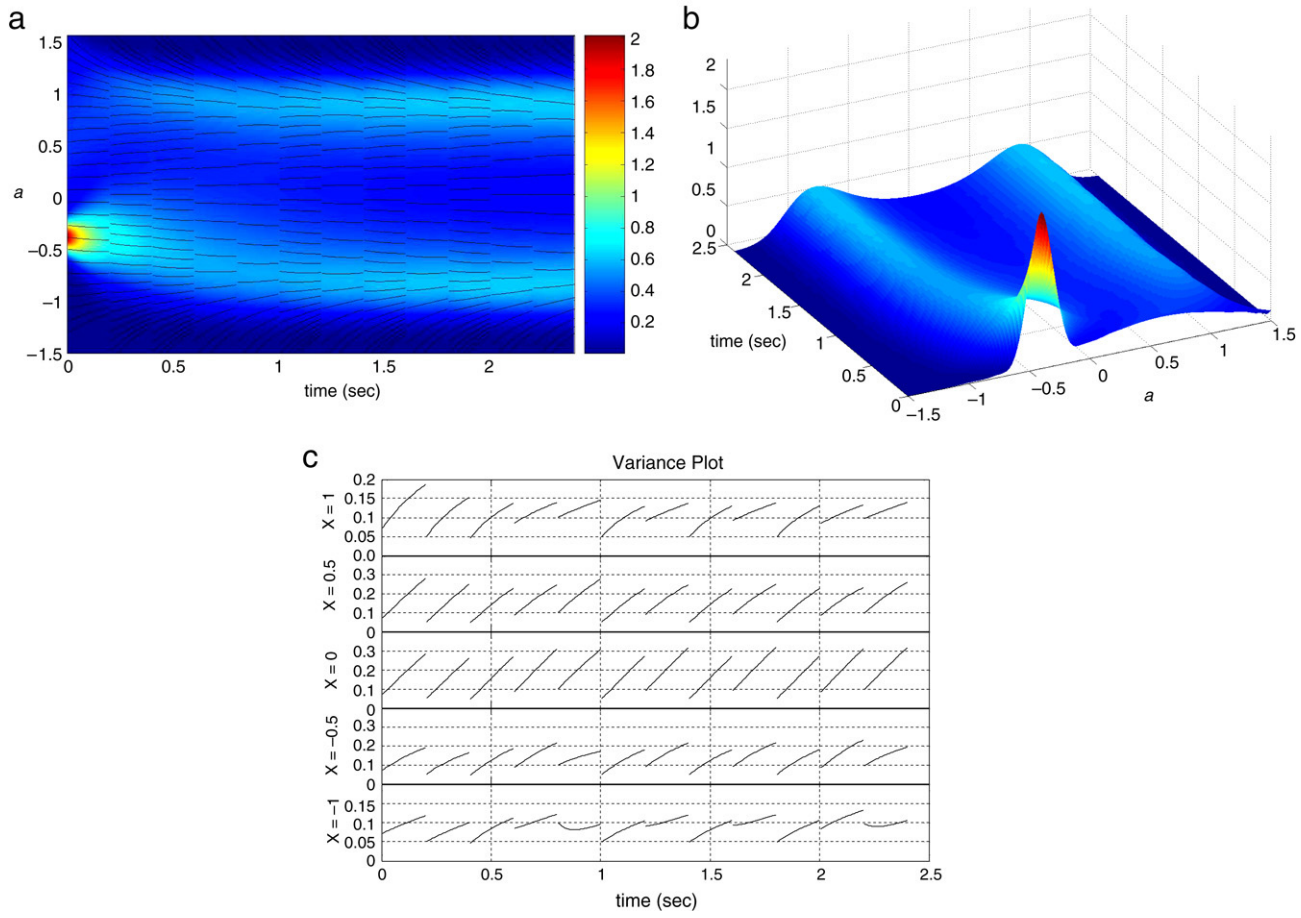

Fig. 3. (a) Response pdf $f_{x(t)}(a)$ and $m_x^j(t)$ curves for Case II with stationary excitation. (b) 3D plot of the response pdf. (c) Variance plots for some kdfs.

re-approximation of the response pdf by kdfs with smaller variances.

Let us now consider our system (4.1) with parameter values as in Case II. Two cases of stochastic excitation will be considered. First we will study the same stationary Gaussian excitation as before, having zero mean and covariance function given by Eq. (11.1). The initial distribution is taken to be bimodal (strongly asymmetric for this case), and is defined as a convex superposition of two Gaussian pdfs with parameters $m_1 = -0.4$, $m_2 = 0.6$, $\sigma_1 = 0.1$, $\sigma_2 = 0.6$, and amplitudes $p_1 = 0.4$, $p_2 = 0.6$, respectively.

Numerical results concerning the evolution of the response pdf, for the time interval $0 < t < 2.4$ s, are presented in Fig. 3. Although the initial pdf has taken to be a strongly asymmetric bimodal one, the eventually resulting response density turns to be a symmetric bimodal pdf, with modes exactly at the stable fixed points, located at $\pm\sqrt{|\mu/k|} = \pm 1$, as expected. The interchange of probability between the kernels (implemented by means of the re-approximation of the response pdf in terms of a new convex superposition of kdfs with smaller variances) takes place approximately every 0.2 s. This is shown in the figure as an apparent discontinuity of the mean-value and variance curves.

From both Figs. 2(a) and 3(a) (see also Fig. 4(a)), we can easily observe a permanent tendency of $m_x^k(t)$-orbits to be attracted by the stable fixed points. This means that there is a continuous inflow of probability mass from the outer region of the phase space ($|a| > 1$) to a strip around the locus of the stable fixed points, which is not stopping even after the response pdf has been reached its stationary form. This apparently paradoxical behaviour should be addressed to the discrepancy between the tail form of the response pdf $f_{x(t)}(a)$, and the tail form of the Gaussian kernels which are used to represent $f_{x(t)}(a)$. This fact reveals the necessity for an asymptotic study of the tail behaviour directly from the differential equation (5.7a), which will permit the construction and use of the kdfs suitably adapted to the specific system, i.e., exhibiting the correct tail behaviour. Such a construction will also facilitate and accelerate the convergence of the numerical solution procedure.

Finally, in Fig. 4 we present numerical results for Case II, with a non-stationary (cyclostationary) Gaussian excitation, with zero mean and covariance function given by

$$C_{YY}(t,s) = \frac{1}{2}\left(1 + 0.2\cos\left(\frac{\pi t}{2}\right)\right)\cos^2(t-s)\,. \tag{11.2}$$

Again the initial distribution is constructed as a superposition of two Gaussian pdfs with parameters $m_1 = -0.4$, $m_2 = 0.6$, $\sigma_1 = 0.3$, $\sigma_2 = 0.7$, and amplitudes $p_1 = 0.4$ and $p_2 = 0.6$, respectively. The evolution of the response probability density function is plotted for the time interval $0 < t <$

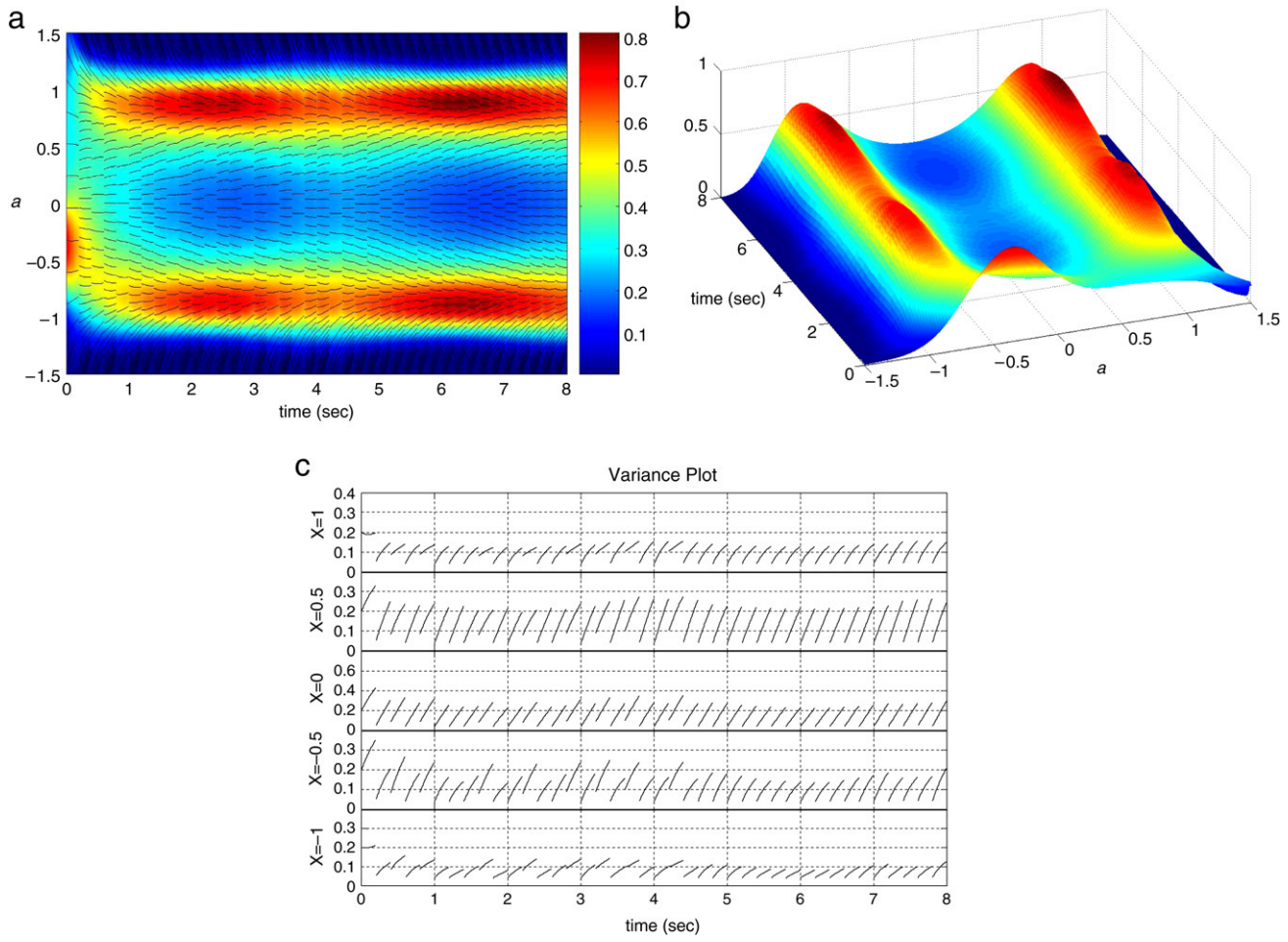

Fig. 4. (a) Response pdf $f_{x(t)}(a)$ and $m_x^j(t)$ curves for Case II with non-stationary excitation. (b) 3D plot of the response pdf. (c) Variance plots for some kdfs.

8.0 s, long enough so that the periodic character of the response becomes clear.

From Fig. 4(a) and (b) we are able to observe that, after a transient state ($0 < t < 1.5$ s), the response density function exhibits a periodic behavior with a period of approximately 4 s, which is the period of the excitation, i.e., the period of the correlation function $C_{YY}(t, s)$ (Eq. (11.2)) with respect to its first argument. Furthermore, it is easily seen that, in this case, a greater amount of kernels is necessary in order to approximate satisfactorily the sought-for pdf, due to the fact that the non-stationary excitation produces a more complicated response.

12. Discussion and conclusions

In this paper new PDEs governing the evolution of the joint, response–excitation, ch.f and pdf of nonlinear dynamical systems under general (correlated) stochastic excitation have been derived. The starting point of our approach is a Hopf-type equation, that governs the joint, response–excitation, characteristic functional, providing a probabilistically complete reformulation, equivalent with the underlying (nonlinear) stochastic differential equation. This 'infinite-dimensional' equation is appropriately reduced (by projection) to linear partial differential equations that govern the response–excitation, characteristic (or probability density) function (see, e.g. Eq. (5.6a) or (5.7a)). The latter equations are supplemented with (non-local) marginal compatibility conditions (see, e.g. Eq. (5.6b) or (5.7b)) and initial conditions (see, e.g. Eq. (5.6c) or (5.7c)), and they can provide us with the evolution of the joint ch.f. (or pdf).

For the numerical solution of these novel PDEs (e.g., either in the form (5.6) or (5.7)) an original, particle-type, method is developed and illustrated through its application to a specific, simple, nonlinear system. The key point of the numerical method is the representation of the joint, pdfs and ch.fs by means of appropriate convex superpositions of kernel density or kernel characteristic functions, respectively. In this way, the non-local marginal compatibility conditions are satisfied *a priori*, and the PDEs governing the evolution of the sought-for pdf and ch.f are eventually transformed to systems of nonlinear ODEs for the kernel parameters.

From the results presented in this work we conclude that the proposed method is able to produce quite satisfactory results for systems subjected to general stochastic excitation. Important aspects of the method are (i) It is a *two-level*, *particle-type* method, separating the fast, inner-cycle (short-term) phase, which describes the particle dynamics separately for each particle, from the slow, outer-cycle (long-term) phase, which accounts for the interchange of probability mass between the

particles and the evolution of the particles' amplitudes. (ii) It can be improved, keeping its two-level, particle-type character, so that to avoid the piece-wise smoothness assumption for the amplitudes p_j, and to ensure the "exact" satisfaction of the PDE, by solving a linear evolution problem in the outer-cycle phase. (iii) It can be generalized to higher dimensional systems. And (iv) It is plainly suitable for parallelized computations, since the nonlinear ODEs describing the evolution of each particle in the inner-cycle phase can be solved independently. In addition, the computationally demanding optimization algorithm (see Appendix) is easily parallelizable.

Acknowledgements

The authors would like to thank the anonymous reviewer for his/her instructive comments and suggestions. Also, they would like to express their gratitude to Dr. Agissilaos Athanassoulis and Ms. Ivi-Sevasti Tsantili, Ph.D. Candidate, who read the manuscript and offered valuable comments and suggestions enhancing the revised version of the paper.

Appendix

In this Appendix a brief outline is given of the optimization algorithm used for the construction of appropriate approximants of the sought-for pdf in terms of kdfs exhibiting a small variance (either given or under specific control). This algorithm is used quite often throughout the numerical solution, i.e. each time the solution procedure switches from the inner-cycle to the outer-cycle and the calculated density is re-approximated by means of kdfs of small variance. It is also used for implementing the initialization, by representing the given initial pdf as a convex superposition of appropriate kdfs. The basic optimization problem is formulated as follows:

Given $f(x)$ and σ_0, find M and $\{p_k, m_k\}_{k=1}^{M}$ such that

$$\int_{-\infty}^{+\infty} \left[f(x) - \sum_{k=1}^{M} \frac{p_k}{\sqrt{2\pi}\sigma_0} \exp\left\{ -\frac{1}{2} \left(\frac{x - m_k}{\sigma_0} \right)^2 \right\} \right]^2 \mathrm{d}x = \min \tag{A.1}$$

under the constraints:

$$p_1 + p_2 + \cdots + p_M = 1 \quad \text{and} \quad p_k \geq 0, \quad \text{for all } k.$$

For the inner-cycle/outer-cycle re-approximation of the sought-for pdf, the integrations can be carried out analytically (since $f(x)$ is already represented as a superposition of Gaussian kernels, with different parameters of course) leading to an explicit linear optimization problem, if M is given. M is obtained by using a variant of an iterative, adaptive procedure, developed by Gavriliadis [22].

For the initial data representation, the optimization procedure is performed quite similarly. However, in this case, the integrations in (A.1) are performed numerically, since, in general, the initial probability distribution may not be analytically described. A detailed description of the solution algorithm will be presented elsewhere.

References

[1] Arnold L, Chueshov I, Ochs G. Stability and capsizing of ships in random sea — a survey. Nonlinear Dynamics 2004;36:135–79.
[2] Athanassoulis GA, Gavriliadis PN. The truncated Hausdorff moment problem solved by using kernel density functions. Probabilistic Engineering Mechanics 2002;17(3):273–91.
[3] Athanassoulis GA, Sapsis ThP. New partial differential equations governing the joint, response–excitation, probability distributions of nonlinear systems, under general stochastic excitation I: Derivation. In: Deodatis G, Spanos PD, editors. Proceedings of this conference, Presented in the 5th conference of computational stochastic mechanics. Rotterdam: Millpress; 2007. p. 27–36.
[4] Athanassoulis GA, Sapsis ThP. Probabilistic responses of dynamical systems under general stochastic excitation. I: New PDEs governing the evolution of joint, response–excitation, probability densities and characteristic functions for Duffing oscillator; 2008 [to be submitted].
[5] Belenky LV, Sevastianov NB. Stability and safety of ships, Vol. II. Risk of capsizing. Amsterdam: Elsevier; 2003.
[6] Bellman R. Methods of nonlinear analysis. New York: Academic Press; 1970.
[7] Beran MJ. Statistical continuum theories. New York: Interscience Publishers; 1968.
[8] Budini A, Caceres M. The generalized wiener process II: Finite systems. Journal of Physics. A. Mathematical and General 1999;32:4005–26.
[9] Budini A, Caceres M. Functional characterization of generalized Langevin equations. Journal of Physics. A. Mathematical and General 2004;37:5959–81.
[10] Bunke H. Gewöhnliche differentialgleichungen mit zufälligen parametern. Berlin: Akademie-Verlag; 1972.
[11] Caceres M, Budini A. The generalized Ornstein–Uhlenbeck process. Journal of Physics. A. Mathematical and General 1997;30:8427–44.
[12] Cai GQ, Lin YK. On exact stationary solutions of equivalent non-linear stochastic systems. International Journal of Non-Linear Mechanics 1988; 23(4):315–25.
[13] Caughey TK. Equivalent linearization techniques. Journal of the Acoustical Society of America 1963;35:1706–11.
[14] Caughey TK. On response of nonlinear oscillators to stochastic excitation. Probabilistic Engineering Mechanics 1986;1:2–4.
[15] Dalecky YuL, Fomin SV. Measures and differential equations in infinite-dimensional space. Dordrecht: Kluwer Academic Publishers; 1991.
[16] Deodatis G. Non-stationary stochastic vector processes: Seismic ground motion applications. Probabilistic Engineering Mechanics 1996;11: 149–68.
[17] Di Paola M, Falsone G. Itô and Stratonovich integrals for delta-correlated processes. Probabilistic Engineering Mechanics 1993;8:197–208.
[18] Di Paola M, Sofi A. Approximate solution of the Fokker–Planck–Kolmogorov equation. Probabilistic Engineering Mechanics 2002;17:369–84.
[19] Egorov AD, Sobolevsky PI, Yanovich LA. Functional integrals : Approximate evaluation and applications. Dordrecht: Kluwer Academic Publishers; 1993.
[20] Feller MN. Infinite dimensional elliptic equations and operators of Levy type. Russian Mathematical Surveys 1986;41:119–70.
[21] Foias C. A functional approach to turbulence. Russian Mathematical Surveys 1974;29:293–326.
[22] Gavriliadis PN. Theoretical and numerical exploitation of the moment problem with applications to the probabilistic prediction of stochastic responses of dynamical systems. Ph.D. thesis. Greece: National Technical University of Athens; 2005.
[23] Gihman II, Skorohod AV. The theory of stochastic processes I. Berlin: Springer; 1974.
[24] Grigoriu M. The Itô and Stratonovich integrals for stochastic differential equations with Poisson white noise. Probabilistic Engineering Mechanics 1998;13(3):175–82.
[25] Grigoriu M. Characteristic function equations for the state of dynamic systems with Gaussian, Poisson and Levy white noise. Probabilistic Engineering Mechanics 2004;19:449–61.

[26] Härdle W. Smoothing techniques. Berlin: Springer-Verlag; 1990.
[27] Hemon P, Santi F. Simulation of a spatially correlated turbulent velocity field using biorthogonal decomposition. Journal of Wind Engineering and Industrial Aerodynamics 2007;95:21–9.
[28] Hönig ChS. Volterra-stieltjes integral equations. Amsterdam: North-Holland; 1975.
[29] Hopf E. Statistical hydromechanics and functional calculus. Journal of Rational Mechanics and Analysis 1952;1:87–123.
[30] Iwankiewicz R. Equations for probability density of response of dynamic systems to a class of non-Poisson random impulse process parametric excitations. In: Deodatis G, Spanos PD, editors. Proceedings of this conference, Presented in the 5th conference of computational stochastic mechanics. Rotterdam: Millpress; 2007. p. 349–56.
[31] Kafali C, Grigoriu M. Non-Gaussian model for spatially coherent seismic ground motions. In: 9 intern. conf. on appl. of statis. and prob. in civil engineering; 2003.
[32] Kollman W. The pdf approach to turbulent flow. Theoretical and Computational Fluid Dyanimics 1990;1:249–85.
[33] Kree P, Soize Ch. Mathematics of random phenomena. Dordrecht: D. Reidel Publ. Co.; 1986.
[34] Lakshmikantham V, Malek S. Generalized quasilinearization. Nonlinear World 1994;1:59–63.
[35] Lewis RM, Kraichnan RH. A space–time functional formalism for turbulence. Communications on Pure and Applied Mathematics 1962;15: 397–411.
[36] Lin YK, Cai GQ. Probabilistic structural dynamics. McGraw-Hill, Inc.; 1995.
[37] Lutes LD. Approximate technique for treating random vibration of hysteretic systems. Journal of the Acoustical Society of America 1970; 48(1):299–306.
[38] Monin AS, Yaglom AM. Statistical fluid mechanics: Mechanics of turbulence, vol. I, II. MIT Press; 1971, 1975.
[39] Pope SB. Turbulent flows. London: Cambridge University Press; 2000.
[40] Pradlwarter HJ. Non-linear stochastic response distributions by local statistical linearization. International Journal of Non-Linear Mechanics 2001;36:1135–51.
[41] Proppe C. Exact stationary probability density functions for non-linear systems under Poisson white noise excitation. International Journal of Non-Linear Mechanics 2003;38(4):557–64.
[42] Proppe C, Pradlwarter HJ, Schuëller GI. Equivalent linearization and Monte Carlo simulation in stochastic dynamics. Probabilistic Engineering Mechanics 2003;18(1):1–15.
[43] Pugachev VS, Sinitsyn . Stochastic differential systems. Chichester (New York): John Wiley & Sons; 1987.
[44] Roberts JB, Spanos PD. Random vibration and statistical linearization. Chichester (New York): John Wiley & Sons; 1990. Re-published by Dover Publications 1990/2003; 2003.
[45] Sapsis ThP, Athanassoulis GA. New partial differential equations governing the joint, response–excitation, probability distributions of nonlinear systems under general stochastic excitation. II: Numerical solution. In: Deodatis G, Spanos PD, editors. Proceedings of this conference, Presented in the 5th conference of computational stochastic mechanics. Rotterdam: Millpress; 2007. p. 587–96.
[46] Schlesinger MF, Swean T. Stochastically excited nonlinear ocean structures. Singapore: World Scientific; 1998.
[47] Scott DW. Multivariate density estimation. Chichester (New York): John Wiley and Sons; 1992.
[48] Simiu E, Scanlan RH. Wind effects on structures. 2nd ed. New York: John Wiley & Sons; 1986.
[49] Sobczyk K. Stochastic differential equations. Dordrecht: Kluwer Academic Publishers; 1991.
[50] Sofi A, Di Paola M, Spanos P. Determining response envelope probability densities for a class of nonlinear oscillators under modulated excitation. In: Deodatis G, Spanos PD, editors. Proceedings of this conference, Presented in the 5th conference of computational stochastic mechanics. Rotterdam: Millpress; 2007. p. 603–12.
[51] Soize C. The Fokker–Planck equation and its explicit steady state solutions. Singapore: World Scientific; 1994.
[52] Soong TT, Grigoriu M. Random vibration of mechanical and structural systems. Englewood Cliffs (NJ): Prentice-Hall; 1993.
[53] Tvrdy M. Linear bounded functionals on the space of regular regulated functions. Tatra Mountains Mathematical Publications 1996;8:203–10.
[54] Tvrdy M. Differential and integral equations in the space of regulated functions. Memoirs on Differential Equations and Mathematical Physics 2002;25:1–104.
[55] Vakhania NN, Tarieladze VI, Chobanyan SA. Probability distributions on banach spaces. Dordrecht: D.Reidel Publ. Co.; 1987.
[56] Vasta M. Exact stationary solution for a class of non-linear systems driven by a non-normal delta-correlated process. International Journal of Non-Linear Mechanics 1995;30(4):407–18.
[57] Volterra V. Theory of functionals and of integral and integro-differential equations. Blackie & Sons Lim.; 1930. Re-published by Dover Publ. Inc. (1959/2005). 1930/1959/2005.
[58] Wilson JF, editor. Dynamics of offshore structures. 2nd ed. Chichester (New York): John Wiley & Sons; 2002.
[59] Zhu WQ, Huang ZL, Suzuki Y. Equivalent non-linear system method for stochastically excited and dissipated partially integrable Hamiltonian systems. International Journal of Non-Linear Mechanics 2003;36: 773–786.