Maximum Entropy Probability Density Principle in Probabilistic Investigations of Dynamic Systems

Jiří Náprstek * [ID] **and Cyril Fischer** [ID]

Institute of Theoretical and Applied Mechanics, Prosecká 809/76, 190 00 Prague 9, Czech Republic; fischerc@itam.cas.cz

* Correspondence: naprstek@itam.cas.cz; Tel.: +420-603-559-765

Abstract: In this study, we consider a method for investigating the stochastic response of a nonlinear dynamical system affected by a random seismic process. We present the solution of the probability density of a single/multiple-degree of freedom (SDOF/MDOF) system with several statically stable equilibrium states and with possible jumps of the snap-through type. The system is a Hamiltonian system with weak damping excited by a system of non-stationary Gaussian white noise. The solution based on the Gibbs principle of the maximum entropy of probability could potentially be implemented in various branches of engineering. The search for the extreme of the Gibbs entropy functional is formulated as a constrained optimization problem. The secondary constraints follow from the Fokker–Planck equation (FPE) for the system considered or from the system of ordinary differential equations for the stochastic moments of the response derived from the relevant FPE. In terms of the application type, this strategy is most suitable for SDOF/MDOF systems containing polynomial type nonlinearities. Thus, the solution links up with the customary formulation of the finite elements discretization for strongly nonlinear continuous systems.

1. Introduction

A seismic wave that propagates horizontally in the upper subsoil layers mostly has a stochastic character, where its predominance increases with the distance from the earthquake epicenter. This fact originates from the filtering of a seismic wave through a continuum with randomly distributed non-homogeneity. A large number of records and data analyses have confirmed this approach. In principle, the seismic wave has three-dimensional characteristics composed of a number of partial waves that propagate with different velocities and other attributes, which change with the distance from the epicenter.

The propagation of a seismic wave through a non-homogeneous continuum and its transformation into a nearly stochastically homogeneous process has been investigated widely in theoretical and experimental studies. These have been conducted under laboratory conditions using special equipment, or directly in situ under the conditions of a real earthquake field. Many previous studies have addressed this issue in recent decades, including seismological investigations (e.g., [1,2]), numerical analyses (e.g., [3]), and theoretical analyses with overlap from seismic to poly-crystalline material features (e.g., [4,5]) or waves generated due to scattering on randomly rough inter-layer limits (e.g., [6]).

If the structure is large in a plane view and its horizontal dimensions are comparable or higher than the typical seismic wavelengths, then the distortion due to excitation processes acting on individual

points of the basement should be considered (see the studies cited above). In particular conditions, either the phase shift of the excitation process in different points of the basement needs to be considered, or different random processes with non-zero cross-correlation are assumed to act on the corresponding points (e.g., see [7,8], where many additional attributes are also discussed).

In general, seismic processes are non-stationary. This fact is well documented and considered in models of the dynamic behavior of structures subjected to this special type of excitation (see [9,10] and monographs devoted to stochastic differential equations (SDEs) and random processes with engineering applications, such as [11,12]). Seismic processes are non-stationary in terms of both their amplitude and frequency, but the rough upper estimate of a structural response can be evaluated by considering a hypothetical stationary random process with suitable parameters [7,8]. Indeed, regardless of their non-stationarity, seismic processes generally have special characteristics, which is clear from the usual shape of their spectral density as a function of the time and frequency. Moreover, it is important to remember the great efforts that have been made to understand the basic mechanisms of these processes and the possibility of simplifying them as stationary events, such as the specific tools employed for their investigation [13,14], empirical mode decomposition analysis [15], and suitable methods for the analysis of seismic signals and data mining (e.g., [16,17]).

To demonstrate the characteristics of these processes and to justify this simplifying hypothesis, Figures 1 and 2 show the displacement records plotted for the Sierra Madre earthquake taken at Altadena, Eaton Canyon Park station on 28 June 1991, with an epicentral distance of 49.3 km for the E–W component (for additional details of the analysis of this record, please refer to [18,19]). The assumption stated above is generally the standard in this area when the dynamic characteristics of seismic processes are considered. However, most recommendations are satisfied by only a certain static equivalent (e.g., see Eurocodes 8 and 9, and [20], as well as related studies). Another study [19] also indicated the discrepancies between the analytical results obtained based on rigorous stochastic dynamics and two standards.
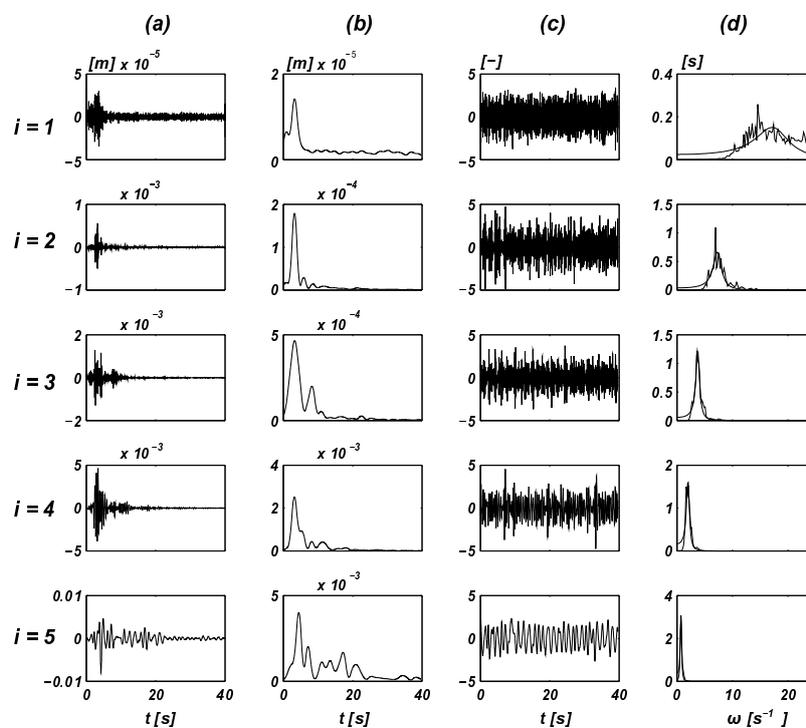

Figure 1. Non-stationary stochastic characteristics of an original record of a seismic shock—decomposition of the Sierra Madre earthquake record (five components i = 1–5). (**a**) Narrow band components $m_i(t) \cdot v_{si}(t)$ [m]. (**b**) Amplitude modulation $m_i(t)$ [m]. (**c**) Stationary process $v_{si}(t)$ [.]. (**d**) Spectral density $\Phi_i(\omega)$ [s] and its approximation (smooth curves).
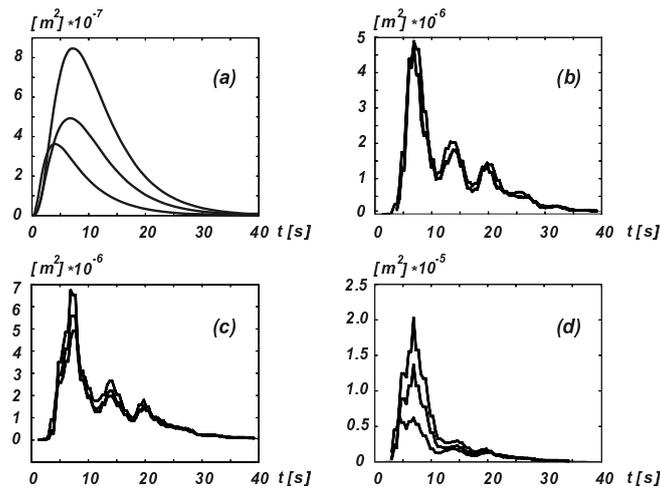
Figure 2. Displacement variance $[m^2]$ in the middle node of a simplified bridge (lumped mass, five degrees of freedom) subjected to excitation in the Sierra Madre earthquake record. (**a**) Simple amplitude modulation—exponential. (**b**) Simple amplitude modulation—B^N spline. (**c**) Split Fourier spectra—B^N spline. (**d**) Wavelet decomposition—B^N spline.

The transmission of vibrations in the contact subsoil structure is usually characterized by significant nonlinearities in terms of both the stiffness and damping. Moreover, various types of nonlinearities emerge from vibration at large amplitudes when non-elastic and hysteretic processes are considered. Geometric nonlinearities cannot be neglected in some configurations (e.g., see [21,22]). The analysis of these nonlinearities is outside the scope of this study, but it is important to be aware that although the stationary part of the random excitation process can be considered Gaussian (e.g., see [1,9,10]). However, the response differs from a Gaussian probability density function (PDF) and it should be investigated in an appropriate manner (e.g., [11]).

By summing up the considerations given above, and assuming that the structure itself can be modeled as a multi-degree of freedom (MDOF) system, the following system of SDEs can be written:

$$\frac{\mathrm{d}x_j(t)}{\mathrm{d}t} = f_j(\mathbf{x},t) + g_{jr}(\mathbf{x},t)w_r(t) , \quad j = 1, \ldots, 2n, \quad n - \text{dynamic degrees of freedom} \tag{1}$$

- $\mathbf{x} = [x_1, x_2, .., x_{2n}]^T$—response components (space variables in the following): (i) x_{2j-1}—displacements; (ii) x_{2j}—velocities ($j = 1, \ldots, n$),
- $w_r(t)$—Gaussian white noise with constant cross-density in terms of the stochastic moments $K_{rs} = \mathcal{E}\{w_r \cdot w_s\}; r, s = 1, m,$ m—number of acting noises
- $\mathcal{E}\{\cdot\}$—mathematical mean value operator in the Gaussian meaning,
- $f_j(\mathbf{x},t), g_{jr}(\mathbf{x},t)$—continuous deterministic functions of state variables $\mathbf{x}$ and time t; $j = 1, 2n$.

The SDE system in Equation (1) can be considered a governing differential system, which allows us to model the various responses of an engineering structure for many types of seismic random processes. Various modifications of the system in Equation (1) can be found in numerous monographs that deal with deterministic or stochastic versions of this basic mathematical model. Rigorous mathematical analyses of its basic properties, limitations of its applicability, and the assumptions that should be satisfied in order for the system to be meaningful have been presented in previous studies (e.g., see [11,12,23,24]). It should be noted that $w_r(t)$ must be Gaussian processes or they need to be prepared using a suitable filtering method, or by considering the specific spectral properties of the processes considered (e.g., see [1,2,23]).

If Equation (1) is linear and the excitation processes are only of additive type, then various modifications of the spectral and correlation type solution methods are applicable because the response of a linear system under additive Gaussian excitation is a vector of Gaussian processes. Therefore,

only the first two stochastic moments (mean value vector and variance square matrix) need to be determined to obtain a full description of the stochastic system response considered. Other dynamic systems produce non-Gaussian responses and they should be investigated using different methods.

The Fokker–Planck Equation (FPE) appears to be the most frequently used tool for solving the cross-PDF of a dynamic system excited by a vector of random processes $w_r(t)$, as shown in Equation (1).

If the input processes w_j, w_r can be considered Gaussian, then the respective FPE for an unknown PDF with the variables $\mathbf{x}, t$ can be associated with Equation (1) as follows:

$$\frac{\partial p(\mathbf{x},t)}{\partial t} = -\frac{\partial}{\partial x_j}\left(\kappa_j(\mathbf{x},t) \cdot p(\mathbf{x},t)\right) + \frac{1}{2}\frac{\partial^2}{\partial x_j \partial x_k}\left(\kappa_{jk}(\mathbf{x},t) \cdot p(\mathbf{x},t)\right) \tag{2}$$

$$\kappa_j(\mathbf{x},t) = \lim_{\Delta t \to 0} \mathcal{E}\{\Delta x_j\}/\Delta t \ ; \ \kappa_{jk}(\mathbf{x},t) = \lim_{\Delta t \to 0} \mathcal{E}\{\Delta x_j \cdot \Delta x_k\}/\Delta t \tag{3}$$

- $\kappa_j(\mathbf{x},t)$—drift coefficients;
- $\kappa_{jk}(\mathbf{x},t)$—diffusion coefficients, $j, k = 1, \dots, n$.

Random inputs affect the system as additive or multiplicative processes. Many previous studies have considered this well-known partial differential equation, which has mostly linear characteristics, although more complex definitions also exist (for a comprehensive explanation, see [11,12,23–25]). The processes $w_r(t)$ can be considered stationary or non-stationary. In principle, they do not need to be perfectly delta-correlated but quite complicated formulae are required for their evaluation (e.g., see [11,23]), and they affect the system as additive or multiplicative processes.

Analytical and semi-analytical methods for obtaining the FPE solution have been described previously. One of the most important was given by [26]. Several studies specifically addressed the analytical or semi-analytical FPE solution procedures (e.g., [27,28]). We note that although there are boundaries between the relevant groups, some of them are rather blurred. Fourier decomposition-based procedures are used particularly widely in theoretical physics, where they are based on a separability $p(\mathbf{x},t) = p(\mathbf{x}) \cdot \varphi(t)$, and the drift and diffusion coefficients should be time independent. A Boltzmann type solution is used frequently as a basic step in the subsequent analysis by employing perturbation techniques. The Galerkin–Petrov method is probably the most general and suitable approach for investigating the majority of problems formulated in terms of the FPE. Many other specific case-oriented methods are based on the idea of completing various potentials, first integrals and their combinations, free parameter fitting, and other operations. Despite the strengths of these methods and the excellent results obtained, they have many disadvantages due to the highly limited dimensionality of FPE, the possibility of configuring the boundary conditions, and problems when analyzing any non-stationary problems.

Numerical methods are a suitable alternative and various generally formulated procedures with respect to large dimensionality have been presented and tested (e.g., see [29]). Nevertheless, the finite element method (FEM) seems to be the most promising approach (e.g., see [30–32] and the references therein). The FEM is affected by several common shortcomings in the same manner as every numerical method but it provides options that are lacking in other methods such as the possibility of analyzing transitions and quasi-periodic processes.

Among the strategies mentioned, one principle has been neglected in recent years, possibly due to the overwhelming dominance of purely numerical methods. This principle is based on the Gibbs entropy of probability and it lies between analytical and numerical methods. Indeed, it can serve as the backbone of a large group of FPE solution methods. We provide some examples of this method in the following.

2. Gibbs Entropy of Probability

Entropy is a physical idea that was introduced into the theory of thermodynamic systems in the 19th century as a quantity for characterizing some macroscopic property, and it is indispensable for

describing the relationships among mechanical, thermic, and other fields. The statistical concept of micro-mechanics absorbed this concept very early, where Gibbs and other theoreticians demonstrated how the entropy with a suitable meaning could be used for statistically describing the uncertainties of internal processes. Subsequent studies by Jaynes [33] and Kullback [34] integrated information theory with statistical mechanics (on micro-level) to allow the possibility of expressing an indeterminacy based on the entropy. In principle, they showed that the actual PDF of a system state supplies the maximum with respect to a certain functional among all other PDF distributions belonging to the admissible states. Based on such a procedure the so-called *b*-value in the Gutenberg–Richter law for earthquakes, which states the (cumulative) number of earthquakes with magnitude greater than *M* occurring in a specified area and time is given by $(N(> M) \approx 10^{-bM})$, was determined in [35] by applying the maximum entropy principle to data analyzed in a new time domain termed natural time [36]. Thus, the response PDF can be found for a particular system provided we have a complete description of the system with the boundary and initial conditions.

This principle led to a revolution in quantum mechanics and thermo-mechanics. The macroscopic properties of a particular system can be described further in terms of statistical parameters based on the properties and state of its microscopic particles. This advance represented a significant step change in solid-state physics and many other areas. Information theory is very closely related to the random characteristics of micro-particles and advanced rapidly as a consequence. This principle has been applied in many disciplines such as data mining and signal processing (as well as stochastic resonance and other areas; e.g., see [37–39]), sea level description [40] (indirect information), and laser technology (e.g., [41]). The maximum entropy principle has proved very useful in studies of earthquake recurrence (e.g., see [42]) as well as physiology and applications in human medicine (e.g., [43]).

The consideration of entropy (as the entropy of probability) in investigations of the dynamics of macroscopic systems in engineering occurred in the late 1980s and early 1990s (e.g., see [44–49]). The delay in the inclusion of entropy compared with solid-state physics and other disciplines may have been due to the somewhat opposite formulation of the problem (from known elements to the unknown PDF of the system response), and thus the use of different mathematical formulation strategies and the subsequent solutions. Projecting entropy into the mathematical formulation means that the unknown PDF is searched for under certain boundary conditions and auxiliary constraints are associated via an unknown Lagrange multiplier. The specific solution procedures appear to be a compromise between (semi-) analytical and numerical methods.

Let us formulate a closed mathematical system that allows us to receive satisfactory information about the PDF of the system response given by Equation (1) together with adequate boundary and initial conditions. The basic principle is associated with the functional introduced by Gibbs [50] in the form intended for the domain of micro-particles, as follows.

$$S = -\kappa_B \int_{(\mathbf{x})} p(\mathbf{x}, t) \cdot \lg \ p(\mathbf{x}, t) \cdot d\mathbf{x}, \tag{4}$$

-   κ_B—Boltzmann constant assigning *S* its physical meaning,
-   $\mathbf{x} = |x_1, x_2, \cdots, x_{2n}|^T$, see Equation (1)

The entropy of the state of the system in the form of Equation (4) was introduced based on a theoretical study of the probability of the state of the gas particle distribution in a closed vessel (see [51]). Except for the multiplication constant, the same expression can be obtained if we examine the rate of indeterminateness in the description of an object or phenomenon in information theory [33].

The quantity in Equation (4) is a measure of the indeterminateness at moment *t* and it characterizes the incompleteness of our information regarding the processes occurring at the microscopic level.

In the case of irreversible processes, the entropy of the state of disequilibrium increases continuously with time. The state of equilibrium for a perfectly isolated system is characterized by the maximum entropy because it holds that:

$$\frac{\partial S}{\partial t} \geq 0. \tag{5}$$

Consequently, the transition to the state of equilibrium is characterized by an increase in indeterminateness and a decrease in the information level. Thus, in terms of Equation (5), the state of equilibrium has an extreme character, as follows.

$$S_{stac.} = max \tag{6}$$

Therefore, the probability density of the state for the equilibrium does not depend on time. The entropy value is maximized compared with all preceding moments for the transition state directed at attaining this equilibrium state. However, the equilibrium state does not need to be unique in the final volume of the phase space because it corresponds to nonlinear systems.

Similar to all other methods, this procedure is not fully universal and it may be recommended as effective for only a special group of systems. This is due to the basic philosophy of the principle and mainly because of the need to obtain a solution with sufficient effort to provide transparent results that allow a physical interpretation. However, some potential problems affect the appropriate techniques used for searching for an extreme of the Gibbs functional. These approaches can work correctly in the case of systems with highly distinct nonlinear normal modes (NNMs) with minimal mutual interaction and the marginal influence of local modes. The problems are closely related to the possibility of effectively employing Boltzmann type solutions (e.g., see [23]) as a starting approximation for subsequent iterative steps.

The practical application of the functional based on Equation (4) to problems in statistical physics is based on the fact that the probability density $p(\mathbf{v}, t)$ is considered to be a known quantity determined in experiments, or based on independent partial considerations according to the specific conditions of the system. For instance, Gibbs introduced the so-called micro-canonic probability distribution for the adiabatic process in gases. However, the statistical mechanics of systems must be based on the assumption that the probability density of a response is an unknown function, particularly for nonlinear systems.

If the input or output are not characterized by the presence of white noise, Equation (6) is not sufficient for determining $p(\mathbf{x}, t)$, and it does not contain any information about the structure of the system. However, every potential system is described by the equations of motion Equation (1), with which $p(\mathbf{x}, t)$ must also comply. Therefore, we can formulate the following variation of the problem as finding the cross-probability density $p(\mathbf{x})$ of a stationary response process where the functional Equation (4):

$$S = - \int_{(\mathbf{x})} p(\mathbf{x}) \cdot \lg \ p(\mathbf{x}) d\mathbf{x}, \tag{7}$$

attains the maximum and the equations of motion Equation (1) comply with the statistical meaning, such as the meaning of the stochastic moments. Thus, we seek the weak solution of the problem in terms of the stochastic meaning, which differs from a weak solution in terms of the classical meaning. In Equation (7), we can use $\kappa_B = 1$ because the entropy will be evaluated only as a relative value when searching for the maximum from a set of admissible functions $p(\mathbf{x})$ and not as an absolute physical quantity. Some attributes considered in extreme search by [52] can provide inspiration, but care should be taken because the functional Equation (7) is not quadratic. Therefore, basic theorems regarding the extreme existence do not hold.

3. Formulation of the Secondary Constraints

As mentioned in Section 2, following Equation (6), the maximum of S should be searched for with respect to the secondary constraints. The secondary constraints on the extreme may be selected in various ways, but they should specify an exact structure for the dynamic system considered in every case. This structure is primarily given by Equation (1). In terms of the stochastic meaning, the most

natural system character specification follows from the FPE, i.e., Equation (2) with a scalar unknown quantity PDF. The coefficients of the FPE include all of the characteristics of Equation (1).

For the stationary case of the FPE, Equation (2) has the form:

$$\frac{\partial\left(\kappa_j(\mathbf{x})p(\mathbf{x})\right)}{\partial x_j} - \frac{1}{2}\frac{\partial^2(\kappa_{jk}(\mathbf{x})\cdot p(\mathbf{x}))}{\partial x_j\partial x_k} = 0, \tag{8}$$

where $\kappa_j(\mathbf{x}), \kappa_{jk}(\mathbf{x}), j, k = 1, \ldots, n$—drift and diffusion coefficients (moments of the 1st and 2nd orders) of the n-dimensional processes, which are now limited to Gaussian white noise independent of time t.

In this case, the drift and diffusion coefficients in Equation (3) can be simplified significantly and written as closed formulae:

$$\kappa_j(\mathbf{x}) = f_j(\mathbf{x}) + \frac{1}{2}K_{rs}\cdot g_{ls}(\mathbf{x})\frac{\partial_{jr}(\mathbf{x},t)}{\partial x_l} \ ; \qquad \kappa_{jk}(\mathbf{x}) = K_{rs}\cdot g_{jr}(\mathbf{x},t)g_{ks}(\mathbf{x},t), \tag{9}$$

where:

- $K_{rs} = \mathcal{E}\{w_r(t)w_s(t)\}$—matrix of white noise intensities (independent of time).
- $\mathcal{E}$—operator of mathematical mean value.

It should be noted with respect to Equation (3) that the drift and diffusion coefficients can be evaluated also in more complicated cases when the spectral densities of input processes are not constant. The following formulae can be used:

$$\kappa_j(\mathbf{x}) = f_j(\mathbf{x}) + \int_{-\infty}^{0} g_{ls}(\mathbf{x}_{t+\tau},t+\tau)\frac{\partial}{\partial \mathbf{x}_l}g_{jr}(\mathbf{x}_t,t)R_{rs}(\tau)\mathrm{d}\tau,$$

$$\kappa_{jk}(\mathbf{x}) = \int_{-\infty}^{\infty} g_{jr}(\mathbf{x}_t,t)g_{ks}(\mathbf{x}_{t+\tau},t+\tau)R_{rs}(\tau)\mathrm{d}\tau. \tag{10}$$

where $R_{rs}(\tau)$ is a general cross-correlation function of processes w_r, w_s. When the correlation function degenerate to Dirac functions $R_{rs} = K_{rs}\delta(\tau)$ as it corresponds with white noises, expressions Equation (10) turn into (9). Application of formulae Equation (10) for processes with bi-modal spectral densities is shown for instance in [11] or [53]. However, from the viewpoint of demonstration of the maximum entropy principle in dynamics this aspect does not play a significant role and the reduced relations (9) are valid.

In the case considered, the probability density must satisfy the boundary conditions as follows:

$$p(\mathbf{x})\Big|_{x_j\to\pm\infty} = 0 \ ; \ \frac{\partial p(\mathbf{x})}{\partial x_j}\bigg|_{x_j\to\pm\infty} = 0 \tag{11}$$

If we substitute according to Equation (1) into coefficients κ_j, κ_{jk} following Equation (9), or into Equation (3) with the time omitted, we obtain the following.

$$\kappa_j = -f_j(x_1,\cdots,x_n)$$
$$\kappa_{jk} = g_{jl}(x_1,\cdots,x_n)g_{lk}(x_1,\cdots,x_n)s_{ll} \tag{12}$$

In these conditions, we can consider Equation (8) as the secondary constraints on the functional Equation (7) extreme. Thus, the problem has been determined and it is possible to begin searching for the stationary point or points, if any exist.

In this particular case, it is best to formulate the secondary constraints in the form of stochastic moments for clarity. The equations can be deduced from the FPE. In the case of a linear system excited by a Gaussian process, if the solution exists, then using only the first and the second moments will

be sufficient for its full description. In the case of nonlinear systems, an infinitely large system of mutually bound equations will emerge and this type of system is generally nonlinear. In general, the series of stochastic moments must be convergent and the cumulants should then be considered instead. Nevertheless, for brevity, let us assume that the stochastic moments are convergent, which allows us to restrict the system using some criterion (e.g., declaring the moments only up to a certain order as independent and the moment of a higher order as their functions, e.g., the Gaussian closure in the simplest case).

Thus, if the number of moments and the number of equations is limited, it is generally possible to find a high number of probability densities $p(\mathbf{x})$ that satisfy these equations. If the number of moments remaining in the problem increases, then the size of the set from which we seek the maximum also decreases. In the limit, this set will contain only one element in an arbitrarily small neighborhood and the functional Equation (7) will only characterize the given state. However, the whole phase space may contain a countable number of such stationary points, which correspond to different equilibrium states according to the different levels of the functional value in Equation (7). In this case, the whole philosophy of the variational solution to $p(\mathbf{x})$ is similar to the philosophy of the direct variational methods used in the theory of elasticity as an example. It should be noted that this overall approach is only applicable close to the stationary point where the state of the system can be considered at least approximately quasi-stationary, and thus time can be considered as a parameter. In addition, it should be noted that the notion of entropy can be extended by expanding the manipulation space ensuring non-stationary processes can also be investigated (e.g., see [54]). However, only stationary problems are considered in the present study and this extension is not addressed.

4. One-Component Systems

4.1. Directly Finding the Extreme—Boltzmann's Solution

The one-dimensional diffusion process is described by a stochastic equation of first order:

$$\dot{x} + f(x) = w(t), \tag{13}$$

where $f(x)$ is a smooth integrable function. The system is excited by Gaussian white noise $w(t)$ with intensity s on the right side. This equation (reduced Langevin type) appears frequently in applied physics and engineering, particularly in earthquake-related problems regarding wave propagation in a significantly pronounced hysteretic continuum where inertia forces are marginal and dissipation forces are dominant.

The coefficients Equation (12) associated with Equation (13) have the form:

$$\kappa_1(x) = f(x) \, , \ \kappa_{11} = s^2 \, , \tag{14}$$

therefore, the FPE reads:

$$\frac{\partial}{\partial x}\left(f(x)p(x)\right) + \frac{1}{2}s\frac{\partial^2 p(x)}{\partial x^2} = 0. \tag{15}$$

Equation (15) with the respective boundary conditions unequivocally satisfies the Boltzmann type solution:

$$p(x) = C\exp\left(-\frac{2}{s}\int f(x)\mathrm{d}x\right). \tag{16}$$

In this process, it is necessary to impose certain conditions on $f(x)$ because $p(x)$ must comply with the requirements imposed on the density of the probability distribution (normalization, existence of moments, etc.).

However, Equation (16) can also be obtained by searching for the maximum of the functional Equation (7) with the secondary constraint Equation (15), which can be satisfied by the weak solution

in terms of the statistical meaning. For odd functions $f(v)$, it is possible to obtain a derivation from Equation (15) for statistical moments:

$$\int_{-\infty}^{\infty} \left(x^k f(x) - \frac{1}{2} s k x^{k-1} \right) p(x) \mathrm{d}x = 0. \tag{17}$$

In addition, it is necessary to respect the normalization constraint:

$$\int_{-\infty}^{\infty} p(x) \mathrm{d}x = 1. \tag{18}$$

Thus, the maximum of the functional Equation (7) is limited by the secondary constraints Equations (17) and (18). Using the Lagrange multipliers, we arrive at the problem of the unconditional extreme of the functional:

$$S = -\int_{-\infty}^{\infty} p(x) \lg p(x) \mathrm{d}x - (\lambda_0 - 1) \int_{-\infty}^{\infty} p(x) \mathrm{d}x - \int_{-\infty}^{\infty} p(x) \sum_{1}^{\infty} \lambda_k \left(x^k f(x) - \frac{1}{2} s k x^{k-1} \right) \mathrm{d}x, \tag{19}$$

where λ_0, λ_k are the unknown Lagrange multipliers. λ_0 is related to the normalization of the PDF.

The annulment of the variation for expression Equation (19) with respect to $p(x)$ using the continuity of the integrand in Equation (19) yields the general expression for the extreme, as follows.

$$p(x) = e^{-\lambda_0} \cdot \exp \left(-\sum_{k=1}^{\infty} \lambda_k \left(x^k f(x) - \frac{1}{2} s k x^{k-1} \right) \right) \tag{20}$$

The coefficient λ_0 can be determined from the normalization constraint and λ_k by re-substituting Equation (20) into Equation (17). In this manner, we arrive at the following relation after several modifications.

$$\sum_{k=1}^{\infty} \lambda_k \left(x^k f(x) - \frac{1}{2} s k x^{k-1} \right) = \frac{2}{s} \left(\frac{1}{1! \cdot 2} f_0' x^2 + \frac{1}{3! \cdot 4} f_0''' x^4 + \frac{1}{5! \cdot 6} f_0^{\mathrm{v}} v^6 + \cdots \right) \tag{21}$$

However, the right-hand side of Equation (21) is the Taylor's series of the indefinite integral for the odd continuous function $f(x)$. Thus, Equation (19) has a form that is entirely identical to Equation (15), which we aimed to prove. Therefore, according to Equation (16), $p(x)$ determines the extreme value of the functional Equation (7) while complying with the secondary constraint Equation (15). In this simple case, we have successfully complied with an infinite number of these conditions (in terms of the moments). Therefore, the selection of the set is reduced to a single element Equation (16), which simultaneously satisfies the equation or the secondary constraint Equation (15) in terms of the statistical meaning as well as the classical solution.

The secondary constraints mentioned several times in this study are symbolically treated in the form of the stationary version of the FPE in Equation (8). This method is preferable given that the basic considerations have general characteristics. The decomposition of the stochastic moments is preferable for obtaining practical solutions to particular problems (although other approaches are also suitable). We consider several applications of this strategy in the following sections of this study. Therefore, it is useful to consider the basic character of convergence when manipulating using this decomposition.

Let us return to the series comprising Equation (17) and the normalization constraint Equation (18). We assume the cubic characteristic $f(x) = ax + bx^3$. The first approximation is obtained for $k = 1$, which means that:

$$p(x) = C \exp \left(-\lambda_1 \left(x \left(ax + bx^3 \right) - \frac{1}{2} s \right) \right). \tag{22}$$

By introducing Equation (22) into Equation (17) for $k = 1$, we obtain an algebraic equation for the multiplier λ_1:

$$\int_{-\infty}^{\infty} \left(x(ax + bx^3) - \frac{1}{2}s \right) C \exp \left(-\lambda_1 \left(x \left(ax + bx^3 \right) - \frac{1}{2}s \right) \right) \mathrm{d}x = 0. \tag{23}$$

The integrals in Equation (23) can be evaluated exactly in cylindric functions, e.g.,

$$\int_{-\infty}^{\infty} \exp \left(-\lambda_1 \left(ax^2 + bx^4 - \frac{1}{2}s \right) \right) \mathrm{d}x = \sqrt{\frac{a}{4b}} \exp \left(\frac{1}{2}\lambda_1 \left(s + \frac{a^2}{4b} \right) \right) \cdot K_{1/4} \left(\frac{\lambda_1 a^2}{8b} \right), \tag{24}$$

where $K_{1/4}(\frac{\lambda_1 a^2}{8b})$ is the Macdonald function. Clearly, in the case of common applications, the integrals in Equation (24) should be evaluated numerically, which causes some difficulties because unknown values or the values of λ_k are parts of the arguments of the exponential functions. Nevertheless, for this example, we can plot (see Figure 3) curves where we illustrate the variance σ_x^2 of the response x for the stochastically linearized case in curve (a), those evaluated at the levels of the first and second approximations $q = 1, 2$ in curves (b) and (c), respectively, and those obtained with the exact Boltzmann solution using Equation (16) in curve (d). Even the very simple approximation obtained using the probability entropy maximum strategy is effective and it converges very rapidly to the exact solution. The results obtained for the second approximation and the Boltzmann solution are almost indistinguishable.
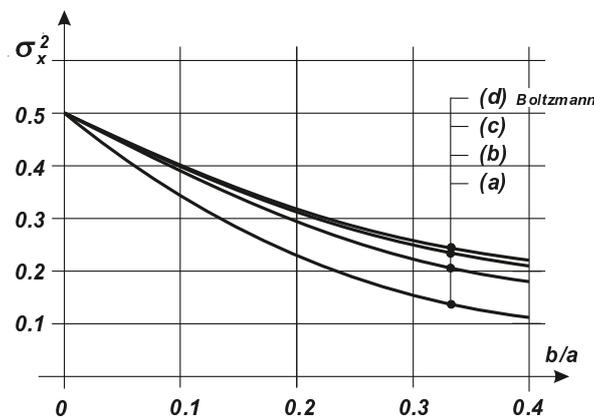

Figure 3. Comparison of the variance σ_x^2 evaluated for: (**a**) system with stochastic linearization; (**b**) system with the maximum entropy of probability, as the first approximation $q = 1$; (**c**) the second approximation $q = 2$; (**d**) Boltzmann solution.

4.2. First Order System with Complex Characteristics

We illustrate the effectiveness of the maximum entropy probability principle for another case with a one-component response process, which is produced by a nonlinear system of first order with relatively complex characteristics:

$$\frac{\mathrm{d}x}{\mathrm{d}t} + ax + bx^2 + cx^3 \lg \frac{|x|}{\alpha} = w(t), \tag{25}$$

where

- $a, b, c > 0, \alpha > 0$ are real coefficients,
- $w(t)$ is a Gaussian white noise with intensity K.

The detailed analysis was reported by Sobczyk and Trebicki [46,47]. A small number of previous studies focused on the principle result obtained with respect to Equation (25). In addition to these studies, mathematical models related to Equation (25) can be found in seismicity research in investigations of the passage of nonlinear waves through a well-saturated medium containing a liquid material. In addition, the aero-elasticity of systems with super-critical damping is related to an equation of this type, as well as the theory of diffusive chemical processes.

The relevant drift and diffusion coefficients are:

$$\kappa_1 = -\left(ax + bx^2 + cx^3 \lg \frac{|x|}{\alpha}\right), \quad \kappa_{11} = s^2. \tag{26}$$

With respect to the Boltzmann general solution for Equation (16), it holds that:

$$p(x_1) = C \exp\left(-\frac{x^2}{s^2}\left(a + \frac{2}{3}bx + \frac{1}{2}cx^2\left(\lg \frac{|x|}{\alpha} - \frac{1}{4}\right)\right)\right), \tag{27}$$

where C is a normalizing constant. Using Equation (27), the exact stochastic moments can be evaluated to demonstrate the fidelity of an exact PDF characterization, as plotted in Figure 4 for the exact PDF, where Equation (27) and the PDFs are determined based on $q = 2, 4, 6$ first stochastic moments. The specific values of the parameters are as follows: $a = 0.5, b = -0.25, c = 5.2, \alpha = 1.0, s = 1.0$. For $q = 2$, a Gaussian distribution is provided and thus the result is unacceptable. Nevertheless $q = 4$ and $q = 6$ exhibit very rapid convergence and they can represent a satisfactory approximation of the PDF according to Equation (27). It should be noted that the exact PDF exhibits approximately bi-modal characteristics. Therefore, the basic approximation for $q = 2$ should pass very quickly from the uni- to bi-modal type of the PDF.
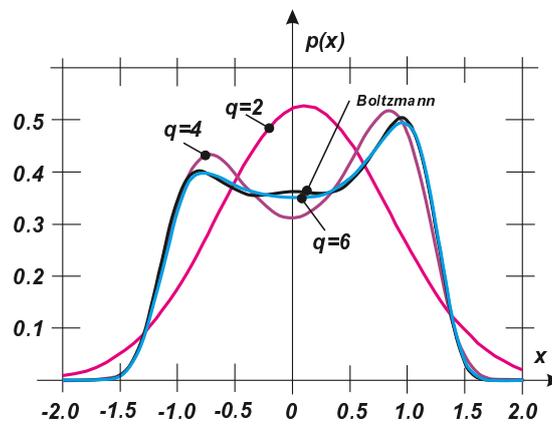

Figure 4. Comparison of the exact PDF (Boltzmann) and the PDFs determined using the given stochastic moments ($q = 2, 4, 6$).

Next, we consider a limited number of equations for the first stochastic moments ($q = 2, 4, 6$). According to the general method for deriving the equations for the stochastic moments of the relevant Ito equations (e.g., [11,12,23]), the following hierarchy of equations for stochastic moments can be deduced:

$$\int_{-\infty}^{\infty} \kappa_1 x^{k-1} dx + \frac{1}{2}(k-1)\int_{-\infty}^{\infty} \kappa_{11}^{1/2} x^{k-2} dx = 0, \quad k = 1, 2, ..., q, \tag{28}$$

where κ_1, κ_{11} are given by Equation (26). The standard codes using a homotopy background were used to evaluate some example results. The values of the parameters were identical to those employed above: $a = 0.5, b = -0.25, c = 5.2, \alpha = 1.0, s = 1.0$. The exact PDF and the PDFs corresponding to two and four given moments are shown in Figure 5. Clearly, the PDF determined using the maximum

entropy principle is very close to the exact PDF even when using four stochastic moment equations, Equation (28). The convergence can be assessed using the objective Kullback–Leibler divergence criterion (see [34,55]):

$$\Delta_{pp^*} = \int_{-\infty}^{\infty} p(x) \cdot \lg \frac{p(x)}{p^*(x)} \mathrm{d}x \tag{29}$$

where

- $p(x)$—the exact PDF,
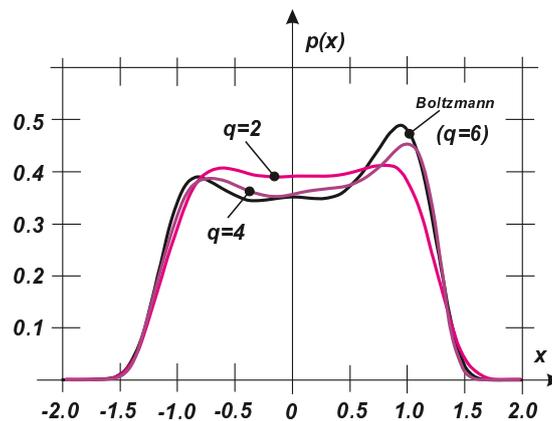- $p^*(x)$—the approximate PDF corresponding to $q = 2, 4, 6$, respectively.

Figure 5. Comparison of the exact PDF (Boltzmann) and the PDFs determined with various numbers of stochastic moment equations ($q = 2, 4, 6$).

The respective divergence parameter Δ_{pp^*} has values of:

$$\Delta_{pp^*}^{(2)} = 0.01053260, \quad \Delta_{pp^*}^{(4)} = 0.00014443, \quad \Delta_{pp^*}^{(6)} = 0.00000278, \tag{30}$$

which demonstrates the rapid convergence of the approximate PDFs to the exact solution of the Boltzmann type. There is no clear difference between the solutions based on the Boltzmann type method and $q = 6$ maximum entropy (see Figure 5), as also shown in Section 4.1 and Figure 3.

5. Two-Component Systems

5.1. System with Diffusion Additive Excitation

We next describe a direct procedure for finding the PDF using the maximum entropy probability principle. We consider Equation (13):

$$\dot{x} + f(x) = \eta(t), \tag{31}$$

where $f(x)$ is a smooth function that is integrable on a limited interval and the process $\eta(t)$ follows from a filtering of white noise through a linear differential filter of the first order:

$$\dot{\eta} + \alpha \eta = w(t)v, \tag{32}$$

where $s = 2\alpha\sigma_\eta^2$ is the intensity of the white noise $w(t)$. The relevant spectral density of the η process (see Figure 6) is given by the formula:

$$\Phi_\eta(\omega) = \frac{\sigma^2}{\pi} \frac{\alpha}{\omega^2 + \alpha^2}. \tag{33}$$
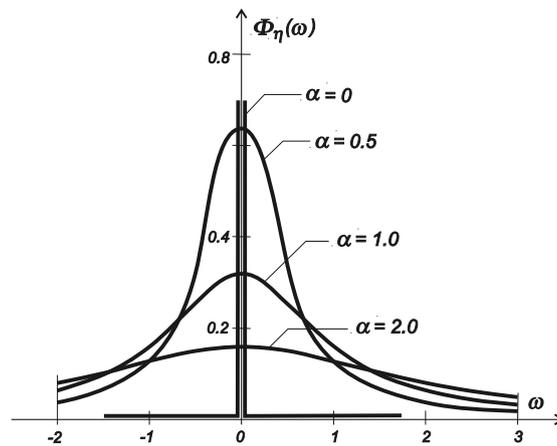
Figure 6. Spectral density $\Phi_\eta(\omega)$ of a diffusion process $\eta(t)$ for various parameters α.

Hence, we can define a two-component random process in terms of the meaning of the system in Equation (1):

$$\begin{aligned}
\dot{x}_1 &= -f(x_1) + x_2 \\
\dot{x}_2 &= -\alpha x_2 + w(t).
\end{aligned} \tag{34}$$

The cross-PDF $p(x_1, x_2)$ of a stationary state can be described with respect to Equation (8) and the coefficients in Equation (9) by the following FPE:

$$\frac{1}{2}s\frac{\partial^2 p}{\partial x_2^2} + \alpha\frac{\partial(x_2 p)}{\partial x_2} + \frac{\partial}{\partial x_1}\left((f(x_1) - x_2)p\right) = 0. \tag{35}$$

The closed form solution of the FPE in Equation (35) cannot be obtained. Therefore, we try to apply probability entropy maximization to the functional S (see Equation (7)).

Let us multiply Equation (35) successively by the factors $x_1^2, x_1 x_2, x_2^2$, and subsequently apply integration on the infinite two-dimensional phase space. After several integration by parts steps with respect to the homogeneous boundary conditions in infinity for both components x_1, x_2, we obtain three equations for the stochastic moments, which should be completed by the fourth equation to represent the constraints on normalization:

$$\iint\limits_{-\infty}^{\infty} (x_1 f(x_1) - x_1 x_2)\, p(x_1, x_2)\mathrm{d}x_1\mathrm{d}x_2 = 0,$$

$$\iint\limits_{-\infty}^{\infty} \left(x_2 f(x_1) + \alpha x_1 x_2 - x_2^2\right) p(x_1, x_2)\mathrm{d}x_1\mathrm{d}x_2 = 0,$$

$$\iint\limits_{-\infty}^{\infty} \left(\tfrac{1}{2}K - \alpha x_2^2\right) p(x_1, x_2)\mathrm{d}x_1\mathrm{d}x_2 = 0, \tag{36}$$

$$\iint\limits_{-\infty}^{\infty} p(x_1, x_2)\mathrm{d}x_1\mathrm{d}x_2 = 1.$$

The extended functional S^* including the secondary and normalization constraints with the corresponding Lagrange multipliers is:

$$S^* = -\iint_{-\infty}^{\infty} p \lg p \, dx_1 dx_2 - \iint_{-\infty}^{\infty} \left[(\lambda_0 - 1) + \lambda_1 \left(x_1 f(x_1) - x_1 x_2 \right) + \right. \tag{37}$$

$$\left. + \lambda_2 \left(x_2 f(x_1) + \alpha x_1 x_2 - x_2^2 \right) + \lambda_3 \left(\tfrac{1}{2} s - \alpha x_2^2 \right) \right] p \, dx_1 dx_2 .$$

According to the principle of maximum entropy, the PDF with the maximum entropy for the functional in Equation (37) has the form:

$$p(x_1, x_2) = C \exp \left[-\lambda_1 \left(x_1 f(x_1) - x_1 x_2 \right) - \lambda_2 \left(x_2 f(x_1) + \alpha x_1 x_2 - x_2^2 \right) - \lambda_3 \left(\tfrac{1}{2} s - \alpha x_2^2 \right) \right], \tag{38}$$

where $C = \exp(-\lambda_0)$.

By substituting Equation (38) back into Equation (36), we obtain four algebraic equations for $\lambda_0 - \lambda_3$, which need to be solved for a particular case of the function $f(x_1)$. This rather complicated algebraic system can be solved effectively by employing the same packages that are commonly used for homotopy continuation methods.

A good initial approximation can serve as a function corresponding to the Gaussian PDF and linear characteristic $f(x_1) = cx_1$:

$$p_0(x_1, x_2) = C \exp \left(-\frac{1}{2(1 - r^2)} \left(\frac{x_1^2}{\sigma_1^2} - \frac{2 r x_1 x_2}{\sigma_1 \sigma_2} + \frac{x_2^2}{\sigma_2^2} \right) \right), \tag{39}$$

where

- σ_1^2, σ_2^2—variances of x_1, x_2 for the adjoint linear system,
- r—correlation of these coordinates.

For the cubic nonlinearity of the system characteristic, $f(x_1) = ax_1 + bx_1^3$ was evaluated fully as an example problem where the following values were used: $a = 1, b = 0.1, \alpha = 0.5, \sigma_1^2 = \sigma_2^2 = s/2\alpha = 1$. The plot in Figure 7 shows the one-dimensional PDFs reduced to coordinate x_1, as usual:

$$p(x_1) = \int_{-\infty}^{\infty} p(x_1, x_2) dx_2. \tag{40}$$

Figure 7 shows the PDF evaluated based on the stochastic linearization as curve (b), the approximation after maximization of the functional S^* with Equation (37) as curve (c) at the level of $q = 2$, and the raw histogram obtained for x_1 by stochastic simulation as the stepped curve in (a). A comparison of curves (a) and (c) clearly shows that the approximation of the second order moments provides acceptable results, whereas the linearization is far from correct.
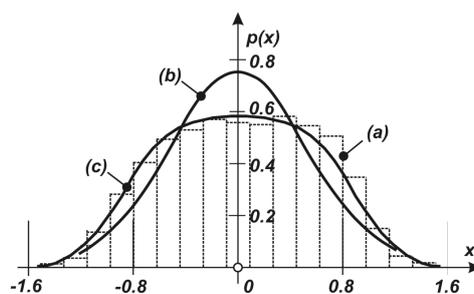

Figure 7. (a) PDF in x_1 following the histogram of the output process x_1. (b) Stochastic linearization. (c) Second moments approximation.

5.2. Dynamic System with a Single Degree of Freedom

The vibrations of a system with one degree of freedom excited by white noise are described by a normal system of Equation (1) type, where $n = 1$ and thus it contains two components x_1, x_2:

$$\begin{aligned}\dot{x}_1 &= x_2 \\ \dot{x}_2 &= (2\omega_b x_2 + \omega_0^2 x_1 + f(x_1)) + w(t),\end{aligned} \tag{41}$$

where the argument t in x_1, x_2 is omitted for brevity, therefore $x_1(t) = x_1$; $\dot{x}_1(t) = x_2(t) = x_2$. Hence, the FPE is:

$$-\frac{\partial(x_2 p)}{\partial x_1} + \frac{\partial}{\partial x_2}\left(\left(2\omega_b x_2 + \omega_0^2 x_1 + f(x_1)\right)p\right) + \frac{1}{2}s\frac{\partial^2 p}{\partial x_2^2} = 0, \tag{42}$$

where:

- $p = p(x_1, x_2)$—cross-probability density of $x_1(t)$, $x_2(t)$
- s—intensity of white noise $w(t)$.

The solution may be obtained for Equation (42) with the boundary conditions in Equation (11) by using the Fourier method in the following form.

$$p(x_1, x_2, t) = C\exp\left(-\frac{4\omega_b}{s}\frac{x_2^2}{2}\right) \cdot \exp\left(-\frac{4\omega_b}{s}\left(\frac{\omega_0^2 x_1^2}{2} + \int f(x_1)\mathrm{d}x_1\right)\right) \tag{43}$$

Thus, the displacement and displacement velocity are statistically independent. However, similarly to the preceding case, Equation (11) can be arrived at by finding the extreme of the functional Equation (7) with the secondary constraint in Equation (42).

If $f(x)$ is an odd function, then based on Equation (42):

$$\iint\limits_{-\infty}^{\infty} x_1^k x_2^l \left(-\frac{\partial(x_2 p)}{\partial x_1} + \frac{\partial}{\partial x_2}\left(\left(2\omega_b x_2 + \omega_0^2 x_1 + f(x_1)\right)p\right) + \frac{1}{2}K\frac{\partial^2 p}{\partial x_2^2}\right)\mathrm{d}x_1\mathrm{d}x_2 = 0, \tag{44}$$

from which and after several adjustments, we obtain:

$$\iint\limits_{-\infty}^{\infty} P_{kl}(x_1, x_2) \cdot p(x_1, x_2)\mathrm{d}x_1\mathrm{d}v_x = 0 \tag{45}$$

$$P_{kl}(x_1, x_2) = kx_1^{k-1}x_2^{l+1} - 2l\omega_b x_1^k x_2^l - l\omega_0^2 x_1^{k+1}x_2^{l-1} - lx_1^k x_2^{l-1}f(x_1) + \frac{1}{2}sl(l-1)x_1^k x_2^{l-2}. \tag{46}$$

The whole functional with an unconstrained extreme has the form:

$$S^* = -\iint\limits_{-\infty}^{\infty} p \cdot \lg\ p \cdot \mathrm{d}x_1\mathrm{d}x_2 - (\lambda_0 - 1)\iint\limits_{-\infty}^{\infty} p\mathrm{d}x_1\mathrm{d}x_2 - \iint\limits_{-\infty}^{\infty} p\sum_{k,l=1}^{\infty}\lambda_{kl}\cdot P_{kl}(x_1, x_2)\mathrm{d}x_1\mathrm{d}x_2, \tag{47}$$

and thus the extreme follows as:

$$p(x_1, x_2) = C\cdot\exp\left(-\sum_{k,l=1}^{\infty}\lambda_{kl}\cdot P_{kl}(x_1, x_2)\right) \tag{48}$$

and λ_0 is included in the normalization constant C. The constants λ_{kl} can be obtained by re-substituting Equation (48) into Equation (46). After analogous operations to those in Section 4.1, a similar procedure to that described in Section 4.1 brings us from Equation (48) to Equation (43).

Next, we demonstrate this procedure based on the Mieses strutted frame with one degree of freedom, as shown in Figure 8. The fact that the system exhibits the snap-through effect does not affect

the numerical aspects of the procedure in terms of the numerical stability or convergence velocity. The governing system reads:

$$
\begin{aligned}
\dot{x}_1 &= x_2 \\
\dot{x}_2 &= \frac{1}{m}\left(-2bx_2 - K(2h^2x_1 - 3hx_1^2 + x_1^3)\right) + \frac{w(t)}{m},
\end{aligned}
\tag{49}
$$

where:

- h—Mieses strutted frame rise,
- $K = EF/L^3$—"longitudinal stiffness" of one bar of the strutted frame,
- m—concentrated mass in the movable node of the strutted frame.
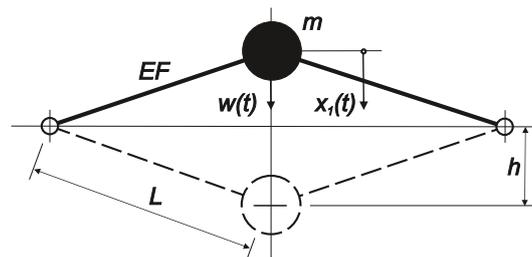

Figure 8. Mieses strutted frame as a nonlinear SDOF system with white noise excitation.

If $w(t)$ is a stationary white noise, the relevant FPE has the form:

$$
-\frac{\partial x_2 p}{\partial x_1} + \frac{\partial}{\partial x_2}\left(\left(\frac{2b}{m}x_2 + \frac{K}{m}(2h^2x_1 - 3hx_1^2 + x_1^3)\right)p\right) + \frac{s}{2m^2}\frac{\partial^2 p}{\partial x_2^2} = 0.
\tag{50}
$$

Equation (50) has an analytic Boltzmann solution (e.g., as described by [12] or [23]):

$$
p(x_1, x_2) = C \cdot \exp\left(-\frac{2bm}{s}\cdot x_2^2\right) \cdot \exp\left(-\frac{4bK}{s}\cdot x_1^2(\tfrac{1}{2}x_1 - h)^2\right),
\tag{51}
$$

where C is a normalization constant (see Equation (48)).

This solution can be obtained based on the maximum of the only functional Equation (7) with the secondary constraints in Equation (50). If we consider $q = 4$ terms of the infinite series in Equation (48), then for the moments, we can write: $\mu_{kl}(k, l = 0, \ldots, 4; k + l \leq 4)$, with 62 equations in total for 15 unknown quantities. As an example, we only select the first two of these equations to obtain the values of the unknown quantities that differ from zero:

$$
\frac{2b}{m} - \frac{s}{m^2}\mu_{20} + \frac{s}{2m^2}\mu_{10}^2 = 0
$$

$$
2\mu_{02} - \frac{2b}{m}\mu_{11} - \frac{4Kh^2}{m}\mu_{20} - \frac{3s}{m^2}\mu_{31} + \frac{2s}{m^2}\mu_{10}\mu_{21} + \frac{2s}{m^2}\mu_{11}\mu_{20} = 0
\tag{52}
$$

$$
\text{etc.}
$$

Equation (52) have the following solution:

$$
\mu_{20} = \frac{2bm}{s}\ ;\ \mu_{02} = \frac{4bKh^2}{s}\ ;\ \mu_{03} = -\frac{4bKh}{s}\ ;\ \mu_{04} = \frac{bK}{s},
\tag{53}
$$

which coincides perfectly with the Boltzmann solution. Other values of μ_{kl} vanish, except for μ_{00}, which serves as the normalization constant, $\exp(\mu_{00}) = C$, and it is also determined after Equation (48) has been cleaned of all terms where μ_{kl} is equal to zero. Equation (49) comprise a typical system

with a marked non-Gaussian bi-modal response. Maximizing the functional in Equation (7) using Equation (47) with the known polynomials $P_{kl}(x_1, x_2)$ to search for the unknown multipliers λ_{kl} yields the final result. The solution to Equation (48) characterizes this situation well and leads to the known classical solution, which confirms the validity of the maximum entropy principle for determining the PDF of a dynamic system. The numerical experiments also confirmed the rapid convergence with $q = 4$ as a satisfactory value to match the Boltzmann solution with Equation (51). Stochastic simulation was also used as an independent tool for verification. Careful numerical processing should be performed due to the stochastic character of the original stochastic differential system in terms of integration stepping and the corrector steps for improving the stability (e.g., see [56]).

6. Dynamic Systems with Many Degrees of Freedom

6.1. General Formulation

The motion of a system of concentrated masses with n degrees of freedom and potential links is described by Equation (1). In order to maximize the functional S according to Equation (7) and with the secondary constraints in Equation (8) in either the direct form of the FPE or the form of stochastic moments, we can investigate any admissible dynamic system response PDF. Nevertheless, as shown in the previous sections, the maximization techniques for particular cases can be complicated and ineffective, especially when $n > 2$ and higher. However, some special classes of dynamic systems allow significant simplification despite a high value of n, particularly when the stochastic moments are convergent and this allows us to approximate the respective PDF with any arbitrary precision.

The original system Equation (1) can be rewritten in the form:

$$\dot{\mathbf{x}}_1 = \mathbf{x}_2$$
$$\mathbf{m}\dot{\mathbf{x}}_2 = -2\varepsilon\mathbf{x}_2 - \frac{\partial \mathcal{U}}{\partial \mathbf{x}_1} + \mathbf{Gw} \tag{54}$$

where:

- $\mathbf{m} \in \mathbb{R}^{n \times n}$—diagonal square matrix of concentrated masses m_j acting in individual degrees of freedom,
- $\boldsymbol{\varepsilon} \in \mathbb{R}^{n \times n}$—diagonal matrix of viscose damping coefficients ε_j,
- $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n$—vectors of displacement or velocity, respectively, in individual degrees of freedom,
- $\mathbf{w}(t) \in \mathbb{R}^m$—vector of white noise with intensity s_j applied as excitation forces
- $\mathbf{G} \in \mathbb{R}^{n \times m}$—rectangular matrix transforming white noise $\mathbf{w}$ into relevant degrees of freedom,
- $\mathcal{U} = \mathcal{U}(\mathbf{x}_1)$—potential energy (scalar function of displacements $\mathbf{x}_1$) of the system.

For example, Equation (54) can model the motion of a system of concentrated masses with large amplitudes if the individual masses are mutually interconnected by massless springs and the damping is affected by a "slightly" viscous environment, thereby allowing the definition of a framework with hinged nodes if the masses of the bars are concentrated in nodes.

The basic idea is obvious, i.e., searching for an alternate system (AS) that is close to the original and that exhibits some special properties to allow radical simplification in terms of degrees of freedom. In particular, this means that subsystems (precisely or approximately) can occur that work more or less independently of the remaining parts of the AS to provide a suitable functional basis (forming a separable functional space in the ideal case). If any can be found, the probabilistic problem can be solved based on the AS using suitable conventional solution methods. The result, such as a one point $p_0(\mathbf{x})$ or multiple point $p_{0i}(\mathbf{x})$ basis, then serves as a zero approximation when using the maximum entropy probability principle in the decomposition of stochastic moments. We outline two possibilities in the next two subsections.

6.2. Stochastically Proportional Systems

Provided that the conditions:

$$m_j = \text{const.} \; ; \; \varepsilon_j / s_j = \eta = \text{const.} \tag{55}$$

are satisfied it is possible to use the methods described in the previous sections to find the maximum of the functional Equation (7) with secondary constraints constructed based on Equation (54). The result is the formula corresponding to the Boltzmann solution (see [23]):

$$p(\mathbf{x}_1, \mathbf{x}_2) = \frac{1}{J} \exp\left(-4\eta \left(\mathcal{U} + \mathcal{K}\right)\right), \tag{56}$$

where $\mathcal{K} = \mathcal{K}(\mathbf{x}_1, \mathbf{x}_2)$ is the kinetic energy of the system.

It should be noted that the satisfied constraints leading to the PDF according to Equation (56) indicate the stochastic independence of the velocities x_2 with respect to the displacements x_1, which confirms the validity of the Heisenberg principle of indeterminacy. A system endowed with the property in Equation (55) can be referred to as a stochastically proportional system (SPS).

A system of equations for the unknown parameters λ with a higher number of degrees of freedom can only be constructed with a computer. The solution is complicated because the system of equations is nonlinear. The system and its solution can be simplified if the system approaches the proportional Hamiltonian system. In this case, it is possible to write the Boltzmann solution and use it as the initial approximation for the gradient method of the solution to the system for the parameters λ. These ideas form the basis of a computer program that generates a nonlinear system of equations for computing the parameters λ for a given geometry of the system (coordinates of nodes, rigidities C_k, masses m_j, damping matrices ε, and incidence tables $\mathbf{T}$) considered as an oriented graph, excitation intensities $\mathbf{K}$ (diagonal in a special case), and the upper limit q of the degree of the polynomial in Equation (48). At present, the program only works for stationary cases, i.e., with constant s_j. Therefore, the system of equations is algebraic and can be solved using standard methods.

The original code was assembled in order to verify the general strategies described above. The test examples are rather hypothetical. In order to fulfill the conditions in Equation (55), damping parameters ϵ_j and white noise intensities s_j were selected and the AS was examined. Next, settings that differed slightly from the conditions in Equation (55) were employed and the entropy probability was maximized.

Two systems were examined as follows.

1. The Mieses strutted frame with one mass and two degrees of freedom, as shown in Figure 9a, while only considering the SDOF system (vertical displacement) described in Section 5.2. At the level of $q = 4$, the algebraic system contained 70 unknown parameters λ and most of them equaled zero. Compared with Equation (55), the results did not differ qualitatively where the stochastic coherence velocity/displacement was still negligible, although the excitation acted only on one degree of freedom.

2. A non-symmetrical strutted frame with two masses and four degrees of freedom, as shown in Figure 9b, at the level of $q = 4$. Again, the velocity and displacement interaction was negligibly small (it was impossible to determine based simply on the approximate character of the overall solution process). The stochastic relationship of the individual displacements varied considerably in terms of the dependence on the excitation intensity. This relationship was small when the excitation was generally small and local snap-through did not occur with high probability. However, it increased steeply locally or globally immediately after overcoming some energy barrier that kept the motion within local limits. The signs of the coefficients λ for the fourth powers of the polynomial in the exponent of the function $p(x_1, x_2)$ were positive. Therefore, the boundary conditions in Equation (11) were satisfied without difficulty.
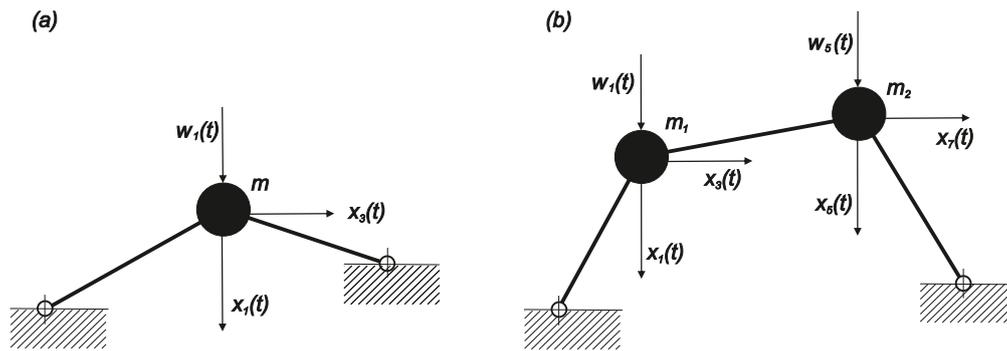
Figure 9. Testing strutted Mieses type frames with white noise excitation: (**a**) system with two degrees of freedom; (**b**) system with four degrees of freedom.

6.3. Transformation of the System with Respect to Nonlinear Normal Modes

Another suitable strategy appears to be using a combination of the maximum entropy probability principle with the preliminary involvement of NNM transformation. NNM was first proposed by Rosenberg [57] and extended subsequently (see [58,59]).

Analogous to the classical linear eigenmodes, the NNM is a tool for expressing the dynamic response of a system in certain generalized coordinates, which are energetically separated as far as possible from each other and thus they exhibit minimal interaction. If introduced correctly, they can be investigated individually and subsequently evaluated in terms of their interaction. At present, some commercial codes for dealing with the NNM are available. However, care should be taken because NNM can be "cheated" due to multivalent meanings. Therefore, local modes should be avoided. In general, a suitable manifold should be composed. If it is constructed successfully, the NNM can provide good quality results when used as zero iteration points for probability entropy maximization. A well separated NNM can lead to the very rapid convergence of the stochastic moments to obtain the final PDF of the system response. This is beneficial and the NNM strategy is actually used in earthquake engineering, and thus subsequent stochastic analysis based on entropy probability maximization can be facilitated by the NNM based on the results of previous activities.

7. Conclusions

Every earthquake event is a random process with a specific character. The occurrence of whatever seismic shock and a relevant ground motion at a particular site is not predictable. Therefore, the seismic process should be unavoidably considered as a stochastic process. This attribute has been adopted in the earthquake engineering many years ago and, therefore, we have a rich material providing wide possibilities to evaluate adequate dynamic characteristics of a stochastic nature.

In order to employ effectively collected data advanced methods of stochastic analysis of dynamic systems must be available. The method based on the maximum entropy probability density principle is worth to be considered. It seems to be very flexible from the viewpoint of structural variability and, moreover, it makes possible to intervene into parametric areas which are inaccessible for other methods. Therefore, this method enables also to serve as a tool for verification of results earned using other methods. It concerns predominantly structures which can be modeled as MDOF systems with concentrated masses, as it is a typical case of structures made of 1D elements (frameworks, trusses, etc.). Furthermore, all systems with well pronounced cyclic structure, easily transformable into state with well separable NNM, or systems close to those enabling closed form solutions of Boltzmann type are appropriate to be analyzed using this way. Many structures of this type can be encountered in industry and traffic engineering.

The problem of finding the probability density for a random movement of a system due to external random excitation can be considered as the problem of finding the constrained maximum of the functional of the entropy of the probability density. In the case of a system with polynomial

nonlinearities, these properties of the extreme can satisfy the secondary constraints derived from the FPE and they may be described best by a multidimensional exponential function with the exponent in the form of a complete polynomial of a certain degree with unknown coefficients. This can be determined from the FPE equation for the system considered by using the stochastic moment decomposition strategy. Our comparison of the results with analytically solvable cases demonstrated their good agreement, with very rapid convergence of the first stochastic moments for the PDF approximation. The advantage of this procedure is the certainty of the positive probability density values in the whole phase space. The disadvantage is the great number of unknown parameters λ, and the need to monitor their positive sign in the highest powers of the phase variables. In the opposite case, the problem loses stochastic stability in terms of the probability. The test cases based on simple MDOF systems indicated good numerical feasibility, although the test cases were located in close proximity to proportional systems or in the neighborhood of the dominant NNM.

Our general analysis and the results obtained for some particularly simple cases demonstrated how the response of a nonlinear system differed significantly from a Gaussian process even though the excitation had Gaussian characteristics. In addition, the principle considered in this investigation allowed us to prove that according to the number of processes with equal dispersal, a very normal process has the probability density with the highest entropy. Thus, the system always tends to produce a Gaussian response if it is permitted by the internal structure of the system. This is also the reason why Gaussian processes occur most frequently in nature.

An open problem comprises the possible existence of a greater number of extremes for a single fixed input and the practical procedure for finding them, particularly with respect to their initial approximation. This problem requires the prior elimination of some parameters λ because their number increases exponentially as both the number of degrees of freedom and the level of q increase. Experience indicates that most of the parameters λ have values that are either equal to zero or insignificant. Thus, the related stochastic moments are either negligible or they vanish. In addition, the problem of how to select the equations remains unsolved, where the number for a given level of q is several times higher than the number of unknown parameters λ. Some of these problems are closely related and they should be solved as a single group.

Obvious difficulties occur due to the increase in the number of independent phase variables x_j as the degrees of freedom increase in the system considered. However, these problems affect all methods for solving the FPE unless they are developed for a narrow class of systems with special properties, or if searching for only some special attributes of the solution. In the maximum entropy probability method, these cases can be considered as systems with well separated generalized coordinates by using the NNM in the preliminary step, or systems that are close to those that can be solved exactly, e.g., by employing the Boltzmann solution or systems with very weak nonlinearity. When searching for suitable classes of systems, this step is related to each specific FPE solution method or the qualitative investigation. An important strength of the strategy based on maximizing the probability entropy is the fact that it affects different classes where it has been proved effective, and thus it complements a set of existing methods for solving or investigating the FPE.

Author Contributions: Conceptualization, Methodology, Analytical investigation: J.N.; Software, Numerical analysis, Validation: C.F.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

FEM   Finite Element Method
FPE   Fokker–Planck Equation
MDOF   Multi-Degree of Freedom
SDOF   Single Degree of Freedom
PDF   Probability Density Function
NNM   Nonlinear Normal Mode
SPS   Stochastically Proportional System
AS   Alternate System

References

1. Askar, A.; Cakmak, A. Seismic waves in random media. *Probab. Eng. Mech.* **1988**, *3*, 124–129. [CrossRef]
2. Langley, R. Wave transmission through one-dimensional near periodic structures: Optimum and random disorder. *J. Sound Vib.* **1995**, *188*, 717–743. [CrossRef]
3. Manolis, G. Acoustic and seismic wave propagation in random media. In *Computational Stochastic Mechanics*; Cheng, A.D., Yang, C.Y., Eds.; Elsevier: Southampton, UK, 1993; Chapter 26, pp. 597–622.
4. Náprstek, J. Wave propagation in semi-infinite bar with random imperfections of density and elasticity module. *J. Sound Vib.* **2008**, *310*, 676–693. [CrossRef]
5. Náprstek, J.; Fischer, C. Planar compress wave scattering and energy diminution due to random inhomogeneity of material density. In Proceedings of the 16th World Conference on Earthquake Engineering, Santiago, Chile, 9–13 January 2017.
6. Náprstek, J. Sound wave scattering at randomly rough surfaces in Fraunhofer domain. In Proceedings of the Engineering Mechanics 2000 ASCE, Austin, TX, USA, 21–24 May 2000; paper N-8; Tassoulas, J., Ed.; University of Texas: Austin, TX, USA, 2000; 8p.
7. Zembaty, Z.; Castellani, A.; Boffi, G. Spectral analysis of the rotational component of eartquake motion. *Probab. Eng. Mech.* **1993**, *8*, 5–14. [CrossRef]
8. Zembaty, Z. Vibrations of bridge structure under kinematic wave excitations. *J. Struct. Eng.* **1997**, *123*, 479–488. [CrossRef]
9. Kanai, K. *Seismic-Empirical Formula for the Seismic Characteristics of the Ground*; Bulletin of the Earthquake Research Institute: Tokyo, Japan, 1957; Volume 35, pp. 309–325.
10. Tajimi, H. A statistical method of determining the maximum response of a building structure during an earthquake. In Proceedings of the 2nd World Conference on Earthquake Engineering, Tokyo, Japan, 11–18 July 1960; Science Council of Japan: Tokyo-Kyoto, Japan, 1960; pp. 781–798.
11. Lin, Y.; Cai, G. *Probabilistic Structural Dynamics*; McGraw-Hill: New York, NY, USA, 1995.
12. Bolotin, V. *Random Vibrations of Elastic Systems (in Russian)*; Nauka: Moskva, Russia, 1979.
13. de Boor, C. *Practical Guide to Splines*; Springer: New York, NY, USA, 1987.
14. Malat, S. Multiresolution approximation and wavelets. *Trans. Am. Math. Soc.* **1989**, *315*, 59–88.
15. Huang, N.E.; Shen, Z.; Long, S.R.; Wu, M.C.; Shih, H.H.; Zheng, Q.; Yen, N.C.; Tung, C.C.; Liu, H.H. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. Lond. A* **1998**, *454*, 903–995. [CrossRef]
16. Zembaty, Z.; Krenk, S. Response spectra of spatial ground motion. In Proceedings of the 10th European Conference on Earthquake Engineering, Vienna, Austria, 28 August–2 September 1994; Flesch, R., Ed.; Balkema: Rotterdam, The Netherlands, 1994; pp. 1271–1275.
17. Fischer, C. Decomposition of the seismic excitation. In Proceedings of the EURODYN'99 Conference, Prague, Czech Republic, 7–10 June 1999; Frýba, L., Náprstek, J., Eds.; Balkema: Rotterdam, The Netherlands, 1999; pp. 1111–1116.
18. Náprstek, J.; Fischer, O. Comparison of classical and stochastic solution to seismic response of structures. In Proceedings of the 11th European Conference on Earthquake Engineering, Paris, France, 6–11 September 1998; CD ROM; Bisch, P., Ed.; Balkema: Rotterdam, The Netherlands, 1998; p. 13.
19. Náprstek, J. Non-stationary response of structures excited by random seismic processes with time variable frequency content. *Soil Dyn. Earthq. Eng.* **2002**, *22*, 1143–1150. [CrossRef]
20. European Committee for Standardization (Ed.) *Eurocode 8. Design Provisions for Earthquake Resistance of Structures*; CEN: Brussels, Belgium, 1996.

21. Tondl, A.; Ruijgrok, T.; Verhulst, F.; Nabergoj, R. *Autoparametric Resonance in Mechanical Systems*; Cambridge University Press: Cambridge, UK, 2000.
22. Náprstek, J. Non-linear auto-parametric vibrations in civil engineering systems. In *Trends in Civil and Structural Engineering Computing*; Topping, B., Neves, L.C., Barros, R., Eds.; Saxe-Coburg Publications: Stirlingshire, UK, 2009; Chapter 14, pp. 293–317.
23. Pugachev, V.; Sinitsyn, I. *Stochastic Differential Systems—Analysis and Filtering*; J. Willey: Chichester, UK, 1987.
24. Arnold, L. *Random Dynamical Systems*; Springer: Berlin/Heidelberg, Germany, 1998.
25. Gichman, I.; Skorochod, A. *Stochasticheskije Differencialnyje Uravnenija (in Russian)*; Naukova dumka: Kiev, Ukraine, 1968.
26. Grasman, J.; Van Herwaarden, O. *Asymptotic Methods for the Fokker–Planck Equation and the exit Problem in Applications*; Springer: Berlin/Heidelberg, Germany, 1999.
27. Náprstek, J. Some properties and applications of eigen functions of the Fokker–Planck operator. In Proceedings of the Engineering Mechanics 2005, Svratka, Czech Republic, 9–12 May 2005; Fuis, V., Krejčí, P., Návrat, T., Eds.; FME TU: Brno, Czech Republic, 2005; p. 12.
28. Weinstein, E.; Benaroya, H. The van Kampen expansion for the Fokker–Planck equation of a Duffing oscillator. *J. Stat. Phys.* **1994**, *77*, 667–679. [CrossRef]
29. Spencer, B.; Bergman, L. On the numerical solution of the Fokker–Planck equation for nonlinear stochastic systems. *Nonlinear Dyn.* **1993**, *4*, 357–372. [CrossRef]
30. Masud, A.; Bergman, L. Application of multi-scale finite element methods to the solution of the Fokker–Planck equation. *Comput. Methods Appl. Mech. Eng.* **2005**, *194*, 1513–1526. [CrossRef]
31. Náprstek, J.; Král, R. Finite element method analysis of Fokker–Planck equation in stationary and evolutionary versions. *Adv. Eng. Softw.* **2014**, *72*, 28–38. [CrossRef]
32. Náprstek, J.; Král, R. Theoretical background and implementation of the finite element method for multi-dimensional Fokker–Planck equation analysis. *Adv. Eng. Softw.* **2017**, *113*, 54–75.
33. Jaynes, E. Information theory and statistical mechanics. *Phys. Rev.* **1957**, *106*, 620–630. [CrossRef]
34. Kullback, S. *Information Theory and Statistics*; Chapman and Hall: New York, NY, USA, 1959.
35. Sarlis, N.V.; Skordas, E.S.; Varotsos, P.A.; Nagao, T.; Kamogawa, M.; Tanaka, H.; Uyeda, S. Minimum of the order parameter fluctuations of seismicity before major earthquakes in Japan. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 13734–13738. [CrossRef] [PubMed]
36. Varotsos, P.A.; Sarlis, N.V.; Skordas, E.S.; Tanaka, H. A plausible explanation of the b-value in the Gutenberg–Richter law from first Principles. *Proc. Jpn. Acad. B* **2004**, *80*, 429–434. [CrossRef]
37. Burg, J. New concepts in power spectra estimation. Presented at the 40th Annual International SEG Meeting, New Orleans, LA, USA, 10 November 1970.
38. Stocks, N. Information transmission in parallel threshold arrays: Suprathreshold stochastic resonance. *Phys. Rev. E* **2001**, *63*, 041114. [CrossRef] [PubMed]
39. McDonnell, M.; Stock, N.; Pearce, C.; Abbott, D. *Stochastic Resonance: From Suprathreshold Stochastic Resonance to Stochastic Signal Quantization*; Cambridge University Press: Cambridge, NY, USA, 2008.
40. Cieslikiewicz, W. *Determination of the Probability Distribution of a Sea Surface Elevation via Maximum Entropy Method (in Polish)*; Rep. Institute of Hydro-Engineering, Polish Acad. Sci.: Gdansk, Poland, 1988.
41. Sargent, M., III; Scully, M.; Lamb, W., Jr. *Laser Physics*; Addison-Wesley: Reading, UK, 1974.
42. Dong, W.; Bao, A.; Shah, H. Use of maximum entropy principle in earthquake recurrence relationship. *Bull. Seismol. Soc. Am.* **1984**, *74*, 725–737.
43. Herzel, H.; Ebeling, W.; Schmitt, A. Entropies of biosequences: The role of repeats. *Phys. Rev. E* **1994**, *6*, 5061. [CrossRef]
44. Rosenblueth, E.; Karmeshu; Hong, H.P. Maximum entropy and discretization of probability distributions. *Probab. Eng. Mech.* **1987**, *2*, 58–63. [CrossRef]
45. Tagliani, A. Principle of Maximum Entropy and probability distributions: definition of applicability field. *Probab. Eng. Mech.* **1989**, *4*, 99–104. [CrossRef]
46. Sobczyk, K.; Trebicki, J. Maximum entropy principle in stochastic dynamics. *Probab. Eng. Mech.* **1990**, *5*, 102–110. [CrossRef]
47. Sobczyk, K.; Trebicki, J. Maximum entropy principle and nonlinear stochastic oscillators. *Phys. A Stat. Mech. Its Appl.* **1993**, *193*, 448–468. [CrossRef]

48. Náprstek, J. Principle of maximum entropy of probability density in stochastic mechanics of elastic systems. In Proceedings of the Dynamics of Machines '95, Prague, Czech Republic, 1–2 February 1995; Dobiáš, I., Ed.; IT ASCR: Prague, Czech Republic, 1995; pp. 115–122.

49. Náprstek, J. Application of the maximum entropy principle to the analysis of non-stationary response of SDOF/MDOF systems. In Proceedings of the 2nd European Non-Linear Oscillations Conference—EUROMECH, Prague, Czech Republic, 9–13 September 1996; Půst, L., Peterka, F., Eds.; IT ASCR: Prague, Czech Republic, 1996; pp. 305–308.

50. Gibbs, J. *Elementary Principles in Statistical Mechanics*; Dover Publications: New York, NY, USA, 1960.

51. Reif, F. *Statistical Physics—Berkeley Physics Course-Vol.5*; Mc Graw Hill: New York, NY, USA, 1965.

52. Michlin, S. *Variational Methods in Mathematical Physics (in Russian)*; Nauka: Moskva, Russia, 1970.

53. Náprstek, J.; Fischer, C. Semi-analytical stochastic analysis of the generalized van der Pol system. *App. Comput. Mech.* **2018**, *12*, 45–58. [CrossRef]

54. Zubarev, D. *Neravnovesnaja Statističeskaja Termodinamika (in Russian)*; Nauka: Moskva, Russia, 1971.

55. Kullback, S.; Leibler, R. On information and sufficiency. *Ann. Math. Stat.* **1951**, *22*, 79–86. [CrossRef]

56. Kloeden, P.; Platen, E. *Numerical Solution of Stochastic Differential Equations*; Springer: Berlin/Heidelberg, Germany, 1992.

57. Rosenberg, R. Normal modes of non-linear dual-mode systems. *J. Appl. Mech.* **1960**, *27*, 263–268. [CrossRef]

58. Vakakis, A.; Manevitch, L.; Mikhlin, Y.; Pilipchuk, V.; Zevin, A. *Normal Modes and Localization in Non-Linear Systems*; Wiley: New York, NY, USA, 1996.

59. Kerschen, G.; Peeters, M.; Golinval, J.; Vakakis, A. Non-linear normal modes, Part I: A useful framework for the structural dynamicist. *Mech. Syst. Signal Process.* **2009**, *23*, 170–194. [CrossRef]