



Linear and Nonlinear Fokker-Planck Equations

Till D. Frank
Department of Psychology and Department of Physics, University of Connecticut, Storrs, CT, USA

Article Outline

- Glossary
- Definition of the Subject
- Introduction
- Time-Dependent Solutions and First-Order Statistics
- Markov Property: Second-Order and Higher-Order Statistics
- Martingales
- Examples
- Summary and Future Directions
- Bibliography

Glossary

Linear Here, linear with respect to a probability density.

Nonlinear Here, nonlinear with respect to a probability density.

Markov process Process for which it is sufficient to have information about the presence in order to make best predictions about the future. Additional information about the past will not improve the predictions.

Martingale process \hat{Z} Process for which the future mean value $\langle Z(t + \Delta t) \rangle$ of a set of realizations $Z^{(i)}$ that is passing at presence t through a certain common state z is the state z : $z : \langle Z(t + \Delta t) \rangle_{Z(t)=z} = z$. Additional information about states z' visited at times t' prior to t is irrelevant.

Definition of the Subject

Let \hat{X} denote a stochastic process defined on the space Ω and the time interval $[t_0, \infty]$, where t_0

denotes the initial time of the process. We assume that the process \hat{X} can be described in terms of a random variable $X \in \Omega$. More precisely, let $X(t)$ denote the time-dependent evolution of the random variable X for $t \geq t_0$. Then, we assume that the process \hat{X} can be described in terms of the infinitely large set of realizations $X^{(i)}(t)$ of $X(t)$ with $i = 1, 2, \dots$. The realizations $i = 1, 2, \dots$ constitute a statistical ensemble. At every time t the probability density P of the process \hat{X} can be computed from the realizations $X^{(i)}(t)$, that is, from the ensemble by means of

$$P(x, t) = \langle \delta(x - X(t)) \rangle, \quad (1)$$

where $\langle \cdot \rangle$ denotes ensemble averaging and $\delta(\cdot)$ is the delta function. We assume that at time t_0 the process is distributed like u . That is, the function $u(x)$ describes the initial probability density of the random variable X and we have $P(x, t_0) = u(x)$. In general, the evolution of P depends on how the process is distributed at initial time t_0 . In order to emphasize this point, we will use in what follows the notation $P(x, t; u)$. That is, we interpret Eq. (1) as a conditional probability density with the constraint given by the initial distribution u :

$$P(x, t; u) = \langle \delta(x - X(t)) \rangle_{\langle \delta(x - X(t_0)) \rangle = u(x)}. \quad (2)$$

We may also say that we study a family of stochastic processes (Frank 2005b). Each family member has a label or name which is given by u . For example, consider three experiments in which the evolution of dust particles in the air is observed for Gaussian, Lévy, and Cauchy initial distributions, respectively. It is known that dust particles perform a so-called Brownian random walk. So we would distinguish the three members $\hat{X}_1, \hat{X}_2, \hat{X}_3$ of our family of Brownian walk processes by the names of their initial distributions: Gauss, Lévy, and Cauchy.

Let us consider a stochastic process \hat{X} whose evolution of its probability density P is defined by a partial differential equation of the form

$$\frac{\partial}{\partial t}P(x, t; u) = \left[-\frac{\partial}{\partial x}D_1(x, t) + \frac{\partial^2}{\partial x^2}D_2(x, t) \right]P(x, t; u), \tag{3}$$

where D_1 and D_2 are functions of state x and time t . The functions D_1 and D_2 are referred to as drift and diffusion coefficients and constitute the Fokker-Planck operator

$$L^0(x, t) = -\frac{\partial}{\partial x}D_1(x, t) + \frac{\partial^2}{\partial x^2}D_2(x, t). \tag{4}$$

The evolution equation (3) is linear with respect to P . In this sense, Eq. (3) is a linear partial differential equation. Irrespective of this feature, the coefficients D_1 and D_2 may depend in a highly nonlinear fashion on the state x . For example, we may have $D_1 = -x + x^3$.

For appropriately chosen coefficients D_1 and D_2 , Eq. (3) describes the probability density P of a Markov process. In this case, Eq. (3) is referred to as a Fokker-Planck equation. More precisely, if \hat{X} is a Markov diffusion process whose probability density P is defined by Eq. (3), then Eq. (3) is called a Fokker-Planck equation. Note that roughly speaking, a Markov *diffusion* process is a Markov process characterized by a partial differential operator that can be truncated after the second-order partial derivative (see section “[Kramers-Moyal Expansion](#)”). In order to distinguish between linear and nonlinear Fokker-Planck equations, we will use the phrase “linear Fokker-Planck equation” instead of “Fokker-Planck equation.”

Let us generalize Eq. (3) by assuming that the drift and diffusion coefficients depend on the probability density P . In this case, Eq. (3) becomes

$$\frac{\partial}{\partial t}P(x, t; u) = \left[-\frac{\partial}{\partial x}D_1(x, t, P(x, t; u)) + \frac{\partial^2}{\partial x^2}D_2(x, t, P(x, t; u)) \right]P(x, t; u). \tag{5}$$

Likewise, the operator (4) is generalized to

$$L(x, t, P(x, t; u)) = -\frac{\partial}{\partial x}D_1(x, t, P(x, t; u)) + \frac{\partial^2}{\partial x^2}D_2(x, t, P(x, t; u)). \tag{6}$$

Equation (5) is nonlinear with respect to $P(x, t; u)$. Since the structure of the differential

operator in the bracket of Eq. (5) is equivalent to the structure of the differential operator (4), evolution equations of the form (5) are frequently called nonlinear Fokker-Planck equations. In this context, it is important to realize that the phrase “nonlinear Fokker-Planck equation” does not necessarily imply that we are dealing with a Markov process. The phrase “nonlinear Fokker-Planck equation” simply means that we are dealing with a nonlinear partial differential equation involving a partial differential operator that exhibits the structure of a Fokker-Planck operator.

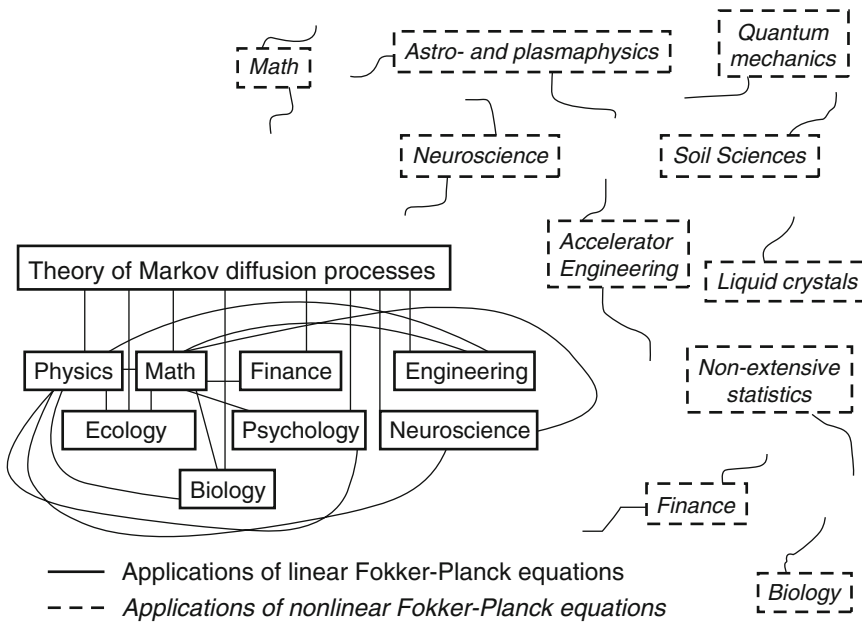
Note again that if an evolution equation of the form (3) is referred to as a Fokker-Planck equation, then it is tacitly assumed that there exists a stochastic process defined by that equation and that this process is a Markov process. Table 1 summarizes how to define linear and nonlinear Fokker-Planck equations by means of structure, existence of solutions, and Markov property.

Linear Fokker-Planck equations are an indispensable tool to describe stochastic processes in a variety of disciplines; see Fig. 1. The theoretical concept of Markov diffusion processes related to linear Fokker-Planck equations is well established. Researchers, applied scientist, technicians, research and development engineers in general, and financial engineers in particular are usually aware that the particular linear Fokker-Planck model they are using belongs to the class of Markov models. That is, the world of linear Fokker-Planck equations is closed and connected.

Nonlinear Fokker-Planck equations are used in a variety of fields that are as diverse as the

Linear and Nonlinear Fokker-Planck Equations, Table 1 Definition of linear and nonlinear Fokker-Planck equations based on structure, existence of solutions, and Markov property

Fokker-Planck equations	Linear	Nonlinear
Structure	Eqs. (3)–(4)	Eqs. (5)–(6)
Solutions	Exist	Do not necessarily exist
Corresponding processes are Markov processes?	Yes	Maybe



Linear and Nonlinear Fokker-Planck Equations, Fig. 1 Connected and disconnected applications of linear and nonlinear Fokker-Planck equations

application fields of linear Fokker-Planck equations. Unfortunately, so far, there is no well-established theory connecting all kinds of nonlinear Fokker-Planck equations. There is not even an academic consent about how to define them at all. This is why in Table 1 we used a very general and less constraining definition for nonlinear Fokker-Planck equations. Concepts of nonlinear Fokker-Planck equations are often developed for particular purposes and are not put into other contexts. That is, theoretical results and other achievements are often tailored to serve special needs and are not discussed in a larger framework. Even worse, so far, a well-established link between linear and nonlinear Fokker-Planck equations that applies to the variety of nonlinearities found in the literature does not exist. In sum, the world of nonlinear Fokker-Planck equations is disconnected. Different types of nonlinear Fokker-Planck equations and different application fields of nonlinear Fokker-Planck equations are often not related to each other and nonlinear Fokker-Planck equations are only loosely connected with their linear “relatives”; see Fig. 1.

Therefore, there is a need for developing a unifying approach to nonlinear Fokker-Planck equations that involves the concept of linear Fokker-Planck equations and applies to all

types of nonlinearities and in doing so applies to all kinds of scientific disciplines. Some first efforts in this regard have been made previously (Acebron et al. 2005; Chavanis 2003, 2004; Frank 2001b, 2002a, b, 2005b; Frank and Daffertshofer 1999, 2001a, b; Kaniadakis 2001a; b; Shiino 2002b, 2003). In the following sections, we will review these efforts, present them in a consistent way, and in doing so make a further effort into this direction.

Introduction

Linear and nonlinear Fokker-Planck equations are widely used to describe stochastic phenomenon; see Fig. 1.

Linear Fokker-Planck equations (Gardiner 1997; Haken 2004; Risken 1989) have been introduced by Fokker (1914) and Planck (1917). In physics, linear Fokker-Planck equations have been used, for example, to describe Brownian motion, that is, the diffusion of dust particles in air or fluid layers (Reif 1965). Linear Fokker-Planck equations have been applied in engineering sciences, for example, to describe fluctuations in electronic circuits (Gardiner 1997). Linear

Fokker-Planck equations have been frequently used in chemistry to model stochastic aspects of chemical reactions (Hänggi et al. 1990; van Kampen 1981). In finance, one of the most important applications of the Fokker-Planck theory is option pricing by means of the so-called Black-Scholes model (Paul and Baschnagel 1999). Linear Fokker-Planck equations of biology systems (Goel and Dyn 1974) have been concerned, for example, with so-called Brownian motors (Hänggi et al. 2005; Reimann 2002). Population diffusion (Okubo and Levin 2001) and group behavior (Ebeling and Sokolov 2004; Mikhailov and Zanette 1999; Schweitzer 2003) in ecological systems and stochastic neuronal processing (Holden 1976) are further examples of application fields of linear Fokker-Planck equations. In psychology, linear Fokker-Planck models have been proposed for decision making (Bogacz et al. 2006; Ratcliff et al. 2004) and group behavior (Schwämmle et al. 2007b).

Many applications of nonlinear Fokker-Planck equation are related to several benchmark models: the Desai-Zwanzig model (Desai and Zwanzig 1978), the liquid crystal model (Doi and Edwards 1988; Hess 1976), the Kuramoto-Shinomoto-Sakaguchi model (Acebron et al. 2005; Kuramoto 1984; Strogatz 2000), the Vlasov model, and the nonlinear diffusion equation (Aronson 1986; Peletier 1981). Let us highlight some of these benchmark models.

Desai-Zwanzig Model

The Desai-Zwanzig model

$$\frac{d}{dt}P(x, t; u) = -\frac{\partial}{\partial x}[h(x) - \kappa \left(x - \int xP(x, t; u) dx\right) + Q\frac{\partial}{\partial x}]P(x, t; u) \quad (7)$$

for $\kappa, Q > 0$ has been proposed by Desai and Zwanzig (1978) and Kometani and Shimizu (1975) to study collective phenomena in self-organizing systems.

- A Lyapunov functional approach to the Desai-Zwanzig model has been introduced by Shiino (1985, 1987) and since then has found several generalizations (Chavanis 2003, 2004;

Dawson and Gärtner 1989; Frank 2001a, 2005b; Frank and Daffertshofer 2001b; Frank et al. 2001; Kaniadakis 2001a, b; Kharchenko and Kharchenko 2005; Schwämmle et al. 2007a; Shiino 2001, 2002a, b; 2003). With such a Lyapunov functional at hand, the stability of stationary probability densities, collective phenomena, and bifurcations can be studied by means of Lyapunov's direct method.

- The original Desai-Zwanzig model and various modifications of it have been discussed (Dawson 1983; Li et al. 1998; Lo 2005).
- The additive noise term in Eq. (7) has been replaced by a multiplicative noise term (Horsthemke and Lefever 1984) in order to study the interplay between the nonlinearity and the multiplicative noise (Birner et al. 2002; Müller et al. 1997; Zaikin et al. 2002).
- Fluctuation-dissipation theorems for stochastic processes described by the Desai-Zwanzig model have been derived (Drozdov and Morillo 1996; Frank 2004c; Morillo et al. 1995).
- The Desai-Zwanzig model has frequently been used as a mean field approximation of spatially distributed systems with diffusive coupling. By means of such a mean field approximation, analytical result has been derived and compared with numerical simulations (Garcia-Ojalvo et al. 1996; Garcia-Ojalvo and Sancho 1999; van den Broeck et al. 1994a, b, 1997).

Liquid Crystal Model

The nonlinear Fokker-Planck equation proposed by Hess (Hess 1976) and Doi and Edwards (1988) reads

$$\frac{\partial}{\partial t}P(\mathbf{x}, t; u) = D_r \mathbf{L} \cdot \left\{ \mathbf{L} + \frac{1}{kT} [\mathbf{L} e(\mathbf{x}, P)] \right\} P(\mathbf{x}, t; u) \quad (8)$$

with $D_r, k, T > 0$ and $\mathbf{L} = \mathbf{x} \times \partial/\partial \mathbf{x}$. The function $e(\mathbf{x}, P)$ describes the self-consistent potential of the Maier-Saupe mean field force. For processes \hat{X} that exhibit cylindrical symmetry $e(\mathbf{x}, P)$ reads

$$e(\theta, P) = -U_0 kT \frac{3\cos^2\theta - 1}{2} \left\langle \frac{3\cos^2\theta - 1}{2} \right\rangle, \tag{9}$$

where θ is related to the unit vector \mathbf{x} by $\mathbf{x} = (\sin\theta \cos \varphi, \sin\theta \sin \varphi, \cos\theta)$. Equation (8) and generalization of it have been extensively studied in the literature (Felderhof 2003; Fialkowiski and Hess 2000; Hütter et al. 2003; Ilg et al. 1999, 2005; Larson and Öttinger 1991) (see also Öttinger (1996) in general and Sect. 6.3.2 in Öttinger (1996) in particular). We will return to this model in section “Liquid Crystal Model.”

Winfree and Kuramoto Model

Winfree’s seminal studies on synchronization among animal populations (Winfree 1967, 2001) supported the interest in the nonlinear Fokker-Planck equation

$$\frac{\partial}{\partial t} P(x, t; u) = \left\{ -\frac{\partial}{\partial x} [h(x) - \kappa \cdot \int \sin(x - y) P(y, t; u) dy] + Q \frac{\partial^2}{\partial x^2} \right\} P(x, t; u), \tag{10}$$

that has been proposed by Kuramoto and co-workers (Kuramoto 1984). In Eq. (10), $h(x)$ is a 2π -periodic function and $\kappa, Q > 0$.

- While the Kuramoto-Shinomoto-Sakaguchi model involves an interaction term $\int a(x, y) P(y, t) dy$, the model originally proposed by Winfree exhibits a coupling term of the form $\int a(y) P(y, t) dy$. Models of this latter kind have also been addressed in Ariaratnam and Strogatz (2001), Li and Hänggi (2001), and Quinn et al. (2007).
- The Kuramoto-Shinomoto-Sakaguchi model describes an ensemble of phase oscillators. The eigenfrequencies of the phase oscillators do not occur in Eq. (10) because Eq. (10) describes an ensemble of phase oscillators exhibiting the same eigenfrequency ω . In this case, the common eigenfrequency ω can then be eliminated by means of a variable transformation into a rotating frame (Frank 2005b). However, in general, we may think of ensembles of coupled phase oscillators with different eigenfrequencies. In this

context, the question arises as to what extent oscillators with different eigenfrequencies synchronize their behavior (Acebron et al. 1998; Arenas and Perez Vincente 1994; Bonilla et al. 1993; Crawford 1995; Kostur et al. 2002; Pikovsky et al. 2001; Sakaguchi 1988; Strogatz and Mirollo 1991).

- Coupled phase oscillator models of the form (10) have been used to describe associative memories (Yamana et al. 1999; Yoshioka and Shiino 2000).
- Just as for the Desai-Zwanzig model, the interplay between multiplicative noise and the nonlinearity of the Kuramoto-Shinomoto-Sakaguchi model has been investigated in several studies (Kim et al. 1997; Park and Kim 1996; Reimann et al. 1999a, b).
- The *sine*-coupling term in Eq. (10) has been replaced by higher-order coupling functions $\sin(2z), \sin(3z), \dots$ (Aonishi and Okada 2002; Daido 1996a, b; Hansel et al. 1993b; Kuramoto 1984). In this context, Daido proposed the so-called order function (Daido 1996a, b) that generalizes the notion of cluster phases and cluster amplitudes (Kuramoto 1984). This order function has also been related to experimental data (Zhai et al. 2005).
- The Kuramoto-Shinomoto-Sakaguchi has found clinical applications in the context of Parkinsonian disease (Tass 2001, 2003; 2006; see also Tass 1999).

Vlasov-Fokker-Planck Model

Vlasov-Fokker-Planck models frequently describe particle systems with electromagnetic interactions between charged particles. A typical example of a Vlasov-Fokker-Planck equation is shown here (Balescu 1975; Klimontovich 1986):

$$\frac{\partial}{\partial t} P(\mathbf{v}, t; u) = -\sum_{i=1}^3 \frac{\partial}{\partial v_i} D_i(\mathbf{v}, P) P + \sum_{i,k=1}^3 \frac{\partial^2}{\partial v_i \partial v_k} D_{ik}(\mathbf{v}, P) P. \tag{11}$$

Equation (11) involves the drift and diffusion coefficients

$$\begin{aligned}
 D_i(\mathbf{v}, P) &= a \frac{\partial}{\partial v_i} \int_{\Omega} \frac{P(\mathbf{v}', t; u)}{|\mathbf{v} - \mathbf{v}'|} d^3 v', \\
 D_{ik}(\mathbf{v}, P) &= b \frac{\partial^2}{\partial v_i \partial v_k} \int_{\Omega} |\mathbf{v} - \mathbf{v}'| P(\mathbf{v}', t; u) d^3 v'.
 \end{aligned}
 \tag{12}$$

Time-dependent solutions (Allen and Victory 1994; MacDonald et al. 1957; Nicholson 1983; Rosenbluth et al. 1957; Takai et al. 1981) for Eq. (12) and generalization of Eq. (12) that account for additional drift forces (Bychenkov et al. 1995; Epperlein et al. 1988) have been studied. In particular, numerical methods using short-time propagators (see section “Short-Time Propagator”) have been developed for Vlasov-Fokker-Planck equations of the form (12) (Donoso and Salgado 2006; Donoso et al. 2005; Soler et al. 1992). Such nonlinear Vlasov-Fokker-Planck equations play important roles in plasma physics (Balescu 1975; Klimontovich 1986; Nicholson 1983) and astrophysics (Binney and Tremaine 1987; Lancellotti and Kiessling 2001). In general, astrophysical problems often require a stochastic description in terms of nonlinear Fokker-Planck equations (Chavanis et al. 2002; Shiino 2003; Sopik et al. 2006). Finally, note that Vlasov-Fokker-Planck models have been used in accelerator physics and accelerator engineering to examine instabilities in particle beams (Frank 2003a, 2006; Heifets 2001, 2003; Shobuda and Hirata 2001; Stupakov et al. 1997; Venturini and Warnock 2002).

Nonlinear Diffusion Equation, Nonextensive Thermostatistics, and Semiclassical Descriptions of Quantum Systems

The nonlinear diffusion equation (Aronson 1986; Peletier 1981) reads

$$\begin{aligned}
 \frac{\partial}{\partial t} P(x, t; u) &= -\frac{\partial}{\partial x} h(x) P(x, t; u) \\
 &\quad + \frac{\partial^2}{\partial x^2} D(P(x, t; u)),
 \end{aligned}
 \tag{13}$$

where $D(P)$ is a diffusion coefficient that depends on the probability density $P(x, t)$ of $X(t)$. In the original version of the nonlinear diffusion

equation, the drift coefficient $h(x)$ vanishes, and the diffusion coefficient is proportional to a power of P . In general, there might be a more complicated dependence of D on P (Crank 1975; Daly and Porporato 2004).

- Since fluid flow through porous materials is an important application of the nonlinear diffusion equation, nonlinear diffusion plays a crucial role in soil sciences (Barenblatt et al. 1990). In biology, nonlinear diffusion equations of the form (13) seem to capture particular aspects of population diffusion (Gurtin and MacCamy 1977; Okubo and Levin 2001).
- The nonlinear diffusion Eq. (13) provides a link to stochastic processes subjected to non-extensive thermostatistics introduced by Tsallis (Abe and Okamoto 2001; Tsallis 1988, 1997, 2004). For $D(P) \propto P^q$, Eq. (13) becomes

$$\begin{aligned}
 \frac{\partial}{\partial t} P(x, t; u) &= -\frac{\partial}{\partial x} h(x) P(x, t; u) \\
 &\quad + Q \frac{\partial^2}{\partial x^2} P^q(x, t; u).
 \end{aligned}
 \tag{14}$$

Plastino and Plastino showed that stationary distributions of Eq. (14) correspond to canonical distribution that can be derived in a nonextensive framework (Plastino and Plastino 1995). Equation (14) has turned out to be a testbed for various analytical and numerical studies (Borland 1998; Chavanis 2003, 2004; Compte and Jou 1996; Drazer et al. 2000; Shiino 2003; Tsallis and Bukman 1996). Alternative nonlinear Fokker-Planck equations related to the Tsallis statistics have been derived from master equations in Curado and Nobre 2003 and Nobre et al. 2004. In addition, Eq. (14) has more recently discussed in finance in the context of a generalized Black-Scholes model for option pricing (Borland 2002, 2008; Borland and Bouchaud 2004; Vellekoop and Nieuwenhuis 2007) and fat tail distributions (Cortines and Riera 2007; Michael and Johnson 2003).

- The nonlinear diffusion Eq. (13) is also related to semiclassical descriptions of quantum

mechanical systems. For an appropriate choice of D , nonlinear Fokker-Planck equations for Fermi-Bose and Einstein-Dirac systems have been derived from Eq. (13) (Frank 2005b; Frank and Daffertshofer 1999). Alternative forms of nonlinear Fokker-Planck equations have been derived from quantum mechanical Boltzmann equations (Kaniadakis 2001a; Kaniadakis and Quarati 1993, 1994), on the basis of energy balance equations (Tsekov 1995, 2001), and by means of other techniques (Frank 2004b, 2005a; Kadanoff 2000; Sopik et al. 2006). We will return to semiclassical quantum mechanical descriptions in section “Semiclassical Description of Quantum Systems.”

In addition, nonlinear Fokker-Planck equations have been turned out to be useful models to describe stochastic aspects of Josephson arrays (Hadley et al. 1988; Wiesenfeld et al. 1996), Landau damping (Strogatz et al. 1992), arrays of semiconductor lasers (Kozyreff et al. 2000), charge density waves (Bonilla 1987), and neurons coupled by Hodgkin-Huxley equations (Han et al. 1995; Hansel et al. 1993a).

Stochastic systems composed of different kinds of interacting subsystems or species have been modeled in terms of multivariate nonlinear Fokker-Planck equations (Dano et al. 2001; Gang et al. 1996; Ichiki et al. 2007). For example, the collective behavior of coupled relaxation oscillators has been studied (Yamaguchi and Shimizu 1984). Networks of neural oscillators as defined by the Wilson-Cowan model, (Schuster and Wagner 1990), the two-dimensional Morris-Lecar system (Han et al. 1995), the FitzHugh-Nagumo equations (Hasegawa 2003; Kanamaru et al. 2001; Park et al. 2004), and the Hindmarsh-Rose equations (Rosenblum and Pikovsky 2004) have been studied.

The dynamics of mean field coupled phase oscillators under the impact of inertia effects (Acebron and Spigler 1998) and related models have attracted considerable attention. Bridge vibrations induced by pedestrian walking have been discussed in this context recently (Eckhardt et al. 2007; Strogatz et al. 2005). Models for circadian rhythms have been examined (Daido 2001).

Solutions of the Kadar-Parisi-Zhang equation have been examined by means of nonlinear Fokker-Planck equations (Giada and Marsili 2000; Marsili and Bray 1996). In doing so, the growth of surfaces and roughening phenomena have been studied.

Wetting processes (de los Santos et al. 2003), interacting Brownian motors (Becker and Engel 2007; Savel’ev et al. 2003), and spatially distributed phase oscillators (Kawamura 2007; Kawamura et al. 2007) have been analyzed by means of the nonlinear Fokker-Planck perspective.

In the mathematical literature, a seminal study on nonlinear Fokker-Planck equations of the Burgers equation type was due to McKean Jr. (1969). In particular, the convergence of stochastic processes described by multivariate linear Fokker-Planck equations to processes described by nonlinear Fokker-Planck equations (Cepa and Lepingle 1997; Dawson 1983, 1993; Ding 1994; Djehiche and Kaj 1995; Fontbona 2003; Gärtner 1988; Greven 2005; Jourdain 2000; McKean 1969; Meleard 1996; Meleard and Coppoletta 1987; Oelschläger 1989; Overbeck 1996; Pilipenko 2005; Rogers and Shi 1993) and martingales of stochastic processes defined by nonlinear Fokker-Planck equations have been addressed (Djehiche and Kaj 1995; Fontbona 2003; Gärtner 1988; Graham 1990; Greven 2005; Jourdain 2000; Meleard 1996; Meleard and Coppoletta 1987; Overbeck 1996). Moreover, the propagation of molecular chaos has been studied (Bonilla 1987; Meleard 1996; Meleard and Coppoletta 1987). The convergence of transient solutions of nonlinear Fokker-Planck equations to stationary ones has been examined by means of functionals that are similar to the Lyapunov functionals introduced by Shiino (see above) (Arnold et al. 1996, 2000, 2001; Carillo et al. 2001, 2008). In addition, from a purely mathematical perspective, nonlinear Fokker-Planck equations should be considered as nonlinear parabolic partial differential equations that have been discussed in several textbooks (Friedman 1969).

In what follows, we will show that there is a common theoretical framework that unifies most of the aforementioned studies on nonlinear Fokker-Planck equations and includes the theory of linear Fokker-Planck equations as a special case. This

common theoretical framework is rooted in the notion of Markov processes and martingales.

Time-Dependent Solutions and First-Order Statistics

Linear Case

Equation (3) defines the evolution of P given an initial distribution $u(x)$. The norm of the probability density P equals unity for all times provided that the norm of $u(x)$ equals unity. That is, if $\int_{\Omega} u(x) dx = 1$ holds, we have $\int_{\Omega} P(x, t; u) dx = 1$ for $t \geq t_0$. We can see this by integrating Eq. (3) with respect to x . For appropriate boundary conditions, it can be shown by partial integration that the right-hand side vanishes which implies that $d[\int_{\Omega} P(x, t; u) dx]/dt = 0$ holds. The formal solution of Eq. (3) reads

$$P(x, t; u) = \exp \left\{ \int_{t_0}^t dz L^0(x, z) \right\} u(x), \quad (15)$$

where L^0 is defined in Eq. (4). Equation (15) can be used to solve Eq. (3) numerically (see Vol. 1, Sect. 6.5 in Haken 2004). Let t_n denote a discrete time point $t_n = t_0 + n\Delta t$ with $n = 0, 1, 2, \dots$, where Δt is the interval of a single time step and should be small. Let us define $P_n(x; u) = P(x, t_n; u)$. Then, we have

$$P_{n+1}(x; u) = \{1 + \Delta t L^0(x, t_n)\} P_n(x; u) \quad (16)$$

with $P_0 = u(x)$. If \hat{X} corresponds to an autonomous processes, then the coefficient D_1 and D_2 do not depend on t . In this case, P_n can be expressed in terms of u as

$$P_n(x; u) = [1 + \Delta t L^0(x)]^n u(x). \quad (17)$$

Numerical solutions converge to exact solutions in the limit $\Delta t \rightarrow 0$.

Nonlinear Case

For appropriately chosen drift and diffusion coefficients, Eq. (5) exhibits time-dependent solutions P . By analogy with the linear case, these solutions are normalized to unity provided that appropriate

boundary conditions hold and that the initial probability density is normalized to unity. Solutions of Eq. (5) are formally defined by

$$P(x, t; u) = \exp \left\{ \int_{t_0}^t dz L(x, z, P(x, z; u)) \right\} u(x). \quad (18)$$

The time-dependent solutions P can be computed numerically by analogy to the linear case discussed above. That is, the probability densities $P_n(x; u) = P(x, t_n; u)$ on the discrete time grid $t_0, t_0 + \Delta t, t_0 + 2\Delta t, \dots$ can be computed from

$$P_{n+1}(x; u) = \{1 + \Delta t L(x, t_n, P_n(x; u))\} P_n(x; u) \quad (19)$$

with $P_0 = u(x)$ and $n = 0, 1, 2, \dots$. If drift and diffusion coefficients do not explicitly depend on t , we find that the operator L still depends implicitly on t because it depends on the time-dependent solution P that in turn depends on t . Consequently, it is not trivial to generalize Eq. (17) to the nonlinear case. If the drift- and diffusion coefficients do not depend explicitly on time t and the process converges to a stationary one, then the nonlinear Fokker-Planck operator L does not depend on time. This implies that the stationary probability density P_{st} satisfies

$$P_{st} = \{1 + \Delta t L(x, P_{st})\} P_{st} + O(\Delta t^2). \quad (20)$$

Finally, note that we do not necessarily need to define the formal solution with respect to the initial probability density u as in Eq. (18). We can solve the nonlinear Fokker-Planck equation on the time interval $[t_0, t]$ by splitting the solution in two intervals $[t_0, t']$ and $[t', t]$. Then, we obtain

$$P(x, t; u) = \exp \left\{ \int_{t'}^t dz L(x, z, P(x, z; u)) \right\} P(x, t'; u). \quad (21)$$

Equation (21) can be solved iteratively by means of Eq. (19) yielding a mapping $T_{\Delta t} : P(x, t; u) = T_{t-t'} [P(x, t'; u)]$ with $\Delta t = t - t'$.

Markov Property: Second-Order and Higher-Order Statistics

Conditional Probability Densities

Let $p(x)$ define the probability density of the time-dependent random variable X at time t given that X assumed at earlier times t', t'', t''', \dots with $t \geq t' > t'' > t''' > \dots$ particular values x', x'', x''', \dots . Then p is defined by

$$p = \langle \delta(x - X(t)) \rangle_{X(t')=x', X(t'')=x'', X(t''')=x''', \dots} \tag{22}$$

In order to point out the information that we need to compute p , we write

$$p = p(x, t | x', t'; x'', t''; x''', t''', \dots). \tag{23}$$

A conditional probability density is a relation that gives us estimates about future events and tells us what we need to know in order to be able to calculate these estimates. In our example given by Eq. (23), we see that we need the information of random values X at times $t' > t'' > t''' > \dots$ in order to make a prediction about the statistics or probability density of X at time t . Alternatively, we may say that the conditional probability density depends on a list of variables. In the context of Markov processes, this alternative viewpoint however gives rise to a problem that will be discussed below.

If \tilde{X} is a Markov process, then the information about the stochastic process available at one particular time t' is sufficient to make predictions about the future $t \geq t'$. Adding more information about how the process evolved in the past before t' does not improve these predictions. That is, the information about the events at time t' is sufficient to make statistical estimates about events at time $t \geq t'$. An alternative definition of a Markov process is that a Markov process exhibits a conditional probability density $p(x, t | \cdot)$ that depends only on one time point prior to t . That is, according to the first definition, we look from time t' into the future, whereas according to the second definition, we look in the opposite direction: we look from time t into the past.

For example, in order to describe the probability density $p(T)$ of the temperature T in Boston (USA) on December 1st, 2007, given that on November

1st, 2007, the temperature was 2°C and on October 1st, 2007, the temperature was 3°C , we would define the conditional probability density $p(T, \text{Dec 1st 2007} | T = 2, \text{Nov 1st 2007}; T = 3, \text{Oct 1st 2007})$. If the temperature T as a function of time t is a Markov process, it is sufficient to know the temperature at November 1st in order to compute the probability density $p(T)$ at December 1st. For example, we would obtain the same function $p(T)$ for the conditions (i) and (ii) with (i) $T = 2^\circ\text{C}$ on Nov 1st and $T = 3^\circ\text{C}$ on Oct 1st and (ii) $T = 2^\circ\text{C}$ on Nov 1st and $T = 5^\circ\text{C}$ on Oct 1st. That is, we would have

$$\begin{aligned} & p(T, \text{Dec 1st 2007} | T = 2, \text{Nov 1st 2007}; \\ & \qquad \qquad \qquad T = 3, \text{Oct 1st 2007}) \\ &= p(T, \text{Dec 1st 2007} | T = 2, \text{Nov 1st 2007}; \\ & \qquad \qquad \qquad T = 5, \text{Oct 1st 2007}) \\ &= p(T, \text{Dec 1st 2007} | T = 2, \text{Nov 1st 2007}). \end{aligned} \tag{24}$$

The information about the October temperature is irrelevant. In this sense, the conditional probability density would depend on the November temperature but would not depend on the October temperature.

A problem that arises in the context of the definition of Markov process is as follows. Suppose that there is a purely deterministic dynamical aspect involved in a stochastic process. In our example about Boston temperatures, we may think of the annual periodic changes of the temperature that are related to the annual changes in distance and declination angle between the earth and sun. Let us assume that distance and declination angle change periodically in a purely deterministic fashion such that the distance and declination angle at November 1st can be computed from the distance and declination angle at January 1st by a simple one-to-one mapping. Then the question arises: does the temperature in Boston on December 1st depend on the distance and declination angle of November 1st as suggested by $p(T, \text{Dec 1st 2007} | T = 25, \text{Nov 1st 2007})$ or does it depend on the distance and declination angle of January 1st. In the former case, we have a Markov conditional probability density. In the latter case, we would need to write p like

$$p(T, \text{Dec 1st 2007} | T = 25, \text{Nov 1st 2007}; \text{distance and angle, Jan 1st 2007})$$

indicated that we are dealing with a non-Markovian process. The situation becomes even worse if we take into consideration that the earth-sun distance and the declination angle at January 1, 2008, can be computed from the information known at November 1 using our simple one-to-one mapping. Therefore, we may say that the temperature estimate for December 1st, 2007, depends on a future event, namely, the earth-sun distance and declination angle given at January 1st, 2008. The conditional probability density would assume the form

$$p(T, \text{Dec 1st 2007} | T = 25, \text{Nov 1st 2007}; \text{distance and angle, Jan 1st 2008})$$

which would suggest again that we are dealing with a non-Markovian and – to a certain extent – noncausal process. We can solve this problem by realizing that purely deterministic relationships in time that represent external driving forces are irrelevant for the distinction between Markov and non-Markov processes. We can completely determine such an external driving force by a parameter set $\{t_0, A_1, A_2, \dots\}$ that describes the initial state of the driving forces. Although this initial state is related to the initial time t_0 of the Markov process, the conditional probability density does not actually depend (i.e., it does not explicitly depend) on t_0 . Likewise the conditional probability density does not actually depend on the parameters $\{A_1, A_2, \dots\}$. The information that we have at time t' includes the information about the driving force at time t' and therefore the information about the driving force at all times $t \in [t_0, \infty)$. Consequently, the information at time t' is sufficient to predict how the driving force will evolve in the future at times $t \geq t'$. There is no need to assess information about events prior to t' or information about events that will happen in the future at times larger than in order to determine the evolution of the deterministic driving force.

Let us summarize. A stochastic process \hat{X} is called a Markov process if information about the process at time t' is sufficient to make predictions about future events. This implies that the conditional probability density p defined in Eq. (23) can be simplified like

$$p(x, t | x', t', x'', t''; x''', t''', \dots) = p(x, t | x', t'). \tag{25}$$

Note that we may say that p depends only on the state x' related to the time t' in the sense that the information at time t' is sufficient to predict how X will be distributed at time $t \geq t'$. We may say it is sufficient to best predict future events where best refers to the fact that adding additional information about the past does not improve our predictions. Let us illustrate this issue by another example. Let $p(x, t | X = \theta)$ denote the probability density of X at time t given that X equals the function θ in the interval $[t_0, t']$ with $t' \leq t$. If X describes a Markov process, we have

$$p(x, t | X = \theta) = p(x, t | \theta(t'), t') = p(x, t | x', t') \tag{26}$$

with $x' = \theta(t')$.

Linear Fokker-Planck Equations

As mentioned in section “[Definition of the Subject](#),” linear Fokker-Planck equations describe Markov processes (Gardiner 1997; Risken 1989). Markov processes related to linear Fokker-Planck equations of the form (4) have conditional probability densities defined by

$$\frac{\partial}{\partial t} p(x, t | x', t') = L^0(x, t) p(x, t | x', t') \tag{27}$$

with $\lim_{t \rightarrow t'} p(x, t | x', t') = \delta(x - x')$. The conditional probability density p is also called the fundamental solution or Green’s function of the Fokker-Planck equation (3). In general, a stochastic process \hat{X} is completely defined in terms of the joint probability density

$$P(x_n, t_n; x_{n-1}, t_{n-1}; \dots; x_0, t_0) = \langle \delta(x_n - X(t_n)) \cdot \delta(x_{n-1} - X(t_{n-1})) \cdots \delta(x_0 - X(t_0)) \rangle, \tag{28}$$

where n can assume arbitrarily large integer numbers. In particular, if \hat{X} is a Markov process, then this joint probability density can be computed from p and u like

$$P(\cdot) = p(x_n, t_n | x_{n-1}, t_{n-1}) \cdot p(x_{n-1}, t_{n-1} | x_{n-2}, t_{n-2}) \cdots p(x_1, t_1 | x_0, t_0) u(x_0). \tag{29}$$

Consequently, the linear Fokker-Planck equation (3) defines completely a Markov process via

the associated evolution equation (27) and the initial distribution u . In particular, the time-dependent probability densities $P(x, t; u)$ and $P(x, t'; u)$ with $t \geq t'$ are related to each other by means of a *linear* functional

$$P(x, t; u) = \int_{\Omega} p(x, t | x', t') P(x', t'; u) dx'. \quad (30)$$

That is, the Green's function p induces a functional that is linear with respect to $P(x', t'; u)$.

Langevin Equations of Linear Fokker-Planck Equations

The stochastic trajectories $X(t)$ of the Markov process \hat{X} defined by Eq. (3) can be computed from the Ito-Langevin equation (Coffey et al. 2004; Gardiner 1997; Risken 1989)

$$\frac{d}{dt}X(t) = D_1(X(t), t) + \sqrt{D_2(X(t), t)}\Gamma(t), \quad (31)$$

where $\Gamma(t)$ denotes a Langevin force normalized to the delta function like $\langle \Gamma(t)\Gamma(t') \rangle = 2\delta(t - t')$. From the Langevin equation (31), it follows again that we are dealing with a Markov process. Information about one reference time t' is sufficient to compute the future behavior of the trajectory $X(t)$ with $t \geq t'$. On a discrete time grid the stochastic trajectories or realizations of \hat{X} can be computed iteratively like (Kloeden and Platen 1992; Risken 1989)

$$X_{n+1} = X_n + \Delta t D_1(X_n, t_n) + \sqrt{\Delta t D_2(X_n, t_n)}\epsilon_n \quad (32)$$

with $X(t_n) = X_n$, $t_n = t_0 + n\Delta t$ and $n = 0, 1, 2, \dots$. Here, ϵ_n are independent Gaussian distributed random numbers with vanishing mean and variance 2. That is, we have $\langle \epsilon_n \rangle = 0$ and $\langle \epsilon_n \epsilon_m \rangle = 2\delta_{nm}$, where δ_{nm} is the Kronecker symbol. Finally, the probability density W of ϵ_n at every step n is given by

$$W(\epsilon_n) = \frac{1}{\sqrt{4\pi}} \exp\left\{-\frac{\epsilon_n^2}{4}\right\}. \quad (33)$$

Note that we do not necessarily need to start the iteration iterative map at t_0 . The Scheme (32) can be started at any time step n . Moreover, in order to

compute the subsequent time steps, it is sufficient to have information about the random variable X at time t_n . Consequently, the sequence $X_n, X_{n+1}, X_{n+2}, \dots$ computed from Eq. (32) describes a trajectory of a Markov process.

Strongly Nonlinear Fokker-Planck Equations

In section “Definition of the Subject,” we pointed out that there is some kind of inconsistency in the definition of linear and nonlinear Fokker-Planck equations. While a linear Fokker-Planck equation defines a stochastic process, a nonlinear Fokker-Planck equation defines at best the evolution of a probability density $P(x, t)$. That is, if solutions of Eq. (5) exist for $u \in U$, then Eq. (5) defines the evolution of first-order statistical properties of a stochastic process \hat{X} such as the time-dependent probability density, the mean, and the variance of the process \hat{X} . In any case, Eq. (5) does not define second- and higher-order statistical quantities such as correlation functions and conditional probability densities (Frank 2004d). In particular, the time-dependent solutions P of Eq. (5) in general cannot be used to construct Green's functions of Markov processes because they do not necessarily correspond to Green's functions of Markov processes (Frank 2003b). Note that this is not a peculiarity of stochastic processes defined by nonlinear Fokker-Planck equations. In fact, time-dependent solutions P of linear Fokker-Planck equations Eq. (3) involving explicitly time-dependent coefficients D_1 and D_2 do not necessarily correspond to Green's functions. Mathematically speaking, let $P(x, t; u = \delta(x - x_0))$ denote the probability density of a process \hat{X} defined by a nonautonomous linear Fokker-Planck equation or by a nonlinear Fokker-Planck equation and let $p(x, t | x', t')$ denote the conditional probability density of that process \hat{X} , then we have (Frank 2003b)

$$P(x, t; u = \delta(x - x'))_{t_0=t'} \text{ is not necessarily equivalent to } p(x, t | x', t'), \quad (34)$$

where $P(x, t; u = \delta(x - x'))_{t_0=t'}$ means that we take the time-dependent solution $P(x, t; u = \delta(x - x_0))$ and replace in this expression x_0 by x' and t_0 by t' .

Let us return to the issue how to define a stochastic process \hat{X} on the basis of a nonlinear Fokker-Planck equation (5). In order to do so, we need to

define appropriate constraints such that out of all possible stochastic processes that exhibit a time-dependent probability density P defined by Eq. (5), one particular process is selected. In what follows, we will discuss one particular set of constraints (Frank 2004d). As we will see, the stochastic processes thus defined exhibit the Markov property.

Let U denote a set of initial probability densities u . That is, U is a set of functions or a space of functions. Let $P(x, t; u)$ denote the solution of the nonlinear Fokker-Planck equation

$$\frac{\partial}{\partial t}P(x, t, u) = L(x, t, P(x, t; u))P(x, t; u) \quad (35)$$

with

$$L(x, t, P(x, t; u)) = -\frac{\partial}{\partial x}D_1(x, t, P(x, t; u)) + \frac{\partial^2}{\partial x^2}D_2(x, t, P(x, t; u)) \quad (36)$$

for an initial distribution $u \in U$. Let us introduce the associated drift and diffusion coefficients \tilde{D}_1 and \tilde{D}_2 by

$$\begin{aligned} \tilde{D}_1(x, t; u) &= D_1(x, t; P(x, t; u)), \\ \tilde{D}_2(x, t; u) &= D_2(x, t; P(x, t; u)). \end{aligned} \quad (37)$$

That is, for any $u \in U$, Eq. (35) is solved analytically or by numerical iteration (19). The solution is substituted into the drift and diffusion coefficients D_1 and D_2 . The coefficients thus obtained are the functions \tilde{D}_1 and \tilde{D}_2 associated to D_1 and D_2 . Let us assume that for all $u \in U$, the evolution equation

$$\frac{\partial}{\partial t}p(x, t|x', t') = \left[-\frac{\partial}{\partial x}\tilde{D}_1(x, t; u) + -\frac{\partial^2}{\partial x^2}\tilde{D}_2(x, t; u) \right] \cdot p(x, t|x', t') \quad (38)$$

has a fundamental solution or Green's function. Then, this solution p and its corresponding initial distribution u define a Markov process. In (Frank 2004d) nonlinear Fokker-Planck equations that induce evolution equations (38) with fundamental solutions were called strongly nonlinear Fokker-Planck equations. Note that nonlinear Fokker-Planck equations (5) do not necessarily exhibit the property of being strongly nonlinear. Note

also that in some applications, it might be worth to define carefully the set U of initial probability densities u such that a nonlinear Fokker-Planck equation under consideration becomes strongly nonlinear.

As indicated above, the time-dependent probability density P of a nonlinear Fokker-Planck equation depends on the initial distribution u . Likewise, the associated coefficients \tilde{D}_1 and \tilde{D}_2 depend on u . As a result, the conditional probability density $p(x, t|x', t')$ depends on u as well. For this reason, the notation $p(x, t|x', t'; u)$ has been suggested. Unfortunately, this notation is likely to cause confusion because one might think that p depends not only on the time t' but also on the initial time t_0 which seems to be incompatible with the notion of a Markov conditional probability density (Frank 2007; McCauley et al. 2006). In fact, this confusion results from the second alternative way to define Markov processes that has been discussed above. The evolution of the function $P(x, t; u)$ is a purely deterministic one. That is, $P(x, t; u)$ represents a deterministic driving force for the purpose of computing the conditional probability density. The distribution u is just a parameter which determines the initial value of this driving force. In this context, note again that the conditional probability density of a Markov process in general depends on parameters and in particular can depend on the initial time t_0 and other parameters $\{A_1, A_2, \dots\}$ that define the initial state of a driving force. Consequently, the notation $p(x, t|x', t'; u)$ does not imply a contradiction with the notion of a Markov process. For example, the nonautonomous Langevin equation

$$\begin{aligned} \frac{\partial}{\partial t}X(t) &= -\gamma X(t) - A \cos(\omega(t - t_0)) \\ &+ \sqrt{Q}\Gamma(t) \end{aligned} \quad (39)$$

with $\gamma, A, \omega, Q > 0$ defines a Markov process that is driven by a harmonic force $-A \cos(\omega(t - t_0))$. That is, the harmonic force has amplitude A at the beginning of the process. The conditional probability density of that process depends on the parameter γ, ω, Q but also on the parameters t_0 and A which correspond to the initial amplitude and time. We have (see Sect. 3.7.3 in Frank 2005b)

$$p = p(x, t | x', t', \gamma, \omega, Q, A, t_0). \quad (40)$$

Nevertheless, in what follows, we will develop a slightly different notation for conditional probability densities p of strongly nonlinear Fokker-Planck equations that is in line with the first definition of Markov processes discussed above and will be helpful to elucidate that the functions p reflect indeed Markov processes.

Let us exploit first the fact that if Eq. (35) is a strongly nonlinear Fokker-Planck equations, then time-dependent solutions $P(x, t; u)$ of Eq. (35) exist for $u \in U$ and are related to their initial probability densities u by a one-to-one mapping T_t . That is, for every t we have $P(x, t; u) = T_t[u(x)]$. For an explicit construction of the map T_t , see, for example, Eq. (18). Likewise, we have $P(x, t'; u) = T_{t'}[u(x)]$. Using the inverse of T , we can map u to P like $u(x) = T_t^{-1}[P(x, t; u)]$. Substituting these expressions into $p(x, t | x', t'; u)$, we obtain $p(x, t | x', t'; T_t^{-1}[P(x, t; u)])$. This result demonstrates that the information about the stochastic process \hat{X} at time t' is sufficient to predict the future at $t > t'$. We can regard the conditional probability density p as a function that does not depend explicitly on u , but it depends explicitly on the state of the driving force P at time t' . In line with this remark, we introduce conditional probability densities of the form $p(x, t | x', t', P(x', t'; u))$.

Let us dwell on the interpretation of a conditional probability density $p(x, t | x', t', P(x', t'; u))$. To this end, we need to discuss briefly the notion of a particular conditional averaging that is important in this context and will become important later on as well. Let us assume that we make observations of realizations of a stochastic process \hat{X} for which the following two conditions hold: (i) $X(t') = x'$ and (ii) the ensemble of all realization is distributed like P at time t' . Next, we average across all observations that we make under these conditions. In doing so, we average under the constraints

$$\begin{aligned} X(t') = x' \text{ and } \langle \delta(x' - X(t')) \rangle \\ = P(x', t'; u). \end{aligned} \quad (41)$$

In order to indicate that such a structured constraint should hold, we will use the notation

$$\langle \cdot \rangle_{X(t')=x'; \langle \delta(x' - X(t')) \rangle = P(x', t'; u)}. \quad (42)$$

In words, Eq. (42) is the instruction to take out of an ensemble with probability density P at time t' only those realizations that assume the value x' at time t' . On the one hand, this constraint induces a trivial situation. We know for sure that $X(t') = x'$ and consequently can replace the random variable $X(t')$ by x' . On the other hand, the averaging procedure may involve the random variable $X(t)$ at a time point t different from t' . Although $X(t')$ is fixed at x' , the random variable $X(t)$ can assume different values at t for different realizations of the process \hat{X} . The conditional probability density $p(x, t | x', t', P(x', t'; u))$ is a special case in which the delta function is averaged under the constraint (41). We have

$$\begin{aligned} p(x, t | x', t', P(x', t'; u)) \\ = \langle \delta(x - X(t)) \rangle_{X(t')=x'; \langle \delta(x' - X(t')) \rangle = P(x', t'; u)}. \end{aligned} \quad (43)$$

Summarizing the results we have derived so far, we see that strongly nonlinear Fokker-Planck equations define Markov processes whose

- Time-dependent probability densities $P(x, t; u)$ are defined by Eq. (35)
- Conditional probability densities $p(x, t | x', t', P(x', t'; u))$ are defined by

$$\frac{\partial}{\partial t} p(x, t | x', t', P') = L(x, t, P) p(x, t | x', t', P') \quad (44)$$

with L given by Eq. (36), $P = P(x, t; u)$, and $P' = P(x', t'; u)$. Note that by multiplying Eq. (44) with $P(x', t'; u)$ and integrating with respect to x' , we get Eq. (35) which in turn defines the evolution of $P(x, t; u)$. Consequently, Eq. (44) defines both the evolution of $P(x, t; u)$ and $p(x, t | x', t', P')$. Note also that the solution of Eq. (44) formally reads

$$\begin{aligned}
 & p(x, t | x', t', P(x', t'; u)) \\
 &= \exp \left\{ \int_{t'}^t dz L(x, z, P(x, z; u)) \right\} \delta(x - x')
 \end{aligned}
 \tag{45}$$

and depends on the evolution of $P(x, z; u)$ for $z \in [t', t]$. In fact, as indicated above, p depends only on $P(x', t', u)$. To see this recall that the formal solution (21) can be obtained by means of the iterative method (19) such that we can write $P(x, z; u) = T_{z-t'} [P(x, t'; u)]$. Substituting this solution into Eq. (45), we get

$$\begin{aligned}
 & p(x, t | x', t', P(x', t'; u)) \\
 &= \exp \left\{ \int_{t'}^t dz L(x, z, T_{z-t'} [P(x, t'; u)]) \right\} \delta(x - x').
 \end{aligned}
 \tag{46}$$

In addition, we find that the solution (45) does not explicitly depend on u .

We arrive at the following conclusion: conditional probability densities $p(x, t | \cdot)$ of Markov processes described by strongly nonlinear Fokker-Planck equations depend only on the value of individual realizations at one prior time point $t' \leq t$ and on the probability density P defined by all realizations at the very same prior time point t' .

Equation (45) can be simplified for stationary Markov processes with operators L that do not depend explicitly on time t . Then the conditional probability density in the stationary case can be computed from

$$\begin{aligned}
 p(x, t | x', t', P_{\text{st}}(x')) &= \exp \{ (t - t') L(x, P_{\text{st}}(x)) \} \\
 &\delta(x - x').
 \end{aligned}
 \tag{47}$$

where $P_{\text{st}}(x)$ denotes a stationary probability density out of a set of stationary probability densities defined by $LP_{\text{st}} = 0$. Note that in this context, P_{st} plays the role of an initial distribution u .

Just as in the linear case, the conditional probability density p in combination with the initial distribution u completely defines the stochastic process \hat{X} . In particular, the joint probability

density $P(x_n, t_n; x_{n-1}, t_{n-1}; \dots x_0, t_0)$ can be computed from p and u like

$$\begin{aligned}
 & P(\cdot) = p(x_n, t_n | x_{n-1}, t_{n-1}, P_{n-1}) \\
 & \cdot p(x_{n-1}, t_{n-1} | x_{n-2}, t_{n-2}, P_{n-2}) \cdot \dots \\
 & \cdot p(x_1, t_1 | x_0, t_0, u) u(x_0),
 \end{aligned}
 \tag{48}$$

with $P_{n-1} = P(x_{n-1}, t_{n-1}; u)$, $P_{n-2} = P(x_{n-2}, t_{n-2}; u)$ and so on.

In particular, the time-dependent probability densities $P(x, t; u)$ and $P(x, t'; u)$ with $t \geq t'$ are related to each other by means of a *nonlinear* functional

$$P(x, t; u) = \int_{\Omega} p(x, t | x', t', P(x', t'; u)) P(x', t'; u) dx',
 \tag{49}$$

where p is defined by Eq. (46). That is, the Green's function p induces a functional that is nonlinear with respect to $P(x', t'; u)$.

Langevin Equations of Strongly Nonlinear Fokker-Planck Equations

The stochastic trajectories $X(t)$ of the Markov process \hat{X} defined by Eq. (44) can be computed from two-layered Langevin equations (see Sect. 3.4 in Frank 2005b) or alternatively from the self-consistent Ito-Langevin equation

$$\begin{aligned}
 \frac{d}{dt} X(t) &= D_1(X(t), t, P(X(t), t; u)) \\
 &+ \sqrt{D_2(X(t), t, P(X(t), t; u))} \Gamma(t),
 \end{aligned}
 \tag{50}$$

where $\Gamma(t)$ denotes the Langevin force introduced earlier. Note that the expression $P(X(t), t; u)$ means that the function $P(x, t; u)$ is evaluated at the state x that is given by the random variable X at time t . That is, we may write $P(X(t), t; u) = P(x, t; u)|_{x=X(t)}$. From the Langevin equation (50), we can read off that we are dealing with a Markov process. Information about one reference time t' in terms of the state $X(t')$ of a realization and the distribution of the ensemble as given by the probability density $P(x, t'; u)$ is sufficient to compute the future behavior of the trajectory $X(t)$ with $t \geq t'$.

The Langevin equation (50) may be implemented on a computer using the iterative map

$$X_{n+1} = X_n + \Delta t D_1(X_n, t_n, P(X_n, t_n; u)) + \sqrt{\Delta t D_2(X_n, t_n, P(X_n, t_n; u))} \epsilon_n \quad (51)$$

with $X(t_n) = X_n, t_n = t_0 + n\Delta t, n = 0, 1, 2, \dots$ and ϵ_n given as statistically independent Gaussian distributed random numbers with vanishing mean and variance 2 (see above). The expression $P(X_n, t; u)$ can be computed from the realizations generated by the iteration map (51). Let $X_n^{(i)}$ denote the i th realization at time step n . Then, the stochastic trajectories $X(t)$ can numerically be computed by simulating an ensemble of realizations $i = 1, \dots, N$ like

$$X_{n+1}^{(i)} = X_n^{(i)} + \Delta t D_1(X_n^{(i)}, t_n, P_n(X_n^{(i)})) + \sqrt{\Delta t D_2(X_n^{(i)}, t_n, P_n(X_n^{(i)}))} \epsilon_n^{(i)}, \quad (52)$$

where $\epsilon_n^{(i)}$ are statistically independent Gaussian random numbers with respect to both indices n and i and P_n is computed from the set $\{X_n^{(1)}, \dots, X_n^{(N)}\}$ of realizations using standard kernel estimators. For example, we may use

$$P_n(x) = \frac{1}{Ns\sqrt{2\pi}} \times \sum_{i=1}^N \exp\left\{-\frac{(x - X_n^{(i)})^2}{2s^2}\right\}, \quad (53)$$

with $s = N^{-1/5} \sigma_e(t_n)$ where $\sigma_e(t_n)$ is the standard deviation of the empirical ensemble $\{X_n^{(1)}, \dots, X_n^{(N)}\}$ (Frank 2005b, 2008; Silverman 1986). Just as in the case of Langevin equations of linear Fokker-Planck equations, the map (52) can be started at any time step n provided that we have information about P_n and X_n . In particular, if we start at a step $n > 0$, we see that the information

about the initial distribution is irrelevant. Consequently, the sequence $X_n, X_{n+1}, X_{n+2}, \dots$ computed from the time-discrete Langevin equation (52) related to the nonlinear Fokker-Planck equation (44) describes a trajectory of a Markov process.

Finally, note that self-consistent Langevin equations can be evaluated analytically in order to determine second-order statistical properties of a stochastic process defined by a strongly nonlinear Fokker-Planck equation (Borland 1998; Kharchenko and Kharchenko 2005).

Short-Time Propagator

The Green's function for short time intervals is frequently called the short-time propagator and can be derived from the time-discrete Ito-Langevin (51). Equation (51) relates the random variable ϵ_n that is distributed like $W(\epsilon_n)$ (see Eq. (33)) to the random variable X_{n+1} . In general, if X_{n+1} is a function of ϵ_n then the probability density $W'(x_{n+1})$ of X_{n+1} is given by

$$W'(x_{n+1}) = W(\epsilon_n) \frac{d\epsilon_n}{dx_{n+1}}. \quad (54)$$

In particular, if X_{n+1} is computed from ϵ_n for a particular value x_n and probability density P , then we obtain the short-time conditional probability density

$$p_s(x_{n+1} | x_n, P(x_n, t_n; u)) = W(\epsilon_n) \frac{d\epsilon_n}{dx_{n+1}}. \quad (55)$$

Equation (51) can be transformed into

$$\epsilon_n = \frac{X_{n+1} - X_n + \Delta t D_1(X_n, t_n, P(X_n, t_n; u))}{\sqrt{\Delta t D_2(X_n, t_n, P(X_n, t_n; u))}}. \quad (56)$$

Substituting Eq. (56) into Eq. (55), we obtain

$$p_s(x_{n+1} | x_n, P(x_n, t_n; u)) = \frac{\exp\left\{-\frac{[x_{n+1} - x_n + \Delta t D_1(x_n, t_n, P(x_n, t_n; u))]^2}{4\Delta t D_2(x_n, t_n, P(x_n, t_n; u))}\right\}}{\sqrt{4\pi\Delta t D_2(x_n, t_n, P(x_n, t_n; u))}}. \quad (57)$$

Using the time-continuous framework and the replacements $n \rightarrow t, x \rightarrow x, n + 1 \rightarrow t = t + \Delta t$ $\rightarrow x$ and likewise $P(x, t'; u) \rightarrow P(x', t'; u) = P'_n$ we obtain the short-time propagator (see Sect. 2.8.1 in Frank 2005b)

$$p_s(x, t | x', t', P') = \frac{\exp\left\{-\frac{[x - x' + \Delta t D_1(x', t', P')]^2}{4\Delta t D_2(x', t', P')}\right\}}{\sqrt{4\pi\Delta t D_2(x', t', P')}}. \tag{58}$$

The short-time propagator has originally been proposed by Wehner and Wolfer (1987) and can be used to solve nonlinear Fokker-Planck equations numerically (Donoso and Salgado 2006; Donoso et al. 2005; Soler et al. 1992). To this end, the short-time propagator is substituted into Eq. (48), and subsequently Eq. (48) is integrated over all variables x_{n-1}, \dots, x_0 . Thus we obtain $P(x, t_n; u)$ for $t_n = n\Delta t$. In the context of stochastic processes described by linear Fokker-Planck equations, the construction of solutions by means of short-time propagators is referred to as path integral approach (Gardiner 1997; Haken 2004; Risken 1989). We will return to a similar path integral approach in section “Semiclassical Description of Quantum Systems.” Expectation values of functions f can be computed from (58) like

$$\begin{aligned} \langle f(X(t)) \rangle_{X(t')=x'; \langle \delta(x - X(t')) \rangle = P(x', t'; u)} \\ = \int_{\omega} f(x) p_s(x, t | x', t', P') dx, \end{aligned} \tag{59}$$

which holds for small intervals $\Delta t = t - t'$. The short-time propagator illustrates again the Markov property of solutions of the strongly nonlinear Fokker-Planck equation (44). The information about x' and P' at time t' is sufficient to make predications in terms of expectation values that the stochastic process will assume at time $t = t' + \Delta t$.

Chapman-Kolmogorov Equation, Kramers-Moyal Expansion, and Drift-Diffusion Estimates

Linear Fokker-Planck equation can be derived using the Kramers-Moyal expansion of the

Chapman-Kolmogorov equation (Gardiner 1997; Risken 1989). The definition of the expansion coefficients in turn can be used to estimate the Kramers-Moyal coefficients in general and the drift and diffusion coefficients of linear Fokker-Planck equations in particular from experimental data (Friedrich and Peinke 1997; Friedrich et al. 2000). We will show in this section that if a stochastic process defined by a nonlinear Fokker-Planck equation can be embedded into a Markov process using the concept of strongly nonlinear Fokker-Planck equations, then we can proceed as in the linear case. Taking a slightly different perspective, we may say that there are Markov processes that involve conditional probability densities of the form $p(x, t | x', t', P(x', t'; u))$ and can be characterized in terms of generalized Kramers-Moyal expansion coefficients.

Chapman-Kolmogorov Equation

Let \tilde{X} denote a stochastic Markov process with conditional probability density $p(x, t | x', t', P(x', t'; u))$. Then as discussed in the previous section, the joint probability $P(x, t; x', t'; x'', t''); u)$ can be expressed by

$$P(x, t; x', t'; x'', t''); u) = p(x, t | x', t', P(x', t'; u)) \cdot p(x', t' | x'', t''); P(x'', t''); u) P(x'', t''); u). \tag{60}$$

Integrating with respect to x' and dividing by $P(x'', t''); u)$ yields the generalized Chapman-Kolmogorov equation

$$\begin{aligned} p(x, t | x'', t''); P(x'', t''); u) \\ = \int_{\Omega} p(x, t | x', t', P(x', t'; u)) \cdot p(x', t' | x'', t''); P(x'', t''); u) dx'. \end{aligned} \tag{61}$$

Note that in what follows, we will use the notation

$$\begin{aligned} P &= P(x, t; u), \\ P' &= P(x', t'; u), \\ P'' &= P(x'', t''); u). \end{aligned} \tag{62}$$

If we need to express probability densities P different from those listed in Eq. (62), we will write down if necessary their arguments explicitly.

For example, we will write $P(x, t'; u)$ to express the probability density $\langle \delta(x - X(t')) \rangle$ for a stochastic process \hat{X} with initial distribution u .

Using the notation of Eq. (62), we can write the joint probability density (60) like

$$P(x, t; x', t'; x'', t''; u) = p(x, t | x', t', P') p(x', t' | x'', t'', P'') \quad (63)$$

and the generalized Chapman-Kolmogorov Eq. (61) becomes

$$p(x, t | x'', t'', P'') = \int_{\Omega} p(x, t | x', t', P') p(x', t' | x'', t'', P'') dx'. \quad (64)$$

Kramers-Moyal Expansion

In this section, the Kramers-Moyal expansion for linear Fokker-Planck equations as discussed in Risken (1989) will be generalized to the nonlinear case. Consider the conditional probability density $p(x, t | x', t', P')$ for $t = t' + \Delta t$. Then, we have

$$p(x, t' + \Delta t | x', t', P') = \int_{\Omega} \delta(y - x) p(y, t' + \Delta t | x', t', P') dy. \quad (65)$$

The variables x and x' denote arbitrary states in Ω . However, let us consider next states y that are close to x' such that is $\epsilon = y - x'$ small. Using $y - x = \epsilon + x' - x$, we obtain

$$p(x, t' + \Delta t | x', t', P') = \int_{\Omega} \delta(x' - x + \epsilon) p(x' + \epsilon, t' + \Delta t | x', t', P') d\epsilon. \quad (66)$$

Use

$$\delta(x' - x + \epsilon) = \delta(x' - x) + \sum_1^{\infty} \frac{\epsilon^n}{n!} \left(\frac{\partial}{\partial x'} \right)^n \delta(x' - x). \quad (67)$$

Then, Eq. (66) becomes

$$p(x, t' + \Delta t | x', t', P') = \delta(x' - x) + \sum_1^{\infty} \int_{\Omega} d\epsilon \frac{\epsilon^n}{n!} p(x' + \epsilon, t' + \Delta t | x', t', P') \left(\frac{\partial}{\partial x'} \right)^n \delta(x' - x). \quad (68)$$

Multiplying Eq. (68) with $p(x'', t'' | x', t', P'')$ and integrating with respect to x' yields on the left-hand side

$$\begin{aligned} \text{LHS} &= \int_{\Omega} p(x, t' + \Delta t | x', t', P') p(x', t' | x'', t'', P'') \\ &dx' = p(x, t' + \Delta t | x'', t'', P'') \end{aligned} \quad (69)$$

and on the right-hand side

$$\begin{aligned} \text{RHS} &= p(x, t' | x'', t'', P'') \\ &+ \sum_1^{\infty} \int_{\Omega} dx' \int_{\Omega} d\epsilon \frac{\epsilon^n}{n!} p(x' + \epsilon, t' + \Delta t | x', t', P') \\ &\cdot p(x', t' | x'', t'', P'') \frac{\partial^n \delta(x' - x)}{\partial x'^n} \\ \text{RHS} &= p(x, t' | x'', t'', P'') \\ &+ \sum_1^{\infty} \int_{\Omega} dx' \delta(x' - x) \frac{\partial^n}{\partial x'^n} (-1)^n \\ &\cdot \int_{\Omega} d\epsilon \frac{\epsilon^n}{n!} p(x' + \epsilon, t' + \Delta t | x', t', P') p(x', t' | x'', t'', P'') \\ \text{RHS} &= p(x, t' | x'', t'', P'') \\ &+ \sum_1^{\infty} \left(-\frac{\partial}{\partial x} \right)^n \int_{\Omega} d\epsilon \frac{\epsilon^n}{n!} p(x + \epsilon, t' + \Delta t | x, t', P(x, t'; u)) \\ &\cdot p(x, t' | x'', t'', P''). \end{aligned} \quad (70)$$

Note that we used the Chapman-Kolmogorov Eq. (61) in order to evaluate the left-hand side (69) and we used Eq. (61) as well as partial integration in order to evaluate the right-hand side (70). Let us define the moments $M_n(x, t, \Delta t, P(x, t; u))$ by

$$M_n(x, t, \Delta t, P) = \int_{\Omega} d\epsilon \frac{\epsilon^n}{n!} p(x + \epsilon, t + \Delta t | x, t, P) \quad (71)$$

or using $\epsilon + x = z$ by

$$M_n(x, t, \Delta t, P) = \int_{\Omega} dz \frac{(z - x)^n}{n!} p(z, t + \Delta t | x, t, P). \quad (72)$$

Combining the left- and right-hand sides given by Eqs. (69) and (70), respectively, we obtain

$$\begin{aligned} &p(x, t' + \Delta t | x'', t'', P'') \\ &= p(x, t' | x'', t'', P'') + \sum_1^{\infty} \left(-\frac{\partial}{\partial x} \right)^n \\ &\cdot M_n(x, t', \Delta t, P(x, t'; u)) p(x, t' | x'', t'', P''). \end{aligned} \quad (73)$$

To improve readability, let us replace t' by t and subsequently t'' by t' . Thus, we obtain

$$\begin{aligned}
 p(x, t + \Delta t | x', t', P') &= p(x, t | x', t', P') \\
 &+ \sum_1^{\infty} \left(-\frac{\partial}{\partial x} \right)^n \\
 &\cdot M_n(x, t, \Delta t, P) p(x, t | x', t', P').
 \end{aligned}
 \tag{74}$$

This is the time-discrete version of the Kramers-Moyal expansion of the generalized Chapman-Kolmogorov Eq. (64). Note that M_n depends on $P(x, t; u)$, whereas p depends on $P(x', t'; u)$. Next, we define the Kramers-Moyal coefficients

$$\begin{aligned}
 D_n(x, t, P) &= \lim_{\Delta t \rightarrow 0} \frac{M_n}{\Delta t} \lim_{\Delta t \rightarrow 0} \\
 &\times \frac{1}{\Delta t} \int_{\Omega} dx \frac{(z-x)^n}{n!} \\
 &\cdot p(z, t + \Delta t | x, t, P).
 \end{aligned}
 \tag{75}$$

Dividing Eq. (74) by Δt and taking the limiting case $\Delta t \rightarrow 0$, Eq. (74) becomes the time-continuous generalized Kramers-Moyal expansion

$$\begin{aligned}
 \frac{\partial}{\partial t} p(x, t | x', t', P) &= \sum_1^{\infty} \left(-\frac{\partial}{\partial x} \right)^n D_n(x, t, P) \\
 &p(x, t | x', t', P').
 \end{aligned}
 \tag{76}$$

Note that by generalizing the Kramers-Moyal expansion to the nonlinear case, we found immediately that the coefficients D_n depend on $P(x, t; u)$, whereas the conditional probability density p depends on $P(x', t'; u)$. Note also that in the special case $D_n = 0$ for $n \geq 3$, the Kramers-Moyal expansion (76) yields the nonlinear Fokker-Planck equation (44). Finally note that since we have $M_n(\Delta t = 0) = 0$ for all n , Kramers-Moyal coefficients can also be defined by

$$\begin{aligned}
 D_n(x, t, P) &= \left. \frac{\partial M_n}{\partial \Delta t} \right|_{\Delta t=0} \\
 &= \int_{\Omega} dx \frac{(z-x)^n}{n!} \left. \frac{\partial}{\partial u} p(z, u | x, t, P) \right|_{u=t}.
 \end{aligned}
 \tag{77}$$

Drift-Diffusion Estimates

The definition of the Kramers-Moyal coefficients can be exploited to extract the drift and diffusion coefficients of nonlinear Fokker-Planck equations

from time series data. Accordingly, the drift coefficient D_1 and the diffusion coefficient D_2 are defined by

$$\begin{aligned}
 D_1(x, t, P) &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int_{\Omega} dx (z-x) p(z, t + \Delta t | x, t, P), \\
 D_2(x, t, P) &= \lim_{\Delta t \rightarrow 0} \frac{1}{2\Delta t} \int_{\Omega} dx \frac{(z-x)^2}{2} p(z, t + \Delta t | x, t, P).
 \end{aligned}
 \tag{78}$$

The limiting case Δt may be approximated by the smallest time step that is available in the data set:

$$\begin{aligned}
 D_1(x, t, P) &\approx \frac{1}{\Delta t} \int_{\Omega} dx (z-x) p(z, t + \Delta t | x, t, P), \\
 D_2(x, t, P) &\approx \frac{1}{2\Delta t} \int_{\Omega} dx \frac{(z-x)^2}{2} p(z, t + \Delta t | x, t, P),
 \end{aligned}
 \tag{79}$$

where Δt denotes now the sampling interval between two data points. Note that on the basis of the alternative definition (77), higher-order approximations can also be defined (Patanarapeelert et al. 2006). The conditional averages can be approximated by empirical conditional averages computed from a finite set of realizations $X^{(1)}(t), X^{(2)}(t), \dots, X^{(N)}(t)$. Thus, we obtain

$$\begin{aligned}
 D_1(x, t, P) &\approx \frac{1}{\Delta t} \frac{1}{\sum_{i \in I(t,x)} 1} \cdot \sum_{i \in I(t,x)} [X^{(i)}(t + \Delta t) - X^{(i)}(t)], \\
 D_2(x, t, P) &\approx \frac{1}{2\Delta t} \frac{1}{\sum_{i \in I(t,x)} 1} \cdot \sum_{i \in I(t,x)} [X^{(i)}(t + \Delta t) - X^{(i)}(t)]^2,
 \end{aligned}
 \tag{80}$$

where $I(t,x)$ is the set of indices i for which $X^{(i)}(t) \approx x$. In the case of Markov processes described by linear Fokker-Planck equations, the argument P in the coefficients can be dropped and the above drift-diffusion estimates reduce to the estimates proposed in Friedrich and Peinke (1997) and Friedrich et al. (2000) that have recently found many applications (Bödeker et al. 2003; Jafari et al. 2002; Sura and Barsugli 2002; Waechter et al. 2004). For Markov processes described by nonlinear Fokker-Planck equations, we need to compute the conditional averages for different probability densities P . To this end, we may vary the initial distribution u of a stochastic process. For a stochastic process with a particular distribution of X at time t , we will obtain the coefficients D_1 and D_2 only for that particular distribution. Using the kernel estimate method mentioned above, we obtain

$$\begin{aligned}
 D_1 \left(x, t, P \approx \frac{1}{Ns\sqrt{2\pi}} \sum_{i=1}^N \exp \left\{ -\frac{(x - X^{(i)}(t))^2}{2s^2} \right\} \right) \\
 \approx \frac{1}{\Delta t} \frac{1}{\sum_{i \in I(t,x)} 1} \\
 \times \sum_{i \in I(t,x)} \left[X^{(i)}(t + \Delta t) - X^{(i)}(t) \right]
 \end{aligned} \tag{81}$$

and

$$\begin{aligned}
 D_2 \left(x, t, P \approx \frac{1}{Ns\sqrt{2\pi}} \sum_{i=1}^N \exp \left\{ -\frac{(x - X^{(i)}(t))^2}{2s^2} \right\} \right) \\
 \approx \frac{1}{2\Delta t} \frac{1}{\sum_{i \in I(t,x)} 1} \\
 \times \sum_{i \in I(t,x)} \left[X^{(i)}(t + \Delta t) - X^{(i)}(t) \right]^2
 \end{aligned} \tag{82}$$

with $s = N^{-1/5} \sigma_e(t_n)$, where $\sigma_e(t_n)$ is the standard deviation of the empirical ensemble $\{X^{(1)}(t), \dots, X^{(N)}(t)\}$. Note that in general the Kramers-Moyal coefficients of Markov processes induced by conditional probability densities of the form $p(x, t|x', t', P')$ can be estimated using

$$\begin{aligned}
 D_n \left(x, t, P \approx \frac{1}{Ns\sqrt{2\pi}} \sum_{i=1}^N \exp \left\{ -\frac{(x - X^{(i)}(t))^2}{2s^2} \right\} \right) \\
 \approx \frac{1}{n! \Delta t} \frac{1}{\sum_{i \in I(t,x)} 1} \\
 \times \sum_{i \in I(t,x)} \left[X^{(i)}(t + \Delta t) - X^{(i)}(t) \right]^n.
 \end{aligned} \tag{83}$$

Alternatively, parametric estimate methods may be used. For example, we may be interested in estimating the exponent q of a Markov process defined by the Plastino-Plastino model (see section “Nonextensive Systems” below)

$$\begin{aligned}
 \frac{\partial}{\partial t} p(x, t|x', t', P') = \left[\frac{\partial}{\partial x} \gamma x + Q \frac{\partial^2}{\partial x^2} P(x, t; u)^{q-1} \right] \\
 p(x, t|x', t', P')
 \end{aligned} \tag{84}$$

with $\gamma, Q, q > 0$. Then, the diffusion coefficient $D_2(P) = QP^{q-1}$ involves the parameter Q and q . Using Eq. (82) and taking the logarithm, we get

$$\begin{aligned}
 \ln Q + (q - 1) \ln \left\{ \frac{1}{Ns\sqrt{2\pi}} \sum_{i=1}^N \exp \left\{ -\frac{(x - X^{(i)}(t))^2}{2s^2} \right\} \right\} \\
 \approx \ln \left\{ \frac{1}{2\Delta t} \frac{1}{\sum_{i \in I(t,x)} 1} \right. \\
 \left. \sum_{i \in I(t,x)} \left[X^{(i)}(t + \Delta t) - X^{(i)}(t) \right]^2 \right\}.
 \end{aligned} \tag{85}$$

For example, at a particular time t , Eq. (85) can be evaluated for different states x_i . In that case, Eq. (85) assumes the form $\ln Q + (q - 1) A_1(x_i) = A_2(x_i)$. Then, the expressions $\ln Q$ and $q - 1$ (and in doing so the parameters Q and q) can be estimated from a linear regression (Frank and Friedrich 2005).

Martingales

Let $Z(t)$ denote a functional of a stochastic process \hat{X} defined for $t \geq t_0$. In what follows, we will put $t_0 = 0$. Then, Z is a martingale of \hat{X} if

$$\langle Z(t) \rangle_{X=\theta} = Z(t') \tag{86}$$

holds for $t \geq t'$, where θ is a realization of the random variable X on the interval $[0, t']$ (see Sect. 1.3 in Karlin and Taylor 1975). That is, the constraint $X = \theta$ means $X(s) = \theta(s)$ holds for $s \in [0, t']$. That is, in the interval $[0, t']$ the trajectory X is fixed. Roughly speaking, a martingale is a random variable for which the best predictor of its future mean value is the present value. With regard to Eq. (86), the prediction of the future mean value is $\langle Z(t) \rangle$, whereas the present value of Z is $Z(t')$. Alternatively, we may say that the information at one time t' about the value of the martingale Z is sufficient to predict the mean value of the martingale Z for future times $t \geq t'$. Note that this alternative point of view is closely related with the first definition of Markov processes discussed in the previous section.

For linear Fokker-Planck equations, there is a close link between martingales and the Markov property. Accordingly, a stochastic process is a Markov process defined by a linear Fokker-Planck equation with drift and diffusion coefficients D_1

and D_2 if and only if a particular random variable Z that involves the Fokker-Planck operator is a martingale (see Sect. 15.1 in Karlin and Taylor 1981). In the mathematical literature, this link has also been studied in the context of nonlinear Fokker-Planck equations (Djehiche and Kaj 1995; Fontbona 2003; Gärtner 1988; Graham 1990; Greven 2005; Jourdain 2000; Meleard 1996; Meleard and Coppoletta 1987; Overbeck 1996).

Our aim in this section is to make the martingale approach more accessible to scientists working in physics, applied mathematicians, and related disciplines. To this end, we will in what follows illustrate this link between martingales and Markov processes defined by strongly nonlinear Fokker-Planck equation by means of standard techniques frequently used in physics.

Theorem 1 Let \hat{X} be a stochastic process with initial probability density $u(x)$ and conditional probability density

$$p(x, t | x', t'; P') = \langle \delta(x - X(t)) \rangle_{X(t')=x'; \langle \delta(x' - X(t')) \rangle = p(x', t'; u)}. \quad (87)$$

Then, \hat{X} is a Markov process defined by the nonlinear Fokker-Planck equation

$$\frac{\partial}{\partial t} p(x, t | x', t'; P') = L(x, t, P) p(x, t | x', t'; P') \quad (88)$$

with

$$L(x, t, P) = - \frac{\partial}{\partial x} D_1(x, t, P) + \frac{\partial^2}{\partial x^2} D_2(x, t, P) \quad (89)$$

if and only if $Z(t)$ defined by

$$Z(t) = f(X(t)) - \int_0^t L_B f[X(z), z, P] dz \quad (90)$$

with

$$L_B(x, t, P) = D_1(x, t, P) \frac{\partial}{\partial x} + D_2(x, t, P) \frac{\partial^2}{\partial x^2} \quad (91)$$

is a martingale of X for smooth functions f . In the context of linear Fokker-Planck equations, the

operator L_B is the Fokker-Planck backwards operator (Gardiner 1997; Risken 1989). Note that in our context, we refer to f as a smooth function if it has continuous second-order derivatives. Note also that above and in what follows, we will frequently use the notation (54). Note finally that in the above theorem the notion $L_B f[X(z), z, P]$ should be interpreted like

$$L_B f[X(z), z, P] = L_B(X(z), z, P) f(X(z)) = \{L_B(x, t, P) f(x)\}_{x=X(z), t=z}. \quad (92)$$

That is, first we carry out the differentiations defined by the operator L_B . Subsequently, we replace in the result the state variable x by the value of the random variable X at time z . Moreover we replace t by z . Let us prove the theorem in two parts.

From Strongly Nonlinear Fokker-Planck Equations to Martingales

Let us prove in this section that a Markov process defined by a strongly nonlinear Fokker-Planck equation exhibits the martingale Z . To this end, we first compute the conditional mean of the random variable Z defined in Eq. (90). Thus, we obtain

$$\begin{aligned} \langle Z(t) \rangle_{X=0} &= \langle f(X(t)) \rangle_{X=0} - \int_0^t dz \langle L_B f[X(z), z, P] \rangle_{X=0} \\ &= \langle f(X(t)) \rangle_{X=0} - \int_0^t dz \langle L_B f[X(z), z, P] \rangle_{X=0} \\ &\quad - \int_0^t ds L_B f[\theta(s), s]. \end{aligned} \quad (93)$$

The Markov property implies that the constraints can be relaxed. That is, for every functional $g(t)$ of $X(t)$ with $t \geq t'$, we have $\langle g(t) \rangle_{X=0} = \langle g(t) \rangle_{X(t')=x'; P'}$ where x' is given by $x' = \theta(t')$. We have indicated here that the average may depend on how the process is distributed at time t' . Consequently, Eq. (93) becomes

$$\begin{aligned} \langle Z(t) \rangle_{X=0} &= \langle f(X(t)) \rangle_{X(t')=x'; \langle \delta(x' - X(t')) \rangle = P'} \\ &\quad - \int_0^t dz \langle L_B f[X(z), z, P] \rangle_{X(t')=x'; \langle \delta(x' - X(t')) \rangle = P'} \\ &\quad - \int_0^t ds L_B f[\theta(s), s, P]. \end{aligned} \quad (94)$$

Multiplying the Fokker-Planck equation (88) with $f(x)$ and integrating with respect to x , we obtain

$$\begin{aligned} & \frac{\partial}{\partial t} \int_{\Omega} f(x)p(x, t|x', t'; P') dx \\ &= \int_{\Omega} f(x)L(x, t, P)p(x, t|x', t'; P') dx. \end{aligned} \tag{95}$$

By means of partial integration, we find that $\int_{\Omega} f(x)L(x, t, P)p(x, t|x', t'; P') dx = \int_{\Omega} p(x, t|x', t'; P')L_B(x, t, P)f(x) dx$. As a result, Eq. (95) can be transformed into

$$\begin{aligned} & \frac{\partial}{\partial t} \int_{\Omega} f(x)p(x, t|x', t', P') dx \\ &= \int_{\Omega} p(x, t|x', t', P')L_B(x, t, P)f(x) dx \\ &= \langle L_B(x, t, P)f(x) \rangle_{X(t')=x'; P'}. \end{aligned} \tag{96}$$

Using Eq. (96), we obtain

$$\begin{aligned} & \langle L_Bf[X(z), z] \rangle_{X(t')=x'; \langle \delta(x-X(t')) \rangle = P'} \\ &= \frac{\partial}{\partial z} \int_{\Omega} f(x)p(x, t|x', t', P') dx. \end{aligned} \tag{97}$$

Consequently, the following integral transformation holds

$$\begin{aligned} I &= \int_{t'}^t dz \langle L_Bf[X(z), z, P] \rangle_{X(t')=x'; P'} \\ &= \int_{t'}^t dz \frac{\partial}{\partial z} \int_{\Omega} dx f(x)p(x, z|x', t', P') \\ &= \langle f(X(t)) \rangle_{X(t')=x'; P'} - f(x'). \end{aligned} \tag{98}$$

Substituting Eq. (98) into Eq. (94), we get

$$\langle Z(t) \rangle_{X=\theta} = f(x') - \int_0^{t'} ds L_Bf[\theta(s), s, P]. \tag{99}$$

By definition, the function $Z(t')$ for $X(s) = \theta(s)$ given in $s \in [0, t']$ reads

$$Z(t') = f(x') - \int_0^{t'} ds L_Bf[\theta(s), s, P]. \tag{100}$$

Consequently, we have our final result

$$\langle Z(t) \rangle_{X=\theta} = Z(t') \tag{101}$$

and the proof is completed.

From Martingales to Strongly Nonlinear Fokker-Planck Equations

Let us prove next that the martingale (90) defines a Markov process of a strongly nonlinear Fokker-Planck equation. Evaluating Eq. (90) by analogy to Eq. (93) gives us

$$\begin{aligned} \langle Z(t) \rangle_{X=\theta} &= \langle f(X(t)) \rangle_{X=\theta} \\ &- \int_{t'}^t dz \langle L_Bf[X(z), z, P] \rangle_{X=\theta} \\ &- \int_0^{t'} ds L_Bf[\theta(s), s, P]. \end{aligned} \tag{102}$$

Substituting this result into Eq. (86) and substituting Eq. (100) into Eq. (86), we see that Eq. (86) becomes

$$\begin{aligned} f(X(t)) \rangle_{X=\theta} &= f(x') \\ &+ \int_{t'}^t dz \langle L_Bf[X(z), z] \rangle_{X=\theta}. \end{aligned} \tag{103}$$

Equation (103) can equivalently be written as

$$\begin{aligned} & \int_{\Omega} f(x)p(x, t|X = \theta) dx \\ &= f(x') + \int_{t'}^t dz \int_{\Omega} dx p(x, z|X = \theta)L_Bf[x, t, P] \end{aligned} \tag{104}$$

with $p(x, z|X = \theta) = \delta(x - X(z))_{X=\theta}$. Using partial integration, we can show that the operator L_B and the differential operator L are related to each other like

$$\begin{aligned} & \int_{\Omega} dx p(x, z|X = \theta)L_Bf[x, t, P] \\ &= \int_{\Omega} dx f(x)L(x, z, P)p(x, z|X = \theta). \end{aligned} \tag{105}$$

Substituting this result into Eq. (104) yields

$$\begin{aligned} 0 &= \int_{\Omega} dx f(x) \{ p(x, t|X = \theta) - \delta(x - x') - \\ & \int_{t'}^t dz L(x, z, P)p(x, z|X = \theta) \}. \end{aligned} \tag{106}$$

This holds for arbitrary smooth functions f . Since f is arbitrary, the expression in the brackets $\{\cdot\}$ of Eq. (106) must vanish, and we obtain

$$p(x, t | X = \theta) = \delta(x - x') + \int_{t'}^t dz L(x, z, P) p(x, z | X = \theta). \tag{107}$$

Differentiating Eq. 107 with respect to t gives us

$$\frac{\partial}{\partial t} p(x, t | X = \theta) = L(x, t, P) p(x, t | X = \theta). \tag{108}$$

Multiplying with the probability density $P(X = \theta)$ and performing a functional integration with respect to the path θ , we obtain

$$\frac{\partial}{\partial t} P(x, t) = L(x, t, P) P(x, t). \tag{109}$$

The formal solutions of Eqs. (108) and (109) read

$$p(x, t | X = \theta) = \exp\left\{ \int_{t'}^t dz L(x, z, P) \right\} \delta(x - x') \tag{110}$$

and

$$P(x, t) = \exp\left\{ \int_{t'}^t dz L(x, z, P) \right\} P(x, t'). \tag{111}$$

We see that a solution of Eq. (108) under the initial condition $p(x, t | X = \theta) = \delta(x - x')$ for $t \rightarrow t'$ with $x' = \theta(t')$ only depends on $\theta(t')$ but does not depend on $\theta(s)$ for $s < t'$. Consequently, X is a Markov process. However, L depends on P . From Eq. (110), it is clear that the conditional probability density p depends on the time-dependent probability density P for $z \in [t', t]$. Since the probability density $P(x, t; u)$ for $t \geq t'$ can be computed from $P(x, t'; u)$ as shown in Eq. (111), we conclude that p depends only on $P(x, t'; u)$ and does not depend on the evolution of P on the whole interval $[t, t']$. Therefore, we have $p(x, t | X = \theta) = p(x, t | x', t', P')$. Substituting this result into Eq. (108), we see that Eq. (108) becomes a strongly nonlinear Fokker-Planck equation

$$\frac{\partial}{\partial t} p(x, t | x', t'; P') = L(x, t, P) p(x, t | x', t'; P'). \tag{112}$$

Examples

Shimizu-Yamada Model

The Shimizu-Yamada model (Shimizu 1974; Shimizu and Yamada 1972) corresponds to the Desai-

Zwanzig model (7) for a linear single-particle force $h(x) = -\gamma x$. The evolution of the conditional probability density p is defined by

$$\frac{\partial}{\partial t} p(x, t | x', t', P') = \left[\frac{\partial}{\partial x} \gamma x + \kappa \left(x - \int_{\Omega} x P(x, t; u) dx \right) + Q \frac{\partial^2}{\partial x^2} \right] \cdot p(x, t | x', t', P') \tag{113}$$

with $\Omega = \mathbb{R}$ and $\gamma, \kappa, Q > 0$. Multiplying Eq. (113) with $P(x', t'; u)$ and integrating with respect to x' yields the evolution equation for $P(x, t; u)$:

$$\frac{\partial}{\partial t} P(x, t; u) = \left[\frac{\partial}{\partial x} \gamma x + \kappa \left(x - \int_{\Omega} x P(x, t; u) dx \right) + Q \frac{\partial^2}{\partial x^2} \right] \cdot P(x, t; u). \tag{114}$$

See also Frank (2004d) and Sect. 3.10 in Frank (2005b). From Eq. (114), it follows that the mean value $m(t) = \int_{\Omega} x P(x, t; u) dx$ decays exponentially like

$$m(t) = m(t_0) \exp\{-\gamma(t - t_0)\} \tag{115}$$

with $m(t_0) = \int_{\Omega} x u(x) dx$. Substituting Eq. (115) into Eqs. (113) and (114), we realize that a solution $P(x, t; u)$ and a Green's function p exists for any initial probability density $u(x)$. Therefore, the Shimizu-Yamada model is a strongly nonlinear Fokker-Planck equation and describes a Markov process.

It can be shown that the conditional probability density $p(x, t | x', t'; u)$ reads (see Frank (2004d) and Sect. 3.10 in Frank (2005b))

$$p(x, t | x', t'; u) = \frac{\exp\left\{ -\frac{[x - g(t, t', t_0, u) - x' m(t, t')]^2}{2K(t, t')} \right\}}{\sqrt{2\pi K(t, t')}} \tag{116}$$

with

$$m(t, t') = \exp\{-(\gamma + \kappa)(t - t')\}, \tag{117}$$

$$K(t, t') = \frac{Q}{\gamma + \kappa} [1 + \exp\{-2(\gamma + \kappa)(t - t')\}], \tag{118}$$

and

$$g = [\exp\{-\gamma(t - t_0)\} - \exp\{-(\gamma + \kappa)t + \gamma t_0 + \kappa t'\}] \cdot \int_{\Omega} xu(x) dx. \tag{119}$$

The mean value $m(t)$ acts as a self-organized driving force of the stochastic process. Since there is a one-to-one mapping of $m(t)$ to $m(t')$ with $t' < t$, we can eliminate the parameter u in $p(x, t|x', t'; u)$ as argued in section “[Strongly Nonlinear Fokker-Planck Equations.](#)” Substituting Eq. (115) into Eq. (119), we obtain

$$g(t, t', P(x, t'; u)) = \exp\{-\gamma(t - t')\} \cdot [1 - \exp\{-\kappa(t - t')\}] \int_{\Omega} xP(x, t'; u) dx \tag{120}$$

or

$$g(t, t', \langle X(t') \rangle) \exp\{-\gamma(t - t')\} \cdot [1 - \exp\{-\kappa(t - t')\}] \langle X(t') \rangle. \tag{121}$$

Consequently, the conditional probability density $p(x, t|x', t', P')$ reads

$$p(x, t|x', t'; P') = \frac{\exp\left\{-\frac{[x - g(t, t', \langle X(t') \rangle) - x'm(t, t')]^2}{2K(t, t')}\right\}}{\sqrt{2\pi K(t, t')}}. \tag{122}$$

Dynamic Takatsuji Model

The dynamic Takatsuji model for the conditional probability density p is defined by

$$\frac{\partial}{\partial t} p(x, t|x', t'; P') = \left[\frac{\partial}{\partial x} (\gamma + c)x - \sqrt{c} \tanh \left(\sqrt{c} \int_{\Omega} xP(x, t; u) dx \right) + Q \frac{\partial^2}{\partial x^2} \right] \cdot p(x, t|x', t'; P') \tag{123}$$

with $x \in \Omega = \mathbb{R}$ and $c, Q > 0, \gamma \in \mathbb{R}$. Likewise, the probability density $P(x, t; u)$ satisfies

$$\frac{\partial}{\partial t} P(x, t; u) = \left[\frac{\partial}{\partial x} (\gamma + c)x - \sqrt{c} \tanh \left(\sqrt{c} \int_{\Omega} xP(x, t; u) dx \right) + Q \frac{\partial^2}{\partial x^2} \right] \cdot P(x, t; u). \tag{124}$$

For details, see Frank (2004e) and Takatsuji (1975). From Eq. (124), it follows that the first moment $M_1(t) = X$ can be computed from

$$\frac{d}{dt} M_1(t) = -(\gamma + c)M_1 + \sqrt{c} \tanh[\sqrt{c}M_1(t)]. \tag{125}$$

For arbitrary initial distribution u , solutions of $M_1(t)$ exist and are smooth functions of t . Substituting these solutions into Eqs. (123) and (124), we see that solutions of Eqs. (123) and (124) in terms of Green’s functions p and probability densities P exist as well. Consequently, the dynamic Takatsuji model belongs to the class of strongly nonlinear Fokker-Planck equations and describes a Markov process.

Since $p(x, t|x', t', P')$ depends on P' , the expected mean value of $X(t)$ of realizations that assume the value x' at time t' depends on the distribution of the ensemble at time t' . Let us illustrate this issue. The conditional mean value under consideration reads

$$\langle X(t) \rangle_{X(t')=x', P'} = \int xp(x, t|x', t', P') dx. \tag{126}$$

Multiplying Eq. (123) with x and integrating with respect to x , we obtain

$$\frac{d}{dt} \langle X(t) \rangle_{X(t')=x', P'} = -(\gamma + c) \langle X(t) \rangle_{X(t')=x', P'} + \sqrt{c} \tanh[\sqrt{c}M_1(t)]. \tag{127}$$

The solution reads

$$\langle X(t) \rangle_{X(t')=x', P'} = x' \exp\{-(\gamma + c)(t - t')\} + \sqrt{c} \int_{t'}^t \tanh[\sqrt{c}M_1(z)] dz, \tag{128}$$

where $M_1(z)$ is the solution of Eq. (125) for the initial value $M_1(t') = \int_{\Omega} xP(x, t'; u) dx$. Let I denote the integral $I = \sqrt{c} \int_{t'}^t \tanh[\sqrt{c}M_1(z)] dz$. Then, I depends on $M_1(t')$, c , γ , t' and t : $I = I(t, t', M_1(t'), c, \gamma)$. Consequently, Eq. (128) can be cast into the form

$$\langle X(t) \rangle_{X(t')=x', P'} = x' \exp\{-(\gamma + c)(t - t')\} + I \left(t, t', \int x'P(x', t'; u) dx', c, \gamma \right) \tag{129}$$

or

$$\langle X(t) \rangle_{X(t')=x', P'} = x' \exp\{-(\gamma + c)(t - t')\} + I(t, t', \langle X(t') \rangle, c, \gamma). \quad (130)$$

Equation (130) illustrates that in order to predict future conditional mean values of a Takatsuji process \hat{X} at times t , it is sufficient to have at one time $t' \leq t$ information about the state value x' of a realization of \hat{X} and the mean value $X(t')$ of all realizations of \hat{X} .

Note that the trajectories $Z(t)$ of martingale processes \hat{Z} induced by the Takatsuji process \hat{X} are given by

$$Z(t) = f(X) - \int_0^t ds \{ [-(\gamma + c)X(s) + \sqrt{c} \tanh(\sqrt{c} \langle X(s) \rangle)] \frac{\partial f}{\partial X(s)} + Q \frac{\partial^2 f}{\partial X^2(s)} \} \quad (131)$$

for arbitrary smooth functions f . We can exploit these martingale processes in order to compute conditional expectations. For example, for $f(y) = y$ from the martingale property (86), it follows that

$$\langle X(t) \rangle_{X(t')=x', P'} = x' - \int_{t'}^t ds (\gamma + c) \langle X(s) \rangle - \sqrt{c} \tanh[\sqrt{c} \langle X(s) \rangle]. \quad (132)$$

Differentiating this relation with respect to t , we obtain Eq. (127) again and so we can compute the conditional expectation (130).

Liquid Crystal Model

Liquid crystals exhibit nematic-isotropic phase transitions (Chandrasekhar 1977; de Gennes and Prost 1993; de Jeu 1980). At high temperatures, the liquid crystal macromolecules exhibit an orientational disorder. The liquid crystal is said to be in the isotropic phase. Below a critical temperature, the macromolecules show some degree of orientational order. The degree of orientational order is often measured by the Maier-Saupe order parameter S (Maier and Saupe 1958).

A nonlinear Fokker-Planck equation that describes the stochastic behavior of the liquid crystal in the isotropic and nematic phases and to a certain extent also describes the order-disorder phase transition was proposed by Doi and Edwards (Doi and Edwards 1988) and Hess (Hess 1976) and is shown above in Eq. (8). Equation (8) describes the random walk of the orientation of liquid crystal molecules, where the orientation is given by a vector \mathbf{x} that points to the surface of a unit sphere. If we are dealing with rod-like molecules, then the orientation corresponds to the primary axis of the molecules along the rod. In particular, for liquid crystals with an axial symmetry, the liquid crystal model can be simplified. The simplified model describes the random walk of the molecule alignment with the symmetry axis. The random variable is defined on $X \in \Omega = [0, 1]$. For sake of simplicity, we will extend the range of definition to the interval $\Omega = [-1, 1]$ and require that distributions are symmetric. For $X = 0$, the molecule has an orientation perpendicular to the symmetry axis. If $X = 1$ or $X = -1$, the molecule points exactly in the direction of the symmetry axis. In this symmetric case, the probability density P of X satisfies (Felderhof 2003)

$$\frac{\partial}{\partial t} P(x, t; u) = \frac{\partial}{\partial x} (1 - x^2) \cdot \left[-\frac{9}{2} \kappa x \left(\int x^2 P(x, t; u) dx - \frac{1}{3} \right) + D_r \frac{\partial}{\partial x} \right] P(x, t; u) \quad (133)$$

with $\kappa, D_r > 0$. Equation (133) as well as the original Eq. (8) are regarded as descriptions for an ensemble of macromolecules that perform rotational Brownian motion (Doi and Edwards 1988). Since Brownian motion is a Markov process, it is reasonable to construct on the basis of Eq. (133) a model for a many-body system that exhibits a Markov process. In line with our discussion in section “[Markov Property: Second-Order and Higher-Order Statistics](#),” we assume that the conditional probability density p satisfies (Frank 2005c)

$$\begin{aligned} \frac{\partial}{\partial t} p(x, t | x', t', P') &= \frac{\partial}{\partial x} (1 - x^2) \\ &\cdot \left[-\frac{9}{2} \kappa x \left(\int x^2 P(x, t; u) dx - \frac{1}{3} \right) + D_r \frac{\partial}{\partial x} \right] \\ &\cdot p(x, t, | x', t', P'). \end{aligned} \tag{134}$$

Note that the expression in the bracket (\cdot) is related to the Maier-Saupe order parameter which reads in the symmetric case

$$S(t) = \frac{1}{2} \left(3 \int x^2 P(x, t; u) dx - 1 \right). \tag{135}$$

Due to the boundary conditions $X \in [-1, 1]$, the order parameter S and consequently the bracket (\cdot) is bounded. This implies that solutions P and p of Eqs. (133) and (134) exist and that the liquid crystal model (133)–(134) describes a Markov process. The self-consistent Ito-Langevin equation of this Markov process reads (Frank 2005c)

$$\begin{aligned} \frac{d}{dt} X(t) &= \frac{9\kappa}{2} (1 - X(t)^2) X(t) \left(\langle X(t)^2 \rangle - \frac{1}{3} \right) \\ &\quad - 2D_r X(t) \\ &\quad + \sqrt{D_r (1 - X(t)^2)} \Gamma(t). \end{aligned} \tag{136}$$

Trajectories $Z(t)$ of martingale processes \hat{Z} of the liquid crystal model are defined by (Frank 2007)

$$Z(t) = f(X(t)) - \int_0^t ds [\cdot] f(X(s)) \tag{137}$$

with

$$\begin{aligned} [\cdot] &= \left[\left(\frac{9\kappa}{2} (1 - X(s)^2) X(s) \left(\langle [X(s)]^2 \rangle - \frac{1}{3} \right) \right. \right. \\ &\quad \left. \left. - 2D_r X(s) \right) \frac{\partial}{\partial X(s)} + D_r \frac{\partial^2}{\partial X^2(s)} \right]. \end{aligned} \tag{138}$$

In the stationary case, the short-time autocorrelation function $C(\Delta t) = X(t)X(t + \Delta t)_{st}$ reads (Frank 2005c)

$$\begin{aligned} C(\Delta t) &= \frac{2S + 1}{3} - \frac{2D_r(1 - S)}{3} \Delta t \\ &\quad + O(\Delta t^2), \end{aligned} \tag{139}$$

where S denotes the order parameter (see above) in the stationary case. That is, we have $S = (3 X_{st}^2 - 1)/2$. Consequently, C depends on S . This has important implications for the hysteresis loop of the nematic-isotropic phase transition. Let us assume that if we decrease the temperature of a liquid crystal, we find the transition from the isotropic to the nematic phase with $S = 0 \rightarrow S > 0$ at the critical temperature $T_{c,low}$. In contrast, if we increase the temperature of a liquid crystal, we find the transition from the nematic to the isotropic phase with $S > 0 \rightarrow S = 0$ at the slightly higher critical temperature $T_{c,high}$. Then, in the temperature interval $[T_{c,low}, T_{c,high}]$, the liquid crystal exhibits two autocorrelation functions

$$C_{isotrope}(\Delta t) = \frac{1}{3} - \frac{2D_r}{3} \Delta t, \tag{140}$$

$$\begin{aligned} C_{nem}(\Delta t) &= \frac{2S(T) + 1}{3} \\ &\quad - \frac{2D_r(1 - S(T))}{3} \Delta t \end{aligned} \tag{141}$$

which hold up to terms of order Δt^2 . Equations (140)–(141) illustrate that we are dealing with a system that exhibits two kinds of Markov processes that we may label “isotropic” and “nematic,” respectively. The modeling approach by means of strongly nonlinear Fokker-Planck equations indicates that these Markov processes are just different members of a family of Markov processes that naturally emerge in the self-organized liquid crystal. That is, the two Markov processes are not related to two different systems but they represent two different “states” of the same self-organizing many-body system.

Let us compute the conditional mean value of molecules that are perpendicular to the symmetry axis. To this end, we consider the random walk of the orientation angle ϕ defined by $X(t) = \sin \phi(t)$. Using the Stratonovich-Langevin equation of Eq. (134) (see Frank 2005c), we obtain a self-consistent Langevin equation for ϕ :

$$\frac{d}{dt}\phi = \frac{9}{4}\kappa \sin(2\phi(t)) \left(\langle \sin^2(\phi(t)) \rangle - \frac{1}{3} \right) - D_r \tan \phi(t) + D_r \Gamma(t). \tag{142}$$

For short time intervals $\Delta t = t - t'$ and appropriate small noise amplitudes D_r , we assume that $\phi(t) \approx 0$ if $\phi(t') \approx 0$. Linearizing Eq. (142) at $\phi = 0$ yields

$$\frac{d}{dt}\phi(t) = \left\{ \frac{9}{4}\kappa \left(\langle \sin^2(\phi(t)) \rangle - \frac{1}{3} \right) - D_r \right\} \phi(t) + D_r \Gamma(t). \tag{143}$$

The conditional expectation value $\int \phi p(\phi, t | \phi', t') P d\phi$ for short time intervals Δt can then be computed from Eq. (143) by averaging both side of Eq. (143) under the constraint $\phi(t') = \phi'$ and $\phi(t')$ distributed like P' . Thus, we obtain

$$\langle \phi(t) \rangle_{\phi(t')=\phi', P'} = \phi' \left[1 + \Delta t \left\{ \frac{9}{4}\kappa \left(\langle \sin^2(\phi(t')) \rangle - \frac{1}{3} \right) - D_r \right\} \right] \tag{144}$$

or

$$\langle \phi(t) \rangle_{\phi(t')=\phi', P'} = \phi' \left[1 + \Delta t \left\{ \frac{3}{2}\kappa S(t') - D_r \right\} \right]. \tag{145}$$

These estimates hold for small intervals Δt , sufficiently small noise amplitudes D_r , and orientation angles $\phi' \approx 0$. Again, in line with our general discussion in the preceding sections, we see that the conditional expectation $\phi(t)_{\phi(t')=\phi', P'}$ can be computed provided that for $t' < t$ the distribution of $\phi(t')$ or at least the order parameter $S(t')$ is known and the angle ϕ' is selected.

Semiclassical Description of Quantum Systems

A stochastic treatment of semiclassical quantum systems by means of nonlinear Fokker-Planck equations that can be cast into the form of Eqs. (5) and (6) has been proposed and analyzed in several studies (Carrillo et al. 2008; Chavanis 2003; Frank and Daffertshofer 1999; Kadanoff

2000; Kaniadakis 2001a; Kaniadakis and Quarati 1993, 1994). Accordingly, a Fermi or Bose particle with mass 1 that moves in a one-dimensional space with velocity v exhibits in the stationary case a Fermi-Dirac or Bose-Einstein distribution of the kinetical energy $E_{\text{kin}} = v^2/2$. The free diffusion of the particle can be described by the nonlinear Fokker-Planck equations (Frank and Daffertshofer 1999)

$$\frac{\partial}{\partial t} P(v, t; u) = \frac{\partial}{\partial v} \gamma v [1 \mp P(v, t, u)] P(v, t; u) + Q \frac{\partial^2}{\partial v^2} P(v, t; u), \tag{146}$$

where the upper sign holds for Fermi particles, the lower for Bose particles. The parameters γ and Q represent damping and fluctuation strength and are related to the temperature T by the fluctuation dissipation theorem $Q/\gamma = 1/(k_B T)$, where k_B is the Boltzmann constant. The stationary probability density $P_{\text{st}}(v)$ of Eq. (146) reads

$$P_{\text{st}}(v) = \frac{1}{\exp\{(E_{\text{kin}} - \mu)/(k_B T)\} \pm 1}, \tag{147}$$

where μ is a normalization constant that can be interpreted as chemical potential. The transient solution $P(v, t; u)$ can be obtained by solving the integral equation (Frank 2007; Meleard and Coppoletta 1987)

$$P(v, t; u) = \int_{\Omega} dv_0 G_B(t, t_0, v, v_0) u(v_0, t_0) + \int_{t_0}^t ds \int_{\Omega} dv' G_B(t, s, v, v') \gamma \frac{\partial}{\partial v'} \times [1 \mp P(v', s; u)] P(v', s; u) \tag{148}$$

with

$$G_B(t, t', v, v') = \frac{1}{\sqrt{2\pi Q(t-t')}} \exp\left\{ -\frac{(v-v')^2}{2Q(t-t')} \right\}, \tag{149}$$

where G_B is the Gaussian propagator of Brownian motion. In the limit $t \rightarrow \infty$, the transient solution

$P(v, t; u)$ approaches $P_{st}(v)$ (Frank and Daffertshofer 2001b; Kaniadakis 2001a). Equation (148) is a useful description for numerical approaches. Using partial integration, Eq. (148) can be written in the form

$$P(v, t; u) = \int_{\Omega} dv_0 G_B(t, t_0, v, v_0) u(v_0, t_0) + \int_{t_0}^t ds \int_{\Omega} dv' G_B(t, s, v, v') \gamma \frac{v - v'}{Q(t - s)} \cdot [1 \mp P(v', s; u)] P(v', s; u). \tag{150}$$

This integral relation can be solved iteratively. In contrast to the iterative procedure discussed in section “Time-Dependent Solutions and First Order Statistics,” there is no need to compute derivatives. That is, we are dealing with some kind of path integral approach here that is similar to the numerical path integral approach involving short-time propagators; see section “Short-Time Propagator.”

In order to describe quantum particles that exhibits a Markov process, we may exploit the approach outlined in section “Markov Property: Second-Order and Higher-Order Statistics.” Accordingly, the Markov conditional probability density of the quantum particle satisfies

$$\frac{\partial}{\partial t} p(v, t | v', t', P') = \left\{ \frac{\partial}{\partial v} \gamma v [1 \mp P(v, t; u)] + Q \frac{\partial^2}{\partial v^2} \right\} \cdot p(v, t | v', t', P'), \tag{151}$$

and the self-consistent Langevin equation reads

$$\frac{d}{dt} v(t) = -\gamma v(t) (1 \mp P(v(t), t; u)) + \sqrt{Q} \Gamma(t). \tag{152}$$

From a martingale perspective, we see that stochastic trajectories $v(t)$ induce for arbitrary smooth functions f the martingale \hat{Z} with trajectories

$$Z(t) = f(v(t)) - \int_{t_0}^t ds \cdot [-\gamma v(s) [1 \mp P(v(s), s; u)] \frac{\partial}{\partial v^2} + Q \frac{\partial^2}{\partial v^2}] f(v) |_{v=v(s)}. \tag{153}$$

Moreover, as far as the Markov short-time propagator is concerned, for small time intervals $t = t' + \Delta t$, the propagator reads

$$p(v, t | v', t', P') = \sqrt{\frac{1}{2\pi Q \Delta t}} \exp \left\{ -\frac{[v - v' + \Delta t \gamma v' [1 \mp P']]^2}{2Q \Delta t} \right\}, \tag{154}$$

and can be computed from the information about the distribution P' of $v(t')$ and the state v that was observed for particular realizations of the process \hat{v} . In the stationary case, the propagator p reads for small time intervals $\Delta t = t - t'$

$$p(v, t | v', t', P_{st}(v')) = \sqrt{\frac{1}{2\pi Q \Delta t}} \exp \left\{ -\frac{[v - v' + \Delta t \gamma v' [1 \mp P_{st}(v')]]^2}{2Q \Delta t} \right\} \tag{155}$$

with P_{st} defined by Eq. (147).

Nonextensive Systems

Nonextensive thermostistical systems have been related to the Tsallis entropy (Abe and Okamoto 2001; Tsallis 1988)

$$S_q = \frac{1}{q - 1} \int_{\Omega} [P(v)^q - P(v)] dv, \tag{156}$$

where q measures the degree of nonextensivity. Diffusion processes in nonextensive thermostistical systems can be regarded as generalized Ornstein-Uhlenbeck processes that satisfy the nonlinear Fokker-Planck equation (Plastino and Plastino 1995) (see also Borland 1998; Chavanis 2003, 2004; Compte and Jou 1996; Drazer et al. 2000; Frank and Daffertshofer 1999, 2000; Shiino 2003; Tsallis and Bukman 1996)

$$\frac{\partial}{\partial t} P(v, t; u) = \frac{\partial}{\partial v} \gamma v P(x, t; u) + Q \frac{\partial^2}{\partial v^2} P(v, t; u)^q, \tag{157}$$

where v is the velocity of a particle with mass 1 that moves in one spatial dimension. In the asymptotic domain $P(v, t; u)$ approaches, a stationary Tsallis distribution

$$P_{st}(v) = \frac{D_{st}}{\left[1 + \gamma(1 - q)v^2 / [2qQD_{st}^{q-1}] \right]^{1/(1-q)}} \tag{158}$$

for $q \in (1/3, 1)$ with $D_{st} = \left[\gamma / (2qQ) z_q^2 \right]^{1/(1+q)}$ and $z_q = \sqrt{\pi / (1 - q)} \Gamma[(1 + q) / (2(1 - q))] /$

$\Gamma[1/(1 - q)]$ (Frank and Daffertshofer 2000). The process described by Eq. (157) is said to evolve in a nonextensive thermodynamic framework because its stationary probability density (158) maximizing the entropy measure (156). As discussed in section “Markov Property: Second-Order and Higher-Order Statistics,” the Markov conditional probability density p satisfies

$$\frac{\partial}{\partial t} p(v, t | v', t', P') = \frac{\partial}{\partial x} \gamma v p(v, t | v', t', P') + Q \frac{\partial^2}{\partial v^2} P(v, t; u)^{q-1} p(v, t | v', t', P'). \tag{159}$$

The self-consistent Langevin equation of the Markov diffusion process reads (see also Borland 1998)

$$\frac{d}{dt} v(t) = -\gamma v(t) + \sqrt{QP(v(t), t; u)^{q-1} \Gamma(t)}. \tag{160}$$

Any stochastic path $v(t)$ computed from Eq. (160) yields for arbitrary smooth functions f the martingale

$$Z(t) = f(v(t)) - \int_0^t ds \left[-\gamma v(s) \frac{\partial}{\partial v(s)} + QP(v(s), s; u)^{q-1} \frac{\partial^2}{\partial v(s)^2} \right] f(v(s)). \tag{161}$$

The autocorrelation $C = v(t)v(t')$ in the transient domain for $u(v_0) = \delta(v - v_0)$ reads (Frank 2004a)

$$C(t, t', v_0, t_0) = M_2(t', t_0, v_0) \exp\{-\gamma(t - t')\} \tag{162}$$

with

$$\begin{aligned} M_2(t', t_0, v_0) &= K(t', t_0) + M_1^2(t', t_0, v_0), \\ K(t', t_0) &= \frac{1}{3q - 1} \left[\frac{2qQ[z_q]^{(1-q)}}{\gamma} \cdot (1 - \exp\{-(1 + q)\gamma(t - t_0)\}) \right]^{2/(1+q)}, \\ M_1(t', t_0, v_0) &= v_0 \exp\{-\gamma(t' - t_0)\}. \end{aligned} \tag{163}$$

The autocorrelation function C depends on t_0 . This is not in contradiction with the Markov property of the underlying process as discussed in section “Markov Property: Second-Order and Higher-Order Statistics.” In particular, we may eliminate the initial condition. Then, Eq. (162) reads

$$C(t, t') = \langle u^2(t') \rangle \exp\{-\gamma(t - t')\}. \tag{164}$$

and holds for arbitrary initial probability densities u .

Linear Nonequilibrium Thermodynamics

Linear and nonlinear Fokker-Planck equations alike can be approached from the principles of linear nonequilibrium thermodynamics (de Groot

and Mazur 1962; Glansdorff and Prigogine 1971; Kondepudi and Prigogine 1998). For stochastic processes to which linear nonequilibrium thermodynamics applies the probability density $P(x, t; u)$ of a process evolves such that the free energy functional $F[P]$ decreases as a function of time t . More precisely, following a study by Compte and Jou (Compte and Jou 1996), it has been proposed that P satisfies the nonlinear Fokker-Planck equations of the form (Chavanis 2004; Frank 2002a, 2005b; Scarfone and Wada 2007)

$$\frac{\partial}{\partial t} P = \frac{\partial}{\partial x} P \tilde{M} \frac{\partial}{\partial x} \frac{\delta F}{\delta P}, \tag{165}$$

where \tilde{M} is an appropriately defined mobility coefficient and $\delta F/\delta P$ denotes the variational

derivative of F . Note that this thermodynamic approach is closely related to the GENERIC approach developed in Espanol et al. (1999), Jelic et al. (2006), Öttinger (2005, 2007), and Öttinger and Grmela (1997).

For example, the Desai-Zwanzig model (8) can be expressed in terms of Eq. (165) for $\tilde{M} = 1$ and (Frank 2005b; Shiino 1987)

$$F = \langle V \rangle + U_{MF} - QS_{BGS}. \quad (166)$$

Here, V is the potential of the force h (i.e., we have $V(x) = -\int h(x) dx$, U_{MF} is the mean field energy given by $U_{MF} = -\kappa\sigma^2/2$ (where σ^2 is the variance of the process), and S_{BGS} is the Boltzmann-Gibbs-Shannon entropy

$$S_{BGS} = -\int_{\Omega} P(x, t; u) \ln P(x, t; u) dx, \quad (167)$$

where we have put the Boltzmann constant equal to unity. The liquid crystal model (8) can be written as Eq. (165) with (Frank 2005c)

$$F = -\frac{\kappa}{2} S^2 - D_r S_{BGS}, \quad (168)$$

where S is the Maier-Saupe order parameter (135). We have $\tilde{M} = 1 - x^2$. Moreover, the expression $-\kappa S^2/2$ is the Maier-Saupe mean field energy. The Kuramoto-Shinomoto-Sakaguchi model (10) can equivalently be expressed in terms of Eq. (165) with $\tilde{M} = 1$ (see Sect. 5.4 in Frank 2005b) using

$$F = \langle V \rangle - \frac{\kappa}{2} r^2 - QS_{BGS}, \quad (169)$$

where r is the cluster phase defined by $r = |\exp\{-iX(t)\}|$. Here, the expression $-\kappa r^2/2$ is a measure for the mean field energy among the phase oscillators described by the model. The Takatsuji model (124) involves a constant mobility coefficient $\tilde{M} = 1$ and the free energy functional (Frank 2005b)

$$F = \frac{\gamma + c}{2} \langle X^2 \rangle - \text{Incosh}(\sqrt{c} \langle X \rangle) - QS_{BGS}. \quad (170)$$

The Plastino-Plastino model (14) related to the nonextensive Tsallis entropy (156) is given by Eq. (165) and $\tilde{M} = 1$ with (Frank 2005b)

$$F = \langle V \rangle - QS_q, \quad (171)$$

where V is the potential of the gradient force h . For an appropriate choice of \tilde{M} , the quantum mechanical nonlinear Fokker-Planck equations (146) can be cast into the form Eq. (165) with

$$F = \langle V \rangle - QS_{FD, BE}, \quad (172)$$

where $S_{FD, BE}$ is the quantum mechanical entropy of the Fermi-Dirac or Bose-Einstein statistics and V is the potential of the function $h(x)$ again. For details, see Frank (2005b) and Frank and Daffertshofer (1999). From the perspective of linear nonequilibrium thermodynamics, linear and nonlinear Fokker-Planck equations can be distinguished by means of the thermodynamic flux (Compte and Jou 1996; Frank 2002a, 2005b, 2007)

$$J = -\tilde{M}P \frac{\partial}{\partial x} \frac{\delta F}{\delta P}. \quad (173)$$

Note that in this approach, the thermodynamic flux is equivalent to the probability current (Frank 2005b). As can be seen from Eq. (173), on the one hand, the flux is associated to the free energy F . On the other hand, from the evolution equation (165), it follows that

$$\frac{\partial}{\partial t} P = -\frac{\partial}{\partial x} J. \quad (174)$$

If J is linear with respect to P , then the corresponding Fokker-Planck equation is linear with respect to P as well. If J is nonlinear with respect to P , then we are dealing with a nonlinear Fokker-Planck equation. The question whether J is linear or not is answered by nature herself (Frank 2007). For the Brownian particle motion, we have $\tilde{M} = \gamma$ and

$$F = \frac{1}{2} \langle v^2 \rangle - QS_{BGS}, \quad (175)$$

which yields

$$J = -\gamma v P - Q \frac{\partial}{\partial x} P. \quad (176)$$

J is linear and the corresponding Fokker-Planck equation is linear as well. For a self-

organizing system, frequently it is found that J is nonlinear because F involves a mean field energy term that is nonlinear with respect to P . For examples, see Eqs. (166), (168), (169), and (170). Likewise, for quantum and nonextensive systems, we find that J is nonlinear because F involves quantum and nonextensive entropies such as $S_{\text{FD,BE}}$ and S_q . For examples, see Eqs. (171) and (172).

Summary and Future Directions

From the previous discussion in section “[Linear Nonequilibrium Thermodynamics](#),” it is clear that modeling approaches based on nonlinear Fokker-Planck equations are rooted in the theory of collective phenomena and self-organization, on the one hand, and in the theory of quantum mechanical and nonextensive systems, on the other. In contrast, linear Fokker-Planck equations are tailored to address the stochastic properties of systems composed of noninteracting subsystems when equating the material subsystem ensemble with the ensemble of statistical realizations. Note that – of course – linear Fokker-Planck equations can also be applied to discuss stochastic properties of self-organizing systems. However, in such cases, either the stochastic behavior of order parameters by means of low-dimensional linear Fokker-Planck equations is discussed (Haken 2004) or linear high-dimensional or even functional Fokker-Planck equations are involved (Gardiner 1997).

We showed in section “[Markov Property: Second-Order and Higher-Order Statistics](#)” that both linear and nonlinear Fokker-Planck equations exhibit Green’s functions and Langevin equations. The fact that a nonlinear evolution equation can give rise to a Green’s function may be counterintuitive because Green’s functions are associated with linearity. In fact, the evolution equation of the Green’s function p is linear with respect to p . The nonlinearity is in the evolution equation for the time-dependent probability density P but not in the evolution equation of the Green’s function p . In this context, we would like to reiterate what we pointed out in

section “[Strongly Nonlinear Fokker-Planck Equations](#)”: time-dependent solutions P do not necessarily correspond to Green’s function p (Frank 2003b).

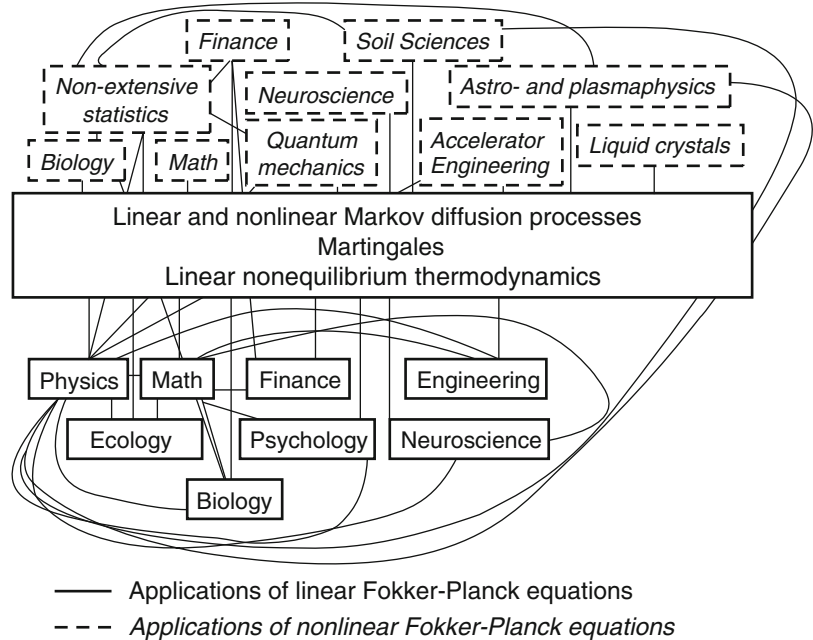
In the mathematical literature, the theory of Markov processes that involve conditional probability densities of the form $p(x, t|x', t', P')$ has been discussed for several decades (see references in section “[Definition of the Subject](#)”). In line with these studies, we suggest to refer to Markov processes with conditional probability densities of the form $p(x, t|x', t', P')$ as *nonlinear* Markov processes or *nonlinear* families of Markov processes (Frank 2004d). Likewise, we suggest to refer to Markov processes whose conditional probability densities do not depend on P' as *linear* Markov processes or *linear* families of Markov processes. Using this terminology, we would say that strongly nonlinear Fokker-Planck equations describe nonlinear Markov diffusion processes, and vice versa nonlinear Markov diffusion processes can be expressed in terms of strongly nonlinear Fokker-Planck equations.

In physics and related disciplines, the relevance of nonlinear Markov processes has to be explored in the future. That is, except for research primarily reported in the mathematical literature, the theory of Markov processes constructed from conditional probability measures of the form $p(x, t|x', t', P')$ is still in its infancy. The Chapman-Kolmogorov equation and the Kramers-Moyal expansion presented in section “[Markov Property: Second-Order and Higher-Order Statistics](#)” provide promising departure points for future studies in this regard.

In the present study, we pointed out that there are a few overarching concepts that apply to linear and nonlinear Fokker-Planck equations alike: the concepts of Markov diffusion processes, martingales, and linear nonequilibrium thermodynamics. Therefore, future studies may change the state of the art illustrated in Fig. 1 into a scenario as shown in Fig. 2. In doing so, a closely connected world of linear and nonlinear Fokker-Planck equations that is governed by a small set of powerful principles could emerge.

Linear and Nonlinear Fokker-Planck Equations,

Fig. 2 Connected applications of linear and nonlinear Fokker-Planck equations



Bibliography

Abe S, Okamoto Y (2001) Nonextensive statistical mechanics and its applications. Springer, Berlin

Acebron JA, Spigler R (1998) Adaptive Frequency Model for Phase-Frequency Synchronization in Large Populations of Globally Coupled Nonlinear Oscillators. *Phys Rev Lett* 81:2229

Acebron JA, Bonilla LL, De Leo S, Spigler R (1998) Breaking the symmetry in bimodal frequency distributions of globally coupled oscillators. *Phys Rev E* 57:5287

Acebron JA, Bonilla LL, Vicente CJP, Ritort F, Spigler R (2005) The Kuramoto model: A simple paradigm for synchronization phenomena. *Rev Mod Phys* 77:137

Allen EJ, Victory HD Jr (1994) A computational investigation of the random particle method for numerical solution of the kinetic Vlasov-Poisson-Fokker-Planck equations. *Physica A: Statistical Mechanics and its Applications* 209:318

Aonishi T, Okada M (2002) Multibranch Entrainment and Slow Evolution among Branches in Coupled Oscillators. *Phys Rev Lett* 88:024102

Arenas A, Perez Vicente CJ (1994) Exact long-time behavior of a network of phase oscillators under random fields. *Phys Rev E* 50:949

Ariaratnam JT, Strogatz SH (2001) Phase Diagram for the Winfree Model of Coupled Nonlinear Oscillators. *Phys Rev Lett* 86:4278

Arnold A, Bonilla LL, Markowich P (1996) *Transp Theor Stat Phys* 25:733

Arnold A, Markowich P, Toscani G (2000) *Transp Theor Stat Phys* 29:571

Arnold A, Markowich P, Toscani G, Unterreiter A (2001) *Commun Part Diff Eq* 26:43

Aronson DG (1986) In: Dobb A, Eckmann B (eds) *Non-linear diffusion problems – lecture notes in mathematics*, vol 1224. Springer, Berlin, pp 1–46

Balescu R (1975) *Equilibrium and nonequilibrium statistical mechanics*. Wiley, New York

Barenblatt GI, Entov VM, Ryzhik VM (1990) *Theory of fluid flows through natural rocks*. Kluwer Academic Publisher, Dordrecht

Becker V, Engel A (2007) Role of interactions in ferrofluid thermal ratchets. *Phys Rev E* 75:031118

Binney J, Tremaine S (1987) *Galactic dynamics*. Princeton University Press, Princeton

Birner T, Lippert K, Müller R, Kühnel A, Behn U (2002) Critical behavior of nonequilibrium phase transitions to magnetically ordered states. *Phys Rev E* 65:046110

Bödeker HU, Röttger MC, Liehr A, Frank TD, Friedrich R, Purwins HG (2003) *Phys Rev E* 67:056220

Bogacz R, Brown E, Moehlis J, Holmes P, Cohen JD (2006) The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychol Rev* 113:700

Bonilla LL (1987) Stable nonequilibrium probability densities and phase transitions for meanfield models in the thermodynamic limit. *J Stat Phys* 46:659

Bonilla LL, Vicente CJP, Rubi JM (1993) *Stat J Phys* 70:921

Borland L (1998) Microscopic dynamics of the nonlinear Fokker-Planck equation: A phenomenological model. *Phys Rev E* 57:6634

Borland L (2002) Option Pricing Formulas Based on a Non-Gaussian Stock Price Model. *Phys Rev Lett* 89:098701

Borland L (2008) Non-Gaussian option pricing. Successes, limitations and perspectives. In: Riccardi C, Roman HE

- (eds) Anomalous fluctuations in complex systems: plasmas, fluids and financial markets. Special Review Book for Research Singpost, Transworld Research Network, KeralaMadison, pp 311–334
- Borland L, Bouchaud JP (2004) A Non-Gaussian Option Pricing Model with Skew. *Quantitative finance* 4:499
- Bychenkov VY, Rozmus W, Tikhonchuk VT (1995) Non-local Electron Transport in a Plasma. *Phys Rev Lett* 75:4405
- Carillo JA, Jüngel A, Markowich PA, Toscani G, Unterreiter A (2001) *Monatshefte für Mathematik* 113:1
- Carrillo JA, Rosado J, Salvarani F (2008) 1D nonlinear Fokker-Planck equations for fermions and bosons. *Appl Math Letters* 21:148
- Cepa E, Lepingle D (1997) Diffusing particles with electrostatic repulsion. *Probab Theory Relat Fields* 107:429
- Chandrasekhar S (1977) *Liquid crystals*. Cambridge University Press, Cambridge
- Chavanis PH (2003) Generalized thermodynamics and Fokker-Planck equations: Applications to stellar dynamics and two-dimensional turbulence. *Phys Rev E* 68:036108
- Chavanis PH (2004) *Physica A* 340:57
- Chavanis PH, Rosier C, Sire C (2002) Thermodynamics of self-gravitating systems. *Phys Rev E* 66:036105
- Coffey WT, Kalmykov YP, Waldron JT (2004) *The Langevin equation*. World Scientific, Singapore
- Compte A, Jou D (1996) Non-extensive behavior of a stock market index at microscopic time scales. *Phys J Math A Gen* 29:4321
- Cortines AAG, Riera R (2007) *Physica A* 377:181
- Crank J (1975) *The mathematics of diffusion*. Clarendon Press, Oxford
- Crawford JD (1995) *Phys Rev Lett* 74:4341
- Curado EMF, Nobre FD (2003) *Phys Rev E* 67:021107
- Daido H (1996a) *Physica D* 91:24
- Daido H (1996b) *Phys Rev Lett* 77:1406
- Daido H (2001) *Phys Rev Lett* 87:048101
- Daly E, Porporato A (2004) *Phys Rev E* 70:056303
- Dano S, Hynne F, De Monte S, d'Ovidio F, Sorensen PG, Westerhoff H (2001) *Faraday Discuss* 120:261
- Dawson DA (1983) *J Stat Phys* 31:29
- Dawson DA (1993) *Lecture notes in mathematics* 1541. In: Dawson DA, Maisonneuve B, Spencer J (eds) . Springer, Berlin, pp 6–16
- Dawson DA, Gärtner J (1989) Large deviations, free energy functional and quasi-potential for a mean field model of interacting diffusions. *American Mathematical Society, Providence*
- de Gennes P, Prost J (1993) *The physics of liquid crystals*. Clarendon Press, Oxford
- de Groot SR, Mazur P (1962) *Non-equilibrium thermodynamics*. North-Holland, Amsterdam
- de Jeu WH (1980) *Physical properties of liquid crystalline materials*. Gordon and Beach, New York
- de los Santos F, Telo da Gamma MM, Munoz MA (2003) *Phys Rev E* 67:021607
- Desai RC, Zwanzig R (1978) *Stat J Phys* 19:1
- Ding X (1994) *Adv Appl Prob* 26:1022
- Djehiche B, Kaj I (1995) *Ann Probab* 23:1414
- Doi M, Edwards SF (1988) *The theory of polymer dynamics*. Clarendon Press, Oxford
- Donoso JM, Salgado JJ (2006) *J Phys A* 39:12587
- Donoso JM, Salgado JJ, Soler M (2005) *J Phys A* 38:9145
- Drazer G, Wio HS, Tsallis C (2000) *Phys Rev E* 61:1417
- Drozдов AN, Morillo M (1996) *Phys Rev E* 54:931
- Ebeling W, Sokolov IM (2004) *Statistical thermodynamics and stochastic theory of nonequilibrium systems*. World Scientific, Singapore
- Eckhardt B, Ott E, Strogatz SH, Abrams DM, McRobie A (2007) *Phys Rev E* 75:021110
- Epperlein EM, Rickard GJ, Bell AR (1988) *Phys Rev Lett* 61:2453
- Espanol P, Serrano M, Öttinger HC (1999) *Phys Rev Lett* 83:4542
- Felderhof BU (2003) *Physica A* 323:88
- Fialkowiski M, Hess S (2000) *Physica A* 282:65
- Fokker AD (1914) *Ann Phys* 43:810
- Fontbona J (2003) *J Funct Analysis* 200:198
- Frank TD (2001a) *Phys Lett A* 280:91
- Frank TD (2001b) *Phys Lett A* 290:93
- Frank TD (2002a) *Phys Lett A* 305:150
- Frank TD (2002b) *Physica A* 310:397
- Frank TD (2003a) *Phys Lett A* 319:173
- Frank TD (2003b) *Physica A* 320:204
- Frank TD (2004a) *Eur Phys J B* 37:139
- Frank TD (2004b) *J Phys A* 37:3561
- Frank TD (2004c) *Phys Lett A* 329:475
- Frank TD (2004d) *Physica A* 331:391
- Frank TD (2004e) *Physica D* 195:229
- Frank TD (2005a) *Math Comput Mod* 42:1057
- Frank TD (2005b) *Nonlinear Fokker-Planck equations: fundamentals and applications*. Springer, Berlin
- Frank TD (2005c) *Phys Rev E* 72:041703
- Frank TD (2006) *Phys Rev ST-AB* 9:084401
- Frank TD (2007) *Physica A* 382:453
- Frank TD (2008) *Physica A* 387:773
- Frank TD, Daffertshofer A (1999) *Physica A* 272:497
- Frank TD, Daffertshofer A (2000) *Physica A* 285:351
- Frank TD, Daffertshofer A (2001a) *Physica A* 292:392
- Frank TD, Daffertshofer A (2001b) *Physica A* 295:455
- Frank TD, Friedrich R (2005) *Physica A* 347:65
- Frank TD, Daffertshofer A, Peper CE, Beek PJ, Haken H (2001) *Physica D* 150:219
- Friedman A (1969) *Partial differential equations*. Holt, Rinehart and Winston, Inc., New York
- Friedrich R, Peinke J (1997) *Phys Rev Lett* 78:863
- Friedrich R, Siegert S, Peinke J, Lück S, Seifert M, Lindemann M, Raethjen J, Deuschl G, Pfister G (2000) *Phys Lett A* 271:217
- Gang H, Haken H, Fagen X (1996) *Phys Rev Lett* 77:1925
- García-Ojalvo J, Sancho JM (1999) *Noise in spatially extended systems*. Springer, New York
- García-Ojalvo J, Parrondo JMR, Sancho JM, van den Broeck C (1996) *Phys Rev E* 54:6918
- Gardiner CW (1997) *Handbook of stochastic methods*, 2nd edn. Springer, Berlin
- Gärtner J (1988) *Math Nachr* 137:197

- Giada L, Marsili M (2000) *Phys Rev E* 62:6015
- Glansdorff P, Prigogine I (1971) *Thermodynamic theory of structure, stability, and fluctuations*. Wiley, New York
- Goel NS, Dyn NR (1974) *Stochastic models in biology*. Academic Press, New York
- Graham C (1990) *Appl Math Optim* 22:75
- Greven A (2005) In: Deuschel J, Greven A (eds) *Interacting stochastic systems*. Springer, Berlin, pp 209–246
- Gurtin ME, MacCamy RC (1977) *Math Biosci* 33:35
- Hadley P, Beasley MR, Wiesenfeld K (1988) *Phys Rev B* 38:8712
- Haken H (2004) *Synergetics: introduction and advanced topics*. Springer, Berlin
- Han SK, Kurrer C, Kuramoto Y (1995) *Phys Rev Lett* 75:3190
- Hänggi P, Talkner P, Borkovec M (1990) *Rev Mod Phys* 62:251
- Hänggi P, Marchesoni F, Nori F (2005) *Ann Phys* 14:51
- Hansel D, Mato G, Meunier C (1993a) *Europhys Lett* 23:367
- Hansel D, Mato G, Meunier C (1993b) *Phys Rev E* 48:3470
- Hasegawa H (2003) *Phys Rev E* 67:041903
- Heifets S (2001) *Phys Rev ST-AB* 4:044401
- Heifets S (2003) *Phys Rev ST-AB* 6:080701
- Hess S (1976) *Naturforschung Z A* 31:1034
- Holden AV (1976) *Models of the stochastic activity of neurons, Lecture notes in biomathematics*, vol 12. Springer, Berlin
- Horsthemke W, Lefever R (1984) *Noise-induced transitions*. Springer, Berlin
- Hütter M, Karlin IV, Öttinger HC (2003) *Phys Rev E* 68:016115
- Ichiki A, Ito H, Shiino M (2007) *Phys E* 40:402
- Ilg P, Karlin IV, Öttinger HC (1999) *Phys Rev E* 60:5783
- Ilg P, Kröger M, Hess S (2005) *Phys Rev E* 71:051201
- Jafari GR, Fazeli SM, Ghasemi F, Allaei SMV, Tabar MRR, Zad AI, Kavei G (2002) *Phys Rev Lett* 91:226101
- Jelic A, Hütter M, Öttinger HC (2006) *Phys Rev E* 74:041126
- Jourdain B (2000) *Methodol Comput Appl Probab* 2:69
- Kadanoff LP (2000) *Statistical physics: statics, dynamics and renormalization*. World Scientific, Singapore
- Kanamaru T, Horita T, Okabe Y (2001) *Phys Rev E* 64:031908
- Kaniadakis G (2001a) *Phys Lett A* 288:283
- Kaniadakis G (2001b) *Physica A* 296:405
- Kaniadakis G, Quarati P (1993) *Phys Rev E* 48:4263
- Kaniadakis G, Quarati P (1994) *Phys Rev E* 49:5103
- Karlin S, Taylor HM (1975) *A first course in stochastic processes*. Academic Press, New York
- Karlin S, Taylor HM (1981) *A second course in stochastic processes*. Academic Press, New York
- Kawamura Y (2007) *Phys Rev E* 76:047201
- Kawamura Y, Nakao H, Kuramoto Y (2007) *Phys Rev E* 75:036209
- Kharchenko DO, Kharchenko VO (2005) *Physica A* 354:262
- Kim S, Park SH, Ryu CS (1997) *Phys Rev Lett* 78:1616
- Klimontovich YL (1986) *Statistical physics*. Harwood Academic Publ, New York
- Kloeden PE, Platen E (1992) *The numerical solution of stochastic differential equations*. Springer, Berlin
- Kometani K, Shimizu H (1975) *J Stat Phys* 13:473
- Kondepudi D, Prigogine I (1998) *Modern thermodynamics*. Wiley, New York
- Kostur M, Luczka J, Schimansky-Geier L (2002) *Phys Rev E* 65:051115
- Kozyreff G, Vladimirov AG, Mandel P (2000) *Phys Rev Lett* 85:3809
- Kuramoto Y (1984) *Chemical oscillations, waves, and turbulence*. Springer, Berlin
- Lancellotti C, Kiessling M (2001) *Astrophys J* 549:L93
- Larson RG, Öttinger HC (1991) *Macromolecules* 24:6270
- Li JH, Hänggi P (2001) *Phys Rev E* 64:011106
- Li JH, Huang ZQ, Xing DY (1998) *Phys Rev E* 58:2838
- Lo CF (2005) *Phys Lett A* 336:141
- MacDonald WM, Rosenbluth MN, Chuck W (1957) *Phys Rev* 107:350
- Maier W, Saupé A (1958) *Z Naturforschung A* 13:564
- Marsili M, Bray AJ (1996) *Phys Rev Lett* 76:2750
- McCauley JL, Gunaratne GH, Bassler KE (2006) *Physica A* 382:445
- McKean HP Jr (1969) In: Aziz AK (ed) *Propagation of chaos for a class of nonlinear parabolic equations, Lectures in differential equations*, vol II. Van Nostrand Reinhold Company, New York, pp 177–193
- Meleard S (1996) *Asymptotic behavior of some interacting particle systems: McKean-Vlasov and Boltzmann models*. In: Graham C, Kurtz TG, Meleard S, Potter PE, Pulvirenti M, Talay D (eds) *Probabilistic models for nonlinear partial differential equations*. Springer, Berlin, pp 42–95
- Meleard S, Coppoletta SR (1987) *Stoch Process Appl* 26:317
- Michael F, Johnson MD (2003) *Physica A* 320:525
- Mikhailov AS, Zanette DH (1999) *Phys Rev E* 60:4571
- Morillo M, Gomez-Ordóñez J, Casado JM (1995) *Phys Rev E* 52:316
- Müller R, Lippert K, Kühnel A, Behn U (1997) *Phys Rev E* 56:2658
- Nicholson DR (1983) *Introduction to plasma theory*. Wiley, New York
- Nobre FD, Curado EMF, Rowlands G (2004) *Physica A* 334:109
- Oelschläger K (1989) *Probab Th Rel Fields* 82:565
- Okubo A, Levin SA (2001) *Diffusion and ecological problems: modern perspectives*. Springer, Berlin
- Öttinger HC (1996) *Stochastic processes in polymeric fluids*. Springer, Berlin
- Öttinger HC (2005) *Beyond equilibrium thermodynamics*. Wiley, New Jersey
- Öttinger HC (2007) *Phys Rev Lett* 99:130602
- Öttinger HC, Grmela M (1997) *Phys Rev E* 56:6633
- Overbeck L (1996) *Ann Probab* 24(2):743
- Park SH, Kim S (1996) *Phys Rev E* 53:3425

- Park K, Lai YC, Liu Z, Nachman A (2004) *Phys Lett A* 326:391
- Patanarapeelert K, Frank TD, Friedrich R, Beek PJ, Tang IM (2006) *Phys Lett A* 360:190
- Paul W, Baschnagel J (1999) *Stochastic processes: from physics to finance*. Springer, Berlin
- Peletier LA (1981) In: Amann H, Bazley N, Kirchgässner K (eds) *Applications of nonlinear analysis in the physical science*. Pitman Advanced Publishing Program, Boston, pp 229–241
- Pikovsky A, Rosenblum M, Kurths J (2001) *Synchronization: a universal concept in nonlinear sciences*. Cambridge University Press, Cambridge
- Pilipenko AY (2005) *Ukrainian Math J* 57:1507
- Planck M (1917) *Sitzungsber. Preuss Akad Wissens, Berlin*, p 324
- Plastino AR, Plastino A (1995) *Physica A* 222:347
- Quinn DA, Rand RH, Strogatz SH (2007) *Phys Rev E* 75:036218
- Ratcliff R, Gomez P, McKoon G (2004) *Psychol Rev* 111:159
- Reif F (1965) *Fundamentals of statistical and thermal physics*. McGraw-Hill Book Company, New York
- Reimann P (2002) *Phys Rep* 361:57
- Reimann P, Kawai R, Van den Broeck C, Hänggi P (1999a) *Europhys Lett* 45:545
- Reimann P, Van den Broeck C, Kawai R (1999b) *Phys Rev E* 60:6402
- Risken H (1989) *The Fokker-Planck equation – methods of solution and applications*. Springer, Berlin
- Rogers LCG, Shi Z (1993) *Probab Theroy Relat Fields* 95:555
- Rosenblum MG, Pikovsky AS (2004) *Phys Rev Lett* 92:114102
- Rosenbluth M, MacDonald WM, Judd DL (1957) *Phys Rev* 107:1
- Sakaguchi H (1988) *Prog Theor Phys* 79:39
- Savel'ev S, Marchesoni F, Nori F (2003) *Phys Rev Lett* 91:010601
- Scarfone MA, Wada T (2007) *Physica A* 384:305
- Schuster HG, Wagner P (1990) *Biol Cybern* 64:77
- Schwämmle V, Nobre FD, Curado E (2007a) *Phys Rev E* 76:041123
- Schwämmle V, Gonzahles MC, Moreira AA, Andrade JS, Hermann HJ (2007b) *Phys Rev E* 75:066108
- Schweitzer F (2003) *Brownian agents and active particles*. Springer, Berlin
- Shiino M (1985) *Phys Lett A* 112:302
- Shiino M (1987) *Phys Rev A* 36:2393
- Shiino M (2001) *J Math Phys* 42:2540
- Shiino M (2002a) *J Math Phys* 43:2654
- Shiino M (2002b) *J Phys Soc* 40:1037
- Shiino M (2003) *Phys Rev E* 67:056118
- Shimizu H (1974) *Prog Theor Phys* 52:329
- Shimizu H, Yamada T (1972) *Prog Theor Phys* 47:350
- Shobuda Y, Hirata K (2001) *Phys Rev E* 64:067501
- Silverman BW (1986) *Density estimation for statistics and data analysis*. Chapman and Hall, London
- Soler M, Martinez FC, Donoso JM (1992) *Stat J Phys* 69:813
- Sopik J, Sire C, Chavanis PH (2006) *Phys Rev E* 74:011112
- Strogatz SH (2000) *Physica D* 143:1
- Strogatz SH, Mirollo RE (1991) *Stat J Phys* 63:613
- Strogatz SH, Morillo RE, Matthews PC (1992) *Phys Rev Lett* 68:2730
- Strogatz SH, Abrams D, McRobie A, Eckhardt B, Ott E (2005) *Nature* 438:43
- Stupakov GV, Breizman BN, Pekker MS (1997) *Phys Rev E* 55:5976
- Sura P, Barsugli J (2002) *Phys Lett A* 305:304
- Takai M, Akiyama H, Takeda S (1981) *J Phys Soc Japan* 50:1716
- Takatsuji M (1975) *Biol Cybern* 17:207
- Tass PA (1999) *Phase resetting in medicine and biology – stochastic modelling and data analysis*. Springer, Berlin
- Tass PA (2001) *Europhys Lett* 55:171
- Tass PA (2003) *Biol Cybern* 89:81
- Tass PA (2006) *Biol Cybern* 94:58
- Tsallis C (1988) *Stat J Phys* 52:479
- Tsallis C (1997) *Phys World* 10(7):42
- Tsallis C (2004) *Physica D* 193:3
- Tsallis C, Bukman DJ (1996) *Phys Rev E* 54:R2197
- Tsekov R (1995) *J Phys A Math Gen* 28:L557
- Tsekov R (2001) *Int J Mol Sci* 2:66
- van den Broeck C, Parrondo JMR, Armero J, Hernandez-Machado A (1994a) *Phys Rev E* 49:2639
- van den Broeck C, Parrondo JMR, Toral R (1994b) *Phys Rev Lett* 73:3395
- van den Broeck C, Parrondo JMR, Toral R, Kawai R (1997) *Phys Rev E* 55:4084
- van Kampen NG (1981) *Stochastic processes in physics and chemistry*. North-Holland, Amsterdam
- Vellekoop M, Nieuwenhuis H (2007) *Quantitative finance* 7:563
- Venturini M, Warnock R (2002) *Phys Rev Lett* 89:224802
- Waechter M, Riess F, Schimmel T, Wendt U, Peinke J (2004) *Eur Phys J B* 41:259
- Wehner MF, Wolfer WG (1987) *Phys Rev A* 35:1795
- Wiesenfeld K, Colet P, Strogatz SH (1996) *Phys Rev Lett* 76:404
- Winfree AT (1967) *Theor J Biol* 16:15
- Winfree AT (2001) *The geometry of biological time*, 2nd edn. Springer, Berlin
- Yamaguchi Y, Shimizu H (1984) *Physica D* 11:213
- Yamana M, Shiino M, Yoshioka M (1999) *Phys J A Math Gen* 32:3525
- Yoshioka M, Shiino M (2000) *Phys Rev E* 61:4732
- Zaikin AA, Garcia-Ojalvo J, Schimansky-Geier L, Kurths J (2002) *Phys Rev Lett* 88:010601
- Zhai Y, Kiss IZ, Daido H, Hudson JL (2005) *Physica D* 205:57