Fluctuations

N. G. van Kampen*

1 Introduction

Originally thermodynamics was only concerned with the macroscopically observed values of physical quantities, which in statistical mechanics are given by averages. However, Brownian motion showed that thermodynamics is not the whole story. In fact, Gibbs statistical treatment led to fluctuations about these averages. (Subsequently Einstein noticed that the relation S = k log W enables one to treat these fluctuations thermodynamically as well). But these are the *instantaneous* values of the physical quantities, and their theory is well known. We are here concerned with the behaviour of the *fluctuations as functions of time*. Their study is much harder, because it involves not only the equilibrium description, but also the evolution in time, which in principle requires the solution of the equations of motion of the physical system considered.

There are several reasons why interest in fluctuations has grown in recent times. First, they form the obstacle in precise measurements and delicate experiments, and techniques to eliminate their effects are of practical importance. Secondly they may be used as a source of information concerning the system; the most striking example being the fluctuations of the electromagnetic field usually referred to as the 3K radiation. Thirdly they may give rise to macroscopic effects such as the Van-der-Waals force.

The proper framework for describing fluctuations is the theory of stochastic processes; they are treated in sections 2-4. and the master equation for Markov processes is derived. Subsequently, in sections 5-8 an expansion of the master equation is given, which makes its application to practical cases possible. The final sections contain an introduction to the theory of stochastic differential equations.

*University of Utrecht, Utrecht, Holland
Re-printed from: Irreversibility in the Many-Body Problem: Sitges International School of Physics, May 1972.
Edited by J. Biel and J. Rae. Plenum Press, New York, 1972. pp. 349-386.

2 Stochastic processes

A *stochastic variable* X is defined by the set of its values x(= set of states = range), and a probability distribution over this set. Once such a stochastic variable is given, any related quantity $Y = f(X)$ is also a stochastic variable. Y may be any kind of mathematical object, in particular a function of an auxiliary variable t, that is, $Y(t) = f(X,t)$. Such $Y(t)$ is called a *stochastic process*. It may be regarded as a collection of sample functions or realizations $y(t) = f(x,t)$, each of which is obtained by assigning to X one of its possible values$x$.

Consider a classical dynamical system; its instantaneous state is given by a point $(q_1, q_2, ..., q_f,\ p_1, p_2, ..., p_f)$ in its phase space. In statistical mechanics one replaces the individual system by an *ensemble*. This amounts to defining a stochastic variable X, whose range consists of all points $x = (q,p)$ in phase space and whose probability distribution is specified by the ensemble. The ensemble may be chosen freely at one particular time $t = 0$ and the equations of motion then determine the probability of finding a particular state at a later (or earlier) time t.

A *dynamical quantity* is a function of the state (q,p). Its value at any time t is determined by the equations of motion together with the initial state x at $t = 0$, and may be written $f(x,t)$. By the introduction of an ensemble, each dynamical variable is turned into a stochastic process $Y(t) = f(X,t)$. In this way stochastic processes enter into physics. So far this is nothing but a reformulation, which does not bring us any nearer to solving the equations of motion. However, it will appear that the theory of stochastic processes suggests certain assumptions, hopefully called approximations, which make it possible to obtain results without having to solve the equations of motion. Some of these results, for instance the theory of Brownian motion, have proved very successful.

In quantum statistical mechanics the possible values

of X are the unit vectors[1] in the Hilbert space of the system , although for describing equilibrium it is sufficient to take only the eigenstates of the Hamiltonian operator. The probability distribution is given by means of a density matrix. The functions $Y(t)$ are the expectation values of the various operators. The theory of stochastic processes does not ask whether the system is classical or quantum mechanical, but, of course, the specific properties of the observables need not be the same in both cases.

3 The Markov assumption

A stochastic process $Y(t)$, defined by means of an X as above, leads to a hierarchy of probability densities. We write

$$P_n(y_1,t_1;y_2,t_2;...;y_n,t_n)\,dy_1dy_2...dy_n$$

for the probability that $Y(t_1)$ lies between y_1 and y_1+dy_1, and also that $Y(t_1)$ lies between y_2 and y_2+dy_2, etc. The P_n are defined for $n=1,2,3,...$ and only for unequal times. Clearly one has

1. $P_n \geq 0$;
2. P_n invariant for interchange of the pair y_i, t_i with the pair y_j, t_j for any i, j ;
3. $\int P_n dy_n = P_{n-1}$, and $\int P_1 dy_1 = 1$.

Kolmogorov [1] has proved that the reverse is also true. Any set of functions P_n obeying these three restrictions defines a stochastic process. For future use we note that the process may also be characterized by the hierarchy of its moments, defined by

$$\mu_n(t_1,t_2,...,t_n) \equiv \langle Y(t_1)\,Y(t_2)\,...Y(t_n)\rangle$$
$$= \int y_1y_2...y_nP_n(y_1,t_1;y_2,t_2;...;y_n,t_n)\,dy_1dy_2...dy_n.$$

The conditional probability density $P_{n|\nu}$ for finding y_1 at t_1, and y_2 at t_2, etc., *given* that one has $Y(\tau_1)=\eta_1$, $Y(\tau_2)=\eta_2$,..., is determined by

$$P_{n|\nu}(y_1,t_1;...;y_n,t_n|\eta_1,\tau_1;...;\eta_\nu,\tau_\nu)\,P(\eta_1,\tau_1;...;\eta_\nu,\tau_\nu)$$
$$= P_{n+\nu}(y_1,t_1;...;y_n,t_n;\eta_1,\tau_1;...;\eta_\nu,\tau_\nu).$$

Clearly, $P_{n|\nu}$is symmetric in the pairs y_i,t_i and also in the pairs η_j,τ_j. It is therefore no restriction to write the τ_j in chronological order, $\tau_1<\tau_2<...<\tau_\nu$. Suppose that all t_i are later than τ_ν; a Markov process is defined as a process for which $P_{n|\nu}$ does not depend on the earlier data $\eta_1,...,\eta_{\nu-1}$, but only on η_ν. More explicitly

$$P_{n|\nu}(y_1,t_1;...;y_n,t_n|\eta_1,\tau_1;...;\eta_\nu,\tau_\nu)$$
$$= P_{n|1}(y_1,t_1;...;y_n,t_n|\eta_\nu,\tau_\nu).$$

It follows that, if $t_1<t_2<...<t_n$,

$$P_n(y_1,t_1;y_2,t_2;...;y_n,t_n)$$
$$= P_1(y_1,t_1)\,P_{1|1}(y_2,t_2|y_1,t_1)\,...P_{1|1}(y_n,t_n|y_{n-1},t_{n-1})$$

This shows that, in the case of a Markov process, the whole hierarchy is determined by only its first member P_1 and the transition probability $P_{1|1}$. The Markov property is a very strong restriction, but it has the effect of making the hierarchy tractable. It often happens that one has reason to believe that a physical quantity, in whose fluctuating behaviour one is interested, is approximately a Markov process. It then becomes possible to find all the stochastic properties of that quantity; the need for solving the actual equations of motion of the whole system has been sidestepped by assuming Markov character. Before demonstrating this technique we still have to develop the formalism somewhat further.

4 The master equation

From the Markov character follows for $t_1<t_2<t_3$

$$P_3(y_1,t_1;y_2,t_2;y_3,t_3)$$
$$= P_2(y_2,t_2|y_1,t_1)\,P_{1|1}(y_3,t_3|y_2,t_2).$$

On integrating over y_2 and dividing by $P_1(y_1,t_1)$ one obtains an integral equation for the transition probability

$$P_{1|1}(y_3,t_3|y_1,t_1)$$
$$= \int P_{1|1}(y_3,t_3|y_2,t_2)\,P_{1|1}(y_2,t_2|y_1,t_1)\,dy_2.$$

This is the *Chapman-Kolmogorov equation.* One easily verifies that for example

$$P_{1|1}(y_2,t_2|y_1,t_1) = [2\pi(t_2-t_1)]^{-\frac{1}{2}}\exp\left[\frac{(y_2-y_1)^2}{2(t_2-t_1)}\right]$$

obeys this equation. Together with

$$P_1(y_1,t_1) = [2\pi t_1]^{-\frac{1}{2}}\exp\left[-\frac{y_1^2}{2t_1}\right]$$

it defines a Markov process for $t>0$, which is called the *Wiener process* and describes the random displacement of a Brownian particle.

[1] More precisely all rays in Hilbert space, that is, all unit vectors disregarding a phase factor.

If one takes $t_3 - t_2 \equiv \Delta t$ small, one expects that the probability that any transition at all has taken place is proportional to Δt. Accordingly we now assume

$$P_{1|1}(y_3, t_2 + \Delta t | y_2, t_2) = A\,\delta(y_3 - y_2) + \Delta t\, W_{t_2}(y_3|y_2).$$

Here W_{t_2} is the transition probability per unit time at the time t_2, and conservation of probability tells us

$$A = 1 - \Delta t \int W_{t_2}(y'|y_2)\, dy'.$$

REMARK: The Wiener process does not obey this assumption. The reason is that it contains in each Δt infinitely many infinitely small transitions. That roughly coincides with the definition of continuous stochastic processes [2]. Our assumption applies to discontinuous processes, in which finite jumps are made, but such that the probability for any jump to take place in Δt is of order Δt. Our contention is that this is the correct description of physical processes, and that the Wiener process can at best be regarded as a first step in an approximating expansion scheme (see section 6).

Substitution leads to the differential form of the Chapman-Kolmogorov equation,

$$\frac{\partial P_{1|1}(y, t|y_1, t_1)}{\partial t} = \int W_t(y|y')\, P_{1|1}(y', t|y_1, t_1)\, dy' - P_{1|1}(y, t|y_1, t_1) \int W_t(y'|y)\, dy'$$

To make the meaning of this result more transparent, we multiply the equation by $P_1(y_1, t_1)$ and integrate over y_1,

$$\frac{\partial P_1(y,t)}{\partial t} = \int \{W_t(y|y')\, P_1(y', t) - W_t(y'|y)\, P_1(y,t)\}\, dy'$$

This is the celebrated *master equation*; it is a linear homogeneous equation describing the evolution of the probability density of a Markov process. Its physical importance derives from the fact that W refers to the short time behaviour of the total system and therefore can often be computed; the master equation then serves to deduce the long time behaviour from this.

Although the Wiener process did not obey the assumption used in deriving the master equation, we see in retrospect that it does obey a master equation with a singular kernel W,

$$\frac{\partial P_1(y,t)}{\partial t} = \frac{1}{2}\frac{\partial^2 P_1(y,t)}{\partial y^2}.$$

This equation can be solved explicitly (Fourier); more generally, the master equation can be solved whenever it is invariant for shifts in y and t, i.e., when y ranges from $-\infty$ to ∞ and $W_t(y|y')$ is a function of $y - y'$ alone. The reason is that in that case the successive jumps are statistically independent of each other. Another solvable master equation is the linear Fokker-Planck equation, which gives rise to the Ornstein-Uhlenbeck process,

$$\frac{\partial P_1(y,t)}{\partial t} = \frac{\partial}{\partial y} y P_1(y,t) + \frac{1}{2}\frac{\partial^2 P_1(y,t)}{\partial y^2}.$$

However, for most master equations an approximation method is needed; this will now be developed.

5 Expansion of the master equation

We suppose that $W_t(y|y')$ is independent of t and rewrite it as a function $W(y'; \xi)$ of the starting point y' and the jump length ξ,

$$\frac{\partial P_1(y,t)}{\partial t} = \int \{W(y - \xi; \xi)\, P_1(y - \xi, t) - W(y; \xi)\, P_1(y,t)\}\, d\xi$$

Expanding the first term on the right in powers of ξ one obtains the *Kramers-Moyal expansion*

$$\frac{\partial P_1(y,t)}{\partial t} = \sum_{\nu=1}^{\infty} \frac{(-1)^\nu}{\nu!} \left(\frac{\partial}{\partial y}\right)^\nu \alpha_\nu(y)\, P_1(y,t),$$

where the α_ν are the moments of the jump probability,

$$\alpha_\nu(y) = \int_{-\infty}^{\infty} \xi^\nu W(y; \xi)\, d\xi.$$

It is tempting to break off after two terms and adopt the resulting nonlinear Fokker-Planck equation

$$\frac{\partial P_1}{\partial t} = -\frac{\partial}{\partial y}\alpha_1(y)\, P_1 + \frac{1}{2}\frac{\partial^2}{\partial y^2}\alpha_2(y)\, P_1$$

as an approximation to the master equation. This equation appears in most textbooks and is often called after Kolmogorov ([3],[4]). However, it is derived by means of the same unphysical assumptions on which the Wiener process is based. Actually this equation is NOT a systematic approximation to the master equation and has caused a great deal of confusion [5]. Our purpose is to provide a systematic expansion in powers of a parameter which obviates these difficulties.

To find the proper expansion parameter we note that the size of a jump ξ is properly expressed in an *extensive* quantity, while the variable y occurring in

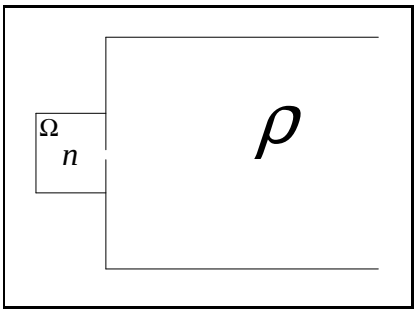

Figure 1: **A small volume Ω attached to an infinite reservoir.**

$W(y;\xi)$ expresses the dependence of the probability on the over-all state of the system and is therefore properly expressed in terms of an *intensive* variable[2]. This suggests that, if Ω is the size of the system, one may make the dependence of W on Ω explicit by setting

$$W(y;\xi) = \Phi\left(\tfrac{y}{\Omega};\xi\right),$$

in such a way that Φ no longer depends implicitly on Ω. This turns out to be true.

As an example, consider an ideal gas in a virtually infinite reservoir with fixed density ρ. A smaller volume Ω is connected with it through a hole, see Figure 1.

The probability that in a given dt the number n of molecules in Ω increases by one unit is $\alpha\rho\,dt$. The probability that n decreases by one unit is $\alpha\frac{n}{\Omega}\,dt$. Hence

$$W(n;\xi) = \alpha\rho\,\delta_{\xi,1} + \alpha\frac{n}{\Omega}\,\delta_{\xi,-1}.$$

Clearly this depends on n only through the intensive variable $\frac{n}{\Omega}$ and no other Ω occurs. This turns out to be the general pattern, although minor modifications are sometimes required. For instance, additional factors Ω may appear in Φ, but they can be absorbed in the time unit. Or it may happen that Φ is not independent of Ω but consists of a power series in Ω^{-1}; this gives rise to additional terms but does not upset the general scheme. Finally it is not necessary that Ω has the physical meaning of a volume. In the case of Brownian motion, Ω is the mass of the particle, its momentum has the role of the extensive quantity, and its velocity is the intensive quantity [3].

[2] *Extensive* properties depend upon the size of the system, for example the mass and volume. *Intensive* properties are independent of the system size, for example the temperature and density. Doubling the system size does not double the temperature.

[3] No relation to the fact that Aristotle also regarded velocity as an intensity.

6 A paradigm

Rather than formulate the expansion method in general terms we demonstrate it on a simple example, which we borrow from the theory of population statistics. Let n be the number of individuals in a population (of bacteria or other organisms). Each individual produces offspring at a rate α and has a probability β per unit time to die. This leads to

$$\dot{n} = \alpha n - \beta n,$$

which is Malthuss law, assuming $\alpha > \beta$. In order to include the struggle for life, Verhulst added for each individual a death probability proportional to the number of individuals present. More precisely, this probability should be taken proportional to the *density* $\frac{n}{\Omega}$, where Ω is the size of the test tube, or the amount of food. This leads to the Malthus-Verhulst equation

$$\frac{dn}{dt} = (\alpha-\beta)\,n - \gamma\frac{n^2}{\Omega},$$

which fits surprisingly well [6]. This is the macroscopic, phenomenological equation or rate equation, and ignores fluctuations.

In order to describe fluctuations one has to introduce the probability $P(n,t)$ for having n individuals at time t. It obeys the master equation

$$\begin{aligned}\dot{P}(n,t) = {}& \alpha(n-1)P(n-1,t) - \alpha n P(n,t)\\ & +\beta(n+1)P(n+1,t) - \beta n P(n,t)\\ & +\tfrac{\gamma}{\Omega}\left\{(n+1)^2 P(n+1,t) - n^2 P(n,t)\right\}.\end{aligned}$$

The Markov character has been introduced by assuming that each individual has fixed probabilities for breeding or dying, regardless of age and gestation period.

The Kramers-Moyal expansion of the master equation is

$$\frac{\partial P(n,t)}{\partial t} =$$

$$= \sum_{\nu=1}^{\infty}\frac{1}{\nu!}\frac{\partial^\nu}{\partial n^\nu}\left\{(-1)^\nu\alpha n + \beta n + \frac{\gamma}{\Omega}n^2\right\}P(n,t).$$

Notice that α and β occur separately, no longer in the combination $\alpha-\beta$ alone. Thus the master equation involves more detailed information about the process than the macroscopic rate equations. This is generally true and explains why the observation of fluctuations may lead to new information, not obtainable from the

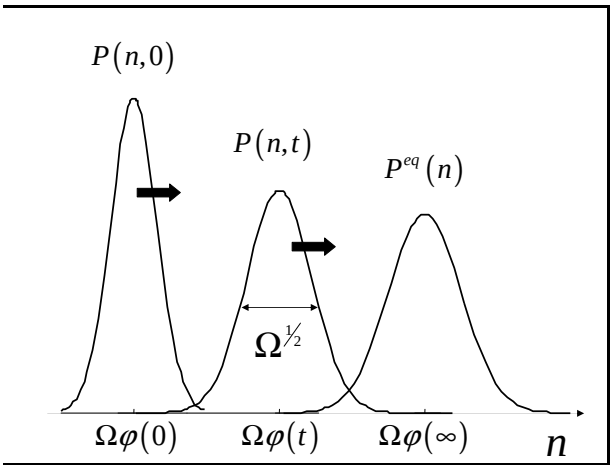

Figure 2: **Fluctuations about the macroscopic trajectory** $\Omega\varphi(t)$**.**

measurement of the rates. For our present purpose, however, it suffices to take $\beta = 0$; moreover, we change the time unit so as to have $\alpha = 1$, and the unit of Ω so that $\gamma = 1$.

7 The Ω-expansion

It is necessary to write the powers of Ω explicitly. We anticipate that n consists of a macroscopic part of order Ω and fluctuations around it of order $\Omega^{1/2}$. Accordingly we transform from the variable n to a new variable x by setting

$$n = \Omega\varphi(t) + \Omega^{1/2}x$$

where $\varphi(t)$ is a function yet to be determined. The rationale is that the probability distribution $P(n,t)$ is represented by a sharp peak, which slides bodily along the n-axis. Its width is of order $\Omega^{1/2}$ at all times, and its position $\Omega\varphi(t)$ is the macroscopically observed value of n (see Figure 2).

The transformation of the probability distribution reads

$$P(n,t) = P\left(\Omega\varphi(t) + \Omega^{1/2}x, t\right) \equiv \Pi(x,t)$$

Actually, there ought to be an additional factor $\Omega^{1/2}$ on the right in order that Π be normalized, but we omit it for convenience.

We write the Kramers-Moyal expansion in the new variables, and order the terms according to the powers of Ω:

$$\frac{\partial\Pi}{\partial t} - \frac{\partial\Pi}{\partial x}\Omega^{1/2}\varphi'(t) = -\Omega^{1/2}\left(\varphi - \varphi^2\right)\frac{\partial\Pi}{\partial x}$$

$$+ (-1+2\varphi)\frac{\partial}{\partial x}x\Pi + \frac{1}{2}\left(\varphi+\varphi^2\right)\frac{\partial^2\Pi}{\partial x^2} + O\left(\Omega^{-1/2}\right).$$

The largest terms are those involving $\Omega^{1/2}$. The fact that the derivative $\partial\Pi/\partial t$ is not one of them means that we are dealing with a singular perturbation problem. However, the two large terms can be caused to cancel by choosing

$$\varphi'(t) = \varphi - \varphi^2.$$

This is the macroscopic rate equation

$$\frac{dn}{dt} = n - \frac{n^2}{\Omega}.$$

Thus we have deduced the macroscopic law from the master equation; it is simply the equation that describes how the peak $P(n,t)$ slides along the n-axis. Note that the macroscopic equation may well be nonlinear, although the master equation is of course linear in $P(n,t)$.

Having disposed of the terms of order $\Omega^{1/2}$, we collect the terms of next order

$$\frac{\partial\Pi}{\partial t} = (-1+2\varphi)\frac{\partial}{\partial x}x\Pi + \frac{\varphi+\varphi^2}{2}\frac{\partial^2\Pi}{\partial x^2}.$$

This equation governs the fluctuations around the macroscopic value $\Omega\varphi(t)$ found in the previous step. It is similar to the equation for the Ornstein-Uhlenbeck process, but its coefficients depend on time. Yet it can be solved explicitly [7].

For $t = \infty$ the coefficients become constant and the equation reduces to

$$\frac{\partial\Pi}{\partial t} = \frac{\partial}{\partial x}x\Pi + \frac{\partial^2\Pi}{\partial x^2}.$$

This is the familiar (linear) Fokker-Planck equation for the equilibrium fluctuations, as used in the standard theory of Rayleigh, Einstein, Smoluchowski, and many others.

The macroscopic equation also has the particular solution $\varphi(t) = 0$. The fluctuations around it obey

$$\frac{\partial\Pi}{\partial t} = -\frac{\partial}{\partial x}x\Pi.$$

This leads to $\langle x\rangle_t = \langle x\rangle_0 e^t$, which demonstrates that the solution $\varphi = 0$ is unstable. Thus the non-existence of unicorns is an unstable situation.

8 Some critical remarks

We have found that the fluctuations are governed in lowest order by a second-order equation, whose first coefficient is linear in x and whose second coefficient is a constant. It turns out that the next order adds a quadratic term in the first coefficient, a linear term in the second coefficient, *and also a third-order derivative* with a constant coefficient; and so on [8]. It follows that *it is inconsistent to use the nonlinear Fokker-Planck equation without including at the same time higher derivatives*. To put it differently, the single parameter Ω governs both the validity of the Fokker-Planck approximation and the influence of the nonlinearity on the fluctuations; it is inconsistent to add nonlinear terms without improving on the Fokker-Planck approximation as well. In particular, it is incorrect to conclude from the first two terms of the Kramers-Moyal series that the equilibrium distribution is

$$P^{eq}(y) = \frac{\text{const.}}{\alpha_2(y)} \exp\left[\int_0^y \frac{2\alpha_1(y')}{\alpha_2(y')} dy'\right].$$

Another way of describing fluctuations was originated by Langevin, who wrote for the velocity v of a Brownian particle

$$\dot{v} = -\gamma v + \eta(t).$$

γ is the friction coefficient and $\eta(t)$ is assumed to be a rapidly fluctuating force, due to the random collisions of the gas molecules. The rapidly fluctuating and irregular character of $\eta(t)$ is taken into account by treating $\eta(t)$ as a random process having the properties

$$\langle \eta(t) \rangle = 0 \qquad \langle \eta(t_1)\eta(t_2) \rangle = C\delta(t_1 - t_2)$$

Of course this does not yet fully specify the stochastic process; for this reason one sometimes completes the definition by stipulating that $\eta(t)$ is a Gaussian process, but that is an *ad hoc* assumption, and fortunately superfluous for deriving the properties of the Brownian particle.

The success of Langevins treatment of Brownian motion has given rise to a popular idea that, for every macroscopically known phenomenon in nature, fluctuations can be taken into account simply by supplementing the macroscopic rate equation by a Langevin term, even though its physical meaning may be obscure. Inasmuch as the present Ω-expansion provides a systematic treatment of the fluctuations, it is now possible to verify this idea *ex post facto*. Our paradigm will serve us to show that the Langevin assumption cannot be correct in the case that the macroscopic equation is nonlinear.

Following the popular idea we write the Langevin equation

$$\frac{dn}{dt} = n - \frac{n^2}{\Omega} + \eta(t).$$

Thus $\langle \eta(t) \rangle$ can be found from

$$\Omega\varphi' + \Omega^{1/2}\frac{d}{dt}\langle x \rangle = \Omega\varphi + \Omega^{1/2}\langle x \rangle - \frac{\left\langle \left(\Omega\varphi + \Omega^{1/2}x\right)^2 \right\rangle}{\Omega} + \langle \eta(t) \rangle$$

On substituting the results of section 7 this reduces to

$$\langle \eta(t) \rangle = \langle x^2 \rangle$$

The fact that $\langle \eta(t) \rangle$ does not vanish may not yet be so bad, but the fact that it depends on the solution itself is fatal to the Langevin approach. Even if one tries to save the Langevin assumption by noting that $\langle \eta(t) \rangle$ has one factor Ω less than the other terms in the above Langevin equation, the same evil reappears when computing C.

9 Stochastic differential equations

This name denotes differential equations in which one or more of the coefficients are stochastic variables. Thus a stochastic differential equation stands for an ensemble of differential equations, each with some values for the coefficients. The solution of a stochastic differential equation with given initial condition (which may be stochastic too) is a stochastic function. The problem is to derive its stochastic properties from those of the coefficients of the equation.

An example is Langevins differential equation for the Brownian particle, in which the coefficient $\eta(t)$ is stochastic. The solution is well known, and can rigorously be proved to be a Wiener process when $\eta(t)$ is assumed to be Gaussian. Note that, whereas the value of η at two different times are uncorrelated, this is not true for the Wiener process; thus the solution of the equation involves more memory than its stochastic coefficient. This is a general feature of stochastic differential equations, since the solutions are obtained by integrations.

Other examples are: the current in a an electric network due to a noisy source ([9],[10]), the response of a control system to random perturbations [11], line broadening [12], and the effect on a photoconductor of correlations in the incoming photon beam [13]. Examples of stochastic *partial* differential equations are: the Brownian motion of a string or drumhead [14] and the propagation of sound waves or radio waves through a turbulent medium [15]. The twinkling of stars and the propagation of ultrasound in the ocean have been treated in the approximation of geometrical optics and belong therefore to stochastic ordinary differential equations [16].

A rigorous solution of a stochastic equation consists of two steps: first solve the equation for y, with each possible value of the coefficients, and subsequently average the resulting y and products thereof. This can be carried out explicitly only in rare cases; one often has to be satisfied with an approximate solution. A much bolder approach to problems that cannot be solved rigorously is the following.

First take an average of the equation as it stands. In the Langevin case this yields a closed equation for $\langle y(t)\rangle$, but in general averages of products of higher powers of y and the coefficients will arise. Next, guided by physical insight and mathematical needs, assume that some of these averages of products may be replaced with the product of the averages, for example $\langle y^2\rangle \to \langle y\rangle^2$. In this way one arrives at a closed equation for $\langle y(t)\rangle$, whose solution hopefully approximates the actual $\langle y(t)\rangle$. Boltzmann's famous *Stosszahlansatz* and later forms of *molecular chaos* or *random phase approximations* are assumptions of that ilk[4]. However, we shall not use such methods. Nor are we concerned with stochastic boundary value problems and the probability distribution of their eigenvalues ([18],[19]).

10 First example

Consider the first-order linear differential equation

$$\dot{y} = ivy, \quad y(0) = 1.$$

Let v be a random real constant with probability density $\varphi(v)$. This example describes a set of harmonic oscillators with different frequencies. For each v the solution is e^{-ivt}, and hence

$$\langle y(t)\rangle = \left\langle e^{-ivt}\right\rangle = \int_{-\infty}^{\infty} \varphi(v)\, e^{-ivt} dv.$$

It is instructive to look at a few special cases; in the following list v_0 and γ are fixed parameters.

1. Lorentz:

$$\varphi(v) = \tfrac{\gamma/\pi}{\gamma^2+(v-v_0)^2}$$
$$\langle y\rangle = e^{-iv_0 t-\gamma t}$$

2. Gaussian:

$$\varphi(v) = (2\pi\gamma)^{1/2}\, e^{-(v-v_0)^2/2\gamma}$$
$$\langle y\rangle = e^{-iv_0 t-\gamma t^2}$$

3. Laplace:

$$\varphi(v) = \tfrac{1}{2}\gamma e^{-\gamma|v-v_0|}$$
$$\langle y\rangle = e^{-iv_0 t}\tfrac{\gamma^2}{\gamma^2+t^2}.$$

In each case the averaged solution tends to zero as $t\to\infty$.This damping is due to the fact that the harmonic oscillators of the ensemble gradually lose the phase coherence they had at $t=0$. In plasma physics this is called phase mixing, in mathematics Riemann-Lebesque theorem. The modulus of y is not subject to phase mixing and does not tend to zero; in fact,

$$|y(t)| = 1, \text{ hence } \left\langle |y(t)|^2\right\rangle = 1.$$

The form of the damping factor is determined by φ. Only for one particular φ does it have the form that corresponds to a complex frequency. It is therefore incorrect to make the Ansatz $\langle y(t)\rangle = e^{-i\omega t}$ and search for eigenfrequencies ω [20]. The problem is *not* invariant for time translations owing to the fixed initial time.

Note that $\langle y(t)\rangle$ is identical with the characteristic function $\chi(t)$ of the distribution $\varphi(v)$. Hence under rather weak restrictions it can have almost any reasonable dependence on t [21]. The fact that $\langle y(t)\rangle = \chi(t)$ suggests the use of the cumulant expansion

$$\langle y(t)\rangle = \left\langle e^{-ivt}\right\rangle = \exp\left[\sum_{m=1}^{\infty} \frac{(-it)^m}{m!}\overline{v^m}\right],$$

[4] J. B. Keller [16] has called this the dishonest method, but that term seems too disparaging for a method which has led to the greatest achievements of statistical mechanics. I propose to restrict this qualification to work in which such an assumption is made but concealed. According to this definition linear response theory should be called dishonest [17].

where $\overline{v^m}$ stands for the m^{th}-cumulant. We shall make ample use of cumulant expansions, but one has to bear in mind that not all probability distributions have one. For instance, the above formula suggests that there exists *no* φ for which $\langle y(t)\rangle$ has the form $e^{-iv_0t-\gamma t}$; but we have seen that that conclusion is wrong.

To characterize the stochastic process y more fully we compute the higher moments,

$$\langle y(t_1)...y(t_n)\,y^*(s_1)...y^*(s_m)\rangle = \\ = \left\langle e^{-iv(t_1+...+t_n-s_1-...-s_m)}\right\rangle \\ = \chi(t_1+...+t_n-s_1-...-s_m).$$

Thus we have succeeded in this simple instance in providing a complete solution of the problem of deriving the stochastic properties of $y(t)$ from those of v. The solution may be written in a condensed form by means of a characteristic functional $\mathrm{X}[\vartheta]$ involving an auxiliary dummy function $\vartheta(t)$,

$$\mathrm{X}[\vartheta] \equiv \left\langle \exp\left[i\int\left\{\vartheta(t)\,y(t)+\vartheta^*(t)\,y^*(t)\right\}dt\right]\right\rangle \\ = \left\langle \exp\left[i\int\left\{\vartheta(t)\,e^{-ivt}+\vartheta^*(t)\,e^{ivt}\right\}dt\right]\right\rangle.$$

11 Second example

Take the same equation with the same initial condition, but let v be a stochastic function of time. Then

$$\langle y(t)\rangle = \left\langle \exp\left[-i\int_0^t v(t')\,dt'\right]\right\rangle.$$

This is the characteristic function belonging to the stochastic variable represented by the integral. It can be written in the form of a cumulant expansion

$$\langle y(t)\rangle = \\ = \exp\left[\sum_{m=1}^{\infty}\frac{(-i)^m}{m!}\int_0^t \overline{v(t_1)\,v(t_2)...v(t_m)}\,dt_1dt_2...dt_m\right],$$

where the cumulants of $v(t)$ are given by the usual cluster expansion [22]

$$\langle v(t_1)\rangle = \overline{v(t_1)}$$

$$\langle v(t_1)\,v(t_2)\rangle = \overline{v(t_1)}\;\overline{v(t_2)} + \overline{v(t_1)\,v(t_2)}$$

$$\langle v(t_1)\,v(t_2)\,v(t_3)\rangle = \overline{v(t_1)}\;\overline{v(t_2)}\;\overline{v(t_3)} + \overline{v(t_1)}\;\overline{v(t_2)\,v(t_3)} \\ +...+\overline{v(t_1)\,v(t_2)\,v(t_3)}\ \text{etc.}$$

The following specialization of this general formula is often used. Suppose $v(t) = v_0 + \varepsilon v_1(t)$ with fixed constant v_0 and random $v_1(t)$ having $\langle v_1(t)\rangle = 0$. Then the first cumulant is simply equal to v_0, and

$$\langle y(t)\rangle = e^{-iv_0t}\times \\ \times\exp\left[\frac{-\varepsilon^2}{2}\int_0^t\int\langle v_1(t_1)\,v_1(t_2)\rangle\,dt_1dt_2+...\right].$$

Note that the expansion in ε occurs in the exponent, which greatly diminishes the danger of secular terms.

If the functions $v_1(t)$ in the ensemble are all, or almost all, so slowly varying that they are practically constant during the time t we are interested in, the present case reduces to that of the previous section. Let us consider the opposite case of rapidly varying $v_1(t)$. That describes an harmonic oscillator whose basic frequency v_0 is perturbed by a rapidly varying influence, for instance an atom undergoing collisions while emitting light.

Accordingly we suppose that the correlation between $v_1(t_1)$ and $v_1(t_2)$ depends on $t_1\ t_2$ alone and vanishes for $|t_1-t_2|>\tau$. Then for $t>>\tau$ the double integral is approximately equal to Dt with [23]

$$D = \int_{-\infty}^{\infty}\langle v_1(0)\,v_1(t)\rangle\,dt.$$

Hence in this case $\langle y(t)\rangle$ does have the form $e^{-iv_0t-\gamma t}$, corresponding to a complex frequency $v_0 - i\gamma \equiv v_0 - \frac{1}{2}i\varepsilon^2 D$. This shows that short lasting collisions give rise to line broadening with Lorentz profile, as was originally shown by Lorentz. The higher orders alternately add corrections to v_0 and γ, but that is only true up to some order n with $n\tau << t$.

A simple trick enables us to find higher moments of y. One has, for instance, taking $0\le t\le s$,

$$\langle y(t)\,y(s)\rangle = \left\langle \exp\left[-i\int_0^t v(t')\,dt' - i\int_0^s v(t')\,dt'\right]\right\rangle \\ = \left\langle \exp\left[-i\int_0^{\infty}\vartheta(t')\,v(t')\,dt'\right]\right\rangle.$$

Here we have defined $\vartheta(t')$ by setting

$$\vartheta(t') = \begin{cases} =2 & 0<t'<t \\ =1 & t<t'<s \\ =0 & s<t'. \end{cases}$$

It is again possible to use the cumulant expansion; if we also use an expression in ε for convenience, we find

$$\langle y(t)\, y(s)\rangle = e^{-iv_0 t - iv_0 s}\times$$

$$\times \exp\left[\frac{-\varepsilon^2}{2}\int_0^\infty\int_0^\infty \vartheta(t')\,\vartheta(t')\,\langle v_1(t')\, v_1(t'')\rangle\, dt'dt''\right].$$

Similar equations hold for the other moments; in particular

$$\langle y(t)\, y^*(s)\rangle = \left\langle \exp\left[-i\int_s^t v(t')\, dt'\right]\right\rangle$$

$$= \exp\left[\frac{-\varepsilon^2}{2}\int_s^t\int \langle v_1(t')\, v_1(t'')\rangle\, dt'dt''\right].$$

12 The general formula for $\langle y(t)\rangle$

As the final complication we suppose that y is a vector and take instead of v a random matrix $\mathbf{V}(t)$,

$$\dot{y} = -i\mathbf{V}(t)\, y, \qquad y(0) = a.$$

The initial value is some fixed vector a. For given $\mathbf{V}(t)$ the solution is

$$y(t) = \left\{1 - i\int_0^t \mathbf{V}(t_1)\, dt_1 - \right.$$

$$\left. -\int_0^t \mathbf{V}(t_1)\int_0^{t_1}\mathbf{V}(t_2)\, dt_2 dt_1 + ...\right\} a.$$

In each term the factors $\mathbf{V}(t)$ appear in chronological order with the latest time on the left. No error is made by writing the factors in arbitrary order, provided one adds a reminder that they should actually be read chronologically:

$$y(t) = \left\{1 - i\int_0^t \mathbf{V}(t_1)\, dt_1 - \right.$$

$$\left. -\frac{1}{2}\int_0^t\int_0^t \lceil \mathbf{V}(t_1)\,\mathbf{V}(t_2)\rceil\, dt_2 dt_1 + ...\right\} a.$$

Where $\lceil\cdot\rceil$ denotes the time-ordered product. More generally

$$y(t) = \left\{1 + \sum_{n=1}^\infty \frac{(-i)^n}{n!}\int_0^t dt_1 ... dt_n\, \lceil \mathbf{V}(t_1) ... \mathbf{V}(t_n)\rceil\right\} a$$

$$= \left[\exp\left\{-i\int_0^t \mathbf{V}(t')\, dt'\right\}\right] a.$$

The last expression is nothing but a symbolic way of writing the previous one: the time ordering has to be performed after expanding the exponential.

Having written the solution of the ordinary differential equation in a suitable way we take the average,

$$\langle y(t)\rangle = \left[\left\langle \exp\left\{-i\int_0^t \mathbf{V}(t')\, dt'\right\}\right\rangle\right] a$$

$$= \left[\exp\left\{\sum_{m=1}^\infty \frac{(-i)^m}{m!}\int_0^t dt_1 ... dt_m\, \overline{\mathbf{V}(t_1) ... \mathbf{V}(t_m)}\right\}\right] a,$$

where the bar denotes the cumulant. This is our general formula for the averaged solution of a stochastic linear differential equation. It is a condensed way of writing the result, but the actual evaluation of the right-hand side is not always easy [5]. It reduces to the result of the previous section when all factors $\mathbf{V}(t)$ commute, not merely at all times but also for all samples in the ensemble.

Let $\mathbf{V}(t) = \mathbf{V}_0 + \varepsilon\mathbf{V}_1(t)$ with fixed $\mathbf{V}_0$ and $\langle \mathbf{V}_1(t)\rangle = 0$. Then the first cumulant can again be written separately,

$$\langle y(t)\rangle = \left[\exp\left\{-i\int_0^t V_0(t_1)\, dt_1 + \right.\right.$$

$$\left.\left. + \sum_{m=2}^\infty \frac{(-i\varepsilon)^m}{m!}\int_0^t dt_1 ... dt_m\, \overline{\mathbf{V}(t_1) ... \mathbf{V}(t_m)}\right\}\right] a.$$

Although $\mathbf{V}_0$ does not depend on time it has been provided with a label t_1, which is required to assign to each factor $\mathbf{V}_0$ its proper place in the time-ordered product. An equivalent expression is [24]

$$\langle y(t)\rangle = e^{i\mathbf{V}_0 t}\left[\exp\left\{\sum_{m=2}^\infty \frac{(-i\varepsilon)^m}{m!}\times\right.\right.$$

$$\left.\left.\int_0^t dt_1 ... dt_m\, \overline{\mathbf{W}_1(t_1) ... \mathbf{W}_1(t_m)}\right\}\right] a,$$

where $\mathbf{W}_1(t)$ stands for the operator $\mathbf{V}_1(t)$ in the interaction representation,

$$\mathbf{W}_1(t) = e^{i\mathbf{V}_0 t}\mathbf{V}_1(t)\, e^{-i\mathbf{V}_0 t}.$$

[5] The equation is mentioned by R. Kubo [12] and a modified form of it was used by J. T. Ubbink [13].

Neglecting third and higher powers of ε in the exponent

$$\langle y(t)\rangle = \exp\left\{ -i\,\mathbf{V}_0 t - \varepsilon^2 \int_0^t e^{i\,\mathbf{V}_0 t_1} dt_1 \times \right.$$

$$\left. \times \int_0^{t_1} \left\langle \mathbf{V}_1(t_1)\, e^{-i\,\mathbf{V}_0(t_1-t_2)} \mathbf{V}_1(t_2) \right\rangle e^{-i\,\mathbf{V}_0 t_2} dt_2 \right\} a.$$

Actually the exponential should still be time-ordered, but that only affects higher orders in ε.

Suppose the correlation depends on t_1 - t_2 alone, so that

$$\left\langle \mathbf{V}_1(t_1)\, e^{-i\,\mathbf{V}_0(t_1-t_2)} \mathbf{V}_1(t_2) \right\rangle = \mathbf{R}(t_1 - t_2).$$

Moreover suppose that the correlation is practically zero for $|t_1 - t_2| > \tau$, so that for t ¿¿ τ

$$\int_0^{t_1} \mathbf{R}(t_1 - t_2)\, e^{-i\,\mathbf{V}_0 t_2} dt_2 = \left[\int_0^{t_1} \mathbf{R}(t')\, e^{-i\,\mathbf{V}_0 t'} dt' \right] e^{-i\,\mathbf{V}_0 t_1}$$

$$\approx \left[\int_0^{\infty} \mathbf{R}(t')\, e^{-i\,\mathbf{V}_0 t'} dt' \right] e^{-i\,\mathbf{V}_0 t_1} = \mathbf{D} e^{-i\,\mathbf{V}_0 t_1}.$$

And, if t is much larger than the periods contained in $\mathbf{V}_0$,

$$\int_0^t e^{i\,\mathbf{V}_0 t_1} dt_1 \mathbf{D} e^{-i\,\mathbf{V}_0 t_1} = t\,\widetilde{\mathbf{D}},$$

where $\widetilde{\mathbf{D}}$ is the diagonal part of $\mathbf{D}$ with respect to $\mathbf{V}_0$. Then

$$\langle y(t)\rangle = e^{-i\,\mathbf{V}_0 t - \varepsilon^2 \widetilde{\mathbf{D}}\, t} a,$$

which is a superposition of waves with complex frequencies.

13 Application to the transmission of waves

An electromagnetic wave propagating in the direction x through the troposphere may roughly be described by the Helmholtz equation

$$\frac{d^2 u}{dx^2} + k_0^2 n^2 u = 0.$$

k_0 is the wave number for the average density of air, and $n^2(x) = 1 + \varepsilon\mu(x)$ is the variation in the refractive index due to density fluctuations. In practice, $\varepsilon\mu \sim 10^{-6}$; the fluctuations extend over some 50 meters and they vary slowly compared to the passage time of a radio signal. The wave length is supposed to be small compared to them, because derivatives of n have been neglected in the equation. This approximation amounts to taking into account the local variations of wave length and amplitude, but not reflection (WKB approximation). Hence one may consider separately the wave going to the right; putting $du/dx + ik_0 nu = y$ one has in the same approximation

$$\frac{dy}{dx} = ik_0 n(x)\, y.$$

One is interested in the effect of many signals transmitted over a fixed distance x much longer than 50 m. Hence $\mu(x)$ is a stochastic function and the problem is the type treated in section (ref 11), except that x is substituted for t. The mean amplitude is attenuated by a factor $e^{-\gamma x}$ with

$$\gamma = \frac{1}{2}\varepsilon^2 k_0^2 \int_{-\infty}^{\infty} \langle \mu(0)\, \mu(x')\rangle\, dx'.$$

The intensity of the signal decreases proportionally to $e^{-2\gamma x}$, the increase in noise is $1 - e^{-2\gamma x}$.

Another problem is the passage of a wave through a space with randomly located scatterers. We take one dimension and y_1,y_2 for the amplitudes of the waves going to the right and to the left. The scatterers are located at random positions ξ_s and are otherwise identical. Each of them is represented by a two-by-two matrix $\mathbf{A}$ such that

$$y(\xi_s + 0) = \mathbf{A}\, y(\xi_s - 0).$$

When we set $\mathbf{A} = 1 - i\varepsilon\mathbf{B}$ the equation for y is

$$\frac{dy}{dx} = \begin{pmatrix} ik_0 & 0 \\ 0 & -ik_0 \end{pmatrix} y - i\varepsilon \sum_s \delta(x - \xi_s)\, \mathbf{B} y.$$

It can be verified that $\mathbf{B}$ is real when the scatterer is elastic and symmetric. The operator on the right has the form $-i\,\mathbf{V} = -i\,\mathbf{V}_0 - i\varepsilon\mathbf{V}_1$, where $\mathbf{V}_0$ is constant and $\mathbf{V}_1(x)$ is random and depends on x. Hence the general formula says

$$\langle y(x)\rangle = \left[\exp\left\{ -i \int_0^x \left[\mathbf{V}_0(x') + \varepsilon \mathbf{V}_1(x') \right] dx' + \right.\right.$$

$$\left.\left. + \sum_{m=2}^{\infty} \frac{(-i\varepsilon)^m}{m!} \int_0^x \overline{\mathbf{V}_1(x_1) \ldots \mathbf{V}_1(x_m)}\, dx_1 \ldots dx_m \right\} \right].$$

In order to compute the cumulants we have to employ a suitable way of describing the random location of the scatterers (ref 9). Let $f_n(x_1, ..., x_n)$ be the probability density for finding a scatterer at each of the locations x_1, ..., x_n, regardless of all other scatterers. Then, if all x_1, ..., x_n are different,

$$f_n(x_1, ..., x_n) = \left\langle \sum_{s_1,...,s_n} \delta(x_1 - \xi_{s_1}) ... \delta(x_1 - \xi_{s_1}) \right\rangle.$$

Subsequently the correlation functions g_n are defined by the usual cluster expansion

$$f_1(x_1) = g_1(x_1)$$

$$f_2(x_1, x_2) = g_1(x_1) g_1(x_1) + g_2(x_1, x_2), \quad etc.$$

We suppose that the distribution is homogeneous in space, so that f_1=g_1=*const.* and f_2 and g_2 only depend on x_1 - x_2.

Returning to the general formula we first note

$$\mathbf{V}_0 + \varepsilon\overline{\mathbf{V}_1(x)} = \mathbf{V}_0 + \varepsilon\langle\mathbf{V}_1(x)\rangle = \mathbf{V}_0 + \varepsilon f_1 \mathbf{B}.$$

Call this constant matrix $\mathbf{W}_0$ and define the interaction representation by

$$\mathbf{W}_1(x) = e^{i\mathbf{W}_0 x}\mathbf{V}_1(x)\, e^{-i\mathbf{W}_0 x}.$$

The general formula may then be written to second order in ε

$$\langle y(x)\rangle = e^{-i(\mathbf{V}_0 + \varepsilon f_1 \mathbf{B})x} \times$$

$$\times \exp\left\{ -\varepsilon^2 \int_0^x \int_0^{x_1} dx_1 dx_2 \overline{\mathbf{W}_1(x_1)\,\mathbf{W}_1(x_1)} \right\} a.$$

The term $\varepsilon f_1 \mathbf{B}$ is the averaged effect of single scatterers, while the double integral describes the additional effect of pairs. With some algebra one finds for it

$$\int_0^x \int_0^{x_1} dx_1 dx_2 g_2(x_1, x_2)\, e^{i\mathbf{W}_0 x_1} \mathbf{B} e^{i\mathbf{W}_0 (x_1 - x_2)} \times$$

$$\times \mathbf{B} e^{i\mathbf{W}_0 x_2} =$$

$$= \int_0^x e^{i\mathbf{W}_0 x_1} \left\{ \int_0^{x_1} g_2(x')\, \mathbf{B} e^{-i\mathbf{W}_0 x'} \mathbf{B} e^{i\mathbf{W}_0 x'} dx' \right\} \times$$

$$\times e^{-i\mathbf{W}_0 x_1} dx_1$$

It is by a similar calculation that Ubbink [13] managed to solve the problem of a photoconductor subject to randomly incident photons.

Various limiting cases of this formula may be recognized. First suppose that the positions ξ_s are uncorrelated; then g_2=0 and the total effect of the scatterers is the sum of their individual effects. Next suppose that the range of $g_2(x)$ is much less than the period of $e^{-i\mathbf{W}_0 x}$; that is, the correlation length is small compared to the wavelength. Then the integral reduces to

$$\left[\int_0^x e^{i\mathbf{W}_0 x_1} \mathbf{B}^2 e^{-i\mathbf{W}_0 x_1} dx_1 \right] \left[\int_0^\infty g_2(x')\, dx' \right].$$

This represents the *coherent* scattering by neighbouring pairs of scatterers. For large x this is proportional to $x\widetilde{\mathbf{B}^2}$ where the tilde denotes the diagonal part of $\mathbf{B}^2$ with respect to $\mathbf{W}_0$. Finally let g_2 have a long range and vary smoothly compared to wavelength. The the integral reduces to incoherent scattering by pairs,

$$x\widetilde{\mathbf{B}^2} \int_0^\infty g_2(x')\, dx'.$$

In all these cases $\langle y(x)\rangle$ depends exponentially on x owing to the assumption that x is much larger than any of the other lengths involved in the problem.

References

[1] A. N. Kolmogorov. *Foundations of the theory of probability.* Chelsea, New York, 1950.

[2] A. Khinchin. *Asymptotische Gesetze der Wahrscheinlichkeitsrechnung.* Chelsea, New York, 1948.

[3] A. T. Bharucha-Reid. *Elements of the Theory of Markov Processes and Their Applications.* Dover, 1997.

[4] I. I. Gichman and A. W. Skorochod. *Stochastische Differentialgleichungen.* Akademie-Verlag, Berlin, 1971.

[5] N. G. van Kampen. Fluctuations in nonlinear systems. In R. E. Burgess, editor, *Fluctuation Phenomena in Solids.* Academic Press, New York, 1965.

[6] A. J. Lotka. *Elements of mathematical biology.* Dover, 1956.

[7] S. Chandrasekhar. Stochastic problems in physics and astronomy. *Review of Modern Physics*, 15:1–89, 1943.

[8] N. G. van Kampen. A power series expansion of the master equation. *Canadian Journal of Physics*, 39:551, 1961.

[9] W. B. Davenport Jr. and W. L. Root. *An introduction to the theory of random signals and noise*. McGraw-Hill, New York, 1958.

[10] R. L. Stratonovic. *Topics in the theory of random noise, I and II*. Gordon and Breach, New York, 1963,1967.

[11] W. M. Wonham. Random differential equations in control theory. In A. T. Barucha-Reid, editor, *Probabilistic methods in applied mathematics*. Academic Press, New York, 1970.

[12] R. Kubo. Stochastic Liouville equations. *Journal of Mathematical Physics*, 4:174–183, 1963.

[13] J. T. Ubbink. The relation between the correlations of the incident photons and the noise in a photoconductor. *Physica*, 52:253–278, 1971.

[14] G. A. van Lear and G. E. Uhlenbeck. The brownian motion of strings and elastic rods. *Physical Review*, 38:1583–1598, 1931.

[15] V. I. Tatarski. *Wave propagation in turbulent media*. McGraw-Hill, New York, 1961.

[16] J. B. Keller. Wave propagation in random media. In *Proceedings of the Symposium on Applied Mathematics - Providence, R. I.*, volume 13, page 145. American Mathematics Association, 1962.

[17] N. G. van Kampen. Case against linear response theory. *Physica Norvegica*, 5:279, 1971.

[18] W. E. Boyce. Random eigenvalue problems. In A. T. Barucha-Reid, editor, *Probabilistic methods in applied mathematics*, volume 1. Academic Press, New York, 1968.

[19] C. E. Porter, editor. *Statistical theory of spectra*. Academic Press, New York, 1965.

[20] P. Chow and J. B. Keller. Wave propagation in a random lattice - 1. *Journal of Mathematical Physics*, 13:1404–1411, 1972.

[21] E. Lukacs. *Characteristic functions*. Griffin, London, 1960.

[22] G. E. Uhlenbeck and G. W. Ford. *Lectures in Statistical Mechanics*. American Mathematical Society, 1963.

[23] G. E. Uhlenbeck and L. S. Ornstein. On the theory of the brownian motion. *Physical Review*, 36:823–841, 1930.

[24] R. P. Feynman. An operator calculus having applications in quantum electrodynamics. *Physical Review*, 84:108–128, 1951.