A simple way of introducing stochastic differential equations

A M Basharov

Moscow Institute of Physics and Technology, 141700 Dolgoprudny, Russia
National Research Centre «Kurchatov Institute», 123182 Moscow, Russia

E-mail: basharov@gmail.com; http://basharov.me

Abstract. The notion of the Ito increment and the stochastic differential equation of the non-Wiener type were introduced using the simple "natural" property of counting process. The properties of the stochastic differential and integral were demonstrated and clarified in a simple and original way.

1. Introduction. SDE paradigm

One of the primary methods for the world knowledge is the research into the system response to small (and not small) perturbations, in particular, the study of the system return to equilibrium or quasi-equilibrium state after its removal from one of this or other condition of the kind. These are known as relaxation processes, and many problems of physics affect the relaxation of various kinds of excitation in the atomic, photonic, phonon and other systems. Similar processes can be found not only in physical, chemical, biological and social systems, but also in such seemingly distant sections as models of financial markets and others.

In the study of the system response to small (and not small) perturbations and/or parameter variations virtually all familiar concepts keep emerging. Indeed, first it is necessary to parameterize the system under study and to identify independent and dependent parameters, such as $y = y(x)$. Then we consider their variations or increments Δy and Δx. In the case of small increments, they appear to be related by direct proportional dependence

$$\Delta y(x) \equiv y(x + \Delta x) - y(x) \cong C_1(x)\Delta x,$$

and the value $C_1(x)$ is the characteristics of the studied system which can be defined experimentally and/or which is obedient to some laws. So, in fact, there arises a concept of derivative since $C_1(x) = y'(x)$ for a «good» function $y(x)$. As a result, the basic laws are expressed by differential equations. For example, the position of the material point of mass m at every point of time t on the axis x is determined as the coordinate x. Therefore we have the parameter dependence $x = x(t)$. Parameter increments define concepts such as velocity v_x and acceleration a_x (projections of velocity and acceleration on the axis x), $\Delta x(t) \cong v_x(t)\Delta t$, $\Delta v_x(t) \cong a_x(t)\Delta t$, and acceleration appears to be related to forces acting on the particle by Newton's second law of motion

$$ma_x = \sum F_x.$$

Assuming that among the forces acting on the particle there can be either deterministic F or random forces $X(t)$, we obtain the differential equation with random values

$$m\ddot{x}(t) = F + X(t) \,. \tag{1}$$

To characterize a random force, we note that when a particle (material point) moves across the environment there arise occasional collisions with environment particles. Let all the environment particles be divided into groups numbered by the index i, such groups that for each collision of the material point concerned with the particles of the i-th group, a material point obtains a value increment of momentum δ_k. Then it is convenient to rewrite the Newton's equation with a random force in the form of

$$\Delta p_x(t) = F\Delta t + \sum_k \delta_k \Delta N(\delta_k, t) \,,$$

where p_x is the momentum of a material point, and $\Delta N(\delta_k, t)$ is the number of collisions of a material point with the k-th group particles of environment per time Δt. Value $\Delta N(\delta_k, t)$ is the increment of some integer random value $N(\delta_k, t)$: $\Delta N(t) = N(\delta_k, t + \Delta t) - N(\delta_k, t)$.

Another example is provided by any electrical circuit. For example, if we consider the definition of electric current, then the charge ΔQ passing over a time Δt across the specified section of the conductor is $\Delta Q(t) = I(t)\Delta t$, where $I(t)$ is the electric current intensity. However, given that the charge is transferred by the elementary carriers of charge q_0, then

$$\Delta Q(t) = q_0 \Delta N(t) \,, \tag{2}$$

where $\Delta N(t)$ is the number of elementary charge carriers passing over a time Δt across the specified section of the conductor.

If the flow of charged particles is quite rare, for example, when driving the elementary charge carriers in electron tubes and the current pulse caused by the charge hitting the anode is very short (small inductance), the equation for the current can be conveniently written as

$$\Delta I(t) = -\alpha I(t)\Delta t + \beta \Delta N(t) \,, \tag{3}$$

where $\Delta N(t) = \sum_i \Delta N_i(t)$. If the time points of charge hitting the anode are designated as values t_k, the equation for the current can also be written as

$$\dot{I}(t) = -\alpha I(t) + \beta \sum_k \delta(t - t_k) \,. \tag{4}$$

The constant α is related to the constant β by the ratio $-\alpha q_0 + \beta = 0$ since before and some time after the charge hitting the anode, the current caused by one elementary charge should be equal to zero

$$\sum \Delta I(t) = -\alpha q_0 + \beta \,.$$

The ratio $-\alpha q_0 + \beta = 0$ remains valid for a more intense flux of electrons if the electrons enter the anode and affect it independently. As an example of equations (1)-(4) we will consider the features of stochastic differential equations (SDEs) at the elementary level and identify the basic concepts arising in this case.

2. Peculiarities of SDEs and stochastic integrals. The Ito increment algebra

We will discuss main differences of random terms of SDEs and stochastic integrals from the ordinary differential equations and the Riemann integrals

In the examples given above a random term was expressed as some value $N(t)$ whose increment

$$\Delta N(t) = N(t + \Delta t) - N(t)$$

provided for an integer number of random events happening in a time interval Δt. At the same time, it is evident that the main incremental rule is fulfilled – the sum of increments during shorter successive intervals is equal to the increment during the entire interval:

$$\Delta N(t) = \sum_{i=1}^{M} \Delta N_i = N(t + \Delta t) - N(t).$$

Here the time interval $[t, t + \Delta t]$, $\Delta t > 0$ is divided by points $t = t_0 < t_1 < t_2 < ... < t_M = t + \Delta t$ into M small subintervals $\Delta t_i = t_i - t_{i-1}$ и $\Delta N_i = N(t_i) - N(t_{i-1})$. Thus, in partitioning the time interval into shorter small subintervals, it is possible to reach the situation when ΔN_i take only two values 0 and 1, which describes only two feasible options – whether some event has happened or not.

Thus, when choosing a sufficiently short interval of time Δt_i, we have an exact equality $\Delta N_i^2 = \Delta N_i$, because for shorter time intervals ΔN_i can take only the values 0 and 1, and $0^2 = 0$ and $1^2 = 1$. Such small time intervals and the corresponding increments of random process $N(t)$ are to be discussed as Ito increments dt and $dN(t)$, respectively. Also, the terms differential or the Ito differential are often used. However, unlike the conventional analysis by which the differential $df(x)$ of function $f(x)$ indicates only that it is a linear (in increment of the argument) part of the increment of the function

$$\Delta f(x) = df(x) + C_2(x)(\Delta x)^2 + C_3(x)(\Delta x)^3 + ..., \tag{5}$$

the Ito differential

$$dN(t) = N(t + dt) - N(t) \tag{6}$$

suggests that the time interval dt is that during this time only one event happens or does not happen, such as an electron hitting the anode. The possibility of simultaneous entrance of two electrons to the anode is neglected.

So, the random process $N(t)$ consisting in counting the number of the events happened over the time t, otherwise the reading process, takes integer values, with increment (6) of the random process satisfying algebra

$$dN(t)dN(t) = dN(t). \tag{7}$$

This algebra determines the main difference between SDEs and integrals containing a random process $N(t)$ and the conventional differential and integral calculus.

Recall that in terms of the function increment, the main point determining an ordinary differential and integral calculus is the fact that for small increments of the argument $|\Delta x| << 1$ we have the following relation

$$|\Delta x| >> |\Delta x|^2 >> |\Delta x|^3 >> ...,$$

so in the expansion of the increment $\Delta f(x)$ of a «good» function $f(x)$,

$$\Delta f(x) = C_1(x)\Delta x + C_2(x)(\Delta x)^2 + ...$$

at $|\Delta x| << 1$ only Δx linear term $\Delta f(x) \approx df(x)$ matters, which is referred to as differential regardless of the value of the argument increment Δx. Obviously, in the case of dependence of the "good" function f on $N(t)$ all the terms should be taken into account in increment Δf

$$\Delta f(N(t)) = f(N(t) + \Delta N(t)) - f(N(t)) = \tfrac{1}{1!} f'(N(t))\Delta N(t) + \tfrac{1}{2!} f''(N(t))(\Delta N(t))^2 + ... =$$
$$= [\tfrac{1}{1!} f'(N(t)) + \tfrac{1}{2!} f''(N(t)) + ...]\Delta N(t) = [f(N(t)+1) - f(N(t))]\Delta N(t). \tag{8}$$

Otherwise, while using increments

$$df(N(t)) = f(N(t) + dN(t)) - f(N(t)) = [f(N(t)+1) - f(N(t))]dN(t). \tag{9}$$

This difference can be seen from the formula for differentiating a function product. For example,

$$dN^2(t) = N^2(t+dt) - N^2(t) = (N(t) + dN(t))^2 - N^2(t) = 2N(t)dN(t) + (dN(t))^2 =$$
$$= 2N(t)dN(t) + dN(t) = (2N(t)+1)dN(t), \tag{10}$$

i.e. in contrast to the ordinary Leibniz formula, in differentiating the function product it is necessary to take into account increment products if these functions are random processes, or functions of random processes.

Another feature of random function $N(t)$ occurs in the concept of a stochastic integral, i.e. an integral over increments of a stochastic process. Let us discuss how to understand the integral

$$\int_{t_0}^{t} N(t')dN(t'). \tag{11}$$

It is natural to represent it as the limit of integral sums $\int_{t_0}^{t} N(t')dN(t') = \lim\limits_{\substack{M \to \infty \\ \max \Delta t_i \to 0}} \sum\limits_{i=1}^{M} N(t_i')\Delta N_i$, where

time interval $[t_0, t]$, $t > t_0$ is divided by points $t = t_0 < t_1 < t_2 < ... < t_M = t$ into M small subintervals, $\Delta N_i = N(t_i) - N(t_{i-1})$, and points $t_i' \in [t_{i-1}, t_i]$.

As is known from the probability theory there are several concepts of limit for random values. We will note a curious fact that is independent of the concept of limit for random values. We will discuss how point t_i' within the partition interval should be chosen. In calculating the integrand at the left-edge of partition intervals

$$t_i' = t_{i-1}, \; i = 1,...,M,$$

then the resulting limit to integral sums will be designated as

$$S_L = \lim\limits_{\substack{M \to \infty \\ \max \Delta t_i \to 0}} \sum\limits_{i=1}^{M} N(t_{i-1})\Delta N_i.$$

It is not hard to understand that this limit will be different from the limit of Riemann sums, when the integrand is evaluated at the right edge of partition intervals

$$S_R = \lim\limits_{\substack{M \to \infty \\ \max \Delta t_i \to 0}} \sum\limits_{i=1}^{M} N(t_i)\Delta N_i = \lim\limits_{\substack{M \to \infty \\ \max \Delta t_i \to 0}} \sum\limits_{i=1}^{M} (N(t_{i-1}) + \Delta N_i)\Delta N_i = S_L + \lim\limits_{\substack{M \to \infty \\ \max \Delta t_i \to 0}} \sum\limits_{i=1}^{M} \Delta N_i = S_L + N(t) - N(t_0).$$

Stochastic integrals defined by the value of the integrand at the left-edge of partition intervals, with account of the relevant definition of the limit to integral comprise a class of the so-called stochastic Ito integrals. The choice of the left-edge of partition interval is useful for the so-called non-anticipating

functions when a would-be behaviour of a random process is statistically independent of its current state. For such functions $F(t)$ the mean value of the stochastic Ito integral has a simple representation

$$< \int_{t_0}^{t} F(t')dN(t') >= \int_{t_0}^{t} < F(t') >< dN(t') > .$$

Selecting the Ito integral as a stochastic integral leads to the unconventional formulas of integration, with the differentiation formulas being consistent with integration, such as the value of (11) can be derived from (10) as

$$\int_{t_0}^{t} N(t')dN(t') = \tfrac{1}{2} \int_{t_0}^{t} dN^2(t') - \tfrac{1}{2} \int_{t_0}^{t} dN(t') = \tfrac{1}{2}(N^2(t) - N^2(t_0)) - \tfrac{1}{2}(N(t) - N(t_0)) .$$

On the other hand, considering the integral as a limit of Riemann sums with values of the integrand at the left-edge of partition intervals, we have

$$\int_{t_0}^{t} N(t')dN(t') = \lim_{\substack{M \to \infty \\ \max \Delta t_i \to 0}} \sum_{i=1}^{M} N(t_{i-1})\Delta N_i = \frac{1}{2} \lim_{\substack{M \to \infty \\ \max \Delta t_i \to 0}} \sum_{i=1}^{M} [(N(t_{i-1}) + \Delta N_i)^2 - N^2(t_{i-1}) - (\Delta N_i)^2] =$$

$$= \frac{1}{2} \lim_{\substack{M \to \infty \\ \max \Delta t_i \to 0}} \sum_{i=1}^{M} [(N(t_{i+1}))^2 - N^2(t_{i-1}) - (\Delta N_i)^2] = \frac{1}{2} \lim_{\substack{M \to \infty \\ \max \Delta t_i \to 0}} [N^2(t) - N^2(t_0) - (N(t) - N(t_0))] =$$

$$= \frac{1}{2}[N^2(t) - N^2(t_0) - (N(t) - N(t_0))].$$

This result does not depend on the concept of limit used. It demonstrates the consistency of algebraic relations (7) with the concept of stochastic Ito integral. Thus, stochastic differential equations and stochastic integrals (and related stochastic integral equations) are somewhat "two sides of the same coin". And the concept of stochastic continuity will allow separating random functions from discontinuous deterministic functions, which can demonstrate the specified features at certain partitions.

Without going too deep into the theory, it is natural to hypothesize the nature of the averaging of increment $dN(t)$. Being a good idealization, random process $N(t)$ should be in some sense homogeneous with the average value $dN(t)$ proportional to the time interval

$$< dN(t) >= \lambda dt . \tag{12}$$

Parameter λ is to be referred to as velocity or intensity of a random process $N(t)$.

Finally, if you give a standard definition for a limit for the partial sums of stochastic integral as the limit in the mean square, it is easy to prove that the integrals of non-anticipating functions take the following form $\int_{t_0}^{t} F(t')dN(t')dt' = 0$, which allows adding an important ratio to the increment algebra (7) and (12)

$$dN(t)dt = 0 . \tag{13}$$

Moreover, the requirement for statistical independence $dN(t)$ and $dN(t')$ in the case of non-overlapping time intervals dt and dt' leads to

$$< dN(t)dN(t') >=< dN(t) >< dN(t') > , \ t \neq t' . \tag{14}$$

Finally, SDEs (4) can be written in terms of the increment of the random process $N(t)$ (with algebra increments (7) (12) - (14)) in the form

$$dI(t) = -\alpha I(t)dt + \beta dN(t).$$ (15)

In terms of the algebra increments (7) (12) - (14) ordinary differential calculus is characterized by the algebra

$$dtdt = 0,$$

if value dt is meant to be a rather small increment, i.e. the Ito increment. Then the formula for the Ito increment of function $f(t)$ takes on the following meaning

$$df(t) = f(t+dt) - f(t) = f'(t)dt + \tfrac{1}{2!}f''(t)(dt)^2 + ... = f'(t)dt.$$

3. SDEs and kinetic equations

Equations (7) and (12)-(14) determine the so-called Poisson process that essentially defines the entire apparatus of the SDEs. Let us see how the SDEs are linked to kinetic equations of the system that underlie the dynamics of open systems.

Let $f(t)$ be a continuously differential function, and $P(n,t)$ is the probability of the fact that by the time t random value $N(t)$ has the meaning n. Then it is possible to introduce the average value

$< f(N(t)) >$ in the form $< f(N(t)) >= \sum\limits_{n=0}^{\infty} f(n)P(n,t)$. Let us consider equation (9):

$df(N(t)) = dN(t)\{f(N(t)+1) - f(N(t))\}$ and average it by means of $P(n,t)$,

$$d < f(N(t)) >= \lambda dt < f(N(t)+1) - f(N(t)) >, \quad \frac{d < f(N(t)) >}{dt} = \lambda < f(N(t)+1) - f(N(t)) >,$$

$$\sum\limits_{n=0}^{\infty} f(n)\frac{\partial P(n,t)}{\partial t} = \lambda \sum\limits_{n=0}^{\infty}\{f(n+1)P(n,t) - f(n)P(n,t)\} = \lambda \sum\limits_{n=1}^{\infty} f(n)P(n-1,t) - \lambda \sum\limits_{n=0}^{\infty} f(n)P(n,t).$$

Supposing $f(t)$ to be an arbitrary function with $f(0) = 0$, we derive the equation

$$\frac{\partial P(n,t)}{\partial t} = \lambda[P(n-1,t) - P(n,t)].$$ (16)

This equation is an example of the kinetic equation derived from the SDE.

The solution of equation (16) by means of the generative function gives rise to the so-called Poisson distribution

$$P(n,t) = e^{-\lambda t}\frac{\lambda^n t^n}{n!},$$ (17)

and value $N(t)$ is also called the Poisson process.

The average value of the number of electrons at the anode $\bar{n}$ and the variance of this number $\sigma^2 = < n^2 > - \bar{n}^2$ for the Poisson distribution prove to be equal to each other. These results are not difficult to obtain from differential algebra (7) and (12) - (14). Since

$$N(t) = \int\limits_{0}^{t} dN(t'),$$

calculation of the average value and the variance of value $N(t)$ can be presented in the following way:

$$< N(t) >= \int_0^t < dN(\tilde{t}) >= \int_0^t \lambda d\tilde{t} = \lambda t \,,$$

$$< [N(t) - \lambda t]^2 >=< [N(t)]^2 > -2\lambda t < N(t) > +(\lambda t)^2 =< [N(t)]^2 > -(\lambda t)^2 \,,$$

$$< [N(t)]^2 >=< \int_0^t dN(\tilde{t}) \int_0^t dN(\tilde{\tilde{t}}) >= \lim < \sum_i \Delta N_i \sum_j \Delta N_j >= \lim \{< \sum_i (\Delta N_i)^2 > +2 \sum_{i>j} < \Delta N_i \Delta N_j >\}$$

$$= \lim \{< \sum_i \Delta N_i > +2\lambda^2 \sum_{i>j} \Delta t_i \Delta t_j\} = \lambda t + \lim \{\lambda^2 \sum_i (\Delta t_i)^2 + 2\lambda^2 \sum_{i>j} \Delta t_i \Delta t_j - \lambda^2 \sum_i (\Delta t_i)^2\} = \lambda t + (\lambda t)^2 .$$

4. Conclusion

The proposed approach to introducing SDEs differs from the conventional ones (see, for example, [1-3]): one is immediately given an idea of non-Wiener type SDEs, which serve the basis for standard Wiener type SDEs derived by means of the central limit theorem [4]. Universality of stochastic processes underlying the SDEs is due to the central limit theorem and different time scales for deterministic and random terms. Their kinetic equations may contain fractional derivatives with respect to spatial variables, reflecting the self-similarity and fractal properties of the medium [4,5]. Moreover, the very random processes are the Markov processes. According to the approach considered above, non-Markov random processes with memory effect arise as the subordinated processes, and their kinetic equations contain terms with fractional time derivative [5]. However, the straightforward derivation of their kinetic equations similar to the ones considered in [4] for Markov processes is unknown to the author.

References

[1] Jacobs K 2010 *Stochastic Processes for Physicists* (Cambridge: CUP)
[2] Breuer H-P and Petruccione F 2002 *Theory of open quantum systems* (Oxford: OUP)
[3] Cont R and Tankov P 2004 *Financial modeling with jump processes* (London: CRC Press)
[4] Basharov A M 2013 *J. Phys.: Conf. Ser.* **478** 012011
[5] Basharov A M 2012 *Nanosystems: Physics, Chemistry, Mathematics* **3** 47