

ARTICLES

A Chemical Langevin Equation with Non-Gaussian Noise[†]Robert Zwanzig[‡]*Laboratory of Chemical Physics, National Institutes of Health, Bethesda, Maryland 20892-0520**Received: September 25, 2000; In Final Form: January 31, 2001*

The noise that appears in the Langevin equation for “well-stirred” chemical reactions is found by considering how one would do a “gedanken” computer simulation of the reaction. The noise is not Gaussian.

1. Introduction

Langevin equations are sometimes used to model the dynamics of complex processes. They separate the rate of change of any quantity into two parts, one systematic, and the other noisy. Typically they require the probability distribution of the rate of change $F(x)$ of some quantity x ,

$$\frac{dx_t}{dt} = F(x_t, t) \quad (1)$$

The main concern here is with chemical kinetics, where x is always an integer and changes discontinuously in time. Then it is more convenient to use the integrated form

$$x_{t+dt} = x_t + \int_0^{dt} F(x_{t+s}, t+s) ds \quad (2)$$

where dt is an infinitesimal time interval. The probability distribution of the integral is required. Conventionally, this distribution is assumed to be Gaussian, determined fully by its first and second moments.

Langevin equations for chemical reactions have been discussed recently.¹ Here we obtain the Langevin equation for a simple chemical reaction. While the following treatment can be made in considerable generality, as an illustration we consider the isomerization reaction



We use A and B for the actual number of molecules of each species. Note that because of conservation of molecules, $A + B = \text{constant}$, and the rate equation reduces to a single equation for A . In this example, the rates of the forward reaction (or U for up because A increases) and reverse reaction (or D for down because A decreases) are

$$k_2 B = U(A) \quad k_1 A = D(A) \quad (4)$$

The notation was chosen to account as well for more complex reactions. In a bimolecular reaction, for example, U and D can

be quadratic functions of A . The deterministic rate equation is

$$\frac{dA}{dt} = U(A) - D(A) \quad (5)$$

This rate equation does not contain any random term. But we know from past experience (as in the fluctuation–dissipation theorem) that in order to achieve a proper thermal equilibrium state, a fluctuating noise $\delta F(t)$ should be added,

$$\frac{dA}{dt} = F(A, t) = U(A) - D(A) + \delta F(t) \quad (6)$$

The deterministic equation has become a Langevin equation. But it still remains to specify the distribution of the noise. Normally, one knows its first and second moments, and one assumes that the distribution is a Gaussian with those two moments. If the noise is Gaussian and white, the probability distribution of A satisfies a Fokker–Planck equation, and the equilibrium distribution can be found by quadrature. In the present case, because of the linearity of the particular rate equation used here, Gaussian noise leads to a Gaussian equilibrium distribution. But we know (see the following section) that the correct distribution is Poisson and not Gaussian. They agree near equilibrium but disagree elsewhere. Often this disagreement is numerically unimportant, but in principle it is still there.

2. The Standard Chemical Master Equation

The derivation of a chemical Langevin equation follows closely the derivation of a chemical master equation.² The kinds of master equations used here are appropriate to “well-stirred” systems,¹ in which the chemical composition is the only slow variable. All other dynamical processes, for example diffusive or heat flow, are assumed to have come to equilibrium much faster than the composition. The master equation describes the dynamics of transitions between states having different compositions. (In ref 1 the transition rates are called “propensities”.) The fundamental quantities in formulating the master equation are probabilities of transitions. In a short enough time interval dt , only two events can occur; A can increase by 1, or decrease by 1, or else it remains unchanged. The probabilities of these events are

[†] Part of the special issue “Bruce Berne Festschrift”.

[‡] E-mail: zwanzig@sunder.niddk.nih.gov. Fax 301 496 0825.

$$\begin{aligned}\text{prob}(A \rightarrow A - 1 \text{ in } dt) &= D(A) dt \\ \text{prob}(A \rightarrow A + 1 \text{ in } dt) &= U(A) dt \\ \text{prob}(A \rightarrow A \text{ in } dt) &= (1 - U(A) - D(A)) dt\end{aligned}\quad (7)$$

The question of how small or large dt can be is a familiar one in nonequilibrium statistical mechanics, arising first in the derivation of the Pauli master equation. In the present case, the interval should be larger than the time required to establish the “well-stirred” state and larger than the duration of an elementary event, for example, a molecular collision, but it should be small enough that one can ignore changes in composition. We denote the probability of a particular composition by

$$P(A, t) = \text{prob}(A \text{ at time } t) \quad (8)$$

Then, when the limit of infinitesimal dt is considered, the master equation is

$$\frac{d}{dt}P(A, t) = (E - 1)D(A) P(A, t) + (E^{-1} - 1)U(A) P(A, t) \quad (9)$$

where E is the unit shift operator, $Ef(n) = f(n + 1)$, and E^{-1} is its inverse.

Its equilibrium solution is the Poisson distribution

$$P(A, eq) = \text{const} \times \frac{1}{A!B!} \left(\frac{k_2}{k_1}\right)^A \quad (10)$$

3. Deriving the Langevin Equation

Let us do a “gedanken” computer simulation of the reaction. (This means a simulation that is done in thought, but not necessarily by a real computer.) At time t , the number of A molecules is the integer A_t . A random number r is selected from a uniform distribution on the interval $(0, 1)$. The probabilities of the various events are accounted for by dividing the interval into three distinct regions. If r falls in the interval $(0, U(A_t) dt)$, then A_t is increased by 1. If r falls in the interval $(1 - D(A_t) dt, 1)$, then A_t is decreased by 1. If r falls in the remaining interval $(U(A_t) dt, 1 - D(A_t) dt)$, then A_t is unchanged. The value of r changes in successive time intervals and is not correlated with values in other intervals. A complete “gedanken” simulation consists of many repetitions of this process; since dt is infinitesimal, an infinite number of repetitions are required to cover a finite time interval.

A single step can be summarized conveniently by using the step function Θ , defined by

$$\begin{aligned}\Theta(x) &= 0 & x < 0 \\ &= 1 & x > 0\end{aligned}\quad (11)$$

Then the changes are given by

$$A_{t+dt} = A_t + \Theta(U(A_t) dt - r) - \Theta(r - 1 + D(A_t) dt) \quad (12)$$

This is a special case of eq 18 in ref 1 but restricted to infinitesimal dt . The single step result has the form of eq 2. It contains the integral of the time derivative of A over an infinitesimal interval dt ,

$$\int_0^{dt} ds \frac{dA_{t+s}}{ds} = A_{t+dt} - A_t \quad (13)$$

and gives an explicit form for the integral of the total force F acting on A ,

$$\int_0^{dt} ds F(t + s) = \Theta(U(A_t) dt - r) - \Theta(r - 1 + D(A_t) dt) \quad (14)$$

Note that this integral can take on only three values, 1 with probability $U dt$, -1 with probability $D dt$, and 0 with probability $1 - U dt - D dt$. This is a clear indication that the distribution of the total integrated force cannot be Gaussian. The average or first moment is the systematic part of the integrated force,

$$\begin{aligned}\langle \int_0^{dt} ds F(t + s) \rangle &= (+1) \times U(A_t) dt + (-1) \times D(A_t) dt \\ &= (U - D) dt\end{aligned}\quad (15)$$

The second cumulant, or the mean squared fluctuation, is

$$\langle (\int_0^{dt} ds \delta F(t + s))^2 \rangle = (U + D) dt - ((U - D) dt)^2 \quad (16)$$

In the limit of infinitesimal dt , the second term can be dropped. That the first term is proportional to dt indicates that the second cumulant of the unintegrated force is delta-correlated in time.

Recall that in solving any Langevin equation by computer, the first step is to select a time interval and then integrate. So the single step of our “gedanken” computer simulation may be regarded as the integration of a Langevin equation over an infinitesimal time interval. The total force F contains a systematic part $(U - D)$ and non-Gaussian fluctuations or noise depending on an “environmental” variable r . We can surely construct a Gaussian distribution with the same mean and second moment; that is what is done in ref 1. But it will lead to an incorrect equilibrium distribution. How much incorrectness is tolerable in practice is a value judgment that I will not make here. A rough analogy is using 3.1416 as a good approximation to π . This is often perfectly adequate. Nevertheless, π is a transcendental number, not a rational fraction. A mathematician finds this interesting but others might not.

In conclusion, in the same sense that a Fokker–Planck equation is associated with a Langevin equation with Gaussian white noise, a chemical master equation is associated with a chemical Langevin equation with non-Gaussian white noise.

References and Notes

- (1) Gillespie, D. T. *J. Chem. Phys.* **2000**, *113*, 297–306
- (2) McQuarrie, D. A. *J. Appl. Probab.* **1967**, *4*, 413.