ARTICLES

The Chemical Langevin and Fokker−Planck Equations for the Reversible Isomerization Reaction†

Daniel T. Gillespie‡

Dan T. Gillespie Consulting, 812 West Vicki Avenue, Ridgecrest, California 93555

Received: July 26, 2001; In Final Form: December 17, 2001

This paper uses the simple reversible isomerization reaction to illustrate and clarify the roles played in chemical kinetics by recently proposed forms for the chemical Langevin equation and chemical Fokker−Planck equation. It is shown that the stationary solution of the chemical Fokker−Planck equation for this model reaction provides, for most purposes, an excellent approximation to the stationary solution of the chemical master equation. It is also shown that, when allowance is made for the stipulated macroscopic nature of the time increment dt in the chemical Langevin equation, the changes in molecular population during dt predicted by that equation for this model reaction closely approximate the changes prescribed by the chemical master equation. The discussion highlights the role of the chemical Langevin equation as not only a potential computational aid but also a conceptual bridge between the stochastic chemical master equation and the traditional deterministic reaction rate equation.

1. Introduction

When molecules of a well-stirred mixture of N molecular species $\{S_1, ..., S_N\}$ interact through M chemical reaction channels $\{R_1, ..., R_M\}$, the molecular population vector $\mathbf{X}(t) \equiv (X_1(t), ..., X_N(t))$, where

$$X_i(t) \equiv \text{the } \textit{number} \text{ of } S_i \text{ molecules in the system at time } t \qquad (i = 1, ..., N) \quad (1)$$

changes *stochastically* because of the inherent randomness of molecular collisions. If the molecules are confined to a fixed volume and kept at constant temperature, straightforward kinetic theory arguments show that for each reaction channel R_j there is a function a_j such that[1]

$$a_j(\mathbf{x})\,\mathrm{d}t \equiv \text{the } \textit{probability}, \text{ given } \mathbf{X}(t) = \mathbf{x}, \text{ that one } R_j \text{ reaction will occur in the system in the next infinitesimal time interval } [t,t+dt) \quad (j = 1, ..., M) \quad (2)$$

This *propensity function* a_j, together with the *state-change vector* $\boldsymbol{\nu}_j \equiv (\nu_{j1}, ..., \nu_{jN})$ as defined by

$$\nu_{ji} \equiv \text{the change in the number of } S_i \text{ molecules produced by one } R_j \text{ reaction} \qquad (j = 1, ..., M;\ i = 1, ..., N) \quad (3)$$

completely characterizes reaction channel R_j. So, for example, if R_j is the reaction $S_1 + S_2 \rightarrow 2S_1$, then $\boldsymbol{\nu}_j = (+1, -1, 0, ..., 0)$ and $a_j(\mathbf{x}) = c_j x_1 x_2$ where c_j in this case is the conventional reaction rate constant k_j divided by the volume of the system.

Using only eqs 2 and 3 and the laws of probability theory, one can prove that the probability $P(\mathbf{x},t|\mathbf{x}_0,t_0)$, that $\mathbf{X}(t)$ will equal $\mathbf{x}$ given $\mathbf{X}(t_0) = \mathbf{x}_0$ for $t \geq t_0$, obeys the *chemical master equation* (CME):[1,2]

$$\frac{\partial}{\partial t}P(\mathbf{x},t|\mathbf{x}_0,t_0) = \sum_{j=1}^{M}[a_j(\mathbf{x} - \boldsymbol{\nu}_j)\,P(\mathbf{x} - \boldsymbol{\nu}_j,t|\mathbf{x}_0,t_0) - a_j(\mathbf{x})\,P(\mathbf{x},t|\mathbf{x}_0,t_0)] \quad (4)$$

Equations 2−4 imply that the system's state point $\mathbf{X}(t)$ performs a "random walk" on the integer lattice in the N-dimensional species population space; in mathematical terms, $\mathbf{X}(t)$ is a *jump Markov process*. But if the molecular population levels happen to be so large that the granularity of the integer lattice is not noticeable, the randomness in the trajectory of $\mathbf{X}(t)$ is often also not noticeable. In that case, the trajectory takes on the character of a *continuous, deterministic* process which is described by the set of ordinary differential equations

$$\frac{\mathrm{d}X_i(t)}{\mathrm{d}t} = \sum_{j=1}^{M}\nu_{ji}a_j(\mathbf{X}(t)) \qquad (i = 1, ..., N) \quad (5)$$

This is the well-known *reaction rate equation* (RRE) of traditional chemical kinetics, although expressed in terms of molecular populations instead of concentrations.

For well-stirred systems the CME (4) has a firm microphysical basis,[1] so for such systems it describes accurately the effects of molecular level randomness. In contrast, the RRE (5), which *also* requires the system to be well-stirred, is a more phenomenological equation; yet we know from experience that it describes most macroscale chemical systems quite well. Just how the CME (4) gets supplanted by the RRE (5) as a chemical

† Part of this work was performed while the author was a Staff Scientist in the Research Department of the Naval Air Warfare Center in China Lake, California.

‡ E-mail: GillespieDT@mailaps.org.

10.1021/jp0128832 CCC: $22.00
Published on Web 04/25/2002

system approaches the "thermodynamic limit" of infinite molecular populations has been the subject of much study and considerable debate for several decades. Recently this issue has become more than merely academic: Biochemists are finding that, inside a living cell, the relatively small molecular population levels of some key reactant enzymes can sometimes cause molecular level randomness to have a dramatic impact on cellular development.[3]

In an attempt to articulate more clearly the relation between the CME (4) and the RRE (5), this writer recently presented in ref 4 arguments showing that, under certain specific conditions, the *jump* Markov process defined by the CME (4) can be decently *approximated* by a *continuous* Markov process that satisfies the following *chemical Langevin equation* (CLE):[4]

$$X_i(t+\bar{\mathrm{d}}t) = X_i(t) + \sum_{j=1}^{M} \nu_{ji} a_j(\mathbf{X}(t))\,\bar{\mathrm{d}}t + \sum_{j=1}^{M} \nu_{ji} a_j^{1/2}(\mathbf{X}(t))\,N_j(t)\,(\bar{\mathrm{d}}t)^{1/2} \qquad (i = 1, ..., N) \quad (6)$$

Here, $N_1(t)$, ..., $N_M(t)$ are M statistically independent, temporally uncorrelated *normal* (or Gaussian) random variables with means 0 and variances 1; and $\bar{\mathrm{d}}t$ is a *positive macroscopically infinitesimal* time increment, which will be defined more precisely in a moment. The CLE (6) evidently tells us how, if we know the state of the system at time t, we can compute the state at the slightly later time $t + \bar{\mathrm{d}}t$; in principle, this is all we need to trace the time evolution of the system.[5]

The definition of the *macroscopic infinitesimal* $\bar{\mathrm{d}}t$ in eq 6 is important, because it defines the *special circumstances* under which that approximate equation is valid: The key requirement is that $\bar{\mathrm{d}}t$ be (i) *small* enough that none of the propensity functions a_j changes in a macroscopically noticeable way during $\bar{\mathrm{d}}t$, yet (ii) *large* enough that each reaction channel R_j fires many more times than once during $\bar{\mathrm{d}}t$. Only to the extent that the system admits a $\bar{\mathrm{d}}t$ satisfying *both* of these conditions will the CLE (6) decently approximate the time evolution of the process $\mathbf{X}(t)$ defined by the CME (4). In cases where it is not possible to find a $\bar{\mathrm{d}}t$ that satisfies both conditions i and ii, the CLE (6) will *not* be a reliable approximation to the CME (4).

This notion of a macroscopically infinitesimal time increment is not at all new in physics or chemistry. For example, the definition of electrical current as the ratio $\mathrm{d}Q \div \mathrm{d}t$, where $\mathrm{d}Q$ is the charge passing in infinitesimal time $\mathrm{d}t$, is meaningful *only* if $\mathrm{d}t$ is a *macroscopic* infinitesimal; because, if $\mathrm{d}t$ were allowed to be arbitrarily close to zero, as for a "true" infinitesimal, we would eventually observe "shot noise" as charge passes by in discrete chunks (on electrons): the ratio $\mathrm{d}Q \div \mathrm{d}t$ would not approach a well-defined limiting value. So, in conventional electrical circuit theory, it is always tacitly understood that the $\mathrm{d}t$ in the ratio $\mathrm{d}Q/\mathrm{d}t$ is *large* enough that very many electrons pass by in time $\mathrm{d}t$. But this "macroscopic" character of the infinitesimal $\mathrm{d}t$ in electrical circuit theory is rarely called out in a notationally explicit way, as we have done in eq 6; indeed, eq 6 appears in ref 4, where it was derived, *without* the overbar on the "d." The reason for the notational emphasis in this paper will become clear later.

Even the derivative in the RRE (5) presumes a macroscopic $\mathrm{d}t$: The change $[X_i(t+\mathrm{d}t) - X_i(t)]$ in the number of S_i molecules between times t and $t + \mathrm{d}t$ approaches zero with $\mathrm{d}t$ not continuously but rather through *discrete* values (and likewise for the concentration of S_i), a behavior that is really not allowed in a differentiable function. Therefore, if we were to write the RRE (5) in *differential* form, replacing the left side by $[X_i(t+\mathrm{d}t) - X_i(t)]/\mathrm{d}t$ and then multiplying through by $\mathrm{d}t$, we should really use some kind of "macroscopic" $\mathrm{d}t$. Notice that if we used $\bar{\mathrm{d}}t$ in that differential form, we would obtain the CLE (6) *except* for the last summation term therein.

As was discussed in ref 4, eq 6 is but one of several different candidates for "the" chemical Langevin equation that have been proposed in the prior literature. The main contribution of ref 4 was to show that eq 6 has the distinction of being *rigorously derivable* from the same premise (2) that underlies the CME (4), by making *specific approximations* that should be valid whenever conditions i and ii hold.

It is known in continuous Markov process theory that every *Langevin equation* for a process $\mathbf{X}(t)$ is accompanied by a unique *Fokker−Planck equation* for the probability density function $P(\mathbf{x},t|\mathbf{x}_0,t_0)$ of that process. The Fokker−Planck equation corresponding to the specific Langevin eq 6 turns out to be[4]

$$\frac{\partial}{\partial t}P(\mathbf{x},t|\mathbf{x}_0,t_0) = -\sum_{i=1}^{N}\frac{\partial}{\partial x_i}\left[\left(\sum_{j=1}^{M}\nu_{ji}a_j(\mathbf{x})\right)P(\mathbf{x},t|\mathbf{x}_0,t_0)\right] + \frac{1}{2}\sum_{i=1}^{N}\frac{\partial^2}{\partial x_i^2}\left[\left(\sum_{j=1}^{M}\nu_{ji}^2 a_j(\mathbf{x})\right)P(\mathbf{x},t|\mathbf{x}_0,t_0)\right] + \sum_{\substack{i,i'=1\\ i<i'}}^{N}\frac{\partial^2}{\partial x_i \partial x_{i'}}\left[\left(\sum_{j=1}^{M}\nu_{ji}\nu_{ji'}a_j(\mathbf{x})\right)P(\mathbf{x},t|\mathbf{x}_0,t_0)\right] \quad (7)$$

Rather amazingly, the time evolution of $P(\mathbf{x},t|\mathbf{x}_0,t_0)$ prescribed by this *chemical Fokker−Planck equation* (CFPE) exactly parallels the time evolution of $\mathbf{X}(t)$ prescribed by the CLE (6).[6] The inherently approximate nature of the CFPE (7) can be discerned from the fact that x in that equation is a *real* variable, whereas x in the exact CME (4) is an *integer* variable. Indeed, since the CFPE (7) is a direct consequence of the CLE (6), it too is ultimately predicated on the system possessing a macroscopically infinitesimal time scale, in the sense of conditions i and ii. In ref 4, plausibility arguments were given suggesting that conditions i and ii will usually be satisfied if the molecular population levels of all the reactant species stay sufficiently large.

Concerns about the correctness of the CLE (6) and the CFPE (7) were recently raised[7] in connection with the simple reversible isomerization reaction,

$$S_1 \underset{c_2}{\overset{c_1}{\rightleftharpoons}} S_2 \quad (8)$$

Assuming that the total number of isomers is a constant x_T, we can put

$$X_1(t) \equiv X(t), \qquad X_2(t) = x_T - X(t)$$

and treat this reaction set as an ($N = 1$, $M = 2$) process with $\mathbf{X}(t) = X(t)$. The propensity functions and the state-change vectors for this univariate process are

$$a_1(x) = c_1 x, \qquad a_2(x) = c_2(x_T - x) \quad (9a)$$

$$\nu_1 = -1, \qquad \nu_2 = +1 \quad (9b)$$

In ref 7, it was questioned whether, for this particular reaction scheme, the stationary solution of the CFPE (7) really matches the stationary solution of the CME (4), and also whether the

infinitesimal increment in $X(t)$ prescribed by the CLE (6) really matches the increment that is implied by the CME (4). In the following sections we shall investigate each of these questions *numerically*. We shall find that this reversible isomerization reaction model actually provides a felicitous illustration of many features of the CME-CLE-CFPE-RRE relationship which was drawn, rather more abstractly, in ref 4.

2. The Stationary Distribution of X(t)

The stationary or "equilibrium" solutions $P_s(\mathbf{x})$ to both the exact CME (4) and the approximating CFPE (7) are defined, whenever they exist, by

$$P_s(\mathbf{x}) = \lim_{(t-t_0)\to\infty} P(\mathbf{x},t|\mathbf{x}_0,t_0) \tag{10}$$

For the univariate process $X(t)$ that describes the reversible isomerization reactions (8) according to eqs 9, it is possible to calculate analytically and exactly the stationary solutions of both the CME (4) and the CFPE (7). In this section we shall obtain those solutions and then compare them, with a view to getting some idea of just how successful the CFPE (7) is at approximating the CME (4).

In Appendix A we review the argument showing that the stationary solution of the CME (4), which we shall designate with a superscript "m", is the *binomial* distribution[8]

$$P_s^m(x) = \frac{x_T!}{x!(x_T-x)!}q^x(1-q)^{x_T-x} \quad (x = 0, 1, ..., x_T) \tag{11a}$$

where

$$q \equiv \frac{c_2}{c_1+c_2} \tag{11b}$$

The mean and variance of P_s^m are therefore given by the standard binomial formulas[8]

$$\langle X\rangle_s^m = x_T q = \frac{x_T c_2}{c_1+c_2} \tag{12a}$$

$$\mathrm{var}\{X\}_s^m = x_T q(1-q) = \frac{x_T c_1 c_2}{(c_1+c_2)^2} \tag{12b}$$

To compute the stationary solution P_s^{FP} of the CFPE (7), we first use eq 9 to evaluate two relevant quantities, namely,

$$\sum_{j=1}^{2} \nu_j a_j(x) = c_2 x_T - (c_1+c_2)x \equiv A(x) \tag{13a}$$

$$\sum_{j=1}^{2} \nu_j^2 a_j(x) = c_2 x_T + (c_1-c_2)x \equiv D(x) \tag{13b}$$

Then, as explained in Appendix A, the stationary solution of the univariate form of the CFPE (7) takes the quadrature form

$$P_s^{FP}(x) = \frac{K}{D(x)}\exp\left(\int^x \frac{2A(x')}{D(x')}\,dx'\right) \tag{14}$$

where K is a normalization constant. The integral in eq 14 can be evaluated for the A and D functions in eq 13 by simply consulting a common table of integrals; however, the two cases $c_1 = c_2$ and $c_1 \neq c_2$ must be treated separately.

For the case $c_1 = c_2$, the integration in eq 14 gives

$$P_s^{FP}(x) = \frac{1}{\sqrt{2\pi(x_T/4)}}\exp\left(-\frac{(x-x_T/2)^2}{2(x_T/4)}\right) \qquad (c_1 = c_2) \tag{15}$$

This is evidently a *normal* (Gaussian) distribution with mean $x_T/2$ and variance $x_T/4$. We note that this mean and variance agree exactly with those of the stationary CME distribution P_s^m in eq 12 for the case $c_1 = c_2$.

For $c_1 \neq c_2$, the integration in eq 14 yields the result

$$P_s^{FP}(x) = K' e^{-\alpha x}[c_2 x_T + (c_1-c_2)x]^{\beta} \qquad (c_1 \neq c_2) \tag{16a}$$

where

$$\alpha \equiv 2\left(\frac{c_1+c_2}{c_1-c_2}\right), \qquad \beta \equiv \frac{4x_T c_1 c_2}{(c_1-c_2)^2} - 1 \tag{16b}$$

and K' is a normalization constant. To the best of this writer's knowledge, the distribution (16a) is not a named type. It is clearly not normal, though, as was claimed in ref 7.[9] But the important question is, do the results for $P_s^{FP}(x)$ in eqs 15 and 16 decently approximate the result for $P_s^m(x)$ in eq 11? Let's look at two examples.

In Figure 1 we compare, for the case $c_1 = c_2$, the stationary solution $P_s^{FP}(x)$ of the CFPE given in eq 15 with the stationary solution $P_s^m(x)$ of the CME given in eq 11 for the three x_T values (a) $x_T = 10$, (b) $x_T = 50$, and (c) $x_T = 500$. In each plot, the exact CME function is shown by *heavy dots* at integer x-values, and the approximating CFPE function is shown as a *solid curve* (which for simplicity has been linearly interpolated between the integer x-values). The agreement in each case appears to be excellent. In Figure 2 we press the comparison a bit harder by plotting the same data on a *logarithmic* vertical scale, and here we can see some differences: In the far tails of the distributions, which all appear to be zero on the linear scales of Figure 1, $P_s^{FP}(x)$ overestimates $P_s^m(x)$. For example, in the $x_T = 500$ case (Figure 2c), the two curves are virtually indistinguishable in the 13-standard-deviation interval $100 < x < 400$, but the equilibrium probability of finding, for instance, only 50 S_1 molecules is pegged by the CFPE to be about 10^{-73} instead of the correct value of about 10^{-85}.

Figures 3 and 4 show a similar treatment of the case $c_1 \neq c_2$, taking for definiteness $c_1 = 1$ and $c_2 = 2$. The heavy dots again show the exact stationary solution $P_s^m(x)$ in eq 11, and the line curves show the approximating stationary solution $P_s^{FP}(x)$, now given by eq 16. The story is much the same as in the $c_1 = c_2$ case: No differences are noticeable between $P_s^m(x)$ and $P_s^{FP}(x)$ on the *linear* scale plots (Figure 3), but on the *logarithmic* scale plots (Figure 4) $P_s^{FP}(x)$ is seen to overestimate the near-zero probabilities of finding numbers of molecules that differ from the mean by many standard deviations.

A clear explanation for the errors in the far tails of $P_s^{FP}(x)$ is difficult to discern, since the CFPE (7) is related only indirectly to conditions i and ii through the CLE (6). But it is probably no accident that these errors are occurring in regions of state space where the number of either S_1 or S_2 molecules is so small that conditions i and ii *cannot* simultaneously be satisfied. For example, in the $x_T = 500$ case described by Figures 1c and 2c, suppose there are $x = 8$ S_1 molecules at a particular instant. During any subsequent time interval that is large enough for both reaction channels to fire many more times than once, in satisfaction of condition ii, the propensity function $a_1(x)$ will almost certainly change from its starting value of $c_1 x = 8$ by a relatively large amount, in violation of condition i; because,

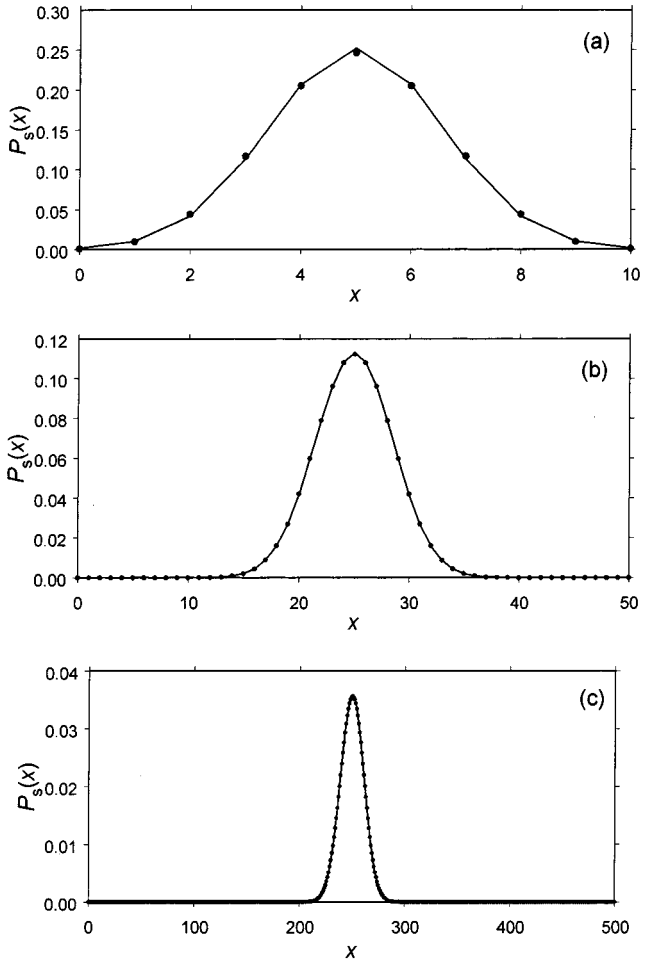

Figure 1. Stationary (equilibrium) probability distribution of $X(t) \equiv X_1(t)$ for the reversible isomerization reaction 8 for the case $c_1 = c_2$ and (a) $x_T = 10$, (b) $x_T = 50$, and (c) $x_T = 500$. In each frame, the *dots* show the *exact* CME function (11), and the line curve shows the *approximating* CFPE function (15).

most of those next reactions will be R_2 reactions, each of which increases x by 1 and hence also $a_1(x)$ by 1. So, when the system is in state $x = 8$, there simply is no time interval that satisfies *both* of conditions i and ii.

Certainly there will be circumstances in which the inaccuracies in the far tails of $P_s^{FP}(x)$ shown in Figures 2 and 4 can pose a problem; for instance, those errors would lead us to overestimate the chances of an extremely large fluctuation from equilibrium. But for most practical purposes, such as making ordinary linear plots such as those in Figures 1 and 3, the stationary solution of the CFPE (7) appears to provide a remarkably good approximation to the stationary solution of the CME (4), at least for the reversible isomerization reaction set (8).

3. The Increment in $X(t)$ in the Next dt

As was mentioned in section 1, the CLE (6) is essentially a formula for "updating" the populations of all the molecular species from the present time t to a slightly later time $t + \bar{\mathrm{d}}t$. In Appendix B, it is shown that the CLE (6) makes the following prediction for the reversible isomerization process: If that process has the value x at time t, i.e., if $X(t) = x$, then the *change* in the process in the next $\bar{\mathrm{d}}t$, namely

$$\Delta(x,\bar{\mathrm{d}}t) \equiv X(t+\bar{\mathrm{d}}t) - x \qquad (17)$$

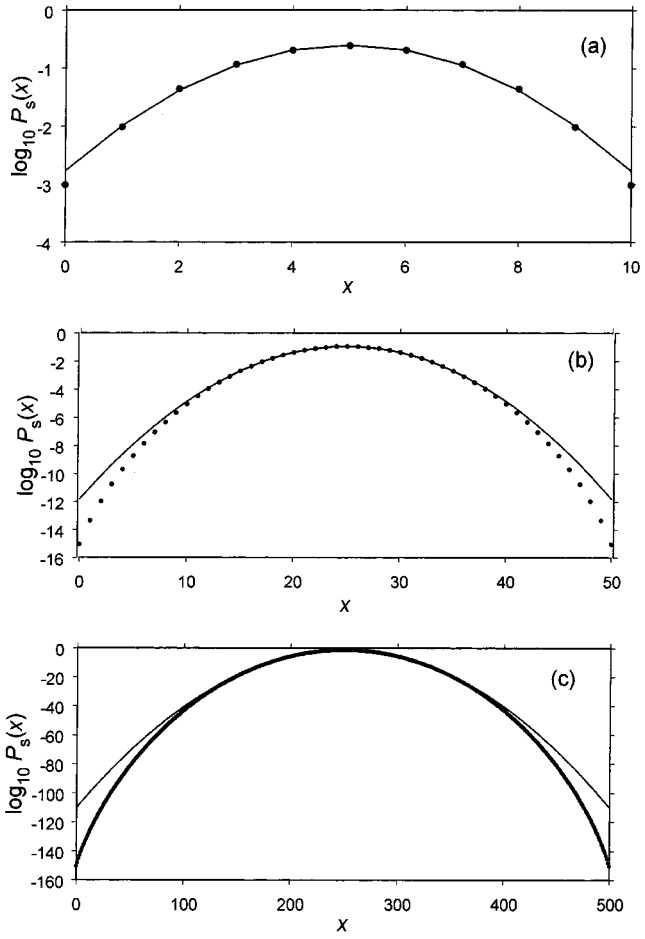

Figure 2. Same data as in Figure 1, but plotted on a logarithmic vertical scale.

will be a *normal* (Gaussian) random variable with *mean* $A(x)\,\bar{\mathrm{d}}t$ and *variance* $D(x)\,\bar{\mathrm{d}}t$:

$$\Delta(x,\bar{\mathrm{d}}t) = \mathcal{N}(A(x)\,\bar{\mathrm{d}}t,\, D(x)\,\bar{\mathrm{d}}t) \qquad \text{(CLE)} \qquad (18)$$

Here, $A(x)$ and $D(x)$ are the functions defined in eq 13. In this section we shall examine the *accuracy* of this prediction of the CLE. But before doing that, we need to discuss a related issue that was raised in ref 7.

Consider a *true* infinitesimal time increment dt, one that is so small that it is very unlikely to span more than one reaction event of any kind. Then in the next dt, one of three things can happen: Either one R_1 reaction will occur, or one R_2 reaction will occur, or no reaction will occur. Equation 2 gives the probabilities of those three eventualities when the system's current state is x, and thereby implies that the state change $\Delta(x,\mathrm{d}t)$ in that next dt will be the *tri-valued integer* random variable

$$\Delta(x,\mathrm{d}t) = \begin{bmatrix} -1, & \text{w/ probability } a_1(x)\,\mathrm{d}t \\ +1, & \text{w/ probability } a_2(x)\,\mathrm{d}t \\ 0, & \text{w/ probability } [1 - a_1(x)\,\mathrm{d}t - a_2(x)\,\mathrm{d}t] \end{bmatrix} \qquad (19a)$$

This result is expressed more compactly in ref 7 in terms of the Heaviside step function Θ and the unit-interval uniform random variable $\mathcal{U}(0,1)$:[10]

$$\Delta(x,\mathrm{d}t) = \Theta(a_2(x)\,\mathrm{d}t - \mathcal{U}(0,1)) - \Theta(\mathcal{U}(0,1) - 1 + a_1(x)\,\mathrm{d}t) \qquad (19b)$$

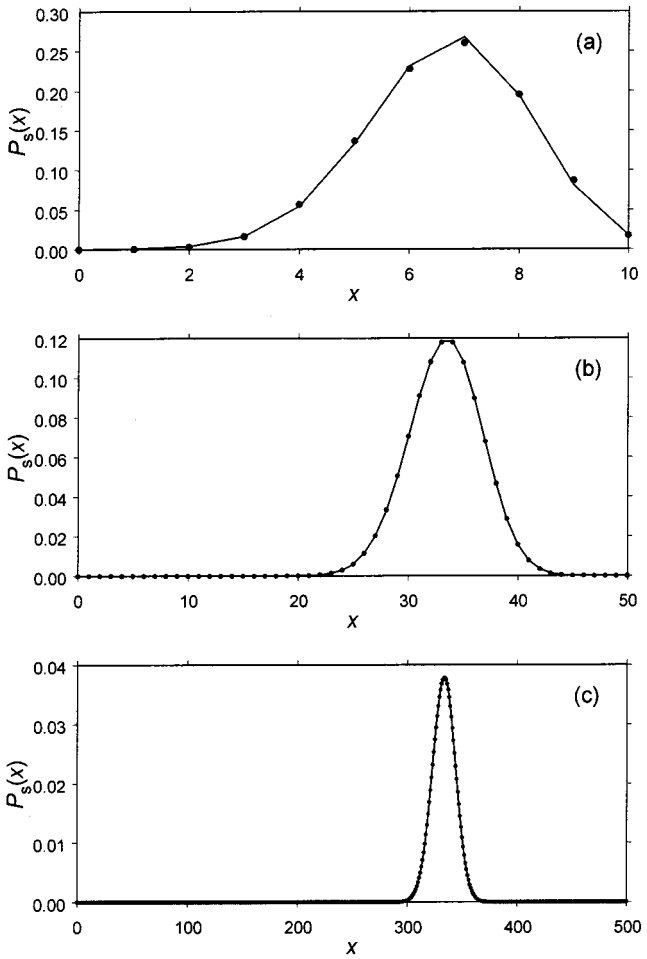

Figure 3. Stationary (equilibrium) probability distribution of $X(t) \equiv X_1(t)$ for the reversible isomerization reaction 8 for the case $c_1 = 1$, $c_2 = 2$, and (a) $x_T = 10$, (b) $x_T = 50$, and (c) $x_T = 500$. The *dots* show the *exact* CME function (11), and the line curve shows the *approximating* CFPE function (16).

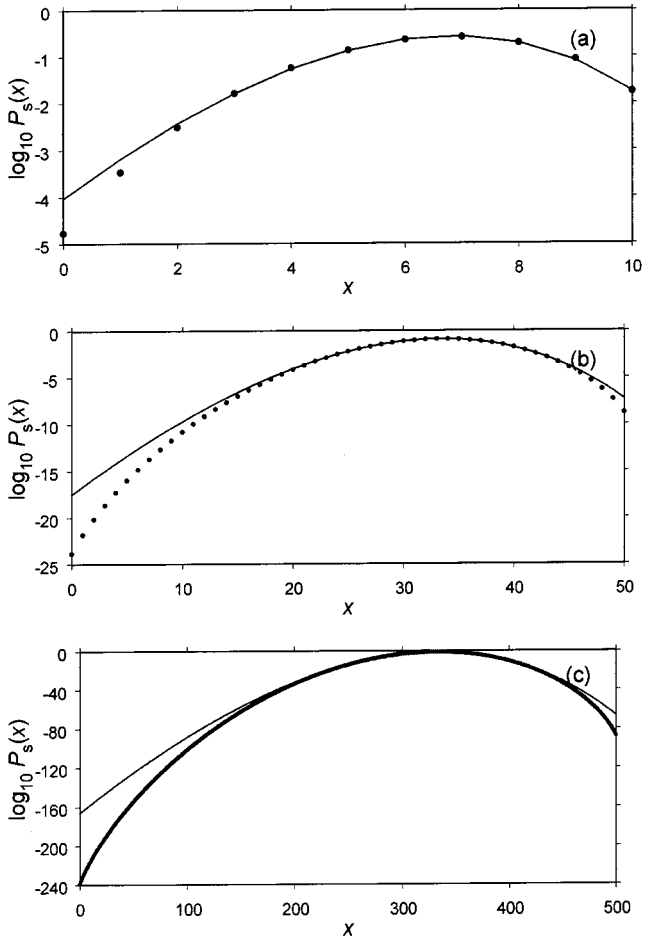

Figure 4. Same data as in Figure 3, but plotted on a logarithmic vertical scale.

It was suggested in ref 7 that the *Gaussian* state increment (18) predicted by the CLE (6) *conflicts* with the *tri-valued* state increment (19) implied by the fundamental premise (2), and that therefore eq 19b might be more deserving of the title "chemical Langevin equation" than eq 6.

If the macroscopic infinitesimal $\bar{d}t$ in eq 18 had been written *without* the overbar on the d, as in fact it was in ref 4 where the CLE (6) was derived (and as it also is in the ordinary reaction rate eq 5 and many other equations of applied science and engineering), one might indeed be tempted to conclude that there is a conflict between eqs 18 and 19. But our notation here makes it clear that *there is no conflict*, because the time increments in eqs 18 and 19 are not the same. The time increment dt in eq 19 is, by definition, so small that there is an insignificant probability that more than one reaction event will occur anywhere inside the system during the next dt. In contrast, the time increment $\bar{d}t$ in eq 18 is, by condition ii, large enough that each of the reaction channels R_1 and R_2 will likely fire many more times than once during the next $\bar{d}t$. *Both* results (18) and (19) are logical consequences of the fundamental premise (2): eq 19 is an *exact* consequence of (2), following by the straightforward argument outlined above, while eq 18 is an *approximate* consequence of (2) which follows, subject to conditions i and ii, by the somewhat more involved argument given in ref 4. We note also that eq 6 has the canonical form of a "Langevin equation" in the technical sense of continuous Markov process theory, while eq 19b does not.

Returning now to the question of the accuracy of formula (18) for $\Delta(x,\bar{d}t)$, and hence also accuracy of the CLE (6), we are immediately faced with two problems: First, how should we go about selecting an appropriate value for $\bar{d}t$? And second, how can we compute the "true" value of $\Delta(x,\bar{d}t)$, in accordance with the CME (4), against which to measure the accuracy of eq 18?

Guidelines for selecting values for $\bar{d}t$ that satisfy conditions i and ii are derived in Appendix C. There it is shown that condition i should be satisfied if

$$\bar{d}t \ll (c_1 + c_2)^{-1} \tag{20a}$$

while condition ii should be satisfied if

$$\bar{d}t \gg \mathrm{Max}\left\{\frac{1}{a_1(x)}, \frac{1}{a_2(x)}\right\} \tag{20b}$$

Simultaneous satisfaction of these conditions evidently hinges on making the right-hand side of (20b) sufficiently small. An inspection of the propensity function formulas (9a) shows that this can always be done if the S_1 population x and the S_2 population $x_T - x$ are *both* sufficiently large.

Once a value for $\bar{d}t$ satisfying conditions (20) has been chosen, which of course requires that values have been specified for the parameters c_1, c_2, x_T, and x, we can straightforwardly compute the normal probability density function prescribed by eq 18 for the random variable $\Delta(x,\bar{d}t)$. But in order to assess the

fidelity of that prediction of the CLE (6), we must determine the *exact* probability density function of $\Delta(x,\bar{d}t)$. That exact density function can be expressed (as a function of δ) as $P(x+\delta,\bar{d}t|x,0)$; so, one way to proceed would be to solve the CME (4). But the CME is typically very difficult to solve, either analytically or numerically. So we shall take a different approach: We shall make very many statistically independent *numerical simulations* of the process $X(t)$. Each simulation will be started at the specified state x at time 0 and will be run to time $\bar{d}t$ to obtain a value $X(\bar{d}t) = x + \delta$. The δ-values thus obtained will constitute a set of *statistically independent samples* of the random variable $\Delta(x,\bar{d}t)$; so a *normalized frequency histogram* of those δ-values will provide an estimate of the corresponding probability density function $P(x+\delta,\bar{d}t|x,0)$. The histogram will of course be subject to uncertainties arising from the necessarily finite number of sample values plotted, but those uncertainties can be easily estimated. This approach is feasible because there exists a simple way of numerically simulating chemical reactions that is *exactly equivalent* to the CME (4): the so-called *stochastic simulation algorithm*.[11]

Let's look at a specific numerical example. We shall take for the reaction probability rate constants $c_1 = 1$ and $c_2 = 2$, and we shall fix the total number of molecules at $x_T = 10^4$. Equations 12a and 12b then predict that the process $X(t)$ will eventually come to equilibrium by fluctuating about its mean $\langle X\rangle_s^m = 6666.7$ with standard deviation $\mathrm{sdev}\{X\}_s^m = 47.1$. Let us choose as our starting state $x = 6430$, a value that is roughly 5 standard deviations below the stationary mean. For these parameter values, conditions (20) will be found to require that any acceptable $\bar{d}t$ should be $\ll 0.33$ and $\gg 1.6 \times 10^{-4}$. We can satisfy both conditions reasonably well by taking $\bar{d}t = 10^{-2}$. Using the definitions (13), we find that the CLE formula (18) becomes

$$\Delta(6430,10^{-2}) = \mathcal{N}(7.1,135.7) \qquad (x_T = 10^4) \qquad (21)$$

Figure 5 shows as the *solid curve* the probability density function of this normal random variable. Superimposed is the normalized frequency histogram of state change values that were *actually observed* in 10^4 exact stochastic simulation runs; each run started in state $x = 6430$ at time 0 and ended at time $\bar{d}t = 10^{-2}$. The vertical error bars indicate the conventional one-standard deviation ($n \pm \sqrt{n}$) uncertainty estimates. Figure 5 shows that the distribution (21) predicted by the CLE (6) provides a remarkably good fit to numerical simulations carried out in exact accord with the CME (4).

To test the assertion of ref 4 that the CLE approximation should become even better as the number of molecules involved is made larger, let us keep the same reaction constants but increase the total number of molecules by a factor of 100, to $x_T = 10^6$. And let us choose as our starting state $x = 660\,000$, which can be shown from eq 12 to be about 14 standard deviations below the stationary mean. For these values, we find from conditions (20) that an acceptable $\bar{d}t$ should be $\ll 0.33$ and $\gg 1.5 \times 10^{-6}$, an acceptance range for $\bar{d}t$ that is about 100 times *larger* than in the previous case. Taking $\bar{d}t = 10^{-3}$, we find that the CLE formula (18) now gives for $\Delta(x,\bar{d}t)$

$$\Delta(660000,10^{-3}) = \mathcal{N}(20,1340) \qquad (x_T = 10^6) \qquad (22)$$

Figure 6 shows the density function for this normal distribution, along with the results of 10^4 exact stochastic simulation runs that were made using the same parameter values. The agreement indeed appears to be as good as, if not better than, the agreement exhibited in Figure 5 for a smaller total number of molecules.

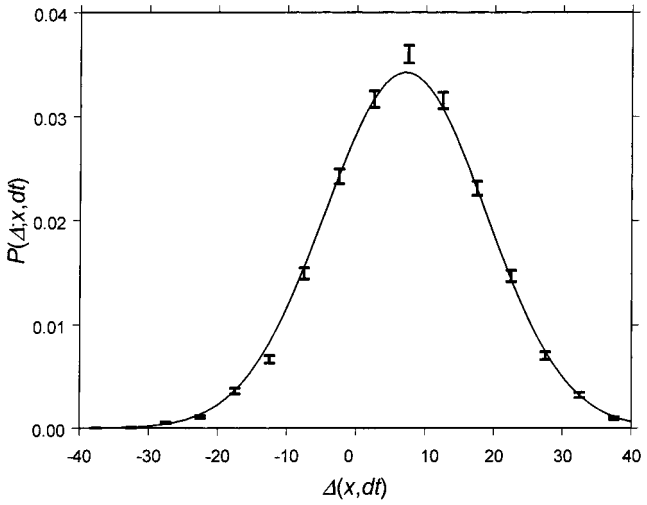

Figure 5. For the case $c_1 = 1$, $c_2 = 2$, and $x_T = 10^4$, a comparison of two different estimates of the probability density function of the *change in state* over a time $dt = 10^{-2}$, starting at $x = 6430$ (about 5 standard deviations below the stationary mean). The solid curve is the prediction (21) of the CLE, a normal (Gaussian) distribution with mean 7.1 and variance 135.7. The frequency histogram displays the results of 10^4 exact stochastic simulation runs, reflecting the prediction of the CME. The vertical interval bars in the frequency histogram indicate the one-standard deviation ($n \pm \sqrt{n}$) uncertainty estimates inferred from the bin occupation numbers n.

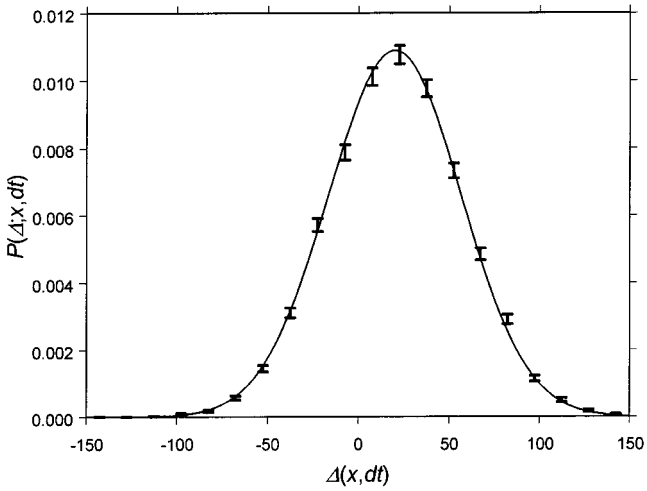

Figure 6. Same as in Figure 5, except the total number of isomers has been increased by 2 orders of magnitude to $x_T = 10^6$, and the state change is measured from the initial state $x = 660\,000$ (about 14 standard deviations below the stationary mean) over a time $\bar{d}t = 10^{-3}$. This larger molecular population level makes it easier to satisfy the approximation conditions (20) required by the CLE.

Besides confirming the essential correctness of the CLE (6), these model calculations also give us a glimpse of the potential utility of that approximate equation. In Appendix C it is shown that the average *total* number of reactions that occur when the system makes the state change $\Delta(x,\bar{d}t)$ is $D(x)\bar{d}t$; this number is also, by eq 18, the estimated variance of $\Delta(x,\bar{d}t)$. So, in the last example considered, eq 22 tells us that a single run typically entails simulating 1340 individual reactions events (a figure that was confirmed in the actual simulations). But Figure 6 shows that the net state change produced by those 1340 reaction events can be accurately approximated by simply drawing a single random number from the normal distribution (22), a computational task that is very easy to accomplish. The message here is that *when conditions i and ii can be satisfied, the CLE (6) can enable substantial gains in simulation efficiency*.[12]

4. Summary and Conclusions

The *chemical master equation* (4) has a sound theoretical basis in kinetic theory for "well-stirred" chemical systems.[1] Reference 4 argued that, under certain specific conditions, the CME can be decently *approximated* by the *chemical Langevin equation* (6) and its companion *chemical Fokker−Planck equation* (7). In this paper we have tested that claim by quantitatively comparing certain predictions of the CME with those of the CLE and CFPE for the reversible isomerization reaction 8.

We found in section 2 that the equilibrium solution of the CFPE approximates that of the CME quite well (see Figures 1 and 3), erring only in predicting precisely how close to zero the CME solution is in the far tails (see Figures 2 and 4), which are regions of the state space where the conditions stipulated in ref 4 are violated. And in section 3 we found that, provided the time infinitesimal $\bar{d}t$ in the CLE (6) is "macroscopic" in the sense stipulated by ref 4, the change in the molecular populations in time dt predicted by the CLE matches very well the change predicted by the CME (see Figures 5 and 6).

The question arises, did we just get lucky here, or may we expect these simple model results to typify those for more complex reaction schemes? Of course, until we explicitly address those other reactions schemes we cannot be sure. But in the past, Langevin-type equations in applied physics have usually been obtained by simply adding to a deterministic, phenomenological rate equation a "random noise term," the form of which has been chosen to ensure some desired thermal equilibrium effect. It must be stressed that the derivation of the CLE (6) given in ref 4 does not use that *ad hoc* strategy; instead, ref 4 starts from premise (2), the same premise that leads *exactly* to the CME (4), and it shows how satisfaction of *two specific conditions* leads *approximately* to the CLE (6). The specific conditions are that the time increment $\bar{d}t$ in the CLE (6) must be (i) small enough that none of the propensity functions changes in a macroscopically noticeable way during $\bar{d}t$, yet (ii) large enough that each reaction channel fires many more times than once during $\bar{d}t$. After establishing the CLE (6), ref 4 then infers the CFPE (7) from some general results in the theory of continuous Markov processes.[6] The logic of the analysis in ref 4 strongly suggests that, so long as conditions i and ii are satisfied, the CLE (6) and the CFPE (7) should provide acceptable approximations to the CME (4) for *any* chemical reaction scenario. If that is true, then the favorable showing of the CLE (6) and CFPE (7) here for the simple reversible isomerization reaction is not just a lucky accident.

A larger question is, why should we bother to concern ourselves with approximating the CME by a CLE/CFPE? There are two reasons. On the practical side, this approximation can open the way to *more efficient computational strategies*. This fact is not so evident in the computations carried out in section 2, where the stationary solutions of the CFPE in eqs 15 and 16 were no easier to obtain than the stationary solution of the CME in eq 11. But, more generally, the art of solving partial differential equations such as the CFPE (7) is much more highly developed than the art of solving differential-difference equations such as the CME (4). A more significant computational advantage afforded by the CLE/CFPE approximation probably lies in the area of *numerical simulation*. Thus, in the last example considered in section 3 it was found that the CLE offers a way of accurately replicating the effect of some 1340 reaction events by simply drawing *one* normal (Gaussian) random number (see Figure 6).[12] Of course, all of these potential computational advantages of the CLE/CFPE approximation are predicated on the satisfaction of conditions i and ii; if the situation is such that those conditions cannot be satisfied, the CLE and the CFPE will have no claim to validity.

The second reason for concerning ourselves with the CLE/CFPE approximation has to do with *conceptual understanding*. As was mentioned at the beginning of this paper, it has never been clear how the CME (4) gets supplanted, in the thermodynamic limit, by the conventional *reaction rate equation* (5). The CLE (6) appears to answer this question. Its derivation from premise (2) under conditions i and ii shows that, when the reactant populations become sufficiently large, the CME (4) segues to the CLE (6) and the CFPE (7). The further connection to the RRE (5) can then be understood from a simple inspection of the two summation terms on the right-hand side of the CLE (6): The first summation term (the *deterministic drift* term containing $\bar{d}t$) is essentially proportional to the propensity functions, while the second term (the *randomly fluctuating* term containing $(\bar{d}t)^{1/2}$) is basically proportional to the *square root* of the propensity functions. Since the propensity functions (eq 9a) for our reactions are themselves proportional to the molecular populations, then the deterministic component of the motion will scale like the molecular populations while the fluctuating component will scale like the *square root* of the molecular populations.[13] Therefore, as the molecular populations increase, the fluctuating component will *decrease* relative to the deterministic component as the inverse square root of the molecular populations. And in the thermodynamic limit of an *infinite* molecular population, the last term in the CLE (6) usually becomes negligible, so the CLE (6) collapses to the RRE (5). *The CLE/CFPE approximation thus appears to be the "bridge" between the CME (4) and the RRE (5).*

The concerns about the CLE (6) raised in ref 7 fail to take cognizance of condition ii, which stipulates the "macroscopic" nature of the infinitesimal time increment in the CLE (6). The alternative increment formula (19b) proposed in ref 7 involves a "true" time infinitesimal and hence amounts to an exact restatement of the fundamental premise (2) for the reversible isomerization process $X(t)$, reaffirming that $X(t)$ is a jump Markov process obeying the master eq 4. Equation 6 by contrast represents a deliberate attempt to *approximate* that *jump* Markov process by a *continuous* Markov process. The results presented here for the simple reversible isomerization reaction model provide concrete evidence of the validity of that approximation.

For the pedagogical purpose of emphasizing the macroscopic nature of the infinitesimal time increment in eq 6, we have used here the notation $\bar{d}t$. But once this point has been clearly understood, it would seem reasonable to revert to the simpler dt notation used in ref 4. That designation for a "macroscopic time infinitesimal" has been commonplace in applied science and engineering for many years; indeed, the dt in the traditional RRE (5) is a macroscopic infinitesimal.

Acknowledgment. I thank Carol Gillespie for her assistance with the numerical computations and figures. This work was supported in part by the Office of Naval Research under Grant No. N0001401WX21040, and in part by the Air Force Office of Scientific Research under Award No. F30602-01-2-0558 and The California Institute of Technology.

Appendix A: Derivation of Eqs 11 and 14

The stationary solution $P_s^m(x)$ of the CME (4) is by definition a time-independent solution of that equation. So for the process defined by eq 9, we have

$$0 = [a_1(x+1)\,P_s^m(x+1) - a_1(x)\,P_s^m(x)] + [a_2(x-1)\,P_s^m(x-1) - a_2(x)\,P_s^m(x)]$$
$$= [a_1(x+1)\,P_s^m(x+1) - a_2(x)\,P_s^m(x)] - [a_1(x)\,P_s^m(x) - a_2(x-1)\,P_s^m(x-1)]$$

Here x is an *integer* variable on the interval $[0,x_T]$. The last equality implies that the quantities in brackets must have a value that is independent of x. A consideration of the last bracketed term for the case $x = 0$ shows that that constant value must in fact be zero; hence, $P_s^m(x)$ satisfies the "detailed balance" condition,

$$a_1(x)\,P_s^m(x) = a_2(x-1)\,P_s^m(x-1)$$

Dividing this by $a_1(x)$ gives a simple recursion relation for $P_s^m(x)$. Iterating that recursion relation yields

$$P_s^m(x) = P_s^m(0)\prod_{k=1}^{x}\frac{a_2(k-1)}{a_1(k)} = P_s^m(0)\left(\frac{c_2}{c_1}\right)^x\frac{x_T!}{x!(x_T-x)!} \qquad (x = 0, 1, ..., x_T) \quad \text{(A1)}$$

where the last step has invoked eq 9a. This result was noted in ref 7. The value of $P_s^m(0)$ can be determined by normalizing $P_s^m(x)$ on the integer interval $0 \le x \le x_T$; however, it is simpler to observe that, with the definition (11b), eq A1 can be written the binomial form (11a).[8]

The CFPE (7) simplifies for a univariate process by losing the cross-derivative terms. With the definitions of the functions A and D in eq 13, the time-independent solution $P_s^{FP}(x)$ of that equation for the process defined by eq 9 will be seen to satisfy

$$0 = \frac{d}{dx}\left[-A(x)\,P_s^{FP}(x) + \frac{1}{2}\frac{d}{dx}(D(x)\,P_s^{FP}(x))\right]$$

where x is now a *real* variable. This shows that the bracketed term is independent of x. Since $P_s^{FP}(x)$ must vanish for sufficiently large and small values of x, the constant value of the bracketed term must be zero. So we are left with the first-order ordinary differential equation,

$$\frac{d}{dx}(D(x)\,P_s^{FP}(x)) = 2A(x)\,P_s^{FP}(x)$$

or

$$\frac{d(D(x)\,P_s^{FP}(x))}{D(x)\,P_s^{FP}(x)} = \frac{2A(x)}{D(x)}dx$$

Integrating this last equation gives the quadrature result (14).

Appendix B: Derivation of Eq 18

For $X(t) = x$, the CLE (6) gives for the $M = 2$ reversible isomerization process

$$\Delta(x,\bar{d}t) \equiv X(t+\bar{d}t) - x = \sum_{j=1}^{2}\nu_j a_j(x)\,\bar{d}t + \sum_{j=1}^{2}\nu_j a_j^{1/2}(x)\,N_j(t)\,(\bar{d}t)^{1/2}$$

Here, $N_1(t)$ and $N_2(t)$ are statistically independent normal random variables with means 0 and variances 1. Denoting the normal random variable with mean m and variance σ^2 by $\mathcal{N}(m,\sigma^2)$, and using the well-known linear combination rule for statistically independent normal random variables, namely we can manipulate the above expression as follows:

$$\alpha\mathcal{N}_1(m_1,\sigma_1^2) + \beta\mathcal{N}_2(m_2,\sigma_2^2) = \mathcal{N}(\alpha m_1+\beta m_2,\alpha^2\sigma_1^2+\beta^2\sigma_2^2)$$

$$\Delta(x,\bar{d}t) = \sum_{j=1}^{2}\nu_j a_j(x)\,\bar{d}t + \sum_{j=1}^{2}[\nu_j a_j^{1/2}(x)\,(\bar{d}t)^{1/2}]\mathcal{N}_j(0,1)$$
$$= \sum_{j=1}^{2}\nu_j a_j(x)\,\bar{d}t + \mathcal{N}(0,\sum_{j=1}^{2}\nu_j^2 a_j(x)\,\bar{d}t)$$
$$= A(x)\,\bar{d}t + \mathcal{N}(0,D(x)\,\bar{d}t)$$

where the last step invokes the definitions (13). A final application of the linear combination rule for normal random variables results in eq 18.

Appendix C: Derivation of Conditions (20)

Condition i essentially requires $\bar{d}t$ to be very small from a *macroscopic* point of view. To get a quantitative handle on this requirement, let us examine the temporal behavior of the *mean* of $X(t)$: If one multiplies the CME (4) by $\mathbf{x}$ and then averages over $\mathbf{x}$, one can show quite generally that

$$\frac{d\langle X(t)\rangle}{dt} = \sum_{j=1}^{M}\nu_j\langle a_j(\mathbf{X}(t))\rangle$$

For the reversible isomerization process $X(t)$, the *linearity* of the two propensity functions (9a) allows this equation to be simplified to

$$\frac{d\langle X(t)\rangle}{dt} = \sum_{j=1}^{2}\nu_j a_j(\langle X(t)\rangle) = c_2x_T - (c_1 + c_2)\langle X(t)\rangle$$

where the last step follows from eq 13a. The solution to this equation for the initial condition $\langle X(t_0)\rangle = x_0$ is easily shown to be

$$\langle X(t)\rangle = \langle X\rangle_s^m + [x_0 - \langle X\rangle_s^m]e^{-(c_1+c_2)(t-t_0)}$$

where $\langle X\rangle_s^m$ is the stationary mean given in eq 12a. This result shows that the mean of $X(t)$ relaxes to its stationary value $\langle X\rangle_s^m$ on a time scale of order $(c_1 + c_2)^{-1}$. So, $(c_1 + c_2)^{-1}$ typifies a "macroscopic" duration of time, and we may take condition i as requiring that $\bar{d}t$ in the CLE (6) should always be very small compared to that value. This gives us condition (20a).

Assuming now that $\bar{d}t$ has been chosen so that condition (20a) is satisfied, then with $X(t) = x$ we can be assured that both propensity functions will remain *approximately constant* at their values $a_j(x)$ throughout the entire interval $[t,t+\bar{d}t]$. So by eq 2, the probability that an R_j reaction will occur during *any true* infinitesimal interval dt inside $[t,t+\bar{d}t]$ will be $a_j(x)dt$. This implies that the total number of R_j reactions occurring during the larger interval $[t,t+\bar{d}t]$ will be a *Poisson* random variable with mean (and variance) $a_j(x)\,\bar{d}t$ [the detailed reasoning behind this conclusion is reviewed in Appendix A of ref 4]. So, viewing condition ii as requiring that the *average* number of R_j reactions occurring in $[t,t+\bar{d}t]$ be very large compared to 1 for *all* reaction

channels R_j, we have the mathematical requirement $a_j(x)\,\bar{\mathrm{d}}t \gg 1$ for $j = 1$ and 2. This gives us condition (20b).

As just noted, when $\bar{\mathrm{d}}t$ satisfies conditions i and ii, $a_j(x)\,\bar{\mathrm{d}}t$ gives the mean number of R_j reactions occurring in the next $\bar{\mathrm{d}}t$. The mean number of reactions of *any* kind occurring in the next $\bar{\mathrm{d}}t$ is therefore

$$[\text{av no. reactions in next } \bar{\mathrm{d}}t] = \sum_{j=1}^{M} a_j(x)\,\bar{\mathrm{d}}t$$

For the reversible isomerization process, the state change vectors of both reaction channels R_j satisfy $\nu_j^2 = 1$ (see eqs 9b); therefore, it follows from the definition (13b) that the average total number of reactions occurring in the next $\bar{\mathrm{d}}t$ can also be written $D(x)\,\bar{\mathrm{d}}t$.

References and Notes

(1) Gillespie, D. T. *Physica A* **1992**, *188*, 404−425.

(2) McQuarrie, D. A. *J. Appl. Prob.* **1967**, *4*, 413−478.

(3) McAdams, H. H. and Arkin, A. P. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 814−819.

(4) Gillespie, D. T. *J. Chem. Phys.* **2000**, *113*, 297−306.

(5) As discussed in ref 4, a *mathematically equivalent* version of the chemical Langevin equation (6), a version that may be more familiar to readers conversant with stochastic differential equations, is

$$\frac{\mathrm{d}X_i(t)}{\mathrm{d}t} = \sum_{j=1}^{M} \nu_{ji} a_j(\mathbf{X}(t)) + \sum_{j=1}^{M} \nu_{ji} a_j^{1/2}(\mathbf{X}(t))\,\Gamma_j(t)$$

Here $\Gamma_1(t)$, ..., $\Gamma_M(t)$ are "Gaussian white noise" processes satisfying $\langle \Gamma_j(t)\,\Gamma_{j'}(t')\rangle = \delta(j,j')\,\delta(t-t')$, where the first delta function is Kronecker's and the second is Dirac's. But this equation contains no information that is not already contained, arguably more clearly, in eq 6; also, it tends to obscure the fact that the "derivative" on the left side does not really exist by ordinary mathematical standards. In any case, we shall not be needing this alternate version of the chemical Langevin equation in our work here.

(6) Equation 7 can be *formally* obtained by Taylor expanding the right side of the CME (4) and then summarily dropping all terms with derivatives of order >2. But that is *not* the way eq 7 is deduced in ref 4: In ref 4, eq 7 is inferred directly from eq 6 by appealing to some general results in continuous Markov process theory [see, e.g.: Gillespie, D. T. *Am. J. Phys.* **1996**, *64*, 1246−1257], and eq 6 in turn is shown to be a direct *approximate* consequence of the fundamental premise (2) whenever conditions i and ii are satisfied.

(7) Zwanzig, R. *J. Phys. Chem. B* **2001**, *105*, 6472−6473.

(8) In ref 7, the distribution (A1), which reduces to eq 11, was incorrectly identified as a *Poisson* distribution. The canonical forms of the binomial and Poisson distributions are given in the table on p 929 of the *Handbook of Mathematical Functions* (Abramowitz, M.; Stegun, I. A.; National Bureau of Standards: Washington, DC, 1964), which also shows the binomial mean and variance formulas (12). It turns out that the binomial distribution (11a) *becomes* a Poisson distribution in the dual limit $x_T \to \infty$, $q \to 0$, with $x_T q$ a finite constant. With the definition (11b), this Poissonian limit is $x_T \to \infty$, $c_2/c_1 \to 0$, with $x_T c_2/c_1$ a finite constant. But we have no need to invoke that special limit here.

(9) The assertion in ref 7 (near the end of its section 1) that "because of the linearity of the particular rate equation used here, Gaussian noise leads to a Gaussian equilibrium distribution" is incorrect. The rate equation in question is a *Langevin* equation, which depends not only on a *drift* function A but also a *diffusion* function D. When A is linear and D is a *constant*, the stationary distribution will indeed be normal (a well-known example being the Ornstein−Uhlenbeck process), and our eq 13 shows that this is the case when $c_1 = c_2$. But if *both* A and D are linear, as eq 13 shows is the case when $c_1 \neq c_2$, the stationary distribution will *not* be normal. The reason is essentially that, if $D(x) \propto x$ and $X(t)$ happens to be normal, then the diffusion term $D^{1/2}(X(t))\,N(t)\,\mathrm{d}t^{1/2}$ in the Langevin equation will contain the product of a normal with the square root of a normal, which is *not* a normal; therefore, $X(t+\mathrm{d}t)$ will not be normal.

(10) By definition, $\Theta(z) = 1$ if $z > 0$ and 0 otherwise. So, if the random variable $\mathscr{U}(0,1)$ takes a value between 0 and $a_2(x)\,\mathrm{d}t$, which will happen with probability $a_2(x)\,\mathrm{d}t$, the first Θ function in eq 19b will be 1 and the second will vanish; hence, with probability $a_2(x)\,\mathrm{d}t$ the expression in eq 19b will be +1, in agreement with eq 19a. If the random variable $\mathscr{U}(0,1)$ takes instead a value between $1 - a_1(x)\,\mathrm{d}t$ and 1, which will happen with probability $a_1(x)\,\mathrm{d}t$, the second Θ function in eq 19b will be 1 and the first will vanish; so, with probability $a_1(x)\,\mathrm{d}t$ the expression in eq 19b will be −1, again in agreement with eq 19a. If neither of these two things happen, both Θ functions in eq 19b will vanish, and we have the third eventuality in eq 19a.

(11) The stochastic simulation algorithm (SSA) is described in: Gillespie, D. T. *J. Comput. Phys.* **1976**, *22*, 403−434; *J. Phys. Chem.* **1977**, *81*, 2340−2361. However, in ref 7 a different simulation procedure was described (though not actually carried out). In that procedure, one chooses a time step Δt, draws a sample value r of the unit-interval uniform random variable $\mathscr{U}(0,1)$, and then effects in the next Δt an R_2 reaction if $r < a_2(x)\,\Delta t$, or an R_1 reaction if $r > 1 - a_1(x)\,\Delta t$, or no reaction otherwise. This is a conceptually straightforward implementation of the fundamental premise (2), but it has the practical drawback that it is *approximate*; it becomes exact only in the limit that Δt becomes infinitesimally small, in which limit the procedure becomes infinitely slow and infinitely consuming of random numbers r. The SSA by contrast is *exact* with respect to premise (2), and hence also with respect to the CME (4). It requires *two* random numbers for each reaction *event*; one of those random numbers determines the time to the next reaction event, and the other determines the identity (index) of that reaction. The SSA does *not* require one to choose a time step size, *nor* does it entail approximating an infinitesimal time interval $\mathrm{d}t$ by a finite time interval Δt. The main limitation of the SSA derives from the fact that it *does* dutifully simulate *every* reaction event that occurs in the system: If the molecular population level of any reactant species happens to be so large that an enormous number of reaction events actually occur per unit of real time, the progress of the SSA in real time will be extremely slow. Of course, this limitation also applies to the other simulation algorithm.

(12) For an extensive discussion of the Langevin and related approximation strategies for accelerating the stochastic simulation algorithm, see: Gillespie, D. T. *J. Chem. Phys.* **2001**, *115*, 1716−1733. Work in this area is ongoing.

(13) This *linear* scaling of the propensity functions with the molecular populations actually obtains for *all* types of reaction in the thermodynamic limit, wherein the molecular populations X_i *and* the system volume Ω approach infinity in such a way that the concentrations X_i/Ω remain finite: The propensity function a_j of an mth order reaction will contain m molecular population factors, but its reaction probability rate constant c_j will contain a factor $\Omega^{-(m-1)}$ (see ref 1), which effectively "cancels" all but one of the population factors in the thermodynamic limit. The CLE (6) thus implies quite generally the well known rule-of-thumb that "fluctuations scale like the square root of the molecular population."