

STOCHASTIC MODELING OF MOLECULAR DYNAMICS*

Jack H. Freed
Baker Laboratory of Chemistry
Cornell University
Ithaca, New York 14853/USA

We outline here a method, due to Stillman and Freed,¹ for the stochastic modeling of the non-Markovian many-body features of diffusing molecules. We believe this approach is a particularly useful one in that it introduces, in a transparent manner, the basic physics of the relevant degrees of freedom and their couplings, and it is not restricted to linear transport laws. Furthermore, the stochastic features of the bath variables are introduced in a simple and physically transparent fashion. Also, this approach permits consideration of either equilibrium or non-equilibrium dynamics in that the modeling allows for physically relevant choice of equilibrium or stationary-state, and it subjects the expressions to the constraints of detailed balance with respect to this state.

In this method the set of relevant dynamical variables is first augmented with stochastic bath variables which are assumed to obey simple Markovian laws. The augmented set then represents a multidimensional Markov process which obeys a stochastic Liouville equation (SLE) that is, in general, incomplete because it ignores the back reaction of the molecule (i.e. the relevant dynamical variables) on the bath variables. The back reaction effects are incorporated into the model by adding terms to the SLE which are obtained by the constraints required for detailed balance. The resulting augmented Fokker-Planck equation (AFPE) properly describes relaxation to thermal equilibrium, and, for the appropriate limiting conditions, reduces to a classical Fokker-Planck equation. Augmented Langevin equations (ALE) may readily be obtained from the AFPE, and because of the constraint of detailed balance, they automatically obey the fluctuation-dissipation theorem.

The AFPE, which is in general complex, can be solved by means of modern computational algorithms. In particular, it has been shown in recent work,^{2,3} that the Lanczos algorithm,⁴ which is closely related to the method of moments,⁵ is a powerful approach. It is intriguing to note that the projection operator of the method of moments, which can be applied as a mathematical technique in any Hilbert space, is closely related to the projection operator that Mori⁶ introduced to obtain a continued fraction representation of the generalized Langevin equation (GLE) for linear physical laws.³

We consider a set of independent dynamical variables $\underline{\Delta}$ whose equation of motion may be written as

$$\frac{d}{dt}\underline{\Delta} = \underline{F}(\underline{\Delta}; \underline{\Xi}(t), \underline{\lambda}) \quad (1)$$

where $\underline{\Xi}(t)$ denotes a set of independent stochastic bath variables and $\underline{\lambda}$ is a set of externally determined parameters such as temperature. In general $\underline{F}(\underline{\Delta}; \underline{\Xi}, \underline{\lambda})$ may be a non-linear function of the variables. The stochastic process for $\underline{\Xi}$ is assumed to be stationary and Markovian with an associated master equation

$$\frac{\partial}{\partial t} P(\underline{\Xi}, t; \lambda) = -\Gamma_{\underline{\Xi}} P(\underline{\Xi}, t; \lambda) \quad (2)$$

the SLE may be written as⁷

$$\frac{\partial}{\partial t} P(\underline{\Delta}, \underline{\Xi}, t; \lambda) = -[\nabla_{\underline{\Delta}} \cdot F(\underline{\Delta}; \underline{\Xi}, \lambda) + \Gamma_{\underline{\Xi}}] P(\underline{\Delta}, \underline{\Xi}, t; \lambda) \quad (3)$$

Here $\nabla_{\underline{\Delta}}$ represents the divergence over the space spanned by $\underline{\Delta}$ and the first term on the rhs of Eq. (3) represents the Liouville equation form of Eq. (1). Also $P(\underline{\Delta}, \underline{\Xi}, t; \lambda)$ is the joint probability distribution in the combined set of variables: $\underline{\Delta}$ and $\underline{\Xi}$. It should be emphasized that Eq. (3) is incomplete in that the back-reaction effects of $\underline{\Delta}$ on the diffusion of $\underline{\Xi}$ do not appear in Eq. (2). Thus the stationary solution of Eq. (3) will, in general, yield the correct Boltzmann distribution only in the limit of infinite temperature. Equivalently we may say that the joint probability density of $\underline{\Delta}$ and $\underline{\Xi}$ as defined by Eq. (3) does not relax to thermal equilibrium. In order to obtain the physically correct stationary solution, additional terms which have been neglected in Eq. (3) need to be found. We note that the SLE will relax to thermal equilibrium if we require that it obey the principle of detailed balance. We therefore seek additional terms to Eq. (3) subject to the constraint that detailed balance be obeyed.

It is convenient to incorporate $\underline{\Delta}$ and $\underline{\Xi}$ into a new set of augmented dynamical variables \underline{q} . The SLE may then be written as

$$\frac{\partial}{\partial t} P(\underline{q}, t; \lambda) = -\Gamma(\underline{q}) P(\underline{q}, t; \lambda) \quad (4)$$

If $\Gamma_{\underline{\Xi}}$ contains only first and second derivative terms, then $\Gamma(\underline{q})$ is of the form of an AFPE:

$$-\Gamma(\underline{q}) = -\sum_i \left(\frac{\partial}{\partial q_i} \right) \cdot K_i(\underline{q}; \lambda) + \frac{1}{2} \sum_{i,k} \left(\frac{\partial^2}{\partial q_i \partial q_k} \right) K_{ik}(\underline{q}; \lambda) \quad (5)$$

where the K_i and K_{ik} are drift and diffusion coefficients, respectively, and except for their time independence are otherwise quite general functions of \underline{q} and λ . $\Gamma_{\underline{\Xi}}$ may in general contain higher order derivative terms or be an integral operator. This will complicate the analysis below without adding any fundamentally new features. Irreversible and reversible drift coefficients may be respectively defined by

$$D_i(\underline{q}; \lambda) \equiv \frac{1}{2} [K_i(\underline{q}; \lambda) + \epsilon_i K_i(\tilde{\underline{q}}; \tilde{\lambda})] \quad (6)$$

$$\text{and } J_i(\underline{q}; \lambda) \equiv \frac{1}{2} [K_i(\underline{q}; \lambda) - \epsilon_i K_i(\tilde{\underline{q}}; \tilde{\lambda})] \quad (7)$$

$$\text{where } \tilde{\underline{q}} = \{\epsilon_1 q_1, \epsilon_2, \dots, \epsilon_n q_n\} \quad (8)$$

and $\epsilon_i = \pm 1$ depending on whether q_i changes sign upon time reversal. The necessary and sufficient conditions for detailed balance D_i are given by Haken as⁸

$$K_{ik}(\underline{q}; \lambda) = \epsilon_i \epsilon_k K_{ik}(\tilde{\underline{q}}; \tilde{\lambda}) \quad (9a)$$

$$D_i - \frac{1}{2} \sum_k \frac{\partial K_{ik}}{\partial q_k} = -\frac{1}{2} \sum_k K_{ik} \frac{\partial \phi}{\partial q_k} \quad (9b)$$

$$\text{and } \sum_i \left(\frac{\partial J_i}{\partial q_i} - J_i \frac{\partial \phi}{\partial q_i} \right) = 0, \quad (9c)$$

where $\phi(\underline{q}; \lambda)$ is the generalized thermodynamic potential defined by the stationary

solution of Eq. (4):

$$P_0(\underline{q}; \underline{\lambda}) = N \exp(-\Phi) \quad (9d)$$

with N a normalization constant. The AFPE is obtained by adding (or modifying) J_i and/or D_i terms to Eq. (5) so that Eqs. (9) are all fulfilled subject to a particular form for $\Phi(\underline{q}; \underline{\lambda})$ which must be determined by physical considerations. Since by Eq. (9c) the choice of additional J_i is not unique, physical considerations are also required here. In this manner, the back-reaction effects of Δ or Ξ are implicitly included in $\Gamma(\underline{q})$, and relaxation to thermal equilibrium is ensured.

It is now possible to generate the ALE from the complete $\Gamma(\underline{q})$.⁸ The ALE may be written for each dynamical variable q_ℓ as

$$\dot{q}_\ell = k_\ell(\underline{q}, \underline{\lambda}) + \sum_{j=1}^m g_{\ell j}(\underline{q}, \underline{\lambda}) \xi_j(t) \quad (10)$$

where the $\xi_j(t)$ are independent Gaussian δ -correlated random functions with

$$\langle \xi_j(t) \rangle = 0 \quad (11a)$$

$$\text{and } \langle \xi_\ell(t+\tau) \xi_j(t) \rangle = \delta_{\ell j} \delta(\tau) \quad (11b)$$

k_ℓ and $g_{\ell j}$ are related to the drift and diffusion coefficients by

$$K_\ell(\underline{q}; \underline{\lambda}) = k_\ell(\underline{q}; \underline{\lambda}) + \frac{1}{2} \sum_{k,j} (\partial g_{\ell j} / \partial q_k) g_{kj} \quad (12a)$$

$$\text{and } K_{\ell k}(\underline{q}; \underline{\lambda}) = \sum_j g_{\ell j}(\underline{q}; \underline{\lambda}) g_{kj}(\underline{q}, \underline{\lambda}) \quad (12b)$$

In general the matrix $K = (K_{\ell m})$ is symmetric. Usually K is also nonnegative definite. If K is nonnegative definite, then there exists a real symmetric nonnegative definite matrix $G = (g_{\ell k})$ such that $G^2 = K$. Then

$$G = U^{-1} [(K_1^0)^{1/2} \delta_{ij}] U \quad (13)$$

where $K_1^0 \delta_{ij}$ is the eigenvalue matrix of K and U is the corresponding eigenvector matrix. Then the $\tilde{K}_\ell(\underline{q}; \underline{\lambda})$ are obtained from Eq. (12a).

The method may be illustrated by a planar model of torque fluctuations.¹ Let γ , $\dot{\gamma}$, and I be the angular orientation, angular momentum and moment of inertia of the rotator respectively, while $T(\gamma, \dot{\gamma}, t)$ is the fluctuating part of the torque and $N(\gamma)$ is the mean-field torque. The Liouville equation for this stochastic process is just:

$$\frac{\partial}{\partial t} P(\gamma, \dot{\gamma}, t) = -[\dot{\gamma} \frac{\partial}{\partial \gamma} + I^{-1} \frac{\partial}{\partial \dot{\gamma}} (N(\gamma) + T(\gamma, \dot{\gamma}, t))] P(\gamma, \dot{\gamma}, t) \quad (14)$$

corresponding to the first term on the rhs of Eq. (3). A simple but useful form for modeling the stochastic properties of $T(\gamma, \dot{\gamma}, t)$ is to set it equal to $IV \sqrt{\frac{KT}{I}} f(\gamma - \phi)$ where $f(\gamma - \phi)$ is periodic in $(\gamma - \phi)$, and regard the angle $\phi(t)$ corresponding to the direction that minimizes the torque acting on the rotator as stochastic. If we choose a simple diffusional model in ϕ we may write:

$$\frac{\partial}{\partial t} P(\phi, t) = -\tau_\phi^{-1} \frac{\partial^2}{\partial \phi^2} P(\phi, t) \quad (15)$$

corresponding to Eq. (2). Eqs. (14) and (15) may then be combined as in Eq. (3), to give the SLE for $P(\gamma, \dot{\gamma}, \phi, t)$. This SLE is incomplete because it is found to violate

Eq. (9c) when we let the equilibrium Eq. (9d) be a Boltzmann distribution in $\dot{\gamma}$ and a Boltzmann distribution in the mean potential $U_N(\gamma)$, where $N(\gamma) = \frac{\partial}{\partial \gamma} U_N(\gamma)$. That is

$$\Phi(\gamma, \dot{\gamma}) = I\dot{\gamma}^2/2kT + U_N(\gamma)/kT \quad (16)$$

One readily finds that we may correct this by introducing a reversible drift term related to the fluctuation in ϕ of form $J_\phi = \sqrt{\frac{I}{kT}} Vg(\gamma-\phi)\dot{\gamma}$ where $f(\gamma-\phi) = \frac{\partial}{\partial \gamma} g(\gamma-\phi)$. This is the back-reaction of the rotator on the bath variable ϕ . Then the complete AFPE operator corresponding to Eqs. (4) and (5) becomes:

$$\Gamma(\gamma, \dot{\gamma}, \phi) = \dot{\gamma} \frac{\partial}{\partial \gamma} + I^{-1} N(\gamma) \frac{\partial}{\partial \gamma} + \sqrt{\frac{kT}{I}} V f(\gamma-\phi) \frac{\partial}{\partial \gamma} + \dot{\gamma} \sqrt{\frac{I}{kT}} V \frac{\partial}{\partial \phi} g(\gamma-\phi) - \tau_\phi^{-1} \frac{\partial^2}{\partial \phi^2} \quad (17)$$

For the simpler form of Eq. (18) with $N(\gamma) = 0$, the spectral densities:

$$j(\omega) = \frac{1}{2\pi} \text{Re} \langle \frac{1}{e} Y | [i\omega + \Gamma]^{-1} | P_0 e^{i\gamma} \rangle \quad (18)$$

(real part of the Fourier transform of the auto-correlation function $\langle e^{-i\gamma(t)} e^{+i\gamma(0)} \rangle$) were found to reduce to the Brownian form in the limit of large τ_ϕ^{-1} . For slower torque diffusion rates more complicated behavior was observed. In particular, for small ω the spectral density approaches the form $j(\omega) = \tau_R / (1 + \epsilon \omega^2 \tau_R^2)$ with $\epsilon > 1$, where τ_R is the Brownian rotational diffusion correlation time. Such a form has been found useful in the interpretation of molecular dynamics studies by ESR⁹⁻¹² and NMR¹³ methods. Care must be taken, however, in interpreting 3D results in terms of planar models because of periodicity effects.

When, however, the equilibrium of Eq. 9d is taken to the instantaneous value of the potential, i.e.

$$\Phi(\gamma, \dot{\gamma}, \phi) = I\dot{\gamma}^2/2kT + U_T(\gamma-\phi)/kT \quad (19)$$

where $U_T(\gamma-\phi) = \sqrt{kTI} Vg(\gamma-\phi)$ and we have let $N(\gamma)=0$ for simplicity, then one finds that Eq. (9b) is not obeyed. One readily finds that we must introduce the irreversible drift coefficient: $D_\phi = \tau_\phi^{-1} T(\gamma-\phi)/kT$, so that the AFPE operator is now:

$$\Gamma(\gamma, \dot{\gamma}, \phi) = \dot{\gamma} \frac{\partial}{\partial \gamma} + I^{-1} T(\gamma-\phi) \frac{\partial}{\partial \gamma} - \tau_\phi^{-1} \left(\frac{\partial^2}{\partial \phi^2} + \frac{\partial}{\partial \phi} \frac{T(\gamma-\phi)}{kT} \right) \quad (20)$$

This form will not reduce to the Brownian form in any limit, since the fluctuating torques do not lead to friction. The equivalent ALE corresponding to Eqs. 17 and 20 may be found in the paper of Stillman and Freed.¹ It was suggested that Eqs. (16) and (17) apply to more rapidly fluctuating torques, while Eqs. (19) and (20) apply to torque components fluctuating at a rate slower than the motion of the planar rotator.^{1, 9, 10, 14} This latter case has been called the slowly-relaxing local structure (SLRS) model,^{10, 14} and it has been useful in analyzing magnetic resonance relaxation studies on molecules in liquid crystals and model membranes. Appropriate 3D forms are given elsewhere.^{1, 14}

We now address a method of solution of the AFPE, Eqs. (4) and (5) appropriate for the general non-linear case (Eqs. 17 and 20 are non-linear). It is frequently convenient to "symmetrize" Γ by the similarity transformation:

$$\tilde{\Gamma} \equiv P_0(q; \lambda)^{-1/2} \Gamma P_0(q; \lambda)^{+1/2} \quad (21)$$

Then any correlation function can be written as:²

$$g(t) \equiv \overline{f^*(t)f(0)} = \langle f | e^{-\Gamma t} P_0 | f \rangle = \langle f P_0^{1/2} | e^{-\tilde{\Gamma} t} | f P_0^{1/2} \rangle \quad (22)$$

where $f(t) = f(\underline{A}(t))$. Therefore the correlation functions can be obtained once the eigenvalue problem for $\tilde{\Gamma}$ is solved. One may first expand the Hilbert space of the variables q in a complete orthonormal basis set so chosen that $\tilde{\Gamma}$ is a complex-symmetric matrix. One chooses $|f P_0^{1/2}\rangle \equiv |z\rangle$ as the starting vector and forms the $(\tilde{\Gamma})^n |z\rangle \equiv |z_{k+1}\rangle$ for $k=0$ to $n-1$. Then by Schmidt orthonormalization the linearly independent vectors $|z_{k+1}\rangle$ are transformed into an n -dimensional basis set represented by $|k\rangle$ for $k=1$ to n . One obtains the recursion relation characteristic of the Lanczos algorithm:²⁻⁴

$$\beta_k |k\rangle = (\tilde{\Gamma} - \alpha_{k-1} 1) |k-1\rangle - \beta_{k-1}^* |k-2\rangle$$

where $\alpha_k = \langle k | \tilde{\Gamma} | k \rangle = \langle k | \tilde{\Gamma}_n | k \rangle$ and $\beta_k = \langle k | \tilde{\Gamma} | k-1 \rangle = \langle k | \tilde{\Gamma}_n | k-1 \rangle$. Here $\tilde{\Gamma}_n$ is the n dimensional approximation of $\tilde{\Gamma}$ in the space spanned by the n $|k\rangle$'s. (That is, we may regard $\tilde{\Gamma}_n = P_n \tilde{\Gamma} P_n$ where P_n is the operator that projects any vector in the full Hilbert space onto the n -dimensional sub-space.) In the $|k\rangle$ representation $\tilde{\Gamma}_n$ is tri-diagonal (i.e. T_n) with k^{th} diagonal element given by α_k and off-diagonal elements given by β_k . Then the n^{th} approximation to the spectral density function of Eq. 22 is

$$j_n(\omega) \equiv \text{Re} \int_0^\infty e^{-i\omega t} g_n(t) dt = \text{Re} \langle \langle | [i\omega 1 + T_n]^{-1} | 1 \rangle \rangle \quad (23)$$

where $|1\rangle$ is the starting vector in the $|k\rangle$ representation. In this form $j_n(\omega)$ can be calculated as a continued fraction, or alternatively T_n can be diagonalized by standard methods. The Lanczos tri-diagonalization is extremely efficient for computations.² Furthermore, by means of Lanczos methods for (complex) non-symmetric matrices one can work directly with the unsymmetrized AFPE operator, Γ , in Eq. 22.

The above examples illustrate some of the applications of our AFPE approach. We believe that this method has application for Stochastic molecular dynamics calculations.¹⁵ In fact, our approach appears to us to be even more convenient than the present use of generalized Langevin equations (GLE), because the ALE have time-independent coefficients rather than the memory kernels of the GLE, and also non-linear couplings are easily included in either the AFPE or the ALE. The absence of memory kernels is due to the fact that the set of relevant variables has been extended to the required level to make this possible; i.e. the remaining "bath" variables are truly "irrelevant." The Mori⁶ concept of a hierarchy of memory functions in which one couples to increasingly faster variables, would have its analogue in the present method. Thus, one would construct a Markov process for the fastest variables, coupling in via the SLE the set of next fastest variables. This, in turn, would serve as the more general Markov process to which the next set of slower variables couple, etc. In this manner, one would go "up the chain" in model building. Lastly, we believe, our method offers the appeal of a direct approach to stochastic modeling such that the physical assumptions being made are quite transparent.

REFERENCES

*Supported by NSF Grants: CHE 8024124 and DMR 8102047

1. A.E. Stillman and J.H. Freed, *J. Chem. Phys.* 72, 550 (1980).
2. G. Moro and J.H. Freed, *J. Chem. Phys.* 74, 3757 (1981).
3. G. Moro and J.H. Freed, *J. Chem. Phys.* 75, 3157 (1981).
4. C. Lanczos, *J. Res. Natn. Bur. Stand.* 45, 255 (1950); 49, 33 (1952), B.N. Parlett, The Symmetric Eigenvalue Problem (Prentice Hall, N.J. 1980).
5. Yu V. Vorobyev, Method of Moments in Applied Mathematics (Gordon and Breach, N.Y. 1965).
6. H. Mori, *Prog. Theor. Phys* 34, 399 (1965).
7. N.G. van Kampen, *Phys. Rept.* 24c, 171 (1976); U. Frisch, in Probabilistic Methods in Applied Mathematics, edited by A.T. Bharucha-Reid (Academic, New York, 1968); A. Brissaud and U. Frisch, *J. Math. Phys.* 15, 524 (1974).
8. H. Haken, *Rev. Mod. Phys.* 47, 67 (1975).
9. J.S. Hwang, R.P. Mason, L.-P. Hwang, and J.H. Freed, *J. Phys. Chem.* 79, 489 (1975).
10. C.F. Polnaszek and J.H. Freed, *J. Phys. Chem.* 79, 2283 (1975).
11. J.S. Hwang, K.V.S. Rao, and J.H. Freed, *J. Phys. Chem.* 80, 1490 (1976); W.-J. Lin and J.H. Freed, *J. Phys. Chem.* 83, 379 (1979).
12. S.A. Zager and J.H. Freed, *J. Chem. Phys.* 77, 3344, 3360 (1982).
13. H.A. Lopes Cardozo, J. Bulthuis, J.H. Freed and W.M.M.J. Bovee, *Chem. Phys. Lett.* 60, 335 (1979).
14. J.H. Freed, *J. Chem. Phys.* 66, 4183 (1977).
15. S.A. Adelman and J.D. Doll, *Accts. Chem. Res.* 10, 378 (1977) and references cited therein.