Draft of L. Ingber, M.F. Wehner, G.M. Jabbour, and T.M. Barnhill, "Application of statistical mechanics methodology to term-structure bond-pricing models," *Mathl. Comput. Modelling* **15** (11), 1991, pp. 77-98.

Application of Statistical Mechanics Methodology to Term-Structure Bond-Pricing Models

Lester Ingber
Science Transfer Corporation
P.O. Box 857
McLean, VA 22101

Michael F. Wehner
B Division
Lawrence Livermore Laboratory
Livermore, CA 94550

George M. Jabbour and Theodore M. Barnhill
Department of Business Administration
The George Washington University
Washington, DC 20052

ABSTRACT

Recent work in statistical mechanics has developed new analytical and numerical techniques to solve coupled stochastic equations. This paper applies the very fast simulated re-annealing and path-integral methodologies to the estimation of the Brennan and Schwartz two-factor term structure model. It is shown that these methodologies can be utilized to estimate more complicated n-factor nonlinear models.

1. CURRENT MODELS OF TERM STRUCTURE

The modern theory of term structure of interest rates is based on equilibrium and arbitrage models in which bond prices are determined in terms of a few state variables. The one-factor models of Cox, Ingersoll and Ross (CIR) [1-4], and the two-factor models of Brennan and Schwartz (BS) [5-9] have been instrumental in the development of the valuation of interest dependent securities. The assumptions of these models include:

- Bond prices are functions of a number of state variables, one to several, that follow Markov processes.
- Investors are rational and prefer more wealth to less wealth.
- Investors have homogeneous expectations.
- No one investor is big enough to affect market prices.
- Taxes and transaction costs are not modeled.
- Information is costless and available to all.
- Assets are traded continuously at equilibrium prices.

However, it should be noted that the "efficient market" hypothesis is not quite universally accepted [10-12].

For one-factor models, the state variable is assumed to be the short-term interest rate which is assumed to follow a diffusion process or a continuous Markov process defined as follows:

$$dr = a(r,t)dt + b(r,t)dz \ ,$$

r = short-term rate,
t = calendar time,
$a(r,t)$ = expected instantaneous change in the short-term rate,
$b(r,t)$ = instantaneous volatility of the process,
dz is a Wiener process.

(1)

Bond prices are assumed to be a function of time and a proxy variable, the short-term rate.

Vasicek [4] assumes that

$$dr = K(\theta - r)dr + \sigma r dz \ . \tag{2}$$

CIR assume

$$dr = K(\theta - r)dr + \sigma r^{1/2} dz \ . \tag{3}$$

In both models, short-term rates are assumed to follow a mean-reverting process; in other words, the short-term rates have a mean of θ, K being the speed of adjustment. σr and $\sigma r^{1/2}$ are the standard deviations of the random component of the process for Vasicek and CIR models, respectively.

BS [5] extended the one factor model. They developed an arbitrage model of equilibrium interest rates based on the assumptions that the entire term structure at any point in time can be expressed as a function of two factors, being the short- and long-term rates of default free instruments. These interest rates are further assumed to follow a joint Wiener stochastic process. This process is of the form:

$$dr = \beta_1(r,l,t)dt + \eta_1(r,l,t)dz_1 \ ,$$

$$dl = \beta_2(r, l, t)dt + \eta_2(r, l, t)dz_2 \ , \tag{4}$$

where r and l are the short- and long-term rates, respectively. β_1 and β_2 are the expected instantaneous rates of change in the short-term and long-term rates respectively. η_1 and η_2 are the instantaneous standard deviations of the processes. dz_1 and dz_2 are Wiener processes, with expected values of zero and variance of dt with correlation coefficient ρ. That is,

$$E[dz_1] = E[dz_2] = 0 \ ,$$

$$E[dz_1^2] = E[dz_2^2] = dt \ , \ E[dz_1 dz_2] = \rho dt \ , \tag{5}$$

where $E[.] \equiv < . >$ is the expectation with respect to the joint Wiener process.

For estimation purposes, BS simplified and reduced this system to

$$dr = (a_1 + b_1(l - r))dt + r\sigma_1 dz_1 \ ,$$

$$dl = l(a_2 + b_2 r + c_2 l)dt + l\sigma_2 dz_2 \ , \tag{6}$$

where $\{a_1, b_1, a_2, b_2, c_2\}$ are parameters to be estimated.

These equations imply that the short-term rate converges towards a long-term rate ($b_1 > 0$). A result found in part of the BS study is $a_1 < 0$, i.e., the short-term rate could be negative. This is a potentially serious problem for this model.

The BS discrete-time representation is

$$\frac{r_{t+1} - r_t}{r_t} = \frac{a_1}{r_t} + b_1(\frac{l_t}{r_t} - 1) + e_{1t} \ ,$$

$$\frac{l_{t+1} - l_t}{l_t} = a_2 + b_2 r_t + c_2 l_t + e_{2t} \ , \tag{7}$$

where e_{1t} and e_{2t} are normally-distributed bivariate random variables with standard deviations of σ_1 and σ_2 respectively with a correlation coefficient of ρ. As discussed below, the above continuous model is only well defined in the differential limit of these discrete equations.

These equations imply that the short-term rate converges towards the long-term rate ($b_1 > 0$). A result found in part of the BS study is that $a_1 < 0$. i.e., the short-term rate could be negative if not subject to an additional external boundary condition. This is a potentially serious problem for this model. BS used the iterative Aitken procedure [13] to estimate the parameters of the model. They found a_1 to be negative (the long-term rate has to be larger than 2% to avoid negative short-term rates). The correlation coefficient ρ, and the coefficients a_1 and b_1 were unstable. In addition, BS found a negative serial correlation in the error terms e_{1t} and e_{2t}, which led them to believe that additional state variables should be added to the model.

Using methods of stochastic calculus [5], BS further derived a partial differential equation for bond prices as the maturity date is approached.

$$\frac{\partial}{\partial \tau} B = (-r + f^r \frac{\partial}{\partial r} + f^l \frac{\partial}{\partial l} + g^{rr} \frac{\partial}{\partial r^2} + g^{rl} \frac{\partial}{\partial r \partial l} + g^{ll} \frac{\partial}{\partial l^2})B$$

$$= AB \ , \tag{8}$$

where the coefficients $\{f, g\}$ depend on r and l, $\tau = T - t$ for t calendar time and T the time of maturity, and A can be considered as a differential operator on B.

It may help to appreciate the importance of the BS methodology by discretizing the above partial differential equation for B, in a "mean-value" limit. That is, at a given calendar time t indexed by s, noting that $\partial/\partial\tau = -\partial/\partial t$, take

$$0 = f^r \frac{\partial B_s}{\partial r} = f^l \frac{\partial B_s}{\partial l} \ ,$$

$$0 = g^{rr} \frac{\partial B_s}{\partial r^2} = g^{rl} \frac{\partial B_s}{\partial r \partial l} = g^{ll} \frac{\partial B_s}{\partial l^2} ,$$

$$B_s - B_{s+1} = -r_s B_s . \tag{9}$$

This yields the popular expectations-hypothesis spot-interest estimate of bond prices, working backwards from maturity,

$$B_s = (1 + r_s)^{-1} B_{s+1} . \tag{10}$$

The important generalization afforded by BS is to include information about r and l and treat them as stochastic variables with drifts and diffusions. Then, this discretized treatment yields

$$B_{s\,rl} = (1 - A_{s\,rlr'l'})^{-1} B_{s+1\,r'l'} , \tag{11}$$

where the operator inverse of the differential operator A has been formally written, and its dependence on intermediate values of r' and l' has been explicitly portrayed. Their discretized calculation of their partial differential equation, and our discretized calculation of the path-integral representation of this model, essentially are mathematical and numerical methods of calculating this evolution of B_s.

In this paper we present an alternative methodology of very fast simulated re-annealing (VFSR) [14] to compute the parameters of the BS model. It is also shown that the VFSR methodology is capable of handling more complicated n-factor non- linear models.

The advantages of using the simulated annealing methodology are: (1) Global minima in parameter space are relatively more certain than with regression fitting. (2) All parameters, including parameters in the noise, are simultaneously and equally treated in the fits, i.e., different statistical methods are not being used to estimate the deterministic parameters, then to go on to estimate noise parameters. (3) Boundary conditions on the variables can be explicitly included in the fitting process, a process not included in standard regression fits. (4) We can efficiently extend our methodology to develop 3-state and higher models, including higher order nonlinearities.

We also present an alternative method of calculating the evolution of bond prices. Our particular non-Monte Carlo path-integral technique has proven to be extremely accurate and efficient for a variety of nonlinear systems [15,16]. The method of path integration is more accurate and efficient for calculating the evolution of B as a function of the stochastic variables r and l. To mention a few advantages: (1) A variable mesh is calculated in terms of the underlying nonlinearities. (2) Initial conditions and boundary conditions typically are more easily implemented with integral, rather that with differential, equations, e.g., by using the method of images. (3) Integration is inherently a "smoothing" process, whereas differentiation is a "sharpening" process. This means that we can handle "stiff" and nonlinear problems with more ease.

We also comment below on how our methodology can be applied to other future-price models under development [17-19].

In Section 2, we give a brief theoretical description of mathematically equivalent representations of multivariate stochastic systems. These methods, just recently developed by mathematical physicists in the last decade in the context of "statistical mechanics," permit the introduction of even more recently developed numerical algorithms. They were proposed for financial systems in a previous paper [20].

In Section 3, we apply this formalism to specific computations of the BS model. Section 4 gives our numerical results. Section 5 presents our conclusions. Appendix A gives a derivation of the stochastic calculus used.

2. DEVELOPMENT OF MATHEMATICAL METHODOLOGY

2.1. Background

Aggregation problems in nonlinear nonequilibrium systems typically are "solved" (accommodated) by having new entities/languages developed at these disparate scales in order to efficiently pass information back and forth [21,22]. This is quite different from the nature of quasi-

equilibrium quasi-linear systems, where thermodynamic or cybernetic approaches are possible. These approaches typically fail for nonequilibrium nonlinear systems.

In the late 1970's, mathematical physicists discovered that they could develop statistical mechanical theories from algebraic functional forms

$$dr/dt = f_r(r,l) + \sum_i \hat{g}_r^i(r,l)\eta_i \ ,$$

$$dl/dt = f_l(r,l) + \sum_i \hat{g}_l^i(r,l)\eta_i \ , \tag{12}$$

where the $\hat{g}$'s and f's are general nonlinear algebraic functions of the variables r and l. These equations represent differential limits of discretized stochastic difference equations, e.g., Wiener noise $dW \rightarrow \eta dt$ [23]. The resulting stochastic differential equations (s.d.e.'s) are referred to as Langevin equations [23-28]. The f's are referred to as the (deterministic) drifts, and the square of the $\hat{g}$'s are related to the diffusions (fluctuations or volatilities). In fact, the statistical mechanics can be developed for any number of variables, not just two. The η's are sources of Gaussian-Markovian noise, often referred to as "white noise." The inclusion of the $\hat{g}$'s, called "multiplicative" noise, recently has been shown to very well mathematically and physically model other forms of noise, e.g., shot noise, colored noise, dichotomic noise [29-32]. Finite-jumps diffusions also can be included [33].

These new methods of nonlinear statistical mechanics only recently have been applied to complex large-scale physical problems, demonstrating that observed data can be described by the use of these algebraic functional forms. Success was gained for large-scale systems in neuroscience, in a series of papers on statistical mechanics of neocortical interactions [34-38], and in nuclear physics [39,40]. This methodology has been used for problems in combat analyses [16,22,41,42]. These methods were suggested for financial markets [20], and this paper is an application of that approach to estimating term structure models.

Thus, now we can investigate various choices of f's and $\hat{g}$'s to test algebraic functional forms. In science, this is a standard phenomenological approach to discovering and encoding knowledge and observed data, i.e., fitting algebraic functional forms which lend themselves to empirical interpretation. This gives more confidence when extrapolating to new scenarios, exactly the issue in building confidence in financial models.

The utility of these algebraic functional forms goes further beyond their being able to fit sets of data. There is an equivalent representation to the Langevin equations, called a "path-integral" representation for the long-time probability distribution of the variables. This short-time probability distribution is driven by a "Lagrangian," which can be thought of as a dynamic algebraic "cost" function. The path-integral representation for the long-time distribution possesses a variational principle, which means that simple graphs of the algebraic cost-function give a correct intuitive view of the most likely states of the variables, and of their statistical moments, e.g., heights being first moments (likely states) and widths being second moments (uncertainties). Like a ball bouncing about a terrain of hills and valleys, one can quickly visualize the nature of dynamically unfolding r and l states.

Especially because we are trying to mathematically model sparse and poor data, different drift and diffusion algebraic functions can give approximately the same algebraic cost-function when fitting short-time probability distributions to data. The calculation of long-time distributions permits a clear choice of the best algebraic functions, i.e., those which best follow the data through a predetermined long epoch of trading. Afterwards, if there are closely competitive algebraic functions, they can be more precisely assessed by calculating higher algebraic correlation functions from the probability distribution.

As discussed previously, the mathematical representation most familiar to other modelers is a system of stochastic rate equations, often referred to as Langevin equations. From the Langevin equations, other models may be derived, such as the times-series model and the Kalman filter method of control theory. However, in the process of this transformation, the Markovian description typically is lost by projection onto a smaller state space [43,44].

2.2. Fitting Parameters

For example, for Θ variables, these coupled stochastic differential equations can be represented equivalently by a short-time conditional probability distribution, P, in terms of the Lagrangian, L:

$$P = \det(\sigma)^{-1/2}(2\pi\Delta t)^{-\Theta/2}\exp(-L\Delta t) \ . \tag{13}$$

The form for the Lagrangian, L, and the determinant of the metric, σ, is

$$L = \sum_G \sum_{G'} \frac{(dM^G/dt - g^G)(dM^{G'}/dt - g^{G'})}{2g^{GG'}} \ ,$$

$$\sigma = \det(g_{GG'}) \ , (g_{GG'}) = (g^{GG'})^{-1} \ ,$$

$$g^{GG'} = \sum_i \hat{g}_i^G \hat{g}_i^{G'} \ , \tag{14}$$

where G and G' run over all Θ variables. Here, the prepoint discretization is used, which hides the Riemannian corrections explicit in the midpoint discretized Feynman Lagrangian; only the latter representation possesses a variational principle useful for arbitrary noise [20,24].

This defines a scalar "dynamic cost function," C, in terms of parameters, e.g., generically represented as $C(\tilde{\alpha})$,

$$C(\tilde{\alpha}) = L\Delta t + \frac{\Theta}{2}\ln(2\pi\Delta t) + \frac{1}{2}\ln\sigma \ , \tag{15}$$

which can be used with the VFSR algorithm, discussed below [14], to find the (statistically) best fit of parameters , e.g., identified by $\{\tilde{\alpha}\}$, to the data. The cost function for a given system is obtained by the product of P's over all data epochs, i.e., a sum of C's is obtained. In the actual VFSR code, C is "normalized" by dividing by the number of epochs. Then, since we essentially are performing a maximum likelihood fit, the cost functions obtained from somewhat different theories or data can provide a relative statistical measure of their likelihood, e.g., $P_{12} \sim \exp(C_2 - C_1)$.

If there are competing mathematical forms, then it is advantageous to utilize the path-integral to calculate the long-time evolution of P [16,22]. Experience has demonstrated that the long-time correlations derived from theory, measured against the observed data, is a viable and expedient way of rejecting models not in accord with observed evidence.

Note that the use of the path integral is *a posteriori* to and independent of the short-time fitting process, and is a subsidiary physical constraint on the mathematical models to judge their internal soundness and suitability for attempts to extrapolate to other trading scenarios.

2.3. Algebraic Complexity Yields Simple Intuitive Results

Consider a multivariate system with variance a general nonlinear function of the variables. The Einstein summation convention helps to compact the equations, whereby repeated indices in factors are to be summed over.

The Itô (prepoint) discretization for a system of stochastic differential equations is defined by

$$\bar{t}_s \in [t_s, t_s + \Delta t] \equiv [t_s, t_{s+1}] \ ,$$

$$M(\bar{t}_s) = M(t_s) \ ,$$

$$dM(\bar{t}_s)/dt = M(t_{s+1}) - M(t_s) \ . \tag{16}$$

The stochastic equations are then written as

$$dM^G/dt = f^G + \hat{g}_i^G \eta^i \ ,$$

$$i = 1, \cdots, \Xi \ ,$$

$$G = 1, \cdots, \Theta \ . \tag{17}$$

The operator ordering (of the $\partial/\partial M^G$ operators) in the Fokker-Planck equation corresponding to this discretization is

$$\frac{\partial P}{\partial t} = VP + \frac{\partial(-g^G P)}{\partial M^G} + \frac{1}{2}\frac{\partial^2(g^{GG'}P)}{\partial M^G \partial M^{G'}} \ ,$$

$$g^G = f^G + \frac{1}{2}\hat{g}_i^{G'}\frac{\partial \hat{g}_i^G}{\partial M^{G'}} \ ,$$

$$g^{GG'} = \hat{g}_i^G \hat{g}_i^{G'} \ . \tag{18}$$

where a "potential" V is present in some systems.

The Lagrangian corresponding to this Fokker-Planck and set of Langevin equations may be written in the Stratonovich (midpoint) representation, corresponding to

$$M(\bar{t}_s) = \frac{1}{2}\left[M(t_{s+1}) + M(t_s)\right] \ . \tag{19}$$

This discretization can be used to define a Feynman Lagrangian L which possesses a variational principle, and which explicitly portrays the underlying Riemannian geometry induced by the metric tensor $g_{GG'}$, calculated to be the inverse of the covariance matrix [20]. More details are given in Appendix A and in another paper on this subject [45].

$$P = \int \cdots \int \underline{D}M \exp(-\sum_{s=0}^{u} \Delta t L_s) \ ,$$

$$\underline{D}M = g_{0_+}^{1/2}(2\pi\Delta t)^{-\Theta/2}\prod_{s=1}^{u} g_{s_+}^{1/2}\prod_{G=1}^{\Theta}(2\pi\Delta t)^{-1/2}dM_s^G \ ,$$

$$\int dM_s^G \rightarrow \sum_{\iota=1}^{N^G}\Delta M_{\iota s}^G \ , M_0^G = M_{t_0}^G \ , M_{u+1}^G = M_t^G \ ,$$

$$L = \frac{1}{2}(dM^G/dt - h^G)g_{GG'}(dM^{G'}/dt - h^{G'}) + \frac{1}{2}h^G{}_{;G} + R/6 - V \ ,$$

$$[\cdots]_{,G} = \frac{\partial[\cdots]}{\partial M^G} \ ,$$

$$h^G = g^G - \frac{1}{2}g^{-1/2}(g^{1/2}g^{GG'})_{,G'} \ ,$$

$$g_{GG'} = (g^{GG'})^{-1} \ ,$$

$$g_s[M^G(\bar{t}_s), \bar{t}_s] = \det(g_{GG'})_s \ , g_{s_+} = g_s[M_{s+1}^G, \bar{t}_s] \ ,$$

$$h^G{}_{;G} = h_{,G}^G + \Gamma_{GF}^F h^G = g^{-1/2}(g^{1/2}h^G)_{,G} \ ,$$

$$\Gamma_{JK}^F \equiv g^{LF}[JK, L] = g^{LF}(g_{JL,K} + g_{KL,J} - g_{JK,L}) \ ,$$

$$R = g^{JL}R_{JL} = g^{JL}g^{JK}R_{FJKL} \ ,$$

$$R_{FJKL} = \frac{1}{2}(g_{FK,JL} - g_{JK,FL} - g_{FL,JK} + g_{JL,FK}) + g_{MN}(\Gamma_{FK}^M\Gamma_{JL}^N - \Gamma_{FL}^M\Gamma_{JK}^N) \ , \tag{20}$$

where R is the Riemannian curvature, and we also have explicitly noted the discretization in the mesh of $M_{\iota s}^G$ by ι, to be discussed further below.

A prepoint discretization for the same probability distribution P, gives a much simpler algebraic form,

$$M(\bar{t}_s) = M(t_s) \ ,$$

$$L' = \frac{1}{2}(dM^G/dt - g^G)g_{GG'}(dM^{G'}/dt - g^{G'}) - V \ , \tag{21}$$

but the Lagrangian L' so specified does not satisfy a variational principle useful for moderate to large noise. Still, this prepoint-discretized form has been quite useful in all systems examined thus far, simply requiring a somewhat finer numerical mesh.

It must be emphasized that the output need not be confined to complex algebraic forms or tables of numbers. Because L possesses a variational principle, sets of contour graphs, at different long-time epochs of the path-integral of P over its variables at all intermediate times, give a visually intuitive and accurate decision-aid to view the dynamic evolution of the scenario. For example, this Lagrangian approach permits a quantitative assessment of concepts usually only loosely defined.

$$\text{``Momentum''} = \Pi^G = \frac{\partial L}{\partial(\partial M^G/\partial t)} \ ,$$

$$\text{``Mass''} = g_{GG'} = \frac{\partial L}{\partial(\partial M^G/\partial t)\partial(\partial M^{G'}/\partial t)} \ ,$$

$$\text{``Force''} = \frac{\partial L}{\partial M^G} \ ,$$

$$\text{``}F = ma\text{''}: \ \delta L = 0 = \frac{\partial L}{\partial M^G} - \frac{\partial}{\partial t}\frac{\partial L}{\partial(\partial M^G/\partial t)} \ , \tag{22}$$

where M^G are the variables and L is the Lagrangian. These physical entities provide another form of intuitive, but quantitatively precise, presentation of these analyses. For example, daily newspapers use this terminology to discuss the movement of security prices.

2.4. Numerical Methodology

Recently, two major computer codes have been developed, which are key tools for the use of this approach to estimate model parameters and price bonds.

The first code, very fast simulated re-annealing (VFSR) [14], fits short-time probability distributions to observed data, using a maximum likelihood technique on the Lagrangian. An algorithm of very fast simulated re-annealing has been developed to fit observed data to a theoretical cost function over a D-dimensional parameter space [14], adapting for varying sensitivities of parameters during the fit. The annealing schedule for the "temperatures" (artificial fluctuation parameters) T_i decrease exponentially in "time" (cycle-number of iterative process) k, i.e., $T_i = T_{i0}\exp(-c_i k^{1/D})$.

Heuristic arguments have been developed to demonstrate that this algorithm is faster than the fast Cauchy annealing [46], $T_i = T_0/k$, and much faster than Boltzmann annealing [47], $T_i = T_0/\ln k$. To be more specific, the kth estimate of parameter α^i,

$$\alpha^i_k \in [A_i, B_i] \ , \tag{23}$$

is used with the random variable x^i to get the $k+1$th estimate,

$$\alpha^i_{k+1} = \alpha^i_k + x^i(B_i - A_i) \ ,$$

$$x^i \in [-1, 1] \ . \tag{24}$$

The generating function is defined as

$$g_T(x) = \prod_{i=1}^{D} \frac{1}{2\ln(1 + 1/T_i)(|x^i| + T_i)} \equiv \prod_{i=1}^{D} g^i_T(x^i) \ ,$$

$$T_i = T_{i0} \exp(-c_i k^{1/D}) \ . \tag{25}$$

Note that the use of C, the cost function given above, is *not* equivalent to doing a simple least squares fit on $M(t + \Delta t)$.

The second code develops the long-time probability distribution from the Lagrangian fit by the first code. A robust and accurate histogram-based (non-Monte Carlo) path-integral algorithm to calculate the long-time probability distribution has been developed to handle nonlinear Lagrangians [15,16,48,49], including a two-variable code for additive and multiplicative cases.

The histogram procedure recognizes that the distribution can be numerically approximated to a high degree of accuracy as sum of rectangles at points M_i of height P_i and width ΔM_i. For convenience, just consider a one-dimensional system. The above path-integral representation can be rewritten, for each of its intermediate integrals, as

$$\begin{aligned} P(M; t + \Delta t) &= \int dM' [g_s^{1/2} (2\pi \Delta t)^{-1/2} \exp(-L_s \Delta t)] P(M'; t) \\ &= \int dM' G(M, M'; \Delta t) P(M'; t) \ , \end{aligned}$$

$$P(M; t) = \sum_{i=1}^{N} \pi(M - M_i) P_i(t)$$

$$\pi(M - M_i) = \begin{cases} 0 \ , \ (M_i - \frac{1}{2} \Delta M_{i-1}) \le M \le (M_i + \frac{1}{2} \Delta M_i) \ , \\ 1 \ , \ \text{otherwise} \ , \end{cases} \tag{26}$$

which yields

$$\begin{aligned} P_i(t + \Delta t) &= T_{ij}(\Delta t) P_j(t) \ , \\ T_{ij}(\Delta t) &= \frac{2}{\Delta M_{i-1} + \Delta M_i} \int_{M_i - \Delta M_{i-1}/2}^{M_i + \Delta M_i/2} dM \int_{M_j - \Delta M_{j-1}/2}^{M_j + \Delta M_j/2} dM' G(M, M'; \Delta t) \ . \end{aligned} \tag{27}$$

T_{ij} is a banded matrix representing the Gaussian nature of the short-time probability centered about the (varying) drift.

This histogram procedure has been extended to two dimensions, i.e., using a matrix T_{ijkl} [16], e.g., essentially similar to the use of the A matrix in the previous section. Explicit dependence of L on time t also can be included without complications. We see no problems in extending it to other dimensions, other than care must be used in developing the mesh in ΔM, which is dependent on the diffusion matrix.

Fitting data with the short-time probability distribution, effectively using an integral over this epoch, permits the use of coarser meshes than the corresponding stochastic differential equation. The coarser resolution is appropriate, typically required, for numerical solution of the time-dependent path-integral: By considering the contributions to the first and second moments of ΔM^G for small time slices θ, conditions on the time and variable meshes can be derived [48]. The time slice essentially is determined by $\theta \le \bar{L}^{-1}$, where $\bar{L}$ is the "static" Lagrangian with $dM^G/dt = 0$, throughout the ranges of M^G giving the most important contributions to the probability distribution P. The variable mesh, a function of M^G, is optimally chosen such that ΔM^G is measured by the covariance $g^{GG'}$, or $\Delta M^G \sim (g^{GG} \theta)^{1/2}$.

2.5. Chaos or Noise?

Given the context of current studies in complex nonlinear systems [50,51], the question can be asked: What if markets have chaotic mechanisms that overshadow the above stochastic considerations? The real issue is whether the scatter in data can be distinguished between being due to noise or chaos. Several studies have been proposed with regard to comparing chaos to simple filtered (colored) noise [J. Theiler, private communication] [51-53].

The previous references must be generalized, such that we must investigate whether scatter in markets' data can be distinguished from multiplicative noise. A previous application of this methodology follows:

One of us (LI) was principal investigator of a US Army Models Committee project, working with a team of Army and Lawrence Livermore National Laboratory personnel to compare, for the first time, large-scale high-fidelity combat computer-model data to exercise data [16,41]. The combat analysis was possible only because now we had recent data on combat exercises from the National Training Center (NTC) of sufficient temporal density to attempt dynamical mathematical modeling. The criteria used to (not) determine chaos in this dynamical system is the nature of propagation of uncertainty, i.e., the variance. For example, following by-now standard arguments [J. Yorke, seminar and private communication], propagation of uncertainty may be considered as (a) diminishing, (b) increasing additively, (c) or increasing multiplicatively. An example of (a) is the evolution of a system to an attractor, e.g., a book dropped onto the floor from various heights reaches the same point no matter what the spread in initial conditions. An example of (b) is the propagation of error in a clock, a cyclic system. Examples of (c) are chaotic systems, of which very few real systems have been shown to belong. An example of (c) is the scattering of a particle in a box whose center contains a sphere boundary: When a spread of initial conditions is considered for the particle to scatter from the sphere, when its trajectories are aligned to strike the sphere at a distance from its center greater that the diameter, the spread in scattering is a factor of about three greater than the initial spread.

In our analysis of NTC data, we were able to fit the short-time attrition epochs (determined to be about 5 minutes from mesh considerations determined by the nature of the Lagrangian) with short-time nonlinear Gaussian-Markovian probability distributions with a resolution comparable to the spread in data. When we did the long-time path-integral from some point at the beginning of the battle, we found that we could readily find a form of the Lagrangian that made physical sense and that also fit the multivariate variances as well as the means at each point in time of the rest of the combat interval. I.e., there was not any degree of hyper-sensitivity to initial conditions that prevented us from "predicting" the long time means and variances of the system. Since the system is dissipative, there is a strong tendency for all moments to diminish in time, but in fact this combat is of sufficiently modest duration (typically 1 to 2 hours) that variances do increase somewhat during the middle of the battle. In summary, this battalion-regiment scale of this particular battle did not seem to possess chaos.

Similar considerations and calculations are planned for these studies of financial markets.

3. BRENNAN-SCHWARTZ MODELS

3.1. Interest Rates

The pioneering Brennan-Schwartz (BS) model [5,7] can be used to illustrate how this methodology is to be implemented numerically.

The BS model is summarized by:

$$dr = [a_1 + b_1(l - r)]dt + r\sigma_1 dz_1 \ ,$$

$$dl = [l(a_2 + b_2 r + c_2 l)]dt + l\sigma_2 dz_2 \ ,$$

$$< dz_i >= 0 \ , \ i = \{1, 2\} \ ,$$

$$< dz_i(t) dz_j(t') >= dt\delta(t - t') \ , \ i = j \ ,$$

$$< dz_i(t) dz_j(t') >= \rho dt\delta(t - t') \ , \ i \neq j \ ,$$

$$\delta(t - t') = \begin{cases} 0 \ , \ , \ \ t \neq t' \ , \\ 1 \ , \ \ \ t = t' \ , \end{cases} \tag{28}$$

where < . > denotes expectations.

These can be rewritten as Langevin equations (in the Itô prepoint discretization)

$$dr/dt = a_1 + b_1(l - r) + \sigma_1 r(\gamma^+ n_1 + \mathrm{sgn}\rho\, \gamma^- n_2) \ ,$$

$$dl/dt = l(a_2 + b_2 r + c_2 l) + \sigma_2 l(\mathrm{sgn}\rho\, \gamma^- n_1 + \gamma^+ n_2) \ ,$$

$$\gamma^\pm = \frac{1}{\sqrt{2}}\,[1 \pm (1 - \rho^2)^{1/2}]^{1/2} \ ,$$

$$n_i = (dt)^{1/2} p_i \ , \tag{29}$$

where p_1 and p_2 are independent [0,1] Gaussian distributions.

The cost function C is defined from the equivalent short-time probability distribution, P, for the above set of equations.

$$P = g^{1/2}(2\pi dt)^{-1/2} \exp(-Ldt)$$

$$= \exp(-C) \ ,$$

$$C = Ldt + \frac{1}{2}\ln(2\pi dt) - \ln(g) \ ,$$

$$L = \frac{1}{2} F^\dagger \underline{g} F \ ,$$

$$F = \begin{pmatrix} dr/dt - (a_1 + b_1(l - r)) \\ dl/dt - l(a_2 + b_2 r + c_2 l) \end{pmatrix},$$

$$g = \det(\underline{g}) \ ,$$

$$k = 1 - \rho^2 \ . \tag{30}$$

$\underline{g}$, the metric in $\{r, l\}$-space, is the inverse of the covariance matrix,

$$\underline{g}^{-1} = \begin{pmatrix} (r\sigma_1)^2 & \rho r l \sigma_1 \sigma_2 \\ \rho r l \sigma_1 \sigma_2 & (l\sigma_2)^2 \end{pmatrix}. \tag{31}$$

As discussed above, the correct mesh for time, dt, in order that P represent the Langevin equations (to order $dt^{3/2}$) is

$$dt \le 1/\bar{L} \ , \tag{32}$$

where $\bar{L}$ is L evaluated with $ds/dt = dl/dt = 0$. If dt is greater than $1/\bar{L}$, then it is inappropriate to use P, and instead the path integral over intermediate states of folded short-time distributions must be calculated. In this context, it should be noted that the correct time mesh for the corresponding differential equations must be at least as small, since typically differentiation is a "sharpening" process. This will be noted in any discipline requiring numerical calculation, when comparing differential and integral representations of the same system.

The VFSR code was checked out with a rather stringent test: Data was "generated" using the BS equations above as a simulation, using a set of parameters given in the 1982 BS paper. Then, the Lagrangian was used as a cost function to search for the parameters used in the generation of the data. A time mesh was established at each point in time using the criteria given above, which turns out to be $dt \sim 1$ day. The code converged to within the statistical accuracy of the generated data. When a time mesh of $dt = 1$ month was used, as did BS, the results did not match the simulated data.

In light of the above, it should not be surprising that computer runs with real Treasury bills and bonds over the same epochs as BS, using end-of-month data, did not agree precisely with BS or with similar runs using daily data. Note we are not using the exact data as BS. However, should we be

concerned about the lack of excellent convergence with BS? As determined from C' above, the $dt \sim 1$ day, possibly 1 week, is appropriate to use to define the short-time probability distribution, data not available in the 1982 BS study.

However, $dt \sim 1$ day, not available for the BS work, shows marked deviations in the development of interest variables for selected trajectories using the BS equations as a simulation of data. This suggests that a two-state model is insufficient to capture faster movements in rates at the daily scale.

When the s.d.e. were permitted to evolve, trajectories developed increasing interest rates. It was noted that when the rates approached 100%, there suddenly was extremely wild growth of these variables. In the context of chaos discussed above, this region will be further investigated.

3.2. Security Prices

BS [5] present arguments recognizing that the stochastic price of a discount bond for a given maturity date T can utilize straightforward stochastic calculus to derive a form in terms of coefficients appearing in their $r - l$ coupled stochastic equations. They use arbitrage arguments on portfolios of bonds with different maturity dates to derive zero risk conditions for the market prices of risks, λ_1 and λ_2, for short-term and long-term interest rates, respectively. By considering l as related to a bond's price, they straightforwardly derive an arbitrage expression for λ_2. Their resulting partial differential equation (p.d.e.) is an equilibrium (mean value) equation for a pure discount-bond price B, at a given time until maturity $\tau = T - t$ and "continuous" coupon payment of c.

The above formulation of interest rates is used by BS to determine the parameters needed to calculate their derived p.d.e. for securities, i.e., bond prices B. Using some notation developed above, with $\{M^G; G = r, l\}$, they obtain

$$\begin{aligned}
\frac{\partial B}{\partial \tau} &= VB + \frac{\partial(-g^G B)}{\partial M^G} + \frac{1}{2}\frac{\partial^2(g^{GG'}B)}{\partial M^G \partial M^{G'}} \ , \\
g^r &= -(\beta_1 - \lambda_1 \eta_1) \\
&= -a_1 - b_1(l - r) + \lambda_1 r \sigma_1 \ , \\
g^l &= -(\beta_2 - \lambda_2 \eta_2) \\
&= -l(\sigma_2^2 + l - r) \ , \\
(g^{GG'}) &= (\underline{g})^{-1} \ , \\
V &= \frac{c}{B} - r \ ,
\end{aligned} \tag{33}$$

where c is the continuous coupon rate for bond B, and λ_1 is an additional parameter to be fit by the data.

The above equation represents a "truly nonlinear" Fokker-Planck equation because of the presence of B in V. However, if c/B is a smooth function, such that

$$\begin{aligned}
\frac{V(M^G;\tau') - V(M^G;\tau)}{\varepsilon} &= \Delta\tau \frac{\partial V}{\partial \tau} + \Delta\tau \int \frac{\delta V}{\delta B}\frac{\partial B(M'^{G'})}{\delta \tau} dM'^{G'} \\
&= O(\Delta\tau^{\nu}) \ ,
\end{aligned} \tag{34}$$

for $\nu > 1$, where $\tau' = \tau + \varepsilon\Delta\tau$, then our numerical path-integral codes may be used here as well [15].

In this formulation, all the above algebraic and numerical methodology can be utilized to define a Lagrangian-like function, L_B, defining the evolution of B, subject to whatever initial conditions and boundary conditions are deemed appropriate, e.g., if considering bonds or options [7,9]. Care must be taken with the discretization in the forward τ direction.

In particular, it was tempting to use our simulated-annealing codes to fit L_B; then our path-integral codes can be used to evolve/predict B for long epochs. The use of L_B as a cost function for fitting data is still reasonable even though B is not a *bona fide* E.g., B is hypothesized to be determined by intense arbitrage, in the derivation of the p.d.e. This is equivalent to stating that the process is confined, within statistical fluctuations, to a maximum likelihood path. This is mathematically articulated in our methodology as stating that the variational principle associated (not independently assumed) with L_B, including the specification of the boundary conditions and initial conditions, implies that we can fit the variables in L_B to data specified by observed values. This is numerically similar to the way we perform a maximum likelihood fit of the Lagrangian associated with the s.d.e. As we show below, this assumption appears to be well supported in our fits of L_B to observed $r - l$ data, in that we get a set parameters very close to those obtained by fitting the L associated with the s.d.e. It should be noted that BS calculated their λ_1 by interpolating fits to bond prices, thereby incorporating lots of pricing information into λ_1.

However, independent of such arguments, once the parameters are established, it is mathematically correct to use our path-integral codes on the B equation, with its initial conditions and boundary conditions, to calculate the long-time evolution of B.

We note that our use of L_B to describe the evolution of B is similar in spirit to recent attempts to describe the evolution of bond prices as a functional of forward rates, for example by Heath, Jarrow and Morton (HJM) [18,19]. Our functional corresponding to theirs is simply the Lagrangian L_B. HJM develop arbitrage arguments to impose equilibrium evolution by randomly varying their interest-rate functional, somewhat similar to the numerical importance-sampling methodology employed when performing numerical Monte Carlo techniques to take advantage of an underlying variational principle [14,54]. The trade-off in developing a theory of stochastic forward pricing is that parabolic equations are unstable for the negative diffusion in calendar time t so defined for financial systems, and so HJM must impose additional constraints to prevent this explosion. The BS technique of developing B as a function of $\tau = T - t$ avoids these problems. The methodology of HJM does not require the separate estimation of market risk factors, e.g., λ_1. Our Lagrangian representation of the BS model, coupled with our arbitrage arguments, also permits us to fit all parameters directly to interest rate data. The methodology of HJM permits the introduction of colored (time delayed) noise, which also can be included in our methodology, albeit with substantial effort. In principle, given a forward rate HJM model for a Lagrangian, then all our methodology could be used here as well.

In many cases it may be of interest to study the stochastic systems of interest rates, e.g., the coupled $r - l$ equations above. However, if only the evolution of B is of interest, then it is more direct, and likely more numerically accurate, to fit the parameters in the cost function C_B (derived from L_B as C was derived from L above). I.e., we fit $\{a_1, b_1, \rho, \sigma_2, \lambda_1\}$ to $\{r, l, B, c\}$ data. Since, as discussed below, we will be dealing with portfolios of pure discount bonds to represent coupon bonds, we consider $c = 0$, thereby requiring only $\{r, l\}$ data. We do not need the parameters $\{a_2, b_2, c_2\}$ because of the arbitrage arguments used to calculate $\beta_2 - \lambda_2\eta_2$ above [3]. This greatly improves the statistical merit of our fits, requiring two less degrees of freedom for the B equation. This approach also may be viewed as an empirical test of the consistency of assumptions used to derive the bond equation. Our VFSR methodology also explicitly include boundary conditions, a very important component of any model even if only implicitly described, since we use the same cost function, but as a function of its variables instead of its parameters, to calculate the evolution of bond prices with the path integral.

The ultimate test of any methodology is to compare theoretical predictions/descriptions with observed data [55]. Assume we already have fit our parameters for the entire epoch of interest. Actual bond prices with coupons may then be evaluated straightforwardly by considering a portfolio of n pure discount bonds with a series of maturity dates T_n equivalent to the dates of payment of coupons and the face value of the actual coupon bond to be modeled. This prescription requires that we integrate back such a portfolio of n pure discount bonds with maturity T_n, to various times $t_i < T$ (including only those bonds in the portfolio with maturity $T_n \geq t_i$). At each of these times, we use the observed values of $r(t_i)$ and $l(t_i)$ to calculate the bond prices $B_n(t_i)$. This portfolio of $\{B_n(t_i)\}$ is then compared to the observed coupon bond $B(t_i)$, i.e., for many such times $\{t_i\}$. For each zero-coupon bond in this portfolio, we start at its time of maturity T_n, enforcing the initial conditions $B_n(r, l; T_n) = 1$, and integrate back to a given time $t < T_n$. We then weight each zero coupon bond by the actual coupon or face value paid on the coupon

bond.

Similarly, for the important purpose of theoretically pricing an actual coupon bond at time t, we use the same methodology, e.g., using the present day's best estimate of $r(t)$ and $l(t)$ to calculate $B(r, l; t)$.

We are not in complete agreement with BS on their use of boundary conditions and numerical implementation [5]. Their use of "natural" boundary conditions, actually more general unrestricted or singular boundary conditions [56], is in part based on their own admittedly *ad hoc* choice of functional forms for r and l diffusions, in both their s.d.e. and p.d.e., and in the r drift in their s.d.e. We believe that the appropriate boundary conditions must be determined by finance considerations as follows:

$r \to 0$ and $l \to 0$ turn out to be natural boundary conditions of their model being inaccessible in any finite time, and therefore do not require or permit any additional specification. At $r = 0$, we should have $B(t + \Delta t) = B(t)$, i.e., reflecting boundary conditions. This is the regular boundary conditions of the model only if the r drift is greater than zero [49,56]. At $l = 0$, the BS model yields a unrestricted boundary conditions implying $B = 0$, and therefore does not require or permit any additional specification. Further examination shows that $l = 0$ is sometimes an entrance condition, sometimes an exit condition, depending on the value of r. Similar analyses show that $r \to \infty$ and $l \to \infty$ also are unrestricted boundary conditions.

For the r drift in the s.d.e. and in the p.d.e. to be non-negative, it is necessary, not sufficient, that $a_1 \geq 0$. We therefore propose that a_1 be constrained to achieve only these values in the fit. As mentioned in the introduction, BS also prefer this constraint on a_1. This may be considered a constraint just as is the functional form of the model. As it turned out, to be discussed below, the value of a_1 we obtained in our fit was essentially zero.

In setting up the path integral for the BS B p.d.e., we use all natural boundary conditions. This implies that there is no freedom to choose or to redundantly impose boundary conditions as we believe did BS. We choose the simple free-space Gaussian short-time propagator, since the evolution of B cannot leave the enclosed boundary conditions, since it is the proper solution deep in the interior of $r - l$ space, and since the drift is benign to the extent that this also is a good solution at the boundaries to order $\Delta\tau^{3/2}$ [15,48,49].

Since the boundary conditions at $r = 0$ and $l = 0$ are mathematically unrestricted if the diffusions vanish at these boundaries, as they do for the BS model, we therefore also propose that the *ad hoc* functional form of the diffusions be relaxed to admit additive noise components, e.g., κ_1 and κ_2. This is especially important for $r = 0$. This not only can be tested by so fitting data, but also mathematically permits the imposition of external boundary conditions more tightly constrained to the financial system under consideration, not being technically restricted by the actual functional forms chosen for the drifts and diffusions. Thus we also tested models for which

$$r\sigma_1 \to r\sigma_1 + \kappa_1 \ ,$$

$$l\sigma_2 \to l\sigma_2 + \kappa_2 \ , \tag{35}$$

in the drifts and diffusions of the bond p.d.e. Because of the induced singular behavior in the l-drift, this transformation still requires unrestricted boundary conditions at $l = 0$, which turn out to be entrance boundary conditions.

We believe it is extremely important to gain this freedom over the functional forms of the drifts and diffusions. For example, our calculations with this model clearly demonstrate that the rather mild nonlinearities of the BS model only permit inflationary evolution, since those were the periods were fit to data and since the functional forms likely cannot accommodate many swings and dips, on time scales of months or years, much longer that of the fluctuations, yet shorter than the period of long-term bonds. This appears to require a higher degree of nonlinearity and/or an increase in the number of independent interest-rate variables.

For future calculations, we propose to include some additive components in the r and l diffusions. We intend to invoke external reflecting boundary conditions for $r = 0$, using the method of images. Similar sets of boundary conditions were used in a previous project [16]. We checked the accuracy of the reflecting boundary conditions using MACSYMA, an algebraic manipulator.

BS procedures used a discretized form of their p.d.e. They compared their theoretical and observed values of B for several values of λ_1, and interpolated to find the best value of λ_1 which fit the data. Our methodology of calculating λ_1 from L_B, simultaneously with the other parameters in the p.d.e., avoids this additional step in their methodology.

4. NUMERICAL RESULTS

4.1. Fits to Interest Rates

Interest rates were developed from Treasury bill and bond yields during the period October 1974 through December 1979, the same period as one of the sets used by BS [7]. Short-term rates were determined from Treasury bills with a maturity of three months (BS used 30-day maturities), and long-term rates were determined from Treasury bonds with a maturity of twenty years (BS used at least 15-year maturities). For monthly runs, we used 63 points of data between 74-10-31 and 79-12-28. For daily runs, we used 1283 points of data between 74-10-31 and 79-12-31. We used yearly rates divided by 12 to fit the parameters.

For daily data, the actual number of days between successive trades was used; i.e., during this time period we had 1282 pieces of daily data and 62 pieces of end-of-month data. Although a rescaling in time only simply scales the deterministic parameters linearly, since that is how they appear in this model, this is not true for ρ. Then we did all subsequent runs using the scale of one day. We used yearly rates divided by 365 to fit the parameters.

The BS parameters also were run through the data, calculating the cost function they give. The single cost function bears the weight of determining all parameters. Typically, three or four significant-figure stability is required to get even one or two significant-figure stability in the parameters. (All runs were performed using double precision for all floating-point variables.) The "cost function" calculated is the sum over all Lagrangians at each short-time epoch (divided by the number of epochs, which doesn't affect its minimum, but helps to compare cost functions over different sets of data). I.e., a maximum probability fit is done by minimizing the cost functions (each the argument of the exponential representing the probability distribution of the variables) over all time epochs. The BS versus our fitted parameters are given in Table 1.

Table 1. BS parameters were fit to data using our Lagrangian representation for their coupled $r - l$ equations, for both end-of-month and daily data between 74-10-31 and 79-12-31. The second column, designated BS Monthly, gives their published 1982 results, using somewhat different data during this period. The third column gives our monthly fits on somewhat different data during this same time period. The fourth column gives daily fits scaled to daily time.

Parameter	BS Monthly	L Monthly	L Daily
a_1	0.0361	$3.02\ 10^{-5}$	$-6.33\ 10^{-9}$
b_1	0.0118	$3.89\ 10^{-4}$	0.0902
σ_1	0.0777	0.0700	0.0132
ρ	0.442	0.534	0.136
a_2	0.169	$9.73\ 10^{-3}$	$2.43\ 10^{-4}$
b_2	0.0089	0.0262	0.0320
c_2	−0.271	−0.707	−0.492
σ_2	0.0243	0.0278	$4.01\ 10^{-3}$

It should be noted that for all periods before October 1974, back through December 1958, using monthly data, BS found $a_1 < 0$, and for the period April 1964 through June 1969 they found $c_2 > 0$.

Fits were performed on a Hewlett Packard 9000-835SE, a "12-MIPS" computer. For example, the fit using the bond Lagrangian took approximately 100 CPU minutes for 1500 acceptance points, representing about 2000 generated points per 100 acceptance points at each re-annealing cycle, in this six-

dimensional parameter space. It was found that once the VFSR code repeated the lowest cost function within two cycles of 100 acceptance points, e.g., typically achieving 3 or 4 significant-figure accuracy in the global minimum of the cost function, by shunting to a local fitting procedure, the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm [57], only several hundred acceptance points were required to achieve 7 or 8 significant-figure accuracy in the cost function. This also provided yet another test of the VFSR methodology.

4.2. Fits to Bond Lagrangian

We examined differences in fitting L_B under varying constraints: In Table 2, we compare the L_B Daily fit with those obtained by admitting some degree of additive noise κ_1 and κ_2.

Table 2. We compare the L_B fit in the second column with those in the third column obtained by admitting some degree of additive noise κ_1 and κ_2. (Absolute values of $\kappa_{1,2}$ were used to constrain them to positive values in the local fits.) Note that BS obtained −0.216 for λ_1 by interpolating among several λ_1's to minimize the deviation of their theoretical bond prices to the observed ones.

Parameter	L_B	κ
a_1	$-2.02\ 10^{-8}$	$-1.01\ 10^{-7}$
b_1	$3.02\ 10^{-4}$	$1.01\ 10^{-3}$
σ_1	0.0173	0.0166
ρ	0.673	0.684
σ_2	$5.36\ 10^{-3}$	$3.84\ 10^{-3}$
λ_1	$-2.42\ 10^{-3}$	−0.0176
κ_1	-	$6.61\ 10^{-8}$
κ_2	-	$3.51\ 10^{-7}$

At this time, we are preparing fits to aggregate portfolios of bonds, to average over particulars of individual bonds, as performed by other investigators. We find that it is quite easy to fit λ_2 to sets of bond prices, after fitting interest-rate parameters using the BS s.d.e., as done by BS. However, we find it more difficult to fit λ_2 to these sets after using the bond Lagrangian to fit the interest-rate parameters, as we have presented here. These results will be reported in a future paper.

5. CONCLUSION

We have demonstrated how mathematical methodologies and numerical algorithms recently developed in the field of statistical mechanics can be brought to bear on term structure models. Specifically, methods of very fast simulated re-annealing can be used to statistically find best global fits of multivariate nonlinear stochastic term structure models, without requiring approximation of the basic models.

We also have argued that other numerical techniques, i.e., the path integral, can be brought to bear to calculate evolution of asset prices, using the term structure models for proxy variables. Another paper in progress will report on more extensive comparisons with observed bond prices.

This new formalism also permits a fresh look at some of these models and affords comparison with other nonlinear stochastic systems.

APPENDIX A: STATISTICAL MECHANICS DERIVATION OF PATH INTEGRAL

This Appendix outlines the derivation of the path integral representation of the nonlinear Langevin equations, via the Fokker-Planck representation. This serves to point out the importance of properly treating nonlinearities, and to emphasize the deceptive simplicity of the Langevin and Fokker-Planck representations of stochastic systems. There are a few derivations in the literature, but the following blend seems to be the most concise. All details may be found in the references given in this paper [23,58-60].

The Stratonovich (midpoint discretized) Langevin equations can be analyzed in terms of the Wiener process $\mathrm{d}W^i$, which can be rewritten in terms of Gaussian noise $\eta^i = \mathrm{d}W^i/\mathrm{d}t$ if care is taken in the limit [23].

$$\mathrm{d}M^G = f^G[t, M(t)]\mathrm{d}t + \hat{g}_i^G[t, M(t)]\mathrm{d}W^i \ ,$$

$$\dot{M}^G(t) = f^G[t, M(t)] + \hat{g}_i^G[t, M(t)]\eta^i(t) \ ,$$

$$\mathrm{d}W^i \to \eta^i \mathrm{d}t \ ,$$

$$M = \{M^G; G = 1, \cdots, \Lambda\} \ ,$$

$$\eta = \{\eta^i; i = 1, \cdots, N\} \ .$$

$$\dot{M}^G = \mathrm{d}M^G/\mathrm{d}t \ ,$$

$$< \eta^j(t) >_\eta = 0 \ ,$$

$$< \eta^j(t), \eta^{j'}(t') >_\eta = \delta^{jj'}\delta(t - t') \ , \tag{A.1}$$

η^i represents Gaussian white noise, and moments of an arbitrary function $F(\eta)$ over this stochastic space are defined by a path-type integral over η^i,

$$< F(\eta) >_\eta = \bar{N}^{-1} \int D\eta F(\eta) \exp(-\frac{1}{2}\int_{t_0}^{\infty} \mathrm{d}t \eta^i \eta^i \ ) \ ,$$

$$\bar{N} = \int D\eta \exp(-\frac{1}{2}\int_{t_0}^{\infty} \mathrm{d}t \eta^i \eta^i \ ) \ ,$$

$$D\eta = \lim_{\nu\to\infty} \prod_{\alpha=0}^{\nu+1} \prod_{j=1}^{N} (2\pi\theta)^{-1/2} \mathrm{d}W_\alpha^j \ ,$$

$$t_\alpha = t_0 + \alpha\theta \ ,$$

$$\frac{1}{2}\int \mathrm{d}t \eta^i \eta^i = \frac{1}{2\theta}\sum_\beta \sum_i (W_\beta^i - W_{\beta-1}^i)^2 \ ,$$

$$< \eta^i >_\eta = 0 \ ,$$

$$< \eta^i(t)\eta^j(t') >_\eta = \delta^{ij}\delta(t - t') \ . \tag{A.2}$$

Non-Markovian sources, $\hat{\eta}$, and their influence throughout this development, can be formally treated by expansions about the Markovian process by defining

$$< F(\hat{\eta}) >_\eta = \bar{N}_\xi^{-1} \int D\hat{\eta} F \ \exp[-\frac{1}{2}\int\int \mathrm{d}t\mathrm{d}t' \hat{\eta}(t)\Delta_\xi^{-1}(t - t')\hat{\eta}(t')] \ ,$$

$$\int \mathrm{d}t\ \Delta_\xi^{-1}(t-t')\Delta_\xi(t'-t'') = \delta(t-t'') \ , \tag{A.3}$$

with ξ defined as an interval centered about the argument of Δ_ξ. Letting $\xi \to 0$ is an unambiguous procedure to define the Stratonovich prescription used below.

In terms of a specific stochastic path η, a solution to Eq. (A.1), $M_\eta^G(t; M_0, t_0)$ with $M_\eta^G(t_0; M_0, t_0) \equiv M_0$, the initial conditions on the probability distribution of M_η is

$$P_\eta[M, t|M_0, t_0] = \delta[M - M_\eta(t; M_0, t_0)] \ . \tag{A.4}$$

Using the conservation of probability condition,

$$P_{\eta,t} + (\dot{M}^G P_\eta)_{,G} = 0 \ ,$$

$$[\cdots]_{,G} = \partial[\cdots]/\partial M^G \ ,$$

$$[\cdots]_{,t} = \partial[\cdots]/\partial t \ , \tag{A.5}$$

the evolution of P_η is written as

$$P_{\eta,t}[M, t|M_0, t_0] = \{[-f^G(t, M) - \hat{g}(t, M)\eta^i]P_\eta\}_{,G} \ . \tag{A.6}$$

To perform the stochastic average of Eq. (A.6), the "functional integration by parts lemma" [28] is used on an arbitrary function $Z(\eta)$ [59],

$$\int D\eta\ \frac{\hat{\delta} Z(\eta)}{\hat{\delta}\eta^i} = 0 \ . \tag{A.7}$$

Applied to $Z = Z' \exp(-\frac{1}{2}\int_{t_0}^{\infty} \mathrm{d}t \eta^i \eta^i)$, this yields

$$< \eta^i Z' >_\eta = < \delta Z'/\delta\eta^i >_\eta \ . \tag{A.8}$$

Applying this to $\hat{F}[M_\eta] = \int \mathrm{d}M\ P_\eta F(M)$,

$$\int \mathrm{d}M\ \frac{\hat{\delta} P_\eta}{\hat{\delta}\eta^i}\ F(M) = \frac{\partial \hat{F}[M_\eta]}{\partial M_\eta^G}\ \frac{\hat{\delta} M_\eta^G}{\hat{\delta}\eta^i}$$

$$= -\frac{1}{2}\int \mathrm{d}M\ F(M)(\hat{g}_j^G \delta^{ij} P^\eta)_{,G} \ , \tag{A.9}$$

where $\hat{\delta}$ designates functional differentiation. The last equation has used the Stratonovich prescription,

$$M_\eta^G(t) = M_0^G + \int \mathrm{d}t' \hat{H}(t-t')\hat{H}(t-t_0)(f^G + \hat{g}_i^G \eta^i) \ ,$$

$$\lim_{t' \to t-0} \frac{\hat{\delta} M_\eta^G(t)}{\hat{\delta}\eta^i(t')} = \frac{1}{2}\ \hat{g}_j^G[t, M_\eta(t)]\delta_{ij} \ ,$$

$$\hat{H}(z) = \begin{cases} 1, \ z \geq 0 \\ 0, \ z < 0 \ . \end{cases} \tag{A.10}$$

Taking the averages $< P_{\eta,t} >_\eta$ and $< \eta^i P_\eta >_\eta$, the Fokker-Planck is obtained from Eq. (A.9). If some boundary conditions are added as Lagrange multipliers, these enter as a "potential" V, creating a Schrödinger-type equation.

$$P_{,t} = \frac{1}{2}(g^{GG'}P)_{,GG'} - (g^G P)_{,G} + VP \ ,$$

$$P = < P_\eta >_\eta \ ,$$

$$g^G = f^G + \frac{1}{2}\,\hat{g}_i^{G'}\,\hat{g}_{i,G'}^G \ ,$$

$$g^{GG'} = \hat{g}_i^G\,\hat{g}_i^{G'} \ ,$$

$$[\cdots]_{,G} = \partial[\cdots]/\partial M^G \ . \tag{A.11}$$

Note that g^G replaces f^G in Eq. (A.1) if the Itô (prepoint discretized) calculus is used to define that equation.

To derive the path integral representation of Eq. (A.11), define operators $\hat{M}^G$, $\hat{p}_G$ and $\hat{H}$,

$$[\hat{M}^G, \hat{p}_{G'}] \equiv \hat{M}^G\,\hat{p}_{G'} - \hat{p}_{G'}\hat{M}^G = i\delta_{G'}^G \ ,$$

$$[\hat{M}^G, \hat{M}^{G'}] = 0 = [\hat{p}_G, \hat{p}_{G'}] \ ,$$

$$P_{,t} = -i\hat{H}P \ ,$$

$$\hat{H} = -\frac{i}{2}\,\hat{p}_G\hat{p}_{G'}g^{GG'} + \hat{p}_G g^G + iV \ , \tag{A.12}$$

and define the evolution operator $U(t,t')$ in terms of "bra" and "ket" probability states of M,

$$\hat{M}^G|M^G> = M^G|M^G> \ ,$$

$$\hat{p}_G|M^G> = -i\partial/\partial M^G|M^G> \ ,$$

$$< M'|M> = \delta(M'-M) \ ,$$

$$< M|p> = (2\pi)^{-1}\exp(ip\cdot M) \ ,$$

$$P[M,t|M_0,t_0] =< M|U(t,t_0)|M_0> \ ,$$

$$\hat{H}(t')U(t',t) = iU(t',t)_{,t'} \ ,$$

$$U(t,t) = 1 \ ,$$

$$U(t_\rho,t_{\rho-1}) \approx 1 - i\theta\hat{H}(t_{\rho-1}) \ , \tag{A.13}$$

where ρ indexes units of θ measuring the time evolution. This is formally integrated to give the path integral in the phase space (p,M),

$$P[M_t|M_0] = \int_{M(t_0)=M_0}^{M(t)=M_t} DM\ Dp\ \exp[\int_{t_0}^{t} \mathrm{d}t'(ip_G M^G - \frac{1}{2}\,p_G p_{G'}g^{GG'} - ip_G g^G + V)\,] \ ,$$

$$DM = \lim_{u\to\infty}\prod_G\prod_{\rho=1}^{u}\mathrm{d}M_\rho^G \ ,$$

$$Dp = \lim_{u\to\infty}\prod_G\prod_{\rho=1}^{u+1}(2\pi)^{-1}\mathrm{d}p_{G\rho} \ ,$$

$$t_\rho = t_0 + \rho\theta \ . \tag{A.14}$$

The integral over each $\mathrm{d}p_{G\rho}$ is a Gaussian and simply calculated. This gives the path integral in coordinate space M, in terms of the prepoint discretized Lagrangian,

$$P[M_t|M_0] = \int DM \prod_{\rho=0}^{u} (2\pi\theta)^{-\Lambda/2} g(M_\rho, t_\rho)^{1/2}$$

$$\times \exp\{ -\frac{1}{2}\theta g_{GG'}(M_\rho, t_\rho)[\Delta_\rho^G/\theta - g^G(M_\rho, t_\rho)]$$

$$\times[\Delta_\rho^{G'}/\theta - g^{G'}(M_\rho, t_\rho)] + \theta V(M_\rho, t_\rho)\} ,$$

$$L_I(\dot{M}^G, M^G, t) = \frac{1}{2}(\dot{M}^G - g^G) g_{GG'}(\dot{M}^{G'} - g^{G'}) - V ,$$

$$g = \det(g_{GG'}) ,$$

$$g_{GG'} = (g^{GG'})^{-1} ,$$

$$\Delta_\rho^G = M_{\rho+1}^G - M_\rho^G . \tag{A.15}$$

This can be transformed to the Stratonovich representation, in terms of the Feynman Lagrangian L possessing a covariant variational principle,

$$P[M_t|M_0] = \int DM \prod_{\rho=0}^{u} (2\pi\theta)^{-\Lambda/2} g(M_\rho + \Delta_\rho, t_\rho + \theta/2)^{1/2}$$

$$\times \exp\{ -\min \int_{t_\rho}^{t_\rho+\theta} \mathrm{d}t' L[M(t'), \dot{M}(t'), t']\} , \tag{A.16}$$

where "min" specifies that Eq. (A.11) is obtained by constraining L to be expanded about that $M(t)$ which makes the action $S = \int \mathrm{d}t' L$ stationary for $M(t_\rho) = M_\rho$ and $M(t_\rho + \theta) = M_{\rho+1}$.

One way of proceeding is to expand Eq. (A.15) and compare to Eq. (A.16), but it is somewhat easier to expand Eq. (A.16) and compare to Eq. (A.15) [60]. It can be shown that expansions to order θ suffice, and that $\Delta^2 = O(\theta)$.

Write L in the general form

$$L = \frac{1}{2} g_{GG'} \dot{M}^G \dot{M}^{G'} - h_G \dot{M}^G + b$$

$$= L^0 + \Delta L ,$$

$$L^0 = \frac{1}{2} g_{GG'}[M(t), t] \dot{M}^G \dot{M}^{G'} ,$$

$$g_{GG'}[M(t), t] = g_{GG'}[M(t), t'] + g_{GG',t'}[M(t), t'](t - t') + O[(t - t')^2] , \tag{A.17}$$

where h_G and b must be determined by comparing expansions of Eq. (A.15) and Eq. (A.16). Only the L^0 term is dependent on the actual $M(t)$ trajectory, and so

$$\int_{t_\rho}^{t_\rho+\theta} \mathrm{d}t\ \Delta L = (\frac{1}{4} g_{GG',t} \Delta^G \Delta^{G'} - h_G \Delta^G - \frac{1}{2} h_{G,G'} \Delta^G \Delta^{G'} + \theta b)|_{(M,t)} , \tag{A.18}$$

where "$|_{(M,t)}$" implies evaluation at (M, t).

The determinant g is expanded as

$$g(M + \Delta, t + \theta/2)^{1/2} \approx g^{1/2}(M, t) \exp[\frac{\theta}{4g} g_{,t} + \frac{1}{2g} \Delta^G g_{,G}$$

$$+\frac{1}{4g}\,\Delta^G\Delta^{G'}(g_{,GG'}+g^{-1}g_{,G}g_{,G'})]|_{(M,t)}\ . \tag{A.19}$$

The remaining integral over L^0 must be performed. This is accomplished using the variational principle applied to $\int L^0$ [58],

$$g_{GH}\ddot{M}^H=-\frac{1}{2}(g_{GH,K}+g_{GK,H}-g_{KH,G})\dot{M}^K\dot{M}^H\ ,$$

$$\ddot{M}^F=-\Gamma^F_{JK}\dot{M}^J\dot{M}^K\ ,$$

$$\Gamma^F_{JK}==g^{LF}[JK,L]=g^{LF}(g_{JL,K}+g_{KL,J}-g_{JK,L})\ ,$$

$$(\frac{1}{2}\,g_{GH}\dot{M}^G\dot{M}^H)_{,t}=0\ ,$$

$$\int_t^{t+\theta}L^0\mathrm{d}t\approx\frac{\theta}{2}\,g_{GH}\dot{M}^G\dot{M}^H|_{(M,t+\theta)}\ . \tag{A.20}$$

Differentiating the second equation in Eq. (A.20) to obtain $\dddot{M}$, and expanding $\dot{M}(t+\theta)$ to third order in θ,

$$\dot{M}(t+\theta)=[\frac{1}{\theta}\,\Delta^G-\frac{1}{2\theta}\,\Gamma^G_{KL}\Delta^K\Delta^L+\frac{1}{6\theta}\,(\Gamma^G_{KL,N}+\Gamma^G_{AN}\Gamma^A_{KL})\Delta^G\Delta^L\Delta^N]|_{(M,t)}\ . \tag{A.21}$$

Now Eq. (A.16) can be expanded as

$$P[M_t|M_0]\mathrm{d}M(t)=\int \underline{D}M\prod_{\rho=0}^{u}\exp[-\frac{1}{2\theta}\,g_{GG'}(M,t)\Delta^G\Delta^{G'}+B]\ ,$$

$$\underline{D}M=\prod_{\rho=1}^{u+1}g_\rho^{1/2}\prod_G(2\pi\theta)^{-1/2}\mathrm{d}M^G_\rho\ . \tag{A.22}$$

Expanding $\exp B$ to $O(\theta)$ requires keeping terms of order Δ, Δ^2, Δ^3/θ, Δ^4/θ, and Δ^6/θ^2. Under the path integral, evaluated at (M,t), and using "$\doteq$" to designate the order of terms obtained from $\int \mathrm{d}\Delta\ \Delta^n\exp(-\frac{1}{2\theta}\,\Delta^2)$,

$$\Delta^G\Delta^H\doteq\theta g^{GH}\ ,$$

$$\Delta^G\Delta^H\Delta^K\doteq\theta(\Delta^G g^{HK}+\Delta^H g^{GH}+\Delta^K g^{GH})\ ,$$

$$\Delta^G\Delta^H\Delta^A\Delta^B\doteq\theta^2(g^{GH}g^{AB}+g^{GA}g^{HB}+g^{GB}g^{HA})\ ,$$

$$\Delta^A\Delta^B\Delta^C\Delta^D\Delta^E\Delta^F\doteq\theta^3(g^{AB}g^{CD}g^{EF}+14\text{ permutations})\ . \tag{A.23}$$

This expansion of $\exp B$ is to be compared to Eq. (A.15), expanded as

$$P[M_t|M_0]\mathrm{d}M(t)\approx\int \underline{D}M\prod_{\rho=0}^{u}\exp(-\frac{1}{2\theta}\,g_{GG'}\Delta^G\Delta^{G'})$$

$$\times[1+g_{GG'}g^G\Delta^{G'}+\theta V+O(\theta^{3/2})]\ , \tag{A.24}$$

yielding identification of h_G and b in Eq. (A.16),

$$h^G=g^{GG'}h_{G'}=g^G-\frac{1}{2}\,g^{-1/2}(g^{1/2}g^{GG'})_{,G'}\ ,$$

$$b = \frac{1}{2}\, h^G h_G + \frac{1}{2}\, h^G{}_{;G} + R/6 - V \ ,$$

$$h^G{}_{;G} = h^G_{,G} + \Gamma^F_{GF} h^G = g^{-1/2}(g^{1/2} h^G)_{,G} \ ,$$

$$R = g^{JL} R_{JL} = g^{JL} g^{FK} R_{FJKL} \ . \tag{A.25}$$

The result is

$$P[M_t|M_{t_0}]\mathrm{d}M(t) = \int \cdots \int \underline{D}M \exp(-\min \int_{t_0}^{t} \mathrm{d}t' L)\delta[M(t_0) = M_0]\delta[M(t) = M_t] \ ,$$

$$\underline{D}M = \lim_{u\to\infty} \prod_{\rho=1}^{u+1} g^{1/2} \prod_G (2\pi\theta)^{-1/2} \mathrm{d}M^G_\rho \ ,$$

$$L(\dot{M}^G, M^G, t) = \frac{1}{2}\,(\dot{M}^G - h^G) g_{GG'} (\dot{M}^{G'} - h^{G'}) + \frac{1}{2}\, h^G{}_{;G} + R/6 - V \ ,$$

$$h^G = g^G - \frac{1}{2}\, g^{-1/2}(g^{1/2} g^{GG'})_{,G'} \ ,$$

$$g_{GG'} = (g^{GG'})^{-1} \ ,$$

$$g = \det(g_{GG'}) \ ,$$

$$h^G{}_{;G} = h^G_{,G} + \Gamma^F_{GF} h^G = g^{-1/2}(g^{1/2} h^G)_{,G} \ ,$$

$$\Gamma^F_{JK} \equiv g^{LF}[JK, L] = g^{LF}(g_{JL,K} + g_{KL,J} - g_{JK,L}) \ ,$$

$$R = g^{JL} R_{JL} = g^{JL} g^{JK} R_{FJKL} \ ,$$

$$R_{FJKL} = \frac{1}{2}\,(g_{FK,JL} - g_{JK,FL} - g_{FL,JK} + g_{JL,FK}) + g_{MN}(\Gamma^M_{FK}\Gamma^N_{JL} - \Gamma^M_{FL}\Gamma^N_{JK}) \ . \tag{A.26}$$

In summary, because of the presence of multiplicative noise, the Langevin system differs in its Itô (prepoint) and Stratonovich (midpoint) discretizations. The midpoint-discretized covariant description, in terms of the Feynman Lagrangian, is defined such that (arbitrary) fluctuations occur about solutions to the Euler-Lagrange variational equations. In contrast, the usual Itô and corresponding Stratonovich discretizations are defined such that the path integral reduces to the Fokker-Planck equation in the weak-noise limit. The term $R/6$ in the Feynman Lagrangian includes a contribution of $R/12$ from the WKB approximation to the same order of $(\Delta t)^{3/2}$ [23].

REFERENCES

1. J.C. Cox, J.E. Ingersoll, Jr., and S.A. Ross, A re-examination of traditional hypotheses about the term structure of interest rates, *J. Finance* **36**, 769-799 (1981).
2. J.C. Cox, J.E. Ingersoll, Jr., and S.A. Ross, An intertemporal general equilibrium model of asset prices, *Econometrica* **53**, 363-384 (1985).
3. J.C. Cox, J.E. Ingersoll, Jr., and S.A. Ross, A theory of the term structure of interest rates, *Econometrica* **53**, 385-407 (1985).
4. O. Vasicek, An equilibrium characterization of the term structure, *J. Finan. Econ.* **5**, 177-188 (1977).
5. M.J. Brennan and E.S. Schwartz, A continuous time approach to the pricing of bonds, *J. Banking Finan.* **3**, 133-155 (1979).
6. M.J. Brennan and E.S. Schwartz, Bond pricing and market efficiency, *Finan. Anal. J.* **Sep-Oct**, 49-56 (1982).
7. M.J. Brennan and E.S. Schwartz, An equilibrium model of bond pricing and a test of market efficiency, *J. Finan. Quant. Anal.* **17**, 301-329 (1982).
8. S.M. Schaefer and E.S. Schwartz, A two-factor model on the term structure: An approximate analytical solution, *J. Finan. Quant. Anal.* **19**, 413-424 (1984).
9. B. Dietrich-Campbell and E. Schwartz, Valuing debt options, *J. Finan. Econ.* **16**, 321-343 (1986).
10. B. B. Mandelbrot, Comments on: 'A subordinated stochastic process model with finite variance for speculative prices,' by Peter K. Clark, *Econometrica* **41**, 157-159 (1973).
11. M. C. Jensen, Some anomalous evidence regarding market efficiency, an editorial introduction, *J. Finan. Econ.* **6**, 95-101 (1978).
12. S. J. Taylor, Tests of the random walk hypothesis against a price-trend hypothesis, *J. Finan. Quant. Anal.* **17**, 37-61 (1982).
13. A.C. Aitken, On least squares and linear combinations of observations, *Proc. Roy. Soc. Edinburgh* **55**, 42-48 (1935).
14. L. Ingber, Very fast simulated re-annealing, *Mathl. Comput. Modelling* **12** (8), 967-973 (1989).
15. M.F. Wehner and W.G. Wolfer, Numerical evaluation of path integral solutions to Fokker-Planck equations. III. Time and functionally dependent coefficients, *Phys. Rev. A* **35**, 1795-1801 (1987).
16. L. Ingber, H. Fujio, and M.F. Wehner, Mathematical comparison of combat computer models to exercise data, *Mathl. Comput. Modelling* **15** (1), 65-90 (1991).
17. T.S.Y. Ho and S.-B. Lee, Term structure movements and pricing interest rate contingent claims, *J. Finance* **41**, 1011-1029 (1986).
18. D. Heath, R. Jarrow, and A. Morton, Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation, (1987).
19. D. Heath, R. Jarrow, and A. Morton, Contingent claim valuation with a random evolution of interest rates, (1989).
20. L. Ingber, Statistical mechanics of nonlinear nonequilibrium financial markets, *Math. Modelling* **5** (6), 343-361 (1984).
21. L. Ingber, Mesoscales in neocortex and in command, control and communications (C^3) systems, in *Systems with Learning and Memory Abilities: Proceedings, University of Paris 15-19 June 1987*, (Edited by J. Delacour and J.C.S. Levy), pp. 387-409, Elsevier, Amsterdam, (1988).
22. L. Ingber, Mathematical comparison of JANUS(T) simulation to National Training Center, in *The Science of Command and Control: Part II, Coping With Complexity*, (Edited by S.E. Johnson and A.H. Levis), pp. 165-176, AFCEA International, Washington, DC, (1989).
23. F. Langouche, D. Roekaerts, and E. Tirapegui, *Functional Integration and Semiclassical Expansions*, Reidel, Dordrecht, The Netherlands, (1982).

24. H. Dekker, Quantization in curved spaces, in *Functional Integration: Theory and Applications*, (Edited by J.P. Antoine and E. Tirapegui), pp. 207-224, Plenum, New York, (1980).
25. H. Grabert and M.S. Green, Fluctuations and nonlinear irreversible processes, *Phys. Rev. A* **19**, 1747-1756 (1979).
26. R. Graham, Covariant formulation of non-equilibrium statistical thermodynamics, *Z. Physik* **B26**, 397-405 (1977).
27. R. Graham, Lagrangian for diffusion in curved phase space, *Phys. Rev. Lett.* **38**, 51-53 (1977).
28. L.S. Schulman, *Techniques and Applications of Path Integration*, J. Wiley & Sons, New York, (1981).
29. L. Ingber, Nonlinear nonequilibrium statistical mechanics approach to C^3 systems, in *9th MIT/ONR Workshop on C^3 Systems: Naval Postgraduate School, Monterey, CA, 2-5 June 1986*, pp. 237-244, MIT, Cambridge, MA, (1986).
30. R.F. Fox, Uniform convergence to an effective Fokker-Planck equation for weakly colored noise, *Phys. Rev. A* **34**, 4525-4527 (1986).
31. C. van der Broeck, On the relation between white shot noise, Gaussian white noise, and the dichotomic Markov process, *J. Stat. Phys.* **31**, 467-483 (1983).
32. P. Colet, H.S. Wio, and M. San Miguel, Colored noise: A perspective from a path-integral formalism, *Phys. Rev. A* **39**, 6094-6097 (1989).
33. C.W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*, Springer-Verlag, Berlin, Germany, (1983).
34. L. Ingber, Statistical mechanics of neocortical interactions. Dynamics of synaptic modification, *Phys. Rev. A* **28**, 395-416 (1983).
35. L. Ingber, Statistical mechanics of neocortical interactions. Derivation of short-term-memory capacity, *Phys. Rev. A* **29**, 3346-3358 (1984).
36. L. Ingber, Statistical mechanics of neocortical interactions. EEG dispersion relations, *IEEE Trans. Biomed. Eng.* **32**, 91-94 (1985).
37. L. Ingber, Statistical mechanics of neocortical interactions: Stability and duration of the 7+−2 rule of short-term-memory capacity, *Phys. Rev. A* **31**, 1183-1186 (1985).
38. L. Ingber and P.L. Nunez, Multiple scales of statistical physics of neocortex: Application to electroencephalography, *Mathl. Comput. Modelling* **13** (7), 83-95 (1990).
39. L. Ingber, Path-integral Riemannian contributions to nuclear Schrödinger equation, *Phys. Rev. D* **29**, 1171-1174 (1984).
40. L. Ingber, Riemannian contributions to short-ranged velocity-dependent nucleon-nucleon interactions, *Phys. Rev. D* **33**, 3781-3784 (1986).
41. L. Ingber, Mathematical comparison of computer models to exercise data, in *1989 JDL C^2 Symposium: National Defense University, Washington, DC, 27-29 June 1989*, pp. 169-192, SAIC, McLean, VA, (1989).
42. L. Ingber and D.D. Sworder, Statistical mechanics of combat with human factors, *Mathl. Comput. Modelling* **15** (11), 99-127 (1991).
43. K. Kishida, Physical Langevin model and the time-series model in systems far from equilibrium, *Phys. Rev. A* **25**, 496-507 (1982).
44. K. Kishida, Equivalent random force and time-series model in systems far from equilibrium, *J. Math. Phys.* **25**, 1308-1313 (1984).
45. L. Ingber, Statistical mechanical aids to calculating term structure models, *Phys. Rev. A* **42** (12), 7057-7064 (1990).
46. H. Szu and R. Hartley, Fast simulated annealing, *Phys. Lett. A* **122** (3-4), 157-162 (1987).
47. S. Kirkpatrick, C.D. Gelatt, Jr., and M.P. Vecchi, Optimization by simulated annealing, *Science* **220** (4598), 671-680 (1983).

48. M.F. Wehner and W.G. Wolfer, Numerical evaluation of path-integral solutions to Fokker-Planck equations. I., *Phys. Rev. A* **27**, 2663-2670 (1983).

49. M.F. Wehner and W.G. Wolfer, Numerical evaluation of path-integral solutions to Fokker-Planck equations. II. Restricted stochastic processes, *Phys. Rev. A* **28**, 3003-3011 (1983).

50. C. Grebogi, E. Ott, and J.A. Yorke, Chaos, strange attractors, and fractal basin boundaries in nonlinear dynamics, *Science* **238**, 632-637 (1987).

51. R. Pool, Is it chaos, or is it just noise?, *Science* **243**, 25-28 (1989).

52. J. Theiler, Correlation dimension of filtered noise, Report 6/29/1988, UC San Diego, La Jolla, CA, (1988).

53. P. Grassberger, Do climatic attractors exist?, *Nature* **323**, 609-612 (1986).

54. N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller, Equation of state calculations by fast computing machines, *J. Chem. Phys.* **21** (6), 1087-1092 (1953).

55. S.G. Brush, Prediction and theory evaluation: The case of light bending, *Science* **246**, 1124-1129 (1989).

56. N.S. Goel and N. Richter-Dyn, *Stochastic Models in Biology*, Academic Press, New York, NY, (1974).

57. D.F. Shanno and K.H. Phua, Minimization of unconstrained multivariate functions, *ACM Trans. Mathl. Software* **2**, 87-94 (1976).

58. K.S. Cheng, Quantization of a general dynamical system by Feynman's path integration formulation, *J. Math. Phys.* **13**, 1723-1726 (1972).

59. F. Langouche, D. Roekaerts, and E. Tirapegui, Discretization problems of functional integrals in phase space, *Phys. Rev. D* **20**, 419-432 (1979).

60. F. Langouche, D. Roekaerts, and E. Tirapegui, Short derivation of Feynman Lagrangian for general diffusion process, *J. Phys. A* **113**, 449-452 (1980).