Ito Processes with Finitely Many States of Memory

McCauley, Joseph L.

University of Houston

16 November 2007

Ito Processes with Finitely Many States of Memory

Joseph L. McCauley
Physics Department
University of Houston
Houston, Tx. 77204-5005
jmccauley@uh.edu

Abstract

We show that Ito processes imply the Fokker-Planck (K2) and Kolmogorov backward time (K1) partial differential eqns. (pde) for transition densities, which in turn imply the Chapman-Kolmogorov equation without approximations. This result is not restricted to Markov processes. We define 'finite memory' and show that Ito processes admit finitely many states of memory. We then provide an example of a Gaussian transition density depending on two past states that satisfies both K1, K2, and the Chapman-Kolmogorov eqn. Finally, we show that transition densities of Black-Scholes type pdes with finite memory are martingales and also satisfy the Chapman-Kolmogorov equation. This leads to the shortest possible proof that the transition density of the Black-Scholes pde provides the so-called 'martingale measure' of option pricing.

Key Words: Ito process, martingale, stochastic differential eqn., Langevin eqn., memory, nonMarkov process, Fokker-Planck eqn., Kolmogorov's backward time eqn., Chapman-Kolmogorov eqn., Black-Scholes eqn.

1. Stochastic processes with finite memory

Much unnecessary confusion has been introduced into the literature under the terms 'nonlinear Markov process' and 'nonlinear Fokker-Planck equation'. We've pointed out that one particular model labeled as a 'nonlinear Markov process', the Shimizu-Yamata model, is an Ito process with finite memory [1]. Here, we prove that assertion in all detail.

An Ito process x is a martingale M(t) plus a drift A(t), x(t)=A(t)+M(t), and is generated locally by an Ito sde (Langevin eqn.)

$$dx = R(x,t)dt + \sqrt{D(x,t)}dB \qquad (1)$$

where B(t) is the Wiener process, a Gaussian process with uncorrelated, stationary increments B(t,T)=B(t+T)-B(t)=B(T), B(t,-T)-B(t)-B(t-T)=B(-T), $\langle B(t,T)B(t,-T)\rangle=0$ so that $\langle B(t+T)B(t)\rangle=\langle B^2(t)\rangle=t$ if T>0. The conditions $\langle B(t+T)B(t)\rangle=\langle B^2(t)\rangle$ and $\langle B(t,T)B(t,-T)\rangle=0$ are actually the same condition and simply reflect that the Wiener process is a martingale [2].

The Langevin eqns. studied in the 1970s in the context of the fluctuation-dissipation theorem for systems near statistical equilibrium are generated locally by correlated noise [3]. Those processes are Gaussian with infinite memory, memory of a continuum of past states traced back to x(0)=0. Fractional Brownian motion (fBm) has infinite memory extending back to $x=-\infty$ but is generated locally by uncorrelated noise, is a stochastic integral over a Wiener process [4,5]. Neither of those processes is an Ito process.

We offer here the mathematical formalism of stochastic processes with finitely many states of memory. This formulation, based as it is on martingales, is applicable to financial markets [6]. Our line of thought is motivated by A. Friedman's treatment of Ito processes and related partial differential equations [7], because one essential result is proven therein while another is listed as an exercise, but the author asserted that he had assumed a Markov process. *We will see that nowhere is the assumption of a Markov process necessary in the derivations.*

Consider an arbitrary stochastic process x(t). The probability density for n points $f_n(x_n,t_n; \ldots; x_1,t_1)$ is defined by

$$f_n(x_n,t_n;\ldots;x_1,t_1) = p_n(x_n,t_n|x_{n-1},t_{n-1};\ldots;x_1,t_1)f_{n-1}(x_{n-1},t_{n-1};\ldots;x_1,t_1) \quad (2)$$

where $p_2(x_n,t_n|x_{n-1},t_{n-1};\ldots;x_o,t_o)$ is the 2-point conditional probability, or transition density, depending on a history of n-1 points [8,9]. That is, p_n is the probability density to observe x_n at time t_n, given that the n-1 points $(x_{n-1},\ldots x_1)$ were observed to have occurred at times $(t_{n-1},\ldots,t_1)$ in the past. The states $(x_{n-1},\ldots x_1)$ are therefore the known part of the history of one trajectory x(t). By 'finite history' we mean that the n-2 points actually were observed at the respective n-2 different times prior to the occurrence of the last observed point x_{n-1}.

The memory in fBm [5] is qualitatively different. All trajectories are 'filtered' through one single point x(0)=0 (required by scaling $x(t)=t^H x(1)$) and there is strong correlation of any point x at present time t with the trajectory's entire past. Since fBm is Gaussian with $B_{kl} = \langle x_k x_l \rangle = \langle x^2(1)\rangle[|t_k|^{2H} + |t_l|^{2H} + |t_k - t_l|^{2H}]$, we have

$$p_n(x_n,t_n|x_{n-1},t_{n-1};\ldots;x_1,t_1) \propto e^{-(B_{nn}^{-1}x_n^2 + \sum_{j=1}^{n-1} B_{nj}^{-1} x_n x_j)/2}, \qquad (3)$$

so that p_n depends on all n of the states x_k considered all the way to n=∞. This is what we mean by 'infinite memory'. This long time autocorrelation follows from stationarity of the increments [2,4,5].

A drift term $A(t)$ can be locally subtracted from a time series generated by an Ito process, but fBm is not a martingale plus a drift and cannot be detrended: the 'trend' $d\langle x(t)\rangle/dt \neq 0$ in fBm is due to the long time autocorrelations that cannot be removed [5]. In a Markov process only the last observed state is remembered, the transition densities obey $p_n = p_2$, $n \geq 2$. Because the 2-point transition density determines the drift and diffusion coefficients $R(x,t)$, $D(x,t)$, those coefficients for a Markov process cannot depend on any state other than the present state (x,t). If other states appear in R and D, then the process has finite memory [2,10]. The latter fact is completely ignored where unjustified claims have been made of 'nonlinear Markov processes' [11].

Following Kolmogorov's definition of a stochastic process [7], one needs the entire hierarchy of transition densities p_n, $n=2,3, \ldots.$, in order to completely specify or completely identify the stochastic process. The hierarchy will truncate at finite n iff. the memory string is finite. For Gaussian processes the conditional density of any and all finite orders n can be constructed once the pair correlations $\langle x(t_k)x(t_l)\rangle$ for times $t_k \neq t_l$ are known [10]. For Markov processes, $p_n = p_2$ for all $n=3,4, \ldots.$, because a Markov process erases all history except that of the last observed point (x_{n-1}, t_{n-1}) [8,9,10]. Our point is that this is far more restrictive than the history dependence that is allowed for a general Ito process.

The fundamental problem in finance, as in empirically based modeling in any science, is how to identify the underlying stochastic dynamics, given some collection of time series representing repeated runs of the same process. In finance, we've shown that we can treat daily trading as a rerun of the same uncontrolled 'experiment' [6]. Hänggi and Thomas pointed out long ago [3] that the equation of motion for a 1-point density $f_1(x,t)$ tells us nothing about the underlying dynamics. E.g., both scaling Gaussian Markov processes and fractional Brownian motion (fBm) have exactly the same 1-point Gaussian density [1,2,3] although their pair correlations are completely different. The transition density for a Markov process obeys a diffusion eqn., the transition density for fBm does not. Therefore the minimal knowledge of dynamics requires either p_2 or at least the discovery of pair correlations $<x(t+T)x(t)>$, nothing less will suffice. Hence, all the literature about 'nonlinear Fokker-Planck pdes' for 1-point densities is irrelevant, it's the time evolution of the 2-point density that reflects the stochastic process, and in an Ito process that time evolution is given by a linear pde, the usual Fokker-Planck pde. Next, we point our that transition densities *do not scale* even if the process is selfsimilar, so that scaling cannot be used to identify a stochastic process.

Hurst exponent scaling, when it exists, is defined by trajectories obeying $x(t)=t^H x(1)$, $x(0)=0$ [2]. Because scaling is confined to 1-point densities, $f_1(x,t)=t^{-H}F(x/t^H)$, scaling exponents H cannot be used to identify an underlying stochastic process. Another way to say it is that, even if we could and would restrict our considerations to trajectories that scale, neither pair correlations, $<x(t+T)x(t)>$ nor 2-point densities $p_2(y,t+T:x,t)$ will scale. That is, the naïve replacement $<x(t)x(s)>=(st)^H<x^2(1)>$, where $<x(s)x(t)>=\int dydxyx\,p_2(y,s|x,t)f_1(x,t)$, is wrong because p_2 does not scale with H: pair correlations destroy the scaling even for drift-free Markov processes. One sees the lack of scaling of pair correlations explicitly for fBm, where it is very easy to construct both the pair correlations $<x(s)x(t)>$ and the Gaussian conditional density p_2 explicitly [5]. One sees lack of scaling of p_2 for martingales (including drift-free Markov processes) because $<x(t+T)x(t)>=<x^2(t)>$ if $T>0$ [2]. So in data analysis (or in modeling) we must extract (or specify) either the pair correlations or else p_2 in order to exhibit any knowledge at all of the nature of the dynamics of the underlying process; scaling, even if it occurs, is irrelevant. In FX markets, e.g., we have verified that the pair correlations are those of a martingale after 10 min. of trading, $\langle x(t+T)x(t)\rangle = \langle x^2(t)\rangle$ if $T \geq 10$min. [6]. We come now to the main point of this article: the existence of 'finite memory' in a general Ito process. Martingales are Ito processes, are the basis for Ito processes.

2. The Chapman-Kolmogorov Equation

By 'finite memory' we mean that the hierarchy of transition densities truncates, that $p_k=p_n$ for $k \geq n$ with n finite so that we obtain from the rule

$$p_{k-1}(x_k,t_k|x_{k-2},t_{k-2};...;x_1,t_1) = \int dx_{k-1} p_k(x_k,t_k|x_{k-1},t_{k-1};...;x_1,t_1) p_{k-1}(x_{k-1},t_{k-1}|x_{k-2},t_{k-2};...;x_1,t_1) \quad (4)$$

the Chapman-Kolmogorov equation

$$p_n(x_n,t_n|x_{n-1},t_{n-1};...;x_1,t_1) = \int dy p_n(x_n,t_n|y,s;x_{n-2},t_{n-2};...;x_1,t_1) p_n(y,s|x_{n-1},t_{n-1};...;x_1,t_1) \quad (5)$$

for the 2-point transition density with a nontrivial history of n-2 points. The process is Markovian iff. n=2. As both Doob and Feller pointed out [12,13], the Chapman-Kolmogorov equation is a necessary but *insufficient* condition for a Markov process. Feller provided a non-Ito process as example where a Chapman-Kolmogorov eqn. holds. Below, we will provide an example of an Ito process with memory that does the job.

If there is one nontrivial state of memory, if e.g. the initial condition is $f_1(x,t_o)=u(x)$ and $x_o=\int xu(x)dx$, then due to memory in the initial data $f_1(x,t)$ [14] we obtain

$$p_2(x_3,t_3|x_2,t_2) = \frac{\int p_3(x_3,t_3|x_2,t_2;x_1,t_1)p_2(x_2,t_2|x_1,t_1)f_1(x_1,t_1)dx_1}{\int p_2(x_2,t_2|x_1,t_1)f_1(x_1,t_1)dx_1}.$$

(6)

Memory appears in (18) if, e.g., at time t_o $f_1(x)=\delta(x-x_o)$ with $x_o\neq0$ [14]. In this case, by the 2-point transition density we must understand $p_2(x,t|y,s) = p_3(x,t|y,s;x_o,t_1)$, it's p_3 that satisfies the Chapman-Kolmogorov eqn. That is, in the simplest case of memory p_3 is required to describe the stochastic process. The main idea is that we are dealing quite generally with Ito sdes and corresponding pdes for transition densities with memory of a finite nr. n-2 of states, so that the 2-point transition density depends on n-2 earlier states.

We proceed with nonstandard derivations of Kolmogorov's two partial differential equations (pdes) and the Chapman-Kolmogorov equation. In order to show that the resulting formalism is not vacuous, we then provide an example of a stochastic process with nontrivial memory satisfying all three of those equations. Some of the details to be presented below can be found scattered disconnected throughout the literature. Their marriage into a unity and the interpretation in terms of finite memory are new, as is the proof that the Shimizu-Yamata model is a nonMarkovian Ito process with one nontrivial state of memory. Our new viewpoint informs us, in particular, that martingale dynamics in particular, with or without memory, is Fokker-Planck dynamics for the case where both x and t vary continuously.

3. Ito processes with finite memory

Consider a diffusive process described by an Ito stochastic differential equation (sde)

$$dx = R(x,t)dt + \sqrt{D(x,t)}dB(t) \qquad (1)$$

with or without finite memory in the drift and diffusion coefficients.

Consider first the Markov case. From

$$x(t+T) = x(t) + \int_t^{t+T} R(x(s),s)ds + \int_t^{t+T} \sqrt{D(x(s),s)}dB(s) \qquad (7)$$

and ignoring the drift, which is $O(T^2)$ for small T, we obtain the mean square fluctuation for small time lags T as

$$\left\langle (x(t+T)-x(t))^2 \right\rangle \approx \int_t^{t+T} \left\langle D(x(s),s) \right\rangle \approx TD(x,t) \qquad (8)$$

on the one hand, since $p_2(y,t+T:x,t) \approx \delta(y-x)$ as T vanishes, but also

$$\left\langle (x(t+T)-x(t))^2 \right\rangle \approx \int dy (y-x)^2 p_2(y,t+|x,t) \qquad (9)$$

on the other. This yields the standard definition

$$D(x,t) \approx \frac{1}{T} \int dy (y-x)^2 p_2(y,t+T|x,t) \quad (10)$$

as T vanishes.

Now, we generalize: from (10) we see that memory of n-2 states in the transition density will appear in the diffusion coefficient, and in the drift coefficient R for the same reason. That is, for a general Ito process we should replace (1) by

$$dx = R(x,t;x_{n-2},t_{n-2};...;x_1,t_1)dt + \sqrt{D(x,t;x_{n-2},t_{n-2};...;x_1,t_1)}dB(t)$$
(11)

where

$$D(x,t:x_{n-2},t_{n-2};...;x_1,t_1) \approx \frac{1}{T}\int dy(y-x)^2 p_n(y,t+T|x,t;x_{n-2},t_{n-2};...;x_1,t_1)$$
. (12)

That is, Ito processes are not restricted to Markov processes but include the generalization to finite memory. We'll provide an explicit example with n=3 below. There, the drift coefficient has memory of one notrivial state but the diffusion coefficient is constant because the process is Gaussian.

4. The meaning of Kolmogorov's first pde

Consider a twice differentiable dynamical variable A(x,t). The sde for A follows from Ito's lemma,

$$dA = \left(\frac{\partial A}{\partial t} + R\frac{\partial A}{\partial x} + \frac{D}{2}\frac{\partial^2 A}{\partial x^2}\right)dt + \sqrt{D}\frac{\partial A}{\partial x}dB, \quad (13)$$

so that

$$A(x(t+T),t+T) = A(x(t),t) + \int_t^{t+T}\left(\frac{\partial A(x(s),s)}{\partial t} + R\frac{\partial A}{\partial x} + \frac{D}{2}\frac{\partial^2 A}{\partial x^2}\right)ds + \int_t^{t+T}\sqrt{D(x(s),s)}\frac{\partial A(x(s),s)}{\partial x}dB(s)$$
. (14)

A martingale is defined by the conditional average $<A(x,t+T)>_c=A(x,t)$. From (14) we see that a backward in time average

$$A(x,t) = \int p^+(x,t:y,t+T)A(y,t+T)dy \qquad (15)$$

is required. We want to obtain the generator for the backward time transition density, which we denote as $p^+(x,t|y,t+T)$. Setting the drift term in (14) equal to zero, yields the backward time diffusion eqn.

$$0 = \frac{\partial A(x,t)}{\partial t} + R(x,t)\frac{\partial A(x,t)}{\partial x} + \frac{D(x,t)}{2}\frac{\partial^2 A(x,t)}{\partial x^2}. \quad (16)$$

We've made no assumption that A is positive. I.e., A is generally not a 1-point probability density, A(x,t) is simply any martingale, and an infinity of martingales can be so constructed depending on the choice of forward time initial conditions specified on A (either an initial value or boundary value problem backward in time is to be solved).

The required transition density is the Green function of (16),

$$0 = \frac{\partial g^+(x,t|y,s)}{\partial t} + R(x,t)\frac{\partial g^+(x,t|y,s)}{\partial x} + \frac{D(x,t)}{2}\frac{\partial^2 g^+(x,t|y,s)}{\partial x^2} \quad (17)$$

where $g^+(x,t|y,t)=\delta(x-y)$ where $t \le s$. The conditions under which g^+ exists, is unique and nonnegative definite are stated in Friedman [7]. Eqn. (17) is called Kolmogorov's first pde (K1) [8]. If g+ is nonnegative and normalizable, the g^+ may be identified as the backward time transition density p^+ for the Ito process.

What does K1 mean? Simply that martingales can be constructed via Ito's lemma.

5. The Fokker-Planck pde with finite memory

Consider next a twice-differentiable dynamical variable A(x(t)). A(x) is *not* assumed to be a martingale. The time evolution of A is given by Ito's lemma [6]

$$dA = (R\frac{\partial A}{\partial x} + \frac{D}{2}\frac{\partial^2 A}{\partial x^2})dt + \sqrt{D}\frac{\partial A}{\partial x}dB. \qquad (18)$$

We can calculate the conditional average of A, conditioned on $x(t_o)=x_o$ at time t_o in $x(t)=x_o+\int R(x,s)ds+\int\sqrt{D}(x,s)dB(s)$, forward in time if we know the transition density $p_2(x,t{:}x_o,t_o))$ forward in time,

$$\left\langle A(x(t))\right\rangle = \int p_2(x,t|x_o,t_o)A(x)dx. \qquad (19)$$

Note that this is not the rule for the time evolution of a 1-point probability density. From

$$\frac{d\left\langle A(x(t))\right\rangle}{dt} = \int \frac{\partial p_2(x,t|x_o,t_o)}{\partial t}A(x)dx \qquad (20)$$

and using

$$\left\langle dA\right\rangle = \left(\left\langle R\frac{\partial A}{\partial x}\right\rangle + \left\langle \frac{D}{2}\frac{\partial^2 A}{\partial x^2}\right\rangle\right)dt \qquad (21)$$

with <dA>/dt defined by (20), we obtain from (21), after integrating twice by parts and assuming that the boundary terms vanish,

$$\int dx A(x)\left[\frac{\partial p_2}{\partial t}+\frac{\partial(Rp_2)}{\partial x}-\frac{1}{2}\frac{\partial^2(Dp_2)}{\partial x^2}\right]=0, \qquad (22)$$

so that the transition density is the Green function of the Fokker-Planck pde [8,7,9,15], or Kolmogorov's second pde (K2)

$$\frac{\partial p_2}{\partial t}=-\frac{\partial(Rp_2)}{\partial x}+\frac{1}{2}\frac{\partial^2(Dp_2)}{\partial x^2}. \qquad (23)$$

So far, no Markovian assumption was made. In particular, no assumption was made that R, D, and hence p_2, are independent of memory of an initial state, or of finitely many earlier states. In particular, if D and R contain n-2 points of memory, and with (6) in mind, then p_2 must be understood as p_n, with $p_k=p_n$ for $k\geq n$. In this case the p_k with $k<n$ obey (4) and p_n obey (5).

For the case where A(x) *is* a martingale then (19) must yield

$$\left\langle A\right\rangle_t=\int p_2(x,t:x_o,t_o)A(x)dx=A(x_o), \qquad (24)$$

and since (24) cannot differ from (15) if the theory is to be consistent, the backward and forward time transition densities p^+ and p_2. Comparing (19) with (5) we see that p^+ and p_2 must be adjoints, $p^+(x,t|y,t+T)=p_n(y,t+T|x,t)$. The Chapman-Kolmogorov eqn. was not used to derive Kolmogorov's two pdes, nor has it been assumed. Next, we will show how an Ito process demands the Chapman-Kolmogorov eqn., and finite memory is allowed.

6. The Chapman-Kolmogorov eqn. for finite memory processes

That a Chapman-Kolmogorov eqn. is possible for finitely many states of memory follows from standard definitions of conditional probability densities. With an unstated, even infinite, number of states in memory the history-dependent 2-point transition densities obey the hierarchy

$$p_{k-1}(x_k,t_k|x_{k-2},t_{k-2};\ldots;x_1,t_1)=\int dx_{k-1}p_k(x_k,t_k|x_{k-1},t_{k-1};\ldots;x_1,t_1)p_{k-1}(x_{k-1},t_{k-1}|x_{k-2},t_{k-2};\ldots;x_1,t_1)\quad.\qquad(25)$$

For stochastic processes where the memory is finite and of number n-2, so that $p_k=p_n$ for all $k\geq n$, then from (25) we obtain the Chapman-Kolmogorov eqn. in the form

$$p_n(x_n,t_n|x_{n-1},t_{n-1};\ldots;x_1,t_1)=\int dy\,p_n(x_n,t_n|y,s;x_{n-2},t_{n-2};\ldots;x_1,t_1)p_n(y,s|x_{n-1},t_{n-1};\ldots;x_1,t_1)\quad.\qquad(26)$$

Next, we show that the pde K1 for an Ito process (21) with finite memory in R and/or D implies both the Fokker-Planck pde and the Chapman-Kolmogorov eqn. (26). This is the reverse of the usual derivation [8], where a Chapman-Kolmogorov equation is assumed and approximations are made to derive Kolmogorov’s two pdes. One realizes in retrospect that a Markov process is not implied there either, all that is expressed is the necessity by the typical derivation, is the necessity, not the sufficiency, of the Chapman-Kolmogorov eqn. for a Fokker-Planck pde. The two pdes describe a Markov process iff. one assumes in addition that there is no memory, or as the Russian translators [8] put it, ‘no aftereffect’.

Consider the linear operators

$$L^{+} = \partial/\partial t + R(x,t)\partial/\partial x + (D(x,t)/2)\partial^2/\partial x^2 \qquad (27)$$

and

$$Lu = -\partial u/\partial t + \partial(R(x,t)u)/\partial x - \partial^2(D(x,t)u/2)/\partial x^2, \qquad (28)$$

acting on a function space of measurable, twice (not necessarily continuously) differentiable functions satisfying boundary conditions at t=∞, and at x=-∞ and x=∞ to be indicated below. Both operators follow from the Ito process (1), but we can start with (27) and then obtain (28) via

$$uL^{+}v - vLu = \frac{\partial}{\partial t}(uv) + \frac{\partial}{\partial x}\left(vRu + \frac{1}{2}uD\frac{\partial v}{\partial x} - v\frac{1}{2}\frac{\partial uD}{\partial x}\right), \qquad (29)$$

which is a form of Green's identity (see also [7] where the operator L is studied in standard elliptic rather than in Fokker-Planck form). With suitable boundary conditions on u,v [4] then L and L^{+} are adjoints of each other:

$$\int_0^\infty dt \int_{-\infty}^\infty (vLu - uL^{+}v)dx = 0. \qquad (30)$$

Starting with an Ito process (1) and K1, we have deduced K2. No Markovian assumption has been made. Again, the formal conditions under which (30) holds are stated in Friedman [7].

Next, let $g^{+}(x,t:\xi,\tau)$ denote the Green function of K1, $L^{+}g^{+}=0$, and let $g(x,t:\xi,\tau)$ denote the Green function of K2, Lg=0. Let $\tau<s<t$ and assume also that $\tau+\varepsilon<s<t-\varepsilon$, which avoids sitting on top of a delta function. Integrating (29) over y from -∞ to ∞ and over s from $\tau+\varepsilon$ to $t-\varepsilon$ with the choices $v(y,s)=g^{+}(y,s:x,t)$ and $u(y,s)=g(y,s:\xi,\tau)$, we obtain [7]

$$\int g(y,t-\varepsilon:\xi,\tau)g^{+}(y,t-\varepsilon:x,t)dy=\int g(y,\tau+\varepsilon:\xi,\tau)g^{+}(\tau+\varepsilon:x,t)dy$$
. (31)

With ε vanishing and using $g(y,\tau:\xi,\tau)=\delta(y-\xi)$, $g^{+}(y,t:x,t)=\delta(y-x)$, we obtain the adjoint condition for the Green functions

$$g(x,t:\xi,\tau)=g^{+}(\xi,\tau:x,t). \qquad (32)$$

Next, apply the same argument but with times $\tau\le t''\le t'\le t$ to obtain (instead of (26))

$$\int g(y,t':\xi,\tau)g(x,t:y,t')dy=\int g(y,t'':\xi,\tau)g(x,t:y,t'')dy. \quad (33)$$

If we let t″ approach τ, then we obtain the Chapman-Kolmogorov eqn.

$$g(x,t:\xi,\tau)=\int g(x,t:y,t')g(y,t':\xi,\tau)dy, \qquad (34)$$

again, *without having made any Markovian assumption*. The implication is that, with suitable boundary conditions on Green functions, an Ito sde implies both K1 and K2 and the Chapman-Kolmogorov eqn.

To show that this formalism is not vacuous finite memory is present, we next provide the simplest example, a Gaussian process with memory of one nontrivial state in the drift coefficient (there is no memory in D because D=constant for a Gausian process). We can offer no example for variable diffusion D(x,t) where the (x,t) dependence is not separable, because even for selfsimilar Markov processes [16] we do not yet know how to calculate a transition density analytically.

7. A Gaussian process with finite memory

Consider first the 2-point transition density for an arbitrary Gaussian process in the form [15]

$$p(x,t:y,s) = \frac{1}{\sqrt{2\pi K(t,s)}} e^{-(x-m(t,s)y-g(t,s))^2/2K(t,s)}. \qquad (35)$$

Until the pair correlation function <x(t)x(s)> α m(t,s) is specified, no particular process is indicated by (35). Processes as wildly different and unrelated as fBm [5], scaling Gaussian Markov processes [5], Ornstein-Uhlenbeck processes [10] and other processes [14] are allowed. Depending on the pair correlation function <x(t)x(s)>, memory, including long time memory, may or may not appear. To obtain fBm, e.g., g=0 and <x(s)x(t)> must reflect the condition for stationary increments [5], which differs from a condition of time translational invariance whereby m, g, and K may depend on (s,t) only in the form s-t. Fortunately, Hänggi and Thomas [14] have stated the conditions for a Gaussian process (30) to satisfy a Chapman-Kolmogorov eqn., namely,

$$\begin{aligned} m(t,t_1) &= m(t,s)m(s,t_1) \\ g(t,t_1) &= g(t,s) + m(t,s)g(s,t_1) \quad . \qquad (36) \\ K(t,t_1) &= K(t,s) + m^2(t,s)K(s,t_1) \end{aligned}$$

Actually, Hänggi and Thomas stated in [14] that (36) is the condition for a *Markov* process, but we will show that the Chapman-Kolmogorov condition (31) is satisfied by at least one Gaussian process with memory.

Consider next the 1-point density $p_1(x,t)$ for a specific Ito process with simple memory in the drift coefficient, the Shimizu-Yamato model [1,12]

$$\frac{\partial p_1}{\partial t} = \frac{\partial}{\partial x}((\gamma+\kappa)x - \kappa\langle x(t)\rangle + \frac{Q}{2}\frac{\partial}{\partial x})p_1 \qquad (37)$$

with initial data $p(x,t_o)=f(x)$ and with $<x(t)>=\int xp_1(x,t)dx$. The parameter Q is the diffusion constant. Since the drift coefficient in (1) is $R=-(\gamma+\kappa)x+\kappa<x(t)>$, and since we can use standard methods to show that

$$\frac{d\langle x\rangle}{dt} = \langle R\rangle = -\gamma\langle x\rangle, \qquad (38)$$

we obtain

$$\langle x(t)\rangle = x_o e^{-\gamma(t-t_o)} \qquad (39)$$

where

$$x_o = \int xf(x)dx. \qquad (40)$$

This provides us with a drift coefficient with initial state memory,

$$R(x,t;x_o,t_o) = -(\gamma+\kappa)x + \kappa x_o e^{-\gamma(t-t_o)}. \qquad (41)$$

Because $\gamma\neq 0$ the memory cannot be eliminated via a simple coordinate transformation $z=x-<x>$.

The Fokker-Planck pde for the transition density $p_2(x,t{:}y,s;x_o,t_o)$ is

$$\frac{\partial p_2}{\partial t} = \frac{\partial}{\partial x}((\gamma+\kappa)x - \kappa x_o e^{-\gamma(t-t_o)} + \frac{Q}{2}\frac{\partial}{\partial x})p_2 \qquad (42)$$

with $p_2(x,t:y,t;x_o,t_o)=\delta(x-y)$. The solution is a Gaussian (35) with 1-state memory where

$$
\begin{aligned}
m(t,s) &= e^{-(\gamma+\kappa)(t-s)} \\
K(t,s) &= \frac{Q}{\gamma+\kappa}(1-e^{-2(\gamma+\kappa)(t-s)}) \qquad . \qquad (43) \\
g(t,s) &= x_o(e^{-\gamma(t-t_o)} - e^{-(\gamma+\kappa)t+\gamma t_o+\kappa s})
\end{aligned}
$$

An easy calculation shows that the Chapman-Kolmogorov conditions (36) are satisfied with finite memory (x_o,t_o). Furthermore, $p^+(y,s:x,t;x_o,t_o)=p_2(x,t:y,s;x_o,t_o)$ satisfies the backward time diffusion pde K1 in the variables (y,s),

$$
0 = \frac{\partial p^+}{\partial s} + R(y,s;x_o,t_o)\frac{\partial p^+}{\partial y} + \frac{Q}{2}\frac{\partial^2 p^+}{\partial y^2} \qquad (44)
$$

with drift coefficient

$$
R(y,s;x_o,t_o) = -(\gamma+\kappa)x + \kappa x_o e^{-\gamma(s-t_o)}. \qquad (45)
$$

This shows that backward time diffusion makes sense in the face of memory. The memory simply yields $p^+(y,t_o:x_o,t_o;x_o,t_o)=\delta(y-x_o)$.

8. Black-Scholes type pdes

Consider more generally Green functions of pdes of the Black-Scholes type

$$
L^+v = \partial v/\partial t + c(x,t)v + R(x,t)\partial v/\partial x + (D(x,t)/2)\partial^2 v/\partial x^2 = 0
$$
(46)

and its adjoint

$$Lu = cu - \partial u/\partial t + \partial(R(x,t)u)/\partial x - \partial^2(D(x,t)u/2)/\partial x^2 . (47)$$

We can prove exactly as in part 6 above that the Green functions of these pdes satisfy the Chapman-Kolmogorov equation (34). The proof appears as an exercise in Friedman [7].

The underlying Ito process is given by (1). Next follows the shortest possible proof that transition densities of the Fokker-Planck pde for stock returns, where R(x,t)=μ-D(x,t)/2 but with the unknown stock interest rate μ replaced by the risk neutral rate r [17], generates martingale option price and thereby provides the so-called 'martingale measure' of financial engineering. A proof via Girsanov's theorem [18] that Black-Scholes predicts fair option prices is therefore superfluous.

Initial value problems of (46), where v(x,T) is specified at a forward time T>t, are solved by a Martingale construction that results in the Feynman-Kac formula [7]. Defining M(s)=v(x,s)I(s), with dv(x,s) given by Ito's lemma and using (46) in Ito's lemma for dv we obtain

$$dM = dvI + vdI = -c(x,s)v(x,s)ds + v(x,s)dI(s) + \sqrt{D(x(s),s)}\frac{\partial v}{\partial x}I(s)dB(s)$$

. (48)

We obtain a martingale M(s)=v(x,s) with the choice

$$I(s) = e^{-\int_s^t c(x(q),q)dq} , \qquad (49)$$

so that the solution of (46) is given by the martingale condition M(t)=<M(T)>,

$$v(x,t) = \left\langle v(x(T),T)e^{\int_t^T c(x(s),s)ds} \right\rangle \qquad (50)$$

where the Feynman-Katz average (50) at time T is calculated using a functional integral based on the Green function $g^+(x,t|y,s)$ of (46) with c=0, i.e., the Green function of K1. This martingale construction for solutions of Black-Scholes type pdes (46) is given in [19][1] using unnecessarily complicated notation, and without the explanation of the connection of the Black-Scholes pde with K1, K2, and the Chapman-Kolmogorov eqn. The result () for Black-Scholes type pdes was derived by Friedman [3] over twenty years before it was rediscovered in financial economics by Duffie [20].

The Feynman-Katz formula is discussed in the financial math literature, but it is not used there to prove that the Black-Scholes transition density generates a martingale option price. With $x=\ln p(t)/p_c$, c=r and R=r-D/2, where r is the risk neutral interest rate, then we obtain the Black-Scholes pde written in the returns variable x. It follows that the initial value problem for pricing a call C(p,t)=v,

$$v(x_T,T) = (pe^{x_T} - K)\theta(pe^{x_T} - K), \quad (51)$$

is solved by

$$e^{-rt}v(x,t) = e^{-rT}\int_{-\infty}^{\infty} dx_T\, v(x_T,T)p_2(x_T,T|x,t) \qquad (52)$$

[1] In [19], eqns. (15.25) and (15.27) are inconsistent with each other, (15.25) cannot be obtained from (15.27) by a shift of coordinate origin because the x-dependent drift and diffusion coefficients break translation invariance. A careful treatment of solving elliptic and parabolic pdes by running an Ito process is provided by Friedman [7].

where $p=p_c e^x$ is the present price at time t and p_c is the consensus price ('value') [21], showing that the risk neutral discounted call price is a martingale. This result was proven in a different way earlier [17].

Acknowledgement

JMC thanks Enrico Scalas, Harry Thomas, Gemunu Gunaratne, and Kevin Bassler for many pleasant and very stimulating discussions via email about various points. Without Harry having sent us references [8,11], and without Enrico having pointed us to Doob's assertion that the Chapman-Kolmogorov eqn. is a necessary but insufficient condition for a Markov process [17], which led to Feller's paper [16], this paper would not have been possible.

References

1. J. L. McCauley, *Markov vs. nonMarkovian processes: A comment on the paper 'Stochastic feedback, nonlinear families of Markov processes, and nonlinear Fokker-Planck equations'* by T.D. Frank, *Physica **A382***, 445, 2007.

2. J.L. McCauley, K.E. Bassler, and G.H. Gunaratne, *Martingales, Detranding Data, and the Efficient Market Hypothesis, Physica **A37***, 202, 2008.

3. P. Hänggi, H. Thomas, H. Grabert, and P. Talkner, *Note on time Evolution of Non-Markov Processes*, *J. Stat. Phys*. **18**: 155 (1978).

4. B. Mandelbrot & J. W. van Ness, SIAM Rev. 10, 2, 422,1968.

5. J. L. McCauley , G.H. Gunaratne, & K.E. Bassler, *Hurst Exponents, Markov Processes, and Fractional Brownian Motion, Physica* ***A379*****,** 1**,** 2007.

6. K.E. Bassler, J. L. McCauley, & G.H. Gunaratne, *Nonstationary Increments, Scaling Distributions, and Variable Diffusion Processes in Financial Markets, PNAS* **104**, 17297, 23 Oct. 2008.

7. A. Friedman, *Stochastic Differential Equations and Applications* (Academic, N.Y., 1975).

8. B. V. Gnedenko, *The Theory of Probability*, tr. by B.D. Seckler (Chelsea, N.Y., 1967).

9. R.L. Stratonovich. *Topics in the Theory of Random Noise*, tr. By R. A. Silverman (Gordon & Breach: N.Y 1963).

10. M.C. Wang & G.E. Uhlenbeck in *Selected Papers on Noise and Stochastic Processes,* ed. N. Wax, Dover: N.Y., 1954.

11. T.D. Frank, *Stochastic feedback, nonlinear families of Markov processes, and nonlinear Fokker-Planck equations, Physica* ***A331***: 391 (2004).

12. J. L. Snell, *A Conversation with Joe Doob*, http://www.dartmouth.edu/~chance/Doob/conversation.html; *Statistical Science* **12**, No. 4, 301, 1997.

13. W. Feller, *The Annals of Math. Statistics* 30, No. 4, 1252, 1959.

14. P. Hänggi and H. Thomas, *Time Evolution, Correlations, and Linear Response of Non-Markov Processes, Zeitschr. Für Physik* **B26**: 85 (1977).

15. L. Arnold, *Stochastic Differential Equations* (Krieger, Malabar, 1992).

16. K.E. Bassler, G.H. Gunaratne, & J. L. McCauley, *Hurst Exponents, Markov Processes, and Nonlinear Diffusion Equations*, *Physica **A*** **369**: 343 (2006).

17. J. L. McCauley, G.H. Gunaratne, & K.E. Bassler, *Martingale Option Pricing*, *Physica **A380***, 351, 2007.

18. R. Durrett, *Brownian Motion and Martingales in Analysis*, Wadsworth, Belmont, 1984.

19. J.M. Steele, *Stochastic Calculus and Financial Applications* (Springer-Verlag, N.Y., 2000).

20. D. Duffie, *An Extension of the Black-Scholes Model of Security Valuation, J. Econ. Theory* **46**,194, 1988.

21. J. L. McCauley, K.E. Bassler, & G.H. Gunaratne, *On the Analysis of Time Series with Nonstationary Increments* in *Handbook of Complexity Research*, ed. B. Rosser, 2008.