22/11/10

Ergodicity:
An Econophysics Primer

Geoffrey Poitras*
Faculty of Business Administration
Simon Fraser University
Vancouver B.C.
CANADA V5A lS6

email: poitras@sfu.ca

ABSTRACT

This paper examines aspects of 'the ergodicity hypothesis' relevant to econophysics. The etymology and history of the concept in 19th century statistical mechanics are reviewed. An explanation of ergodicity is provided that uses Sturm-Liouville theory to decompose the transition probability density of a one-dimensional diffusion process subject to regular upper and lower reflecting barriers. The decomposition divides the transition density of an ergodic process into a limiting stationary density which is independent of time and initial condition, and a power series of time and boundary dependent transient terms. Properties of the quartic exponential stationary density are considered and used to assess the role of the ergodicity assumption in the *ex post / ex ante* quandary confronting important theories in financial economics.

Keywords: Ergodicity; Ludwig Boltzmann; Econophysics; quartic density; stationary distribution.

* The author is Professor of Finance at Simon Fraser University. The financial support of the SSHRC in funding this paper is gratefully acknowledged.

Ergodicity:
A Econophysics Primer

Schinckus [1, p.3816] accurately recognizes that the positivist philosophical foundations of econophysics depends fundamentally on empirical observation: "The empiricist dimension is probably the first positivist feature of econophysics". Following McCauley [2] and others`, this concern with empiricism focuses on the identification of macro-level statistical regularities that are typically characterized by scaling laws. Unfortunately, this empirically driven ideal is confounded by the 'non-repeatable' experiment that characterizes observed economic data. There is quandary posed by having only a single observed *ex post* time path to estimate the distributional parameters for the ensemble of *ex ante* time paths. Resolution of this quandary highlights the importance of the ergodicity assumption in defining the boundary between mainstream economics and econophysics. This paper examines aspects of 'the ergodicity hypothesis' relevant to econophysics. The etymology and history of the concept in 19th century statistical mechanics are reviewed. An explanation of ergodicity is provided that uses Sturm-Liouville theory to decompose the transition probability density of a one-dimensional diffusion process subject to regular upper and lower reflecting barriers. The decomposition divides the transition density of an ergodic process into a limiting stationary density which is independent of time and initial condition, and a power series of time and boundary dependent transient terms. Properties of the quartic exponential stationary density are considered and used to assess the role of the ergodicity assumption in the *ex post / ex ante* quandary confronting important theories in mainstream financial economics.

1. A Brief History of Ergodic Theory

The Encyclopedia of Mathematics [3] defines ergodic theory as the "metric theory of dynamical systems. The branch of the theory of dynamical systems that studies systems with an invariant measure and related problems." This modern definition implicitly identifies the birth of ergodic theory with proofs of the mean ergodic theorem by von Neumann [4] and the pointwise ergodic theorem by Birkhoff [5]. These early proofs have had significant impact in a wide range of modern subjects. For example, the notions of invariant measure and metric transitivity used in the proofs are fundamental to the measure theoretic foundation of modern probability theory (Doob [6]; Mackey [7]). Building on Kolmogorov [8], a seminal contribution to probability theory, in the years immediately following it was recognized that the ergodic theorems generalize the strong law of large numbers. Similarly, the equality of ensemble and time averages – the essence of the mean ergodic theorem – is necessary to the concept of a strictly stationary stochastic process. Ergodic theory is the basis for the modern study of random dynamical systems, e.g., Arnold [9]. In mathematics, ergodic theory connects measure theory with the theory of transformation groups. This connection is important in motivating the generalization of harmonic analysis from the real line to locally compact groups.

From the perspective of modern mathematics, statistical physics or systems theory, Birkhoff [5] and von Neumann [4] are excellent starting points for a history of ergodic theory. Building on the ergodic theorems, subsequent developments in these and related fields have been dramatic. These contributions mark the solution to a problem in statistical mechanics and thermodynamics that was recognized sixty years earlier when Ludwig Boltzmann (1844-1906) introduced the ergodic hypothesis to permit the theoretical phase space average to be interchanged with the measurable time average. From the perspective of both econophysics and mainstream economics, the selection of the less formally correct and rigorous contributions of Boltzmann are a more auspicious beginning for a history of the ergodic hypothesis. Problems of interest in mathematics are generated by a range of subjects, such as physics, chemistry, engineering and biology. The formulation and solution of physical problems in, say, statistical mechanics will have mathematical features which are unnecessary in, say, mainstream economics. For example, in statistical mechanics, points in the phase space are often multi-dimensional functions representing the mechanical state of the system, hence the desirability of a group-theoretic interpretation of the ergodic hypothesis. From the perspective of mainstream economics, such complications are largely irrelevant and an alternative history of ergodic theory that captures the etymology and basic physical interpretation is more revealing than a history that focuses on the relevance for mathematics. This arguably more revealing history begins with the formulation of the problems that von Neumann and Birkhoff were able to solve.

Mirowski [10, esp. ch.5] establishes the importance of 19th century physics in the development of the neoclassical economic system advanced by Jevons, Walras and Menger during the marginalist revolution of the 1870's. As such, neoclassical economic theory inherited essential features of mid-19th century physics: deterministic rational mechanics; conservation of energy; and the non-atomistic continuum view of matter that inspired the energetics movement later in the 19th century.[1] It was during the transition from rational to statistical mechanics during the last third of the century that Boltzmann made the contributions that led to the transformation of theoretical physics from the microscopic mechanistic models of Rudolf Clausius (1822-1888) and James Maxwell (1831-1879) to the macroscopic probabilistic theories of Josiah Gibbs (1839-1903) and Albert Einstein (1879-1955).[2] Coming largely after the start of the marginalist revolution in economics, this fundamental transformation in theoretical physics and mathematics had little impact on the progression of mainstream economic theory until the appearance of contributions on continuous time finance in the

1970's.[3] The deterministic mechanics of the energistic model was well suited to the subsequent axiomatic formalization of neoclassical economic theory which culminated in the von Neumann and Morgenstern expected utility approach to modeling uncertainty and the Bourbaki inspired Arrow-Debreu general equilibrium theory, e.g., Weintraub [11].

Having descended from the deterministic rational mechanics of mid-19th century physics, defining works of neoclassical economics, such as Hicks [12] and Samuelson [13], do not capture the probabilistic approach to modeling systems initially introduced by Boltzmann and further clarified by Gibbs.[4] Mathematical problems raised by Boltzmann were subsequently solved using tools introduced in a string of later contributions by the likes of the Ehrenfests and Cantor in set theory, Gibbs and Einstein in physics, Lebesque in measure theory, Kolmogorov in probability theory, Weiner and Levy in stochastic processes. Boltzmann was primarily concerned with problems in the kinetic theory of gases, formulating dynamic properties of the stationary Maxwell distribution – the velocity distribution of gas molecules in thermal equilibrium. Starting in 1871, Boltzmann took this analysis one step further to determine the evolution equation for the distribution function. The mathematical implications of this analysis still resonate in many subjects of the modern era. The etymology for "ergodic" begins with an 1884 paper by Boltzmann, though the initial insight to use probabilities to describe a gas system can be found as early as 1857 in a paper by Clausius and in the famous 1860 and 1867 papers by Maxwell.[5]

The Maxwell distribution is defined over the velocity of gas molecules and provides the probability for the relative number of molecules with velocities in a certain range. Using a mechanical model that involved molecular collision, Maxwell [14] was able to demonstrate that, in thermal equilibrium, this distribution of molecular velocities was a 'stationary' distribution that would not change shape due to ongoing molecular collision. Boltzmann aimed to determine whether the Maxwell distribution would emerge in the limit whatever the initial state of the gas. In order to study the dynamics of the equilibrium distribution over time, Boltzmann introduced the probability distribution of the relative time a gas molecule has a velocity in a certain range while still retaining the notion of probability for velocities of a relative number of gas molecules. Under the ergodic hypothesis, the average behavior of the macroscopic gas system, which can objectively be measured over time, can be interchanged with the average value calculated from the ensemble of unobservable and highly complex microscopic molecular motions at a given point in time. In the words of Weiner [15, p.1]: "Both in the older Maxwell theory and in the later theory of Gibbs, it is necessary to make some sort of logical transition between the average behavior of all dynamical systems of a given family or ensemble, and the historical average of a single system."

II. Ergodicity in Mainstream Economics

The ergodic hypothesis associated with the statistical mechanics of the kinetic gas model is distinct from the various concepts of ergodicity encountered in modern economics; if only because the complex microscopic interactions of individual gas molecules have to obey the second law of thermodynamics which has no corresponding concept in economics.[6] Despite differences in physical interpretation, there are two fundamental difficulties associated with the ergodic hypothesis in Boltzmann's statistical mechanics – reversibility and recurrence – that have a rough similarity to notions arising in heterodox critiques of mainstream economics.[7] These difficulties are compounded by different applications of the ergodic hypothesis arising in mainstream economics. In econometrics, ergodicity is necessary for both strict and covariance stationarity of a stochastic process. In addition, the economic models being tested may also have ergodic restrictions, e.g., the models employ rational expectations or Markovian dynamics. In financial economics, ergodicity is required for the stochastic differential equations used in option pricing models. Examples in other economic subjects are readily available, e.g., whenever a 'mean-reverting' process is used.

Even though the formal solutions proposed were inadequate by standards of modern mathematics, the thermodynamic model introduced by Boltzmann to explain the dynamic properties of the Maxwell distribution is a pedagogically useful starting point to develop the implications of ergodicity in economics. To be sure, von Neumann [4] and Birkhoff [5] correctly specify ergodicity using Lebesque integration – an essential analytical tool unavailable to Boltzmann – but the analysis is too complex to be of much value to all but the most mathematically specialized economists. The physical intuition of the kinetic gas model is lost in the generality of the results. Using Boltzmann as a starting point, the large number of mechanical and complex molecular collisions could correspond to the large number of microscopic, atomistic competitors and consumers interacting to determine the macroscopic market price.[8] In this context, it is variables such as the asset price or the interest rate or the exchange rate, or some combination, that is being measured over time and ergodicity would be associated with the properties of the transition density generating the macroscopic variables. Ergodicity can fail for a number of reasons and there is value in determining the source of the failure.

Halmos [16, p.1017] is a helpful starting point to sort out the differing notions of ergodicity that can arise in range of subjects: "The ergodic theorem is a statement about a space, a function and a transformation". In mathematical terms, ergodicity or 'metric transitivity' is a property of 'indecomposable', measure preserving transformations. Because the transformation acts on points in the space, there is a fundamental connection to the method of measuring relationships such as distance or volume in the space. In von Neumann [4] and Birkhoff [5], this is accomplished using the notion of Lebesque measure: the admissible functions are either integrable (Birkhoff) or square integrable (von Neumann). In contrast to, say, statistical mechanics where spaces and functions account for the complex physical interaction of large numbers of particles, economics can often specify the space in a mathematically convenient fashion. For example, in the case where there is a single random variable, then the space is "superfluous" (Mackey [7, p.182]) as the random variable is completely described by the distribution. Multiple random variables can be handled by assuming the random variables are discrete with finite state spaces. In effect, conditions for an 'invariant measure' can often be assumed in economics in order to focus attention on "finding and studying the invariant measures" (Arnold [9, p.22]) where, in the terminology of econometrics, the invariant measure corresponds to the stationary distribution or likelihood function.

The mean ergodic theorem of von Neumann [4] provides an essential connection to the ergodicity hypothesis in econometrics. It is well known that, in the Hilbert and Banach spaces common to econometric work, the mean ergodic theorem corresponds to the strong law of large numbers. In statistical applications where strictly stationary distributions are assumed, the relevant ergodic transformation, U^*, is the unit shift operator: $U^* \Psi[x(t)] = \Psi[U^* x(t)] = \Psi[x(t+1)]$; $[(U^*)^k] \; \Psi[x(t)] = \Psi[x(t+k)]$; and $\{(U^*)^{-k}\} \; \Psi[x(t)] = \Psi[x(t\text{-}k)]$ with k being an integer and $\Psi[x]$ the strictly stationary distribution for x that in the strictly stationary case is replicated at each t. Significantly, this reversible transformation is independent of initial time and state. Because this transformation imposes strict stationarity on $\Psi[x]$, U^* will only work for certain ergodic processes. In effect, the ergodic requirement that the transformation be measure preserving is weaker than the strict stationarity of the stochastic process required for U^*. The implications of the reversible ergodic transformation U^*, are described by Davidson [17, p.331]: "In an economic world governed entirely by ergodic processes ... economic relationships among variables are timeless, or ahistoric in the sense that the future is merely a statistical reflection of the past". In effect, mainstream economics requires that the real world distribution for $x(t)$ be sufficiently similar to those for both $x(t+k)$ or $x(t-k)$, i.e., the ergodic transformation U^* is reversible.

Critiques of mainstream economics that are rooted in the insights of *The General Theory* recognize the distinction between fundamental uncertainty and objective probability. It is unlikely that Keynes gained much exposure to Birkhoff [5] and von Neumann [4] or the substantive extensions and applications that appeared in the years following these contributions. As a consequence, the definition of ergodic theory in heterodox criticisms of mainstream economics lacks formal precision, e.g., "There is no reason to presume that structures will remain stable; the economic system is nonergodic" (Dow [18, p.387]). Ergodic theory is implicitly seen as another piece of the mathematical formalism inspired by Hilbert and Bourbaki and captured in the Arrow-Debreu general equilibrium model of mainstream economics. Yet, the 19th century statistical mechanics that inspired ergodic theory is grounded in real world problems; in particular, Birkhoff [5] and von Neumann [4] formally solved the Boltzmann problem of incorporating dynamic phase transitions – from gas to liquid and from liquid to solid – where the mechanical model governing the ergodic process changes abruptly. Though there are a variety of ergodic transformations that incorporate such possibilities, heterodox critiques of mainstream economics do not explore such transformations.

III. The Phenomenological Approach to Ergodicity

Important economic variables, such as wage rates, prices, incomes, interest rates and the like, have relatively innocuous sample paths compared to some types of variables encountered in subjects such as physics, chemistry or biology. There is an impressive range of mathematical and statistical models that, seemingly, could be applied to almost any physical or economic situation. If the process can be verbalized, then a model can be specified. This begs the question: are there transformations – ergodic or otherwise – that capture the basic 'stylized facts' of observed economic data? Significantly, the random instability in the observed sample paths identified in, say, financial time series is consistent with the stochastic bifurcation of an ergodic process, e.g., Chiarella et al. [19]. The associated *ex ante* stationary densities are multimodal and irreversible, a situation where the mean calculated from past values of a single, non-experimental *ex post* realization the process is not informative about the mean for future values.

Boltzmann was concerned with demonstrating that the Maxwell distribution emerged in the limit as $t \rightarrow \infty$ for systems with large numbers of particles. The limiting process for t requires that the system run long enough that the initial conditions do not impact the stationary distribution. At the time, two fundamental criticisms were aimed at this general approach: reversibility and recurrence. In the context of mainstream economics, reversibility relates to the use of past values of the process to forecast future values.[9] Recurrence relates to the properties of the long run average which involves the ability and length of time for an ergodic process to return to its stationary state. For Boltzmann, both these criticisms have roots in the difficulty of reconciling the second law of thermodynamics with the ergodicity hypothesis. Using a mathematical framework where ergodicity requires the transition density of the process to be decomposable into the sum of a stationary density and a mean zero transient term that captures the impact of the initial condition of the system on the individual sample paths, irreversibility relates to properties of the stationary density and non-recurrence to the behavior of the transient term. The objective of this section is to examine these properties for irreversible ergodic processes.

Because the particle movements in a kinetic gas model are contained within an enclosed system, e.g., a vertical glass tube, classical Sturm-Liouville (S-L) methods can be applied to obtain solutions for the transition densities. These results for the distributional implications of imposing regular reflecting boundaries on diffusion processes are representative of the “phenomenonological approach” to random systems theory which: “studies qualitative changes of the densites of invariant measures of the Markov semigroup generated by random dynamical systems induced by stochastic differential equations” (Crauel et al.[20, p.27]). Because the initial condition of the system is explicitly recognized, ergodicity in these models takes a different form than that associated with the unit shift transformation which is central to mainstream economics. The ergodic transition densities are derived as solutions to the forward differential equation associated with one-dimensional diffusions. The transition densities contain a transient term that is dependent on the initial condition of the system and boundaries imposed on the state space. Irreversibility is introduced by employing multi-modal stationary densities.

The distributional implications of boundary restrictions, derived by modeling the random variable as a diffusion process subject to reflecting barriers, have been studied for many years, e.g., Feller [21]. The diffusion process framework is useful because it imposes a functional structure that is sufficient for known partial differential equation (PDE) solution procedures to be used to derive the relevant transition probability densities. Wong [22] demonstrated that with appropriate specification of parameters in the PDE, the transition densities for popular stationary distributions such as the exponential, uniform, and normal distributions can be derived using S-L methods. This paper proposes that the S-L framework does have sufficient generality to permit a formalization that captures key stylized facts raised by empirical difficulties arising from non-experimental economic time series. In turn, the framework suggests a method of generalizing mainstream economic theory to encompass the nonlinear dynamics of diffusion processes. In other words, within the more formal mathematical framework of econophysics, it is possible to reformulate the ergodicity assumption to permit a stochastic generalization of mainstream economic theory.

The use of the diffusion model to represent the nonlinear dynamics of stochastic processes is found in a wide range of subjects. Physical restrictions such as the rate of observed genetic mutation in biology or character of heat diffusion in engineering often determine the specific formalization of the diffusion model. Because physical interactions can be complex, mathematical results for diffusion models are pitched at a level of generality sufficient to cover such cases.[10] Such generality is usually not required in economics. In this vein, it is possible to exploit mathematical properties of bounded state spaces and one dimensional diffusions to overcome certain analytical problems that can confront continuous time Markov solutions. The key construct in the S-L method is the ergodic transition probability density function U which is associated with the random (economic) variable x at time t ($U = U[x, t \mid x_0]$) that follows a regular, time homogeneous diffusion process. While it is possible to allow the state space to be an infinite open interval $I_o = (a,b\!: \infty \leq a < b \leq \infty)$, a finite closed interval $I_c = [\,a,b\!: -\infty < a < b < +\infty\,]$ or the specific interval $I_s = [0 = a < b < \infty)$ are applicable to economic variables.[11] Assuming that U is twice continuously differentiable in x and once in t and vanishes outside the relevant interval, then U obeys the forward equation (e.g., Gihhman and Skorohod [23, p.102-4]):

$$\frac{\partial^2}{\partial x^2}\left\{ B[x]\ U \right\} - \frac{\partial}{\partial x}\{A[x]\ U\} = \frac{\partial U}{\partial t} \qquad (1)$$

where: $B[x]$ ($= \frac{1}{2}\sigma^2[x] > 0$) is the one half the infinitesimal variance and $A[x]$ the infinitesimal drift of the process. $B[x]$ is assumed to be twice and $A[x]$ once continuously differentiable in x. Being time homogeneous, this formulation permits state, but not time, variation in the drift and variance parameters.

If the diffusion process is subject to upper and lower reflecting boundaries that are regular and fixed ($-\infty < a < b < \infty$), the "Sturm-Liouville problem" involves solving (1) subject to the separated boundary conditions:[12]

$$\frac{\partial}{\partial x} \{ B[x] \ U[x,t] \} \Big|_{x=a} - A[a] \ U[a,t] = 0 \tag{3}$$

$$\frac{\partial}{\partial x} \{ B[x] \ U[x,t] \} \Big|_{x=b} - A[b] \ U[b,t] = 0 \tag{4}$$

And the initial condition:

$$U[x,0] = f[x_0] \qquad where: \quad \int_a^b f[x_0] = 1 \tag{5}$$

and $f[x_0]$ is the continuous density function associated with x_0 where $a \leq x_0 \leq b$. When the initial starting value, x_0, is known with certainty, the initial condition becomes the Dirac delta function, $U[x,0] = \delta[\ x - x_0\ ]$, and the resulting solution for U is referred to as the 'principal solution'. Within the framework of the S-L method, a stochastic process has **the ergodic property** when the transition density satisfies:[13]

$$\lim_{t \to \infty} U[x,t \mid x_0] = \int_a^b f[x_0] \quad U[x,t \mid x_0] \ dx_0 = \Psi[x]$$

Important special cases occur for the principal solution ($f[x_0] = \ \delta[\ x - x_0\ ]$) and when $f[x_0]$ is from a specific class such as the Pearson distributions. To allow for phenomena such as phase transitions, the time invariant stationary density $\Psi[x]$ is permitted to 'decompose' into a finite number of indecomposable sub-densities, each of which is time invariant. Such irreversible processes are ergodic, because each of the sub-densities obeys the ergodic theorem, but not strictly stationary because the multi-modal stationary density is split into ergodic sub-densities with different means. Such irreversible ergodic processes have the property that the mean calculated from past values of the process are not generally informative enough about the means of the *ex ante* sub-densities.

In order to more accurately capture the *ex ante* properties of economic time series, there are some potentially restrictive features in the S-L framework that can be identified. For example, time homogeneity of the process eliminates the need to explicitly consider the location of t_0.[14] Time homogeneity is a property of both U and U^* and, as such, is consistent with 'ahistorical' mainstream economic theorizing. In the case of U^*, the time homogeneous and reversible stationary distribution governs the dynamics of $x(t)$. Significantly, while U is time homogeneous, there are some U consistent with irreversible processes. The relevant issue for econophysics is to determine which concept – time homogeneity or reversibility – is inconsistent with economic processes that capture: ratchet effects in wages; liquidity traps in money markets; structural shifts; and, collapsing conventions. In the S-L framework, the initial state of the system (x_0) is known and the ergodic transition density provides information about how a given point x_0 shifts t units along a trajectory.[15] For econometric applications employing the strictly stationary transformation U^*, the location of x_0 is irrelevant while U incorporates x_0 as an initial condition associated with the solution of a partial differential equation.

IV. Density Decomposition Results[16]

In general, solving the forward equation (1) for U subject to (3), (4) and some admissible form of (5) is difficult, e.g., Feller [21], Risken [24]. In such circumstances, it is expedient to restrict the problem specification to permit closed form solutions for the transition density to be obtained. Wong [22] provides an illustration of this approach. The PDE (1) is reduced to an ODE by only considering the strictly stationary distributions arising from the Pearson system. Restrictions on the associated $\Psi[x]$ are constructed by imposing the fundamental ODE condition for the unimodal Pearson system of distributions:

$$\frac{d\Psi[x]}{dx} = \frac{e_1 x + e_0}{d_2 x^2 + d_1 x + d_0} \Psi[x]$$

The transition probability density U for the ergodic process can then be reconstructed by working back from a specific closed form for the stationary distribution using known results for the solution of specific forms of the forward equation. In this procedure, the d_0, d_1, d_2, e_0 and e_1 in the Pearson ODE are used to specify the relevant parameters in (1). The U for important stationary distributions that fall within the Pearson system, such as the normal, beta, central t, and exponential, can be derived by this method.

The solution procedure employed by Wong [22] depends crucially on restricting the PDE problem sufficiently to apply classical S-L techniques. Using S-L methods, various studies have generalized the set of solutions for U to cases where the stationary distribution is not a member of the Pearson system or U is otherwise unknown, e.g., Linetsky [25]. In order to employ the separation of variables technique used in solving S-L problems,(1) has to be transformed into the canonical form of the forward equation. To do this, the following function associated with the invariant measure is introduced:

$$r[x] = B[x] \exp\left[-\int_a^x \frac{A[s]}{B[s]} \, ds \right]$$

Using this function, the forward equation can be rewritten in the form:

$$\frac{1}{r[x]}\ \frac{\partial}{\partial x}\left\{ p[x]\ \frac{\partial U}{\partial x} \right\} + q[x]\ U = \frac{\partial U}{\partial t} \qquad (6)$$

$$where: \qquad p[x] = B[x]\ r[x] \qquad q[x] = \frac{\partial^2 B}{\partial x^2} - \frac{\partial A}{\partial x}$$

Equation (6) is the canonical form of equation (1). The S-L problem now involves solving (6) subject to appropriate initial and boundary conditions.

Because the methods for solving the S-L problem are ODE-based, some method of eliminating the time derivative in (1) is required. Exploiting the assumption of time homogeneity, the eigenfunction expansion approach applies separation of variables, permitting (6) to be specified as:

$$U[x,t] = e^{-\lambda\, t}\ \varphi[x] \qquad (7)$$

Where $\varphi[x]$ is only required to satisfy the easier-to-solve ODE:

$$\frac{1}{r[x]}\ \frac{d}{dx}\left[p[x]\ \frac{d\varphi}{dx} \right] + [q[x] + \lambda]\ \varphi[x] = 0 \qquad (1')$$

Transforming the boundary conditions involves substitution of (7) into (3) and (4) and solving to get:

$$\frac{d}{dx}\{B[x]\ \varphi[x]\ \}\Big|_{x=a} - A[a]\ \varphi[a] = 0 \qquad (3')$$

$$\frac{d}{dx}\{B[x]\ \varphi[x]\ \}\Big|_{x=b} - A[b]\ \varphi[b] = 0 \qquad (4')$$

Significant analytical advantages are obtained by making the S-L problem 'regular' which involves assuming that $[a,b]$ is a closed interval with $r[x]$, $p[x]$ and $q[x]$ being real valued and $p[x]$ having a continuous derivative on $[a,b]$; and, $r[x] > 0$, $p[x] > 0$ at every point in $[a,b]$. 'Singular' S-L problems arise where these conditions are violated due to, say, an infinite state space or a vanishing coefficient in the interval [a,b]. The separated boundary conditions (3) and (4) ensure the problem is self-adjoint (Berg and McGregor [26, p.91]).

The S-L problem of solving (6) subject to the initial and boundary conditions admits a solution only for certain critical values of λ, the eigenvalues. Further, since equation (1) is linear in U, the general solution for (7) is given by a linear combination of solutions in the form of eigenfunction expansions. Details of these results can be found in Hille [27, ch. 8], Birkhoff and Rota [28, ch. 10] and Karlin and Taylor [29]. When the S-L problem is self-adjoint and regular the solutions for the transition probability density can be summarized in the following (see n.19):

Proposition: Ergodic Transition Density Decomposition

The regular, self-adjoint Sturm-Liouville problem has an infinite sequence of real eigenvalues, $0 = \lambda_0 < \lambda_1 < \lambda_2 < \ldots$ with:

$$\lim_{n \to \infty} \lambda_n = \infty$$

To each eigenvalue there corresponds a unique eigenfunction $\varphi_n \equiv \varphi_n[x]$. Normalization of the eigenfunctions produces:

$$\psi_n[x] = \left[\int_a^b r[x] \ \varphi_n^2 \ dx \right]^{-1/2} \varphi_n$$

The $\psi_n[x]$ eigenfunctions form a complete orthonormal system in $L_2[a,b]$. The unique solution in $L_2[a,b]$ to (1), subject to the boundary conditions (3)-(4) and initial condition (5) is, in general form:

$$U[x,t] = \sum_{n=0}^{\infty} c_n \ \psi_n[x] \ e^{-\lambda_n t} \qquad (8)$$

$$where: \qquad c_n = \int_a^b r[x] \ f[x_0] \ \psi_n[x] \ dx$$

Given this, the transition probability density function for x at time t can be reexpressed as the sum of a stationary limiting equilibrium distribution associated with the $\lambda_0 = 0$ eigenvalue, that is linearly independent of the boundaries, and a power series of transient terms, associated with the remaining eigenvalues, that are boundary and initial condition dependent:

$$U[x,t \mid x_0] \;=\; \Psi[x] \;+\; T[x,t \mid x_0] \tag{9}$$

$$where: \qquad \Psi[x] \;=\; \frac{r[x]^{-1}}{\int_a^b r[x]^{-1}\, dx} \tag{10}$$

Using the specifications of λ_n, c_n, and ψ_n, the properties of $T[x,t]$ are defined as:

$$T[x,t \mid x_0] \;=\; \sum_{n=1}^{\infty} c_n\, e^{-\lambda_n t}\, \psi_n[x] \;=\; \frac{1}{r[x]} \sum_{n=1}^{\infty} e^{-\lambda_n t}\, \psi_n[x]\, \psi_n[x_0] \tag{11}$$

$$with: \qquad \int_a^b T[x,t \mid x_0]\, dx \;=\; 0 \qquad and \qquad \lim_{t \to \infty} T[x,t \mid x_0] \;=\; 0$$

This Proposition provides the general solution to the regular, self-adjoint S-L problem of deriving U when the process is subject to regular reflecting barriers. Taking the limit as $t \to \infty$ in (9), it follows from (10) and (11) that the transition density of the stochastic process satisfies the ergodic property. Considerable effort has been given to determining the convergence behavior of different processes. The distributional impact of the initial conditions and boundary restrictions enter through $T[x,t]$ ($= T[x,t \mid x_0]$). From the restrictions on $T[x,t]$ in (11), the total mass of the transient term is zero so the mean ergodic theorem still applies. The transient only acts to redistribute the mass of the stationary distribution, thereby causing a change in shape. The specific degree and type of alteration depends on the relevant assumptions made about the parameters and initial functional forms.

The theoretical advantage obtained by imposing regular reflecting barriers on the diffusion state space for the forward equation is that an ergodic decomposition of the transition density is assured.

The relevance of bounding the state space and imposing regular reflecting boundaries can be illustrated by considering the well known solution (e.g., Cox and Miller [30, p.209]) for U involving a constant coefficient standard normal variate $Y(t) = (\{x - x_0 - \mu t\}/\sigma)$ over the unbounded state space $I_o = (\infty \leq x \leq \infty)$. In this case the forward equation (1) reduces to: $\frac{1}{2}\{\partial^2 U / \partial Y^2\} = \partial U / \partial t$. By evaluating these derivatives, it can be verified that the principal solution for U is:

$$U[x,t \mid x_0] = \frac{1}{\sigma\sqrt{(2\pi t)}} \exp\left[-\frac{(x - x_0 - \mu t)^2}{2\sigma^2 t} \right]$$

and as $t \to -\infty$ or $t \to +\infty$ then $U \to 0$ and the stochastic process is nonergodic because it does not possess a non-trivial stationary distribution. The mean ergodic theorem fails: if the process runs long enough, then U will evolve to where there is no discernible probability associated with starting from x_0 and reaching the neighborhood of a given point x. The absence of a stationary distribution raises a number of questions, e.g., whether the process has unit roots. Imposing regular reflecting boundaries is a certain method of obtaining a stationary distribution and a discrete spectrum (Hansen and Schienkman [31, p.13]). Alternative methods, such as specifying the process to admit natural boundaries where the parameters of the diffusion are zero within the state space, can give rise to continuous spectrum and raise significant analytical complexities. At least since Feller [21], the search for useful solutions, including those for singular diffusion problems, has produced a number of specific cases of interest. However, without the analytical certainty of the S-L framework, analysis proceeds on a case by case basis.

One possible method of obtaining a stationary distribution without imposing both upper and lower boundaries is to impose only a lower (upper) reflecting barrier and construct the stochastic process such that positive (negative) infinity is non-attracting, e.g., Linetsky [25]; Aït-Sahalia [32]. This can be achieved by using a mean-reverting drift term. In contrast, Cox and Miller [30, p.223-5] use the Brownian motion, constant coefficient forward equation with $x_0 > 0$, $A[x] = \mu < 0$ and $B[x] = ½\sigma^2$ subject to the lower reflecting barrier at $x = 0$ given in (2) to solve for both the U and the stationary density. The principal solution is solved using the 'method of images' to obtain:

$$U[x,t \mid x_0] = \frac{1}{\sigma\sqrt{2\pi t}} \left\{ \exp-\left(\frac{(x - x_0 - \mu t)^2}{2\sigma^2 t} \right) + \exp-\left(\frac{4x_0\mu t - (x - x_0 - \mu t)^2}{2\sigma^2 t} \right) \right\}$$
$$+ \frac{1}{\sigma\sqrt{2\pi t}} \left\{ \frac{2\mu}{\sigma^2} \exp\left(\frac{2\mu x}{\sigma^2} \right) \left(1 - N\left[\frac{x + x_0 + \mu t}{\sigma\sqrt{t}} \right] \right) \right\}$$

where $N[x]$ is again the cumulative standard normal distribution function. Observing that $A[x] = \mu > 0$ again produces $U \rightarrow 0$ as $t \rightarrow + \infty$, the stationary density for $A[x] = \mu < 0$ has the Maxwell form:

$$\Psi[x] = \frac{2|\mu|}{\sigma^2} \exp-\left(\frac{2|\mu|x}{\sigma^2} \right)$$

Though x_0 does not enter the solution, combined with the location of the boundary at $x = 0$, it does implicitly impose the restriction $x > 0$. From the Proposition, $T[x,t \mid x_0]$ can be determined as $U[x,t \mid x_0] - \Psi[x]$.

Following Linetsky [25], Veerstraeten [33] and others, the analytical procedure used to determine U involves specifying the parameters of the forward equation and the boundary conditions and then solving for $\Psi[x]$ and $T[x,t]$. Wong [22] uses a different approach, initially selecting a stationary distribution and then solving for U using the restrictions of the Pearson system to specify the forward equation. In this approach, the functional form of the desired stationary distribution determines the appropriate boundary conditions. While application of this approach has been limited to the restricted class of distributions associated with the Pearson system, it is expedient when a known stationary distribution, such as the standard normal distribution, is of interest. More precisely, let:

$$\Psi[x] = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{x^2}{2}\right], \qquad I_o = (-\infty < x < \infty)$$

In this case, the boundaries of the state space are non-attracting and not regular. Solving the Pearson equation gives: $d\Psi[x]/dx = -x\,\Psi[x]$ and a forward equation of the OU form:

$$\frac{\partial^2 U}{\partial x^2} + \frac{\partial}{\partial x} xU = \frac{\partial U}{\partial t}$$

Following Wong [22, p.268] Mehler's formula can be used to express the solution for U as:

$$U[x,t \mid x_0] = \frac{1}{\sqrt{2\pi(1 - e^{-2t})}} \exp\left[\frac{-(x - x_0\, e^{-t})^2}{2(1 - e^{-2t})}\right]$$

Given this, as $t \to -\infty$ then $U \to 0$ and as $t \to +\infty$ then U achieves the stationary standard normal distribution.

V. The Quartic Exponential Distribution

The roots of bifurcation theory can be found in the early solutions to certain deterministic ordinary differential equations. Consider the deterministic dynamics described by the pitchfork bifurcation ODE:

$$\frac{dx}{dt} = -x^3 + \rho_1\, x + \rho_0$$

where ρ_0 and ρ_1 are the 'normal' and 'splitting' control variables, respectively (e.g., Cobb [34, 35]). While ρ_0 has significant information in a stochastic context, this is not usually the case in the deterministic problem so $\rho_0 = 0$ is assumed. Given this, for $\rho_1 \leq 0$, there is one real equilibrium ($\{dx/dt\} = 0$) solution to this ODE at $x = 0$ where "all initial conditions converge to the same final point exponentially fast with time" (Caudel and Flandoli [36, p.260]). For $\rho_1 > 0$, the solution bifurcates into three equilibrium solutions $x = \{ 0, \pm\sqrt{\rho_1} \}$, one unstable and two stable. In this case, the state space is split into two physically distinct regions (at $x = 0$) with the degree of splitting controlled by the size of ρ_1. Even for initial conditions that are 'close', the equilibrium achieved will depend on the sign of the initial condition. Stochastic bifurcation theory extends this model to incorporate Markovian randomness. In this theory, "invariant measures are the random analogues of deterministic fixed points" (Arnold [9, p.469]). Significantly, ergodicity now requires that the component densities that bifurcate out of the stationary density at the bifurcation point be invariant measures, e.g., Crauel et al. [20, sec.3]. As such, the ergodic bifurcating process is irreversible in the sense that past sample paths (prior to the bifurcation) cannot reliably be used to generate statistics for the future values of the variable (after the bifurcation).

It is well known that the introduction of randomness to the pitchfork ODE changes the properties of the equilibrium solution, e.g., (Arnold [9, sec.9.2]). It is no longer necessary that the state space for the principal solution be determined by the location of the initial condition relative to the bifurcation point. The possibility for randomness to cause some paths to cross over the bifurcation point depends on the size of volatility of the process, σ, which measures the non-linear signal to white noise ratio. Of the different approaches to introducing randomness (e.g., multiplicative noise), the simplest approach to converting from a deterministic to a stochastic context is to add a Weiner process ($dW(t)$) to the ODE. Augmenting the diffusion equation to allow for σ to control the relative impact of non-linear drift versus random noise produces the "pitchfork bifurcation with additive noise" (Arnold [9, p.475]) which in symmetric form is:

$$dX(t) \;=\; (\rho_1 \; X(t) \;-\; X(t)^3) \; dt \;+\; \sigma \; dW(t)$$

In economic applications, e.g., Aït-Sahlia [32], this diffusion process is referred to as the double well process. While consistent with the common use of diffusion equations in econometrics and other areas of mainstream economics, the dynamics of the pitchfork process captured by $T[x,t \mid x_0]$ have been "forgotten" (Arnold [9, p.473]).

Mainstream economics is married to the transition probability densities associated with unimodal stationary distributions. More flexibility in the shape of the stationary distribution can be achieved using a higher order exponential density, e.g., Fisher [37], Cobb et al. [38], Caudel and Flandoli [36]. Increasing the degree of the polynomial in the exponential comes at the expense of introducing additional parameters resulting in a substantial increase in the analytical complexity, typically defying a closed form solution for the transition densities. However, at least since Elliott [39], it has been recognized that the solution of the associated regular S-L problem will still have a discrete spectrum, even if the specific form of the eigenfunctions and eigenvalues in $T[x,t \mid x_0]$ are not precisely determined (Horsthemke and Lefever [40, sec. 6.7]). Inferences about transient stochastic behavior can be obtained by examining the solution of the deterministic non-linear dynamics. In this process, attention initially focuses on the properties of the higher order exponential distributions.

To this end, assume that the stationary distribution is a fourth degree or "general quartic" exponential:

$$\Psi[x] = K \exp[-\Phi[x]] = K \exp[-(\beta_4 x^4 + \beta_3 x^3 + \beta_2 x^2 + \beta_1 x)]$$

where: K is a constant determined such that the density integrates to one; and, $\beta_4 > 0$.[17] Following Fisher [37], the class of distributions associated with the general quartic exponential admits both unimodal and bimodal densities and nests the standard normal as a limiting case where $\beta_4 = \beta_3 = \beta_1 = 0$ and $\beta_2 = ½$ with $K = 1/(\sqrt{2\pi})$. The stationary distribution of the bifurcating double well process is a special case of the symmetric quartic exponential distribution:

$$\Psi[y] = K_S \exp[-\{\beta_2 (x - \mu)^2 + \beta_4 (x - \mu)^4\}] \qquad where \quad \beta_4 \geq 0$$

where μ is the population mean and the symmetry restriction requires $\beta_1 = \beta_3 = 0$. Such multi-modal stationary densities have received scant attention in mainstream economics. To see why the condition on β_1 is needed, consider change of origin $X = Y - \{\beta_3 / 4\, \beta_4\}$ to remove the cubic term from the general quartic exponential (Matz [41, p.480]):

$$\Psi[y] = K_Q \exp[-\{\kappa (y - \mu_y) + \alpha (y - \mu_y)^2 + \gamma (y - \mu_y)^4\}] \qquad where \quad \gamma \geq 0$$

The substitution of *y* for *x* indicates the change of origin which produces the following relations between coefficients for the general and specific cases:

$$\kappa = \frac{8\beta_1\beta_4^2 - 4\beta_2\beta_3\beta_4 + \beta_3^3}{8\beta_4^2} \qquad \alpha = \frac{8\beta_2\beta_4 - 3\beta_3^2}{8\beta_4} \qquad \gamma = \beta_4$$

The symmetry restriction $\kappa = 0$ can only be satisfied if both β_3 and $\beta_1 = 0$. Given the symmetry restriction, the double well process further requires $-\alpha = \gamma = \sigma = 1$. Solving for the modes of Ψ[y] gives $\pm\sqrt{\{|\alpha| / (2\gamma)\}}$ which reduces to ± 1 for the double well process, as in Ait-Sahlia [32, Figure 6B, p.1385].

INSERT FIGURE 1 HERE

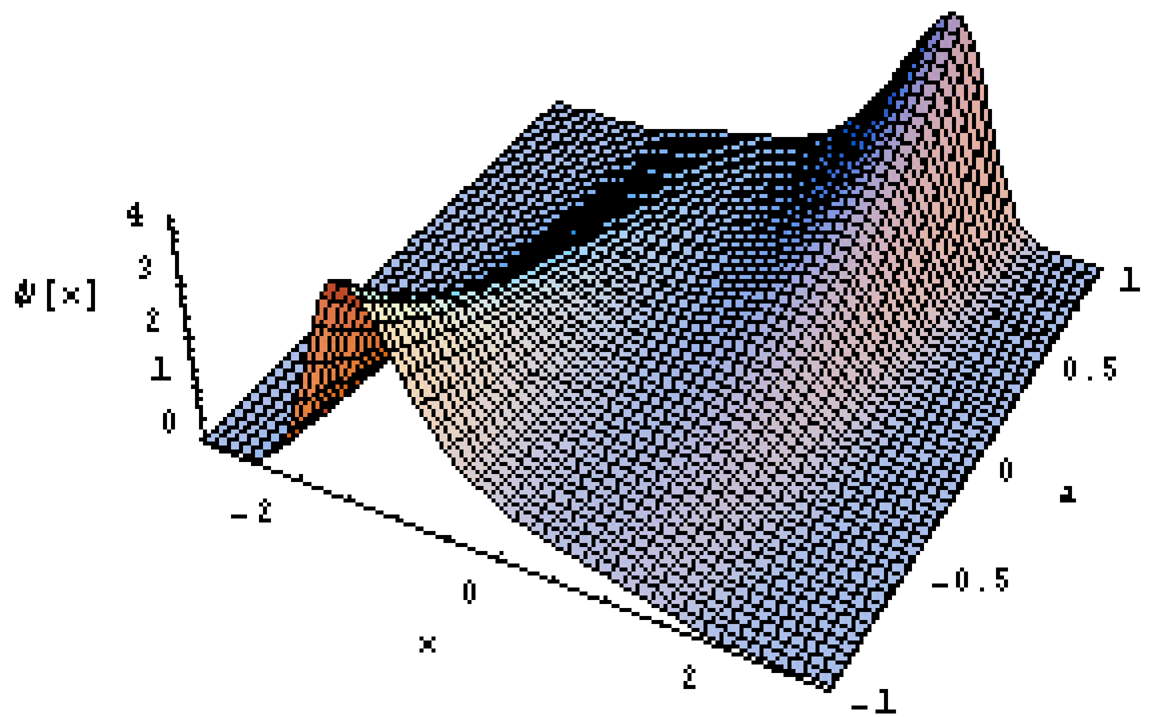
As illustrated in Figure 1, the selection of a_i in the stationary density $\Psi_i[x] = K_Q\, exp\{ -(.25\, x^4 - .5\, x^2 - a_i\, x) \}$ defines a family of general quartic exponential densities, where a_i is the selected value of κ for that specific density.[18] The coefficient restrictions on the parameters α and γ dictate that these values cannot be determined arbitrarily. For example, given that β_4 is set at .25, then for $a_i = 0$, it follows that $\alpha = \beta_2 = 0.5$. 'Slicing across' the surface in Figure 1 at $a_i = 0$ reveals a stationary distribution that is equal to the double well density. Continuing to slice across as a_i increases in size, the bimodal density becomes progressively more asymmetrically concentrated in positive x values. Though the location of the modes does not change, the amount of density between the modes and around the negative mode decreases. Similarly, as a_i decreases in size the bimodal density becomes more asymmetrically concentrated in positive x values. While the stationary density is bimodal over a_i ε {-1,1}, for $|a_i|$ large enough the density becomes so asymmetric that only a unimodal density appears. For the general quartic, asymmetry arises as the amount of the density surrounding each mode (the sub-density) changes with a_i. In this, the individual stationary sub-densities have a symmetric shape. To introduce asymmetry in the sub-densities, the reflecting boundaries at a and b that bound the state space for the regular S-L problem can be used to introduce positive asymmetry in the lower sub-density and negative asymmetry in the upper sub-density.

Following Chiarella et al. [19], the stochastic bifurcation process has a number of features which are consistent with the *ex ante* behavior of a securities market driven by a combination of chartists and fundamentalists. In particular, because the stationary distributions are multi-modal and depend on forward parameters – such as κ, α, γ and a_i in Figure 1 – that are not known on the decision date, the rational expectations models employed in mainstream economics are uninformative. What use is the forecast provided by $E[x(T)]$ when it is known that there are other $x(T)$ values that are more likely to occur? A mean estimate that is close to the bifurcation point would even be unstable. In a multi-modal world, complete fundamental uncertainty – where nothing is known about the evolution of economic variables – is replaced by uncertainty over unknown parameter values that can change due, say, to the collapse of a market convention. The associated difficulty of calculating a mean value forecast or other econometric estimates from past data is compounded by the presence of transients that originate from boundaries and initial conditions. For example, the presence of a recent structural break can be accounted for by appropriate selection of x_0. Of particular relevance to a comparison of econophysics theories with theories arising in mainstream financial economics is the fundamental dependence of investment decisions on x_0 which is not captured by the reversible ergodic processes employed in mainstream economics. The theoretical tools available in econophysics are able to demonstrate this fundamental dependence by exploiting properties of *ex ante* bifurcating ergodic processes to generate *ex post* sample paths that provide a better approximation to the sample paths of observed financial data.

VI. Conclusion

The concept of ergodicity is central to the distinction between mainstream economics and econophysics. This paper demonstrates that if economic observations are generated by bifurcating ergodic processes, then the calculation of time averages based on a sufficiently long enough set of past data can not be expected to provide a statistically reliable estimate of any *ex ante* time or space averages that will be observed in a sufficiently distant future calendar time. In other words, to deal with the problem of making statistical inferences from 'non-experimental' data, mainstream economics employs reversible ergodic transformations. The possibility of irreversible ergodic processes is not recognized or, it seems, intended. Significantly, a type of fundamental uncertainty is inherent in bifurcating processes as illustrated in the selection of a_i in Figure 1. A semantic connection can be established between the subjective uncertainty about encountering a future bifurcation point and, say, the possible collapse of an asset price bubble due to a change in Keynesian convention about market valuations. Examining the quartic exponential stationary distribution associated with a bifurcating ergodic process, it is apparent that this distribution nests the Gaussian distribution as a special case. In this sense, the models available in econophysics represent a stochastic generalization of the ergodic processes employed in mainstream economic theory.

Figure 1:* **Family of Stationary Densities for $\Psi_i[x] = K_Q\, exp\{ -(.25\, x^4 - .5\, x^2 - a_i\, x) \}$**



* Each of the continuous values for a signifies a different stationary density. For example, at $a = 0$ the density is the double well density which symmetric about zero and with modes at ±1.

Bibliography

[1] Schinckus, C. (2010), "Is econophysics a new discipline? The neopositivist argument", Physica A 389: 3814-3821.

[2] McCauley, J. (2004), Dynamics of markets : econophysics and finance, Cambridge, UK: Cambridge University Press.

[3] Encyclopedia of Mathematics (2002), "Ergodic Theory", New York: Springer Verlag / SpringerLink.

[4] von Neumann, J. (1932), "Proof of the Quasi-ergodic Hypothesis", Proceedings of the National Academy of Sciences 18: 255-63.

[5] Birkhoff, G. (1931), "Proof of the Ergodic Theorem", Proceedings of the National Academy of Sciences 17: 656-60.

[6] Doob, J. (1953), Stochastic Processes, New York: Wiley.

[7] Mackey, G. (1974), "Ergodic Theory and its Significance for Statistical Mechanics and Probability Theory", Advances in Mathematics 12: 178-268.

[8] Kolmogorov, A. (1933), Foundations of the Theory of Probability (in German), Springer: Berlin; English Edition, New York: Chelsea (1950).

[9] Arnold, L. (1998), Random Dynamical Systems, New York: Springer-Verlag.

[10] Mirowski, P. (1989), More Heat than Light, Economics as Social Physics: Physics as nature's economics, Cambridge, UK: Cambridge University Press.

[11] Weintraub, E.R. (2002), How Economics Became a Mathematical Science, Durham, NC: Duke University Press.

[12] Hicks, J. (1939), Value and Capital, Clarendon Press.

[13] Samuelson, P. (1947), Foundations of Economic Analysis, Boston, MA: Harvard University Press.

[14] Maxwell, J.C. (1867), "On the Dynamical Theory of Gases", Philosophical Transactions of the Royal Society 157: 49-88

[15] Weiner, N. (1939), "The Ergodic Theorem", Duke Mathematical Journal: 1-18.

[16] Halmos, P. (1949), "Measurable Transformations", Bulletin of the American Mathematical Society (Nov.): 1015-34.

[17] Davidson, P. (1991), "Is Probability Theory Relevant for Uncertainty? A Post Keynesian Perspective", Journal of Economic Perspectives 5:129-43.

[18] Dow, S. (2005), "Axioms and Babylonian thought: a reply", Journal of Post Keynesian Economics 27 (Spring): 385-91.

[19] Chiarella, C., X-Z He, D. Wang and M. Zheng (2008), "The stochastic bifurcation behaviour of speculative financial markets", Physica A 387: 3837-3846.

[20] Crauel, H., P. Imkeller, and M. Steinkamp (1999), "Bifurcations of One-Dimensional Stochastic Differential Equations", chapter 2 in H. Crauel and M. Gundlach (eds.), Stochastic Dynamics, New York: Springer Verlag.

[21] Feller, W. (1954), "Diffusion Processes in One Dimension", Transactions of the American Mathematical Society 77: 1-31.

[22] Wong, E. (1964), "The Construction of a Class of Stationary Markoff Processes", in Proceedings of Symposium on Applied Mathematics R. Bellman (ed.), Providence, RI: American Mathmatical Society.

[23] Gihman, I. and A. Skorohod (1979), The Theorv of Stochastic Processes (vol. II, III), New York: Springer-Verlag.

[24] Risken, H. (1989), The Fokker-Planck Equation, Methods of Solution and Applications, New York: Springer-Verlag.

[25] Linetsky, V. (2005), "On the Transitions Densities for Reflected Diffusions", Advances in Applied Probability 37: 435-60.

[26] Berg, P. and J. McGregor (1966), Elementary Partial Differential Equations, San Francisco: Holden-Day.

[27] Hille, E. (1969), Lectures on Ordinary Differential Equations, London: Addison-WeS-Ley.

[28] Birkhoff, G. and G. Rota (1989), Ordinary Differential Equations (4th ed.), New York: Wiley.

[29] Karlin, S. and H. Taylor (1981), A Second Course in Stochastic Processes, New York: Academic Press.

[30] Cox, D. and H. Miller (1965), The Theory of Stochastic Processes, London: Chapman and Hall.

[31] Hansen, L., J. Scheinkman and N. Touzi (1998), “Spectral Methods for Identifying Scalar Diffusions”, Journal of Econometrics 86: 1-32.

[32] Aït-Sahalia, Y. (1999), “Transition Densities for Interest Rates and Other Nonlinear Diffusions”, Journal of Finance 54: 1361-95.

[33] Veerstraeten, D. (2004), “The Conditional Probability Density Function for a Reflected Brownian Motion”, Computational Economics 24: 185-207.

[34] Cobb, L. (1978), “Stochastic Catastrophe Models and Multimodal Distributions”, Behavioral Science 23: 360-74.

[35] Cobb, L. (1981), “The Multimodal Exponential Families of Statistical Catastrophe Theory”, in C. Taillie, G. Patil and B. Baldessari (eds.), Statistical Distributions in Scientific Work (vol.4), Dordrecht, Holland: Reidel Press.

[36] Crauel, H. and F. Flandoli (1998), “Additive Noise Destroys a Pitchfork Bifurcation”, Journal of Dynamics and Differential Equations 10: 259-74.

[37] Fisher, R.A. (1921), “On the mathematical foundations of theoretical statistics”, Philosophical Transactions of the Royal Society 222A: 309-68.

[38] Cobb, L., P. Koppstein and N. Chen (1983), “Estimation and Moment Recursion Relations for Multimodal Distributions of the Exponential Family”, Journal of the American Statistical Association 78: 124-30.

[39] Elliott, J. (1955), “Eigenfunction expansions associated with singular differential operators”, Transactions of the American Mathematical Society 78: 406-25.

[40] Horsthemke, W. and R. Lefever (1984), Noise-Induced Transitions, New York: Springer-Verlag.

[41] Matz, A. (1978), “Maximum Likelihood Parameter Estimation for the Quartic Exponential Distribution”, Technometrics 20: 475-84.

[42] Jevons, W.S. (1877), The Principles of Science (2nd ed.), New York: Dover Press reprint (1958).

[43] Brush, S. (1983), Statistical Physics and the Atomic Theory of Matter: From Boyle and Newton to Landau and Onsager, Princeton, NJ: Princeton University Press.

[44] Ceragnani, C. (1998), Ludwig Boltzmann: The Man who Trusted Atoms, Oxford, UK: Oxford University Press.

[45] Brush, S. (1976), The Kind of Motion We Call Heat: A History of the Kinetic Theory of Gases in the 19th Century, Amsterdam: North Holland.

[46] Gallivotti, G. (1995), "Ergodicity, Ensembles, Irreversibility in Boltzmann and Beyond", Journal of Statistical Physics 78:1571-1589.

[47] Ehrenfest, P. (1911), *Begriffliche Grundlagen der statistischen Auffassung in der Mechanik*, Leipzig: B.G. Teubner.

[48] Reed and Simon

[49] Georgescu-Roegen, N. (1971), The entropy law and the economic process, Cambridge, MA: Harvard University Press.

[50] Davidson, P. (1988), "A technical definition of uncertainty and the long-run non-neutrality of money", Cambridge Journal of Economics 12: 329-37.

[51] Nicola, P. (1997), "Real Time Tatonnements and Firm Behaviour", Rivista Internazionale di Scienze Sociali. 105: 403-18.

[52] Ball, C. and A. Roma (1998), "Detecting Mean Reversion Within Reflecting Barriers: Applications to the European Exchange Rate Mechanism", Applied Mathematical Finance 5: 1-15.

[53] Poitras, G. and J. Heaney (2008), "Transition Density Decomposition and Generalized Pearson Distributions", available from www.sfu.ca/~poitras/Econth_3$.pdf.

[54] De Jong, F. (1994), "A univariate analysis of European monetary system exchange rates using a target zone model", Journal of Applied Econometrics 9: 31-45.

NOTES

1. In rational mechanics, once the initial positions of the particles of interest, e.g., molecules, are known, the mechanical model fully determines the future evolution of the system. This scientific and philosophical approach is often referred to as Laplacian determinism.

2. Boltzmann and Max Planck were vociferous opponents of energetics. The debate over energetics was part of a larger intellectual debate concerning determinism and reversibility. Jevons [42, p.738-9] reflects the entrenched determinist position of the marginalists: "We may safely accept as a satisfactory scientific hypothesis the doctrine so grandly put forth by Laplace, who asserted that a perfect knowledge of the universe, as it existed at any given moment, would give a perfect knowledge of what was to happen thenceforth and for ever after. Scientific inference is impossible,

unless we may regard the present as the outcome of what is past, and the cause of what is to come. To the view of perfect intelligence nothing is uncertain." What Boltzmann, Planck and others had observed in statistical physics was that, even though the behavior of one or two molecules can be completely determined, it is not possible to generalize these mechanics to the describe the macroscopic motion of molecules in large, complex systems, e.g., Brush [43, esp. ch.II].

3. This ignores developments in econometrics that commenced in the 1950's. These developments were concentrated on discrete time models that featured additive errors with strictly stationary distributions. In other words, probabilistic implications were incorporated by solving a deterministic model and then adding an error to the postulated theoretical relationship. This static probabilistic approach to modeling uncertainty has difficulty determining the non-linear dynamics that are captured by models associated with statistical mechanics.

4. As such, Boltzmann was part of the larger: "Second Scientific Revolution, associated with the theories of Darwin, Maxwell, Planck, Einstein, Heisenberg and Schrödinger, (which) substituted a world of process and chance whose ultimate philosophical meaning still remains obscure" (Brush [43, p.79]). This revolution superceded the: "First Scientific Revolution, dominated by the physical astronomy of Copernicus, Kepler, Galileo, and Newton, ... in which all changes are cyclic and all motions are in principle determined by causal laws." As such, the irreversibility and indeterminism of the Second Scientific Revolution replaces the reversibility and determinism of the First.

5. There are many interesting sources on these points which provide citations for the historical papers that are being discussed. Cercignani [44, p.146-50] discusses the role of Maxwell and Boltzmann in the development of the ergodic hypothesis. Maxwell [14] is identified as "perhaps the strongest statement in favour of the ergodic hypothesis". Brush [45] has a detailed account of the development of the ergodic hypothesis. Gallavotti [46] traces the etymology of "ergodic" to the 'ergode' in an 1884 paper by Boltzmann. More precisely, an ergode is shorthand for 'ergomonode' which is a 'monode with given energy' where a 'monode' can be either a single stationary distribution taken as an ensemble or a collection of such stationary distributions with some defined parameterization. The specific use is clear from the context. Boltzmann proved that an ergode is an equilibrium ensemble and, as such, provides a mechanical model consistent with the second law of thermodynamics. It is generally recognized that the modern usage of 'the ergodic hypothesis' originates with Ehrenfest [47].

6. The second law of thermodynamics is the universal law of increasing entropy – a measure of the randomness of molecular motion and the loss of energy to do work. First recognized in the early 19^{th} century, the second law maintains that the entropy of an isolated system, not in equilibrium, will necessarily tend to increase over time. Entropy approaches a maximum value at thermal equilibrium. A number of attempts have been made to apply the entropy of information to problems in economics, with mixed success. In addition to the second law, physics now recognizes the zeroth law of thermodynamics that "any system approaches an equilibrium state" (Reed and Simon [48, p.54]). This implications of this law for theories in economics was explored by Georgescu-Roegen [49].

7. Heterodox critiques are associated with views considered to originate from within economics. Such critiques are seen to be made by 'economists', e.g., Post Keynesian economists, institutional economists, radical political economists and so on. Because such critiques take motivation from the theories of mainstream economics, these critiques are distinct from econophysics. Following Schinckus [1, p.3818]: "Econophysicists have then allies within economics with whom they should become acquainted."

8. This interpretation of the microscopic collisions differs from Davidson [50, p.332]: "If there is only one actual economy, and we do not possess, never have possessed and conceptually never will possess an *ensemble* of economics worlds, then even a definition of probability distribution functions is questionable." In this context, points in the phase space at time t represent individual realizations of different macroscopic outcomes for the economic system at t. This interpretation of the ensembles is closer to Gibbs than Maxwell. Precisely how to interpret the ensembles in an economic context has not been closely examined. One exception is Nicola [51].

9. The connection of the reversibility and recurrence concepts used in this paper with the actual arguments made during the Boltzmann debates is somewhat tenuous. For example, the assumption that the diffusion process is regular deals with the version of the recurrence problem that concerned Boltzmann. The objective of introducing these concepts is pedagogy rather than historical accuracy.

10. The phenomenological approach is not without difficulties. For example, the restriction to Markov processes ignores the possibility of invariant measures that are not Markov. In addition, an important analytical construct in bifurcation theory, the Lyapunov exponent, can encounter difficulties with certain invariant Markov measures. Being primarily concerned with the properties of the stationary distribution, the phenomenological approach is not well suited to analysis of the dynamic paths around a bifurcation point. And so it goes.

11. A diffusion process is 'regular' if starting from any point in the state space I, any other point in I can be reached with positive probability (Karlin and Taylor [29, p.158]). This condition is distinct from other definitions of regular that will be introduced: 'regular boundary conditions' and 'regular S-L problem'.

12. The classification of boundary conditions is typically an important issue in the study of solutions to the forward equation. Important types of boundaries include: regular; exit; entrance; and natural. Also important in boundary classification are: the properties of attainable and unattainable; whether the boundary is attracting or non-attracting; and whether the boundary is reflecting or absorbing. In the present context, regular, attainable, reflecting boundaries are usually being considered, with a few specific extensions to other types of boundaries. In general, the specification of boundary conditions is essential is determining whether a given PDE is self-adjoint

13. Heuristically, if the ergodic process runs long enough, then the stationary distribution can be used to estimate the constant mean value. This definition of ergodic is appropriate for the one-dimensional diffusion cases considered in this paper. Other combinations of transformation, space and function will produce different requirements. Various theoretical results are available for the

case at hand. For example, the existence of an invariant Markov measure and exponential decay of the autocorrelation function are both assured.

14. For ease of notation it is assumed that $t_0 = 0$. In practice, solving (1) combined with (3)-(5) requires *a* and *b* to be specified. While *a* and *b* have ready interpretations in physical applications, e.g., the heat flow in an insulated bar, determining these values in economic applications can be more challenging. Some situations, such as the determination of the distribution of an exchange rate subject to control bands (e.g., Ball and Roma [52]), are relatively straight forward. Other situations, such as profit distributions with arbitrage boundaries or output distributions subject to production possibility frontiers, may require the basic S-L framework to be adapted to the specifics of the modeling situation.

15. The mathematics at this point are heuristic. More appropriate would be to observe that U^* is the special case where $U = \Psi[x]$. This would require discussion of how to specify the initial and boundary conditions to ensure that this is the solution to the forward equation.

16. A more detailed mathematical treatment can be found in the working paper Poitras and Heaney [53, see also de Jong [54].

17. In what follows, except where otherwise stated, it is assumed that $\sigma = 1$. Hence, the condition that K be a constant such that the density integrates to one incorporates the $\sigma = 1$ assumption. Allowing $\sigma \neq 1$ will scale either the value of K or the β's from that stated.

18. A number of simplifications were used to produce the 3D image in Figure 1: x has been centered about μ; and, $\sigma = K_Q = 1$. Changing these values will impact the specific size of the parameter values for a given x but will not change the general appearance of the density plots.