OPERATOR METHODS, ABELIAN PROCESSES AND DYNAMIC CONDITIONING

Claudio Albanese

Independent Consultant

15. December 2006

OPERATOR METHODS, ABELIAN PROCESSES AND DYNAMIC CONDITIONING

CLAUDIO ALBANESE

ABSTRACT. A mathematical framework for Continuous Time Finance based on operator algebraic methods offers a new direct and entirely constructive perspective on the field. It also leads to new numerical analysis techniques which can take advantage of the emerging massively parallel GPU architectures which are uniquely suited to execute large matrix manipulations.

This is partly a review paper as it covers and expands on the mathematical framework underlying a series of more applied articles. In addition, this article also presents a few key new theorems that make the treatment self-contained. Stochastic processes with continuous time and continuous space variables are defined constructively by establishing new convergence estimates for Markov chains on simplicial sequences. We emphasize high precision computability by numerical linear algebra methods as opposed to the ability of arriving to analytically closed form expressions in terms of special functions. Path dependent processes adapted to a given Markov filtration are associated to an operator algebra. If this algebra is commutative, the corresponding process is named Abelian, a concept which provides a far reaching extension of the notion of stochastic integral. We recover the classic Cameron-Dyson-Feynman-Girsanov-Ito-Kac-Martin theorem as a particular case of a broadly general block-diagonalization algorithm. This technique has many applications ranging from the problem of pricing cliquets to target-redemption-notes and volatility derivatives. Non-Abelian processes are also relevant and appear in several important applications to for instance snowballs and soft calls. We show that in these cases one can effectively use block-factorization algorithms. Finally, we discuss the method of dynamic conditioning that allows one to dynamically correlate over possibly even hundreds of processes in a numerically noiseless framework while preserving marginal distributions.

CONTENTS

1. Introduction

The goal of this paper is to attempt to consolidate and present a number of mathematical methods developed over several years by myself and collaborators while addressing concrete problems in derivative pricing theory. The results scattered across a number of papers which are collected here have been complemented with a rigorous *ab initio* treatment and a few key theorems which make the framework mathematically self-contained. This results in a quite comprehensive approach to the theory of Stochastic Processes and Mathematical Finance which is novel in that it is fully constructive and perhaps has applications beyond the realm of Financial Engineering.

There are several traditions of Constructive Mathematics. One attempts to re-derive classical results of real and functional analysis based on a restrictive constructivist logic according to which no mathematical object can be considered unless one specifies explicitly how to construct it, see (Bridges and Richman 1987) and (Bishop 1967). Along another tradition, Constructive Field Theory, see (Glimm and Jaffe 1987), aimed at establishing the existence of interacting quantum field theories by providing a constructive procedure for computing n-point functions and demonstrating that they satisfy a set of axioms. Measure theoretic probability and the related theory of stochastic processes, (Doob 1953), does not seem to be understandable constructively. The PDE approach in (Feller 1961) and the harmonic analysis approach in (Bochner 1955) are instead essentially constructive but do not delve into the theory of stochastic integrals and path dependent processes and into lattice discretization schemes.

The main motivation that guided this research is the creation of an engineering framework for exotic financial derivatives. Efficient computability on current hardware has been and remains throughout this article our key motivating concern. To this end, we work towards an algebraization of Probability Theory that reduces all calculations to matrix manipulations which can be performed efficiently and in particular to matrix multiplications. Similarly to the standard framework of algebraic topology, (Spanier 1966), we consider processes taking values in separable topological spaces and approximate continuous domains by means of simplicial sequences. To establish convergence in the continuous limit, we directly estimate convergence rates for probability transition kernels in the continuous space limit following an approach similar in spirit to Constructive Lattice Field Theory, see (Glimm and Jaffe 1987). Similarly to constructive field theory, sets of axioms on n-point functions are used to identify processes and renormalization group transformations are used to control the continuous limit. Following (Naimark 1959), the approach is grounded upon the algebraic theory of integration on locally compact Hausdorff separable topological spaces.

Calculations with stochastic processes are carried out using operator methods developed in Quantum Mechanics, (Landau and Lifshits 1977) and systematized in Mathematical Physics references such as (Reed and Simon 1980). In Finance, operator methods have been developed along two independent and non-overlapping streams of research, one by Ait-Sahalia, Hansen and Scheinkman who focused on econometric estimations in a series of papers reviewed in (Ait-Sahalia *et al.* 2005), see also (Ait-Sahalia 1996), (Hansen *et al.* 1998), (Hansen and Scheinkman 1995). The second stream of research is by the author and collaborators who instead worked on derivative pricing for path dependent and correlation derivatives, see (Albanese *et al.* 2005-2006b), (Albanese and Chen 2004a), (Albanese and Chen 2004b), (Albanese and Kusnetsov 2005),

(Albanese and Lawi 2004), (Albanese *et al.* 2006*a*), (Albanese and Trovato 2006), (Albanese and Trovato 2005), (Albanese and Vidler 2007), (Albanese and Jones 2007) and (Albanese and Osseiran 2007). In this paper, we attempt to systematize the mathematical framework of pricing theory in the operator formalism from our own viewpoint, reserving to future work the task of pursuing overlaps with the econometric literature.

The references quoted above are all relevant to our undertaking and provided motivations on many levels. However, in an effort to keep this writing self-contained, we are not going to assume any previous knowledge of the reader.

To ground the mathematical framework, we obtain sharp pointwise convergence estimates for probability kernels and its derivatives. More precisely, we show that probability kernels converge pointwise at rates of order $O(h^2)$, where h is the lattice spacing. The result applies to a large class of diffusion processes and extensions thereof including smooth space inhomogeneities, regime switching, finite activity jumps and some degree of time inhomogeneities. We also prove similar convergence results for the fast exponentiation method, our preferred numerical method for exponentiating Markov generators, showing that errors in this scheme are also of order $O(h^2)$, in the sense of pointwise convergence for the probability kernel. In the particular case of Brownian motion, we show that similar $O(h^2)$ pointwise error estimates apply also to derivatives of the probability kernels and that the power 2 in the $O(h^2)$ bounds is actually sharp.

The interest in convergence estimates was prompted by the desire of understanding the mechanisms behind the empirically observed smoothness and robustness in the calculation of price sensitivities with the methods in our applied papers. We also observed empirically that the fast exponentiation algorithm is stable under single precision floating point arithmetics. We find that key to a high precision numerical framework for sensitivities is handling the time coordinate either as continuous or very finely discretized. A sufficiently fine discretization is defined as one for which explicit differentiation schemes are stable and typically correspond to a hourly time scale in applications. Typically, weekly time steps would permit stable implicit differentiation schemes but would not allow for as much stability in the calculation of price sensitivities and the probability kernels we require to evaluate and manipulate. This motivates us to avoid implicit differentiation schemes on coarse time intervals.

Measure changes and time changes are defined constructively and a version of the Fundamental Theorem of Finance is re-obtained. The Cameron-Martin-Girsanov's theorem, see (Cameron and Martin 1949), and Ito's lemma, see (Ito 1949), are proved twice in different ways with operator methods. We also derive the Feynman-Kac formula, see (Feynman 1948) and (Kac 1948). One of the key results is an extension of the Feynman-Kac-Ito formula in three different directions. This formula concerns the characteristic function of a stochastic integral over a diffusion process. In our formalism, this formula becomes a block-diagonalization algorithm for large matrices associated to path-dependent processes. The extension we discuss (i) covers Markov processes more general than diffusions, (ii) allows for a class of path-dependent processes we name *Abelian* which extends the notion of stochastic integral and (iii) generalizes the harmonic analysis framework to include extensions of trigonometric Fourier transforms. The theory of Abelian processes finds numerous practical applications to path dependent options and is applicable to the great majority of path-dependent payoffs, from volatility swaps to cliquets, range accruals, lookback options, target redemption notes and more. We also give a version of Dyson's formula to accelerate the pricing of path-dependent options given by Abelian processes by means of a moment expansion. Non-Abelian processes are more difficult to handle but we single out a class admitting block-factorizations (as opposed to a block-diagonalizing transformation) and which are also amenable to numerical analysis by matrix algebra. Finally, we illustrate the method of dynamic conditioning that allows one to correlate possibly numerous processes by means of kernel manipulations while preserving marginals and not incurring into dimensional explosion.

The mathematical methods in this article are particularly efficient as they lend themselves to transparent hardware acceleration on the emerging multi-core GPU hardware platforms. These massively parallel architectures are based on low-cost technologies that have been developed

for the games and high definition markets and are uniquely suited to implement BLAS Level 3 routines such as matrix-matrix multiplications with high efficiency. See also (Goto and van de Geijn to appear) for a state of the art account on matrix-matrix multiplication software on CPUs.

The paper is organized as follows. In Section 2, we introduce the notion of simplicial sequence which is key to devising approximation schemes for continuous valued process by means of a sequence of Markov chains. In Section 3, we consider a general definition of path functional and in Section 4 we give a general description of a stochastic process by means of an n-point function. Markov Processes are introduced in Section 5, martingales and monotonic processes in Section 6. In Section 7, we derive the Fundamental Theorem of Arbitrage Free Pricing Theory. The classical results on weak convergence of Markov generators by Bernstein, Bochner, Kyntchine and Levy are re-obtained in Section 8. Time homogeneous Markov Processes, fast exponentiation and spectral methods are described in Section 9. In Section 10, we carry out a constructive analysis of Brownian motion and prove convergence estimates. In Section 11, we study the spectrum of diffusion generators. Sharp pointwise kernel estimates are extended to general diffusion processes in Section 12. In Section 13, we give estimates for the convergence rate of time discretisation schemes of the type we advocate for applications, i.e. based on fast exponentiation. Section 14 reviews the derivation of hypergeometric Brownian motion and particular cases such as the CEV model. In Section 15, we study stochastic integrals and obtain Ito's formula for diffusion processes and of Girsanov's theorem. Section 16 contains a derivation of the Feynman-Kac formula for bridges over general Markov processes. The general notion of Abelian process in continuous time is introduced in Section 17. In Section 18, we discuss the discrete time case. In this section, we introduce also the notion of non-resonant block-diagonalisation scheme which provides a numerically useful extension of Fourier analysis based on trigonometric functions. Dyson's formula and moment expansions are in Section 19, covering the uni-variate case, and Section 20, covering the multivariate case. These two sections include applications to exotic volatility derivatives. Block factorizations and applications to snowballs and soft calls are in Section 21. Dynamic conditioning and multi-factor correlation modeling is discussed in Section 22. Conclusions end the paper.

2. Measure Theory on Simplicial Sequences

Let $d > 0$ be an integer, consider the space $\mathbb{R}^d$ and the sequence of lattices $h_m \mathbb{Z}^d$ where $h_m = 2^{-m}$.

Definition 1. (Lattices.) *If $A \subset h_m \mathbb{Z}^d$, the convex hull of A in $\mathbb{R}^d$ is denoted by $\mathrm{Hull}(A)$. The interior of $A \subset h_m \mathbb{Z}^d$ is denoted with $\mathrm{Int}(A)$ and is defined as the set of all sites $x \in A$ contained in A along with each one its neighbors in $h_m \mathbb{Z}^d$ at distance h_m.*

Definition 2. (Simplicial Sequences.) *A bounded simplicial sequence is given by an integer $m_0 > 0$, and a sequence of subsets $A_m \in h_m \mathbb{Z}^d$ defined for all $m > m_0$ such that*

- *$\mathrm{Hull}(A_m) \subset \mathrm{Hull}(A_{m'})$ whenever $m < m'$.*
- *for all $m > 0$ and all internal points $x \in \mathrm{Int}(A_m)$ there is a $M > 0$ and an $\epsilon > 0$ such that, for all $m' > M$ and for all $y \in h_{m'} \mathbb{Z}^d$ with $d(x, y) < \epsilon$ we have that $y \in A_{m'}$.*

We follow a constructivist logic paradigm according to which in order to identify a set or a sequence of sets one has to explicitly state how to construct it, possibly with a recursive algorithm, and it must be possible to decide each step of the recursion in a finite number of logical steps.

Definition 3. (Lattice Functions.) *Let $A = (A_m)$, $m \geq m_0$ be a bounded simplicial sequence. A real valued simplicial function, denoted by $f : A \to \mathbb{R}$, is defined as a sequence of functions $f_m : A_m \to \mathbb{R}$ such that $f_{m'}(x) = f_m(x)$ for all $m' > m \geq m_0$ and all $x \in \mathrm{Int}(A_m)$. The function $f : A \to \mathbb{R}$ is said uniformly bounded if there is a constant $c > 0$ such that $|f_m(x)| < c$ for all $x \in A_m$. The function $f : A \to \mathbb{R}$ is uniformly continuous if it is uniformly bounded and for all $\epsilon > 0$ there is a $\delta > 0$ such that if $x, y \in h_m \mathbb{Z}$ and $d(x, y) < \delta$ we have that $|f(x) - f(y)| < \epsilon$.*

Definition 4. (Equivalence of Lattice Functions.) *Let $A = (A_m)$, $m \geq m_0$, be a bounded simplicial sequence and $f = (f_m)$ and $g = (g_m)$ are two uniformly continuous real valued functions on A. We say that these series provide equivalent representations of the same function in case, for all $m > 0$ and all all $x \in h_m \mathbb{Z}$ we have that $\lim_{m' \to \infty} |f_{m'}(x) - g_{m'}(x)| = 0$.*

Let $\mathcal{C}(A)$ be the set of all continuous functions on A endowed with the natural structure of C^* algebra given by the operations of sum, multiplication by a scalar and pointwise multiplication.

Definition 5. (Integrals.) *Let $A = (A_m)$, $m \geq m_0$, be a bounded simplicial sequence. An integral is given by a sequence $I = (I_m)$ where I_m is a linear functional on the linear space of functions $f_m : A_m \to \mathbb{R}$ such that*

- *$I_m(f_m) \geq 0$ whenever $f_m(x) \geq 0$ for all $x \in A_m$.*
- *The following limit exists for all continuous functions $f = (f_m) \in \mathcal{C}(A)$:*

$$(2.1) \qquad I(f) \equiv \lim_{m \to \infty} I_m(f_m)$$

- *The functionals $I_m(f_m)$ defined above satisfies the following bound:*

$$(2.2) \qquad I_m(f_m) \leq c ||f_m||_\infty$$

for some constant $c > 0$ and all functions f_m.

An integral is said to correspond to a probability measure if $I_m(1) = 1$.

Definition 6. (Equivalent Integrals.) *Let $A = (A_m)$, $m \geq m_0$, be a bounded simplicial sequence and let (I_m) and (J_m) be two integrals. We say that these series provide equivalent representations of the same integral I if $\lim_{m \to \infty} |I_m(f_m) - J_m(f_m)| = 0$ for all uniformly continuous functions (f_m).*

Given an integral I on the simplicial sequence A, for all $p \geq 1$, one defines the following semi-norm on the space of continuous functions $\mathcal{C}(A)$:

$$(2.3) \qquad ||f||_p = (I(|f|^p))^{1/p}.$$

Let $L^p(A; I)$ be the linear space obtained by completing $\mathcal{C}(A)$ with respect to the semi-norm $||f||_{I,p}$ and identifying equivalence classes of functions at zero distance. More precisely, $L^p(A; I)$ is the linear space of the Cauchy sequences (f_m) in $\mathcal{C}(A)$ with respect to the norm $||f||_{I,p}$ modulo the linear space of the Cauchy sequences converging to a limit of zero L^p-norm.

Definition 7. (Summable Functions.) *A function is called summable if it is in $L^1(A; I)$ and square summable if it is in $L^2(A; I)$.*

Theorem 1. (Monotonic Sequences.) *Let f_k be a non-decreasing sequence of summable functions and consider the function $f = \lim_{k \to \infty} f_k$ and the limit $\bar{I} = \lim_{k \to \infty} I(f_k)$. Then either $\bar{I} = \infty$ or f is summable and $\bar{I} = I(f)$.*

Proof. If $\bar{I} < \infty$, then whenever $k > j$ we have that $||f_k - f_j||_1 = I(f_k - f_j)$. Since the sequence $I(f_k)$ is uniformly increasing, it is Cauchy. Hence the sequence f_k is Cauchy in $L^1(A; I)$. Its limit f therefore belongs to $L^1(A; I)$ since this space is complete. $\square$

Definition 8. (Measurable Functions.) *A function is called measurable if it can be represented as the limit of a non-decreasing sequence of summable functions.*

Theorem 2. (Dominated Convergence.) *Let f_k be a sequence of measurable functions, suppose there is a summable function g such that $|f_k(x)| \leq |g(x)|$ and suppose also that the limit $f(x) = \lim_{k \to \infty} f_k(x)$ exists in a pointwise sense. Then f is a summable function and $I(f) = \lim_{k \to \infty} I(f_k)$.*

Proof. Consider the sequence u_k constructed iteratively so that $u_1 = \min(u_1, f)$ and $u_k = \max(u_{k-1}, \min(f_k, f))$. The sequence u_k is uniformly non-decreasing and converges to $f(x)$. Furthermore, $I(|u_k|) \leq I(|g|)$ for all k. Hence f is summable and $I(f) = \lim_{k \to \infty} I(f_k)$. $\square$

Let $\mathcal{V}$ be the linear space of all functions which can be expressed as the limit of a non-decreasing sequence of continuous functions $f(x) = \lim_{k \to \infty} f_k(x)$ such that the following norm is finite

$$(2.4) \qquad \qquad ||f||_\infty = \lim_{k \to \infty} ||f_k||_\infty.$$

$L^\infty(A; I)$ is defined as the completion of the linear space $\mathcal{V}$ with respect to the uniform norm.

Definition 9. (Essentially Bounded Functions.) *A function is called essentially bounded if it is in $L^\infty(A; I)$.*

Definition 10. (Absolute Continuity.) *Let I and J be two integrals on the simplicial sequence $A = (A_m)$. I is said absolutely continuous with respect to J if these two integrals admit simplicial representations $I = (I_m)$ and $J = (J_m)$ and there exists a $J-$summable function $g = (g_m)$ such that $I_m(f_m) = J_m(g_m f_m)$ for all $I-$summable functions $f = (f_m)$.*

3. Path Functionals

To introduce processes we need to specify the notion of measure on a set of paths. One could possibly introduce simplicial sequences in space-time but we refrain from doing so and consider instead the time coordinate as a continuous variable. In the time direction, we are only going to be using Riemann integrals for piecewise smooth functions, so that the underlying time discretisation one might possibly imagine is quite straightforward and keeping track of it would only lead to notational complexities.

Let us consider a bounded simplicial sequence $A = (A_m)$ and let $[T, T'] \subset \mathbb{R}$ be a fixed finite time interval. Let us consider the sequence $\Lambda = (\Lambda_m)$ where $\Lambda_m = A_m \times [T, T']$.

Given an integer $q \geq 0$, let $\mathcal{H}_m^q$ denote the set of all functions $y. : [T, T'] \to A_m$ which are constant on a family of q mutually disjoint sub-intervals of the form $[t_{k_i}, t_{k_{i+1}}), i = 0, ..q$ with $t_{k_0} = 0$ and $t_{k_{q+1}} = T'$, spanning the interval $[T, T']$. Notice that $(\mathcal{H}_m^q)$ for q fixed is itself a finite dimensional manifold with boundaries. In fact, a function $y \in \mathcal{H}_m^q$ is characterized by the ordered sequence $t_0 = T \leq t_1 \leq .... \leq t_{q+1} = T'$ of time points between which y_t is constant and a set of values $y_0, ...y_{q-1}$ such that, $y_s = y_i$ for all $s \in [t_i, t_{i+1})$ and for all $i = 0, ..q - 1$.

The path spaces $\mathcal{H}_m^q$ for $q = 0, 1, ...$ can be regarded as nested into each other $\mathcal{H}_0(T, T') \subset \mathcal{H}_1(T, T') \subset ....$ In fact, a path y_t in $\mathcal{H}_m^q$ is also a path in $\mathcal{H}_{m'}(T, T')$ if $m < m'$. Let $\mathcal{H}(T, T') = \cup_{m=0,1,..} \mathcal{H}_m^q$ be the union of all path spaces containing paths with a finite number of jumps.

Definition 11. (Function Algebras.) *Let us denote with $\mathbb{C}^q(\Lambda)$, the C^* algebra of all uniformly continuous sequences of simplicial functions $F = F_m : \mathcal{H}_m^q \to \mathbb{R}$ endowed with the operations of sum, multiplication and with respect to the uniform norm defined as follows:*

$$(3.1) \qquad \qquad ||F||_\infty = \lim_{m \to \infty} \sup_{y. \in \mathcal{H}_m^q} |F(y.)|.$$

Definition 12. (Path Functionals.) *A path functional is a sequence $F = (F^q), q = 0, 1, ...$ of continuous simplicial functions $F^q = (F_m^q) \in C^q(\Lambda)$ satisfying the following mutual compatibility condition:*

$$(3.2) \qquad \qquad F_m^{q'}(y.) = F_m^q(y.), \quad \text{for all} \quad y. \in \mathcal{H}_m^q \quad \text{and all} \quad q' > q.$$

Definition 13. (Non-anticipatory Path Functionals.) *Let $F(y., t) = (F^q(y., t)) \; t \in [T, T']$ be a one-parameter family of path functionals. One says that this is a non anticipatory path functional if*

$$(3.3) \qquad \qquad F_m^q(y., t) = F_m^q(y'., t)$$

whenever $y_s = y'_s$ for all $s \leq t$.

Intuitively, non-anticipatory path functionals are indifferent to information about the realization of the path $y.$ in the argument at future times $s > t$. An elementary example of non-anticipatory path functional is given by a function of two arguments

$$(3.4) \qquad \qquad F_m^q(y., t) = F_0(y_t, t)$$

where $F_0 : A_\infty \times [T, T'] \to \mathbb{R}$ is a piecewise smooth one-parameter family of functions. Applications typically call for functions $F_0(y, t)$ which are piecewise smooth in t for each $y \in \Lambda$ with possibly a discrete set of jump discontinuities. We follow the usual convention according to which if jumps occur, then the discontinuity is of *cadlag* type, i.e. right continuous and with a left limit.

A less elementary example of non-anticipatory path functional one often encounters is given by integrals of the form:

$$(3.5) \qquad F_m^q(y., t) \equiv \int_T^t ds_1 ... \int_{s_{k-1}}^t ds_n a(t; s_1, ...s_k) F_1(y_{s_1}, s_1)...F_k(y_{s_k}, s_k).$$

where the $F_j : A_\infty \times [T, T'] \to \mathbb{R}$, $j = 1, ...k$, are one-parameter families of lattice functions and also $a(t; s_0, ...s_k)$ is a function of the time coordinates. Applications typically call for functions $F_j(y, t)$ which are piecewise smooth in t for each $y \in \Lambda$ with possibly a discrete set of jump discontinuities. Also in this case, we follow the convention according to which at the points of jump discontinuity the function is *cadlag*. The same regularity assumptions will be postulated for the functions $a(t; s_0, ...s_k)$ with respect to each of its arguments. Although the regularity assumption for these path functionals is sufficient for applications, they are not strictly needed and can be relaxed by taking limits.

4. n-POINT FUNCTIONS

Definition 14. (Filtered Probability Spaces.) *Consider the C^* algebra $\mathcal{P}$ generated by the functionals of form (3.4) and (3.5) by taking linear combinations of finite products and completing the resulting normed space. A filtered probability space upon the lattice Λ is defined as a bounded linear functional on the C^*-algebra $\mathcal{P}$.*

A constructive definition of a stochastic process can be given with various degrees of generality. We don't aim here at the utmost generality but instead at pedagogical simplicity and at ultimately explaining how to frame the theory of Markov processes. As a step toward this goal, we introduce here a family of time ordered $n-$point functions

$$(4.1) \qquad \Phi_m(y_1, t_1; ...; y_n, t_n | y_0)$$

defined for $y_1, ...y_n \in A_m$ and $t_1, ...t_n \in [T, T']$ which is assumed to be piecewise differentiable in the time coordinates. In addition, we assume that the following properties hold for all $T < t_1 \le ... \le t_n \le T'$:

(SP1) $\quad \Phi(y_1, t_1; ...; y_n, t_n | y_0) \ge 0 \quad \forall y_0, y_1, ..y_n \in \Lambda,$

(SP2) $\quad \displaystyle\sum_{y_1, ...y_n \in \Lambda} \Phi(y_1, t_1; ...; y_n, t_n | y_0) = 1 \quad \forall y_0 \in \Lambda,$

(SP3) $\quad \Phi(y_1, t_1; ...; y_i, t_i; y_{i+1}, t_{i+1}; ...; y_n, t_n | y_0) = 0$

$\qquad\qquad\qquad$ if $t_i = t_{i+1}$ for some $i = 1...n-1$ and if $y_i \ne y_{i+1}$

(SP4) $\quad \displaystyle\sum_{y_i} \Phi(y_1, t_1; ...; y_{i-1}, t_{i-1}; y_i, t_i; y_{i+1}, t_{i+1}; ...; y_n, t_n | y_0) =$

$\qquad\qquad \Phi(y_1, t_1; ...; y_{i-1}, t_{i-1}; y_{i+1}, t_{i+1}; ...; y_n, t_n | y_0) \quad \forall y_1, ..y_n \in \Lambda.$

Notice that for each fixed starting point y_0 and fixed sequence $T < t_1 \le ... \le t_n \le T'$, the function $\Phi_m(y_1, t_1; ...; y_n, t_n | y_0)$ is a probability distribution function in the arguments $y_1, ...y_n$. This is interpreted as the probability distribution density for paths starting from the site y_0 at time T and achieving the values $y_1, ...y_n$ at times $t_1, ...t_n$.

Definition 15. (Stochastic Processes.) *An adapted (stochastic) process is given by a non-anticipatory path functional $F \in \mathcal{P}$ and a measure on P defined by a family of $n-$point functions $(\Phi)_{n=1,2...}$*

Given an adapted process of the form $F_t \equiv F(y_t, t) \in \mathcal{P}$ and given a $t > T$, the expectation subject to the initial condition $y_T = y_0$ of the value attained by the process F_t at time t is given by

$$(4.2) \qquad E\big[F_t | y_T = y_0\big] = \sum_{y_1 \in \Lambda} \Phi(y_1, t | y_0) F(y_1, t).$$

For an adapted process of the form (3.5) instead, the conditional expectation is given by
(4.3)
$$E\big[F_t | y_T = y_0\big] = \sum_{y_1, ..y_n} \int_T^t ds_1 ... \int_{s_{n-1}}^t ds_n \Phi(y_1, s_1; ...; y_n, s_n | y_0) a(t; s_1, ...s_n) F_1(y_1, s_1)...F_n(y_n, s_n).$$

Also higher moments can be computed. To keep expressions simple, consider a path functional of the form

$$(4.4) \qquad F_t \equiv \int_T^t ds_1 a(t; s_1) F_1(y_{s_1}, s_1).$$

The variance of this process at a future point in time conditional to the starting point at time T is given by
(4.5)
$$E\big[F_t^2 | y_T = y_0\big] = 2 \sum_{y_1, y_2} \int_T^t ds_1 \int_{s_1}^t ds_2 \Phi(y_1, s_1; y_2, s_2 | y_0) a(t; s_1) a(t; s_2) F_1(y_1, s_1) F_1(y_2, s_2).$$

Notice that the factor 2 compensates for the time-ordering needed to recast the expression in such a way that we can apply to it a 2-point function. This expression and its generalizations are used extensively in the applications to path-dependent options discussed below.

More generally, consider two path functionals F_t and $G_t \in L^1(\mathcal{H}(T, T'), \mu)$, where F_t is defined as in (3.5) and

$$(4.6) \qquad G_t \equiv \int_T^t ds_1 ... \int_{s_{q-1}}^t ds_q b(t; s_1, ...s_q) G_1(y_{s_1}, s_1)...G_q(y_{s_q}, s_q).$$

Then we can compute a mixed moment as follows:

$$E\big[F_t G_t | y_T = y_0\big] = \sum_{y_1, ..y_{n+q}} \int_T^t ds_1 ... \int_{s_{n+q-1}}^t ds_{n+q} \quad \Phi(y_1, s_1; ....; y_{n+q}, s_{n+q})$$

$$\sum_\pi a(t; s_{\pi(1)}, ...s_{\pi(n)}) F_1(y_{\pi(1)}, s_{\pi(1)})...F_n(y_{\pi(n)}, s_{\pi(n)})$$

$$b(t; s_{\pi(n+1)}, ...s_{\pi(n+q)}) G_1(y_{\pi(n+1)}, s_{\pi(n+1)})...G_q(y_{\pi(n+q)}, s_{\pi(n+q)}),$$

where the sum ranges over all the permutations π of the time ordered sequence $s_1 \leq ... \leq s_{n+q}$ such that $s_{\pi(1)} \leq ... \leq s_{\pi(n)}$ and $s_{\pi(n+1)} \leq ... \leq s_{\pi(n+q)}$. Higher conditional moments of any path functional of the form above can be evaluated in a similar way.

n-point functions $\Phi(y_1, s_1; ...; y_n, s_n | y_0)$ are conditioned to the starting point y_0. It is straightforward to define variations of these n-point functions which reflect conditioning on an initial stretch of a path y_s for $s \in [T, t)$, where $t > T$. Conditioning to a past history is equivalent to restricting the integral over each manifold $\mathcal{H}(T, T')$ to a sub-manifold which is part of its boundary. In general, this results in rather clumsy expressions which are difficult to compute. In the next section we specialize to the Markovian case for the underlying lattice process and in this case there are no memory effects and conditioning is more straightforward.

Definition 16. (Radon-Nykodim Derivative.) *Consider two sequences of n-point functions on the simplicial sequence A_m $\Phi_m^1(y_1, t_1; ...; y_n, t_n | y_0)$ and $\Phi_m^2(y_1, t_1; ...; y_n, t_n | y_0)$, where $m = 1, 2, ...$. The Radon-Nykodim derivative of Φ^1 with respect to Φ^2 is given by the path functional defined as follows:*

$$(4.7) \qquad \rho(y_.) = \lim_{n \to \infty} \frac{\Phi_m^1(y_{t_1}, t_1; ...; y_{t_n}, t_n | y_0)}{\Phi_m^2(y_{t_1}, t_1; ...; y_{t_n}, t_n | y_0)}$$

where $t_i = T + \frac{i}{n}(T' - T)$. Notice that the Radon-Nykodim derivative might possibly be infinite on some paths.

The measure in path space given by (Φ^1_m) is said to be absolutely continuous with respect to (Φ^2_m) if the Radon-Nykodim derivative is finite and summable.

Finally, a word on stopping times.

Definition 17. (Stopping Times.) *A stopping time is an adapted process $\tau_t = \tau(y., t)$ which can take only two values, by convention 0 and 1.*

Stopping times are often used in conjunction with other adapted processes F_t to construct stopped versions of it. If $F_t = F(y., t)$ is an adapted process and $R_t = R(y., t)$ is a second process, then a stopped version of F_t corresponds to the adapted process of function $\bar{F}(y., t)$, where $\bar{F}(y., t) = F(y., t)$ if $\tau_t = 0$ and $\bar{F}(y., t) = R(y., t)$ if $\tau_t = 1$.

5. Markov Processes

Definition 18. (Markov Propagator.) *A Markov propagator on the simplicial sequence $A_m, m \geq m_0$ is defined as a sequence of functions $U_m(y_1, t_1; y_2, t_2)$ where $y_1, y_2 \in A_m$ and $T \leq t_1 \leq t_2 \leq T'$ satisfying the following Chapman-Kolmogorov axioms:*

(CK1) $U_m(y_1, t_1; y_2, t_2) \geq 0 \quad \forall y_1, y_2 \in \Lambda, \quad \forall t_1 \leq t_2 \in \mathbb{R},$

(CK2) $U_m(y_1, t_1; y_2, t_1) = \delta_{y_1 y_2} \quad \forall y_1, y_2 \in \Lambda, \quad \forall t_1 \in \mathbb{R},$

(CK3) $\displaystyle\sum_{y_2 \in \Lambda} U_m(y_1, t_1; y_2, t_2) U_m(y_2, t_2; y_3, t_3) = U_m(y_1, t_1; y_3, t_3)$

$$\forall y_1, y_3 \in \Lambda, \quad \forall t_1 \leq t_2 \leq t_3 \in \mathbb{R}.$$

Given a Markov propagator $U_m(y_1, t_1; y_2, t_2)$, one can define a sequence of $n-$point function having all the necessary properties by setting

$$(5.1) \qquad \Phi_m(y_1, t_1; ...; y_n, t_n | y_0) = \prod_{j=1..n} U_m(y_{j-1}, t_{j-1}; y_j, t_j).$$

where $t_0 = T$.

Definition 19. (Markov Process.) *A filtered probability space generated by a Markov propagator is called Markovian and the corresponding stochastic process is called Markov process.*

Definition 20. (Markov Generator.) *If the matrix elements $U_m(y_1, t_1; y_2, t_2)$ are differentiable functions of the time parameter t_2 in a right neighborhood of t_1, then one defines the Markov generator at time t_1 as the following right derivative:*

$$(5.2) \qquad \mathcal{L}_m(y_1, y_2; t_1) = \lim_{t_2 \downarrow t_1} \frac{d}{dt_2} U_m(y_1, t_1; y_2, t_2)$$

Proposition 1. *If $\mathcal{L}_m(y_1, y_2; t_1)$ is a Markov generator, then for all pairs $y_1, y_2 \in A_m$ the following two properties hold:*

$$(5.3) \qquad (MG1) \quad \mathcal{L}_m(y_1, y_2; t) \geq 0 \quad \text{if } y_1 \neq y_2,$$

$$(5.4) \qquad (MG2) \quad \mathcal{L}_m(y_1, y_1; t) = - \sum_{y_2 \neq y_1} \mathcal{L}_m(y_1, y_2; t).$$

Viceversa, if $\mathcal{L}_m(y_1, y_2; t)$ is a differentiable one-parameter family of matrices satisfying conditions (A) and (B) above, then the differential equation (5.2) admits one and only one solution satisfying the initial condition $U_m(y_1, t; y_2, t) = \delta_{y_1 y_2}$.

The propagator $U_m(y_1, t_1; y_2, t_2)$ defined by the differential equation in (5.2) can be represented by means of a so-called *path-exponential* defined as follows. Let $N > 0$ be an integer and let us consider the product

$$(5.5) \qquad U^N_m(y_1, t_1; y_2, t_2) = \left(1 + \delta t_N \mathcal{L}_m(t_1)\right) \cdot .... \cdot \left(1 + \delta t_N \mathcal{L}_m(t_N)\right)$$

where $\delta t_N = \frac{t_2 - t_1}{N}$. If N is so large that

$$(5.6) \qquad (\delta t)_N < \min_{y \in \Lambda, t \in [t_1, t_2]} \mathcal{L}_m(y, y; t)^{-1},$$

then the operator product $U_m^N(y_1, t_1; y_2, t_2)$ in equation (5.5) is a probability kernel. Passing to the limit $N \to \infty$, we find

$$(5.7) \qquad \lim_{N \to \infty} U_m^N(y_1, t_1; y_2, t_2) = U_m(y_1, t_1; y_2, t_2).$$

By expanding the product in (5.5) and passing to the limit, we arrive at the following:

Theorem 3. (Dyson expansion.) *The probability kernel can be represented as the following convergent series:*
(5.8)

$$U_m(y_1, t_1; y_2, t_2) = \left(1 + \sum_{n=1}^{\infty} \int_{t_1}^{t_2} ds_1 \int_{s_1}^{t_2} ds_2 ... \int_{s_{n-1}}^{t_2} ds_n \mathcal{L}_m(s_1) \mathcal{L}_m(s_2) .... \mathcal{L}_m(s_n) \right) (y_1, y_2).$$

By differentiating with respect to the two time coordinates in the Dyson expansion, we also find the following two equations:

Theorem 4. (Forward and backward equations.) *The probability kernel satisfies the backward equation*

$$(5.9) \qquad \frac{\partial}{\partial t_1} U_m(t_1; t_2) + \mathcal{L}_m(t_1) U_m(t_1; t_2) = 0$$

as well as the forward equation

$$(5.10) \qquad \frac{\partial}{\partial t_2} U_m(t_1; t_2) = U_m(t_1; t_2) \mathcal{L}_m(t_2).$$

A handy notation for this expansion is given by the following:

Definition 21. (Path-ordered Exponential.) *The equation (5.8) is written as a path ordered exponential*

$$(5.11) \qquad U_m(y_1, t_1; y_2, t_2) = \text{Pexp}\left(\int_{t_1}^{t_2} \mathcal{L}_m(s) ds \right)(y1, y2)$$

where the operator P formally acts as follows:

$$(5.12) \qquad \text{P}\left(\int_{t_1}^{t_2} \mathcal{L}_m(s) ds \right)^n = n! \int_{t_1}^{t_2} ds_1 \int_{s_1}^{t_2} ds_2 ... \int_{s_{n-1}}^{t_2} ds_n \mathcal{L}_m(s_1) \mathcal{L}_m(s_2) .... \mathcal{L}_m(s_n).$$

Using the Dyson expansion for the path-ordered exponential, one finds a path-integral representation of the probability kernel. Let us set the following definition:

Definition 22. (Symbolic Path.) *A symbolic path $\gamma = \{\gamma_0, \gamma_1, \gamma_2, ...\}$ is an infinite sequence of sites in $h_m \mathbb{Z}$ such that $\gamma_j \neq \gamma_{j-1}$ for all $j = 1, ...$ Let Γ_m be the set of all symbolic paths in $h_m \mathbb{Z}$.*

Theorem 5. (Path-Integral Representation.) *The propagator admits the following representation:*
(5.13)

$$U_m(x, T; y, T') = \sum_{q=1}^{\infty} \sum_{\gamma \in \Gamma_m : \gamma_0 = x, \gamma_q = y} \int_T^{T'} ds_1 \int_{s_1}^{T'} ds_2 ... \int_{s_{q-1}}^{T'} ds_q$$

$$(5.14) \qquad \exp\left(\int_0^{s_1} \mathcal{L}_m(\gamma_0, \gamma_0; v_0) dv_0 \right) \prod_{j=1}^{q} \mathcal{L}_m(\gamma_{j-1}, \gamma_j; s_j) \exp\left(\int_{s_j}^{s_{j+1}} \mathcal{L}_m(\gamma_j, \gamma_j; v_j) dv_j \right)$$

where $t_{q+1} = T$.

Notice that the total mass of the sector of path space $\mathcal{H}_m^q(y_0, T; T')$ consisting of lattice paths originating from $y_0 \in \Lambda$ at time T and attaining at most q different values by time T' is given by

$$\mu(\mathcal{H}_m^q(y_0, T; T')) = \sum_{\gamma \in \Gamma_m : \gamma_0 = x} \int_T^{T'} ds_1 \int_{s_1}^{T'} ds_2 ... \int_{s_{q-1}}^{T'} ds_q \exp\left(\int_T^{s_1} \mathcal{L}_m(\gamma_0, \gamma_0; v_0) dv_0\right)$$

$$\prod_{j=1}^q \mathcal{L}_m(\gamma_{j-1}, \gamma_j; s_j) \exp\left(\int_{s_j}^{s_{j+1}} \mathcal{L}_m(\gamma_j, \gamma_j; v_j) dv_j\right).$$

where $s_{m+1} = T'$. If $\lambda_m = \max_{x \in A_m} |\mathcal{L}_m(x, x)|$ and $c = \sup_{t \in [T, T']} ||\mathcal{L}_m(t)||$ is the operator norm of the matrix $\mathcal{L}_m(t)$, then we have that

$$(5.15) \qquad \mu\left[\mathcal{H}_m^q(y_0, T; T')\right] \leq \frac{c^q (T' - T)^q}{q!} e^{-\lambda_m (T' - T)}$$

This expression reaches a maximum as a function of q at $q \approx c(T' - T)$ and then declines at super-exponential rate. Hence, by discretizing the space coordinate, we ensure that the probability mass of the set of paths with q changes decreases faster than any exponential in the limit as $q \to \infty$. This convergence is however not uniform as $m \to \infty$ and $h_m \to 0$.

Definition 23. (Inverse lattice.) *Let us consider the following lattice:*

$$(5.16) \qquad B_m = \left\{-\frac{2^{m-1}\pi}{L} + \frac{k\pi}{L}, k = 0, .. 2^m - 1\right\}$$

also called Brillouin zone or inverse lattice with respect to A_m.

Definition 24. (Pseudo-differential Symbols.) *The symbol of a Markov generator is defined as follows:*

$$(5.17) \qquad \hat{\mathcal{L}}_m(x, p; t) = \sum_{y \in A_m} \mathcal{L}_m(x, y) e^{ip(y-x)}$$

where $p \in B_m$.

Two particularly important special examples of Markov processes are given by monotonic processes and diffusions.

Definition 25. (Monotonic process.) *A Markov process of generator $\mathcal{L}_m(x, y)$ on the simplicial sequence A_m is said monotonic non-decreasing if*

$$(5.18) \qquad \mathcal{L}_m(x, y) = 0 \quad \text{whenever} \quad y < x.$$

and monotonic non-increasing if

$$(5.19) \qquad \mathcal{L}_m(x, y) = 0 \quad \text{whenever} \quad y > x.$$

Definition 26. (Diffusion Process.) *A diffusion process is a Markov process with generator of the form*

$$(5.20) \qquad \mathcal{L}_m(x, y; t) = \mu_m(x; t) \nabla_{h_m}(x, y) + \frac{\sigma_m(x; t)^2}{2} \Delta_{h_m}(x, y)$$

where

$$(5.21) \qquad \nabla_h(x, y) = \frac{\delta_{y, x+h} - \delta_{y, x-h}}{2h} \quad \text{and} \quad \Delta_h(x, y) = \frac{\delta_{y, x+h} + \delta_{y, x-h} - 2\delta_{y, x}}{h^2}.$$

and $\mu(x; t) = (\mu_m(x; t))$ and $\sigma(x; t) = (\sigma_m(x; t))$ are two simplicial functions which, for simplicity, we assume smooth in both arguments.

Proposition 2. (i) *The symbol of a diffusion process is given by*

$$(5.22) \qquad \hat{\mathcal{L}}_m(x, p; t) = \mu_m(x; t) \frac{\sin ph}{ih} + \frac{\sigma_m(x; t)^2}{2} \frac{\cos ph - 1}{h^2}.$$

(ii) *The path integral representation for a diffusion process has the form*

$$(5.23) \qquad U_m(x, T; y, T') = \sum_{q=1}^{\infty} 2^{-q} \sum_{\substack{\gamma \in \Gamma_m : \gamma_0 = x, \gamma_q = y \\ |\gamma_j - \gamma_{j-1}| = 1 \ \forall j \geq 1}} W_m(\gamma, q; T, T')$$

where

$$W_m(\gamma, q; T, T') = \int_T^{T'} ds_1 \int_{s_1}^{T'} ds_2 ... \int_{s_{q-1}}^{T'} ds_q$$

$$(5.24) \qquad e^{\int_{s_q}^{T'} \mathcal{L}_m(y, y, v_q) dv_q} \prod_{j=0}^{q-1} \left(e^{\int_{s_j}^{s_{j+1}} \mathcal{L}_m(\gamma_j, \gamma_j, v_j) dv_j} 2 \mathcal{L}_m(\gamma_j, \gamma_{j+1}) \right)$$

and $s_0 = T$.

6. Martingales and Monotonic Processes

In this section we introduce the notion of piecewise smooth Markov process which covers a large family of models useful for applications. In this context, we define martingales and monotonic Markov processes.

Definition 27. (Piecewise Smooth Markov Processes.) *Consider the time interval $[T, T']$, the simplicial sequence A_m, $m \geq m_0$ and a finite number of time points $T = t_0 < t_1 < .. < t_n = T'$. A piecewise smooth Markov process is given by a family of Markov generators $\mathcal{L}_m^i(y_1, y_2; t)$ defined on each half open time interval $[t_i, t_{i+1})$, where $i = 0, 1, ...n - 1$. In correspondence to each time point $t_i, i = 1, 2, ...n$, one also defines a mapping operator $U_m^i(x, y)$ such that*

$$(6.1) \qquad \qquad (MA1) \quad U_m^i(y_1, y_2) \geq 0$$

$$(6.2) \qquad \qquad (MA2) \quad \sum_{y_2} U_m^i(y_1, y_2) = 1 \ \ \forall y_1 \in \Lambda.$$

The Markov propagator for any pair of time points $t_i \leq s < s' < t_{i+1}$ is defined as follows:

$$(6.3) \qquad \qquad U_m(y_1, s; y_2, s') = \text{Pexp} \left(\int_s^{s'} \mathcal{L}_m^i(v) dv \right) (y_1, y_2).$$

Moreover, if $s' = t_{i+1}$, then

$$(6.4) \qquad U_m(y_1, s; y_2, t_{i+1}) = \sum_{y_3 \in \Lambda} \text{Pexp} \left(\int_s^{t_{i+1}} \mathcal{L}_m^i(v) dv \right) (y_1, y_3) U_m^i(y_3, y_2).$$

More general Markov propagators are obtained by taking products of the ones above.

Definition 28. (Attainable Sets.) *Let U_m be a piecewise smooth Markov process, $m \geq m_0$, $y \in A_m$ and $t \in [T, T']$. The attainable set $\mathcal{D}_m(U, t, y) \subset A_m$ is defined as follows: if $t \in (t_i, t_{i+1})$ for some $i = 0, ..n - 1$, then $\mathcal{D}_m(U, t, y)$ is the set of $\bar{y} \in \Lambda$ such that $\mathcal{L}_i(y, \bar{y}; t) > 0$. If instead $t = t_i$ for some $i = 1, ..n - 1$, then $\mathcal{D}_m(U, t_i, y)$ is defined as the set of the $\bar{y} \in \Lambda$ such that $U_i(y, \bar{y}; t) > 0$.*

Definition 29. (Equivalent Markov Processes.) *Two piecewise smooth Markov propagators U and U' are called equivalent if their attainable sets $\mathcal{D}_m(U, t, y)$ and $\mathcal{D}_m(U', t, y)$ are equal for all $t \in [T, T']$ and all $y \in \Lambda$. If $\mathcal{D}_m(U, t, y)$ is a subset of $\mathcal{D}_m(U', t, y)$ for all $t \in [T, T']$ and all $y \in \Lambda$, then one says that U_m is absolutely continuous with respect to U_m'.*

Definition 30. (Measure Changes.) *Let $U_m, m \geq m_0$ be a family of piecewise smooth Markov propagators. A measure change is characterized by a family of positive, non-zero functions $G_m^{yt}(y') \geq 0$ indexed by $y \in A_m$ and $t \in [T, T']$, which is strictly positive for all $y' \in \mathcal{D}_m(U, y, t)$*

and zero otherwise. A measure change function defines a transformation of a Markov generator into an equivalent one according to the following formula:

$$(6.5) \qquad \mathcal{L}'(y, y'; t) = \frac{1}{G_m^{yt}(y)} \mathcal{L}(y, y'; t) G_m^{yt}(y') - \frac{1}{G_m^{yt}(y)} (\mathcal{L}(t) G_m^{yt})(y) \delta_{yy'}.$$

Notice that the specification of the function $G_m^{yt}(y')$ at the point $y' = y$ is immaterial in the sense that it does not affect the measure change transformation.

Definition 31. (Time Changes.) *A measure change is called time change if there is a function $\phi(y, t)$ such that $G_m^{yt}(y') = \phi(y, t)$ for all $y \in A_m$, $y' \in \mathcal{D}_m(U, y, t)$ and $t \in [T, T']$. The time change is called state independent if $\phi(y, t) \equiv \phi(t)$ is a function of the time coordinate only.*

Theorem 6. (Deterministic Time Changes.) *If $\phi(t)$ defines a state independent time change corresponding to the measure change function $G_m^{yt}(y') = \phi(t)$, then*

$$(6.6) \qquad U'_m(y, t; y', t') = U_m(y, \lambda(t); y', \lambda(t'))$$

where

$$(6.7) \qquad \lambda(t) = \int_T^t \phi(s) ds.$$

The following is a particularly interesting special case of measure change:

Definition 32. (Numeraire Changes.) *Consider a smooth (as opposed to just piece-wise smooth) Markov process of generator $\mathcal{L}_m(y; t)$. Let $g_m(y'; t)$ be a function satisfying the equation*

$$(6.8) \qquad \frac{\partial g_m(y; t)}{\partial t} + (\mathcal{L}_m g_m)(y; t) = 0.$$

The measure change given by the function $G_m^{yt}(y') = g_m(y'; t)$ is called numeraire change and the Markov generator transforms as follows:

$$(6.9) \qquad \mathcal{L}'_m(y, y'; t) = \frac{g_m(y'; t)}{g_m(y; t)} \mathcal{L}_m(y, y'; t) + \frac{1}{g_m(y; t)} \frac{\partial g_m(y; t)}{\partial t} \delta_{yy'}.$$

Notice that if $g_m(t)$ denotes the multiplication operator of kernel $g_m(y; t) \delta_{yy'}$, then equation (13.4) can be written more compactly as follows:

$$(6.10) \qquad \mathcal{L}'_m(t) = \frac{1}{g_m(t)} \mathcal{L}_m(t) g_m(t) + \frac{1}{g_m(t)} \frac{\partial g_m(t)}{\partial t}.$$

Theorem 7. (Numeraire Changes.) *If g_m satisfies equation (13.3) and defines a numeraire change, then*

$$(6.11) \qquad U'_m(y, t; y', t') = \frac{g_m(y', t')}{g_m(y, t)} U_m(y, t; y', t').$$

Let $F_t = F(y., t)$ be an adapted process in the time interval $[T, T']$. Let us consider a fixed time $t \in [T, T']$. Recall that, since F_t is non-anticipatory, we have that $F(y., t) = F(y'., t)$ for all pairs of paths such that $y_s = y'_s$ for all $s \leq t$. If $\bar{y} \in \mathcal{D}_m(U', t, y)$, let $\tilde{y}. = \text{Ext}_t(y., y_t)$ be the constant extension path such that $\tilde{y}_s = y_s$ for all $s < t$ and $\tilde{y}_s = y_t$ for all $s \geq t$. With probability one, we have that

$$(6.12) \qquad \lim_{\delta t \downarrow 0} F(y., t + \delta t) = F(\text{Ext}_t(y., \bar{y}), t)$$

for some $\bar{y} \in \mathcal{D}_m(U, t, y)$.

Definition 33. (Monotonic Processes.) *Let U be a piecewise smooth Markov propagator and let F_t be an adapted process given by the non-anticipatory path functional $F(y., t)$. F_t is said to be increasing at time t if*

  (i) *For all $\bar{y} \in \mathcal{D}_m(U, t, y)$ we have that*

$$(6.13) \qquad F(\text{Ext}_t(y., \bar{y}), t) - F(y, t) \geq 0.$$

(ii) *We have that*

$$(6.14) \qquad \left.\frac{\partial F(\mathrm{Ext_s}(y_., y_{t-0}), s)}{\partial s}\right|_{s=t} \geq 0$$

where $y_{t-0} = \lim_{\delta t \downarrow 0} y_{t-\delta t}$.

(iii) *For all $y \in A_m$ and all $t \in [T, T']$, either the inequality in (i) holds in a strict sense for at least one $\bar{y} \in \mathcal{D}_m(U, t, y)$ or the inequality in (ii) holds in a strict sense.*

In case the property (iii) fails but the other two still hold, the process is called non decreasing.

Theorem 8. (Monotonic Processes.) *Let U_m and U'_m be two piecewise smooth Markov propagators and let $F(y_., t)$ be a non-anticipatory path functional. If U_m and U'_m are equivalent and if $F(y_., t)$ regarded as an adapted process under the path measure generated by U_m is increasing (non-decreasing), then also $F(y_., t)$ regarded as an adapted process under U'_m is increasing (non-decreasing).*

Proof. This theorem descends from the fact that the definition of monotonicity depends only on the attainable sets. $\qquad \square$

Definition 34. (Martingale Processes.) *A process F_t is a martingale if for all times t we have that*

$$(6.15) \quad \frac{\partial}{\partial t}\tilde{F}(\mathrm{Ext_t}(y_., y_{t-0}), t) + \sum_{\bar{y} \in \mathcal{D}(U, t, y_t)} \mathcal{L}(y_t, \bar{y}, t)\big(F(\mathrm{Ext_t}(y_., \bar{y}), t) - F(\mathrm{Ext_t}(y_., y_{t-0}), t)\big) = 0$$

The process F_t is called an equivalent martingale if there is a second piecewise smooth Markov propagator U' for which the non-anticipatory path functional $F(y_., t)$ is a martingale process.

Theorem 9. (Equivalent Martingales.)

(i) *If F_t is an increasing adapted process then it is not an equivalent martingale. Otherwise stated, if F_t is an equivalent martingale then it is not increasing.*

(ii) *If F_t is a non-decreasing adapted process and it is also an equivalent martingale, then it is a constant process with $F(y_., t) = \mathrm{const}$ for all $t \in [T, T']$.*

Proof. This theorem descends from the fact that monotonicity properties are preserved by measure changes. $\qquad \square$

Martingales are particularly useful as they can be constructed by taking expectations.

Let $\Phi(y_.)$ be a continuous path-functional. From the modeling viewpoint, such a functional can represent future cash flow streams. For instance, one choice could be

$$(6.16) \qquad \Phi(y_.) = \Phi_0(y_{t'})$$

where Φ_0 is a continuous univariate function and $t' \in [T, T']$ is fixed. In a more general example, one may consider a path functional of the form

$$(6.17) \qquad \Phi(y_.) \equiv \int_T^{T'} ds_1 ... \int_{s_{n-1}}^{T'} ds_n a(t; s_1, ...s_n) F_1(y_{s_1}, s_1)...F_n(y_{s_n}, s_n)$$

with $a(t; s_1, ...s_n) = 0$ for $t \in [T, t']$. The path conditioned expectation of $\Phi(y_.)$ is the non-anticipatory path functional $F_t = F(y_., t)$ such that

$$(6.18) \qquad F(y_., t) = \int_{A(y_., t)} \Phi(z_.)\mu[dz_.]$$

where the integral is restricted to the set $A(y_., t) = \{z_. | z_s = y_s \forall s \leq t\}$. The intersection of the set $A(y_., t)$ with each of the spaces $\mathcal{H}_{mn}^q$ is a compact, finite dimensional submanifold with boundaries. To denote path conditioning, we also use the following notation:

$$(6.19) \qquad F_t = F(y_., t) = E[\Phi(y_.)|y_s, s \leq t] = E_t[\Phi(y_.)].$$

Proposition 3. *The process F_t in (6.19) is a martingale for $t < t_1$.*

Proof. Equation (6.15) descends from the backward equation in (5.9). $\qquad \square$

7. The Fundamental Theorem of Finance

In this section, we derive the Fundamental Theorem of Finance in a context general enough to encompass most cases of practical relevance.

Definition 35. (Financial Model.) *Let $A_m, m \geq m_0$, be a simplicial sequence. A financial model is given by a family $A_m^k(y., t), k = 1, ...n$ of adapted processes modeling asset prices and a non-decreasing adapted process $B_m(y., t)$ modeling the money-market account. For notational convenience, we set $A_m^0(y., t) = B_m(y., t)$. Let us introduce also the discounted asset price process defined as follows:*

$$(7.1) \qquad \tilde{A}_m^k(y., t) = \frac{A_m^k(y., t)}{B_m(y., t)}.$$

Definition 36. (Trading Strategies.) *Given a financial model with n assets, a trading strategy is given by a family of adapted processes $\zeta_m^k(y., t), k = 0, ...n$. The value process of a strategy is the adapted process*

$$(7.2) \qquad \Pi_m(y., t) = \sum_{k=0}^{n} \zeta_m^k(y., t) A_m^k(y., t)$$

The discounted value process instead is given by

$$(7.3) \qquad \tilde{\Pi}_m(y., t) = \sum_{k=0}^{n} \zeta_m^k(y., t) \tilde{A}_m^k(y., t)$$

Definition 37. (Self-Financing Condition.) *An adapted trading strategy is called self-financing if the following two conditions hold:*

$$(7.4) \qquad (SF1) \qquad \sum_{k=0}^{n} \frac{\partial \zeta_m^k}{\partial t}(\text{Ext}_t(y., y_{t-0}), t) \tilde{A}_m^k(\text{Ext}_t(y., y_{t-0}), t) = 0$$

and if, for all $\bar{y} \in \mathcal{D}_m(U, t, y_t)$, we have that

$$(7.5) \qquad (SF2) \qquad \sum_{k=0}^{n} \left(\zeta_m^k(\text{Ext}_t(y., \bar{y}), t) - \zeta_m^k(\text{Ext}_t(y., y_{t-0}), t) \right) \tilde{A}_m^k(\text{Ext}_t(y., \bar{y}), t) = 0.$$

Proposition 4. *If $\zeta_m^k(y., t)$ is a self-financing trading strategy, then the corresponding discounted value process satisfies*

$$(7.6) \qquad \frac{\partial \tilde{\Pi}_m(\text{Ext}_t(y., y_{t-0}), t)}{\partial t} = \sum_{k=0}^{n} \zeta_m^k(\text{Ext}_t(y., y_{t-0}), t) \frac{\partial \tilde{A}_m^k(\text{Ext}_t(y., y_{t-0}), t)}{\partial t}$$

and for all $\bar{y} \in \mathcal{D}_m(U, t, y_t)$, we have that

$$(7.7) \quad \tilde{\Pi}_m(\text{Ext}_t(y., \bar{y}), t) - \tilde{\Pi}_m(\text{Ext}_t(y., y_{t-0}), t)$$

$$(7.8) \qquad = \sum_{k=0}^{n} \zeta_m^k(\text{Ext}_t(y., \bar{y}), t) \left(\tilde{A}_m^k(\text{Ext}_t(y., \bar{y}), t) - \tilde{A}_m^k(\text{Ext}_t(y., y_{t-0}), t) \right).$$

Definition 38. (Arbitrage Strategies.) *A self-financing strategy is called arbitrage at time t if the corresponding discounted value process $\tilde{\Pi}_m(y., t)$ is increasing at time t.*

Theorem 10. (Fundamental Theorem of Finance.) *If there is an equivalent measure with respect to which all discounted base asset price processes are martingales, then*

- (i) *The discounted value process of any self-financing trading strategy under the same equivalent measure is a martingale.*
- (ii) *There is no arbitrage.*

Conversely, if there is no arbitrage than there exists a measure change under which all discounted asset price processes become martingales.

Proof. The first part of the theorem is a simple consequence of the definitions put forward. The converse instead requires a proof.

Let $\mathcal{V}_m(U, t, y_t)$ be the vector space spanned by functions of the form $v(\bar{y})$ where $\bar{y} \in \mathcal{D}_m(U, t, y_t)$. Let K be the cone in $\mathcal{V}_m(U, t, y_t)$ made up by all the vectors with non-negative components. Also, let us introduce the following vector $\xi_m^k \in \mathcal{V}_m(U, t, y_t)$:

(7.9)

$$\xi_m^k(\bar{y}) = \delta_{\bar{y}, y_{t-0}} \frac{\partial \tilde{A}_m^k(\mathrm{Ext}_t(y., y_{t-0}), t)}{\partial t} + \left(1 - \delta_{\bar{y}, y_{t-0}}\right)\left(\tilde{A}_m^k(\mathrm{Ext}_t(y., \bar{y}), t) - A_m^k(\mathrm{Ext}_t(y., y_{t-0}), t)\right).$$

Notice that, in case $k = 0$ and the asset is the money market account, then $\xi_m^0(\bar{y}) = 0$ for all $\bar{y} \in \mathcal{D}_m(U, t, y_t)$.

Suppose there is no arbitrage and fix a time t. Assuming there does not exist a trading strategy with a strictly increasing value process, for all vectors $(v^k)_{k=1,\ldots n} \in \mathcal{V}_m(U, t, y_t)$ there are two elements $y_+, y_- \in \mathcal{D}_m(U, t, y_t)$ such that

(7.10)
$$\sum_{k=1}^n v^k \xi_m^k(\bar{y}_+) > 0 \quad \text{while} \quad \sum_{k=1}^n v^k \xi_m^k(\bar{y}_-) < 0.$$

If in equation (7.10) we set $v^k = \delta_{k1}$, we conclude that the vector $\left(\xi_m^1(\bar{y})\right)_{\bar{y} \in \mathcal{D}_m(U, t, y_t)}$ has both positive and negative components. Hence, the the hyperplane $\Pi^1 \subset \mathcal{V}_m(U, t, y_t)$ orthogonal to the vector $\xi_m^1(\bar{y}) \in \mathcal{V}_m(U, t, y_t)$ intersects K.

Let P_1 be the orthogonal projection operator onto the hyperplane Π^1 and let us consider the vector

(7.11)
$$(P_1 \xi_m^2)(\bar{y}) = \xi_m^2(\bar{y}) - \frac{\sum_{\bar{z} \in \mathcal{D}_m(U, t, y_t)} \xi_m^2(\bar{z}) \xi_m^1(\bar{z})}{\sum_{\bar{z} \in \mathcal{D}_m(U, t, y_t)} \xi_m^1(\bar{z}) \xi_m^1(\bar{z})} \xi_m^1(\bar{y}).$$

for $i = 2, \ldots n$. Due to absence of arbitrage, there are two elements $y_+, y_- \in \mathcal{D}_m(U, t, y_t)$ such that

(7.12)
$$(P_1 \xi_m^2)(\bar{y}_+) > 0 \quad \text{while} \quad (P_1 \xi_m^2)(\bar{y}_-) < 0.$$

Hence, the vector $P_1 \xi_m^2$ is transversal to the octant K of positive vectors. As a consequence, the hyperplane $\Pi_2 \subset \Pi_1 \mathcal{V}$ is orthogonal to both vectors ξ_m^1 and ξ_m^2, also intersects K.

The argument above can be iterated n times, leading to the conclusion that there exists a strictly positive function $g_m^{yt}(\bar{y}) > 0, \bar{y} \in \mathcal{D}_m(U, t, y_t)$, such that

(7.13)
$$\sum_{\bar{y} \in \mathcal{D}_m(U, t, y_t)} \xi_m^k(\bar{y}) g_m^{yt}(\bar{y}) = 0$$

for all $k = 0, \ldots n$. In particular, this implies that there exists a measure change function $G_m^{yt}(\bar{y})$ such that

(7.14)
$$\sum_{y \in h\mathbb{Z}^d} \left(\tilde{A}_m^k(\mathrm{Ext}_t(y., \bar{y}), t) - A_m^k(\mathrm{Ext}_t(y., y_{t-0}), t)\right) \mathcal{L}(y, \bar{y}; t) G_m^{yt}(\bar{y}) = 0.$$

$\square$

8. Weak Convergence of Markov Generators

Consider a one dimensional Markov process defined on the simplicial sequence $A_m, m \geq m_0$ and the time interval $[T, T']$. Many different specifications of Markov generators on A_m may correspond to the same limit. In this section we identify a canonical sequence of generators under a few regularity hypotheses which imply the existence of a weak limit in distribution sense for the generator. This is a necessary first step to single out the general form of an admissible Markov generator. In the following sections, we then investigate convergence under the much finer criteria of pointwise convergence for probability kernels and their derivatives.

First consider the case when the limiting domain A_∞ is bounded. Without restricting generality, let us suppose that $A_\infty = [-L, L]$.

Let $\mathcal{L}_m$ be a sequence of Markov generators. The first assumption we make is that the first two moments are finite, or more specifically

Hypothesis MG1. *The sequences*

$$(8.1) \qquad \mu_m(x,t) \equiv \sum_{y \in A_m} \mathcal{L}_m(x,y;t)(y-x), \qquad \sigma_m(x,t) \equiv \sqrt{\sum_{y \in A_m} \mathcal{L}_m(x,y;t)(y-x)^2}$$

are uniformly bounded in absolute value as $m \to \infty$ and converge to limits for all $x \in A_m$ and all $t \in [T, T']$, i.e. the following limits exist:

$$(8.2) \qquad \mu(x,t) \equiv \lim_{m \to \infty} \mu_m(x,t), \qquad \sigma(x,t) \equiv \lim_{m \to \infty} \sigma_m(x,t).$$

Notice that, due to the dominated convergence theorem, the functions $\mu_m(x,t)$ and $\sigma_m(x,t)$ regarded as piecewise constant functions on $\mathcal{A}_\infty$ converge weakly to the corresponding limits.

Hypothesis MG2. *The following limits exist and are finite for all $m \geq m_0$ and all pairs $x, y \in A_m$ such that $|x - y| \geq 2h_m$:*

$$(8.3) \qquad \lambda_m(x,y;t) = \lim_{m' \to \infty} \sum_{z \in A_{m'} : y - h_m < z \leq y + h_m} \mathcal{L}_{m'}(x,z;t).$$

The family of functions $\lambda_m(x,y)$ can be represented in terms of the so called Levy measures specified as follows:

Theorem 11. (Levy measures.) *For all $x \in A_m$ there exists a measure $\nu_{xt}(d\xi)$ in $[-L, L]$ (called Levy measure) with the following two properties:*

(i)

$$(8.4) \qquad \int_{-L}^{L} \nu_{xt}(d\xi)\xi^2 < \infty$$

(ii) *For all $y \in A_m$ with $|x - y| \geq 2h_m$ the following representation is valid:*

$$(8.5) \qquad \lambda_m(x,y) = \lim_{\varepsilon \downarrow 0} \int_{y - h_m/2 + \varepsilon}^{y + h_m/2} \nu_{xt}(d\xi).$$

Definition 39. (Finite activity jumps.) *A Markov process is said to have finite activity jumps if for all $m \geq m_0$ and all $x \in A_m$ we have that*

$$(8.6) \qquad \int_{-L}^{L} \nu_{xt}(d\xi) < \infty$$

Notice that given $MG1$ and $MG2$ we have

$$(8.7) \qquad \limsup_{m \to \infty} \sum_y \lambda_m(x,y;t)(y-x)^2 \leq \sum_y \mathcal{L}_m(x,y;t)(y-x)^2 < \infty.$$

More generally, if $\phi \in \mathcal{D}(\mathbb{R})$ is a test function such that $\phi(0) = 0$ and $\phi'(0) = 0$, then we have that

$$(8.8) \qquad \limsup_{m \to \infty} \sum_y \lambda_m(x,y;t)\phi(y-x) < \infty.$$

Although the above sequence is bounded, the limit may not exist in general. We thus need to stipulate this as a separate assumption:

Hypothesis MG3. *For all test functions $\phi \in \mathcal{D}(\mathbb{R})$ such that $\phi(0) = \phi'(0) = 0$ and all $x \in A_m$, the sequence $\sum_y \lambda_{m'}(x,y;t)\phi(y-x)$ admits a limit as $m' \to \infty$.*

Let us introduce the following notations:

$$(8.9) \qquad \mu_m^0(x;t) = \sum_y \lambda_m(x,y;t)(y-x), \qquad \sigma_m^0(x;t) = \sqrt{\sum_y \lambda_m(x,y;t)(y-x)^2}.$$

Notice that the sequence $\sigma_m^0(x,t)$ is non-negative and is uniformly bounded as a function of m. More precisely

$$(8.10) \qquad\qquad 0 \le \sigma_m^0(x,t) \le \sigma_m(x,t).$$

On the other hand, we cannot conclude in general that $\mu_m^0(x,t)$ is necessarily uniformly bounded. In fact, the difference $\mu^m(x,t) - \mu_m^0(x,t)$ could diverge as $m \to \infty$ while being compensated by terms concentrated at $y = x \pm 1$ which in turn also diverge in the limit as $m \to \infty$ while keeping the total drift $\mu^m(x,t)$ uniformly bounded. An exception to this general situation is found in case the following condition is satisfied:

Hypothesis MG4. *For all $m \ge m_0$ and all $x \in A_m$ we have that*

$$(8.11) \qquad\qquad \int_{-L}^{L} |x| \nu_{xt}(d\xi) < \infty$$

Two particularly important situations in which this condition $MG4$ holds are given in the following:

Theorem 12. (MG4.) *Assuming $MG1, MG2, MG3$, if the Markov process is either monotonic or has finite activity jumps, then $MG4$ holds.*

Under the three assumptions $MG1, MG2, MG3$, the sequence $\mathcal{L}_m$ of Markov generators can be mapped into an equivalent canonical sequence of Markov generators $\mathcal{L}^C$. More precisely, we set

$$(8.12) \qquad\qquad \mathcal{L}_m^C(x,y;t) = \lambda(x,y;t)$$

in case $|x-y| \ge 2h_m$. Furthermore, we set

$$(8.13) \qquad \mathcal{L}_m^C(x,x\pm h_m;t) = \tilde{\mu}_m(x;t)\nabla_{h_m}(x,x\pm 1) + \frac{\tilde{\sigma}_m(x;t)^2}{2}\Delta_{h_m}(x,x\pm 1)$$

where

$$(8.14) \qquad \tilde{\mu}_m(x;t) = \mu_m(x) - \mu_m^0(x), \qquad \tilde{\sigma}_m(x;t)^2 = \sigma_m(x;t)^2 - \sigma_m^0(x;t)^2.$$

and the operators ∇_h and Δ_h are defined as in equation (5.21). Finally,

$$(8.15) \qquad\qquad \mathcal{L}_m^C(x,x;t) = \sum_{y \in A_m, y \ne x} \mathcal{L}_m^C(x,y;t)$$

Theorem 13. (Canonical Representations of Markov Generators.) *For all smooth functions of compact support $\phi \in \mathcal{D}(A_\infty)$ and all $x \in A_m$ we have that*

$$(8.16) \qquad\qquad \lim_{m' \to \infty} \sum_{y \in A_{m'}} \mathcal{L}_{m'}^C(x,y)\phi(y) = \lim_{m' \to \infty} \sum_{y \in A_{m'}} \mathcal{L}_{m'}(x,y)\phi(y).$$

A compact representation for a canonical generator summarizing these definition is obtained by considering the symbol as specified in Definition (24).

Theorem 14. (Levy-Khintchine representation.) *Under the hypothesis $MG1, MG2, MG3$ above and in case $A_\infty = [-L,L]$ is bounded, the symbol of a generator in canonical form can be expressed as follows:*

$$(8.17)$$
$$\hat{\mathcal{L}}_m^C(x,p;t) = i\mu(x;t)\frac{\sin ph}{h} + \tilde{\sigma}(x;t)^2\frac{\cos ph - 1}{h^2} + \sum_{y \in A_m}\left(e^{ip(y-x)} - 1 - i\frac{\sin ph}{h}(y-x)\right)\lambda_m(x,y;t).$$

The following limit converges in the weak topology:

$$(8.18) \qquad \lim_{m \to \infty} \hat{\mathcal{L}}_m^C(x, p; t) = i\mu(x; t)p - \frac{\tilde{\sigma}^2(x; t)}{2}p^2 + \int_{-L}^{L} (e^{ip\xi} - 1 - ip\xi)\nu_{xt}(d\xi),$$

where $\nu_{xt}(d\xi)$ is the Levy measure. Moreover, if also MG4 is satisfied, then the symbol of a generator in canonical form can be expressed as follows:

$$(8.19) \qquad \hat{\mathcal{L}}_m^C(x, p; t) = i\tilde{\mu}(x; t)\frac{\sin ph}{h} + \tilde{\sigma}(x; t)^2 \frac{\cos ph - 1}{h^2} + \sum_{y \in A_m} \left(e^{ip(y-x)} - 1 \right) \lambda_m(x, y; t).$$

The following limit converges in the weak topology:

$$(8.20) \qquad \lim_{m \to \infty} \hat{\mathcal{L}}_m^C(x, p; t) = i\tilde{\mu}(x; t)p - \frac{\tilde{\sigma}^2(x; t)}{2}p^2 + \int_{-L}^{L} (e^{ip\xi} - 1)\nu_{xt}(d\xi).$$

9. Fast Exponentiation and Spectral Methods

Numerical analysis of pricing models in the operator formalism depend on the ability to compute the propagator for a given generator $\mathcal{L}(t)$. Time homogenous Markov generators $\mathcal{L}_m(y_1, y_2)$ represent a privileged special case of particular importance. In this case, the associated propagator solving the differential equation in (5.2) is given by the matrix exponential

$$(9.1) \qquad U_m(y_1, s; y_2, s') = \exp((s' - s)\mathcal{L}_m)(y_1, y_2).$$

Matrix exponentiation can be defined in several equivalent ways, such as by Taylor expansion

$$(9.2) \qquad \exp(t\mathcal{L}_m) = \sum_{j=0}^{\infty} \frac{t^j}{j!} \mathcal{L}_m^j$$

or by means of Neper's formula

$$(9.3) \qquad \exp(t\mathcal{L}_m) = \lim_{N \to \infty} \left(1 + \frac{t}{N}\mathcal{L}_m \right)^N$$

We find that the most efficient and robust method for exponentiating Markov generators is the so called *fast exponentiation* algorithm. Let us fix a Markov generator $\mathcal{L}(x, y)$ and a time horizon t. Let $\delta t > 0$ be the largest time interval for which both of the following properties are satisfied:

$$(\text{FE1}) \qquad \min_{y \in \Lambda}(1 + \delta t \mathcal{L}(y, y)) \geq 1/2$$

$$(\text{FE1}) \qquad \log_2 \frac{t}{\delta t} = n \in \mathbb{N}.$$

To compute $e^{t\mathcal{L}}(x, y)$, we first define the elementary propagator

$$(9.4) \qquad u_{\delta t}(x, y) = \delta_{xy} + \delta t \mathcal{L}(x, y)$$

and then evaluate in sequence $u_{2\delta t} = u_{\delta t} \cdot u_{\delta t}$, $u_{4\delta t} = u_{2\delta t} \cdot u_{2\delta t}$, ... $u_{2^n \delta t} = u_{2^{n-1}\delta t} \cdot u_{2^{n-1}\delta t}$.

As we show in the next few sections, this algorithm approximates probability kernels with errors with respect to the continuous time kernel density which, in fairly general cases, are of order $O(h_m^2 |\log h_m|^3)$. This is the same order by which the continuous time kernel density differs from its continuous limit also according to estimates below. In the case of Brownian motion, a sharp convergence estimates of ordered $O(h_m^2)$ can also be proven.

Matrix multiplication is accomplished numerically by invoking the routine `dgemm` in Level-3 BLAS. Very efficient, processor specific version of `dgemm` are now available along with implementations on massively parallel GPU chipsets. It turns out that the standard measure of algorithmic complexity as the number of floating point operations required to accomplish a certain task, is not simply proportional and scales non-linearly with respect to execution time. Using blocking and cache optimizations and distributing the load across many cores, execution time for medium to large matrices appears to scales much better than the naive n^3 scaling one would obtain by triple looping (Goto and van de Geijn to appear).

A more general method that allows to compute not only exponentials but also other functions of a Markov generator is full diagonalization. Unfortunately, unless the Markov generator is symmetrizable, diagonalization algorithms can possibly run into serious instabilities within double precision arithmetics because of the phenomenon of pseudo-spectrum (Trefethen and Embree 2006). This makes it impossible to diagonalize exactly within the limits of double-precision arithmetics and forces one to resort to expedients such as for instance spectral truncations. Since the fast exponentiation method has much better scaling properties than full diagonalization and is entirely stable, we recommend that it be used in all situations where a matrix exponentiation is required. However, it often arises the necessity to define and compute also other functions of a given time independent Markov generator and for this purpose diagonalization can be usefully employed, especially when the target matrix is symmetrizable.

Not all Markov generators are diagonalizable but most are and it is safe to take diagonalizability for granted in numerical applications. To make this statement more precise, one sets out the following definitions:

Definition 40. (Generic Properties.) *A dense G_δ of a topological space is a countable intersection of dense open subsets. If a property is valid on a dense G_δ one says that it is valid generically, or that it is generic. Instead, if the same topological set is endowed of a measure and a property is valid on a full measure set of parameters, one says that it is valid almost surely with respect to that particular measure.*

We have that

Proposition 5. *Markov generators are diagonalizable both generically and almost surely.*

Let $u_n(x)$ and λ_n, $n = 1, ..N$, be eigenfunctions and eigenvalues of the Markov generator $\mathcal{L}$, i.e.

$$(9.5) \qquad \mathcal{L}u_n = \lambda_n u_n.$$

Let U be the matrix whose columns are given by the vectors $u_n(x)$ and let Λ be the diagonal matrix with the eigenvalues λ_n on the diagonal. Hence

$$(9.6) \qquad \mathcal{L}U = U\Lambda.$$

We have that $\mathcal{L} = U\Lambda U^{-1}$ and $e^{t\mathcal{L}} = Ue^{t\Lambda}U^{-1}$. This equation expressed in components reads as follows:

$$(9.7) \qquad e^{t\mathcal{L}}(x, x') = \sum_{n=1}^{N} e^{\lambda_n t} u_n(x) v_n(y)$$

where $v_n(y)$ is the $n-$th row vector of the inverse matrix $V = U^{-1}$.

One may extend the above definition to other functions of a Markov generator. If $\psi(\lambda)$ is a function, one may define $\psi(\mathcal{L})$ as the operator whose matrix is given by

$$(9.8) \qquad \psi(\mathcal{L})(x, x') = \sum_{n=1}^{N} \psi(\lambda_n) u_n(x) v_n(y)$$

An important example of functional calculus is found to express the Markov generator of processes obtained by stochastic time change, whereby the time-change process has independent, uniformly distributed increments. Processes in this class are called *Bochner subordinators*. Because of time and space homogeneity, Bochner subordinators can be constructed starting from a process on simplicial sequence $h_m\mathbb{Z}$ and are characterized by a Markov generator of the form $\mathcal{L}_m(x, y) = \ell_{h_m}(y - x)$.

Theorem 15. (Bochner Subordinators.) *If the limit $\lim_{h\downarrow 0} \ell_h = \ell$ exists in weak sense on $\mathcal{D}'(\mathbb{R})$ then the limit kernel has the following form*

$$(9.9) \qquad \int \ell(x)\phi(x)dx = \mu\phi'(0) + \int_0^\infty (\phi(x) - \phi(0))\nu(dx)$$

where $\nu(dx)$ is a positive measure supported on $\mathbb{R}_+$ and is such that

$$(9.10) \qquad \int_0^\infty x\nu(dx) < \infty.$$

The *characteristic function* of the process defined as the Fourier transform of the Markov generator is thus

$$(9.11) \qquad \epsilon(k) = \lim_{h\downarrow 0} \sum_x e^{ikx}\ell_h(x) = i\mu k + \int_0^\infty (e^{ikx} - 1)\nu(dx).$$

The Laplace transform instead is given by

$$(9.12) \qquad \phi(\lambda) \equiv -\epsilon(i\lambda) = -\lim_{h\downarrow 0} \sum_x e^{-\lambda x}\ell_h(x) = \mu\lambda + \int_0^\infty (1 - e^{-\lambda x})\nu(dx).$$

A function of this form is called *Bernstein function.*

Bernstein functions may be used to express Laplace transform of the transition probability kernel of time-homogeneous monotonic processes as follows:

$$(9.13) \qquad \int_0^\infty e^{t\mathcal{L}}(0,x)e^{-\lambda x}dx = e^{-t\phi(\lambda)}.$$

In turn, we may also express the transition probability kernel in terms of the characteristic function, i.e.

$$(9.14) \qquad e^{t\mathcal{L}}(0,x) = \int_{-\infty}^\infty e^{it\epsilon(k)+ikx}\frac{dk}{2\pi}.$$

Notice that since $\phi(-ik) = -\epsilon(k)$, the last formula may also be interpreted as the Fourier-Mellin inversion of the previous one.

Example 1. (Poisson Process) The Poisson process corresponds to

$$(9.15) \qquad \phi_P(\lambda; c) = c(1 - e^{-\lambda}) = c\int_0^\infty (1 - e^{-\lambda t})\delta(t - 1)dt.$$

Example 2. (Stable Process) The stable subordinator with index $\alpha \in (0,1)$ is given by

$$(9.16) \qquad \phi_S(\lambda; \alpha) = \lambda^\alpha = \frac{\alpha}{\Gamma(1-\alpha)} \int_0^\infty (1 - e^{-\lambda t})t^{-1-\alpha}dt.$$

Example 3. (Gamma Process) The Gamma subordinator with variance rate $\nu > 0$ is given by

$$(9.17) \qquad \phi_{VG}(\lambda; \nu) = \frac{1}{\nu}\log(1 + \nu\lambda) = \frac{1}{\nu} \int_0^\infty (1 - e^{-\lambda t})t^{-1}e^{-t/\nu}dt.$$

To add jumps to a diffusion process one can use the method of independent stochastic subordination. Namely let F_t be the diffusion process with drift function $\mu(F)$ and volatility function $\sigma(F)$ and let $\mathcal{L}^d(y_1, y_2; \mu, \sigma)$ be the corresponding Markov generator. Consider the generator

$$(9.18) \qquad \mathcal{L}_{VG}(\mu, \sigma, \nu) = -\phi_{VG}(-\mathcal{L}^d(\mu, \sigma); \nu).$$

The propagator for $\mathcal{L}^j(\mu, \sigma, \nu)$ satisfies the following equation:
(9.19)

$$\exp\left(t\mathcal{L}_{VG}(\mu, \sigma, nu)\right)(y_1, y_2) = E_0\left[\exp\left(T_t^\nu\mathcal{L}^d(\mu, \sigma)\right)(y_1, y_2)\right] = \sum_{n=1}^N e^{-\phi_{VG}(-\lambda_n;\nu)t}u_n(x)v_n(y)$$

where T_t^ν are the paths of the monotonic process in Example 3, the $u_n(x)$ are the eigenfunctions of the diffusion generator $\mathcal{L}^d(\mu, \sigma)$, λ_n are the corresponding eigenvalues and the functions $v_n(x)$ are defined as in (9.7). Hence, the generator $\mathcal{L}_{VG}(\mu, \sigma, \nu)$ identifies the time-changed process with paths $F_{T_t^\nu}$.

To model asymmetric jumps one can follow several strategies. A simple one is to specify the two different variance rates ν_+ and ν_- for the up and down jumps and compute separately two Markov generators

$$(9.20) \qquad \mathcal{L}_{VG}(\mu, \sigma, \nu_\pm) = -\phi_{VG}(-\mathcal{L}^d(\mu, \sigma); \nu_\pm).$$

The new generator for our process with asymmetric jumps is obtained by combining the two generators above

$$(9.21) \qquad \mathcal{L}_{VG}(y_1, y_2; \mu, \sigma, \nu_+, \nu_-) = \begin{cases} \mathcal{L}_{VG}(y_1, y_2; \mu, \sigma, \nu_-) & \text{if } F(y_1) < F(y_2) \\ \mathcal{L}_{VG}(y_1, y_2; \mu, \sigma, \nu_+) & \text{if } F(y_1) > F(y_2) \\ -\sum_{y_2 \neq y_1} \mathcal{L}_{VG}(y_1, y_2; \mu, \sigma, \nu_+, \nu_-). \end{cases}$$

The construction can also be localized. If $\phi_x(\lambda)$ is a family of Bernstein functions indexed by the state variable $x \in \Lambda_m$, then one can consider the Markov generator of matrix

$$(9.22) \qquad \tilde{\mathcal{L}}(x, y) = -\phi_x(-\mathcal{L}^d(\mu, \sigma))(x, y).$$

This construction allows one to model state dependent jumps.

10. Construction of Brownian Motion

This section is based on work in collaboration with Alex Mijatovic, see (Albanese and Mijatovic 2006).

Let $A = (A_m), m = m_0, m_0+1, \dots$ be a simplicial sequence converging to an interval $\lim_{m \to \infty} A_m = [-L, L] \subset \mathbb{R}$, where $0 < L < \infty$. Let μ and σ be two constants.

The generator of a Brownian motion on A_m has the form

$$(10.1) \qquad \mathcal{L}_m = \mu \nabla_{h_m} + \frac{1}{2} \sigma^2 \Delta_{h_m}$$

for all interior points $x \in Int(A_m)$. We assume that m_0 is large enough so that

$$(10.2) \qquad \frac{\sigma^2}{2h_m^2} > \frac{|\mu|}{2h_m}$$

for all $m \geq m_0$.

If x is a boundary point, i.e. $x \in \partial(A_m)$, then several definitions of the operator $\mathcal{L}_m$ are possible depending on the choice of boundary conditions. *Absorbing* boundary conditions correspond to the choice

$$(10.3) \qquad \mathcal{L}_m(x, y) = 0 \quad \forall y \in A_m.$$

Reflecting boundary conditions correspond to

$$(10.4) \qquad \mathcal{L}_m(x, x) = \mu \nabla_{h_m}(x, x) + \frac{1}{2} \sigma^2 \Delta_{h_m}(x, x)$$

$$(10.5) \qquad \mathcal{L}_m(x, y) = -\mathcal{L}_m(x, x)$$

where $y \in Int(A_m)$ is the closest point to x in the interior of A_m, while $\mathcal{L}_m(x, y) = 0$ for all other points y. *Periodic* boundary conditions are implemented by setting

$$(10.6) \qquad \mathcal{L}_m(x, y) = \mu \nabla_{h_m}(x, y) + \frac{1}{2} \sigma^2 \Delta_{h_m}(x, y)$$

if $y = x$ or $y \in Int(A_m)$ is the closest point to x in the interior of A_m, and also

$$(10.7) \qquad \mathcal{L}_m(x, x') = \mu \nabla_{h_m}(x, y) + \frac{1}{2} \sigma^2 \Delta_{h_m}(x, y)$$

where $x' \in \partial(A_m)$ is the boundary point at the opposite extreme of the simplex A_m. Finally, *mixed* boundary conditions can also be defined by taking a convex linear combination of the generators satisfying one of the three boundary conditions above.

Theorem 16. (Convergence Estimates for Brownian Motion.) *Consider a Brownian motion defined as the process with generator in (10.1) where μ and σ are constants satisfying the bounds in (10.2) for all $m \geq m_0$ and we assume periodic boundary conditions. Let us consider the kernels*

$$(10.8) \qquad U_m(x, 0; y, T) \equiv e^{T\mathcal{L}_m}(x, y)$$

and

$$(10.9) \qquad U_m^{\delta t}(x, 0; y, T) \equiv \left(1 + \delta t \mathcal{L}_m\right)^{\frac{T}{\delta t}}(x, y).$$

If δt is such that $0 < \delta t \leq \frac{1}{2}\frac{h^2}{2\sigma^2 - h|\mu|}$ and m_0 is large enough, then there is a constant $c > 0$ such that for all $m' \geq m \geq m_0$ we have the following inequalities:

(i)

$$(10.10) \qquad |h_m^{-1} U_m(x, 0; y, T) - h_{m'}^{-1} U_{m'}(x, 0; y, T)| \leq ch_m^2$$

(ii)

$$(10.11) \qquad |h_m^{-1} \nabla_{h_m} U_m(x, 0; y, T) - h_{m'}^{-1} \nabla_{h_{m'}} U_{m'}(x, 0; y, T)| \leq ch_m^2$$

(iii)

$$(10.12) \qquad |h_m^{-1} \Delta_{h_m} U_m(x, 0; y, T) - h_{m'}^{-1} \Delta_{h_{m'}} U_{m'}(x, 0; y, T)| \leq ch_m^2$$

(iv)

$$(10.13) \qquad |h_m^{-1} U_m(x, 0; y, T) - h_m^{-1} U_m^{\delta t}(x, 0; y, T)| \leq ch_m^2$$

Proof. It suffices to show this inequality for $m' = m + 1$. Let us assume for simplicity that $L = 2^{m_0} h_{m_0}$ and $A_m = \{x = h_m i, i = 0, \dots 2^m\}$ for all $m \geq m_0$. The argument extends to more general lattice geometry but the consideration of these more general cases would obscure the simplicity of the proof with needless detail and will thus be omitted.

Let B_m be the Brillouin zone as defined in equation (5.16). Let $\mathcal{F}_m : \ell^2(A_m) \to \ell^2(B_m)$ be the Fourier transform operator defined so that:

$$(10.14) \qquad \hat{f}(p) \equiv \mathcal{F}_m(f)(p) = h_m \sum_{x \in A_m} f(x) e^{-ipx}$$

for all $p \in B_m$. The inverse Fourier transform is given by

$$(10.15) \qquad \mathcal{F}_m^{-1}(\hat{f})(x) = \frac{1}{2L} \sum_{p \in B_m} \hat{f}(p) e^{ipx}.$$

The Fourier transformed generator is diagonal and is given by the operator of multiplication by

$$(10.16) \qquad \hat{\ell}_m(p) = \mathcal{F}_m \mathcal{L}_m \mathcal{F}_m^{-1}(p, p) = -i\mu \frac{\sin h_m p}{h_m} + \sigma^2 \frac{\cos h_m p - 1}{h_m^2}.$$

We have

$$(10.17) \qquad h_m^{-1} U_m(x, 0; y, T) = \frac{1}{2L} \sum_{p \in B_m} e^{T\hat{\ell}_m(p)} e^{ip(y-x)}.$$

Using this Fourier series representation, we find

$$\left| h_m^{-1} U_m(x, 0; y, T) - h_{m+1}^{-1} U_{m+1}(x, 0; y, T) \right|$$

$$\leq \frac{1}{2L} \left| \sum_{p \in B_m} \left(e^{T\hat{\ell}_m(p)} - e^{T\hat{\ell}_{m+1}(p)} \right) e^{ip(y-x)} \right| + \frac{1}{2L} \left| \sum_{p \in B_{m+1} \setminus B_m} e^{T\hat{\ell}_{m+1}(p)} e^{ip(y-x)} \right|.$$

$$(10.18)$$

Let

$$(10.19) \qquad K_m = \sqrt{\frac{|\log h_{m+1}|}{\sigma^2 T}}.$$

If h_m is small enough, i.e. if m_0 is sufficiently large, we have that

$$\frac{1}{2L}\left|\sum_{p\in B_m, |p|\geq K_m} e^{T\hat{\ell}^m(p)} e^{ip(y-x)}\right| \leq \frac{1}{2L}\sum_{p\in B_m, |p|\geq K_m} e^{T\Re(\hat{\ell}^m(p))} \leq c\exp\left(T\sigma^2\frac{\cos h_m K_m - 1}{h_m^2}\right) \leq ch_m^2.$$

$$(10.20)$$

where $\Re(a)$ denotes the real part of $a \in \mathbb{C}$ and c denotes a generic constant. Similarly

$$\frac{1}{2L}\left|\sum_{p\in B_{m+1}, |p|\geq K_m} e^{T\hat{\ell}_{m+1}(p)} e^{ip(y-x)}\right| \leq \frac{1}{2L}\sum_{p\in B_m, |p|\geq K_m} e^{T\Re(\hat{\ell}_{m+1}(p))}$$

$$\leq c\exp\left(T\sigma^2\frac{\cos h_{m+1}K - 1}{h_{m+1}^2}\right) \leq ch_{m+1}^2$$

$$(10.21)$$

Since

$$(10.22) \qquad \frac{1}{2}h^2 p^3 - \frac{1}{8}h^4 p^5 \leq \frac{\sin hp}{h} - \frac{\sin 2hp}{2h} \leq \frac{1}{2}h^2 p^3$$

and

$$(10.23) \qquad -\frac{1}{8}h^2 p^4 \leq \frac{\cos hp - 1}{h^2} - \frac{\cos 2hp - 1}{(2h)^2} \leq -\frac{1}{8}h^2 p^4 + \frac{1}{48}h^4 p^6.$$

we find that if $|p|\leq \frac{\sqrt{2}}{h}$ then

$$(10.24) \qquad |\hat{\ell}_m(p) - \hat{\ell}_{m+1}(p)|\leq \frac{\mu}{4}h^2|p|^3 + \frac{\sigma^2}{16}h^2 p^4.$$

Moreover, since

$$(10.25) \qquad -\frac{1}{2}p^2 \leq \frac{\cos hp - 1}{h} \leq -\frac{1}{2}p^2 + \frac{1}{24}h^2 p^4$$

we conclude that in case $|p|\leq h^{-1}\sqrt{\frac{2}{3}}$, the following inequality holds:

$$(10.26) \qquad \frac{\cos hp - 1}{h} \leq -\frac{1}{4}p^2$$

Hence, if m_0 is large enough, we find

$$\frac{1}{2L}\left|\sum_{p\in B_m, |p|\leq K}\left(e^{T\hat{\ell}_m(p)} - e^{T\hat{\ell}_{m+1}(p)}\right)e^{ip(y-x)}\right|$$

$$\leq \frac{1}{2L}\sum_{p\in B_m, |p|\leq K} e^{-\frac{1}{4}p^2}\left(e^{\frac{\mu T}{4}h_m^2|p|^3 + \frac{\sigma^2 T}{16}h_m^2 p^4} - 1\right)$$

$$(10.27) \qquad \leq \frac{1}{2L}\sum_{p\in B_m, |p|\leq K} e^{-\frac{1}{4}p^2}\left(\frac{\mu T}{4}h_m^2|p|^3 + \frac{\sigma^2 T}{16}h_m^2 p^4\right) \leq ch_m^2$$

for some constant $c > 0$ independent of m. This concludes the proof of the bound in (10.10).

To estimate the sensitivity in (10.11) notice that

$$(10.28) \qquad h_m^{-1}\nabla U_m(x,0;y,T) = \frac{1}{L}\sum_{p\in B_m} e^{T\hat{\ell}_m(p)}\frac{\sin ph_m}{h_m} e^{ip(y-x)}.$$

and

$$\left| h_m^{-1} U_m(x,0;y,T) - h_{m+1}^{-1} U_{m+1}(x,0;y,T) \right|$$
$$\leq \frac{1}{2L} \left| \sum_{p \in B_m} \left(e^{T\hat{\ell}_m(p)} \frac{\sin ph_m}{h_m} - e^{T\hat{\ell}_{m+1}(p)} \frac{\sin ph_{m+1}}{h_{m+1}} \right) e^{ip(y-x)} \right|$$
$$+ \frac{1}{2L} \left| \sum_{p \in B_{m+1} \setminus B_m} e^{T\hat{\ell}_{m+1}(p)} e^{ip(y-x)} \right|.$$

(10.29)

Let

(10.30)
$$K_m = 2\sqrt{\frac{|\log h_{m+1}|}{\sigma^2 T}}.$$

If h_m is small enough, we have that

$$\frac{1}{2L} \left| \sum_{p \in B_m, |p| \geq K_m} e^{T\hat{\ell}_m(p)} \frac{\sin ph_m}{h_m} e^{ip(y-x)} \right| \leq \frac{1}{2L} \sum_{p \in B_m, |p| \geq K_m} e^{T\Re(\hat{\ell}^m(p))} \frac{\sin ph_m}{h_m}$$
$$\leq c \left| \frac{\sin Kh_m}{h_m} \right| \exp\left(T\sigma^2 \frac{\cos Kh_m - 1}{h_m^2} \right) \leq ch_m^2.$$

(10.31)

where c denotes a generic constant. Similarly

(10.32)
$$\frac{1}{2L} \left| \sum_{p \in B_{m+1}, |p| \geq K_m} e^{T\hat{\ell}_{m+1}(p)} e^{ip(y-x)} \right| \leq ch^2.$$

If m is large enough, we also find

$$\frac{1}{2L} \left| \sum_{p \in B_m, |p| \leq K_m} \left(\frac{\sin ph_m}{h_m} e^{T\hat{\ell}_m(p)} - \frac{\sin ph_{m+1}}{h_{m+1}} e^{T\hat{\ell}_{m+1}(p)} \right) e^{ip(y-x)} \right|$$
$$\leq \frac{1}{2L} \sum_{p \in B_m, |p| \leq K_m} \left| \frac{\sin ph_m}{h_m} \right| e^{-\frac{1}{4}p^2} \left(e^{\frac{\mu T}{4} h_m^2 |p|^3 + \frac{\sigma^2 T}{16} h_m^2 p^4} - 1 \right)$$
$$+ e^{-\frac{1}{4}p^2} \left| \frac{\sin ph_{m+1}}{h_{m+1}} - \frac{\sin ph_m}{h_m} \right| \leq ch_m^2$$

(10.33)

for some constant $c > 0$ independent of m. This concludes the proof of the bound in (10.11). The bound in (10.12) can be derived in a similar way.

Finally, consider the following Fourier representation for the discretized kernel

(10.34)
$$h_m^{-1} U_m^{\delta t}(x,0;y,T) = \frac{1}{L} \sum_{p \in B_m} \left(1 + \delta t \hat{\ell}_m(p) \right)^{\frac{T}{\delta t}} e^{ip(y-x)}.$$

Consider the formula

(10.35)
$$\left(1 + \delta t \hat{\ell}_m(p) \right)^{\frac{T}{\delta t}} = \exp\left(T \log\left(1 + \hat{\ell}_m(p) \right) \right).$$

and let's represent the difference between kernels in (10.13) as follows:

$$|h_m^{-1}U_m(x,0;y,T) - h_m^{-1}U_m^{\delta t}(x,0;y,T)|$$

$$\leq \frac{1}{2L}\left| \sum_{p\in B_m} \left(\exp\left(T\hat{\ell}_m(p)\right) - \exp\left(\frac{T}{\delta t}\log\left(1+\delta t\hat{\ell}_m(p)\right)\right) e^{ip(y-x)} \right) \right|$$

$$\leq \frac{1}{2L} \sum_{p\in B_m, p\leq K_m} e^{-\frac{1}{4}p^2}\left| \exp\left(\frac{T}{\delta t}\log\left(1+\delta t\hat{\ell}_m(p)\right) - T\hat{\ell}_m(p)\right) - 1 \right|$$

$$\frac{1}{2L}\sum_{p\in B_m, p\geq K_m} \left| \exp\left(T\hat{\ell}_m(p)\right) \right| + \frac{1}{2L}\sum_{p\in B_m, p\geq K_m} \left| \exp\left(\frac{T}{\delta t}\log\left(1+\delta t\hat{\ell}_m(p)\right)\right) \right|$$

(10.36)

where K_m is chosen as in (10.19). The very same bounds above lead to the conclusion that this difference is $\leq ch_m^2$.

$\square$

11. Kernel Convergence Estimates for Diffusion Processes

Diffusions are a particularly important class of Markov processes which generalize Brownian motions to allow for space dependent drifts and volatility. In this section we find kernel convergence estimates for the one-dimensional case following (Albanese 2007b).

Let $A = (A_m), m = m_0, m_0+1, ...$ be a simplicial sequence converging to an interval $\lim_{m\to\infty} A_m = [-L, L] \subset \mathbb{R}$, where $0 < L < \infty$. Let $\mu(x)$ and $\sigma(x)$ be smooth functions defined in a neighborhood of $[-L, L]$. The generator of a diffusion on A_m has the form

$$(11.1) \qquad\qquad \mathcal{L}_m = \mu(x)\nabla_{h_m} + \frac{1}{2}\sigma(x)^2\Delta_{h_m}.$$

We assume that m_0 is large enough so that

$$(11.2) \qquad\qquad \frac{\sigma^2(x)}{2h_m^2} > \frac{|\mu(x)|}{2h_m}$$

for all $m \geq m_0$ and all $x \in A_m$. The definition of the generators at boundary points can be extended by imposing one of the boundary conditions in the previous section, i.e. reflecting, absorbing, periodic or mixed.

Theorem 17. (Convergence Estimates for Diffusions.) *Consider a diffusion process defined as in (11.1) where $\mu(y)$ and $\sigma(y)$ are smooth functions satisfying the bounds in (11.2) for all $m \geq m_0$. Assume that boundary conditions are either periodic or absorbing. Then there is a constant $c > 0$ such that*

$$(11.3) \qquad\qquad |h_m^{-1}U_m(x,0;y,T) - h_{m'}^{-1}U_{m'}(x,0;y,T)| \leq ch_m^2$$

for all $m' \geq m$ and all $y \in A_m$.

It suffices to establish the above inequality in the case $m' = m + 1$. In fact, given this particular case, the general statement can be derived with an iterative argument. Let $h = h_{m+1}$ so that $h_m = 2h$.

Recall that the path integral representation for a diffusion process has the form in equation (11.4). In the special case of a time-homogeneous process, this expansion reads as follows:

$$(11.4) \qquad U_m(x,0;y,T) = \sum_{q=1}^{\infty} 2^{-q} \sum_{\substack{\gamma\in\Gamma_m : \gamma_0 = x, \gamma_q = y \\ |\gamma_j - \gamma_{j-1}| = 1 \ \forall j\geq 1}} W_m(\gamma, q, T)$$
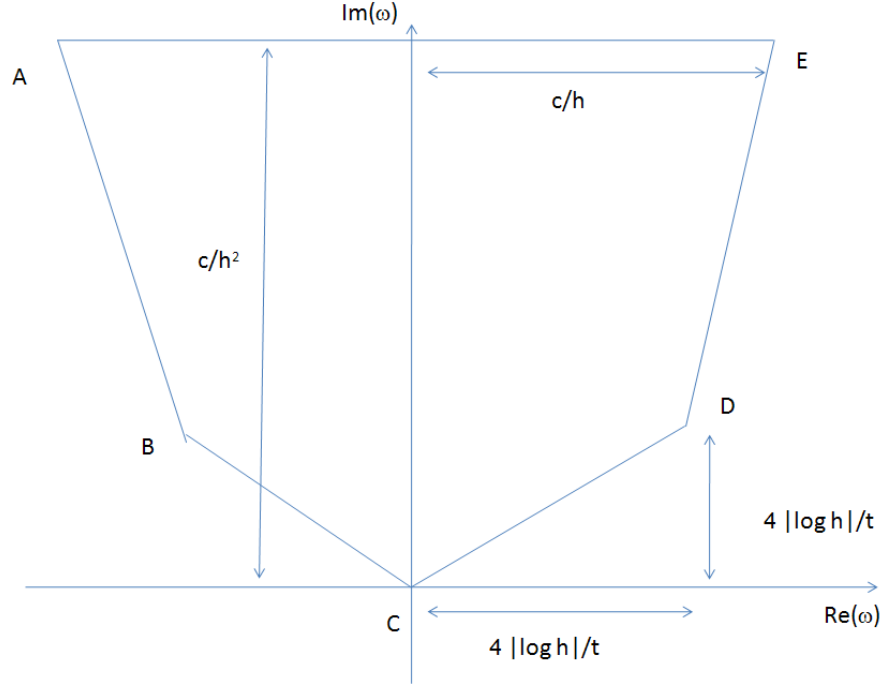
FIGURE 1. Contour of integration for the integral in (11.9). $\mathcal{C}_+$ is the countour joining the point D to the points E, A, B. $\mathcal{C}_-$ is the countour joining the point B to C to D.

where

(11.5)

$$W_m(\gamma, q, T) = \int_0^T ds_1 \int_{s_1}^T ds_2 ... \int_{s_{q-1}}^T ds_q e^{(T-s_q)\mathcal{L}_m(y,y)} \prod_{j=0}^{q-1} \left(e^{(s_{j+1}-s_j)\mathcal{L}_m(\gamma_j,\gamma_j)} \mathcal{L}_m(\gamma_j, \gamma_{j+1}) \right)$$

where $s_0 = 0$.

Let us introduce the following two constants characterizing the volatility function:

(11.6) $$\Sigma_0 = \inf_{x \in A_m} \sigma(x), \quad \Sigma_1 = \sup_{x \in A_m} \sqrt{\sigma(x)^2 + h_m|\mu(x)|}.$$

and let

(11.7) $$M = \sup_{x \in A_m} |\mu(x)|.$$

Since our interval is bounded, we have that $\Sigma_0 > 0$ and $\Sigma_1, M < \infty$.

Let us introduce the following Green's function:

(11.8) $$G_m(x, y; \omega) = \int_0^\infty U_m(x, 0; y, t)e^{-i\omega t}dt = \frac{1}{\mathcal{L} + i\omega}(x, y).$$

The propagator can be expressed as the following contour integral

(11.9) $$U_m(x, 0; y, T) = \int_{\mathcal{C}_-} \frac{d\omega}{2\pi} G_m(x, y; \omega)e^{i\omega T} + \int_{\mathcal{C}_+} \frac{d\omega}{2\pi} G_m(x, y; \omega)e^{i\omega T}.$$

Here, $\mathcal{C}_+$ is the contour joining the point D to the points E, A, B in Fig. 1, while $\mathcal{C}_-$ is the contour joining the point B to C to D. By design, each point ω on the upper path $\mathcal{C}_+$ is separated from the spectrum of $\mathcal{L}$.

Lemma 1. *For m sufficiently large, there is a constant $c > 0$ such that*

$$(11.10) \qquad \left| \int_{\mathcal{C}_+} \frac{d\omega}{2\pi} G_m(x, y; \omega) e^{i\omega T} \right| \leq ch^3.$$

Proof. The proof is based on the geometric series expansion

$$(11.11) \qquad G_m(\omega) = \frac{1}{\mathcal{L}^m + i\omega} = \sum_{j=0}^{\infty} \frac{1}{\frac{1}{2}\sigma^2 \Delta^m + i\omega} \left[\mu \nabla^m \frac{1}{\frac{1}{2}\sigma^2 \Delta^m + i\omega} \right]^j$$

whose convergence for $\omega \in \mathcal{C}_+$ can be established by means of a Kato-Rellich relative bound, see (Kato 1976). More precisely, for any $\alpha > 0$, one can find a $\beta > 0$ such that the operators ∇^m and Δ^m satisfy the following relative bound estimate:

$$(11.12) \qquad ||\nabla^m f||_2 \leq \alpha ||\Delta^m f||_2 + \beta ||f||_2.$$

for all periodic functions f and all $m \geq m_0$. This bound can be derived by observing that ∇^m and Δ^m can be diagonalized simultaneously by a Fourier transform, as done in the previous section, and by observing that for any $\alpha > 0$, one can find a $\beta > 0$ such that

$$(11.13) \qquad \left| \frac{\sin h_m p}{h_m} \right| \leq \alpha \left| \frac{\cos h_m p - 1}{h_m^2} \right| + \beta$$

for all $m \geq m_0$ and all $p \in B_m$.

Under the same conditions, we also have that

$$(11.14) \qquad ||\mu \nabla^m f||_2 \leq \frac{2M\alpha}{\Sigma_0^2} \left| \left| \frac{1}{2}\sigma^2 \Delta^m f \right| \right|_2 + \beta ||f||_2.$$

Hence

$$(11.15) \quad \left| \left| \mu \nabla^m \frac{1}{\frac{1}{2}\sigma^2 \Delta^m + i\omega} f \right| \right|_2 \leq \frac{2M\alpha}{\Sigma_0^2} \left| \left| \frac{1}{2}\sigma^2 \Delta^m \frac{1}{\frac{1}{2}\sigma^2 \Delta^m + i\omega} f \right| \right|_2 + \beta \left| \left| \frac{1}{\frac{1}{2}\sigma^2 \Delta^m + i\omega} f \right| \right|_2 < 1$$

where the last inequality holds if $\omega \in \mathcal{C}_+$, if α is chosen sufficiently small and if m is large enough. In this case, the geometric series expansion converges in (11.11) converges in L^2 operator norm. The uniform norm of the kernel $|G_m(x, y; \omega)|$ is pointwise bounded from above by h_m^{-1}.

Since the points B and D have imaginary part equal at height $4\frac{|\log h_m|}{T}$, the integral over the contour $\mathcal{C}_+$ converges also and is bounded from above by ch_m^3 in uniform norm.

$\square$

Lemma 2. *If $q \geq \frac{e^2 \Sigma_1^2 T}{2h_m^2}$ we have that*

$$(11.16) \qquad W_m(\gamma, q; 0, T) \leq \sqrt{\frac{q}{2\pi}} \exp\left(-\frac{\Sigma_0^2 T}{2} - q \right).$$

Proof. Let us define the function

$$(11.17) \qquad \phi(t) = \frac{\Sigma_1^2}{2h_m^2} e^{-\frac{\Sigma_0^2 t}{2h_m^2}} 1(t \geq 0)$$

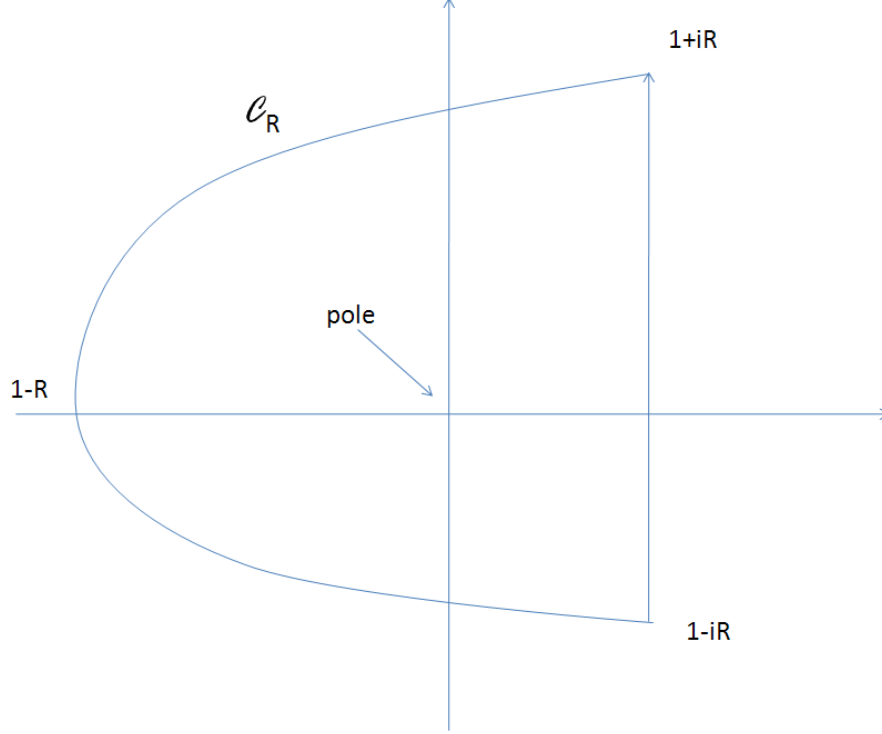where $1(t \geq 0)$ is the characteristic function of $\mathbb{R}_+$. We have that

$$(11.18) \qquad W_m(\gamma, q; 0, T) \leq \phi^{\star q}(T)$$

where $\phi^{\star q}$ is the q−th convolution power, i.e. the q−fold convolution product of the function ϕ by itself. The Fourier transform of $\phi(t)$ is given by

$$(11.19) \qquad \hat{\phi}(\omega) = \frac{\Sigma_1^2}{2h_m^2} \int_0^{\infty} e^{-i\omega t - \frac{\Sigma_0^2 t}{2h_m^2}} dt = \frac{\Sigma_1^2}{2i\omega h_m^2 + \Sigma_0^2}.$$

The convolution power is given by the following inverse Fourier transform:

$$(11.20) \qquad \phi^{\star q}(T) = \int_0^{\infty} \hat{\phi}(\omega)^q e^{i\omega T} = \left(\frac{\Sigma_1}{\Sigma_0} \right)^{2q} \int_{-\infty}^{\infty} \left(1 + \frac{2i\omega h_m^2}{\Sigma_0^2} \right)^{-q} e^{i\omega T} \frac{d\omega}{2\pi}.$$

FIGURE 2. Contour of integration $\mathcal{C}_R$ for the integral in (11.21).

Introducing the new variable $z = 1 + \frac{2i\omega h_m^2}{\Sigma_0^2}$, the integral can be recast as follows

$$(11.21) \qquad \phi^{\star q}(T) = \frac{\Sigma_0^{2-2q}\Sigma_1^{2q}}{4\pi i h_m^2} \lim_{R \to \infty} \int_{\mathcal{C}_R} z^{-q} \exp\left(\frac{\Sigma_0^2 T}{2h_m^2}(z-1)\right) dz$$

where $\mathcal{C}_R$ is the contour in Fig. 2. Using the residue theorem and noticing that the only pole of the integrand is at $z = 0$, we find

$$(11.22) \qquad \phi^{\star q}(T) = \frac{1}{(q-1)!}\left(\frac{\Sigma_1^2 T}{2h_m^2}\right)^q \exp\left(\frac{-\Sigma_0^2 T}{2h_m^2}\right).$$

Making use of Stirling's formula $q! \approx \sqrt{2\pi} q^{q+\frac{1}{2}} e^{-q}$, we find

$$(11.23) \qquad \phi^{\star q}(T) \approx \sqrt{\frac{q}{2\pi}} \exp\left(-\frac{\Sigma_0^2 T}{2h_m^2} + q\log\frac{\Sigma_1^2 T}{2h_m^2} + q(1 - \log q)\right).$$

If $\log q \geq \log\frac{\Sigma_1^2 T}{2h_m^2} + 2$, then we arrive at the bound in (11.16).

$\square$

Definition 41. (Decorating Paths.) *Let $m \geq m_0$ and let $\gamma = \{y_0, y_1, y_2, ...\}$ be a symbolic sequence in Γ_m. A decorating path around γ is defined as a symbolic sequence $\gamma' = \{y_0, y_1', y_2', ...\}$ with $y_i' \in h_{m+1}\mathbb{Z}$ containing the sequence γ as a subset and such that if $y_j' = y_i$ and $y_k' = y_{i+1}$, then all elements y_n' with $j < n < k$ are such that $|y_n' - y_j'| \leq h_{m+1}$. Let $\mathcal{D}_{m+1}(\gamma)$ be the set of all decorating sequences around γ. The decorated weights are defined as follows:*

$$(11.24) \qquad \tilde{W}_m(\gamma, q; 0, T) = \sum_{q'=q}^{\infty} \sum_{\substack{\gamma' \in \mathcal{D}_{m+1}(\gamma) \\ \gamma'_{q'} = \gamma_q}} W_{m+1}(\gamma', q'; 0, T).$$

Finally, let us introduce also the following Fourier transform:

$$(11.25) \quad \hat{W}_m(\gamma, q; \omega) = \int_0^\infty W_m(\gamma, q; 0, t) e^{i\omega t} dt, \quad \hat{\tilde{W}}_m(\gamma, q; \omega) = \int_0^\infty \tilde{W}_m(\gamma, q; 0, t) e^{i\omega t} dt.$$

Definition 42. (Notations.) *In the following, we set $h = h_{m+1}$ so that $h_m = 2h$. We also use the Landau notation $O(h^n)$ to indicate a function $f(h)$ such that $h^{-n} f(h)$ is bounded in a neighborhood of (0).*

Lemma 3. *Let $x, y \in A_m$ and let $\mathcal{C}_-$ be an integration contour as in Fig. 1. Then*

$$(11.26) \quad \left| \left(\int_{\mathcal{C}_-} 2G_{m+1}(x, y; \omega) - G_m(x, y; \omega) \right) e^{i\omega T} \frac{d\omega}{2\pi} \right| = O(h^3).$$

Proof. We have that

$$(11.27)$$

$$2G_{m+1}(x, y; \omega) - G_m(x, y; \omega) = \sum_{q=1}^\infty 2^{-q} \sum_{\substack{\gamma \in \Gamma_m : \gamma_0 = x, \gamma_q = y \\ |\gamma_j - \gamma_{j-1}| = 1 \forall j \geq 1}} \left(2\hat{\tilde{W}}_m(\gamma, q; \omega) - \hat{W}_m(\gamma, q; \omega) \right).$$

The number of paths over which the summation is extended is

$$(11.28) \quad N(\gamma, q; x, y) \equiv \sharp\{\gamma \in \Gamma_m : \gamma_0 = x, \gamma_q = y, |\gamma_j - \gamma_{j-1}| = 1 \forall j \geq 1\} = \binom{q}{\frac{q}{2} + k}$$

where $k = \frac{|y - x|}{h_m}$. Applying Stirling's formula we find

$$(11.29) \quad N_\gamma \lesssim 2^q \sqrt{\frac{2}{\pi q}}.$$

Hence

$$\left| \int_{\mathcal{C}_-} \left(2G_{m+1}(x, y; \omega) - G_m(x, y; \omega) \right) e^{i\omega T} \frac{d\omega}{2\pi} \right|$$

$$\leq c \sum_{q=1}^\infty \sqrt{\frac{1}{q}} \max_{\substack{\gamma \in \Gamma_m : \gamma_0 = x, \gamma_q = y \\ |\gamma_j - \gamma_{j-1}| = 1 \forall j \geq 1}} \left| \int_{\mathcal{C}_-} \left(2\hat{\tilde{W}}_m(\gamma, q; \omega) - \hat{W}_m(\gamma, q; \omega) \right) e^{i\omega T} \frac{d\omega}{2\pi} \right|.$$

$$(11.30)$$

for some constant $c \approx \sqrt{\frac{2}{\pi}} > 0$. It suffices to extend the summation over q only up to

$$(11.31) \quad q_{\max} \equiv \frac{e^2 \Sigma_1^2 T}{2h^2}.$$

To resum beyond this threshold, one can use the previous lemma. More precisely, we have that

$$\left| \int_{\mathcal{C}_-} \left(2G_{m+1}(x, y; \omega) - G_m(x, y; \omega) \right) e^{i\omega T} \frac{d\omega}{2\pi} \right|$$

$$\leq c \sqrt{q_{\max}} \max_{\substack{q, \gamma \in \Gamma_m : \gamma_0 = x, \gamma_q = y \\ |\gamma_j - \gamma_{j-1}| = 1 \forall j \geq 1}} \left| \int_{\mathcal{C}_-} \left(2\hat{\tilde{W}}_m(\gamma, q; \omega) - \hat{W}_m(\gamma, q; \omega) \right) e^{i\omega T} \frac{d\omega}{2\pi} \right|.$$

$$(11.32)$$

Let $v(x) = \sigma(x)^2$. To evaluate the resummed weight function, let us form the matrix

$$(11.33) \quad \bar{\mathcal{L}}(x; h) = \begin{pmatrix} -\frac{v(x+h)}{h^2} & \frac{v(x+h)}{2h^2} - \frac{\mu(x+h)}{2h} & 0 \\ \frac{v(x)}{2h^2} + \mu(x)/(2h) & -\frac{v(x)}{h^2} & \frac{v(x)}{2h^2} - \frac{\mu(x)}{2h} \\ 0 & \frac{v(x-h)}{2h^2} + \frac{\mu(x-h)}{2h} & -\frac{v(x-h)}{h^2} \end{pmatrix}$$

and decompose it as follows:

$$(11.34) \qquad \bar{\mathcal{L}}(x;h) = \frac{1}{h^2}\bar{\mathcal{L}}_0(x) + \frac{1}{h}\bar{\mathcal{L}}_1(x) + \bar{\mathcal{L}}_2(x) + h\bar{\mathcal{L}}_3(x) + O(h^2).$$

where

$$(11.35) \qquad \bar{\mathcal{L}}_0(x) = \begin{pmatrix} -v(x) & \frac{1}{2}v(x) & 0 \\ \frac{1}{2}v(x) & -v(x) & \frac{1}{2}v(x) \\ 0 & \frac{1}{2}v(x) & -v(x) \end{pmatrix},$$

$$(11.36) \qquad \bar{\mathcal{L}}_1(x) = \begin{pmatrix} -v'(x) & \frac{1}{2}v'(x) - \frac{1}{2}\mu(x) & 0 \\ \frac{1}{2}\mu(x) & 0 & -\frac{1}{2}\mu(x) \\ 0 & -\frac{1}{2}v'(x) + \frac{1}{2}\mu(x) & v'(x) \end{pmatrix},$$

$$(11.37) \qquad \bar{\mathcal{L}}_2(x) = \begin{pmatrix} -\frac{1}{2}v''(x) & \frac{1}{4}v''(x) - \frac{1}{2}\mu'(x) & 0 \\ 0 & 0 & 0 \\ 0 & \frac{1}{4}v''(x) - \frac{1}{2}\mu'(x) & -\frac{1}{2}v''(x) \end{pmatrix}.$$

and

$$(11.38) \qquad \bar{\mathcal{L}}_3(x) = \begin{pmatrix} -\frac{1}{6}v'''(x) & \frac{1}{12}v'''(x) - \frac{1}{4}\mu''(x) & 0 \\ 0 & 0 & 0 \\ 0 & -\frac{1}{12}v'''(x) + \frac{1}{4}\mu''(x) & \frac{1}{6}v'''(x) \end{pmatrix}.$$

Let us introduce the sign variable $\tau = \pm 1$, the functions

$$(11.39) \qquad \phi_0(t,x,\tau) \equiv 2\mathcal{L}_m(x, x+2\tau h)e^{t\mathcal{L}_m(x,x)}1(t \geq 0)$$

$$(11.40) \qquad \phi_1(t,x,\tau) \equiv 2\mathcal{L}_{m+1}(x+\tau h, x+2\tau h)e^{t\bar{\mathcal{L}}(x;h)}(x, x+\tau h)1(t \geq 0)$$

and their Fourier transforms

$$\hat{\phi}_0(\omega,x,\tau) = \left(\frac{v(x)}{4h^2} + \tau\frac{\mu(x)}{2h}\right)\left(\frac{v(x)}{4h^2} + i\omega\right)^{-1}$$

$$\hat{\phi}_1(\omega,x,\tau) = \left(\frac{v(x)}{h^2} + \tau\frac{\mu(x)+v'(x)}{h} + \frac{v''(x)+\mu'(x)}{2} + \left(\frac{v'''(x)}{6} + \frac{\mu''(x)}{2}\right)\tau h + O(h^2)\right)$$

$$(11.41)$$

$$< x|\left(-\bar{\mathcal{L}}(x;h) + i\omega\right)^{-1}|x+\tau h > .$$

where

$$(11.42) \qquad |x> = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \text{and} \quad |x+\tau h> = \begin{pmatrix} \delta_{\tau,1} \\ 0 \\ \delta_{\tau,-1} \end{pmatrix}.$$

We also require the functions

$$(11.43) \qquad \psi_0(t,x) \equiv e^{t\mathcal{L}_m(x,x)}1(t \geq 0), \qquad \psi_1(t,x) \equiv e^{t\bar{\mathcal{L}}(y;h)}(x,x)1(t \geq 0)$$

and the corresponding Fourier transforms

$$(11.44) \qquad \hat{\psi}_0(\omega,x) = \left(\frac{v(x)}{4h^2} + i\omega\right)^{-1}, \quad \hat{\psi}_1(\omega,x) = < x|\left(-\bar{\mathcal{L}}(x;h) + i\omega\right)^{-1}|x> .$$

If γ is a symbolic sequence, then

$$(11.45) \qquad \hat{W}_m(\gamma,q;\omega) = \hat{\psi}_0(\omega,\gamma_q)\prod_{j=0}^{q-1}\hat{\phi}_0(\omega;\gamma_j, \text{sgn}(\gamma_{j+1} - \gamma_j))$$

$$(11.46) \qquad \hat{\tilde{W}}_m(\gamma,q;\omega) = \hat{\psi}_1(\omega,\gamma_q)\prod_{j=0}^{q-1}\hat{\phi}_1(\omega;\gamma_j, \text{sgn}(\gamma_{j+1} - \gamma_j)).$$

Let us estimate the difference between the functions $\hat{\phi}_1(\omega, x, \tau)$ and $\hat{\phi}_2(\omega, x, \tau)$ assuming that ω is in the contour $\mathcal{C}_-$ in Fig. 2. Retaining only terms up to order up to $O(h^3)$, we find

$$\hat{\phi}_0(\omega, x, \tau) = 1 + \frac{2\mu(x)\tau h}{v(x)} - \frac{4i\omega h^2}{v(x)} - 8\mu(x)\frac{i\omega\tau h^3}{v(x)^2} - \frac{16\omega^2 h^4}{v(x)^2} + O(h^5).$$

(11.47)

A lengthy but straightforward calculation which is best carried out using a symbolic manipulation program, gives

$$\hat{\phi}_1(\omega, x, \tau) = 1 + \frac{2\mu(x)\tau h}{v(x)} - \frac{4i\omega h^2}{v(x)} - \big[8\mu(x) - v'(x)\big]\frac{i\omega\tau h^3}{v(x)^2}$$
$$+ r(x) \cdot h^3\tau + i\omega h^4 p(x) - \frac{14\omega^2 h^4}{v(x)^2} + O(h^5)$$

(11.48)

where

$$r(x) = \frac{1}{2v(x)^3}\Big[\mu''(x)v(x) - 4\mu(x)^3 + 2v'(x)\mu(x)^2 - 2v'(x)v(x)\mu'(x)$$
$$- \big(2\mu(x)\mu'(x) + v''(x)v(x) - 2v'(x)^2\big)\mu(x)\Big].$$

$$p(x) = \frac{1}{v(x)^3}\Big[4\mu(x)^2 - 2v'(x)\mu(x) + 4v(x)\mu'(x) + v''(x)v(x) - 2v'(x)^2\Big].$$

(11.49)

We have that

$$\sum_{j=0}^{q-1}\Big(\log\hat{\phi}_0(\omega; \gamma_j, \operatorname{sgn}(\gamma_{j+1} - \gamma_j)) - \log\hat{\phi}_1(\omega; \gamma_j, \operatorname{sgn}(\gamma_{j+1} - \gamma_j))\Big)$$

$$= \sum_{j=0}^{q-1}\Big(\frac{i\omega v'(\gamma_j)}{v(\gamma_j)^2} + r(\gamma_j)\Big)h^3\operatorname{sgn}(\gamma_{j+1} - \gamma_j) + \big(|\omega|\|p\|_\infty + 2|\omega|^2\|v^{-2}\|_\infty\big)O(h^4 q)$$

$$= i\omega h^2\log\Big(\frac{v(\gamma_q)}{v(\gamma_0)}\Big) + h^2\big(R(\gamma_q) - R(\gamma_0)\big) + \big(|\omega|\|p\|_\infty + 2|\omega|^2\|v^{-2}\|_\infty\big)O(h^4 q)$$

(11.50)

where $R(x)$ is a primitive of $r(x)$, i.e.

(11.51)
$$R(x) = \int^x r(z)dz.$$

We conclude that there is a constant $c > 0$ such that

(11.52)
$$\left|\int_{\mathcal{C}_-}\Big(\prod_{j=0}^{q-1}\hat{\phi}_0(\omega; \gamma_j, \operatorname{sgn}(\gamma_{j+1} - \gamma_j)) - \prod_{j=0}^{q-1}\hat{\phi}_1(\omega; \gamma_j, \operatorname{sgn}(\gamma_{j+1} - \gamma_j))\Big)e^{i\omega T}\frac{d\omega}{2\pi}\right| \le ch^2.$$

for all $q \le q_{\max}$. Here we use the decay of $e^{i\omega T}$ in the upper half of the complex ω plane to offset the ω dependencies in the integrand. Similar calculations lead to the following expansions:

(11.53)
$$\hat{\psi}_0(\omega, y) = \frac{4h^2}{v(y)} + O(\omega h^4), \quad \hat{\psi}_1(\omega, y) = \frac{2h^2}{v(y)} + O(\omega h^4) = \frac{1}{2}\hat{\psi}_0(\omega, y) + O(\omega h^4).$$

Since $q < ch^{-2}$ and $\omega \le |\log h|$, we find

(11.54)
$$\left|\int_{\mathcal{C}_-}\Big(2G_{m+1}(x, y; \omega) - G_m(x, y; \omega)\Big)e^{i\omega T}\frac{d\omega}{2\pi}\right| \le cq_{\max}^{\frac{1}{2}}h^4 \le ch^3.$$

This completes the proof of the Lemma and of the Theorem. $\qquad\square$

12. Convergence of Time Discretization Schemes

In this section we analyze the convergence of the time discretized kernel that is obtained by means of fast exponentiation.

Let $A = (A_m), m = m_0, m_0 + 1, ...$ be a simplicial sequence converging to an interval $\lim_{m \to \infty} A_m = [0, L] \subset \mathbb{R}$, where $0 < L < \infty$. Let $\mu(x)$ and $\sigma(x)$ be functions on $[0, L]$ satisfying all the conditions in the previous section and consider the generator of a diffusion on A_m of the form

$$(12.1) \qquad \mathcal{L}_m = \mu(x)\nabla_{h_m} + \frac{1}{2}\sigma(x)^2 \Delta_{h_m}.$$

Theorem 18. (Convergence Estimates for Fast Exponentiation.) *Let $\delta t > 0$ and consider the discretized kernel*

$$(12.2) \qquad U_m^{\delta t}(x, 0; y, T) = (1 + \delta t \mathcal{L}_m)^{\frac{T}{\delta t}} (x, 0; y, T).$$

where $\mathcal{L}_m$ is the operator in (12.1) and δt_m is so small that

$$(12.3) \qquad \min_{x \in A_m} 1 + \delta t_m \mathcal{L}_m(x, x) > 0$$

Assume that boundary conditions are either periodic or absorbing and that the ratio $\frac{T}{\delta t} = N$ is an integer. Then there is a constant $c > 0$ such that

$$(12.4) \qquad |h_m^{-1} U_m(x, 0; y, T) - h_m^{-1} U_m^{\delta t}(x, 0; y, T)| \leq ch_m^2$$

for all $m \geq m_0$ and all $y \in A_m$.

Proof. A Dyson expansion can also be obtained for the time-discretized kernel and has the form

$$U_m^{\delta t}(y_1, 0; y_2, T) = \sum_{q=1}^{\infty} \sum_{\gamma \in \Gamma_m : \gamma_0 = x, \gamma_q = y} \sum_{k_1=1}^{N} \sum_{k_2=k_1+1}^{N} \cdots \sum_{k_q=k_{q-1}+1}^{N}$$

$$(12.5) \qquad \left(1 + \delta t \mathcal{L}_m(\gamma_0, \gamma_0)\right)^{k_1 - 1} (\delta t)^q \prod_{j=1}^{q} \mathcal{L}_m(\gamma_{j-1}, \gamma_j)\left(1 + \delta t \mathcal{L}_m(\gamma_j, \gamma_j)\right)^{k_{j+1} - k_j - 1}$$

where $t_{q+1} = T$ and $k_{q+1} = N$. In this case, the propagator can be expressed through a Fourier integral as follows:

$$(12.6) \qquad U_m^{\delta t}(y_1, 0; y_2, T) = \int_{-\frac{\pi}{\delta t}}^{\frac{\pi}{\delta t}} G_m^{\delta t}(y_1, y_2; \omega) e^{i\omega t} \frac{d\omega}{2\pi}$$

where

$$(12.7) \qquad G_m^{\delta t}(y_1, y_2; \omega) = \delta t \sum_{j=0}^{\frac{T}{\delta t}} U_m^{\delta t}(y_1, 0; y_2, j\delta t) e^{-i\omega j \delta t}.$$

The propagator can also be represented as the limit

$$(12.8) \qquad U_m^{\delta t}(y_1, 0; y_2, T) = \lim_{H \to \infty} \int_{\mathcal{C}_H} G_m^{\delta t}(y_1, y_2; \omega) e^{i\omega t} \frac{d\omega}{2\pi}$$

where $\mathcal{C}_H$ is the contour in Fig. 3. This is due to the fact that the integral along the segments BC and DA are the negative of each other, while the integral over CD tends to zero exponentially fast as $\Im(\omega) \to \infty$, where $\Im(\omega)$ is the imaginary part of ω. Using Cauchy's theorem, the contour in Fig. 3 can be deformed into the contour in Fig. 1. To estimate the discrepancy between the time-discretized kernel and the continuous time one, one can thus compare the Green's function along such contour. Again, the only arc that requires detailed attention is the arc BCD, as the integral over rest of the contour of integration can be bounded from above as in the previous section.

Let $h = h_m$ and let us introduce the two functions

$$(12.9) \qquad \phi_0(t, x, \tau) \equiv 2\mathcal{L}_m(x, x + \tau h)e^{t\mathcal{L}_m(x,x)} 1(t \geq 0),$$

$$(12.10) \qquad \phi_{\delta t}(j, x, \tau) \equiv 2\mathcal{L}_m(x, x + \tau h)\left(1 + \delta t \mathcal{L}_m(x, x)\right)^{j-1}.$$
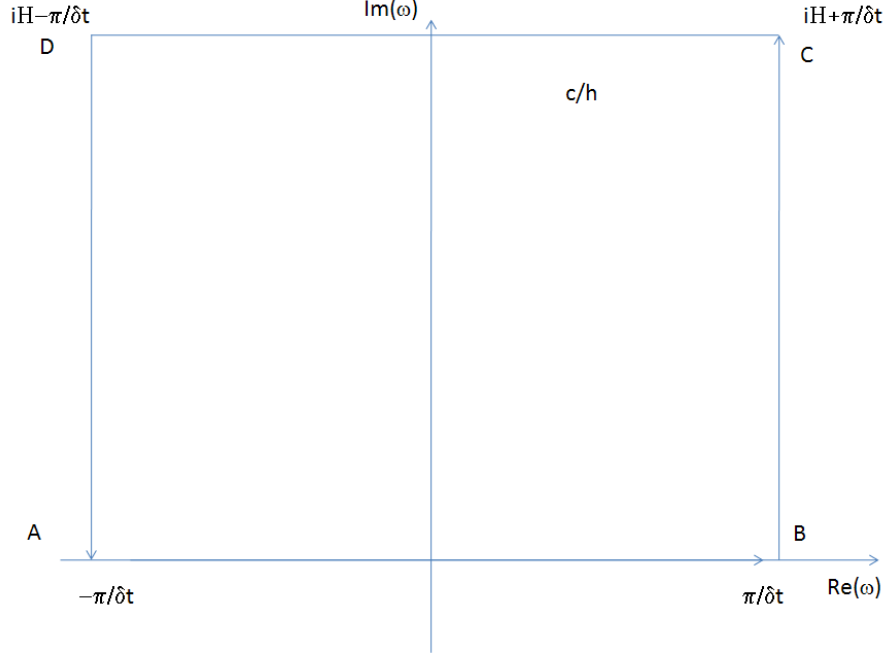
FIGURE 3. Contour of integration for the integral in (12.8).

and the corresponding Fourier transforms

$$(12.11) \qquad \hat{\phi}_0(\omega, x, \tau) = \int_0^\infty \phi_0(t, x, \tau) e^{-i\omega t} \frac{d\omega}{2\pi} = \left(\frac{v(x)}{h^2} + \tau \frac{\mu(x)}{h} \right) \left(\frac{v(x)}{h^2} + i\omega \right)^{-1}$$

$$(12.12) \qquad \hat{\phi}_{\delta t}(\omega, x, \tau) = \sum_{j=0}^{\frac{t}{\delta t}} \phi_{\delta t}(j, x, \tau) e^{-i\omega j \delta t} = \left(\frac{v(x)}{h^2} + \tau \frac{\mu(x)}{h} \right) \left(e^{i\omega \delta t} - 1 + \delta t \frac{v(x)}{h^2} \right)^{-1}.$$

We have that

$$\hat{\phi}_{\delta t}(\omega, x, \tau) = \left(\frac{v(x)}{h^2} + \tau \frac{\mu(x)}{h} \right) \left(i\omega + \frac{v(x)}{h^2} - \frac{\omega^2}{2} \delta t + O(\delta t^2) \right)^{-1}$$

$$(12.13) \qquad\qquad = \hat{\phi}_0(\omega, x, \tau) + \frac{\omega^2}{2v(x)} h^2 \delta t + O(h^2 \delta t^2). = \hat{\phi}_0(\omega, x, \tau) + O(h^4),$$

where the last step uses the fact that $\delta t = O(h^2)$.

Let us also introduce the functions

$$(12.14) \qquad \psi_0(t, x, \tau) \equiv e^{t \mathcal{L}_m(x,x)} 1(t \geq 0), \qquad \psi_{\delta t}(j, x, \tau) \equiv \sum_{k=1}^{j} \left(1 + \delta t \mathcal{L}_m(x, x) \right)^{j-1}.$$

and the corresponding Fourier transforms

$$(12.15) \qquad \hat{\psi}_0(\omega, x, \tau) = \left(\frac{v(x)}{h^2} + i\omega \right)^{-1}, \qquad \hat{\psi}_{\delta t}(\omega, x, \tau) = \left(e^{i\omega \delta t} - 1 + \delta t \frac{v(x)}{h^2} \right)^{-1}.$$

Again we find that

$$(12.16) \qquad\qquad\qquad \hat{\psi}_0(\omega, x, \tau) = \hat{\psi}_{\delta t}(\omega, x, \tau) + O(h^4).$$

If γ is a symbolic sequence, then let us set

(12.17)
$$\hat{W}_m(\gamma, q; \omega) = \hat{\psi}_0(\omega, \gamma_q) \prod_{j=0}^{q-1} \hat{\phi}_0(\omega; \gamma_j, \operatorname{sgn}(\gamma_{j+1} - \gamma_j))$$

(12.18)
$$\hat{W}_m^{\delta t}(\gamma, q; \omega) = \hat{\psi}_{\delta t}(\omega, \gamma_q) \prod_{j=0}^{q-1} \hat{\phi}_{\delta t}(\omega; \gamma_j, \operatorname{sgn}(\gamma_{j+1} - \gamma_j)).$$

We have that
(12.19)

$$G_m^{\delta t}(x, y; \omega) - G_m(x, y; \omega) = \sum_{q=1}^{\infty} 2^{-q} \sum_{\substack{\gamma \in \Gamma_m : \gamma_0 = x, \gamma_q = y \\ |\gamma_j - \gamma_{j-1}| = 1 \forall j \geq 1}} \left(\hat{W}_m^{\delta t}(\gamma, q; \omega) - \hat{W}_m(\gamma, q; \omega) \right).$$

The integration over the contour in Fig. 1 can again be split into an integration over the countour $\mathcal{C}_-$ and an integration over $\mathcal{C}_+$. The integral over $\mathcal{C}_+$ can be bounded from above thanks to Lemma 3. Furthermore, we have that

$$\left| \int_{\mathcal{C}_-} \left(G_m^{\delta t}(x, y; \omega) - G_m(x, y; \omega) \right) e^{i\omega T} \frac{d\omega}{2\pi} \right|$$

$$\leq c\sqrt{q_{\max}} \max_{\substack{q, \gamma \in \Gamma_m : \gamma_0 = x, \gamma_q = y \\ |\gamma_j - \gamma_{j-1}| = 1 \forall j \geq 1}} \left| \int_{\mathcal{C}_-} \left(\hat{W}_m^{\delta t}(\gamma, q; \omega) - \hat{W}_m(\gamma, q; \omega) \right) e^{i\omega T} \frac{d\omega}{2\pi} \right|.$$

(12.20) $\leq ch^3$

$\square$

13. Hypergeometric Brownian Motion

This section is based on work in collaboration with Joe Campolieti, Peter Carr and Alex Lipton, see (Albanese *et al.* 2001).

In this section we pose the problem of constructing driftless diffusion models which reduce to a given diffusion by means of a combination of a measure change and a coordinate transformation.

Consider a Markov process on the simplicial sequence $A_m \subset \mathbb{R}^d$ with generator $\mathcal{L}_m(x, x'; t)$. Let ρ be a real valued parameter and suppose $f_m^1(x)$ and $f_m^2(x)$ are two linearly independent solutions of the equation

(13.1)
$$\sum_{x' \in A_m} \mathcal{L}_m(x, x'; t) f_m)(x') = \rho f_m(x)$$

for all $x \in Int(A_m)$. Notice that on the boundary $\partial(A_m)$ these equations may fail and, in actual applications, they will indeed as a rule fail. Consider the function

(13.2)
$$g_m(x; t) = e^{-\rho t} \big(c_1 f_m^1(x) + c_2 f_m^2(x) \big)$$

for some choice of constants c_1, c_2 such that this function is strictly positive. This function satisfies the equation

(13.3)
$$\frac{\partial g_m(x; t)}{\partial t} + (\mathcal{L}_m g_m)(x; t) = 0.$$

Hence $g_m(x; t)$ defines a measure change and one can construct a new Markov generator by setting

(13.4)
$$\mathcal{L}_m^g(x, x'; t) = \frac{g_m(x'; t)}{g_m(x; t)} \mathcal{L}_m(x, x'; t) + \frac{1}{g_m(x; t)} \frac{\partial g_m(x; t)}{\partial t} \delta_{xx'}.$$

Consider the linear fractional transformation

(13.5)
$$Y_m(x) = \frac{c_3 f_m^1(x) + c_4 f_m^2(x)}{c_1 f_m^1(x) + c_2 f_m^2(x)}.$$

for some choice of constants c_3, c_4.

Theorem 19. (Linear Fractional Transformations.) *The process $Y(x_t)$ satisfies the martingale condition*

$$(13.6) \qquad \lim_{s \downarrow 0} \frac{1}{s} E_t \left[\frac{Y_m(x_{t+s}) - Y_m(x_t)}{s} \right] = 0.$$

for all $x_t \in Int(A_m)$.

Proof. We have that

$$\lim_{s \downarrow 0} \frac{1}{s} E_t \left[\frac{Y_m(x_{t+s}) - Y_m(x_t)}{s} \right] = \sum_{x' \in A_m} \mathcal{L}_m^g(x, x') \big(Y_m(x') - Y_m(x) \big)$$

$$= \frac{1}{g_m(x,t)} \left(\sum_{x' \in A_m} \mathcal{L}_m(x, x') \big(c_3 f_m^1(x') + c_4 f_m^2(x') \big) \right) + \frac{1}{g_m(x,t)^2} \frac{\partial g_m(x;t)}{\partial t} \big(c_3 f_m^1(x) + c_4 f_m^2(x) \big)$$

$$= e^{\rho t} \frac{\partial}{\partial t} \left[\frac{e^{-\rho t} (c_3 f_m^1(x) + c_4 f_m^2(x))}{g_m(x,t)} \right] = 0$$

$$(13.7)$$

$\square$

This Theorem provides a general methodology for constructing a process which is nearly a martingale out of a Markov process. In the particular case of one-dimensional diffusion processes, the construction gives rise to families with up to 7 adjustable parameters of analytically solvable diffusions with drift equal to zero within the interior of the domain of definition. What makes the case of one-dimensional diffusions special is the fact that the function $Y_m(x)$ is invertible in the limit as $m \to \infty$, i.e. either monotone increasing or monotone decreasing in this limit.

More specifically, consider the diffusion with Markov generator:

$$(13.8) \qquad \mathcal{L}_m = \delta(x) \nabla_{h_m} + \frac{1}{2} \nu(x)^2 \Delta_{h_m}.$$

where $x \in \mathbb{R}$. Two important cases are the CIR process (reducing to the Bessel equation in the continuous limit) for which

$$(13.9) \qquad \delta(x) = (\lambda_0 + \lambda_1 x), \quad \nu(x) = \nu_0 \sqrt{x}$$

with $x \in \mathbb{R}_+$ and the Jacobi process (reducing to a gaussian hypergeometric polynamials of type $_2F_1$) for which

$$(13.10) \qquad \delta(x) = (\lambda_0 + \lambda_1 x), \quad \nu(x) = \nu_0 \sqrt{x(1 - x)}$$

and $x \in [0, 1]$.

The construction above admits a continuous limit if the simplicial functions $f_m^j(x), j = 1, 2, 3, 4$ converge to twice differentiable functions $f^j(x), j = 1, 2, 3, 4$ and satisfy the equation

$$(13.11) \qquad \mathcal{L}f = \rho f$$

within the interior of the domain $A = \lim_{m \to \infty} A_m$. Here

$$(13.12) \qquad \mathcal{L} = \delta(x) \frac{\partial}{\partial x} + \frac{1}{2} \nu(x)^2 \frac{\partial^2}{\partial x^2}.$$

Theorem 20. (Invertibility.) *Let $f^j(x), j = 1, 2, 3, 4$ be functions satisfying equation 13.12 in the interior of the corresponding domain of definition and let*

$$(13.13) \qquad Y(x) = \frac{c_3 f^1(x) + c_4 f^2(x)}{c_1 f^1(x) + c_2 f^2(x)}.$$

for some choice of constants c_1, c_2, c_3, c_4 such that the denominator in this equation has no zeros. Then we have that

$$(13.14) \qquad Y(x) = \int^x \frac{W(y)}{g(y)^2} dy + \text{const}$$

where

$$(13.15) \qquad W(x) = \pm \sigma_0 \exp\left(-\int^x \frac{2\delta(y)}{\nu(y)^2} dy \right)$$

where either sign would be allowed and σ_0 is a positive constant. In particular, the function $Y(x)$ is invertible.

Proof. Let us introduce the Wronskian

$$(13.16) \qquad W(x) = \frac{dh(x)}{dx} g(x) - \frac{dg(x)}{dx} h(x)$$

such that

$$(13.17) \qquad g(x) = c_1 f^1(x) + c_2 f^2(x), \qquad h(x) = c_3 f^1(x) + c_4 f^2(x).$$

A direct calculation shows that

$$(13.18) \qquad \frac{d}{dx} Y(x) = W(x) Y(x)$$

and

$$(13.19) \qquad \frac{d}{dx} W(x) = -\frac{2\delta(x)}{\nu(x)^2} W(x).$$

Hence, (13.15) gives the solution to the equation in (13.19). $\qquad\qquad\qquad \square$

Let $X(y)$ be the inverse of the function $Y(x)$, we have that

$$(13.20) \qquad \frac{dy}{\sigma(y)} = \pm \frac{dX(y)}{\nu(X(y))}$$

where

$$(13.21) \qquad \sigma(y) = Y'(X(y)) \nu(X(y)) = \frac{\sigma_0 \nu(X(y)) \exp\left(-2 \int^{X(y)} \frac{\delta(s)ds}{\nu(s)^2} \right)}{g(X(y),\rho)^2}$$

Theorem 21. (Kernel mapping.) *In the limit $m \to \infty$, the propagator density for the process Y_t with zero drift and volatility as in (13.21) is given by*

$$(13.22) \qquad U(y_0, 0; y, t) = \frac{\nu(X(y))}{\sigma(y)} \frac{g(X(y),\rho)}{g(X(y_0),\rho)} e^{-\rho t} u(X(y_0), 0; X(y), t)$$

where $u(x_0, 0; x, t)$ is the propagator density for the process of generator (13.12).

Let us consider in detail the case of the CIR model in (13.9). If $\lambda_1 = 0$, then the pricing kernel for the state variable is expressed in terms of modified Bessel functions as follows:

$$(13.23) \qquad u(x,t;x_0,0) = \left(\frac{x}{x_0} \right)^{\frac{1}{2}\left(\frac{2\lambda_0}{\nu_0^2}-1\right)} \frac{e^{-2(x+x_0)/\nu_0^2 t}}{\nu_0^2 t/2} I_{\frac{2\lambda_0}{\nu_0^2}-1} \left(\frac{4\sqrt{xx_0}}{\nu_0^2 t} \right) \ .$$

The generating function is

$$(13.24) \qquad \hat{v}(x,\rho) = x^{\frac{1}{2}\left(1-\frac{2\lambda_0}{\nu_0^2}\right)} \left[q_1 I_{\frac{2\lambda_0}{\nu_0^2}-1}\left(\sqrt{\frac{8\rho x}{\nu_0^2}} \right) + q_2 K_{\frac{2\lambda_0}{\nu_0^2}-1}\left(\sqrt{\frac{8\rho x}{\nu_0^2}} \right) \right],$$

with arbitrary constants q_1, q_2. Here $I_\nu(z)$ is the modified Bessel function of order ν and $K_\nu(z)$ is the associated McDonalds function. In this case we obtain a dual family with 6 adjustable parameters.

In case $\lambda_1 < 0$. the propagator density for the state variable x can still be expressed in terms of modified Bessel functions as follows:

$$(13.25) \qquad u(x,t;x_0,0) = c_t \left(\frac{xe^{-\lambda_1 t}}{x_0} \right)^{\frac{1}{2}\left(\frac{2\lambda_0}{\nu_0^2}-1\right)} \exp\left[-c_t(x_0 e^{\lambda_1 t} + x) \right] I_{\frac{2\lambda_0}{\nu_0^2}-1}\left(2c_t \sqrt{xx_0 e^{\lambda_1 t}} \right),$$

where $c_t \equiv 2\lambda_1/(\nu_0^2(e^{\lambda_1 t} - 1))$. For a derivation see (Giorno *et al.* 1988). The general solution of equation (13.1) reduces to Whittaker's equation and generating functions have the general form

$$(13.26) \qquad \hat{v}(x, \rho) = x^{-\lambda_0/\nu_0^2} e^{-\lambda_1 x/\nu_0^2} \left[q_1 W_{k,m}\left(-\frac{2\lambda_1}{\nu_0^2} x \right) + q_2 M_{k,m}\left(-\frac{2\lambda_1}{\nu_0^2} x \right) \right]$$

for arbitrary constants q_1, q_2. Here $W_{k,m}(\cdot)$ and $M_{k,m}(\cdot)$ are Whittaker functions which can also be expressed in terms of confluent hypergeometric functions or in terms of Kummer functions.(Abramowitz and Stegun 1972) This construction gives rise to a dual family with 7 free parameters where

$$(13.27) \qquad k = \frac{\lambda_0}{\nu_0^2} + \frac{\rho}{\lambda_1} \ , \ m = \frac{\lambda_0}{\nu_0^2} - \frac{1}{2} \ .$$

The 7 parameter family which reduces to the CIR model has a local volatility function defined on either an interval or on a half line and behaves asymptotically as the CEV volatility on one hand and as a quadratic model on the other. This hybrid shape allows for greater flexibility.

Next, we show that classic exact solutions in the literature, namely quadratic and CEV models, can all be rediscovered as particular cases of our general formula for the Bessel family where we make use of the above solutions to the underlying x space process with $\beta = \frac{1}{2}$, $\lambda_1 = 0$ and $\lambda \equiv \lambda_0$. Without loss of generality, we can fix $\nu_0 = 2$. Let's specialize further to the case where

$$(13.28) \qquad Y(x) = \bar{y} - a \frac{K_{\frac{\lambda}{2}-1}(\sqrt{2\rho x})}{I_{\frac{\lambda}{2}-1}(\sqrt{2\rho x})}$$

which leads to a transformed process $y_t = Y(x_t)$ with volatility

$$(13.29) \qquad \sigma(y) = \frac{a}{\sqrt{X(y)} \left[I_{\frac{\lambda}{2}-1}(\sqrt{2\rho X(y)}) \right]^2},$$

where $x = X(y)$ is the inverse of the function in equation (13.28). In this family, a and ρ are positive, $\bar{y}$ is arbitrary and $\lambda > 2$. The function $Y(x)$ maps the half line $x \in [0, \infty)$ into $y \in (-\infty, \bar{y}]$, where $Y(x)$ is a strictly monotonically increasing function with $dY(x)/dx = \sigma(Y(x))/\nu(x)$. This solution region can be inverted so that $y \in [\bar{y}, \infty)$. This is accomplished by either replacing a by $-a$ in equation (13.28) or by applying a linear change of variables that maps y into $2\bar{y} - y$. In this special case, we make use of the generating function in equation (13.24), with the choice $q_2 = 0$, and formula (13.22) reduces to

$$(13.30) \quad U(y, t; y_0, 0) = \frac{e^{-\rho t - (X(y) + X(y_0))/2t}}{at} \frac{X(y) \left[I_{\frac{\lambda}{2}-1}(\sqrt{2\rho X(y)}) \right]^3}{I_{\frac{\lambda}{2}-1}(\sqrt{2\rho X(y_0)})} I_{\frac{\lambda}{2}-1}\left(\frac{\sqrt{X(y)X(y_0)}}{t} \right).$$

We note that this density integrates exactly to unity in y space (i.e. no absorption).

Example 4. (Quadratic volatility models) Pricing kernels for quadratic volatility models are readily obtained as a subset of the above general family with the special choice of parameter $\lambda = 3$. After making the substitution $y \to 2\bar{y} - y$ and setting $a = (\bar{y} - \bar{\bar{y}})/\pi$ the transformation function $Y(x)$ becomes

$$(13.31) \qquad Y(x) = \bar{y} + \frac{(\bar{y} - \bar{\bar{y}})}{\pi} \frac{K_{\frac{1}{2}}(\sigma_0 \sqrt{x}/2)}{I_{\frac{1}{2}}(\sigma_0 \sqrt{x}/2)} = \bar{y} + \frac{(\bar{y} - \bar{\bar{y}})}{\exp(\sigma_0 \sqrt{x}) - 1}$$

where $\sigma_0 > 0$. Here, we assume that $\bar{y} > \bar{\bar{y}}$. The inverse transformation $X(y)$ is given by

$$(13.32) \qquad X(y) = (1/\sigma_0^2) \log^2[1 + (\bar{y} - \bar{\bar{y}})/(y - \bar{y})],$$

and the volatility function $\sigma(y)$ is obtained by insertion into equation (13.29) while using the Bessel function of order $\frac{1}{2}$,

$$(13.33) \qquad \sigma(y) = \frac{\sigma_0}{(\bar{y} - \bar{\bar{y}})}(y - \bar{y})(y - \bar{\bar{y}}).$$

Inserting the expression (13.32) into equation (13.30), one obtains the pricing kernel

$$(13.34) \quad U(y,t;y_0,0) = \frac{2e^{-\sigma_0^2 t/8}}{\sigma(y)\sqrt{2\pi t}} \sqrt{\frac{(y_0-\bar{y})(y_0-\bar{\bar{y}})}{(y-\bar{y})(y-\bar{\bar{y}})}} e^{-(\phi(y)^2+\phi(y_0)^2)/2\sigma_0^2 t} \sinh\left(\frac{\phi(y_0)\phi(y)}{\sigma_0^2 t}\right)$$

where $\phi(y) \equiv \log((y-\bar{\bar{y}})/(y-\bar{y}))$. In the special case of a volatility function with a double root, i.e.

$$(13.35) \qquad\qquad\qquad \sigma(y) = \sigma_0(y-\bar{y})^2$$

the pricing kernel is computed by taking the limit as $\bar{\bar{y}} \to \bar{y}$, and one finds

$$U(y,t;y_0,0) = \frac{1}{\sigma_0\sqrt{2\pi t}} \frac{(y_0-\bar{y})}{(y-\bar{y})^3} \left[e^{-\left((y-\bar{y})^{-1}-(y_0-\bar{y})^{-1}\right)^2/2\sigma_0^2 t} \right.$$

$$(13.36) \qquad\qquad\qquad\qquad \left. -e^{-\left((y-\bar{y})^{-1}+(y_0-\bar{y})^{-1}\right)^2/2\sigma_0^2 t} \right].$$

Example 5. (Lognormal models) The pricing kernel for the log-normal Black-Scholes model with $\sigma(y) = \sigma_0 y$ is a particular case of the above formula for the quadratic model. The derivative with respect to y of the quadratic volatility function in (13.33), evaluated at $y = \bar{y}$, is σ_0. Taking the limit $\bar{\bar{y}} \to -\infty$ (or $\bar{\bar{y}} << \bar{y}$), while holding the other parameters fixed, one obtains $\sigma(y) = \sigma_0(y-\bar{y})$. The pricing kernel in (13.34) gives the kernel for the log-normal model in the limit $\bar{\bar{y}} \to -\infty$, i.e.

$$(13.37) \quad U(y,t;y_0,0) = \frac{1}{(y-\bar{y})\sigma_0\sqrt{2\pi t}} \exp\left[-\left(\log((y_0-\bar{y})/(y-\bar{y})) - \frac{\sigma_0^2}{2}t \right)^2 \Big/ 2\sigma_0^2 t \right].$$

Example 6. (CEV model) The constant-elasticity-of-variance (CEV) model is recovered in the limiting case as $\rho \to 0$. Assume $\lambda > 2$ and let $\theta > 0$ be defined so that $\lambda = \theta^{-1} + 2$. The transformation $y = Y(x)$

$$(13.38) \qquad\qquad\qquad Y(x) = \bar{y} + (\sigma_0^2 x)^{-(2\theta)^{-1}}$$

has inverse $x = X(y)$ given by

$$(13.39) \qquad\qquad\qquad X(y) = \sigma_0^{-2}(y-\bar{y})^{-2\theta},$$

for any constant $\bar{y}$. The volatility function for this model is

$$(13.40) \qquad\qquad\qquad \sigma(y) = \frac{\sigma_0}{|\theta|}(y-\bar{y})^{1+\theta}.$$

In the limit $\rho \to 0$, the Laplace transform $\hat{v}(X(y),0) = 1$, which implies that the numeraire change is trivial in this case. The pricing kernel can be evaluated by substitution into the general formula (13.22), and after collecting terms, it turns out to be

$$U(y,t;y_0,0) = \frac{|\theta|}{\sigma_0^2 t} \frac{(y_0-\bar{y})^{\frac{1}{2}}}{(y-\bar{y})^{\frac{3}{2}+2\theta}} e^{-\left((y-\bar{y})^{-2\theta}+(y_0-\bar{y})^{-2\theta}\right)/2\sigma_0^2 t}$$

$$(13.41) \qquad\qquad\qquad I_{\frac{1}{2\theta}}\left(\frac{((y-\bar{y})(y_0-\bar{y}))^{-\theta}}{\sigma_0^2 t} \right).$$

This formula was derived in the case $\theta > 0$, for which the limiting value $y = \bar{y}$ is not attained and the density is easily shown to integrate to unity (i.e. no absorption occurs and the density also vanishes at the endpoint $y = \bar{y}$). We note that the same formula solves the propagator equation for $\theta < 0$, leading to the same Bessel equation of order $\pm(2\theta)^{-1}$. In the range $\theta < 0$, however, the properties of the above pricing kernel are generally more subtle. In particular, one can show that the density integrates to unity for all values $\theta < -1/2$, hence no absorption occurs for $\theta \in (-\infty, -1/2)$. The boundary conditions for the density can be shown to be vanishing at $y = \bar{y}$ (i.e. paths do not attain the lower endpoint) for all $\theta < -1$. In contrast, for $\theta \in (-1, -1/2)$ the density becomes singular at the lower endpoint $y = \bar{y}$ (hence this corresponds to the case that the density has an integrable singularity for which paths can also attain the lower endpoint, but are not absorbed). For the special case of $\theta = -1/2$ the formula gives rise to absorption.

[Note that only for the range $\theta \in (-1/2, 0)$ the above pricing kernel is not useful since it gives rise to a density that has a non-integrable singularity at $y = \bar{y}$. In this case, however, another solution that is integrable is obtained by only replacing the order $(2\theta)^{-1}$ by $-(2\theta)^{-1}$ in the Bessel function. The latter solution for the density does not integrate to unity and hence gives rise to absorption which can be of use to price options in a credit setting.] The special case of $\theta = -1$ gives a nonzero constant value at the lower endpoint, and recovers the Wiener process with reflection and no absorption on the interval $[\bar{y}, \infty)$ with

$$(13.42) \qquad U(y, t; y_0, 0) = \frac{1}{\sigma_0 \sqrt{2\pi t}} \left(e^{-(y-y_0)^2/2\sigma_0^2 t} + e^{-(y+y_0-2\bar{y})^2/2\sigma_0^2 t} \right).$$

14. Stochastic Integrals for Diffusion Processes

This section provides a new derivation of a theorem by Cameron, Feynman, Girsanov, Ito, Kac and Martin, see (Cameron and Martin 1949), (Girsanov 1960), (Ito 1949), (Feynman 1948) and (Kac 1948). This result is at the center of stochastic calculus and in the following sections, we derive far reaching extensions and applications.

Theorem 22. (Cameron-Feynman-Girsanov-Ito-Kac-Martin.) *Consider a diffusion process on the simplicial sequence $A_m \subset \mathbb{R}$ and of generator*

$$(14.1) \qquad \mathcal{L}_m = \mu(x, t)\nabla_{h_m} + \frac{\sigma(x, t)^2}{2}\Delta_{h_m}.$$

Consider also the process given by the integral

$$(14.2) \qquad I_t = \int_0^t a(x_s, s)dx_s + b(x_s, s)ds$$

where $a(x, t)$ and $b(x, t)$ are smooth functions in both arguments. Let us introduce also the function $\phi(x, t) = \int_0^x a(y, t)dy$. We have that

(i) **Ito's Lemma in integral form.**

$$(14.3) \qquad I_t = \phi(x_t, t) - \phi(x_0, 0) + J_t + O(h).$$

where

$$(14.4) \qquad J_t = \int_0^t \left(b(x_s, s) - \frac{1}{2}\sigma(x_s, s)^2 a'(x_s, s) - \dot{\phi}(x_s, s) \right)ds.$$

where $\dot{\phi}(x, t) = \frac{\partial}{\partial t}\phi(x, t)$ and $a'(x, t) = \frac{\partial}{\partial x}a(x, t)$.

(ii) **Feynman-Kac formula.** *The characteristic function of J_t on the bridge leading from x to y is given by*

$$(14.5)$$
$$E_0\left[e^{ipJ_t}\delta(x_t = y)|x_0 = x\right] = P\exp\left(\int_0^t \left(\mathcal{L}(s) + ipb(s) - \frac{ip}{2}\sigma^2(s)a'(s) - ip\dot{\phi}(s) + O(h) \right) \right)(x, y).$$

(iii) **Ito's Lemma in differential form.** *Let $\phi(x, t)$ be a smooth function in both arguments. Then we have that*

$$\lim_{s \downarrow 0} E_t\left[s^{-1}(\phi(x_{t+s}, t+s) - \phi(x_t, t))|x_t = x \right]$$

$$(14.6) \qquad = \frac{\partial \phi}{\partial t}(x, t) + \mu(x, t)\frac{\partial \phi}{\partial x}(x, t) + \frac{\sigma(x, t)^2}{2}\frac{\partial^2 \phi}{\partial x^2}(x, t) + O(h_m)$$

and

$$(14.7) \qquad \lim_{s \downarrow 0} E_t\left[s^{-1}(\phi(x_{t+s}, t+s) - \phi(x_t, t))^2|x_t = x \right] = \sigma(x, t)^2\left(\frac{\partial \phi}{\partial t}(x, t) \right)^2 + O(h_m).$$

For all $n \geq 3$ instead we have that

$$(14.8) \qquad \lim_{h \downarrow 0}\lim_{s \downarrow 0} E_t\left[s^{-1}(\phi(x_{t+s}, t+s) - \phi(x_t, t))^n|x_t = x \right] = 0.$$

Proof. Consider a discretization of the process $I_t = (\Delta I)n_t$. This can be accomplished by introducing a simplicial complex of one more dimension. Eventually in the proof, we shall take the limit as $\Delta I \to 0$ while leaving $h > 0$ constant.

Consider the lifted generator

$$\tilde{\mathcal{L}}(x, n; x', n'; t) = \left(\frac{\sigma(x,t)^2}{2h^2} + \frac{\mu(x,t)}{2h} \right) \delta_{x',x+h} \delta\left(n' - n - \left[\frac{a(x,t)h}{\Delta I} \right] \right)$$

$$+ \left(\frac{\sigma(x,t)^2}{2h^2} - \frac{\mu(x,t)}{2h} \right) \delta_{x',x-h} \delta\left(n' - n + \left[\frac{a(x,t)h}{\Delta I} \right] \right)$$

$$(14.9) \qquad - \frac{\sigma(x,t)^2}{h^2} \delta_{x'x} \delta(n - n') + \frac{b(x,t)}{\Delta I} \left(\delta_{n',n+1} - \delta_{n'n} \right)$$

where $[a]$ stands for the nearest integer to a. The partial Fourier transform in the n variable of this kernel is

$$\tilde{\mathcal{L}}_p(x, x'; t) = \sum_n \tilde{\mathcal{L}}(x, 0; x', n; t)e^{-i\Delta I pn}$$

$$= \left(\frac{\sigma(x,t)^2}{2h^2} + \frac{\mu(x,t)}{2h} \right) \exp\left(-i \left[\frac{ha(x,t)}{\Delta I} \right] \Delta I p \right) \delta_{x',x+h}$$

$$+ \left(\frac{\sigma(x,t)^2}{2h^2} - \frac{\mu(x,t)}{2h} \right) \exp\left(i \left[\frac{ha(x,t)}{\Delta I} \right] \Delta I p \right) \delta_{x',x-h}$$

$$(14.10) \qquad - \frac{\sigma(x,t)^2}{h^2} \delta_{x'x} + \frac{b(x,t)}{\Delta I} \left(e^{ip\Delta I} - 1 \right)$$

In the limit as $\Delta I \to 0$ we find

$$\tilde{\mathcal{L}}_p(x, x'; t) = \frac{\sigma(x,t)^2}{2} \Delta_h + \left(\mu(x,t) - ip\sigma(x,t)^2 a(x,t) \right) \nabla_h$$

$$(14.11) \qquad -ipa(x,t)\mu(x,t) - \frac{1}{2}p^2\sigma(x,t)^2 a(x,t)^2 - ipb(x,t) + O(h).$$

Introducing the function $\phi(x,t)$ defined above, we find that

$$(14.12) \; e^{ip\phi(x,t)}\mathcal{L}(x, x'; t)e^{-ip\phi(x',t)} = \tilde{\mathcal{L}}_p(x, x'; t) + ipb(x,t)\delta_{xx'} - \frac{ip}{2}\sigma(x,t)^2 a'(x,t)\delta_{xx'} + O(h).$$

This equality can be rearranged as follows:

$$e^{-ip\phi(x,t)}\tilde{\mathcal{L}}(x, x'; t)e^{ip\phi(x',t)} + ip\dot{\phi}(x,t)\delta_{xx'} =$$

$$(14.13) \qquad \mathcal{L}_p(x, x'; t) - ipb(x,t)\delta_{xx'} + \tfrac{ip}{2}\sigma(x,t)^2 a'(x,t)\delta_{xx'} + ip\dot{\phi}(x,t)\delta_{xx'} + O(h).$$

Hence

$$P \exp\left(\int_0^t \tilde{\mathcal{L}}_p(s)ds \right)(x, x') =$$

$$e^{ip(\phi(x,0) - \phi(x',t))} P \exp\left(\int_0^t ds \big(\mathcal{L}(s) - ip\big(b(s) - \tfrac{1}{2}\sigma(s)^2 a'(s) - \dot{\phi}(s) + O(h) \big) \big) \right)(x, x').$$

(14.14)

The joint kernel is thus given by

$$P \exp\left(\int_0^t ds\tilde{\mathcal{L}}(s) \right)(x, I; x', I') =$$

$$\int \frac{dp}{2\pi} e^{ip[I' - I - \phi(x',t) + \phi(x,0)]} P \exp\left(\int_0^t ds \big(\mathcal{L}(s) - ip\big(b(s) - \tfrac{1}{2}\sigma(s)^2 a'(s) - \dot{\phi}(s) + O(h) \big) \big) \right)(x, x').$$

(14.15)

This formula proves both statements in the Theorem. $\square$

When in Definition (30) we introduced the notion of measure change, convergence and smoothness in the continuum limit of the measure change function $G_y(y'; t)$ was not emphasized. However, this concept is important, as it is stressed by Girsanov's theorem below, of which we give two independent proofs.

Theorem 23. (Girsanov.) *Consider two lattice diffusions on the simplicial sequence A_m with generators*

$$(14.16) \qquad \mathcal{L}_m = \mu(x,t)\nabla_h + \frac{\sigma(x,t)^2}{2}\Delta_h, \quad \bar{\mathcal{L}}_h = \bar{\mu}(x,t)\nabla_h + \frac{\sigma(x,t)^2}{2}\Delta_h,$$

respectively. Here $\mu(x,t), \bar{\mu}(x,t), \sigma(x,t)$ are assumed to be smooth functions. Then the Markov generators $\mathcal{L}_m$ and $\bar{\mathcal{L}}_m$ are related by the smooth measure change with function

$$(14.17) \qquad G_m^{xt}(x') = \exp\left(\frac{(\bar{\mu}(x,t) - \mu(x,t))}{\sigma(x,t)^2}(x' - x)\right)$$

and the Radon-Nykodym derivative is given by

$$(14.18) \quad \rho(x.) = \exp\left\{ \int_T^{T'} \frac{\bar{\mu}(x_t,t) - \mu(x_t,t)}{\sigma(x_t,t)^2} dy_t - \frac{\bar{\mu}(x_t,t)^2 - \mu(x_t,t)^2}{2\sigma(x_t,t)^2} dt + O(h) \right\}.$$

First Proof. Let $v(x,t) = \bar{\mu}(x,t) - \mu(x,t)$ and let $a(x,t) = \frac{v(x,t)}{\sigma(x,t)^2}$. Consider the measure change with

$$(14.19) \qquad G_m^{xt}(x') = e^{a(x,t)(x'-x)}.$$

The off-diagonal elements of the transformed Markov generator are
$$(14.20)$$
$$\bar{\mathcal{L}}(x, x\pm h; t) = \left(\frac{\sigma(x,t)^2}{2h^2} \pm \frac{\mu(x,t)}{2h}\right)e^{\pm\frac{v(x,t)}{\sigma(x,t)^2}h} = \frac{\sigma(x,t)^2}{2h^2} \pm \frac{\bar{\mu}(x,t)}{2h} + \frac{v(x,t)^2}{4\sigma(x,t)^2} + \frac{\mu(y,t)v(x,t)}{2\sigma(x,t)^2} + O(h).$$

The diagonal elements instead change as follows

$$(14.21) \qquad \bar{\mathcal{L}}(x,x;t) = \mathcal{L}(x,x;t) - \frac{v(x,t)^2}{2\sigma(x,t)^2} - \frac{\mu(x,t)v(x,t)}{\sigma(x,t)^2} + O(h).$$

Second Proof. Let us consider two generators differing by the drift term

$$(14.22) \qquad \mathcal{L}(t) = \frac{\sigma(x,t)^2}{2}\Delta_h + \mu(x,t)\nabla_h, \qquad \bar{\mathcal{L}}(t) = \frac{\sigma(x,t)^2}{2}\Delta_h + \bar{\mu}(x,t)\nabla_h$$

and consider the formula

$$E\left[e^{\int_0^t a(x_s,s)dx_s + b(x_s,s)ds}\delta(x_t = y)|x_0 = x\right] =$$
$$e^{(\phi(y,t)-\phi(x,0))}P\exp\left(\int_0^t ds\left(\mathcal{L}(s) + b(s) - \tfrac{1}{2}\sigma(s)^2 a'(s) + O(h)\right)\right)(x,y).$$
$$(14.23)$$

To derive Girsanov's theorem, let us ask for what choice of the functions $a(x,t), b(x,t)$ the right-hand side equals $e^{t\bar{\mathcal{L}}}(x,y;t)$ up to terms of order $O(h)$. Since
$$(14.24)$$
$$e^\phi\left(\mathcal{L}+b-\frac{1}{2}\sigma^2 a'\right)e^{-\phi} = \frac{\sigma(x)^2}{2}\Delta_h + (a(x)\sigma(x)^2 + \mu(x))\nabla_h + \frac{\sigma(x)^2}{2}a(x)^2 + \mu(x)a(x) + b(x) + O(h)$$

we see that the right hand side equals $\bar{\mathcal{L}}$ up to terms of order $O(h)$ if

$$(14.25) \qquad a(x) = \frac{\bar{\mu}(x) - \mu(x)}{\sigma(x)^2}, \qquad\qquad b(x) = \frac{\mu(x)^2 - \bar{\mu}(x)^2}{2\sigma(x)^2}.$$

15. Markov Bridges

A first simple application of this general framework leads to an extension of the results in the previous section to the case of a general Markov generator.

Consider a Markov process on the simplicial sequence $A_m = h_m \mathbb{Z} \cap [-L, L] \subset \mathbb{R}$. According to theorem (14) and assuming that hypothesis *MG1, MG2, MG3, MG4* hold, the symbol of a Markov generator in canonical form can be expressed as follows:

$$\hat{\mathcal{L}}_m(x, p; t) = i\tilde{\mu}(x; t)\frac{\sin ph}{h} + \tilde{\sigma}(x; t)^2 \frac{\cos ph - 1}{h^2} + \sum_{y \in A_m} \left(e^{ip(y-x)} - 1 - i\frac{\sin ph}{h}(y - x) \right) \tilde{\lambda}(x, y; t).$$

Although not necessary for the validity of the calculation and the convergence in the limit $h \to 0$, we restrict to the special case where the generator has the form

$$(15.1) \qquad \mathcal{L}_m(x, y; t) = i\mu(x; t)\nabla_h + \sigma(x; t)^2 \Delta_h + h_m \lambda_m(x, y; t)$$

where the functions $\mu(x; t)$ and $\sigma(x; t)$ are smooth in both arguments and $\lambda_m(x, y; t) \geq 0$ is smooth and non-negative for $x \neq y$ and we have that

$$(15.2) \qquad \lambda_m(x, x; t) = - \sum_{y \neq x \in A_m} \lambda_m(x, y; t).$$

Consider a stochastic integral of the form

$$(15.3) \qquad I_t = \int_0^t a(x_s, s)dx_s + b(x_s, s)ds$$

where $a(x, t)$ and $b(x, t)$ are smooth functions in both arguments. Let $\phi(x, t) = \int_0^x a(y, t)dy$. The following result holds:

Theorem 24. (Markov Bridges.) *The joint probability distribution function for I_t and the underlying process on the bridge leading from x to x' is given by*

$$P \exp\left(\int_0^t \tilde{\mathcal{L}}_m(s)ds \right)(x, I; x', I')$$

$$= \int \frac{dp}{2\pi} e^{ip[I' - I - \phi(x') + \phi(x)]} P \exp\left(\int_0^t \left(\mathcal{L}_m(s) - ipb(s) + \frac{ip}{2}\sigma^2 a'(s) - \dot{\phi}(s) + \kappa_{m,p}(s) + O(h)\right)ds \right)(x, x').$$

$$(15.4)$$

where $\kappa_{m,p}(t)$ is the operator of matrix elements

$$(15.5) \qquad \kappa_{m,p}(x, x'; t) = h_m \lambda_m(x, x'; t) \exp\left(- ip\big[a(x, t)(x' - x) + \phi(x, t) - \phi(x', t)\big] \right).$$

Proof. Consider again a discretization of the process $I_t = (\Delta I)n_t$ and the lifted generator

$$\tilde{\mathcal{L}}_m(x, n; x', n') = \mathcal{L}_m(x; x')\delta\left(n' - n - \left[\frac{a(x)(x' - x)}{(\Delta I)}\right] \right) + \frac{b(x)}{(\Delta I)}\big(\delta_{n', n+1} - \delta_{n'n}\big)$$

$$(15.6)$$

where $[a]$ stands for the nearest integer to a. The partial Fourier transform in the n variable of this kernel is

$$\tilde{\mathcal{L}}_{m,p}(x, x') = \sum_n \tilde{\mathcal{L}}(x, 0; x', n)e^{-i(\Delta I)pn}$$

$$(15.7) \qquad\qquad = \mathcal{L}_m(x; x') \exp\left(- i(\Delta I)p\left[\frac{a(x)(x' - x)}{(\Delta I)}\right] \right) + \frac{b(x)}{(\Delta I)}\big(e^{ip(\Delta I)} - 1\big)$$

In the limit as $(\Delta I) \to 0$ we find

$$\tilde{\mathcal{L}}_{m,p}(x, x') = \frac{\sigma(x)^2}{2}\Delta_h + \big(\mu(x) - ip\sigma(x)^2 a(x)\big)\nabla_h - ipa(x)\mu(x) - \frac{1}{2}p^2\sigma(x)^2 a(x)^2 - ipb(x)$$

$$+ h_m \lambda_m(x, x'; t)e^{-ipa(x)(x'-x)} + O(h)$$

$$(15.8) \qquad\qquad\qquad\qquad\qquad .$$

Introducing the function $\phi(x,t)$ defined above, we find that

$$e^{ip\phi(x,t)}\mathcal{L}_m(x,x';t)e^{-ip\phi(x',t)} = \tilde{\mathcal{L}}_{m,p}(x,x';t) + ipb(x,t)\delta_{xx'} - \frac{ip}{2}\sigma(x,t)^2 a'(x,t)\delta_{xx'}$$

(15.9) $$+ h_m\lambda_m(x,x';t)\exp\left(-ip\big[a(x,t)(x'-x) + \phi(x,t) - \phi(x',t)\big]\right) + O(h_m).$$

Hence

$$\exp\big(t\tilde{\mathcal{L}}_{m,p}\big)(x,x';t)$$

$$= ^{ip(\phi(x,t)-\phi(x',t))} P\exp\left(\int_0^t \big(\mathcal{L}_m(s) - ipb(s) + \tfrac{ip}{2}\sigma(s)^2 a'(s) - \dot{\phi}(s) + \kappa_{m,p}(s) + O(h_m)\big)ds\right)(x,x').$$

(15.10)

The joint kernel is thus given by (16.14). $\square$

16. Abelian Processes in Continuous Time

This section is based on work in collaboration with Harry Lo and Alex Mijatovic, see (Albanese *et al.* 2006*a*).

Abelian processes are path-dependent processes which provide an extension of the notion of stochastic integral for which one can extend the Feynman-Kac theorem and Ito's formula in Section (14).

Let $A_m, m \geq m_0$ be a simplicial sequence, consider a time interval $[T, T']$ and a stochastic process described by the Markov generator $\mathcal{L}_m(y_1; y_2; t)$. To numerically exponentiate, it is crucial in most application examples to first reduce dimensionality by means of block diagonalisations, or else the multiplication of the lifted generators would not be feasible. It turns out that this is possible for the payoffs of the type above as these two fall in the general class outlined by the following definition:

Definition 43. (Markov Generator.) *Let $\mathcal{L}(y_1; y_2; t)$ be a Markov generator and let us consider a lifting of the form*

(16.1) $$\bar{\mathcal{L}}(y_1, k_1; y_2, k_2; t) = \mathcal{L}(y_1; y_2; t)\delta_{k_1 k_2} + \mathcal{A}(y_1, k_1; y_2, k_2; t).$$

where $k = 0...K$. This lifting is called Abelian if the operators $A(y_1; y_2; t)$ defined for each pair $y_1, y_2 \in \Lambda$ and all times $t \in [T, T']$ as the linear operators of matrix elements

(16.2) $$A(y_1; y_2; t)_{k_1, k_2} \equiv \mathcal{A}(y_1, k_1; y_2, k_2; t)$$

are mutually commuting in the sense that

(16.3) $$A(y_1, y_2)A(y_3, y_4) = A(y_3, y_4)A(y_1, y_2)$$

for all $y_1, y_2, y_3, y_4 \in \Lambda$ and all $t \in [T, T']$.

Lemma 4. *If all the matrices $A(y_1; y_2; t)_{k_1, k_2}$, $y_1, y_2 \in \Lambda$, $t \in [T, T']$ are mutually commuting and if furthermore they are all diagonalizable, then they can be diagonalized simultaneously at any given point in time, i.e. there exists a time dependent matrix $V(k, i; t), i = 0, ..K$ such that $V(t)^{-1}\mathcal{A}(y_1; y_2; t)V(t) = \Lambda(y_1; y_2; t)$, where $\Lambda(y_1; y_2; t)$ is a diagonal matrix of the form $(\lambda(y_1; y_2; t)\delta_{i_1, i_2})$ for any $t \in [T, T']$.*

Proof. Fix a $t \in [T, T']$. If $y_1, y_2, y_3, y_4 \in \Lambda$ and $A(y_1; y_2; t)\psi(t) = \lambda(y_1; y_2; t)\psi(t)$ for some vector $\psi(t)$, then

(16.4) $$A(y_3; y_4; t)A(y_1; y_2; t)\psi = A(y_1; y_2; t)A(y_3; y_4; t)\psi(t) = \lambda(y_1; y_2; t)A(y_3; y_4; t)\psi(t).$$

Hence if $\psi(t)$ is an eigenvector of $A(y_1; y_2; t)$ also $A(y_3; y_4; t)\psi(t)$ is an eigenvector of $A(y_1; y_2; t)$. If $\psi(t)$ is an eigenvector of multiplicity one, this shows that $\psi(t)$ is also an eigenvector of $A(y_3; y_4; t)$ for all y_3, y_4. Otherwise, we conclude that the eigenspace of $A(y_1; y_2; t)$ of eigenvalue λ is invariant under $A(y_3, y_4)$ for any y_3, y_4. Iterating the argument above to this eigenspace one can constructively obtain a common set of invariant eigenspaces shared by all the operators

$\mathcal{A}(y_1; y_2; t)$ for $y_1, y_2 \in \Lambda$. Hence, for any given $t \in [T, T']$, these operators can be diagonalized simultaneously by the matrix whose column vectors span bases for all the common eigenspaces. $\square$

Let $V(k, i; t), i = 0, ..K$ be a matrix which diagonalizes simultaneously all members of the family of operators $T(k_1, k_2; y_1 t)$, so that $V(t)^{-1} T(y_1; t) V(t) = \Lambda(y_1; t)$ is diagonal. Consider the lifted matrix $\mathcal{V}(k, y_1; i, y_2; t) = V(k; i; t) \delta_{y_1 y_2}$ and the transformed lifted Markov generator

$$(16.5) \qquad (\mathcal{V}(t)^{-1} \bar{\mathcal{L}} \mathcal{V}(t))(y_1, i_1; y_2, i_2; t) = \mathcal{L}(y_1; y_2; t) \delta_{i_1 i_2} + \Lambda(y_1; t)_{i_1} \delta_{y_1, y_2} \delta_{i_1 i_2}.$$

Since this matrix is block-diagonal, the problem of exponentiating it is reduced to the problem of exponentiating K blocks separately. This reduces the overall numerical complexity and makes it comparable with the complexity of evaluating propagators on K different time points. As a further simplification, to exponentiate this block-diagonal matrix it is necessary to hold in memory only the blocks, and not the entire matrix. Notice that the blocks are in general complex matrices. Hence to fast exponentiate them one needs to invoke the complex matrix-matrix multiplication routine `sgemm` or `cgemm` of Level-3 BLAS, depending on whether the block-diagonalizing transformation is real or complex valued.

Example 7. (Stochastic Integrals) Stochastic integrals over diffusion processes provide examples of Abelian process. This was already used in the previous section but it is worthwhile here to stress the link between the computability of the characteristic functions for the path dependent process on a bridge with the block-diagonalizability under partial Fourier transforms of the lifted generator. Working directly with Markov generators, one can also generalize the results in the previous sections.

Consider the integral

$$(16.6) \qquad I_t = I(y_\cdot, t) \equiv \int_T^t \left(\phi(y_s; s) + \chi(y_{s-0}, y_s; s) \lim_{t \downarrow 0} \frac{\psi(y_s; s) - \psi(y_{s-t}; s - t)}{t} \right) ds$$

where $\phi(y; s)$, $\psi(y; s)$ and $\chi(y_1; y_2; s)$ are continuous functions. Consider the problem of finding its distribution on a bridge for the underlying Markov process y_t. Introducing the discretisation

$$(16.7) \qquad I_t \approx (\Delta I) m_t$$

where $m_t \in [0, ..N]$ and $(\Delta I) > 0$, the lifted generator for the pair of processes $(y_t, (\Delta I) m_t)$ is

$$\bar{\mathcal{L}}(y_1, m_1; y_2, m_2; t) = \mathcal{L}(y_1; y_2; t) \delta \big(m_2 - m_1 - [(\Delta I)^{-1} \chi(y_1; y_2; t)(\psi(y_2; t) - \psi(y_1; t))] \big)$$

$$+ \frac{\delta_{y_1, y_2}}{(\Delta I)} \phi(y_1; t) \delta_{m_1 + 1, m_2}.$$

$$(16.8)$$

Theorem 25. (Abelian Bridges.) *The joint kernel of I_t and the underlying process on the bridge leading from y_1 to y_2 is given by*

$$P e^{\int_0^t \tilde{\mathcal{L}}_m(s)}(y_1, I_1; y_2, I_2; s) = \int \frac{dp}{2\pi} e^{ip(I_2 - I_1)} P \exp \left(\int_0^t \big(\mathcal{L}_m(s) - ip\phi(s) + \kappa_{m,p}(s) + O(h) \big) \right)(x, x').$$

$$(16.9)$$

where $\kappa_{m,p}(t)$ is the operator with matrix elements

$$(16.10) \qquad \kappa_{m,p}(y_1, y_2; t) = \mathcal{L}_m(y_1, y_2; t) \left(\exp \big(- ip\chi(y_1, y_2, t)(\psi(y_2, t) - \psi(y_1, t)) \big) - 1 \right).$$

Notice that this theorem extends the result in the previous section on Markov bridges as it includes multifactor processes such as processes with stochastic volatility and, more generally, regime switching.

Example 8. (Double Liftings) There are situations that emerge in practice where one has to track two integrals over paths, one of form (16.6) and a similar one of form

$$(16.11) \qquad I'_t = I'(y_\cdot, t) \equiv \int_T^t \left(\phi'(y_s, s) + \chi'(y_{s-0}, y_s, s) \lim_{t \downarrow 0} \frac{\psi'(y_s, s) - \psi'(y_{s-t}, s - t)}{t} \right) ds$$

In this case one can introduce a similar approximation

(16.12) $$I'_t \approx \bar{I}'_t = (\Delta I')n_t$$

where $n_t \in [0, ..N]$ and a double lifting for the generator

$$\bar{\mathcal{L}}(y_1, m_1, n_1; y_2, m_2, n_2; t) =$$

$$\mathcal{L}(y_1; y_2; t)\delta\left(m_2 - m_1 - \left[\frac{\chi(y_1; y_2; t)(\psi(y_2; t) - \psi(y_1; t))}{(\Delta I)}\right]\right)\delta\left(n_2 - n_1 - \left[\frac{\chi'(y_1; y_2; t)(\psi'(y_2; t) - \psi'(y_1; t))}{(\Delta I')}\right]\right)$$

$$+ \frac{\delta_{y_1, y_2}}{(\Delta I)}\phi(y_1; t)\delta_{m_1+1, m_2}\delta_{n_1, n_2} + +\frac{\delta_{y_1, y_2}}{(\Delta I')}\phi'(y_1; t)\delta_{n_1+1, n_2}\delta_{m_1, m_2}.$$

(16.13)

Theorem 26. (Multifactor Bridges.) *The joint kernel of I_t, I'_t and the underlying process on the bridge leading from y_1 to y_2 is given by the double Fourier transform*

$$e^{t\tilde{\mathcal{L}}_m}(y_1, I_1; y_2, I_2) = \int \frac{dp}{2\pi}e^{ip(I_2 - I_1)} \int \frac{dp'}{2\pi}e^{ip'(I'_2 - I'_1)}$$

$$P\exp\left(\int_0^t \left(\mathcal{L}_m(s) - ip\phi(s) - ip'\phi'(s) + \kappa_{m, p, p'}(s) + O(h)\right)\right)(x, x').$$

(16.14)

where $\kappa_{m, p, p'}(t)$ is the operator with matrix elements
(16.15)
$$\kappa_{m, p, p'}(y_1; y_2; t) = \mathcal{L}_m(y_1; y_2; t)\left(e^{-ip\chi(y_1; y_2; t)(\psi(y_2; t) - \psi(y_1; t))} - 1\right)\left(e^{-ip'\chi'(y_1; y_2; t)(\psi'(y_2; t) - \psi'(y_1; t))} - 1\right).$$

Example 9. (The Max Process) Consider the path dependent quantity given by the maximum attained by a given Markov process, i.e.

(16.16) $$K_t = K(y_t, t) \equiv \max_{[T, t]} \chi(y_s; s)ds$$

Let's introduce the approximation

(16.17) $$K_t \approx \bar{K}_t = \alpha k_t,$$

where α is a constant and k_t is an integer value process m_t whose paths take values in $k_t \in \{0...M\}$. The dynamics for the joint process (y_t, k_t) is defined by the lifted generator $\bar{\mathcal{L}}$ on $\bar{\Lambda}$ such that:

(16.18) $$\bar{\mathcal{L}}(y_1, k_1; y_2, k_2; t) = \mathcal{L}(y_1; y_2; t)\delta_{k_1, k_2} + \mathcal{A}(y_1)_{k_1, k_2}\delta_{y_1 y_2}$$

where

(16.19) $$\mathcal{A}(y_1)_{k_1, k_2} = \begin{cases} A & \text{if} \quad \chi(y_1) > \alpha k_1 \quad \text{and} \quad k_2 = [\chi(y_1)/\alpha] \\ 0 & \text{otherwise.} \end{cases}$$

where $[a]$ stands for the nearest integer to a and $A > 0$ is a fixed number. Typically, A is chosen to be a large number as the approximation converges in the limit at $A \to \infty$ and $\alpha \to 0$. A direct calculation shows that the operators $\mathcal{A}(y_1)_{k_1, k_2}$ commute and hence the maximum of the underlying process is itself an Abelian path dependent process.

17. Abelian Processes in Discrete Time and non-Resonance Conditions

This section is based on work in collaboration with Manlio Trovato (Albanese and Trovato 2005) and Paul Jones, see (Albanese and Jones 2007).

An important class of path-dependent options requires computing the joint distribution of the underlying lattice process and of a discrete sum of the following form:

(17.1) $$J_t = J(y_t, t) \equiv \sum_{i=1}^{n} \psi(y_{t_{i-1}}, y_{t_i}; t_i)$$

where n is an integer, $t_i = i\Delta T$ and $t = n\Delta T$. Suppose that the Markov generator $\mathcal{L}$ is time homogenous in the interval $[t_0, t_n]$ and consider the elementary propagator

$$(17.2) \qquad U_i(y_1, y_2) = P \exp\left(\int_{t_i}^{t_{i+1}} \mathcal{L}(s)ds \right)(y_1, y_2).$$

To find the joint transition probability, one can again discretize the variable J_t so that

$$(17.3) \qquad J_t \approx (\Delta J)n_t.$$

As opposed to lifting the generator as in the continuous time case discussed in the previous section, here we lift the elementary propagator itself and form the joint propagator

$$(17.4) \qquad \tilde{U}_i(y_1, k_1; y_2, k_2) = U_i(y_1, y_2)\delta\big(k_1 - k_2 + [\psi_i(y_1, y_2)(\Delta J)^{-1}]\big).$$

This operator can be block-diagonalized by means of a partial Fourier transform (Albanese and Trovato 2005).

Theorem 27. *Consider the Fourier transform operator $\bar{U}_{i,p}$ of matrix elements*

$$(17.5) \qquad \tilde{U}_{i,p}(y_1; y_2) = U_i(y_1, y_2)e^{-ip\psi_i(y_1, y_2)}.$$

Then we have that

$$(17.6) \qquad \left(P \prod_{i=0}^{n-1} \tilde{U}_i \right)(y_1, J_1; y_2, J_2) = \int \frac{dp}{2\pi} e^{ip(J_2 - J_1)} \left(P \prod_{i=0}^{n-1} \tilde{U}_{i,p} \right)(y_1, y_2)$$

This case also falls under a more general class of Abelian liftings.

Definition 44. (Propagator Liftings.) *Let $U_i(y_1, y_2)$ be a family of Markov propagators indexed by $i = 0, 1, ...n-1$ defined over the time intervals $[t_i, t_{i+1}]$. Consider a propagator lifting of the form*

$$(17.7) \qquad \bar{U}_i(y_1, k_1; y_2, k_2) = U_i(y_1, y_2)Q_i(y_1, y_2)_{k_1 k_2}.$$

where $k_1, k_2 = 0...K$. This lifting is called Abelian if the operators $Q_i(y_1, y_2)$ satisfy the following commutation condition:

$$(17.8) \qquad Q_i(y_1, y_2)Q_i(y_1', y_2') - Q_i(y_1', y_2')Q_i(y_1, y_2) = 0$$

for all $y_1, y_2, y_3, y_4 \in A_m$ and all time intervals $i = 0, ...n-1$.

Let $V(k, j), j = 0, ..K$ be a matrix that simultaneously diagonalizes all operators of the family $Q(y_1, y_2)_{k_1 k_2}$ so that $V^{-1}Q(y_1, y_2)V = \Lambda(y_1, y_2)$ is diagonal. Consider the lifted matrix $\mathcal{V}(k, y_1; j, y_2) = V(k, j)\delta_{y_1 y_2}$ and the transformed lifted Markov generator

$$(17.9) \qquad (\mathcal{V}^{-1}\bar{U}_i\mathcal{V})(y_1, j_1; y_2, j_2) = \bar{U}_i(y_1; y_2)\Lambda_i(y_1, y_2)\delta_{j_1 j_2}.$$

Since this matrix is block-diagonal, the evaluation of the time-ordered product

$$(17.10) \qquad P \prod_{i=0}^{n-1} \bar{U}_i = \mathcal{V}\left(P \prod_{i=0}^{n-1} (\mathcal{V}^{-1}\bar{U}_i\mathcal{V}) \right)\mathcal{V}^{-1}$$

involves only multiplying blocks whose dimension equals the size of the lattice A_m. The reader will recognize that the formula in (17.10) is yet another generalization and extension of the Cameron-Feynman-Girsanov-Ito-Kac-Martin formulas discussed above.

The non-singular linear transformation that accomplished the block-diagonalization can be a Fourier transform or a more general transformation. Several possibilities are open and the optimal choice depends on the objective. When using transforms more general than the Fourier transform, one has to keep present that the simultaneous diagonalization of the matrices $Q(y_1, y_2)_{k_1 k_2}$ above can possibly result in a numerically ill-conditioned problem. An example of this phenomenon arises when one attempts to use a non-homogeneous discretization of the path process coordinate. To seize the benefit of inhomogeneous lattices, one needs to be careful when discretizing and avoid resonances.

Consider a lattice for the path dependent process defined as follows $\Omega = \{\omega_1, \omega_1, ...\omega_K\}$ where $\omega_1 < \omega_2 < ... < \omega_K$. Consider the shift operator $\mathcal{R}$ of matrix elements

$$(17.11) \qquad\qquad R_{k_1 k_2} = (1 - p_{k_1})\delta_{k_1 k_2} + p_{k_1}\delta_{k_2, k_1 + 1},$$

if $k_1 < k_2$ and $R_{KK} = 0$. Here we assume that $0 < p_k < 1$, for $k = 1, ..n$. The eigenvalues of $\mathcal{R}$ are given by the diagonal elements $\rho_k = 1 - p_k$. Diagonalizing this sort of matrix is potentially a seriously ill-conditioned especially if the matrix is large and the eigenvalues are degenerate of nearly degenerate. For practical application, one thus needs to achieve non resonance conditions and keep the eigenvalues ρ_i as far apart from each other as possible.

Given a non-resonant choice of transition probabilities such as the one above, one can then fix a ω_0 and a $\Delta\omega > 0$ and determine the lattice Ω in such a way that

$$(17.12) \qquad\qquad \sum_{k_2=1}^{K} R_{k_1 k_2}\omega_{k_2} = \omega_{k_1} + \Delta\omega$$

for all $k_1 = 1, ...K - 1$.

To obtain a non-resonant spectrum, one may choose the lattice ω_k so that

$$(17.13) \qquad\qquad \omega_k = \omega_0 \cdot Z^k$$

where $\omega_0 > 0$ and $Z > 1$. For this discretization to be acceptable, the resulting transition probabilities

$$(17.14) \qquad\qquad p_k = \frac{\Delta\omega}{\omega_0 Z^k (Z - 1)}$$

ought to be between 0 and 1. Hence Z must satisfy the additional constraint

$$(17.15) \qquad\qquad Z(Z - 1) > \frac{\Delta\omega}{\omega_0}.$$

We interpolate between the transition kernels $\mathcal{R}$ to obtain the relevant kernels of all possible values of $\psi(x_1, x_2)$. We construct the relevant kernels Λ_ψ as follows:

$$Q(y_1, y_2) = ([\psi(y_1, y_2)] + 1 - \psi(y_1, y_2))\mathcal{R}^{[\psi(y_1, y_2)]} + (\psi(y_1, y_2) - [\psi(y_1, y_2)])\mathcal{R}^{[\psi(y_1, y_2)]+1},$$
$$(17.16)$$

for all y_1 and y_2 in A_m, where $[\psi(y_1, y_2)]$ represents the integer part of the functional ψ at (y_1, y_2). In the representation in which the operator $\mathcal{R}$ is diagonal, the operators $Q(y_1, y_2)$ are also diagonal and the lifted propagator is block-diagonal.

18. Univariate Moment Expansions on Bridges

This section is based on work in collaboration with Adel Osseiran, see (Albanese and Osseiran 2007).

A second methodology that applies to most Abelian path dependent options is is based on obtaining only a few moments of an Abelian process on any given bridge with respect to the underlying Markov process, as opposed to reconstructing the entire conditional distribution.

Consider a time interval $[T, T']$ and a Markov generator $\mathcal{L}(y_1, y_2; t)$. Consider again the integral I_t in (16.6). Let's introduce the following one parameter family of deformed Markov operators parameterized by $\epsilon \in \mathbb{R}$

$$(18.1) \qquad\qquad \mathcal{L}_\epsilon(y_1, y_2; t) = \mathcal{L}(y_1, y_2; t) + \epsilon V(y_1, y_2; t)$$

where

$$(18.2) \qquad V(y_1, y_2; t) = \phi(y_1; t)\delta_{y_1, y_2} + \mathcal{L}(y_1, y_2; t)\chi(y_1, y_2; t)(\psi(y_2; t) - \psi(y_1; t))$$

Theorem 28. (Dyson Formula.) *We have that*

$$(18.3) \qquad \left(\frac{d}{d\epsilon}\right)^n \bigg|_{\epsilon=0} P\exp\left(\int_T^t \mathcal{L}_\epsilon(s)ds\right)(y_1, y_2) = E_T\left[I_t^n\delta(y_t - y_2)|y_T = y_1\right].$$

Proof. Consider Neper's formula for the propagator

$$(18.4) \qquad P \exp\left(\int_T^t \mathcal{L}_\epsilon(s) ds \right) = \lim_{N \to \infty} P \prod_{i=1}^N \left(1 + \frac{t-T}{N} (\mathcal{L}(t_i) + \epsilon \phi(t_i)) \right)^N$$

where $t_i = T + \frac{it}{N}$. By collecting similar powers in ϵ, one finds Dyson's formula

(18.5)

$$P \exp\left(\int_T^t \mathcal{L}_\epsilon(s) ds \right) = \exp((t-T)\mathcal{L}+$$

$$(18.6) \qquad \epsilon \int_T^t ds_1 \left(e^{(s_1-T)\mathcal{L}} V(s_1) e^{(t-s_1)\mathcal{L}} \right) +$$

$$(18.7) \qquad \sum_{n=2}^\infty \epsilon^n \int_T^t ds_1 ... \int_{s_{n-1}}^t ds_n \left(e^{(s_1-T)\mathcal{L}} V(s_1) e^{(s_2-s_1)\mathcal{L}} .... V(s_n) e^{(t-s_n)\mathcal{L}} \right).$$

The time-ordered integrals above are proportional to conditional moments, i.e.

$$(18.8) \qquad P \exp\left(\int_T^t \mathcal{L}_\epsilon(s) ds \right)(y_1, y_2) = \sum_{n=0}^\infty \frac{\epsilon^n}{n!} E_T \left[I_t^n \delta(y_t - y_2) | y_T = y_1 \right].$$

Here, the factorials originate from the time ordering. $\qquad \square$

A technique which is numerically stable in many situations of practical relevance is to numerically differentiate with respect to ϵ the deformed propagators $P \exp\left(\int_T^t \mathcal{L}_\epsilon(s) ds \right)(y_1, y_2)$ and evaluate at $\epsilon = 0$. This technique can be used to obtain the first moments of I_t on any given bridge for the underlying Markov process. In most applications, we find that two moments suffice to extrapolate the probability distribution function to sufficient accuracy. To do so, it is convenient to choose from among the probability distribution functions which are analytically tractable. For instance, starting from the first two moments only, one can use either the log-normal or the chi-square distribution.

The standard chi-square distribution is given by

$$f(x) = \frac{1}{2\Gamma\left(\frac{a}{2}\right)} \left(\frac{x}{2}\right)^{a/2-1} e^{-x/2}$$

where a is the number of degrees of freedom. The first and second (raw) moments of this distribution are

$$E[x] = a, \qquad E[x^2] = a(a+2)$$

To match the pre-assigned first and second moments m_1, m_2, respectively, one can pass to the new variable

$$\xi = \frac{m_1}{a} x$$

and chose

$$a = \frac{2m_1^2}{m_2 - m_1^2}$$

Let ξ be a log-normally distributed random variable with probability distribution function

$$f(x; \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-(\ln x - \mu)^2/2\sigma^2}$$

The first two moments are given by

$$E[x] = e^{\mu + \sigma^2/2} \qquad \text{and} \qquad E[x^2] = e^{2\mu + 2\sigma^2}$$

The parameters μ and σ can be reconstructed from two pre-assigned first and second moments, m_1 and m_2 respectively, by setting

$$\mu = \log\left(\frac{m_1^2}{\sqrt{m_2}} \right) \qquad \text{and} \qquad \sigma^2 = \log\left(\frac{m_2}{m_1^2} \right).$$

The log-normal and the chi-square distributions allow one to match the first two moments. More general distributions available in closed form allow one to match higher moments also. The Pearson Type III distribution for instance has a probability distribution function given by

$$f(x) = \frac{1}{b\Gamma(p)} \left(\frac{x-a}{b}\right)^{p-1} e^{-(x-a)/b}$$

and extends over the half line $[a, +\infty)$. The special case of this distribution when $a = 0$, $b = 2$ and p is half of an integer, gives the Chi-Squared distribution. In general, the moments are given by

$$E[x] = a + bp$$
$$E[x^2] = (a + bp)^2 + b^2 p$$
$$E[x^3] = (a + bp)^3 + 3b^2 p(a + bp) + 2b^3 p$$

and matching these with our pre-assigned moments m_1, m_2 and m_3 and computing the values of a, b and p in terms of these moments we get

$$a = m_1 - \frac{2(m_2 - m_1^2)^2}{m_3 + 2m_1^3 - 3m_1 m_2}$$

$$b = \frac{m_3 + 2m_1^3 - 3m_1 m_2}{2(m_2 - m_1^2)}$$

$$p = \frac{4(m_2 - m_1^2)^3}{(m_3 + 2m_1^3 - 3m_1 m_2)^2}$$

Example 10. (Volatility contracts) As an example, consider variance and volatility swaps. Variance swaps of maturity T' and time of issuance T have a payoff given by

$$\min\left(\int_T^{T'} v(y_{s-0}, y_s)ds, f \cdot SR^2\right) - SR^2$$

where SR is the swap rate and f is a factor, a typical value being $f = 6.2$. Here, $v(y_1, y_2, t)$ is the instantaneous variance defined as follows:

$$(18.9) \qquad\qquad v(y_1, y_2) = \mathcal{L}(y_1, y_2; t) \log^2\left(\frac{S(y_2)}{S(y_1)}\right)$$

if y_1 is such that $x(y_1) \neq 0$. Otherwise, if $x(y_1) = 0$ then $v(y_1, t) = \infty$. The variance swap is said to be at equilibrium if its price is 0. The payoff of a volatility swap is

$$\min\left(\sqrt{\int_T^{T'} v(y_s, s)ds}, SR\right) - SR$$

Here, SR is the volatility swap rate.

To price these contracts it suffices to evaluate the distribution of realized variance on a bridge, i.e. of the functional

$$(18.10) \qquad\qquad RV(y_2) = \delta(y_{T'} - y_2)\int_T^{T'} v(y_s, s)ds$$

By approximating the distribution of $RV(y)$ with the chi-squared distribution, and using the CDF

$$F(x; a) = \frac{\gamma\left(\frac{a}{2}, \frac{x}{2}\right)}{\Gamma\left(\frac{a}{2}\right)}$$

where $\Gamma(z)$ and $\gamma(z, a)$ are the gamma and incomplete gamma functions respectively, i.e:

$$\Gamma(z) = \int_0^\infty s^{z-1}e^{-s}ds\,, \quad \gamma(z, a) = \int_0^a s^{z-1}e^{-s}ds$$

we find

$$E_t\left[\min(RV, RV_{\max})\,\delta(y_{T'} - y_2)\right] = \frac{m_1}{a}\left[K(1 - F(K; a)) + aF(K; a + 2)\right]$$

and

$$E_t\left[\min(\sqrt{RV}, \sqrt{RV_{\max}})\,\delta(y_{T'} - y_2)\right] = \sqrt{\frac{m_1}{a}}\left[\sqrt{K}(1 - F(K; a)) + \frac{\sqrt{2}\gamma\left(\frac{a+1}{2}, \frac{K}{2}\right)}{\Gamma\left(\frac{a}{2}\right)}\right]$$

where

$$a = a(y_1, y_2) = \frac{2m_1(y_1, y_2)^2}{m_2(y_1, y_2) - m_1(y_1, y_2)^2}, \quad K = K(y_1, y_2) = \frac{a(y_1, y_2)}{m_1(y_1, y_2)}RV_{\max}$$

Since the dependency on the swap rate in both cases is non-linear, an exact calculation requires using a root finder.

Using instead the log-normal distribution to extrapolate, we find

$$E_t[\min(RV, RV_{\max})\,\delta(y_{T'} - y_2)]$$
$$= e^{\mu+\sigma^2/2}\mathcal{N}\left(\frac{\log(RV_{\max})-\mu-\sigma^2/2}{\sigma}\right) + RV_{\max}\mathcal{N}\left(\frac{\log(RV_{\max})-\mu}{\sigma}\right).$$

(18.11)

For volatility swaps instead, we find

$$E_t\left[\min(\sqrt{RV}, \sqrt{RV_{\max}})\,\delta(y_{T'} - y_2)\right]$$
$$= e^{\frac{1}{8}(4\mu+\sigma^2)}\mathcal{N}\left(\frac{\log(RV_{\max})-\mu-\sigma^2/2}{\sigma}\right) + \sqrt{RV_{\max}}\mathcal{N}\left(\frac{\log(RV_{\max})-\mu}{\sigma}\right)$$

(18.12)

where μ and σ are specified above in terms of m_1 and m_2.

Finally, in case the Pearson Type 3 distribution is used, we can express these expectations in terms of the chi-square cumulative distribution function $F(x; a)$ as follows:

$$E_t[\min(RV, RV_{\max})\,\delta(y_{T'} - y_2)] = (a + bp - RV_{\max})F_{\mathrm{Chi}}\left(2p, 2\frac{RV_{\max} - a}{b}\right)$$

(18.13)
$$-b\left(\frac{RV_{\max} - a}{b}\right)^p\frac{e^{-\frac{RV_{\max}-a}{b}}}{\Gamma(p)} + (2RV_{\max} - a - bp).$$

19. Bivariate Moment Expansions on Bridges

This section is based on work in collaboration with Adel Osseiran, see (Albanese and Osseiran 2007).

Consider a time interval $[T, T']$ and a Markov generator $\mathcal{L}(y_1, y_2; t)$. We are interested in the joint distribution on any given bridge for the underlying process of two stochastic integrals

$$(19.1) \qquad I_{1,t} = I(y., t) \equiv \int_T^t \left(\phi_1(y_s; s) + \chi_1(y_{s-0}, y_s; s)\lim_{t\downarrow 0}\frac{\psi_1(y_s; s) - \psi_1(y_{s-t}; s - t)}{t}\right)ds$$

and

$$(19.2) \qquad I_{2,t} = I(y., t) \equiv \int_T^t \left(\phi_2(y_s; s) + \chi_2(y_{s-0}, y_s; s)\lim_{t\downarrow 0}\frac{\psi_2(y_s; s) - \psi_2(y_{s-t}; s - t)}{t}\right)ds.$$

To handle this problem using the moment method, we introduce the following two-parameter family of deformed Markov operators parameterized by $\epsilon, \epsilon' \in \mathbb{R}$

$$(19.3) \qquad \mathcal{L}_{\epsilon_1, \epsilon_2}(y_1, y_2; t) = \mathcal{L}(y_1, y_2; t) + \epsilon_1 V_1(y_1, y_2; t) + \epsilon_2 V_2(y_1, y_2; t)$$

where

$$(19.4) \qquad V_1(y_1, y_2; t) = \phi_1(y_1; t)\delta_{y_1, y_2} + \mathcal{L}(y_1, y_2; t)\chi_1(y_1, y_2; t)(\psi_1(y_2; t) - \psi_1(y_1; t))$$

and

$$(19.5) \qquad V_2(y_1, y_2; t) = \phi_2(y_1; t)\delta_{y_1, y_2} + \mathcal{L}(y_1, y_2; t)\chi_2(y_1, y_2; t)(\psi_2(y_2; t) - \psi_2(y_1; t))$$

Theorem 29. (Dyson Formula.) *We have that*

$$(19.6) \quad \left(\frac{\partial^{n_1+n_2}}{\partial \epsilon_1^{n_1} \partial \epsilon_2^{n_2}} \right) \Bigg|_{\epsilon_1=0, \epsilon_2=0} P \exp \left(\int_T^t L_{\epsilon_1, \epsilon_2}(s) ds \right) (y_1, y_2) = E_T \left[I_{1,t}^{n_1} I_{2,t}^{n_2} \delta(y_t - y_2) | y_t = y_1 \right].$$

Proof. The proof is a simple extension of the proof in the univariate case and will be left to the reader. $\qquad \square$

Consider a bi-variate log-normally distributed variable

$$\left(\begin{array}{c} Y_1 \\ Y_2 \end{array} \right) = \left(\begin{array}{c} \log(X_1) \\ \log(X_2) \end{array} \right) \sim N \left(\left(\begin{array}{c} \mu_1 \\ \mu_2 \end{array} \right), \left(\begin{array}{c} \sigma_1 \\ \sigma_2 \end{array} \right) \right)$$

whereby X_1 and X_2 are correlated normally distributed variables of joint distribution

$$(19.7) \qquad f(x_1, x_2) = \frac{1}{2\pi \sigma_1 \sigma_2 \sqrt{1 - \rho^2} x_1 x_2} \cdot$$

$$\exp \left\{ -\frac{1}{2(1-\rho^2)} \left[\left(\frac{\log x_1 - \mu_1}{\sigma_1} \right)^2 - 2\rho \frac{(\log x_1 - \mu_1)(\log x_2 - \mu_2)}{\sigma_1 \sigma_2} + \left(\frac{\log x_2 - \mu_2}{\sigma_2} \right)^2 \right] \right\},$$

where ρ is a correlation parameter. Both X_1 and X_2 are log-normally distributed with

$$E[X_i] = e^{\mu_i + \sigma_i^2/2}, \quad \text{and} \quad E[X_i^2] = e^{2\mu_i + 2\sigma_i^2}, \quad i = 1, 2.$$

Matching these with the pre-assigned moments, and solving for μ_i and σ_i we find

$$\mu_i = \log \left(E \left[I_t^{(i)} \right]^2 \Big/ \sqrt{E \left[\left(I_t^{(i)} \right)^2 \right]} \right) \quad \text{and} \quad \sigma_i^2 = \log \left(E \left[\left(I_t^{(i)} \right)^2 \right] \Big/ E \left[\left(I_t^{(i)} \right) \right]^2 \right)$$

Moreover, the mixed term is

$$E[X_1 X_2] = E \left[e^{Y_1 + Y_2} \right] = e^{\mu_1 + \mu_2 + \frac{1}{2}(\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2)} = E[X_1] E[X_2] e^{\rho \sigma_1 \sigma_2}$$

which gives ρ (in terms of the pre-assigned moments):

$$(19.8) \qquad \rho = \frac{1}{\sigma_1 \sigma_2} \log \frac{E \left[I_t^{(1)} I_t^{(2)} \right]}{E \left[I_t^{(1)} \right] E \left[I_t^{(2)} \right]}$$

Example 11. (Conditional Variance Swaps) The payoff of a conditional variance swap is given by the ratio

$$(19.9) \qquad \frac{\int_T^{T'} v(y_s, s) \mathbf{1}(L < S(y_s) < H) ds}{\int_T^{T'} \mathbf{1}(L < S(y_s) < H) ds} - SR^2$$

To apply the moment method to this case, the first thing to note is that essentially we are modelling the two integrals appearing in the payoff at the same time. We are going to need a Bi-variate distribution to do this. Firstly let's write:

$$I_t^{(1)} = \int_T^{T'} v(y_s, s) \mathbf{1}(L < S(y_s) < H) ds$$

$$I_t^{(2)} = \int_T^{T'} \mathbf{1}(L < S(y_s) < H) ds$$

and in order to compute the expectation

$$(19.10) \qquad E \left[\frac{I_t^{(1)}}{I_t^{(2)}} \right]$$

we'll need the following expectations:

$$E \left[I_t^{(1)} \right], \quad E \left[I_t^{(2)} \right], \quad E \left[\left(I_t^{(1)} \right)^2 \right], \quad E \left[\left(I_t^{(2)} \right)^2 \right], \quad \text{and} \quad E \left[I_t^{(1)} I_t^{(2)} \right]$$

To tackle this problem consider the operator

$$(19.11) \qquad \mathcal{L}_{\epsilon_1,\epsilon_2}(y_1,y_2) = \mathcal{L}(y_1,y_2) + \epsilon_1 \phi(y_1)\delta_{y_1 y_2} + \epsilon_2\,\psi(y_1)\delta_{y_1 y_2}$$

where

$$(19.12) \qquad \phi(y_1,t) = \sum_{y_2} \mathcal{L}(y_1,y_2;t)\log^2\left(\frac{S(y_2)}{S(y_1)}\right)\mathbf{1}(L < S(y_1) < H)$$

and

$$(19.13) \qquad \psi(y_1) = \mathbf{1}(L < S(y_1) < H).$$

We have that

$$\left.\frac{\partial}{\partial\epsilon_1}\right|_{\epsilon_1=0} e^{((t'-t)\mathcal{L}_{\epsilon_1,\epsilon_2})}(y_1,y_2) = E\left[I_t^{(1)}\,\Big|\,y_t = y_1, y_{t'} = y_2\right]$$

and

$$\left.\frac{\partial^2}{\partial\epsilon_1^2}\right|_{\epsilon_1=0} e^{((t'-t)\mathcal{L}_{\epsilon_1,\epsilon_2})}(y_1,y_2) = E\left[\left(I_t^{(1)}\right)^2\,\Big|\,y_t = y_1, y_t' = y_2\right]$$

similarly (but with respect to ϵ_2) for

$$E\left[I_t^{(2)}\,\Big|\,y_t = y_1, y_{t'} = y_2\right] \quad \text{and} \quad E\left[\left(I_t^{(2)}\right)^2\,\Big|\,y_t = y_1, y_{t'} = y_2\right]$$

The joint expectation will involve the mixed derivative:

$$(19.14) \qquad E\left[I_t^{(1)}I_t^{(2)}\right] = \left.\frac{\partial^2}{\partial\epsilon_1\partial\epsilon_2}\right|_{\epsilon_1,\epsilon_2=0} e^{((t'-t)\mathcal{L}_{\epsilon_1,\epsilon_2})}(y_1,y_2)$$

and once computed, we make use of these expectations to match a bivariate distribution. For simplicity of notation we leave out the conditional part of these expectations, noting that all the moments we obtain are conditional to the initial and final points.

To evaluate the expectation $E\left[I_t^{(1)}\big/I_t^{(2)}\right]$, let us notice that

$$E\left[\frac{X_1}{X_2}\right] = E\left[e^{Y_1-Y_2}\right] = E\left[e^{Y_1+Y_2'}\right]$$

where $Y_2' = -Y_2$ is also be normally distributed $\sim N(-\mu_2, \sigma_2^2)$. Hence

$$E\left[\frac{X_1}{X_2}\right] = e^{\mu_1+\frac{1}{2}\sigma_1^2}\, e^{-\mu_2+\frac{1}{2}\sigma_2^2}\, e^{-\rho\sigma_1\sigma_2}$$

and the expectation (19.10) is given by

$$(19.15) \qquad E\left[\frac{I_t^{(1)}}{I_t^{(2)}}\right] = E\left[I_t^{(1)}\right]\cdot\frac{E\left[\left(I_t^{(2)}\right)^2\right]}{E\left[I_t^{(2)}\right]^3}\cdot\frac{E\left[I_t^{(1)}\right]E\left[I_t^{(2)}\right]}{E\left[I_t^{(1)}I_t^{(2)}\right]} = \frac{E\left[I_t^{(1)}\right]^2 E\left[\left(I_t^{(2)}\right)^2\right]}{E\left[I_t^{(2)}\right]^2 E\left[I_t^{(1)}I_t^{(2)}\right]}$$

moreover

$$E\left[\left(\frac{X_1}{X_2}\right)^2\right] = e^{2\mu_1+2\sigma_1^2}\, e^{-2\mu_2+2\sigma_2^2}\, e^{-4\rho\sigma_1\sigma_2} = E\left[\frac{X_1}{X_2}\right]^4 e^{-2\mu_1}e^{2\mu_2}$$

so

$$(19.16) \qquad E\left[\left(\frac{I_t^{(1)}}{I_t^{(2)}}\right)^2\right] = \frac{E\left[I_t^{(1)}\right]^4 E\left[\left(I_t^{(2)}\right)^2\right]^3 E\left[\left(I_t^{(1)}\right)^2\right]}{E\left[I_t^{(2)}\right]^4 E\left[I_t^{(1)}I_t^{(2)}\right]^4}$$

To compute a payoff we will need the expectation $E\left[\min\left(\frac{X_1}{X_2}, CV_{\mathrm{Max}}\right)\right]$ for the payoff of the conditional variance swap.

$$(19.17)\ E\left[\min\left(\frac{X_1}{X_2}, CV_{\mathrm{Max}}\right)\right] = \int_0^\infty \int_0^{X_2 CV_{\mathrm{Max}}} \left(\frac{X_1}{X_2} - CV_{\mathrm{Max}}\right) f(X_1, X_2) dX_1 dX_2 + CV_{\mathrm{Max}}$$

This double integral will need to be evaluated numerically. As the first one has finite bounds and the second has an infinite upper bound it would make sense to use two types of Gaussian quadratures: a Gauss-Legendre quadrature on the inner integral and a Gauss-Laguerre quadrature on the outer one.

20. BLOCK FACTORIZATIONS

Although most path dependent processes emerging in applications to Finance are Abelian, some aren't. If the Abelian property fails, the propagator cannot be block-diagonalized by the methods discussed above and also moment methods generally fail. In several relevant situations one can still achieve a numerically viable framework by using block-factorizations instead of diagonalizations.

Let $U_i(y_1, y_2)$ be a family of Markov propagator indexed by $i = 0, 1, ... n - 1$ defined over the time intervals $[t_i, t_{i+1}]$. Consider a propagator lifting of the form

$$(20.1) \qquad \bar{U}_i(y_1, k_1; y_2, k_2) = U_i(y_1, y_2) Q_i(y_1, y_2)_{k_1 k_2}.$$

where $k_1, k_2 = 0...K$.

Definition 45. (Block-Factorization.) *We say that the propagators in (20.1) admit a block-factorization if they can be represented in the form:*

$$(20.2) \qquad \bar{U}_i = \Pi_i \cdot \left(U_i \otimes \mathbb{I} \right)$$

Here $\Pi_i, i = 0, ... n - 1$ is a family of permutation matrices, i.e. matrices with the property that for each pair (y_1, k_1) we have that

$$(20.3) \qquad \Pi_i(y_1, k_1; y_2, k_2) = 0$$

for all pairs (y_2, k_2) except for one single pair $(Y(y_1, k_1), K(y_1, k_1))$ for which we have instead

$$(20.4) \qquad \Pi_i(y_1, k_1; Y(y_1, k_1), K(y_1, k_1)) = 1.$$

Furthermore, the tensor product notation in equation (20.2) stands for the operator with the following matrix elements:

$$(20.5) \qquad \left(U_i \otimes \mathbb{I} \right)(y_1, k_1; y_2, k_2) = U_i(y_1, y_2) \delta_{k_1 k_2}.$$

If a block-factorization exists, then an efficient backward induction algorithm can be setup. Namely, if v is a vector, then we can value the following path ordered products

$$(20.6) \qquad v_i = P \prod_{j=i}^{n-1} \bar{U}_j v_n = P \prod_{j=i}^{n} \Pi_j \cdot \left(U_j \otimes \mathbb{I} \right) v_n$$

iteratively in i from $i = n - 1$ down to $i = 0$. At the first step one applies the operator $\left(U_{n-1} \otimes \mathbb{I} \right)$ to v_n, followed by the permutation Π_{n-1} to obtain v_{n-1} and then iterate. Due to the tensor product structure of the first operator, it is convenient to apply by regarding the vector v_i as a matrix of elements $v_i(y, k)$. This representation makes it possible to leverage on numerical efficiencies and to re-interpret a high dimensional BLAS level 2 matrix-vector multiplication as a lower dimensional BLAS level 3 matrix-matrix multiplication. Applying a permutation operator is then quite straightforward from the numerical viewpoint and is an operation whose complexity scales linearly with respect to the dimension of the vector v.

Example 12. (Snowballs) Here we follow (Albanese 2007a). Consider the case of a valuing a snowballs for which the structured coupon at time $T_i = T + (\Delta T)i$ has the following form:

$$(20.7) \qquad C_{T_i} = f C_{T_{i-1}} + \Phi_i(y_{T_{i-1}})$$

where the factor f is a fixed parameter and $\Phi_i(y_{T_{i-1}})$ is a given function. Since the coupon amount at a given time affects the coupon amou/nt in the next period, the process is not Abelian, in fact it is path-dependent. However, block-factorizations are still possible.

Consider discretizing the coupon dimension in intervals ΔC, so that a generic coupon can be approximated as follows

$$(20.8) \qquad C_{T_i} = (\Delta C) k_{T_i}$$

where $k_{T_i} = 0, 1, \dots N - 1$ is a discrete time, integer value process. A strategy to implement the backward induction scheme is to organize the payoff function $v(y, k)$ at maturity in a matrix with N columns, each one indexed by the state variable y and representing the price conditional to the discretized value of the previous coupon paid. One can then iterate backward by applying the pricing kernel to this matrix of conditional pricing functions. After each step in the iteration, one needs to reshuffle the pricing functions in such a way that the conditioning relation on each column is satisfied. More precisely, we set

$$(20.9) \quad v_{i-1}(y_1, k) = \sum_{y_2} U_i(y_1, y_2) v_i(y_2, [(f(\Delta C) \cdot k_1 + \Phi_i(y_1))/(\Delta C)]) = \left[\Pi_i \cdot (U_i \otimes \mathbb{I}) v_i \right](y_1, k)$$

where

$$(20.10) \qquad \Pi_i(y_1, k_1; y_2, k_2) = \delta(y_2 - y_1)\delta(k_2 - K(y_1, k_1))$$

and

$$(20.11) \qquad K(y_1, k_1) = [(f(\Delta C) \cdot k_1 + \Phi_i(y_1))/(\Delta C)].$$

Notice that this backward induction scheme can accommodate callability provisions.

Example 13. (Soft Calls) Soft calls provide another interesting example of a non-Abelian process. In this case, a call decision depends on the fraction of time the process spends above a given barrier level during a fixed time period prior to the decision itself. By restricting times at which one makes observations to the discrete sequence of time points $t_i = T + i\delta t, i = 0, \dots M$, the pricing function takes the form

$$(20.12) \qquad v_i(y, \overrightarrow{\sigma})$$

where

$$(20.13) \qquad \overrightarrow{\sigma} = (\sigma_1, \dots, \sigma_N) \in \{0, 1\}^N.$$

Here the variable σ_j equals 1 if $y_{t_{i-N+j}}$ satisfies the barrier condition and 0 otherwise. We model this dependency by means of a generic function notation

$$(20.14) \qquad \sigma_j = \Sigma_i(y_{t_{i-N+j}}).$$

The sequences $\overrightarrow{\sigma}$ can be arranged in an ordered sequence of 2^N integers so that the data structure $v_i(y, \overrightarrow{\sigma})$ appears as a matrix with 2^N columns. A backward induction scheme involves evaluating

$$(20.15) \qquad v_{i-1}(y_1, \overrightarrow{\sigma}) = \sum_{y_2} U_i(y_1, y_2) v_i(y_2, \Phi_i(\overrightarrow{\sigma}, y_1))$$

where

$$(20.16) \qquad \Phi_i(\overrightarrow{\sigma}, y_1) = \{\sigma_2, \dots \sigma_{N-1}, \Sigma(t_i, t_{i-1}, y_1)\}$$

This form is block-factorized as

$$(20.17) \qquad v_{i-1}(y_1, \overrightarrow{\sigma}) = \left[\Pi_i \cdot (U_i \otimes \mathbb{I}) v_i \right](y_1, k)$$

where Π_i is the permutation operator of matrix

$$(20.18) \qquad \Pi_i(y_1, \overrightarrow{\sigma}_1; y_2, \overrightarrow{\sigma}_2) = \delta(y_2 - y_1)\delta((\overrightarrow{\sigma}_2)_N - \Sigma_i(y_1)) \prod_{j=1}^{N-1} \delta((\overrightarrow{\sigma}_2)_j - (\overrightarrow{\sigma}_1)_{j+1}).$$

More general types of useful block-factorizations can be imagined, although they might be numerically less efficient. An example is set forth by the following definition:

Definition 46. (State Dependent Block-Factorization.) *We say that the propagators in (20.1) admit a state dependent block-factorization if they can be represented in the form:*

$$(20.19) \qquad \bar{U}_i = \Pi_i \cdot \bigoplus_{k=1}^{K} U_{i,k}$$

Here $\Pi_i, i = 0, \dots n - 1$ is a family of permutation matrices as above. The direct sum notation in equation (20.19) stands for the operator with the following matrix elements:

$$(20.20) \qquad \big(\bigoplus_{k_0=1}^{K} U_{i,k_0} \big)(y_1, k_1; y_2, k_2) = U_{i,k_1}(y_1, y_2)\delta_{k_1 k_2}.$$

A backward induction in this case is still a lower dimensional problem but since the lower dimensional propagator is state dependent this reduces to a sequence of matrix-vector multiplications as opposed to a single matrix-matrix multiplication. The scheme is thus numerically less efficient. Still it would be useful in cases for instance such as GARCH type models and extensions, see (Engle 1982).

21. Dynamic Conditioning

This section is based on work in collaboration with Alicia Vidler, see (Albanese and Vidler 2007).

Modeling correlations between several processes has sometimes been considered a problem affected by the so-called curse of dimensionality which causes the numerical complexity to explode exponentially, see (Bellman 1957). This motivates the recourse to Monte Carlo methods while calibrating using closed form solutions.

Here we deviate substantially from this framework and aim at building models viable from the engineering deployment viewpoint and specified semi-parametrically or even non-parametrically without any restriction imposed by analytical tractability. The method of fast exponentiation allows one to calibrate single name models without recurring to analytic solvability thanks to providing smooth sensitivities. Although Monte Carlo schemes could in principle be employed in a model constructed non-parametrically and calibrated with operator methods, in many circumstances important for applications this can be avoided by using the technique of dynamic conditioning discussed in this section. This technique can be regarded as a dynamic copula whereby the single factor marginal distributions are preserved and the numerical complexity grows linearly with the number of factors, similarly to what happens with Monte Carlo methods. But, unlike Motecarlo methods, dynamic conditioning is numerically noiseless as there is no sampling, no variance reduction scheme is needed and even features such as callability are numerically treatable.

Dynamic conditioning is based on constructing a hierarchy of conditioning relations. At the base of the hierarchy we have continuous time lattice models for single factor marginal distributions. Next, one introduces a binomial process for each risk factor to condition the corresponding Markov chain. Next one finds a hierarchy of binomial processes to condition the former conditioning lattices. See figure (4).

To describe the technique of dynamic conditioning in detail, consider a particular risk factor described by a lattice process whose filtered probability space is engendered by a Markov propagator $U(y_1, t_1; y_2, t_2)$. For each such single factor, we introduce a conditioning binomial process $h_t \in \mathbb{Z}$, which is constant over the time intervals $[T_i, T_{i+1})$ where $T_i = T + i\Delta T$, $i = 0, 1, 2, \dots N$ and $\Delta T = (T' - T)/N$. At initial time we set $h_T = 0$ while for all $i > 0$ we have that $h_{T_i} - h_{T_{i-1}} = \pm 1$. The elementary propagator for the process h_t across neighboring time points T_i is

$$(21.1) \qquad V(h_i, T_i; h_{i+1}, T_{i+1}) = q_1(h_i, i)\delta_{h_i+1, h_{i+1}} + q_{-1}(h_i, i)\delta_{h_i-1, h_{i+1}}$$
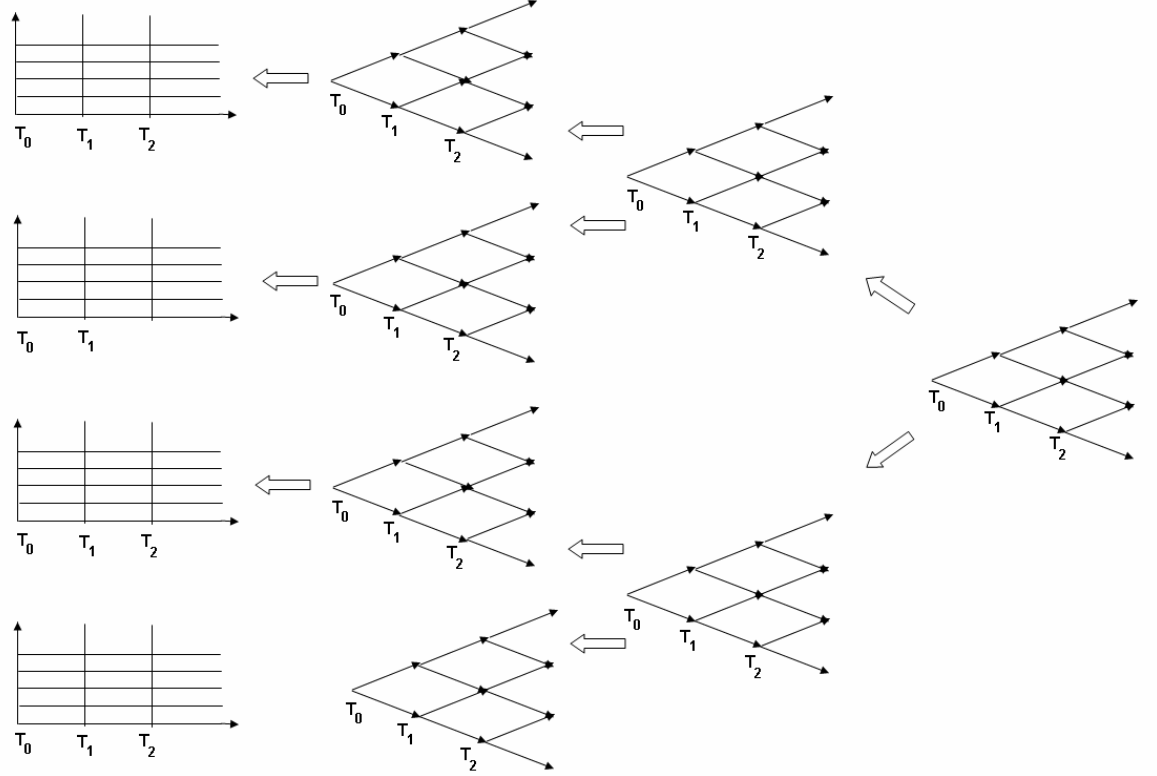
FIGURE 4. Dynamic conditioning scheme: a binomial process conditions each risk factor and is in turn conditioned by another binomial process. The latter is also conditioned by a hierarchy of binomial processes.

where $q_1(h_i, i), q_{-1}(h_i, i) > 0$ and $q_1(h_i, i) + q_{-1}(h_i, i) = 1$. On a general time interval $[T_i, T_j]$, we have that

$$(21.2) \qquad V(h_i, T_i; h_j, T_j) = \sum_{h_t : h_{T_i} = h_i, h_{T_j} = h_j} \prod_{k=i}^{j-1} V(h_{T_k}, T_k; h_{T_{k+1}}, T_{k+1}).$$

Next we define a lifted conditional propagator $\bar{U}(y_i, h_i, T_i; y_j, h_j, T_j)$ in such a way to preserve unconditional transition probabilities from the starting time at T, i.e. so that

$$(21.3) \qquad \sum_{h_j} \bar{U}(y_0, 0, T; y_j, h_j, T_j) = U(y_0, T; y_j, T_j).$$

To do so, one strategy is to define two propagators for each node (h, T_i), namely to choose a pair of operators $U_1(y_i, h_i, T_i; y_{i+1}, h_i + 1, T_{i+1})$ and $U_{-1}(y_i, h_i, T_i; y_{i+1}, h_i - 1, T_{i+1})$ so that

$$U(y_i, T_i; y_{i+1}, T_{i+1}) = q_1(h, i)U_1(y_i, h, T_i; y_{i+1}, h + 1, T_{i+1}) +$$
$$q_{-1}(h, i)U_{-1}(y_i, h, T_i; y_{i+1}, h - 1, T_{i+1}).$$

The operator $\bar{U}(y_0, 0, T; y_j, h_j, T_j)$ satisfying (21.3) is defined as the following sum:

$$\bar{U}(y_0, h_0, T_0; y_j, h_j, T_j) = \sum_{h_t : h_T = 0, h_{T_j} = h_j} \sum_{y_1, ..y_{j-1}} \prod_{k=1..j} q_{h_{T_k} - h_{T_{k-1}}}(h_{T_{k-1}}) \times$$
$$U_{h_{T_k} - h_{T_{k-1}}}(y_{k-1}, h_{T_{k-1}}, T_{k-1}; y_k, h_{T_k}, T_k).$$

Since the operators U_1 and U_{-1} do not commute with each other, if not in trivial situations, each path h_t in the summation on the right hand side of this equation represents a different operator. In many situations one can however avoid the numerical complexities of a non-recombining

tree by finding modified versions of such operators so that the kernels

$$U_{h_.}(y_0, y_k) = \sum_{y_1,..y_{j-1}} \prod_{k=1..j} q_{h_{T_k} - h_{T_{k-1}}}(h_{T_{k-1}}) \bar{U}(y_{k-1}, h_{T_{k-1}}, T_{k-1}; y_k, h_{T_k}, T_k)$$

are all equal to each other, for all paths $h_t : h_T = 0, h_{T_j} = h_j$ and at least one single fixed starting point $y_0 = \bar{y}_0$. This will be referred to as *conditional recombination property*.

The reasons why we focus on the initial point $y_0 = \bar{y}_0$ only are manifold. Firstly, in applications one needs to condition marginal distributions only to spot values, as the price of options in the hypothetical case asset prices were different is not known. Secondly, if we insisted on the same property being valid for all initial starting points we would end up with a seriously ill posed problem of difficult solution. Because of these reasons, we settle for the more modest objective of conditional recombination.

Namely, on each node (h_i, T_i) with $i > 0$ we define

$$(21.4) \qquad \bar{U}(y_{i-1}, h_{i-1}, T_{i-1}; y_i, h_i, T_i) = U_{h_i - h_{i-1}}(y_{i-1}, h_{i-1}, T_{i-1}; y_i, h_i, T_i)$$

in case $h_i = \pm i$. Otherwise, we determine this operator so that

$$\bar{U}(\bar{y}_0, h_0, T_0; y_i, h_i, T_i) =$$
$$\sum_{y_{i-1} \in \Lambda} \bar{U}(\bar{y}_0, h_0, T_0; y_{i-1}, h_{i-1}, T_{i-1}) \bar{U}(y_{i-1}, h_{i-1}, T_{i-1}; y_i, h_i, T_i)$$

for all $y_i \in \Lambda$ and a fixed $\bar{y}_0$. This can be achieved in more than one way. As a guideline, it is advisable to deviate the least possible from the definition (21.4).

As a next step in the construction, consider a second binomial process c_t which is piecewise constant across neighboring time points T_i. The dynamics of c_t is by construction correlated to that of h_t. More precisely, the propagator for the joint process is

$$W(h_i, c_i, T_i; h_{i+1}, c_{i+1}, T_{i+1}) =$$
$$q_{++}(h_i, c_i, i)\delta_{h_i+1,h_{i+1}}\delta_{c_i+1,c_{i+1}} + q_{+-}(h_i, c_i, i)\delta_{h_i+1,h_{i+1}}\delta_{c_i-1,c_{i+1}}$$
$$+ q_{-+}(h_i, c_i, i)\delta_{h_i-1,h_{i+1}}\delta_{c_i+1,c_{i+1}} + q_{--}(h_i, c_i, i)\delta_{h_i-1,h_{i+1}}\delta_{c_i-1,c_{i+1}}$$

where $q_{\pm\pm}(h_i, c_i, i) \geq 0$ and

$$(21.5) \qquad q_{++}(h_i, c_i, i) + q_{+-}(h_i, c_i, i) + q_{-+}(h_i, c_i, i) + q_{--}(h_i, c_i, i) = 1.$$

Due to the conditional recombination property, if $y_0 = \bar{y}_0$, then the conditional propagators resum with a simple formula

$$\sum_{h_t : h_T = 0, h_{T_j} = h_j} \sum_{y_1,..y_{j-1}} \prod_{k=1..j} W(h_i, c_i, T_i; h_{i+1}, c_{i+1}, T_{i+1}) \times$$
$$q_{h_{T_k} - h_{T_{k-1}}}(h_{T_{k-1}}) \bar{U}_{h_{T_k} - h_{T_{k-1}}}(y_{k-1}, h_{T_{k-1}}, T_{k-1}; y_k, h_{T_k}, T_k)$$
$$= \sum_{h_j} W(0, 0, T_0; h_j, c_j, T_j) \bar{U}(y_0, 0, T_0; y_j, h_j, T_j)$$
$$\equiv \tilde{U}(y_0, 0, T_0; y_j, c_j, T_j).$$

As a next step, the construction above is repeated for a number of different risk factors associated to N lattice processes $y_t^{(\alpha)}$, where $\alpha = 1, ..N$ all correlated to the process c_t, possibly in different ways. Then, conditioned to starting all processes at fixed lattice points $y_0^{(\alpha)} = \bar{y}_0^{(\alpha)}$ and conditioned to $c_{T_j} = c_j$, the multi-factor propagator factorizes into the product of conditional single factor propagators

$$(21.6) \qquad \prod_{\alpha=1...N} \tilde{U}(y_0^{(\alpha)}, 0, T_0; y_j^{(\alpha)}, c_j, T_j)$$

This is the key formula which we use to correlate processes while ensuring that numerical complexity increases only linearly with the number of factors N.

The construction can obviously be iterated and we can have a whole hierarchy of binomial processes conditioning each other to model multi-factor correlations.

22. Conclusion

We reviewed a new framework for Mathematical Finance and the theory of stochastic processes based on operator methods. The framework grew over the years through applied research and solving specific concrete problems in derivative pricing. As wrong ideas were weeded out, a coherent framework emerged and is now summarized in this review paper. From the numerical viewpoint, the methods rely on the ability to multiply efficiently large matrices as the main computational engine. Technologies for matrix manipulation are currently being developed at great pace. The emerging multi-core GPU chipsets will soon provide affordable teraflop performance for algorithms based on matrix-matrix manipulations such as ours. To both set the theory of stochastic processes on firm theoretical grounds and justify the empirically observed convergence rates, we derive pointwise kernel convergence estimates of a new type which explain the observed smoothing properties of the method. We introduce the notion of Abelian process to generalize the concept of stochastic integral and extend the classic Cameron-Feynman-Girsanov-Ito-Kac-Martin theorem in multiple directions. This theorem is here reinterpreted as a block-diagonalization scheme for large matrices corresponding to Abelian processes. We outline solution methods for Abelian path dependent options, a class we identified and which comprises the great majority of the path-dependent options currently traded. Important non-Abelian processes are also considered and discussed by means of a weaker but still effective method of block-factorizations. Furthermore, we also discuss a method for dynamic conditioning which applies to correlation products such as baskets and hybrid derivatives.

References

Abramowitz, M. and I.A. Stegun (1972). *Handbook of Mathematical Functions*. Dover, New York.

Ait-Sahalia, Y. (1996). Nonparametric Pricing of Interest Rate Derivative Securities. *Econometrica*.

Ait-Sahalia, Y., L.P. Hansen and J. A. Scheinkman (2005). Operator Methods for Continuous Time Markov Processes. *Journal of Economic Theory*.

Albanese, C. (2007a). Callable Swaps, Snowballs and Videogames. *preprint*.

Albanese, C. (2007b). Kernel Convergence Estimates for Diffusions with Continuous Coefficients. *arXiv:0711.0132v1 [math.NA]*.

Albanese, C. and A. Kusnetsov (2005). Transformations of Markov Processes and Classification Scheme for Solvable Driftless Diffusions. *preprint*.

Albanese, C. and A. Mijatovic (2006). Convergence Rates for Diffusions on Continuous-Time Lattices. *preprint, available at www.level3finance.com*.

Albanese, C. and A. Osseiran (2007). Moment Methods for Exotic Volatility Derivatives. *preprint*.

Albanese, C. and A. Vidler (2007). A Structural Credit-Equity Model for Bespoke CDOs. *Wilmott Magazine*.

Albanese, C. and M. Trovato (2005). A Stochastic Volatility Model for Callable cms Swaps and Translation Invariant Path Dependent Derivatives. *preprint*.

Albanese, C. and M. Trovato (2006). An Interest Rate Process with Stochastic Monetary Policy, Volatility Cube Calibration and Callable CMS Spread Range Accruals. *to be released*.

Albanese, C. and P. Jones (2007). Non-resonant block diagonalizations for abelian processes. *preprint*.

Albanese, C. and S. Lawi (2004). Laplace Transforms for Integrals of Markov Processes. *To appear in Markov Processes and Functional Analysis*.

Albanese, C. and X.O. Chen (2004a). Credit Barrier Models on a Continuous Time Lattice. *to appear in Quantitative Finance*.

Albanese, C. and X.O. Chen (2004b). Pricing Equity Default Swaps. *preprint*.

Albanese, C., G. Campolieti, P. Carr and A. Lipton (2001). Black-Scholes goes hypergeometric. *Risk* **14(12)**, 99–103.

Albanese, C., H. Lo and A. Mijatovic (2006a). Spectral Methods for Volatility Derivatives. *preprint*.

Albanese, C., O. Chen, A. Dalessandro and A. Vidler (2005-2006b). Dynamic Credit Correlation Modelling. *preprint*.

Bellman, R.E. (1957). *Dynamic Programming*. Princeton University Press.

Bishop, E. (1967). *Foundations of Constructive Analysis*. McGraw-Hill.

Bochner, S. (1955). *Harmonic Analysis and the Theory of Probability*. University of California Press.

Bridges, D. and F. Richman (1987). *Varieties of Constructive Mathematics*. Cambridge University Press.

Cameron, R.H. and W.T Martin (1949). Transformations of Wiener Integrals by Non-Linear Transformations. *Trans. Amer. Math. Soc.*

Doob, J.L. (1953). *Stochastic Processes*. John Wiley and Sons.

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity, with Estimates of the Variance of United Kingdom Inflation. *Econometrica* **50**, 987–1007.

Feller, W. (1961). *An Introduction to Probability Theory and its Applications*. Wiley, New York.

Feynman, R.P. (1948). Space-Time Approach to Non-Relativistic Quantum Mechanics. *Rev. Modern Phys.*

Giorno, V., A.G. Nobile, L.M. Ricciardi and L. Sacerdote (1988). Some remarks on the rayleigh process. *Journal of Applied Probability*.

Girsanov, I. V. (1960). On Transforming a Certain Class of Stochastic Processes by Absolutely Continuous Substitution of Measures. *Theory of Prob. and Appl.* pp. 285–301.

Glimm, J. and A. Jaffe (1987). *Quantum Physics. A Functional Integral Point of View*. Springer.

Goto, K. and R. van de Geijn (to appear). Anatomy of High-Performance Matrix Multiplication. *ACM Trans. Math. Soft.*

Hansen, L. P. and J. Scheinkman (1995). Back to the Future: Generating Moment Implications for Continuous-Time Markov Processes. *Econometrica*.

Hansen, L. P., J. Scheinkman and N. Touzi (1998). Spectral Methods for Identifying Scalar Diffusion Processes. *Journal of Econometrics*.

Ito, K. (1949). On a Formula Concerning Stochastic Differentials. *Trans. Amer. Math. Soc.*

Kac, M. (1948). On Some Connections Between Probability Theory and Differential Equations. *Rev. Modern Phys.*

Kato, T. (1976). *Perturbation theory for linear operators*. Springer-Verlag, Berlin.

Landau, L. D. and E. M. Lifshits (1977). *Quantum Mechanics: Non-relativistic Theory (3rd ed.)*. Pergamon, London.

Naimark, M.A. (1959). *Normed Rings*. Noordhoff, Groningen.

Reed, M. and B. Simon (1980). *Methods of Modern Mathematical Physics, Vol I-IV*. Academic Press.

Spanier, E.S. (1966). *Algebraic Topology*. McGraw-Hill.

Trefethen, Lloyd N. and Mark Embree (2006). *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators*. Princeton University Press.