



Centrum voor Wiskunde en Informatica

REPORTRAPPORT

SEN

Software Engineering



Software ENgineering

On the structure of Gaussian pricing models and
Gaussian Markov functional models

C.D.D. Neumann

REPORT SEN-R0219 NOVEMBER 30, 2002

CWI is the National Research Institute for Mathematics and Computer Science. It is sponsored by the Netherlands Organization for Scientific Research (NWO).

CWI is a founding member of ERCIM, the European Research Consortium for Informatics and Mathematics.

CWI's research has a theme-oriented structure and is grouped into four clusters. Listed below are the names of the clusters and in parentheses their acronyms.

Probability, Networks and Algorithms (PNA)

Software Engineering (SEN)

Modelling, Analysis and Simulation (MAS)

Information Systems (INS)

Copyright © 2001, Stichting Centrum voor Wiskunde en Informatica

P.O. Box 94079, 1090 GB Amsterdam (NL)

Kruislaan 413, 1098 SJ Amsterdam (NL)

Telephone +31 20 592 9333

Telefax +31 20 592 4199

ISSN 1386-369X

On the Structure of Gaussian Pricing Models and Gaussian Markov Functional Models

C.D.D. Neumann

CWI, P.O. Box 94079

1090 GB Amsterdam

The Netherlands

neumann@cwi.nl

ABSTRACT

This article investigates the structure of Gaussian pricing models (that is, models in which future returns are normally distributed). Although much is already known about such models, this article differs in that it is based on a formulation of the theory of derivative pricing in which numeraire invariance is manifest, extending earlier work on this subject. The focus on symmetry properties leads to a deeper insight in the structure of these models. The central idea is the construction of the most general class of derived Gaussian tradables given a set of underlying tradables which are themselves Gaussian. These derived tradables are called “generalized power tradables” and they correspond to portfolios in which the fraction of total value invested in each asset is a deterministic function of time. Applying this theory to Gaussian HJM models, the new tradables give an explicit description of the interdependence of bonds implicit in such models. Given this structure, a simple condition is derived under which these models allow a description in terms of an M -factor Markov functional model, as introduced by Hunt, Kennedy and Pelsser. Finally, conditions are derived under which these Gaussian Markov functional models are time homogeneous (bond volatilities depending only on the time to maturity). This result is linked to recent results by Björk and Gombani.

2000 Mathematics Subject Classification: 58J35, 58J70, 60H10, 91B24

Keywords and Phrases: option pricing, interest rate modeling, markov functional models, local scale invariance, partial differential equations

Note: Work carried out under theme SEN4 “Evolutionary Systems and Applied Algorithmics”

1. INTRODUCTION

Models in which future returns are normally distributed, known as Gaussian models, are a cornerstone in the theory of derivative pricing, despite their many well known deficiencies. The mathematical tractability of these models makes it possible to find analytical expressions for the prices of a large range of derivative contracts. For example, any contract which gives the right to exchange one asset for another at a fixed future date (e.g. plain vanilla European puts and calls) has a price in the familiar Black-Scholes-Merton form. Given the huge popularity of these models, it is important to have a clear picture of their structure. Although Gaussian market models can never be incomplete, they can be overcomplete, which implies that there exist deterministic relations between prices of different assets in the market. If this is the case, it is necessary to impose conditions on the model to exclude arbitrage possibilities. This is an issue that has been addressed in many articles. On the other hand, it is interesting to characterize the nature of these deterministic relations, given that the market is arbitrage free.

This article offers an investigation into the structure of Gaussian models from an alternative point of view. It builds on a formulation of pricing theory in which numeraire invariance (a local scale invariance) is manifest. This formulation was developed in a series of articles [HN01a, HN01b, HNV02].

The core of the formalism is the idea that pricing problems should be formulated only in terms of self-financing objects which are called tradables. From simple dimensional analysis one then finds that derivative prices must be homogeneous functions of degree one in these tradables. The consequent use of tradables as a proper parametrization of every object in a model that has a price makes numeraire invariance manifest, and pricing problems more transparent.

Following this line of thought, the main tool in the investigation of Gaussian models is the introduction of the most general class of derived Gaussian tradables that can be constructed given a set of underlying tradables which are themselves Gaussian. These derived tradables will be called “generalized power tradables”, and they correspond to a portfolio in which the fraction of total value invested in each of the underlying tradables is a deterministic function of time. In general, these new tradables are strongly path-dependent. Given these new tradables, the structure of Gaussian models is analyzed by looking at the subspace of linear relations between volatility vectors of tradables which respect numeraire invariance (“proper relations”). It is shown that, under some technical conditions, a Gaussian model can be parametrized by a basic subset of tradables for which no proper relations exist between their volatility vectors (they are “properly independent”). All tradables in the model can be expressed as generalized power tradables constructed from this basic set.

The theory is especially useful when applied to interest rate models, in particular to Gaussian Heath-Jarrow-Morton models, because it gives a clear picture of the deterministic relations between bonds that are implicit in such models. This makes it possible to derive a simple condition under which the current price of each tradable in the model depends only on the current prices of the basic set of tradables, not on their paths. The model then becomes a Markov functional model, in the sense of Hunt, Kennedy and Pelsser. Among these, models which are time homogeneous are of special interest. In these models, the bond volatilities are a function of the time to maturity only. It is shown that there exists an $M(M + 3)/2$ parameter family of time homogeneous models that allow a description as an M -factor Gaussian Markov Functional model.

2. REVIEW OF THE TRADABLE FORMALISM

The material in this section is a review from results appearing in Ref. [HN01a]. Details and proofs can be found in the original paper. Consider a market consisting of $N + 1$ tradables which is complete, and denote the price at time t of asset i , in some fixed but arbitrary numeraire, by $x_i(t)$ where $i = 0, \dots, N$. The prices in the market are assumed to be driven by K uncorrelated standard Wiener processes and satisfy, under the objective measure

$$\frac{dx_i(t)}{x_i(t)} = \mu_i(x, t)dt + \sum_{k=1}^K \sigma_i^k(x, t)dW_t^k \quad (2.1)$$

for functions μ and σ which guarantee existence and uniqueness of the processes at all times, where $x = (x_0, \dots, x_N)$. The central idea of the tradable formalism is that every object in the theory that has the dimension of value should be expressed in terms of the underlying tradables only, which is always possible because of the assumption of market completeness. Take for example the following model

$$\frac{dx_0}{x_0} = 0, \quad \frac{dx_1}{x_1} = \mu dt + (\sigma x_1)^\alpha dW$$

where μ, σ are constants and x_0 is taken as numeraire: both x_0 and x_1 are expressed in units of x_0 , which represents the unit of value. In other words, x_0 and x_1 have dimension $[x_0]$. It is simple to see that the parameter σ has implied dimension $[x_0]^{-1}$ so not everything in this model with the dimension of value is expressed in terms of the underlying tradables only. But it is clear how to correct this:

make the substitution $\sigma \rightarrow \sigma/x_0$. This leads to the equivalent model

$$\frac{dx_0}{x_0} = 0, \quad \frac{dx_1}{x_1} = \mu dt + \left(\sigma \frac{x_1}{x_0} \right)^\alpha dW$$

where both μ and σ are now dimensionless (with respect to the value dimension, they still have a time dimension). Now if the model Eq. (2.1) is formulated according to this principle, then the functions $\mu_i(x, t)$ and $\sigma_i^k(x, t)$ are necessarily homogeneous of degree zero in the variables x . Similarly, the value $V(x, t)$ of any derivative security depending on x must be homogeneous of degree one. By the Euler formula this is equivalent to

$$\sum_{i=0}^N x_i \frac{\partial V(x, t)}{\partial x_i} = V(x, t)$$

On the other hand, Itô's lemma gives

$$dV = (\mathcal{L}V)dt + \sum_{i=0}^N \frac{\partial V(x, t)}{\partial x_i} dx_i$$

where

$$\mathcal{L}V \equiv \frac{\partial V(x, t)}{\partial t} + \frac{1}{2} \sum_{i,j=0}^N \sum_{k=1}^K \sigma_i^k(x, t) \sigma_j^k(x, t) x_i x_j \frac{\partial^2 V(x, t)}{\partial x_i \partial x_j}$$

So if a claim price $V(x, t)$ satisfies the PDE $\mathcal{L}V = 0$ it is the homogeneity property of $V(x, t)$ which ensures that this claim can be replicated by a self-financing portfolio using delta hedging. The solutions of this PDE are invariant under simultaneous shifts of the volatility vectors $\sigma_i = (\sigma_i^1, \dots, \sigma_i^K)$ by an arbitrary vector λ

$$\sigma_i(x, t) \rightarrow \sigma_i(x, t) - \lambda(x, t) \quad \text{for all } i$$

Such transformations correspond to numeraire changes. In fact, taking $\lambda(x, t)$ equal to $\sigma_j(x, t)$ for some j corresponds to taking x_j as numeraire. This article will be focused on market models in which returns are Gaussian. In such models the volatility functions only depend on time, not on the tradables

$$\sigma_i(x, t) = \sigma_i(t) \quad \text{for all } i$$

In this case, asset prices are log-normally distributed, and it is possible to write down a very elegant formula for the price of a European type claim at time t , defined by a payoff $V(x, T)$ at maturity T . One first determines a singular value decomposition of the time integrated covariance matrix Σ_{ij}

$$\Sigma_{ij} \equiv \sum_{k=1}^K \int_t^T \sigma_i^k(u) \sigma_j^k(u) du = \sum_{r=1}^R \theta_i^r(t, T) \theta_j^r(t, T)$$

Here the dimension R of the vectors $\theta_i = (\theta_i^1, \dots, \theta_i^R)$ equals the rank of the matrix Σ_{ij} . The price of the claim is then given by

$$V(x, t) = \int V(x_0 \phi(z - \theta_0(t, T)), \dots, x_N \phi(z - \theta_N(t, T)), T) d^R z$$

where $z = (z_1, \dots, z_R)$ and

$$\phi(z) = \frac{1}{(\sqrt{2\pi})^R} \exp\left(-\frac{1}{2} \sum_{r=1}^R z_r^2\right)$$

3. LINEAR RELATIONS BETWEEN VOLATILITY VECTORS

In this section linear relations that might exist between volatility vectors in a Gaussian market model are examined. Such relations can in general be time-dependent. Let's start by looking at a fixed point in time t . Of course, every linear relation can be written in the form

$$\sum_{i=0}^N c^i(t) \sigma_i(t) = 0$$

for some vector $c(t) = (c^0(t), \dots, c^N(t))$. The set of all such relations forms a linear space L in a natural way. Linear relations are not necessarily invariant under numeraire changes. Indeed, consider a simultaneous shift of the volatility vectors $\sigma_i(t) \rightarrow \hat{\sigma}_i(t) = \sigma_i(t) - \lambda(t)$. The shifted vectors satisfy

$$\sum_{i=0}^N c^i(t) \hat{\sigma}_i(t) = - \left(\sum_{i=0}^N c^i(t) \right) \lambda(t)$$

so the relation will be numeraire independent if and only if

$$\sum_{i=0}^N c^i(t) = 0 \tag{3.1}$$

Vectors c satisfying this condition form a linear subspace $L_p \subset L$ which will be called the space of *proper* relations. By construction, the dimension of this subspace is a numeraire-independent quantity. A set of vectors for which no proper relations exist will be called *properly independent*. By abuse of terminology, a set of tradables will also be called properly independent if their volatility vectors are.

Relations which are not in L_p will be called *improper* relations. Now let $c_1(t), c_2(t)$ be two linearly independent improper relations. Then the linear combination

$$\left(\sum_{i=0}^N c_2^i(t) \right) c_1(t) - \left(\sum_{i=0}^N c_1^i(t) \right) c_2(t)$$

corresponds to a proper relation. This shows that

$$\dim L - \dim L_p \leq 1$$

So either $L = L_p$ or L is generated by L_p and exactly one improper relation. Whether or not an improper relation exists depends on the choice of a numeraire. In fact, it is always possible to find a numeraire such that an improper relation does exist. Simply choose a vector $c(t)$ not satisfying Eq. (3.1), then solve

$$\sum_{i=0}^N c^i(t) \hat{\sigma}_i(t) = \sum_{i=0}^N c^i(t) (\sigma_i(t) - \lambda(t)) = 0$$

for $\lambda(t)$. The result is

$$\lambda(t) = \frac{\sum_{i=0}^N c^i(t) \sigma_i(t)}{\sum_{i=0}^N c^i(t)}$$

Later in this article it will become clear that such a numeraire is related to a derivative security in the tradables. In general, it is a good idea to use a numeraire of this kind, since this tends to simplify calculations. Obviously, the simplest such choice is taking one of the tradables itself as numeraire.

Conversely, it is always possible to construct a numeraire in such a way that no improper relations exist. The simplest way to accomplish this is to extend the volatility vectors as follows:

$$(\sigma_i^1(t), \dots, \sigma_i^K) \rightarrow (\sigma_i^1(t), \dots, \sigma_i^K, 1)$$

This amounts to the introduction of an additional Wiener process. Obviously, the new volatility vectors describe the same market in a different numeraire, one that is not related to a derivative security in the tradables.

In the following sections the following trivial result will be used. Suppose that $\dim L_p = N - M$ for some M . Then there exists a subset of $M + 1$ volatility vectors which is properly independent. Relabeling the tradables if need be, one may assume that this is the set $\{\sigma_0(t), \dots, \sigma_M(t)\}$. The volatility vector of every tradable x_j where $j = 0, \dots, N$ can be expressed as a linear combination of vectors in this set as follows:

$$\sigma_j(t) = \sum_{i=0}^M \eta_j^i(t) \sigma_i(t) \quad (3.2)$$

where the coefficient vectors $\eta_j(t) = (\eta_j^0(t), \dots, \eta_j^M(t))$ are subject to

$$\sum_{i=0}^M \eta_j^i(t) = 1$$

The number M will be called the *rank* of the model at time t . This number equals the minimum number of independent Wiener processes necessary to describe the market model. It is a numeraire independent quantity and in general $M \leq K$.

As mentioned before, linear relations can be time-dependent. In fact, the rank M might in general not be constant in time. For the sake of simplicity, we will restrict our attention to models in which this rank *is* constant. Furthermore it will be assumed that there exists a fixed subset of tradables $\{x_0, \dots, x_M\}$ such that for all t in a time interval $[0, T^*]$ (the lifetime of some contract to be priced) every volatility vector $\sigma_j(t)$ can be written in the form of Eq. (3.2). In other words, the volatility vectors corresponding to this fixed subset are properly independent for $t \in [0, T^*]$. In the next section it will be shown that under these conditions every tradable can be considered as a (possibly path-dependent) derivative security whose value is completely fixed by its initial value and the basic set of tradables $\{x_0, \dots, x_M\}$.

4. GENERALIZED POWER TRADABLES

In this section a class of derived tradables is introduced that will be called generalized power tradables. The starting point is a properly independent set of $M + 1$ tradables $\{x_0, \dots, x_M\}$ satisfying

$$\frac{dx_i(t)}{x_i(t)} = \mu_i(t)dt + \sum_{k=1}^K \sigma_i^k(t) dW_t^k \quad (4.1)$$

The idea behind the new tradables is that they are constructed as a portfolio where at every point in time t a fraction $\eta^i(t)$ of the total value of the portfolio V is invested in tradable x_i . The fractions $\eta^i(t)$ are deterministic functions of time, only subject to the condition that

$$\sum_{i=0}^M \eta^i(t) = 1 \quad (4.2)$$

The hedge parameters can be expressed in terms of these fractions as

$$\phi^i(t) = \frac{\eta^i(t)V(t)}{x_i(t)}$$

It is then obvious that

$$\sum_{i=0}^M \phi^i(t)x_i(t) = V(t)$$

If the portfolio is to be self-financing, it should also satisfy

$$dV(t) = \sum_{i=0}^M \phi^i(t)dx_i(t)$$

and this is equivalent to

$$\frac{dV(t)}{V(t)} = \sum_{i=0}^M \eta^i(t) \frac{dx_i(t)}{x_i(t)} = \sum_{i=0}^M \eta^i(t) \left(\mu_i(t)dt + \sum_{k=1}^K \sigma_i^k(t)dW_t^k \right) \quad (4.3)$$

This shows that V is again a Gaussian tradable, and in fact it is the most general Gaussian tradable that can be constructed from the given set of underlying Gaussian tradables. Obviously, if one takes $\eta(t) = \eta_j(t)$ as defined in Eq. (3.2), then the generalized power tradable V has exactly the same volatility vector as x_j . But this implies that the two must be equal up to a multiplicative constant, for else there would be arbitrage opportunities. The constant of proportionality follows from specification of the values of the tradables at some fixed time. This shows explicitly that in a Gaussian model with rank M prices of all tradables can be expressed in terms of the price paths of a subset of $M + 1$ basic tradables. It also shows that the model is arbitrage free if and only if for each proper relation $c(t)$ the drift terms satisfy

$$\sum_{i=0}^N c^i(t)\mu_i(t) = 0$$

It is a simple exercise to show that this condition is necessary and sufficient for the existence of functions $\gamma^k(t)$ such that for all $i, j = 0, \dots, N$

$$\mu_i(t) - \mu_j(t) = \sum_{k=1}^K (\sigma_i^k(t) - \sigma_j^k(t))\gamma^k(t)$$

This is a symmetric definition for market prices of risk $\gamma^k(t)$ [HNV02]. So the existence of market prices of risk is necessary and sufficient to ensure the absence of arbitrage opportunities, a fact that is well known [MR97].

Now let's turn to the construction of an explicit formula for the value of the generalized power tradable V . Using Itô's lemma, Eq. (4.3) can be rewritten as (omitting t dependencies to simplify the notation)

$$d \ln V + \frac{1}{2} \frac{\langle V, V \rangle}{V^2} dt = \sum_{i=0}^M \left(d(\eta^i \ln x_i) - \frac{\partial \eta^i}{\partial t} \ln x_i dt + \frac{1}{2} \eta^i \frac{\langle x_i, x_i \rangle}{x_i^2} dt \right)$$

which in turn becomes

$$d \ln V = \sum_{i=0}^M \left(d(\eta^i \ln x_i) - \frac{\partial \eta^i}{\partial t} \ln x_i dt \right) + \frac{1}{4} \sum_{i,j=0}^M \sum_{k=1}^K \eta^i \eta^j (\sigma_i^k - \sigma_j^k)^2 dt$$

In this form, the equation can be integrated to give

$$V(t) = V(0)e^{\xi(t) - \Xi(t)} \prod_{i=0}^M \left(\frac{x_i(t)}{x_i(0)} \right)^{\eta^i(t)}$$

where

$$\xi(t) = \frac{1}{4} \sum_{i,j=0}^M \sum_{k=1}^K \int_0^t \eta^i(s) \eta^j(s) (\sigma_i^k(s) - \sigma_j^k(s))^2 ds$$

$$\Xi(t) = \sum_{i=0}^M \int_0^t \frac{\partial \eta^i(s)}{\partial s} \ln(x_i(s)) ds$$

It now becomes clear that the price $V(t)$ will in general depend upon the price history up to time t of the underlying tradables x_i via the term $\Xi(t)$. This path-dependence vanishes if and only if the functions $\eta^i(t)$ are constants in time. In this case, the tradable reduces to a power tradable, as introduced in [HN01b]. This justifies the name “generalized power tradable” for the general tradable.

It turns out that generalized power tradables play an important role in the valuation of geometric Asian options. Indeed, consider a market with two basic tradables x_0 and x_1 (for example a bond and a stock). Take

$$\eta^0(t) = \frac{t}{T}, \quad \eta^1(t) = \frac{T-t}{T}$$

Then the value of V at time T becomes

$$V(0) \exp \left(\xi(T) + \frac{1}{T} \int_0^T \ln \frac{x_1(s)}{x_0(s)} ds \right) \frac{x_0(T)}{x_0(0)}$$

and this equals, up to a deterministic factor, the continuously sampled geometric average of $x_1(s)/x_0(s)$ for $0 < s < T$ times $x_0(T)$. More details about the valuation of geometric average Asian options can be found in [HN01b].

5. GAUSSIAN HEATH-JARROW-MORTON MODELS

The most obvious application of this theory is in the field of interest rate modeling. Gaussian models are still very popular since they can be related directly to Black formulae, which constitute the market standard for valuation of plain-vanilla instruments. Of course, one has to make a choice which instruments, available in the market, will be the basic tradables. One can think of discount bonds, coupon bonds or swaps, depending on the type of contracts one would like to price. In this article, we will restrict our attention to the modeling of discount bond prices, but note that the theory is equally applicable to other choices.

Let $P(t, T)$ be the price at time t of a discount bond which pays one unit of currency, say dollars, at time T . The price is expressed in terms of some fixed but arbitrary numeraire. The most general Gaussian model for these bonds is given by

$$\frac{dP(t, T)}{P(t, T)} = \mu(t, T) + \sum_{k=1}^K \sigma^k(t, T) dW_t^k \quad (5.1)$$

where K is some positive integer. It is in fact a K -factor Gaussian Heath-Jarrow-Morton model. Since bonds cease to exist at maturity, the functions $\mu(t, T)$ and $\sigma(t, T)$ are defined only when $t \leq T$.

From the discussion in the previous sections it follows that in this model at most $K + 1$ bonds can be properly independent at any time. We will use similar assumptions as before, i.e. there exists a fixed set of $M + 1$ bonds ($M \leq K$) characterized by a set of maturities $\{T_0, \dots, T_M\}$ whose volatility vectors are properly independent for $t \in [0, T^*]$ (where $T^* \leq T_i$ for all i) and such that every $\sigma(t, T)$ can be written as

$$\sigma(t, T) = \sum_{i=0}^M \eta^i(t, T) \sigma(t, T_i)$$

where the vectors $\eta(t, T)$ satisfy

$$\sum_{i=0}^M \eta^i(t, T) = 1$$

Again using the results from the previous section one can write

$$P(t, T) = P(0, T) e^{\xi(t) - \Xi(t)} \prod_{i=0}^M \left(\frac{P(t, T_i)}{P(0, T_i)} \right)^{\eta^i(t, T)} \quad (5.2)$$

where

$$\begin{aligned} \xi(t) &= \frac{1}{4} \sum_{i,j=0}^M \sum_{k=1}^K \int_0^t \eta^i(s, T) \eta^j(s, T) (\sigma^k(s, T_i) - \sigma^k(s, T_j))^2 ds \\ \Xi(t) &= \sum_{i=0}^M \int_0^t \frac{\partial \eta^i(s, T)}{\partial s} \ln(P(s, T_i)) ds \end{aligned} \quad (5.3)$$

This explicitly shows the strong relations between bonds that are implicit in the use of a Gaussian HJM-model. The no-arbitrage conditions can be written as

$$\mu(t, T) = \sum_{i=0}^M \eta^i(t, T) \mu(t, T_i)$$

This is equivalent to the existence of functions $\gamma^k(t)$, market prices of risk, such that for all S, T one has

$$\mu(t, S) - \mu(t, T) = \sum_{k=1}^K (\sigma^k(t, S) - \sigma^k(t, T)) \gamma^k(t) \quad (5.4)$$

6. CASH, CASH BOND AND FORWARD RATES

According to the principles of the tradable formalism, every contract specification should be expressible in terms of tradable objects only. Of course, many contracts make reference to cash payments, which are expressed in terms of money. However, money is not a tradable object, since it is not self-financing. The way out is to observe that the bond $P(t, T)$ has a value of 1\$ at time T . So references to cash payments at time T should be replaced by references to the bond $P(t, T)$.

On the other hand, one can define a process which has a value of one unit of currency for all t by $C(t) \equiv P(t, t)$. This process is well defined in terms of the basic bonds, using Eq. (5.2)

$$C(t) = P(0, t) e^{\xi(t) - \Xi(t)} \prod_{i=0}^M \left(\frac{P(t, T_i)}{P(0, T_i)} \right)^{\eta^i(t, t)}$$

Note that $C(t)$ is not constant when an arbitrary numeraire is used. Closely related to $C(t)$ is the cash bond or money market account $B(t)$. This object is by definition a tradable, obtained by constantly reinvesting in bonds with the shortest time to maturity. By a no arbitrage argument it follows that it has to satisfy

$$\frac{dB(t)}{B(t)} = \mu(t, t)dt + \sum_{k=1}^K \sigma^k(t, t)dW_t^k \quad (6.1)$$

when its value is expressed in the same numeraire as the bonds. Unlike cash, the cash bond never appears in real contract specifications. It is of theoretical interest only. There is a simple relation between $B(t)$ and $C(t)$, which in fact defines the short rate $r(t)$

$$\frac{B(t)}{B(0)} = \exp\left(\int_0^t r(u)du\right) \frac{C(t)}{C(0)}$$

From this it follows that $C(t)$ satisfies

$$\frac{dC(t)}{C(t)} = (\mu(t, t) - r(t))dt + \sum_{k=1}^K \sigma^k(t, t)dW_t^k \quad (6.2)$$

In the original formulation of the HJM approach, not bond prices but forward rates are the objects being modeled. How is this related to the tradable approach? A numeraire-independent definition for forward rates is

$$B(t, T) = \exp\left(-\int_t^T f(t, u)du\right) C(t)$$

(with a factor $C(t)$ in the right hand side). Assume that the forward rates satisfy

$$df(t, T) = a(t, T)dt + \sum_{k=1}^K b^k(t, T)dW_t^k$$

Then consider the following object

$$I(t) = -\int_t^T f(t, u)du = \log \frac{P(t, T)}{C(t)}$$

On the one hand it satisfies [MR97]

$$dI(t) = \left(f(t, t) - \int_t^T a(t, u)du\right) dt - \sum_{k=1}^K \left(\int_t^T b^k(t, u)du\right) dW_t^k$$

On the other hand, using Eq. (5.1), Eq. (6.2) and Itô's lemma

$$\begin{aligned} dI(t) &= \left(\mu(t, T) - \mu(t, t) + r(t) - \frac{1}{2} \sum_{k=1}^K (\sigma^k(t, T)^2 - \sigma^k(t, t)^2)\right) dt \\ &+ \sum_{k=1}^K (\sigma^k(t, T) - \sigma^k(t, t)) dW_t^k \end{aligned}$$

Using the equality $r(t) = f(t, t)$ one finds the equalities

$$\begin{aligned} - \int_t^T a(t, u) du &= \mu(t, T) - \mu(t, t) - \frac{1}{2} \sum_{k=1}^K (\sigma^k(t, T)^2 - \sigma^k(t, t)^2) \\ - \int_t^T b^k(t, u) du &= \sigma^k(t, T) - \sigma^k(t, t) \end{aligned}$$

It is an easy check that the right hand sides are numeraire-invariant quantities, and consequently, the same holds for the left hand sides. From these equations it follows that the no-arbitrage condition Eq. (5.4) implies

$$a(t, T) = \sum_{k=1}^K b^k(t, T) \left(\int_t^T b^k(t, u) du - \sigma^k(t, t) + \gamma^k(t) \right)$$

with $\gamma^k(t)$ market prices of risk. This is the well known HJM no-arbitrage condition. Furthermore, by differentiation with respect to T one obtains

$$a(t, T) = -\partial_T \mu(t, T) + \sum_{k=1}^K \sigma^k(t, T) \partial_T \sigma^k(t, T), \quad b^k(t, T) = -\partial_T \sigma^k(t, T)$$

7. GAUSSIAN MARKOV FUNCTIONAL MODELS

Of special interest among interest rate models are so-called Markov functional models [HKP00]. These are models in which values of bonds (relative to each other) can be expressed as a functional of a Markov process, i.e.

$$\frac{P(t, T)}{P(t, T_0)} = \Psi(\alpha(t), t, T, T_0) \tag{7.1}$$

where Ψ is a deterministic function, and $\alpha(t) = (\alpha_1(t), \dots, \alpha_M(t))$ follows an M dimensional Markov process. The advantage of a Markov functional model is that it can be implemented efficiently, since it is only necessary to track the process $\alpha(t)$. In this section we will consider under what conditions a Gaussian HJM model can be described as a Markov functional model. Observe that by inversion, the process $\alpha(t)$ can be inferred from the prices of $M + 1$ different bonds (barring degenerate cases), characterized by a set of maturities $\{T_0, \dots, T_M\}$

$$\alpha(t) = f \left(t, \frac{P(t, T_1)}{P(t, T_0)}, \dots, \frac{P(t, T_M)}{P(t, T_0)} \right)$$

Inserting this in Eq. (7.1), it is possible to write

$$P(t, T) = \hat{\Psi}(P(t, T_0), \dots, P(t, T_M), t, T) \tag{7.2}$$

for some deterministic function $\hat{\Psi}$ which is homogeneous of degree one in the bond prices. In other words, every bond can be expressed as a derivative security in terms of a basic set of $M + 1$ bonds, and this derivative security is only allowed to depend on the prices of these basic bonds at time t ; there can be no path-dependency. If all bonds are to be Gaussian, it follows from the results of the previous sections that this derivative security must be a generalized power tradable. To exclude path-dependency, it is necessary that the path-dependent term Eq. (5.3) vanishes, which is the case iff the $\eta^i(t, T)$ are constant in t for all i and T . All in all, this shows that a Gaussian HJM model allows a description as an M dimensional Markov functional model if and only if the volatility vectors of bonds can be written in the form

$$\sigma(t, T) = \sum_{i=0}^M \eta^i(T) \sigma(t, T_i) \tag{7.3}$$

for functions $\eta(T)$ satisfying

$$\sum_{i=0}^M \eta^i(T) = 1 \quad (7.4)$$

and

$$\eta^i(T_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \quad \text{for all } i, j \quad (7.5)$$

This generalizes the result of [Ca94]. Given these relations, one has the following explicit form for Eq. (7.2)

$$P(t, T) = P(0, T) e^{\xi(t)} \prod_{i=0}^M \left(\frac{P(t, T_i)}{P(0, T_i)} \right)^{\eta^i(T)} \quad (7.6)$$

where

$$\xi(t) = \frac{1}{4} \sum_{i,j=0}^M \eta^i(T) \eta^j(T) \sum_{k=1}^K \int_0^t (\sigma^k(s, T_i) - \sigma^k(s, T_j))^2 ds$$

So every bond can be expressed as a power tradable in terms of the set of basic bonds.

8. TIME HOMOGENEOUS MODELS

A useful class of Gaussian interest rate models are the so-called time homogeneous models. For these models, the volatility vectors only depend on the time to maturity $\tau = T - t$, i.e.

$$\sigma(t, T) = \rho(T - t) = \rho(\tau)$$

or equivalently

$$(\partial_t + \partial_T)\sigma(t, T) = 0 \quad (8.1)$$

8.1 Markov conditions

An interesting question is under what conditions a time homogeneous model allows a description as an M -factor Markov functional model. This class of models was characterized in [BG99]. Here we re-derive their result. The claim is that a model belongs to this class if and only if there exist M real constants β_1, \dots, β_M such that the vector $\rho(\tau)$ satisfies

$$\frac{\partial^{M+1} \rho(\tau)}{\partial \tau^{M+1}} + \sum_{i=1}^M \beta_i \frac{\partial^i \rho(\tau)}{\partial \tau^i} = 0 \quad (8.2)$$

Indeed, such a model should satisfy both Eq. (7.3) and Eq. (8.1). By combining these two equations one gets

$$\sum_{i=0}^M \left(\eta^i(T) \frac{\partial \sigma(t, T_i)}{\partial t} + \frac{\partial \eta^i(T)}{\partial T} \sigma(t, T_i) \right) = 0 \quad (8.3)$$

for all t, T . In particular the equation holds for the set of times $T \in \{T_0, \dots, T_M\}$. So one can write, making use of Eq. (7.5)

$$\frac{\partial \sigma(t, T_i)}{\partial t} + \sum_{j=0}^M A_{ij} \sigma(t, T_j) = 0, \quad A_{ij} = \left. \frac{\partial \eta^j(T)}{\partial T} \right|_{T=T_i} \quad (8.4)$$

Now let $p(\lambda)$ be the characteristic polynomial of the matrix A

$$p(\lambda) = \det(\lambda \mathbf{1} - A) = \sum_{i=0}^{M+1} \beta_i \lambda^i$$

for some constants β_j with $\beta_{M+1} = 1$. It is well known that $p(A) = 0$. This fact can be used to derive a differential condition on the volatility vectors $\rho(\tau)$ (using Eqs. (7.3) and (8.4))

$$\begin{aligned} \sum_{i=0}^{M+1} \beta_i \frac{\partial^i \rho(\tau)}{\partial \tau^i} &= \sum_{i=0}^{M+1} (-1)^i \beta_i \frac{\partial^i \sigma(t, T)}{\partial t^i} \\ &= \sum_{i=0}^{M+1} \sum_{j=0}^M (-1)^i \beta_i \eta^j(T) \frac{\partial^i \sigma(t, T_j)}{\partial t^i} \\ &= \sum_{i=0}^{M+1} \sum_{j,k=0}^M \beta_i (A^i)_{jk} \eta^j(T) \sigma(t, T_k) = 0 \end{aligned}$$

It remains to consider Eq. (7.4). In view of Eq. (7.5) it is equivalent to the condition

$$\sum_{j=0}^M \frac{\partial \eta^j(T)}{\partial T} = 0 \quad \text{for all } T$$

Again, this holds for $T \in \{T_0, \dots, T_M\}$. Using the definition of A one gets

$$\sum_{j=0}^M \frac{\partial \eta^j(T)}{\partial T} \Big|_{T=T_i} = \sum_{j=0}^M A_{ij} = 0 \quad \text{for all } i$$

This shows that the matrix A must be degenerate. Consequently the coefficient β_0 in $p(\lambda)$ is zero. Therefore the volatility vectors in an M -factor time homogeneous Gaussian Markov functional model must satisfy Eq. (8.2). It is straightforward to show that the converse also holds. Indeed, Eq. (8.2) can be solved for general constants β_j . To do this, one first determines the roots of the characteristic polynomial

$$p(\lambda) = \lambda^{M+1} + \sum_{j=1}^M \beta_j \lambda^j = \prod_k (\lambda - \alpha_k)^{n_k} \quad (8.5)$$

For every root α_k with multiplicity n_k one finds the following linearly independent solutions of the equation

$$\tau^i e^{\alpha_k \tau}, \quad \text{for } 0 \leq i < n_k$$

Let's denote these solutions, in some particular order, by $\psi_0(\tau), \dots, \psi_M(\tau)$, such that $\psi_0(\tau) = 1$, which is a solution because $\alpha = 0$ is always a root. One can now check explicitly that any linear combination of these solutions can always be written in the form of Eq. (7.3). This is left as an exercise to the reader.

8.2 Finding $\eta(T)$

In this section an explicit solution for the vector $\eta(T)$ is constructed. It turns out to be completely fixed by the constants β_1, \dots, β_M . By substituting Eq. (8.4) back into Eq. (8.3) one finds

$$\sum_{i=0}^N \left(\frac{\partial \eta^i(T)}{\partial T} - \sum_{j=0}^N \eta^j(T) A_{ji} \right) \sigma(t, T_i) = 0$$

for all t, T . Without loss of generality one may assume that the $\sigma(t, T_i)$ are linearly independent. This implies that

$$\frac{\partial \eta^i(T)}{\partial T} - \sum_{j=0}^N \eta^j(T) A_{ji} = 0$$

Now an argument similar to that used in the previous section shows that $\eta(T)$ satisfies an equation identical in form to Eq. (8.2)

$$\frac{\partial^{M+1} \eta(T)}{\partial T^{M+1}} + \sum_{i=1}^M \beta_i \frac{\partial^i \eta(T)}{\partial T^i} = 0 \quad (8.6)$$

Therefore, each component of $\eta(T)$ is a linear combination of the functions $\psi_i(T)$. They are uniquely determined by Eq. (7.5). Indeed, a little linear algebra leads to

$$\eta^i(T) = \frac{\omega(T_0, \dots, T_{i-1}, T, T_{i+1}, \dots, T_M)}{\omega(T_0, \dots, T_M)}$$

where ω is given by the determinant

$$\omega(T_0, \dots, T_M) = \begin{vmatrix} \psi_0(T_0) & \dots & \psi_M(T_0) \\ \vdots & \ddots & \vdots \\ \psi_0(T_M) & \dots & \psi_M(T_M) \end{vmatrix}$$

Note that the solutions $\eta^i(T)$ are invariant under a change of basis in the space of solutions of Eq. (8.6), as they should be.

8.3 General time homogeneous model

The only numeraire changes $\sigma(t, T) \rightarrow \sigma(t, T) + \lambda(t)$ which respect the time homogeneity property are those for which $\lambda(t)$ is constant. This corresponds to the fact that the constant function $\psi_0(\tau) = 1$ is always a solution of Eq. (8.2). In order to classify time homogeneous models, it is useful to fix this remaining freedom. One way to do this is to demand that

$$\sigma(t, t) = \rho(0) = 0$$

In view of Eq. (6.1), this amounts to the choice of the money market account as numeraire. Obviously, under this restriction one is left with the following set of linearly independent solutions of Eq. (8.2)

$$\bar{\psi}_i(\tau) = \psi_i(\tau) - \psi_i(0)$$

where $1 \leq i \leq M$. Every component of the volatility vector $\rho(\tau)$ must be a linear combination of these basis functions

$$\rho^k(\tau) = \sum_{m=1}^M c_{km} \bar{\psi}_m(\tau)$$

for some constants c_{km} . So the most general M -factor time homogeneous Gaussian Markov functional model is described by an SDE of the form

$$\frac{dP(t, T)}{P(t, T)} = \mu(t, T)dt + \sum_{k=1}^K \sum_{m=1}^M c_{km} \bar{\psi}_m(T-t) dW_t^k \quad (8.7)$$

where the drift terms $\mu(t, T)$ are restricted by no-arbitrage conditions. They are not necessarily time homogeneous. On the other hand, they are irrelevant for derivative pricing. The description Eq. (8.7) is not yet unique. To make it unique, introduce a new basis in the space of Wiener processes

$$\bar{W}_t^m \equiv \sum_{k=1}^K c_{km} W_t^k$$

In terms of these, the model simply becomes

$$\frac{dP(t, T)}{P(t, T)} = \mu(t, T)dt + \sum_{m=1}^M \bar{\psi}_m(T-t)d\bar{W}_t^m$$

The processes \bar{W}_t^m are uniquely described by their covariance matrix

$$\langle \bar{W}_t^i, \bar{W}_t^j \rangle = \sum_{k=1}^K c_{ki}c_{kj} \equiv C_{ij}$$

Let's count the number of degrees of freedom. It equals the sum of M (choice of the β_i 's) and $M(M+1)/2$ (choice of a symmetric $M \times M$ matrix C_{ij}). This is

$$\frac{M(M+3)}{2}$$

9. SOME APPLICATIONS

9.1 One factor, the Vasicek model

For a one factor model, the characteristic polynomial Eq. (8.5) has two roots. One is always zero, the other will be called α . This leads to the following solutions of Eq. (8.2)

$$\psi_0(\tau) = 1, \quad \psi_1(\tau) = \begin{cases} \tau & \text{if } \alpha = 0 \\ e^{\alpha\tau} & \text{otherwise} \end{cases}$$

We will concentrate on the case $\alpha \neq 0$, since the case $\alpha = 0$ can be obtained as a limit of this. Fixed to the money market account numeraire, the solution becomes

$$\bar{\psi}_1(\tau) = e^{\alpha\tau} - 1$$

One recognizes the well known (extended) Vasicek model. It is usually parametrized like

$$\sigma(t, T) = \frac{\sigma(e^{\alpha(T-t)} - 1)}{\alpha} \tag{9.1}$$

with $\alpha < 0$. The vector $\eta(T)$ is given by

$$\eta(T) = \{\eta^0(T), \eta^1(T)\} = \left\{ \frac{e^{\alpha T_1} - e^{\alpha T}}{e^{\alpha T_1} - e^{\alpha T_0}}, \frac{e^{\alpha T} - e^{\alpha T_0}}{e^{\alpha T_1} - e^{\alpha T_0}} \right\}$$

and it is straightforward to check that the volatility vector can be written in the form of Eq. (7.3), i.e. in terms of the volatility vectors of two basic bonds with maturity T_0 and T_1 respectively

$$\sigma(t, T) = \eta^0(T)\sigma(t, T_0) + \eta^1(T)\sigma(t, T_1)$$

The functional relation for a general bond is given by Eq. (7.6), where

$$\xi(t) = \frac{\sigma^2}{4\alpha^3}(1 - e^{-2\alpha t})(e^{\alpha T_1} - e^{\alpha T})(e^{\alpha T} - e^{\alpha T_0})$$

In the limit $\alpha \rightarrow 0$ the model reduces to the Ho-Lee model, with

$$\begin{aligned}\sigma(t, T) &= \sigma(T - t) \\ \eta(T) &= \left\{ \frac{T_1 - T}{T_1 - T_0}, \frac{T - T_0}{T_1 - T_0} \right\} \\ \xi(t) &= \frac{\sigma^2}{2} t(T_1 - T)(T - T_0)\end{aligned}$$

If two factors are used in the construction, it is possible to model the empirically observed volatility hump. See e.g. [RC99].

9.2 Pricing an American option

In this section we consider a concrete example, the pricing of an American put option under stochastic interest rates. The aim is to show explicitly how this problem can be expressed entirely in terms of a reduced set of tradables. It can then be solved numerically. The put option is defined as follows. It gives the owner the right to sell one stock at any time t up to maturity T_0 for a strike price of K dollars, thereby terminating the contract. The profit is then

$$KC(t) - S(t) \tag{9.2}$$

where $C(t)$ is the cash process. This leads to a well-defined free boundary problem. The underlying model is as follows. The stock price $S(t)$ satisfies

$$\frac{dS(t)}{S(t)} = \mu_S dt + \sigma_S dW_1$$

while the bonds are modeled by the extended Vasicek model

$$\frac{dP(t, T)}{P(t, T)} = \mu(t, T)dt + \sigma(t, T)dW_2$$

In both cases the dynamics are defined with respect to the money market account numeraire under the objective measure. The function $\sigma(t, T)$ is defined by Eq. (9.1). To simplify matters, we will assume that W_1 and W_2 are not correlated.

Now it was shown in the previous section that all bonds can be written as a function of two basic bonds, for which we will use $P_0(t) \equiv P(t, T_0)$ and $P_1(t) \equiv P(t, T_1)$ with $T_1 > T_0$. This ensures that both bonds will exist during the entire lifetime of the option. It will be clear that the value of the put option can be written as a function of the stock and these two basic bonds. This follows from the fact that the early exercise function Eq. (9.2), which essentially defines the contract, can be expressed in terms of this set of tradables (and the initial term structure) using the relation

$$C(t) = P(t, t) = P(0, t)e^{\xi(t)} \left(\frac{P_0(t)}{P_0(0)} \right)^{\eta^0(t)} \left(\frac{P_1(t)}{P_1(0)} \right)^{\eta^1(t)}$$

with

$$\begin{aligned}\eta(t) &= \left\{ \frac{e^{\alpha T_1} - e^{\alpha t}}{e^{\alpha T_1} - e^{\alpha T_0}}, \frac{e^{\alpha t} - e^{\alpha T_0}}{e^{\alpha T_1} - e^{\alpha T_0}} \right\} \\ \xi(t) &= \frac{\sigma^2}{4\alpha^3} (1 - e^{-2\alpha t})(e^{\alpha T_1} - e^{\alpha t})(e^{\alpha t} - e^{\alpha T_0})\end{aligned}$$

It remains to derive the pricing PDE. For this, it is useful to rewrite the model, choosing one of the tradables as numeraire. We will use P_0 for this purpose. The reduced model now becomes

$$\begin{aligned}\frac{dS}{S} &= \mu_1(t)dt + \sigma_S dW_1 - \sigma(t, T_0)dW_2 \\ \frac{dP_0}{P_0} &= 0, \quad \frac{dP_1}{P_1} = \mu_2(t)dt + (\sigma(t, T_1) - \sigma(t, T_0))dW_2\end{aligned}$$

where $\mu_1(t), \mu_2(t)$ are certain drift functions which are irrelevant for the pricing problem. According to the discussion in section 2, this leads to the following pricing operator

$$\mathcal{L}V = \frac{\partial V}{\partial t} + \frac{1}{2}a_0(t)S^2 \frac{\partial^2 V}{\partial S^2} + a_1(t)SP_1 \frac{\partial^2 V}{\partial S \partial P_1} + \frac{1}{2}a_2(t)P_1^2 \frac{\partial^2 V}{\partial P_1^2}$$

where

$$\begin{aligned}a_0(t) &= \sigma_S^2 + \sigma(t, T_0)^2 \\ a_1(t) &= \sigma(t, T_0)(\sigma(t, T_0) - \sigma(t, T_1)) \\ a_2(t) &= (\sigma(t, T_1) - \sigma(t, T_0))^2\end{aligned}$$

The boundary conditions for an American put $V(S, P_0, P_1, t)$ are [MR97]

$$\begin{aligned}V(S, P_0, P_1, T_0) &= (KC(T_0) - S(T_0))^+ = (KP_0(T_0) - S(T_0))^+ \\ \mathcal{L}V(S, P_0, P_1, t) &\leq 0 \\ V(S, P_0, P_1, t) &\geq (KC(t) - S(t))^+ \\ (\mathcal{L}V(S, P_0, P_1, t) = 0) \vee (V(S, P_0, P_1, t) = (KC(t) - S(t))^+)\end{aligned}$$

This concludes the description of the problem in the reduced set of tradables. To numerically solve the problem, it is useful to reduce the dimension of the pricing PDE, making use of the homogeneity of $V(S, P_0, P_1, t)$. This is done by introducing new variables

$$s \equiv \frac{S}{P_0}, \quad p \equiv \frac{P_1}{P_0}, \quad v(s, p, t) \equiv \frac{V(S, P_0, P_1, t)}{P_0} = V(s, 1, p, t)$$

The pricing operator then simplifies to

$$\mathcal{L}v = \frac{\partial v}{\partial t} + \frac{1}{2}a_0(t)s^2 \frac{\partial^2 v}{\partial s^2} + a_1(t)sp \frac{\partial^2 v}{\partial s \partial p} + \frac{1}{2}a_2(t)p^2 \frac{\partial^2 v}{\partial p^2}$$

10. OUTLOOK

A possible extension to the theory presented here is to look at models in which prices are driven not only by diffusion but also by jump processes. It was shown in [HNV02] that it is possible to define power tradables in this setting if drifts, volatilities and jump sizes are deterministic functions of time. This suggests that the theory can be extended to such models. This is work in progress.

References

- [BG99] T. Björk and A. Gombani, *Minimal realizations of interest rate models*, Finance and Stochastics, **3**, (1999), 4, 413–432.
- [Ca94] A. Carverhill, *When is the spot rate Markovian?* Mathematical Finance, **4**, 305–312.
- [HN01a] J.K. Hoogland and C.D.D. Neumann, *Local Scale Invariance and Contingent Claim Pricing I*, IJTAF, **4**, (2001), 1, 1–21.
- [HN01b] J.K. Hoogland and C.D.D. Neumann, *Local Scale Invariance and Contingent Claim Pricing II: Path-dependent Contingent Claims*, IJTAF, **4**, (2001), 1, 23–43.
- [HNV02] J.K. Hoogland, C.D.D. Neumann and M.H. Vellekoop, *Symmetries in jump-diffusion models with applications in option pricing and credit risk*, Working paper, <http://www.cwi.nl/~neumann>.
- [HKP00] P. Hunt, J. Kennedy and A. Pelsser, *Markov-Functional Interest Rate Models*, Finance and Stochastics, **4**, (2000), 4, 391–408.
- [MR97] M. Musiela and M. Rutkowski, *Martingale Methods in Financial Modelling*, Springer-Verlag, (1997).
- [RC99] P. Ritchken and I. Chuang, *Interest Rate Option Pricing With Volatility Humps*, Review of Derivatives Research, **3**, (1999), 237–262.