THE HEATH-JARROW-MORTON THEOREM

EDITED March/Laurence

This chapter describes the fundamentals of term-structure modeling from the point of view of Arbitrage Pricing Theory. The main question that we address is how to construct a risk-neutral probability measure for the evolution of the the term-structure of interest rates. The term-structure is represented by the collection of discount factors

$$P_t^T \;=\; \mathbf{E}_t^P \left\{ e^{-\int_t^T r_s\, ds} \right\} \;,\quad T \;>\; t \;, \tag{1}$$

or, equivalently, by the instantaneous forward rates

$$f(t;T) \;=\; \lim_{\Delta t \,\to\, 0} \frac{1}{\Delta t} \left(\frac{\mathbf{E}_t^P \left\{ e^{-\int_t^T r_s\, ds} \right\}}{\mathbf{E}_t^P \left\{ e^{-\int_t^{T+\Delta t} r_s\, ds} \right\}} \;-\; 1 \right) \;=\; -\; \frac{\partial log(P_t^T)}{\partial T} \;. \tag{2}$$

The main theoretical result of the chapter is the Heath-Jarrow-Morton (HJM) charaterization of the risk-neutral measure for the evolution of forward rates. This can be seen as the analogue of the Black-Scholes-Merton characterization of the risk-neutral measures of for equities, in the fixed-income context. Following the derivation of the HJM theorem, we present a few examples which can be solved in closed-form. We then discuss "multifactor" models: i.e., the modeling of correlations between different sectors of the forward rate curve.

1. The Heath-Jarrow-Morton Theorem

In addition to the basic postulates of Arbitrage Pricing Theory we assume that

(i) agents have the highest credit quality and can borrow and lend at the same rates; and

(ii) at any time, there is a liquid market in zero-coupon bonds of arbitrary forward maturities (**"yield-curve completeness"**).

The last assumption is new. It postulates that the market is "complete" with respect to (optionless) loans of arbitrary maturities. The expression "arbitrary maturities" is understood here as meaning a *continuum of maturities* $T \leq T_{max}$. From a practical point of view, yield-curve completeness is tantamount to assuming that there exist enough debt instruments so that investors can lend and borrow money for arbitrary maturities $T, \ t < T < T_{max}$ at interests rates which are known at the date the loan is made (t). This assumption is well-suited for modeling mature debt markets, such as the Eurodollar/LIBOR swaps market, the U.S. Government bond market, and the debt markets of European Community member countries, all of which have liquid instuments with maturities spanning from a few months to over 15 years (interest-rate futures, bonds, bond futures, swaps, FRAs, bank loans, etc). The existence of such a large class of debt instruments allows traders to know, at each point in time and with high precision, any term rate on the yield curve. More precisely, although zero-coupon bonds may not actually be traded for all maturities, it is possible to *synthetize* them using other debt instruments and thus to price them by no-arbitrage considerations. In constrast, then concept of a continuous yield curve does not apply to markets where there are only a few traded instruments, like the debt markets in emerging economies. In the latter cases, forward prices may not be readily derived from those of existing instruments. Therefore, yield-curve completeness is not appropriate for modeling *any* debt market. Nevertheless, as we shall see, it is a powerful modeling tool.

Our goal is to characterize the evolution of the discount factors P_t^T , $t < T < T_{max}$ in a no-arbitrage economy. For this purpose, let us assume that there are ν risk-factors in this economy, represented by uncorrelated Brownian motions $W_i(t)$, $t > 0$, $i = 0, ...\nu$, and that P_t^T satisfy

(3) $$\frac{dP_t^T}{P_t^T} \ = \ \sum_{i=1}^{\nu} \sigma_i(t;T)\, dW_i(t) \ + \ \mu_i(t;T)\, dt \qquad t \ \leq T \ .$$

Here $\mu_i(t;T)$ and $\sigma_i(t;T)$ are, respectively, the instantaneous means and variances, labelled by the maturity date (T). We make no *a priori* assumptions regarding the dependence of these parameters on the underlying risk factors, rates, etc. In other words, the instantaneous mean and variance are allowed to depend on state-variables, interest rates, forward rates, and so on. The only restriction imposed is that these functions be adapted processes, i.e. that $\sigma_i(t;T)$ and $\mu_i(t;T)$ are known given the information available at time t. Essentially, we assume here the general setting of continuous-time finance, in which prices are Itô processes. We do not discuss the case in which there are discountinuities (jumps) in asset prices. Another restriction imposed on the dynamics of prices, implicit in equation (3), is the fact that there are finitely many (ν) risk factors. We postpone the discussion of how to determine ν and the specific forms of the parameters $\sigma_i(t;T)$.

A simple observation, which leads to the characterization of the risk-neutral measure, is the fact that discount factors are prices of zero-coupon bonds. From the assumption of yield-curve completeness, zeros are traded securities (or, more precisely, synthetizable using tradeable securities, which has the same consequences). Hence, we must have, under a risk-neutral measure,

$$\mu_i(t;T) \;=\; r_t \quad ,$$

where r_t is the short-term rate, *i.e.,*

$$\frac{dP_t^T}{P_t^T} \;=\; \sum_{i=1}^{\nu} \sigma_i(t;T)\, dW_i(t) \;+\; r_t\, dt \qquad t\; \leq T \;. \tag{4}$$

Here, we used the fact that zero-coupon bonds pay no dividends before maturity. Notice that the short rate r_t is itself determined by the discount curve, since

$$r_t \;=\; \lim_{\Delta t\,\to\,0} \frac{1}{\Delta t} \left(\frac{1}{P_t^{t+\Delta t}} \;-\; 1 \right) \;=\; -\; \left[\frac{\partial log(P_t^T)}{\partial T} \right]_{T\,=\,t} \tag{5}$$

To express the risk-neutral measure in more "transparent" form, we shall work with forward rates instead of discount factors. In view of equation (3), we shall first derive an SDE for $log(P_t^T)$. Applying Itô's Formula to this variable, we have

$$d\,log(P_t^T) \;=\; \sum_{i=1}^{\nu} \sigma_i(t;T)\, dW_i(t) \;+\; r_t\, dt \;-\; \frac{1}{2} \sum_{i=0}^{\nu} \sigma_i(t;T)\, \sigma_i(t;T)\, dt \;, \quad t\; \leq T \;.$$

Differentiating this equation with respect to T, and taking into equation account (2), we obtain

$$-\; d\,f(t;T) \;=\; \sum_{i=1}^{\nu} \dot\sigma_i(t;T)\, dW_i(t) \;-\; \left(\sum_{i=0}^{\nu} \dot\sigma_i(t;T)\, \sigma_i(t;T) \right) dt \;, \quad t\; \leq T \;, \tag{7}$$

where $\dot\sigma_i(t;T)$ represents the derivative of $\sigma_i(t;T)$ with respect to T. Using the fact that $-W_i(\cdot)$ is also distributed like a Brownian motion, we obtain in this way a stochastic differential equation that describes the evolution of the forward-rate curve under the risk-neutral measure, namely,

$$(8) \qquad d\,f(t;\,T) \;=\; \sum_{i=1}^{\nu} \dot\sigma_i(t;\,T)\,dW_i(t) \;+\; \left(\sum_{i=0}^{\nu} \dot\sigma_i(t;\,T) \int\limits_t^T \dot\sigma_i(t;\,s)\,ds \right) dt \quad t\ \le T\ ,$$

This equation is generally known as the **Heath-Jarrow-Morton equation** who first derived it in a series of celebrated papers (cf. HJM 1989, 1992).

The coefficient $\dot\sigma_i(t;\,T)$ represents the instantaneous standard deviation of the forward rate $f(t;\,T)$ contributed by the i^{th} factor. More precisely. these coefficients determine the *instantaneous covariance of forward rates*, since equation ( ) implies that

$$(9) \qquad \mathbf{Cov}\,(d\,f(t;\,T_1),\ d\,f(t;\,T_2)\,) \;=\; \sum_{i=1}^{\nu} \dot\sigma_i(t;\,T_1)\,\dot\sigma_i(t;\,T_2)\ dt$$

The HJM equation provides an exact relation that must hold between the instantaneous standard deviations and drifts of forward rates in the absence of arbitrage, assuming yield-curve completeness. It can be viewed as the analogue, in the context of fixed-income, of the characterization of the risk-neutral measure for stocks and commodities (drift= riskless rate minus dividend rate) derived by Black & Scholes [] and Merton []. We shall also refer to this charaterization as the **Heath-Jarrow-Morton theorem**.

Using the HJM equation, we can derive a differential equation that describes the evolution of the short rate r_t under the risk-neutral measure. For this purpose, let us write the HJM equation in integral form, *viz.*,

(10)

$$f(t;\,T) \;=\; f(0;\,T) \;+\; \int\limits_0^t \left(\sum_{i=1}^{\nu} \dot\sigma_i(s;\,T) \int\limits_s^T \dot\sigma(s;\,u)\,du \right) ds \;+\; \int\limits_0^t \sum_{i=1}^{\nu} \dot\sigma_i(s;\,T)\,dW_i(s)\ .$$

Recalling that $r_t \;=\; f(t;\,t)$, we have

$$(11) \quad r_t \;=\; f(0;\,t) \;+\; \int\limits_0^t \left(\sum_{i=1}^{\nu} \dot\sigma_i(s;\,t) \int\limits_s^t \dot\sigma(s;\,u)\,du \right) ds \;+\; \int\limits_0^t \sum_{i=1}^{\nu} \dot\sigma_i(s;\,t)\,dW_i(s)\ .$$

This equation shows that r_t is an Itô process under the risk-neutral measure, satisfying the stochastic differential equation

(12)
$$dr_t \;=\; \sigma_r(t)\; dZ(t) \;+\; \mu_r(t)\, dt\ ,$$

with mean

(13)
$$\mu_r(t) \;=\; \dot{f}(0;\, t) + \sum_{i=1}^{\nu} \int_0^t ds \left[\left(\dot{\sigma}_i(s;\, t)\right)^2 \;+\; \ddot{\sigma}_i(s;\, t) \int_s^t \dot{\sigma}_i(s;\, u)\, du \right] ds + \sum_{i=1}^{\nu} \ddot{\sigma}_i(s;\, t)\, dW_i(s)$$

(here $(\bullet)^{\cdot\cdot} \;=\; \frac{\partial^2 \bullet}{\partial T^2}$) and with standard deviation

$$\sigma_r(t) \;=\; \sqrt{\sum_{i=1}^{\nu} \left(\dot{\sigma}_i(t;\, t)\right)^2}\ .$$

These equations show, in particular, that the drift of the short rate is a function of the *past history* of forward rate volatility and thus possibly of the forward rates themselves, according to the specification of the volatility used. The latter observation leads to the interesting questions regarding the practical implementation of such models that will be discussed later.

The HJM theorem is interesting for the following reasons:

- it focusses on the forward rate curve and on the instantaneous covariance of forward rates, as opposed to more abstract state variables;

- It shows that the "risk premia", or "market prices of risk" $(\lambda_1(t), ...\lambda_\nu(t))$ introduced in the APT framework, are completely determined by the initial curve $f(0;\, T)$ and the instantaneous covariance matrix $\mathbf{Cov}\{df(t;\, T)\, df(t;\, T')\,\}$.

The HJM theorem shows that it is possible to model *directly* the evolution of the (observable) forward rates, as opposed to modeling the evolution of (often unobservable) "fundamental" state variables $(X_1(t),\, ...,\, X_N(t))$ from which the curve is *derived.* This represents a significant advantage. Moverover, from a computational point of view, the reader should recall that in a "state-variables" model, asset prices discount factors are *functions* of the state variables, obtained by solving Fokker-Plank partial differential equations. Since the cashflows of fixed-income securities are themselves functions of the zero-coupon bond prices, pricing derivative securities in the "state-variable" framework typically requires first solving auxilliary PDEs to determine the values of the payoffs, followed by another PDE for computing the present value of the payoffs. On the other hand, in the HJM framework, there is no need to perform auxiliary calculations (for pricing European style derivatives): the state-variables *are* the forward rates. This make the HJM paradigm very useful.

Summarizing, the "HJM approach" to term-structure modeling consists in specifying a structure for the instantaneous covariance of forward rates,($\dot{\sigma}_i(t;\, T)\,,\quad i = 1,\, 2,\, ...\nu$). Under the assumption of yield-curve completeness, this gives rise to a stochastic differential equation for the forward rate curve in a risk-neutral world. This equation can be used to calculate the present value of any cashflow that can be expressed as a function of the forward rate curve at any time in the future. Let us mention, however, that there are very few "HJM models" (i.e. models for the instantaneous covariance of forward rates) that give rise to analytic solutions for the HJM equations. In practice, the computation of expected cashflows is usually done by Monte Carlo simulation.

2. The Ho-Lee model

The simplest illustration of the HJM theorem arises in the case of a single risk-factor ($\nu = 1$) and a constant variance of forward rates, *i.e.*

$$\dot{\sigma}(t;\, T) \;=\; \sigma \;=\; \text{constant.}$$

In this case, the HJM equation becomes

$$df(t;\, T) \;=\; \sigma\, dW(t) \;+\; \left(\sigma \int_t^T \sigma\, ds \right) dt \;=\; \sigma\, dW(t) \;+\; \sigma^2\; (T \;-\; t)\; dt$$

This equation admits an explicit solution, obtained by integrating with respect to t, namely

$$f(t;T) \;=\; f(0;\, T) \;+\; \sigma\, W(t) \;+\; \sigma^2 \left(T\, t \;-\; \frac{1}{2}\, t^2 \right) \tag{14}$$

This model is known as the **Ho-Lee model** [].

Let us examine the corresponding dynamics of the foward rate curve implied by equation (14). Clearly, the deformations of the curve associated with the Ho-Lee model correspond to a "shift" of the initial forward rates by the linear shock

$$L(T) = \;\sigma\, W(t) \;+\; \sigma^2 \left(T\, t \;-\; \frac{1}{2}\, t^2 \right)$$

$$= \sigma\, W(t) \;+\; \sigma^2\, \frac{t^2}{2} \;+\; \sigma^2\, t\, (T \;-\; t)\;\; . \tag{15}$$

The slope of the perturbation is, from ( ), equal to $\sigma^2\, t$. In particular, the shock amplitude increases with maturity. The value of the line at the shortest maturity $T = t$ represents the shock to the short rate, which is given by

$$r_t \;=\; f(0;\, t) \;+\; \sigma\, W(t) \;+\; \frac{\sigma^2\, t^2}{2}\;.$$

Another way to express the structure of the forward rate curve under the Ho-Lee model is through the equation

$$f(t;T) \;=\; r_t \;+\; \sigma^2 t \cdot (T \;-\; t) \;+ f(0;T) \;-\; f(0;t) \;.$$

In particular, forward rates are perfectly correlated with the short term rate (and thus with each other). This is a common feature of all one-factor models ($\nu = 1$) since only one source of randomness drives all rates.

Perfect correlation is a limitation of one-factor models which makes them of little use for pricing complex derivatives such as options on the slope of the forward rate curve. In the extreme case of the Ho-Lee model, the difference between two forward rates observed at time t is, in fact,

$$f(t;T_1) \;-\; f(t;T_2) \;=\; f(0;T_1) \;-\; f(0;T_2) \;+\; \sigma^2 t\,(T_1 \;-\; T_2) \;,$$

a *deterministic* function. Market participants generally believe that forward rates are not perfectly correlated: an option on the difference between two forward rates (or two bond yields) should have a positive premium. Econometric analysis also points to an imperfect correlation between forward and spot rates (see Scott []). For instance, the correlation between spot 6-months LIBOR and 20-years forward 6-month LIBOR is believed to be approximately 40%. (This statement is about the "real world" probability, as opposed to the risk-neutral measure, but the price evidence usually supports imperfect correlation as well.) Another limitation of Ho-Lee, which we shall address in the next section, is the fact that shocks to the forward curve increase with maturity.

The Ho-Lee model is analytically tractable; this is its main theoretical interest. Recalling (eg. (4), (8)), the relation between zero-coupon bond volatility and forward rate volatilities, we find that the volatility of a zero coupon bond with maturity T is given by

$$\sigma(t;T) \;=\; \int_t^T \sigma\, ds \;=\; (T-t)\,\sigma \;.$$

Since the risk-neutral drift of a zero coupon bond is the short-term rate, we obtain the stochastic differential equation

$$\frac{d\,P_t^T}{P_t^T} \;=\; (T-t)\,\sigma\; dW(t) \;+\; \left(f(0;\,t) \;+\; \sigma\,W(t) \;+\; \frac{\sigma^2\,t^2}{2} \right)\, dt \;.$$

In the Ho-Lee model, the prices of zero-coupon bonds follow a lognormal diffusion with a volatility proportional to the to the time-to-maturity. An explicit expression for the price of a zero coupon bond maturing at date T given the forward rate furve at date 0 is

$$P_t^T \;=\; e^{-\int_t^T f(t;s)\,ds}$$

$$= e^{-\int_t^T \left(f(0;s) \, + \, \sigma W(s) \, + \, \sigma^2 \left(s\,t \, - \, \frac{1}{2} t^2 \right) \right) ds}$$

$$= P_0^T \cdot e^{-\int_t^T \sigma W(s)\, ds} \cdot e^{- \frac{\sigma^2 \, T \, t \, (T - t)}{2}}$$

This formula can be used, in turn, to compute the dynamics of swap rates and zero-coupon bonds, which can then be used to price interest rate derivatives according to this model. Details of option pricing calculations using the Ho-Lee model will be studied in a separate lecture.

3. Mean reversion: the modified Vasicek model

Another example of HJM dynamics that deserves attention is the case when $\nu = 1$ and forward rates are Gaussian with

$$\dot{\sigma}(t; T) \;=\; \sigma\, e^{-\,\kappa\,(T-t)}$$

where κ is a positive constant The idea behind this is that the volatility of forward rates is allowed to decrease with maturity. Let us investigate the consequences of this volatility assumption. The HJM evolution of forward rates is

$$df(t;T) \;=\; \sigma\, e^{-\kappa\,(T-t)}\, dW(t) \;+\; \left(\sigma\, e^{-\kappa\,(T-t)} \int_t^T \sigma\, e^{-\kappa\,(s-t)}\, ds \right) dt$$

$$= \sigma\, e^{-\kappa\,(T-t)}\, dW(t) \;+\; \frac{\sigma^2}{\kappa} \left(e^{-\kappa\,(T-t)} \;-\; e^{-2\,\kappa\,(T-t)} \right) dt \;.$$

Integrating with respect to t, we obtain

$$f(t;T) \;=\; f(0;T) \;+\; \sigma \int_0^t e^{-\kappa\,(T-s)}\, dW(s) \;-\; \frac{\sigma^2}{2\,\kappa^2} \left(1 \;-\; e^{-\kappa\,(T-t)} \right)^2$$

$$+\; \frac{\sigma^2}{2\,\kappa^2} \left(1 \;-\; e^{-\kappa\, T} \right)^2 \;.$$

Introducing the notation

$$X(t) \;=\; \int_0^t e^{-\kappa\,(t-s)}\; dW(s)\;,$$

and using the fact that

$$\int_0^t e^{-\kappa\,(T-s)}\; dW(s) \;=\; e^{-\kappa(T-t)}\, X(t)\;,$$

we obtain the following formulas for $f(t;\,T)$ and r_t:

$$f(t;\,T) \;=\; f(0;\,T) \;+ e^{-\kappa\,(T-t)}\, X(t) \;-\; \frac{\sigma^2}{2\,\kappa^2}\,\left(1 \;-\; e^{-\kappa\,(T-t)}\,\right)^2 \;+\; \frac{\sigma^2}{2\,\kappa^2}\,\left(1 \;-\; e^{-\kappa\,T}\,\right)^2\;,$$

and

$$r_t \;=\; f(0;\,t) \;+\; \sigma\; X(t) \;+\; \frac{\sigma^2}{2\,\kappa^2}\,\left(1 \;-\; e^{-\kappa\,t}\,\right)^2\;.$$

This is known as the **modified Vasicek model** (Vasicek (1977), Hull and White (1989), Heath-Jarrow-Morton (1992)). The formulas show that the forward rates and the instantaneous short rate are linear functions of the same Gaussian process $X(t)$: once again we observe perfect correlation of forward rates.

Nevertheless, there is an important difference with respect to the Ho-Lee model ($\kappa \;=\; 0$). In the latter case, the short rate tends to infinity as $t \;\to\infty$, whereas the short rate remains finite when κ is positive. More precisely, the short rate process behaves asymptotically for $t \;\gg 1$ like

$$r_t \;\approx\; f(0;\,t) \;+\; \sigma\; X(t) \;+\; \frac{\sigma^2}{2\,\kappa^2}\;;$$

which is a Gaussian random variable with mean

$$\mu_\infty(t) \;=\; f(0;\,t) \;+\; \frac{\sigma^2}{2\,\kappa^2}$$

and variance

$$\sigma_\infty^2 \;=\; \sigma^2\,\mathbf{E}\left\{\left(\int_0^t e^{-\kappa\,(t-s)}\;dW(s)\right)^2\right\}$$

$$=\; \sigma^2 \int_0^t e^{-2\,\kappa\,(t-s)}\,ds$$

$$\approx\; \frac{\sigma^2}{2\kappa} \qquad\qquad (\;t \;\gg\; 1)\;.$$

In the modified Vasicek model, short-rate fluctuations have a non-trivial *asymptotic probability distribution*. This fact is known as **mean-reversion** of the spot rate and κ is called the **rate of mean reversion**. At the level of forward-rates, notice that, for large T, we have

$$f(t;T) \;\approx\; f(0;T) \;+\; e^{\kappa\,(T-t)}\,X(t)\;,\quad T \;\gg\; 1\;,$$

The effect of having an exponentially decayng volatility of forward rates is to generate yield curve shocks that have exponentially decreasing amplitudes as the maturity dates increases: the "front end" more volatile than the "back end". The short-term rate is asymptotically in long-term equilibrium about the mean μ_∞. Notice that the variance of the asymptotic distribution is inversely proportional to the rate of mean reversion. Thus, for large κ , the equilibrium distribution of the short rate is sharply concentrated about the forward rate curve, with a small upward bias. For small rates of mean reversion, the short rate has wide fluctuations about the forward curve and a large positive bias.

As in the Ho-Lee model, zero-coupon bond prices are lognormal and can be computed explicitly. Thus, the model is quite tractable for computing option prices, as we shall see later on.

4. Multi-factor models: specifying the correlations of forward rates

In this section we discuss the specification of the correlation of forward rates. This question, including the determination of the number of factors (ν) , has been a central issue in "pure" interest-rate modeling. It has consequences in terms of pricing derivative securities that depend on the slope of the forward curve, such as swaptions and bond options.

Let us begin with a brief digression, considering a debt market with only a few traded instruments (e.g. three instruments: a short-term discount bond, 5- and and 30-year bonds). Here, we do not assume yield-curve completeness. In this case, the appropriate model has three factors and it is natural to use as state-variables the three bond prices, B_1, B_2, B_3. For instance, we can write

$$dB_i(t) \;=\; B_i(t)\,[\;\sigma_i(t)\,dZ_i \;+\; \mu_i\;dt\;]\;,\;\; i=1,\,2,\,...3\;, \tag{16}$$

where $B_i(t)$ represents the price of the i^{th} bond and Z_i are (possibly correlated) Brownian motions. To specify the "structure of the fluctuations ", we need to model the volatilities of the bonds and their correlations. Notice that the above equations can be rewritten in the form

$$dB_i(t) \;=\; B_i(t)\left[\;\sigma_i(t)\sum_{j=1}^{3}\beta_{ij}(t)\,dW_j \;+\; \mu_i\;dt\;\right]\;,\;\; i=1,\,2,\,...3\;,$$

where W_j are *independent* Brownian motions, and β_{ij} are coefficients such that

$$\sum_{j=1}^{3}\beta_{ij}^2 \;=\; 1\;.$$

In other words, we represent the "correlated" Brownian motions Z_i as linear combinations of independent ones (W_j):

$$dZ_i \;=\; \sum_{j=1}^{3}\beta_{ij}\,dW_j\;.$$

In this formulation, the instantaneous correlation of the returns of bonds i and k is given by

$$\mathbf{Corr}_t\left\{\frac{dB_i\,dB_k}{B_i\,B_k}\right\} \;=\; \mathbf{E}\,\{\,dZ_i\,dZ_k\,\} \;=\; \left(\sum_{j=1}^{3}\beta_{ij}\,\beta_{kj}\right)\;dt\;.$$

Specifying the correlation structure implies specifying the matrix β_{ij} (which may be time-dependent and may depend also on prices, for example) and the three volatilities. Notice, however, that the correlation structure may be such that the model has lower dimensionality: for instance, a correlation of 100% between the prices of bonds 1 and 2 can be modeled by choosing $\beta_{1\,j}$ and $\beta_{2\,j}$ so that

$$\sum_{j=1}^{3} \beta_{1j}\, \beta_{2j} \;=\; 1\ .$$

There are basically two paradigms for determining volatility/correlation parameters: historical data and data implied from current market prices. Let us describe the "historical" approach for estimating correlations. (Estimating "implied correlations" – a much difficult question – is beyond the scope of this lecture.)

Suppose that we generate from the 3 time-series of bond prices, an estimator for the instantaneous covariance matrix of returns, $\mathbf{Corr}\left\{ \frac{dB_i\, dB_k}{B_i\, B_k} \right\}$. Let $\widehat{C}_{ij}$ represent this estimator, in annualized terms (so we can omit dt from the calculations to follow). The specifics of how this matrix is constructed from data is an econometric problem that will not concern us here. (Let us mention, for the sake of definiteness, the simplest possible estimator – the Pearson sample correlation – which can be written as

$$\widehat{C}_{i\,j} \;=\; \frac{1}{\Delta t} \cdot \frac{\sum\limits_{\alpha=1}^{\mathrm{N}} \left(\frac{\Delta B_i^\alpha}{B_i^\alpha} \;-\; \overline{\frac{\Delta B_i}{B_i}} \right) \left(\frac{\Delta B_j^\alpha}{B_j^\alpha} \;-\; \overline{\frac{\Delta B_j}{B_j}} \right)}{\sqrt{\sum\limits_{\alpha=1}^{\mathrm{N}} \left(\frac{\Delta B_i^\alpha}{B_i^\alpha} \;-\; \overline{\frac{\Delta B_i}{B_i}} \right)^2} \; \sqrt{\sum\limits_{\alpha=1}^{\mathrm{N}} \left(\frac{\Delta B_j^\alpha}{B_j^\alpha} \;-\; \overline{\frac{\Delta B_j}{B_j}} \right)^2}}\ ,$$

where N represents the sample size, Δt is the time increment over which returns are computed, $\frac{\Delta B_i^\alpha}{B_i^\alpha}$ $\alpha = 1,\, 2 \ldots \mathrm{N}$ is the sequence of returns corresponding to bond i and $\overline{\frac{\Delta B_i}{B_i}}$ represents the corresponding sample mean.)

Since $\widehat{C}_{ij}$ is a symmetric non-negative matrix, it can be written in the form

$$\widehat{C}_{ij} \;=\; \sum_{k=1}^{3} \lambda_k^2\, v_i^{(k)}\, v_j^{(k)}\ ,$$

where $\lambda_k^2\,,\ k = 1, 2, 3$ are the eigenvalues and $\mathbf{v}^{(k)} \;=\; \left(v_1^{k)},\, v_2^{k)},\, v_3^{k)} \right)$ $k = 1, 2, 3$ are orthonormal eigenvectors. In this case, a reasonable estimator for β_{ij} is given by

$$\hat{\beta}_{ij} \;=\; \lambda_j\ \ v_i^{(j)}\ .$$

Indeed, the prescription

$$dZ_i \;=\; \sum_{j=1}^{3} \lambda_j\ \ v_i^{(j)}\ dW_j\ ,$$

with independent W_j leads to

$$\mathbf{E}\left\{dZ_i\, dZ_k\ \right\} \;=\; \mathbf{E}\left\{\left(\sum_{j=1}^{3}\ \lambda_j\ \ v_i^{(j)}\ dW_j\right)\ \left(\sum_{j=1}^{3}\ \lambda_j\ \ v_k^{(j)}\ dW_j\right)\right\}$$

$$=\ \sum_{k=1}^{3}\ \lambda_j^2\ \ v_i^{(j)}\, v_k^{(j)}\ dt\ =\ \widehat{C}_{i\,k}\, dt\ ,$$

which is consistent with the estimated correlation. The eigenvectors $\mathbf{v}^{(k)}$ represent the principal directions of change for the vector returns; they are called the **principal components** of the correlation matrix.

Several remarks are in order. First, notice that that the principal component analysis can lead in some cases to eivenvalues λ_k^2 which have very small magnitude, so that they may be considered insignificant from a statistical point of view. If this is the case, it makes sense to work with a lower-dimensional model based on one or two Brownian motions W_j. The **effective dimension** of the model, i.e. the number of principal components that reproduce the correlation matrix with good accuracy, is determined by the *rank* of the estimator of the correlation matrix.

The second remark concerns the distinction between *covariance* and *correlation* of returns. In the approach described here, we chose to model the correlation matrix. Of course, we could have chosen to estimate the covariance matrix using historical data instead. Estimating the covariance matrix and the corresponding eigenvalues and eigenvectors would lead to the same effective dimension. However, working with the covariance matrix is consistent with specifying the volatilities as well as the correlations. On the other hand, working with the correlation matrix gives us the freedom to specify the volatilities in a different way, e.g. using the implied volatilities of options. As we mentioned earlier, it is generally more difficult to "imply" correlations from market prices.

We now turn to the HJM model. We shall apply a similar principal component analysis (PCA) to estimate the effective dimension ν of and the correlation of forward rates.

Notice that the HJM model assumes a *continuum* of forward rates. This is not a serious modeling problem because the actual number of forward rates is actually finite, albeit large. This means that we have to "discretize" the HJM equations before making the PCA. Accordingly, let $f(t;T_i)\,,\ i\,=\,1,\,2\,..M$ be a discrete set of forward rates that is representative of the forward-rate curve as a whole. The Heath-Jarrow-Morton equations for these rates are

$$df(t;\,T_i)\ =\ \sum_{k=1}^{\nu}\ \dot{\sigma}_k(t;\,T_i)\, dW_k\ +\ \mu_i(t;T)\, dt$$

where ν is the (unknown) effective dimension and μ_i is the risk-neutral drift. In practice, and according to the market studied, this discretization may involve a between M =10 and M =120 dates T_i, the latter corresponding to 3-month forward rates from today to 30 years from today. We can rewrite the last equation in the form

$$df(t;\, T_i) \;=\; \sigma_f(t;\, T_i) \, \sum_{k=1}^{\nu} \beta_k(t;\, T_i)\, dW_k \;+\; \mu_i(t; T)\, dt \;, \tag{17}$$

where $\sigma_f(t;\, T_i) \;=\; \sqrt{\sum_{k=1}^{\nu} \dot{\sigma}_k^2(t; T_i)}$ represents the standard deviation of the i^{th} forward rate. The quantities $\beta_k(t;\, T_i)$ determine the correlation structure. Specifically, we have

$$\mathbf{Corr}_t \,\{ df(t;\, T_i)\, df(T;\, T_j) \,\} \;=\; \sum_{k=1}^{\nu} \beta_k(t;\, T_i)\, \beta_k(t;\, T_j)\; dt \;\equiv\; C_{i\,j}\, dt \;.$$

We consider, as before, a statistical estimator of the annualized correlation matrix, which we denote by $\widehat{C}_{i\,j}$. Diagonalization of this matrix leads to

$$\widehat{C}_{i\,j} \;=\; \sum_{k=1}^{M} \lambda_k^2\, v_i^{(k)}\, v_j^{(k)} \;,$$

for some sequence of eigenvalues λ_1^2, λ_2^2, ... λ_M^2 and orthonormal eigenvectors $\mathbf{v}^{(k)}$. This suggests taking

$$\hat{\beta}_k(t;\, T_i) \;=\; \lambda_k\, v_i^{(k)} \quad , \quad i = 1,\, 2,\, \ldots M$$

as estimators for $\beta_k(t;\, T_i)$ and the model for the correlation of the M rates follows.

An analysis of the importance of the eigenvalues λ_k, or more precisely, of the *relative importance* of these factors, by means of the quotients

$$\frac{\lambda_k^2}{\sum_j \lambda_j^2} \;, \qquad k \;=\; 1,\, 2,\, \ldots$$

can be used to determine the effective dimension of the model. In a pioneering paper, Litterman and Scheinkman (1980) were the first to conduct such study using U.S. Treasury bond data. Their work, as well as other that followed, show that a dimension of $\nu = 3$ is sufficient to explain the correlation structure to more than 99%. Specifically, Litterman and Sheinkman demonstrated, for the data considered, that if the eigenvalues are arranged in decreasing order, i.e., $\lambda_1 > \lambda_2 > ...$, then

$$\frac{\lambda_k^2}{\sum_k \lambda_k^2} < 0.01 \qquad \text{for } k \geq 4 .$$

We note,however, that the effective dimension varies considerably with the market under consideration. For instance, European central banks have actively managed short-term interest rates in recent years in view of the impending European Monetary Union. This has resulted in forward curves that are very different from the U.S. Treasury and U.S. LIBOR curves. We are not aware of statistical studies regarding the *stationarity* of the correlation of forward rates, i.e. that the correlation factors are functions of the time to maturity $T_i - t$.

From a modeling standpoint, the next step is to "interpolate" this result to the continuum of maturities $t < T < T_{max}$. Implicit in this assumption, is the fact that the forward rate curve is "smooth" in T. This is a belief held by most market practitioners (again, only in "mature" markets).

One way to justify the smoothness of the forward-rate curve is as follows. Suppose that the forward rate curve has a jump discontinuity at a given maturity T_0, at which, for instance, $f(t; T_0 - \epsilon) > f(t; T_0)$. In this case, investors can simultaneously lend forward for a short period ΔT at time $T_0 - \epsilon$ and borrow forward at date T for delivery at date $T_0 + \Delta T - \epsilon$ at a cheaper rate. The two loans offset each other perfectly and the investor can realize a profit after unwinding the trade. Notice that this argument assumes that an instantaneous forward rate corresponds to an actual rate for a (finite) period $\Delta T \ll 1$. This is therefore not a rigorous arbitrage argument, but it gives a strong indication as to why such jumps cannot exist in practice. Notice also that supply/demand for such loans would have the effect of lowering the higher forward rate and raising the lower one until the discontinuity dissapears. In the same vein, we note that sharp "kinks" in the curve would be conductive to arbitrage opportunities of a similar kind. In fact, if the differences between two succesive forwards

$$f(t; T) - f(t; T + \epsilon) , \quad \epsilon \text{ small}$$

exhibit a sharp discontinuity across some date T_0, investors can construct forward positions immediately before and after this date that would generate a riskless profit.

A possible approach to generating a "continuous time" correlation structure based upon the principal component analysis and the paradigm of of smoothness of the forward rate curve is to think of a particular eigenvector

$$\mathbf{v}^{(k)} = \left(v_1^{(k)}, v_2^{(k)}, ... v_M^{(k)}\right)$$

as representing a discrete sampling at points $x_i = T_i - t$ of a smooth function defined on the positive real line, i.e.,

$$\lambda_k \, v_i^{(k)} \; = \; \phi_k(T_i \; - \; t) \; , \quad k \; = \; 1, 2 \,..\, \nu \quad , \quad i = 1, \, 2, \, ... \, M \; .$$

Smooth functions ϕ_k can then be generated by a standard interpolation procedure, such as polynomial interpolation, spline functions, etc., under the normalization constraint

$$\sum_{k=1}^{\nu} \left(\phi_k(x) \right)^2 \; = \; 1.$$

The resulting HJM equation for the evolution of the curve has the form

$$df(t; T) \; = \; \sigma_f(t \, T) \, \sum_{k=1}^{\nu} \, \phi_k(T \, - \, t) \, dW_k \; + \; \mu(t; T) \, dt$$

where $\mu(t; T)$ is the HJM drift that makes the probability risk-neutral. The functions $\phi_k(x)$ are called the **correlation factors** of the model. The correlations between instantaneous forward rates obtained in this way are

$$\mathbf{Corr}_t \; \{ \; df(t; T) \, df(t; T') \; \} \; = \; \sum_{k=0}^{\nu} \phi_k(T \, - \, t) \, \phi_k(T' \, - \, t) \; , \qquad T, T' \; > \; t \; .$$

The reader should note that the correlation function generated in this way is time-homogeneous, in the sense that it depends on the differences $T - t$ and $T' - t$. This is a consequence of the simple statistical method used to estimate the correlation. Given the relative paucity of correlation data (compared with, volatility data, that is embedded in option prices), and the instability of correlation data due to different market conditions, the assumption of a time-homogeneous correlation is perhaps not sufficient to capture some market effects. Nevertheless, the procedure outlined here is based on market data and also constitutes a starting point for building more sophisticated models.

What can be said about the shapes of the correlation factors generated in this way? An important observation can be made in the case when all forward rates are *positively correlated*. This is generally true in the US dollar market. Under this additional assumption, we can use the fact that $\widehat{C}_{ij}$ which is a positive operator, i.e. an operator that (via matrix multiplication) sends vectors with non-negative coefficients into vectors with non-negative coefficients. It follows from a well-know theorem in matrix theory (Krein's Theorem), that the vector with largest eigenvalue is unique and has positive entries. Hence, *if forward rates are positively correlated, the factor with largest eigenvalue corresponds to a shock to the curve curve such that all forward rates change in the same direction.* By the same token, since the remaining eigenvectors are orthogonal to the latter, they must contain positive as well as negative entries. therefore *the remaining correlation factors correspond to shocks such that some sectors of the curve vary in opposite directions.* This qualitative behavior of the correlation factors is consistent with the Litterman-Scheinkman econometric results for the U.S. Dollar markets and with traders beliefs. Indeed, it is a widely held fact that the principal component, corresponding to this "parallel" shift in forward rates, explains at least 80% of the U.S. yield curve movement. The remaining correlation components are associated with secondary motions of the curve, such as "tilting" and "bending", the latter having the least importance in terms of the size of its eigenvalue. The interplay between different correlation components is analyzed in a simple example in the next section.

5. Example: a three-factor correlation structure

We present an example of a term-structure model which has three modes of fluctuation for the term-structure of interest rates and satisfies some of the stylized facts about the term structure mentioned above. This kind of model, in which the correlation factors are built from elementary functions (exponentials), is sometimes used by practitioners, by fitting the model to a set of correlations. Consider, for instance, the correlation factors

$$\phi_1 = \quad 0.8$$

$$\phi_2(x) \;=\; 0.5 \;-\; e^{-(0.1)\,x}$$

$$\phi_3(x) \;=\; \sqrt{1-(\phi_1(x))^2 \;-\; (\phi_2(x))^2} \;=\; \sqrt{0.11 \;+\; e^{-(0.1)\,x} - \; e^{-(0.2)\,x}} \;.$$

The first correlation factor coresponds to a "parallel shift". It contributes 64% of the variance of the curve (uniformly across maturities). The second correlation factor vanishes for

$$x_0 \;=\; 10\;log2 \;\approx\; 3.01 \quad \text{(yrs.)}$$

and varies from $\phi_1(0) \;=\; 0.5$ to $\phi(\infty) \;=\; -.05$. A "shock" to the curve along this factor will contribute to a "tilt" in the curve, since forward rates with maturities lying before and after x_0 move in opposite directions; its contribution to the variance of the forward rates is at most 25%. Finally, the third component satisifes $\phi_3(0) \;=\; \phi_3(\infty) \;=\; \sqrt{.11} \;=\; 0.332$ and has a global maximum at x_0 with $\phi_3(x_0) \;=\; 0.6$. It corresponds to a "bending mode" since it will move up the center of the curve keeping the ends fixed. It contributes between 11% and 25% of the variance of the rates, according to which maturity we consider.

This model is consistent with the following (hypothetical) table of correlations between forward rates:

	T=0	T=1	T=3	T=10	T=15
T=0					
T=1					
T=3					
T=10					
T=15					

Notice that the correlation between the spot rates and the forward rate 15 years hence is .43 (check this again). This number is approximately consistent with data on U.S. rates.

6. Conclusion: which volatility model to pick?

The simplicity of the HJM approach lies in the fact that the evolution of the curve in the risk-netural world can be described in terms of a stochastic differential equation. However, a closer inspection of the material of this chapter shows that there exist infinitely many models which are consistent with the HJM equations. There is no such thing as *a* "Heath-Jarrow-Morton model". HJM is a *paradigm* for thinking about the evolution of the collection of interest rates for different maturities in terms of a continuous curve of instantaneous forward rates. This approach centers the modeling effort in the selection of appropriate functions $\dot{\sigma}_i(t T),\ i = 1, 2, ...\nu$ to describe the fluctuations of the curve. As we have shown, the specification of these parameters implies "taking a view" of the nature of the correlation of forward rates of different maturities and on the number and shapes of the correlation factors.

Sicne we have already discussed the specification of correlations, let us turn to the issue of specification of the volatility $\sigma_f(t; T)$ in equation (17). Here, several choices present themselves, all of which lead to arbitrage-free dynamics. For instance, we can choose $\sigma_f(t; T)$ of the form

$$\sigma_f(t; T) \ = \ \tilde{\sigma} \, (\, t, \ T, \ f(t; \ T) \ ) \ ,$$

where $\tilde{\sigma}(\cdot,\ \cdot,\ \cdot)$ is a deterministic function. "Natural" choices for this function are

$$\sigma_f(t;\,T)\ = \sigma_0(t;\,T)\ (f(t;\,T)\,)^{\alpha} \tag{18}$$

where $\sigma_0(t;\,T)$ is a deterministic function of calendar time and the time-to-maturity and α is a positive exponent. The value $\alpha\ =\ 0$ corresponds to a Gaussian model and $\alpha\ =\ 1$ to a "lognormal" specification of volatility. Notice however that $\alpha\ =\ 1$ does not give rise to a lognormal rate under the risk-neutral measure because the corresponding HJM drift, which is (recall (8))

$$\mu_f(t;\,T)\ =\ \sum_{k=1}^{\nu}\ \sigma_0(t;\,T)\,\phi_k(T\,-\,t)\int_t^T f(t;\,s)\,\sigma_0(t;\,s)\,\phi_k(s\,-\,t)\,ds\ ,$$

depends on the forward rate curve. Functional forms of the type (18) are proposed and studied in Heath-Jarrow-Morton (1992) and in Morton (1988). It is fair to say, at this point, that *the only class of models of the type (18) for which the HJM stochastic differential equation can be solved in closed form are the Gaussian models.* There are however some drawbacks with Gaussian models, the most obvious one being that the model allows for interest rates to be negative with positive probability.

The functional form (18) with $\alpha = 1$ presents some interesting advantages for calibrating the model to (Black '76) implied volatility data for interest interest rate caps; this point will be discussed in further chapters. On the other hand it presents a serious pathology: forward rates can become infinite ("blow up") in finite time!

This pathology can be demonstrated by the following calculation, for the case $\nu = 1,\ \alpha = 1$ with $\sigma_0(t;\,T) = \sigma_0\ =$ constant. The solution of the corresponding HJM equation

$$\frac{df(t;\,T)}{f(t;\,T)}\ =\ \sigma_0\,dZ\ +\ \left(\sigma_0^2\int_t^T f(t;\,s)\,ds\right)\,dt$$

satisfies, formally,

$$f(t;\,T)\ =\ f(0;\,T)\,e^{\sigma_0\,Z(t)\ -\ \frac{\sigma_0^2\,t}{2}\ +\ \sigma_0^2\int_0^t ds\int_s^T f(s;\,u)\,du}$$

$$=\ f(0;\,T)\cdot\ M(t)\ \cdot\ e^{\int_0^t ds\ \sigma_0^2\int_s^T f(s;\,u)\,du}\ .$$

Assume, for simplicity, that $f(0; T) = f_0 =$ constant. Differentiating both sides of the last equation with respect to T, we obtain

$$\frac{\partial}{\partial T} f(t; T) = f(t; T)\, \sigma_0^2 \int_0^t f(s; T)\, ds \tag{19}$$

$$= \frac{\sigma_0^2}{2} \frac{\partial}{\partial t} \left(\int_0^t f(s; T)\, ds \right)^2 .$$

Hence, integration with respect to t, yields the differential equation

$$\frac{\partial}{\partial T} \int_0^t f(s; T)\, ds = \frac{\sigma_0^2}{2} \left(\int_0^t f(s; T)\, ds \right)^2 .$$

Setting

$$X(T) = \int_0^t f(s; T)\, ds ,$$

we conclude from equation (19) that

$$X(T) = \frac{2\, X(t)}{2 - \sigma_0^2\, X(t)\, (T - t)} .$$

Notice that $X(T)$ is a monotone function of T and that it blows up as T increases (along every Brownian path Z). More formally, suppose that that there exists some date $t_0 > 0$ such that the HJM equation admits a finite solution for all $s < t_0$ and all forward dates T such that $s < T < \infty$. We shall see that this leads to a contradiction. In fact this would imply that $X(t_0) < +\infty$ with probability 1. Suppose, then, that we select a constant k such that that the inequality $X(t_0) \geq k$ holds with nonzero probability $p(k)$. But then, by the above equation, we conclude that

$$X(T) > \frac{2\, k}{2 - \sigma_0^2\, k\, (T - t_0)}$$

with probability at least $p(k)$. Therefore, if we choose

$$T_0 \ = \ \frac{2}{\sigma_0^2 \, k} \ + \ t_0 \ ,$$

we conclude that for all $T \ \geq \ T_0$, we have $X(T) \ = \ + \, \infty$. This implies that some rate on the curve $\{f(t_0;\, T) \, , \ T \ > \ t\}$ must be infinite – this is a contradiction. Hence, the solution of the above model blows up instantaneously: *for every* $t \ > \ 0$ *the probability that the forward rates blow up is finite.*

Further computations along these lines, which we omit for the sake of brevity, show that the solutions of the HJM equation with the simple volatility model ( ) blow up in finite time if $\alpha \ \geq \ 1$ and reach zero in finite time if $\alpha \ < \ 1$. This example shows that care must be taken when modeling the volatility process to avoid undesirable properties for the risk-neutral measure.

Specifications of the volatility of forward rates of the form (18) have other drawbacks: why should the volatility of $f(t;T)$ be a function of the *single* rate $f(t;T)$ insead of a dependency on the *entire curve* $f(t;\, \theta) \ , \quad \theta \ \in \ [t, \, T_{max}]$? The answer to this question is simplicity fo the corresponding SDE. In practice, the choice of a volatility/ correlation structure is dictated by two main considerations:

- the model should satisfy "stylized" facts about the term structure, such as having enough degrees of freedom to model a rich enough set of forward curve scenarios;
- it should be computationally tractable;
- it should be relatively easy to calibrate to option prices.

These requirements are difficult to reconcile in practice. To gain a perspective on computational tractability of models of this type, notice that, in practice, $f(t;\, \bullet)$ is a high-dimensional vector (which can have dimension as large as 120 if we consider a 30-year forward rate curve with "instantaneous" forward rates replaced by 3-month forward rates). Thus, the state-space on which the curve lies is of very high dimension. This precludes, in most cases, the use of PDE/lattice models, given that current computers can handle 3 spatial dimensions at most, leaving Monte Carlo simulation as the only viable alternative for computing theoretical prices. However, lattice models are a necessity if we have to compute the values of interest-rate derivatives that have time-optionality, such as callable bonds, American-style bond options or American swaptions.

The issue of calibration is intimately related to computational tractability. For applications to the U.S./LIBOR market, it is desirable to have a model that prices a sufficiently rich class of caps and swaptions consistently with the market. This requirement often involves an interative search in "parameter space" (think of the calculation of implied volatility for BS) where computational efficiency is a necessity.

These considerations lead us to search for *reduced representations* of the forward-rate curve which correspond to an evolution in a low-dimensional space. This problem can be phrased as follows: assume that the number of factors is small (e.g. $\nu \ = \ 2$ or 3. Since the "dimension of the noise" is small, we would like to describe the forward rate curve as a function of a small number of parameters $\mathbf{X}(t) = (X_1(t),\ X_2(t),\ \ldots X_M(t))$, i.e.,

$$f(t;\ T) = F\left(\mathbf{X}(t),\ t,\ T\ \right) \ ,$$

where $F(\cdot,\ \cdot,\ \cdot)$ is a deterministic function, M is a number equal to or comparable to ν and $\mathbf{X}(t)$ is a diffusion process (a Markovian Itô process). In this case, the vector $\mathbf{X}$ acts as a "coordinate system" which allows to represent the curve in terms of a low dimensional object. Since an $\mathbf{R}^M$-valued diffusion process can be represented numerically in terms of a lattice model, volatility formulations that lead to a reduced representation are ideal in terms of computational tractability.

The Ho-Lee model and the Modified Vasicek models are examples of term-structure models that admit reduced representations (with M=1). In the Ho-Lee case, the underlying state variable can be taken to be $X(t) = W(t)$ and in the Modified Vasicek model we can take

$$X(t) = \int_0^t e^{-\kappa\,(t-s)}\,dW(s) \ ,$$

which is a diffusion process, since it satisfies the stochastic differential equation

$$dX = -\,\kappa\,X\,dt + dW \ .$$

The study of term-structure models which admit a reduced representation, also known as **Markov term-structure models** will be carried out in the next chapter.