the stable density above with $\alpha = 1/2$, so this case is also straightforward. Perhaps surprisingly, it is also fairly easy to sample from other stable distributions even though their densities are unknown. An important tool in sampling from stable distributions is the following representation: if V is uniformly distributed over $[-\pi/2, \pi/2]$ and W is exponentially distributed with mean 1, then

$$\frac{\sin(\alpha V)}{(\cos(V))^{1/\alpha}} \left(\frac{\cos((1-\alpha)V)}{W} \right)^{(1-\alpha)/\alpha}$$

has a symmetric α-stable distribution; see p.42 of Samorodnitsky and Taqqu [316] for a proof. As noted there, this reduces to the Box-Muller method (see Section 2.3.2) when $\alpha = 2$. Chambers, Mallows, and Stuck [79] develop simulation procedures based on this representation and additional transformations. Samorodnitsky and Taqqu [316], pp.46-49, provide computer code for sampling from an arbitrary stable distribution, based on Chambers et al. [79].

Feller [119], p.336, notes that the Lévy process generated by a symmetric stable distribution can be constructed through a random time change of Brownian motion. This also follows from the observation in Samorodnitsky and Taqqu [316], p.21, that a symmetric stable random variable can be generated as the product of a normal random variable and a positive stable random variable, a construction similar to (3.86).

3.6 Forward Rate Models: Continuous Rates

The distinguishing feature of the models considered in this section and the next is that they explicitly describe the evolution of the full term structure of interest rates. This contrasts with the approach in Sections 3.3 and 3.4 based on modeling the dynamics of just the short rate $r(t)$. In a setting like the Vasicek model or the Cox-Ingersoll-Ross model, the current value of the short rate determines the current value of all other term structure quantities — forward rates, bond prices, etc. In these models, the state of the world is completely summarized by the value of the short rate. In multifactor extensions, like those described in Section 3.3.3, the state of the world is summarized by the current values of a finite number (usually small) of underlying factors; from the values of these factors all term structure quantities are determined, at least in principle.

In the framework developed by Heath, Jarrow, and Morton [174] (HJM), the state of the world is described by the full term structure and not necessarily by a finite number of rates or factors. The key contribution of HJM lies in identifying the restriction imposed by the absence of arbitrage on the evolution of the term structure.

At any point in time the term structure of interest rates can be described in various equivalent ways — through the prices or yields of zero-coupon bonds or par bonds, through forward rates, and through swap rates, to name just a few examples. The HJM framework models the evolution of the term structure through the dynamics of the forward rate curve. It could be argued that forward rates provide the most primitive description of the term structure (and thus the appropriate starting point for a model) because bond prices and yields reflect averages of forward rates across maturities, but it seems difficult to press this point too far.

From the perspective of simulation, this section represents a departure from the previous topics of this chapter. Thus far, we have focused on models that can be simulated exactly, at least at a finite set of dates. In the generality of the HJM setting, some discretization error is usually inevitable. HJM simulation might therefore be viewed more properly as a topic for Chapter 6; we include it here because of its importance and because of special simulation issues it raises.

3.6.1 The HJM Framework

The HJM framework describes the dynamics of the forward rate curve $\{f(t,T), 0 \leq t \leq T \leq T^*\}$ for some ultimate maturity T^* (e.g., 20 or 30 years from today). Think of this as a curve in the maturity argument T for each value of the time argument t; the length of the curve shrinks as time advances because $t \leq T \leq T^*$. Recall that the forward rate $f(t,T)$ represents the instantaneous continuously compounded rate contracted at time t for riskless borrowing or lending at time $T \geq t$. This is made precise by the relation

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

between bond prices and forward rates, which implies

$$f(t,T) = -\frac{\partial}{\partial T} \log B(t,T). \tag{3.87}$$

The short rate is $r(t) = f(t,t)$. Figure 3.14 illustrates this notation and the evolution of the forward curve.

In the HJM setting, the evolution of the forward curve is modeled through an SDE of the form

$$df(t,T) = \mu(t,T)\, dt + \sigma(t,T)^\top\, dW(t). \tag{3.88}$$

In this equation and throughout, the differential df is with respect to time t and not maturity T. The process W is a standard d-dimensional Brownian motion; d is the number of *factors*, usually equal to 1, 2, or 3. Thus, while the forward rate curve is in principle an infinite-dimensional object, it is driven by a low-dimensional Brownian motion. The coefficients μ and σ in (3.88) (scalar and $\Re^d$-valued, respectively) could be stochastic or could depend on current

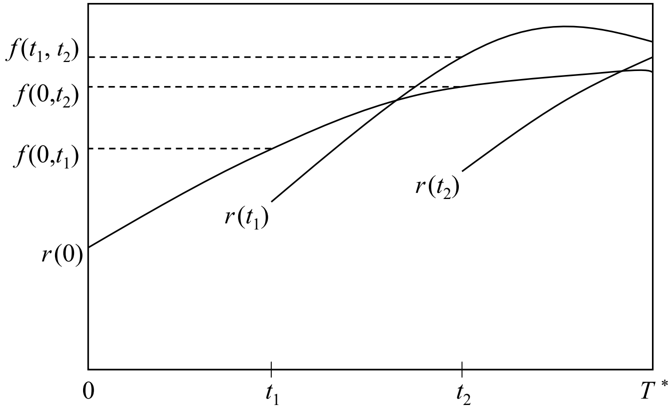

Fig. 3.14. Evolution of forward curve. At time 0, the forward curve $f(0,\cdot)$ is defined for maturities in $[0,T^*]$ and the short rate is $r(0) = f(0,0)$. At $t > 0$, the forward curve $f(t,\cdot)$ is defined for maturities in $[t,T^*]$ and the short rate is $r(t) = f(t,t)$.

and past levels of forward rates. We restrict attention to the case in which μ and σ are deterministic functions of t, $T \geq t$, and the current forward curve $\{f(t,u), t \leq u \leq T^*\}$. Subject to technical conditions, this makes the evolution of the curve Markovian. We could make this more explicit by writing, e.g., $\sigma(f,t,T)$, but to lighten notation we omit the argument f. See Heath, Jarrow, and Morton [174] for the precise conditions needed for (3.88).

We interpret (3.88) as modeling the evolution of forward rates under the risk-neutral measure (meaning, more precisely, that W is a standard Brownian motion under that measure). We know that the absence of arbitrage imposes a condition on the risk-neutral dynamics of asset prices: the price of a (dividend-free) asset must be a martingale when divided by the numeraire

$$\beta(t) = \exp\left(\int_0^t r(u)\,du\right).$$

Forward rates are not, however, asset prices, so it is not immediately clear what restriction the absence of arbitrage imposes on the dynamics in (3.88). To find this restriction we must start from the dynamics of asset prices, in particular bonds. Our account is informal; see Heath, Jarrow, and Morton [174] for a rigorous development.

To make the discounted bond prices $B(t,T)/\beta(t)$ positive martingales, we posit dynamics of the form

$$\frac{dB(t,T)}{B(t,T)} = r(t)\,dt + \nu(t,T)^\top dW(t), \quad 0 \leq t \leq T \leq T^*. \tag{3.89}$$

The bond volatilities $\nu(t,T)$ may be functions of current bond prices (equivalently, of current forward rates since (3.87) makes a one-to-one correspondence between the two). Through (3.87), the dynamics in (3.89) constrain the evolution of forward rates. By Itô's formula,

$$d \log B(t,T) = [r(t) - \tfrac{1}{2}\nu(t,T)^\top \nu(t,T)]\,dt + \nu(t,T)^\top dW(t).$$

If we now differentiate with respect to T and then interchange the order of differentiation with respect to t and T, from (3.87) we get

$$\begin{aligned} df(t,T) &= -\frac{\partial}{\partial T} d \log B(t,T) \\ &= -\frac{\partial}{\partial T}[r(t) - \tfrac{1}{2}\nu(t,T)^\top \nu(t,T)]\,dt - \frac{\partial}{\partial T}\nu(t,T)^\top dW(t). \end{aligned}$$

Comparing this with (3.88), we find that we must have

$$\sigma(t,T) = -\frac{\partial}{\partial T}\nu(t,T)$$

and

$$\mu(t,T) = -\frac{\partial}{\partial T}[r(t) - \tfrac{1}{2}\nu(t,T)^\top \nu(t,T)] = \left(\frac{\partial}{\partial T}\nu(t,T)\right)^\top \nu(t,T).$$

To eliminate $\nu(t,T)$ entirely, notice that

$$\nu(t,T) = -\int_t^T \sigma(t,u)\,du + \text{constant}.$$

But because $B(t,T)$ becomes identically 1 as t approaches T (i.e., as the bond matures), we must have $\nu(T,T) = 0$ and thus the constant in this equation is 0. We can therefore rewrite the expression for μ as

$$\mu(t,T) = \sigma(t,T)^\top \int_t^T \sigma(t,u)\,du; \tag{3.90}$$

this is the risk-neutral drift imposed by the absence of arbitrage. Substituting in (3.88), we get

$$df(t,T) = \left(\sigma(t,T)^\top \int_t^T \sigma(t,u)\,du\right) dt + \sigma(t,T)^\top\,dW(t). \tag{3.91}$$

This equation characterizes the arbitrage-free dynamics of the forward curve under the risk-neutral measure; it is the centerpiece of the HJM framework.

Using a subscript $j = 1, \ldots, d$ to indicate vector components, we can write (3.91) as

$$df(t,T) = \sum_{j=1}^d \left(\sigma_j(t,T) \int_t^T \sigma_j(t,u)\,du\right) dt + \sum_{j=1}^d \sigma_j(t,T)\,dW_j(t). \tag{3.92}$$

This makes it evident that each factor contributes a term to the drift and that the combined drift is the sum of the contributions of the individual factors.

In (3.91), the drift is determined once σ is specified. This contrasts with the dynamics of the short rate models in Sections 3.3 and 3.4 where parameters of the drift could be specified independent of the diffusion coefficient without introducing arbitrage. Indeed, choosing parameters of the drift is essential in calibrating short rate models to an observed set of bond prices. In contrast, an HJM model is automatically calibrated to an initial set of bond prices $B(0,T)$ if the initial forward curve $f(0,T)$ is simply chosen consistent with these bond prices through (3.87). Put slightly differently, calibrating an HJM model to an observed set of bond prices is a matter of choosing an appropriate initial condition rather than choosing a parameter of the model dynamics. The effort in calibrating an HJM model lies in choosing σ to match market prices of interest rate *derivatives* in addition to matching bond prices.

We illustrate the HJM framework with some simple examples.

Example 3.6.1 *Constant σ.* Consider a single-factor ($d = 1$) model in which $\sigma(t,T) \equiv \sigma$ for some constant σ. The interpretation of such a model is that each increment $dW(t)$ moves all points on the forward curve $\{f(t,u), t \le u \le T^*\}$ by an equal amount $\sigma\, dW(t)$; the diffusion term thus introduces only parallel shifts in the forward curve. But a model in which the forward curve makes only parallel shifts admits arbitrage opportunities: one can construct a costless portfolio of bonds that will have positive value under every parallel shift. From (3.90) we find that an HJM model with constant σ has drift

$$\mu(t,T) = \sigma \int_t^T \sigma\, du = \sigma^2(T-t).$$

In particular, the drift will vary (slightly, because σ^2 is small) across maturities, keeping the forward curve from making exactly parallel movements. This small adjustment to the dynamics of the forward curve is just enough to keep the model arbitrage-free. In this case, we can solve (3.91) to find

$$\begin{aligned} f(t,T) &= f(0,T) + \int_0^t \sigma^2(T-u)\, du + \sigma W(t) \\ &= f(0,T) + \tfrac{1}{2}\sigma^2[T^2 - (T-t)^2] + \sigma W(t). \end{aligned}$$

In essentially any model, the identity $r(t) = f(t,t)$ implies

$$dr(t) = df(t,T)\bigg|_{T=t} + \frac{\partial}{\partial T} f(t,T)\bigg|_{T=t} dt.$$

In the case of constant σ, we can write this explicitly as

$$dr(t) = \sigma\, dW(t) + \left(\frac{\partial}{\partial T} f(0,T)\bigg|_{T=t} + \sigma^2 t \right) dt.$$

Comparing this with (3.50), we find that an HJM model with constant σ coincides with a Ho-Lee model with calibrated drift. □

Example 3.6.2 *Exponential* σ. Another convenient parameterization takes $\sigma(t,T) = \sigma\exp(-\alpha(T-t))$ for some constants $\sigma, \alpha > 0$. In this case, the diffusion term $\sigma(t,T)\,dW(t)$ moves forward rates for short maturities more than forward rates for long maturities. The drift is given by

$$\mu(t,T) = \sigma^2 e^{-\alpha(T-t)} \int_t^T e^{-\alpha(T-u)}\,du = \frac{\sigma^2}{\alpha}\left(e^{-2\alpha(T-t)} - e^{-\alpha(T-t)}\right).$$

An argument similar to the one used in Example 3.6.1 shows that the short rate in this case is described by the Vasicek model with time-varying drift parameters.

This example and the one that precedes it may be misleading. It would be incorrect to assume that the short rate process in an HJM setting will always have a convenient description. Indeed, such examples are exceptional. □

Example 3.6.3 *Proportional* σ. It is tempting to consider a specification of the form $\sigma(t,T) = \tilde{\sigma}(t,T)f(t,T)$ for some deterministic $\tilde{\sigma}$ depending only on t and T. This would make $\tilde{\sigma}(t,T)$ the volatility of the forward rate $f(t,T)$ and would suggest that the distribution of $f(t,T)$ is approximately lognormal. However, Heath et al. [174] note that this choice of σ is inadmissible: it produces forward rates that grow to infinity in finite time with positive probability. The difficulty, speaking loosely, is that if σ is proportional to the level of rates, then the drift is proportional to the rates *squared*. This violates the linear growth condition ordinarily required for the existence and uniqueness of solutions to SDEs (see Appendix B.2). Market conventions often presuppose the existence of a (proportional) volatility for forward rates, so the failure of this example could be viewed as a shortcoming of the HJM framework. We will see in Section 3.7 that the difficulty can be avoided by working with simple rather than continuously compounded forward rates. □

Forward Measure

Although the HJM framework is usually applied under the risk-neutral measure, only a minor modification is necessary to work in a forward measure. Fix a maturity T_F and recall that the forward measure associated with T_F corresponds to taking the bond $B(t,T_F)$ as numeraire asset. The forward measure P_{T_F} can be defined relative to the risk-neutral measure P_β through

$$\left(\frac{dP_{T_F}}{dP_\beta}\right)_t = \frac{B(t,T_F)\beta(0)}{\beta(t)B(0,T_F)}.$$

From the bond dynamics in (3.89), we find that this ratio is given by

$$\exp\left(-\tfrac{1}{2}\int_0^t \nu(u,T_F)^\top \nu(u,T_F)\,du + \int_0^t \nu(u,T_F)^\top\,dW(u)\right).$$

By the Girsanov Theorem, the process W^{T_F} defined by

$$dW^{T_F}(t) = -\nu(t, T_F)^\top dt + dW(t)$$

is therefore a standard Brownian motion under P_{T_F}. Recalling that $\nu(t,T)$ is the integral of $-\sigma(t,u)$ from $u = t$ to $u = T$, we find that the forward rate dynamics (3.91) become

$$\begin{aligned} df(t,T) &= -\sigma(t,T)^\top \nu(t,T)\, dt + \sigma(t,T)^\top [\nu(t,T_F)^\top dt + dW^{T_F}(t)] \\ &= -\sigma(t,T)^\top [\nu(t,T) - \nu(t,T_F)]\, dt + \sigma(t,T)^\top dW^{T_F}(t) \\ &= -\sigma(t,T)^\top \left(\int_T^{T_F} \sigma(t,u)\, du \right) dt + \sigma(t,T)^\top dW^{T_F}(t), \end{aligned} \tag{3.93}$$

for $t \le T \le T_F$. Thus, the HJM dynamics under the forward measure are similar to the dynamics under the risk-neutral measure, but where we previously integrated $\sigma(t,u)$ from t to T, we now integrate $-\sigma(t,u)$ from T to T_F. Notice that $f(t,T_F)$ is a martingale under P_{T_F}, though none of the forward rates is a martingale under the risk-neutral measure.

3.6.2 The Discrete Drift

Except under very special choices of σ, exact simulation of (3.91) is infeasible. Simulation of the general HJM forward rate dynamics requires introducing a discrete approximation. In fact, each of the two arguments of $f(t,T)$ requires discretization. For the first argument, fix a time grid $0 = t_0 < t_1 < \cdots < t_M$. Even at a fixed time t_i, it is generally not possible to represent the full forward curve $f(t_i,T)$, $t_i \le T \le T^*$, so instead we fix a grid of maturities and approximate the forward curve by its value for just these maturities. In principle, the time grid and the maturity grid could be different; however, assuming that the two sets of dates are the same greatly simplifies notation with little loss of generality.

We use hats to distinguish discretized variables from their exact continuous-time counterparts. Thus, $\hat{f}(t_i,t_j)$ denotes the discretized forward rate for maturity t_j as of time t_i, $j \ge i$, and $\hat{B}(t_i,t_j)$ denotes the corresponding bond price,

$$\hat{B}(t_i,t_j) = \exp\left(-\sum_{\ell=i}^{j-1} \hat{f}(t_i,t_\ell)[t_{\ell+1} - t_\ell] \right). \tag{3.94}$$

To avoid introducing any more discretization error than necessary, we would like the initial values of the discretized bonds $\hat{B}(0,t_j)$ to coincide with the exact values $B(0,t_j)$ for all maturities t_j on the discrete grid. Comparing (3.94) with the equation that precedes (3.87), we see that this holds if

$$\sum_{\ell=0}^{j-1} \hat{f}(0,t_\ell)[t_{\ell+1} - t_\ell] = \int_0^{t_j} f(0,u)\, du;$$

i.e., if

$$\hat{f}(0, t_\ell) = \frac{1}{t_{\ell+1} - t_\ell} \int_{t_\ell}^{t_{\ell+1}} f(0, u)\, du, \qquad (3.95)$$

for all $\ell = 0, 1, \ldots, M-1$. This indicates that we should initialize each $\hat{f}(0, t_\ell)$ to the *average* level of the forward curve $f(0, T)$ over the interval $[t_\ell, t_{\ell+1}]$ rather than, for example, initializing it to the value $f(0, t_\ell)$ at the left endpoint of this interval. The discretization (3.95) is illustrated in Figure 3.15.

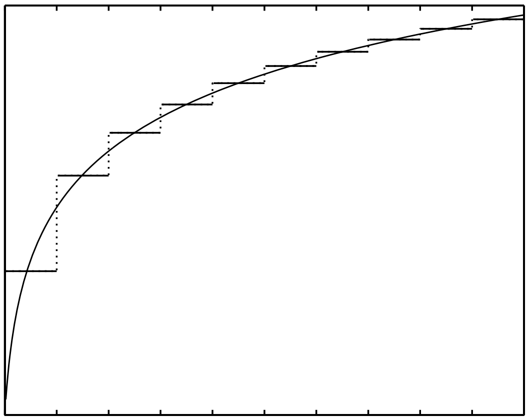

Fig. 3.15. Discretization of initial forward curve. Each discretized forward rate is the average of the underlying forward curve over the discretization interval.

Once the initial curve has been specified, a generic simulation of a single-factor model evolves like this: for $i = 1, \ldots, M$,

$$\begin{aligned} \hat{f}(t_i, t_j) &= \hat{f}(t_{i-1}, t_j) + \\ &\hat{\mu}(t_{i-1}, t_j)[t_i - t_{i-1}] + \hat{\sigma}(t_{i-1}, t_j)\sqrt{t_i - t_{i-1}} Z_i, \ j = i, \ldots, M, \end{aligned} \qquad (3.96)$$

where $Z_1, \ldots, Z_M$ are independent $N(0, 1)$ random variables and $\hat{\mu}$ and $\hat{\sigma}$ denote discrete counterparts of the continuous-time coefficients in (3.91). We allow $\hat{\sigma}$ to depend on the current vector $\hat{f}$ as well as on time and maturity, though to lighten notation we do not include $\hat{f}$ as an explicit argument of $\hat{\sigma}$.

In practice, $\hat{\sigma}$ would typically be specified through a calibration procedure designed to make the simulated model consistent with market prices of actively traded derivative securities. (We discuss calibration of a closely related class of models in Section 3.7.4.) In fact, the continuous-time limit $\sigma(t, T)$ may never be specified explicitly because only the discrete version $\hat{\sigma}$ is used in the simulation. But the situation for the drift is different. Recall that in deriving (3.91) we chose the drift to make the model arbitrage-free; more precisely, we chose it to make the discounted bond prices martingales. There are many ways one might consider choosing the discrete drift $\hat{\mu}$ in (3.96) to approximate the continuous-time limit (3.90). From the many possible approximations, we choose the one that preserves the martingale property for the discounted bond prices.

Recalling that $f(s,s)$ is the short rate at time s, we can express the continuous-time condition as the requirement that

$$B(t,T)\exp\left(-\int_0^t f(s,s)\,ds\right)$$

be a martingale in t for each T. Similarly, in the discretized model we would like

$$\hat{B}(t_i,t_j)\exp\left(-\sum_{k=0}^{i-1}\hat{f}(t_k,t_k)[t_{k+1}-t_k]\right)$$

to be a martingale in i for each j. Our objective is to find a $\hat{\mu}$ for which this holds. For simplicity, we start by assuming a single-factor model.

The martingale condition can be expressed as

$$\begin{aligned}\mathsf{E}&\left[\hat{B}(t_i,t_j)e^{-\sum_{k=0}^{i-1}\hat{f}(t_k,t_k)[t_{k+1}-t_k]}|Z_1,\ldots,Z_{i-1}\right]\\ &= \hat{B}(t_{i-1},t_j)e^{-\sum_{k=0}^{i-2}\hat{f}(t_k,t_k)[t_{k+1}-t_k]}.\end{aligned}$$

Using (3.94) and canceling terms that appear on both sides, this reduces to

$$\mathsf{E}\left[e^{-\sum_{\ell=i}^{j-1}\hat{f}(t_i,t_\ell)[t_{\ell+1}-t_\ell]}|Z_1,\ldots,Z_{i-1}\right] = e^{-\sum_{\ell=i}^{j-1}\hat{f}(t_{i-1},t_\ell)[t_{\ell+1}-t_\ell]}.$$

Now we introduce $\hat{\mu}$: on the left side of this equation we substitute for each $\hat{f}(t_i,t_\ell)$ according to (3.96). This yields the condition

$$\begin{aligned}\mathsf{E}&\left[e^{-\sum_{\ell=i}^{j-1}\left(\hat{f}(t_{i-1},t_\ell)+\hat{\mu}(t_{i-1},t_\ell)[t_i-t_{i-1}]+\hat{\sigma}(t_{i-1},t_\ell)\sqrt{t_i-t_{i-1}}Z_i\right)[t_{\ell+1}-t_\ell]}|Z_1,\ldots,Z_{i-1}\right]\\ &= e^{-\sum_{\ell=i}^{j-1}\hat{f}(t_{i-1},t_\ell)[t_{\ell+1}-t_\ell]}.\end{aligned}$$

Canceling terms that appear on both sides and rearranging the remaining terms brings this into the form

$$\begin{aligned}\mathsf{E}&\left[e^{-\sum_{\ell=i}^{j-1}\hat{\sigma}(t_{i-1},t_\ell)\sqrt{t_i-t_{i-1}}[t_{\ell+1}-t_\ell]Z_i}|Z_1,\ldots,Z_{i-1}\right]\\ &= e^{\sum_{\ell=i}^{j-1}\hat{\mu}(t_{i-1},t_\ell)[t_i-t_{i-1}][t_{\ell+1}-t_\ell]}.\end{aligned}$$

The conditional expectation on the left evaluates to

$$e^{\frac{1}{2}\left(\sum_{\ell=i}^{j-1}\hat{\sigma}(t_{i-1},t_\ell)[t_{\ell+1}-t_\ell]\right)^2[t_i-t_{i-1}]},$$

so equality holds if

$$\tfrac{1}{2}\left(\sum_{\ell=i}^{j-1}\hat{\sigma}(t_{i-1},t_\ell)[t_{\ell+1}-t_\ell]\right)^2 = \sum_{\ell=i}^{j-1}\hat{\mu}(t_{i-1},t_\ell)[t_{\ell+1}-t_\ell];$$

i.e., if

$$\hat{\mu}(t_{i-1},t_j)[t_{j+1}-t_j] = \tfrac{1}{2}\left(\sum_{\ell=i}^{j}\hat{\sigma}(t_{i-1},t_\ell)[t_{\ell+1}-t_\ell]\right)^2 - \tfrac{1}{2}\left(\sum_{\ell=i}^{j-1}\hat{\sigma}(t_{i-1},t_\ell)[t_{\ell+1}-t_\ell]\right)^2. \quad (3.97)$$

This is the discrete version of the HJM drift; it ensures that the discretized discounted bond prices are martingales.

To see the connection between this expression and the continuous-time drift (3.90), consider the case of an equally spaced grid, $t_i = ih$ for some increment $h > 0$. Fix a date t and maturity T and let $i, j \to \infty$ and $h \to 0$ in such a way that $jh = T$ and $ih = t$; each of the sums in (3.97) is then approximated by an integral. Dividing both sides of (3.97) by $t_{j+1} - t_j = h$, we find that for small h the discrete drift is approximately

$$\frac{1}{2h}\left[\left(\int_t^T \sigma(t,u)\,du\right)^2 - \left(\int_t^{T-h}\sigma(t,u)\,du\right)^2\right] \approx \tfrac{1}{2}\frac{\partial}{\partial T}\left(\int_t^T \sigma(t,u)\,du\right)^2,$$

which is

$$\sigma(t,T)\int_t^T \sigma(t,u)\,du.$$

This suggests that the discrete drift in (3.97) is indeed consistent with the continuous-time limit in (3.90).

In the derivation leading to (3.97) we assumed a single-factor model. A similar result holds with d factors. Let $\hat{\sigma}_k$ denote the kth entry of the d-vector $\hat{\sigma}$ and

$$\hat{\mu}_k(t_{i-1},t_j)[t_{j+1}-t_j] = \tfrac{1}{2}\left(\sum_{\ell=i}^{j}\hat{\sigma}_k(t_{i-1},t_\ell)[t_{\ell+1}-t_\ell]\right)^2 - \tfrac{1}{2}\left(\sum_{\ell=i}^{j-1}\hat{\sigma}_k(t_{i-1},t_\ell)[t_{\ell+1}-t_\ell]\right)^2,$$

for $k = 1, \ldots, d$. The combined drift is given by the sum

$$\hat{\mu}(t_{i-1},t_j) = \sum_{k=1}^{d}\hat{\mu}_k(t_{i-1},t_j).$$

A generic multifactor simulation takes the form

$$\hat{f}(t_i,t_j) = \hat{f}(t_{i-1},t_j) + \hat{\mu}(t_{i-1},t_j)[t_i - t_{i-1}] + \sum_{k=1}^{d}\hat{\sigma}_k(t_{i-1},t_j)\sqrt{t_i - t_{i-1}}Z_{ik}, \quad j = i, \ldots, M, \quad (3.98)$$

where the $Z_i = (Z_{i1}, \ldots, Z_{id})$, $i = 1, \ldots, M$, are independent $N(0, I)$ random vectors.

We derived (3.97) by starting from the principle that the discretized discounted bond prices should be martingales. But what are the practical implications of using some other approximation to the continuous drift instead of this one? To appreciate the consequences, consider the following experiment. Imagine simulating paths of $\hat{f}$ as in (3.96) or (3.98). From a path of $\hat{f}$ we may extract a path

$$\hat{r}(t_0) = \hat{f}(t_0, t_0), \quad \hat{r}(t_1) = \hat{f}(t_1, t_1), \quad \ldots \quad \hat{r}(t_M) = \hat{f}(t_M, t_M),$$

of the discretized short rate $\hat{r}$. From this we can calculate a discount factor

$$\hat{D}(t_j) = \exp\left(-\sum_{i=0}^{j-1} \hat{r}(t_i)[t_{i+1} - t_i]\right) \tag{3.99}$$

for each maturity t_j. Imagine repeating this over n independent paths and let $\hat{D}^{(1)}(t_j), \ldots, \hat{D}^{(n)}(t_j)$ denote discount factors calculated over these n paths. A consequence of the strong law of large numbers, the martingale property, and the initialization in (3.95) is that, almost surely,

$$\frac{1}{n}\sum_{i=1}^{n} \hat{D}^{(i)}(t_j) \to \mathsf{E}[\hat{D}(t_j)] = \hat{B}(0, t_j) = B(0, t_j).$$

This means that if we simulate using (3.97) and then use the simulation to price a bond, the simulation price converges to the value to which the model was ostensibly calibrated. With some other choice of discrete drift, the simulation price would in general converge to something that differs from $B(0, t_j)$, even if only slightly. Thus, the martingale condition is not simply a theoretical feature — it is a prerequisite for internal consistency of the simulated model. Indeed, failure of this condition can create the illusion of arbitrage opportunities. If $\mathsf{E}[\hat{D}^{(1)}(t_j)] \neq B(0, t_j)$, the simulation would be telling us that the market has mispriced the bond.

The errors (or apparent arbitrage opportunities) that may arise from using a different approximation to the continuous-time drift may admittedly be quite small. But given that we have a simple way of avoiding such errors and given that the form of the drift is the central feature of the HJM framework, we may as well restrict ourselves to (3.97). This form of the discrete drift appears to be in widespread use in the industry; it is explicit in Andersen [11].

Forward Measure

Through an argument similar to the one leading to (3.97), we can find the appropriate form of the discrete drift under the forward measure. In continuous time, the forward measure for maturity T_F is characterized by the requirement that $B(t,T)/B(t,T_F)$ be a martingale, because the bond maturing at T_F is the numeraire asset associated with this measure. In the discrete approximation, if we take $t_M = T_F$, then we require that $\hat{B}(t_i,t_j)/\hat{B}(t_i,t_M)$ be a martingale in i for each j. This ratio is given by

$$\frac{\hat{B}(t_i,t_j)}{\hat{B}(t_i,t_M)} = \exp\left(\sum_{\ell=j}^{M-1} \hat{f}(t_i,t_\ell)[t_{\ell+1}-t_\ell]\right).$$

The martingale condition leads to a discrete drift $\hat{\mu}$ with

$$\hat{\mu}(t_{i-1},t_j)[t_{j+1}-t_j] = \frac{1}{2}\left(\sum_{\ell=j+1}^{M-1} \hat{\sigma}(t_{i-1},t_\ell)[t_{\ell+1}-t_\ell]\right)^2 - \frac{1}{2}\left(\sum_{\ell=j}^{M-1} \hat{\sigma}(t_{i-1},t_\ell)[t_{\ell+1}-t_\ell]\right)^2. \quad (3.100)$$

The relation between this and the risk-neutral discrete drift (3.97) is, not surprisingly, similar to the relation between their continuous-time counterparts in (3.91) and (3.93).

3.6.3 Implementation

Once we have identified the discrete form of the drift, the main consideration in implementing an HJM simulation is keeping track of indices. The notation $\hat{f}(t_i,t_j)$ is convenient in describing the discretized model — the first argument shows the current time, the second argument shows the maturity to which this forward rate applies. But in implementing the simulation we are not interested in keeping track of an $M \times M$ matrix of rates as the notation $\hat{f}(t_i,t_j)$ might suggest. At each time step, we need only the vector of current rates. To implement an HJM simulation we need to adopt some conventions regarding the indexing of this vector.

Recall that our time and maturity grid consists of a set of dates $0 = t_0 < t_1 < \cdots < t_M$. If we identify t_M with the ultimate maturity T^* in the continuous-time model, then t_M is the maturity of the longest-maturity bond represented in the model. In light of (3.94), this means that the last forward rate relevant to the model applies to the interval $[t_{M-1},t_M]$; this is the forward rate with maturity argument t_{M-1}. Thus, our initial vector of forward rates consists of the M components $\hat{f}(0,0),\hat{f}(0,t_1),\ldots,\hat{f}(0,t_{M-1})$, which is consistent with the initialization (3.95). At the start of the simulation we will represent this vector as $(f_1,\ldots,f_M)$. Thus, our first convention is to use 1 rather than 0 as the lowest index value.

As the simulation evolves, the number of relevant rates decreases. At time t_i, only the rates $\hat{f}(t_i,t_i),\ldots,\hat{f}(t_i,t_{M-1})$ are meaningful. We need to specify how these $M-i$ rates should be indexed, given that initially we had a vector of M rates: we can either pad the initial portion of the vector with irrelevant data or we can shorten the length of the vector. We choose the latter and represent the $M-i$ rates remaining at t_i as the vector $(f_1,\ldots,f_{M-i})$. Thus, our second convention is to index forward rates by *relative* maturity rather than absolute maturity. At time t_i, f_j refers to the forward rate $\hat{f}(t_i, t_{i+j-1})$. Under this convention f_1 always refers to the current level of the short rate because $\hat{r}(t_i) = \hat{f}(t_i, t_i)$.

Similar considerations apply to $\hat{\mu}(t_i, t_j)$ and $\hat{\sigma}_k(t_i, t_j)$, $k = 1,\ldots,d$, and we adopt similar conventions for the variables representing these terms. For values of $\hat{\mu}$ we use variables m_j and for values of $\hat{\sigma}_k$ we use variables $s_j(k)$; in both cases the subscript indicates a relative maturity and in the case of $s_j(k)$ the argument $k = 1,\ldots,d$ refers the factor index in a d-factor model. We design the indexing so that the simulation step from t_{i-1} to t_i indicated in (3.98) becomes

$$f_j \leftarrow f_{j+1} + m_j[t_i - t_{i-1}] + \sum_{k=1}^{d} s_j(k)\sqrt{t_i - t_{i-1}}Z_{ik}, \quad j = 1,\ldots,M-i.$$

Thus, in advancing from t_{i-1} to t_i we want

$$m_j = \hat{\mu}(t_{i-1}, t_{i+j-1}), \quad s_j(k) = \hat{\sigma}_k(t_{i-1}, t_{i+j-1}). \tag{3.101}$$

In particular, recall that $\hat{\sigma}$ may depend on the current vector of forward rates; as implied by (3.101), the values of all $s_j(k)$ should be determined before the forward rates are updated.

To avoid repeated calculation of the intervals between dates t_i, we introduce the notation

$$h_i = t_i - t_{i-1}, \quad ,i = 1,\ldots,M.$$

These values do not change in the course of a simulation so we use the vector $(h_1,\ldots,h_M)$ to represent these same values at all steps of the simulation.

We now proceed to detail the steps in an HJM simulation. We separate the algorithm into two parts, one calculating the discrete drift parameter at a fixed time step, the other looping over time steps and updating the forward curve at each step. Figure 3.16 illustrates the calculation of

$$\hat{\mu}_k(t_{i-1}, t_j) =$$
$$\frac{1}{2h_j}\left[\sum_{k=1}^{d}\left(\sum_{\ell=i}^{j}\hat{\sigma}_k(t_{i-1}, t_\ell)h_{\ell+1}\right)^2 - \sum_{k=1}^{d}\left(\sum_{\ell=i}^{j-1}\hat{\sigma}_k(t_{i-1}, t_\ell)h_{\ell+1}\right)^2\right]$$

in a way that avoids duplicate computation. In the notation of the algorithm, this drift parameter is evaluated as

$$\frac{1}{2(t_{j+1} - t_j)}[B_{\text{next}} - B_{\text{prev}}],$$

and each $A_{\text{next}}(k)$ records a quantity of the form

$$\sum_{\ell=i}^{j} \hat{\sigma}_k(t_{i-1}, t_\ell) h_{\ell+1}.$$

```
Inputs: s_j(k), j = 1,...,M − i, k = 1,...,d as in (3.101)
        and h_1,...,h_M (h_ℓ = t_ℓ − t_{ℓ−1})

A_prev(k) ← 0, k = 1,...,d
for j = 1,...,M − i
    B_next ← 0
    for k = 1,...,d
        A_next(k) ← A_prev(k) + s_j(k) * h_{i+j}
        B_next ← B_next + A_next(k) * A_next(k)
        A_prev(k) ← A_next(k)
    end
    m_j ← (B_next − B_prev)/(2h_{i+j})
    B_prev ← B_next
end
return m_1,...,m_{M−i}.
```

Fig. 3.16. Calculation of discrete drift parameters $m_j = \hat{\mu}(t_{i-1}, t_{i+j-1})$ needed to simulate transition from t_{i-1} to t_i.

Figure 3.17 shows an algorithm for a single replication in an HJM simulation; the steps in the figure would naturally be repeated over many independent replications. This algorithm calls the one in Figure 3.16 to calculate the discrete drift for all remaining maturities at each time step. The two algorithms could obviously be combined, but keeping them separate should help clarify the various steps. In addition, it helps stress the point that in propagating the forward curve from t_{i-1} to t_i, we first evaluate the $s_j(k)$ and m_j using the forward rates at step $i-1$ and then update the rates to get their values at step i.

To make this point a bit more concrete, suppose we specified a single-factor model with $\hat{\sigma}(t_i, t_j) = \tilde{\sigma}(i, j)\hat{f}(t_i, t_j)$ for some fixed values $\tilde{\sigma}(i, j)$. This makes each $\hat{\sigma}(t_i, t_j)$ proportional to the corresponding forward rate. We noted in Example 3.6.3 that this type of diffusion term is inadmissible in the continuous-time limit, but it can be (and often is) used in practice so long as the increments h_i are kept bounded away from zero. In this model it should be clear that in updating $\hat{f}(t_{i-1}, t_j)$ to $\hat{f}(t_i, t_j)$ we need to evaluate $\tilde{\sigma}(i-1, j)\hat{f}(t_{i-1}, t_j)$ before we update the forward rate.

Since an HJM simulation is typically used to value interest rate derivatives, we have included in Figure 3.17 a few additional generic steps illustrating how

```
Inputs: initial curve (f_1, ..., f_M) and intervals (h_1, ..., h_M)

D ← 1, P ← 0, C ← 0.
for i = 1, ..., M − 1
    D ← D * exp(−f_1 * h_i)
    evaluate s_j(k), j = 1, ..., M − i, k = 1, ..., d
            (recall that s_j(k) = σ̂_k(t_{i−1}, t_{i+j−1}))
    evaluate m_1, ..., m_{M−i} using Figure 3.16
    generate Z_1, ..., Z_d ~ N(0, 1)
    for j = 1, ..., M − i
        S ← 0
        for k = 1, ..., d  S ← S + s_j(k) * Z_k
        f_j ← f_{j+1} + m_j * h_i + S * sqrt(h_i)
    end
    P ← cashflow at t_i (depending on instrument)
    C ← C + D * P
end
return C.
```

Fig. 3.17. Algorithm to simulate evolution of forward curve over $t_0, t_1, \ldots, t_{M-1}$ and calculate cumulative discounted cashflows from an interest rate derivative.

a path of the forward curve is used both to compute and to discount the payoff of a derivative. The details of a particular instrument are subsumed in the placeholder "cashflow at t_i." This cashflow is discounted through multiplication by D, which is easily seen to contain the simulated value of the discount factor $\hat{D}(t_i)$ as defined in (3.99). (When D is updated in Figure 3.17, before the forward rates are updated, f_1 records the short rate for the interval $[t_{i-1}, t_i]$.) To make the pricing application more explicit, we consider a few examples.

Example 3.6.4 *Bonds.* There is no reason to use an HJM simulation to price bonds — if properly implemented, the simulation will simply return prices that could have been computed from the initial forward curve. Nevertheless, we consider this example to help fix ideas. We discussed the pricing of a zero-coupon bond following (3.99); in Figure 3.17 this corresponds to setting $P \leftarrow 1$ at the maturity of the bond and $P \leftarrow 0$ at all other dates. For a coupon paying bond with a face value of 100 and a coupon of c, we would set $P \leftarrow c$ at the coupon dates and $P \leftarrow 100 + c$ at maturity. This assumes, of course, that the coupon dates are among the $t_1, \ldots, t_M$. □

Example 3.6.5 *Caps.* A *caplet* is an interest rate derivative providing protection against an increase in an interest rate for a single period; a *cap* is a portfolio of caplets covering multiple periods. A caplet functions almost like a call option on the short rate, which would have a payoff of the form

$(r(T) - K)^+$ for some strike K and maturity T. In practice, a caplet differs from this in some small but important ways. (For further background, see Appendix C.)

In contrast to the instantaneous short rate $r(t)$, the underlying rate in a caplet typically applies over an interval and is based on discrete compounding. For simplicity, suppose the interval is of the form $[t_i, t_{i+1}]$. At t_i, the continuously compounded rate for this interval is $\hat{f}(t_i, t_i)$; the corresponding discretely compounded rate $\hat{F}$ satisfies

$$\frac{1}{1 + \hat{F}(t_i)[t_{i+1} - t_i]} = e^{-\hat{f}(t_i,t_i)[t_{i+1}-t_i]};$$

i.e.,

$$\hat{F}(t_i) = \frac{1}{t_{i+1} - t_i}\left(e^{\hat{f}(t_i,t_i)[t_{i+1}-t_i]} - 1\right).$$

The payoff of the caplet would then be $(\hat{F}(t_i) - K)^+$ (or a constant multiple of this). Moreover, this payment is ordinarily made at the end of the interval, t_{i+1}. To discount it properly we should therefore simulate to t_i and set

$$P \leftarrow \frac{1}{1 + \hat{F}(t_i)[t_{i+1} - t_i]}(\hat{F}(t_i) - K)^+; \tag{3.102}$$

in the notation of Figure 3.17, this is

$$P \leftarrow e^{-f_1 h_{i+1}}\left(\frac{1}{h_{i+1}}\left(e^{f_1 h_{i+1}} - 1\right) - K\right)^+.$$

Similar ideas apply if the caplet covers an interval longer than a single simulation interval. Suppose the caplet applies to an interval $[t_i, t_{i+n}]$. Then (3.102) still applies at t_i, but with t_{i+1} replaced by t_{i+n} and $\hat{F}(t_i)$ redefined to be

$$\hat{F}(t_i) = \frac{1}{t_{n+i} - t_i}\left(\exp\left(\sum_{\ell=0}^{n-1} \hat{f}(t_i, t_{i+\ell})[t_{i+\ell+1} - t_{i+\ell}]\right) - 1\right).$$

In the case of a cap consisting of caplets for, say, the periods $[t_{i_1}, t_{i_2}]$, $[t_{i_2}, t_{i_3}]$, $\ldots$, $[t_{i_k}, t_{i_{k+1}}]$, for some $i_1 < i_2 < \cdots < i_{k+1}$, this calculation would be repeated and a cashflow recorded at each t_{i_j}, $j = 1, \ldots, k$. □

Example 3.6.6 *Swaptions.* Consider, next, an option to swap fixed-rate payments for floating-rate payments. (See Appendix C for background on swaps and swaptions.) Suppose the underlying swap begins at t_{j_0} with payments to be exchanged at dates $t_{j_1}, \ldots, t_{j_n}$. If we denote the fixed rate in the swap by R, then the fixed-rate payment at t_{j_k} is $100R[t_{j_k} - t_{j_{k-1}}]$, assuming a principal or *notional* amount of 100. As explained in Section C.2 of Appendix C, the value of the swap at t_{j_0} is

$$\hat{V}(t_{j_0}) = 100 \left(R \sum_{\ell=1}^{n} \hat{B}(t_{j_0}, t_{j_\ell})[t_{j_\ell} - t_{j_{\ell-1}}] + \hat{B}(t_{j_0}, t_{j_n}) - 1 \right).$$

The bond prices $\hat{B}(t_{j_0}, t_{j_\ell})$ can be computed from the forward rates at t_{j_0} using (3.94).

The holder of an option to enter this swap will exercise the option if $\hat{V}(t_{j_0}) > 0$ and let it expire otherwise. (For simplicity, we are assuming that the option expires at t_{j_0} though similar calculations apply for an option to enter into a *forward* swap, in which case the option expiration date would be prior to t_{j_0}.) Thus, we may view the swaption as having a payoff of $\max\{0, \hat{V}(t_{j_0})\}$ at t_{j_0}. In a simulation, we would therefore simulate the forward curve to the option expiration date t_{j_0}; at that date, calculate the prices of the bonds $\hat{B}(t_{j_0}, t_{j_\ell})$ maturing at the payment dates of the swaps; from the bond prices calculate the value of the swap $\hat{V}(t_{j_0})$ and thus the swaption payoff $\max\{0, \hat{V}(t_{j_0})\}$; record this as the cashflow P in the algorithm of Figure 3.17 and discount it as in the algorithm.

This example illustrates a general feature of the HJM framework that contrasts with models based on the short rate as in Sections 3.3 and 3.4. Consider valuing a 5-year option on a 20-year swap. This instrument involves maturities as long as 25 years, so valuing it in a model of the short rate could involve simulating paths over a 25-year horizon. In the HJM framework, if the initial forward curve extends for 25 years, then we need to simulate only for 5 years; at the expiration of the option, the remaining forward rates contain all the information necessary to value the underlying swap. Thus, although the HJM setting involves updating many more variables at each time step, it may also require far fewer time steps. □

3.7 Forward Rate Models: Simple Rates

The models considered in this section are closely related to the HJM framework of the previous section in that they describe the arbitrage-free dynamics of the term structure of interest rates through the evolution of forward rates. But the models we turn to now are based on *simple* rather than continuously compounded forward rates. This seemingly minor shift in focus has surprisingly far-reaching practical and theoretical implications. This modeling approach has developed primarily through the work of Miltersen, Sandmann, and Sondermann [268], Brace, Gatarek, and Musiela [56], Musiela and Rutkowski [274], and Jamshidian [197]; it has gained rapid acceptance in the financial industry and stimulated a growing stream of research into what are often called *LIBOR market models*.

3.7.1 LIBOR Market Model Dynamics

The basic object of study in the HJM framework is the forward rate curve $\{f(t,T), t \leq T \leq T^*\}$. But the instantaneous, continuously compounded forward rates $f(t,T)$ might well be considered mathematical idealizations — they are not directly observable in the marketplace. Most market interest rates are based on simple compounding over intervals of, e.g., three months or six months. Even the instantaneous short rate $r(t)$ treated in the models of Sections 3.3 and 3.4 is a bit of a mathematical fiction because short-term rates used for pricing are typically based on periods of one to three months. The term "market model" is often used to describe an approach to interest rate modeling based on observable market rates, and this entails a departure from instantaneous rates.

Among the most important benchmark interest rates are the London Inter-Bank Offered Rates or LIBOR. LIBOR is calculated daily through an average of rates offered by banks in London. Separate rates are quoted for different maturities (e.g., three months and six months) and different currencies. Thus, each day new values are calculated for three-month Yen LIBOR, six-month US dollar LIBOR, and so on.

LIBOR rates are based on *simple* interest. If L denotes the rate for an accrual period of length δ (think of δ as 1/4 or 1/2 for three months and six months respectively, with time measured in years), then the interest earned on one unit of currency over the accrual period is δL. For example, if three-month LIBOR is 6%, the interest earned at the end of three months on a principal of 100 is $0.25 \cdot 0.06 \cdot 100 = 1.50$.

A *forward* LIBOR rate works similarly. Fix δ and consider a maturity T. The forward rate $L(0,T)$ is the rate set at time 0 for the interval $[T, T+\delta]$. If we enter into a contract at time 0 to borrow 1 at time T and repay it with interest at time $T+\delta$, the interest due will be $\delta L(0,T)$. As shown in Appendix C (specifically equation (C.5)), a simple replication argument leads to the following identity between forward LIBOR rates and bond prices:

$$L(0,T) = \frac{B(0,T) - B(0,T+\delta)}{\delta B(0,T+\delta)}. \tag{3.103}$$

This further implies the relation

$$L(0,T) = \frac{1}{\delta}\left(\exp\left(\int_T^{T+\delta} f(0,u)\,du\right) - 1\right) \tag{3.104}$$

between continuous and simple forwards, though it is not necessary to introduce the continuous rates to build a model based on simple rates.

It should be noted that, as is customary in this literature, we treat the forward LIBOR rates as though they were risk-free rates. LIBOR rates are based on quotes by banks which could potentially default and this risk is presumably reflected in the rates. US Treasury bonds, in contrast, are generally considered to have a negligible chance of default. The argument leading to (3.103) may not hold exactly if the bonds on one side and the forward rate on the other reflect different levels of creditworthiness. We will not, however, attempt to take account of these considerations.

Although (3.103) and (3.104) apply in principle to a continuum of maturities T, we consider a class of models in which a finite set of maturities or *tenor dates*

$$0 = T_0 < T_1 < \cdots < T_M < T_{M+1}$$

are fixed in advance. As argued in Jamshidian [197], many derivative securities tied to LIBOR and swap rates are sensitive only to a finite set of maturities and it should not be necessary to introduce a continuum to price and hedge these securities. Let

$$\delta_i = T_{i+1} - T_i, \quad i = 0, \ldots, M,$$

denote the lengths of the intervals between tenor dates. Often, these would all be equal to a nominally fixed interval of a quarter or half year; but even in this case, day-count conventions would produce slightly different values for the fractions δ_i.

For each date T_n we let $B_n(t)$ denote the time-t price of a bond maturing at T_n, $0 \le t \le T_n$. In our usual notation this would be $B(t, T_n)$, but writing $B_n(t)$ and restricting n to $\{1, 2, \ldots, M+1\}$ emphasizes that we are working with a finite set of bonds. Similarly, we write $L_n(t)$ for the forward rate as of time t for the accrual period $[T_n, T_{n+1}]$; see Figure 3.18. This is given in terms of the bond prices by

$$L_n(t) = \frac{B_n(t) - B_{n+1}(t)}{\delta_n B_{n+1}(t)}, \quad 0 \le t \le T_n, \quad n = 0, 1, \ldots, M. \tag{3.105}$$

After T_n, the forward rate L_n becomes meaningless; it sometimes simplifies notation to extend the definition of $L_n(t)$ beyond T_n by setting $L_n(t) = L_n(T_n)$ for all $t \ge T_n$.

From (3.105) we know that bond prices determine the forward rates. At a tenor date T_i, the relation can be inverted to produce

$$B_n(T_i) = \prod_{j=i}^{n-1} \frac{1}{1 + \delta_j L_j(T_i)}, \quad n = i+1, \ldots, M+1. \tag{3.106}$$

However, at an arbitrary date t, the forward LIBOR rates do not determine the bond prices because they do not determine the discount factor for intervals shorter than the accrual periods. Suppose for example that $T_i < t < T_{i+1}$ and we want to find the price $B_n(t)$ for some $n > i+1$. The factor

$$\prod_{j=i+1}^{n-1} \frac{1}{1 + \delta_j L_j(t)}$$

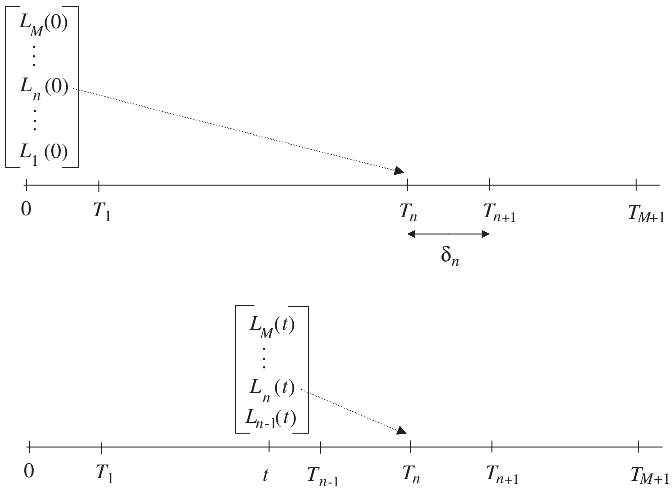

Fig. 3.18. Evolution of vector of forward rates. Each $L_n(t)$ is the forward rate for the interval $[T_n, T_{n+1}]$ as of time $t \leq T_n$.

discounts the bond's payment at T_n back to time T_{i+1}, but the LIBOR rates do not specify the discount factor from T_{i+1} to t.

Define a function $\eta : [0, T_{M+1}) \to \{1, \ldots, M+1\}$ by taking $\eta(t)$ to be the unique integer satisfying

$$T_{\eta(t)-1} \leq t < T_{\eta(t)};$$

thus, $\eta(t)$ gives the index of the next tenor date at time t. With this notation, we have

$$B_n(t) = B_{\eta(t)}(t) \prod_{j=\eta(t)}^{n-1} \frac{1}{1+\delta_j L_j(t)}, \quad 0 \leq t < T_n; \tag{3.107}$$

the factor $B_{\eta(t)}(t)$ (the current price of the shortest maturity bond) is the missing piece required to express the bond prices in terms of the forward LIBOR rates.

Spot Measure

We seek a model in which the evolution of the forward LIBOR rates is described by a system of SDEs of the form

$$\frac{dL_n(t)}{L_n(t)} = \mu_n(t)\,dt + \sigma_n(t)^\top dW(t), \quad 0 \leq t \leq T_n, \quad n = 1, \ldots, M, \tag{3.108}$$

with W a d-dimensional standard Brownian motion. The coefficients μ_n and σ_n may depend on the current vector of rates $(L_1(t), \ldots, L_M(t))$ as well as the current time t. Notice that in (3.108) σ_n is the (proportional) volatility because we have divided by L_n on the left, whereas in the HJM setting (3.91) we took $\sigma(t,T)$ to be the absolute level of volatility. At this point, the distinction is purely one of notation rather than scope because we allow $\sigma_n(t)$ to depend on the current level of rates.

Recall that in the HJM setting we derived the form of the drift of the forward rates from the absence of arbitrage. More specifically, we derived the drift from the condition that bond prices be martingales when divided by the numeraire asset. The numeraire we used is the usual one associated with the risk-neutral measure, $\beta(t) = \exp(\int_0^t r(u)\,du)$. But introducing a short-rate process $r(t)$ would undermine our objective of developing a model based on the simple (and thus more realistic) rates $L_n(t)$. We therefore avoid the usual risk-neutral measure and instead use a numeraire asset better suited to the tenor dates T_i.

A simply compounded counterpart of $\beta(t)$ works as follows. Start with 1 unit of account at time 0 and buy $1/B_1(0)$ bonds maturing at T_1. At time T_1, reinvest the funds in bonds maturing at time T_2 and proceed this way, at each T_i putting all funds in bonds maturing at time T_{i+1}. This trading strategy earns (simple) interest at rate $L_i(T_i)$ over each interval $[T_i, T_{i+1}]$, just as in the continuously compounded case a savings account earns interest at rate $r(t)$ at time t. The initial investment of 1 at time 0 grows to a value of

$$B^*(t) = B_{\eta(t)}(t) \prod_{j=0}^{\eta(t)-1} [1 + \delta_j L_j(T_j)]$$

at time t. Following Jamshidian [197], we take this as numeraire asset and call the associated measure the *spot measure*.

Suppose, then, that (3.108) holds under the spot measure, meaning that W is a standard Brownian motion under that measure. The absence of arbitrage restricts the dynamics of the forward LIBOR rates through the condition that bond prices be martingales when *deflated* by the numeraire asset. (We use the term "deflated" rather than "discounted" to emphasize that we are dividing by the numeraire asset and not discounting at a continuously compounded rate.) From (3.107) and the expression for B^*, we find that the deflated bond price $D_n(t) = B_n(t)/B^*(t)$ is given by

$$D_n(t) = \left(\prod_{j=0}^{\eta(t)-1} \frac{1}{1 + \delta_j L_j(T_j)} \right) \prod_{j=\eta(t)}^{n-1} \frac{1}{1 + \delta_j L_j(t)}, \quad 0 \le t \le T_n. \tag{3.109}$$

Notice that the spot measure numeraire B^* cancels the factor $B_{\eta(t)}(t)$ used in (3.107) to discount between tenor dates. We are thus left in (3.109) with an expression defined purely in terms of the LIBOR rates. This would not have been the case had we divided by the risk-neutral numeraire asset $\beta(t)$.

We require that the deflated bond prices D_n be positive martingales and proceed to derive the restrictions this imposes on the LIBOR dynamics (3.108). If the deflated bonds are indeed positive martingales, we may write

$$\frac{dD_{n+1}(t)}{D_{n+1}(t)} = \nu_{n+1}(t)^\top \, dW(t), \quad n = 1, \ldots, M,$$

for some $\Re^d$-valued processes ν_{n+1} which may depend on the current level of $(D_2, \ldots, D_{M+1})$ (equivalently, of $(L_1, \ldots, L_M)$). By Itô's formula,

$$d \log D_{n+1}(t) = -\tfrac{1}{2}\|\nu_{n+1}(t)\| \, dt + \nu_{n+1}^\top(t) \, dW(t).$$

We may therefore express ν_{n+1} by finding the coefficient of dW in

$$d \log D_{n+1}(t) = -\sum_{j=\eta(t)}^{n} d \log(1 + \delta_j L_j(t));$$

notice that the first factor in (3.109) is constant between maturities T_i. Applying Itô's formula and (3.108), we find that

$$\nu_{n+1}(t) = -\sum_{j=\eta(t)}^{n} \frac{\delta_j L_j(t)}{1 + \delta_j L_j(t)} \sigma_j(t). \tag{3.110}$$

We now proceed by induction to find the μ_n in (3.108). Setting $D_1(t) \equiv B_1(0)$, we make D_1 constant and hence a martingale without restrictions on any of the LIBOR rates. Suppose now that $\mu_1, \ldots, \mu_{n-1}$ have been chosen consistent with the martingale condition on D_n. From the identity $D_n(t) = D_{n+1}(1 + \delta_n L_n(t))$, we find that $\delta_n L_n(t) D_{n+1}(t) = D_n(t) - D_{n+1}(t)$, so D_{n+1} is a martingale if and only if $L_n D_{n+1}$ is a martingale. Applying Itô's formula, we get

$$\begin{aligned}
&d(L_n D_{n+1}) \\
&= D_{n+1} \, dL_n + L_n \, dD_{n+1} + L_n D_{n+1} \nu_{n+1}^\top \sigma_n \, dt \\
&= \left(D_{n+1} \mu_n L_n + L_n D_{n+1} \nu_{n+1}^\top \sigma_n\right) \, dt + L_n D_{n+1} \sigma_n^\top \, dW + L_n \, dD_{n+1}.
\end{aligned}$$

(We have suppressed the time argument to lighten the notation.) To be consistent with the martingale restriction on D_{n+1} and $L_n D_{n+1}$, the dt coefficient must be zero, and thus

$$\mu_n = -\sigma_n^\top \nu_{n+1};$$

notice the similarity to the HJM drift (3.90). Combining this with (3.110), we arrive at

$$\mu_n(t) = \sum_{j=\eta(t)}^{n} \frac{\delta_j L_j(t) \sigma_n(t)^\top \sigma_j(t)}{1 + \delta_j L_j(t)} \tag{3.111}$$

as the required drift parameter in (3.108), so

$$\frac{dL_n(t)}{L_n(t)} = \sum_{j=\eta(t)}^{n} \frac{\delta_j L_j(t) \sigma_n(t)^\top \sigma_j(t)}{1 + \delta_j L_j(t)} \, dt + \sigma_n(t)^\top \, dW(t), \quad 0 \le t \le T_n, \tag{3.112}$$

$n = 1, \ldots, M$, describes the arbitrage-free dynamics of forward LIBOR rates under the spot measure. This formulation is from Jamshidian [197], which should be consulted for a rigorous and more general development.

Forward Measure

As in Musiela and Rutkowski [274], we may alternatively formulate a LIBOR market model under the forward measure P_{M+1} for maturity T_{M+1} and take the bond B_{M+1} as numeraire asset. In this case, we redefine the deflated bond prices to be the ratios $D_n(t) = B_n(t)/B_{M+1}(t)$, which simplify to

$$D_n(t) = \prod_{j=n+1}^{M} (1 + \delta_j L_j(t)). \tag{3.113}$$

Notice that the numeraire asset has once again canceled the factor $B_{\eta(t)}(t)$, leaving an expression that depends solely on the forward LIBOR rates.

We could derive the dynamics of the forward LIBOR rates under the forward measure through the Girsanov Theorem and (3.112), much as we did in the HJM setting to arrive at (3.93). Alternatively, we could start from the requirement that the D_n in (3.113) be martingales and proceed by induction (backwards from $n = M$) to derive restrictions on the evolution of the L_n. Either way, we find that the arbitrage-free dynamics of the L_n, $n = 1, \ldots, M$, under the forward measure P_{M+1} are given by

$$\frac{dL_n(t)}{L_n(t)} = -\sum_{j=n+1}^{M} \frac{\delta_j L_j(t)\sigma_n(t)^\top \sigma_j(t)}{1 + \delta_j L_j(t)}\, dt + \sigma_n(t)^\top\, dW^{M+1}(t), \quad 0 \le t \le T_n, \tag{3.114}$$

with W^{M+1} a standard d-dimensional Brownian motion under P_{M+1}. The relation between the drift in (3.114) and the drift in (3.112) is analogous to the relation between the risk-neutral and forward-measure drifts in the HJM setting; compare (3.90) and (3.93).

If we take $n = M$ in (3.114), we find that

$$\frac{dL_M(t)}{L_M(t)} = \sigma_M(t)^\top\, dW^{M+1}(t),$$

so that, subject only to regularity conditions on its volatility, L_M is a martingale under the forward measure for maturity T_{M+1}. Moreover, if σ_M is deterministic then $L_M(t)$ has lognormal distribution $LN(-\bar{\sigma}_M^2(t)/2, \bar{\sigma}_M^2(t))$ with

$$\bar{\sigma}_M(t) = \sqrt{\frac{1}{t}\int_0^t \|\sigma_M(u)\|^2\, du}. \tag{3.115}$$

In fact, the choice of M is arbitrary: each L_n is a martingale (lognormal if σ_n is deterministic) under the forward measure P_{n+1} associated with T_{n+1}.

These observations raise the question of whether we may in fact take the coefficients σ_n to be deterministic in (3.112) and (3.114). Recall from Example 3.6.3 that this choice (deterministic proportional volatility) is inadmissible in the HJM setting, essentially because it makes the HJM drift quadratic in the current level of rates. To see what happens with simple compounding, rewrite (3.112) as

$$dL_n(t) = \sum_{j=\eta(t)}^{n} \frac{\delta_j L_j(t) L_n(t) \sigma_n(t)^\top \sigma_j(t)}{1 + \delta_j L_j(t)}\, dt + L_n(t)\sigma_n(t)^\top\, dW(t) \tag{3.116}$$

and consider the case of deterministic σ_i. The numerators in the drift are quadratic in the forward LIBOR rates, but they are stabilized by the terms $1 + \delta_j L_j(t)$ in the denominators; indeed, because $L_j(t) \geq 0$ implies

$$\left| \frac{\delta_j L_j(t)}{1 + \delta_j L_j(t)} \right| \leq 1,$$

the drift is linearly bounded in $L_n(t)$, making deterministic σ_i admissible. This feature is lost in the limit as the compounding period δ_j decreases to zero. Thus, the distinction between continuous and simple forward rates turns out to have important mathematical as well as practical implications.

3.7.2 Pricing Derivatives

We have noted two important features of LIBOR market models: they are based on observable market rates, and (in contrast to the HJM framework) they admit deterministic volatilities σ_j. A third important and closely related feature arises in the pricing of interest rate caps.

Recall from Example 3.6.5 (or Appendix C.2) that a cap is a collection of caplets and that each caplet may be viewed as a call option on a simple forward rate. Consider, then, a caplet for the accrual period $[T_n, T_{n+1}]$. The underlying rate is L_n and the value $L_n(T_n)$ is fixed at T_n. With a strike of K, the caplet's payoff is $\delta_n(L_n(T_n) - K)^+$; think of the caplet as refunding the amount by which interest paid at rate $L_n(T_n)$ exceeds interest paid at rate K. This payoff is made at T_{n+1}.

Let $C_n(t)$ denote the price of this caplet at time t; we know the terminal value $C_n(T_{n+1}) = \delta_n(L_n(T_n) - K)^+$ and we want to find the initial value $C_n(0)$. Under the spot measure, the deflated price $C_n(t)/B^*(t)$ must be a martingale, so

$$C_n(0) = B^*(0)\mathsf{E}^*\left[\frac{\delta_n(L_n(T_n) - K)^+}{B^*(T_{n+1})}\right],$$

where we have written E^* for expectation under the spot measure. Through $B^*(T_{n+1})$, this expectation involves the joint distribution of $L_1(T_1), \ldots, L_n(T_n)$, making its value difficult to discern. In contrast, under the forward measure P_{n+1} associated with maturity T_{n+1}, the martingale property applies to $C_n(t)/B_{n+1}(t)$. We may therefore also write

$$C_n(0) = B_{n+1}(0)\mathsf{E}_{n+1}\left[\frac{\delta_n(L_n(T_n) - K)^+}{B_{n+1}(T_{n+1})}\right],$$

with E_{n+1} denoting expectation under P_{n+1}. Conveniently, $B_{n+1}(T_{n+1}) \equiv 1$, so this expectation depends only on the marginal distribution of $L_n(T_n)$. If we take σ_n to be deterministic, then $L_n(T_n)$ has the lognormal distribution $LN(-\bar{\sigma}_n^2(T_n)/2, \bar{\sigma}_n^2(T_n))$, using the notation in (3.115). In this case, the caplet price is given by the *Black formula* (after Black [49]),

$$C_n(0) = \mathrm{BC}(L_n(0), \bar{\sigma}_n(T_n), T_n, K, \delta_n B_{n+1}(0)),$$

with

$$\begin{aligned}&\mathrm{BC}(F, \sigma, T, K, b) = \\ &b\left(F\Phi\left(\frac{\log(F/K) + \sigma^2 T/2}{\sigma\sqrt{T}}\right) - K\Phi\left(\frac{\log(F/K) - \sigma^2 T/2}{\sigma\sqrt{T}}\right)\right)\end{aligned} \tag{3.117}$$

and Φ the cumulative normal distribution. Thus, under the assumption of deterministic volatilities, caplets are priced in closed form by the Black formula.

This formula is frequently used in the reverse direction. Given the market price of a caplet, one can solve for the "implied volatility" that makes the formula match the market price. This is useful in calibrating a model to market data, a point we return to in Section 3.7.4.

Consider, more generally, a derivative security making a payoff of $g(L(T_n))$ at T_k, with $L(T_n) = (L_1(T_1), \ldots, L_{n-1}(T_{n-1}), L_n(T_n), \ldots, L_M(T_n))$ and $k \geq n$. The price of the derivative at time 0 is given by

$$\mathsf{E}^*\left[\frac{g(L(T_n))}{B^*(T_k)}\right]$$

(using the fact that $B^*(0) = 1$), and also by

$$B_m(0)\mathsf{E}_m\left[\frac{g(L(T_n))}{B_m(T_k)}\right]$$

for every $m \geq k$. Which measure and numeraire are most convenient depends on the payoff function g. However, in most cases, the expectation cannot be evaluated explicitly and simulation is required.

As a further illustration, we consider the pricing of a swaption as described in Example 3.6.6 and Appendix C.2. Suppose the underlying swap begins at T_n with fixed- and floating-rate payments exchanged at $T_{n+1}, \ldots, T_{M+1}$. From equation (C.7) in Appendix C, we find that the forward swap rate at time t is given by

$$S_n(t) = \frac{B_n(t) - B_{M+1}(t)}{\sum_{j=n+1}^{M+1} \delta_j B_j(t)}. \tag{3.118}$$

Using (3.107) and noting that $B_{\eta(t)}(t)$ cancels from the numerator and denominator, this swap rate can be expressed purely in terms of forward LIBOR rates.

Consider, now, an option expiring at time $T_k \leq T_n$ to enter into the swap over $[T_n, T_{M+1}]$ with fixed rate T. The value of the option at expiration can be expressed as (cf. equation (C.11))

$$\sum_{j=n+1}^{M+1} \delta_j B_j(T_k)(R - S_n(T_k))^+.$$

This can be written as a function $g(L(T_k))$ of the LIBOR rates. The price at time zero can therefore be expressed as an expectation using the general expressions above.

By applying Itô's formula to the swap rate (3.118), it is not difficult to conclude that if the forward LIBOR rates have deterministic volatilities, then the forward swap rate cannot also have a deterministic volatility. In particular, then, the forward swap rate cannot be geometric Brownian motion under any equivalent measure. Brace et al. [56] nevertheless use a lognormal approximation to the swap rate to develop a method for pricing swaptions; their approximation appears to give excellent results. An alternative approach has been developed by Jamshidian [197]. He develops a model in which the term structure is described through a vector $(S_0(t), \ldots, S_M(t))$ of forward swap rates. He shows that one may choose the volatilities of the forward swap rates to be deterministic, and that in this case swaption prices are given by a variant of the Black formula. However, in this model, the LIBOR rates cannot also have deterministic volatilities, so caplets are no longer priced by the Black formula. One must therefore choose between the two pricing formulas.

3.7.3 Simulation

Pricing derivative securities in LIBOR market models typically requires simulation. As in the HJM setting, exact simulation is generally infeasible and some discretization error is inevitable. Because the models of this section deal with a finite set of maturities from the outset, we need only discretize the time argument, whereas in the HJM setting both time and maturity required discretization.

We fix a time grid $0 = t_0 < t_1 < \cdots < t_m < t_{m+1}$ over which to simulate. It is sensible to include the tenor dates $T_1, \ldots, T_{M+1}$ among the simulation dates. In practice, one would often even take $t_i = T_i$ so that the simulation evolves directly from one tenor date to the next. We do not impose any restrictions on the volatilities σ_n, though the deterministic case is the most widely used. The only other specific case that has received much attention takes $\sigma_n(t)$ to be the product of a deterministic function of time and a function of $L_n(t)$ as proposed in Andersen and Andreasen [13]. For example, one may take $\sigma_n(t)$ proportional to a power of $L_n(t)$, resulting in a CEV-type of volatility. In either this extension or in the case of deterministic volatilities, it often suffices to restrict the dependence on time to piecewise constant functions that change values only at the T_i. We return to this point in Section 3.7.4.

Simulation of forward LIBOR rates is a special case of the general problem of simulating a system of SDEs. One could apply an Euler scheme or a higher-order method of the type discussed in Chapter 6. However, even if we restrict ourselves to Euler schemes (as we do here), there are countless alternatives. We have many choices of variables to discretize and many choices of probability measure under which to simulate. Several strategies are compared both theoretically and numerically in Glasserman and Zhao [151], and the discussion here draws on that investigation.

The most immediate application of the Euler scheme under the spot measure discretizes the SDE (3.116), producing

$$\begin{aligned}\hat{L}_n(t_{i+1}) = \hat{L}_n(t_i) + \mu_n(\hat{L}(t_i), t_i)\hat{L}_n(t_i)[t_{i+1} - t_i] \\ + \hat{L}_n(t_i)\sqrt{t_{i+1} - t_i}\sigma_n(t_i)^\top Z_{i+1}\end{aligned} \tag{3.119}$$

with

$$\mu_n(\hat{L}(t_i), t_i) = \sum_{j=\eta(t_i)}^{n} \frac{\delta_j \hat{L}_j(t_i)\sigma_n(t_i)^\top \sigma_j(t_i)}{1 + \delta_j \hat{L}_j(t_i)}$$

and $Z_1, Z_2, \ldots$ independent $N(0, I)$ random vectors in $\Re^d$. Here, as in Section 3.6.2, we use hats to identify discretized variables. We assume that we are given an initial set of bond prices $B_1(0), \ldots, B_{M+1}(0)$ and initialize the simulation by setting

$$\hat{L}_n(0) = \frac{B_n(0) - B_{n+1}(0)}{\delta_n B_{n+1}(0)}, \quad n = 1, \ldots, M,$$

in accordance with (3.105).

An alternative to (3.119) approximates the LIBOR rates under the spot measure using

$$\begin{aligned}\hat{L}_n(t_{i+1}) &= \hat{L}_n(t_i)\times \\ &\exp\left(\left[\mu_n(\hat{L}(t_i), t_i) - \tfrac{1}{2}\|\sigma_n(t_i)\|^2\right][t_{i+1} - t_i] + \sqrt{t_{i+1} - t_i}\sigma_n(t_i)^\top Z_{i+1}\right).\end{aligned} \tag{3.120}$$

This is equivalent to applying an Euler scheme to $\log L_n$; it may also be viewed as approximating L_n by geometric Brownian motion over $[t_i, t_{i+1}]$, with drift and volatility parameters fixed at t_i. This method seems particularly attractive in the case of deterministic σ_n, since then L_n is close to lognormal. A further property of (3.120) is that it keeps all $\hat{L}_n$ positive, whereas (3.119) can produce negative rates.

For both of these algorithms it is important to note that our definition of η makes η right-continuous. For the original continuous-time processes we could just as well have taken η to be left-continuous, but the distinction is important in the discrete approximation. If $t_i = T_k$, then $\eta(t_i) = k+1$ and the sum in each $\mu_n(\hat{L}(t_i), t_i)$ starts at $k+1$. Had we taken η to be left-continuous, we would have $\eta(T_i) = k$ and thus an additional term in each μ_n. It seems intuitively more natural to omit this term as time advances beyond T_k since L_k ceases to be meaningful after T_k. Glasserman and Zhao [151] and Sidenius [330] both find that omitting it (i.e., taking η right-continuous) results in smaller discretization error.

Both (3.119) and (3.120) have obvious counterparts for simulation under the forward measure P_{M+1}. The only modification necessary is to replace $\mu_n(\hat{L}(t_i), t_i)$ with

$$\mu_n(\hat{L}(t_i), t_i) = -\sum_{j=n+1}^{M} \frac{\delta_j \hat{L}_j(t_i)\sigma_n(t_i)^\top \sigma_j(t_i)}{1 + \delta_j \hat{L}_j(t_i)}.$$

Notice that $\mu_M \equiv 0$. It follows that if the σ_M is deterministic and constant between the t_i (for example, constant between tenor dates), then the log Euler scheme (3.120) with $\mu_M = 0$ simulates L_M without discretization error under the forward measure P_{M+1}. None of the L_n is simulated without discretization error under the spot measure, but we will see that the spot measure is nevertheless generally preferable for simulation.

Martingale Discretization

In our discussion of simulation in the HJM setting, we devoted substantial attention to the issue of choosing the discrete drift to keep the model arbitrage-free even after discretization. It is therefore natural to examine whether an analogous choice of drift can be made in the LIBOR rate dynamics. In the HJM setting, we derived the discrete drift from the condition that the discretized discounted bond prices must be martingales. In the LIBOR market model, the corresponding requirement is that

$$\hat{D}_n(t_i) = \prod_{j=0}^{n-1} \frac{1}{1 + \delta_j \hat{L}_j(t_i \wedge T_j)} \tag{3.121}$$

be a martingale (in i) for each n under the spot measure; see (3.109). Under the forward measure, the martingale condition applies to

$$\hat{D}_n(t_i) = \prod_{j=n}^{M} \left(1 + \delta_j \hat{L}_j(t_i)\right); \tag{3.122}$$

see (3.113).

Consider the spot measure first. We would like, as a special case of (3.121), for $1/(1 + \delta_1 \hat{L}_1)$ to be a martingale. Using the Euler scheme (3.119), this requires

$$\mathsf{E}\left[\frac{1}{1+\delta_1(\hat{L}_1(0)[1+\mu_1 t_1+\sqrt{t_1}\sigma_1^\top Z_1)]}\right]=\frac{1}{1+\delta_1\hat{L}_1(0)},$$

the expectation taken with respect to $Z_1 \sim N(0,I)$. However, because the denominator inside the expectation has a normal distribution, the expectation is infinite no matter how we choose μ_1. There is no discrete drift that preserves the martingale property. If, instead, we use the method in (3.120), the condition becomes

$$\mathsf{E}\left[\frac{1}{1+\delta_1(\hat{L}_1(0)\exp([\mu_1-\|\sigma_1\|^2/2]t_1+\sqrt{t_1}\sigma_1^\top Z_1))}\right]=\frac{1}{1+\delta_1\hat{L}_1(0)}.$$

In this case, there is a value of μ_1 for which this equation holds, but there is no explicit expression for it. The root of the difficulty lies in evaluating an expression of the form

$$\mathsf{E}\left[\frac{1}{1+\exp(a+bZ)}\right],\qquad Z\sim N(0,1),$$

which is effectively intractable. In the HJM setting, calculation of the discrete drift relies on evaluating far more convenient expressions of the form $\mathsf{E}[\exp(a+bZ)]$; see the steps leading to (3.97).

Under the forward measure, it is feasible to choose μ_1 so that $\hat{D}_2$ in (3.122) is a martingale using an Euler scheme for either L_1 or $\log L_1$. However, this quickly becomes cumbersome for $\hat{D}_n$ with larger values of n. As a practical matter, it does not seem feasible under any of these methods to adjust the drift to make the deflated bond prices martingales. A consequence of this is that if we price bonds in the simulation by averaging replications of (3.121) or (3.122), the simulation price will not converge to the corresponding $B_n(0)$ as the number of replications increases.

An alternative strategy is to discretize and simulate the deflated bond prices themselves, rather than the forward LIBOR rates. For example, under the spot measure, the deflated bond prices satisfy

$$\begin{aligned}\frac{dD_{n+1}(t)}{D_{n+1}(t)}&=-\sum_{j=\eta(t)}^{n}\left(\frac{\delta_j L_j(t)}{1+\delta_j L_j(t)}\right)\sigma_j^\top(t)\,dW(t)\\&=\sum_{j=\eta(t)}^{n}\left(\frac{D_{j+1}(t)}{D_j(t)}-1\right)\sigma_j^\top(t)\,dW(t).\end{aligned}\tag{3.123}$$

An Euler scheme for $\log D_{n+1}$ therefore evolves according to

$$\hat{D}_{n+1}(t_{i+1})=\hat{D}_{n+1}(t_i)\exp\left(-\tfrac{1}{2}\|\hat{\nu}_{n+1}(t_i)\|^2[t_{i+1}-t_i]+\sqrt{t_{i+1}-t_i}\hat{\nu}_{n+1}(t_i)^\top Z_{i+1}\right)\tag{3.124}$$

with

$$\hat{\nu}_{n+1}(t_i) = \sum_{j=\eta(t_i)}^{n} \left(\frac{\hat{D}_{j+1}(t_i)}{\hat{D}_j(t_i)} - 1 \right) \sigma_j(t_i). \tag{3.125}$$

In either case, the discretized deflated bond prices are automatically martingales; in (3.124) they are positive martingales and in this sense the discretization is arbitrage-free. From the simulated $\hat{D}_n(t_i)$ we can then *define* the discretized forward LIBOR rates by setting

$$\hat{L}_n(t_i) = \frac{1}{\delta_n} \left(\frac{\hat{D}_n(t_i) - \hat{D}_{n+1}(t_i)}{\hat{D}_{n+1}(t_i)} \right),$$

for $n = 1, \ldots, M$. Any other term structure variables (e.g., swap rates) required in the simulation can then be defined from the $\hat{L}_n$.

Glasserman and Zhao [151] recommend replacing

$$\left(\frac{\hat{D}_{j+1}(t_i)}{\hat{D}_j(t_i)} - 1 \right) \quad \text{with} \quad \min \left\{ \left(\frac{\hat{D}_{j+1}(t_i)}{\hat{D}_j(t_i)} \right)^+ - 1, 0 \right\}. \tag{3.126}$$

This modification has no effect in the continuous-time limit because $0 \leq D_{j+1}(t) \leq D_j(t)$ (if $L_j(t) \geq 0$). But in the discretized process the ratio $\hat{D}_{j+1}/\hat{D}_j$ could potentially exceed 1.

Under the forward measure P_{M+1}, the deflated bond prices (3.113) satisfy

$$\begin{aligned} \frac{dD_{n+1}(t)}{D_{n+1}(t)} &= \sum_{j=n+1}^{M} \frac{\delta_j L_j(t)}{1 + \delta_j L_j(t)} \sigma_j(t)^\top \, dW^{M+1}(t) \\ &= \sum_{j=n+1}^{M} \left(1 - \frac{D_{j+1}(t)}{D_j(t)} \right) \sigma_j^\top(t) \, dW^{M+1}(t). \end{aligned} \tag{3.127}$$

We can again apply an Euler discretization to the logarithm of these variables to get (3.124), except that now

$$\hat{\nu}_{n+1}(t_i) = \sum_{j=n+1}^{M} \left(1 - \frac{\hat{D}_{j+1}(t_i)}{\hat{D}_j(t_i)} \right) \sigma_j(t_i),$$

possibly modified as in (3.126).

Glasserman and Zhao [151] consider several other choices of variables for discretization, including (under the spot measure) the normalized differences

$$V_n(t) = \frac{D_n(t) - D_{n+1}(t)}{B_1(0)}, \quad n = 1, \ldots, M;$$

these are martingales because the deflated bond prices are martingales. They satisfy

$$\frac{dV_n}{V_n} = \left[\left(\frac{V_n + V_{n-1} + \cdots + V_1 - 1}{V_{n-1} + \cdots + V_1 - 1}\right)\sigma_n^\top + \sum_{j=\eta}^{n-1}\left(\frac{V_j}{V_{j-1} + \cdots + V_1 - 1}\right)\sigma_j^\top\right] dW,$$

with the convention $\sigma_{M+1} \equiv 0$. Forward rates are recovered using

$$\delta_n L_n(t) = \frac{V_n(t)}{V_{n+1}(t) + \cdots + V_{M+1}(t)}.$$

Similarly, the variables

$$\delta_n X_n(t) = \delta_n L_n(t) \prod_{j=n+1}^{M} (1 + \delta_j L_j(t))$$

are differences of deflated bond prices under the forward measure P_{M+1} and thus martingales under that measure. The X_n satisfy

$$\frac{dX_n}{X_n} = \left(\sigma_n^\top + \sum_{j=n+1}^{M} \frac{\delta_j X_j \sigma_j^\top}{1 + \delta_j X_j + \cdots + \delta_M X_M}\right) dW^{M+1}.$$

Forward rates are recovered using

$$L_n = \frac{X_n}{1 + \delta_{n+1} X_{n+1} + \cdots + \delta_M X_M}.$$

Euler discretizations of $\log V_n$ and $\log X_n$ preserve the martingale property and thus keep the discretized model arbitrage-free.

Pricing Derivatives

The pricing of a derivative security in a simulation proceeds as follows. Using any of the methods considered above, we simulate paths of the discretized variables $\hat{L}_1, \ldots, \hat{L}_M$. Suppose we want to price a derivative with a payoff of $g(L(T_n))$ at time T_n. Under the spot measure, we simulate to time T_n and then calculate the deflated payoff

$$g(\hat{L}(T_n)) \cdot \prod_{j=0}^{n-1} \frac{1}{1 + \delta_j \hat{L}_j(T_j)}.$$

Averaging over independent replications produces an estimate of the derivative's price at time 0. If we simulate under the forward measure, the estimate consists of independent replications of

$$g(\hat{L}(T_n)) \cdot B_{M+1}(0) \prod_{j=1}^{n-1} (1 + \delta_j \hat{L}_j(T_j)).$$

Glasserman and Zhao [151] compare various simulation methods based, in part, on their discretization error in pricing caplets. For the case of a caplet over $[T_{n-1}, T_n]$, take $g(x) = \delta_{n-1}(x - K)^+$ in the expressions above. If the σ_j are deterministic, the caplet price is given by the Black formula, as explained in Section 3.7.2. However, because of the discretization error, the simulation price will not in general converge exactly to the Black price as the number of replications increase. The bias in pricing caplets serves as a convenient indication of the magnitude of the discretization error.

Figure 3.19, reproduced from Glasserman and Zhao [151], graphs biases in caplet pricing as a function of caplet maturity for various simulation methods. The horizontal line through the center of each panel corresponds to zero bias. The error bars around each curve have halfwidths of one standard error, indicating that the apparent biases are statistically significant. Details of the parameters used for these experiments are reported in Glasserman and Zhao [151] along with several other examples.

These and other experiments suggest the following observations. The smallest biases are achieved by simulating the differences of deflated bond prices (the V_n in the spot measure and the X_n in the forward measure) using an Euler scheme for the logarithms of these variables. (See Glasserman and Zhao [151] for an explanation of the modified V_n method.) An Euler scheme for $\log D_n$ is nearly indistinguishable from an Euler scheme for L_n. Under the forward measure P_{M+1}, the final caplet is priced without discretization error by the Euler schemes for $\log X_n$ and $\log L_n$; these share the feature that they make the discretized rate $\hat{L}_M$ lognormal.

The graphs in Figure 3.19 compare discretization biases but say nothing about the relative variances of the methods. Glasserman and Zhao [151] find that simulating under the spot measure usually results in smaller variance than simulating under the forward measure, especially at high levels of volatility. An explanation for this is suggested by the expressions (3.109) and (3.113) for the deflated bond prices under the two measures: whereas (3.109) always lies between 0 and 1, (3.113) can take arbitrarily large values. This affects derivatives pricing through the discounting of payoffs.

3.7.4 Volatility Structure and Calibration

In our discussion of LIBOR market models we have taken the volatility factors $\sigma_n(t)$ as inputs without indicating how they might be specified. In practice, these coefficients are chosen to calibrate a model to market prices of actively traded derivatives, especially caps and swaptions. (The model is automatically calibrated to bond prices through the relations (3.105) and (3.106).) Once the model has been calibrated to the market, it can be used to price less liquid

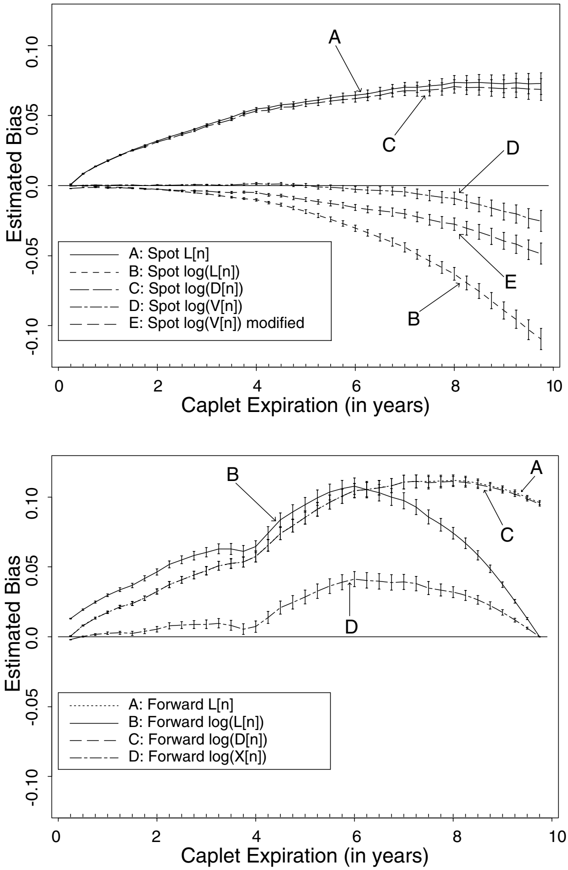

Fig. 3.19. Comparison of biases in caplet pricing for various simulation methods. Top panel uses spot measure; method A is an Euler scheme for L_n and methods B–E are Euler schemes for log variables. Bottom panel uses the forward measure P_{M+1}; method A is an Euler scheme for L_n and methods B–D are Euler schemes for log variables.

instruments for which market prices may not be readily available. Accurate and efficient calibration is a major topic in its own right and we can only touch on the key issues. For a more extensive treatment, see James and Webber [194] and Rebonato [303]. Similar considerations apply in both the HJM framework and in LIBOR market models; we discuss calibration in the LIBOR setting because it is somewhat simpler. Indeed, convenience in calibration is one of the main advantages of this class of models.

The variables $\sigma_n(t)$ are the primary determinants of both the level of volatility in forward rates and the correlations between forward rate. It is often useful to distinguish these two aspects and we will consider the overall level of volatility first. Suppose we are given the market price of a caplet for the interval $[T_n, T_{n+1}]$ and from this price we calculate an implied volatility v_n by inverting the Black formula (3.117). (We can assume that the other parameters of the formula are known.) If we choose σ_n to be any deterministic $\Re^d$-valued function satisfying

$$\frac{1}{T_n}\int_0^{T_n} \|\sigma_n(t)\|^2\,dt = v_n^2,$$

then we know from the discussion in Section 3.7.2 that the model is calibrated to the market price of this caplet, because the model's caplet price is given by the Black formula with implied volatility equal to the square root of the expression on the left. By imposing this constraint on all of the σ_j, we ensure that the model is calibrated to all caplet prices. (As a practical matter, it may be necessary to infer the prices of individual caplets from the prices of caps, which are portfolios of caplets. For simplicity, we assume caplet prices are available.)

Because LIBOR market models do not specify interest rates over accrual periods shorter than the intervals $[T_i, T_{i+1}]$, it is natural and customary to restrict attention to functions $\sigma_n(t)$ that are constant between tenor dates. We take each σ_n to be right-continuous and thus denote by $\sigma_n(T_i)$ its value over the interval $[T_i, T_{i+1})$. Suppose, for a moment, that the model is driven by a scalar Brownian motion, so $d = 1$ and each σ_n is scalar valued. In this case, it is convenient to think of the volatility structure as specifed through a lower-triangular matrix of the following form:

$$\begin{pmatrix} \sigma_1(T_0) & & & \\ \sigma_2(T_0) & \sigma_2(T_1) & & \\ \vdots & \vdots & \ddots & \\ \sigma_M(T_0) & \sigma_M(T_1) & \cdots & \sigma_M(T_{M-1}) \end{pmatrix}.$$

The upper half of the matrix is empty (or irrelevant) because each $L_n(t)$ ceases to be meaningful for $t > T_n$. In this setting, we have

$$\int_0^{T_n} \sigma_n^2(t)\,dt = \sigma_n^2(T_0)\delta_0 + \sigma_n^2(T_1)\delta_1 + \cdots + \sigma_n^2(T_{n-1})\delta_{n-1},$$

so caplet prices impose a constraint on the sums of squares along each row of the matrix.

The volatility structure is *stationary* if $\sigma_n(t)$ depends on n and t only through the difference $T_n - t$. For a stationary, single-factor, piecewise constant volatility structure, the matrix above takes the form

$$\begin{pmatrix} \sigma(1) & & & \\ \sigma(2) & \sigma(1) & & \\ \vdots & \vdots & \ddots & \\ \sigma(M) & \sigma(M-1) & \cdots & \sigma(1) \end{pmatrix}$$

for some values $\sigma(1), \ldots, \sigma(M)$. (Think of $\sigma(i)$ as the volatility of a forward rate i periods away from maturity.) In this case, the number of variables just equals the number of caplet maturities to which the model may be calibrated. Calibrating to additional instruments requires introducing nonstationarity or additional factors.

In a multifactor model (i.e., $d \geq 2$) we can think of replacing the entries $\sigma_n(T_i)$ in the volatility matrix with the norms $\|\sigma_n(T_i)\|$, since the $\sigma_n(T_i)$ are now vectors. With piecewise constant values, this gives

$$\int_0^{T_n} \|\sigma_n(t)\|^2 \, dt = \|\sigma_n(T_0)\|^2 \delta_0 + \|\sigma_n(T_1)\|^2 \delta_1 + \cdots + \|\sigma_n(T_{n-1})\|^2 \delta_{n-1},$$

so caplet implied volatilities continue to constrain the sums of squares along each row. This also indicates that taking $d \geq 2$ does not provide additional flexibility in matching these implied volatilities.

The potential value of a multifactor model lies in capturing correlations between forward rates of different maturities. For example, from the Euler approximation in (3.120), we see that over a short time interval the correlation between the increments of $\log L_j(t)$ and $\log L_k(t)$ is approximately

$$\frac{\sigma_k(t)^\top \sigma_j(t)}{\|\sigma_k(t)\| \, \|\sigma_j(t)\|}.$$

These correlations are often chosen to match market prices of swaptions (which, unlike caps, are sensitive to rate correlations) or to match historical correlations.

In the stationary case, we can visualize the volatility factors by graphing them as functions of time to maturity. This can be useful in interpreting the correlations they induce. Figure 3.20 illustrates three hypothetical factors in a model with $M = 15$. Because the volatility is assumed stationary, we may write $\sigma_n(T_i) = \sigma(n-i)$ for some vectors $\sigma(1), \ldots, \sigma(M)$. In a three-factor model, each $\sigma(i)$ has three components. The three curves in Figure 3.20 are graphs of the three components as functions of time to maturity. If we fix a time to maturity on the horizontal axis, the total volatility at that point is given by the sums of squares of the three components; the inner products of these three-dimensional vectors at different times determine the correlations between the forward rates.

Notice that the first factor in Figure 3.20 has the same sign for all maturities; regardless of the sign of the increment of the driving Brownian motion, this factor moves all forward rates in the same direction and functions approximately as a parallel shift. The second factor has values of opposite signs at short and long maturities and will thus have the effect of tilting the forward curve (up if the increment in the second component of the driving Brownian motion is positive and down if it is negative). The third factor bends the forward curve by moving intermediate maturities in the opposite direction of long and short maturities, the direction depending on the sign of the increment of the third component of the driving Brownian motion.

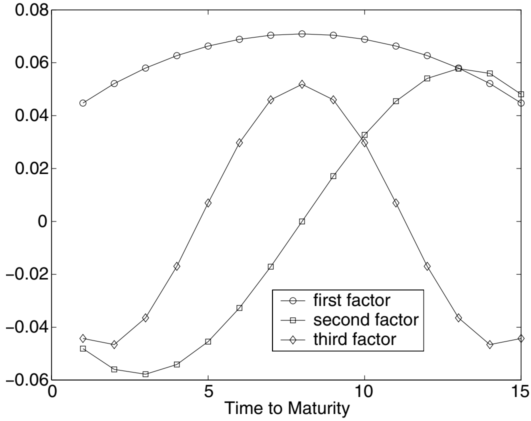

Fig. 3.20. Hypothetical volatility factors.

The hypothetical factors in Figure 3.20 are the first three principal components of the matrix

$$0.12^2 \exp((-0.8\sqrt{|i-j|})), \quad i,j = 1,\ldots,15.$$

More precisely, they are the first three eigenvectors of this matrix as ranked by their eigenvalues, scaled to have length equal to their eigenvalues. It is common in practice to use the principal components of either the covariance matrix or the correlation matrix of changes in forward rates in choosing a factor structure. Principal components analysis typically produces the qualitative features of the hypothetical example in Figure 3.20; see, e.g., the examples in James and Webber [194] or Rebonato [304].

An important feature of LIBOR market models is that a good deal of calibration can be accomplished through closed form expressions or effective approximations for the prices of caps and swaptions. This makes calibration fast. In the absence of formulas or approximations, calibration is an iterative procedure requiring repeated simulation at various parameter values until the model price matches the market. Because each simulation can be quite time consuming, calibration through simulation can be onerous.