Chapter 8

Term Structure Models

In these notes we put the interest rate curve in motion. Chapter 7 built today's curve out of today's quotes, which is enough to price anything linear; here we ask how the whole curve may diffuse without admitting arbitrage, and price the products whose value depends on the answer. We work in a single curve framework, so that there is one interest rate in the market and the projection and discount curves of chapter 7 coincide.

8.1 What We Are Modelling

We take from chapter 7 the discount factors $P(t,T)$, the instantaneous forward rate

$$f(t,T) = -\frac{\partial}{\partial T}\ln P(t,T), \qquad P(t,T) = e^{-\int_t^T f(t,s)ds},$$

the short rate $r(t) = f(t,t)$, and the initial curve $f(0,\cdot)$ inferred from market instruments. All of that was a description of the curve at one instant.

The difference between interest rates and equities now becomes the whole problem. In chapter 5 the state of the market was a number S_t and we gave it a diffusion. Here the state is the curve $T \mapsto f(t,T)$, an entire function, and we have to give *that* a diffusion. Two questions arise immediately and neither has an analogue in the equity case. What drift is consistent with no arbitrage, when every point of the curve is the price of a traded bond and they must all be consistent with one another? And can the resulting infinite dimensional object be made to depend on finitely many state variables, so that anything can actually be computed?

The first question is answered completely by the Heath-Jarrow-Morton framework, which is the subject of the next section. The second is answered only under a restriction on the volatility, which gives the Hull-White model here and the quasi-Gaussian models of chapter 12.

8.2 Heath-Jarrow-Morton Framework

The first question of the previous section — which drifts are consistent with no arbitrage — has a complete answer, and the answer is that there is no freedom at all. Once the volatility of the curve is chosen, the drift is determined. This is the Heath-Jarrow-Morton drift condition, and everything in chapter 12 and chapter 14 that concerns interest rates rests on it.

The starting point, and why it is not an assumption

Take as given a filtered probability space carrying a Brownian motion W, and suppose each discount bond follows

$$\frac{dP(t,T)}{P(t,T)} = r(t)\,dt + \nu(t,T)\,dW_t, \tag{8.1}$$

with $\nu(t,T)$ any adapted process. The volatility is free; the drift is not.

Chapter 6 establishes that a measure under which every traded asset discounted by the money market account is a martingale exists if and only if there is no arbitrage. So (8.1) is the general form of an arbitrage-free bond price, not a special one, and ν is where all the modelling freedom lives.

Two constraints on ν come free and both matter below. A bond pays one unit at its own maturity, so $P(T,T) = 1$ with certainty, which forces

$$\nu(t,t) = 0. \tag{8.2}$$

A bond about to mature cannot be volatile. And $\nu(t,T)$ is a volatility of a *price*, indexed by the maturity of the bond, so the family $\{\nu(t,\cdot)\}$ is called the volatility term structure.

The drift condition

Theorem 8.1 (Heath, Jarrow and Morton)**.** *Let bond prices follow* (8.1) *under the risk-neutral measure and set*

$$\sigma(t,T) = -\frac{\partial \nu}{\partial T}(t,T). \tag{8.3}$$

Then the instantaneous forward rate satisfies

$$df(t,T) = \sigma(t,T)\left(\int_t^T \sigma(t,s)\,ds\right)dt \;+\; \sigma(t,T)\,dW_t. \tag{8.4}$$

The drift is a function of the volatility alone: no choice remains.

Proof. Apply Itô's lemma to $\ln P(t,T)$ using (8.1). The logarithm of a process with proportional volatility ν picks up the usual variance correction,

$$d\ln P(t,T) = \Big(r(t) - \tfrac{1}{2}\nu^2(t,T)\Big)dt + \nu(t,T)\,dW_t.$$

Now differentiate in the maturity. Since $f(t,T) = -\partial_T \ln P(t,T)$ and the two derivatives commute,

$$df(t,T) = -\frac{\partial}{\partial T}\,d\ln P(t,T) = \frac{\partial}{\partial T}\Big[\tfrac{1}{2}\nu^2(t,T)\Big]dt - \frac{\partial \nu}{\partial T}(t,T)\,dW_t,$$

the short rate $r(t)$ having no T in it and so contributing nothing. Carrying out the derivative of the square and substituting (8.3),

$$df(t,T) = \nu(t,T)\,\frac{\partial \nu}{\partial T}(t,T)\,dt + \sigma(t,T)\,dW_t = -\nu(t,T)\,\sigma(t,T)\,dt + \sigma(t,T)\,dW_t.$$

It remains to identify $-\nu$. Integrating (8.3) in the maturity from t to T and using (8.2) to fix the constant,

$$-\nu(t,T) = \int_t^T \sigma(t,s)\,ds, \tag{8.5}$$

which is (8.4). □

Remark (Where each hypothesis went)**.** The martingale measure supplied the drift rP, and with it the fact that the drift of the *bond* carries no maturity dependence at all. That is what makes $r(t)$ disappear on differentiating, and it is the whole reason a condition emerges: the forward rate's drift is left holding only the variance correction.

The variance correction $-\frac{1}{2}\nu^2$ is what makes the condition non-trivial. Without Itô's second-order term the forward rate would be driftless.

And $P(T,T) = 1$ fixed the constant of integration in (8.5). Without it ν would be determined only up to a function of t, and the drift with it. A bond maturing at par is what closes the system.

Structure (The bonds are not independent assets)**.** The result is surprising on first meeting — an entire curve of assets and no drift may be chosen — and stops being surprising once one notices what the objects are.

Bonds of different maturities are not separate securities that happen to be correlated. They are all functions of one object, the curve, and a bond is the integral of the curve over an interval. So the family $\{P(t,T)\}_T$ is a single degree of freedom per Brownian motion dressed as a continuum, and demanding that every member of it be a martingale after discounting is demanding one condition of one object, over and over. That the conditions are compatible at all is the content; that they leave no freedom follows.

The equity case has no analogue because there is nothing to be consistent with. A stock's drift is pinned by the same argument — it is r under the martingale measure — but a stock is one asset and the condition is one equation. Here it is an equation for every maturity, and (8.4) is what the whole system collapses to.

Remark (Testing it, and what failure looks like)**.** The drift condition has a consequence sharp enough to check: a model built with it reprices the curve it started from. Simulating the whole forward surface under (8.4) and averaging $\exp(-\int_0^T r)$ returns $P(0,T)$, and nothing in the simulation is told what that number is.[1]

Setting the drift to zero instead — which is what one writes down if the convexity term is forgotten — breaks it. The drift is second order in the volatility, so at one per cent volatility the error is under a basis point of discount factor at one year and grows to half a per cent at ten.[2] It also has a sign: without the convexity drift the forwards are too low and the discount factors too high. So the failure is the uncomfortable kind — nearly invisible on the short instruments a desk would check first, and compounding out to the maturities where the exotic actually sits.

The same equation in time to maturity

Equation (8.4) indexes the curve by maturity date, which is how the market quotes it, and is the wrong coordinate for almost every other purpose. A bond's maturity is a fixed date; a modeller cares about the ten year point, which is a different bond every day.

[1] `checked at one, three and five years`.
[2] `measured, and one-signed`.

Definition 8.2 (Musiela parametrisation)**.** Write the curve by time to maturity rather than by maturity date,

$$r_t(x) = f(t, t + x), \qquad x \geq 0. \tag{8.6}$$

Calculation 8.3 (The transport term)**.** Holding x fixed means letting the maturity date advance with the clock, so the chain rule contributes an extra term. Writing $\bar{\sigma}(t, x) = \sigma(t, t + x)$,

$$dr_t(x) = \left(\frac{\partial r_t}{\partial x}(x) + \bar{\sigma}(t, x) \int_0^x \bar{\sigma}(t, y)\, dy \right) dt \;+\; \bar{\sigma}(t, x)\, dW_t. \tag{8.7}$$

The new piece is $\partial_x r_t$, and it is not a modelling term: it is the statement that if nothing happens, tomorrow's nine year rate is today's ten year rate. The curve slides along itself.

Structure (Rolling down is a transport equation)**.** Equation (8.7) is a stochastic partial differential equation on a space of curves, and the operator $\partial/\partial x$ appearing in its drift is the generator of translation. In the absence of volatility it reduces to $\partial_t r = \partial_x r$, whose solution is $r_t(x) = r_0(t + x)$ — the initial curve, read further along. That is precisely the carry and roll-down that chapter 22 computes as a trading quantity, arriving here as the deterministic part of the dynamics.

The parametrisation earns its keep twice more in these notes. Chapter 12 needs it because the criterion for a finite dimensional realisation is a statement about the Lie algebra generated by the drift and volatility vector fields, and the transport term is one of those fields — bracketing against it is what differentiates the maturity profile and produces the quasi-exponential condition. And it is the right coordinate for any question about stationarity, since a model with time-homogeneous $\bar{\sigma}$ is one whose dynamics look the same on every date, which is not visible in the maturity-date coordinate at all.

What is left to choose

If $\sigma(t, T)$ is deterministic then (8.4) makes $f(t, T)$ Gaussian and bond prices lognormal, which is the Gaussian HJM class and the setting for everything in this chapter.

One Brownian motion has been used for the whole curve, and that is a restriction rather than a simplification. Every point of the curve is driven by the same dW_t, in proportion to its own $\sigma(t, T)$, so the increments of any two forward rates are perfectly correlated: the curve can shift and steepen as it shifts, but it cannot twist. Chapter 12 measures what that costs on a spread option and shows the cost is large. Relaxing it means several Brownian motions with their own volatility term structures, and (8.4) generalises with a sum over factors and nothing else changed.

8.3 Hull-White Model

Consider one-factor HJM diffusion. Instrument pricing is much more tractable if we restrict the volatility term-structure to be deterministic, separable and of the form:

$$\sigma(t, T) = \sigma(t) e^{-\int_t^T \alpha(s) ds}$$

Hull-White in general permits both $\alpha(s)$ and $\sigma(t)$ to be time-dependent. We take both to be constant, $\alpha(s) = \alpha$ and $\sigma(t) = \sigma$, which is the version usually meant by the Hull-White (or extended Vasicek) model and the version used in practice, since a fitted $\sigma(t)$ makes the future volatility term structure implied by the model non-stationary. Nothing structural in what follows depends on σ being constant, only on $\sigma(t,T)$ being deterministic and separable, so the derivation goes through for a deterministic $\sigma(t)$ with the closed forms replaced by integrals against σ^2.

In this case,

$$\begin{aligned}
df(t,T) =&\sigma e^{-\alpha(T-t)}\left(\int_t^T \sigma e^{-\alpha(u-t)}du\right)dt + \sigma e^{-\alpha(T-t)}dW_t \\
=&\sigma^2 e^{-\alpha(T-t)}\left(\int_t^T e^{-\alpha(u-t)}du\right)dt + \sigma e^{-\alpha(T-t)}dW_t \\
=&\frac{\sigma^2}{\alpha}e^{-\alpha(T-t)}\left(1-e^{-\alpha(T-t)}\right)dt + \sigma e^{-\alpha(T-t)}dW_t
\end{aligned}$$

$$\begin{aligned}
f(t,T) =&f(0,T) + \int_0^t \frac{\sigma^2}{\alpha}e^{-\alpha(T-u)}\left(1-e^{-\alpha(T-u)}\right)du + \int_0^t \sigma e^{-\alpha(T-u)}dW_u \\
=&f(0,T) + \frac{\sigma^2}{\alpha}\int_0^t \left(e^{-\alpha(T-u)} - e^{-2\alpha(T-u)}\right)du + \sigma\int_0^t e^{-\alpha(T-u)}dW_u \\
=&f(0,T) + \frac{\sigma^2}{\alpha^2}\left(e^{-\alpha(T-t)} - e^{-\alpha T} - \frac{1}{2}(e^{-2\alpha(T-t)} - e^{-2\alpha T})\right) + \sigma\int_0^t e^{-\alpha(T-u)}dW_u
\end{aligned}$$

Using the Itô isometry, $f(t,T)$ therefore is normally distributed as

$$f(t,T) \sim N\left(f(0,T) + \frac{\sigma^2}{\alpha^2}\left(\frac{e^{\alpha t}-1}{e^{\alpha T}} - \frac{e^{2\alpha t}-1}{2e^{2\alpha T}}\right), \frac{\sigma^2}{2\alpha}\frac{e^{2\alpha t}-1}{e^{2\alpha T}}\right)$$

where $N(\mu,\sigma^2)$ is a normal distribution with mean μ and variance σ^2.

We then have

$$\begin{aligned}
\int_t^T f(t,s)ds =& \int_t^T \left(f(0,s) + \frac{\sigma^2}{\alpha^2}\left(\frac{e^{\alpha t}-1}{e^{\alpha s}} - \frac{e^{2\alpha t}-1}{2e^{2\alpha s}}\right)\right)ds + \sigma\int_t^T\int_0^t e^{-\alpha(s-u)}dW_u ds \\
=& \int_t^T f(0,s)ds + \frac{\sigma^2}{\alpha^3}\left((e^{\alpha t}-1)(e^{-\alpha t} - e^{-\alpha T}) - \frac{1}{4}(e^{2\alpha t}-1)(e^{-2\alpha t} - e^{-2\alpha T})\right) \\
&+ \sigma\int_t^T\int_0^t e^{-\alpha(s-u)}dW_u ds \\
=& \int_t^T f(0,s)ds + \frac{\sigma^2}{\alpha^3}\left((1-e^{-\alpha t})(1-e^{-\alpha(T-t)}) - \frac{1}{4}(1-e^{-2\alpha t})(1-e^{-2\alpha(T-t)})\right) \\
&+ \sigma\int_t^T\int_0^t e^{-\alpha(s-u)}dW_u ds \\
=& \int_t^T f(0,s)ds + \frac{\sigma^2}{\alpha^3}\left((1-e^{-\alpha t})(1-e^{-\alpha(T-t)}) - \frac{1}{4}(1-e^{-2\alpha t})(1-e^{-2\alpha(T-t)})\right) \\
&+ \frac{\sigma}{\alpha}(e^{-\alpha t} - e^{-\alpha T})\int_0^t e^{\alpha u}dW_u
\end{aligned}$$

Therefore,

$$\int_t^T f(t,s)ds \sim N\Big(\int_t^T f(0,s)ds + \frac{\sigma^2}{\alpha^3}\big((1-e^{-\alpha t})(1-e^{-\alpha(T-t)}) - \tfrac{1}{4}(1-e^{-2\alpha t})(1-e^{-2\alpha(T-t)})\big), \\ \frac{\sigma^2}{2\alpha^3}(1-e^{-\alpha(T-t)})^2(1-e^{-2\alpha t})\Big) \tag{8.8}$$

and the bond price

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

is log-normally distributed.

The Hull-White model can be specified completely in terms of the short rate diffusion as well, and is therefore called a short-rate model. Let us see how this re-parametrization works.

$$\begin{aligned} r(t) =& f(t,t) \\ =& f(0,t) + \frac{\sigma^2}{\alpha^2}\left(1-e^{-\alpha t} - \frac{1}{2}(1-e^{-2\alpha t})\right) + \sigma\int_0^t e^{-\alpha(t-u)}dW_u \\ =& f(0,t) + \frac{\sigma^2}{\alpha^2}\left(1-e^{-\alpha t} - \frac{1}{2}(1-e^{-2\alpha t})\right) + \sigma e^{-\alpha t}\int_0^t e^{\alpha u}dW_u \\ =& f(0,t) + \frac{\sigma^2}{2\alpha^2}\big(1-e^{-\alpha t}\big)^2 + \sigma e^{-\alpha t}\int_0^t e^{\alpha u}dW_u \end{aligned}$$

Notice here that the stochastic evolution of both $\int_t^T f(t,s)ds$ and $r(t)$ is determined by the term $\int_0^t e^{\alpha u}dW_u$ and therefore the bond yields $\ln P(t,T)$ can be specified as a linear function of $r(t)$ with coefficients of this linear relationship being deterministic functions of t and T. Models with such linear relationship between short rate and yield term structure are known as affine term structure models.

$$\begin{aligned} P(t,T) =& \exp\big(A_1(t,T) + B_1(t,T)\int_0^t e^{\alpha u}dW_u\big) \\ =& \exp\big(A_2(t,T) + B_2(t,T)f(t,T)\big) \\ =& \exp\big(A_3(t,T) + B_3(t,T)r(t)\big) \end{aligned}$$

where $A_i(t,T)$ and $B_i(t,T)$ are deterministic functions of t and T.

Getting back to short rate diffusion specification, we have

$$\begin{aligned} dr(t) =& \left(\frac{\partial}{\partial t}f(0,t) + \frac{\sigma^2}{\alpha}(e^{-\alpha t} - e^{-2\alpha t}) - \sigma\alpha e^{-\alpha t}\int_0^t e^{\alpha u}dW_u\right)dt + \sigma e^{-\alpha t}e^{\alpha t}dW_t \\ =& \left(\frac{\partial}{\partial t}f(0,t) + \frac{\sigma^2}{\alpha}(e^{-\alpha t} - e^{-2\alpha t}) - \alpha\left(r(t) - (f(0,t) + \frac{\sigma^2}{2\alpha^2}\big(1-e^{-\alpha t}\big)^2)\right)\right)dt + \sigma dW_t \\ =& \left(\frac{\partial}{\partial t}f(0,t) + \frac{\sigma^2}{\alpha}(e^{-\alpha t} - e^{-2\alpha t}) + \alpha(f(0,t) + \frac{\sigma^2}{2\alpha^2}\big(1-e^{-\alpha t}\big)^2) - \alpha r(t)\right)dt + \sigma dW_t \\ =& \big(\theta(t) - \alpha r(t)\big)dt + \sigma dW_t \end{aligned}$$

where we have introduced the parameter $\theta(t)$ which is specified of terms of other deterministic parameters α, σ and initial instantaneous forward rate curve $f(0,\cdot)$. From this specification, we have another interpretation of α as the mean-reversion strength of short rate.

8.4 Future-Forward Basis

Here is the first place a model is unavoidable. Chapter 7 priced a forward rate agreement off today's curve alone, with no assumption about how rates move. A futures contract on the same rate cannot be priced that way, and the difference between the two prices is a pure model quantity. We work it through for LIBOR futures and forwards. LIBOR itself has been decommissioned, but the calculation is the instructive one and the logic transfers unchanged to futures on other benchmark rates.

Recall from chapter 7 the reference rate

$$L(t,T,T+\delta) = \frac{1}{\delta}\left(\frac{P(t,T)}{P(t,T+\delta)} - 1\right),$$

and the forward rate agreement, which pays $\delta(L(T,T,T+\delta) - R)/(1+\delta L(T,T,T+\delta))$ at T. We write FRA for the value of R making it worth zero. First we confirm what chapter 7 asserted, that this is the forward rate, and identify the measure in which that is the natural statement.

$$\begin{aligned}
0 =& P(0,T)\mathbb{E}^T[\frac{\delta(L(T,T,T+\delta) - \mathrm{FRA}(0))}{1+\delta L(T,T,T+\delta)}] \\
=& P(0,T)\mathbb{E}^T[\delta(L(T,T,T+\delta) - \mathrm{FRA}(0))P(T,T+\delta)] \\
=& P(0,T)\mathbb{E}^{T+\delta}[\delta(L(T,T,T+\delta) - \mathrm{FRA}(0))P(T,T+\delta)\frac{P(T,T)P(0,T+\delta)}{P(0,T)P(T,T+\delta)}] \\
=& P(0,T+\delta)\mathbb{E}^{T+\delta}[\delta(L(T,T,T+\delta) - \mathrm{FRA}(0))] \\
\mathrm{FRA}(0) =& \mathbb{E}^{T+\delta}[L(T,T,T+\delta)]
\end{aligned}$$

Notice that $L(t,T,T+\delta)$ is a ratio of two prices plus a constant, and scaled by a constant. Therefore it is a martingale in the measure with $P(t,T)$ as numeraire i.e. T-forward measure. Therefore

$$\mathrm{FRA}(t) = \mathbb{E}_t^{T+\delta}[L(T,T,T+\delta)] = L(t,T,T+\delta)$$

is known as forward LIBOR rate.

Definition 8.4 (Eurodollar Futures)**.** Eurodollar futures are daily margined contracts that at maturity T settle at value $1 - L(T,T+\delta)$ which is published at T.

Eurodollar futures are settled in market at value $1 - L(T,T+\delta)$ but for ease of comparison we will take futures value at T $\mathrm{Fut}(T)$ to be $L(T,T+\delta)$. We can get the true futures value from $1 - \mathrm{Fut}(0)$. Nothing is lost in the substitution: the margin condition below sets a price to zero and sees the quote only through $\frac{\partial}{\partial s}\mathrm{Fut}(s)$, so it is blind to both the sign and the additive constant, and $\mathbb{E}^{\mathbb{P}}[1-L] = 1 - \mathbb{E}^{\mathbb{P}}[L]$ carries the martingale property between the two conventions. That is a fact about $L \mapsto 1 - L$ being affine, and it fails for a contract settling on anything else — the Treasury futures of chapter 22 settle against a bond price, convex in its yield, so there the implied yield is not a martingale even though the price is.

Futures contracts on LIBOR are different from forwards in two major ways.

1. Futures contracts are daily margined. That is, the buyer does not pay for the contract initially, except for initial margin which is usually much smaller than the futures price. Then the movement in futures prices are settled daily; if the futures price goes up, the buyer has cash put in his margin account and if the futures price goes down, the buyer has cash withdrawn from this account.

2. Futures contracts settle at T, the start of the period, for the rate that has just fixed, and *without discounting*. A forward rate agreement also settles at T by market convention — but with the $1/(1+\delta L)$ factor of chapter 7, which is exactly $P(T, T+\delta)$, so its payoff is worth the same as $\delta(L-R)$ paid at $T+\delta$. The calculation below shows that equivalence rather than assuming it.

The futures price $\mathrm{Fut}(t)$ is such that price of future cashflows from daily margining is 0 i.e. for all t

$$
\begin{aligned}
0 =& \frac{1}{D_t}\mathbb{E}_t^{\mathbb{P}}\left[\int_t^T D_s \frac{\partial}{\partial s}\mathrm{Fut}(s) ds\right] \\
0 =& \mathbb{E}_t^{\mathbb{P}}\left[\int_t^T D_s \frac{\partial}{\partial s}\mathrm{Fut}(s) ds\right]
\end{aligned}
$$

where $\frac{\partial}{\partial s}\mathrm{Fut}(s)ds$ is the cashflow from margin in the time interval $(s, s+ds)$. Differentiating with respect to t,

$$
\begin{aligned}
0 =& \mathbb{E}_t^{\mathbb{P}}\left[-D_t \frac{\partial}{\partial t}\mathrm{Fut}(t)\right] \\
=& -D_t \mathbb{E}_t^{\mathbb{P}}\left[\frac{\partial}{\partial t}\mathrm{Fut}(t)\right] \\
0 =& \mathbb{E}_t^{\mathbb{P}}\left[\frac{\partial}{\partial t}\mathrm{Fut}(t)\right]
\end{aligned}
$$

Integrating and using iterated conditioning, we get

$$
\begin{aligned}
\mathrm{Fut}(0) =& \mathbb{E}_0^{\mathbb{P}}[\mathrm{Fut}(T)] \\
=& \mathbb{E}_0^{\mathbb{P}}[L(T, T, T+\delta)]
\end{aligned}
$$

which says that the current futures price is the expected settlement price of the future in risk neutral measure.

Remark (Quoting in price or in rate)**.** The basis derived below is positive in rate terms, so the futures rate sits above the forward rate; in the market's price quote the ordering reverses and the future trades below the forward.

The future-forward basis is largely sensitive to the covariance between rate fixing and money market discount factor.

$$\begin{aligned}
\mathrm{Fut}(0)-\mathrm{FRA}(0) =& \mathbb{E}^{\mathbb{P}}[L(T,T,T+\delta)] - \mathbb{E}^{T+\delta}[L(T,T,T+\delta)] \\
=& \mathbb{E}^{\mathbb{P}}[L(T,T,T+\delta)] - \mathbb{E}^{\mathbb{P}}[L(T,T,T+\delta)\frac{P(T,T+\delta)D_T}{P(0,T+\delta)D_0}] \\
=& \frac{P(0,T+\delta)\mathbb{E}^{\mathbb{P}}[L(T,T,T+\delta)] - \mathbb{E}^{\mathbb{P}}[L(T,T,T+\delta)P(T,T+\delta)D_T]}{P(0,T+\delta)} \\
=& \frac{\mathbb{E}^{\mathbb{P}}[P(T,T+\delta)D_T]\mathbb{E}^{\mathbb{P}}[L(T,T,T+\delta)] - \mathbb{E}^{\mathbb{P}}[L(T,T,T+\delta)P(T,T+\delta)D_T]}{P(0,T+\delta)} \\
=& -\frac{\mathrm{Cov}^{\mathbb{P}}(P(T,T+\delta)D_T, L(T,T,T+\delta))}{P(0,T+\delta)} \\
=& -\frac{\mathrm{Cov}^{\mathbb{P}}(P(T,T+\delta)D_T, \frac{1}{P(T,T+\delta)})}{\delta P(0,T+\delta)}
\end{aligned}$$

where $\mathrm{Cov}^{\mathbb{P}}(X,Y)$ is the covariance of X and Y in risk neutral measure.

Note the sign. The covariance itself is negative: when rates rise the fixing $L(T,T,T+\delta)$ rises, while both the bond price $P(T,T+\delta)$ and the money market discount factor D_T fall, so the Radon-Nikodym density $P(T,T+\delta)D_T/P(0,T+\delta)$ moves against the fixing. With the leading minus sign the basis is therefore positive, and the futures rate sits above the forward rate. This is the precise sense in which daily margining is worth something to the buyer of a future: the contract pays the fixing undiscounted, so the cash it throws off is largest exactly when it is worth least to reinvest.

Worked case 8.5 (The basis in the Hull-White model)**.** The covariance above is exact but is not a number until a model is chosen. Hull-White is the model this chapter has already built, so the computation is available without anything new, and it shows how a basis is extracted once a model is fixed. The theory does not rest on it: a reader may take expression 8.9 off the last line and carry on.

Both legs are simply compounded yields, $\delta L(T,T,T+\delta) = 1/P(T,T+\delta)-1$ and $\delta L(0,T,T+\delta) = P(0,T)/P(0,T+\delta)-1$, and $\mathbb{E}^{T+\delta}[L(T,T,T+\delta)] = L(0,T,T+\delta)$ because the forward rate is a martingale in its own forward measure. The constants cancel and

$$\mathrm{Fut}(0)-\mathrm{FRA}(0) = \frac{1}{\delta}\left(\mathbb{E}^{\mathbb{P}}\Big[\frac{1}{P(T,T+\delta)}\Big] - \frac{P(0,T)}{P(0,T+\delta)}\right),$$

so everything turns on $\mathbb{E}^{\mathbb{P}}[1/P(T,T+\delta)] = \mathbb{E}^{\mathbb{P}}[e^{\int_T^{T+\delta} f(T,s)ds}]$. Expression 8.8 makes the exponent normal with mean μ and variance V, and a lognormal has mean $e^{\mu+V/2}$. Setting $t=T$ there, replacing the upper limit T by $T+\delta$, and abbreviating $u = 1-e^{-\alpha T}$ and $v = 1-e^{-\alpha\delta}$ so that $1-e^{-2\alpha T} = u(2-u)$ and $1-e^{-2\alpha\delta} = v(2-v)$,

$$\mu = \int_T^{T+\delta} f(0,s)ds + M, \qquad M = \frac{\sigma^2}{4\alpha^3}uv\big(4-(2-u)(2-v)\big) = \frac{\sigma^2}{4\alpha^3}uv\big(2u+2v-uv\big)$$
$$V = \frac{\sigma^2}{2\alpha^3}u(2-u)v^2$$

where M names the part of the mean that today's curve does not already contain. The curve supplies the rest, since $P(0,T)/P(0,T+\delta) = e^{\int_T^{T+\delta} f(0,s)ds} = 1+\delta L(0,T,T+\delta)$, and therefore

$$\mathrm{Fut}(0) - \mathrm{FRA}(0) = \frac{1}{\delta}\frac{P(0,T)}{P(0,T+\delta)}\big(e^{M+\frac{V}{2}}-1\big) = \left(\frac{1}{\delta} + L(0,T,T+\delta)\right)\big(e^{M+\frac{V}{2}}-1\big).$$

Since $u, v \in (0,1)$ both $2u+2v-uv$ and $u(2-u)$ are positive, so M and V are, and the basis is strictly positive: the futures rate always sits above the forward rate, as the covariance argument above anticipated.

The Ho-Lee model is the $\alpha \to 0$ limit, in which the forward rate volatility $\sigma(t,T) = \sigma$ no longer decays with maturity. Then $u \to \alpha T$ and $v \to \alpha\delta$, so $M \to \frac{\sigma^2}{2}T\delta(T+\delta)$ and $V \to \sigma^2 T\delta^2$, giving $M + \frac{V}{2} \to \frac{\sigma^2}{2}T\delta(T+2\delta)$. Both are of order $\sigma^2 T^2$ and hence small, so $e^{M+V/2}-1 \approx M+V/2$; and $\delta L(0,T,T+\delta) \ll 1$, so the leading factor is approximately $1/\delta$. The Ho-Lee basis is therefore

$$\mathrm{Fut}(0) - \mathrm{FRA}(0) \approx \frac{\sigma^2}{2}T(T+2\delta) \tag{8.9}$$

Remark. The basis is often quoted in textbooks as $\frac{1}{2}\sigma^2 T_1 T_2$ with $T_1 = T$ and $T_2 = T+\delta$, which is not what expression 8.9 says. The discrepancy is instructive, and it is entirely a question of which compounding convention the rate is quoted in.

The term M on its own measures the drift of the yield $\int_T^{T+\delta} f(T,s)ds$ away from today's forward yield, i.e. the basis on the *continuously compounded* rate $R_c = -\frac{1}{\delta}\ln P(T,T+\delta)$:

$$\mathbb{E}^{\mathbb{P}}[R_c] - \frac{1}{\delta}\int_T^{T+\delta} f(0,s)ds = \frac{M}{\delta} \to \frac{\sigma^2}{2}T(T+\delta) = \frac{\sigma^2}{2}T_1T_2$$

This is exactly the textbook expression. In Hull-White it can be written in the familiar form, with $B(t,T) = \frac{1}{\alpha}(1-e^{-\alpha(T-t)})$ so that $B(0,T) = u/\alpha$ and $B(T,T+\delta) = v/\alpha$,

$$\frac{M}{\delta} = \frac{\sigma^2}{4\alpha\delta}B(T,T+\delta)\Big(B(T,T+\delta)\big(1-e^{-2\alpha T}\big) + 2\alpha B(0,T)^2\Big)$$

LIBOR, however, is simply compounded, and $\frac{1}{\delta}(\frac{1}{P(T,T+\delta)}-1)$ is a convex function of the yield. That convexity contributes a further $\frac{1}{2}\sigma^2 T\delta$ to the basis and turns $T+\delta$ into $T+2\delta$. The two agree to relative order δ/T, so for a three month rate the textbook formula understates the LIBOR basis by about 11% at $T=2$ and about 5% at $T=5$; mean reversion pulls in the opposite direction and is the larger correction at realistic α.

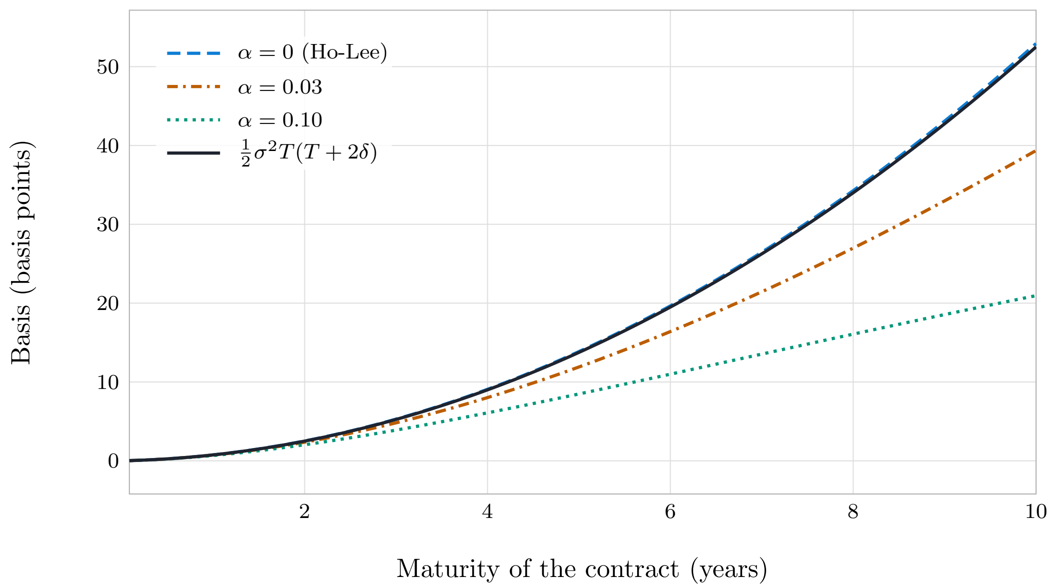

Figure 8.1: The basis against the maturity of the contract, for a three month rate. It grows like $T(T+2\delta)$ rather than like T^2 — doubling the maturity from one year to two multiplies it by 3.3, not by 4, because the accrual period is still a quarter of the maturity — and only well beyond the accrual does the growth look quadratic. Mean reversion pulls the curve down, and increasingly so with maturity, since it is the accumulated decay that matters. At the front end the basis is under a basis point and can be ignored; at ten years it is not.

Drawn by `basis` and `ho_lee_basis`.

The picture explains why futures are used in the belly of the curve and swaps beyond it, as chapter 7 described without saying why. The convexity correction is negligible where the futures are liquid and becomes both large and model-dependent exactly where they stop being used.

8.5 Swaptions

Chapter 7 introduced the swap, its annuity

$$\mathrm{Ann}(t) = \Sigma_{i=1}^{N_r} \delta_i^r P(t, \tau_i^r),$$

and the par swap rate

$$S(t) = \frac{P(t, \tau_0) - P(t, \tau_{N_p}^p)}{\mathrm{Ann}(t)},$$

together with the observations that the annuity is a price and is strictly positive, so it is a numeraire, and the swap rate is a ratio of a price to it, so by chapter 6

$$S(t) = \mathbb{E}_t^{\mathrm{Ann}}[S(T)]$$

is a martingale in the annuity measure. We also had the value of a receiver swap struck at R as $\text{Ann}(0)(R - S(0))$, and of a payer swap as $\text{Ann}(0)(S(0) - R)$.

None of that needed a model. What follows does, and the annuity measure is what makes it manageable.

Definition 8.6 (Swaption)**.** A swaption is an option to enter a swap with swaption strike as the swap fixed rate. A call option is an option to enter a payer swap and a put option is an option to enter a receiver swap. The swap start date is the option exercise date.

Price of call swaption with exercise date T and strike K is therefore

$$\begin{aligned}\text{Swo}(0) =& P(0,T)\mathbb{E}^T[\max(\text{Ann}(T)(S(T)-K),0)]\\ =& P(0,T)\mathbb{E}^{\text{Ann}}[\max(\text{Ann}(T)(S(T)-K),0)\frac{\text{Ann}(0)P(T,T)}{\text{Ann}(T)P(0,T)}]\\ =& \text{Ann}(0)\mathbb{E}^{\text{Ann}}[\max(S(T)-K,0)]\\ =& \text{Ann}(0)\mathbb{E}^{\text{Ann}}[(S(T)-K)^+]\end{aligned}$$

The same swaption inside a one-factor model

That was the market's calculation. A model has to produce the same price from its own dynamics, and for a one-factor short rate model there is a decomposition that makes it a closed form rather than a numerical integral. It is Jamshidian's.

Start by seeing the swaption as an option on a bond. At expiry T the floating leg of the swap is worth par, so entering a receiver struck at K is worth

$$\sum_i c_i\, P(T,\tau_i) \;-\; 1, \qquad c_i = \delta_i K \;\;(i<N), \qquad c_N = 1 + \delta_N K,$$

the difference between a coupon bond and par. A receiver swaption is therefore a call on that coupon bond struck at 1, and a payer swaption a put.

Theorem 8.7 (Jamshidian's decomposition)**.** *Suppose every discount bond at T is a strictly decreasing function of a single state x. Let x^* be the state at which the coupon bond equals the strike, and write $P_i^* = P(T,\tau_i)$ evaluated at x^*. Then, for every x,*

$$\Big(\sum_i c_i P(T,\tau_i) - 1\Big)^+ \;=\; \sum_i c_i\big(P(T,\tau_i) - P_i^*\big)^+. \tag{8.10}$$

An option on a coupon bond is a portfolio of options on the zeros.

Proof. The coupon bond is a positive combination of decreasing functions and so is decreasing, which gives a unique x^* and, by definition, $\sum_i c_i P_i^* = 1$.

Now fix x. If $x < x^*$ then every bond exceeds its own value at x^*, since each is decreasing, so every bracket on the right is positive and the sum is $\sum_i c_i P(T,\tau_i) - \sum_i c_i P_i^*$, which is the left side. If $x > x^*$ then every bracket is negative, the right side is zero, and the coupon bond is below par so the left side is zero too. The two sides agree at every state, not merely in expectation.[3] □

[3] checked pointwise across the state.

Each term of (8.10) is an option on a single discount bond, and in a Gaussian model the bond is lognormal under the appropriate forward measure, so each is a Black formula. The swaption is a sum of them.

Remark (The hypothesis is the one factor, and it is also the defect)**.** Everything rested on the bonds being decreasing functions of *one* number, which is what makes x^* a point.

With two factors the exercise boundary is a curve rather than a point. Two states on it give the same coupon bond value and different individual bond prices, so no single set of strikes P_i^* can serve — whichever state supplied them, the other would decompose wrongly.[4] The decomposition is unavailable, and a multi-factor model prices its swaptions by simulation or by a grid.

This says something about the trick. It works because the model can only move the whole curve as a function of one number — the very property chapter 12 measures the cost of, and the reason a one-factor model misprices anything sensitive to the curve's shape. The closed form is not a reward for a good model. It is a consequence of the restriction that makes the model wrong.

Remark (Where this stops)**.** The swaption is priced and no smile has been needed: one volatility, one strike, Black's formula in the annuity measure.

References

- Kirikos, G., & Novak, D. (1997). Convexity Conundrums: Presenting a treatment of swap convexity in the Hull-White framework. *RISK-LONDON-RISK MAGAZINE LIMITED-*, 10, 60–61.

- Schlenkrich, S., & Ursachi, I. (2015). Multi-Curve Pricing of Non-Standard Tenor Vanilla Options. Available at SSRN 2695011.

- Heath, D., Jarrow, R., & Morton, A. (1992). Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation. *Econometrica*, 60(1), 77–105.

- Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5(2), 177–188.

- Hull, J., & White, A. (1990). Pricing interest-rate-derivative securities. *Review of Financial Studies*, 3(4), 573–592.

- Jamshidian, F. (1989). An exact bond option formula. *Journal of Finance*, 44(1), 205–209. The decomposition of Theorem 8.7. Not to be confused with the same author's 1997 paper, cited in chapter 13 for the spot Libor measure, which is an unrelated result.

[4] `measured`.