Chapter 7

Short Rate Models

→ B[3](Chapters 16–17), MR[20](Chapter 12), etc

7.1 Generalities

Short rate models are the classical interest rate models. As in the last section we fix a stochastic basis \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})\), where \(\mathbb{P}\) is considered as objective probability measure, and a \(d\)-dimensional Brownian motion \(W\).

We assume that

- the short rates follow an Itô process

\[
dr(t) = b(t) \, dt + \sigma(t) \, dW(t)
\]

determining the savings account \(B(t) = \exp\left(\int_0^t r(s) \, ds\right)\),

- all zero-coupon bond prices \((P(t, T))_{t \in [0,T]}\) are adapted processes (with \(P(T, T) = 1\) as usual),

- no-arbitrage: there exists an EMM \(\mathbb{Q}\) of the form \(d\mathbb{Q}/d\mathbb{P} = \mathcal{E}_\infty(\gamma \cdot W)\), such that

\[
\frac{P(t, T)}{B(t)}, \quad t \in [0, T],
\]

is a \(\mathbb{Q}\)-martingale for all \(T > 0\).

According to the last chapter, the existence of an ELMM for all \(T\)-bonds excludes arbitrage among every finite selection of zero-coupon bonds, say
CHAPTER 7. SHORT RATE MODELS

\( P(t, T_1), \ldots, P(t, T_n) \). To be more general one would have to consider strategies involving a continuum of bonds. This can be done (see [4] or the forthcoming book by R.Carmona and M.Tehranchi) but is beyond the scope of this course.

For convenience we require \( \mathbb{Q} \) to be an EMM (and not merely an ELMM) because then we have

\[
P(t, T) = \mathbb{E}_\mathbb{Q} \left[ e^{-\int_t^T r(s) \, ds} \mid \mathcal{F}_t \right] \tag{7.1}
\]

(compare this to the last section). Let \( \tilde{W} = W - \int \gamma \, dt \) denote the Girsanov transformed \( \mathbb{Q} \)-Brownian motion.

**Proposition 7.1.1.** Under the above assumptions, the process \( r \) satisfies under \( \mathbb{Q} \)

\[
dr(t) = (b(t) + \sigma(t) \cdot \gamma(t)) \, dt + \sigma(t) \, d\tilde{W}(t). \tag{7.2}
\]

Moreover, if the filtration \( (\mathcal{F}_t) \) is generated by the Brownian motion \( W \), for any \( T > 0 \) there exists a predictable \( \mathbb{R}^d \)-valued process \( \sigma^\gamma(t, T), t \in [0, T] \), such that

\[
\frac{dP(t, T)}{P(t, T)} = r(t) \, dt + \sigma^\gamma(t, T) \, d\tilde{W}(t) \tag{7.3}
\]

and hence

\[
\frac{P(t, T)}{B(t)} = P(0, T) \mathcal{E}_t \left( \sigma^\gamma \cdot \tilde{W} \right). \tag{7.4}
\]

**Proof.** Exercise (proceed as in the Completeness Lemma 6.3.2). \( \square \)

It follows from (7.3) that the \( T \)-bond price satisfies under the objective probability measure \( \mathbb{P} \)

\[
\frac{dP(t, T)}{P(t, T)} = (r(t) - \gamma(t) \cdot \sigma^\gamma(t, T)) \, dt + \sigma^\gamma \, dW(t).
\]

This illustrates again the role of the market price of risk, \(-\gamma\), as the excess of instantaneous return over \( r(t) \) in units of volatility.

In a general equilibrium framework, the market price of risk is given endogenously (as it is carried out in the seminal paper by Cox, Ingersoll and Ross (85) [8]). Since our arguments refer only to the absence of arbitrage between primary securities (bonds) and derivatives, we are unable to identify the market price of risk. In other words, we started by specifying the \( \mathbb{P} \)-dynamics of the short rates, and hence the savings account \( B(t) \). However,
the savings account alone cannot be used to replicate bond payoffs: the model is incomplete. According to the Completeness Theorem 6.3.3, this is also reflected by the non-uniqueness of the EMM (the market price of risk). A priori, $\mathbb{Q}$ can be any equivalent probability measure $\mathbb{Q} \sim \mathbb{P}$. A short rate model is not fully determined without the exogenous specification of the market price of risk.

It is custom (and we follow this tradition) to postulate the $\mathbb{Q}$-dynamics $(d\mathbb{Q}/d\mathbb{P} = \mathcal{E}_\infty(\gamma \cdot W))$ of $r$ which implies the $\mathbb{Q}$-dynamics of all bond prices by (7.1). All contingent claims can be priced by taking $\mathbb{Q}$-expectations of their discounted payoffs. The market price of risk (and hence the objective measure $\mathbb{P}$) can be inferred by statistical methods from historical observations of price movements.

### 7.2 Diffusion Short Rate Models

We fix a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{Q})$, where now $\mathbb{Q}$ is considered as martingale measure. We let $W$ denote a $d$-dimensional $(\mathbb{Q}, \mathcal{F}_t)$-Brownian motion.

Let $\mathcal{Z} \subset \mathbb{R}$ be a closed interval, and $b$ and $\sigma$ continuous functions on $\mathbb{R}_+ \times \mathcal{Z}$. We assume that for any $\rho \in \mathcal{Z}$ the stochastic differential equation (SDE)

$$dr(t) = b(t, r(t)) \, dt + \sigma(t, r(t)) \, dW(t)$$

(7.4) admits a unique $\mathcal{Z}$-valued solution $r = r^\rho$ with

$$r(t) = \rho + \int_0^t b(u, r(u)) \, du + \int_0^t \sigma(u, r(u)) \, dW(u)$$

and such that

$$\exp \left( - \int_0^T r(u) \, du \right) \in L^1(\mathbb{Q})$$

(7.5) for all $0 \leq t \leq T$. Notice that (7.5) is always satisfied if $\mathcal{Z} \subset \mathbb{R}_+$.

Sufficient for the existence and uniqueness is Lipschitz continuity of $b(t, r)$ and $\sigma(t, r)$ in $r$, uniformly in $t$. If $d = 1$ then Hölder continuity of order $1/2$ of $\sigma$ in $r$, uniformly in $t$, is enough. A good reference for SDEs is the book of Karatzas and Shreve [15] on Brownian motion and stochastic calculus.
Condition (7.5) allows us to define the \( T \)-bond prices

\[
P(t, T) = \mathbb{E}_Q \left[ \exp \left( - \int_t^T r(u) \, du \right) \mid \mathcal{F}_t \right].
\]

It turns out that \( P(t, T) \) can be written as a function of \( r(t), t \) and \( T \). This is a general property of certain functionals of Markov process, usually referred to as Feynman–Kac formula. In the following we write

\[
a(t, r) := \frac{\|\sigma(t, r)\|^2}{2}
\]

for the diffusion term of \( r(t) \).

**Lemma 7.2.1.** Let \( T > 0 \) and \( \Phi \) be a continuous function on \( Z \), and assume that \( F = F(t, r) \in C^{1,2}([0, T] \times Z) \) is a solution to the boundary value problem on \( [0, T] \times Z \)

\[
\left\{ \begin{array}{l}
\partial_t F(t, r) + b(t, r) \partial_r F(t, r) + a(t, r) \partial_r^2 F(t, r) - r F(t, r) = 0 \\
F(T, r) = \Phi(r).
\end{array} \right.
\]

Then

\[
M(t) = F(t, r(t)) e^{-\int_0^t r(u) \, du}, \quad t \in [0, T],
\]

is a local martingale. If in addition either

1. \( \partial_r F(t, r(t)) e^{-\int_0^t r(u) \, du} \sigma(t, r(t)) \in \mathcal{L}^2[0, T] \), or

2. \( M \) is uniformly bounded,

then \( M \) is a true martingale, and

\[
F(t, r(t)) = \mathbb{E}_Q \left[ \exp \left( - \int_t^T r(u) \, du \right) \Phi(r(T)) \mid \mathcal{F}_t \right], \quad t \leq T.
\]

**Proof.** We can apply Itô’s formula to \( M \) and obtain

\[
dM(t) = \left( \partial_t F(t, r(t)) + b(t, r(t)) \partial_r F(t, r(t)) \\
+ a(t, r) \partial_r^2 F(t, r(t)) - r(t) F(t, r(t)) \right) e^{-\int_0^t r(u) \, du} dt \\
+ \partial_r F(t, r(t)) e^{-\int_0^t r(u) \, du} \sigma(t, r(t)) dW(t) \\
= \partial_r F(t, r(t)) e^{-\int_0^t r(u) \, du} \sigma(t, r(t)) dW(t).
\]
Hence $M$ is a local martingale.

It is now clear that either Condition 1 or 2 imply that $M$ is a true martingale. Since

$$M(T) = \Phi(r(T)) e^{-\int_0^T r(u) \, du}$$

we get

$$F(t, r(t)) e^{-\int_0^t r(u) \, du} = M(t) = \mathbb{E}_Q \left[ \exp \left( -\int_0^T r(u) \, du \right) \Phi(r(T)) \mid \mathcal{F}_t \right].$$

Multiplying with $e^{\int_0^t r(u) \, du}$ yields the claim.

We call (7.6) the term structure equation for $\Phi$. Its solution $F$ gives the price of the $T$-claim $\Phi(r(T))$. In particular, for $\Phi \equiv 1$ we get the $T$-bond price $P(t, T)$ as a function of $t$, $r(t)$ (and $T$)

$$P(t, T) = F(t, r(t); T).$$

**Remark 7.2.2.** Strictly speaking, we have only shown that if a smooth solution $F$ of (7.6) exists and satisfies some additional properties (Condition 1 or 2) then the time $t$ price of the claim $\Phi(r(T))$ (which is the right hand side of (7.7)) equals $F(t, r(t))$. One can also show the converse that the expectation on the right hand side of (7.7) conditional on $r(t) = r$ can be written as $F(t, r)$ where $F$ solves the term structure equation (7.6) but usually only in a weak sense, which in particular means that $F$ may not be in $C^{1,2}([0, T] \times \mathbb{Z})$. This is general Markov theory and we will not prove this here.

In any case, we have found a pricing algorithm. Is it computationally efficient? Solving PDEs numerically in more than three dimensions causes difficulties. PDEs in less than three space dimensions are numerically feasible, and the dimension of $Z$ is one. The nuisance is that we have to solve a PDE for every single zero-coupon bond price function $F(\cdot, \cdot; T)$, $T > 0$. From that we might want to derive the yield or even forward curve. If we do not impose further structural assumptions we may run into regularity problems. Hence

short rate models that admit closed form solutions to the term structure equation (7.6), at least for $\Phi \equiv 1$, are favorable.
7.2.1 Examples

This is a (far from complete) list of the most popular short rate models. For all examples we have \( d = 1 \). If not otherwise stated, the parameters are real-valued.

1. Vasicek (1977): \( \mathcal{Z} = \mathbb{R} \),
   \[
   dr(t) = (b + \beta r(t)) \, dt + \sigma \, dW(t),
   \]

2. Cox–Ingersoll–Ross (CIR, 1985): \( \mathcal{Z} = \mathbb{R}_+ \), \( b \geq 0 \),
   \[
   dr(t) = (b + \beta r(t)) \, dt + \sigma \sqrt{r(t)} \, dW(t),
   \]

3. Dothan (1978): \( \mathcal{Z} = \mathbb{R}_+ \),
   \[
   dr(t) = \beta r(t) \, dt + \sigma r(t) \, dW(t),
   \]

4. Black–Derman–Toy (1990): \( \mathcal{Z} = \mathbb{R}_+ \),
   \[
   dr(t) = \beta(t) r(t) \, dt + \sigma(t) r(t) \, dW(t),
   \]

5. Black–Karasinski (1991): \( \mathcal{Z} = \mathbb{R}_+ \), \( \ell(t) = \log r(t) \),
   \[
   d\ell(t) = (b(t) + \beta(t) \ell(t)) \, dt + \sigma(t) \, dW(t),
   \]

6. Ho–Lee (1986): \( \mathcal{Z} = \mathbb{R} \),
   \[
   dr(t) = b(t) \, dt + \sigma \, dW(t),
   \]

7. Hull–White (extended Vasicek, 1990): \( \mathcal{Z} = \mathbb{R} \),
   \[
   dr(t) = (b(t) + \beta(t) r(t)) \, dt + \sigma(t) \, dW(t),
   \]

8. Hull–White (extended CIR, 1990): \( \mathcal{Z} = \mathbb{R}_+ \), \( b(t) \geq 0 \),
   \[
   dr(t) = (b(t) + \beta(t) r(t)) \, dt + \sigma(t) \sqrt{r(t)} \, dW(t).
   \]
7.3 Inverting the Yield Curve

Once the short rate model is chosen, the initial term structure

\[ T \mapsto P(0, T) = F(0, r(0); T) \]

and hence the initial yield and forward curve are fully specified by the term structure equation (7.6).

Conversely, one may want to invert the term structure equation (7.6) to match a given initial yield curve. Say we have chosen the Vasicek model. Then the implied \( T \)-bond price is a function of the current short rate level and the three model parameters \( b, \beta, \sigma \)

\[ P(0, T) = F(0, r(0); T, b, \beta, \sigma). \]

But \( F(0, r(0); T, b, \beta, \sigma) \) is just a parametrized curve family with three degrees of freedom. It turns out that it is often too restrictive and will provide a poor fit of the current data in terms of accuracy (least squares criterion).

Therefore the class of time-inhomogeneous short rate models (such as the Hull–White extensions) was introduced. By letting the parameters depend on time one gains infinite degree of freedom and hence a perfect fit of any given curve. Usually, the functions \( b(t) \) etc are fully determined by the empirical initial yield curve.

7.4 Affine Term Structures

Short rate models that admit closed form expressions for the implied bond prices \( F(t, r; T) \) are favorable.

The most tractable models are those where bond prices are of the form

\[ F(t, r; T) = \exp(-A(t, T) - B(t, T)r), \]

for some smooth functions \( A \) and \( B \). Such models are said to provide an affine term structure (ATS). Notice that \( F(T, r; T) = 1 \) implies

\[ A(T, T) = B(T, T) = 0. \]

The nice thing about ATS models is that they can be completely characterized.
Proposition 7.4.1. The short rate model \((7.4)\) provides an ATS only if its diffusion and drift terms are of the form
\[
a(t, r) = a(t) + \alpha(t)r \quad \text{and} \quad b(t, r) = b(t) + \beta(t)r, \tag{7.8}
\]
for some continuous functions \(a, \alpha, b, \beta\). The functions \(A\) and \(B\) in turn satisfy the system
\[
\begin{align*}
\partial_t A(t, T) &= a(t)B^2(t, T) - b(t)B(t, T), \quad A(T, T) = 0, \tag{7.9} \\
\partial_t B(t, T) &= \alpha(t)B^2(t, T) - \beta(t)B(t, T) - 1, \quad B(T, T) = 0. \tag{7.10}
\end{align*}
\]
Proof. We insert \(F(t, r; T) = \exp(-A(t, T) - B(t, T)r)\) in the term structure equation \((7.6)\) and obtain
\[
a(t, r)B^2(t, T) - b(t, r)B(t, T) = \partial_t A(t, T) + (\partial_t B(t, T) + 1)r. \tag{7.11}
\]
The functions \(B(t, \cdot)\) and \(B^2(t, \cdot)\) are linearly independent since otherwise \(B(t, \cdot) \equiv B(t, t) = 0\), which trivially would lead to be above results with \(a(t) = \alpha(t) \equiv 0\). Hence we can find \(T_1 > T_2 > t\) such that the matrix
\[
\begin{pmatrix}
B^2(t, T_1) & -B(t, T_1) \\
B^2(t, T_2) & -B(t, T_2)
\end{pmatrix}
\]
is invertible. Hence we can solve \((7.11)\) for \(a(t, r)\) and \(b(t, r)\), which yields \((7.8)\). Replace \(a(t, r)\) and \(b(t, r)\) by \((7.8)\), so the left hand side of \((7.11)\) reads
\[
a(t)B^2(t, T) - b(t)B(t, T) + (\alpha(t)B^2(t, T) - \beta(t)B(t, T)) r.
\]
Terms containing \(r\) must match. This proves the claim. \(\square\)

The functions \(a, \alpha, b, \beta\) in \((7.8)\) can be further specified. They have to be such that \(a(t, r) \geq 0\) and \(r(t)\) does not leave the state space \(\mathcal{Z}\). In fact, it can be shown that every ATS model can be transformed via affine transformation into one of the two cases

1. \(\mathcal{Z} = \mathbb{R}\): necessarily \(\alpha(t) = 0\) and \(a(t) \geq 0\), and \(b, \beta\) are arbitrary. This is the (Hull–White extension of the) Vasicek model.

2. \(\mathcal{Z} = \mathbb{R}_+\): necessarily \(a(t) = 0\), \(\alpha(t) \geq 0\) and \(b(t) \geq 0\) (otherwise the process would cross zero), and \(\beta\) is arbitrary. This is the (Hull–White extension of the) CIR model.

Looking at the list in Section 7.2.1 we see that all short rate models except the Dothan, Black–Derman–Toy and Black–Karasinski models have an ATS.
7.5 Some Standard Models

We discuss some of the most common short rate models.
→ B[3](Section 17.4), BM[6](Chapter 3)

7.5.1 Vasicek Model

The solution to \( dr = (b + \beta r) dt + \sigma dW \)
is explicitly given by (→ exercise)
\[
r(t) = r(0)e^{bt} + \frac{b}{\beta} \left( e^{bt} - 1 \right) + \sigma e^{bt} \int_0^t e^{-\beta s} dW(s).
\]

It follows that \( r(t) \) is a Gaussian process with mean
\[
E[r(t)] = r(0)e^{bt} + \frac{b}{\beta} \left( e^{bt} - 1 \right)
\]
and variance
\[
Var[r(t)] = \sigma^2 e^{2bt} \int_0^t e^{-2\beta s} ds = \frac{\sigma^2}{2\beta} \left( e^{2bt} - 1 \right).
\]

Hence
\[
\mathbb{Q}[r(t) < 0] > 0,
\]
which is not satisfactory (although this probability is usually very small).

Vasicek assumed the market price of risk to be constant, so that also the objective \( \mathbb{P} \)-dynamics of \( r(t) \) is of the above form.

If \( \beta < 0 \) then \( r(t) \) is mean-reverting with mean reversion level \( b/|\beta| \), see Figure 7.1, and \( r(t) \) converges to a Gaussian random variable with mean \( b/|\beta| \) and variance \( \sigma^2/(2|\beta|) \), for \( t \to \infty \).

Equations (7.9)–(7.10) become
\[
\frac{\partial}{\partial t} A(t, T) = \frac{\sigma^2}{2} B^2(t, T) - bB(t, T), \quad A(T, T) = 0,
\]
\[
\frac{\partial}{\partial t} B(t, T) = -\beta B(t, T) - 1, \quad B(T, T) = 0.
\]

The explicit solution is
\[
B(t, T) = \frac{1}{\beta} \left( e^{\beta(T-t)} - 1 \right)
\]
Figure 7.1: Vasicek short rate process for $\beta = -0.86$, $b/|\beta| = 0.09$ (mean reversion level), $\sigma = 0.0148$ and $r(0) = 0.08$.

and $A$ is given as ordinary integral

$$A(t, T) = A(T, T) - \int_t^T \partial_s A(s, T) \, ds$$

$$= -\frac{\sigma^2}{2} \int_t^T B^2(s, T) \, ds + b \int_t^T B(s, T) \, ds$$

$$= \frac{\sigma^2}{4\beta^3} \left( 4e^{\beta(T-t)} - e^{2\beta(T-t)} - 2\beta(T-t) - 3 \right) + \frac{b e^{\beta(T-t)} - 1 - \beta(T-t)}{\beta^2}.$$  

We recall that zero-coupon bond prices are given in closed form by

$$P(t, T) = \exp \left( -A(t, T) - B(t, T)r(t) \right).$$

It is possible to derive closed form expression also for bond options (see Section 7.6).

### 7.5.2 Cox–Ingersoll–Ross Model

It is worth to mention that, for $b \geq 0$,

$$dr(t) = (b + \beta r(t)) \, dt + \sigma \sqrt{r(t)} \, dW(t), \quad r(0) \geq 0,$$
7.5. **SOME STANDARD MODELS**

has a unique strong solution $r \geq 0$, for every $r(0) \geq 0$. This also holds when
the coefficients depend continuously on $t$, as it is the case for the Hull–White
extension. Even more, if $b \geq \sigma^2/2$ then $r > 0$ whenever $r(0) > 0$.

The ATS equation (7.10) now becomes non-linear

$$\partial_t B(t, T) = \frac{\sigma^2}{2} B^2(t, T) - \beta B(t, T) - 1, \quad B(T, T) = 0.$$  

This is called a Riccati equation. It is good news that the explicit solution is known

$$B(t, T) = \frac{2 \left( e^{\gamma(T-t)} - 1 \right)}{(\gamma - \beta) \left( e^{\gamma(T-t)} - 1 \right) + 2\gamma}$$

where $\gamma := \sqrt{\beta^2 + 2\sigma^2}$. Integration yields

$$A(t, T) = -\frac{2b}{\sigma^2} \log \left( \frac{2\gamma e^{(\gamma-\beta)(T-t)/2}}{(\gamma - \beta) \left( e^{\gamma(T-t)} - 1 \right) + 2\gamma} \right).$$

Hence also in the CIR model we have closed form expressions for the bond
prices. Moreover, it can be shown that also bond option prices are explicit(!)
Together with the fact that it yields positive interest rates, this is mainly the
reason why the CIR model is so popular.

### 7.5.3 Dothan Model

Dothan (78) starts from a drift-less geometric Brownian motion under the
objective probability measure $\mathbb{P}$

$$dr(t) = \sigma r(t) dW^\mathbb{P}(t).$$

The market price of risk is chosen to be constant, which yields

$$dr(t) = \beta r(t) dt + \sigma r(t) dW(t)$$

as $\mathbb{Q}$-dynamics. This is easily integrated

$$r(t) = r(s) \exp \left( (\beta - \sigma^2/2) (t - s) + \sigma (W(t) - W(s)) \right), \quad s \leq t.$$  

Thus the $\mathcal{F}_s$-conditional distribution of $r(t)$ is lognormal with mean and variance ($\to$ exercise)

$$\mathbb{E}[r(t) | \mathcal{F}_s] = r(s) e^{\beta(t-s)}$$

$$\text{Var}[r(t) | \mathcal{F}_s] = r^2(s) e^{2\beta(t-s)} \left( e^{\sigma^2(t-s)} - 1 \right).$$
The Dothan and all lognormal short rate models (Black–Derman–Toy and Black–Karasinski) yield positive interest rates. But no closed form expressions for bond prices or options are available (with one exception: Dothan admits an “semi-explicit” expression for the bond prices, see BM[6]).

A major drawback of lognormal models is the explosion of the bank account. Let $\Delta t$ be small, then

$$
\mathbb{E}[B(\Delta t)] = \mathbb{E}\left[\exp\left(\int_{0}^{\Delta t} r(s) \, ds\right)\right] \approx \mathbb{E}\left[\exp\left(\frac{r(0) + r(\Delta t)}{2}\Delta t\right)\right].
$$

We face an expectation of the type

$$
\mathbb{E}[\exp(\exp(Y))]
$$

where $Y$ is Gaussian distributed. But such an expectation is infinite. This means that in arbitrarily small time the bank account grows to infinity in average. Similarly, one shows that the price of a Eurodollar future is infinite for all lognormal models.

The idea of lognormal rates is taken up later by Sandmann and Sondermann (1997) and many others, which finally led to the so called market models with lognormal LIBOR or swap rates.

### 7.5.4 Ho–Lee Model

For the Ho–Lee model

$$
\, dr(t) = b(t) \, dt + \sigma \, dW(t)
$$

the ATS equations (7.9)–(7.10) become

$$
\partial_t A(t, T) = \frac{\sigma^2}{2} B^2(t, T) - b(t) B(t, T), \quad A(T, T) = 0,
$$

$$
\partial_t B(t, T) = -1, \quad B(T, T) = 0.
$$

Hence

$$
B(t, T) = T - t,
$$

$$
A(t, T) = -\frac{\sigma^2}{6} (T - t)^3 + \int_{t}^{T} b(s)(T - s) \, ds.
$$
The forward curve is thus
\[
f(t, T) = \partial_T A(t, T) + \partial_T B(t, T) + \sigma^2 \frac{(T - t)^2}{2} + \int_t^T b(s) \, ds + r(t).
\]
Let \( f^*(0, T) \) be the observed (estimated) initial forward curve. Then
\[
b(s) = \partial_s f^*(0, s) + \sigma^2 s.
\]
gives a perfect fit of \( f^*(0, T) \). Plugging this back into the ATS yields
\[
f(t, T) = f^*(0, T) - f^*(0, t) + \sigma^2 t (T - t) + r(t).
\]
We can also integrate this expression to get
\[
P(t, T) = e^{-\int_t^T f^*(0, s) \, ds + f^*(0, t) (T - t) - \frac{\sigma^2}{2} t (T - t)^2 - (T - t) r(t)}.
\]
It is interesting to see that
\[
r(t) = r(0) + \int_0^t b(s) \, ds + \sigma W(t) = f^*(0, t) + \sigma^2 t^2 + \sigma W(t).
\]
That is, \( r(t) \) fluctuates along the modified initial forward curve, and we have
\[
f^*(0, t) = \mathbb{E}[r(t)] - \frac{\sigma^2 t^2}{2}.
\]

### 7.5.5 Hull–White Model

The Hull–White (1990) extensions of Vasicek and CIR can be fitted to the initial yield and volatility curve. However, this flexibility has its price: the model cannot be handled analytically in general. We therefore restrict ourselves to the following extension of the Vasicek model that was analyzed by Hull and White 1994

\[
dr(t) = \left( b(t) + \beta r(t) \right) \, dt + \sigma \, dW(t).
\]
In this model we choose the constants \( \beta \) and \( \sigma \) to obtain a nice volatility structure whereas \( b(t) \) is chosen in order to match the initial yield curve.

Equation (7.10) for \( B(t, T) \) is just as in the Vasicek model
\[
\partial_s B(t, T) = -\beta B(t, T) - 1, \quad B(T, T) = 0
\]
with explicit solution

\[ B(t, T) = \frac{1}{\beta} \left( e^{\beta(T-t)} - 1 \right). \]

Equation (7.9) for \( A(t, T) \) now reads

\[ A(t, T) = -\sigma^2 + \int_t^T B^2(s, T) \, ds + \int_t^T b(s) B(s, T) \, ds \]

We consider the initial forward curve (notice that \( \partial_T B(s, T) = -\partial_s B(s, T) \))

\[ f^*(0, T) = \partial_T A(0, T) + \partial_T B(0, T) r(0) \]

\[ = \frac{\sigma^2}{2} \int_0^T \partial_s B^2(s, T) \, ds + \int_0^T b(s) \partial_T B(s, T) + \partial_T B(0, T) r(0) \]

\[ = -\frac{\sigma^2}{2\beta^2} \left( e^{\beta T} - 1 \right)^2 + \int_0^T b(s) e^{\beta(T-s)} \, ds + e^{\beta T} r(0). \]

The function \( \phi \) satisfies

\[ \partial_T \phi(T) = \beta \phi(T) + b(T), \quad \phi(0) = r(0). \]

It follows that

\[ b(T) = \partial_T \phi(T) - \beta \phi(T) \]

\[ = \partial_T \left( f^*(0, T) + g(T) \right) - \beta \left( f^*(0, T) + g(T) \right). \]

Plugging in and performing some calculations eventually yields

\[ f(t, T) = f^*(0, T) - e^{\beta(T-t)} f^*(0, t) - \frac{\sigma^2}{2\beta^2} \left( e^{\beta(T-t)} - 1 \right) \left( e^{\beta(T-t)} - e^{\beta(T+t)} \right) + e^{\beta(T-t)} r(t). \]

### 7.6 Option Pricing in Affine Models

We show how to price bond options in the affine framework. The discussion is informal, we do not worry about integrability conditions. The procedure has to be carried out rigorously from case to case.
Let \( r(t) \) be a diffusion short rate model with drift
\[
b(t) + \beta(t)r,
\]
diffusion term
\[
a(t) + \alpha(t)r
\]
and ATS
\[
P(t, T) = e^{-A(t,T)B(t,T)r(t)}.
\]
Let \( \lambda \in \mathbb{C} \), and \( \phi \) and \( \psi \) be given as solutions to
\[
\begin{align*}
\partial_t \phi(t, T, \lambda) &= a(t)\psi^2(t, T, \lambda) - b(t)\psi(t, T, \lambda) \\
\phi(T, T, \lambda) &= 0 \\
\partial_t \psi(t, T, \lambda) &= \alpha(t)\psi^2(t, T, \lambda) - \beta(t)\psi(t, T, \lambda) - 1 \\
\psi(T, T, \lambda) &= \lambda.
\end{align*}
\]
This looks much like the ATS equations (7.9)–(7.10), and indeed, by plugging the right hand side below in the term structure equation (7.6), one sees that
\[
\mathbb{E} \left[ e^{-\int_0^T r(s)\,ds}e^{-\lambda r(T)} \mid \mathcal{F}_t \right] = e^{-\phi(t,T,\lambda) - \psi(t,T,\lambda)r(t)}.
\]
In fact, we have
\[
\begin{align*}
\phi(t, T, 0) &= A(t, T) \quad \text{and} \quad \psi(t, T, 0) = B(t, T).
\end{align*}
\]
Now let \( t = 0 \) (for simplicity only). Since discounted zero-coupon bond prices are martingales we obtain for \( T \leq S \) (\( \rightarrow \) exercise)
\[
\begin{align*}
\mathbb{E} \left[ e^{-\int_0^T r(s)\,ds}e^{-\lambda r(T)} \mid \mathcal{F}_S \right] &= \mathbb{E} \left[ e^{-\int_0^T r(s)\,ds}e^{-A(T,S)B(T,S)r(T)}e^{-\lambda r(T)} \right] \\
&= e^{-A(T,S)}\mathbb{E} \left[ e^{-\int_0^T r(s)\,ds}e^{-\lambda + B(T,S))r(T)} \right] \\
&= e^{-A(T,S) - \phi(0,T,\lambda + B(T,S)) - \psi(0,T,\lambda + B(T,S))r(0)}.
\end{align*}
\]
But
\[
\frac{dQ^S}{dQ} = \frac{e^{-\int_0^S r(s)\,ds}}{P(0, S)}
\]
defines an equivalent probability measure \( Q^S \sim Q \) on \( \mathcal{F}_S \), the so called \( S \)-forward measure. Hence we have shown that the (extended) Laplace transform of \( r(T) \) with respect to \( Q^S \) is
\[
\mathbb{E}_{Q^S} [e^{-\lambda r(T)}] = e^{A(0,S) - A(T,S) - \phi(0,T,\lambda + B(T,S)) + (B(0,S) - \psi(0,T,\lambda + B(T,S)))r(0)}.
\]
 Chapters 7. Short Rate Models

By Laplace (or Fourier) inversion, one gets the distribution of $r(T)$ under $\mathbb{Q}^S$. In some cases (e.g. Vasicek or CIR) this distribution is explicitly known (e.g. Gaussian or chi-square). In general, this is done numerically.

We now consider a European call option on a $S$-bond with expiry date $T < S$ and strike price $K$. Its price today ($t = 0$) is

$$\pi = \mathbb{E} \left[ e^{-\int_0^T r(s) \, ds} \left( e^{-A(T,S) - B(T,S)r(T)} - K \right)^+ \right].$$

The payoff can be decomposed according to

$$\left( e^{-A(T,S) - B(T,S)r(T)} - K \right)^+ = e^{-A(T,S) - B(T,S)r(T)} 1_{\{r(T) \leq r^*\}} - K 1_{\{r(T) \leq r^*\}}$$

where

$$r^* = r^*(T, S, K) := -\frac{A(T, S) + \log K}{B(T, S)}.$$

Hence

$$\pi = \mathbb{E} \left[ e^{-\int_0^T r(s) \, ds} 1_{\{r(T) \leq r^*\}} \right] - K \mathbb{E} \left[ e^{-\int_0^T r(s) \, ds} 1_{\{r(T) \leq r^*\}} \right]$$

$$= P(0, S) \mathbb{Q}^S[r(T) \leq r^*] - K P(0, T) \mathbb{Q}^T[r(T) \leq r^*].$$

The pricing of the option boils down to the computation of the probability of the event $\{r(T) \leq r^*\}$ under the $S$- and $T$-forward measure.

7.6.1 Example: Vasicek Model ($a$, $b$, $\beta$ const, $\alpha = 0$).

We obtain ($\rightarrow$ exercise)

$$\pi = P(0, S) \Phi \left( \frac{r^* - \ell_1(T, S, r(0))}{\sqrt{\ell_2(T)}} \right) - K P(0, T) \Phi \left( \frac{r^* - \ell_1(T, T, r(0))}{\sqrt{\ell_2(T)}} \right)$$

where

$$\ell_1(T, S, r) := \frac{1}{\beta^2} \left( \beta \left( e^{\beta T} (b + \beta r) - b \right) - a \left( 2 - e^{\beta(S-T)} - 2e^{\beta T} + e^{\beta(S+T)} \right) \right)$$

$$\ell_2(T) := \frac{a}{\beta} \left( e^{2\beta T} - 1 \right)$$

and $\Phi(x)$ is the cumulative standard Gaussian distribution function.

A similar closed form expression is available for the price of a put option, and hence an explicit price formula for caps. For $\beta = -0.86$, $b/|\beta| = 0.09$
(mean reversion level), $\sigma = 0.0148$ and $r(0) = 0.08$, as in Figure 7.1, one gets the ATM cap prices and Black volatilities shown in Table 7.1 and Figure 7.2 (→ exercise). In contrast to Figure 2.1, the Vasicek model cannot produce humped volatility curves.

Table 7.1: Vasicek ATM cap prices and Black volatilities.

<table>
<thead>
<tr>
<th>Maturity</th>
<th>ATM prices</th>
<th>ATM vols</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.00215686</td>
<td>0.129734</td>
</tr>
<tr>
<td>2</td>
<td>0.00567477</td>
<td>0.106348</td>
</tr>
<tr>
<td>3</td>
<td>0.00907115</td>
<td>0.0915455</td>
</tr>
<tr>
<td>4</td>
<td>0.0121906</td>
<td>0.0815358</td>
</tr>
<tr>
<td>5</td>
<td>0.01503</td>
<td>0.0743607</td>
</tr>
<tr>
<td>6</td>
<td>0.017613</td>
<td>0.0689651</td>
</tr>
<tr>
<td>7</td>
<td>0.0199647</td>
<td>0.0647515</td>
</tr>
<tr>
<td>8</td>
<td>0.0221081</td>
<td>0.0613624</td>
</tr>
<tr>
<td>10</td>
<td>0.025847</td>
<td>0.0562337</td>
</tr>
<tr>
<td>12</td>
<td>0.028963</td>
<td>0.0525296</td>
</tr>
<tr>
<td>15</td>
<td>0.0326962</td>
<td>0.0485755</td>
</tr>
<tr>
<td>20</td>
<td>0.0370565</td>
<td>0.0443967</td>
</tr>
<tr>
<td>30</td>
<td>0.0416089</td>
<td>0.0402203</td>
</tr>
</tbody>
</table>

Figure 7.2: Vasicek ATM cap Black volatilities.
CHAPTER 7. SHORT RATE MODELS