CONTINUOUS TIME FINANCE
G63.2792, Spring 2004
Wednedays 7:10-9pm
WWH 1302
updated Feb. 17, 2004

Instructor: Robert V. Kohn. Office: 612 Warren Weaver Hall. Phone: 212-998-3217. Email: kohn@cims.nyu.edu. Web: www.math.nyu.edu/faculty/kohn. Office hours: Wednesday 4:30-5:30 and by appointment.

Teaching Assistant: Oana Fainarea. Office: 608. Phone: 212-998-3138.
Email: fainarea@cims.nyu.edu Office hours: Tuesday 4-5 and Tues 6-7.

Special Dates: First lecture Jan. 21. No class Feb. 11 (I'm out of town). No lecture March 17 (spring break). Last lecture April 28. Final exam: May 5.

Prerequisites: Derivative Securities and Stochastic Calculus, or equivalent.

Content: This is a "second course" in arbitrage-based pricing of derivative securities, continuing where the "first course" Derivative Securities left off. The first 1/3 of the semester will be devoted to the Black-Scholes model and its generalizations (equivalent martingale measures; the martingale representation theorem; the market price of risk; applications including change of numeraire and the analysis of quantos). The next 1/3 will be devoted to interest rate models (the Heath-Jarrow-Morton approach and its relation to short-rate models; applications including mortgage-backed securities). The last 1/3 will address more advanced topics, including the volatility smile/skew and approaches to accounting for it (underlyings with jumps, local volatility models, and stochastic volatility models).

Course requirements: There will be several homework sets, one every couple of weeks, probably 6 in all. Collaboration on homework is encouraged (homeworks are not exams) but registered students must write up and turn in their solutions individually. There will be one in-class final exam.

Lecture notes: Lecture notes and homework sets will be handed out, and also posted on my web-site as they become available.

Books: We will not follow any single textbook. However I strongly recommend

- M. Baxter and A. Rennie, *Financial calculus: an introduction to derivative pricing*, Cambridge University Press, 1996.

It correlates strongly with the material we'll cover in the first 2/3 of the semester. I'll also draw material from the following books, which will be on reserve in the CIMS library:

- M. Avellaneda and P. Laurence, *Quantitative modeling of derivative securities: from theory to practice*, Chapman and Hall, 2000
- D. Lamberton and B. Lapeyre, *Introduction to stochastic calculus applied to finance*, Chapman and Hall, 1996
- R. Korn and E. Korn, *Option pricing an portfolio optimization*, American Mathematical Society, 2001
- S. Neftci, *An introduction to the mathematics of financial derivatives*, second edition, Academic Press, 2000.
- D. Brigo and F. Mercurio, *Interest rate models: theory and practice*, Springer, 2001.
- R. Rebonato, *Interest-rate models: understanding, analysing, and using models for exotic interest-rate options*, second edition, John Wiley and Sons, 1998.

You'll notice there's nothing on the volatility smile/skew in this list. That's because I haven't yet decided what sources to use for the final 1/3 semester. The reserve list will be augmented as appropriate.

Continuous Time Finance Notes, Spring 2004 – Section 1. *1/21/04*
Notes by Robert V. Kohn, Courant Institute of Mathematical Sciences. For use in connection with the NYU course Continuous Time Finance.

This section discusses risk-neutral pricing in the continuous-time setting, for a market with just one source of randomness. In doing so we'll introduce and/or review some essential tools from stochastic calculus, especially the *martingale representation theorem* and *Girsanov's theorem*. For the most part I'm following Chapter 3 of Baxter & Rennie.

****

Let's start with some general orientation. Our focus is on the pricing of derivative securities. The most basic market model is the case of a lognormal asset with no dividend yield, when the interest rate is constant. Then the asset price S and the bond price B satisfy

$$dS = \mu S\,dt + \sigma S\,dw, \quad dB = rB\,dt$$

with μ, σ, and r all constant. We are also interested in more sophisticated models, such as:

(a) Asset dynamics of the form $dS = \mu(S,t)S\,dt + \sigma(S,t)S\,dw$ where $\mu(S,t)$ and $\sigma(S,t)$ are known functions. In this case S is still *Markovian* (the statistics of dS depend only on the present value of S). Almost everything we do in the lognormal case has an analogue here, except that explicit solution formulas are no longer so easy.

(b) Asset dynamics of the form $dS = \mu_t S\,dt + \sigma_t S\,dw$ where μ_t and σ_t are stochastic processes (depending only on information available by time t; more technically: they should be $\mathcal{F}_t$-measurable where $\mathcal{F}_t$ is the sigma-algebra associated to w). In this case S is *non-Markovian*. Typically μ and σ might be determined by separate SDE's. Stochastic volatility models are in this class (but they use two sources of randomness, i.e. the SDE for σ makes use of another, independent Brownian motion process; this makes the market incomplete.)

(c) Interest rate models where r is not constant, but rather random; for example, the spot rate r may be the solution of an SDE. (An interesting class of path-dependent options occurs in the modeling of mortgage-backed securities: the rate at which people refinance mortagages depends on the history of interest rates, not just the present spot rate.)

(d) Problems involving two or more sources of randomness, for example options whose payoff depends on more than one stock price [e.g. an option on a portfolio of stocks; or an option on the max or min of two stock prices]; quantos [options involving a random exchange rate and a random stock process]; and incomplete markets [e.g. stochastic volatility, where we can trade the stock but not the volatility].

My goal is to review background and to start relatively slowly. Therefore I'll focus in this section on the case when there is just one source of randomness (a single, scalar-valued Brownian motion).

Option pricing on a binomial tree is easy. The subjective probability is irrelevant, except that it determines the stock price tree. The risk-neutral probabilities (on the same tree!) are determined by

$$E_t^{\mathrm{RN}}[e^{-r\delta t}s_{t+\delta t}] = s_t.$$

If the interest rate r is constant, then the price of a contingent claim with payout $f(S_T)$ at time T is obtained by taking discounted expected payoff

$$V_0 = e^{-rT}E^{\mathrm{RN}}[f(S_T)].$$

The binomial tree method works even if interest rates are not constant. They can even be random (they can vary from node to node; all that matters is that the interest earned starting from any node be the same whether the stock goes up or down from that node). But if interest rates are not constant then we must discount correctly; and if they are random then the discounting must be inside the expectation:

$$V_0 = E^{\mathrm{RN}}\left[e^{-\int_0^T r(s)\,ds}f(S_T)\right]$$

If the payout is path-dependent (or if the subjective process is non-Markovian, i.e. its SDE has history-dependent coefficients) then we cannot use a recombining tree; in this case the binomial tree method is OK in concept but hard to use in practice.

Recall what lies behind the binomial-tree pricing formula: the binomial-tree market is complete, i.e. every contingent claim is replicatable. The prices given above are the initial cost of a replicating portfolio. So they are forced upon us (for a given tree) by the absence of arbitrage. (My Derivative Securities notes demonstrated this "by example," but see Chapter 2 of Baxter & Rennie for an honest yet elementary proof.)

The major flaw of this binomial-tree viewpoint: the market is not really a binomial tree. Moreover, a similar argument with a trinomial tree would not give unique prices (since a trinomial market is not complete), though it's easy to specify a trinomial model which gives lognormal dynamics in the limit $\delta t \to 0$. Similar but more serious: what to do when there are two stocks that move independently? (Using a binomial approximation for each gives an incomplete model, since each one-period subtree has four branches but just three tradeables.) Thus: if the market model we'd really like to use is formulated in continuous time, it would be much better to formulate the theory as well in continuous time (viewing the time-discrete models as numerical approximation schemes).

In the continuous-time setting, options can be priced using the Black-Scholes PDE. The solution of the Black-Scholes PDE tells us how to choose a replicating portfolio (namely: hold $\Delta = \frac{\partial V}{\partial S}(S_t, t)$ units of stock at time t); so the price we get using the Black-Scholes PDE is forced upon us by the absence of arbitrage, just as in the analysis of binomial trees. Let's be clear about what this means: a continuous-time *trading strategy* is a pair of stochastic processes, ϕ_t and ψ_t, each depending only on information available by time t; the associated (time-dependent) "portfolio" holds ϕ_t units of stock and ψ_t units of bond at time t. Its value at time t is thus

$$W_t = \phi_t S_t + \psi_t B_t.$$

(I call this W not V, because it represents the investor's wealth at time t as he pursues this trading strategy.) It is self-financing if

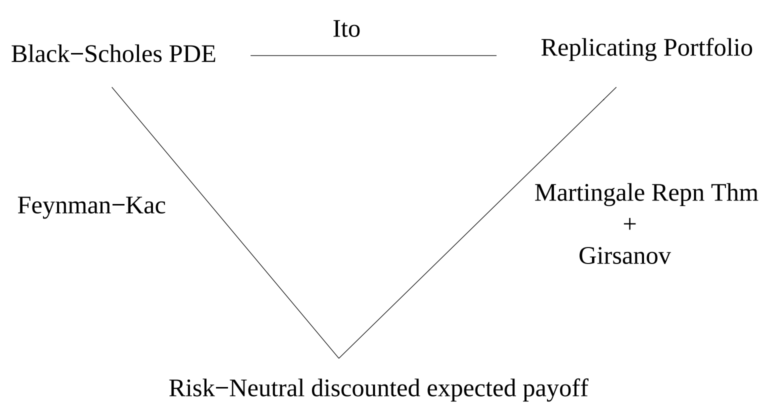
$$dW = \phi\, dS + \psi\, dB.$$

Our assertion is that if $V(S,t)$ solves the Black-Scholes PDE, and its terminal value $V(S,T)$ matches the option payoff, then the associated trading strategy

$$\phi_t = \frac{\partial V}{\partial S}(S_t, t), \quad \psi_t = (V(S_t,t) - \phi_t S_t)/B_t$$

is self-financing, and its value at time t is $V(S_t,t)$ for every t. (The assertion about its value is obvious; the fact that it's self-financing is a consequence of Ito's formula applied to $V(S_t,t)$, together with the Black-Scholes PDE and the fact that $dB = rBdt$.) Thus $V(S_0,0)$ is the time-0 value of a trading strategy that replicates the option payoff at time T.

We can recognize the option value as its "risk-neutral expected discounted payoff" by noticing that the Black-Scholes PDE is linked to a suitably-defined "risk neutral" diffusion by the Feynman-Kac formula. Indeed, suppose V solves the Black-Scholes PDE and S solves the stochastic differential equation $dS = rS\,dt + \sigma S\,dw$; for simplicity let's suppose the interest rate r is constant. Then $e^{r(T-t)}V(S_t,t)$ is a martingale, since its stochastic differential is $e^{r(T-t)}$ times $-rV\,dt + V_S dS + (1/2)V_{SS}\sigma^2 S^2 dt + V_t = \sigma S V_S\, dW$. So the expected value of $e^{r(T-t)}V(S_t,t)$ at time T (the risk-neutral expected payoff) equals the value of this expression at time 0, namely $e^{rT}V(S_0,0)$, as asserted. (The preceding argument works, with obvious modifications, even if r is time-dependent or S-dependent. In that setting the discount factor $\exp(r(T-t))$ must be replaced by $\exp(\int_t^T r\,ds)$.)



The major shortcoming of this Black-Scholes PDE viewpoint is this: it is limited to Markovian evolution laws ($dS = \mu S\,dt + \sigma S\,dw$, where μ and σ may depend on S and t, but not on the full history of events prior to time t) and path-independent options.

Beyond this rather practical restriction, there's also a conceptual issue. The probabilistic viewpoint is in many ways more natural than the PDE-based one; for example, it gives the analogue of our binomial tree discussion; it provides the basis for Monte Carlo simulation.

Moreover, for some purposes (e.g. understanding change of numeraire) the probabilistic framework is the only one that's really clear. So it's natural to seek a continuous-time understanding directly analogous to what we achieved with binomial trees. (It was also natural to postpone this till now, since such an understanding requires sufficient command of stochastic calculus.)

In summary: consider the triangle shown in the Figure, which shows three alternative ways of thinking about the pricing of derivative securities in a continuous-time framework. The course Derivative Securities focused mainly on two legs of the triangle (connecting the Black-Scholes PDE with the replicating portfolio, and with a suitable discounted expected risk-neutral payoff). Now we'll seek an equally clear understanding of the third leg.

****

The main tools we need from stochastic calculus are the *martingale representation theorem* and *Girsanov's theorem*. We now discuss the former, for diffusions involving a single source of randomness, i.e. diffusions of the form $dY = \mu_t\, dt + \sigma_t\, dw$, where μ_t and σ_t may be random (but depend only on information available at time t). Recall that an $\mathcal{F}_t$-adapted stochastic process M_t is called a *martingale* if

(a) it is integrable, i.e. $E[|Y_t|] < \infty$ for all t, and

(b) its conditional expectations satisfy $E[M_t|\mathcal{F}_s] = M_s$ for all $s < t$.

The expectation is of course with respect to a measure on path space. When M solves an SDE of the form $dM = \mu\, dt + \sigma\, dw$ where w is Brownian motion then the expectation we have in mind is the one associated with the underlying Brownian motion (and the sigma-algebra $\mathcal{F}_t$ is the one generated by this Brownian motion). Soon we'll discuss Girsanov's theorem, which deals with changing the measure (and, as a result, changing the stochastic differential equation). Then we must be careful which measure we're using – calling M a P-martingale if conditions (a) and (b) hold when the expectation is taken using measure P. But for now let's suppose there's only one measure under discussion, so we need not specify it.

The integrability condition is important to a probabilist (it's easy to construct processes satisfying (b) but not (a) – these are called "local martingales" – and many of the theorems we want to use have counterexamples if M is just a local martingale). However the processes of interest in finance are always integrable, and my purpose is to get the main ideas without unnecessary technicality, so I won't fuss much over integrability.

There are two different ways of constructing martingales defined for $0 \leq t \leq T$. One is to take an $\mathcal{F}_T$-measurable random variable X (e.g. an option payoff – which is now permitted to be path-dependent, i.e. to depend on all information up to time T) and take its conditional expectations:

$$M_t = E[X|\mathcal{F}_t]. \tag{1}$$

This is a martingale due to the "tower property" of conditional expectations. The second method of constructing martingales is to solve a stochastic differential equation of the form

$$dM = \phi_t\, dw \tag{2}$$

where ϕ_t is $\mathcal{F}_t$-adapted (i.e. it depends only on information available by time t).

Of course in considering (b), we need some condition on ϕ to be sure M is integrable. An obvious sufficient condition is that $E[\int_0^T \psi_s^2\, ds] < \infty$, since this assures us that $E[M^2(t)] < \infty$. In finance (2) typically takes the form $dM = \sigma_t M\, dw$. In this case one can show that the "Novikov condition" $E\left[\exp\left(\frac{1}{2}\int_0^T \sigma_s^2\, ds\right)\right] < \infty$ is sufficient to assure integrability. This fact is not easy to prove with full rigor; but it is certainly intuitive, since the SDE $dM = \sigma_t M\, dw$ has the explicit solution

$$M_t = M_0 e^{\int_0^t \sigma_s\, dw - \frac{1}{2}\int_0^t \sigma_s^2\, ds}$$

(To see this: apply Ito's formula to the right hand side to verify that the process defined by this formula satisfies $dM = \sigma_t M\, dw$.)

In its simplest form, the martingale representation theorem says that *any martingale can be expressed in the form (2).* Moreover the associated ϕ_t is unique. Thus, for example, a martingale created via conditional expectations as in (1) can alternatively be described by an SDE.

We need a slightly more sophisticated version of the martingale representation theorem. Rather than expressing M in terms of the Brownian motion w, we'll need to represent it in terms of *another* martingale, say N. The feasibility of doing this is obvious: if N and M are both martingales, then our simplest form of the martingale representation theorem says they can both be represented in the form (2):

$$dM = \phi^M\, dw \quad \text{and} \quad dN = \phi^N\, dw.$$

So eliminating dw gives a representation of the form

$$dM = \psi_t\, dN \tag{3}$$

with $\psi = \phi^M/\phi^N$. This representation is the more sophisticated version we wanted. Of course we cannot divide by 0: we assumed, in deriving (3), that $\phi^N \neq 0$ with probability 1. Like our simpler version (2), the representation (3) is unique, i.e. the density ψ is uniquely determined by M and N.

****

The martingale representation theorem is all we need to connect risk-neutral expectations with self-financing portfolios. This connection works even for path-dependent options, and even for stock processes whose drift and volatility are random (but $\mathcal{F}_t$-measurable). It also permits the risk-free rate to be random. To demonstrate the strength of the method we work at this level of generality. Thus our market has a stock whose price satisfies

$$dS = \mu_t S\, dt + \sigma_t S\, dw, \quad S(0) = S_0$$

and a bond whose value satisfies

$$dB = r_t B\, dt, \quad B(0) = 1$$

where μ_t, σ_t, and r_t are all $\mathcal{F}_t$-measurable. Our goal is to price an option with payoff X, where X is any $\mathcal{F}_T$-measurable random variable.

To get started, we need one more assumption. We suppose there is a measure Q on path space such that the ratio S_t/B_t is a Q-martingale. This is the "risk-neutral measure." The whole point of Girsanov's theorem is to prove existence of such a Q (and to describe it). For now, we simply assume Q exists.

Here's the story: given such a Q, we'll show that every payoff X is replicatable. Moreover if V_t is the value at time t of the replicating portfolio then V_t is characterized by

$$V_t/B_t = E_Q[X/B_T|\mathcal{F}_t]. \tag{4}$$

In particular, the initial cost of the replicating portfolio is the expected discounted payoff $V_0 = E_Q[X/B_T]$. This is therefore the value of the option.

The proof is surprisingly easy. Let V_t be the process defined by (4). Then V_t/B_t is a Q-martingale, by (1) applied to X/B_T. Also S_t/B_t is a Q-martingale, by hypothesis. Therefore by the martingale representation theorem there is a process ϕ_t such that

$$d(V/B) = \phi_t\, d(S/B).$$

We shall show that the trading strategy which holds ϕ_t units of stock and $\psi_t = (V_t - \phi_t S_t)/B_t$ units of bond at time t is self-financing and replicates the option.

Easiest first. It replicates the option because $V_t = \phi_t S_t + \psi_t B_t$ (by the choice of ψ) and $V_T = X$ (by (4), noting that X/B_T is $\mathcal{F}_T$ measurable so taking its conditional expectation relative to $\mathcal{F}_T$ doesn't change it). Thus $V_T = X$, which is exactly what we mean when we say it replicates the option.

OK, now a little work. We must show that the proposed trading strategy is self-financing, i.e. that $dV = \phi_t\, dS + \psi_t\, dB$. Recall from Ito's lemma that when X and Y are diffusions, $d(XY) = X\,dY + Y\,dX + dX\,dY$. Also note that if the SDE for one of the two processes has no "dw" term then $dX\,dY = 0$ and the formula becomes $d(XY) = X\,dY + Y\,dX$. We apply this twice: since $V = (V/B)B$ and $d(V/B) = \phi\, d(S/B)$ we have

$$dV = d((V/B)B) = B\phi\, d(S/B) + (V/B)\, dB; \tag{5}$$

and since $S = (S/B)B$ we have

$$\phi\, dS = B\phi\, d(S/B) + \phi(S/B)\, dB. \tag{6}$$

Now, from the definition of ψ we have

$$\psi\, dB = [(V/B) - \phi(S/B)]\, dB. \tag{7}$$

Adding the last two equations gives

$$\phi\, dS + \psi\, dB = B\phi\, d(S/B) + (V/B)\, dB.$$

Comparing this with (5), we conclude that the portfolio is self-financing. The proof is now complete.

Question: where did we use the hypothesis that $dB = r_t B dt$? Answer: when we applied Ito's formula in (5) and (6). Remarkably, however, this hypothesis was not necessary! The assertion remains true (though the proof needs some adjustment) even if B is a volatile diffusion process. Thus B need not be a the value of a bond! We'll return to this when we discuss *change of numeraire*. But let's check now that the preceding argument works even if B is a volatile diffusion (i.e. if it solves an SDE with a nonzero dw term). In this case (5) must be replaced by

$$dV = d((V/B)B) = B\phi\, d(S/B) + (V/B)\, dB + d(V/B)\, dB$$

and (6) by

$$\phi\, dS = B\phi\, d(S/B) + \phi(S/B)\, dB + \phi\, d(S/B)\, dB.$$

Equation (7) remains unchanged:

$$\psi\, dB = [(V/B) - \phi(S/B)]\, dB.$$

Adding the last two equations, and remembering that $\phi\, d(S/B) = d(V/B)$, we conclude as before that $dV = \phi\, dS + \psi\, dB$, so the portfolio is still self-financing.

Our argument tells us how to replicate (and therefore price) any option, assuming only the existence of a "risk-neutral measure" Q with respect to which S/B is a martingale. We have in effect shown that our simple market (with one source of randomness) is complete. It also follows that the value P_t of *any* tradeable must be such that P_t/B_t is a Q-martingale (for the same measure Q). Indeed, we can replicate the payoff P_T starting at time t at cost $B_t E_Q[P_T/B_T|\mathcal{F}_t]$. To avoid arbitrage, this had better be the market price of the same payoff, namely P_t. Thus $P_t/B_t = E_Q[P_T/B_T|\mathcal{F}_t]$, as asserted.

****

Now what about the existence of Q? This is the point of Girsanov's theorem. Restricted to the case of a single Brownian motion, it says the following. Consider a measure P on paths, and suppose w is a P-Brownian motion (i.e. the induced measure on $w_t - w_s$ is Gaussian for each $s < t$, with mean 0 and variance $t - s$, and $w_t - w_s$ is independent of $w_s - w_r$ for $r < s < t$.). Consider in addition an adapted process γ_t, and the associated martingale

$$M_t = e^{-\int_0^t \gamma_s\, dw - \frac{1}{2}\int_0^t \gamma_s^2\, ds}.$$

(As noted earlier in these notes, γ_t must satisfy some integrability condition to be sure M_t is a martingale; the Novikov condition $E_P\left[\exp\left(\frac{1}{2}\int_0^T \gamma_s^2\, ds\right)\right] < \infty$ is sufficient.) Finally, suppose we're only interested in behavior up to time T. Then the measure Q defined by

$$E_Q[X] = E_P[M_T X] \quad \text{for every } \mathcal{F}_T\text{-measurable random variable } X$$

(in other words $dQ/dP = M_T$) has the property that

$$\tilde{w}_t = w_t + \int_0^t \gamma_s\, ds \quad \text{is a } Q\text{-Brownian motion.} \tag{8}$$

Moreover, we have the following formula for conditional probabilities taken with respect to Q:

$$E_Q[X_t|\mathcal{F}_s] = E_P[(M_t/M_s)X_t|\mathcal{F}_s] = M_s^{-1}E_P[(M_tX_t|\mathcal{F}_s] \tag{9}$$

whenever X is $\mathcal{F}_t$-measurable and $s \leq t \leq T$.

It's easy to state the theorem, but not so easy to get one's head around it. The discussion in Baxter & Rennie is very good: they explain in particular how this generalizes our familiar binomial-tree change of measure from the "subjective probabilities" to the "risk-neutral" ones. I also recommend the short Section 3.1 of Steele (on importance sampling, for Gaussian random variables), to gain intuition about change-of-measure in the more elementary setting of a single Gaussian random variable. Baxter & Rennie give an honest proof of (8) for the special case γ_t = constant in Section 3.4 (note that the solution of Problem 3.9 is in the back of the book). Avellaneda & Laurence give essentially the same argument in their Section 9.3, but they present it in greater generality (for nonconstant γ_t). Moreover Avellaneda & Laurence explain the formula (9) for conditional probabilities in the appendix to their Chapter 9. I recommend reading all these sources to gain an understanding of the theorem and why it's true. Here I'll not repeat that material; rather, I simply make some comments:

(a) In finance, the main use of Girsanov's theorem is the following: suppose S solves the SDE $dS = \alpha_t\,dt + \beta_t\,dw$ where w is a P-Brownian motion. Then it also solves the SDE $dS = (\alpha_t - \beta_t\gamma_t)\,dt + \beta_t\,d\tilde{w}$ where $d\tilde{w}$ is a Q-Brownian motion. Indeed, (8) can be restated as $d\tilde{w} = dw + \gamma_t\,dt$, and with this substitution the two SDE's are identical.

(b) Why is there a measure Q such that S/B is a Q-martingale? Well, if S and B solve SDE's then so does S/B (by Ito's lemma). Suppose this SDE has the form $d(S/B) = \alpha_t\,dt + \beta_t\,dw$. Then we can eliminate the drift by choosing $\gamma_t = \alpha_t/\beta_t$. In other words, applying Girsanov with this choice of γ_t gives $d(S/B) = \beta_t\,d\tilde{w}$ with $\tilde{w}$ a Q-Brownian motion. Thus S/B is a Q-martingale.

(c) Our statement of Girsanov's theorem selected (arbitrarily) an initial time 0. This choice should be irrelevant, since Brownian motion is a Markov process (so it can be viewed as starting at any time s with initial value w_s). Also, our statement selected (arbitrarily) a final time T, which should also be irrelevant. These choices are indeed unimportant, as a consequence of (i) the fact that M_t is a P-martingale, and (ii) relation (9), combined with the observation that

$$M(t)/M(s) = e^{-\int_s^t \gamma_\rho\,dw - \frac{1}{2}\int_s^t \gamma_\rho^2\,d\rho}.$$

Bottom line: we have shown in great generality that there is a "risk-neutral" measure Q with respect to which S/B is a martingale, and using it we have shown how to replicate (and therefore price) an arbitrary (even path-dependent) contingent claim.

What good is this? First of all it tells us that Monte-Carlo methods can be used even for path-dependent options and non-Markovian diffusions (whose SDE's have coefficients depending on prior history). But besides that, it is even useful in relatively simple settings (e.g. constant drift, volatility, and interest rates) for mastering concepts like change-of-numeraire. We'll discuss such applications next. (If you want to read ahead, look at Chapter 4 of Baxter & Rennie.)

Continuous Time Finance Notes, Spring 2004 – Section 2, *Jan. 28, 2004*
Notes by Robert V. Kohn, Courant Institute of Mathematical Sciences. For use in connection with the NYU course Continuous Time Finance.

In Section 1 we discussed how Girsanov's theorem and the martingale representation theorem tell us how to price and hedge options. This short section makes that discussion concrete by applying it to (a) options on a stock which pays dividends, and (b) options on foreign currency. We close with a brief discussion of Siegel's paradox. For topics (a) and (b) see also Baxter and Rennie's sections 4.1 and 4.2 – which are parallel to my discussion, but different enough to be well worth reading and comparing to what's here. For a discussion of Siegel's paradox and some related topics, see chapter 1 of the delightful book *Puzzles of Finance: Six Practical Problems and their Remarkable Solutions* by Mark Kritzman (J. Wiley & Sons, 2000, available as an inexpensive paperback).

****

Options on a stock with dividend yield. You probably already know from Derivative Securities how to price an option on a stock with continuous dividend yield q. If the stock is lognormal with volatility σ and the risk-free rate is (constant) r then the "risk-neutral process" is $dS = (r-q)S\,dt + \sigma S\,dw$, and the time-0 value of an option with payoff $f(S_T)$ and maturity T is $e^{-rT}E_{\mathrm{RN}}[f(S_T)]$. From the SDE we get $dE_{\mathrm{RN}}[S]/dt = (r-q)E_{\mathrm{RN}}[S]$, so $E_{\mathrm{RN}}[S](T) = e^{(r-q)T}S_0$. This is the *forward price*, i.e. the unique choice of k such that a forward with strike k and maturity T has initial value 0. (Proof: apply the pricing formula to $f(S_T) = S_T - k$.) When the option is a call, we get an explicit valuation formula using the fact that if X is lognormal with mean $E[X] = F$ and volatility s (defined as the standard deviation of $\log X$), then

$$E[(X-K)_+] = FN(d_1) - KN(d_2) \tag{1}$$

with

$$d_1 = \frac{\ln(F/K) + s^2/2}{s}, \quad d_2 = \frac{\ln(F/K) - s^2/2}{s}.$$

There is of course a similar formula for a call (easily deduced by put-call parity).

What does it mean that the "risk-neutral process is $dS = (r-q)S\,dt + \sigma S\,dw$?" What happens when the dividends are paid at discrete times? We can clarify these points by using the framework of Section 1. Remember the main points: (a) there is a unique "equivalent martingale measure" Q, obtained by an application of Girsanov's theorem; and (b) this Q is characterized by the property that the value V_t of any tradeable asset satisfies $V_t/B_t = E[V_T/B_T \,|\, \mathcal{F}_t]$ where B_t is the value of a risk-free money-market account, i.e. it solves $dB = rB\,dt$ with $B(0) = 1$. Put differently: V/B is a Q-martingale.

Suppose the subjective stock process is $dS = \mu S\,dt + \sigma S\,dw$, and it pays dividends at constant rate q. If we hold the stock, we receive the dividend yield as well. So the stock itself is not a tradeable asset; but the stock *with dividends reinvested* is a tradeable. The value of this asset is $X_t = S_t D_t$ where $dD = qD\,dt$. Since the SDE for D has no dw term, Ito's formula becomes the ordinary Leibniz rule $dX = SdD + DdS$. Thus, after a bit of manipulation, the SDE for X is

$$dX = (\mu + q)X\,dt + \sigma X\,dw \tag{2}$$

A similar calculation gives the SDE satisfied by its discounted value $Y = X/B$:

$$dY = (\mu + q - r)Y\,dt + \sigma Y\,dw.$$

The preceding SDE's all use the original w, which is a Brownian motion in the original, subjective probability.

Now we apply Girsanov's theorem. Remember how it works: Girsanov changes the drift but not the volatility. The risk-neutral measure Q has the property that Y is a Q-martingale. So the SDE for Y must be

$$dY = \sigma Y\,d\tilde{w}$$

where $\tilde{w}$ is a Q-Brownian motion. To get the SDE for S, we observe that recall that $S_t = Y_t B_t / D_t$ so

$$dS = (r - q)S\,dt + \sigma S\,d\tilde{w}.$$

This is the "risk-neutral process" alluded to above, and the "risk-neutral" expectations E_{RN} are simply to be taken using the measure Q.

It was not important for the preceding calculations that the interest rate (or the dividend rate) be constant. For example, q could be a function of S. If r is function of time however the option value is not $e^{-rT}E_{\mathrm{RN}}[f(S_T)]$ but $\exp\left(-\int_0^T r(s)\,ds\right)E_{\mathrm{RN}}[f(S_T)]$. (If r is random the discount factor must be brought inside the expectation; but normally the randomness of r would be independent of, or at least not fully correlated with, that of S; we'll discuss problems with more than one source of randomness soon.) Conclusion: the equation for the risk-neutral process is actually quite general; the specialization to constant volatility, constant risk-free rate, and constant dividend rate is needed only if we desire an explicit formula for the value of the option.

What if the dividends are paid at discrete times? Suppose at each time $T_1, T_2, \ldots$ the stock pays a dividend equal to fraction q of the stock price. Then S_t is lognormal between the dividend dates (say, $dS = \mu S\,dt + \sigma S\,dw$) but at each dividend date S must decrease by exactly the value of the dividend, i.e. at T_j it jumps from S_{T_j} to $(1-q)S_{T_j}$. But as above, the stock itself is not a tradeable; rather, the convenient tradeable is the stock with all dividends reinvested. Its value X now satisfies $dX = \mu X\,dt + \sigma X\,dw$. The risk-neutral Q is defined by the property that $Y = X/B$ is a martingale; a calculation parallel to the one done above shows that

$$dX = r\,X dt + \sigma X\,d\tilde{w}$$

where $\tilde{w}$ is a Q-martingale. Now, if N dividend dates occur between time 0 and time T, then the final-time stock price is $S_T = (1-q)^N X_T$. Thus under the risk-neutral probability S_T is lognormal, with mean $(1-q)^N e^{rT} S_0$ and volatility $\sigma\sqrt{T}$. We can again get explicit prices for calls by using the formula (1).

****

Options on an exchange rate. You probably also learned in Derivative Securities how to price an option on a foreign currency rate. Here's the basic setup: suppose the US dollar risk-free rate is r; the British pound risk-free rate is q, and the dollar value of one pound is lognormal, i.e. the exchange rate C_t with units dollars/pound satisfies $dC = \mu C\, dt + \sigma C\, dw$. To a dollar investor the pound looks like a "stock with continuous dividend yield q." So no further work is needed: from our calculation with continuous dividend yield, the risk-free process is

$$dC = (r - q)C\, dt + \sigma C\, d\tilde{w}. \tag{3}$$

Here $\tilde{w}$ is a Q-Brownian motion, where Q is the dollar investor's risk-neutral measure. A typical application would be to price an option giving its holder the right to buy a British pound at time T for K dollars. Its payoff (to the dollar investor) is $(C_T - K)_+$. Since C_T is lognormal under the risk-neutral probability, with mean $e^{(r-q)T}C_0$ and volatility $\sigma\sqrt{T}$, we get an explicit price from (1).

But now a new issue arises. What about an investor who thinks in pounds? He can do a similar calculation of course. Will his calculation be consistent with that of the dollar investor? We expect so, since an inconsistency would lead to arbitrage. But let's check the consistency explicitly.

We start by spelling out the pound investor's calculation. His money-market discount factor is D_t not B_t (here $dD = qD\, dt$ and $dB = rB\, dt$). His exchange rate is $1/C$ not C. By Ito, if $dC = \mu C\, dt + \sigma C\, dw$ then

$$d(1/C) = -C^{-2}dC + C^{-3}dCdC = (-\mu + \sigma^2)(1/C)\, dt - \sigma(1/C)\, dw.$$

What is the pound investor's risk-neutral measure? Well, it isn't Q! We find it by considering the obvious stochastic tradeable: the value in pounds of the dollar money-market account. Arguing as for (2), its value $\overline{X} = BC^{-1}$ satisfies

$$d\overline{X} = (r - \mu + \sigma^2)\overline{X}\, dt - \sigma\overline{X}\, dw$$

and the associated discounted value $\overline{Y} = \overline{X}/D$ satisfies

$$d\overline{Y} = (r - q - \mu + \sigma^2)\overline{Y}\, dt - \sigma\overline{Y}\, dw.$$

Thus pound-investor's risk-neutral measure $\overline{Q}$ has the property that

$$d\overline{Y} = -\sigma\overline{Y}\, d\overline{w}$$

where $\overline{w}$ is a $\overline{Q}$-Brownian motion. In terms of $\overline{w}$ the SDE for $C^{-1} = D\overline{Y}/B$ is

$$d(1/C) = (q - r)(1/C)\, dt - \sigma(1/C)\, d\overline{w} \tag{4}$$

and by Ito's lemma the SDE for C is therefore

$$dC = (r - q + \sigma^2)C\, dt + \sigma C\, d\overline{w}.$$

Comparing this with (3) we see that

$$d\tilde{w} = d\overline{w} + \sigma \, dt.$$

Recall that $\overline{w}$ is a $\overline{Q}$-Brownian motion, while $\tilde{w}$ is a Q-Brownian motion. Thus $\overline{Q}$ is indeed different from Q. From Girsanov's theorem (with $\overline{Q}$ and $\overline{w}$ in place of P and w) we know the relation between them:

$$\frac{dQ}{d\overline{Q}} = e^{-\sigma\overline{w}(T) - \frac{1}{2}\sigma^2 T} \tag{5}$$

with the usual convention that $\overline{w}(0) = \tilde{w}(0) = 0$.

OK, now we're ready to check for consistency. Consider an option whose payoff in dollars is f. (For example: an option to buy a pound for K dollars at time T, whose payoff is $f = (C_T - K)_+$.) Note that its payoff in pounds is $C_T^{-1} f$. To the dollar investor its value at time 0 is

$$e^{-rT} E_Q[f] \quad \text{dollars.}$$

To the pound investor its value at time 0 is

$$e^{-qT} E_{\overline{Q}}[C_T^{-1} f] \quad \text{pounds.}$$

To show the prices are consistent, we must verify that

$$C_0 e^{-qT} E_{\overline{Q}}[C_T^{-1} f] = e^{-rT} E_Q[f].$$

Since

$$E_Q[f] = E_{\overline{Q}}\left[\frac{dQ}{d\overline{Q}} f\right]$$

it will be sufficient to show that

$$C_0 e^{-qT} C_T^{-1} = e^{-rT} \frac{dQ}{d\overline{Q}}. \tag{6}$$

But solving the SDE (4) we have

$$C_T^{-1} = C_0^{-1} e^{(q - r - \frac{1}{2}\sigma^2)T - \sigma\overline{w}(T)},$$

which makes the left hand side of (6) explicit in terms of $\overline{w}(T)$. Equation (5) makes the right hand side explicit. Examining the resulting expressions we find that the desired consistency condition (6) is indeed valid.

****

It is worth digressing to mention *Siegel's paradox*. Recall that under the dollar investor's risk-neutral process

$$dC = (r - q)C \, dt + \sigma C \, d\tilde{w},$$

and that the forward rate (in dollars per pound) is the expected risk-neutral mean

$$F = E_Q[C_T] = e^{(r-q)T} C_0.$$

The pound investor has a similar calculation: under his risk-neutral process

$$d(1/C) = (q - r)(1/C)\,dt - \sigma(1/C)\,d\overline{w}$$

so his forward rate (in pounds per dollar) is

$$\overline{F} = E_{\overline{Q}}[C_T^{-1}] = e^{(q-r)T}C_0^{-1}.$$

They are consistent: $\overline{F} = 1/F$, as must be the case.

This might at first seem surprising, since the function $1/x$ is convex, which implies by Jensen's inequality that

$$(E[C])^{-1} < E[C^{-1}]$$

when both means are taken with respect to the same probability distribution and C is not deterministic. But the means defining F and $\overline{F}$ involve two different probability distributions: Q and $\overline{Q}$.

What about the subjective probability distribution P. Is it plausible that $E_P[C_T] = e^{(r-q)T}C_0$? By now we know the answer: it's possible, though unlikely, since this holds only for a particular choice of the drift μ. What would the financial implication be? Well, consider the following investment strategy: (a) borrow X dollars at time 0, at interest rate r; (b) convert them to X/C_0 pounds; (c) invest these pounds at interest rate q; (d) liquidate the position at time T. The amount the investor realizes by liquidation is $X[(C_T/C_0)e^{qT} - e^{rT}]$. So his expected outcome is 0 exactly if $E_P[C_T] = e^{(r-q)T}C_0$, in other words *if the forward exchange rate is an unbiased estimate of the future spot exchange rate.*

Siegel's paradox is the observation that the forward exchange rate cannot, in general, be an unbiased estimate of the future spot exchange rate. More precisely: this cannot be true simultaneously for the dollar investor and the pound investor, since $(E_P[C])^{-1} < E_P[C^{-1}]$. If the property holds for the dollar investor, then it is false for the pound investor. This is also clear from our calculation of the risk-neutral measures Q and $\overline{Q}$: if $\sigma \neq 0$ then they are different, so they can't both be equal to the subjective probability.

Continuous Time Finance Notes, Spring 2004 – Section 3, *Feb. 4, 2004*
Notes by Robert V. Kohn, Courant Institute of Mathematical Sciences. For use in connection with the NYU course Continuous Time Finance.

Announcement: There will be no class on Wednesday, February 11.

These notes wrap up our treatment of the probabilistic framework for option pricing. We conclude the discussion of problems with one source of randomness by discussing the "market price of risk" and the use of alternative numeraires (leading to consideration of equivalent martingale measures other than the risk-neutral one). Then we discuss the analogous theory for complete markets with multiple sources of randomness. Finally we discuss two important examples where there are two sources of randomness: (a) exchange options, and (b) quanto options. Except for exchange options, this material is covered in Sections 4.4-4.5 and 6.1-6.4 of Baxter & Rennie (however my discussion of quantos is organized differently). The corresponding part of Hull (5th edition) is Chapter 21 (my discussion of exchange options is from there). Hull is relatively terse, but well worth reading.

****

The market price of risk. We continue a little longer the hypothesis that the market has just one source of randomness (a scalar Brownian motion w). Recall what we achieved in Section 1: let r_t be the risk-free rate, and B the associated money-market instrument (characterized by $dB = r_t B\, dt$ with $B(0) = 1$); let S be tradeable, and assume S solves an SDE of the form $dS = \mu_t S\, dt + \sigma_t S\, dw$. We showed that if S_t/B_t is a martingale relative to Q then any payoff X at time T can be replicated by self-financing portfolio which invests in just S and B, and the time-t value of this portfolio V_t is characterized by $V_t = B_t E_Q[X/B_T \mid \mathcal{F}_t]$.

We also showed that Q exists (and is unique), by an application of Girsanov's theorem. Reviewing this: we need S/B to be a Q-martingale. Under the original probability we have

$$d(S/B) = (\mu_t - r_t)(S/B)\, dt + \sigma_t (S/B)\, dw.$$

According to Girsanov's theorem changing the measure (without changing the sets of measure zero) can change the drift but not the volatility. So Q is characterized by

$$d(S/B) = \sigma_t (S/B)\, d\tilde{w}$$

where $\tilde{w}$ is a Q-Brownian motion. Comparing these equations we see that

$$dw + \frac{\mu_t - r_t}{\sigma_t}\, dt = d\tilde{w}.$$

The ratio

$$\lambda_t = \frac{\mu_t - r_t}{\sigma_t}$$

is called the *market price of risk*, by analogy with the Capital Asset Pricing Model (notice that λ is the ratio of excess return to volatility). It is important because it determines Q in terms of P, namely:

$$dQ/dP = \exp\left(-\int_0^T \lambda_t \, dw - \tfrac{1}{2} \int_0^T \lambda_t^2 \, dt \right).$$

This Q is called the "risk-neutral measure."

Finally, we observed that if $\overline{S}$ is *any* tradeable in this market, its value $\overline{S}_t$ must have the property that $\overline{S}_t/B_t$ is a Q-martingale (for the *same* risk-neutral measure Q). This shows that *all tradeables must have the same market price of risk.*

It's enlightening to give another proof that all tradeables have the same market price of risk. Suppose S and $\overline{S}$ are distinct tradeables, with

$$dS = \mu_t S \, dt + \sigma_t S \, dw, \quad d\overline{S} = \overline{\mu}_t \overline{S} \, dt + \overline{\sigma}_t \overline{S} \, dw.$$

Consider the following trading strategy: at time t, hold $\overline{\sigma}_t \overline{S}$ units of S and $-\sigma_t S$ units of $\overline{S}$. The value of the associated portfolio is

$$V = \overline{\sigma}_t \overline{S} S - \sigma_t S \overline{S} = (\overline{\sigma}_t - \sigma_t) S \overline{S}.$$

The portfolio is risk-free, since its increments are

$$dV = \overline{\sigma}_t \overline{S} \, dS - \sigma_t S \, d\overline{S} = (\overline{\sigma}_t \mu_t - \sigma_t \overline{\mu}_t) S \overline{S} \, dt.$$

So it must rise at the risk-free rate: $dV = r_t V dt$, which gives

$$(\overline{\sigma}_t \mu_t - \sigma_t \overline{\mu}_t) = (\overline{\sigma}_t - \sigma_t) r_t$$

Algebraic rearrangement gives

$$\frac{\overline{\mu}_t - r_t}{\overline{\sigma}_t} = \frac{\mu_t - r_t}{\sigma_t},$$

which is the desired conclusion. This argument has the advantage of being in some sense more elementary. However it does not reveal the true meaning of the market price of risk – which is that it expresses the risk-neutral measure Q in terms of the original measure P.

****

Equivalent martingale measures. We observed in Section 1 (the top half of page 7) that the theory just summarized works equally well when B is replaced by any tradeable. In other words, the "numeraire" can be any tradeable N. To price options using N in place of B, we need a measure $\overline{Q}$ such that S/N is a $\overline{Q}$-martingale. Then the value V of payoff X has the property that V/N is a $\overline{Q}$-martingale. And any tradeable $\overline{S}$ has the property that $\overline{S}/N$ is a $\overline{Q}$-martingale. In particular, the value at time 0 of payoff X is

$$V_0 = N_0 E_{\overline{Q}}[X/N_T]. \tag{1}$$

The replicating portfolio uses S and N. If N is not the risk-free money-market fund B then $\overline{Q}$ is different from Q; it is called "forward risk-neutral with respect to N." To see why, observe that the contract to exchange S for k units of N at time T is valueless at time 0 when $k = E_{\overline{Q}}[S/N_T]$. Thus: the $\overline{Q}$ mean of S/N_T is the forward price of S, in units of N.

The risk-neutral measure Q, and more generally the forward risk-neutral measure $\overline{Q}$ associated with any tradeable N, are called *equivalent martingale measures*. We can price options using any of them; the choice is up to us, and is a matter of convenience.

For markets with a single source of randomness the main application of (1) is the justification of Black's formula. In that case $N_t = B(t,T)$ is the value at time t of a zero-coupon bond worth 1 dollar at time T, and (1) becomes

$$V_t = B(0,T)E_{\overline{Q}}[X].$$

Thus, by using the measure $\overline{Q}$ rather than Q, we can arrange that the discount factor be outside the expectation.

We'll give another application of (1) below, to the pricing of an exchange option. That example however involves two sources of randomness. So this is a convenient time to discuss how the theory generalizes to problems with multiple sources of randomness.

****

Complete markets with n sources of randomness. The most important examples involve two sources of randomness, for example an "exchange option" (whose payoff is $(S - \overline{S})_+$, where S and $\overline{S}$ are two stocks) or a "quantos call" (whose payoff in dollars is $(S - K)_+$, where S is the stock price of British Petroleum in pounds). Later we'll specialize to these cases. But to get started let's suppose there are n sources of randomness, and n (independent) tradeables $S^1, \ldots, S^n$. Each individually satisfies an SDE of the form

$$dS_i = \mu_i S_i\,dt + \sigma_i S_i\,dw_i \tag{2}$$

however the Brownian motions w_i may (and often will) be *correlated*. There are two (ultimately equivalent) ways to handle this. The more straightforward one is to express everything in terms of n independent Brownian motions $z_1, \ldots, z_n$:

$$dS_i = \mu_i S_i\,dt + S_i \sum_j \sigma_{ij}\,dz_j \tag{3}$$

where $\Sigma = (\sigma_{ij})$ is an $n \times n$ matrix. (Our n tradeables are "independent" if the matrix Σ is invertible.) Notice that (3) is equivalent to (2) when σ_i is the length of the ith row of Σ and $dw_i = \sigma_i^{-1} \sum_j \sigma_{ij} dz_j$. (Notice that w_i is a Brownian motion!). Sometimes however it is convenient to stick with the correlated Brownian motions w_i. In that case, when we apply Ito's formula, we must replace the shorthand $dw\,dw = dt$ by

$$dw_i\,dw_j = \rho_{ij}\,dt$$

with

$$\rho_{ij} = \sigma_i^{-1}\sigma_j^{-1}\sum_k \sigma_{ik}\sigma_{jk}\,dt.$$

Notice that along the diagonal $\rho_{ii} = 1$, consistent with the fact that each w_i is individually a Brownian motion process.

The n-factor version of the martingale representation theorem is the direct analogue of the 1-factor version: if M is a martingale (for the σ-algebra $\mathcal{F}_t$ generated by our n Brownian motions) then there exist adapted processes $\phi_1, \ldots, \phi_n$ such that

$$dM = \sum_j \phi_j dz_j.$$

If $N_1, \ldots, N_n$ are also martingales and $dN_j = \sum_k \psi_{jk} dz_k$ then (provided the matrix ψ_{jk} is invertible) we deduce the existence of representation

$$dM = \sum_j \phi'_j dN_j.$$

These representations are unique.

The n-factor version of Girsanov's theorem is also directly analogous to the one-factor version. For any adapted, vector-valued process $\gamma = (\gamma_1, \ldots, \gamma_n)$ there is a measure Q such that

$$\tilde{z} = z + \int_0^t \gamma\,ds$$

is an n-dimensional Brownian motion under Q (i.e. each component is an independent Q-Brownian motion). Its density relative the the original measure P (with respect to which z is an n-dimensional Brownian motion) is

$$dQ/dP = \exp\left(-\int_0^T \sum_i \gamma_i\,dz_i - \tfrac{1}{2}\int_0^T |\gamma|^2\,dt\right).$$

Notice that the relation between $\tilde{z}$ and z can be written as

$$d\tilde{z} = dz + \gamma dt.$$

Our discussion of option pricing carries over to the n-factor setting with essentially no change. We won't repeat everything, but let characterize the risk-neutral measure Q. To price options using the tradeables S_i described by (3) we need S_i/B to be a martingale under Q. Under the original probability we have

$$d(S_i/B) = (\mu_i - r)(S_i/B)\,dt + (S_i/B)\sum_j \sigma_{ij}\,dz_j.$$

According to Girsanov's theorem changing the measure (without changing the sets of measure zero) can change the drift but not the volatility. So Q is characterized by

$$d(S_i/B) = (S_i/B)\sum_j \sigma_{ij}\,d\tilde{z}_j.$$

Thus we need

$$d\tilde{z} = dz + \Sigma^{-1}(\mu - r1)\, dt$$

using vector notation, with $\Sigma = (\sigma_{ij})$ and $1 = (1, \ldots, 1)$. The vector-valued process

$$\lambda = \Sigma^{-1}(\mu - r1)$$

is again called the *market price of risk*. It characterizes the risk-neutral measure, which is unique. So as in the one-factor case, the market price of risk is independent of the particular choice of n independent tradeables used to value options and construct replicating portfolios.

A brief digression about incomplete markets. We assumed the existence of n independent tradeables so we could solve uniquely for λ and Q. In some settings however there are fewer independent tradeables than sources of randomness. A simple example would be an option whose payoff depends the values of two underlyings $S_1(T)$ and $S_2(T)$, if only S_1 is tradeable. A more fundamental example is any *stochastic volatility* model, since volatility is never tradeable (at least, not in the literal sense). In such settings the linear system giving $d\tilde{z}$ in terms of dz is underdetermined (we have n unknowns, but only as many equations as we have independent tradeables). Therefore there is more than one possible measure Q, i.e. more than one possible market price of risk. It is still true that every tradeable must have the *same* market price of risk (see Hull's Appendix 21B for a proof). So it is still true that all tradeables must be priced using the same measure Q. But this market is not complete; the measure Q cannot be determined from absence of arbitrage alone; most payoffs cannot be replicated. Pricing and hedging in an incomplete market requires an entirely different viewpoint. We'll discuss some incomplete market models in the last 1/3 of the semester.

Returning now to the complete setting: As in the one-factor case, we need not use the money-market account as our numeraire; rather, we can use any tradeable N. The associated equivalent martingale measure has the property $\overline{Q}$ has the property that S/N is a $\overline{Q}$-martingale for every tradeable S. In particular the value at time 0 of an option with payoff X is $V_0 = N_0 E_{\overline{Q}}[X/N_T]$.

What must we do in practice to value and hedge an option in this multifactor setting? If the risk-free rate r and the volatilities σ_{ij} are constant then the risk-neutral process

$$dS_i = rS_i\, dt + S_i \sum_j \sigma_{ij}\, d\tilde{z}_j$$

is easy to understand: the vector $(\log S_1, \ldots, \log S_n)$ has a multivariate normal distribution. One can therefore value any European option by a suitable integration using the multivariate normal distribution. In greater generality – if σ_{ij} depend only on the state variables S_i and time – one can value options by solving the multivariate analogue of the Black-Scholes PDE. The hedge is, as usual, determined by using $\partial V/\partial S_i$ units of the ith stock.

In some important special cases we can get more explicit valuation formulas. The rest of this section is devoted to two examples of this kind.

****

Exchange options. Suppose now there are two sources of randomness, and we are interested in two stocks S_1 and S_2. Let's use the representation (2), and let ρ be the correlation between w_1 and w_2:

$$dw_1\, dw_2 = \rho\, dt.$$

Consider the "exchange option," which gives the holder the right to exchange S_1 for S_2 at time T. Its payoff is $(S_2 - S_1)_+$.

The key to valuing this option is to use the equivalent martingale measure $\overline{Q}$ which is forward risk-neutral with respect to S_1. It has the property that S/S_1 is a $\overline{Q}$-martingale for every tradeable S. The value it assigns to our option at time 0 is

$$V_0 = S_1(0) E_{\overline{Q}}[(S_2/S_1 - 1)_+]$$

since

$$\frac{(S_2 - S_1)+}{S_1} = (S_2/S_1 - 1)_+.$$

By Ito's formula, the SDE satisfied by $X = S_2/S_1$ (under the original measure) is

$$dX = S_2 d(S_1^{-1}) + S_1^{-1} dS_2 + dS_2\, dS_1^{-1}.$$

Since

$$d(S_1^{-1}) = -S_1^{-2}\, dS_1 + S_1^{-3}\, dS_1\, dS_1 = (\sigma_1^2 - \mu_1) S_1^{-1}\, dt - \sigma_1 S_1^{-1}\, dw_1$$

a bit of manipulation gives

$$dX = (\text{stuff})\, dt - \sigma_1 X\, dw_1 + \sigma_2 X\, dw_2.$$

Under $\overline{Q}$ the process must be drift-free, with the same volatility; therefore

$$dX = -\sigma_1 X\, d\tilde{w}_1 + \sigma_2 X\, d\tilde{w}_2$$

where $\tilde{w}_1$ and $\tilde{w}_2$ are independent $\overline{Q}$-Brownian motions. Now assume that σ_1, σ_2 and ρ are constant, and observe that $\sigma_1 \tilde{w}_1 + \sigma_2 \tilde{w}_2$ has mean value 0 and variance $\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2$. A moment's reflection reveals that it is in fact $\hat{\sigma}$ times a $\overline{Q}$ Brownian motion, where $\hat{\sigma} = (\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)^{1/2}$. So $X_T = S_2(T)/S_1(T)$ is lognormal under $\overline{Q}$, with mean value $S_2(0)/S_1(0)$ and volatility $\hat{\sigma}$. By Equation (1) of Section 2 we conclude that

$$E_{\overline{Q}}[(S_2/S_1 - 1)_+] = \frac{S_2(0)}{S_1(0)} N(d_1) - N(d_2)$$

where

$$d_1 = \frac{\ln[S_2(0)/S_1(0)] + \frac{1}{2}\hat{\sigma}^2 T}{\hat{\sigma}\sqrt{T}}, \quad d_2 = \frac{\ln[S_2(0)/S_1(0)] - \frac{1}{2}\hat{\sigma}^2 T}{\hat{\sigma}\sqrt{T}}.$$

The value of the exchange option is therefore

$$V_0 = S_2(0) N(d_1) - S_1(0) N(d_2).$$

The options with payoff $\min(S_1, S_2)$ and $\max(S_1, S_2)$ are easily priced using this result. Indeed, we have

$$\min(S_1, S_2) = S_2 - (S_2 - S_1)_+$$

and

$$\max(S_1, S_2) = S_1 + (S_2 - S_1)_+$$

so the value of the min payoff is $S_2(0) - V_0$ and the value of the max payoff is $S_1(0) + V_0$.

****

Quantos. A quanto option is an option on a stock price quoted in a foreign currency. For example, if S is the price of a stock in British pounds, a a quanto call on this underlying with strike K has payoff $(S_T - K)_+$ *dollars.* Clearly we are again dealing with two sources of randomness: the stock price S (in pounds) and the exchange rate C (expressed as dollars per pound). It is again convenient to use viewpoint (2): we suppose S and C satisfy

$$dS = \mu_S S\,dt + \sigma_S S\,dw_S$$

and

$$dC = \mu_C C\,dt + \sigma_C C\,dw_C$$

and we suppose the Brownian motions w_S and w_C have correlation ρ. As in Section 2, the risk-free rate in pounds is u and the risk-free rate in dollars is r, so the pound money market account D and the dollar money market account B satisfy $dD = uD\,dt$ and $dB = rB\,dt$.

The pound investor sees tradeables D (the pound money market account), S (the stock), and BC^{-1} (the dollar money market account, valued in pounds). His risk-neutral measure – let's call it $\overline{Q}$, reserving the notation Q for the dollar investor's risk-neutral measure – has the property that S/D and B/CD are $\overline{Q}$-martingales. We could work out what $\overline{Q}$ is if we wanted to. But we don't need to, so we won't.

The dollar investor sees tradeables CD (the pound money market account, valued in dollars), CS (the stock, valued in dollars), and B. His risk-neutral measure Q has the property that CD/B and CS/B are Q-martingales. Since the volatilities are unchanged when we pass from the subjective measure to $\overline{Q}$ or Q, the fact that CD/B is a Q-martingale tells us that

$$d(CD/B) = \sigma_C (CD/B)\,d\tilde{w}_C$$

where $\tilde{w}_C$ is a Q-Brownian motion.

We're interested in an option whose dollar payoff is a function of S_T (for example a quanto call, with payoff $f(S_T) = (S_T - K)_+$). So the crucial question is: what is the risk-neutral process satisfied by S? Since changing the measure leaves the volatility invariant, it must take the form

$$dS = \alpha S\,dt + \sigma_S S\,d\tilde{w}_S$$

where $\tilde{w}_S$ is a Q-Brownian motion, whose correlation with $\tilde{w}_C$ is ρ, i.e. the same as the correlation between w_S and w_C. Our task is to find α. Our tool is the fact that CS/B is a Q-martingale. Since $CS/B = (S/D)(CD/B)$, Ito's formula gives

$$\begin{aligned} d(SC/B) &= (S/D)\, d(CD/B) + (CD/B)\, d(S/D) + d(S/D)\, d(CD/B) \\ &= (SC/B)\sigma_C\, d\tilde{w}_C + (SC/B)[(\alpha - u)\, dt + \sigma_S S\, d\tilde{w}_S] + (SC/B)\sigma_S\sigma_C\, d\tilde{w}_S\, d\tilde{w}_C \\ &= (SC/B)(\alpha - u + \rho\sigma_S\sigma_C)\, dt + (SC/B)\sigma_C\, d\tilde{w}_C + (SC/B)\sigma_S\, d\tilde{w}_S \end{aligned}$$

since $\tilde{w}_S$ and $d\tilde{w}_S\, d\tilde{w}_C = \rho\, dt$. For this to be a Q-martingale the drift must vanish, so

$$\alpha = u - \rho\sigma_S\sigma_C.$$

Thus if r is constant, the time-0 value (in dollars) of the quanto with payoff $f(S_T)$ (dollars) is

$$e^{-rT} E_Q[f(S_T)]. \tag{4}$$

If σ_C, σ_S, and ρ are constant then the distribution of S_T (viewed using Q) is lognormal with mean value

$$E_Q[S_T] = e^{(u-\rho\sigma_S\sigma_C)T} S_0$$

(this is the quanto forward rate, i.e. the choice of k for which payoff $S_T - k$ has initial value 0) and volatility $\sigma_S\sqrt{T}$. For a quanto put or a quanto call, we easily deduce a Black-Scholes-like formula (for example: use eqn 1 of Section 2 for the case of a call).

Continuous Time Finance Notes, Spring 2004 – Section 4, *Feb. 18, 2004*
Notes by Robert V. Kohn, Courant Institute of Mathematical Sciences. For use in connection with the NYU course Continuous Time Finance.

This section begins our discussion of interest rate models. You should already have some familiarity with basic terminology (e.g. bond prices, instantaneous forward rates), instruments (e.g. caps and captions, swaps and swaptions), and their valuation using Black's formula. If you need to review these topics, Hull is excellent; see also my Derivative Securities lecture notes, sections 10 and 11 (warning: the notation there is slightly different from the present notes). Today we'll get started by discussing relatively simple one-factor models of the short rate (Vasicek, Cox-Ingersoll-Ross). Next time we'll discuss the Hull-White (also known as modified Vasicek) model, a richer one-factor short rate model that can be calibrated to an arbitrary initial yield curve. After that we'll turn to the Heath-Jarrow-Morton theory.

I've adopted Baxter & Rennie's notation. However my pedagogical strategy is different from theirs. I'm starting with short-rate models, because they're easier. Baxter & Rennie start with HJM, specializing afterward to short-rate models, because it's more efficient. The Vasicek model is by now quite standard; treatments can be found (with slightly different organization and viewpoints) in Brigo & Mercurio, Lamberton & Lapeyre, and Avellaneda & Laurence. For a good survey of the "big picture," I recommend reading chapter 1 of R. Rebonato, *Modern pricing of interest-rate derivatives: the Libor market model and beyond* (2nd edition, 2002, on reserve in the CIMS library).

****

Basic terminology. The time-value of money is expressed by the *discount factor*

$$P(t,T) = \text{value at time } t \text{ of a dollar received at time } T.$$

This is, by its very definition, the price at time t of a zero-coupon (risk-free) bond which pays one dollar at time T. If interest rates are stochastic then $P(t,T)$ will not be known until time t; its evolution in t is random, and can be described by an SDE. Note however that $P(t,T)$ is a function of *two* variables, the initiation time t and the maturity time T. The dependence on T reflects the *term structure* of interest rates; we expect $P(t,T)$ to be relatively smooth as a function of T, since interest is being averaged over the time interval $[t,T]$. We usually take the convention that the present time is $t=0$ – thus what is observable now is $P(0,T)$ for all $T>0$.

There are several equivalent ways to represent the time-value of money. The *yield* $R(t,T)$ is defined by

$$P(t,T) = e^{-R(t,T)(T-t)};$$

it is the unique constant interest rate that would have the same effect as $P(t,T)$ under continuous compounding. Evidently

$$R(t,T) = -\frac{\log P(t,T)}{T-t}.$$

The *instantaneous forward rate* $f(t,T)$ is defined by

$$P(t,T) = e^{-\int_t^T f(t,\tau)\,d\tau};$$

fixing t, it is the deterministic time-varying interest rate that describes all the loans starting at time t with various maturities. Clearly

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

The *instantaneous interest rate* $r(t)$, also called the short rate, is

$$r(t) = f(t,t);$$

it is the rate earned on the shortest-term loans starting at time t. Notice that the yields $R(t,T)$ and the instantaneous forward rate $f(t,T)$ carry the same information as the entire family of discount factors $P(t,T)$; however the short rate $r(t)$ carries much less information, since it is a function of just one variable.

Why is $f(t,T)$ called the instantaneous forward rate? To explain, we start by observing that the ratio

$$P(0,T)/P(0,t)$$

has an important interpretation: it is the discount factor (for time-t borrowing, with maturity T) that can be locked in at time 0, at no cost, by a combination of market positions. In fact, consider the following portfolio:

(a) long a zero-coupon bond worth one dollar at time T (present value $P(0,T)$), and

(b) short a zero-coupon bond worth $P(0,T)/P(0,t)$ at time t (present value $-P(0,T)$).

Its present value is 0, and its holder pays $P(0,T)/P(0,t)$ at time t and receives one dollar at time 0. Thus the holder of this portfolio has "locked in" $P(0,T)/P(0,t)$ as his discount rate for borrowing from time t to time T. This ratio is called the *forward term rate* at time 0, for borrowing at time t with maturity T. Similarly, $P(t,T_2)/P(t,T_1)$ is the forward term rate at time t, for borrowing at time T_1 with maturity T_2. The associated yield is

$$-\frac{\log P(t,T_2) - \log P(t,T_1)}{T_2 - T_1}.$$

In the limit $T_2 - T_1 \to 0$ we get $-\partial \log P(t,T)/\partial T = f(t,T)$.

What's our goal? As usual, we want a no-arbitrage-based framework for pricing and hedging options. A simple example would be a call option on a zero-coupon bond. If the option's holder has the right to purchase a zero-coupon bond at time t, paying one dollar at maturity T, for price K, then his payoff at time t is $(P(t,T) - K)_+$. The task of pricing a caplet is almost the same: consider an option that caps the term interest rate $\mathcal{R}$ for lending between times T_1 and T_2 at rate R_0. For a loan with principal one dollar, the option's payoff at time T_2 is

$$\Delta t \max\{\mathcal{R} - R_0, 0\}$$

where $\Delta t = T_2 - T_1$ and $\mathcal{R}$ is the actual term interest rate in the market at time T_1 (defined by $P(T_1, T_2) = 1/(1 + \mathcal{R}\Delta t)$). The discounted value of this payoff at time T_1 is

$$(1 + R_0 \Delta t) \max \left\{ \frac{1}{1 + R_0 \Delta t} - \frac{1}{1 + \mathcal{R}\Delta t}, 0 \right\}$$

since

$$1 - \frac{1 + R_0 \Delta t}{1 + \mathcal{R}\Delta t} = \frac{(\mathcal{R} - R_0)\Delta t}{1 + \mathcal{R}\Delta t}.$$

So the caplet is equivalent to $(1 + R_0 \Delta t)$ put options on a zero-coupon bond at time T_1, paying one dollar at maturity T_2, with strike price $1/(1 + R_0 \Delta t)$. A cap consists of a collection of caplets; it is equivalent by this argument to a portfolio of puts on zero-coupon bonds.

Let's step back a minute. If this is the first time you think about interest-based instruments, you may wonder how the interest rate can be simultaneously risk-free and random. For short-rate models this is easy to understand. We already saw the continuous-time version on HW1, problem 1: if, under the risk-neutral probability, the short rate r solves a stochastic differential equation of the form $dr = \alpha(r,t)\, dt + \beta(r,t)\, dw$ then the value at time t of a dollar received at time T is

$$P(t,T) = E\left[e^{-\int_t^T r(s)\, ds} \mid \mathcal{F}_t \right].$$

Moreover $P(t,T)$ can be determined by solving the PDE

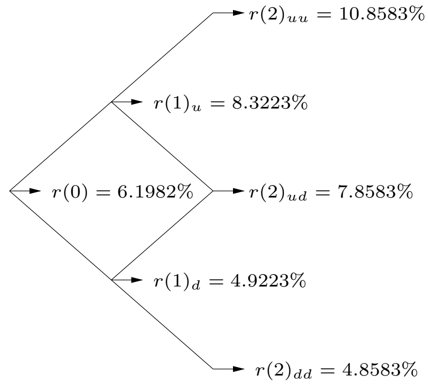
$$V_t + \alpha V_r + \tfrac{1}{2}\beta^2 V_{rr} - rV = 0$$

with final-time condition $V(T,r) = 1$ for all r at $t = T$; the value of $P(t,T)$ is then $V(t, r(t))$.

To make this very concrete, let's briefly discuss how something similar can be done with a binomial tree. (My treatment follows one in the textbook by Jarrow & Turnbull.) The basic idea is shown in the figure: each node of the tree is assigned a risk-free rate, different from node to node; it is the one-period risk-free rate for the binomial subtree just to the right of that node.

What probabilities should we assign to the branches? It might seem natural to start by figuring out what the subjective probabilities are. But why bother – what we need for option pricing are the risk-neutral probabilities. Moreover recall that there is always some freedom in designing a binomial tree to fit a specified stochastic process; the continuum limit is governed by the central limit theorem, so what matters is that we get the means and variances right. We can therefore take the convention that the risk-neutral probability of going up at any branch is $q = 1/2$, then choose the "drifts" and "volatilities" implicit in the nodes to mimic the desired stochastic process. This procedure leads to the Black-Derman-Toy framework for valuing interest-based instruments; briefly, the approach involves

- restricting attention to the risk-neutral interest rate process;
- assuming the risk-neutral probability is $q = 1/2$ at each branch; and



- choosing the interest rates at the various nodes so that the values of $P(0,T)$ obtained using the tree match those observed in the marketplace for all T.

The last bullet – the *calibration* of the tree to market information – is of course crucial. It will be discussed at length in the class Interest Rate and Credit Models. Here we won't try to do any calibration. But let's be sure we understand how the tree determines $P(0,T)$ for all T. As an example let's determine $P(0,3)$, the value at time 0 of a dollar received at time 3, for the tree shown in the figure. We take the convention that the interest rates shown in the tree are rates per time period (equivalently: we suppose the time period is $\Delta t = 1$).

Consider first time period 2. The value at time 2 of a dollar received at time 3 is $P(2,3)$; it has a different value at each time-2 node. These values are computed from the fact that

$$P(2,3) = e^{-r\Delta t}[\tfrac{1}{2}P(3,3)_{\rm up} + \tfrac{1}{2}P(3,3)_{\rm down}] = e^{-r\Delta t}$$

so

$$P(2,3) = \begin{cases} e^{-r(2)_{uu}} = .897104 & \text{at node } uu \\ e^{-r(2)_{ud}} = .924425 & \text{at node } ud \\ e^{-r(2)_{dd}} = .952578 & \text{at node } dd. \end{cases}$$

Now we have the information needed to compute $P(1,3)$, the value at time 1 of a dollar received at time 3. Applying the rule

$$P(1,3) = e^{-r\Delta t}[\tfrac{1}{2}P(2,3)_{\rm up} + \tfrac{1}{2}P(2,3)_{\rm down}]$$

at each node gives

$$P(1,3) = \begin{cases} e^{-r(1)_u}(\tfrac{1}{2}\cdot .897104 + \tfrac{1}{2}\cdot .924425) = .838036 & \text{at node } u \\ e^{-r(1)_d}(\tfrac{1}{2}\cdot .924425 + \tfrac{1}{2}\cdot .952578) = .893424 & \text{at node } d. \end{cases}$$

Finally we compute $P(0,3)$ by applying the same rule:

$$\begin{aligned} P(0,3) &= e^{-r\Delta t}[\tfrac{1}{2}P(1,3)_{\rm up} + \tfrac{1}{2}P(1,3)_{\rm down}] \\ &= e^{-r(0)}[\tfrac{1}{2}\cdot .838036 + \tfrac{1}{2}\cdot .893424] = .8137. \end{aligned}$$

****

General orientation. Let's briefly review our understanding of options on equities in a constant-interest-rate environment. The underlying is assumed to solve a stochastic differential equation; in the simplest one-factor setting the model is determined by specifying the drift and volatility. They determine a unique risk-neutral measure, and (equivalently) a unique market price of risk. Every option – indeed, every tradeable – has a unique price, forced upon us by the existence of replicating portfolios and the principle of no-arbitrage. Do the prices we get this way match the marketplace? Well, that depends what we use for the volatility. We often use a constant volatility, e.g. the implied volatility for at-the-money options; the prices obtained this way do *not* match market prices exactly (the implied volatility of an option depends, in practice, on its strike and maturity). Though imperfect, the theory is nevertheless useful, because it gives approximate hedge portfolios. To do better, one can make the theory more complicated – e.g. by looking for a time-and-price-dependent volatility that prices all European options (with all strikes and maturities) correctly at time 0; the resulting "local vol" model would then be used to price more exotic options (still at time 0).

In modeling interest rates we have a similar tradeoff between simplicity and accuracy. There are three basic viewpoints:

(a) *Simple short rate models.* Historically these came first. A standard example, which we'll discuss in detail, is the Vasicek model, which assumes that under the risk-neutral probability the short rate solves

$$dr = (\theta - ar)\, dt + \sigma\, dw \tag{1}$$

with θ, a, and σ constant and $a > 0$. This mean-reverting diffusion is known to physicists as an Ornstein-Uhlenbeck process. (We don't have to start with the risk-neutral process. It is equivalent to assume the subjective short-rate process has the form (1) and the market price of risk is constant. Then the risk-neutral process has the form (1) as well, with a new value of θ.)

The advantage of such a model is that it leads to explicit formulas. Moreover for some short-rate models – including Vasicek – we can see relatively easily that Black's formula is valid, since $P(t,T)$ has lognormal statistics under the forward risk-neutral measure.

The disadvantage of such a model is that it has just a few parameters. So there's no hope of calibrating it to match the entire yield curve $P(0,T)$ observed in the marketplace at time 0. For this reason Vasicek and its siblings are rarely used in practice. The main purpose of discussing it is to warm up.

(b) *Richer short-rate models.* The conceptual simplicity of a short-rate model is attractive. But we want to calibrate it to the full yield curve $P(0,T)$ observed at time 0. So the model must depend on a function of one variable. A standard example, which we'll discuss in detail soon, is the extended Vasicek model, also called the Hull-White model:

$$dr = (\theta(t) - ar)\, dt + \sigma\, dw \tag{2}$$

where a and σ are still constant but θ is a function of t. We'll show in due course that when θ satisfies

$$\theta(t) = \frac{\partial f}{\partial T}(0,t) + af(0,T) + \frac{\sigma^2}{2a}(1 - e^{-2at}) \tag{3}$$

the Hull-White model correctly reproduces the entire yield curve at time 0.

Hull-White retains many of the advantages of the basic Vasicek model. It still leads to explicit formulas; moreover it is still consistent with Black's formula (this is important, as a practical matter, since Black's formula is widely used in practice). Also it can be approximated by a recombining trinomial tree (very convenient for numerical use).

However Hull-White still has a major disadvantage: it gives us little freedom in modeling the *evolution* of the yield curve. Tomorrow's yield curve will be different from today's; but if we insist on getting the yield curve right at time 0, Hull-White we can only play with a and σ to try and fit this evolution. That, of course, is very constraining.

(c) *One-factor Heath-Jarrow-Morton.* We'd like a theory that can be calibrated precisely to the time-0 yield curve, but also permits a rich family of possible assumptions about the evolution of the yield curve. The one-factor HJM framework achieves precisely this goal. Rather than work in terms of a short-rate, it specifies the evolution of the instantaneous forward rate $f(t,T)$ by solving an SDE in t:

$$d_t f(t,T) = \alpha(t,T)\,dt + \sigma(t,T)\,dW_t. \tag{4}$$

The initial data $f(0,T)$ should naturally be taken from market data. The volatility $\sigma(t,T)$ in (4) must be specified – it is what determines the model. We'll see presently that Vasicek corresponds to the choice $\sigma(t,T) = \sigma e^{-a(T-t)}$. It turns out that the drift $\alpha(t,T)$ cannot be specified independently – the fact that all bonds (with all maturities) have the same market price of risk will give us an expression for α in terms of σ. (This is the analogue, for HJM, of the fact that for equities the risk-neutral process always has drift r.)

Every theory has advantages and disadvantages, and HJM is no exception. For one thing, it gives us too much freedom: how, in practice, should we choose $\sigma(t,T)$? Another disadvantage is the difficulty of using HJM numerically: only a few special cases (mainly corresponding to familiar short-rate models such as Hull-White and Black-Derman-Toy) give Markovian short-rates, which can be modelled using recombining trees.

The preceding viewpoints are in order of increasing complexity, and also in historical order. It is remarkable how recent these theories are. The fundamental work on equity-based options (e.g. the papers by Black, Scholes, and Merton) was done in the mid 70's, and the use of simple short-rate models like Vasicek dates from about the same time. But the development of Hull-White and its siblings (e.g. Black-Derman-Toy) dates from about 1990. And the fundamental papers by Heath, Jarrow, and Morton were published in the same year. In brief: the modern viewpoint on interest-rate modeling is not quite 15 years old!

The preceding discussion has focused entirely on one-factor models. It is of course possible (and valuable) to consider models with multiple sources of randomness (multifactor short-rate and multifactor HJM models). But let's not get carried away. We have already bitten off a lot; it's time to chew and swallow it.

****

The Vasicek model. Vasicek's model (1) has the advantage that we can work out almost everything explicitly. Moreover many of the methods used to analyze it extend straightforwardly to the Hull-White model (2).

An explicit formula for $r(t)$. It is easy to solve (1) explicitly: we have

$$d(e^{at}r) = e^{at}\,dr + ae^{at}r\,dt = \theta e^{at}\,dt + e^{at}\sigma\,dw,$$

so

$$e^{at}r(t) = r(0) + \theta\int_0^t e^{as}\,ds + \sigma\int_0^t e^{as}dw(s).$$

which simplifies to

$$r(t) = r(0)e^{-at} + \frac{\theta}{a}(1-e^{-at}) + \sigma\int_0^t e^{-a(t-s)}\,dw(s). \tag{5}$$

That calculation could have started at any time; thus

$$r(t) = r(s)e^{-a(t-s)} + \frac{\theta}{a}(1-e^{-a(t-s)}) + \sigma\int_s^t e^{-a(t-\tau)}\,dw(\tau). \tag{6}$$

We observe from (5) that $r(t)$ is Gaussian (the stochastic integral is Gaussian, because each term of the approximating Riemann sum is Gaussian, and sums of Gaussians are Gaussian.) Its mean is clearly

$$E[r(t)] = r(0)e^{-at} + \frac{\theta}{a}(1-e^{-at}),$$

and the variance is

$$Var\,[r(t)] = \sigma^2 E\left[\left(\int_0^t e^{-a(t-s)}\,dw(s)\right)^2\right] = \sigma^2\int_0^t e^{-2a(t-s)}\,ds = \frac{\sigma^2}{2a}(1-e^{-2at}).$$

We can use the preceding calculation to see that $P(t,T)$ has lognormal statistics. Indeed, by definition

$$P(t,T) = E\left[e^{-\int_t^T r(s)\,ds}\,|\,\mathcal{F}_t\right]. \tag{7}$$

Using (6) (with t and s interchanged) we have

$$P(t,T) = A(t,T)e^{-B(t,T)r(t)} \tag{8}$$

where

$$B(t,T) = \int_t^T e^{-a(s-t)}\, ds, \quad \text{and} \quad A(t,T) = E\left[e^{-\int_t^T \left\{\frac{\theta}{a}(1-e^{-a(s-t)})+\sigma\int_t^s e^{-a(s-\tau)}\,dw(\tau)\right\} ds}\right].$$

Since $A(t,T)$ and $B(t,T)$ are deterministic and $r(t)$ is Gaussian, (8) shows that $P(t,T)$ is lognormal.

An explicit formula for P(t,T). One approach is to evaluate $A(t,T)$ and $B(t,T)$ from the definitions given above. The calculation of B is easy; that of A is less trivial – an amusing exercise in stochastic integration. You'll find the details of an equivalent calculation on pages 128-129 of Lamberton & Lapeyre.

Here let's implement another approach, which has the nice feature of being somewhat more general. (For example, a similar calculation works for the Cox-Ingersoll-Ross model $dr = (\theta - ar)\,dt + \sigma\sqrt{r}\,dw$, though the distribution of r is *not* Gaussian in this case.) Recall that $P(t,T) = V(t, r(t))$ where V solves the PDE

$$V_t + (\theta - ar)V_r + \tfrac{1}{2}\sigma^2 V_{rr} - rV = 0$$

with final-time condition $V(T,r) = 1$ for all r at $t = T$. Motivated by the preceding discussion, let's look for a solution of the form

$$V = A(t,T)e^{-B(t,T)r}.$$

To satisfy the PDE, A and B should satisfy

$$A_t - \theta AB + \tfrac{1}{2}\sigma^2 AB^2 = 0 \quad \text{and} \quad B_t - aB + 1 = 0$$

with final-time conditions

$$A(T,T) = 1 \quad \text{and} \quad B(T,T) = 0.$$

Solving for B first, then A, we get

$$B(t,T) = \frac{1}{a}(1 - e^{-a(T-t)})$$

and

$$A(t,T) = \exp\left[\left(\frac{\theta}{a} - \frac{\sigma^2}{2a^2}\right)(B(t,T) - T + t) - \frac{\sigma^2}{4a}B^2(t,T)\right].$$

The desired formula for $P(t,T)$ is of course

$$P(t,T) = A(t,T)e^{-B(t,T)r(t)}.$$

Evaluating the associated term structure, and its volatility. Vasicek has only three parameters, so of course its term structure is rather special. For later comparison with HJM, let's identify it by calculating $f(0,T)$. The calculation is slightly tedious but entirely straightforward, using the definition $f(0,T) = -\partial \log P(0,T)/\partial T$ and our formula for $P(0,T)$. The final result is:

$$f(0,T) = \frac{\theta}{a} + e^{-aT}(r_0 - \frac{\theta}{a}) - \frac{\sigma^2}{2a^2}(1 - e^{-aT})^2.$$

Note that this is consistent with (3).

An HJM model is characterized by the volatility of the instantaneous forward rate. We'll show in due course that Vasicek is a special case of HJM. Preparing for this, let's identify now the volatility of f, i.e. let's find $\sigma(t,T)$ such that

$$d_t f(t,T) = (\text{stuff})\, dt + \sigma(t,T)\, dw.$$

This is easy: we have $\log P(t,T) = \log A(t,T) - B(t,T)r(t)$, so

$$f(t,T) = -\partial_T \log A(t,T) + \partial_T B(t,T) r(t),$$

so by Ito's formula

$$\sigma(t,T) = \sigma \partial_T B(t,T) = \sigma e^{-a(T-t)}.$$

Validity of Black's formula. To show Black's formula is valid for options on zero-coupon-bonds, we must show that $P(t,T)$ is lognormal under the forward-risk-neutral measure. (Remember: this is the measure under which tradeables normalized by $P(t,T)$ are martingales.) We already know it is lognormal under the risk-neutral measure, but here we're interested in a different measure, associated with a different numeraire.

Let's review how change-of-numeraire works in the one-factor setting. The risk-neutral measure is associated with the risk-free money-market account B as numeraire (by definition $dB = rB\, dt$ with $B(0) = 1$). Let N be another numeraire, and suppose $\overline{Q}$ is the associated equivalent martingale measure. We can only use tradeables as numeraires, so the risk-neutral process for N is

$$dN = rN\, dt + \sigma_N N\, dw$$

where w is a Brownian motion under the risk-neutral measure. By Ito we have

$$d(B/N) = B\, d(N^{-1}) + N^{-1}\, dB$$

(there is no $dBd(N^{-1})$ term since $dB = rB\, dt$). A bit of algebra gives

$$d(B/N) = (B/N)\sigma_N^2\, dt - (B/N)\sigma_N\, dw.$$

When we use N as numeraire, the associated martingale measure $\overline{Q}$ is characterized by the fact that B/N is a $\overline{Q}$-martingale, i.e.

$$d(B/N) = -(B/N)\sigma_N\, d\overline{w}$$

where $\overline{w}$ is a $\overline{Q}$-Brownian motion. It follows that

$$d\overline{w} = -\sigma_N\, dt + dw.$$

What's the point of all this? We need to write the SDE for the Vasicek process under the forward-risk-neutral measure, i.e. when the numeraire is $P(t,T)$. Recall that $P(t,T) = A(t,T)e^{-B(t,T)r(t)}$, so from Ito the volatility of $P(t,T)$ is $-B(t,T)\sigma$. Therefore the preceding calculation gives

$$d\overline{w} = \sigma B(t,T)\,dt + dw.$$

We conclude that the SDE for the interest rate is

$$dr = (\theta - ar)\,dt + \sigma\,dw = [\theta - ar - \sigma^2 B(t,T)]\,dt + \sigma\,d\overline{w}$$

where $\overline{w}$ is a Brownian motion under the forward-risk-neutral measure. This SDE can be solved much as we did before (the fact that the drift depends on time doesn't disturb the calculation at all). Explicit formulas are a bit tedious to obtain – but they're not needed. From the structure of the calculation, we see quite easily (just as in the risk-neutral case) that the short rates and bond prices are lognormal.

****

Aside from being too restrictive (not enough parameters), Vasicek has one other awkward feature: since $r(t)$ is Gaussian, there is a positive probability that it is negative. This is of course unrealistic: interest rates must be positive. The Cox-Ingersoll-Ross model mentioned above was introduced to fix this problem: one can show that if $\theta \geq \sigma^2/2$ then r stays positive (with probability one). The form of $P(t,T)$ can be found by a separation-of-variables calculation much like the one done above. The statistics of $r(t)$ can also be characterized. See e.g. Avellaneda and Laurence for some analysis of this model.

Continuous Time Finance Notes, Spring 2004 – Section 5, *Feb. 25, 2004*
Notes by Robert V. Kohn, Courant Institute of Mathematical Sciences. For use in connection with the NYU course Continuous Time Finance.

This section discusses the Hull-White model. The continuous-time analysis is not much more difficult than Vasicek – everything is still quite explicit. The basic paper is Rev. Fin. Stud. 3, no. 4 (1990) 573-592, downloadable via JSTOR; my treatment is much simpler though because I keep the parameters a and σ constant rather than letting them (as well as θ) vary with time.

The real importance of Hull-White is that while it's rich enough to match any forward curve, it's also simple enough to be approximated by a (recombining, trinomial) tree. This topic is covered very clearly in Sections 23.11-23.12 of Hull (5th edition), so I won't cover it separately in these notes.

****

The Hull-White model. Also sometimes known as "extended Vasicek," this model assumes that the risk-neutral process for the short rate has the form

$$dr = (\theta(t) - ar)\, dt + \sigma\, dw \tag{1}$$

where a and σ are constant but θ is a function of t. (Actually the 1990 paper by Hull and White also considers taking $a = a(t)$ and $\sigma = \sigma(t)$.) We'll show that

(a) for a given choice of $\theta(t)$, the situation is a lot like Vasicek;

(b) there is a unique choice of θ that matches the term structure observed in the marketplace at $t = 0$.

Solving for $r(t)$. The calculation is entirely parallel to Vasicek: we have

$$d(e^{at}r) = e^{at}\, dr + ae^{at}r\, dt = \theta(t)e^{at}\, dt + e^{at}\sigma\, dw,$$

so

$$e^{at}r(t) = r(0) + \int_0^t \theta(s)e^{as}\, ds + \sigma \int_0^t e^{as} dw(s).$$

which simplifies to

$$r(t) = r(0)e^{-at} + \int_0^t \theta(s)e^{-a(t-s)}\, ds + \sigma \int_0^t e^{-a(t-s)}\, dw(s). \tag{2}$$

That calculation could have started at any time; thus

$$r(t) = r(s)e^{-a(t-s)} + \int_s^t \theta(\tau)e^{-a(t-\tau)}\, ds + \sigma \int_s^t e^{-a(t-\tau)}\, dw(\tau).$$

Notice that $r(t)$ is still Gaussian.

Solving for $P(t,T)$. We use the same PDE method that worked for Vasicek. We know that $P(t,T) = V(t, r(t))$ where V solves the PDE

$$V_t + (\theta(t) - ar)V_r + \tfrac{1}{2}\sigma^2 V_{rr} - rV = 0$$

with final-time condition $V(T,r) = 1$ for all r at $t = T$. Let's look for a solution of the form

$$V = A(t,T)e^{-B(t,T)r(t)}. \tag{3}$$

To satisfy the PDE, A and B should satisfy

$$A_t - \theta(t)AB + \tfrac{1}{2}\sigma^2 AB^2 = 0 \quad \text{and} \quad B_t - aB + 1 = 0$$

with final-time conditions

$$A(T,T) = 1 \quad \text{and} \quad B(T,T) = 0.$$

The equation for B doesn't involve θ, so the solution is the same as for Vasicek:

$$B(t,T) = \frac{1}{a}\left(1 - e^{-a(T-t)}\right). \tag{4}$$

The equation for A is different only in the fact that θ is no longer constant; not surprisingly, the θ-dependent part of the solution formula requires doing an integration:

$$A(t,T) = \exp\left[-\int_t^T \theta(s)B(s,T)\,ds - \frac{\sigma^2}{2a^2}(B(t,T) - T + t) - \frac{\sigma^2}{4a}B(t,T)^2\right]. \tag{5}$$

Determining θ from the term structure at time 0. Our goal is to demonstrate that following relation between the infinitesimal forward rate and the function $\theta(t)$:

$$\theta(t) = \frac{\partial f}{\partial T}(0,t) + af(0,t) + \frac{\sigma^2}{2a}(1 - e^{-2at}). \tag{6}$$

When we get to HJM we'll find a simple proof of this relation. But we can also prove it now, by using the explicit representation of $P(t,T)$ given by (3)-(5). Recall that $f(t,T) = -\partial \log P(t,T)/\partial T$. We have

$$-\log P(0,T) = \int_0^T \theta(s)B(s,T)\,ds + \frac{\sigma^2}{2a^2}(B(0,T) - T) + \frac{\sigma^2}{4a}B(0,T)^2 + B(0,T)r_0.$$

Differentiating and using that $B(T,T) = 0$ and $\partial_T B - 1 = -aB$, we get

$$f(0,T) = \int_0^T \theta(s)\partial_T B(s,T)\,ds - \frac{\sigma^2}{2a}B(0,T) + \frac{\sigma^2}{2a}B(0,T)\partial_T B(0,T) + \partial_T B(0,T)r_0.$$

Differentiating again, we get

$$\begin{aligned}\partial_T f(0,T) \quad &= \quad \theta(T) + \int_0^T \theta(s)\partial_{TT}B(s,T)\,ds - \frac{\sigma^2}{2a}\partial_T B(0,T) \\ &\quad + \frac{\sigma^2}{2a}[(\partial_T B(0,T))^2 + B(0,T)\partial_{TT}B(0,T)] + \partial_{TT}B(0,T)r_0.\end{aligned}$$

Combining these equations, and using the fact that $a\partial_T B + \partial_{TT} B = 0$, we get

$$af(0,T) + \partial_T f(0,T) = \theta(T) - \frac{\sigma^2}{2a}(aB + \partial_T B) + \frac{\sigma^2}{2a}[aB\partial_T B + (\partial_T B)^2 + B\partial_{TT} B].$$

Substituting the formula for B and simplifying, we finally get

$$af(0,T) + \partial_T f(0,T) = \theta(T) - \frac{\sigma^2}{2a}(1 - e^{-2aT}),$$

which is equivalent to (6).

A convenient representation. Looking at (6), it seems at first that we must use the *differentiated* term structure $\partial_T f(0,T)$ to calibrate the model. That would be unfortunate, because differentiation amplifies the effect of observation-error. Actually, we can make do with f alone. Indeed, let's look for a representation of the form

$$r(t) = \alpha(t) + x(t) \tag{7}$$

where $\alpha(t)$ is deterministic and $x(t)$ solves

$$dx = -ax\,dt + \sigma\,dw \quad \text{with } x(0) = 0.$$

A brief calculation reveals that if

$$\alpha' + a\alpha = \theta \quad \text{and } \alpha(0) = r_0$$

then $\alpha(t)+x(t)$ solves the SDE for $r(t)$, and has the right initial condition, so (by uniqueness) it equals $r(t)$. The ODE for α is easy to solve: we have $(e^{at}\alpha)' = e^{at}\theta$, so

$$\alpha(t) = r_0 e^{-at} + \int_0^t e^{-a(t-s)}\theta(s)\,ds.$$

Substituting (6) on the right, we get

$$\alpha(t) = r_0 e^{-at} + \int_0^t \partial_s[e^{-a(t-s)} f(0,s)] + \frac{\sigma^2}{2a} e^{-a(t-s)}(1 - e^{-2as})\,ds.$$

This simplifies to

$$\alpha(t) = f(0,t) + \frac{\sigma^2}{2a^2}(1 - e^{-at})^2.$$

Thus the decomposition (7) expresses r as the sum of two terms: a deterministic $\alpha(t)$ reflecting the term structure at time 0, and a random process $x(t)$ that's entirely independent of market data.

Validity of Black's formula. The situation is exactly the same as for Vasicek. The SDE for the interest rate under the forward-risk-neutral measure is

$$dr = [\theta(t) - ar - \sigma^2 B(t,T)]\,dt + \sigma\,d\overline{w}$$

where $d\overline{w}$ is a Brownian motion under this measure. This is simply a version of Hull-White with a different choice of θ. So bond prices are lognormal and Black's formula is valid.

****

The trinomial tree version of Hull-White. This topic is discussed quite clearly in Hull's book *Options, Futures, and other Derivatives* (5th edition), sections 23.11 and 23.12. Please read it there.

Continuous Time Finance Notes, Spring 2004 – Section 6, *March 3, 2004*
Notes by Robert V. Kohn, Courant Institute of Mathematical Sciences. For use in connection with the NYU course Continuous Time Finance.

The first bit of this section addresses the pricing of swaptions – something we didn't get around to before. The rest is a supplement to Hull's discussion of the trinomial tree version of Hull-White.

****

Pricing swaptions using Hull-White. We explained in Section 4 how a caplet is equivalent to a put option on a zero coupon bond. A similar argument shows that a floorlet is equivalent to a call option on a zero coupon bond. So we can easily derive formulas for the prices of caplets and floorlets using Black's formula (problem 1 of HW 3 covers the case of a caplet). Caps and floors are simply portfolios of caplets and floorlets, so we've priced them too. But what about swaptions?

The first observation is general: the task of pricing a swaption is identical to that of pricing a suitable option on a coupon bond. To be specific, let's suppose the underlying swap exchanges the floating rate for fixed rate k, the interest payments being at times $T_1, \ldots, T_N$ with a return of principal at T_N. (The holder of the swap receives the fixed rate and pays the floating rate.) Consider the associated swaption, which gives the holder the right to enter into this swap at time T_0. For simplicity assume the time intervals $T_j - T_{j-1}$ are all the same length Δt. The value of the underlying swap at time T_0 is then

$$P(T_0, T_N) + k\Delta t \sum_{j=1}^{N} P(T_0, T_j) - 1$$

times the notional principal. Indeed, the first term is the value at time T_0 of the principal payment at T_N; the second term is the value at time T_0 of the coupon payments; and the third term is the value at time T_0 of a (short position in a) bond which pays the floating rate. Therefore the payoff of the swaption at time T_0 is

$$(P(T_0, T_N) + k\Delta t \sum_{j=1}^{N} P(T_0, T_j) - 1)_+ \ .$$

This is identical to the payoff of a call option on a coupon bond (with interest rate k and payments at times T_j) with strike 1.

The next observation is special to Hull-White (well, it's a bit more general than that: the argument works for any one-factor short-rate model). We claim that a call or put on a coupon bond is equivalent to a suitable portfolio of calls or puts options on zero coupon bonds. To explain why, let's focus on the case of a call. Recall that $P(t, T) = A(t, T) \exp[-B(t, T)r(t)]$. The key point is that $P(t, T)$ is really a function of three variables:

t, T, and $r(t)$, and it is *monotone* in the third argument $r(t)$. So there is a unique "critical value" r_* for the short rate at time T_0 such that an option with payoff

$$X = \left(P(T_0, T_N) + k\Delta t \sum_{j=1}^{N} P(T_0, T_j) - K \right)_+$$

is in the money (at time T_0) precisely if $r(T_0) < r_*$. Moreover we can write X as a sum of call payoffs on zero-coupon bonds,

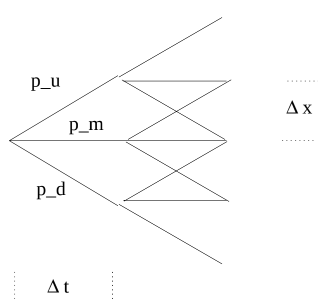
$$X = \left(P(T_0, T_N) - A(T_0, T_N)e^{-B(T_0,T_N)r_*} \right)_+ + k\Delta t \sum_{j=1}^{N} \left(P(T_0, T_j) - A(T_0, T_j)e^{-B(T_0,T_j)r_*} \right)_+$$

since each term is in the money at time T_0 exactly if $r(T_0) < r_*$. Option prices are additive – the value of a portfolio is the sum of the values of its component instruments – and we know how to price options on zero-coupon bonds. So we also know how to price swaptions.

****

The trinomial tree approximation to an additive random walk. You may have been exposed to trinomial trees as a scheme for pricing equity-based options. The setup is in some ways similar to the more familiar binomial trees, except we are not using the trinomial tree to *find* the risk-neutral dynamics; rather we are simply using it to *simulate* the risk-neutral dynamics – or, equivalently, to evaluate expected final-time payoffs by working backward through the tree. (One advantage of the trinomial tree: it amounts to an explicit finite-difference method for the backward Kolmogorov PDE. Second advantage: it lets us be sure there's always a node at a specified value of the stock price; this is important e.g. for valuing barrier options.)

To see the main issues, let's consider the trinomial tree approximation to the scaled Brownian motion process $dx = \sigma\, dw$ with $x(0) = 0$. The tree is determined once we choose Δx, Δt, and the probabilities p_u, p_m, p_d of going up, staying constant, or going down at each node (see the figure). To understand what choices are appropriate, recall that we identify



the continuous-time limit of a (binomial or trinomial) tree by applying the central limit theorem. Therefore the statistics of the continuous-time limit depend only on the means and variances of the one-period increments. Thus to get the continuous-time limit $dx = \sigma\, dw$ we require

$$p_u = p_d$$

so the one-period increment has mean value 0, and

$$p_u(\Delta x)^2 + p_d(\Delta x)^2 = \sigma^2 \Delta t$$

so the one-period increment has variance $\sigma^2 \delta t$. Of course we also require the probabilities all to be positive and

$$p_u + p_m + p_d = 1.$$

Since $p_u = p_d$ let's simplify the notation by calling it p. The condition for the variance requires

$$p = \frac{\sigma^2}{2} \frac{\Delta t}{(\Delta x)^2}$$

and positivity of $p_m = 1 - 2p$ requires

$$\sigma^2 \Delta t / (\Delta x)^2 < 1.$$

Notice that we have a range of possible choices: any p between 0 and 1/2 is OK.

To value an option we work backward in the tree. This is equivalent to finding the expected final-time payoff $E[f(x(T))]$ for the stochastic process associated with the tree. We claim this is equivalent to a standard (explicit) finite-difference approximation of the backward Kolmogorov equation – which in the present setting is $u_t + \frac{1}{2}\sigma^2 u_{xx} = 0$. Indeed, when we work backward in the tree, we determine $u(x,t)$ at time t and nodal value x by

$$u(x,t) = pu(x + \Delta x, t + \Delta t) + pu(x - \Delta x, t + \Delta t) + (1 - 2p)u(x, t + \Delta t),$$

which amounts after some reorganization to

$$\frac{u(x, t + \Delta t) - u(x,t)}{\Delta t} + \frac{1}{2}\sigma^2 \frac{u(x + \Delta x, t + \Delta t) + u(x - \Delta x, t + \Delta t) - 2u(x, t + \Delta t)}{(\Delta x)^2} = 0.$$

We required above that the probabilities in the tree be positive. This is natural, since they are probabilities. But is it also necessary? Yes indeed! This is the condition that the numerical scheme be *stable*. Indeed, if the probabilities are positive then it's easy to see that an option with small final-time payoff has small value. (More: we have the "maximum principle" $E[|f(x_T)|] \leq \max |f|$.) If the probabilities go negative then we lose this property, and working backward in the tree becomes wildly unstable.

Hull likes to take $p = 1/6$ (in other words $\sigma^2 \Delta t = (1/3)(\Delta x)^2$). The reason is that with this choice, the trinomial tree is an unexpectedly good approximation to the PDE (in this constant-volatility, zero-drift setting). To see why, let $u(x,t)$ be the solution of the PDE (with specified final-time value $u(x,T) = f$). By Taylor expansion about the point $(x, t{+}\Delta t)$ we have

$$u(x, t + \Delta t) - u(x,t) = u_t(x, t + \Delta t)\Delta t - \frac{1}{2} u_{tt}(x, t + \Delta t)(\Delta t)^2 + \ldots$$

and

$$u(x+\Delta x, t+\Delta t)+u(x-\Delta x, t+\Delta t)-2u_x(x, t+\Delta t) = u_{xx}(x, t+\Delta t)(\Delta x)^2+\frac{1}{12}u_{xxxx}(x, t+\Delta t)(\Delta x)^4+\ldots$$

Therefore the error terms in our trinomial tree (finite difference) numerical scheme are, at leading order,

$$-(1/2)u_{tt}\Delta t + \frac{\sigma^2}{24}u_{xxxx}(\Delta x)^2.$$

But differentiating the PDE we have $u_{tt} = (1/4)\sigma^4 u_{xxxx}$ so the preceding expression is equal to u_{xxxx} times

$$-\frac{\sigma^4}{8}\Delta t + \frac{\sigma^2}{24}(\Delta x)^2.$$

The special choice $\sigma^2\Delta t = (1/3)(\Delta x)^2$ makes this vanish. For this special choice of parameters, the leading order error is associated with the *next* even term of the Taylor expansion, i.e. it is of order $(\Delta x)^4 \sim (\Delta t)^2$ rather than of order $(\Delta x)^2 \sim \Delta t$.

In the trinomial tree version of Hull-White we have a nonzero drift. I doubt that $p = 1/6$ is any longer so special, because the preceding argument should be disturbed by the drift. But anyway $p = 1/6$ is no worse than any other choice.

****

Why, when doing Hull-White, must we do something different when x is very large or very small? We focus now on the trinomial tree approximation to

$$dx = -ax\,dt + \sigma\,dw$$

with initial data $x(0) = 0$. Suppose we use the standard branching, i.e. the process goes up, stays the same, or goes down with probabilities p_u, p_m, p_d respectively. The condition that the mean be right is now

$$p_u\Delta x - p_d\delta x = -ax\Delta t \tag{1}$$

and the condition that the variance be right is

$$p_u(\Delta x)^2 + p_d(\Delta x)^2 = a^2x^2(\Delta t)^2 + \sigma^2\Delta t. \tag{2}$$

These equations determine p_u and p_d; they in turn determine $p_m = 1 - p_u - p_d$.

Working backward in this tree gives, as before, a finite-difference numerical scheme for solving the backward Kolmogorov PDE, which is now $u_t - axu_x + \frac{1}{2}\sigma^2u_{xx} = 0$. Indeed, when working backward we determine $u(x,t)$ by the formula

$$u(x,t) = p_u u(x+\Delta x, t+\Delta t) + p_m u(x, t+\Delta t) + p_d u(x-\Delta x, t+\Delta t).$$

Since $p_m = 1 - p_u - p_d$ this amounts to

$$[u(x, t+\Delta t)-u(x,t)]+p_u[u(x+\Delta x, t+\delta t)-u(x, t+\Delta t)]+p_d[u(x-\Delta x, t+\delta t)-u(x, t+\Delta t)] = 0.$$

Writing δx for the increment of x (taking values Δx with probability p_u, $-\Delta x$ with probability p_d, and 0 with probability p_m), and using the Taylor expansions

$$u(x+\delta x, t+\Delta t) = u(x) + u_x \delta x + \tfrac{1}{2} u_{xx} (\delta x)^2 + \ldots, \quad u(x, t+\Delta t) = u(x,t) + u_t \Delta t + \ldots$$

we see that

$$u_t \Delta t + u_x E[\delta x] + \tfrac{1}{2} u_{xx} E[(\delta x)^2] = \text{higher order terms.}$$

Thus by getting the mean and variance of δx right, we are designing a finite-difference scheme for the backward Kolmogorov PDE.

Let's make (1) and (2) more explicit. Following Hull, we fix $\Delta x = \sigma\sqrt{3\Delta t}$. After N timesteps there are $2N+1$ spatial nodes, at $x = 0, \pm\Delta x, \ldots, \pm N\Delta x$; therefore we may set $x = j\Delta x$ with $-N \leq j \leq N$. A bit of algebra gives

$$p_u = \frac{1}{6} + \frac{a^2 j^2 (\Delta t)^2 - aj\Delta t}{2}, \quad p_d = \frac{1}{6} + \frac{a^2 j^2 (\Delta t)^2 + aj\Delta t}{2}, \quad p_m = \frac{2}{3} - a^2 j^2 (\Delta t)^2.$$

These probabilities must be *positive*. Writing $\xi = aj\Delta t$, the conditions for positivity are

$$\frac{1}{3} + \xi^2 - \xi > 0, \quad \frac{1}{3} + \xi^2 + \xi > 0, \quad \xi^2 < \frac{2}{3}.$$

The first two inequalities impose no conditions (the polynomials $1/3 + \xi^2 \pm \xi = 0$ have no real roots) but the last one requires

$$aj\Delta t < \sqrt{2/3} \approx .816.$$

We are in trouble after N timesteps, where $aN\delta \approx .816$.

****

Why does it work to truncate the tree? Fixing this problem is surprisingly easy. We truncate the tree at suitably chosen values of j, and use a different trinomial branching scheme at the top and bottom spatial nodes (see the figure, which however truncates the tree much earlier than would normally be done). To see that this works, consider the situation at the top spatial node $j_{\max}$. The branching scheme there (Hull's figure 21.7c) gets the mean and variance right when

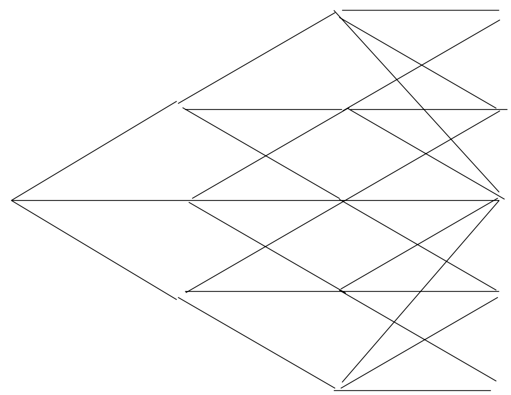
$$(-p_m - 2p_d)\Delta x = -ax\Delta t, \quad p_m (\Delta x)^2 + p_d (2\Delta x)^2 = a^2 x^2 (\Delta t)^2 + \sigma^2 \Delta t$$

which gives (using $\Delta x = \sigma\sqrt{3\Delta t}$ as before, and writing $x = j\Delta x$)

$$p_u = \frac{7}{6} + \frac{a^2 j^2 (\Delta t)^2 - 3aj\Delta t}{2}, \quad p_d = \frac{1}{6} + \frac{a^2 j^2 (\Delta t)^2 - aj\Delta t}{2}, \quad p_m = -\frac{1}{3} - a^2 j^2 (\Delta t)^2 + 2aj\Delta t.$$

For these to be positive we need $\xi = aj\Delta t$ to satisfy

$$\frac{7}{3} + \xi^2 - 3\xi > 0, \quad \frac{1}{3} + \xi^2 - \xi > 0, \quad -\frac{1}{3} - \xi^2 + 2\xi > 0.$$



The first two conditions are always satisfied (the polynomials have no real roots), and the last is satisfied provided ξ lies between $1-\sqrt{2/3} = .184$ and $1+\sqrt{2/3} = 1.816$. It is natural to truncate the tree as early as possible (this minimizes the number of nodes). Therefore we should choose $j_{\max}$ so that

$$aj_{\max}\Delta t \text{ is slightly larger than } .184.$$

The situation at the bottom of the tree is symmetric, so we need not discuss it separately.

****

How do we calibrate the tree version of Hull-White to a given initial term structure? Recall that in continuous time, the solution of

$$dr = (\theta(t) - ar)\,dt + \sigma\,dw, \quad r(0) = r_0$$

was expressed as a sum of two terms: $r = \alpha(t) + x(t)$, where

$$dx = -ax\,dt + \sigma\,dw, \quad x(0) = 0$$

is independent of the initial term structure, and

$$\alpha(t) = f(0,t) + \frac{\sigma^2}{2a^2}\left(1 - e^{-at}\right)^2.$$

This decomposition suggests one way of proceeding: we may model $r(t)$ by the trinomial tree obtained by adding $\alpha(t_j)$ to each nodal value of x at time $t_j = j\Delta t$. This is reasonable – but at finite Δt it only gets the initial term structure *approximately* correct.

A more popular alternative is to choose the interest-rate tree so that it gives exactly the observed values for $P(0, j\Delta t)$ for each $j = 1, 2, 3, \ldots$. We do this by adding well-chosen values $\tilde{\alpha}_j$ to the values of x at times $j\delta t$, $j = 0, 2, 3, \ldots$. The values $\tilde{\alpha}_j$ can be chosen sequentially: choosing $\tilde{\alpha}_0$ to be the initial short rate r_0 assures that the tree predicts $P(0, \Delta t) = e^{-r_0\Delta t}$ as desired; there is a unique value of $\tilde{\alpha}_1$ such that the tree correctly reproduces $P(0, 2\Delta t)$; etc. (This procedure is spelled out in more detail in Hull's book.)

Continuous Time Finance Notes, Spring 2004 – Section 7, *March 10, 2004*
Notes by Robert V. Kohn, Courant Institute of Mathematical Sciences. For use in connection with the NYU course Continuous Time Finance.

Brief announcements concerning the rest of this semester: HW 4 will be available after spring break, and will be due March 31. HW5 will be due April 14. HW6 will be due April 28. The final exam will be in the normal class slot on May 5. You may bring two sheets of your own notes (both sides of each page, any font) to the exam, but you may not use books, my notes, HW solutions, etc.

****

The March 10 lecture covered one-factor HJM, following closely the treatment in Baxter and Rennie. So I'll just outline what was done in class; please read the book for details.

The essential concept of HJM is to model the evolution of the *entire* term structure, rather than to introduce a specific model for the short rate. The advantage of this framework is that the resulting models are consistent – almost by their very definition – with the initial term structure observed in the market. The main disadvantage is that (except for some special cases, which correspond to short rate models like Hull-White) the HJM framework is difficult to calibrate to market data, and difficult to use for the actual pricing and hedging of instruments. Still, the approach is conceptually attractive, providing a general framework for thinking about interest-based instruments analogous to the familiar diffusion-based framework for thinking about equities. So no modern discussion of interest rates could be complete without touching on this topic.

In Section 5.2 Baxter and Rennie discuss the special case when $\sigma(t,T) = \sigma$ is constant:

$$d_t f(t,T) = \alpha(t,T)\,dt + \sigma\,dw. \tag{1}$$

Integrating the SDE leads easily to explicit formulas for $f(t,T)$ and (setting $T = t$) for $r(t)$. Differentiating the latter one gets the SDE for the short rate

$$dr = \left[\partial_T f(0,t) + \alpha(t,t) + \int_0^t \partial_T \alpha(s,t)\,ds\right] dt + \sigma\,dw. \tag{2}$$

Further integrations give explicit formulas for the values of the money-market fund $B(t) = \exp[\int_0^t r(s)\,ds]$ and the bond $P(t,T) = \exp[-\int_t^T f(t,u)\,du]$. Combining these – and using the fact that

$$\int_0^t \int_0^u \alpha(s,u)\,ds\,du = \int_0^t \int_s^t \alpha(s,u)\,du\,ds$$

(since the 2D region of integration is the triangle in the (s,u) plane where $0 \le s \le u$ and $0 \le u \le t$) we get the explicit formula $P(t,T)/B(t) = e^X$ where

$$X = -\int_0^T f(0,u)\,du - \int_0^t \int_s^T \alpha(s,u)\,du\,ds - \sigma(T-t)w(t) - \sigma \int_0^t w(s)\,ds.$$

Notice that

$$d_t X = \left[- \int_t^T \alpha(t,u)\, du \right] dt - \sigma(T-t)\, dw.$$

By Ito we have $d(e^X) = e^X\, dX + \frac{1}{2} e^X\, dX\, dX$, and this gives

$$d[P(t,T)/B(t)] = [P(t,T)/B(t)] \left\{ \left(\tfrac{1}{2}\sigma^2 (T-t)^2 - \int_t^T \alpha(t,u)\, du \right) dt - \sigma(T-t)\, dw \right\}.$$

At this point Baxter and Rennie change to the risk-neutral measure. But it's more transparent in my view to simply assume the original equation holds in the risk-neutral measure. Then the $P(t,T)/B(t)$ must be a martingale, so the drift in the preceding SDE must vanish:

$$\tfrac{1}{2}\sigma^2 (T-t)^2 = \int_t^T \alpha(t,u)\, du.$$

This must be true for all maturities T at once. So we can differentiate in T to conclude that

$$\alpha(t,T) = \sigma^2 (T-t).$$

In particular: the drift term α in (1) is not something we can choose; it is entirely determined by the volatility.

This special case of HJM is equivalent to the limit $a = 0$ limit of Hull-White (which is know as the Ho-Lee model). To see this, observe that the short rate equation (2) reduces to

$$dr = [\partial_T f(0,t) + \sigma^2 t]\, dt + \sigma\, dw.$$

To see this is the limit of Hull-White when $a \to 0$, recall that Hull-White says

$$dr = (\theta(t) - ar)\, dt + \sigma\, dw$$

with

$$\theta(t) = \partial_T f(0,t) + a f(0,t) + \frac{\sigma^2}{2a}(1 - e^{-2at}).$$

The limit of the right hand side as $a \to 0$ is indeed $\partial_T f(0,t) + \sigma^2 t$ as asserted.

****

The situation is not much different for the general one-factor HJM. Start with

$$d_t f(t,T) = \alpha(t,T)\, dt + \sigma(t,T)\, dw \tag{3}$$

instead of (1). Do all the same integrations as before. Now the integrations involving w cannot be done explicitly; but they don't require any new work, because they're formally analogous to the integrations involving α. Starting exactly as for the simple case done above, we find that

$$f(t,T) = f(0,T) + \int_0^t \sigma(s,T)\, dw(s) + \int_0^t \alpha(s,T)\, dw$$

and

$$r(t) = f(t,t) = f(0,t) + \int_0^t \sigma(s,t)\,dw(s) + \int_0^t \alpha(s,t)\,dw$$

whence the short rate SDE is

$$dr(t) = \left[\partial_T f(0,t) + \int_0^t \partial_T \sigma(s,t)\,dw(s) + \alpha(t,t) + \int_0^t \partial_T \alpha(s,t)\,ds\right] dt + \sigma(t,t)\,dw(t).$$

(Note: usually the drift term in this SDE is not a function of t and $r(t)$ alone. Indeed, even if σ and α are deterministic, the drift term still involves a stochastic integral. Therefore the short rate process is usually non-Markovian.)

Further integrations give once again explicit formulas for the values of the money-market fund $B(t) = \exp[\int_0^t r(s)\,ds]$ and the bond $P(t,T) = \exp[-\int_t^T f(t,u)\,du]$. Combining these – and using the fact that

$$\int_0^t \int_0^u \sigma(s,u)\,dw(s)\,du = \int_0^t \int_s^t \sigma(s,u)\,du\,dw(s)$$

we get the explicit formula $P(t,T)/B(t) = e^X$ where

$$X = -\int_0^T f(0,u)\,du - \int_0^t \int_s^T \alpha(s,u)\,du\,ds - \int_0^t \int_s^T \sigma(s,u)\,du\,dw(s).$$

It is convenient to introduce the notation

$$\Sigma(s,T) = -\int_s^T \sigma(s,u)\,du.$$

Then direct calculation gives

$$d_t X = \left[-\int_t^T \alpha(t,u)\,du\right] + \Sigma(t,T)\,dw(t).$$

By Ito we have $d(e^X) = e^X\,dX + \frac{1}{2}e^X\,dX\,dX$, and this gives

$$d[P(t,T)/B(t)] = [P(t,T)/B(t)]\left\{\left(\tfrac{1}{2}\Sigma^2(t,T) - \int_t^T \alpha(t,u)\,du\right) dt + \Sigma(t,T)\,dw\right\}.$$

Let's assume as before that we are working from the start in the risk-neutral measure. Then P/B must be a martingale, so the drift term must vanish:

$$\tfrac{1}{2}\Sigma^2(t,T) = \int_t^T \alpha(t,u)\,du.$$

This holds for every T, so we may differentiate with respect to T. This gives

$$\alpha(t,T) = \Sigma(t,T)\partial_T \Sigma(t,T).$$

Recalling the definition of Σ, this amounts to the statement that

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

****

How can we get the mean-reverting Hull-White model from HJM? Well, from our analysis of Hull-White, we know that its instantaneous forward rates satisfy

$$d_t f(t,T) = (\text{stuff})\, dt + \sigma e^{-a(T-t)}\, dw$$

under the risk-neutral measure. So we can expect one-factor HJM to specialize to Hull-White when $\sigma(t,T) = \sigma_0 e^{-a(T-t)}$ with σ_0 constant. One of the problems on HW4 asks you to verify this.

Continuous Time Finance Notes, Spring 2004 – Section 8, *March 24, 2004*
Notes by Robert V. Kohn, Courant Institute of Mathematical Sciences. For use in connection with the NYU course Continuous Time Finance.

This section begins by wrapping up our discussion of HJM – discussing its strengths and weaknesses. Then we turn to the "Libor Market Model" – which though relatively new is rapidly becoming the method of choice for many purposes. Our discussion of the Libor Market Model follows Hull's treatment (Section 24.3 of the 5th edition).

****

Some comments on the HJM approach to interest rate modelling:

HJM is a "framework" not a "model". When modeling equity-based derivatives, our usual framework is to assume the underlying solves a diffusion process $ds = \mu(s,t)\,dt + \sigma(s,t)\,dw$. The market is complete for any choice of the functions $\mu(s,t)$ and $\sigma(s,t)$ (with some minor restrictions, for example $\sigma(s,t) > 0$). But what to choose for μ and σ? The choice of μ is of course irrelevant for option pricing, because under the risk-neutral measure the underlying satisfies $ds = r\,dt + \sigma(s,t)\,dw$. But the choice of σ is crucial. We commonly choose $\sigma(s,t) = \sigma_0 s$ with σ_0 constant, i.e. we commonly assume the underlying has *lognormal* dynamics. (This is not the only possibility; in a week or two we'll discuss "local vol" models, i.e. the idea of using the skew/smile of implied volatilities to infer an appropriate choice of $\sigma(s,t)$.)

The situation with HJM is analogous. For one-factor HJM, the framework is $d_t f(t,T) = \alpha(t,T)\,dt + \sigma(t,T)\,dw$. Working under the risk-neutral measure (i.e. assuming w is a Brownian motion in the risk-neutral measure) the drift α is completely specified by σ. The choice of σ is again crucial. It is no longer so obvious how to choose it; we saw how to choose $\sigma(t,T)$ to get either Ho-Lee or Hull-White. But many other choices are possible. Calibration to data is not easy, because when σ is not deterministic (e.g. if it depends explicitly on $f(t,T)$) we have no exact pricing formulas.

It is natural to suppose σ is a function of f as well as t and T. Indeed, if σ is a deterministic function of t and T then $f(t,T)$ is Gaussian, so there is a positive probability that it is negative. We tolerated this in Hull-White, and one often tolerates it in HJM for the same reason – namely analytical tractability of the model. However a model that permits negative interest rates cannot be the last word. Amongst short-rate models the simplest way to keep the short rate positive is to use a state-dependent volatility: for example the Cox-Ingersoll-Ross model assumes $dr = (\theta(t) - ar)\,dt + \sigma_0\sqrt{r}\,dw$ (compare this to Hull-White: $dr = (\theta(t) - ar)\,dt + \sigma_0\,dw$). Our discussion of HJM (in particular: our explanation how the volatility determines the drift) did *not* assume σ was deterministic. It applies with no change if, for example, σ has the form $\sigma_0(t,T)f^\alpha$ for some α. [Note however that problems 5 and 6 on HW4 implicitly assume $\sigma(t,T)$ is deterministic.]

It is tempting to suppose f is lognormal, but this leads to problems. The standard method for pricing a cap using Black's formula assumes the forward term rate is lognormal (more on this below). HJM considers the infinitestimal forward rate not the forward term rate, but one might hope it could also be lognormal. However there's a problem. If we assume $\sigma(t,T) = \sigma_0 f(t,T)$, the HJM theory tells us the risk-neutral process has a non-constant drift term

$$\alpha(t,T) = \sigma(t,T)\int_t^T \sigma(t,u)\,du = \sigma_0^2 f(t,T)\int_t^T f(t,u)\,du.$$

Since the drift is random (it depends on f) the solution will not be lognormal. But worse: the drift is *quadratic* in f. As a result, one can show that the associated HJM stochastic differential equation experiences finite-time blowup with positive probability. (Heuristically: it blows up for the same reason that the solution of $dy = y^2 dt$ becomes infinite in finite time; for an honest analysis of this issue for HJM see Section 13.6 of Avellaneda and Laurence.) By the way, other simple hypotheses on σ can also lead to problems. For example if we set $\sigma(t,T) = \sigma_0 f^\alpha(t,T)$ with $0 < \alpha < 1$ we must worry about the possibility that f reaches 0 in finite time.

Multifactor HJM is no more difficult than one-factor HJM. For the reasons summarized just above, it is common to assume $\sigma(t,T)$ is deterministic. But it is also common to permit more than one factor, e.g. to assume

$$d_t f(t,T) = \alpha(t,T)\,dt + \sum_{i=1}^n \sigma_i(t,T)\,dw_i \tag{1}$$

where $w_1, \ldots, w_n$ are independent Brownian motions under the risk-neutral measure. The advantage of a multifactor model over a one-factor model is simple: when there is only one factor, the prices of bonds with different maturities, say $P(t,T)$ and $P(t,S)$, are perfectly correlated. This is obviously not the case in the real world (a change in the short-term interest environment will, in practice, not have an entirely predictable effect on the yields of 30-year treasuries.) Permitting more one factor captures this effect. In practice one usually uses just a few factors (two or three); the coefficients $\sigma_i(t,T)$ may be chosen for analytical tractability (Baxter-Rennie gives an example in Section 5.7), or they may be chosen using a principal-component analysis of historical data (Section 13.4 of Avellaneda-Laurence discusses how this works and what it produces). Whatever the form of σ_i, they determine the drift in (1) by essentially the same argument we used in the one-factor case:

$$\alpha(t,T) = \sum_{i=1}^n \sigma_i(t,T)\int_t^T \sigma_i(t,u)\,du.$$

Pricing using HJM must usually be done computationally, via Monte Carlo simulation. Short-rate trees are not practical (except for a few special cases, like Hull-White) because the short-rate process is usually non-Markovian (so the associated tree cannot recombine). One might consider a tree approximation for the evolution of $f(t,T)$ with T fixed, if e.g. $\sigma(t,T)$ is deterministic; but remember the drift term couples all maturities, so this requires considering as many trees as you have maturities – not a happy thought. By the way, correlations between different maturities are more important when considering swaptions than when considering caps, because a cap is a sum of its caplets (each of which can be valued independently) but a swaption is an option on a basket of payments, occuring at different times – so the correlations between the interest rates at each payment date are relevant. (Our trick for pricing swaptions using Hull-White avoided this issue, but the argument – due to Jamshidian – only works for one-factor short-rate models.)

****

The real problem with HJM however is that it does not make sufficient contact with (a) *market practice*, and (b) *market data*. Indeed: the simplest and most widely-used method of valuing caps is Black's formula. There's plenty of data concerning the implied vols associated with this formula. An ideal theory would be one that's consistent with Black's formula and easy to calibrate using the associated implied vols. (The other widely-quoted instruments are swaptions. One could just as well request a theory that's consistent with Black's formula for swaptions, and easy to calibrate using the associated implied vols. Unfortunately one must choose: there is a version of the Libor Market Model that's consistent with Black's formula for caps, and one that's consistent with Black's formula for swaptions, but none that has both features simultaneously. We'll focus here on caps as the basic instrument.)

The Libor Market Model is the "ideal theory" just requested. It can be viewed as a special case of HJM, but it is much easier to derive it from scratch. As usual, I focus on the one-factor version because it captures all the main ideas in the most transparent form. The following discussion follows Section 24.3 of Hull's 5th edition in both notation and content.

Let's start by reviewing how Black's formula is used to price caplets. A caplet is an option on the term interest rate available in the marketplace at time T_1 for a loan with maturity $T_2 = T_1 + \delta$. The payment is made at the end; therefore if the cap rate is K the holder receives

$$L\delta(R-K)_+$$

at time T_2, where L is the notional principal and R is the term rate, defined by

$$\frac{1}{1+\delta R} = P(T_1, T_2).$$

Since we want to value this instrument at a time $t < T_1$, it is natural to consider the associated *forward* term rate $F(t)$ for lending at T_1 with maturity T_2. It is defined, for any $t \leq T_1$, by

$$\frac{1}{1+\delta F(t)} = \frac{P(t,T_2)}{P(t,T_1)};$$

notice that $F(T_1) = R$. It is immediate from the definition that

$$F(t) = \frac{1}{\delta}\frac{(P(t,T_1) - P(t,T_2)}{P(t,T_2)}$$

so $F(t)$ is the price of a tradeable divided by $P(t, T_2)$. Therefore $F(t)$ is a martingale under the forward measure associated to time T_2. (Indeed: by definition, this measure has the property that the value of any tradeable at time t normalized by $P(t, T_2)$ is a martingale under it.) Thus – always using the forward measure – the mean of $F(t)$ is independent of t. In particular, the mean of $F(T_1)$ is known in advance: it equals $F(t)$ at time t. Since we are working in the forward measure and the option is a tradeable, the option value satisfies

$$\frac{\text{option value at } t}{P(t, T_2)} = \text{expected payoff at time } T_2$$

i.e.

$$\text{option value at } t = P(t, T_2) \cdot L\delta E[(F(T_1) - K)_+].$$

When we use Black's formula, we evaluate the expectation by making the *assumption* that $F(T_1)$ has lognormal statistics (under the forward measure associated with maturity T_2). The expectation is then given by a Black-Scholes-like formula, whose only inputs are the the mean $E[F(T_1)] = F(t)$ and the standard deviation of $\log F(T_1)$. In summary: to be consistent with Black's formula, a theory must predict that each term rate has lognormal statistics under the forward measure associated to its maturity time; nothing else is needed.

OK, now the Libor Market Model. We want to consider many caplets with different maturities; for simplicity let's suppose the maturities of interest are all multiples of a single parameter δ. So the present time is $t = 0$, and we're only interested in term rates for lending at time $t_k = k\delta$ with maturity $t_{k+1} = (k + 1)\delta$. We now have many forward rates; let's distinguish them notationally, writing

$$F_k(t) = \frac{1}{\delta} \frac{(P(t, t_k) - P(t, t_{k+1})}{P(t, t_{k+1})} \tag{2}$$

for the term rate associated with (t_k, t_{k+1}). The basic hypothesis we need for Black's formula to hold is that these forward rates evolve with lognormal statistics; if there is only one factor this means

$$dF_k(t) = \zeta_k(t) F_k(t)\, dw_k \tag{3}$$

for each k, where w_k is Brownian motion under the martingale measure associated with maturity t_{k+1} and $\zeta_k(t)$ is a *deterministic* function of time. (We should of course choose $\zeta_k(t)$ to match the implied volatilities of forward term rates, obtained from market prices of caplets via Black's formula.)

But do we have the right to do this? Our experience with HJM suggests that the answer is yes – in that setting the volatility $\sigma(t, T)$ was ours to choose. But it's hard to be sure when we use a different measure for each forward rate. In HJM we checked consistency by writing everything in the risk-neutral measure; this corresponds of course to using the money-market fund as numeraire. In the present setting the money-market fund isn't a natural object, because we're working with term rates and a discrete set of maturities. Rather, the natural object is the *rolling CD*, which earns interest during the time interval (t_k, t_{k+1}) at term rate $F_k(t_k)$.

Every numeraire has an associated martingale measure. The one associated to the rolling CD has the property that the value of any option divided by the value of the rolling CD is a martingale. Hull calls this the *rolling forward risk-neutral measure.*

Our goal is now to express the evolution of all the forward rates F_k under the rolling-forward risk-neutral measure. To this end, we recall the change-of-numeraire calculation done on page 9 of Section 4: if Q is the risk-neutral measure and N is any tradeable, define σ_N by

$$dN = rN\,dt + \sigma_N N\,dw$$

where w is Brownian motion under the risk-neutral measure Q. Then the equivalent martingale measure Q_N associated with numeraire N is characterized by the property that

$$dw_N = -\sigma_N\,dt + dw$$

is a martingale under Q_N.

In the present setting we are interested in many numeraires, but none of them is the money market fund. No problem: just apply the preceding result to any pair of numeraires and subtract. We find that if M and N are two numeraires, then their martingale measures are related by the fact that

$$dw_N = (\sigma_M - \sigma_N)\,dt + dw_M$$

is a Brownian motion under Q_N if w_M is a Brownian motion under Q_M. Thus: when we change numeraire, we introduce a drift equal to the difference of the volatilities of the two numeraires.

We want to apply this with $M = P(t, t_k)$ and N equal to the rolling CD. Let's write $v_k(t)$ for the volatility of $P(t, t_k)$. What is the volatility of the rolling CD? Well, for any time t there is a unique "next maturity date" $t_{m(t)}$, and the rolling CD provides a known payment on this maturity date $t_{m(t)}$. So the volatility of the rolling CD is precisely the volatility of $P(t, t_{m(t)})$. In short: the volatility of the rolling CD at time t is

$$v_{m(t)}(t), \quad \text{where } m(t) \text{ is the first maturity date} \geq t.$$

OK, let's pull this together. Applying the change of measure result, the SDE (3) becomes

$$dF_k(t) = \zeta_k(t)[v_{m(t)}(t) - v_{k+1}(t)]F_k(t)\,dt + \zeta_k(t)F_k(t)\,dz$$

where z is Brownian motion in the rolling forward risk-neutral world.

Finally, we need to know the relation between $v_{m(t)}(t)$ and $v_{k+1}(t)$. We get it from the relation between bond prices and forward rates (2). Rearranging then taking the logarithm gives

$$\log P(t, t_k) - \log P(t, t_{k+1}) = \log(1 + \delta F_k(t)).$$

Applying Ito's formula, using the definition

$$d_t P(t, t_k) = rP(t, t_k)dt + v_k(t)P(t, t_k)\,dw = (\text{stuff})\,dt + v_k(t)P(t, t_k)\,dz$$

and matching the dz terms we get

$$v_k(t) - v_{k+1}(t) = \frac{\delta F_k(t)\zeta_k(t)}{1+\delta F_k(t)}.$$

Thus the laws (3) can be expressed simultaneously in a single measure – the rolling risk-neutral measure – as

$$dF_k = \zeta_k F_k \left(\sum_{i=m(t)}^{k} \frac{\delta F_i(t)\zeta_i(t)}{1+\delta F_i(t)} \right) dt + \zeta_k F_k dz. \tag{4}$$

This SDE is *equivalent* to the separate equations (3). It expresses the fact that our hypotheses are consistent. Indeed, to satisfy our hypothesis that $dF_k = \zeta_k(t)dw_k$ with w_k a Brownian motion under the forward measure associated with maturity t_k for each k, we need merely solve the single SDE (4), under a single probability measure (associated with the rolling CD). Then the relations (3) are all recovered by reversing the changes of measure done above.

Notice that (4) is simply a discrete analogue of the HJM equation

$$d_t f(t,T) = \alpha(t,T)\,dt + \sigma(t,T)\,dw$$

with

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

and $\sigma(t,T) = \zeta_i(t)f_i(t)$ when $T \approx t_i$. Indeed: formally as $\delta \to 0$ we expect $F_k(t) \approx f(t,t_k)$ since f is the infinitesimal forward rate. And formally as $\delta \to 0$ we have $1+\delta F_i(t) \approx 1$ while

$$\sum_{i=m(t)}^{k} \delta F_i(t)\zeta_i(t) \approx \int_t^T f(t,u)\zeta(t,u)\,du.$$

Thus the Libor Market Model is essentially the discrete analogue of HJM with lognormal forward rates.

Question: HJM with lognormal statistics was problematic due to blowup. Why don't we have the same trouble here? Answer: at finite δ the drift term

$$\zeta_k(t)F_k(t) \sum_{i=m(t)}^{k} \frac{\delta F_i(t)\zeta_i(t)}{1+\delta F_i(t)}$$

is *not* quadratic in F; rather it has linear growth for large values of F_k due to the denominator $1+\delta F_i(t)$. So blowup is not expected (and indeed does not occur).

The fact that the model makes sense is nice of course. But the real point is that it is easy to make this model reproduce the predictions of Black's formula, by simply choosing the functions $\zeta_k(t)$ appropriately. That's how the model is used in practice: choose the $\zeta_k(t)$ (typically piecewise constant) to match market data on caps; then use the theory to price more exotic or complicated instruments for which no Black-like formula is possible.

I recommend reading Hull's section 24.3, which goes much farther than the preceding discussion. In particular he addresses (a) the multifactor version of the model; (b) analytic approximations for the prices of swaptions [useful for calibration, along with data on caps]; (c) some typical applications e.g. to mortgage-backed securities.

Continuous Time Finance Notes, Spring 2004 – Section 9, *March 31, 2004*
Notes by Robert V. Kohn, Courant Institute of Mathematical Sciences. For use in connection with the NYU course Continuous Time Finance.

This section begins the third segment of the course: a discussion of the volatility skew/smile – its sources and consequences. We start with a general discussion of the phenomenon, and its relation to "fat tails," following Hull's Chapter 15. There are three main approaches to modeling the volatility skew/smile quantitatively: (a) local vol models, (b) jump-diffusion models, and (c) stochastic vol models. We have time for just a very brief introduction to each; you'll learn much more in the course Case Studies. Today we focus on local vol models, explaining why it is "in principle" easy but in practice quite difficult to extract a local volatility function from market data on calls. The heart of the matter is "Dupire's equation."

****

The implied volatility skew/smile. Consider call options on an underlying which earns no dividends. We assume the interest rate r is constant. We write

$$C(S_0, K, T) = \text{market price of a call option with strike } K \text{ and maturity } T$$

where S_0 is the spot price and the current time is $t = 0$. Now define

$$C_{BS}(S_0, K, \sigma, T) = \text{Black-Scholes value of the call, using constant volatility } \sigma.$$

Then the *implied volatility* $\sigma_I(S_0, K, T)$ is defined by the equation

$$C(S_0, K, T) = C_{BS}(S_0, K, \sigma_I(S_0, K, T), T).$$

Since the Black-Scholes value of a call is a monotone function of σ, the implied volatility is well-defined. If the constant-vol Black-Scholes model were "correct," i.e. if it gave the actual market values of call options, then σ_I would be constant, independent of S_0, K, and T.

In fact however σ_I is not constant. The "volatility skew/smile" refers to its dependence on K. Typically, for equities, σ_I decreases as K increases. For foreign exchange the typical behavior is different: σ_I is smallest when $K \approx S_0$ so its graph looks like a "smile."

The definition of implied vol depends on the choice of payoff. But if we used puts rather than calls we would get the *same* implied vols, by put-call parity. (We use here the fact that put-call parity is model-independent!).

Hull discusses at length how the skew/smile reflects "fat tails" in the risk-neutral probability distribution. No need to repeat his discussion here. But note what underlies it. Since prices are linear, the value of any option with maturity T is a linear function of its payoff at time T. We recognize from the relation

$$\text{option value} = e^{-rT} E_{RN}[\text{payoff}]$$

that (for a complete market model) this linear relation is expressed by integration against the risk-neutral probability density times e^{-rT}. In other words, if the payoff is $f(S_T)$ and the risk-neutral probability density of S_T at time T (given price S_0 at time 0) is $p(\xi, T; S_0)$ then

$$\text{option value} = e^{-rT} \int_{-\infty}^{\infty} f(\xi) p(\xi, T; S_0)\, d\xi.$$

The volatility smile associated with foreign currency rates reflects the fact that the marketplace uses a p with "fatter tails" than a lognormal distribution. This is reasonable, since empirical studies suggest that in the real world, large moves are more likely than a lognormal distribution would suggest. (Of course the empirical studies are of subjective, not risk-neutral, probabilities; but change of measure changes the drift, not the volatility.) The skew associated with a stock can be attributed to a different effect: the volatility tends to increase when the stock price decreases, since a large decrease in the stock price is a sign that the company is in trouble.

Implied volatilities depend on the time-to-maturity too. Thus we can speak of the "term structure of implied volatility," i.e. the dependence of the skew/smile on maturity.

What's the goal here? Well, it's natural to ask how the basic Black-Scholes model can be changed to make it consistent with the observed prices of calls and puts of all strikes and maturities. More practically: it's natural to ask whether we can use information about the implied vol skew/smile for improved pricing and/or hedging. Concerning pricing: don't expect too much – we use European option prices to deduce the implied volatility, so we can't expect to get anything new about the prices of European options. But we can hope that a model calibrated using European options will do better than basic Black-Scholes for valuing other, more exotic options (Americans, barriers, etc). And we can hope that an improved model will provide improved hedges, even for Europeans.

OK, the goal is clear. But what changes of Black-Scholes should we consider? There are basically three different approaches:

(a) *Local volatility models.* This approach relaxes the hypothesis that σ be constant, permitting it instead to be a (deterministic) function of S and t. It has the advantage of staying very close to the original Black-Scholes framework: the market is still complete, options can still be hedged, etc. One disadvantage: the function $\sigma(S, t)$ is difficult to extract from market data in a stable way.

(b) *Jump-diffusion models.* Why must the underlying execute a diffusion process? It is obvious that unexpected news can make the market react abruptly. This is best modelled by permitting the underlying to jump. This approach has advantage of being intuitively plausible. One disadvantage: it involves two sources of randomness (the jumps and the diffusion). So hedging cannot be achieved by trading the underlying alone.

(c) *Stochastic volatility models.* Returning briefly to (a), let's suppose the underlying follows a diffusion process, i.e. there are no jumps. Why must the volatility be deterministic? Market data are at least consistent with the idea that volatility is itself random. Typically the process used to model volatility is mean-reverting. We are thus led to study a system of two coupled SDE's, for the price of the underlying and the value of the volatility. This approach has the advantage of capturing relatively well the observed time-series behavior of asset dynamics. Again, it has the disadvantage of using two sources of randomness, so hedging cannot be achieved by trading the underlying alone.

Each of these approaches has its proponents; most people believe (b) and (c) are closer to the truth than (a). This is not a purely academic question: the different approaches give, in many cases, different prices (for exotics) and different hedging strategies (even for Europeans). A recent review of jump-diffusion and stochastic vol models, focusing on how well they fit the observed skew/smile and term structure of implied volatility, is: S.R. Das and R.K. Sundaram, *Of smiles and smirks: a term structure perspective,* J. Fin. Quant. Anal. 34 (1999), 211-239 (available through JSTOR).

****

Local volatility models. As noted above, the local vol framework pursues the idea that the only flaw in the Black-Scholes framework was the assumption that σ was constant. Put differently: the local vol framework assumes the underlying solves an SDE of the form

$$dS = rS\,ds + \sigma(S,t)S\,dw \tag{1}$$

under the risk-neutral measure, where $\sigma(S,t)$ is some deterministic function of S and t. The idea is to deduce the form of $\sigma(S,t)$ from market data on European options (e.g. the implied volatilities of calls, of various strikes and maturities). Once we know $\sigma(S,t)$ we can price and hedge options by using familiar means – e.g. by solving the associated Black-Scholes PDE numerically. (For Europeans this would reproduce the prices used to calibrate the model, but it would give us also the hedge; for American or barrier options we would get predictions of prices as well as information on hedging.)

The good news is: the idea is internally consistent. Indeed: if the underlying really *does* follow a diffusion of the form (1), then the prices of European call options (of all strikes and maturities) provide, in principle, sufficient information to determine $\sigma(S,t)$.

The bad news is: this program is rather difficult to pursue in practice. To see the difficulty in its simplest form, suppose $\sigma = \sigma(t)$ depends only on time. Then $S(T)$ is lognormal, and its drift and volatility depend only on $\int_0^T \sigma^2(s)\,ds$. So from the prices of options (of all maturities) we would get the value of $\int_0^T \sigma^2(s)\,ds$ as a function of the maturity T. But to get $\sigma(t)$ itself we would have to *differentiate* the results of our observations with respect to T. Alas, differentiation is very unstable. Real market data is noisy; therefore are our "observations" of $\int_0^T \sigma^2(s)\,ds$ are imperfect and/or noisy. The process of differentiation amplifies the noise a lot – leading, if we're not careful, to large errors in the estimation of σ.

I said "if we're not careful." What does it mean to be careful? Well, suppose you have determined, by noisy measurements, a function $f(T)$ which is your best estimate of $\int_0^T \sigma^2(s)\,ds$. A good way to find $\sigma(s)$ is to minimize, over some reasonable class of candidates, the error between your measurement and the prediction of that candidate. In practice our "measurements" would not be continuous functions of time at all; rather they would be given at selected maturities $\{T_j\}_{j=1}^N$. So one could minimize (numerically)

$$\sum_{j=1}^{N} \left| f(T_j) - \int_0^{T_j} \sigma^2(s)\, ds \right|^2 \tag{2}$$

over an appropriate (not-too-high-dimensional) family of candidate functions $\sigma(s)$, defined e.g. by splines.

Let's turn now to the more general setting where $\sigma = \sigma(S,t)$. Clearly the task of finding $\sigma(S,t)$ from market data will be at least as difficult as the one discussed just a moment ago. Let's nevertheless ask: can we understand the *mathematical form* of the problem in a direct, transparent way – as we achieved above for the case when σ depends only on t? Remarkably, the answer is yes, in the following sense. Suppose the underlying solves (1), and suppose the stock price today is S_0. Let $C(K,T)$ be resulting value of a call with strike K and maturity T. Then C solves *Dupire's equation*

$$C_T - \tfrac{1}{2}\sigma^2(K,T)K^2 C_{KK} + rKC_K = 0, \tag{3}$$

for all $T > 0$ and $K > 0$, with initial condition

$$C(K,0) = (S_0 - K)_+$$

and boundary condition

$$C(0,T) = S_0.$$

Thus: market data on calls gives us the solution of the Dupire PDE, a PDE in "strike and maturity space", whose "diffusion coefficient" is $\frac{1}{2}K^2$ times the unknown function σ^2 evaluated at "stock price" K and "time" T. (Dupire's equation was first derived in B. Dupire, *Pricing with a smile*, Risk 7(1) 18-20, 1994. Analogous binomial-tree discussions were given in E. Derman and I. Kani, *Riding on a smile*, Risk 7(2) 32-39, 1994, and in M. Rubinstein, *Implied binomial trees,* J. Finance 49, 771-818, 1994. The binomial-tree version is presented by Hull in Section 20.4.)

Formally, we can give a formula for σ by reorganizing Dupire's equation:

$$\sigma^2(K,T) = \frac{C_T + rKC_K}{\frac{1}{2}K^2 C_{KK}}.$$

This formula is useless in practice, since our estimates of C_T, C_K, and C_{KK} will be hopelessly inaccurate. But Dupire's equation can still be used to estimate local vol from market data, by using a more robust scheme analogous to (2): see e.g. L. Jiang, Q. Chen, L. Wang, and J. Zhang, *A new well-posed algorithm to recover implied local volatility*, Quant. Finance 3 (2003) 451-457; also Y. Achdou and O. Pironneau, *Volatility smile by multilevel least square*, Int'l J. Theoret. Appl. Finance 5, 619-643, 2002.

Let's explain why Dupire's equation is valid. Easy things first: the boundary and initial conditions are obvious. Indeed, a call with strike $K = 0$ is equivalent to the stock itself, so its value is S_0; and a call with maturity $T = 0$ has value equal to its payoff, namely $(S_0 - K)_+$.

There are two key ingredients to the derivation of Dupire's equation:

(i) The forward Kolmogorov equation for the probability density of the underlying. Let $p(\xi, \tau)$ be the probability that the underlying has value ξ at time τ, given that it has value S_0 at time 0. Then p solves the forward Kolmogorov PDE

$$p_t - (\tfrac{1}{2}\sigma^2(\xi, t)\xi^2 p)_{\xi\xi} + r(\xi p)_\xi = 0$$

for $t > 0$, with a "delta-function" at $\xi = S_0$ as initial data. (I suppose you learned about the forward Kolmogorov PDE in Stochastic Calculus; for a review of this topic, see Section 1 of my PDE for Finance lecture notes.)

(ii) The fact that the second derivative of the call payoff is a delta function. Thus: differentiating the formula

$$C(K, T) = e^{-rT} E[(S - K)_+] = e^{-rT} \int (\xi - K)_+ p(\xi, T)\, d\xi$$

twice with respect to K gives $C_{KK}(K, T) = e^{-rT} p(K, T)$.

Starting from these ingredients, the argument is easy. Differentiating with respect to T the equation $e^{rT} C_{KK}(K, T) = p(K, T)$, we get

$$C_{KKT} + rC_{KK} = e^{-rT} p_T.$$

Combining this with the forward Kolmogorov equation and remembering that $e^{-rT} p(K, T) = C_{KK}$ we get

$$C_{KKT} + rC_{KK} = (\tfrac{1}{2}\sigma^2(K, T)K^2 C_{KK})_{KK} - r(KC_{KK})_K.$$

Integrate once in K to get

$$C_{KT} + rC_K - (\tfrac{1}{2}\sigma^2(K, T)K^2 C_{KK})_K + rKC_{KK} = a(T)$$

where the right hand side is an unknown function of T alone. Now, $KC_{KK} = (KC_K)_K - C_K$, so the preceding equation can be rewritten as

$$C_{KT} - (\tfrac{1}{2}\sigma^2(K, T)K^2 C_{KK})_K + r(KC_K)_K = a(T).$$

Therefore we can integrate again, obtaining

$$C_T - \tfrac{1}{2}\sigma^2(K, T)K^2 C_{KK} + rKC_K = a(T)K + b(T) \tag{4}$$

where $b(T)$ is another unknown function of T alone. Finally, observe that as $K \to \infty$ the value of a call tends to zero – and so do its derivatives C_T, C_K, and C_{KK} – since the probability density $p(\xi, \tau)$ decays as $\xi \to 0$. So the left hand side of (4) tends to zero as $K \to \infty$. Therefore the right hand side must do the same. This implies $a(T) = b(T) = 0$ for all T. Thus finally

$$C_T - \tfrac{1}{2}\sigma^2(K, T)K^2 C_{KK} + rKC_K = 0,$$

which is Dupire's equation.

****

A different viewpoint, which I won't pursue at any length, is the following. Rather than considering only *constant* σ (corresponding to geometric Brownian motion) or considering *arbitrary* $\sigma(S, t)$ (as we have done above), one can consider simple functional forms for $\sigma(S, t)$ with just a few free parameters. Exact option pricing formulas can be given for some classes of such σ. If we're lucky, there will be a choice of the parameters for which the associated smile/skew resembles what is seen in the marketplace. The problem with this approach is that its output depends strongly on the specific class of σ's considered. So one can question whether its predictive value is much greater than the original constant-volatility framework. A good survey of work in this direction can be found in A. Lipton's book *Mathematical Methods for Foreign Exchange: a Financial Engineer's Approach*, World Scientific, 2001.

Continuous Time Finance Notes, Spring 2004 – Section 10, *April 7, 2004*
Notes by Robert V. Kohn, Courant Institute of Mathematical Sciences. For use in connection with the NYU course Continuous Time Finance.

Jump-diffusion models. Merton was the first to explore option pricing when the underlying follows a jump-diffusion model. His 1976 article "Option pricing when underlying stock returns are discontinuous" (reprinted as Chapter 9 of his book *Continuous Time Finance*) is a pleasure to read. My notes follow it – but the article contains much more information than I'm presenting here. Much has happened since 1976 of course; a recent reference is A. Lipton, *Assets with jumps*, *RISK*, Sept. 2002, 149-153.

I will begin with an introduction to jump-diffusions. Then I'll discuss option pricing using such models. This cannot be done using absence of arbitrage alone: when the underlying can jump the market is not complete, since there are two sources of noise (the diffusion and the jumps) but just one tradeable (the underlying). How, then, can we price options? Merton's proposal (still controversial) was to assume that the extra randomness due to jumps is uncorrelated with the market – i.e. its β is zero. This means it can be made negligibile by diversification, and (by the Capital Asset Pricing Model) only the *average* effect of the jumps is important for pricing.

The analogue of the Black-Scholes PDE for a jump-diffusion model is an *integrodifferential* equation. You may wonder how one could ever hope to solve it. In the constant-coefficient setting the Fourier transform is a convenient tool. That's beyond the scope of this course. I've nevertheless included a discussion of the Fourier transform and its use in this setting, as enrichment reading for those who have sufficient background.

****

Jump-diffusion processes. The standard (constant-volatility) Black-Scholes model assumes that the logarithm of an asset price is normally distributed. In practice however the observed distributions are not normal – they have "fat tails," i.e. the probability of a very large positive or negative change is (though small) much larger than permitted by a Gaussian. The jump-diffusion model provides a plausible mechanism for explaining the fat tails and their consequences.

A one-dimensional diffusion solves $dy = \mu\,dt + \sigma\,dw$. (Here μ and σ can be functions of y and t.) A jump-diffusion solves the same stochastic differential equation most of the time, but the solution occasionally jumps.

We need to specify the statistics of the jumps. We suppose the occurrence of a jump is a Poisson process with rate λ. This means the jumps are entirely independent of one another. Some characteristics of Poisson processes:

(a) The probability that a jump occurs during a short time interval of length Δt is $\lambda\Delta t + o(\Delta t)$.

(b) The probability of two or more jumps occuring during a short time interval of length Δt is negligible, i.e. $o(\Delta t)$.

(c) The probability of exactly n jumps occuring in a time interval of length t is $\frac{(\lambda t)^n}{n!}e^{-\lambda t}$.

(d) The mean waiting time for a jump is $1/\lambda$.

We also need to specify what happens when a jump occurs. Our assumption is that a jump takes y to $y + J$. The jump magnitudes are independent, identically distributed random variables. In other words, each time a jump occurs, its size J is selected by drawing from a pre-specified probability distribution.

This model is encapsulated by the equation

$$dy = \mu\, dt + \sigma\, dw + JdN$$

where N counts the number of jumps that have occurred (so it takes integer values, starting at 0) and J represents the random jump magnitude. Ito's Lemma can be extended to this setting: if $v(x,t)$ is smooth enough and y is as above then $v(y(t),t)$ is again a jump-diffusion, with

$$d[v(y(t),t)] = (v_t + \mu v_x + \tfrac{1}{2}\sigma^2 v_{xx})dt + \sigma v_x dw + [v(y(t)+J,t) - v(y(t),t)]dN.$$

All the terms on the right are evaluated at $(y(t),t)$, as usual. In writing the jump term, we're trying to communicate that while the occurence of a jump in $v(y(t),t)$ is determined by N (i.e. by the presence of a jump in y) the size of the jump depends on $y(t)$ and the form of v.

Now consider the expected final-time payoff

$$u(x,t) = E_{y(t)=x}\left[w(y(T))\right]$$

where $w(x)$ is an arbitrary "payoff" (later it will be the payoff of an option). It is described as usual by a backward Kolmogorov equation

$$u_t + \mathcal{L}u = 0 \text{ for } t < T, \quad \text{with} \quad u(x,T) = w(x) \text{ at } t = T. \tag{1}$$

The operator $\mathcal{L}$ for our jump-diffusion process is

$$\mathcal{L}u = \mu u_x + \tfrac{1}{2}\sigma^2 u_{xx} + \lambda E\left[u(x+J,t) - u(x,t)\right].$$

The expectation in the last term is over the probability distribution of jumps. The proof of (1) follows the standard strategy used for diffusions without jumps (see e.g. Section 1 of my PDE for Finance lecture notes for a review of this topic). Let u solve (1), and apply Ito's formula. This gives

$$\begin{aligned} u(y(T),T) - u(x,t) \;&=\; \int_0^T (\sigma u_x)(y(s),s)\, dw + \int_0^T (u_s + \mu u_x + \tfrac{1}{2}\sigma^2 u_{xx})(y(s),s)\, ds \\ &\quad + \int_0^T [u(y(s)+J,s) - u(y(s),s)]dN. \end{aligned}$$

Now take the expectation. Only the jump term is unfamiliar; since the jump magnitudes are independent of the Poisson jump occurence process, we get

$$E\left([u(y(s)+J,s) - u(y(s),s)]dN\right) = E\left([u(y(s)+J,s) - u(y(s),s)]\right)\lambda ds.$$

Thus when u solves (1) we get

$$E[u(y(T),T)] - u(x,t) = 0.$$

This gives the result, since $u(y(T),T) = w(y(T))$ from the final-time condition on u.

A similar argument shows that the discounted final-time payoff

$$u(x,t) = E_{y(t)=x}\left[e^{-r(T-t)}w(y(T))\right]$$

solves

$$u_t + \mathcal{L}u - ru = 0 \text{ for } t < T, \quad \text{with} \quad u(x,T) = w(x) \text{ at } t = T,$$

using the same operator $\mathcal{L}$.

What about the probability distribution? It solves the forward Kolmogorov equation,

$$p_s - \mathcal{L}^* p = 0 \text{ for } s > 0, \quad \text{with} \quad p(z,0) = p_0(z)$$

where p_0 is the initial probability distribution and $\mathcal{L}^*$ is the adjoint of $\mathcal{L}$. (See Section 1 of my PDE for Finance lecture notes for an explanation why this must be so.) What is the adjoint of the new jump term? For any functions $\xi(z), \eta(z)$ we have

$$\int_{-\infty}^{\infty} E[[\xi(z+J) - \xi(z)]\,\eta(z)\,dz = \int_{-\infty}^{\infty} \xi(z)E[[\eta(z-J) - \xi(z)]\;dz =$$

since $\int_{-\infty}^{\infty} E[\xi(z+J)]\eta(z)\,dz = \int_{-\infty}^{\infty} \xi(z)E[\eta(z-J)]\,dz$. Thus

$$\mathcal{L}^* p = \tfrac{1}{2}(\sigma^2 p)_{zz} - (\mu p)_z + \lambda E\left[p(z-J) - p(z)\right],$$

i.e. the probability distribution satisfies

$$p_s - \tfrac{1}{2}(\sigma^2 p)_{zz} + (\mu p)_z - \lambda E\left[p(z-J,s) - p(z,s)\right] = 0.$$

****

Hedging and the risk-neutral process. I'd like to say that the time-0 value of an option with payoff $w(S)$ should be the discounted final-time payoff under "the risk-neutral dynamics." This is not obvious; indeed, it is a modeling hypothesis not consequence of arbitrage. I now explain the meaning and logic of this hypothesis, following Merton.

It is convenient to focus on the stock price itself not its logarithm. If (as above) J represents a typical jump of $\log S$ then the stock price leaps from S to $e^J S$, i.e. the jump in stock price is $(e^J - 1)S$. So (applying the jump-diffusion version of Ito to $S = e^y$, with $dy = \mu dt + \sigma dw$) the stock dynamics is

$$dS = (\mu + \tfrac{1}{2}\sigma^2)Sdt + \sigma S dw + (e^J - 1)SdN.$$

The associated risk-neutral process is determined by two considerations:

(a) it has the same volatility and jump statistics – i.e. it differs from the subjective process only by having a different drift; and

(b) under the risk-neutral process $e^{-rt}S$ is a martingale, i.e. $dS - rSdt$ has mean value 0.

We easily deduce that the risk-neutral process is

$$dS = (r - \lambda E[e^J - 1])Sdt + \sigma Sdw + (e^J - 1)SdN. \tag{2}$$

Applying Ito once more, we see that under the risk-neutral dynamics $y = \log S$ satisfies

$$dy = (r - \tfrac{1}{2}\sigma^2 - \lambda E[e^J - 1])dt + \sigma dw + JdN$$

Thus the formalism developed in the preceding subsection can be used to price options; we need only set $\mu = r - \frac{1}{2}\sigma^2 - \lambda E[e^J - 1]$.

But is this right? And what are its implications for hedging? To explain, let's examine what becomes of the standard demonstration of the Black-Scholes PDE in the presence of jumps. Assume the option has a well-defined value $u(S(t), t)$ at time t. Suppose you try to hedge it by holding a long position in the option and a short position of Δ units of stock. Then over a short time interval the value of your position changes by

$$\begin{aligned} d[u(S(t),t)] - \Delta dS &= u_t dt + u_S([\mu + \tfrac{1}{2}\sigma^2]Sdt + \sigma Sdw) + \tfrac{1}{2}u_{SS}\sigma^2 S^2 dt \\ &\quad + [u(e^J S(t), t) - u(S(t), t)]dN \\ &\quad - \Delta([\mu + \tfrac{1}{2}\sigma^2]Sdt + \sigma Sdw) - \Delta(e^J - 1)SdN. \end{aligned}$$

There are two sources of randomness here – the Brownian motion dw and the jump process dN – but only one tradeable. So the market is incomplete, and there is no choice of Δ that makes this portfolio risk-free.

But consider the choice $\Delta = u_S(S(t), t)$. With this choice the randomness due to dw cancels, leaving only the uncertainty due to jumps:

$$\text{change in portfolio value} = (u_t + \tfrac{1}{2}\sigma^2 S^2 u_{SS})dt + \{[u(e^J S(t), t) - u(S(t), t)] - u_S(e^J S - S)\}dN.$$

To make progress, we must assume something about the statistics of the jumps. Merton's suggestion (still controversial) was to assume they are uncorrelated with the marketplace. The impact of such randomness can be eliminated by diversification. Put differently: according the the Capital Asset Pricing Model, for such an investment (whose β is zero) only the mean return is relevant to pricing. So the mean return on our hedge portfolio should be the risk-free rate:

$$(u_t + \tfrac{1}{2}\sigma^2 S^2 u_{SS})dt + \lambda E[u(e^J S(t), t) - u(S(t), t) - (e^J S - S)u_S]dt = r(u - Su_S)dt. \tag{3}$$

After rearrangement, this is precisely the backward Kolmogorov equation describing the discounted final-time payoff under the risk-neutral dynamics (2):

$$u_t + (r - \lambda E[e^J - 1])Su_S + \tfrac{1}{2}\sigma^2 S^2 u_{SS} - ru + \lambda E[u(e^J S, t) - u(S, t)] = 0.$$

A final remark about the experience of the investor who follows this hedge rule. If the option value is convex in S (as we expect for a call or a put) then the term in (3) associated with the jumps is positive:

$$E[u(e^J S(t), t) - u(S(t), t) - (e^J S - S)u_S] \geq 0.$$

So in the absence of jumps the value of the hedge portfolio (long the option, short $\Delta = u_S$ units of stock) rises a little slower than the risk-free rate. Without accounting for jumps, the investor who follows this hedge appears to be falling behind (relative to a cash investment at the risk-free rate). But due to convexity the net effect of the jumps is favorable – exactly favorable enough that the investor's long-term (mean) experience is risk-free.

****

Let's review why we're doing this. Certainly not for pricing European options, which are relatively liquid – their prices are visible in the marketplace. Interpreting their prices using standard Black-Scholes (deducing an implied volatility from the option price) one obtains a result in contradiction to the model: the implied volatility depends on maturity and strike price (the dependence on strike price is often called the "volatility smile"). The introduction of jumps provides a plausible family of models that's better able to fit the market data. But it introduces headaches of modeling and calibration (e.g. how to choose the distribution of jumps?). If the goal were merely to price Europeans, there would be no reason to bother – their prices are visible in the marketplace.

So why are we doing this? Three reasons. One is the desire for a consistent theoretical framework. The second, more practical, is the need to hedge (not simply to price) options – the Delta predicted by a jump-diffusion model is different from that of the Black-Scholes framework. A third reason is the need to price and hedge exotic options (e.g. barriers) which are less liquid. The basic idea: calibrate your jump-diffusion model using the market prices of European options, then use it to price and hedge barrier options.

****

Solution via Fourier transform. To assess what a jump-diffusion model says about "fat tails" we need a scheme for solving the forward Kolmogorov equation. And to value options (e.g. to calibrate the model to market prices) we need a scheme for solving the backward Kolmogorov equation. In the constant-coefficient setting these integrodifferential equations can be solved using the Fourier transform. The rest of these notes explain how. (This material is offered for enrichment purposes only; it will not be presented in lecture, and will not be required for homeworks or the exam.)

We henceforth focus on

$$dy = \mu dt + \sigma dw + J dN$$

with μ and σ *constant*. To be specific let's focus on the *forward* Kolmogorov equation, which is now

$$p_s - \tfrac{1}{2}\sigma^2 p_{zz} + \mu p_z - \lambda E\left[p(z - J, s) - p(z, s)\right] = 0 \tag{4}$$

and let us solve it with initial condition $p(0) = \delta_{z=0}$. (This will give us the fundamental solution, i.e. $p(z, s)$ = probability of being at z at time s given that you started at 0 at time 0.)

Why is the Fourier transform useful in this setting? Basically, because the forward equation is a mess – nonlocal in space (due to the jump term) and involving derivatives too (the familiar terms). But when we take its Fourier transform in space we get a simple, easy-to-solve ODE. The result is a simple expression not for the probability distribution itself, but rather for its Fourier transform – what a probabilist would call the *characteristic function* of the distribution.

Most students in this class will be relatively unfamiliar with the Fourier transform. Here's a brief summary of what we'll use:

(a) Given a function $f(x)$, its Fourier transform $\hat{f}(\xi) = \mathcal{F}[f](\xi)$ is defined by

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x) e^{ix\xi}\, dx.$$

Notice that even if f is real-valued, $\hat{f}$ is typically complex-valued.

(b) Elementary manipulation reveals that the translated function $f(x - a)$ has Fourier transform

$$\mathcal{F}[f(x-a)](\xi) = e^{i\xi a}\hat{f}(\xi)$$

(c) Integration by parts reveals that the Fourier transform takes differentiation to multiplication:

$$\mathcal{F}[f_x](\xi) = -i\xi\hat{f}(\xi)$$

(d) It is less elementary to prove that the Fourier transform takes convolution to multiplication: if $h(x) = \int_{-\infty}^{\infty} f(x-y)g(y)\, dy$ then

$$\hat{h}(\xi) = \hat{f}(\xi)\hat{g}(\xi).$$

(e) Another less elementary fact is Plancherel's formula:

$$\int_{-\infty}^{\infty} \overline{f} g\, dx = \frac{1}{2\pi}\int_{-\infty}^{\infty} \overline{\mathcal{F}[f]}\mathcal{F}[g]\, d\xi$$

where $\overline{f}$ is the complex conjugate of f.

(f) The Fourier transform is invertible, and

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ix\xi}\hat{f}(\xi)\, d\xi$$

(g) The Fourier transform of a Gaussian is again a Gaussian. More precisely, for a centered Gaussian with variance σ^2,

$$\mathcal{F}\left[\frac{1}{\sqrt{2\pi}\sigma}e^{-x^2/2\sigma^2}\right] = e^{-\sigma^2\xi^2/2}.$$

OK, let's use this tool to solve the forward equation (4). We apply the Fourier transform (in space) to each term of the equation. The first three terms are easy to handle using property (c). For the jump term, let the jump J have probability density ω, that

$$E\left[p(z-J)-p(z)\right] = \int [p(z-J)-p(z)]\omega(J)\,dJ = \int p(z-J)\omega(J)\,dJ - p(z).$$

By property (d) this the Fourier transform of this term is $(\hat{\omega}-1)\hat{p}$. Thus in the Fourier domain the equation becomes

$$\hat{p}_s + \tfrac{1}{2}\sigma^2\xi^2\hat{p} - i\xi\mu\hat{p} - \lambda(\hat{\omega}-1)\hat{p} = 0$$

with initial condition $\hat{p}(0) = \int e^{i\xi z}\delta_{z=0}\,dz = 1$. Writing the equation in the form

$$\hat{p}_s = K(\xi)\hat{p}$$

with

$$K(\xi) = -\tfrac{1}{2}\sigma^2\xi^2 + i\xi\mu + \lambda(\hat{\omega}(\xi)-1)$$

we recognize immediately that the solution is

$$\hat{p}(\xi, s) = e^{sK(\xi)}. \tag{5}$$

The probability density itself can of course be obtained by taking the inverse Fourier transform:

$$p(z,s) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-iz\xi + sK(\xi)}\,d\xi. \tag{6}$$

Option pricing. What use is this? Well, one can explore what the tails of the distribution look like, as you vary the jump rate λ and jump distribution ω. Or one can use it to price options. Recall that the value of an option should be the discounted expected payoff relative to the risk-neutral probability distribution. Assume for the moment that the under the risk-neutral distribution the stock price is e^y where y is a jump-diffusion as above. Then the time-0 value of an option with payoff w is

$$e^{-rT}E_{y(0)=\ln S_0}[w(e^y)]$$

if the time-0 stock price is S_0. If $w(S)$ vanishes for S near 0 and ∞ (corresponding to y very large negative or positive) then the option price is easy to express. Let $v(y) = w(e^y)$ be the payoff as a function of y, and let $x = \ln S_0$. By translation invariance, the probability density of y is $p(z-x,T)$ where p is given by (6). Therefore

$$\begin{aligned} E_{y(0)=\ln S_0}[w(e^y)] &= \int p(z-x,T)v(z)\,dz \\ &= \frac{1}{2\pi}\int \overline{\mathcal{F}[p(z-x,T)]}\mathcal{F}[v]\,d\xi \\ &= \frac{1}{2\pi}\int e^{-i\xi x}\hat{p}(-\xi,T)\hat{v}(\xi)d\xi. \end{aligned} \tag{7}$$

In the last step we used that $\mathcal{F}[p(z-x,T)] = e^{i\xi x}\hat{p}(\xi,T)$, and the fact that the complex conjugate of $\hat{p}(\xi,T) = \int e^{i\xi z}p(z,T)\,dz$ is $\int e^{-i\xi z}p(z,T)\,dz = \hat{p}(-\xi,T)$ since p is real. Equation (7) reduces the task of option pricing to the calculation of two Fourier transforms (those of ω and v) followed by a single integration (7).

The hypothesis that the payoff $w(S)$ vanishes near $S = 0$ and $S = \infty$ is inconvenient, because neither a put nor a call satisfies it. Fortunately this hypothesis can be dispensed with. Consider for example the call $w(S) = (S - K)_+$, for which $v(y) = (e^y - K)_+$. Its Fourier transform is not defined on the real axis, because the defining integral diverges as $y \to \infty$. But $e^{-\alpha y}v(y)$ decays near ∞ for $\alpha > 1$. So its Fourier transform in y is well-defined. This amounts to examining the Fourier transform of v along the line $\Im\xi = \alpha$ (here $\Im\xi$ is the imaginary part of ξ) since

$$\mathcal{F}[e^{-\alpha y}v(y)](\xi_1) = \int_{-\infty}^{\infty} e^{i(\xi_1+i\alpha)y}v(y)\,dy.$$

Fortunately, the Plancherel formula isn't restricted to integrating along the real axis in Fourier space; one can show that

$$\int_{-\infty}^{\infty} \overline{f}g\,dx = \frac{1}{2\pi}\int_{-\infty}^{\infty} \overline{\mathcal{F}[f]}(\xi_1+i\alpha)\mathcal{F}[g](\xi_1+i\alpha)\,d\xi_1$$

if the Fourier transforms of f and g exist (and are analytic) at $\Im\xi = \alpha$. Using this, an argument similar to that given above shows that

$$E_{y(0)=\ln S_0}[w(e^y)] = \int p(z-x,T)v(z)\,dz = \frac{1}{2\pi}\int e^{i(-\xi_1+i\alpha)x}\hat{p}(-\xi_1+i\alpha,T)\hat{v}(\xi_1+i\alpha)d\xi_1.$$

By the way, in the case of the call the Fourier transform of v is explicit and easy:

$$\hat{v}(\xi) = \int_{\ln K}^{\infty} e^{iy\xi}(e^y - K)\,dy = -\frac{K^{1+i\xi}}{\xi^2 - i\xi}$$

by elementary integration. Here $\xi = \xi_1 + i\alpha$ is any point in the complex plane such that the integral converges (this requires $\alpha > 1$).

Continuous Time Finance Semester Review, Spring 2004

The final exam will be Wednesday May 5, in the usual class location and timeslot. You may use two sheets of notes (any font, both sides) but no other books, notes, or materials. Expect the exam to be about 1/3 on Segment 1, 1/2 on Segment 2, and 1/6 on Segment 3. All material covered through 4/14 or on any homework is fair game; there will be no questions on stochastic volatility (the 4/21 lecture). This Semester Review is intended to help you see the "big picture" and to give you some idea what kind of questions might be on the exam. (Not everything here will be on the exam, and there may be things on the exam which are not mentioned here.)

SEGMENT 1: Continuous-time methods in the equity-based setting (Sections 1-3 and Homeworks 1-2)

Topics included: understanding pricing and hedging via the Black-Scholes PDE and via the martingale representation theorem; understanding the market price of risk and its relation to the existence and uniqueness of the risk-neutral measure; forward measures associated with risky numeraires. Applications including pricing and hedging of exchange options and quantos.

Possible exam questions: (a) Price or hedge an option similar to one on the homework. (b) What is the SDE for the dollar-to-pound exchange rate under the dollar investor's RN measure? What is it under the pound investor's RN measure? (c) Explain why all tradeables must have the same market price of risk. (d) Consider two tradeables whose prices S and N solve the SDE's .... What is the SDE for S under the forward risk neutral measure associated with N?

SEGMENT 2: Interest-based derivatives (Sections 4-8 and Homeworks 3-5)

Topics related to Vasicek-Hull-White short-rate model included: pricing and hedging via PDE's (Feynman-Kac formula); derivation and use of $P(t,T) = A(t,T)\exp[-B(t,T)r(t)]$; lognormality of $P(t,T)$; capacity to match any term structure; pricing and hedging of options on bonds via Black's formula; relevance of this for caplets and swaps; trinomial trees. Also discussed more advanced models: Heath-Jarrow-Morton and the Libor Market Model.

Possible exam questions: (a) Price or hedge an option similar to one on the homework. (b) Consider the following formula for the value of a caplet ...: explain how it amounts to a formula for the solution of a certain PDE. (c) Explain why, in the context of the Hull-White model, $P(t,T)$ is lognormal. (d) What SDE does $P(t,T)$ satisfy under the RN measure? What about under the forward risk-neutral measure? (e) When constructing a trinomial tree for the SDE $dx = -ax\,dt + \sigma\,dw$ we usually use .... What system of linear equations should be solved to find the probabilities p_u, p_m, and p_d? Why is it necessary to truncate the tree? (f) Explain the statement "HJM models usually lead to non-Markovian short-rate processes."

SEGMENT 3: Approaches to the volatility skew/smile (Sections 9-10 and Homework 6)

Topics included: relevance of "fat tails"; simple formula for implied vol when σ depends only on t; the Dupire equation; jump-diffusion models. [Stochastic vol too but it won't be on the exam.]

Possible exam questions: (a) Like problem 2 of HW6, (b) Consider a jump diffusion model in which all jumps are positive: what would the implied vol skew/smile look like? (c) Explain why $c_{KK}(K,T) = e^{-rT}p(K,T)$.

Continuous Time Finance, Spring 2004 – Homework 1
Distributed 1/28/04, due 2/4/04

(1) In the Section 1 notes, we proved that if V solves the Black-Scholes PDE with final-value f, then $V(S_0, 0) = e^{-rT}E[f(S_T)]$ where S solves the SDE $dS = rS\,dt + \sigma S\,dw$ with initial value $S(0) = S_0$. Let's do something similar for a stochastic interest rate. Suppose the spot rate r_t solves a diffusion of the form $dr = \alpha\,dt + \beta\,dw$ with $r(0) = r_0$, where $\alpha = \alpha(r,t)$ and $\beta = \beta(r,t)$ are fixed functions of r and t. Consider the function $U(r,t)$ defined by solving $U_t + \alpha U_r + \frac{1}{2}\beta^2 U_{rr} - rU = 0$ with final value $U(r,T) = 1$. Show that

$$U(r_0, 0) = E\left[e^{-\int_0^T r(s)\,ds}\right].$$

[Comment: if the SDE for r is the risk-neutral process, then $U(r_0, 0)$ is the value of a zero-coupon bond that pays one dollar at time T. Hint: show that $U(r(t), t)\exp\left(-\int_0^t r(s)\,ds\right)$ is a martingale.]

(2) Consider a non-dividend-paying stock whose share price satisfies $dS = \mu S\,dt + \sigma S\,dw$, and assume for simplicity that the risk-free rate r is constant. Consider a European option with maturity T and payoff $f(S_T)$. We now have two apparently different ways to price and hedge it:

(a) Using the Black-Scholes PDE. The value at time t is $V(S_t, t)$ where V solves the Black-Scholes PDE with final-value f, and the hedge portfolio consists of $\phi_t = \frac{\partial V}{\partial S}(S_t, t)$ stock and $(V(S_t, t) - \phi_t S_t)/B_t$ units of the risk-free bond whose value at time t is B_t.

(b) Using the Girsanov's theorem and the martingale representation theorem. This means we must find a risk-neutral measure Q with respect to which S_t/B_t is a martingales; then the option value at time t is $V_t = B_t E_Q[f(S_T)/B_T \mid \mathcal{F}_t]$, and the hedge ratio ϕ_t is determined by the martingale representation theorem, which tells us that $d(V/B) = \phi_t\, d(S/B)$ for some ϕ_t.

Show these two frameworks are consistent. In other words, show that the value and hedge defined by (a) satisfy the properties asserted by (b).

(3) Consider the discussion in the Section 2 notes concerning options on foreign exchange rates.

(a) What PDE should the dollar investor solve to value an option with payoff $f(C_T)$? How does it determine the hedge portfolio?

(b) What PDE does the pound investor solve to value the same option? How is his hedge portfolio related to that of the dollar investor?

(c) Use these results to give another proof that the two investors price the option consistently.

Continuous Time Finance, Spring 2004 – Homework 2
Distributed 2/4/04, due 2/18/04

(1) Consider a market with one source of randomness, a scalar Brownian motion w. Suppose the price S of a stock satisfies

$$dS = \mu(S,t)S\,dt + \sigma(S,t)S\,dw$$

where μ and σ are known functions of S and t. Let r be the risk-free rate, assumed constant for simplicity. We know that all tradeables must have the same market price of risk

$$\lambda = \frac{\mu - r}{\sigma}.$$

Let's check this for the special case of a European option on S with maturity T and payoff $f(S_T)$. Use the Black-Scholes PDE to verify that this option's market price of risk is the same as that of the underlying.

(2) Let S_1 and S_2 be stocks with constant drift and volatility

$$dS_1 = \mu_1 S\,dt + \sigma_1 S\,dw_1, \quad dS_2 = \mu_2 S\,dt + \sigma_2 S\,dw_2,$$

and assume that w_1 and w_2 have (constant) correlation ρ. The risk-free rate is r (also constant). The Section 3 notes discuss the pricing of an exchange option with payoff $(S_2(T) - S_1(T))_+$. Find the trading strategy that replicates this payoff. (Give your answers in terms of cumulative normal distribution functions.)

(3) Now let's consider the same two stocks as in Problem 2, but an option with a general payoff $f(S_1(T), S_2(T))$.

(a) Identify the risk-neutral SDE for S_1 and S_2.

(b) Show that the value of the option at time t is $V(S_1(t), S_2(t), t)$ where $V(x,y,t)$ solves

$$V_t + rxV_x + ryV_y + \tfrac{1}{2}\sigma_1^2 x^2 V_{xx} + \tfrac{1}{2}\sigma_2^2 y^2 V_{yy} + \rho\sigma_1\sigma_2 xyV_{xy} - rV = 0$$

with final-time condition

$$V(x,y,T) = f(S_1, S_2, T).$$

(c) Show that if the payoff has the special form $g(S_2/S_1)S_1$ it suffices to solve the simpler, one-space-dimensional PDE $U_t + \frac{1}{2}\hat{\sigma}^2 x^2 U_{xx} = 0$ with final-time condition $U(x,T) = g(x)$. (Here, as in Section 3, $\hat{\sigma} = (\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)^{1/2}$.) How does U determine the value of the option?

(4) Let S be the price of a stock in pounds, and let C be the exchange rate in dollars per pound. Assume

$$dS = \mu_S S\,dt + \sigma_S S\,dw_S, \quad dC = \mu_C C\,dt + \sigma_C C\,dw_C,$$

where the drifts and volatilities are constant. Let r and u be the risk-free rates in dollars and pounds respectively, also constant. Consider a quanto call, whose payoff in dollars is $(S_T - K)_+$. The Section 3 notes discuss how to price it. Find the trading strategy that replicates it. (Give your answers in terms of cumulative normal distribution functions.)

(5) For the same stock and exchange rate processes as in Problem 4, consider a general quanto whose payoff in dollars is $f(s_T)$. What PDE should you solve to price it? How does the solution of the PDE determine a dollar investor's replicating portfolio?

Continuous Time Finance, Spring 2004 – Homework 3
Distributed 2/27/04, due 3/10/04

(1) Assume the Vasicek model $dr = (\theta - ar)\,dt + \sigma\,dw$ for the risk-neutral short rate process. Consider a call option with maturity T and strike K, on a zero-coupon bond with maturity $S > T$. Its payoff at time T is $(P(T,S) - K)_+$. Show using Black's formula that the value of this option at time t is

$$P(t,S)N(d_1) - KP(t,T)N(d_2)$$

where

$$d_1 = \frac{1}{\sigma_p} \log \frac{P(t,S)}{P(t,T)K} + \tfrac{1}{2}\sigma_p, \quad d_2 = \frac{1}{\sigma_p} \log \frac{P(t,S)}{P(t,T)K} - \tfrac{1}{2}\sigma_p$$

with

$$\sigma_p = \sigma \left(\frac{1 - e^{-2a(T-t)}}{2a} \right)^{1/2} B(T,S).$$

(The function $B(T,S)$ is the one from the representation $P(t,T) = A(t,T)e^{-B(t,T)r(t)}$.)

(2) Since Vasicek is a one-factor model, the call option of Problem 1 can be replicated by a self-financing trading strategy using any pair of tradeables.

(a) What trading strategy produces a replicating portfolio using tradeables $P(t,T)$ and $P(t,S)$?

(b) What trading strategy produces a replicating portfolio using tradeables $P(t,S)$ and the money market fund $B(t)$?

(c) What trading strategy produces a replicating portfolio using two bonds $P(t,T_1)$ and $P(t,T_2)$, where T_1 and T_2 are arbitrary (distinct) times greater than T?

(3) The call option of Problem 1 can be viewed as an exchange option involving the zero-coupon bonds $S_1 = KP(t,T)$ and $S_2 = P(t,S)$; indeed, its payoff at time T is $(S_2 - S_1)_+$. We discussed exchange options in Section 3. Can the method we applied to exchange options be used to solve Problem 1? Explain.

(4) Suppose instead of Vasicek we use Hull-White for the short-rate: $dr = (\theta(t) - ar)\,dt + \sigma\,dw$.

(a) Consider a call on a zero-coupon bond, as in Problem 1. Show that the valuation formula given in Problem 1 remains valid.

(b) Does this mean that the value of the option doesn't depend on $\theta(t)$? Explain.

(5) Suppose we didn't know the answer to Problem 1, and we decided to value the call by solving a PDE instead of by using Black's formula. What PDE would we have to solve? (Be sure to specify the final-time condition). How does the solution determine the value of the option? (Note that by solving Problem 1, we have given a formula for the solution of this PDE!)

Continuous Time Finance, Spring 2004 – Homework 4
Posted 3/19/04, due 3/31/04

(1) To solve Problem 5 of HW3 you needed to know that if $dr = (\theta - ar)\,dt + \sigma\,dw$ then the function $v(x,t)$ defined by

$$v(x,t) = E_{r(t)=x}\left[e^{-\int_t^T r(s)\,ds} f(r(T))\right] \tag{1}$$

solves

$$v_t + (\theta - ax)v_x + \tfrac{1}{2}\sigma^2 v_{xx} - xv = 0$$

for $t < T$, with final-time condition $v(x,T) = f(x)$. This is a special case of the Feynman-Kac formula. Give a self-contained proof, using the method of HW1, problem 1. (You should assume that the PDE has a unique solution with this final-time condition; your task is to prove that the solution of the PDE satisfies (1).)

(2) The Section 6 notes explain how a trinomial tree can be used to approximate the random walk $dx = \sigma\,dw$, and how working backward in this tree amounts to a standard finite-difference scheme for solving the backward Kolmogorov equation $u_t + \frac{1}{2}\sigma^2 u_{xx} = 0$. Let's try to do something similar for the "geometric Brownian motion with drift" process $dy = \mu y\,dt + \sigma y\,dw$, whose backward Kolmogorov equation is $v_t + \mu y v_y + \frac{1}{2}\sigma^2 y^2 v_{yy} = 0$. Assume the time interval is Δt, and at time $t = n\Delta t$ the tree has nodes at $-n\Delta y, \ldots, n\Delta y$. The process on the tree goes from (y,t) to $(y+\Delta y, t+\Delta t)$ with probability p_u, to $(y, t+\Delta t)$ with probability p_m, and to $(y-\Delta y, t+\Delta t)$ with probability p_d.

(a) How must p_u, p_m, and p_d be chosen to get the means and variances right? What are the conditions for them to be positive?

(b) What is wrong with this scheme?

(3) A better trinomial approximation of "geometric brownian motion with drift" is obtained by recognizing that if $dy = \mu y\,dt + \sigma y\,dw$ then $y = e^z$ with $dz = (\mu - \frac{1}{2}\sigma^2)\,dt + \sigma\,dw$.

(a) Consider a trinomial tree process which goes from (z,t) to $(z+\Delta z, t+\Delta t)$ with probability p_u, to $(z, t+\Delta t)$ with probability p_m, and $(z-\Delta z, t+\Delta t)$ with probability p_d. How must p_u, p_m, and p_d be chosen to match the means and variances of the z process? What are the conditions for them to be positive?

(b) Working backward in this tree amounts to a finite-difference scheme for solving the backward Kolmogorov PDE $w_t + (\mu - \frac{1}{2}\sigma^2)w_z + \frac{1}{2}\sigma^2 w_{zz}$ with specified final-time data at $t = T$. In what sense can this also be viewed as a scheme for solving the PDE $v_t + \mu y v_{yy} + \frac{1}{2}\sigma^2 y^2 v_{yy} = 0$?

(Note: The "trinomial tree" scheme for valuing options uses this tree for the z process, with $\mu = r$. However the option value is the *discounted* payoff; this introduces a discount factor of $e^{-r\Delta t}$ at each timestep, and a term $-rw$ in the PDE.)

(4) As we discussed in class, the general one-factor HJM model stipulates

$$d_t f = \alpha(t,T)\,dt + \sigma(t,T)\,dw \tag{2}$$

in the risk-neutral measure. We may choose the volatility $\sigma(t,T)$ arbitrarily, but it determines the drift $\alpha(t,T)$ through the formula

$$\alpha(t,T) = \sigma(t,T)\int_t^T \sigma(t,u)\,du. \tag{3}$$

The associated short rate is

$$r(t) = f(0,t) + \int_0^t \sigma(s,t)\,dw(s) + \int_0^t \alpha(s,t)\,ds$$

which solves the SDE

$$dr = \left[\partial_T f(0,t) + \int_0^t \partial_T\sigma(s,t)\,dw(s) + \alpha(t,t) + \int_0^t \partial_T\alpha(s,t)\,ds\right]dt + \sigma(t,t)\,dw(t). \tag{4}$$

Let's verify that when $\sigma(t,T) = \sigma e^{-a(T-t)}$ (with σ constant) we recover the Hull-White model:

(a) Show that $\alpha(t,T) = \frac{\sigma^2}{a}e^{-a(T-t)}\left(1 - e^{-a(T-t)}\right)$.

(b) Show that the SDE (4) reduces in this case to $dr = (\theta(t) - ar)\,dt + \sigma\,dw$ with

$$\theta(t) = \partial_T f(0,t) + af(0,t) + \frac{\sigma^2}{2a}\left(1 - e^{-2at}\right).$$

(5) This problem revisits HW3, problem 1, using the general one-factor HJM theory $d_t f(t,T) = \alpha(t,T)\,dt + \sigma(t,T)\,dw$ rather than Vasicek-Hull-White. Well, not the most general theory: you must assume for this problem that $\sigma(t,T)$ is a given, deterministic function of t and T (whereas the general HJM framework permits it to be random, provided it depends only on time-t information). Besides the formulas (2)-(3), you'll need the fact that

$$d_t[P(t,T)/B_t] = [P(t,T)/B_t]\Sigma(t,T)\,dw \tag{5}$$

where B_t is the money-market account and

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

(a) Show that for $t \le \tau \le T \le S$, the random variable $\ln[P(\tau,S)/P(\tau,T)]$ is normal under the risk-neutral measure, and its variance (given information at time t) is

$$\int_t^\tau (\Sigma(u,S) - \Sigma(u,T))^2\,du.$$

(b) To apply Black's formula, we need the statistics of $\ln[P(\tau, S)/P(\tau, T)]$ under the forward measure, not the risk-neutral measure. (The forward measure is the one for which $V_t/P(t,T)$ is a martingale whenever V_t is the value of a tradeable.) Show that if w is Brownian motion under the risk-neutral measure and $\overline{w}$ is Brownian motion under the forward measure then

$$d\overline{w} = -\Sigma(t,T)\,dt + dw.$$

(Hint: specialize the calculation on page 9 of the Section 4 notes to the case at hand.)

(c) Use the result of (b) to show that $\ln[P(\tau, S)/P(\tau, T)]$ is also normal under the forward measure, and its variance is the *same* under the forward and risk-neutral measures.

(d) Consider a call option with maturity T and strike K, on a zero-coupon bond with maturity $S > T$. Its payoff at time T is $(P(T,S) - K)_+$. Show that its value at time t is

$$P(t,S)N(d_1) - KP(t,T)N(d_2)$$

where

$$d_1 = \frac{\ln[\frac{P(t,S)}{P(t,T)K}] + \frac{1}{2}s^2}{s}, \quad d_2 = d_1 - s$$

where s is defined by

$$s^2 = \int_t^T (\Sigma(u,S) - \Sigma(u,T))^2\,du.$$

(6) This problem revisits HW3, problem 2, using the general one-factor HJM theory. Consider the call option valued in problem 5.

(a) What trading strategy produces a replicating portfolio using tradeables $P(t,S)$ and $P(t,T)$?

(b) What trading strategy produces a replicating portfolio using tradeables $P(t,S)$ and the money market fund B_t?

(c) What trading strategy produces a replicating portfolio using two bonds $P(t,T_1)$ and $P(t,T_2)$, where T_1 and T_2 are arbitrary (distinct) values greater than T?

Continuous Time Finance, Spring 2004 – Homework 5
Posted 4/3/04, due 4/14/04 (TYPO IN THE SDE FOR C CORRECTED 4/7/04)

In problem 3 of HW1 we considered options on a foreign exchange rate, assuming the interest rate in each currency was constant. Now we have more sophisticated interest rate models; let's see how they work in this setting. Let $C(t)$ be the exchange rate in dollars/pound, and consider an option that gives a dollar investor the right to buy pounds at exchange rate K at time T; its payoff (to the dollar investor) is

$$(C(T) - K)_+ \quad \text{dollars at time } T. \tag{1}$$

Use subscripts D, P, and C to distinguish analogous dollar, pound, and exchange-rate objects: for example

$$P_D(t,T) = \text{dollar value at time } t \text{ of a zero-coupon bond worth one dollar at time } T.$$

Use Hull-White models for the dollar and pound short rates:

$$dr_D = (\theta_D(t) - a_D r_D)\,dt + \sigma_D dw_D$$

where w_D is a Brownian motion under the dollar investor's risk-neutral measure; and

$$dr_P = (\theta_P(t) - a_P r_P)\,dt + \sigma_P dw_P$$

where w_P is a Brownian motion under the pound investor's risk-neutral measure. Assume the exchange rate has constant drift and volatility:

$$dC = \mu_C C\,dt + \sigma_C C dw_C$$

where w_C is a Brownian motion under some (subjective) probability. The Brownian motions may be correlated: assume

$$dw_D\,dw_P = \rho_{DP}\,dt, \quad dw_D\,dw_C = \rho_{DC}\,dt, \quad dw_P\,dw_C = \rho_{PC}\,dt,$$

where ρ_{DP}, ρ_{DC}, and ρ_{PC} are constant.

(a) What is the value (to the dollar investor, at time $t < T$) of the payoff (1)? (Make your answer as explicit as possible.)

(b) Describe a trading strategy for the dollar investor that replicates this payoff. (Again, be as explicit as possible.)

(c) Is a similar analysis possible if we use one-factor HJM models for the interest rates rather than Hull-White?

[Extra credit: consider the analogous question for quanto call, whose value to the dollar investor is $(S(T) - K)_+$ at time T, where S is the price of a stock in pounds. This is of course the stochastic-interest-rate analogue of our discussion of quantos, in Section 3 and problem 4 of HW2.]

Continuous Time Finance, Spring 2004 – Homework 6
Distributed 4/14/04, due 4/28/04

(1) We studied the Dupire equation, for calls on a non-dividend-paying stock. This problem asks you to derive the analogous equation for calls on a foreign currency rate. Since the letter C will be used for the call price, we use S for the foreign currency rate. Recall that under the (domestic investor's) risk-free measure it evolves by

$$dS = (r - q)S\,dt + \sigma(S,t)S\,dw$$

where r is the domestic risk-free rate and q is the foreign risk-free rate. Assume r and q are constant, and assume $\sigma(S,t)$ is a deterministic function of S and t. Let

$$C(K,T) = e^{-rT}E\left[(S_T - K)_+\right]$$

be the time-zero value of a call with strike K and maturity T under this model. Show that it solves

$$C_T = \tfrac{1}{2}K^2\sigma^2(K,T)C_{KK} + (q - r)KC_K - qC$$

for $T > 0$ and $K > 0$, with initial condition

$$C(K,0) = (S_0 - K)_+$$

and boundary condition

$$C(0,T) = e^{-qT}S_0$$

where S_0 is the time-zero spot exchange rate.

(2) Consider scaled Brownian motion with drift and jumps: $dy = \mu dt + \sigma dw + JdN$, starting at $y(0) = 0$. Assume the jump occurences are Poisson with rate λ, and the jump magnitudes J are Gaussian with mean 0 and variance δ^2. Find the probability distribution of the process y at time t. (*Hint*: don't try to solve the forward Kolmogorov PDE. Instead observe that you know, for any n, the probability that n jumps will occur before time t; and after conditioning on the number of jumps, the distribution of y is a Gaussian whose mean and variance are easy to determine. Assemble these ingredients to give the density of y as an infinite sum.) [*Comment*: Using essentially the same idea, Merton gave an explicit formula for the value of an option when y is the logarithm of the stock price under the subjective measure.]

Continuous Time Finance Final Exam
Spring 2004 – Professor Kohn
May 5, 2004

This was administered as a closed-book exam, however students were permitted to bring two sheets of notes (both sides, any font). Problems 1-10 require relatively short answers and are worth 10 points each. Problems 11 and 12 require longer answers and are worth 25 points each. Thus the total possible score is 150.

PART I: SHORT ANSWER QUESTIONS (10 points each).

(1) Is the following statement true or false? *"If the interest rate r is constant then the market price of risk must also be constant."*

(2) Suppose S is the price of a risky tradeable. Can S^2 also be the price of a tradeable?

(3) In the Cox-Ingersoll-Ross short-rate model, the SDE for the short rate under the risk-neutral probability is $dr = (\theta - ar)\,dt + \sigma\sqrt{r}dw$, where σ and a are constant and $\theta(t)$ is a deterministic function of time. The prices of zero-coupon bonds under this model satisfy $P(t,T) = V(r(t),t)$ where $V(x,t)$ solves the PDE

$$V_t + \underline{\quad\quad} V_x + \underline{\quad\quad} V_{xx} + \underline{\quad\quad} V = 0, \quad \text{with} \quad V(x,t) = \underline{\quad} \text{ at } t = \underline{\quad}.$$

Fill in the blanks.

(4) Consider options on an underlying whose actual volatility is a function of time alone, i.e. whose subjective process is

$$dS = \mu(S,t)S\,dt + \sigma(t)S\,dw$$

where $\sigma(t)$ is a deterministic function of time. Assume the interest rate r is constant. What is the Black-Scholes implied volatility $\sigma_I(K,T)$?

(5) Consider the one-factor HJM model under which infinitesimal forward rates $f(t,T)$ solve

$$d_t f(t,T) = \alpha(t,T)\,dt + \sigma(t,T)\,dw$$

under the risk-neutral measure. Recall that the SDE satisfied by $P(t,T)$ is then $d_t P(t,T) = rP(t,T)\,dt + \Sigma(t,T)P(t,T)\,dw$ where $\Sigma(t,T) = -\int_t^T \sigma(t,u)\,du$. Fill in the blank: if

$$d\overline{w} = dw + \underline{\quad\quad}\,dt$$

then $\overline{w}$ is a Brownian motion under the forward measure associated with maturity T (i.e. the measure with respect to which the price of any tradeable divided by $P(t,T)$ is a martingale).

(6) Consider a one-factor HJM model as in Problem 5. What is the SDE for the associated short rate? Explain the statement: "HJM models usually lead to non-Markovian short rates."

(7) In discussing Libor Market Model we considered a discrete set of maturity dates $T_1, T_2, \ldots$. Therefore it was natural to consider as numeraire the "rolling-forward CD", i.e. an account which is reinvested at each maturity date T_k, at the term rate for investing from T_k to T_{k+1}. What is the relation between the volatility of the rolling-forward CD and the volatilities of zero-coupon bonds?

(8) Let's consider constructing a trinomial tree for the process $dx = -ax^3\,dt + \sigma\,dw$, with time increment Δt and spatial increment Δx. Suppose that from the node at time $t_j = j\Delta t$ and spatial value $x_k = k\Delta x$ the process goes up to $x_{k+1} = (k+1)\Delta x$ with probability p_u, goes down to $x_{k-1} = (k-1)\Delta x$ with probability p_d, and stays the same (i.e. remains at $x_k = k\Delta x$) with probability p_m. What linear system should be solved to find p_u, p_m, and p_d? (You need not actually solve it. Do not assume any special relation between Δx and Δt.) Do you think it will be necessary to truncate this tree?

(9) Consider the jump-diffusion process

$$dy = -ay\,dt + \sigma dw + J\,dN,$$

with $y(0) = 0$. Assume a and σ are constant; the occurrence of a jump is a Poisson event with rate λ, and the jumps are drawn independently from a specified distribution with mean m. What ODE should you solve to find the mean value $\overline{y} = E[y]$? (You need not actually solve it.)

(10) Consider a local vol model, in which the risk-neutral dynamics of the underlying is $dS = rSdt + \sigma(S,t)Sdw$ for some deterministic function $\sigma(S,t)$. Assume the interest rate r is constant, and the present time is $t = 0$. Let $P(K,T)$ be the value of a put option with strike K and maturity T, and let $\rho(\xi, T)$ be the probability (under the risk-neutral measure) that the underlying has value ξ at time T. Give a formula for P_{KK} in terms of ρ.

PART II: LONGER-ANSWER QUESTIONS (25 points each).

(11) This problem concerns the pricing and hedging of a quanto option. The underlying is a Euro tradeable, whose price (in Euros) satisfies

$$dS = \mu_S S\,dt + \sigma_S S\,dw_S$$

under the subjective measure. The exchange rate C in dollars per Euro satisfies

$$dC = \mu_C C\,dt + \sigma_C C\,dw_C.$$

The dollar risk-free rate is r and the Euro risk-free rate is u. Assume all parameters $(\mu_S, \mu_C, \sigma_S, \sigma_C, r, u)$ are constant, and $dw_S\,dw_C = \rho\,dt$. Notice that the natural dollar tradeables are the stock valued in dollars (value SC), the Euro money market account valued in dollars (value CD where $D = e^{ut}$), and the dollar money market account (value $B = e^{rt}$).

Let's price and hedge the option whose payoff at time T is $f(S_T)$ dollars.

(a) What is the SDE for S under the dollar investor's risk-neutral measure?

(b) What PDE should the dollar investor solve to price the option?

(c) How can a dollar investor replicate this option, by trading only in the stock (valued in dollars), the Euro money market account (valued in dollars), and the dollar money market account?

(12) This problem concerns the pricing and hedging of certain interest-based exchange options using the Hull-White model. So we assume short rate satisfies $dr = (\theta(t) - ar)\,dt + \sigma\,dw$ where a and σ are constant and θ is a deterministic function of t. Consider the option whose payoff is $[P(T,T_2) - KP(T,T_1)]_+$ at time $t = T$, where $T < T_1 < T_2$ and K is a fixed constant.

(a) Explain the sense in which this is an "exchange option."

(b) What is the SDE for $P(t,T_2)/P(t,T_1)$ under the forward measure associated with maturity T_1?

(c) Value the option using the result of (b) and Black's formula.