m3f33chVII

Ch. VII: MATHEMATICAL FINANCE IN CONTINUOUS TIME

§1. Geometric Brownian Motion (GBM)

As before, we write B for standard Brownian motion. We write $B_{\mu,\sigma}$ for Brownian motion with *drift* μ and *diffusion coefficient* σ: the path-continuous Gaussian process with independent increments such that

$$B_{\mu,\sigma}(s+t) - B_{\mu,\sigma}(s) \text{ is } N(\mu t, \sigma^2 t).$$

This may be realised as

$$B_{\mu,\sigma}(t) = \mu t + \sigma B(t).$$

Consider the process

$$X_t = f(t, B_t) := x_0 . \exp\{(\mu - \frac{1}{2}\sigma^2)t + \sigma B_t\} : \quad \log X_t = const + (\mu - \frac{1}{2}\sigma^2)t + \sigma B_t. \tag{$*$}$$

Here, since

$$f(t,x) = x_0 . \exp\{(\mu - \frac{1}{2}s^2)t + \sigma x\},$$

$$f_1 = (\mu - \frac{1}{2}\sigma^2)f, \qquad f_2 = \sigma f, \qquad f_{22} = \sigma^2 f.$$

By Itô's Lemma (Ch. VI: $dX_t = U_t dt + V_t dB_t$ and f smooth implies $df = (f_1 + U f_2 + \frac{1}{2}V^2 f_{22})dt + V f_2 dB_t$) we have (taking $U = 0, \quad V = 1, \quad X = B$),

$$dX_t = df = [(\mu - \frac{1}{2}\sigma^2)f + \frac{1}{2}\sigma^2 f]dt + \sigma f dB_t :$$

$$dX_t = \mu f dt + \sigma f dB_t = \mu X_t dt + \sigma X_t dB_t :$$

X satisfies the SDE

$$dX_t = X_t(\mu dt + \sigma dB_t) : \qquad dX_t/X_t = \mu dt + \sigma dB_t, \tag{GBM}$$

and is called *geometric Brownian motion* (GBM). We turn to its economic meaning, and the role of the two parameters μ and σ, below. It will be used to model *price processes* in the Black-Scholes model of VII.2. But note that in $(*)$, *log-prices* $\log X_t$ *are normally distributed.*

Note that for $\mu = 0$, (GBM) shows that X is a martingale (see VII.3, in connection with Girsanov's theorem).

We recall the model of Brownian motion from Ch. VI. It was developed (by Brown, Einstein, Wiener, ...) in *statistical mechanics*, to model the irregular, random motion of a particle suspended in fluid under the impact of collisions with the molecules of the fluid.

The situation in *economics* and *finance* is analogous: the price of an asset depends on many factors (a share in a manufacturing company depends on, say, its own labour costs, and raw material prices for the articles it manufactures. Together, these involve, e.g., foreign exchange rates, labour costs, transport costs, etc. – all of which respond to the unfolding of events – economic data/political events/the weather/technological change/labour, commercial and environmental legislation/ ... in time. There is also the effect of individual transactions in the buying and selling of a traded asset on the asset price. The analogy between the buffeting effect of molecules on a particle in the statistical mechanics context on the one hand, and that of this continuous flood of new price-sensitive information on the other, is highly suggestive. The first person to use Brownian motion to model price movements in economics was Bachelier in his celebrated thesis of 1900.

Bachelier's seminal work was not definitive (indeed, not correct), either mathematically (it was pre-Wiener) or economically. In particular, Brownian motion itself is inadequate for modelling prices, as
(i) it attains negative levels, and
(ii) one should think in terms of *return*, rather than prices themselves.
However, one can allow for both of these by using *geometric*, rather than ordinary, Brownian motion as one's basic model. This has been advocated in economics from 1965 on by Samuelson[1] – and was Itô's starting-point for his development of Itô or stochastic calculus in 1944 – and is now standard.

Returning now to (GBM), the SDE above for geometric Brownian motion driven by Brownian noise, we can see how to interpret it. We have a risky asset (stock), whose price at time t is X_t; $dX_t = X(t + dt) - X(t)$ is the change in X_t over a small time-interval of length dt beginning at time t; dX_t/X_t is the gain per unit of value in the stock, i.e. the *return*. This is a sum of two components:
(i) a deterministic component μdt, equivalent to investing the money risklessly in the bank at interest-rate μ (> 0 in applications), called the *under-*

[1] Paul A. Samuelson (1915-2009), American economist; Nobel Prize in Economics, 1970

lying return rate for the stock,
(ii) a random, or noise, component σdB_t, with *volatility* parameter $\sigma > 0$ and driving Brownian motion B, which models the market uncertainty, i.e. the effect of noise. Note that dB_t is a Brownian increment, so is normally distributed. So: *returns are normally distributed.*
Return intervals.

That both log-prices and returns are normally distributed just reflects

$$\log(1+x) \sim x \qquad (x \to 0),$$

or equivalently (as in II.1),

$$\left(1 + \frac{x}{n}\right) \to e^x \qquad (n \to \infty).$$

We can recognise this as being bound up with the passage from discrete time (time-interval Δt, small but finite, as in V) to continuous time (time-interval dt, infinitesimal, and the SDE for GBM as above). Now in investment, there are many possible time-scales, corresponding to how often we observe prices; we single out the main three (cf. [BK, §2.9]).
1. *Long (macroscopic).*

Here we are investing over a time-scale of months (say – or years), and observe prices daily (say). As the price-change over the month is the sum of price-changes over the days, and these are independent (as Brownian increments are), the reason we get normality is the *Central Limit Theorem (CLT)*: if we sum many independent random variables with finite mean and variance, we get normality (in the limit) after centring and scaling. This is the phenomenon of *aggregational Gaussianity.* Note that Gaussian (normal) tails are *extremely thin* ('minus log-density' grows quadratically). The 'rule of thumb' is that 16 trading days suffice here.
2. *Intermediate (mesoscopic).*

If our investment time-frame is, say, a day (there are 'day traders' out there!), aggregational Gaussianity does not set in, and the tails observed are *much fatter* – typically, 'minus log-density' grows linearly. One model commonly used here is that of *hyperbolic distributions* (see e.g. [BK, §2.12]).
3. *Short (microscopic).*

With the development of the Internet and the intensive computerisation of trading, high-frequency data – '*tick data*' – is available; here the interval may be of the order of seconds or much smaller. Here, the picture is different again: the tails are *much fatter still*: tails decay like a power, so 'minus

log-density' grows logarithmically. Distributions used include Student t and stable (see e.g. [BK, §2.9]).
Note. The world's most famous investor, Warren Buffett, the Sage of Omaha, famously invests *right*, and over a time-frame of *many years*.

§2. The Black-Scholes Model

For this section only, it is convenient to be able to use the 'W for Wiener' notation for Brownian motion/Wiener process, thus liberating B for the alternative use 'B for bank [account]'. Thus our driving noise terms will now involve dW_t, our deterministic [bank-account] terms dB_t.

We now consider an investor constructing a trading strategy in continuous time, with the choice of two types of investment:
(i) riskless investment in a bank account paying interest at rate $r > 0$ (the *short rate* of interest): $B_t = B_0 e^{rt} \quad (t \geq 0)$ [we neglect the complications involved in possible failure of the bank – though *banks do fail* – witness Barings 1995, or AIB 2002!];
(ii) risky investment in stock, one unit of which has price modelled as above by $GMB(\mu, \sigma)$. Here the volatility $\sigma > 0$; the restriction $0 < r < \mu$ on the short rate r for the bank and underlying rate μ for the stock are economically natural (but not mathematically necessary); the stock dynamics are thus

$$dS_t = S_t(\mu dt + \sigma dW_t).$$

Notation. Later, we shall need to consider several types of risky stock - d stocks, say. It is convenient, and customary, to use a *superscript* i to label stock type, $i = 1, \cdots, d$; thus $S^1, \cdots, S^d$ are the risky stock prices. We can then use a superscript 0 to label the bank account, S^0. So with one risky asset as above, the dynamics are

$$\begin{aligned} dS_t^0 &= rS_t^0 dt, \\ dS_t^1 &= \mu S_t^1 dt + \sigma S_t^1 dW_t. \end{aligned}$$

We shall focus on pricing at time 0 of options with expiry time T; thus the index-set for time t throughout may be taken as $[0, T]$ rather than $[0, \infty)$.

We proceed as in the discrete-time model of V.1. A *trading strategy* H is a vector stochastic process

$$H = (H_t : 0 \leq t \leq T) = ((H_t^0, H_t^1, \cdots, H_t^d)) : 0 \leq t \leq T)$$

which is *previsible*: each H_t^i is a previsible process (so, in particular, $(\mathcal{F}_{t-})$-adapted) [we may simplify with little loss of generality by replacing previsibility here by *left-continuity* of H_t in t]. The vector $H_t = (H_t^0, H_t^1, \cdots, H_t^d)$ is the *portfolio* at time t. If $S_t = (S_t^0, S_t^1, \cdots, S_t^d)$ is the vector of *prices* at time t, the *value* of the portfolio at t is the scalar product

$$V_t(H) := H_t.S_t = \Sigma_{i=0}^d H_t^i S_t^i.$$

The *discounted value* is

$$\tilde{V}_t(H) = \beta_t(H_t.S_t) = H_t.\tilde{S}_t,$$

where $\beta_t := 1/S_t^0 = e^{-rt}$ (fixing the scale by taking the initial bank account as 1, $S_0^0 = 1$), so

$$\tilde{S}_t = (1, \beta_t S_t^1, \cdots, \beta_t S_t^d)$$

is the vector of discounted prices.

Recall that
(i) in V.1 H is a self-financing strategy if $\Delta V_n(H) = H_n.\Delta S_n$, i.e. $V_n(H)$ is the martingale transform of S by H,
(ii) stochastic integrals are the continuous analogues of mg transforms.
We thus define the strategy H to be *self-financing*, $H \in SF$, if

$$dV_t = H_t.dS_t = \Sigma_0^d H_t^i dS_t^i.$$

The discounted value process is

$$\tilde{V}_t(H) = e^{-rt} V_t(H)$$

and the interest rate is r. So

$$d\tilde{V}_t(H) = -re^{-rt}dt.V_t(H) + e^{-rt}dV_t(H)$$

(since e^{-rt} has finite variation, this follows from integration by parts,

$$d(XY)_t = X_t dY_t + Y_t dX_t + \frac{1}{2}d\langle X, Y\rangle_t$$

– the quadratic covariation of a finite-variation term with any term is zero)

$$= -re^{-rt}H_t.S_t dt + e^{-rt}H_t.dS_t = H_t.(-re^{-rt}S_t dt + e^{-rt}dS_t) = H_t.d\tilde{S}_t$$

($\tilde{S}_t = e^{-rt}S_t$, so $d\tilde{S}_t = -re^{-rt}S_t dt + e^{-rt}dS_t$ as above).

Summarising: for H self-financing,

$$dV_t(H) = H_t.dS_t, \qquad d\tilde{V}_t(H) = H_t.d\tilde{S}_t,$$

$$V_t(H) = V_0(H) + \int_0^t H_s dS_s, \qquad \tilde{V}_t(H) = \tilde{V}_0(H) + \int_0^t H_s d\tilde{S}_s.$$

Now write $U_t^i := H_t^i S_t^i / V_t(H) = H_t^i S_t^i / \Sigma_j H_t^j S_t^j$ for the *proportion* of the value of the portfolio held in asset $i = 0, 1, \cdots, d$. Then $\Sigma U_t^i = 1$, and $U_t = (U_t^0, \cdots, U_t^d)$ is called the *relative portfolio*. For H self-financing,

$$dV_t = H_t.dS_t = \Sigma H_t^i dS_t^i = V_t \Sigma \frac{H_t^i S_t^i}{V_t}.\frac{dS_t^i}{S_t^i} : \qquad dV_t = V_t \Sigma U_t^i dS_t^i / S_t^i.$$

Dividing through by V_t, this says that the return dV_t/V_t is the weighted average of the returns dS_t^i/S_t^i on the assets, weighted according to their proportions U_t^i in the portfolio – as one would expect.

Note. Having set up this notation (that of [HP]) – in order to be able if we wish to have a basket of assets in our portfolio – we now prefer – for simplicity – to specialise back to the simplest case, that of one risky asset. Thus we will now take $d = 1$ until further notice.

§3. The (continuous) Black-Scholes formula (BS): derivation via Girsanov's Theorem

The Sharpe ratio.

There is no point in investing in a risky asset with mean return rate μ, when cash is a riskless asset with return rate r, unless $\mu > r$. The excess return $\mu - r$ (the investor's reward for taking a risk) is compared with the risk, as measured by the volatility σ, via the *Sharpe ratio*

$$\theta := (\mu - r)/\sigma,$$

also written λ, and also known as the *market price of risk*. This is important, both here (see below), in CAPM (II.3), and in asset allocation decisions.

Consider now the Black-Scholes model, with dynamics

$$dB_t = rB_t dt, \qquad dS_t = \mu S_t dt + \sigma S_t dW_t.$$

The discounted asset prices $\tilde{S}_t := e^{-rt}S_t$ have dynamics given, as before, by

$$\begin{aligned} d\tilde{S}_t &= -re^{-rt}S_t dt + e^{-rt}dS_t = -r\tilde{S}_t dt + \mu\tilde{S}_t dt + \sigma\tilde{S}_t dW_t \\ &= (\mu - r)\tilde{S}_t dt + \sigma\tilde{S}_t dW_t = \sigma\tilde{S}_t(\theta dt + dW_t). \end{aligned}$$

We summarise the main steps briefly as (a) - (f) below:
(a) Dynamics are given by GBM, $dS_t = \mu S dt + \sigma S dW_t$ (VI.1).
(b) Discount: $d\tilde{S}_t = (\mu - r)\tilde{S}dt + \sigma\tilde{S}dW_t = \sigma\tilde{S}(\theta dt + dW_t)$ (above).

We work with the discounted stock price $\tilde{S}_t$. We would like this to be a *martingale*, as in Ch. V, where we passed from P-measure to Q- (or P^*)-measure, so as to make *discounted asset prices martingales*. Girsanov's theorem (below) accomplishes this, in our new continuous-time setting: it maps P to P^* (or Q), and μ to r, so θ to 0. This kills the dt term on the right in (b). If we then integrate $d\tilde{S}_t = \sigma\tilde{S}dW_t$, we get an Itô integral, so a *martingale*, on the right. Assuming this for now:
(c) Use Girsanov's Theorem to change μ to r, so $\theta := (\mu - r)/\sigma$ to 0: under P^*, $d\tilde{S}_t = \sigma\tilde{S}dW_t$.
(d) This and $d\tilde{V}_t(H) = H_t d\tilde{S}_t$ (where V is the value process and H the trading strategy replicating the payoff h – VII.2) give $d\tilde{V}_t(H) = H_t.\sigma\tilde{S}_t dW_t$ (VII.2 above). Integrate: $\tilde{V}_t$ is a P^*-mg, so has constant E^*-expectation.
(e) This gives the Risk-Neutral Valuation Formula (RNVF), as in V.4.
(f) From RNVF, we can obtain BS, by integration, as in V.6.

It remains to state and discuss *Girsanov's theorem*. We cannot prove it in full (only the finite-dimensional approximation below) – this is technical Measure Theory. But we must expect this in this chapter: in discrete time (Ch. V) we could prove everything; here in continuous time, we can't.

Consider first ([KS], §3.5) independent $N(0,1)$ random variables $Z_1, \cdots, Z_n$ on $(\Omega, \mathcal{F}, P)$. Given a vector $\mu = (\mu_1, \cdots, \mu_n)$, consider a new probability measure $\tilde{P}$ on $(\Omega, \mathcal{F})$ defined by

$$\tilde{P}(d\omega) = \exp\{\Sigma_1^n \mu_i Z_i(\omega) - \frac{1}{2}\Sigma_1^n \mu_i^2\}.P(d\omega).$$

This is a positive measure as $\exp\{.\} > 0$, and integrates to 1 as $\int \exp\{\mu_i Z_i\}dP = E[e^{\mu_i Z_i}] = \exp\{\frac{1}{2}\mu_i^2\}$ (normal MGF – Problems 4 (bivariate normal), or Problems 8 Q1), so is a probability measure. It is also *equivalent* to P (has the same null sets), again as the exponential term is positive (the exponential on the right is the *Radon-Nikodym derivative* $d\tilde{P}/dP$). Also

$$\tilde{P}(Z_i \in dz_i, \quad i = 1, \cdots, n) = \exp\{\Sigma_1^n \mu_i z_i - \frac{1}{2}\Sigma_1^n \mu_i^2\}.P(Z_i \in dz_i, \quad i = 1, \cdots, n)$$

($Z_i \in dz_i$ means $z_i \leq Z_i \leq z_i + dz_i$, so here $Z_i = z_i$ to first order)

$$= (2\pi)^{-\frac{1}{2}n} \exp\{\Sigma\mu_i z_i - \frac{1}{2}\Sigma\mu_i^2 - \frac{1}{2}\Sigma z_i^2\}\Pi dz_i = (2\pi)^{-\frac{1}{2}n} \exp\{-\frac{1}{2}\Sigma(z_i - \mu_i)^2\}dz_1 \cdots dz_n.$$

This says that if the Z_i are independent $N(0,1)$ under P, they are independent $N(\mu_i, 1)$ under $\tilde{P}$. Thus the effect of the *change of measure* $P \mapsto \tilde{P}$, from the original measure P to the *equivalent* measure $\tilde{P}$, is to *change the mean*, from $0 = (0, \cdots, 0)$ to $\mu = (\mu_1, \cdots, \mu_n)$.

This result extends to infinitely many dimensions – i.e., stochastic processes. We quote (Igor Vladimirovich GIRSANOV (1934-67) in 1960):

Theorem (Girsanov's Theorem). Let $(\mu_t : 0 \leq t \leq T)$ be an adapted process with $\int_0^T \mu_t^2 dt < \infty \quad a.s.$ such that the process L with

$$L_t := \exp\{\int_0^t \mu_s dW_s - \frac{1}{2}\int_0^t \mu_s^2 ds\} \qquad (0 \leq t \leq T)$$

is a martingale. Then, under the probability P_L with density L_T relative to P, the process W^* defined by

$$W_t^* := W_t - \int_0^t \mu_s ds, \qquad (0 \leq t \leq T)$$

is a standard Brownian motion (so W is BM + $\int_0^t \mu_s ds$).

Here, L_t is the *Radon-Nikodym derivative* of P_L w.r.t. P on the σ-algebra $\mathcal{F}_t$. In particular, for $\mu_t \equiv \mu$, *change of measure* by introducing the RN derivative $\exp\{\mu W_t - \frac{1}{2}\mu^2\}$ corresponds to a *change of drift* from 0 to μ.
Exponential martingale.

The martingale condition in Girsanov's theorem is satisfied in the case $\mu_t \equiv \mu$ is constant. For, write

$$M_t := \exp\{\mu W_t - \frac{1}{2}\mu^2 t\}.$$

This is a martingale. For, if $s < t$,

$$\begin{aligned} E[M_t|\mathcal{F}_s] &= E[\exp\{\mu(W_s + (W_t - W_s)) - \frac{1}{2}\mu^2(s + (t-s))\}|\mathcal{F}_s] \\ &= \exp\{\mu W_s - \frac{1}{2}\mu^2 s\}.E[\exp\{\mu(W_t - W_s) - \frac{1}{2}\mu^2(t-s)], \end{aligned}$$

as the conditioning has no effect on the second term, by independent increments of Brownian motion. The first term on the right is M_s. The second term is 1. For (normal MGF), if $Z \sim N(0,1)$,

$$E[\exp\{\mu Z\}] = \exp\{\frac{1}{2}\mu^2\};$$

$$W_t - W_s = \sqrt{t-s}Z, \qquad Z \sim N(0,1)$$

(properties of BM). Combining, M is a mg, as required. //

So the case μ_t constant $= \mu$ of Girsanov's theorem passes between BM and BM + μt. The argument above uses this with $\mu - r$ for μ.

Girsanov's Theorem (or the Cameron-Martin-Girsanov Theorem: R. H. Cameron and W. T. Martin, 1944, 1945) is formulated in varying degrees of generality, and proved, in [KS, §3.5], [RY, VIII].

Stochastic exponential.

The SDE for GBM, $dS_t/S_t = \mu dt + \sigma dW_t$, with solution $S_t = S_0 \exp\{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t\}$ as above, is a special case of the *Doléans-Dade exponential* (or *stochastic exponential*: Cathérine Doléans-Dade (1942-2004)). It extends from Brownian motion to semi-martingales M, when it is written $\mathcal{E}(M)$.

Theorem (Risk-Neutral Valuation Formula, RNVF). The no-arbitrage price of the claim $h(S_T)$ is given by

$$F(t,x) = e^{-r(T-t)} E^*_{t,x}[h(S_T)|\mathcal{F}_t],$$

where $S_t = x$ is the asset price at time t and P^* is the measure under which the asset price dynamics are given by

$$dS_t = rS_t dt + \sigma S_t dW_t.$$

Proof (Step (e) in the above: (a) – (d) are already done). Change measure from P, corresponding to $GBM(\mu,\sigma)$, to P^*, corresponding to $GBM(r,\sigma)$, by Girsanov's Theorem. Then as above, $d\tilde{S}_t = \sigma \tilde{S}_t dW_t$. So by VII.2, $d\tilde{V}_t = H_t d\tilde{S}_t = H_t.\sigma \tilde{S}_t dW_t$, where V is the value process following strategy H to replicate payoff h. Integrating, V_t is a P^*-martingale, as it is an Itô integral. So it has constant expectation. So if $S_t = x$ is the asset price at time t,

$$E^*_{t,x}[\tilde{V}_t(H)|\mathcal{F}_t] = E^*_{t,x}\tilde{V}_T(H) = e^{-rT} E^*_{t,x} h(S_T):$$

$$F(t,x) = E^*_{t,x} V_t(H) = e^{-r(T-t)} E^*_{t,x} h(S_T). \qquad //$$

Theorem ((Continuous) Black-Scholes Formula, BS).

$$F(t,S) = S\Phi(d_+) - e^{-r(T-t)}K\Phi(d_-), \qquad d_\pm := [\log(S/K) + (r \pm \frac{1}{2}\sigma^2)(T-t)]/\sigma\sqrt{T-t}.$$

Proof (Step (f) in the above). After the change of measure $P \mapsto P^*$, $\mu \mapsto r$ by Girsanov's Theorem, S_t has P^*-dynamics as in $GBM(r,\sigma)$:

$$dS_t = rS_t dt + \sigma S_t dW_t, \qquad S_t = s, \qquad (*)$$

with W a P^*-Brownian motion. So (VII.1) we can solve this explicitly:

$$S_T = s\exp\{(r - \frac{1}{2}\sigma^2)(T-t) + \sigma(W_T - W_t)\}.$$

Now $W_T - W_t$ is normal $N(0, T-t)$, so $(W_T - W_t)/\sqrt{T-t} =: Z \sim N(0,1)$:

$$S_T = s\exp\{(r - \frac{1}{2}\sigma^2)(T-t) + \sigma Z\sqrt{T-t}\}, \qquad Z \sim N(0,1).$$

So by the Risk-Neutral Valuation Formula, the pricing formula is

$$F(t,x) = e^{-r(T-t)} \int_{-\infty}^{\infty} h(s\exp\{(r - \frac{1}{2}\sigma^2)(T-t) + \sigma(T-t)^{\frac{1}{2}}x\}).\frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}}dx.$$

For a general payoff function h, there is no explicit formula for the integral, which has to be evaluated numerically. But we can evaluate the integral for the basic case of a European call option with strike-price K:

$$h(s) = (s-K)^+.$$

Then

$$F(t,x) = e^{-r(T-t)} \int_{-\infty}^{\infty} \frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}}[s\exp\{(r - \frac{1}{2}\sigma^2)(T-t) + \sigma(T-t)^{\frac{1}{2}}x\} - K]_+ dx.$$

We have already evaluated such integrals in Chapter V, where we obtained the BS formula from the binomial model by a passage to the limit. Completing the square in the exponential as before gives the result, as in V.6. //

4. Related results.

The Black-Scholes PDE and the Feynman-Kac formula.

The original proof of the Black-Scholes formula was via the *Black-Scholes PDE*; this approach is closely linked with the *Feynman-Kac formula*, which links PDEs with SDEs. For details, see the website for last year's course.

Risk-neutral measure.

We call P^* the *risk-neutral* probability measure. It is equivalent to P (by Girsanov's Theorem, which gives the Radon-Nikodym derivative showing equivalence), and is a martingale measure (as the discounted asset prices are P^*-martingales, by above), i.e. P^* (or Q) is the *equivalent martingale measure (EMM).*

Fundamental Theorem of Asset Pricing (FTAP). The above continuous-time result may be summarised just as the FTAP in discrete time: to get the no-arbitrage price of a contingent claim, take the discounted expected value under the equivalent mg (risk-neutral) measure.

Completeness.

In discrete time, we saw that absence of arbitrage corresponded to *existence* of risk-neutral measures, completeness to *uniqueness*. We have obtained existence and uniqueness here (and so completeness), by appealing to Girsanov's Theorem, which we have not proved in full. Completeness questions are linked to the Representation Theorem for Brownian Martingales, below.

Theorem (Representation Theorem for Brownian Martingales). Let $(M_t : 0 \leq t \leq T)$ be a square-integrable martingale with respect to the Brownian filtration $(\mathcal{F}_t)$. Then there exists an adapted process $H = (H_t : 0 \leq t \leq T)$ with $E \int H_s^2 ds < \infty$ such that

$$M_t = M_0 + \int_0^t H_s dW_s, \qquad 0 \leq t \leq T.$$

That is, all Brownian martingales may be represented as stochastic integrals with respect to Brownian motion.

We refer to, e.g., [KS], [RY] for proof.

The economic relevance of the Representation Theorem is that it shows (see e.g. [KS, I.6], and below) that the Black-Scholes model is *complete* – that

is, that EMMs are unique, and so that *Black-Scholes prices are unique* (we know this already, from FTAP/RNVF above). Mathematically, the result is purely a consequence of properties of the Brownian filtration. The desirable mathematical properties of BM are thus seen to have hidden within them desirable economic and financial consequences of real practical value.
Hedging.

To find a hedging strategy $H = (H_t^0, H_t)$ (H_t^0 for cash, H_t for stock) that replicates the value process $V = (V_t)$, itself given by RNVF (VII.3):

$$V_t = H_t^0 + H_t S_t = E^*[e^{-r(T-t)} h | \mathcal{F}_t].$$

Now

$$M_t := E^*[e^{-rT} h | \mathcal{F}_t]$$

is a martingale (indeed, a uniformly integrable mg: IV.4, V.2) under the filtration $\mathcal{F}_t$, that of the driving BM in (GBM) (VII.1, VII.2), and the filtration is unchanged by the Girsanov change of measure (we quote this). So by the Representation Theorem for Brownian Martingales, there is some adapted process $K = (K_t)$ with

$$M_t = M_0 + \int_0^t K_s dW_s \qquad (t \in [0, T]).$$

Take

$$H_t := K_t / (\sigma \tilde{S}_t), \qquad H_t^0 := M_t - H_t \tilde{S}_t.$$

Then

$$dM_t = K_t dW_t = \frac{K_t}{\sigma \tilde{S}_t} . \sigma \tilde{S}_t dW_t = H_t d\tilde{S}_t,$$

and the strategy given by K is self-financing, by VII.2. This is of limited practical value:
(a) the Representation Theorem does not give $K = (K_t)$ explicitly – it is merely an existence proof;
(b) we already know that, as Brownian paths have infinite variation, exact hedging in the Black-Scholes model is too rough to be practically possible.

So to hedge in practice, we need to go back to discrete time, where we can compute things and where such roughness questions do not arise. But this is familiar by now (and is why we have Chapters IV, V in discrete time and Chapters VI, VII in continuous time). We need to go back and forth at will between *continuous* time – where we can do *calculus*, in particular, Itô

calculus – and *discrete* time – where we can *calculate*, using *computers*.
Comments on the Black-Scholes formula.
1. The Black-Scholes formula transformed the financial world. Before it (see Ch. II), the expert view was that asking what an option is worth was (in effect) a silly question: the answer would necessarily depend on the attitude to risk of the individual considering buying the option. It turned out that – at least approximately (i.e., subject to the restrictions to perfect – frictionless – markets, including No Arbitrage – an over-simplification of reality) there *is* an option value. One can see this in one's head, without doing any mathematics, if one knows that the Black-Scholes market is *complete* (above). So, every contingent claim (option, etc.) can be *replicated*, by a suitable combination of cash and stock. Anyone can price this: (i) count the cash, and count the stock; (ii) look up the current stock price; (iii) do the arithmetic.
2. The programmable pocket calculator was becoming available around this time. Every trader immediately got one, and programmed it, so that he could price an option (using the BS model!) in real time, from market data.
3. The missing quantity in the Black-Scholes formula is the *volatility*, σ. But, the price is continuous and strictly increasing in σ (options like volatility!). So there is *exactly one* value of σ that gives the price at which options are being currently traded. This – the *implied volatility* – is the value that the market currently judges σ to be, and the one that traders use.
4. Because the Black-Scholes model is the benchmark model of mathematical finance, and gives a value for σ at the push of a button, it is widely used.
5. This is *despite* the fact that no one actually believes the Black-Scholes model! It is an over-simplified approximation to reality. Indeed, Fischer Black himself famously once wrote a paper called *The holes in Black-Scholes*.
6. This is an interesting example of theory and practice interacting!
7. Black and Scholes has considerable difficulty in getting their paper published! It was ahead of its time. When published, and its importance understood, it changed its times.
8. Black-Scholes theory and its developments, plus the internet (a global network of fibre-optic cables – using *photons* rather than *electrons*), were important contributory factors to *globalization*. Enormous sums of money can be transported round the world at the push of a button, and are every day. This has led to *financial contagion* – "one country's economic problem becomes the world's economic problem". (The Ebola virus comes to mind here.) The resulting problems of *systemic stability* are very important, and still largely unsolved; they dominate the agenda at international meetings.

§5. Infinite time-horizon; American puts

We sketch here the theory of the American option (one can exercise at any time), over an infinite time-horizon. We deal first with a *put* option (see VII.6 below under *Real options* for the corresponding 'call option') – giving the right to sell at the strike price K, at any time τ of our choosing. This τ has to be a *stopping time*: we have to take the decision whether or not to stop at τ based on information already available – no access to the future, no insider trading. As above, we pass to the risk-neutral measure.

Recall that for American puts over a *finite* time-horizon, we used the *Snell envelope*, which is also the *least supermartingale majorant*. Now for the Snell envelope, we need the finite time-horizon T to begin, as we build up the Snell envelope by working *backwards in time* from T – *dynamic programming*. Although this does not apply when $T = \infty$, the least supermartingale majorant *does*. We quote this. We refer for the theory here, and for the pricing argument below, to the standard work on this subject, Peskir & Shiryaev:
[PS] G. PESKIR and A. N. SHIRYAEV: *Optimal stopping and free-boundary problems*. Birkhäuser, 2006.

Under the risk-neutral measure, the SDE for GBM becomes

$$dX_t = rX_t dt + \sigma X_t dB_t. \qquad (GBM_r)$$

To evaluate the option, we have to solve the *optimal stopping problem*

$$V(x) := \sup_{\tau} E_x[e^{-r\tau}(K - X_\tau)^+]$$

where the sup is taken over all stopping times τ and $X_0 = x$ under P_x.

The process X satisfying (GBM_r) – a *diffusion* – is specified by a second-order linear differential operator, called its (infinitesimal) *generator*,

$$L_X := rxD + \frac{1}{2}\sigma^2 x^2 D^2, \qquad D := \partial/\partial x.$$

Now the closer X gets to 0, the less likely we are to gain by continuing. This suggests that our best strategy is to stop when X gets too small: to stop at $\tau = \tau_b$, where

$$\tau_b := \inf\{t \geq 0 : X_t \leq b\},$$

for some $b \in (0, K)$ (the only range in which we would want to exercise an option to sell at K). This gives the following *free boundary problem* for the *unknown value function* $V(x)$ and the *unknown point* $b \in (0, K)$:

$$L_X V = rV \qquad \text{for } x > b; \tag{i}$$

$$V(x) = (K - x)_+ = K - x \qquad \text{for } x = b; \tag{ii}$$

$$V'(x) = -1 \qquad \text{for } x = b \text{ (smooth fit)}; \tag{iii}$$

$$V(x) > (K - x)^+ \qquad \text{for } x > b; \tag{iv}$$

$$V(x) = (K - x)^+ \qquad \text{for } 0 < x < b. \tag{v}$$

Writing $d := \sigma^2/2$ ('d for diffusion'), (i) is

$$dx^2 V'' + rxV' - rV = 0. \tag{i^*}$$

This ODE is *homogeneous*. So (Euler's theorem): use trial solution:

$$V(x) = x^p.$$

Substituting gives a quadratic for p:

$$p^2 - (1 - \frac{r}{d})p - \frac{r}{d} = 0.$$

Trial solution: $V(x) = x^p$. Substituting gives a quadratic for p:

$$p^2 - (1 - \frac{r}{d})p - \frac{r}{d} = 0 : \qquad (p - 1)(p + r/d) = 0.$$

One root is $p = 1$; the other is $p = -r/d$. So the general solution is $V(x) = C_1 x + C_2 x^{-r/d}$. But $V(x) \leq K$ for all $x \geq 0$ (an option giving the right to sell at price K cannot be worth more than K!): $V(x)$ is bounded. Taking x large ($x < b$ is covered by (v)), we must have $C_1 = 0$. So (with $C := C_2$)

$$V(x) = Cx^{-r/d}; \qquad V'(x) = -\frac{r}{d}.C^{-r/d-1}. \tag{$*$}$$

From (ii),

$$Cb^{-r/d} = K - b,$$

while from (iii),

$$-\frac{r}{d}.Cb^{-r/d}.\frac{1}{b} = -1 : \qquad Cb^{-r/d} = \frac{bd}{r}.$$

Equating, this gives C and b:

$$\frac{bd}{r} = K - b, \qquad K = b(1 + d/r), \quad b = K/(1 + d/r).$$

Then $(*)$ and (iii) give

$$C = \frac{d}{r}\Big(\frac{K}{1 + d/r}\Big)^{1+r/d}.$$

So

$$\begin{aligned} V(x) &= \frac{d}{r}\Big(\frac{K}{1 + d/r}\Big)^{1+r/d} x^{-r/d} \qquad \text{if } x \in [b, \infty) \\ &= K - x \qquad \text{if } x \in (0, b]. \end{aligned}$$

This is in fact the full and correct solution to the problem; see [PS], §25.1.

The 'smooth fit' in (iii) is characteristic of free boundary problems. For a heuristic analogy: imagine trying to determine the shape of a rope, tied to the ground on one side of a convex body, stretched over the body, then pulled tight and tied to the ground on the other side. We can see on physical grounds that the rope will be:
straight to the left of the convex body;
continuously in contact with the body for a while, then
straight to the right of the body, and
there should be no kink in the rope at the points where it makes and then leaves contact with the body. This corresponds to 'smooth fit' in (iii).

6. Real options (Investment options).

For background and details, see e.g. Peskir & Shiryaev [PS], and
[DP] Avinash K. DIXIT and Robert S. PINDYCK: *Investment under uncertainty*. Princeton University press, 1994.

The options considered above concern financial *derivatives* (so called because they derive from the underlying fundamentals such as stock). We turn now to options of another kind, concerned with business decision-making. Typically, we shall be concerned with the decision of whether or not to make a particular investment, and if so, when. Because these options concern the real economy (of manufacturing, etc.) rather than financial markets such as the stock market, such options are often called *real options*. But because they typically concern investment decisions, they are also often called *investment* *options*. There is a good introductory treatment in [DP].

The key features are as follows. We are contemplating making some major investment – buying or building a factory, drilling an oil well, etc. While if the decision goes wrong it may be possible to recoup some of the cost, much or most of it will usually be irrecoverable (a *sunk cost* – as with an oil well). So the investment is *irreversible* – at least in part. Just as stock prices are uncertain – so we model them as random, using some stochastic process – here too, the future profitability of the proposed investment is *uncertain*. Finally, we do not have to act now, or indeed at all. So we have an open-ended – or infinite – time-horizon, $T = \infty$.

We may choose to delay investment,
(a) to gather more information, to help us assess the project, or
(b) to continue to generate interest on the capital we propose to invest.
So we must recognize, and feed into the decision process, the value of *waiting for further information*. When we commit ourselves and make the decision to invest, it is not just the sunk cost that we lose – we lose the valuable option to wait for new information.

This situation is really that of an *American call* option with an infinite time-horizon. With such an American *call*, we have the right to buy at a specified price at a time of our choosing (or indeed, not to buy). Following Dixit & Pindyck [DP, Ch. 5], we formulate an *optimal stopping problem*, and solve it as a *free boundary problem*, using the *principle of smooth fit*.

We suppose the cost of the investment is I, and that the value of the project is given by a GBM, $X = (X_t) \sim GBM(\mu, \sigma)$ (the value of a project is uncertain for the same reasons that stock prices are uncertain; we model them both as stochastic processes; GBM is the default option here, just as in the BS theory of Ch. IV). If we invest at time τ, we want to maximize

$$V(X) := \max_{\tau} \; E[(X_\tau - I)e^{-r\tau}],$$

with r the riskless rate (discount rate) as before. Now if $\mu \leq 0$ the value of the project will fall, so we should invest immediately if $X_0 > I$ and not invest if not. If $\mu > r$, the growth of X will swamp the investment cost I and more than offset the discounting, so we should invest and there is no point in waiting. So we take $\mu \in (0, r]$. We invest iff the value x^* at the time of investment is *large enough*; finding x^* is part of the problem; x^* is a *free boundary* (between the continuation region and the investment region).

We need the following four conditions:

$$\frac{1}{2}\sigma^2 x^2 V''(x) + \mu x V'(x) - rV = 0, \qquad (i)$$

$$V(0) = 0, \qquad (ii)$$

$$V(x^*) = x^* - I, \qquad (iii)$$

$$V'(x^*) = 1 \qquad \text{(smooth pasting)}. \qquad (iv)$$

For (i): this comes from the *generator* of the diffusion $GBM(r, \sigma)$ (cf. the *SDE* for $GBM(r, \sigma)$, and Black-Scholes *PDE*, VI.2); for details, see [DP Ch. 5], or Peskir & Shiryaev [PS, Ch. III]. For (ii) ("Nothing will come of nothing"): the GBM does not hit 0, but if it approaches 0, so will the value of the project, so (ii) follows from this by continuity). For (iii), this is the *value-matching condition*: on investment, the firm receives the net pay-off $x^* - I$. For (iv) (*smooth pasting*): think of a rope stretched tightly over a convex surface.

Again, the ODE (i) is *homogeneous* (cf. *Euler's theorem*). So we use a trial solution $V(x) = Cx^p$. So (i) gives that p satisfies the *fundamental quadratic*

$$Q(p) := \frac{1}{2}\sigma^2 p(p-1) + \mu p - r = 0.$$

The product of the roots is negative, and $Q(0) = -r < 0$, $Q(1) = \mu - r < 0$. So one root $p_1 > 1$ and the other $p_2 < 0$. The general solution is $V(x) = C_1 x^{p_1} + C_2 x^{p_2}$, but from $V(0) = 0$, $C_2 = 0$, so $V(x) = C_1 x^{p_1}, = Cx^{p_1}$ say. With x^* the critical value at which it is optimal to invest, (iii) and (iv) give

$$V(x^*) = x^* - I, \qquad V'(x^*) = 1.$$

From these two equations, we can find C and x^*. The second is

$$V'(x^*) = Cp_1 (x^*)^{p_1 - 1} = 1, \qquad C = (x^*)^{1-p_1}/p_1.$$

Then the first gives

$$C(x^*)^{p_1} = x^* - I, \qquad x^*/p_1 = x^* - I, \qquad x^* = \frac{p_1}{(p_1 - 1)} I.$$

The main feature here is the factor

$$q := p_1/(p_1 - 1) > 1$$

by which the value must exceed the investment cost I before investment should be made (q is used because this is related to "Tobin's q" in Economics). One can check that q increases with σ (the riskier the project, the more reluctant we are to invest), and also q increases with r (as then investing our capital risklessly becomes more attractive). Then the critical threshold above which it is optimal to invest is

$$x^* = qI.$$

Also

$$C = (qI)^{1-p_1}/p_1, \qquad V(x) = (qI)^{1-p_1} x^{p_1}/p_1.$$

The results above show that the traditional *net present value* (NPV – accountancy-based) approach to valuing real options is misleading – see [DP]. This is no surprise: our methods (*arbitrage pricing technique*, etc.) are superior to NPV!

7. Stochastic volatility (SV).

The Black-Scholes theory above – in discrete or continuous time – has involved the volatility – the parameter that describes the sensitivity of the stock price to new information, to the market's assessment of new information. Volatility is so important that it has been subjected to intensive scrutiny, in the light of much real market data. Alas, such detailed scrutiny reveals that volatility is not really constant at all – the Black-Scholes theory over-simplifies reality. (This is hardly surprising: real financial markets are more complicated than the contents of this course, as they involve *investor psychology*, rather than straight mathematics!) One way out is to admit that volatility is *random* (stochastic), and then try to model the stochastic process generating it. Volatility exhibits *clustering*, linked to *mean reversion*, so *Ornstein-Uhlenbeck* models are useful here. Such *stochastic volatility models* are topical today.

Stylised facts.

There are a number of *stylised facts* in mathematical finance. E.g.:

(i). Financial data show *skewness*. This is a result of the asymmetry between profit and loss (large losses are lethal!; large profits are just nice to have).

(ii). Financial data have much *fatter tails* than the normal (Gaussian). We have discussed this in II.5.

(iii) Financial data show *volatility clustering*. This is a result of the economic and financial environment, which is extremely complex, and which moves

between good times/booms/upswings and bad times/slumps/downswings. Typically, the market 'gets stuck', staying in its current state for longer than is objectively justified, and then over-correcting. As investors are highly sensitive to losses (see (i) above), downturns cause widespread nervousness, which is reflected in higher volatility. The upshot is that good times are associated with periods of growth but low volatility; downturns spark extended periods of high volatility (and economic stagnation, or shrinkage).
ARCH and GARCH.

We turn to models that can incorporate such features.

The model equations are (with Z_t ind. $N(0,1)$)

$$X_t = \sigma_t Z_t, \qquad \sigma_t^2 = \alpha_0 + \sum_1^p \alpha_i X_{i-1}^2, \qquad (ARCH(p))$$

while in $GARCH(p,q)$ the σ_t^2 term becomes

$$\sigma_t^2 = \alpha_0 + \sum_1^p \alpha_i X_{i-1}^2 + \sum_1^q \beta_j \sigma_{t-j}^2. \qquad (GARCH(p,q))$$

The names stand for (generalised) autoregressive conditionally heteroscedastic (= variable variance). These are widely used in Econometrics, to model *volatility clustering* – the common tendency for periods of high volatility, or variability, to cluster together in time. They were introduced in 1987 by Robert Engle (1942) and C. W. J. (Sir Clive) Granger (1934-2009), who received the Nobel Prize for this in 2003. From Granger's obituary (*The Times*, 1.6.2009): "Following Granger's arrival at UCSD in La Jolla, he began the work with his colleague Robert F. Engle for which he is most famous, and for which they received the Bank of Sweden Nobel Memorial Prize in Economic Sciences in 2003. They developed in 1987 the concept of cointegration. Cointegrated series are series that tend to move together, and commonly occur in economics. Engle and Granger gave the example of the price of tomatoes in N. and S. Carolina .... Cointegration may be used to reduce non-stationary situations to stationary ones, which are much easier to handle statistically and so to make predictions for. This is a matter of great economic importance, as most macroeconomic time series are non-stationary, so temporary disturbances in, say, GDP may have a long-lasting effect, and so a permanent economic cost. The Engle-Granger approach helps to separate out short-term effects, which are random and unpredictable, from long-term effects, which reflect the underlying economics. This is invaluable for macroeconomic policy formulation, on matters such as interest rates, exchange rates, and the relationship between incomes and consumption."

Volatility Modelling

In the standard Black-Scholes theory we have developed, volatility σ is *constant*. Thus a graph of volatility against strike K (or stock price S) should be flat. But typically it isn't, and displays curvature. Such volatility curves often turn upwards at both ends ('volatility smile'); there may well be asymmetry ('volatility smirk').

As above, it may be useful to model volatility stochastically, and use an SV model. However, the driving noise in this model will have a volatility of its own ('vol of vol'), etc. Practitioners often use *computer graphics* to represent *volatility surfaces* – the three-dimensional equivalents of graphs, where e.g. σ is graphed against K and S. The subject is too big to pursue further here; there is a good account (mixing theory with practice) in
J. GATHERAL: *The volatility surface: A practitioner's guide*. Wiley 2006.

Volatility is rough.

This is the title of an influential paper by Gatheral, Jaisson and Rosenbaum in 2014. The message there is that (log-)volatility is not only rough, it is *rougher than Brownian motion.* The reasons are the obvious ones: high-frequency trading, and order splitting. There is a family, *fractional Brownian motion*, with a parameter controlling the roughness (called the Hurst index): BM is 'in the middle'. This is highly topical today: there is a lot going on in this area.

Postscript.

1. One recent book on Financial Mathematics describes the subject as being composed of three strands:
arbitrage – the core economic concept, which we have used throughout;
martingales – the key probabilistic concept (Ch. III on);
numerics. Finance houses in the City use *models*, which they need to *calibrate to data* – a task involving both statistical and numerical skills, and in particular an ability to *programme*.

2. You will probably already have experience with at least one general/mathematical programming language (e.g., Matlab, Python) (if not: get it, a.s.a.p.!), and for Statistics, R. You may also know some Numerical Analysis, the theory behind computation. You may have encountered *simulation*, also known as *Monte Carlo*, and/or a branch of Probability and Statistics called *Markov Chain Monte Carlo (MCMC)* – computer-intensive methods for numerical solutions to problems too complicated to solve analytically. The leaders of R & D teams in the City need to be expert at both stochastic modelling (e.g., to propose new products), and simulation (to evaluate how these perform). Most of the ones I know use Matlab for this. At a lower level, quantitative analysts (quants) working under them need expertise in a computer language; C++ is the industry standard. If you are thinking of a career in Mathematical Finance, learn C++, as soon as possible, and for academic credit.

3. This course deals with *equity markets* – with *stocks*, and financial derivatives of them – options on stocks, etc. The relevant mathematics is *finite-dimensional*. Lurking in the background are *bond markets* ('money markets': bonds, gilts etc., where *interest rates* dominate), and the relevant options – *interest-rate derivatives*, and *foreign exchange* between different currencies ('forex'). The resulting mathematics (which is highly topical, and so in great demand in the City!) is *infinite-dimensional*, and so much harder than the equity-market theory we have done. However, the underlying principles are basically the same. One has to learn to walk before one learns to run, and equity markets serve as a preparation for money markets.

4. The aim of this lecture course is simple. It is to familiarize the student with the basics of Black-Scholes theory, as the core of modern finance, and with the mathematics necessary to understand this. The motivation driving the ever-increasing study of this material is the financial services industry and the City. I hope that any of you who seek City careers will find this introduction to the subject useful in later life. NHB, 2017