ull5a.tex **Week 5: am, 24.10.2018**

Stochastic integration (continued).

One can define stochastic integration in much greater generality.

1. *Integrands.* The natural class of integrands X to use here is the class of *predictable* processes. These include the left-continuous processes to which we confine ourselves above.

2. *Integrators.* One can construct a closely analogous theory for stochastic integrals with the Brownian integrator B above replaced by a *continuous local martingale* integrator M (or more generally by a *local martingale*: see below). The properties above hold, with D replaced by

$$E[(\int_0^t X_u dM_u)^2] = E[\int_0^t X_u^2 d\langle M\rangle_u].$$

See e.g. [KS], [RY] for details.

One can generalise further to *semimartingale* integrators: these are processes expressible as the sum of a local martingale and a process of (locally) finite variation. Now C is replaced by: stochastic integrals of local martingales are local martingales. See e.g. [RW1] or Meyer (1976) for details.

§6. Stochastic Differential Equations (SDEs) and Itô's Lemma

Suppose that U, V are adapted processes, with U locally integrable (so $\int_0^t U_s ds$ is defined as an ordinary integral, as in Ch. II), and V is left-continuous with $\int_0^t E[V_u^2]du < \infty$ for all t (so $\int_0^t V_s dB_s$ is defined as a stochastic integral, as in §5). Then

$$X_t := x_0 + \int_0^t U_s ds + \int_0^t V_s dB_s$$

defines a stochastic process X with $X_0 = x_0$. It is customary, and convenient, to express such an equation symbolically in differential form, in terms of the *stochastic differential equation (SDE)*

$$dX_t = U_t dt + V_t dB_t, \qquad X_0 = x_0. \qquad (SDE)$$

Now suppose that $f : \mathbb{R}^2 \to \mathbb{R}$ is a function, continuously differentiable once in its first argument (which will denote time), and twice in its second argument (space): $f \in C^{1,2}$. The question arises of giving a meaning to the stochastic differential $df(t, X_t)$ of the process $f(t, X_t)$, and finding it.

Recall the Taylor expansion of a smooth function of several variables, $f(x_0, x_1, \cdots, x_d)$ say. We use suffices to denote partial derivatives: $f_i := \partial f / \partial x_i$, $f_{i,j} := \partial^2 f / \partial x_i \partial x_j$ (recall that if partials not only exist but are continuous, then the order of partial differentiation can be changed: $f_{i,j} = f_{j,i}$, etc. – *Clairaut's theorem*). Then for $x = (x_0, x_1, \cdots, x_d)$ near u,

$$f(x) = f(u) + \Sigma_{i=0}^d (x_i - u_i) f_i(u) + \frac{1}{2} \Sigma_{i,j=0}^d (x_i - u_i)(x_j - u_j) f_{i,j}(u) + \cdots$$

In our case (writing t_0 in place of 0 for the starting time):

$$f(t, X_t) = f(t_0, X(t_0)) + (t - t_0) f_1(t_0, X(t_0)) + (X(t) - X(t_0)) f_2 + \frac{1}{2}(t - t_0)^2 f_{11} +$$

$$(t - t_0)(X(t) - X(t_0)) f_{12} + \frac{1}{2}(X(t) - X(t_0))^2 f_{22} + \cdots,$$

which may be written symbolically as

$$df(t, X(t)) = f_1 dt + f_2 dX + \frac{1}{2} f_{11} (dt)^2 + f_{12} dt dX + \frac{1}{2} f_{22} (dX)^2 + \cdots.$$

In this, we
(i) substitute $dX_t = U_t dt + V_t dB_t$ from above,
(ii) substitute $(dB_t)^2 = dt$, i.e. $|dB_t| = \sqrt{dt}$, from §4:

$$df = f_1 dt + f_2 (U dt + V dB) + \frac{1}{2} f_{11} (dt)^2 + f_{12} dt (U dt + V dB) + \frac{1}{2} f_{22} (U dt + V dB)^2 + \cdots$$

Now using $(dB)^2 = dt$,

$$\begin{aligned} (U dt + V dB)^2 &= V^2 dt + 2UV dt dB + U^2 (dt)^2 \\ &= V^2 dt + \text{higher-order terms}: \end{aligned}$$

$$df = (f_1 + U f_2 + \frac{1}{2} V^2 f_{22}) dt + V f_2 dB + \text{higher-order terms}.$$

Summarising, we obtain *Itô's Lemma*, the analogue for the Itô or stochastic calculus of the chain rule for ordinary (Newton-Leibniz) calculus:

Theorem (Itô's Lemma). If X_t has stochastic differential

$$dX_t = U_t dt + V_t dB_t, \qquad X_0 = x_0,$$

and $f \in C^{1,2}$, then $f = f(t, X_t)$ has stochastic differential

$$df = (f_1 + U f_2 + \frac{1}{2} V^2 f_{22}) dt + V f_2 dB_t.$$

That is, writing f_0 for $f(0, x_0)$, the initial value of f,

$$f(t, X_t)) = f_0 + \int_0^t (f_1 + U f_2 + \frac{1}{2} V^2 f_{22}) dt + \int_0^t V f_2 dB.$$

This important result may be summarised as follows: use Taylor's theorem formally, together with the rule

$$(dt)^2 = 0, \qquad dt dB = 0, \qquad (dB)^2 = dt.$$

Itô's Lemma extends to higher dimensions, as does the rule above:

$$df = (f_0 + \Sigma_{i=1}^d U_i f_i + \frac{1}{2} \Sigma_1^d V_i^2 f_{ii}) dt + \Sigma_1^d V_i f_i dB_i$$

(where U_i, V_i, B_i denote the ith coordinates of vectors U, V, B, f_i, f_{ii} denote partials as above); here the formal rule is

$$(dt)^2 = 0, \qquad dt dB_i = 0, \qquad (dB_i)^2 = dt, \qquad dB_i dB_j = 0 \qquad (i \neq j).$$

Corollary. $E]f(t, X_t)] = f_0 + \int_0^t E[f_1 + U f_2 + \frac{1}{2} V^2 f_{22}] dt$.

Proof. $\int_0^t V f_2 dB$ is a stochastic integral, so a martingale, so its expectation is constant ($= 0$, as it starts at 0). //

Note. Powerful as it is in the setting above, Itô's Lemma really comes into its own in the more general setting of semimartingales. It says there that if X is a semimartingale and f is a smooth function as above, then $f(t, X(t))$ is also a semimartingale. The ordinary differential dt gives rise to the finite-variation part, the stochastic differential gives rise to the martingale part. This closure property under very general non-linear operations is very powerful and important.

Example: The Ornstein-Uhlenbeck Process.

The most important example of a SDE for us is that for geometric Brownian motion (VI.1 below). We close here with another example.

Consider now a model of the velocity V_t of a particle at time t ($V_0 = v_0$), moving through a fluid or gas, which exerts
(i) a frictional drag, assumed proportional to the velocity,
(ii) a noise term resulting from the random bombardment of the particle by the molecules of the surrounding fluid or gas. The basic model is the SDE

$$dV = -\beta V dt + c dB, \qquad (OU)$$

whose solution is called the *Ornstein-Uhlenbeck* (velocity) process with *relaxation time* $1/\beta$ and *diffusion coefficient* $D := \frac{1}{2}c^2/\beta^2$. It is a stationary Gaussian Markov process (not stationary-increments Gaussian Markov like Brownian motion), whose limiting (ergodic) distribution is $N(0, \beta D)$ – the *Maxwell-Boltzmann distribution* – and whose limiting correlation function is $e^{-\beta|\cdot|}$.

If we integrate the OU velocity process to get the OU *displacement process*, we lose the Markov property (though the process is still Gaussian). Being non-Markov, the resulting process is much more difficult to analyse.

The OU process is the prototype of processes exhibiting *mean reversion*, or a *central push*: frictional drag acts as a restoring force tending to push the process back towards its mean. It is important in many areas, including
(i) statistical mechanics, where it originated;
(ii) mathematical finance, where it appears in the *Vasicek model* for the term-structure of interest-rates (the mean represents the 'natural' interest rate);
(iii) *stochastic volatility* models, where the volatility σ itself is now a stochastic process σ_t, subject to an SDE of OU type.

Theory of interest rates.

This subject (see MATL481 next semester) dominates the mathematics of *money markets*, or *bond markets*. These are more important in today's world than stock markets, but are more complicated, so we must be brief here. The area is crucially important in *macro-economic policy*, and in political decision-making, particularly after the financial crisis ("credit crunch"). Government policy is driven by fear of speculators in the bond markets (rather than aimed at inter-governmental cooperation against them). The mathematics is infinite-dimensional in principle (at each time-point t we have a whole *yield curve* over future times), but reduces to finite-dimensionality in practice: bonds are only offered at discrete times, with a *tenor* structure (a finite set of maturity times).

Mean reversion is used in models, to reflect the underlying 'natural interest rate', from which deviations may occur due to short-term pressures (pre-Crash – these may be longer-lasting nowadays, as we see post-Crash).

Chapter VI. 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

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

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. V: $dX_t = U_t dt + V_t dB_t$ 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

$$dX_t = X_t(\mu dt + \sigma dB_t), \qquad (GBM)$$

and is called *geometric Brownian motion* (GBM). We turn to its economic meaning, and the role of the two parameters μ and σ, below.

We recall the model of Brownian motion from Ch. V. 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 – domestic and foreign, 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 was 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 has now become standard:
SAMUELSON, P. A. (1965): Rational theory of warrant pricing. *Industrial Management Review* **6**, 13-39,
SAMUELSON, P. A. (1973): Mathematics of speculative prices. *SIAM Review* **15**, 1-42.
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 as-

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

set (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 *underlying 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.
Justification. For a treatment of this and other diffusion models via microeconomic arguments, see
[FS] FÖLLMER, H. & SCHWEIZER, M. (1993): A microeconomic approach to diffusion models for stock prices. *Mathematical Finance* **3**, 1-23.
Note. Observe the decomposition of what we are modelling into two components: a systematic component and a random component (driving noise). We have met such decompositions elsewhere – e.g. regression, and the Doob decomposition.

§2. The Black-Scholes Model

For the purposes of 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 given by

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

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 IV.1. A *trading strategy* H is a vector stochastic process

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

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 IV.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 martingale 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)

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

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

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 known as the *market price of risk*. This is important, both here (see below), in CAPM (I.3, Week 1a), 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. IV, 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 – VI.2) give $d\tilde{V}_t(H) = H_t.\sigma \tilde{S}_t dW_t$ (VI.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 IV.4.
(f) From RNVF, we can obtain BS, by integration, as in IV.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. IV) 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 4b 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)

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

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. This is *Girsanov's theorem*, below (Igor Vladimirovich GIRSANOV (1934-67) in 1960). As this involves a martingale condition, we pause to note that this is satisfied in the case that concerns us, when the drift μ_t is constant, $\mu_t \equiv \mu$. This involves the exponential mg of Problems 4b Q3:

Exponential martingale.

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, if $Z \sim N(0,1)$,

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

(normal MGF). Also,

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

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

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.

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) is formulated in varying degrees of generality, and proved, in [KS, §3.5], [RY, VIII].