Umeå University, January, 2017

Stochastic Differential Equations and Stochastic Optimal Control for Economists: Learning by Exercising

by Karl-Gustaf Löfgren

These notes originate from my own efforts to learn and use Ito-calculus to solve stochastic differential equations and stochastic optimization problems. Although the material contains theory and, at least, sketches of proofs, most of the material consists of exercises in terms of problem solving. The problems are borrowed from textbooks that I have come across during my own attempts to become an amateur mathematician. However, my Professors Björk, Nyström and Öksendal have done extremely work to help me.

Tomas Björk and Bernt Öksendal have made my help with the books Arbitrage Theory in Continuous Time and Stochastic Differential Equations, and they are excellent. Kaj Nyström is a fantastic mathematician and teacher, and he has been one of the very best to help me with the booaks from Björk and Öksendal.

To my teachers Professor Tomas Björk, Professor Kaj Nyström and Professor Bernt Öksendal

Keywords: Stochastic differential equations, stochastic optimal control and finance.

JEL-codes: C61, C73

Two of my friends in economics are Professors Thomas Aronsson, Umeå and Chuang-Zhong Li, Uppsala. Both of them can use stochastic differential and stochastic optimal controls. They have helped me with both economics and stochastic differential equations in twenty years.
Thanks from
Karl-Gustaf

Table of contents

Chapter 0: Introductory remarks

These notes originate from my own efforts to learn and use Ito-calculus to solve stochastic differential equations and stochastic optimization problems. Although the material contains theory and, at least, sketches of proofs, most of the material consists of exercises in terms of problem solving. The problems are borrowed from textbooks that I have come across during my own attempts[1] to become an amateur mathematician. I have learnt a lot from the following texts; Åström (1970), *Introduction to Stochastic Control Theory*, Mangel (1985), *Decision and Control in Uncertain Resource Systems, Malliaris and Brock (1991), Stochastic Methods in Mathematics and Finance, Björk (1994), Stokastisk kalkyl och kapitalmarknadsteori, Björk (1998/2004,2009), Arbitrage Theory in Continuous Time, and Öksendal (2003), Stochastic Differential Equations: An Introduction with Applications.* My pedagogical favorite is the book by Björk (1998/2004) and theory, and for that matter the exercises, are best developed in Öksendahl (2000/ 2003).

The audience I have had in mind is graduate students in Economics and possibly also Finance, who want some insights into the application of Ito calculus to applied problems in their own field. My hope is that the notes can be used as complement to a more comprehensive textbook in a shorter graduate course on the topic. The recommended background textbooks are Björk (1998/2004) and Öksendal (2003) where most of the stringency is.

The text is structured as follows: It starts with a fairly short chapter on continuous time stochastic processes where the reader gets a first taste of Brownian motion (or, which amounts to the same, Wiener Processes) and Ito Calculus.

The second chapter on stochastic integrals deals more explicitly with Ito calculus and the properties of stochastic integrals. Martingales are introduced and it is shown under what

---

[1] Most problems are collected from exercise sets that do not contain explicit solutions. I will typically indicate from where the problem was collected. However, many problems are generic in the sense that they pop up in all text-books. The reason is often that the equation and/or the optimization problem are possible to solve analytically, at the same time as they have relevance for applied work. I have benefited a lot from the problem demonstations conducted by my Math teacher Professor Kaj Nyström, Dept of Mathematics, Umeå University and now Uppsala University. Without Kaj, these notes would not exist. Any errors and the mathematical amateurism should be blamed on the author.

conditions a stochastic integral or, which amounts to the same thing, a stochastic differential equation, is a Martingale. In particular, it is demonstrated how Ito Calculus and Martingales can be used to compute mathematical expectation. The chapter, like all chapters, ends with a section with solutions to exercises that illustrate the new technicalities that have been introduced.

The third chapter deals with how to solve stochastic differential equations, and how the Dynkin - Ito - Kolmogorov operator combined with Feynman-Kac-representation theorem can be used to solve partial differential equations. Again, the text is complemented with problems that are solved at the end of the chapter.

The forth chapter introduces stochastic optimal control by using a workhorse model in terms of a stochastic optimal growth problem. We introduce the relevant theorems connected with the Hamilton-Jacobi-Bellman equation, and we, in particular, solve a fair number of stochastic optimal control problems.

In chapter five, we apply results from diffusion theory like Dynkin’s formula, Feynman-Kacs’ formula, and certain variational inequalities to solve optimal stopping problems. Again, solving exercises is the most important learning device.

Finally, in Chapter 6 we give the reader a taste of Financial Economics by essentially deriving one of the most well known results; Black and Schools formula. This is done by introducing the notion of an absolutely continuous probability measure, and Girsanov’s theorem.

Chapter 1: Continuous-time stochastic processes

This chapter contains a brief, and rather non-stringent, introduction of some of the mathematical tools that are necessary for the preceding analysis.

A stochastic process is a variable, $X(t)$, that evolves over time in a way that is - at least - to some extent random. In economic modeling, continuous-time stochastic processes are typically used in Capital Theory and Financial Economics. The most widely studied continuous time process is a Brownian motion. The name originates from the English botanist Robert Brown who in 1827 observed that small particles immersed in a liquid exhibit ceaseless irregular motions. Einstein (1905) is generally given credit for the precise mathematical formulation of the Brownian motion process (the paper was instrumental for understandung that the atom exist), but an even earlier equivalent formulation was set down by Louis Bachelier (1900) in his theory of stock option pricing.

A stochastic process $X(t)$ is characterized by its distribution function $G(x, t)$:

$$Prob\{X(t) \leq x\} = G(x,t) \tag{1.1}$$

According to equation (1.1) the probability of finding the process $X(t)$ not above some level x at time t is given by the value of the (possibly time dependent) distribution function evaluated at x. If the derivative $\partial G(x,t)/\partial x = g(x,t)$ exists, it can be used to characterize $X(t)$ as follows:

$$\begin{aligned} &Prob\{x \leq X(t) \leq x+dx\} = G(x+dx,t) - G(x,t) = \\ &= \left\{G(x,t) + \frac{\partial G}{\partial x}(x,t)dx + O(dx)\right\} - G(x,t) = g(x,t)dx + O(dx) \end{aligned} \tag{1.2}$$

The second equality in equation (1.2) follows from a first order Taylor expansion of $G(\cdot)$ around the point x. Here $O(dx)$ denotes terms that are of higher order than dx and, therefore, can be ignored when dx is small. More specifically, a term is of order $O(dx)$, *if*

$\lim_{dx \to 0} O(dx)/dx = 0$. The function $g(x, t) = \partial G(x, t)/\partial x$ is the density function evaluated at $X = x$.

A Brownian motion, $B(t)$, or a Wiener process, is a stochastic process with the following properties:

(i) the sample paths of $B(t)$ are continuous

(ii) $B(0) = 0$

(iii) the increment $B(t+\tau) - B(\tau)$ is normally distributed with mean zero and variance $\sigma^2 t$.

(iv) if (t, τ) and (t^1, τ^1) are disjoint intervals, then the increments $B(\tau) - B(t)$, and $B(\tau^1) - B(t^1)$ are independent random variables.

Let $dB(t) = B(t + dt) - B(t)$. Then, if we denote the standard normal density function by $\phi(\cdot)$, the normality of the increments implies that

$$\Pr\text{ob}[\beta \le dB \le \beta + d\beta] = \frac{1}{\sqrt{\sigma^2 dt}}\phi(\beta)d\beta = \left(\frac{1}{\sqrt{2\pi\sigma^2 dt}}\right)\exp\left(\frac{-\beta^2}{2\sigma^2 dt}\right)d\beta \tag{1.3}$$

for a sufficiently small $d\beta$. Moreover, the first two moments of the distribution are

$$E\{dB\} = 0 \qquad E\{(dB)^2\} = \sigma^2 dt \tag{1.4}$$

The variance of the increment dB is of order dt (proportional to the small interval dt). This gives rise to mathematical complications. To see this, dividing both sides of the expression for the variance by $(dt)^2$ we obtain

$$E\left\{\left(\frac{dB}{dt}\right)^2\right\} = \frac{\sigma^2}{dt} \to \infty \qquad \text{as } dt \to 0 \tag{1.5}$$

Meaning that $B(t)$ is not differentiable, but nevertheless everywhere continuous.

The fourth condition on the increments of a Brownian motion process is frequently referred to as the Markov property. This reflects a kind of lack of memory, in the sense that that the past history of a process does not influence its future position. The requirement of independent increments, however, is more restrictive than to require that the"future" state only depends on the present state, which is the true Markov property.

There are other special features of a Brownian process. To exemplify, let the capital stock $K(t)$ follow a Brownian motion, i.e., $E[dK] = 0$ and $E(dK^2) = \sigma_K^2 dt$, where $\sigma_K^2 dt$ is the variance of the increments in the capital stock. Let the production function be $Y = F\left[t, K(t)\right]$. Estimating dY at (t, K) for changes dt and dK by a second order Taylor expansion yields

$$dY = \frac{\partial F}{\partial t}dt + \frac{\partial F}{\partial K}dK + \frac{1}{2}\left[\frac{\partial^2 F}{\partial t^2}(dt)^2 + 2\frac{\partial^2 F}{\partial t \partial K}dtdK + \frac{\partial^2 F}{\partial K^2}(dK)^2\right] \tag{1.6}$$

Since $K(t)$ is stochastic so is Y, and the differential dY therefore makes sense in terms of moments or distributions. Taking expectations of (1.6) conditional on $K(t) = k$ gives

$$E\left\{dY \middle| K(t) = k\right\} = \frac{\partial F(t,k)}{\partial t}dt + \frac{1}{2}\frac{\partial^2 F(t,k)}{\partial K^2}\sigma_K^2 dt + O(dt) \tag{1.7}$$

The first second order derivative within brackets in (1.6) is merged in the term O(dt), while the second vanishes because $E(dK) = 0$. The third term within brackets, which contains the second derivative of the production function times the variance of dK, is introduced since it can be shown that $E[dK^2]$ (under Ito Calculus) is of order dt rather than $(dt)^2$. Therefore, the expected change in production over the short interval dt consists of two terms. The first can be interpreted as technological progress, and the second measures the effect of an additional unit of capital on the marginal product of capital, which is scaled by $E\left[dK^2\right] = \sigma_K^2 dt$. This term is presumably non-positive, since production functions are usually assumed to be strictly concave. The interpretation is that the uncertainty of K is greater, the longer the time horizon. The expected value of a change in a strictly concave function is thus reduced by an amount that increases with time - a consequence of Jensen's inequality $E\left[f(x)\right] \underset{(>)}{<} f\left(E(x)\right)$ for a strictly concave (convex) function.

The Brownian motion induces a new calculus', and one of them is known as the Ito calculus after its inventor; see Ito (1944, 1946)[2] . This is expressed by a first differential that is generated by second order terms:

$$dY = \frac{\partial F}{\partial t}dt + \frac{\partial F}{\partial K}dK + \frac{1}{2}\frac{\partial^2 F}{\partial K^2}\sigma_K^2 dt + \mathrm{O}(dt) \tag{1.8}$$

where $\lim_{dt \to 0} \mathrm{O}(dt)/dt = 0$.

Equation (1.8), which measures the first order differential of a function containing a stochastic variable that follows a Brownian motion process, is frequently referred to as Ito's lemma (or Ito's formula).

We can be more precise about the stochastic process by specifying the following general Brownian motion process[3]

$$dK = a(K,t)dt + b(K,t)dB \tag{1.9}$$

Here $a(K, t)$ and $b(K, t)$ are known non-random functions, which are usually referred to as the drift and variance components of the process; dB is the increment of the process, and it holds that $E[dB] = 0$ and $E[dB^2] = dt$. This means that dB can be represented by $dz = \varepsilon\sqrt{dt}$ where $\varepsilon \sim N(0, 1)$. Substitution of (1.9) into (1.8) now gives

$$dY = \left[\frac{\partial F}{\partial t} + a(K,t)\frac{\partial F}{\partial K} + \frac{1}{2}b^2(K,t)\frac{\partial^2 F}{\partial K^2}\right]dt + b(K,t)\frac{\partial F}{\partial K}dB + O(dt) \tag{1.10}$$

Note that

[2] An alternative way of defining a stochastic integral under Brownian motion was introduced by Stratonovich (1966). It results in a more conventional, but perhaps less practical, calculus.

[3] A process whose trend and volatility are functions of the state is often referred to as a diffusion process.

$$dK^2 = a^2dt^2 + 2abdtdB + b^2dB^2$$
$$= b^2dB^2 + O(dt) = b^2dt + O(dt) \qquad (1.11)$$

since $dtdz = \varepsilon\, dt^{3/2} \propto dt^{3/2}$, and $dz^2 = \varepsilon^2\, dt \propto dt$ (the sign $\propto$ means ”proportional to”).

To introduce a more specific example, let $Y = \ln K$, and let dK follow a Brownian motion of the following shape[4]

$$dK = \alpha K dt + \sigma K dB \qquad (1.12)$$

We now have

Moreover, $a(K, t) = \alpha K$ and $b(K, t) = \sigma K$, which substituted into (1.10) yields[5]

$$dY = (\alpha - \frac{\sigma^2}{2})dt + \sigma dB \qquad (1.13)$$

However, over any finite interval ΔT, the change in ln K is normally distributed with mean $(\alpha - \frac{\sigma^2}{2})\Delta T$ and variance $\sigma^2\Delta T$. Again, the reason why the expected value of the change in production grows more slowly than the drift in the capital accumulation equation is the strict concavity of the production function.

---

[4] Samuelson (1965) called this specific process geometric Brownian motion with drift.

[5]Terms of magnitude O(dt) are ignored in equation (1.13).

Chapter 2: Continuous stochastic processes and Ito integrals

2.1 Technicalities ending up in a Brownian motion

We have already in Chapter 1, perhaps prematurely, defined a Wiener process (Brownian motion). As the reader may remember the increments of the Wiener process are assumed to be independent normally distributed stochastic variables with mathematical expectation zero and variance $\sigma^2 (t\text{-}s)$.

Introducing this definition just like that is, of course, a rough short-cut. A formal definition of a stochastic process, for example, requires measure theory and other concepts from general probability theory. These concepts are hardly necessary to solve stochastic optimization problems in practice. However, some of them can be worthwhile to have seen. Two of the fundamental concepts are introduced in the following definition[6]

Definition 2.1: (σ - algebra and measurable space) If Ω is a given set, then a σ - algebra F on Ω is a family F *of subsets of Ω with the following properties*

(i) $\emptyset \in F$

(ii) $\mathrm{F} \in F \Rightarrow \mathrm{F}^C \in F$ *where* $\mathrm{F}^C = \Omega/\mathrm{F}$ is the complement of F in Ω

(iii) $A_1, A_s, \ . \ . \ . \in F \Rightarrow \bigcup_{i=1}^{\infty} A_i \in F$

The pair (Ω, F) is called a measurable space. A probability measure P on a measurable space (Ω, F) is a function P: $F \rightarrow [0,1]$ such that

a) $P(\emptyset) = 0$, $P(\Omega) = 1$

b) If $A_1, A_s, \ . \ . \ . \in F$ and $\{A_i\}_i^{\infty}$ are disjoint sets ($A_i \cap A_j = 0$), then $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$

The triplet (Ω, F, P) is called a probability space.

The first condition (i) in the definition means that the empty set as well as set Ω belongs to the σ - algebra F . Condition (ii) tells us that if a subset F (event) belongs to the σ - algebra

[6] The definitions are borrowed from Öksendal (2003).

then its complement also belongs to the σ - algebra. Finally, if the members of an infinite sequence of subsets (events) A_i belong to *F*, then the union of the sequence of sets, A_i, belong to *F*. The fact that the space (Ω, F) is measurable, means in a probability context, that a set *A* that belong to *F* are interpretable as events, and *P(A)* is the probability that event *A* occurs. In particular, if *P*(F)=1, we say that F occurs with probability one, or almost surely (a.s.).

A particular important σ algebra is the Borel σ algebra. Given a family *U* of subsets of Ω, the Borel σ algebra is defined as the smallest σ algebra, H_U, that contains U. It is formally defined as

$$H_U = \bigcap \{H\,; H \ \sigma \ algebra \ of \ \Omega, U \subset H\}$$

The elements of the Borel algebra are called Borel sets.

A random variable X is a *F* measurable function $X : R^n \to R$. The random variable induces a probability measure *G(x) on* R^n called the distribution of *X*. Moreover, given that the first integral in equation 2.1 below converges, one has

$$E[f(X)] = \int_{\Omega} f(X(\omega))dP(\omega) = \int_{R^n} f(x)dG(x) \tag{2.1}$$

which is the mathematical expectation of $f(X)$ with respect to *P*. Here $\omega \in \Omega$ is a random variable $\omega \to X(\omega)$, which can be interpreted as an event in the $\Omega -$space.

We conclude the section of formal definitions by introducing the definition of stochastic process

Definition 2.2: (Stochastic process) A stochastic process is a parameterized collection of random variables

$$\{X_t\}_{t \in T}$$

defined on a probability space (Ω, F, P) *with values in* R^n.

The parameter space T is typically the half-line $[0,\infty)$ interpreted as time. For t fixed we have a *stochastic variable* $\omega \to X_t(\omega)$, $\omega \in \Omega$. If we fix ω we can define the function $t \to X_\omega(t)$, where $t \in T$. The latter function is called the time path of X_t. In this context ω can be looked upon as an experiment. An alternative way of writing a *stochastic process* is to write it as $(t,\omega) \to X(t,\omega)$, *where* $(t,\omega) \in Tx\Omega$. This is a convenient way of writing the stochastic process, since it typically has to be measurable in (t,ω).

The founder of modern probability theory Kolmogorov has shown under what conditions there exists a probability space (Ω, U, P) and a stochastic process $\{X_t\}$ on Ω, $X_t : \Omega \to R^n$) such that $P(X_{t_1} \in U_1, X_{t_2} \in U_2 .... X_{t_k} \in U_k) = \nu_{t_1,t_2...t_k}(U_1 x U_2 x...x U_k)$, where $\nu_{t_1,t_2...t_k}(U_1 x U_2 x...x U_k)$ is the finite dimensional distribution of $\{X_t\}_{t \in T}$.
$U_i, i = 1....k$, are Borel sets (events). Here we skip these conditions, and for that matter the proof. The Theorem that can be found in Öksendahl (2003) Chapter 2 and it is referred to as Kolmogorov's extension theorem.

An important representative of a stochastic process is The Brownian motion or Wiener process touched upon already in chapter 1. Brown observed that pollen grains suspended in liquid performed an irregular motion that was later explained by the random collision with molecules in the liquid. Mathematically, it turned out to be convenient to model this process as a stochastic process, $B_t(\omega)$, satisfying the conditions stipulated in Definition 1.1 above. Here $B_t(\omega)$ is interpreted as the position of the grain ω at time *t*.
To construct the Brownian motions process on R (the real line) we use Kolmogorov's (extension) theorem and fix $x \in R$ and define

$p(t,x,y) = (2\pi t)^{-0.5} \exp(-\frac{(x-y)^2}{2t})$ for $y \in R, t > 0$, which is the density of the univariate normal distribution. Now, if $0 \le t_1 \le t_2 \le .... \le t_k$, we define a measure on R by

$$\nu_{t_1,t_2...t_k}(U_1 x U_2 x...x U_k) = \int_{U_1 x U_2 .. x . U_k} p(t_1 x, x_1) p(t_2 - t_1, x_1, x_2) ... p(t_k - t_{k-1}, x_{k-1}, x_k) dx_1 dx_2 ... dx_k \quad (2.2)$$

From Kolmogorov's Theorem and the properties of the chosen distribution it follows that

$$\int\limits_{U_1 x U_2 x \ldots x U_k} p(t,x,x_1)p(t_2-t_1,x_1,x_2)\ldots p(t_k-t_{k-1,}x_{k-1},x_k)dx_i dx_2,\ldots d_k =$$

$$= P(X_{t_1} \in U_1, X_{t_2} \in U_2 \ldots . X_{t_k} \in U_k) \qquad (2.3)$$

This process is called *a Brownian motion process starting at* x*. Note that the probability* $P(X_0 = x) = 1$. This definition does unfortunately not result in a unique process and not even in a continuous process. Kolmogorov has, however, a solution also to this problem. Given that the process satisfies a growth condition on the mathematical expectation of the increments, there exists a continuous version of the stochastic process. The Brownian motion process satisfies this condition.

We are now moving to some exercises. The problems are borrowed from *Öksendahl*(2003), but not all the solutions.

Exercises

(Öksendal2003)):

2.1 Let X: $\Omega \rightarrow R$ *be a random variable. The distribution function F of X is defined by F(x)=P(X*$\leq x$*)*

Prove that F has the following properties

(i) $0 \leq F \leq 1$, $\lim\limits_{x \to -\infty} F(x) = 0$, $\lim\limits_{x \to \infty} F(x) = 1$

(ii) F *is increasing (=nondecreasing)*

(iii) F is right continuous, i.e. F(x)=$\lim\limits_{h \to 0} F(x+h)$ *for h>0.*

Solution: The first two statements are more or less trivial but perhaos not so easy to prove in a stringent manner. I guess that one can use that density function has support on the rweal line. To prove (iii) we note that F(x)$\leq$*F(x+h) for all h>0 from (ii). The limit also exists since F is non-decreasing and bounded from below. Call the limit* $\tilde{F}(x)$*. Now assume that* $\tilde{F}(x_0) > F(x_0)$ *, which means that*

$$F(x_0+h) - F(x_0) > \delta \textit{ iff } P(x_0 < x \leq x_0 + h) > \delta$$

Now, when $h \to 0$, *we obtain* $\widehat{F}(x_0) - F(x_0) = P(x_0 < x \le x_0) > 0$ *which is a contradiction.*

2.2 Let g: $\mathrm{R} \to \mathrm{R}$ *be measurable such that* $E[|g(X|] < \infty$. *Prove for the case when g is simple (a step function) that E[g(X)] =* $\int_{-\infty}^{\infty} g(x) dF(x)$

Solution: Since g is simple we can write E[g(x)] = $\sum_{i=1}^{\infty} \varepsilon_i \mathrm{X}_{a_i}^{a_{i+1}}(x)$ *for constants* ε_i, a_i.

Hence E[g(x)] = $\sum_{i=1}^{\infty} \int_{\Omega} \varepsilon_i \mathrm{X}_{a_i}^{a_{i+1}}(x(\omega)) dP(\omega)$ *(A)*

Now let $B_i = \{\omega : x(\omega) \in (a_i, a_{i+1}]\}$, *where* B_i *measurable. Hence,*

$$A = \sum_{i=1}^{\infty} \varepsilon_i \int_{B_i} dP(\omega) = \sum_{i=1}^{\infty} \varepsilon_i P(B_i) = \sum_{i=1}^{\infty} \varepsilon_i P(\omega; a_i < x(\omega) \le a_{i+1}) =$$

$$\sum_{i=1}^{\infty} \varepsilon_i [F(a_{i+1}) - F(a_i)] = \int_{-\infty}^{\infty} g(x) dF(x)$$

2.3 Find the density of $P(B_t^2 \le \sigma)$ *where* B_t *is Brownian motion.*

Solution: $\sigma \le 0$ *implies P(•)=0 Assume* $\sigma > 0$. *This means that*

$$P(B_t^2 \le \sigma) = P[-\sqrt{\sigma} < B_t \le \sqrt{\sigma}] = \frac{1}{\sqrt{2\pi t}} \int_{-\sqrt{\sigma}}^{\sqrt{\sigma}} e^{-x^2/2t} dx = \text{(from}$$

$$\text{symmetry)} \frac{2}{\sqrt{2\pi t}} \int_{0}^{\sqrt{\sigma}} e^{-x^2/2} dx = [\text{when } y = x^2] = \frac{1}{\sqrt{2\pi t}} \int_{0}^{\sigma} e^{\frac{-y}{2t}} y^{-0.5} dy$$

i.e., $B^2(t)$ *has density* $\dfrac{1}{\sqrt{2\pi t}} \dfrac{1}{\sqrt{y}} e^{\frac{-y}{2t}}$.

2.4 Let X, Y: $\Omega \to R$ *be two independent bounded random variables. Prove that E[XY] = E[X]E[Y]* $*$.

Solution: Put $|X| \le M$, $|Y| \le N$, $x = X + M$ *and y=Y+N. Assume that the result is true for x,y.- This implies that E[(X+M)(Y+N)]=E[X+M]E[Y+N]*

$\Leftrightarrow E[XY] + NE[X] + ME[Y] + MN = E[X]E[Y] + NE[X] + ME[Y] + MN$ *implying that E[XY]=E[X]E[Y]; in other words,* $*$ *holds for X and Y.*

Now assume that $0 \le x < M, 0 \le y < N$ *and choose K large,* $a_j = \frac{M}{K} j$ *for j=0,1,…,k.*

Define $\rho_K(\omega) = \sum_{j=0}^{k-1} a_j \mathrm{X}_{F_j}(\omega)$ *where*

$F_j = \mathrm{X}^{-1}([a_j, a_{j+1}])$

$$\mathrm{X}_{F_j}(\omega) = \begin{Bmatrix} 1 \mathit{if}\ \omega \in F_j \\ 0\, \mathit{otherwise} \end{Bmatrix}$$

Now, we define $E[x - \rho_K] = \int_\Omega [x - \rho(\omega)] dP(\omega) = (\mathit{with} \Omega = \cup F_j, F_i \cap F_j = \phi, j \ne i)$

$$= \sum \int [(x - \rho_K(\omega)] dP(\omega) \le \{(x - a_j) \ge 0\, \mathit{on}\ F_j\ \mathit{by\, construction}\} \le \sum (a_{j+1} - a_j) P(F_j) = \frac{M}{K}$$

For $K \to \infty$, *E[x-* ρ_K *]* $\to 0$.

In the same spirit we construct $\varphi_k(\omega)$, *such that* $E[y - \varphi_k(\omega)] \to 0$ *when* $k \to \infty$

Now consider

E[$\rho_K(\omega)\varphi_k(\omega)] = \sum_j \sum_i a_j b_i P(F_j \cup G_i) = (\mathit{by\, indep}) = \sum_j \sum_i a_j b_i P(F_j) P(G_i) = E[\rho_K] E[\varphi_k]$

Now

E[XY]=

$$E[(X - \rho_K)Y] - E[\rho_K Y] = E[(X - \rho_K)Y] + E[\rho_K(Y - \varphi_k)] + E[\rho_K \varphi_k] = E[(X - \rho_K)Y] + E[\rho_K(Y - \varphi_k)] +$$

$$+ E[\rho_K] E[\varphi_k]$$

For $K \to \infty, k \to \infty$ *we obtain E[XY]=* $\lim_{K \to \infty} E[\rho_K] \lim_{k \to \infty} E[\varphi_k] = E[X]E[Y]$

2.5 Let B_t *be a Brownian motion on R,* $B_0 = 0$. *Prove that*

$$E[e^{iuB_t}] = \exp(-\frac{1}{2} u^2 t) \quad \mathit{all}\ u \in R$$

Here i= $\sqrt{-1}$ *. Use that since Brownian motion is a Gaussian process it holds that:*

$E[e^{iuB_t}] = e^{-\frac{1}{2}u^2c+iuM}$ *See Öksendal (2003) p 13.*

Solution: Identification of coefficients yields that $E[B_t^2] = c = t$ *and M=* $E[B_t]$ *=0.*

2.6 Let B_t *be a Brownian motion and fix* $t_0 \geq 0$*. Prove that* $\tilde{B}_t = B_{t_0+t} - B_{t_0}$ *,* $t \geq 0$ *is a Brownian motion.*

Solution: What we have to show is that we can write the new process, $\tilde{B}_t$*, as equation 2.3 above Assume that .* $B_0 = x_0$*. Now*

$$P(B_{t_1} \in F_1.......B_{t_k} \in F_k) = \int_{-\infty}^{\infty} P[(B_{t_0+t_1} - B_{t_0}) \in F_1.......(B_{t_0+t_k} - B_{t_0}) \in F_k / B_{t_0} = x_0] p(t_0; x_0, x) dx_0 =$$

$$\int_{-\infty}^{\infty} \int_{F_1+x_0,....F_k+x_0} p(t_1, x_0, x_1).......p(t_k - t_{k-1}, x_{k-1}, x_k) p(t_0, x, x_0) dx_1....dx_k dx_0 = (y = x - x_0) =$$

$$\int_{-\infty}^{\infty} \int_{F_1..F_k} p(t_1, 0, y_1)....p(t_k - t_{k-1}, y_{k-1}, y_k) p(t_0, x, x_0) dy_1...dy_k dx_o =$$

$$\int_{F_1..F_k} p(t_1, 0, y_1)....p(t_k - t_{k-1}, y_{k-1}, y_k) dy_1...dy_k$$

2.6 Extra: Reflect on what this tells us about the importance of starting the Brownian motion process at zero.

2.7 Let B_t *be an n-dimensional Brownian motion starting at zero and let* $U \in R^{nxn}$ *be a constant orthogonal matrix, i.e.* $UU^T = I$ *. Prove that* $\tilde{B}_t = UB_t$ *is also a Brownian motion process.*

Solution: $P[\tilde{B}_{t_1} \in F_1....\tilde{B} \in F_k] = P^0[B_{t_1} \in U^{-1}(F_1),....B_{t_k} \in U^{-1}(F_k)]$ *=(where*

$$U^{-1}(F_j) = [U^{-1}y, y = (y_1....y_n) \in F_j]) = \int_{U^{-1}(F_1)...U^{-1}(F_k)} p(t_1, 0, x_1)....p(t_k - t_{k-1}, x_{k-1}, x_k) dx_1...dx_k \, *$$

Let $z = Ux$ *be n-dimensional vectors. Since* $UU^T = I$ *, det U=1, implying that* $dz = dx$

Moreover, $p(t,x_{k-1},x_k)=\frac{1}{(2\pi t)^{0.5}}e^{-\frac{|x_k-x_{k-1}|^2}{2t}}$ *. Since* $UU^T=I$ *it is also true that*

$|x_k-x_{k-1}|^2=|z_k-z_{k-1}|^2$. *This proves the claim that* $*=P[B_{t_1}\in F_1,...B_{t_k}\in F_k]$.

2.7Extra: Why is orthogonality important to preserve Brownian motion under linear transformations?

2.8. Let B_t *be a one dimensional Brownian motion. For c>0 prove that* $B_t^*=\frac{1}{c}B_{c^2t}$ *is also a Brownian motion.*

Solution: Let $B_{t_0}=x,\ B_t^*\in F_1\Leftrightarrow\frac{1}{c}B_{c^2t}\in F_1.\quad cF_1=\{y;y=cy^*,y^*\in F_1\}$

Now $P^*(B_{t_1}^*\in F_1....B_{t_k}^*\in F_k)=\int_{cF_1...cF_k}p(c^2t_1,x,x_1)....p(c^2(t_k-t_{k-1}),x_{k-1},x_k)dx_1..dx_k$

$x_j=cy_j\to[dx_1,...,dx_k]=[cdy_1,...,cdy_k]$

This means that

$p(c^2(t_k-t_{k-1}),cy_{k-1},cy_k)=\frac{1}{c}p((t_k-t_{k-1}),y_{k-1},y_k)$ *(normal distribution)*

We obtain

$$P^{*x/c}(B_{t_1}^*\in F_1.....B_{t_k}^*\in F_k)=P^x(B_{t_1}\in F_1....B_{t_k}\in F_k)=\int_{F_1...F_k}p(t_1,x/c,x_1)....p(t_k-t_{k-1},x_{k-1},x_k)dx_1..dx_k$$

2.2 Stochastic integrals-Ito and Stratonovich

From chapter one, we know that that a Brownian motion process is continuos but nowhere differentiable. This will lead to some complications that are related to how we define stochastic integrals. Say that we start from the following stochastic differential

$$\begin{aligned}dX(t)&=\mu(t,X(t))dt+\sigma(t,X(t))dB(t)\\X(0)&=x_0\end{aligned}\tag{2.4}$$

which can be interpreted as a stochastic integral equation of the following shape

$$X(t)=x_0+\int_0^t\mu(s,X(s))ds+\int_0^t\sigma(s,X(s))dB(s)\tag{2.5}$$

Since the first integral contains no stochastic components it can, given continuity of $\mu(s)$, be interpreted as a Rieman integral, but the second integral is more problematic. The reason is, loosely speaking, that the normally distributed increments mean that the $B(t) - trajectories$ are locally unbounded[7] . This implies that we need are more restrictive definition of the stochastic part of the integral equation.

In other words, we will have to find a suitable definition of integrals of the type

$$\int_0^t g(s)dB(s) \tag{2.6}$$

and also develop the corresponding differential calculus. Hence, we will have to deal with differential equations of a "new type". To this end, we introduce the following definition:

Definition 2.3: Let F_t^X *denote the "information" generated by the process X(s) on the interval* $[0,t]$. $A \in F_t^X$ *means that we can decide whether A has occurred in the interval* $[0,t]$.

If the stochastic variable Y *can be completely determined given observations of the tradjetory* $\{X(s); 0 \le s \le t\}$, *then we write* $Y \in F_t^X$ *. If Y(t) is a stochastic process such that* $Y(t) \in F_t^X$ *for all* $t \ge 0$, *we say that* Y *is adapted to the filtration* $\{F_t^X\}_{t \ge 0}$.

A couple of examples will explain the contents of the definition. For example[8], if we define the event *A by* $A = \{X(s) \ne 1 \; for \; s \le 7\}$, *then* $A \in F_7^X$

For a stochastic variable

$$Y = \frac{d}{dT}\{\int_0^T X(s)ds\},\; Y \in F_T^X,\; i.e. Y_T \text{ is adapted to the filtration.}$$

Now let $B(s)$ be a Brownian motion and *Y(t) a process defined by*

$$Y(t) = \sup_{s \le t} B(s)$$

---

[7] The normal distribution has support on $(-\infty, +\infty)$.

[8] The examples are borrowed from Björk (1998, 2008).

then Y(t) is adapted to the filtration $\{F_t^B\}_{t\geq 0}$. *Finally, this is not true for*

$Y(t) = \sup_{s\leq t+1} B(s)$.*Why?*

To create a well defined stochastic integral, like the one in equation (2.6), we need some integrability conditions, i.e., a condition that makes the integral well defined. To this end consider a Brownian motion process, and another stochastic process *g* and assume that the following conditions are fulfilled[9]:

(i) $\int_a^b E\{g^2(s)\}ds < \infty$, *(the process g(s) belongs to* L^2 *)*[10]

(ii) The process g is adapted to the F_t^B filtration

More formally we say that the process *g belongs to the class* $L^2[a,b]$, *if the conditions (i) and (ii) are fulfilled.* For a simple process *g*, i.e. there exists points in time $a = t_0 < t_1 < ..... < t_n = b$ such that $g(s) = g(t_k)$ for $s \in [t_{k,}t_{k+1})$, it is straightforward to define the stochastic integral as

$$\int_a^b g(s)dB(s) = \sum_{k=0}^{n-1} g(t_k)[B(t_{k+1}) - B(t_k)] \tag{2.7}$$

but this is not the only way to do it. In expression (2.7), the integral is defined by forward increments. However, it may seem equally reasonable to use backward increments, i. e.

$$\int_a^b g(s)dB(s) = \sum_{k=0}^{n-1} g(t_{k+1})[B(t_{k+1}) - B(t_k)] \tag{2.8}$$

---

[9] Here I lean on Björk (1998)

[10] In Mathematics, L^p spaces are function spaces defined using natural generalizations of p-norms for finite dimensional vector spaces . The *L* stands for Lebesque that was one of the inventor. Riesz was another. For p=2 the resulting norm is the standadard Euclidian norm (measure of vector length), and p=1 gives the Manhattan distance

but the two definitions will yield approximations with very different properties.

To see this, taking expectations of the expression in (2.7) yields

$$E\{\int_a^b g(s)dB(s)\} = \sum_{k=0}^{n-1} E\{g(t_k)[B(t_{k+1} - B(t_k)]\} = 0 \tag{2.9}$$

Since Brownian motion has independent increments with mean zero we have $E\{g(t_k)[B(t_{k+1} - B(t_k)]\} = E[g(t_k)]E[B(t_{k+1}) - B(t_k)] = E[g(t_k)]\times 0 = 0$.

If g is not simple the definition of the stochastic integral will be much more tricky. First, we approximate g by a simple process, g_n, such that

$$\int_a^b E\{[g_n(s) - g(s)]^2\}ds \to 0 \tag{2.11}$$

For each n the integral $\int_a^b g_n(s)dB(s)$ is a well defined stochastic variable Z_n. It remains to prove that there exists a stochastic variable Z such that $Z_n \to Z$ in L^2 as $n \to \infty$. This is, indeed, possible, and we define

$$\int_a^b g(s)dB(s) = \lim_{n\to\infty} \int_a^b g_n(s)dB(s) \tag{2.12}$$

as the Ito–integral. It has some very convenient properties. Three of them are

(i) $\quad E\{\int_a^b g(s)dB(s)\} = 0$

(ii) $\quad E\{[\int_a^b g(s)dB(s)]^2\} = \int_a^b E\{[g(s)]^2 ds\}$ (2.13)

(iii) $\quad \int_a^b g(s)dB(s)$ is F_b^B measurable

The first property we proved for the case when g is simple. The second property is called the Ito-isometry. Loosely speaking, it transforms a certain Ito integral into a Rieman integral. The third property is e.g. handy when one wants to prove that certain stochastic processes are martingales.

The following Corollary follows from the Ito- isometry. If

$$E\{\int_S^T [f_n(t,\omega) - f(t,\omega)]^2 dt \to 0, when\, n \to \infty, then$$

$$\int_S^T f_n(t,\omega) dB_t(\omega) \to \int_S^T f(t,\omega) dB_t(\omega)\, in\, L^2\; as\, n \to \infty$$

This result can e.g., be used to find an explicit solution to an Ito integral.

Let us now calculate the integral

$$\int_0^t B(s)dB(s)$$

by making use of both forward and backward increments, respectively[11]. To start with we will show that the sum of squares of the increments converge to t , when the increments shrink to zero.

Define: $S_n = \sum_{k=0}^{n-1} [B(t_{k+1}) - B(t_k)]^2 = \sum_{k=0}^{n-1} (\Delta B_k)^2$

Claim 2.1 $\lim_{n\to\infty} S_n(t) = t \;\; in \;\; L^2$

Proof: Put $\Delta t = t_{k+1} - t_k, i.e. n\Delta t = t$. From the definition of Brownian motion it follows that

$E[\Delta B_k^2] = \Delta t$, which gives $E[S_n(t)] = \sum_{k=0}^{n-1} E[\Delta B^2] = \sum_0^{n-1} \Delta t = t$.

The variance of the sum, $Var[S_n(t)] = \sum_{k=0}^{n-1} E[(\Delta B)^2] = 2\sum_{k=0}^{n-1} (\Delta t_k)^2 = 2n({t}/{n})^2 = {2t^2}/{n}$.

Hence $\lim_{n\to\infty} E[(S_n - t)^2] \to 0$ proving that $S_n(t)$ converges to t in L^2 **.**

---

[11] The Claim comes from Björk (1994).

In other words, we can write $\int_0^t (dB(s))^2 = t$ .Now back to the sums in equations (2.7) and (2.8), which we write

$$I_n^f = \sum_k B(t_k)[B(t_{k+1}) - B(t_k)]$$

$$\bullet$$

$$I_n^b = \sum_k B(t_{k+1})[B_{k+1}(t_{k+1}) - B(t_k)]$$

Hence we have that $I_n^b + I_n^f = B^2(t)$, and $I_n^b - I_n^f = \sum_k (\Delta B_k)^2 = S_n$

From the Claim it follows that $I_n^b - I_n^f \to t$, and from $\bullet$ it follows that

$$I_n^f \to I^f \qquad (I_n^f + I_n^b) - (I_n^b - I_n^f) = B^2(t) - S_n = 2I_n^f$$

$$I_n^b \to I^b \qquad (I_n^f + I_n^b + I_n^b - I_n^f = B^2(t) + S_n = 2I_n^b)$$

i.e. the sums converge in L^2 , and

$$I^f = \frac{B^2(t)}{2} - \frac{t}{2}$$

$$I^b = \frac{B^2(t)}{2} + \frac{t}{2}$$

The forward increments yield the result from an Ito integral, while the backward increments yields $I^f + t$

Martingales

The conditional expectation given the information at time t , F_t is written as $E[Y|F_t]$.

The following results on conditional expectations are useful:

(i) If Y and Z are stochastic variables and Z is F_t measurable then

$E[ZY|F_t] = ZE[Y|F_t]$

(ii) If Y is a stochastic variable and $s < t$, then $E[Y|F_s] = E\{E[Y|F_t]|F_s\}$ (the law of iterated expectations).

The proof of (i) is simple, since the fact that Z is F_t measurable means that it is known at *t* . The law of iterated expectations is a version of the “law of total probability”. A mathematical expectation $E[y]$ can be written

$$E[y] = \int yf(y)dy = \int y \int f(x,y)dxdy = \int y \int f(y|x)f(x)dxdy = \int f(x)\int yf(y|x]dydx = E_x\{E[y|x]\}$$

which has some similarity to (ii).However, the law of iterative expectations is a consequence of the “tower property”, which in our notation can be written

$$E\{E[Y|F_t]|F_s\} = E[Y|F_s] = E\{E[Y|F_s]|F_t]$$ (The Tower Property)

where $F_s \subset F_t$. The right equality follows since $E[Y|F_s]$ is F_s measurable and hence F_t measurable. To prove the left equality, let $A \in F_s$. Then since A is also in F_t, we have

$$E\{E[E[(Y|F_t)|F_s]|I_A\} = E\{E[Y|F_t]|I_A\} = E\{YI_A\} = E\{E[Y|F_s]|I_A\}$$

Since both sides are F_s measurable the equality follows. Here I_A is an indicator function. We are conditioning on A which belong to both filtrations. To remember the tower property one can memorize that the smaller “set” always dominates.

One can prove the Tower property by introducing $\sigma -$ algebras explicitly into the analysis. The following Claim does the job (Björk1994):

Claim 2.2 (*Iterated Expectations) Assume that G and H are σ -algebras with $G \supseteq H$, then the following is true*

(i) $E[E(X|G)|H] = E[X|H]$

(ii) In particular $E(X) = E[E(X|G)]$

Proof: *To prove (i) define* $Z = E[X|H]$. *That Z is H measurable follows directly from the definition. It remains to prove that* $E[X|G] = \int_H ZdP$ *is H measurable.*

We have that $\int_H E[X|G]dP = \int_H XdP = \int_H E[X|H]dP = \int_H ZdP$

The first equality follows since $H \subseteq G$ *implies* $H \in G$. *The second follows because X is H measurable and the last by definition. Finally (ii) follows from (i) since H is the "underlying"* $\sigma-$ *algebra, i.e.* $E[X] = E[X|H]$.

Note that (ii) is the last equality in the Tower property.

We are now ready to define a martingale.

Definition 2.3(F_t *- martingale): A stochastic process X(t) is called a* F_t *matingale if the following conditions hold*

(i) *X (t) is adapted to the filtration* $\{F_t\}_{t\geq 0}$

(ii) *For all t* $E[|X(t)|] < \infty$

(iii) *For all s and t, with t>s the following relation holds:* $E\{X(t)|F_s\} = X(s)$

We will now prove that given the integrability condition $g \in L^2$, every stochastic integral is a martingale. To start, we introduce an extension of the following result: $E\{\int_a^b g(s)dB(s)\} = 0$. It also holds that for any process $g \in L^2$ that

$$E\{\int_s^t g(s)dB(s)|F_s^B\} = 0 \tag{2.13}$$

where the notation F_s^B means that the process *B* is known up to time s.

Now exercise 2.9 shows that every Ito-integral is a martingale.

Exercises (Björk (1998, 2008) and Öksendal2003)):

2.9 Prove that for $g \in L^2(s,t)$, *the process defined by X(t)=* $\int_0^t g(\tau)dB(\tau)$ *is a* F_t^B *martingale.*

Solution: Pick an s<t and write

$E\{X(t)\left|F_s^B\right.\} = E\{\int_0^t g(\tau)dB(\tau)\left|F_s^B\right.\} = E\{\int_0^s g(\tau)dB(\tau)\left|F_s^B\right.\} + E\{\int_s^t g(\tau)dB(\tau)\left|F_s^B\right.\}$ *. The first integral is* F_s^B *measurable so we ca take away the expectation sign, and the second integral is zero from (2.13). Hence, we have*

$$E\{X(t)\left|F_s^B\right.\} = \int_0^s g(\tau)dB(\tau) + 0 = X(s),$$

2.10 Check whether the following processes X(t) are matingales w. r, t, $\{F_t\}$

(i) *X(t)=B(t)+4t*

(ii) *X(t)=* $B^2(t)$

(iii) *X(t)=* $t^2 B(t) - 2\int_0^t sB(s)ds$

(iv) *X(t)=* $B_1(t)B_2(t)\, where [B_1(t), B_2(t)]$ *is 2 dimensional Brownian motion.*

Solution: (i) is a martingales iff E[X(t)-X(s)] $\left|F_s\right.$ *]=0 This is not the case since* $E[B(t) - B(s) + 4(t-s)\left|F_t\right.] = 4(t-s)$.

(ii) $E[B^2(t) - B^2(s)\left|F_s\right.] = E\{[B(t) - B(s)]^2\left|F_s\right.\} + 2E\{B(s)[B(t) - B(s)]\left|F_s\right.] = E\{\Delta B_{t-s}^2\} = t - s$,

i.e. is not a martingale

(iii) $E[X(t) - X(s)\left|F_s\right.\} = E\{t^2 B(t) - s^2 B(s)\left|F_s\right.\} - 2E\{[\int_0^t \tau B(\tau)d\tau - \int_0^s \tau B(\tau)d\tau]\left|F_s\right.\} = .$

$$E\{[t^2B(t) - s^2B(s)]\left|F_s\right.\} - 2E\{\int_s^t \tau B(\tau)d\tau\left|F_s\right.\} = E\{t^2[B(t) - B(s)] + [t^2 - s^2]B(\tau)\left|F_s\right.\} - 2\int_s^t \tau E[B(s)\left|F_s\right.\} =$$

$$= (t^2 - s^2)B(s) - (t^2 - s^2)B(s) = 0.$$

(iv) X(t)-X(s)= $B_1(t)B_2(t) - B_1(s)B_2(s) = [B_1(t) - B_2(s)]B_2(t) + B_1(s)B_2(t) - B_1(s)B_2(s) =$

$[B_1(t) - B_1(s)]B_2(t) + B_1(s)[B_2(t) - B_2(s)] = [B_1(t) - B_1(s)][B_2(t) - B_2(s)] + B_2(t)[B_1(t) - B_1(s)] +$

$B_1(s)[B_2(t) - B_2(s)]$

Taking expectations conditional on F_s *yields zero. A martingale!*

2.11. Prove that M(t)= $B^2(t)-t$ *is a* F_t *martingale*

Solution: $M(t)-M(s)=B^2(t)-B^2(s)-(t-s)=$

$[B(t)-B(s)]^2+2B(t)B(s)-2B^2(s)-(t-s)=[B(t)-B(s)]^2+2B(s)[B(t)-B(s)]-(t-s)$.

Taking expectations conditional on F_s *yields zero, q.e.d.*

It should also be clear from the above analysis that a sufficient condition for a process *X(t)* to be a martingale is that the stochastic differential has no dt term, i.e. $dX(t)=g(t)dB(t)$. It is much harder to show that this condition is also necessary, but this is according to Björk (1998), indeed, true.

The Stratonovich integral

So far, we have only dealt with one way to define a stochastic integral: the Ito integral. However, by starting from an elementary function *g(s),* we can define an integral

$$\int_0^T g(s)\circ dB_s=\sum_{k=0}^{n-1} g(t_k^*)[B(t_{k+1})-B(t_k)] \qquad (2.14)$$

where $t_k^*=\dfrac{t_k+t_{k+1}}{2}$, and $\circ\, dB_s$ denotes a Stratonovich differential.

In other words, we measure the value of *g(t) in the middle of the interval*, instead of at the beginning. For a simple function g this does not make any difference, but in a more general case it does. This means that the Stratonovich (1966) integral will, since it, loosely speaking, looks a little into the future, give results that are different from the Ito integral. Starting from the stochastic differential (2.4) one can show that the Stratonovich solution *X(t)* of the integral equation

$$X(t)=x_0+\int_0^t \mu(s,X(s))ds+\int_0^t \sigma(s,X(s))\circ dB(s) \qquad (2.15)$$

or

$x(0)=x_0$

also solves the following “modified” Ito equation

$$X(t) = x_0 + \int_0^t \mu(s, X(s))ds + \frac{1}{2}\int_0^t \frac{\partial\sigma(s, X(s))}{\partial X}\sigma(s, X(s))ds + \int_0^t \sigma(s, X(s))dB(s) \qquad (2.16)$$

or

$$dX(t) = [\mu(t, X(t)) + \frac{1}{2}\frac{\partial\sigma(t, X(t))}{\partial X}\sigma(t, X(t))]dt + \sigma(t, X(t))dB(t) \qquad (2.16')$$

This means that the result calculated in one integral can be transformed into the other. A disadvantage in the calculus resulting from a Stratonovich integral is that the integral is not generally a martingale. Note that, *if* σ *is independent of* x *the two integrals will coincide.* To see how the relationship between the Stratonovich and Ito integral can be used we solve the following exercise

2. 12.a) Transform the following Stratonovich differential into an Ito differential equation:
$dX(t) = \gamma X(t)dt + \alpha X \circ dB(t)$.

b) Move the following Ito differential equations into Stratanovich differential equations:
$dX(t) = rX(t)dt + \alpha X(t)dB(t)$

Solution: a) From the relationship between Ito and Stratonovich differentials we have that

$$\gamma X(t) \circ dB(t) = \frac{1}{2}\sigma^2 X(t)dt + \sigma X(t)dB(t)$$

Inserting this expression into the Stratonovich differential yields

$$dX(t) = (\gamma + \frac{\sigma^2}{2})X(t)dt + \sigma X(t)dB(t)$$

b) Left as an exercise to the reader.

c) Show that $\int_0^t B(s) \circ dB(s) = \frac{B^2(t)}{2}$*, ie.,Stradonovich integral of Brownian motion does not depend on time as a separate argument.Use 2.15 , 2.16 and Claim 1 that tells us that* $\int_0^t B(s)dB(s) = \frac{1}{2}[B^2(t) - t]$.

d) Could you have guessed the result in c?

2.3 Ito calculus-the one dimensional case

In the introductory chapter we handled a special case of the following problem: Given that *X(t))*solves a stochastic differential equation, and $Y(t) = f(t, X(t)),$ what is the dynamics of the $Y(t)-process$. Given that the increments are Brownian motion, we fix *s and t with* $s<t$ and define:

$$\Delta t = t - s$$
$$\Delta B = B(t) - B(s)$$

Since the increments of Brownian motion are normally distributed, $N(0,\sigma)$, it follows that

(i) $E\{\Delta B) = 0$

(ii) $E\{\Delta B^2\} = \Delta t = \text{var}\{\Delta B\}$

(iii) $Var\{\Delta B^2\} = 2[\Delta t]^2$

The $Var(\Delta B^2)$ is of order $O(\Delta B)$ implying that it goes to zero when $\Delta \to 0$. This means that a more mathematically talented person than the author would guess that $[dB(t)]^2 = dt$. This indeed turns out to be true, as was shown above, i.e.

$$\int_0^t [dB(s)]^2 = t$$

or equivalently

$$\left[dD(t)\right]^2 = dt$$

We are now ready to introduce Ito' s formula.

Theorem 2.1 (Ito's Lemma) *Assume that the process X(t) has a stochastic differential given by* $dX(t) = \mu(t)dt + \sigma(t)dB(t)$, $\mu(t)$ *and* $\sigma(t)$ *are adapted processes and f(t,x(t))* $\in C^2$.
Define $Z(t) = f(t, X(t))$, *then*

$$dZ(t) = df = [\frac{\partial f(t,X(t))}{\partial t} + \mu(t)\frac{\partial f(t,X(t))}{\partial X} + \frac{1}{2}\sigma^2(t)\frac{\partial^2 f}{\partial X^2}]dt + \sigma^2(t)\frac{\partial f}{\partial x}dB(t).$$

Remark: B(t) is one dimensional.

An intuitive proof is obtained by a second order Taylor expansion of $f(t, X(t))$ using that $dt^2 = O(dt)$, inserting the stochastic differential equation for $dX(t)$ and using that $dtdB = 0$ and $dB^2(t) = dt$. See equations (1.10) and (1.11) above.

Again, we are ready for some exercises. We start by illustrating how Ito's formula can be used to calculate mathematical expectation.

Exercises (from Björk and Öksendal):

2.13 Compute $E\{B^4(t)\}$, $B(0)=0$

Solution: Put X(t)=B(t) and Z(t)=f(t,X(t))= $X^4(t)$. *Clearly, by Ito's formula* $dZ = 4X^3dx + 6X^2dx^2 = 4B^3(t)dB(t) + 6B^2(t)dt$. *Integrating yields* $z(t) = 0 + 4\int_0^t B^2(s)dB(s) + 6\int_0^t B^2(s)ds$. *Finally taking expectations*

E{z(t)}=0+ $6\int_0^t E\{B^2(s)\}ds = 6\int_0^t sds = 3t^2$.

2.14 Compute E{ $e^{\alpha B(t)}$ *}, B(0)=0*

Solution: Put X(t)=B(t), and Z(t)= $e^{\alpha X(t)}$ *and compute* $dZ = \frac{\alpha^2}{2}e^{\alpha B(t)}dt + \alpha e^{\alpha B(t)}dB(t)$ *using Ito's formula. Integrating yields* $Z(t) = 1 + \frac{\alpha^2}{2}\int_0^t Z(s)ds + \alpha\int_0^t Z(s)dB(s)$. *After taking expectations we get*

$E\{Z(t)\} = 1 + \frac{\alpha^2}{2}\int_0^t E\{Z(s)\}ds$, *since the stochastic integral disappears. This equation is of type* $m(t) = 1 + \frac{\alpha^2}{2}\int_0^t m(s)ds$ *which results in a differential equation of the following type*

$\frac{dm(t)}{dt} = \frac{\alpha^2}{2}m(t)$ *m(0)=1. The solution is* $m(t) = m(0)e^{\alpha^2 t/2}$

2.15 Compute the stochastic integral $\int_0^t B(s)dB(s)$ *by putting Z(t)=* $B^2(t)$ *Why the particular relation? Well, a vague idea is that the integral equals* $B^2(t)/2$.

Solution: Put X(t)=B(t) and apply Ito's formula on $Z(t) = X(t)^2$. *This yields* $dZ(t) = dt + 2B(t)dB(t)$. *Integration gives*

$$Z(t) = B^2(t) = t + 2\int_0^t B(t)dB(t) \Leftrightarrow \int_0^t B(t)dB(t) = \frac{1}{2}[B^2(t) - t]$$

Compare the result from equation 2.7.

2.4 The n-dimensional Ito formula

To move to more then one dimension, we write the stochastic differential equations in the following manner:

$$dX_i(t) = \mu_i(t)dt + \sum_{j=1}^{d} \sigma_{ij} dB_j(t) \qquad i = 1,....,n \tag{2.17}$$

where $B_1,....,B_n$ are independent Wiener-processes. The drift vector is $\mu' = [\mu_1,...,\mu_n]$, and the $n \times d$ dimensional diffusion matrix is

$$\boldsymbol{\sigma} = \begin{bmatrix} \sigma_{11},......, \sigma_{1d} \\ . \\ \sigma_{n1},.......,\sigma_{nd} \end{bmatrix} \tag{2.18}$$

In vector notation we can now write the $\mathbf{x(t)}$ dynamics in the following manner

$$\mathrm{dX(t)} = \mu(\mathrm{t})\,\mathrm{dt} + \sigma(\mathrm{t})\,\mathrm{dB}(t) \tag{2.19}$$

where

$$B'(t) = [B_1(t),....B_d(t)] \tag{2.20}$$

Define a new process by

$$Z(t) = f(t, X(t))$$

Following the same idea as in the derivation of the Ito formula in the one dimensional case, and using the extended multiplication rule that $dB_i dB_j = 0$ when $i \neq j$, we obtain

$$dZ = \{ \frac{\partial f(\cdot)}{\partial t} + \sum_{i=1}^{n} \mu_i(t) \frac{\partial f}{\partial X_i} + \frac{1}{2} tr[\boldsymbol{\sigma}' \mathbf{H} \boldsymbol{\sigma}] \} dt + \sum_{i=1}^{n} \frac{\partial f}{\partial X_i} \sigma_i dB_i(t) \qquad (2.21)$$

where H is the Hessian matrix, i.e. the matrix of second order derivatives with respect to $X(t)$ of $f(t, X)$. For more detail see Björk (1998) Chapter 3 and Öksendal (2003) Chapter 3

We end the chapter by introducing a pack of exercises.

+Exercises (from Öksendal (2003)):

2.16 Use Ito's formula to write the following stochastic processes on the standard form $dX(t) = \alpha(t,\omega)dt + \beta(t,\omega)dB(t)$

a) $X(t) = B^2(t)$, $B(t)$ *is one* dim *ensional*

b) $X(t) = 2 + t + e^{B(t)}$, $B(t)$,1dim

c) $X(t) = B_1^2(t) + B_2^2(t)$, B_1, B_2, 2 dim

d) $X(t) = [t_0 + t, B(t)]$, $B(t)$,1dim

e) $X(t) = [B_1(t) + B_2(t) + B_3(t), B_2^2 - B_1(t)B_3(t)]$

Solutions: a) $dX(t) = 2B(t)dB(t) + \frac{2}{2} dB^2(t) = 2B(t)dB(t) + dt$.

b) $dX(t) = dt + e^{B(t)}dB(t) + \frac{e^{B(t)}}{2} dB^2(t) = (1 + \frac{e^{B(t)}}{2})dt + e^{B(t)}dB(t)$

c) Model $f(X_1, X_2) = X_1^2 + X_2^2$. $dX(t) = 2B_1(t)dB_1(t) + \frac{2}{2} dB_1^2(t) + 2B_2(t)dB_2 + \frac{2}{2} dB_2^2 =$

$= 2dt + 2[B_1(t)dB_1(t) + B_2(t)dB_2(t)]$

d) $dX(t) = [dt, dB(t)]$

e)Model $[X_1, X_2] \rightarrow [dX_1, dx_2]$, $\{ dX_1(t) = dB_1(t) + dB_2(t) + dB_3(t)$,

$dX_2 = dt - B_3(t)dB_1(t) + 2B_2(t)dB_2(t) - B_1(t)dB_3(t) \}$

2.17 Use Ito's formula to prove that $\int_0^t B^2(s)dB(s) = \frac{1}{3} B^3(t) - \int_0^t B(s)ds$

Solution: Clearly $\frac{1}{3} d[B^3(t)] = B(t)dt + B^2(t)dB(t)$. *Now* $\frac{1}{3} d[B^3(t)] - B(t)dt = B^2(t)dB(t)$ *and integrating yields the desired result,*

q.e.d..

2.18 Let $X(t), Y(t)$ *Ito processes in R, prove that*

$d(X(t)Y(t)) = X(t)dY(t) + Y(t)dX(t) + dX(t)dY(t)$

Solution: Idea let f(X(t)Y(t))=X(t)Y(t) and use Ito's formula to get

$$df(X,Y) = f_1 dX + f_2 dY + \frac{1}{2}[f_{11}dX^2 + f_{22}dY^2] + f_{12}dXdY = YdX + XdY + dXdY$$

This can be written;as a formula for integration by parts.

To see this we write this expression in differential form and use the form of d(X(t,Y(t)). We obtain X(t)dY(t)=X(t)dY(t)+Y(t)dX(t)+dX(t)dY(t)-Y(t)dX(t)-dX(t)dY(t)= X(t)dY(t).

2.19 Let $\theta(t,\omega) = (\theta_1(t,\omega),....\theta_n(t,\omega))$ *with* $\theta_k \in V(0,T)$ *all k=1…n. Here V(0,T) is the class of functions that fulfills conditions under which the Ito integral is well defined. This means in particular that* $E[\int_0^T \theta_k(t)^2 dt] < \infty$ *and that* $\theta(t,\omega)$ *is* F_t *measurable. See Öksendahl(2003) Chapter 3 page 25.*

Now define Z(t)=exp{ $\int_0^t \theta(t,\omega)dB(s) - \frac{1}{2}\int_0^t \theta^2(s,\omega)ds$ *}, where B(t) is n dimensional Brownian motion and* $\theta^2 = \theta x \theta$

a) *Prove that dZ(t)=Z(t)* $\theta(t,\omega)dB(t)$

b) *Prove that* $[Z(t)]_{t \le T}$ *is a martingale if* $Z(t)\theta_k \in V(0,T)$.

Solution a: Let h(t)= $\int_0^t \theta(s,\omega)B(t) - \frac{1}{2}\int_0^t \theta^2(s,\omega)ds$ *implying that Z(t)=* $e^{h(t)}$.

Now dZ=z(dh+ $\frac{1}{2}dh^2$*), dh=* $\theta dB - \frac{1}{2}\theta^2 dt$ *and* $dh^2 = \theta^2 dt$. *Substitutions now yield the desired result.*

b) Under the above assumptions the Ito integral is well defined and we can write

$Z(t) = z(0) + \int_0^t Z(s)\theta(s,\omega)dB(s)$ *and the result now follows since the Ito integral is a martingale.*

2.20 Let B(t) be n dimensional Brownian motion and let f: $R^n \rightarrow R$ *be* C^2. *Use Ito's formula to prove that* $f(B(t))=f(B(0))+\int_0^t gradf(B(s))dB(s)+\frac{1}{2}\int_0^t \Delta f(B(s))ds$ *where* $\Delta=\sum_1^n \frac{\partial^2}{\partial x_i^2}$ *is the Laplace operator.*

Solution: Follows directly from Ito's Lemma by noting that $dB_j dB_i = 0$ *for* $i \neq j$ *and* $dB_i dB_i = dt, \forall i$

2.21 Use Ito's formula to prove that the following stochastic processes are martingales.

a) $X(t)=e^{\frac{1}{2}t}\cos B(t)$

b) $X(t)=e^{\frac{1}{2}t}\sin B(t)$

c) $X(t)=(B(t)+t)e^{-(B(t)+\frac{1}{2}t)}$

Solution:

a) $dX(t)=\frac{1}{2}e^{\frac{t}{2}}\cos B(t)dt-e^{\frac{t}{2}}\sin B(t)dB(t)-\frac{1}{2}e^{\frac{t}{2}}\cos B(t)dt=-e^{\frac{t}{2}}\sin B(t)dB(t)$ *by Ito's formula. This is an Ito integral*

$X(t)=x(0)-\int_o^t e^{\frac{s}{2}}\sin B(s)dB(t)$ *which is a martingale.*

b) Again use Ito's formula to get:

$X(t)=x(0)+\int_0^t e^{\frac{s}{2}}\cos B(s)dB(s)$ *which is a martingale.*

c) $\frac{dX(t)}{dt}=\exp[-(B(t)+\frac{t}{2})]-\frac{1}{2}(B(t)+t)\exp[-(B(t)+\frac{t}{2})],$

$$\frac{\partial X(t)}{dB(t)}=\exp-(B(t)+\frac{t}{2})-(B(t)+t)\exp[-(B(t)+\frac{t}{2})]$$

$$\frac{1}{2}\frac{\partial^2 X(t)}{\partial B^2(t)}=-\frac{1}{2}\{\exp[-(B(t)+\frac{t}{2})]+\exp[-(B(t)+\frac{t}{2})]+(B(t)+t)\exp[-(B(t)+\frac{t}{2})]$$

Canceling terms give $X(t)=x(0)+\int_0^t \frac{\partial X}{\partial B(t)}dB(t)$, *which is a martingale.*

2.22 In each of the processes below find the process $f(t,\omega) \in V(0,T)$ *such that*

$$F(\omega) = E(F\ ) + \int_0^t f(t,\omega)dB(t) \text{ for a given } F \qquad *$$

The formula $*$ *is called Ito's Representation Theorem and it tells us, loosely, that any* F_t *measurable Martingale has an Ito representation. In other words it is the converse of the result in 2.9 that every Ito integral is a Martingale. See Öksendal (2003) page51.*

a) $F(\omega) = B(T,\omega)$

b) $F(\omega) = \int_0^T B(t,\omega)dt$

c) $F(\omega) = B^2(T,\omega)$

d) $F(\omega) = B^3(T,\omega)$

e) $F(\omega,T) = e^{B(T,\omega)}$

Here B(t) is 1-dimensional Brownian motion.

Solution:

a) $E\{F(\omega)\} = E\{B(T,\omega)\} = E\{\int_0^T dB(t)\} = 0$

implying that

$$F(\omega) = 0 + \int_0^T dB(t) = B(T,\omega)$$

b) Here also $E[F(0)] = 0$. *To see this we calculate the value of the integral by partial integration, since the integral is a Riemann integral:*

$\int_0^T B(t)dt = B(T,\omega)T - \int_0^T tdB(t,\omega) = \int_0^T (T-t)dB(t,\omega)$. *Hence* $F(\omega) = 0 + \int_o^T (T-t)dB(t,\omega)$.

c) We start by using Ito's Lemma to get $d(B^2(t)) = 2B(t)dB(t) + 1$. *Integrating yields*

$F(\omega) = B^2(T,\omega) = 2\int_0^T B(t)dB(t) + T$, *so* $E\{F(\omega)\} = T$, *and* $F(\omega) = E[F] + 2\int_0^T B(t)dB(t)$.

d) Applying Ito's lemma on $B^3(t)$ *yields*

$$d[B^3(t)] = 3B^2(t)dB(t) + 3B(t)dt,$$

$$\text{int}\,egratingyields\;\; B^3(T) = 3[\int_0^T B^2(t)dB(t) + \int_0^T B(t)dt] = 3[\int_0^T B^2(t)dB(t) + TB(t) - \int_0^T tdB(t)]$$

$$= 3[\int_0^T B^2(t)dB(t) + \int_o^T (T-t)dB(t)] = F(\omega)$$

e) Standard Ito calculations yield: $d(e^{B(T)}) = e^{B(t)}dB(t) + \frac{1}{2}e^{B(t)}dt$. *Integrating yields*

$$e^{B(t)} = 1 + \int_0^t e^{B(s)}dB(s) + \frac{1}{2}\int_0^t e^{B(s)}ds$$

Now put H(T)= $E\{e^{B(T)}\} = 1 + \frac{1}{2}E\{\int_0^T e^{B(s)}ds\}$ *This means that* $H^{'}(T) = \frac{1}{2}H(T)$ *implying that*

$H(T) = Ce^{\frac{1}{2}T} = e^{\frac{1}{2}T}$ *since H(0)=1. Moreover, since* $E\{e^{B(T)}\} = E[F] = e^{\frac{1}{2}T}$ *, we need a representation that contains both the exponential function at T and a stochastic integral integrated over the interval [0,T]. We try* $Y(T) = e^{B(T)-\frac{1}{2}T}$. *Again using Ito calculus we get*

$dY(T) = Y(T)dB(T)$, *which after integration reads* $Y(T) = y(0) + \int_0^T Y(s)dB(s)$, where y(0)=1.

This in turn means that $Y(T)e^{\frac{1}{2}T} = 1 + \int_0^T e^{B(s)-\frac{1}{2}s}dB(s)$. *Inserted into the formula for the Representation Theorem we obtain* $F(\omega) = e^{\frac{1}{2}T} + \int_0^T e^{B(s)-\frac{1}{2}(s-T)}dB(s)$

f) Sticking to the same procedure we use Ito calculus to obtain

$d(\sin B(t)) = \cos B(t)dB(t) - \frac{1}{2}\sin B(t)dt$. *We integrate to get*

$$\sin B(T) = -\frac{1}{2}\int_0^T \sin B(s)ds + \int_0^T \cos B(s)dB(s)$$

Taking expectation we obtain $E\{\sin B(T)\} = -\frac{1}{2}E\{\int_0^T \sin B(s)ds\}$

To solve explicitly for the expected value we put $H(T) = E\{\sin B(T)\}$ *and note that*

$H^{'}(T) = -\frac{1}{2}E\{\sin B(T)\}$. *So* $H^{'}(T) + H(T) = 0$ *implying that H(T)=* Ce^{-T} *Since sinB(0)=sin(0)=0, we get C=0, and E(F)=0. Now we need an informed guess about the* *stochastic process that will represent* $F(\omega)=\sin B(T,\omega)$ *A good guess is to put* $Y(T)=e^{T/2}\sin B(T)$ *and use Ito's lemma to get*

$$d[e^{t/2}\sin B(t)]=\frac{1}{2}e^{t/2}\sin B(t)dt+\frac{1}{2}e^{t/2}\cos B(t)-\frac{1}{2}e^{t/2}\sin B(t)dt=e^{t/2}\cos B(t)dB(t),$$

or $\sin B(T)e^{T/2}=\int_0^T e^{t/2}\cos B(t)dB(t)$*, i.e.,* $\sin B(T,\omega)=\int_0^T e^{(t-T)/2}\cos B(t)dB(t)$.

Chapter 3: Stocastic Differential Equations (SDE,s)

In section 2.2 above we introduced the stochastic differential

$$\begin{aligned} dX(t)&=\mu(t,X(t))dt+\sigma(t,X(t))dB(t)\\ X(0)&=x_0 \end{aligned} \tag{3.1}$$

As discussed there the Ito interpretation of (3.1) is that a solution of this equation satisfies the stochastic integral equation

$$X(t)=x_0+\int_0^t \mu(s,X(s))ds+\int_0^t \sigma(s,X(s))dB(s)$$

In this chapter we will be concerned with how we can solve the equation, but also, to a lesser extent, whether one can find existence and uniqueness theorems for such equations. Moreover, we discuss the properties of the solutions.

If we start with the existence and uniqueness results, there exist theorems that take care of this problem. For details the reader is referred to Björk (1998) Chapter 4, and Öksendahl (2003) Chapter 5. The latter reference contains a proof. Loosely speaking, there are two important conditions that are needed in the proof of the existence result.The first is a growth condition involving the coefficients of the differential equation in equation in equation 3.1. It ensures that the solution $X(t,\omega)$ does not explode, i.e. does not approach infinity in finite time. An often used example of an "explosion" is the solution to the following ordinary differential equation (ODE).

$$\frac{dx(t)}{dt} = x^2(t)$$
$$x(0) = 1$$

In terms of equation (3.1), this corresponds to the case $\mu(t, x(t)) = x^2(t)$, which does not satisfy the growth condition. The equation has the unique solution $x(t) = [1-t]^{-1}$ on $0 \le t < 1$ This means that the solution cannot be defined for all *t*.

The second condition is a so called Lipschitz condition[12] on the same coefficients, which, as a matter of fact, are functions. This condition guarantees a unique solution. The example that violates the existence of a solution is

$$\frac{dx(t)}{dt} = 3x(t)^{2/3}$$
$$x(0) = 0$$

This equation has for each $\bar{t} > 0$ a solution

$$X(t) = \begin{Bmatrix} 0 \ for\, t \le \bar{t} \\ (t-\bar{t})^3 \ for\, t > \bar{t} \end{Bmatrix}$$

The reason is that $\mu(t, x(t)) = 3x(t)^{2/3}$ does not fulfill the Lipschitz condition at zero.

3.1 Some important SDE;s

The unique solution of the SDE, $X(t,\omega)$, has a continuous trajectory, it is a Markov process, and it is $F_t^{B(t)}$ adapted, which means that it is a functional of the tracjetory on the interval *[0,t]*. More formally, a SDE induces a transformation of the class of continuous function on the space $[0,\infty)$ into itself, where a Brownian trajectory $B(t,\omega)$ is mapped into the corresponding $X(t)$ trajectory. To put it bluntly, it is complicated and it is rare that an explicit solution can be found.

[12] A Lipschitz condition is a smoothness condition on functions that is stronger than continuity.

An example of a stochastic differential equation that can be solved, and which has important applications in Economics is the Geometric Brownian Process. Say that we start from a growth model

$$\frac{dX(t)}{dt} = g(t)X(t) \qquad (3.2)$$
$$X(0) = x_0$$

where $g(t) = g + \sigma W(t)$, g and σ constants greater than zero, and $W(t)$ is white noise. By multiplying through by dt we can give the equation an Ito interpretation of the following shape

$$dX(t) = gX(t)dt + \sigma X(t)dB(t) \qquad (3.3)$$
$$X(0) = x_0$$

An ordinary differential equation $\dot{X}(t) = aX(t)$, can be solved by writing $dX / X = adt$ integrating to get $\ln X(t) = at + x_0$, and by taking antilogarithms to get $X(t) = x_0 e^{at}$. So, why not put $Z(t) = \ln X(t)$, implying that $e^{Z(t)} = X(t)$. To take it from there we use Ito's lemma and equation (3.3) to obtain

$$dZ(t) = \frac{dX(t)}{X(t)} - \frac{dX^2(t)}{2X^2(t)} = gdt + \sigma dB(t) - \frac{\sigma^2}{2}dt \qquad (3.4)$$

Since the right hand side of the equation does not depend on $Z(t)$ we can integrate to get

$$Z(t) = \ln x_o + (g - \sigma^2 / 2)t + \sigma B(t) \qquad (3.5)$$

or

$$X(t) = x_o \exp\{(g - \sigma^2 / 2)t + \sigma B(t) \qquad (3.6)$$

The above derivation is not stringent. The calculations just presented presupposes that $X(t) > 0$, otherwise the logarithm would not be defined. Moreover, they presupposes that a solution exists. One way out would be to start from equation (3.6) and show that it satisfies

(3.3). To this end we prove the following result (the particular idea is borrowed from Björk (1998), but the idea is probably around elsewhere):

Proposition 3.1: *The solution to the stochastic differential equation (SDE)*

$$dX(t) = gX(t))dt + \sigma X(t))dB(t)$$
$$X(0) = x_0$$

is given by

$X(t) = x_0 \exp[(g - \frac{\sigma^2}{2})t + \sigma B(t)]$, *and* $E[X(t)] = e^{gt}$

Proof: To prove the first claim we start from $X(t) = e^{Z(t)}$, *where*

$Z(t) = \ln x_0 + (g - \frac{\sigma^2}{2})t + \sigma B(t)$. *This means that* $dZ(t) = (g - \frac{\sigma^2}{2})dt + \sigma dB(t)$. *Using Ito's lemma we get* $dX(t) = e^{Z(t)}dZ(t) + \frac{1}{2}e^{Z(t)}dZ^2(t) = gX(t)dt + \sigma X(t)dB(t)$.

To prove the second claim we use the Ito interpretation of the differential equation to write

$X(t) = x_0 + \int_0^t gX(s))ds + \int_0^t \sigma X(s))dB(s)$. *Taking expectations yields*

$E[X(t)] = x_0 + \int_0^t \sigma E[X(s)]dX(s)$. *Defining* $m(t) = E[X(t)]$, *we can write*

$m(t) = x_0 + \sigma\int_0^t m(s)ds$ *Differentiating with respect to t yields* $\dot{m}(t) = \sigma m(t)$, *m(0)=* x_0, *which can be solved to obtain* $E[X(t)] = m(t) = x_0 e^{gt}$ *,q.e.d..*

Remark 1: Note that in the second part of the proof, we use a trick similar to that was used in applying the Ito Representation theorem on martingales in exercise 2.21.

Remark 2: The first part of the claim in Proposition 3.1 means that the solution $X(t)$ *will indeed remain positive, which is helpful in many economic applications. Prices and capital stocks are likely to remain non-negative.*

The second claim in Proposition 3.1 is valuable when one wants to evaluate the expected discounted value of an income stream. Say that we want to calculate $E[\int_0^\infty X(t)e^{-rt}dt]$, were $X(t)$ follows a Geometric Brownian Motion. From Proposition 3.1 we know that $E[X(t)] = x_0 e^{gt}$. If $g < r$ we get $E[\int_0^\infty X(t)e^{-rt}dt] = \frac{x_0}{r-g}$. More generally, suppose that $F(x(t)) = x^\theta(t),$ where $X(t)$ follows a Geometric Brownian motion, and calculate $E\{\int_0^\infty F[X(t)]e^{-rt}dt]\}$. From Ito calculus we obtain $dF = [\theta g + \frac{1}{2}(\theta(\theta-1)\sigma^2]Fdt + \theta\sigma FdB$. Using the calculation idea in Proposition 3.1 we obtain $E[F(X(t))] = F(x_0)\exp[\theta g + \frac{1}{2}(\theta(\theta-1)\sigma^2]t = F(x_0)e^{\Theta t}$

and

$$E\{\int_0^\infty F[x(t)]e^{-rt}dt]\} = x_0^\theta[r-\Theta]^{-1}$$

provided $r > \Theta = \theta g + \frac{1}{2}[\theta(\theta-1)]\sigma^2$.

The linear SDE

The following stochastic differential equation that is linear in the Brownian increments can also be solved by alluding to ODE methods[13]

$$\begin{aligned} dX(t) &= \alpha X(t)dt + \sigma dB(t) \\ X(0) &= x_0 \end{aligned} \tag{3.7}$$

To find the solution we recall that the ODE $\dot{X}(t) = \alpha X(t) + U(t)$, with $X(0) = x_0$ has the solution $X(t) = x_0 e^{\alpha t} + \int_0^t e^{\alpha(t-s)}U(s)ds$. A “wild guess” is that the solution to (3.7) is

[13] ODE=ordinary differential equations .

$$X(t) = x_0 e^{at} + \sigma \int_0^t e^{\alpha(t-s)} dB(s) \qquad (3.8)$$

To prove this we proceed along the lines we used in Proposition 1. Let us write $X(t) = Y(t) + Z(t)R(t)$, where $Y(t) = x_0 e^{\alpha t}$, $Z(t) = \sigma e^{\alpha t}$, and $R(t) = \int_0^t e^{-\alpha s} dB(s)$. Differentiate to get $dX(t) = dY(t) + Z(t)dR + r(t)dZ(t) + dR(t)dZ(t)$. The last term vanishes since $dtdB(t) = 0$ in Ito calculus, and the other terms end up as equation (3.7).

The reason why the wild guess works is the linear structure, which means that the second order term in Ito's formula vanishes. A similar result can be formulated for a more general situation where $X(t)$ is vector valued[14]. The equation in equation (3.7) is called the Ornstein-Uhlenbeck process.

Exercise:

3.1What would the Stratonovich solution of the Ornstein-Uhlenbeck equation look like?

Weak and strong solutions

The solutions we have studied so far are so called strong solutions. What characterizes a strong solution is that the Brownian motion process is given in advance and the solution $X(t)$ is F_t^B adapted. If we only are given the functions $\mu(t, X(t))$ and $\sigma(t, X(t))$, and ask for a pair of processes $[(\tilde{X}(t), \tilde{B}(t)), H_t]$ on some probability space $[\Omega, H, P]$ that satisfies equation 3.1, then the solution $(\tilde{X}(t), \tilde{B}(t))$ is called a weak solution. Here H_t is the filtration of the sigma algebra H on a given set Ω . The pair (Ω, H) is called a measurable space and P is a probability measure on this measurable space. The triplet (Ω, H, P) is called a probability space. A strong solution is a weak solution.

In a modern proof of one of the most famous theorems in Financial Economics, the Black and Scholes Theorem - one changes, starting from a SDE like 3.1, probability measure using a famous result by Girsanov, which result in a new SDE and a new Brownian motion process.

---

[14] See e.g. Björk (1998) p 57.

The solution to the new equation is an important example of a weak solution. For details see Chapter 6.

Exercises (from Björk(1998) and Öksendal(2003):

3.2 Verify that $X(t) = e^{B(t)}$ *solves* $dX(t) = \frac{1}{2}X(t)dt + X(t)dB(t)$

Solution: Use Ito's formula on $X(t)$ *to get*

$$dX(t) = e^{B(t)}dB(t) + \frac{1}{2}e^{B(t)}(dB(t))^2 = X(t)dB(t) + \frac{1}{2}X(t)dt$$

3.3 Verify that $X(t) = \frac{B(t)}{1+t}$ *solves* $dX(t) = \frac{-X(t)dt}{1+t} + \frac{dB(t)}{1+t}$

Solution: Ito's formula gives $dX(t) = \frac{dB(t)}{1+t} - \frac{B(t)dt}{(1+t)^2}t = \frac{dB(t)}{1+t} - \frac{X(t)dt}{1+t}$

3.4 Let $B = (B_1,...,B_n)$ *be n-dimensional Brownian motion, and let* $\alpha = (\alpha_1,...,\alpha_n)$. *Solve the stochastic differential equation* $dX(t) = rX(t)dt + X(t)[\sum_{k=1}^{n}\alpha_k dB_k(t)],\ x(0) > 0$. *Note here that X(t) is one dimensional. This means that we can use the same trick as we used when we solve the Geometric Brownian Motion model, i.e., we put* $Y(t) = \ln X(t)$, *and use Ito's formula. The result is* $X(t) = x(0)\exp[(r - \frac{1}{2}\sum_{k=1}^{n}\alpha_k)t + \sum_{i=1}^{n}\alpha_k B_k(t)]$

3.5 Solve the mean reverting Ornstein-Uhlenbeck process $dX(t) = [m - X(t)]dt + \sigma dB(t)$ *and find E(X(t)), and VarX(t).*

Solution: Work from $X(t)e^t$ *and use the equation and Ito's formula to get*

$$d(e^t X(t) = me^t dt + \sigma e^t dB(t)$$

Integration yields $e^t X(t) = \int_0^t me^s ds + \int_0^t \sigma e^s ds$

Or $X(t) = x_0 e^{-t} + \int_0^t me^{s-t}ds + \int_0^t \sigma e^{s-t}dB(s) = x_0 e^{-t} + me^{-t}(e^t - 1) + \int_0^t \sigma e^{s-t}dB(t)$ *Taking expectations yields* $E(X(t)) = e^{-t}E(x_0) + m(1 - e^{-t})$

Var[X(t)] =

$$e^{-2t}\,\mathrm{var}\,x_0+\sigma^2e^{-2t}E\{[\int_0^t e^s\,dB(s)]^2\}=e^{-2t}\,\mathrm{var}(x_0)+\sigma^2e^{-2t}\int_0^t e^{2s}ds=e^{-2t}\,\mathrm{var}\,x_0+\frac{\sigma^2}{2}(1-e^{-2t})$$

where the first equality follows from the Ito isometry.

3.6 Suppose X(t) satisfies the SDE $dX(t)=\alpha X(t)dt+\sigma X(t)dB(t)$. *Let Y(t)=* $X(t)^{\beta}$ *, and compute dY(t) to find out which SDE Y(t) satisfies.*

Solution: The equation for X(t) is Geometric Brownian motion which means that $X(t)\neq 0$. *(Why is this important here?). Ito's formula yields*

$$dY=\beta X^{\beta-1}dX+\frac{1}{2}\beta(\beta-1)X^{\beta-2}dX^2=[\alpha\beta X^{\beta}+\frac{1}{2}\beta(\beta-1)\sigma^2X^{\beta}]dt+\sigma\beta X^{\beta}dB=$$

$$=\beta Y[\alpha+\frac{1}{2}\beta(\beta-1)\sigma^2]dt+\beta\sigma YdB.$$

3.7 Suppose that X(t)satisfies the SDE $dX(t)=\alpha X(t)dt+\sigma X(t)dB_1(t)$, *and Y(t) satisfies* $dY(t)=\gamma Y(t)dt+\delta Y(t)dB_2(t)$ *where* B_1, B_2 *are one dimensional Brownian motion. If we interpret X(t)as nominal income and Y(t) as the price index, we can interpret* $z(t)=X(t)/Y(t)=f(X(t),Y(t))$ *as real income . Determine the SDE for real income.*

Solution: Again we are dealing with Geometric Brownian motion, so division will work everywhere. A Taylor expansion of z= f(X,Y)yields

$$dz=\frac{\partial f}{\partial x}dX+\frac{\partial f}{\partial y}dY+\frac{1}{2}\frac{\partial^2 f}{\partial X^2}dX^2+\frac{\partial^2 f}{\partial X\partial Y}dXdY+\frac{\partial^2 f}{\partial y^2}dY^2$$ *. Moreover,*

$$\frac{\partial f}{\partial X}=\frac{1}{Y},\ \frac{\partial f}{\partial Y}=\frac{-X}{Y^2},\ \frac{\partial^2 f}{\partial X^2}=0,\ \frac{\partial^2 f}{\partial X\partial Y}=\frac{-1}{Y^2},\ \text{and}\ \frac{\partial^2 f}{\partial Y^2}=\frac{2X}{Y^3}$$ *, and* $dXdY=0, dY^2=\delta^2Y^2dt$.

Substitutions for dX, dY, and dY^2 *yield dz(t)=* $(\alpha-\gamma+\delta^2)zdt+\sigma zdB_1-\delta zdB_2$.

The Stratanovich integral modeled in an Ito world

To compare a solution pertaining to an Ito integral with a solution of a corresponding Stratanovich integral we will look upon the geometric Brownian motion model in connection with Proposition 3.1 and interpret it as a model for population growth. The Ito version of the stochastic differential equation has the following shape:

$$dP(t)=nP(t)dt+\alpha P(t)dB(t), P(0)=P_0$$

The Stratonovich interpretation of the equation is written:

$$d\bar{P}(t)=n\bar{P}(t)dt+\alpha\bar{P}(t)\circ dB(t), \bar{P}(0)=P_0$$

The first term in each equation is the growth trend with growth rate *n* the second term is the stochastic component. The variable $P(t)$ is the size of the population at time t.

We will start by solving the Ito version in a slightly different manner than earlier. To that end we divide through by $P(t)$ to get

$$\frac{dP(t)}{P(t)} = ndt + \alpha dB(t),\ B(0) = 0$$

Integration yields

$$\int_0^t \frac{dP(s)}{P(s)} = nt + \alpha B(t)$$

To evaluate the integral on the left hand side we again use Ito's formula on $g(t,x) = \ln x,\ x > 0$.

$$d[\ln P(t)] = \frac{dP(t)}{P(t)} - \frac{1}{2}\frac{dP^2(t)}{P^2(t)} = (use\,Ito\,calculus) =$$

$$\frac{dP(t)}{P(t)} - \frac{\alpha^2}{2}dt.$$

Hence,

$$\frac{dP(t)}{P(t)} = d\ln P(t) + \frac{\alpha^2}{2}dt$$

This means that after integrations we get

$$\ln P(t) + \frac{\alpha^2}{2}t = nt + \alpha B(t) + \ln(P_0)$$

After taking anti-logarithms we obtain

$$P(t) = \exp[(n - \frac{\alpha^2}{2})t + \alpha B(t)]$$

To find the corresponding Stratonovich solution we can use equations (2.15) and (2.16) to derive the modified Ito equation

$$\frac{d\overline{P}(t)}{\overline{P}(t)} = ndt + \frac{\alpha^2}{2}dt + \alpha dB(t),\ B(0) = 0$$ ‘

We also introduce the $\ln\overline{P}(t)$ function in the same manner as before since we have moved into an Ito world. This gives

$$\ln\overline{P}(t) + \frac{\alpha^2}{2}t = nt + \frac{\alpha^2}{2}t + \alpha B(t) + \ln(P_0)$$

Taking anti-logarithms yields

$$\bar{P}(t) = P_0 \exp[nt + \alpha B(t)]$$

which is the solution of the Stratonovich equation.

In connection with Proposition 3.1 it has already been shown that the Ito integral gives the mathematical expectation[15] $E[P(t)] = P_0 e^{nt}$. In the Ito modified Stratonovich case putting $Z(t) = e^{\alpha B(t)}$ and applying Ito's formula one obtains

$$dZ(t) = \alpha e^{\alpha B(t)} dB(t) + \frac{1}{2}\alpha^2 e^{\alpha B(t)} dt$$

Integrating yields

$$Z(t) = Y(0) + \alpha \int_o^t e^{\alpha B(s)} dB(s) + \frac{1}{2}\alpha^2 \int_0^t e^{\alpha B(s)} dt$$

Taking expectations the first integral disappears and we are left with

$$E[Z(t)] = Y_0 + \frac{1}{2}\alpha^2 \int_0^t E[Z(s)]ds \qquad *$$

Putting $z(t) = E[Z(t)]$ and differentiating $*$ yields

$$\dot{z}(t) = \frac{1}{2}\alpha^2 z(t)$$

The solution is $z(t) = E[Z(0)]e^{\frac{1}{2}\alpha^2} = e^{\frac{1}{2}\alpha^2}$ an after substitution we have that $E[\bar{P}(t)] = P_0 e^{(n + \frac{\alpha^2}{2})t}$.

It should be clear that the two solutions give different qualitative results. One can prove that the Ito solution goes to infinity with t if $n > \frac{\alpha^2}{2}$, and that it converges to zero if $n < \frac{\alpha^2}{2}$. For $n = \frac{\alpha^2}{2}$ if will fluctuate between large and small values. The Stratanovich solution goes to zero if $n < 0$, to infinity when $n > 0$ and fluctuates if $n = 0$. The proof needs something called[16] "The law of iterated logarithms". It is of course the Brownian motion process that complicated the behavior in the limit.

---

[15] One can also write $E[N(0)]$ treating the starting point as a stochastic variable.

[16] See Öksendal (2003) p 66.

The bottom line is that whether to use the Ito or Stratanovich integral depends on what process one wants to model. Since Ito calculus is quite elegant compared to how the Stratanovich integral can be treated this is tends, ceteris paribus, to lean in the Ito direction. We will return to this problem once again in chapter 4.

3.2 Stochastic differential equations and partial differential equations

In this section we will introduce the reader to the close and important connection that exist between a certain class of partial differential equations (PDE;s) and stochastic differential equations. This connection can be used to simplify the solution of the PDE;s. We will start by introducing an operator *A*, which is known as the Dynkin-Ito operator. An operator is a symbol that induces a rule that is applied to a variable or a function.

An example is $T(x) =$ subtract 2 from *x*. So $T(6) = 4$ and $T^2(6) = T(T(6)) = 2.$ A derivative can also be viewed as an operator. For example, $\frac{d}{dx}[f(x)]$ tells us to compute the first derivative of $f(x)$. And $\frac{d^2}{dx^2}[f(x)]$ results in the second derivative of $f(x)$, or the derivative of the first derivative of $f(x)$. The operation we are going to introduce tells us to take a series of derivatives of a given function $f(\mathbf{x}) \in C^2(R^n)$. This means that the function is twice continuously differentiable in ***x*.**

Definition 3.1 Given the SDE in equation (2.4) with $\mathbf{X} \in R^n$ *, the partial differential operator A of* ***X*** *is defined for any function f(x) with* $f \in C^2(R^n), by$

$$Af(t,\mathbf{X}) = \sum_{i=1}^{n} \mu(t,\mathbf{X}) \frac{\partial f(t,\mathbf{X})}{\partial \mathbf{X}_i} + \frac{1}{2}\sum_i \sum_j C_{ij} \frac{\partial^2 f(t,\mathbf{X})}{\partial \mathbf{X}_i \partial \mathbf{X}_j}$$

C(t,x)= $\sigma(t,\mathbf{X}(t))\sigma'(t,\mathbf{X}(t))$ *with element* C_{ij} ***.*** *The operator is known as the Dynkin-Ito operator.*

The operator means that the Ito formula can be written

$$df(t,\mathbf{X}(t)) = \{\frac{\partial f}{\partial t} + Af\}dt + gradf(\mathbf{X}(t))\sigma d\mathbf{B}(t) \qquad (3.9)$$

A PDE boundary problem

In this section we will derive a stochastic representation formula known as the *Feynman-Kac representation formula* which can be used to solve certain PDE;s. Starting from three scalar functions $\mu(t,x(t)), \sigma(t,x(t)),$ and $\phi(x)$ the task is to find a function F , which satisfies a boundary value problem on $[0,T]xR$

$$\begin{aligned}&\frac{\partial F(t,x)}{\partial t}+\mu(t,x)\frac{\partial F}{\partial x}+\frac{1}{2}\sigma^2(t,x)\frac{\partial^2 F}{\partial x^2}=0\\&F(T,x)=\phi(x)\end{aligned} \tag{3.10}$$

This problem is a boundary value problem in the sense that at time *T, the solution F coincides with the function* $\phi(x)$. We will now produce this solution in terms of a solution to a SDE that is related to (3.9) in a natural manner. The SDE is by now well known to us. It reads

$$\begin{aligned}&dX(t)=\mu(t,X(t))dt+\sigma(t,X(t))dB(t)\\&X(t)=x_t\end{aligned} \tag{3.1a}$$

If we use our operator we can now rewrite the boundary value problem in the following manner

$$\begin{aligned}&\frac{\partial F(t,x)}{dt}+AF(t,x)=0\\&F(T,x)=\phi(x)\end{aligned} \tag{3.10a}$$

Now, we apply Ito's formula (3.9) on the process $F(s,X(s))$ with $X(s)$ one-dimensional and integrate equation (3.9) forwards, ($F = f$). This gives

$$F(T,X(T))=F(t,x_t)+\int_t^T\{\frac{\partial F(s,X(s))}{\partial s}+AF(s,X(s))\}ds+\int_t^T\sigma(s,X(s))\frac{\partial F(s,X(s))}{\partial X}dB(t)$$

(3.11)

Since F solves the boundary problem the time integral vanishes on account of equation (3.10a). Moreover, if the process inside the Ito integral fulfills the condition (i) in section 2.2, it vanishes under the *expectation operator*. This leaves us with the result

$$F(t, x_t) = E^{tx_t}[\phi(X(T)] \tag{3.12}$$

where expectations are taken at t, given the initial value $X(t) = x_t$

This result is a version of a Theorem by Feynman and Kac. A related boundary value problem that appears over and over again in connection with the valuation of financial instruments within Financial Economics is the following

$$\begin{aligned} &\frac{\partial F(t,x)}{\partial t} + \mu(t,x)\frac{\partial F(t,x)}{\partial x} + \frac{1}{2}\sigma^2(t,x)\frac{\partial^2 F(t,x)}{\partial x^2} + rF = 0 \\ &F(T,x) = \phi(x) \end{aligned} \tag{3.13}$$

In Financial Economics *r* represents the interest rate, but, more generally, it is a real number. To modify our previous representation result we use the ODE idea of an integrating factor by multiplying the equation by the factor e^{rt} and apply Ito's formula to the process $Z(s, X(s)) = e^{rs}F(s, X(s))$, where $X(s)$ solves the SDE in equation (3.1a) We obtain the following result

Proposition 3.2 *Assume that F solves the boundary value problem*

$$\begin{aligned} &\frac{\partial F(t,x)}{\partial t} + \mu(t,x)\frac{\partial F(t,x)}{\partial x} + \frac{1}{2}\sigma^2(t,x)\frac{\partial^2 F(t,x)}{\partial x^2} + rF = 0 \\ &F(T,x) = \phi(x) \end{aligned}$$

Assume further that the process $\sigma(t, X(t))\frac{\partial F(t, X(t))}{\partial X}$ *is in* L^2, *whereX(t) is defined below.*

Then

F has the representation

$$F(t,x) = e^{r(T-t)}E^{t,x_t}[\phi(X(T)]$$

where X(t) satisfies the SDE

$$\begin{aligned} &dX(t) = \mu(t, X(t))dt + \sigma(t, X(t))dB(t) \\ &X(t) = x_t \end{aligned}$$

Proof: Left as an exercise to the reader (see exercise3.12.)

In the exercises below, we will illustrate how this result can be used to find explicit solutions to PDE problems. It is worth mentioning that formally the PDE problem in Proposition 3.2 will have infinitely many solutions, but only one that is “practically relevant”. The representation method just presented will give us this solution.

Exercises (from Björk(1998,2008):

3.8 Solve the PDE

$$\frac{\partial F(t,x)}{\partial t}+\frac{1}{2}\sigma^2\frac{\partial^2 F(t,x)}{\partial x^2}=0$$
$$F(T,x)=x^2$$

Solution: From Proposition 3.2 we immediately get $F(t,x_t)=E^{t,x_t}[X(T)^2]$, *where* $dX(s)=0\cdot dt+\sigma dB(s)$.

Now integrating the stochastic differential equation yields $x(T)=\sigma[B(T)-B(t)]+x(t)$, *and* $X(T)^2=x(t)^2+2x(t)\sigma[B(T)-B(t)]+\sigma^2[B(T)-B(t)]^2$. *Taking expectations yields* $E^{t,x_t}[x(T)^2]=x(t)^2+\sigma^2[T-t]$. *Since* $X(T)$ *has the distribution* $N[x,\sigma\sqrt{T-t}]$ *we can write F(t,x(t))=var(X(T))+* $\{E[X(T)]\}^2$.

Exercise 3.9

Use the stochastic representation result in order to solve the following boundary value problem in the domain [0,T]xR

$$\frac{\partial F(t,x(t))}{\partial t}+\mu x(t)\frac{\partial F(t,x(t))}{\partial t}+\frac{1}{2}\sigma^2\frac{\partial F(t,x(t))}{\partial t}=0$$
$$F(T,x)=\ln(x^2)$$

where $\mu, and \sigma$ *are given constants.*

Solution: Proposition 3.2 gives $F(t,x(t))=E_{tx}[\ln x(T)^2]$

and X(t) solves the differential equation

$$dX(t)=\mu X(t)dt+\sigma X(t)dB(t)$$

This equation (Geometric Brownian Motion) has the well known solution

$$X(T)=x(t)e^{(\mu-\frac{\sigma^2}{2})(T-t)+\sigma B(T)}$$

and

$X(T)^2 = x^2(t)e^{2[(\mu-\frac{\sigma^2}{2})+\sigma B(T)]}$ *implying that* $\ln[x(T)^2] = 2(\mu - \frac{1}{2}\sigma^2)(T-t) + 2\sigma B(T) + \ln x^2(t)$.

Taking expectations yields $F(t, x(t)) =$ $E^{t,x_t}\{\ln[x(T)^2]\} = 2(\mu - \frac{1}{2}\sigma^2)(T-t) + \ln x(t)$. *Check the solution by substituting it into the PDE.*

3.10 Prove that the problem boundary value problem

in the domain [0,T]xR

$$\frac{\partial F(t,x(t))}{\partial t} + \mu x(t)\frac{\partial F(t,x(t))}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial F(t,x(t))}{\partial t} + k(t,x) = 0$$
$$F(T,x) = \Phi(x)$$

has the stochastic representation formula $F(t,x(t)) = E_{t,x}\{\Phi(X(T)\} + \int_t^T E_{t,x}[k(s,X(s)]ds$, *and X(s) has the dynamics*

$$dX(t) = \mu(t,X(t))dt + \sigma(t,X(t))dB(t)$$
$$X(t) = x_t$$

Solution: Follow the ideas leading to equation (3.12)in the main text.

3.11 Use the result in the previous exercise to solve

$$\frac{\partial F(t,x(t)}{\partial t} + \frac{1}{2}x(t)^2\frac{\partial^2 F}{\partial x^2} + x(t) = 0$$
$$F(T,x(T)) = \ln[x(T)^2]$$

Solution: Here x(s) has the dynamics dX(s)=X(s)dB(s). Integrating yields

$X(T) = \int_t^T X(s)dB(s) + x(t)$. *Taking expectations we obtain* $E^{t,x_t}[x(T)] = x(t)$. *From the result of the previous exercise we know that* $F(t,x(t)) = E^{t,x_t}[\ln x(T)^2] + \int_t^T E^{t,x_t}[X(s)]ds$

By using that E[X(s)]=x(t) for all $s \geq t$ *(a martingale property), we get*

$\int_t^T E[X(s)]ds = x(t)[T-t]$. *In an attempt to determine* $E^{t,x_t}[\ln X^2(T)]$, *we put*

$Y(t) = \ln X(t)^2$ *Using the stochastic differential equation above and Ito's lemma on Y(t), we obtain dY(t)=2dB(t)-dt. Integrating gives* $y(T) = 2\int_t^T dB(s) + (t-T) + \ln x(t)^2$ *After taking* *expectation we have* $E^{t,x_t}[y(T)] = [t-T] + \ln x(t)^2$. *To sum up the solution to the boundary value problem is*

$F(t,x) = (t-T) + \ln x(t)^2 + x(t)(T-t)$

Check the solution by substituting it into the PDE.

3.12 Consider the following boundary value problem
$$\frac{\partial F(t,x)}{\partial t} + \mu(t,x)\frac{\partial F(t,x)}{\partial x} + \frac{1}{2}\sigma^2(t,x)\frac{\partial^2 F(t,x)}{\partial x^2} + r(t,x(t))F = 0$$
$$F(T,x) = \phi(x)$$

Prove that it has a stochastic representation formula of the form

$F(t,x(t)) = E_{t,x}[\phi(X(T)e^{\int_t^T t(s,X(s))ds}]$ *by considering the process*

$Z(s) = F(s,X(s))e^{\int_t^T r(s,X(s))ds}$ *(Note that this problem is a generalized version of Proposition 3.2.)*

Solution: Let X(t) be a process that solves

$dX(t) = \mu(t,X(t))dt + \sigma(t,X(t))dB(t)$
$X(t) = x_t$

Now $Z(s) = F(s,X(s))e^{\int_t^s r(\tau,X(\tau))d\tau}$, and by putting *A(s)=F(s,x(s)), and G(s)=* $e^{\int_t^s r(\tau,X(\tau))d\tau}$ *we can writeZ(s)=A(s)G(s). Ito's Lemma gives*

$dA(s) = (\frac{\partial F}{ds} + \mu\frac{\partial F}{\partial x} + \frac{1}{2}\sigma^2\frac{\partial^2 F}{\partial x^2})dt + \sigma\frac{\partial F}{\partial x}dB(s)$, *and* $dG(s) = G(s)r(s,X(s))ds$

Ito Lemma also gives dZ=AdG+GdA+dGdA=AGrds+GdA(s). After integration we have

(=0)

$z(T) - z(t) = \int_t^T A(s)G(s)r(s,X(s))ds + \int_t^T G(s)dA(s)$

Since z(t)=F(t,x(t)), and Z(T)= $F(T,x(T))\exp[\int_t^T r(s,X(s))ds] = \phi(X(T)\exp[\int_t^T r(s,X(s))ds$

it suffices to prove that E[Z (T)-z(t)]=0. To this end we note that since F solves the original PDE it holds that $dA(s) = -r(s, x(s))F(s, x(s))ds + \sigma(s, x(s))\frac{\partial F(s, x(s))}{\partial x}dB(s)$. *Substituting this into the expression for z(T)-z(t) and taking expectations proves the claim.*

Chapter 4: Stochastic optimal control (SOC)

4.1 The Hamilton-Jacobi-Bellman equation

This chapter will draw on the presentation of SOC in a recent monograph by Aronsson et.al[17]. (2004).The mathematical technicalities are presented in connection with a stochastic version of a Ramsey model[18]. The particular version of the model used here was introduced by Merton (1975). The Ramsey problem is to optimize the present utility value of the consumption stream over time by at each instant of time optimally choosing consumption and net investment subject to a convex production possibility set.

Merton treats the asymptotic properties of both the neoclassical growth model developed by Solow (1956) and Swan (1956), as well as the Ramsey (1928) optimal growth model, when the growth of the labor force follows a geometric Brownian motion process. We will concentrate on the Ramsey model and deal with both one and two sector versions. We will, in

[17] It is perhaps superfluous to mention that ideas there have been borrowed from the material that is listed in the introductory section.

[18] The deterministic version was introduced by the English logician and mathematician Frank Plumpton Ramsey, who also wrote three papers in Economics. At least two of them became classical contribution to Economics. The model under consideration here is published in Ramsey (1928). He died at the age of 27 from kidney decease.

particular, show how the Hamiltonian-Jacobi-Bellman (HJB) equation comprises deterministic optimal control as a special case.

Let $F(K, L) \in C^2(R^{2+})$ be a linear homogeneous net production function (i.e., depreciation has been accounted for), where K denotes units of capital input and L denotes units of labor input. The capital stock evolves according to

$$\dot{K}(t) = F(K(t), L(t)) - C(t) = L(t)F(k(t), 1) - C(t) \tag{4.1}$$

Let $k = K/L$, assume that $L(t) = L(0)e^{nt}$, $L(0) > 0$, $0 < n < 1$and differentiate totallay with respect to time. Using the linear homogeneity of the production function, it follows that

$$\dot{k}(t) = f(k(t)) - c(t) - nk(t) \tag{4.2}$$

where $f(k)$ is net output per capita and n is the growth rate of the population. Equation (4.2) is a variation of the Solow neoclassical differential equation of capital stock growth under certainty. Note that $dL/dt = nL$ or $dL = nLdt$.

Now, suppose that the growth of the labor force is described by the stochastic differential equation

$$dL = nL(t)dt + \sigma L(t)dB(t) \tag{4.3}$$

The stochastic part is $dB(t)$, where $B = B(t)$ is a Brownian motion process defined on some probability space. The drift of the process is governed by the expected rate of labor growth per unit of time, n. In other words, over a short interval of time, dt, the proportionate change of the labor force (dL/L) is normally distributed with mean ndt and variance $\sigma^2 dt$.

We are now ready to transform the uncertainty about the growth in the labor force into uncertainty about growth of the capital labor ratio $k = K/L$. We use Ito's lemma. To this end define

$$k(t) = \frac{K(t)}{L} = Z(L,t) \tag{4.4}$$

to obtain

$$dk = \frac{\partial Z}{\partial t}dt + \frac{\partial Z}{\partial L}dL + \frac{1}{2}\frac{\partial^2 Z}{\partial L^2}dL^2 \tag{4.5}$$

where

$$\frac{\partial Z}{\partial t} = f(k) - c \qquad dL = nLdt + \sigma LdB$$

$$\frac{\partial Z}{\partial L} = -\frac{K(t)}{L^2} = -\frac{k}{L} \qquad (dL)^2 = \sigma^2 L^2 dt \tag{4.6}$$

(4.6)

$$\frac{\partial^2 Z}{\partial L^2} = 2\frac{K(t)}{L^3} = \frac{2k}{L^2} \tag{4.7}$$

After substitutions into (4.5)

$$dk = [f(k) - c - (n - \sigma^2)k]dt - k\sigma dB \tag{4.8}$$

In other words, we have translated uncertainty with respect to the growth rate of the labor force into uncertainty with respect to capital per unit of labor and, indirectly, to uncertainty with respect to output per unit of labor, $y(t)=f(k(t))$.*We have in the first three chapters denoted stochastic processes by capital letters. Since the growth model is set up in rates, we find it here convenient to use lower case letters.*

We are now ready to formulate a variation of Merton's (1975) version of the stochastic Ramsey problem. The main difference lies in Merton's assumption that saving is a fixed proportion of production, and the control problem consists of choosing an optimal saving function. Here we choose an optimal consumption function.

Let $u(c(t))$ be a twice continuously differentiable and strictly concave utility function, where $c(t)$ denotes per capita consumption. The optimization problem is to find an optimal consumption policy, and the stochastic Ramsey problem can be written

$$\max_{c(t)} E_0 \left\{ \int_0^T u\big(c(t)\big) e^{-\theta t} dt \right\}; \tag{4.9}$$

subject to

$$dk(t) = [f(k(t)) - c(t) - (n - \sigma^2)k(t)]dt - \sigma k(t) dB(t) \qquad k(0) = k_0 \tag{4.10}$$

$$c(t) \geq 0 \quad \forall t$$

E_0 denotes that mathematical expectations are taken conditional on the information available at time zero. Note also, that given the state of the economy, by choosing $c(t)$, one indirectly chooses expected net investment.

The formulation of the optimization problem in equations (4.9) and (4.10) is incomplete in at least two respects. First of all, we have to specify the information on which the choice of the control function is based (this is not required in the deterministic Ramsey problem). In most contexts it is realistic to assume that the control process $c(t)$ is allowed to be conditioned solely on past observed values of the state process $k(t)$. In such a case, mathematicians would say that the control process is adapted to the state process. One special case is a control function of the form $c(t) = c(t, k(t))$, where $c(t)$ is a deterministic function. This is a feedback control law. More specifically, we condition the control on the state of the system at time t, i.e. it does not depend on the starting point (s, k). It is also called a Markov control.

Given that we have chosen a control law we can substitute it into the stochastic differential equation to obtain (4.10). An admissible control is also required to imply that the above stochastic differential equation has a unique solution.

A second problem with the above formulation is that there are restrictions on the control process but not on the capital stock (the state variable). Under such circumstances the problem is likely to be nonsensical. If the utility function is unbounded in $c(t)$, the consumer can increase his utility to any given level by consuming an appropriately large amount at every t. However, this would mean that the capital stock, or wealth, goes below zero, and when it does, goods would not be produced, only consumed. This problem can be handled in different ways, but one of the most elegant is the following. Define T as the first exit time from the solvency set $G = \{k_\tau(\omega); k_\tau > 0\}$, i.e. $T = \inf\{\tau > s; k_\tau(\omega) \notin G\} \leq \infty$. In other words, the process is stopped when the capital stock per capita becomes non-positive (when bankruptcy occurs). This formulation ensures that when the consumer holds no wealth, all activity is terminated. Before introducing the necessary condition for an optimal consumption path (control process), we define the value function and the optimal value function. The former is defined by

$$J(0,k_0,c) = E_0\{\int_0^{\mathrm{T}} u(c(k(t))e^{-\theta t}dt\}$$

given the dynamics in (4.10'). Here the bottom and top indexes $0, c$ denote that the path starts at zero and is driven by the control function $c(t,k(t))$. The optimal value function is defined by (skipping the top index c)

$$V(0,k_0) = \sup_c J(0,k_0,c)$$

In other words, $J(0,k_0,c)$ is the expected utility of using the control function *c(t,k(t))* over th e time interval $[0,\mathrm{T}]$, given the starting point k_0 at time zero. The optimal value function gives the expected maximum utility over the same interval, starting at the initial condition. If we start at t with initial condition k_t, the optimal value function is $V(t,k_t)$. Here I use the supremum norm rather than the maximum to be a little snobbish. It is now time to introduce the following theorem:

Theorem 4.1 *(The Hamilton – Jacobi - Bellman or HJB equation) Assume that*

- *the optimal value function V is in C^{12}*

- *if an optimal control* c^* *exists, then*

  the following will hold

  (i) V satisfies the equation $\frac{\partial V(t,k)}{\partial t} + \sup_c\{u(c^*(t))e^{-\theta t} + h\frac{\partial V(t,k)}{\partial k} + \frac{1}{2}\frac{\partial^2 V(t,k)}{\partial k^2}\sigma^2 k^2\} = 0$

  where $h(k,c,n,\sigma^2) = f(k) - c - (n-\sigma^2)k$ *;*

  (ii) *the transversality conditions is V(T,k)=0*

  (iii) *for each* $(t,k) \in D$ *the supremum in the HJB is attained by* $c^* = c^*(t,k)$

  where D=[0,T]$\times R_+$.

A sketch of a proof is outlined in an Appendix 4A[19].

That the value function belongs to C^{12} means that it is once continuously differentiable in time and twice continuously differentiable in $k(t)$. Theorem 4.1 is a necessary condition for an optimal path. It is, however, interesting and important to know that the *HJB equation* also acts as a sufficient condition for optimum. The result is typically referred to as the *Verification Theorem.* A little informally, it tells us that if there is another function $V^\circ(t,k(t))$, and an admissible optimal control $c^\circ(t,k(t))$ that solve the HJB-equation, then these functions coincides with the optimal value function and the optimal control. In other words, having found a solution to the HJB equation, means that one has "found" the optimal solution. The qualification is that one has to assume that the integrability condition for the stochastic integrals is fulfilled. The proof is available in Björk (1998, 2008), Chapter 14/19 and in Öksendahl (2003), Chapter 11.

4.2 Relating the HJB-equation to Deterministic Optimal Control

The only non-autonomous time dependence (time as a separate argument) in the above problem is introduced through the discount factor. This means that the problem under an infinite planning horizon can be rewritten on a more convenient form. Starting from the value function in present value

$$V(t,k_t) = \underset{c}{\mathrm{Max}}\, E_t\left\{\int_t^T u[c(\tau)]e^{-\theta\tau}d\tau\right\}; \qquad (4.11)$$

[19] A formal proof is found in Öksendahl (2003), Chapter 11.

subject to equation (4.10), which is the Brownian motion equation for the capital stock, and k(t) = k_t. We have

$$e^{\theta t}V\left(t,k_t\right)=\underset{c}{Max}\,E_t\{\int_t^{T}u[c(\tau)]e^{-\theta(\tau-t)}d\tau\}=W(t,k_t) \qquad (4.11')$$

$$k(t)=k_t$$

where $W(t,k_t)$ is the optimal current value function . For the case when $T=\infty$ it is straightforward to prove that for a Markov control $c=c(k(t))$. The observation does also hold for the construction with T as the first exit time from the solvency set[20].

Observation 1: $V(t,k_t)=V(0,k_t)e^{-\theta t}$

Proof: $V(t,k_t)=e^{-\theta t}\max_c E[\int_t^{\infty}u(c(\tau)e^{-\theta(\tau-t)}d\tau]=e^{-\theta t}\max_c E[\int_0^{\infty}u(c(\tau')e^{-\theta\tau'}]d\tau'=e^{-\theta t}V(0,k_t)$

Observation 1 means that the current value function, $W(k_t)=V(0,k_t)$, does not depend on the starting point. This implies that

$$V_t=\frac{d}{dt}\left[e^{-\theta t}W\right]=-\theta e^{-\theta t}W$$

and the HJB equation can be rewritten in the following manner

$$\theta W=\underset{c}{Max}\left[u(c(t)+W_k h(k,c;\sigma^2,n)+\frac{1}{2}\sigma^2k^2W_{kk}\right] \qquad (4.12)$$

where $W_k = dW(\cdot)/dk$ and $W_{kk} = d^2W(\cdot)/dk^2$. We can now define a co-state variable $p(t)$ as

$$p(t) = W_k\,(k(t)), \qquad (4.13.)$$

and its derivative

---

[20] See Li and Löfgren (2012).

$$\frac{dp(t)}{dk} = W_{kk}(k(t)) \tag{4.14}$$

Given the optimal consumption policy, (4.12) can be written (neglecting the top index on *k(t)*, and the time index to save notational clutter) as

$$\theta W(k_t)) = u(c^*) + ph(k,c^*;\sigma^2,n) + \frac{1}{2}\frac{dp}{dk}\sigma^2 k^2 = H^{c^*}(k,p,\frac{dp}{dk}) \tag{4.12'}$$

The function $H^{c^*}(\cdot)$ can be interpreted as a "generalized" Hamiltonian in current value terms (see below). The Hamiltonian plays a key role in deterministic optimal control theory, DOC. We can now calculate how the variables *k* and *p* develop over time along an optimal path. Using (4.10') and the definition of H^{c^*}, we obtaind

$$dk = h(k,c^*;\sigma^2,n)dt + \sigma kdB = H_p^{c^*}(k,p,\frac{dp}{dk})dt + \sigma kdB \tag{4.15}$$

where $H_p^{c^*} = dH^{c^*}/dp$. Equation (4.15) describes how *k* develops over time under the optimal consumption policy. To find the corresponding condition for *p*, we use Ito's lemma and derive

$$d\tilde{p} = \left[V_{kt} + V_{kk}h + \frac{1}{2}V_{kkk}\sigma^2 k^2\right]dt - V_{kk}\sigma kdB \tag{4.16}$$

where $\tilde{p}(t) = p(t)e^{-\theta t}$. As in the case of perfect certainty, it is often convenient to relate $d\tilde{p}$ to derivatives of "the Hamiltonian". Using the expression $-V_t = H^{c^*}e^{-\theta t} = \theta We^{-\theta t}$ to compute V_{kt}, we can, after substitutions, rewrite (4.16) to read

$$d\tilde{p} = -H_k^{c^*}e^{-\theta t}dt - \sigma ke^{-\theta t}W_{kk}dB \tag{4.17}$$

where $H_k^{c^*} = dH^{c^*}/dk$. Next, since $d\tilde{p} = (dp - \theta pdt)e^{-\theta t}$, (4.17) is easily transformed to current value terms, i.e.,

$$dp - \theta p dt = - H_k^{c^*} dt - \sigma k W_{kk} dB \tag{4.18}$$

Let us now interpret equations (4.12’), (4.15) and (4.18). In so doing, we relate them to their counterparts under certainty. Equation (4.12’) clearly implies that the generalized Hamiltonian in current value is maximized with respect to the control variable, since it is directly proportional to the optimal value function.

.

The interpretation of the generalized current value Hamiltonian is that it is the sum of the instantaneous utility, the expected infinitesimal increment of capital valued at its marginal expected current value, plus the valuation of the risk associated with a given investment. The stochastic differential equation (4.15) reveals how capital evolves over time along the optimal path, while equation (4.18) is the corresponding stochastic differential equation for the development of the co-state variable over time which, by definition, is interpreted as the derivative of the optimal value function with respect to the state variable.

In the deterministic case $\sigma = 0$ and

$$H^{c^*} = u(c^*(t)) + ph(k^*(t), c^*(t); 0, n) = \theta \int_t^{\infty} u(c^*(t)) e^{-\theta(s-t)} ds \tag{4.19}$$

which shows that the maximized (deterministic) Hamiltonian is directly proportional to future utility along the optimal path. Equation (4.19) is the main result in Weitzmann’s (1976) classical paper on the welgfare significance of comprehensive NNP (Net National Product); a special case of the appropriate welfare measure under uncertainty. Moreover, since the Brownian increments are eliminated from the equations for *dk* and *dp*, the time derivatives *dk*/*dt* and *dp*/*dt* are well defined. Hence, we have

$$\begin{aligned} \frac{dk^*(t)}{dt} &= h(k^*(t), c^*(t); 0, n) = f(k^*(t)) - nk^*(t) - c^* \\ \frac{dp(t)}{dt} - \theta p(t) &= -\frac{dH^{c^*}(t)}{dk} \end{aligned} \tag{4.20}$$

Readers that are familiar with Pontryagin's maximum principles, or DOC, recognize that the conditions for an optimal path follow directly from the HJB-equation.

To sum up, we have shown how the Hamilton – Jacobi - Bellman equation from stochastic control theory can be used to derive the appropriate welfare measure under uncertainty, which turns out to be analogous to its deterministic counterpart. A generalized Hamiltonian is directly proportional to the expected future utility along the optimal path. Not surprisingly, but neatly, the stochastic welfare measure collapses to the corresponding deterministic measure, when $\sigma \to 0$. More generally, but less precisely, deterministic optimal control theory is a special case of stochastic optimal control theory.

A stochastic local welfare criterion[21]

Equation (4. 19) above is a global welfare indicator in the sense that the current value Hamiltonian is proportional tho the value function along an optimal path; the discounted value of all future utilities. There is also a corresponding deterministic local welfare indicator in terms of the time derivative of the optimal value function. It is called genuine saving (GS) and has for more than 10 years ago been a statistics that is published for a vast number of countries. Here we will generalize the concept into an Ito world

To derive a local welfare measure like GS we start the optimal value function

$$W(s,k) = \underset{c}{Max}\, E_t''\{\int_s^T u(c(\tau)e^{-\theta(\tau-s)}d\tau\} = E_s\{\int_s^T u(c^*(\tau)e^{-\theta(t-s)}d\tau\} \tag{4.21}$$

Differentiating with respect to time (the lower intergration level) yields

$$\frac{\partial W}{\partial s} = \dot{W}(k)) = -u(c^*(k)) + \theta W(k) \tag{4.22}$$

Now, using equation (4.12), i.e., the HJB-equation for the time autonomous problem we get, after substituting for $\theta W(k)$,

[21] This section also builds on Li and Löfgren (2012)

$$\dot{W}(s) = p(s)h(c^*(k(s));\sigma^2, n) + \frac{1}{2}\frac{dp(s)}{dk}\sigma^2 k^2 =$$
$$W_k(k)[f(k(s) - c^*(k(s)) - (n - \sigma^2)] + \frac{1}{2}W_{kk}(k)\sigma^2 k^2 \qquad (4.23)$$

The interpretation of the co-state variable $p(s)$ is the derivative of the optimal value function with respect to the initial capital stock, and $h(\bullet)$ is the drift in net investment along the optimal path. The second term in the expression originates from Ito calculus and the sign of this second order derivative of the value function with respect to the capital stock, W_{kk}, or, what amounts to the same, the derivative of the co-state variable (the shadow utility value of net investment) from an increase in the capital stock. For a "well behaved" maximization problem this entity should be negative. For $\sigma = 0$ equation (4.23) collapses to the static GS measure. This means that we would under a stochastic Ramsey problem expect that a positive net investment value would not be enough to indicate a local welfare improvement. Net investment has to be large enough to compensate for the variance component. In the variance component we interpret $-W_{kk}(k(t))$ as the price of risk, and $\sigma^2 k^2$ as the "quantity of risk".

The reason why this particular component appears is that an Ito integral is constructed from forward increments. An alternative way of constructing a stochastic integral is the Stratanovich integral[22], which picks the middle of the increment to weigh the components of the sums that approximates the integral. For a whole economy, where risk cannot be diversified away the Ito integral seems reasonable. However, if risk can be diversified a stochastic integral which leaves out the risk component in expressions like (4.23) is more relevant.

To find the solution in the general time autonomous case with *n* consumption goods and *m* capital goods the above procedure can be generalized. We will only have to change into a general HJB-equation. The derivative of the value function will look like the one in equation (4.22). In other words, we are left with the following result.

[22] The seminal reference is Stratonovich (1966).

Observation 2: *In a stochastic time autonomous Ramsey problem with n consumption goods and m capital goods the derivative of the value function with respect to time is given by the following expression* $\dot{W}(k(s)) = HJB - u(\mathbf{c}^*)$.

If the problem is not time autonomous extra first order terms will be added in the HJB equation and change the time derivative accordingly. An example would be exogenous technological progress, which would add net value to the GS component. Another example would be negative externalities that would deduct net value from the GS component. Finally, a Markov control may seem overly specific. A more general control would be to allow the control at time *t* to be conditioned on the whole process from start up to *t*, i.e., the control function is $F_t - adapted$. Such controls are called closed loop or feed back controls. Under an integrability condition and a smoothness condition on the set *G* it is possible to show that the optimal value function for the Markov control coincides with the optimal control for the open loop control for any starting point in *G* . Hence, the Markov control is not particular restricted[23].

[23] See e.g. Öksendal (2003) Theorem 11.2.3.

4.3 A two sector model[24]

We now augment the stochastic version of the Ramsey model with a stochastic pollution equation and a pollution externality. More specifically, we introduce a stochastic version of the model in Brock (1977). Hence the stochastic population growth which generated the stochastic Ramsey problem is retained.

We modify the objective function to read

$$U(0)=\int_0^\infty u(c(t),x(t))e^{-\theta t}dt \tag{4.24}$$

In other words, we insert the stock of pollution, $x(t)$, as an additional argument in the utility function. The marginal utility of pollution, $u_x(\bullet)$, is, of course, assumed to be negative. The evolution of the capital stock per capita obeys the stochastic differential equation

$$\begin{aligned} dk(t) &= [f(k(t),g(t))-c(t)-(n-\sigma_1^2)]k(t)dt-\sigma_1 k(t)dB_1(t) \\ k(0) &= k_0 \end{aligned} \tag{4.25}$$

where $g(t)$ is interpreted as the input of energy per capita, and $B_1(t)$ is one dimensional Brownian motion. The stock of pollution evolves according to

$$\begin{aligned} dx(t) &= g(t)x(t)dt-\sigma_2\gamma dB_2 \\ x(0) &= x_o \end{aligned} \tag{4.26}$$

This means that *x(t)* follows a geometric Brownian motion process with drift. Here $B_2(t)$ is one dimensional Brownian motion.The shape of the process is chosen to keep $x(t)$ positive. The reader knows at this stage that the solution has the form

$$x(t)=x_0 e^{\left(\int_0^t g(s)ds-\frac{\sigma_2^2\gamma^2 t}{2}-\sigma_2\gamma B_2\right)} \tag{4.27}$$

---

[24] Although we are dealing with two dimensional vectors we will not switch to vector notation.

As in section 4.2 we assume that the control process is adapted to the state process and we choose to allow a feed back control. If we define

$$y(t) = \begin{bmatrix} c(t) \\ g(t) \end{bmatrix} \qquad \kappa(t) = \begin{bmatrix} k(t) \\ x(t) \end{bmatrix} \qquad B(t) = \begin{pmatrix} B_1(t) \\ B_2(t) \end{pmatrix} \tag{4.28}$$

the control process can be written in the following manner $y(t) = y(t, \kappa(t))$, where $y(t)$ is a deterministic control function. By substituting the control functions into the stochastic differential equations (4.25) and (4.26) we obtain

$$d\kappa(t) = \begin{bmatrix} dk(t) \\ dx(t) \end{bmatrix} = \begin{bmatrix} h(c(t,k,x)g(t,k,x),k;\sigma_1^2,n) \\ e(t,k,x) \end{bmatrix} dt - \begin{bmatrix} \sigma_1 k \\ \\ \sigma_2 \gamma \end{bmatrix} dB \tag{4.29}$$

$$\kappa(0) = \begin{bmatrix} k_o \\ x_0 \end{bmatrix}$$

where $e = g(t)x(t)$

As in the preceding section an admissible control is required to imply that the above system of stochastic differential equations has a unique solution. We also require that $y(t) \geq 0$. Moreover, to avoid a nonsensical solution, we assume that T is the first exit from the solvency set.

Hence, we can write the optimal value function as

$$V(0, s_0) = \sup_y \int_0^{\mathrm{T}} u(c(t), x(t)) e^{-\theta t} dt \tag{4.30}$$

which is optimized subject to equations (4.29). Again, the value function will satisfy a HJB equation similar to that in Theorem 4.1 above. The generalized HJB equation can be written

$$\frac{\partial V(t,\kappa(0))}{\partial t}+\sup_{y}\{u(c(t),x(t))e^{-\theta t}+L^{y}V(t,\kappa(0)\}=0 \;\; \forall(t,\kappa)\in D \tag{4.31}$$

with tranversality conditions $V(T,\kappa)=0$. Here L^{y} is a partial differential operator, which will be explained below. Next, start from equation (4.26) and write compactly in vector notation

$$\mathbf{d}\kappa_{\mathbf{t}}^{\mathbf{y}}=\mathbf{a}^{\mathbf{y}}(\mathbf{t},\kappa)dt-\boldsymbol{\sigma}^{y}(\mathbf{t},\kappa)d\mathbf{B}(\mathbf{t}) \tag{4.29'}$$

where the top index denotes that the process is driven by the control function *y(t)* or a fixed vector *y*. To clarify $\boldsymbol{\sigma}^{\mathbf{y}}(\mathbf{t},\boldsymbol{\kappa})=\boldsymbol{\sigma}^{\mathbf{y}}(\mathbf{t},\boldsymbol{\kappa},\mathbf{y}(\mathbf{t},\boldsymbol{\kappa}))$, if the process is driven by a control function, and if the control vector is fixed, y is substituted for $y(t,\boldsymbol{\kappa})$.
We now define a matrix[25]

$$\mathrm{M}^{y}=\boldsymbol{\sigma}^{\mathbf{y}}(\mathbf{t},\boldsymbol{\kappa},\mathbf{y})\boldsymbol{\sigma}^{\mathbf{y}}(t,\boldsymbol{\kappa},\mathbf{y})' \tag{4.32}$$

where the prime denotes the transpose of a vector. The partial differential operator, L^{y}, can now be defined as

$$L^{u}=\frac{\partial}{\partial t}+\sum_{i=1}^{2}a^{u}(t,\kappa)\frac{\partial}{\partial\kappa_{i}}+\frac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{2}M_{ij}^{u}\frac{\partial^{2}}{\partial\kappa_{i}\partial\kappa_{j}}$$

with an obvious modification for a case with *n* stochastic differential equations. For the present case with one SDE's, after applying the operator we have the *HJB*

$$\sup_{u}\{u(c(t)x(t))e^{-\theta t}+\frac{\partial V(t,X)}{\partial t}+\frac{\partial V(t,X)}{\partial X}h(t,X)+\frac{1}{2}\sigma^{2}X^{2}\frac{\partial^{2}V(t,X)}{\partial X^{2}}=0$$

---

[25] Here σ^{y} is a 2×1 vector.

$f(\bullet)$ is a given bounded continuous function and $\rho > 0$, and the *inf is taken over* $[t, T^{+}]$ and *u(t)* is a homogeneous Markov control, i.e. $u = u(x(t))$. It is decently ease to prove that the optimal value function can be written as $V(s,x) = e^{-\rho s}\Gamma(x),$ where $\Gamma(x) = V(0,x).$ The reason is that that we have an autonomous problem, i.e. it does not deal with clock time expression corresponding to equation (4.12), i.e.,

$$-\frac{\partial V(t,s)}{\partial t} = \sup_{y \in A} H(t,\kappa, y, \tilde{p}, \frac{\partial \tilde{p}}{\partial \kappa}) = H^{*}(t,\kappa,\tilde{p},\frac{\partial \tilde{p}}{\partial \kappa}) \tag{4.35}$$

Here H^{*} is the generalized present value "Hamiltonian", and $\tilde{p}(t) = (\tilde{p}_k, \tilde{p}_X) = (\frac{\partial V}{\partial k}, \frac{\partial V}{\partial x})$ defines the stochastic co-state variables in present value. Provided that the increments dB_1 and dB_2 are uncorrelated, a multidimensional analogue of the co-state stochastic differential equations has the following shape

$$d\tilde{p}_k = -H_k^{*}dt - \frac{\partial^2 V}{\partial k^2}\sigma_1 k dB_1 - \frac{\partial^2 V}{\partial k \partial x}\sigma_2 \gamma x dB_2 \tag{4.36}$$

$$d\tilde{p}_x = -H_x^{*}dt - \frac{\partial^2 V}{\partial x^2}\sigma_2 \gamma x dB_2 - \frac{\partial^2 V}{\partial x \partial k}\sigma_1 k dB_1$$

The derivation of the general form of the stochastic co-state equations follows the reasoning in Section 4.3, where there is only one state variable. Thus, we can begin by defining the co-state variable as the derivative of the optimal value function with respect to the state variables and take the first differential using Ito's Lemma. The resulting expression contains a term which is the cross derivative of the optimal value function with respect to time and the state variable. The shape of this derivative can be obtained by taking the first derivative of the HJB-equation with respect to the state variable, again using Ito calculus. Substituting the resulting expression for the cross derivative into the original co-state differential equation and canceling terms gives the result in (4.33). For details, see Appendix A. The calculations in the n-state variable case are straightforward, although somewhat messy[26]. They are therefore omitted.

4.4 Stochastic cost-benefit rules[27]

The form of the co-state equation (4.36) contains the key to the shape of a cost-benefit rule under Brownian motion. Since the co-state variable measures the marginal contribution to the value function due to an increase in the state variable, they can be used to derive a cost benefit rule. The trick is to introduce an artificial or, rather, an unnecessary state variable in terms of a parameter that describes a project. In the model above, the parameter γ could represent a project that improves the assimilative capacity of the environment. Since it is a constant, we can write its differential equation as $d\gamma = 0,\ \gamma(0) = \gamma$. This gives us three stochastic differential equations, one of which is highly deterministic. We can nevertheless elicit a current value co-state variable by defining it as the partial derivative of the optimal value function, i.e., $\tilde{p}_\gamma = \Omega_\gamma = \partial V / \partial \gamma$. We can then use the general form of the co-state equation in (4.36) to write

$$d\Omega_\gamma = -H_\gamma^*(\cdot)dt + \frac{\partial^2 V}{\partial \gamma^2}\sigma_3 dB_3 - \frac{\partial^2 V}{\partial \gamma \partial k}\sigma_1 k dB_1 - \frac{\partial^2 V}{\partial \gamma \partial x}\sigma_2 \gamma x dB_2 \qquad (4.37)$$

However, $\sigma_3 \equiv 0$ by assumption, and we can integrate (4.37) over the interval (t, t_1) to get

$$\Omega_\gamma(t_1) = \Omega_\gamma(t) - \int_t^{t_1} H_\gamma^*(\cdot)ds - \int_t^{t_1} \frac{\partial^2 V}{\partial \gamma \partial k}\sigma_1 k dB_1 - \int_t^{t_1} \frac{\partial^2 V}{\partial \gamma \partial x}\sigma_2 \gamma x dB_2 \qquad (4.38)$$

Since $\Omega_\gamma(T) = 0$, according to the tranversality condition we obtain the cost benefit rule as

$$\Omega_\gamma(t) = \int_t^T H_\gamma^*(\tau)d\tau + \int_t^T \frac{\partial^2 V}{\partial \gamma \partial k}\sigma_1 k dB_1 + \int_t^T \frac{\partial^2 V}{\partial \gamma \partial x}\sigma_2 \gamma x dB_2 \qquad (4.39)$$

---

[26] The n-dimensional case is not difficult to guess, however.

[27] See also Aronsson, Löfgren and Nyström (2003).

Taking mathematical expectations of both sides and using the fact that the last two integrals are Ito-integrals, we have

$$E(p_\gamma) = E\{\int_t^T H_\gamma^*(\tau) d\tau\} \tag{4.40}$$

which is a close analogue to the corresponding dynamic cost-benefit rule in the deterministic case[28], only the expectation operator differs .

Project uncertainty can be introduced in this context by specifying the differential equation for the project state variable as

$$d\gamma = \sigma_3 dB_3, \quad \gamma(0) = \gamma \tag{4.41}$$

All terms in equation (4.37) are relevant, and equation (4.39) will contain one more Ito integral. In mathematical expectations the answer will look the same as in equation (4.40).

[28] See e.g. Aronsson et. al. (2004) Chapter 2.

4.5 Additional comments on the solution of the HJB-equation

Theorem 4.1 and its multidimensional analogue comprise a necessary condition, since the theorem states that if $\hat{y}$ is an optimal control, then the value function fulfills the HJB equation, and $\widehat{V}$ realizes the supremum in equation (4.30). The formal proof of this theorem is rather involved and omitted here; an intuitive informal sketch can be found in e.g. Björk (1998, 2009) and in Appendix A. The proof of a slightly more general theorem, where the ad hoc assumption that a solution exists is relaxed, may be found in Öksendal (2003, Chapter 11).

An important aspect is that the HJB equation also acts as a sufficient condition for optimum. The so called "verification theorem" states that if there are two more functions *V(t,x)* and *y(t,x)* , where *V(t,x)* solves the HJB equation and the admissible control function *y(t,x)* implies that the infimum is attained in equation (4.31), then the former function is identical to the optimal value function, while the latter is the optimal control function. The proof is accessible, and can be studied in both Björk (1998, 2009) and Öksendal (2003).

Surprisingly, the fact that we have restricted the control function to be a feedback (or Markov) control is not so restrictive. One can show that it typically coincides with the optimal control conditioned on the whole history of the state process[29]

Technically and schematicallly, one handles the solution of a stochastic control problem in the following manner. Treat the HJB equation as a partial differential equation for an unknown function V and fix the an arbitrary point $(t, X(t)$ and solve the static optimization problem

$$\max_{y \in A}[f(t,K,y)e^{-\rho t} + L^{y}V(t,K)]$$

The optimal solution will depend on the arbitrarily point and the function V . We can use $\hat{y} = \hat{y}(t,K,V)$ and substitute this into the *HJB* and solve for the the value function V .The last step entails the hardest problems. However, it is often helpful borrowing the form of the function $f(\bullet)$ as a blue print for the form of $V(\bullet)$.

---

[29] See Theorem 11.2.3 in Öksendahl (2003).

$f(\bullet)$ is a given bounded continuous function and $\rho > 0$, and the max *f is taken over* $[t, T^{+}]$ and $y(t)$ is a homogeneous Markov control, i.e. $y = y(x(t))$. It is decently ease to prove that the optimal value function can be written as $V(s,x) = e^{-\rho s}\Gamma(x),$ where $\Gamma(x) = V(0,x).$ The reason is that we have an autonomous problem, i.e. it does not deal with clock time.

The following exercise illustrates the solution process as well as the derivation of the cost benefit rule. We consider the following stochastic control problem osv

$$V(t,x) = \min_{c} E_t[\int_t^{\infty} (x^2(s) + c^2(s))e^{-\rho s} ds$$

where the underlying process is given by

$$\begin{aligned} dX(s) &= c(s)ds + \sigma\gamma dB(s) \\ X(t) &= x_t \end{aligned}$$

To derive the cost-benefit rule- the derivative of the value function with resprct to the project parameter, we define $\tilde{p}_\gamma = V_\gamma$, and we also calculate $E_t(\tilde{p}_\gamma)$, where the sub index *t* indicates that the process starts at time *t*. The problem can be approached in two ways: either explicitly solve the stochastic optimal control problem and develop all expressions explicitly before carrying out the calculation or, more simply use the cost benefit rule in equation (4.37). The second approach means less work.

We start with the first approach. The HJB-equation becomes

$$-\frac{\partial V}{\partial t} = \min_{c}[e^{-\rho t}(x^2 + c^2) + c\frac{\partial V(t,x)}{\partial x} + \frac{1}{2}\sigma^2\gamma^2\frac{\partial^2 V(t,x)}{\partial x^2}]$$

Minimizing with respect to the control variable gives

$$c = -\frac{1}{2}e^{\rho t}\frac{\partial V(t,x)}{\partial x}$$

Inserting the expression for the control variable into the HJB equation, we obtain

$$0 = e^{-\rho t}x^2 - \frac{1}{4}e^{\rho t}(\frac{\partial V(t,x)}{\partial x})^2 + \frac{\partial V(t,x)}{\partial t} + \frac{1}{2}\sigma^2\gamma^2\frac{\partial^2 V(t,x)}{\partial x^2})$$

By using separation of variables as a blueprint for the value function,

write $V(t,x) = e^{-\rho t}\phi(x),\ and\ \phi(x) = ax^2 + b$. We may then solve for the parameters. If we solve for our guess we get $c = -ax$. Substituting this into the *HJB*-ekvation we obtain:

$$x^2(1-\rho a - a^2) + \sigma^2\gamma^2 - \rho b = 0 \qquad (*)$$

This is only possible if:

$a^2 + \rho a - 1 = 0$ and $b = \sigma^2\gamma^2 a/\rho$

The positive root is:

a= $\frac{1}{2}[-\rho + \sqrt{\rho^2 + 1}]$

If the control is only allowed to assume non negative values we may conclude, by referring to Theorem 11.2.2. in Öksendahl (2003) and in Björk Theorem 19.6 (2009), that we have found the unique solution to the stochastic optimal control problem under consideration. In fact, the parameter *a* does not depend on γ. The optimal value function is given by

$$V(t,x) = e^{-\rho t}(x^2 + \frac{\sigma^2\gamma^2}{\rho})a$$

and the minimized present value Hamiltonian becomes

$$H^* = e^{-\rho t}[x^2(1-a^2) + \sigma^2\gamma^2 a]$$

Therefore,

$$\tilde{p}_\gamma = 2e^{-\rho t}\frac{\sigma^2}{\rho}\gamma a = \int_t^\infty H^*_\gamma(s)ds$$

In this particular case, taking expectations makes no difference.

Now moving to the second approach, by definition

$$-\frac{\partial V(t,x)}{\partial t} = H^*(t,x,\gamma,V_x,V_{xx})$$

Differentiation with respect to γ gives

$$H_{\gamma}^{*}=\sigma^{2}\gamma V_{xx}=2e^{-\rho t}\sigma^{2}\gamma a$$

where we have used the explicit solution. Therefore according to our cost-benefit result

$$E_{t}(\tilde{p}_{\gamma})=E_{t}[\int_{t}^{\infty}H_{\gamma}^{*}(s)ds=2e^{-\rho t}\frac{\sigma^{2}}{\rho}\gamma a\ .$$

Exercises (The exercises are borrowed from Öksendal (2003)): We continue to use the notation inf (sup) that Öksendal uses, but nothing essential would be changed if we write min(max) instead.

4.1 Produce the Hamilton-Jacobi-Bellman equation for the problem

$$\Phi(s,x)=\inf_{u}E^{s,x}\{\int_{s}^{\infty}e^{-\rho t}[g(X(t))+u(t)^{2}]dt\}$$

where $dX(t)=u(t)dt+dB(t)$, $u(t),X(t),B(t)\in R$

$X(s)=x_{s}$, *where* $\rho>0$, *and g is a real bounded continuous function.*

Solution: The HJB equation is obtained by applying the operator L^{u} *on the value function. One obtains*

$$h(u)=\inf_{u}\{L^{u}\Phi(s,x)+e^{-\rho t}[g(x)+u^{2}]=\inf_{u}\{\frac{\partial\Phi}{\partial s}+u\frac{\partial\Phi}{\partial x}+\frac{1}{2}\frac{\partial^{2}\Phi}{\partial x^{2}}+e^{-\rho t}[g(x)+u^{2}]\}=0$$

For a fixed (s,x) optimum is obtained as the solution to

$$h^{'}(u)=\frac{\partial\Phi}{\partial x}+2e^{-\rho s}u=0\ \textit{i.e.,}$$

$$u=-\frac{\partial\Phi}{\partial x}\frac{e^{\alpha s}}{2}$$

4.2 Consider the stochastic control problem

$$\Phi(s,x)=\inf_u E^{sx}\{\int_t^\infty e^{-\rho s} f(u(s),X(s))ds\}$$
$$dx(t)=b(u(t),X(t))dt+\sigma(u(t),X(t))dB(t)$$
$$X(t)\in R^n, u(t)\in R^k, B(t)\in R^m$$

f is a given bounded continuous real function and $\rho>0$, *and the inf is taken over* $[0,\infty)$, *and u is a time homogenous Markov control, i.e., u=u(x(t)). Prove that* $\Phi(s,x)=e^{-\rho s}\xi(x)$ *, where* $\xi(x)=\Phi(0,x)$

Solution: The claim is a general version of Observation 1 in section 4.2, which was proved above. The reason why the "trick" works is that the only explicit time dependence (clock time matters) is through the discount factor. Note that the coefficients in the stochastic differential equation are also independent of clock time. Such a SDE is called an Ito diffusion. The technicalities are left to the reader. If the coefficients had an explicit time dependence, it would be transferred to the HJB-equation by adding a new variable and modifying the operator that is used to derive theHJB equation. Prove the same result

for $T=\inf[t>0;k(t)=0]\Lambda\infty$.!!

4.3 Define $dX(t)=ru(t)X(t)dt+\alpha u(t)X(t)dB(t)$, $x(t),u(t),B(t)\in R$

and

$$\Phi(s,x)=\sup_u E^{s,x}\{\int_s^\infty e^{-\rho t} f(X(t))dt\}$$

where r,α,ρ *are constants and* $\rho>0$.

a) *Show that the HJB can be written* $\sup_u\{e^{-\rho t} f(x)+\frac{\partial\Phi}{\partial t}+ru(t)\frac{\partial\Phi}{\partial x}+\frac{1}{2}\sigma^2u^2\frac{\partial^2\Phi}{\partial x^2}\}=0$

Deduce that $\frac{\partial^2\Phi}{\partial x^2}\le 0$

b) *Assume that* $\frac{\partial^2\Phi}{\partial x^2}<0$. *Prove that* $u^*(t,x)=\frac{-r\frac{\partial\Phi}{\partial x}}{\alpha^2x\frac{\partial^2\Phi}{\partial x^2}}$ *and that*

$$2\alpha^2(e^{-\rho t}f+\frac{\partial\Phi}{\partial t})\frac{\partial^2\Phi}{\partial x^2}-r^2(\frac{\partial\Phi}{\partial x})^2=0$$

c) Assume that $\frac{\partial^2\Phi(s,x)}{\partial x^2}=0.$ *Prove that* $\frac{\partial\Phi}{\partial x}=0$ *and* $e^{-.\rho t}f(x)+\frac{\partial\Phi}{\partial t}=0$

d) Assume that $u^*(t)=u^*(x(t))$ *and that b) holds. Prove that* $\Phi(t,x)=e^{-\rho t}\xi(x)$ *and*

$$2\alpha^2(f-\rho\xi)\xi''-r^2(\varsigma')^2=0$$

Solution:

a) The HJB can bewritten

$$h(u)=\sup_u\{L^u\Phi(t,x)+e^{-\rho t}f(x)\}=0=\sup_u\{\frac{\partial\Phi}{\partial t}+rux\frac{\partial\Phi}{\partial x}+\frac{1}{2}\alpha^2u^2x^2\frac{\partial^2\Phi}{\partial x^2}+e^{-\rho t}f(x)\}$$

For the supremum to be finite $g(u^*)=u^*\{rx\frac{\partial\Phi}{\partial x}+\frac{1}{2}(u^*\sigma^2x^2\frac{\partial^2\Phi}{\partial x^2}\}$ *has to be finite. If the second order derivative of the optimal value function is positive, it is possible to put* u^* *either "equal" to* $+\infty$ *or* $-\infty$ *depending on the sign of the first derivative of the value function and no supremum would exist.*

b) Differentiate the g(u) with respect to u, putting the derivative equal to zero and solving proves the claim regarding the shape of u^*. *Plugging* u^* *into the HJB proves the second claim.*

c) Now assume that $\frac{\partial^2\Phi}{\partial x^2}=0$ *The HJB-equation takes the form* $\sup_u\{L^u\Phi+e^{-\rho s}f(x)\}=0=$

$=\frac{\partial\Phi}{\partial s}+e^{-\rho s}f(x)+\sup_u rxu\frac{\partial\Phi}{\partial x}$. *Clearly, the supremum will only exist if* $\frac{\partial\Phi}{\partial x}=0$. *Hence,*

$$\frac{\partial\Phi}{\partial s}+e^{-\rho s}f(x)=0$$

d) Start from $2\alpha^2(e^{-\rho t}f+\frac{\partial\Phi}{\partial t})\frac{\partial^2\Phi}{\partial x^2}-r^2(\frac{\partial\Phi}{\partial x})^2=0$. *We have proved that*

$\Phi(t,x)=\Phi(0,x)e^{-\rho t}=\xi(x)e^{-\rho t}$. *Now* $\frac{\partial\Phi}{\partial t}=-\rho e^{-\rho t},\frac{\partial\Phi}{\partial x}=e^{-\rho t}\xi_x(x),$ *and* $\Phi_{xx}=e^{-\rho t}\xi_{xx}(x)$.

Substitutions prove the claim.

4.4 Let X(t) denote wealth at time t. Suppose you at any time t, you have a choice between two investments. A risky investment where the unit price $p_1(t) = p_1(t,\omega)$ *satisfies the equation* $dp_1 = \alpha_1 p_1 dt + \sigma_1 p_1 dB(t)$ *A safer less risky investment where the unit price is* $p_2(t)$ *and satisfies*

$dp_2 = \alpha_2 p_2 dt + \sigma_2 p_2 d\tilde{B}$ *, where* α_i, σ_i *are positive constants such that* $\alpha_1 > \alpha_2$ *and* $\sigma_1 > \sigma_2$*, and* $B(t), \tilde{B}(t)$ *are independent one dimensional Brownian motion.*

a) Let $u(t,\omega)$ *denote the fraction of wealth which is placed in the riskier investment at time t. Show that*

$$dX(t) = X(t)(\alpha_1 u + \alpha_2(1-u))dt + X(t)(\sigma_1 u dB(t) + \sigma_2(1-u)d\tilde{B}(t))$$

b) Assuming that u is a Markov control u=u(t,x(t)), find the generator A^u *of* $(t, x^u(t))$

c) Write down the HJB equation for the control problem $\Phi(t,x) = \sup_u E^{tx}\{x_T^u)^\gamma$ *, where* $T = \min(t_1, \tau_0)$*,* $\tau_0 = \inf[t > s; x(t) = 0]$*,* t_1 *is a given future time (constant), and* $\gamma \in (0,1)$.

d) Find the optimal control for the problem in c)

Solution: Since the price processes are Geometric Brownian motion we know that p_i *(t)>0 for all t. Hence we are allowed to write:* $dp_i / p_i = \alpha_i dt + \sigma_i dB(t)$ *all i. For a given number of assets the change in the value of the portfolio is*

$$dX(t) = q_1 dp_1 + q_2 dp_2 = uX(t)\frac{dp_1}{p_1} + (1-u)X(t)\frac{dp_2}{p_2}.$$

Substitutions now give the answer. Note that dX(t) is Geometric Brownian motion so X(t)>0.

b) Let y(t)=f(t,x(t)). Taylor expansion gives $dy(t) = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x}dx + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}dx^2 =$

$= A^u f dt + L_1 dB(t) + L_2 d\tilde{B}(t)$ *The second equality is a result of substituting the equation for dx. Here* $A^u f = \frac{\partial f}{\partial t} + (u\alpha_1 + (1-u)\alpha_2)x(t)\frac{\partial f}{\partial x} + [u^2\sigma^2 x^2 + (1-u)^2\sigma^2 x^2]\frac{1}{2}\frac{\partial^2 f}{\partial x^2}$.

c) Let G={(s,x);0<t <T, 0<t< ∞ *}, where T is the first exit time from G. The HJB now becomes*

$\sup_u\{A^u\Phi(s,x)] = 0$ *for* $(s,x) \in G$*. Since x(t) >0 all t the exit time is* t_1.

$\Phi(s,x) = x^{\gamma}$ *for* $(s,x) \in \partial G$ *where*

$$A^{u}\Phi = \frac{\partial \Phi}{\partial s} + (u\alpha_1 + (1-u)\alpha_2)x(s)\frac{\partial \Phi}{\partial x} + (u^2\sigma^2 + (1-u)^2\sigma^2]\frac{x^2}{2}\frac{\partial^2 \Phi}{\partial x^2}.$$

d) Define h(u)= $A^{u}\Phi$ *for given (s,x). Putting* $h^{'}(u) = 0$ *and solving for u yields*

$u^{*} = \frac{1}{\sigma_1 + \sigma_2}[\frac{(\alpha_1 - \alpha_2)\Phi_x}{x\Phi_{xx}} + \sigma_2^2]$ *After inserting* u^{*} *into the HJB equation we are supposed to solve for* Φ *.in the interior of G. A guess based on "experience" and the value of the objective function on* ∂G *(the border of G), we put* $\Phi(s,x) = e^{\lambda(t-s)}x^{\gamma}$ *(x=0 is uninteresting). Here* λ *is a constant to be determined. Now* $\Phi_x = \gamma \Phi x^{-1}$ *and* $\Phi_{xx}\gamma(\gamma-1)\Phi x^{-2}$. *By inserting this into* u^{*}*, we obtain* $u^{*} = \frac{1}{\sigma_1 + \sigma_2}[\frac{(\alpha_1 - \alpha_2)\gamma}{\gamma - 1} + \sigma_2^2]$ *, implying that the optimal budget share is constant. To determine* λ *we again use the HJB*

$A^{u}\Phi = -\lambda\Phi + (u^{*}\alpha_1 + (1-u^{*})\alpha_2)\gamma\Phi + [(u^{*})^2\sigma^2 + (1-u^{*})^2\sigma^2]\frac{1}{2}\gamma(\gamma-1)\Phi = 0$. *Solve for* λ *and the problem is solved.*

4.6 Dynkin's formula[30]

The exit time from the open set *G* in problem 4.4 above is a *stopping time* (to be defined), and for stopping times there exist (under an Ito diffusion) an elegant result that can be used to calculate mathematical expectation at the exit time for twice continuously differentiable functions with compact support on R^n. It is called Dynkin's formula and it can, among other things, be used to solve problems like 4.4 above.

Ito diffusion

To this end, we introduce the time homogenous diffusion process

$$\begin{aligned} \mathbf{dX} &= \mathbf{a(X)}dt + \boldsymbol{\sigma}\mathbf{(X)}d\mathbf{B(t)} \\ \mathbf{X(0)} &= \mathbf{x_0} \end{aligned} \qquad \textit{(Ito diffusion)}$$

where $\mathbf{B(t)}$ is *m*-dimensional Brownian motion $\mathbf{a(X)}: R^n \to R^n$, and $\boldsymbol{\sigma}\mathbf{(X)}: R^n \to R^{nxm}$.

---

[30] The section is based on Öksendahl (2003).

An important property of the above diffusion process is that the future behavior of the process given what has happened up to time *t is the same as starting the process at* $\mathbf{x}_t$. *This is called the Markov property.* This is proved by Öksendal (2003) chapter 7.

Stopping time

A stopping time is a random time which is a function of an event ω, i.e., $\tau(\omega)$. More specifically we have:

Definition 4.2: Let $\{M_t\}$ *be an increasing family of* $\sigma-$*algebras (subsets of the event space* Ω*). A function* $\tau:\Omega\to[0,\infty)$ *is called a stopping time with respect to* $\{M_t\}$ *if*

$\{\omega;\tau(\omega)\le t\}\in M_t$ *for all* $t\ge 0$.

In plain words this means that it should be possible to decide whether or nor $\tau\le t$ has happened or not on the basis of the knowledge about M_t. Also, if $H\subset R^n$ is any set we define the first exit time from H, τ_H, as follows

$$\tau_H=\inf[t>0;\mathbf{x}(t)\notin H]$$

Subject to some technicalities this is also a stopping time according to *Definition 4.2.* We are now ready to introduce Dynkin's formula.

Theorem 4.2(Dynkins formula) *Let* $f(\mathbf{x(t)})\in C_0^2$, *and suppose* τ *is a stopping time,* $E[\tau]<\infty$. *Then*

$$E^x\{f(\mathbf{X(}\tau))\}=f(\mathbf{x}(0))+E^x\{\int_0^\tau Af(\mathbf{X}(s))ds\},$$

where $A=\sum_{i=1}^{2}\mathbf{a}\ (\mathbf{x})\frac{\partial}{\partial x_i}+\frac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{2}M_{ij}(\mathbf{x})\frac{\partial^2}{\partial x_i\partial x_j}$, and $\mathbf{M}=\boldsymbol{\sigma}\ (\mathbf{X})\boldsymbol{\sigma}\ (\mathbf{X})'$.

Note that the coefficients correspond to the Ito diffusion introduced above, and the function depend only on the vector $\mathbf{x(t)}$ and not *t, so* $\frac{\partial f}{\partial t}=0$. **.** The reason is that the underlying stochastic process is an Ito diffusion. C_0^2 means that f is twice continuously differentiable with compact support[31] on R^n.

Exercises (Öksendal (2003)):

4.5 *Solve the problem in 4.4 by using Dynkin's formula to obtain the optimal control.*

Solution: By noting that

$$E^{s,x}\{X^{\gamma}(t_1)\} = x_s^{\gamma} + \int_s^{t_1} \{[u\alpha_1 + (1-u)\alpha_2]X\gamma X^{\gamma-1} + [u^2\sigma_1^2 + (1-u)^2\sigma_2^2]X^2\gamma(\gamma-1)X^{\gamma-2}\frac{1}{2}\}dt$$

Clearly, since the left hand side will be a maximum (supremum) if the integrand is at its maximum we choose u to accomplish this. That is, we maximize

$$h(u) = [u\alpha_1 + (1-u]\alpha_2)]\gamma x^{\gamma} + \{[u^2\sigma_1^2 + (1-u)^2\sigma_2^2]\gamma(\gamma-1)x^{\gamma}\}\frac{1}{2}\}$$

with respect to u. This yields the same solution as in example 4.4 above. To find the optimal value function we will have to proceed as in 4.4.

4.6 Solve the portfolio problem $\max E\{\ln x(t_1)\}$ *subject to the same stochastic differential equation as in the two previous problems. Use Dynkin's formula to obtain the optimal control, and use the corresponding trick as in the previous problems to solve for the optimal value function.*

Solution: Note that with $\ln x = \phi(x)$, $\phi_x = 1/x$ *and* $\phi_{xx} = -1/x^2 = -\phi_x / x$, *resulting in a similar analysis as in 4.5.*

4.7 Consider the following stochastic control problem:

$$\Psi(s,x) = \inf_u E^{s,x}\{\int_s^{\infty} e^{-\rho t} X(t)^2 dt$$

where $dX(t) = \alpha u(t)dt + dB(t)$. *Here* α, ρ *are given positive constants and the control is restricted to the closed set U=* $[-1,1]$

a) Show that the HJB equation is

[31] Compact support means that the function is defined on a compact set in R^n and vanishes outside that set..

$$\inf_u \{e^{-\rho s}x^2 + \frac{\partial\Psi}{\partial s} + \alpha u\frac{\partial\Psi}{\partial x} + \frac{1}{2}\frac{\partial^2\Psi}{\partial x^2}\} = 0$$

b) If $\Psi \in C^2$ *and* u^* *exists, show that*

$u^*(x) = -sign\,\alpha x$, *where sign z=* $\begin{pmatrix} 1 if\ z > 0 \\ -1\ if\ z \le 0 \end{pmatrix}$

Hint: Explain why $x(t)$ *>0 implies that* $\frac{\partial\Psi}{\partial x} > 0$ *and x(t)<0 implies that* $\frac{\partial\Psi}{\partial x} < 0$.

Solution:

a) Should be trivial. If not, restart from the beginning of chapter 4

b) From the HJB-equation it is clear that there is no interior solution for the control. From the objective function it is clear that the absolute value of x(t) should be kept as small as possible. If x(t)>0, this can only be done if the control is put equal to minus one, and this implies that $\frac{\partial\Psi}{\partial x}$ *has to be positive to create an inf for given s and given x. A similar reasoning holds for x(t)<0.*

4.8 Find the optimal control function for the problem:

$$\max_{c(t)} E\{\int_0^\infty e^{-rt}c^a(t)dt\}$$

subject to

$$dX(t) = (bx(t) - c(t))dt + hX(t)dB(t)$$
$$X(0) = x_0$$

Solution: In this time autonomous problem the HJB equation can be written as in section 4.2 above

$$r\Psi(x) = \max_c\{c^a + [bx - c]\Psi_x(x) + \frac{h^2x^2}{2}\Psi_{xx}(x)\}$$

The maximization with respect to c yields

$$c = (\Psi_x a^{-1})^{(a-1)^{-1}}$$

Now we try the following form of the value function

$$\Psi(x) = Ax^a \text{ yielding } c = A^{\frac{1}{a-1}} x$$

We then use the suggested form of the value function and the HJB equation to solve for A to obtain

$$c(t) = [(r - ba)/(1 - a) + a\frac{h^2}{2}]x$$

4.9 Assume that total wealth at time zero is W(0)=W_0 *. It can be allocated over time to current consumption c(t), a risky asset with expected return a, and a safe asset with return r (a>r). The variance of the risky asset is* σ^2 *. The change in wealth is given by the stochastic differential equation*

$$dW(t) = [r(1 - w)W + awW - c]dt + wW\sigma dB(t)$$

W(0)=W_0,

where w is the share of wealth that is allocated to the risky asset

The objective function of the consumer is

$$\Psi(W_0) = \max_{c,w} \int_o^\infty e^{-\theta t} c^b(t) b^{-1} dt$$

subject to the stochastic differential equation for wealth. Find the optimal controls !

Solution: Note that we have two control variables, and one state variable. The problem is also autonomous so we can use the current value version of the the HJB –equation. This means that we have

$$\theta\Psi(W) = \max_{c,w}\{c^b b^{-1} + [r(1 - w)W + awW - c]\Psi_W(W) + \frac{w^2 W^2 \sigma^2}{2}\Psi_{WW}(W)$$

Calculus gives the maximizing values of c and w in terms of the parameters of the problem, the state W and the unknown function Ψ .

One has $c(t) = \Psi_x (W(t))^{\frac{1}{b-1}}$ *w(t)=* $\Psi_W (W)(r-a)/\sigma^2 W \Psi_{WW}(W)$.

Now put $\Psi(W) = AW^b$

Solve for A along the format used in the previous exercise to get

$$Ab = \{[\theta - rb - (r-a)^2 b / 2\sigma^2 (1-b)]/1-b)\}^{b-1}$$

$$c = W(Ab)^{1/(b-1)},\ \ w = (a-s)/(1-b)\sigma^2$$

Note that like in problem 4.4 the budget share w is constant over time.

4.10 control problem Consider the stochastic

$dX(t) = au(t)dt + u(t)dB(t)\ \ X_0 = x > 0$

dt=1dt

where $B(t) \in R, u(t) \in R$ *and* $a \in R$ *is a given constant and* $\Phi(s,x) = \sup_u E^{s,x}[(X(\tau)^\gamma]$ *and* *0<* γ *<1 is a constant and* $\hat{\tau} = \inf\{t > 0, X(t) = 0\} \Lambda (T-s)$. *T being a given future time. Show that this problem has the optimal control*

$u^*(t,x) = ax/(1-\gamma)$ *with corresponding optimal performance*

$$\Phi(s,x) = x^\gamma \exp[\frac{a^2(t_1 - s)\gamma}{2(1-\gamma)}]$$

G{(s,x);x>0 and s<T} Moreover, $\Phi(s,x) = \Phi(y) = \sup_u E^y[g(Y_{\tau G})]$

where $g(y) = g(s,x) = x^\gamma$ *and* $\tau_G = \inf[t > 0; Y(t) \notin G] = \hat{\tau}$

Here we apply Theorem 11.2.2 and look for a function ϕ *such that*

$\sup_{v \in H}\{f^y(y) + (L^y \phi(y)\} = 0$ *for* $y \in G$, *where in the case* $f^\upsilon(y) = 0$.

$$L^\upsilon \phi(y) = L^\upsilon \phi(s,x) = \frac{\partial \phi}{\partial s} + a\upsilon \frac{\partial \phi}{\partial x} + \frac{1}{2}\upsilon^2 \frac{\partial^2 \phi}{\partial x^2}$$

If we guess that $\frac{\partial^2\phi}{\partial x^2}$ *<0 then the maximum of the* $\upsilon \to L^{\upsilon}\phi(s,x)$ *is attained at*

$$\upsilon = \upsilon^*(s,x) = \frac{-a\frac{\partial\phi}{\partial x}}{\frac{\partial^2\phi}{\partial x^2}}$$

Trying a function ϕ *of the form* $\phi(s,x) = f(s)x^{\gamma}$ *and the attainuse the f. Substituted this gives*

$$\upsilon^*(s,x) = \frac{ax}{1-\gamma}$$

And $f^{'}(s)x^{\gamma} + \frac{a^2x}{1-\gamma} f(s)\gamma x^{\gamma-1} + \frac{1}{2}(\frac{ax}{1-\gamma})^2 f(s)\gamma(\gamma-1)x^{\gamma-2} = 0$ *or* $f^{'}(s) + \frac{a^2\gamma}{2(1-\gamma)} f(s) = 0$

The terminal codition is $\phi(y) = g(y)$ *for* ∂G *and* $f(T) = 1$. *Moreover*

$$f(s) = exp[\frac{a^2\ \gamma}{2(1-\gamma)}](T-s) \qquad s \le T\ .$$

Finally $\Phi(s,x) = \phi(s,x) = f(s)x^{\gamma}$ *and* $u^*(s,x) = \upsilon^*(s,x) = \frac{ax}{1-\gamma}$ *in a Markov control.*

Appendix to Chapter 4

A. The following is an "engineer's" derivation of the key (the HJB) equation in Theorem 4.1

$$V\big(k(t),t\big) = \sup_{\mathrm{c}} E_t\{\int_t^{\infty} u[(c(\tau)]e^{-\theta\tau}d\tau\}$$

$$= \sup_c E_t\{\int_t^{t+\Delta t} u(c(\tau))e^{-\theta\tau}d\tau + \sup_{\mathrm{c}} E_{t+\Delta t}\{\int_{t+\Delta t}^{\infty} u(c(\tau))e^{-\theta\tau}d\tau\}$$

$$= \sup_{\mathrm{c}} E_t\left[\int_t^{t+\Delta t} u(c(\tau))e^{-\theta\tau}d\tau + V\big(k(t+\Delta t),t+\Delta t\big)\right] \tag{A.1}$$

$$= \sup_{\mathrm{c}} E_t\big[u\big(c(t)\big)e^{-\theta t}\Delta t + V\big(k(t),t\big) +$$

$$V_k\big(k(t),t\big)\Delta k + V_t\big(k(t),t\big)\Delta t + \frac{1}{2}V_{kk}\big(k(t),t\big)(\Delta k)^2$$

$$+ V_{kt}\big(k(t),t\big)\Delta k\Delta t + \frac{1}{2}V_{tt}\big(k(t),t\big)\Delta t^2 + O(\Delta t)\big]$$

The first equality follows by definition. The second equality in (A.1) follows from Bellman's principle of optimality -"every part of the optimal path must be optimal". The third equality is a consequence of the definition of a value function, while the fourth equality follows from the Taylor expansion of the value function, which implies assuming that $V(\cdot)$ has continuous partial derivatives of all orders less than three. If the stochastic differential equation for capital is approximated by

$$\begin{aligned}\Delta k &= \left[f(k) - (n - \sigma^2)k - c \right]\Delta t - \sigma k \Delta z + O(\Delta t) \\ &= h(k,c;\sigma^2,n)\Delta t - \sigma k \Delta z + O(\Delta t)\end{aligned} \tag{A.2}$$

we can substitute for Δk in (A.1), and use the multiplication rules for Ito-calculus - in particular $(\Delta z)^2 = \Delta t$ - to obtain the first order differential as

$$\begin{aligned}&\sup_{c} E_t \{ \left[u(c)e^{-\theta t} + V_k h + V_t + \frac{1}{2} V_{kk} \sigma^2 k^2 \right] \Delta t \\ &- V_k \sigma \Delta z + O(\Delta t) \} = 0\end{aligned} \tag{A.3}$$

Note that the value function at time t appears on both sides of equation (A.1), so netting out creates the zero in the RHS of equation (A.3). Passing through the expectation parameter, dividing both sides by Δt, and taking the limit as $\Delta t \rightarrow 0$ we obtain

$$0 = \sup_{c} \left[u(c)e^{-\theta t} + V_t + V_k h + \frac{1}{2} V_{kk} \sigma^2 k^2 \right] \tag{A.4}$$

This equation is known as the Hamilton-Jacobi-Bellman equation of stochastic control theory, and it is typically written as

$$-V_t = \sup_c \left[u(c)e^{-\theta t} + V_K h + \frac{1}{2} V_{kk} \sigma^2 k^2 \right] \qquad \text{(A.4')+}$$

Turning to the shape of the stochastic co-state variables in section 4.4, recall that we have written the vector of co-state variables in present value terms as follows:

$$p = (\tilde{p}_{k,} \tilde{p}_x) = (V_k, V_x) \qquad \text{(A.5)}$$

Using Ito's formula on $\tilde{p}_k$ we obtain

$$\begin{aligned} \tilde{p}_k &= \{V_{kt} + V_{kk} h + V_{kx} e + \frac{1}{2}[V_{kkk}\sigma_1^2 + V_{xxx}\sigma_2^2\gamma^2 x^2\} dt \\ &-V_{kk}\sigma_1 k dw_1 - V_{kx}\sigma_2 \gamma x dw_2 \end{aligned} \qquad \text{(A.6)}$$

Since $V_{kt} = V_{tk}$ it follows from the HJB equation that

$$-V_{tk} = H_k^* + V_{kk} h + V_{kx} e + \frac{1}{2}[V_{kkk}\sigma_1^2 k^2 + V_{xxx}\sigma_2^2 \gamma^2 x^2] \qquad \text{(A.7)}$$

which inserted into the equation for the co-state, yields the first equation in (4.33). The co-state equation for the state variable $x(t)$ follows analogously.

B. It is the important to note that the first form is a necessary condition for the HJB. It says that if V is an optimal value function, and if the $\hat{u}$ is the control satisfies the HJB, and $\hat{u}(t,x)$ realizes the supremum in the equation.

However, there is a *verification theorem* that the HJB-equation also acts as a sufficient condition for the optimal control problem. The theorem is in fact esasy to prove. It has been used repeatedly. It looks like this:

Theorem (*Verification theorem) If we have two functions H(t,x) and g(t,x), such that*

- *H is sufficiently integrable, and solves the HJB equation*

$$\frac{\partial H(t,x)}{\partial t} + \sup_{u \in U}\{F(t,x,u) + A^u H(t,x)\} = 0$$
$$\forall (t,x) \in (0,T) x R^n$$

- *The function g is än admissible control law.*
- *For each fixed (t,x) the supremum in the expression*

$$\sup_{u \in U}\{F(t,x,u) + A^u H(t,x)\}$$

is attained by the choice u=g(t,x).

Then the following hold:

1. *The optimal value function V to the control problem is given*

V(t,x)=H(t,x).

2. *There exist an optimal control law* $\hat{u}$, *and in fact* $\hat{u}(t,x) = g(t,x)$. *(T Björk 2009 for both Theorem and Proof)*

Proof: *Assume that* $H(t,x)$ *and g(t.x) are given as above.*

H is intergrable, and solves the HJB equation

$$\frac{\partial H}{\partial t}(t,x) + \sup_{u \in U}\{F(t,x,u) + A^u H(t,x)\} = 0$$
and $\forall (t,x) \in (0,T) x R^n$

Now choose an arbitrary a control law $u \in U$ *and fix a point (t,x). We define* x^u *on the time interval [t, T) at the solution to the equation*

$$dx_s^u = \mu^u(s,x)ds + \sigma(s,x_s^u)dW_s$$
$$x_t = x$$

Inserting the process X^u *into the function H and using the Ito formula we obtain*

$$H(T,x_T^u) = H(t,x) + \int_t^T \{\frac{\partial H}{\partial t}(s,X_s^u) + (A^u H(s,X_s^u)\}ds + \int_t^T \nabla_x H(s,X_s^u)\sigma^u(s,X_s^u)dW_s$$

Since H the HJB equation we see that

$$\frac{\partial H(t,x)}{\partial x} + F(t,x,u) + A^u H(t,x) \le 0$$

For all $u \in U$, *and thus we have, for each s and P-a.s.,the inequality*

$$\frac{\partial H(t,x^u)}{\partial t} + \mathrm{A}^u H(s,X_s^u) \le -F(s,X_s^u)$$

From the boundary condition for the HJB equation we also have $H(T, X_s^u) = \Phi(X_T^u)$ *so we obtain the inequality*

$$H(t,x) \geq E_{t,x}\int_t^T F^u(s, X_s^u)ds + \Phi(X_T^u) - \int_t^T \nabla_x H(s, X_s^T)\sigma_s^u(s, X_s^u)dW_s$$

Taking expectation, and assuming enough integrality, we make the stochastic integral vanish, leaving us with the inequality

$$H(t,x) \geq E_{tx}[\int_t^T F^u(s, X_s^u)ds + \Phi(X_T^u)] = J(t,x,u)$$

Since the control law u *was arbitrarily*

$$H(t,x) \geq \sup_{u \in U} J(t,x,u) = V(t,x)$$

To obtain the reverse inequality we choose the specific law $u(t,x) = g(t,u)$. *Going through, the same calculations as above, and using the fact that by assumption we have:*

$$\frac{\partial H}{\partial t}(t,x) + F^g(t,x) + \mathrm{A}^g H(t,x) = 0$$

we obtain the equality

$$H(t,x) \geq E_{tx}[\int_t^T F^g(s, X_s^u)ds + \Phi(X_T^g)] = J(t,x,g)$$

On the other hand we have the "trivial" inequality

$$V(t,x) \geq J(t,x,g)$$

So using $H(t,x) \geq V(t,x) \geq J(t,x,g) = H(t,x,g)$

This shows that,

$$H(t,x) = V(t,x) = J(t,x,g),$$

which shows that $H(t,x) = V(t,x)$ *and g is the optimal control laws.*

Chapter 5: Optimal stopping[32]

In connection with the introduction of Dynkin's Lemma in chapter 4 we touched upon the optimal stopping problem. In this chapter we will try to convey some of the key properties of the optimal stopping problem. We start by defining the optimal stopping problem in the simplest case.

Problem 5.1: (The optimal stopping problem). Let $\mathbf{X}(t)$ *be an Ito diffusion on* R^n *and let* $g(\bullet)$ *, the objective function, be a given function on* R^n *, satisfying*

(i) $g(\mathbf{y}) \geq 0$ *for* $\mathbf{y} \in R^n$

(ii) g *is continuous.*

Find a stopping time $\tau^* = \tau^*(\mathbf{x}, \omega)$ *for* $\{\mathbf{X}_t\}$ *such that*

$$E^x[g(\mathbf{X}_{\tau^*})] = \sup_{\tau} E^x[g(\mathbf{X}_\tau)] \text{ for all } \mathbf{x} \in R^n \qquad (5.1)$$

The supremum is taken over all stopping times τ *for* $\{\mathbf{X_t}\}$

Like in chapter 4 an Ito diffusion is a stochastic differential equation

$$\begin{aligned} &\mathbf{dX(t) = b(X(}t\mathbf{))dt + \boldsymbol{\sigma}(X(}t\mathbf{))dB(t)} \\ &\mathbf{X(0) = x_0} \end{aligned} \qquad (5.2)$$

The coefficients are not functions of time, and the diffusion is called time homogenous. The expectation E^x is taken with respect to the probability law P^x for the process $\{\mathbf{X_t}\}$ starting at $\mathbf{x}$. We may regard $\mathbf{X(t)}$ as the state of an experiment ω at time t. For each t we have the option of stopping the experiment obtaining the reward $g(\mathbf{X}_t)$, or continuing the process hoping for a higher reward by stopping later. Technically the stopping times we are looking for are stopping times in the sense of *Definition 4.2*. The decision whether $\tau \leq t$ or not, should only depend on the Brownian motion process up to t. Loosely speaking, equation (5.1) means that among all stopping times we are looking for the one that is the best in the long run, i.e., if the experiment is repeated over and over again. The optimal value function is defined as:

---

[32] The chapter relies heavily on Öksendal (2003).

$$g^*(\mathbf{x}) = E^{\mathbf{x}}[g(\mathbf{X}_{\tau^*})]$$

Our objective is to outline how a solution to this problem can be found. To this end we introduce the following definition:

Definition 5.2: A measurable function $f : R^n \to [0,\infty)$ *is called supermeanvalued with respect to* $\mathbf{X}_t$ *if*

$$f(\mathbf{x}) \geq E^{\mathbf{x}}[f(\mathbf{X}_\tau)]$$

for all stopping times τ *and all* $\mathbf{x} \in R^n$ *. If* f *is (lower-semi) continuous*[33]*, then* f *is called superharmonic.*

Check that if $f \in C^2(R^n)$ it follows from Dynkin's formula in Theorem 4.2 that f is superharmonic if and only if $Af \leq 0$, where A is the characteristic operator of $\mathbf{X_t}$

Other key concepts are introduced by our next definition.

Definition 5.3: Let h *be a real measurable function on* R^n *. If* f *is a superharmonic (supermeanvalued) function and* $f \geq h$ *we say that* f *is a superharmonic (smv) majorant of* h *with respect to* X_t *. The function* $\bar{h}(\mathbf{x}) = \inf_f f(\mathbf{x}); \mathbf{x} \in R^n$ *, where the infimum is taken over all superharmonic majorants* f *of* h *, is called the least superharmonic majorant of* h *.*

---

[33] Semicontinuity is a mathematical property of real valued functions which is weaker than continuity. A real valued function f is (lower) semicontinuous at a point x_0 iff for every ε >0 there exist a neigbourhood U of x_0 such that $f(x) \geq f(x_0) - \varepsilon$ for all $x \in U$. It is upper semicontinuous if $f(x) \leq f(x_0) + \varepsilon$. Equivalently , this can be expressed as $\liminf_{x \to x_0} f(x) \geq f(x_0)$ and $\limsup_{x \to x_0} f(x) \leq f(x_0)$. The function f is lower(upper) semicontinuous if it is lower (upper) semicontinuous at every point of its domain. A continuous function is both upper and lower semicontinuous.

One can also define a least superharmonic majorant in the following manner.

Definition 5.2a: Suppoese there exists a function $\widehat{h}$ *such that*

(i) $\widehat{h}$ *is a superharmonic majorant of h.*

(ii) f *is any other superharmonic majorant of* h *, and* $\widehat{h} \leq f$ *.*

Then $\widehat{h}$ *is called the least superharmonic majorant of h.*

One can prove that if g is non-negative (or lower bounded) and (lower semi-) continuous then $\widehat{g}$ exists and $\widehat{g} = \overline{g}$.

We now relate the superharmonic majorant to the optimal objective function. Let $g \geq 0$ and let f be a superharmonic majorant of g .Then, if τ is a stopping time

$$f(\mathbf{x}) \geq E^{\mathbf{x}}[f(\mathbf{X}_\tau)] \geq E^{\mathbf{x}}[g(\mathbf{X}_\tau)]$$

The first equality follows since f is superharmonic, and the second since the expectation of the objective function and its majorant are evaluated at the same $\mathbf{X}_\tau$. This means that

$$f(\mathbf{x}) \geq \sup_\tau E^{\mathbf{x}}[g(\mathbf{X}_\tau)] = g^*(\mathbf{x}) \tag{5.3}$$

Therefore, we always have

$$\widehat{g}(\mathbf{x}) \geq g^*(\mathbf{x}) \text{ for all } \mathbf{x} \in R^n$$

i.e., the least superharmonic majorant of the optimal value function is at least as large as the optimal value function. Surprisingly, one can also show that the converse equality also holds, which implies

$$\widehat{g}(\mathbf{x}) = g^*(\mathbf{x}) \tag{5.4}$$

The existence theorem for optimal stopping tells us this. More precisely:

Theorem 5.1 *Let* g^* *denote the optimal value function and* $\hat{g}$ *the least superharmonic majorant of a value function,* g *, then* $\hat{g}(\mathbf{x}) = g^*(\mathbf{x})$. *Moreover, define*

$D = \{\mathbf{x}; g(\mathbf{x}) < g^*(\mathbf{x})\}$

the continuation set , i.e., the set of **x** *such that the experiment (search, game) is continued. Given a finite stopping time,* τ_D *, defined by Def.4.2, and modulo some integrability condition with respect to the probability measure, then*

$$g^*(\mathbf{x}) = E^x[g(\mathbf{X}_{\tau_D})]$$

and $\tau_D = \tau^*$ *is an optimal stopping time.*

The details are available in Öksendal 2003 ch. 10. The theorem, given that we know that it is enough that the value function is (lower semi-)continuous for a superharmonic majorant to exist, gives conditions for the existence of an optimal stopping rule. It also indicates how the rule can be found. Öksendal introduces a procedure to construct the least superharmonic majorant. Moreover, given that there exist an optimal stopping time, it can be shown to be unique.

To get a hold on whether the process is in the continuation region in a situation where the value function is twice continuously differentiable ($g \in C^2(R^n)$), we can use Dynkin's formula to prove that the set $U = \{\mathbf{x}; A\mathbf{x} > 0\}$ is a subset of D.

To prove this choose $\mathbf{x} \in U$ and let τ_0 be the first exit time from a bounded open set $W, \mathbf{x} \in W, and\, W \subset U$. From Dynkin's formula it now follows for $u > 0$ that

$$E^x[g(\mathbf{X}_{\tau_0 \wedge u}) = g(\mathbf{x}) + \int_0^{\tau_0 \wedge u} Ag(\mathbf{X}(s))ds > g(\mathbf{x}) \quad \tau_0 \wedge u = \min(\tau_0, u)$$

The result means that $g(\mathbf{x}) < g^*(\mathbf{x})$ and therefore $x \in D.$

Hence, given $A\mathbf{x} > 0$, we know that we are in the continuation set. Since *U* is a subset of *D*, there are (possibly) cases when $U \neq$`D and it is optimal to proceed beyond *U* before stopping.

The following definition and result are connected to the existence of an exit time, and since we have discussed superharmonic functions, it may be interesting to know the definition of a harmonic function.

Definition 5.3: Let f be a locally bounded, measurable function on D. Then f is called harmonic in D if

$f(\mathbf{x}) = E^x[f(\mathbf{X}_{\tau_U})]$

for all $\mathbf{x} \in D$ *and all bounded open sets U with* $\bar{U} \subset D$.

Here τ_U is the exit time from *U*. We are now ready for a lemma.

Lemma5.1:

a) Let $f \in C^2(D)$ *be harmonic in D, then* $Af = 0$ *in D*

b) Conversely, suppose that $f \in C^2(D)$ *and* $Af = 0$ *in D, then f is harmonic.*

The first result follows direcly from the formula for A, and the second follows by making use of Dynkin's formula. See Öksendal (2003) Chapter 9, section 9.2. Harmonic functions are important for the solution of what is called the *Stochastic Dirichlet problem*[34]. They are not indispensable for solving optimal stopping problems.

5.1 Find the stochastic solution f(t,x) of the boundary problem

$$K(x)e^{-\rho t} + \frac{\partial f}{\partial t} + \alpha x \frac{\partial f}{\partial x} + \frac{1}{2}\beta^2 x^2 \frac{\partial^2 f}{\partial x^2} = 0 \qquad x > 0, \quad 0 < t < T$$

$$f(T, x) = e^{-\rho t}\phi(x)$$

Let $D = \{(t, x); 0 < t < T, x > 0\}$

$Y_t^{s,x} = (s + t, x_t^x)$, *where* $dx_t = \alpha x_t dt + \beta x_t dB_t \quad x_t = x$

Let

[34] To an economist interested in the History of Economic Theory, it may be interesting to know that the mathematician Johan Peter Dirichlet (born in Belgium) was a classmate and friend of the famous French economist Augustin Cournot when both of them studied mathematics at Sorbonne.

$$g(t,x) = e^{-\rho t} K(x)$$
$$h(t,x) = e^{-\rho t} \phi(x)$$

Let τ_D *be the exit-timefor the process* $Y_t^{s,x}$ *from D. The solution to the stochcastic Dirichlet problem which is only one simple-problem*

$$f(s,x) = E^{s,x}[h(Y_{\tau_D})] + E^{s,x}[\int_0^{\tau_D} g(Y_t)dt]$$

We know that $0 < x_t^x \ \forall t \geq 0 \ if \ x > 0$. *This means that* $\tau_D^{s,x} = T - s$

Hence, $f(s,x) = E^{s,x}[e^{-\rho T}\phi(x_{T-s})] + E^{s,x}[\int_0^{T-s} e^{-\rho t} K(x_t)dt]$

5.1 A generalized objective function

We have already noted that the existence theorem still holds if we relax the continuity assumption by introducing lower semi-continuity instead of continuity in the problem sketched in the introduction. It is also relatively straightforward to relax the non-negativity of the objective function g. If g is bounded from below, say $g \geq -L$, where $L > 0$ is a constant, we apply the theory to

$$g_1 = g + L \geq 0$$

and back out the true optimal value function by noting that $g^*(\mathbf{x}) = g_1^*(\mathbf{x}) - L$. There is also medicine for the case when g is not bounded from below, but the reader has to visit the "Pharmacy" in Öksendahl (2003), Chapter 10.

What we will deal with is, however, the case when the objective function is time inhomogeneous. Most dynamic problem in Economics involves discounting, and this introduces a time argument in the value function. A typical shape of the objective function could be

$$g = g(t,x) : RxR \rightarrow [0,\infty), \ g \ \textit{is continuous}$$

The optimization problem would then be to find $g_0(x) = \sup_\tau E^x[g(\tau, X_\tau)] = E^x[g(\tau^*, X_{\tau^*}]$. In order to transform this problem into the original time homogenous problem, we use the Ito diffusion in equation 5.2, and to save some notational clutter we assume that the process X_t is one dimensional. The multidimensional case is hopefully straightforward for the reader. If not, consult Öksendal Chapter 10. Let us now define the Ito diffusion $Y_t = Y_t^{s,x}$ in R^2 by

$$Y_t = \begin{pmatrix} s+t \\ X_t \end{pmatrix} : t \geq 0 \tag{5.5}$$

This means that

$$dY_t = \begin{pmatrix} 1 \\ b(X_t) \end{pmatrix} dt + \begin{pmatrix} 0 \\ \sigma(X_t) \end{pmatrix} dB_t \tag{5.6}$$

If $X_0 = x$ we can in the problem at hand put *s+t=0 and write*

$$g_0(x) = g(0,x) = \sup_\tau E^{(o,x)}\{g(Y_\tau) = E^{(0,x)} g(Y_\tau^*) \tag{5.7}$$

The independent time argument will have consequences for the characteristic operator for Y_t, which is now given by

$$\widehat{A}_X f(s,x) = \frac{\partial f(s,x)}{\partial s} + A_X f(s,x) \tag{5.8}$$

where $f \in C(R^2)$. Here A_X is the characteristic operator working on X_t.

Exercises (Öksendal (2003)):

5.2 Let $X_t = B_t$ be one dimensional Brownian motion and let the objective function be

$$g(s,x) = e^{-\alpha t + \beta B(t)}$$

Solve for the optimal stopping time τ *, and the optimal value function.*

Solution: Since the X_t *evolve according to* $dX(t) = dB(t)$ *the characteristic operator of the process becomes* $Ag = (-\alpha + \frac{1}{2}\beta^2)g$ *. Hence if* $2\alpha \le \beta^2$*, we have that* $Ag \le 0$ *for* $t \ge 0$ *and the process will remain outside the continuation region for all t. Hence, it is optimal to stop immediately and* $g = g^*$.

If $2\alpha > \beta^2$ *we have that the set* $U = \{(s,x); Ax > 0\} = R^2$*. Since* $U \subset D$ *we must have that the continuation region* $D = R^2$ *and the optimal stopping time does not exist.*

It is near at hand to guess that the reward for waiting for ever approaches infinity but this has to be proved.

The following theorem describes an algorithm to construct the least superharmonic majorant.

Theorem 5.2 *Let* $g = g_0$ *be a non-negative (lower semi)-continuous function on* R^n *and define inductively*

$$g_n(x) = \sup_{t \in S_n} E^x[g_{n-1}(X(t))]$$

where

$S_n = \{k2^{-n} \le k \le 4^n\}$ *n=1,2....*

Then $g_n \uparrow \hat{g}$ *and* $\hat{g}$ *is the least superharmonic majorant of g and* $\hat{g} = \bar{g}$

Remark: Note that S_n *expands as n increases. This is the reason why the supremum is non-decreasing with n.*

We will now return to our example and show that the reward for waiting approaches infinity when $2\alpha < \beta^2$.

5.2a Show that when $2\alpha < \beta^2$ *then* $g^* = \infty$

Solution: To construct the least superharmonic majorant we use Theorem 5.2 We start from

$$\sup_{t \in S_n} E^{s,x}[g(Y_t)] = \sup_{t \in S_n} E[e^{-\alpha(s+t)+\beta B(t)}]$$

To solve for the expectation we turn back to the methods in Chapter 3 and use Itos lemma on Z(t)= $e^{\beta B(t)}$ *. After integration and taking expectations one obtains*

$E[Z(t)] = E[Z(0)] + \frac{1}{2}\beta^2\int_0^t E[Z(s)]ds$ *. Differentiating with respect to t we obtain*

$\frac{dE[Z(t)]}{dt} = \frac{\beta^2}{2}E[Z(t)]$*, with* $E[Z(0))] = e^{\beta x}$ *. The solution can hence be written*

$E[Z(t)] = e^{\beta x + \frac{\beta^2 t}{2}}$ *. This means that we can write*

$$\sup_{t\in S_n} E^{s,x}[g(Y_t)] = \sup_{t\in S_n} E[e^{-\alpha(s+t)+\beta B(t)}] = \sup_{t\in S_n}[e^{-\alpha(s+t)}e^{\beta x + \frac{\beta^2 t}{2}}]$$
$$= \sup_{t\in S_n} g(s,x)e^{(-\alpha+\frac{\beta^2}{2})t} = g(s,x)e^{(-\alpha+\frac{\beta^2}{2})2^n}$$

Note that $2^n \in S_n$ is the lower bound of S_n, and therefore generates supremum (the least upper bound). Hence, for $n \to \infty$, $g_n \to \infty$.

5.2 The optimal stopping problem involving an integral

The objective function for many optimal stopping problems in Economics will naturally involve an integral. In this section we will illustrate how this problem can be handled. To keep the notation as simple as possible we start from the one dimensional version of the Ito diffusion in equation (5.2), i.e.

$$dX_t = b(X_t)dt + \sigma(X_t)dB(t) \quad X_0 = x \tag{5.9}$$

$$X_0 = x$$

The instantaneous objective function $f : R \to [0,\infty)$ is continuous with at most linear growth, and let $g : R \to [0,\infty)$ be a continuous function that measures the value that is obtained at the stopping time τ . The growth condition on the function f is there to keep the optimization problem bounded. The optimal stopping problem is formulated as: Find $\Phi(x)$ and τ^* such that

$$\Phi(x) = \sup_{\tau} E^x[\int_0^{\tau} f(X(t)dt + g(X(\tau))] = E^x[\int_0^{\tau^*} f(X(t)dt + g(X(\tau))] \tag{5.10}$$

To extend the Ito diffusion we use a trick similar to the one that was used in connection with the time inhomgenous problem above. Define an Ito diffusion by

$$dY_t = \begin{pmatrix} dX_t \\ dW_t \end{pmatrix} = \begin{pmatrix} b(X_t) \\ f(X_t) \end{pmatrix} dt + \begin{pmatrix} \sigma(X_t) \\ 0 \end{pmatrix} dB_t \qquad (5.11)$$
$$Y_0 = y = (x, \omega)$$

The optimization problem in (5.9) and (5.10) can now be rewritten in the following manner

$$\Phi(x) = \sup_{\tau} E^{(x,0)}[W_\tau + g(X_\tau)] = \sup_{\tau} E^{(x,0)}[\tilde{h}(Y_\tau)] \qquad (5.12)$$
$$\tilde{h}(y) = \tilde{h}(x,\omega) = g(x) + \omega,\ y \in R^2, \omega = f(x)$$

The connection between the characteristic operators A_X and A_Y is in the following

m $A_Y\phi(y) = A_Y\phi(x,\omega) = A_X\phi(x,\omega) + f(x)\partial\phi/\partial\omega,\ \phi \in C^2(R^2)$ (5.13)

The second term in the right hand side of 5.13 is analogous to the term $\frac{\partial f}{\partial t}$ in equation (5.8). The reason is that $f(X_t)$ is the drift coefficient of the process dW_t. Note that in equation (5.6) the corresponding coefficient equals one. In this case we get $\tilde{h}(y) = \tilde{h}(y,\omega) = g(y) + f(x)$, which means that we obtain

$$A_Y\tilde{h}(x,\omega) = A_X g(x) + f(x) \qquad (5.14)$$

since $\frac{\partial \tilde{h}}{\partial \omega} = 1$. Moreover, we have to modify the set U into

$$U = \{x; A_X g(x) + f(x) > 0\} \qquad (5.15)$$

Which belongs to the continuation region, $U \subset D$, from Dynkin’s formula.

The following exercise from Öksendahl shows how what we just have learnt can be used.

Exercises:

5.3 What can we say bout the solution to the following optimal stopping problem (i.e., solve it).

$$\Phi(x) = \sup_{\tau} E^{x}[\int_{0}^{\tau} \theta e^{-\rho t} X(t)dt + e^{-\rho\tau} X(\tau)]$$

$$dX(t) = \alpha X(t)dt + \beta X(t)dB(t)$$
$$X(0) = x$$

where $\theta, \rho > 0$

This problem does not only contain an integral, but is also time inhomogeneous. We start by dealing with time inhomogeneity by introducing the transformation

$$dY_t = \begin{pmatrix} 1 \\ \alpha X(t) \end{pmatrix} dt + \begin{pmatrix} 0 \\ \beta X(t) \end{pmatrix} dB_t \quad Y(0) = (s, x)$$

Next we take care of the integral by writing

$$dZ(t) = \begin{pmatrix} dY(t) \\ dW(t) \end{pmatrix} = \begin{pmatrix} 1 \\ \alpha X(t) \\ e^{-\rho t} X(t) \end{pmatrix} dt + \begin{pmatrix} 0 \\ \beta X(t) \\ 0 \end{pmatrix} dB(t) \quad Z(0) = (s, x, \omega)$$

A little reflection reveals that

$$\tilde{h}(s, x, \omega) = g(s, x) + \omega$$
$$g(s, x) = e^{-\rho s} x,\ \omega = \theta e^{-\rho s} x = f(s, x)$$

and

$$A_Z \tilde{h}(s, x, \omega) = \frac{\partial \tilde{h}}{\partial s} + \alpha x \frac{\partial \tilde{h}}{\partial x} + \frac{\beta^2 x^2}{2} \frac{\partial^2 \tilde{h}}{\partial x^2} + \theta e^{-\rho s} \frac{\partial \tilde{h}}{\partial \omega} = (-\rho + \alpha + \theta) e^{-\rho s} x$$

Hence $U = \{(s, x, \omega); A_Z \tilde{h}(s, x, \omega) > 0\} = \begin{pmatrix} R^3 \text{ if } \rho < \alpha + \theta \\ \varnothing, \text{ if } \rho \geq \alpha + \theta \end{pmatrix}$ $\quad\varnothing$ *is the empty set.*

This means $U = D = R^3$ *for* $\rho < \alpha + \theta$, *which implies that* τ^* *does not exist. The process will never leave the continuation region D. For* $\rho \geq \alpha + \theta$, $\tau^* = 0$. *What remains is to determine* *the optimal value function* $\Phi(s,x,\omega)$. *For* $\tau^*=0$ *it is clear that we get what we started from, i.e.,* $\Phi(s,x,\omega)=e^{-\rho s}x+\omega$. *When* τ^* *does not exist it is tempting to assume* $\Phi(s,x,\omega)\to\infty$. *However, this depends on whether the integral in (5.16) converges or not, when* $\tau\to\infty$. *From Chapter 3 , we know that* $E^x[X(t)]=xe^{\alpha t}$ *when the process is geometric Brownian motion like in the example. This means that when* $\alpha<\rho<\theta+\alpha$ *, the integral converges and* $\Phi(s,x,\omega)=\theta xe^{-\rho s}[\rho-\theta]^{-1}+\omega$. *If* $\rho\le\alpha$ *,then* $\Phi(s,x,\omega)\to\infty$. *(Intuition: if you interpret* ρ *as the discount rate, and x grows faster than the discount rate the integrand grows at rate* $\alpha-\rho>0$*).*

5.3 The Brekke-Öksendal Verification Theorem

If the objective function *g* in our original optimal stopping problem in equations (5.1)-(5.3) is twice continuously differentiable, then, under certain ("normal") conditions, the optimal value function g^* is continuously differentiable. This property is referred to as the high contact principle. Brekke and Öksendal (1991) have proved a sufficiency condition of high contact type for the optimal stopping problem, which makes it "easy" (possible!) to verify that a given candidate for $g^*(\Phi^*)$ is actually the solution. The idea is that the type of the continuation set as well as the optimal value function can be guessed parametrically and then verified to work by determining the parameters by using the continuity and differentiability conditions.

The Theorem is a little messy, but the use we can make from it is worth the pain. We start by fixing a domain $V\subset R^n$ and we introduce the Ito diffusion

$$d\mathbf{Y}(t)=b(\mathbf{Y}(t)dt+\sigma(\mathbf{Y}(t))d\mathbf{B}(t)$$
$$\mathbf{Y}(0)=\mathbf{y}$$

in R^n. We define $T=T(y,\omega)=\inf\{t>0;\mathbf{Y}_t\notin V\}$

In an economic application the set V can be interpreted as the solvency set (bankrupty occurs at *T).*

Let $f:R^n\to R, g:R^n\to R$

be a continuous functions. f satisfying

$$E^y[\int_0^T |f(\mathbf{Y}(t))| dt < \infty$$

for all $y \in R^n$. Moreover, there is an integrability condition on $g(Y(\tau))$ for all stopping times in the solvency set that this author does not quite understand. We can interpret f as the profit rate and g as a bequest function.

Now, consider the following problem: Find $\Phi(\mathbf{y})$ and $\tau \le T$ such that

$$\Phi(\mathbf{y}) = \sup_{\tau \le T} J^T(\mathbf{y}) = J^{\tau^*}(\mathbf{y})$$

where $J^\tau(\mathbf{y}) = E^y[\int_0^\tau f(\mathbf{Y}(t)dt] + g(\mathbf{Y}(\tau))]$ for all $\tau \le T$

We are now ready to formulate a "loose version" of Theorem 10.4.1 in Öksendal

Claim 5.1 *Suppose that we can find a function* $\phi : \bar{V} \to R$ *(from the closure of V to R)* such that

(i) $\phi \in C^1(V) \cap C(\bar{V})$

(ii) $\phi \ge g$ on V and $\phi = g$ on ∂V *(the boundary of V)*

Define D$= \{\mathbf{y} \in V; \phi(\mathbf{y}) > g(\mathbf{y})\}$

Suppose $\mathbf{Y_t}$ *spends zero time on* ∂D *for all starting points, and that* ∂D *is a Lipschitz surface, i.e. locally it is the graph of a function from* $R^{n-1} \to R$ *that fulfills a Lipschitz condition.*

Moreover, suppose the following holds

(iii) $\phi \in C^2(V / \partial D)$*, and the second order derivatives are bounded near the boundary of D.*

(iv) $L\phi + f \le 0$ *on the part of V that does not belong to D (*$\{V / D\}$*)*

(v) $L\phi + f = 0$ *on D*

(vi) $\tau_D = \inf\{t > 0; \Upsilon_t \notin D\} < \infty$ *almost surely for* $\mathbf{y} \in V$

(vi) the family $\{\phi(\mathbf{Y}_\tau); \tau \le \tau_D\}$ *fulfills an integrabilty condition on V.*

Then $\phi(\mathbf{y}) = \Phi(\mathbf{y}) = \sup_{\tau \le T} J^T(\mathbf{y}) = J^{\tau^*}(\mathbf{y})$ *, and* $\tau^* = \tau_D$ *is an optimal stopping time for this problem.*

The Claim is a sufficient condition for an optimal stopping problem, since if we can find a function ϕ that fulfills the listed conditions; we know that we have found the solution of the optimal stopping problem. Condition *(i)* tells us that the function has to be differentiable on the solvency set. Condition *(ii)* tells how the optimal value function is qualitatively related to the solvency set. *D* is the continuation region.

Condition *(iii)* requires that the optimal value function is twice continuously differentiable on the solvency set, minus the boundary of the continuation set. Conditions *(iv)* and *(v)* are restrictions on the objective function outside and inside the continuation region. L is the partial differential operator which coincides with A_Y for twice continuous differentiable functions with compact support. Compact support means that the function vanishes outside some bounded set. ϕ is defined on the closure of $\bar{V}$, and is twice continuously differentiable except on the boundary of the continuation set, where the process spends zero time. So the compact support assumption holds on $\bar{V}$ minus the boundary of *D*. So in "practice" $L = L_Y = A_Y$.

We will now use Claim 5.1 to solve some exercises borrowed from Öksendahl (2003).

Exercises (Öksendal2003and Nyström 2001):

5.4 Find the supremum g^* *and the optimal stopping time* τ^* *for the optimal stopping problems below (B(t) is 1-dimensional B-motion)*

a) $g^*(x) = \sup_\tau E^x[B^2(\tau)]$

b) $g^*(x) = \sup_\tau E^x[e^{-B^2(t)}]$

Solution:

a) the idea is to look for the least superharmonic majorant $\tilde{g}$ *of the function* x^2. *Brownian motion has the generator* $L = \frac{1}{2}\frac{\partial^2}{\partial x^2}$. *This means that the generator of* g *equals*

Lg=1 implying that U=R=D and the process stays in the continuaion region for ever, i.e. $\tau^* \to \infty$. *Moerover, the least superharmoninc majorant* $\tilde{g} \geq x^2$ *for all x. Hence* $g^* = \infty$.

b) Again, the operator of Brownian motion is $L = \frac{1}{2}\frac{\partial^2}{\partial x^2}$. This means that the operator of e^{-x^2} is $Lg = 2x^2e^{-x^2} \geq 0$. *Hence, the only way to leave the set U is x=0.In other words, the optimal stopping time is* $\tau^* = \inf_{\tau}[t > 0, B(t) = 0]$. *At* $B(t) = 0$ *a little thought reveals that the least superharmonic majorant is* $g^* = 1$. *One can in fact show that the only non-negative(* B_t *) superharmonic functions in R and* R^2 *are constants. See Öksendal problem 10.2.*

5.4 Find g^*, τ^* *such that*

$$g^*(s, x) = \sup_{\tau} E^{(s,x)}[e^{-\rho(s+t)}B_{\tau}^2] = E^{(s,x)}[e^{-\rho(s+t)}B_{\tau^*}^2]$$

where B_t *is one dimensional Brownian motion,* $\rho > 0$ *constant. The working strategy is to guess a continuation region. The shape of the continuation region can often be guessed from the shape of U.*

Solution: The first thing we do is to apply the partial differential operator on the objective function. The generator of the diffusion (s+t, B(t)) is

$L = \frac{\partial}{\partial t} + \frac{1}{2}\frac{\partial^2}{\partial x^2}$, *which results in*

$$Lg = -\rho e^{-\rho t}x^2 + e^{-\rho} = (1 - \rho x^2)e^{-\rho t}$$

From this we can conclude that $U = \{(t, x) : Lg > 0\} = \{(t, x); \frac{-1}{\sqrt{\rho}} < x < \frac{1}{\sqrt{\rho}}\}$. *Moreover, the continuation region is by definition* $D = \{(s, x); g^*(s, x) > g(s, x)\}$. *From theory we know that* $U \subset D$. *Hence, given the shape of U it is near at hand to assume that*

$D = \{(s,x); -x_0 < x < x_0\}$ *for some* $x_0 \geq \frac{1}{\sqrt{\rho}}$. *We are looking for a*

$\phi(t,x) = \begin{pmatrix} h(t,x) \;\; for\, x_{abs} < x_0 \\ h(t,x) = e^{-\rho t}x^2 \;\; for\;\; x_{abs} \geq x_0 \end{pmatrix}$ *, where h solves (according to (v) in Claim 5.1)*

$\left(\frac{\partial}{\partial t} + \frac{1}{2}\frac{\partial^2}{\partial x^2}\right)h = 0,\; x \in D$ *and* $h(t,x) = e^{-\rho t}x^2$ *for* $x_{abs} \geq x_0$ *. Now we have to guess the form of h(s,x), and why not choose h(s,x)=* $e^{-\rho x}A(x)$*, which gives*

$$A^{"}(x) - 2\rho A(x) = 0, x \in D$$
$$A(x) = x^2 \; x \notin D$$

Hence for $x \in D$ *we have A(x)=* $C_1 e^{\sqrt{2\rho}x} + C_2 e^{-\sqrt{2\rho}x}$*. The two boundary conditions for D gives*

$x_0^2 = C_1 e^r + C_2 e^{-r}$ *and* $x_0^2 = C_1 e^{-r} + C_2 e^r$ *implying that* $C_1 = C_2 = \frac{x_0^2}{e^r + e^{-r}} = \nu(x_0)$*. Moreover,*

A(x)= $\nu(x_0)[e^{\sqrt{2\rho}x} + e^{-\sqrt{2\rho}x}]$*,* $\phi(t,x) = e^{-\rho t}\nu(x_0)[C_1 e^{\sqrt{2\rho}x} + C_2 e^{-\sqrt{2\rho}x}],\; x \in D$*, and*

$\phi(t,x) = e^{-\rho t}x^2, x \notin D$

From the construction $\phi(t,x)$ *is globally continuous by*

$$\frac{\partial \phi}{\partial x} = e^{-\rho t}\nu(x_0)\sqrt{2\rho}[e^{-\sqrt{2\rho}x} + e^{\sqrt{2\rho}x}] \quad x_{abs} < x_0$$
$$\frac{\partial \phi}{\partial x} = 2xe^{-\rho t} \quad x_{abs} \geq x_0$$

Now it remains to determine x_0*, and to do that we use contditions (i) in Claim 5.1., which means that* $\phi_x(t,x)$ *has to be globally continuous on* $\bar{V}$*. The problem is the border of the continuation region, i.e., at* $x_0(-x_0)$*. From the symmetry of D, we check at the upper boundary where* $x = x_0$*.Continuity requires that*

$\sqrt{2\rho}x_0 \dfrac{x_0(e^r - e^{-r})}{e^r + e^{-r}} = 2x_0 \;\; r = \sqrt{2\rho}x_0$. *From "forgotten" knowledge about trigonometry this can be written*[35] $r \bullet tgh(r) = \sqrt{2\rho}x_0 \tan h(r) = 2 \;\; *$

This equation cannot be solved explicitly, but by putting

F(r) = $r \bullet tgh(r)$ *and differentiating with respect to r, we find that for* $r \geq \frac{1}{2}$ *and* $F(\frac{1}{2}) < 2$, *implying that the solution to* $*, r^*$, *is unique and hence* $x_0 = \dfrac{r^*}{\sqrt{2\rho}}$. *Finally,* $\tau^* = \tau_D$.

[If $\dfrac{\partial\phi}{\partial x}$ *has global continuity and* $r = \sqrt{2\rho}x_0$ *and.* $\dfrac{x_0^2}{e^r + e^{-r}}\sqrt{2}\sqrt{\rho}[e^r - e^{-r}] = 2x_0$. *The tanh(x)* $= \dfrac{e^x - e^{-x}}{e^x + e^{-x}}$ *and we can take f(t)=rtanh(r)=2 and*

$f_r(r) = \tan h(r) + r[(e^r + e^{-r})^2 - (e^r - e^{-r})^2](e^r + e^{-r})^{-1}$ *If* $\dfrac{\partial\phi}{\partial x}$ *has global continuity and* $r = \sqrt{2\rho}x_0$ *and.* $\dfrac{x_0^2}{e^r + e^{-r}}\sqrt{2}\sqrt{\rho}[e^r - e^{-r}] = 2x_0$. *The tanh(x) =* $\dfrac{e^x - e^{-x}}{e^x + e^{-x}}$ *and we can take f(t)=rtanh(r)=2 and* $f_r(r) = \tan h(r) + r[(e^r + e^{-r})^2 - (e^r - e^{-r})^2](e^r + e^{-r})^{-1}$ *We note that* $r \geq \frac{1}{2}$, $f\,(\frac{1}{2}) < 2$ *and* $f_r(r) > 0$ *and.we have shown that the solution* r_0*.and* $x_0 = r_0 / \sqrt{2p}$ *].*

The steps we followed in the solution were:

(i) Apply the generator L on the objective function and try to determine U

(ii) Use U as a possible model for D, expressed in terms of parameters to be determined by using the differentiability and continuity conditions in Claim 5.1. Find a candidate for an optimal value function inside and outside D. A tentative model inside D is often a general version of the function outside *D.*

---

[35] tanh=tangenshyperbolicus.

(iii) Apply the variational inequalities on the proposed optimal value function to find a differential equation on D.

(iv) Solve it by using the restriction that it has to be globally continuous. This helps to determine the constants that will be functions of the parameters characterizing D.

(v) Determine the parameter(s) of *D* such that the derivative of the objective function is globally continuous. The main problem is to glue it together at the boundary of *D.*

(vi) Now the objective function will fulfill all conditions in Claim 5.1 and the region D is determined explicitly or implicitly by the parameters, or by the implicitly determined parameters.

(vii) The first exit time from D will be the optimal stopping time.

Now, a slightly more complicated problem

5.5 Solve the optimal stopping problem $\gamma^*(x) = \sup_\tau E^{(0,x)}[\int_0^\tau e^{-\rho t} B^2(t)dt + e^{-\rho t} B^2(\tau)]$

where $Y_t^{(sx)} = \{s+t, B_t^x\}, t \geq 0, g(t,x) = f(t,x) = e^{-\rho t} x^2$ *We also note that* $\gamma^*(x) = \gamma^*(0,x)$

Solution: Like in problem 5.3 we want to use Claim 5. 1 to solve the problem, i.e., to find a C^1 *function* ϕ *such that* $\phi \geq g$ *(majorant of* g*) and*
$L\phi + f = 0$ *in* $D = \{\phi(s,x) > g(s,x)\}$ *and* $L\phi + f \leq 0$ *on V minus D.*

Now $Lg + f = (\frac{\partial}{\partial s} + \frac{1}{2}\frac{\partial^2}{\partial x^2})g + f = (1 + (1-\rho)x^2)e^{-\rho s}$*, and we define* $U = \{(s,x); Lg + f > 0\} = \{(s,x); 1 + (1-\rho)x^2 > 0\}$ *If* $\rho \in [0,1]$*, then* $U = R_+ x R = D$ *and* τ^* *does not exist. Hence, we assume that* $\rho > 1$ *and in the continuation region the function* $\phi(s,x)$ *that we are looking for has to fulfill* $L\phi + f = 0$ *or*
$\frac{\partial \phi}{\partial s} + \frac{1}{2}\frac{\partial^2 \phi}{\partial x^2} = -e^{-\rho s} x^2$ *So why not try* $\phi(s,x) = e^{-\rho s} A(x)$*! This implies that*

$\frac{1}{2}A''(x) - \rho A(x) = -x^2 \bullet$. *If we try* $A(x) = C_1x^2 + C_2x + C_3$ *and insert it into* $\bullet$ *we get*

$C_1 - \rho(C_1x^2 + C_2x + C_3) = -x^2$. *Solving for the constants yields*

$C_1 = 1/\rho, C_2 = 0,$ *and* $C_3 = 1/\rho^2$ *(since* $C_1 - \rho C_3 = 0$*). Hence, our candidate for*

$g^*(s,x) = \phi(s,x) = e^{-\rho s}(x^2/\rho + 1/\rho^2)$ *Moreover,* $\gamma(0,x) = (x^2/\rho + 1/\rho)$. *We can now specify the set U by the parameterization* $U = \{(s,x); x_{abs} < (\frac{1}{\rho - 1})^{1/2}\}$ *We can also guess D by putting* $D = \{(s,x); \phi(s,x) > g(s,x) = \{(s,x); -x_0 < x < x_0\}$ *and solving*

$$L\phi + f = 0 \, in \, D \rightarrow A''(x) - 2\rho A(x) = -2x^2$$
$$\phi = g \, at \, \partial D \rightarrow A(x) = x^2, x_{abs} = x_0$$

The solution $\gamma(0,x) = (x^2/\rho + 1/\rho)$ *is the particular solution when s=0. The general solution has the form*

$$A(x) = C_1e^{\sqrt{2\rho}x} + C_2e^{-\sqrt{2\rho}x} + \frac{x^2}{\rho} + \frac{1}{\rho}$$

If we define $\phi \equiv g \, in \, R^2$ *minus D, we can use the conditions at the boundary of D to determine the constants and the resulting function will be globally continuous. We get*

$$x_0^2 = C_1e^{\sqrt{2\rho}x_0} + C_2e^{-\sqrt{2\rho}x_0} + \frac{1}{\rho} + \frac{x^2}{\rho}$$
$$x_0^2 = C_1e^{-\sqrt{2\rho}x_0} + C_2e^{\sqrt{2\rho}x_0} + \frac{1}{\rho} + \frac{x^2}{\rho}$$

From symmetry we get $C_1 = C_2 = [(1 - \frac{1}{\rho})x_0^2 - \frac{1}{\rho^2}](e^{\sqrt{2\rho}x_0} + e^{-\sqrt{2\rho}x_0})^{-1}$

Finally we obtain a C^1 *function by solving for* x_0 *in the equation*

$$2x_0 = \frac{2x_0}{\rho} + \sqrt{2\rho}[(1 - \frac{1}{\rho})x_0^2 - \frac{1}{\rho^2}]\tan h(\sqrt{2\rho}x_0$$

which can be shown to have a unique solution for x_0.

From the Claim5.1 it now follows that

$\gamma(s,x)=\phi(s,x), D=\{(s,x);\phi(s,x)>g(s,x)\}$, *and* $\tau^*=\tau_D$

5.6 Solve the optimal stopping problem

$$g^*(s,x)=\sup_{\tau} E^{(s,x)}[e^{-\rho(s+\tau)}B_\tau^+]$$
$$B_t \in R \text{ and } x^+=\max[x,0]$$

Solution: For this problem it is only possible to find the derivatives of $g(s,x)=e^{-\rho s}x$ *for* $x>0$. *We obtain* $Lg=-\rho e^{-\rho s}$ *which is not helpful to design the continuation set from the set U. However, even starting points outside the set U can belong to the continuation region. For example, the set* $W=\{(s,x);\exists\tau \text{ with } g(s,x)<E^{s,x}[g(s+\tau,X_\tau)]\}$ *belongs to* $D=\{(s,x);g(s,x)<g^*(s,x)\}$. *The reason is that* $g^*(s,x)=\sup_{\tau>0} E^{(s,x)}[g(s,x+\tau)]\geq E^{(s,x)}[g(s,x+\tau)]>g(s,x)$. *(Why is the last inequality true?)Moreover, one can show that the continuation region, for an objective function like the one here, is independent of translations of time. The reason is similar to the reason behind Observation 1 in chapter 4. Hence, if the continuation region is connected it has to have the shape* $D(x_0)=\{(s,x);x<x_0\}$. *However, if it also consists of a set G disjoint from W, then* $g(s,x)=g^*(s,x)$ *and G would be empty,* $E^{s,x}(g(s+\tau,x)<g(s,x)\,\forall\tau$ *.Note that* $U\subset W$.

Hence, we proceed with $D(x_0)=\{(s,x);0<x<x_0\}$ *and look for a function*

$$\phi(s,x)=h(s,x) \text{ for } x<x_0$$

$$\phi(s,x)=e^{-\rho s}x^+ \text{ for } x\geq x_0$$

h is chosen as the solution to Lh=0 for $x<x_0$, *and* $h(s,x)=e^{-\rho s}x^+$ *for* $x\geq x_0$.

We obtain $h(s,x)=e^{-\rho s}[C_1e^{\sqrt{2\rho}x}+C_2e^{-\sqrt{2\rho}x}]$. *To obtain* $h(s,-\infty)=0$, *we have to put* $C_2=0$. $h(s,x_0)=x_0$ *gives* $C_1=x_0e^{-x_0\sqrt{2\rho}}$ *for* $x_0>0$ *For* $x_0\leq 0, h(x,x)=0$*.This gives*

$h(s,x) = x_0 e^{-\rho s} e^{(x-x_0)\sqrt{2\rho}}$ *The function* $\phi(s,x)$ *is now globally continuous. It remains to pick* x_0 *in an optimal manner. Maximization at fixed x gives* $x_0 = \dfrac{1}{\sqrt{2\rho}}$ *. Finally the smooth pasting condition means that the derivative* $\dfrac{\partial \phi}{\partial x} = e^{-\rho s}\sqrt{2\rho}\, x_0 e^{(x-x_0)\sqrt{2\rho}}$, $x < x_0$ *and fortunate,* $x_0 = \dfrac{1}{\sqrt{2\rho}}$ *makes the derivative continuous(???). Hence*

$\tau^* = \tau_D$.

5.8 Suppose that the price P_t *of oil follows a geometric Brownian motion*

$dP_t = \alpha P_t dt + \beta P_t dB_t,\ P_0 = p$

Let the extraction be given by the equation $dQ_t = -\lambda Q_t dt,\ Q_0 = q$ *. Here* λ *is the constant extraction rate. If we stop the extraction at* $\tau = \tau(\omega)$, *then the expected total profit is given by*

$$J^\tau(s,p,q) = E^{(s,p,q)}[\int_0^\tau (\lambda P(t)Q(t) - K)e^{-\rho(s+t)}dt + e^{-\rho(s+t)}g(P(\tau),Q(\tau))$$

Here $\rho > 0$ *is the discount factor, K fixed extraction cost over the interval dt, and* $g(\bullet)$ *a bequest function.*

5.8a Write down the characteristic operator A of the diffusion process

$$dX(t) = \begin{pmatrix} dt \\ dP(t) \\ dQ(t) \end{pmatrix} \qquad X(0) = (s,p,q)\,,$$

and the variational inequalities in Claim5.1 corresonding to the optimal stopping problem $G^*(s,p,q) = \sup_\tau J^\tau(s,p,q) = J^{\tau^*}(s,p,q)$.

Solution: The characteristic operator can be written

$$A_x = \frac{\partial}{\partial s} + \alpha p \frac{\partial}{\partial p} + \frac{1}{2}\beta^2 p^2 \frac{\partial^2}{\partial p^2} - \lambda q \frac{\partial}{\partial q}$$

If we assume that the function $g(\bullet) = pq$ *(quite reasonable!) and put*

$$f(t,p,q) = (\lambda pq - K)e^{-\rho t}, \tilde{g}(t,p,q) = g(p,q)e^{-\rho t}$$

we can write the objective finction in the following manner

$$J^{\tau}(s,p,q) = E^{(s,p,q)}[\int_0^{\tau} f(X(t))dt + \tilde{g}(X(\tau))]$$

Now $D = \{(s,p,q); \phi(s,p,q) > \tilde{g}(s,p,q)\}$ *and the variational inequalities are*

$$A_x\phi + f = 0 \ on D$$

$$A_x\phi + f \le 0 in CD$$

5.8

Find the domain U and conclude that if $\rho \ge \alpha$, *then* $\tau^* = 0$ *and* $g^*(s,p,q) = e^{-\rho s}pq$.

Show that for $\rho < \alpha$ *then,* $D \supset \{(s,p,q); pq > K/(\alpha - \rho)\}$

Solution: Formally U is defined as $U = \{(s,p,q); A_X\tilde{g} + f > 0\}$. *Since* $g(p,q) = pq$, *we have that* $\tilde{g}(s,p,q) = e^{-\rho s}pq$ *. Applying the operator yields*

$$A_x\tilde{g} + f = e^{-\rho s}(\alpha - \rho - \lambda)pq + (\lambda pq - K)e^{-\rho s} = e^{-\rho s}[(\alpha - \rho)pq - K]$$

For $\alpha \le \rho$ *implies* $A_x\tilde{g} + f < 0$ *on* $R^{+\,3}$ *and U=D and* $\tau^* = 0$. *The economic interpretation is that the drift in price is lower than the rate of discount, and it is optimal to sell the mine at the ruling price.*

If $\alpha > \rho$ *it it is straightforward to see that* $U = \{(s,p,q); pq > K(\alpha - \rho)^{-1}\}$.

5.7c Solve the optimal stopping problem that has been introduced for,

$\alpha > \rho$.

Solution: Given the geometry of U is reasonable to guess that

$D = \{(s, p, q); pq > y_0, p, q > 0\}$

The advise in Öksendal is to put

$$\phi(s,p,q) = \begin{pmatrix} e^{-\rho s} pq,\ 0 < pq \le y_0 \\ e^{-\rho s}\gamma(pq),\ pq = y_0 \end{pmatrix}$$

Now we use the variational inequalities on ϕ *to get*

$$A_x\phi + f = 0\, on\, D$$
$$\phi(s,p,q) = e^{-\rho s} pq,\ pq = y_0$$

We get $A_x\phi + f = \dfrac{\partial\phi}{\partial s} + \alpha p \dfrac{\partial\phi}{\partial p} + \dfrac{1}{2}\beta^2 p^2 \dfrac{\partial^2\phi}{\partial p^2} - \lambda q \dfrac{\partial\phi}{\partial q} + (\lambda pq - K)e^{-\rho s}$

With $\phi = \psi(pq)e^{-\rho s}$ *we obtain after putting pq=r*

$\dfrac{1}{2} r^2\beta^2\psi''(r) + (\alpha - \lambda) r\psi'(r) - \rho\psi(r) + \lambda r - K = 0$ *for* $\{r; y_0 < r\}$*, i.e., in D and*

$\psi(r) = r,\ r = y_0$

We are looking for $\psi(r) = C_1 r^{\gamma_1} + C_2 r^{\gamma_2} + Ar + B$

We can now determine A and B by indefying coefficients

$$(\alpha - \lambda)Ar - \rho(Ar + B) = K - \lambda r$$
$$A = \frac{-\lambda}{\alpha - \lambda - \rho}, B = \frac{-K}{\rho}$$

The roots γ_1, γ_2 *are now determined by the second order equation in* γ *(put* $\psi(r) = r^\gamma$ *and work with the homogeneous equation).*

$$\frac{1}{2}\beta^2\gamma(\gamma - 1) + (\alpha - \lambda)\gamma - \rho = 0$$

We get $\gamma = \frac{1}{\beta^2}(\lambda - \alpha + \frac{\beta^2}{2}) \pm \sqrt{\frac{2\rho}{\beta^2} + \frac{1}{\beta^4}(\alpha - \lambda - \frac{\beta^2}{2})^4}$

We note that we have two real roots, $\gamma_1 > 0, \gamma_2 < 0.$ *We can write the differential equation, including the shape of the particular solution as*

$$\psi(r) = C_1 r^{\gamma_1} + C_2 r^{\gamma_2} + \frac{\lambda r}{\lambda + \rho - \alpha} - \frac{K}{\rho}.$$

If $\alpha < \rho + \lambda$ *it follows that the last term is increasing in* $r = pq$*, which is reasonable. However, if we want to avoid a profit that is exponentially increasing in r we put* $C_1 = 0$. *To fix global continuity of the profit function we solve for* C_2 *by putting* $\psi(y_0) = y_0$,

or $C_2 y_0^{\gamma_2} + \frac{\lambda r}{\lambda + \rho - \alpha} = y_0$. *This yields*

$$C_2 = y_0^{-\gamma_2}[\frac{\lambda y_0}{\lambda + \rho - \alpha} + \frac{K}{\rho}]$$

We can now write the candidate for the optimal value function the following manner

$$\phi(s,p,q) = \begin{pmatrix} e^{-\rho s} h(y_0) r^{\gamma_2} + \frac{\lambda r}{\lambda + \rho - \alpha} - \frac{K}{\rho} \; on \, D \\ e^{-\rho s} r \; in \, R_+^3 / D \end{pmatrix}$$

where $h(y_0) = C_2$. *It now remains to determine* y_0 *such that* $\phi \in C_1(R_+^3)$. *Hence we solve*

$$\gamma y_0^{\gamma-1} y_0^{-\gamma}[\frac{\lambda y_0}{\lambda + \rho - \alpha} + \frac{K}{\rho}] + \frac{\lambda r}{\lambda + \rho - \alpha} = 1 \Leftrightarrow$$

$$\frac{\gamma K}{\rho}\frac{1}{y_0} + \frac{\gamma(\rho - \alpha)}{\lambda + \rho - \alpha} + \frac{\lambda}{\lambda + \rho - \alpha} = 1 \Leftrightarrow$$

$$\frac{\gamma K}{\rho}\frac{1}{y_0} = \frac{\lambda + \rho - \alpha - \lambda - \gamma(\rho - \alpha)}{\lambda + \rho - \alpha} \Leftrightarrow$$

$$\frac{\gamma K}{\rho}\frac{1}{y_0} = \frac{(1 - \gamma}{\lambda + \rho - \alpha} \; and$$

$$y_0 = \frac{\gamma K}{\rho} \frac{(\lambda + \rho - \alpha)}{(1-\gamma)(\rho - \alpha)}$$

Since $\gamma < 0,\ \alpha > \rho \rightarrow y_0 > 0 \Leftrightarrow \lambda + \rho > \alpha$

If $\alpha > \lambda + \rho$ *the value function increases without bound and* $\tau^* \rightarrow \infty$ *,since* $y_0 < 0$ *and the process will stay in the continuation region D for ever.*

Chapter 6: A Taste of Financial Economics-Black and Scholes Formula[36]

In this final chapter we will give the reader a taste of a modern part of Financial Economics. We will try to do this in the same spirit as the material presented in Chapter 0. The analysis will be related to a classical theoretical problem[37] in Financial Economics called option pricing. The underlying stochastic processes will be Brownian motion, and we apply Ito calculus. The idea is to price the option to buy a financial instrument (sometimes called a contingent claim) that gives the holder the right (but not the obligation) to buy (sell) at a time T in the future an underlying asset (say a share) at a given price $\bar{P}_T$. The price is called the strike price, and if you buy the right to buy an underlying financial instrument at a given price at time T, you buy a call option. If you buy the right to sell the instrument, you buy a put option.

This pricing problem was solved in a very influential article by Black and Scholes (197+) and the resulting formula is called Black and Scholes formula.

We have already in Chapters 4 and 5 solved optimal portfolio selection problems, and in Chapter 3 we introduced the Feynman –Kac–representation formula which will be one of the means for solving our pricing problem. In this chapter we will introduce yet another important theoretical tool which has to do with *Absolutely Continuous Measure Transformations* or what is also called Girsanov theory.

6.1 A self-financing portfolio

In this section we will introduce informally the shape of the budget constraint or, which amounts to the same thing the portfolio dynamics. We will derive it for a portfolio where there are no additions or subtractions from external sources. Such a portfolio is called a *self-financed portfolio.* For simplicity we will assume that there are only two assets; one risky asset, and one safe.

---

[36] This section relies heaviliy on Björk (1994, 1998 and 2009).

[37] Bachelier (1900) introduced the problem.

The price of the safe asset (a bond) is $X_0(t)$, and the price of the risky asset is $X_1(t)$. In vector form the price process is written $X(t) = (X_0(t), X_1(t))$. Moreover, $\theta(t) = (\theta_0(t), \theta_1(t))$ is the vector of the number of bonds and shares of the risky asset held during the period $[t, t + \Delta t)$. $\theta(t)$ is called the portfolio. $c(t)$ is the rate of money spent at time t on consumption at time t , and $V(t)$ is the value of the portfolio at time *t*. Time *t* is the start of period $[t, t + \Delta t)$, and we start from the portfolio $\theta(t - \Delta(t))$. The price vector $X(t)$ can be observed at time *t* and we choose a new portfolio $\theta(t)$ to be held during the period $[t, t + \Delta(t))$. We also choose consumption $c(t)$ to be held over the same period, i.e. both consumption and the new portfolio are held constant over the period.

The value of the portfolio at time *t* is

$$V(t) = \theta(t - \Delta(t))X(t) \tag{6.1}$$

The cost of the new portfolio at time *t* is $\theta(t)X(t)$, and the cost of consumption is $c(t)\Delta(t)$. This means that a discrete version of the budget equation can be written

$$X(t)\Delta\theta(t) + c(t)\Delta(t) = 0 \tag{6.2}$$

where $\Delta\theta(t) = (\theta(t) - \theta(t - \Delta(t))$. Note that this is a self-financing portfolio since there are no additions or subtractions from external sources. We need a continuous time version of it to be able to use Ito calculus. The problem with equation (6.2) is that it is formulated in backward increments in stead of forward increments which are the base for the Ito integral (differentials). In order to get forward differences we reformulate the budget constraint by adding and subtracting the term $X(t - \Delta(t))\Delta\theta(t)$ to the left hand side of equation (6.2). Collecting terms yields

$$X(t - \Delta(t))\Delta\theta(t) + \Delta X(t)\Delta\theta(t) + c(t)\Delta(t) = 0 \tag{6.3}$$

where $\Delta X(t) = X(t) - X(t - \Delta(t))$. If we now let $\Delta(t) \to 0$, we obtain

$$X(t)d\theta(t) + dX(t)d\theta(t) + c(t)dt = 0 \tag{6.4}$$

In the same spirit, letting $\Delta(t)$ go to zero in the value equation (6.1) we get

$$V(t) = \theta(t)X(t) \qquad (6.5)$$

and using Ito calculus yields

$$dV(t) = \theta(t)dX(t) + X(t)d\theta(t) + dX(t)d\theta(t) \qquad (6.6)$$

From (6.4) we get $dX(t)d\theta(t) = -(X(t)d\theta(t) + c(t)dt$ which substituted into (6.6) yields

$$dV(t) = \theta(t)dX(t) - c(t)dt \qquad (6.7)$$

which is the dynamics of a *self-financing portfolio* in continuous time.

We also need to determine with what information the portfolio is chosen. In the general case we would like the price process to be F_t^X adapted, i.e. we have information about the price process up to period *t.* We make the same assumption about the consumption process and the *portfolio strategy* $\theta(t)$. However, we will be restricted to the following form

$$\theta(t) = \theta(t, X(t)) \qquad (6.8)$$

which we by now know is called a Markovian strategy. The portfolio strategy does not depend on the history of the process.

The following definition is vital for pricing options in financial markets.

Definition 6.1: An arbitrage possibility in a financial market is a self- financed portfolio such that

(i) $V^{\theta}(0) = 0$

and

(ii) $V^{\theta}(T) > 0$*, with probability 1.*

The market is arbitrage free if arbitrage possibilities are lacking.

We now assume thast the rate of return (interest) on the safe asset is r .The following proposition shows under what circumstances that a self-financed portfolio is arbitrage free.

Proposition 6.1 *Assume that there exists a self-financed portfolio* θ *such that*
$dV^{\theta}(t) = k(t)V^{\theta}(t)dt$
where k(t) is F_t *adapted. Then* $k(t) = r$ *for all t, or there exists an arbitrage possibility*

We will omit the proof (the author has lost his notes), but the intuition should be clear from the special case $k(t) = k$. If $r > k$, it is profitable to sell the portfolio and put the money into a bank account, and if the opposite holds you can profit from borrowing money and invest in the portfolio.

6.2 The Black and Scholes formula

We will discuss the pricing of an option in a very simple setting. We assume that there are only two assets, one risky and one safe asset.

The option (contingent claim) is denoted Ψ , which at the exercise date is worth $\Phi[X(T)]$. We assume

(i) It is traded in a market
(ii) It has a price process $\Pi(t,\Psi) = F(t, X(t))$ which is twice continuously differentiable.
(iii) The market $(A(t), X(t), \Pi(t,\Psi))$ is arbitrage free. $A(t)$ is the a safe asset and $X(t)$ is a risky asset

Note that the price process depends on time and the underlying risky asset.
In the standard version of Black and Scholes model the contingent claim is a European call option that is worth $\Phi(T) = \max[X(T) - \bar{P}_T, 0]$. The holder of the option will use his right to buy the share iff $X(T) > \bar{P}_T$. An American option is a contingent claim where the exercise

date can be freely chosen in the interval [0,T].

The assets follow the processes

$$\begin{aligned} dX(t) &= \alpha X(t)dt + \sigma X(t)dB(t) \\ dA(t) &= rA(t)dt \end{aligned} \tag{6.9}$$

Here α is the local mean rate of the return on the risky asset (the stock), σ is the volatility, r is the return of the safe asset (the bond), and $B(t)$ is a Brownian motion process under a probability measure P.

Now, given (ii) above we ask what the price process of the option will look like. We do this in terms of an exercise and a discussion that result in a proposition that contains some remaining fog. The fog is lifted in section 6.3.

Exercise(Björk (1998,2009):

6.1 *Show that the price process of the option satisfies the following stochastic differential equation:* $d\Pi = \alpha_\Pi(t)\Pi(t)dt + \sigma_\Pi(t)\Pi(t)dB(t)$

where $\alpha_\Pi(t) = \frac{1}{F}\left(\frac{\partial F}{\partial t} + \alpha X\frac{\partial F}{\partial X} + \frac{1}{2}X^2\sigma^2\frac{\partial^2 F}{\partial X^2}\right)$

and $\sigma_\Pi = \frac{1}{F}\sigma X\frac{\partial F}{\partial X}$.

Solution: Use Ito's lemma on the price function to get

$\partial\Pi = \left(\frac{\partial F}{\partial t} + \alpha X\frac{\partial F}{\partial X} + \frac{1}{2}X^2\sigma^2\frac{\partial^2 F}{\partial X^2}\right)dt + \sigma X\frac{\partial F}{\partial X}dB$, *since* $F = \Pi$ *the result now follows by substitutions.*

Now we use that the market is assumed to be arbitrage free and create a self-financed portfolio consisting of the risky asset and the option. The idea is to use the result in Proposition 6.1 to restrict the self-financing portfolio to become arbitrage free. We need arbitrage freeness to nail down the price function of the option.

We start by the portfolio dynamics of the suggested self-financed portfolio.

$$dV^{\theta} = \theta_X dX + \theta_{\Pi} d\Pi = V^{\theta}(\frac{\theta_X X}{V^{\theta}}\frac{dX}{X} + \frac{\theta_{\Pi}\Pi}{V^{\theta}}\frac{d\Pi}{\Pi})$$

Define

$z_X = \frac{\theta_X X}{V^{\theta}}$ and $z_{\Pi} = \frac{\theta_{\Pi}\Pi}{V^{\theta}}$, Now substituting for $\frac{dX}{X}$ and $\frac{d\Pi}{\Pi}$ by using the differential equations for dX and $d\Pi$ yields

$$dV^{\theta} = V^{\theta}[z_X(\alpha dt + \sigma dB) + z_{\Pi}(\alpha_{\Pi} dt + \sigma_{\Pi} dB)] =$$
$$V^{\theta}[(\alpha z_X + \alpha_{\Pi} z_{\Pi})dt + (\sigma z_X + \sigma_{\Pi} z_{\Pi})dB]$$

We now use Proposition 6.1 to make the portfolio arbitrage free by putting $\sigma z_X + \sigma_{\Pi} z_{\Pi} = 0$, and $\alpha z_X + \alpha_{\Pi} z_{\Pi} = r$. Moreover, the budget shares sum to one, $z_X + z_{\Pi} = 1$. The idea is to solve for the budget shares (z_X, z_{Π}). The system is, however, over determined so one equation has to be written as a linear combination of the other two. To accomplish this we solve the budget share equation and the equation that neutralizes Brownian motion for (z_X, z_{Π}), and plug the result into the remaining equation. One obtains:

$$\alpha\sigma_{\Pi} - \sigma\alpha_{\Pi} = r(\sigma_{\Pi} - \sigma) \tag{6.10}$$

Finally, using the definitions of σ_{Π} and α_{Π} in exercise 6.1, we obtain a deterministic PDE with the following shape

$$\frac{\partial F}{\partial t} + rx\frac{\partial F}{\partial x} + \frac{1}{2}x^2\sigma^2\frac{\partial^2 F}{\partial x^2} - rF = 0 \tag{6.11}$$

with the boundary condition

$$F(T, x) = \Phi(x) \tag{6.12}$$

We are now ready to sum up what we have learnt.

Proposition 6.2 *(Black and Scholes equation) Given that the market for the safe and unsafe assets are specified by equation 6.9 and that we want to price a contingent claim of type* $\Psi = \Phi(X(T))$ *, then the only pricing function of the form* $\Pi(t,\Psi) = F(t, X(t))$ *, which is consistent with no arbitrage is when F is the solution to the boundary value problem 6.11-6.12 in the domain* $[0,T]xR_+$.

A few observations are valuable. First of all the price of the contingent claim will be a function of the underlying asset. Moreover, the deterministic PDE that determine the pricing equation does not contain the mean rate of return of the risky asset, but the safe interest rate. In other words, the pricing formula for the claim is independent of the rate of return of the underlying asset. This seems rather strange. Moreover, we would like to solve for the pricing function, and the idea that comes to our mind is to use Feyman-Kac-representation formula that we introduced in Chapter 3. However, the underlying market equation for the risky asset does not represent the pricing formula, since its drift is α , rather than *r*. However, if $\alpha = r$, we know from chapter 3 that the solution is given by

$$F(t,x) = e^{-r(T-t)} E^P_{t,x}[\Phi(X(T)] \tag{6.13}$$

where the mathematical expectation is taken with respect to the probability measure *P*, the Brownian motion process obeys this probability measure. This is an uninteresting special case and we have to find a solution in the general case. With the help of a deep mathematical theorem due to Kolmogorov’s student Girsanov, some of the fog will hopefully disappear.

6.3 Girsanov’s theorem and risk free valuation

What we will do in this section is to show how we can move from the objective probability measure *P* into another probability measure Q where the drift term will change, but the diffusion term will stay the same. This can be used in the Black and Scholes framework to transfer the *P*-dynamics into a *Q*-dynamics that is arbitrage free, it drifts according to the risk free interest rate, and the new model is represented in the PDE for the price equation under the probability measure Q . Hence we can use the Feynman-Kac-theorem to solve for the price of the claim under the Q dynamics, like in equation 6.13 above. The probability measures that can be transformed in this way must be related to each other in a special manner. The measure Q has to be *absolutely continuous* with respect to P .

Definition 6.2: Given a probability space (Ω, F, P) *, a probability measure Q is said to be absolutely continuous with respect to P, if*

$$P(A) = 0 \Rightarrow Q(A) = 0$$

and one writes $Q << P$*. If both* $Q << P$ *and* $P << Q$ *and we say that P and Q are equivalent. One writes* $Q \approx P$*.*

Here A denotes an event, F is a σ-algebra on the given event set Ω.

Exercise:

6.2 Prove that if P and Q are equivalent, then $P(A) = 0 \Leftrightarrow Q(A) = 0$ *and*
$P(A) = 1 \Leftrightarrow Q(A) = 1$*. (Björk)*

Proof: Use the definition, the event A and its complement.

The exercise tells us that P and Q agree on which events are impossible and those which are certain. Otherwise they can assign totally different probabilities to the same events. The following theorem gives a necessary and sufficient condition for $Q << P$.

Theorem 6.1 *We have* $Q << P$ *if and only if* $Q(A) = \int_A L(\omega) dP(\omega),$ *all* $A \in R$.

L *is called the Radon-Nikodym derivative or the likelihood ratio. One writes* $dQ = LdP$ *or*

$$\frac{dQ}{dP} = L.$$

We are interested in what an absolute continuous transformation does to the dynamics of a Brownian motion process. We start from a filtered probability space (Ω, F, P) with the filtration F_t; $t \geq 0$. Choose an F_t adapted process $y(t)$ and define the likelihood process

$$\begin{aligned} dL(t) &= \rho(t)L(t)dB(t) \\ L(0) &= 1 \end{aligned} \tag{6.14}$$

Exercise

6.3 Show that the solution of (6.14) is $L(t) = e^{[\int_0^t g(\tau)d\tau - \frac{1}{2}\int_0^t g^2(s)ds]}$

Solution: Let $L(t) = e^{z(t)}$, *where* $dz(t) = y(t)dB(t) - \frac{1}{2}y^2(t)dt, z(0) = 0$. *Moreover,*

$$dL(t) = L(t)[dz + \frac{1}{2}(dz)^2)] = L(t)[y(t)dB(t) - \frac{1}{2}y^2(t)dt + \frac{1}{2}y^2(t)dt] = L(t)y(t)dB(t).$$

We now introduce the following result that is a version of more general results first proved by Girsanov.

Theorem 6.2 *Given* $y(t), B(t),$ *and* $L(t)$ *as in 6.14, define a new measure* $dQ = L(t)dP(t)$ *on* F_t. *Then, the process B(t) , which under P is a standard Brownian motion can be written as* $dB(t) = y(t)dt + d\tilde{B}(t)$, *where* $d\tilde{B}(t)$ *is a standard Brownian motion process under Q.*

Now, substituting $dB(t) = y(t)dt + d\tilde{B}(t)$ into the Black and Scholes model yields the Q-dynamics

$$\begin{aligned} dX(t) &= [\alpha + y(t)\sigma]X(t)dt + \sigma X(t)d\tilde{B}(t) \\ dA(t) &= rA(t)dt \end{aligned} \tag{6.15}$$

From general arbitrage theory it follows that this model is arbitrage free iff there exists an absolutely continuous measure transformation of P into some measure Q such that under Q the ratio $X(t)/A(t)$ is a Q martingale, i.e. has no drift component. This can be shown to imply that under Q the local rate of return of $X(t)$ should equal the interest rate. Hence we have to put

$$r = \alpha + y(t)\sigma \tag{6.16}$$

which implies that we choose the “Girsanov kernel” in the following manner

$$y(t) = \frac{-(\alpha - r)}{\sigma} \qquad (6.17)$$

The economic interpretation of $-y(t)$ is the price of risk or the Sharp ratio.

We are now ready to use Feynman-Kac theorem from Chapter 3 to prove the following result.

Proposition 6.2, *The arbitrage free price* $\Pi(t, X(t)) = F(t, X(t))$ *of the contingent claim* $\Psi = \Phi(X(T))$ *is given by* $F(t,x) = e^{-r(T-t)}E^{Q}_{t,x}[\Phi(X(T))]$. The Q dynamics is given by

$$dX(u) = rX(u)du + \sigma X(u)d\tilde{B}(t)$$
$$X(t) = X_t$$

$\tilde{B}(t)$ *is a Brownian motion process under probability measure Q.*

Remark1: We have not yet said much about conditions on the Girsanov kernel function y(t). However, from exercise 6.3 it looks like the stochastic variable L(t) is a Martingale under the probability measure P. The reason is that the dynamics of L(t) has no drift term. There is, however, conditions on y(t) that have to be fulfilled. A sufficient condition is called Novikov's condition. It reads $E^P[e^{(\frac{1}{2}\int_0^T y^2(t,\omega)dt}] < \infty$.

Remark2: Note that under the new measure Q *individuals' attitudes to risk do no longer matter. The reason is that when the market is arbitrage free the portfolio drifts according to the risk free interest rate. One talks about risk neutral valuation.*

Exercise:

6.4 Prove that if y(t) satisfies Novikov's condition, then $E^P[L(t)] = 1$ *for all* $t \in [0,T]$.

Solution: Since $L(t)$ *is a Martingale it holds that* $E^P[L(t)|F_s] = L(s),$ *and* $L(0) = 1$.

6.5 Prove that $y(t)$ *in the Black and Scholes model fulfils Novikov's condition.*

Solution: Integrate.

How can we solve the Black and Scholes Model analytically for a European Call Option? Let us at least sketch the general idea. The details are available in e.g. Björk 1998.We can solve the differential equation for $X(T)$ under the Q dynamics to get

$$X(T) = x_t \exp\{(r - \frac{1}{2}\sigma^2)(T-t) + \sigma[\tilde{B}(T) - \tilde{B}(t)]\} = xe^Y \tag{6.18}$$

Y has the distribution $N[(r-\frac{1}{2}\sigma^2)(T-t), \sigma\sqrt{T-t}]$. Now we obtain, from 6.13 and 6.15, the pricing formula

$$F(t, x(t)) = e^{-r(T-t)} \int_{-\infty}^{\infty} \Phi(xe^y) f(y) dy \tag{6.19}$$

where $\Phi[X(T)] = \max[X(T) - \bar{P}_T, 0]$ for an European Call Option. Defining $r - \frac{\sigma^2}{2} = \tilde{r}$ and $T - t = \tau$, we can write $X(T) = x_t e^{\tilde{r}\tau + \sigma\sqrt{\tau}Z}$, where Z has distribution $N(0,1)$ Inserrting the new information into (6.19) yields a new shape to the integral which now reads

$$F(\bullet)e^{r\tau} = \int_{-\infty}^{+\infty} \max[xe^{\tilde{r}\tau + \sigma\sqrt{\tau}z} - \bar{P}_T, 0] g(z) dz \tag{6.20}$$

where $g(z) = \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}}$ is the density of the *N(0,1)* distribution. Moreover, the integral vanishes for $z < z_0$, where $z_0 = \frac{\ln {P_{\bar{T}}}/{x} - \tilde{r}\tau}{\sigma\sqrt{\tau}}$. The reason is that the net value of the option is negative and there will be no trade. Hence, one can write

$$F(\bullet)e^{r\tau} = \int_{z_0}^{+\infty} \max[xe^{\tilde{r}\tau + \sigma\sqrt{\tau}z} - \bar{P}_T, 0] g(z) dz \tag{6.21}$$

With a few of more tricks the integral can be further simplified. The tricks are not trivial to this author, so if you fail look up the details in Björk (1998/2009). Öksendal is also interested in Option Pricing. He has the details in chapter 12. The six and five editions are come from

2003 and 2000. However, Kaj Nyström has also worked hard in Öksensendal (2000) with 2001 for the book. The exercises in chapter 12 are quick and elegant.

Exercise

6.6 *Let* B_t *be a 1-dim Brownsk movement and* $F(w)$ *, find* $z \in R,\ \phi(t,w) \in V(0,T)$ *such that*

. $F(w) = z + \int_0^T \phi(t,w)dB(t)$ i

(i)

$F(w) = B^2(T,w)$

Ito's formula gives $dB^2(t) = 2B(t)dB(t) + dt$ *and we get*

$$F(w) = T + 2\int_0^{BT} B(t)dB(t)$$

(ii)

$$dB^3(t) = 3B^2(t)dB(t) + 3B(t)dt$$
$$= B^2(t)dB(t) + 3d(tB(t) - 3tdB(t)$$

And

$$B^3(T) = 3\int_0^T B^2(t)dB(t) + 3TB(T) - 3\int_0^T tdB(t)$$

$$= 3\int_0^T B^2(t)dB(t) - 3\int_o^T tdB(t) + 3T\int_0^T dB(t)$$

which give the representation formula.

(iii)

Let us now use theorem 12.3.3

$F(w) = h(B(T,w))$ *h(x)=* e^x

The theorem 12.3.3 help by $F(w) = h(B(T,w))$ *and* $h(x) = e^x$ *. The theorem 12.3.3 use*

$$F(w) = E^y[h(B(T))] + \int_0^T \frac{\partial}{\partial z} E^2[h(B(T-t))]_{z=B(t)} dB(t)$$

We note that $E^{y}[h(B(s))],\ y\in R, s\in[0,T]$ *and have to be calculated. Itos formula hold.* $d[h(B(s)) = h(B(s))]dB(s)+\frac{1}{2}h(B(s))]$ *Let* $f(s)=h(B(s))=E^{y}(h(B(s))]$ *and give* $f^{'}(s)=\frac{1}{2}f(s),\ f(0)=e^{y}$.

We get that $f(s)=e^{s/2+y}$ *d.v.s is* $E^{y}[h(B(s))]=e^{s/2+y}$ *Hence, we know that*

$\frac{\partial}{\partial y}E^{y}[h(B(T-t)]_{y=B(t)}=e^{(T-t)/2+B(t)}$ *and we get*

$$F(w)=e^{T/2+y}+\int_{0}^{T}e^{(T-t)/2+B(t)}dB(t)$$

6.7

Let B_t *be a 1-dim Brownsk movement and* $F(w)$ *, find* $z\in R,\ \phi(t,w)\in V(0,T)$ *such that*

$$.\,F(w)=z+\int_{0}^{T}\phi(t,w)dB(t)\ \mathrm{i}$$

(i)

$$F(w)=B^{2}(T,w)$$

Ito's formula gives $dB^{2}(t)=2B(t)dB(t)+dt$ *and we get*

$$F(w)=T+2\int_{0}^{BT}B(t)dB(t)$$

(ii)

$$dB^{3}(t)=3B^{2}(t)dB(t)+3B(t)dt$$
$$=B^{2}(t)dB(t)+3d(tB(t)-3tdB(t)$$

And

$$B^{3}(T)=3\int_{0}^{T}B^{2}(t)dB(t)+3TB(T)-3\int_{0}^{T}tdB(t)$$

$$=3\int_{0}^{T}B^{2}(t)dB(t)-3\int_{o}^{T}tdB(t)+3T\int_{0}^{T}dB(t)$$

which give the representation formula.

(iii)

Let us now use theorem 12.3.3

$F(w)=h(B(T,w))$ $\qquad$ *h(x)=* e^{x}

The theorem 12.3.3 help by $F(w)=h(B(T,w))$ *and* $h(x)=e^x$ *. The theorem 12.3.3 use*

$$F(w)=E^y[h(B(T))]+\int_0^T \frac{\partial}{\partial z}E^2[h(B(T-t))]_{z=B(t)}dB(t)$$

We note that $E^y[h(B(s))],\ y\in R, s\in[0,T]$ *and have to be calculated. Itos formula hold.* $d[h(B(s))=h(B(s))]dB(s)+\frac{1}{2}h(B(s))]$ *Let* $f(s)=h(B(s))=E^y(h(B(s))]$ *and give* $f^{'}(s)=\frac{1}{2}f(s),\ f(0)=e^y$.

We get that $f(s)=e^{s/2+y}$ *d.v.s is* $E^y[h(B(s))]=e^{s/2+y}$ *Hence, we know that*

$$\frac{\partial}{\partial y}E^y[h(B(T-t)]_{y=B(t)}=e^{(T-t)/2+B(t)}$$ *and we get*

$$F(w)=e^{T/2+y}+\int_0^T e^{(T-t)/2+B(t)}dB(t).$$

6.8

Suppose the market is given by

$$dX_0=\rho X_0(t)dt \quad X_0(0)=1$$
$$dX_1=(m-X_1(t))dt+\sigma dB(t) \quad X_1(0)=X_1$$
$$\rho, m>0$$

a) We are searching $E_Q(\xi(T)F$ *where* $F=X_1(T,w)$ *and we read*

$$\bar{X}_1=\xi(t)X_1(t)$$

$$d\bar{X}_1(t)=\xi(t)(dX_1-\rho X_1 dt)=\xi(t)[m-(1+\rho)X_1(t)]dt+\sigma dB(t)$$

Moreover,

$$d\bar{B}(t)=\frac{1}{\sigma}(m-(1+\rho)X_1)dt+dB(t)$$

We get

$$d\bar{X}_1(t)=\xi(t)\sigma d\bar{B}(t)$$

where $\bar{B}(t)$ *Browns motion under* σ *can get*

$$\bar{X}_1(t)=\bar{X}(0)+\int_0^T \xi(t)\sigma d\bar{B}(t)$$

We get $E_Q[\xi F] = F_Q[\bar{X}_1] = \bar{X}_1(0) = X_1(t) = x_1$

b) With a) gets $\xi F(w) = x_1 + \int_0^T d\,\bar{X}_1(t)$

Take $\theta_1(t) = 1$ *and* θ_0 *so that the portfolio becomes financed.*

∵

6.9

Let $dX_0 = e^{X_0(t)}dt \quad X_0(0) = 1, \rho > 0$. *We use .* $F(w) = B(T,w)$ *We are calculating* $E_Q(\xi(t)F)$ *on a number of markets and fine a hedge* $\theta = (\theta_0(t), \theta_1(t))$.

a) $dX_1(t) = \alpha X_1(t)dt + \beta X_1(t)dB(t)$, $n = (\alpha - \rho)/\beta$ *and ..* $d\,\bar{B} = ndt + dB$. *According to Girsanov's Theorem is* $\bar{B}$ *a Browns movement:*

$$\xi(T)F(w) = e^{-\rho T}B(T,w) = e^{-\rho T}\,\bar{B}(T,w) - e^{-\rho T}nT =$$

$$-e^{-\rho t}n + e^{-\rho T}\int_0^T d\,\bar{B}(t) =$$

Under Q it holds $d\,\bar{X}_1(t) = \beta\,\bar{X}_1(t)d\,\bar{B}$ *and =* $-e^{-\rho T}nT + e^{-\rho T}\int_0^T d\,\bar{X}_1(t)/\beta\,\bar{X}_1$ *and we get*

$$E_Q[\xi(T)F] = -e^{-\rho T}nT \quad .$$

b) Find the replicating portfolio $\theta(t) = (\theta_0(t), \theta_1(t))$ for this the claim. We have that

$$V^{\theta_1}(t) - V^{\theta_0}(t) = \int_0^t \theta_0(s)dX_0(s) + \int_0^t \theta_1(s)dX_1(s)$$

Let $f(t) = e^{-\rho(T-t)}\beta^{-1}$

Then $\theta_1(t) = f(t)X_1^{-1}(t)$ and we get that

$$\theta_0(t)X_0(t) + \theta_1(t)X_1(t) = V^\theta + \int_0^t \theta_0(s)dX_0(t) + \int_0^t \theta_1(s)dX_1(s)$$

which gives

$$\theta_0(t)X_0(t) + f(t) = V^0(0) + \int_0^t f(s)(\alpha ds + \beta dB(s))$$

Let

$$Y_0(t) = \theta_0(t)X_0(t)$$

$$dY_0(t) + f^1(t)dt = \theta_0(t)\rho X_0(t) + df(t)dt + f(t)\beta dB(t)$$

$$dY_0(t) - \rho Y_0(t) = (\alpha f(t) - f^{'}(t))dt + f(t)\beta dB(t)$$

$$d(Y_0\xi) = (\alpha f(t) - f^{'}(t))\xi dt + f(t)\xi\beta dB(t)$$

$$(\alpha f(t) - f^{'}(t))\xi(t)) = \frac{e^{-\rho t}}{\beta}[\alpha e^{-\rho(T-t)} - \rho e^{-\rho(T-t)}] - e^{-\rho(T-t)}n$$

$f(t)\xi(t) = e^{-\rho T}$ and

$$\theta_0(t)X_0(t)\xi = \theta_0(0) + nTe^{-\rho T} + B(T)e^{-\vartheta T}$$

Hence,

$$\theta_0(t) = \theta_0(0) + (nT + B(T))e^{-\rho T}$$

Appendix

A Sketch of Girsanov's Theorem and Tomas Björk

Define a measure $Q = Q(t)$ by

$$dQ = L(t) / dP$$

$$\int_B dQ = Q(B) = \int_B L(t) \; P \quad \forall B \in R$$

The measure transformation is generate a likelihood process defined by

$$L(T) = \frac{dQ}{dP} \qquad \text{on } F_t$$

and L is a P-martingale. Moreover, W is also a martingale (suitably integral), and it is natural to define L as the solution of the SDE

$$dL = \rho(t)L(t)W(t)^P$$
$$L_0 = 1$$

For some choice of the process ρ.

The transformation from P to a new measure Q by the following idea;

- Choose an arbitrary adapted process ρ .
- Define a likelihood process L by

$$dL(t) = \rho(t) \; L(t)dW(t)^P$$
$$L(0) = 1$$

- Define a new measure Q by setting

$$dQ = L(t)dP(t)$$

on F_t for all $t \in [0,T]$.

The Ito formula we can easily see that we can express L as

$$L(t) = e^{\int^T \rho_s dW_s^P - \frac{1}{2}\int_0^t \rho_{sds}^2}$$

where L is nonnegative, which is necessary to act as a likelihood process. If ρ is integrable also clear enough it is a martingale and the initial conditional $L(0) = 1$ guarantees that $E^P[L(t)] = 1$.

To see what the dynamics of W^P are under Q, we recall that if is a process X has the dynamics

$dX(t) = \mu(t)dt + \sigma(t)dW(t)^P$

The drift is μ and σ is a conditional drift and quadratic variation processes respective. A bit more precisely but is still heuristically, we have

$E^P[dX(t)/F_t] = \mu(t)dt$

$E^P[(dX(t))^2 / F_t] = \sigma^2(t)dt$

Here we have an informal interpretation $dX(t) = X_{t+dt} - X_t$. Define the process $X(t) = W^P$, i.e. $\mu = 0$ and $\sigma = 1$ under P. Our task is to compute the drift and diffusion under Q and for that we will use the Abstract Bayes' Theorem. Since L is a $P-$ martingale, and recalling that $dX(t) \in F_{t+dt}$, we obtain

$$E^Q[dX(t)/F_t] = \frac{E^P[L(t+d(t))dX(t)/F_t]}{E^P[L(t+dt)/F_t]} = \frac{E^P[L(t+d(t)]dX(t)/F_t]}{L(t)}$$

$$= \frac{E^P[L(t)dX(t) + dL(t)dX(t)/F_t]}{L(t)} = \frac{E^P[L(t)dX(t)/F_t]}{L(t)} + \frac{E^P[dL(t)dX(t)/F_t]}{L(t)} \text{ that}$$

Since L is adapted (so is $L(t) \in F_t$) and X has zero drift under P, we have

$$\frac{E^P[L(t)dX(t)/F_t]}{L(t)} = L(t).\frac{E^P[dX(t)/F_t]}{L(t)} = E^P[dX(t)/F_t] = 0 \bullet dt$$

Furthermore we have

$$dL(t)dX(t) = L(t)\rho(t)dW(0 \bullet dt + dW^P(t)) = L(t)\rho(t)(dW(t)^P)^2 = L(t)\rho(t)dt$$

Using the fact that under $L(t)\rho(t) \in F_t$ we get

$$\frac{E^P[dL(t)dX(t)/F_t]}{L(t)} = \frac{L(t)\rho(t)}{L(t)}dt = \rho(t)dt$$

Moreover, the fact that under P we have $dX^2 = dt$ and easily compute the quadratic variation of X and Q as

$$E^Q[(dX_t)^2 / F_t] = \frac{E^P[L(t+dt)(dX^2)/F_t]}{L(t)} = \frac{E^P[L(t+dt)/F_t]}{L(t)}$$

$$\frac{E^P[L(t+dt)/F_t]}{L(t)}dt = \frac{L\ (t)}{L(t)}dt = dt$$

Summing up

$E^Q[dX(t)/F_t] = \rho_t dt$

$E^Q[(dX(t))^2 / F_t] = 1 \bullet dt$

Or in other words: We see that we should be able to write the $P-Wiener(Brown)$ process W^P as

$$dW^P(t) = \rho(t)dt + dW(t)^Q$$

where W^Q is a $Q-Wiener$ process this is precisely the content of Girsanov Theorem, which we now formulate.

Theorem: (The Girsanov-(Björk) Theorem) *Let W^P be a d-dimensional standard P-Wiener process on $(Q, F, P, \underline{F})$ and let ρ be any d-dimension adapted column vector process. Choose a fixed T and define the process L on $[0,T]$ by*

$$dL(t) = \rho * (t)L(t)dW^P(t)$$
$$L(0) = 1$$

Assume that

$$E^P[L(T)] = 1$$

and define the new probability measure on F_t by

$$L(T) = \frac{dQ}{dP} \quad on\, F_T \ .$$

Then

$$dW_t^P = \rho_t dt + dW_t^Q$$

when W_t^Q is a $Q-Wiener$ process (One can also define $W^Q(t) = W^P(t) - \int_0^t \rho(s)ds$)

The proof is difficult.

References

Aronsson T., Löfgren, K.G., and Backlund K. (2004) *Welfare measurement in Imperfect Markets,* Cheltenham: Edward Elgar

Aronsson T., Löfgren K.G., and Nyström K. (2003) Stochastic Cost benefit Rules: A Back of the Lottery Ticket Calculation Method, *Umeå Economic Studies, No 606.*

Bachelier L. (1900) Theorie de la Speculation, *Annales l'Ecole Normale Superieure 17, 21-86.*

Björk T., (1994) *Stokastisk Kalkyl och Kapitalmarknadsteori, del 1,* Stockholm: Matematiska Institutionen KTH

Björk T., (1998/2004/2009) *Arbitrage Theory in Continuous Time,* Oxford: Oxford University Press

Black F and Shooles M (1973) The Pricing of Options and Corporate Liabilities, *Journal of Political Economy 81, 659-83.*

Einstein, A (1956) *Investigation on the Theory of Brownian Motion,* New York: Dover (Contains his 1905 paper).

Li C.Z. and Löfgren K.G. (2012) Genuine Saving under Stochastic Growth, *Letters in Spatial and Resource Sciences 5, 167-174.*

Li , C-Z. and K-G. Löfgren (2013),Genuine Saving Measurement and Uncertainty and its Implication for Depleteable Resource Management. *Environmental Economics 3.*

Malliaris A.G. and Brock W.A. (1991) *Stochastic Methods in Economics and Finance,* Amsterdam: North Holland.

Mangel M. (1985) *Decision and Control in Uncertain Resource Systems,* London: Academic Press.

Merton R. (1975) An Asymptotic Theory of Growth under Uncertainty, *Review of Economic Studies 42, 375-93.*

Ramsey F.P. (1928) A Mathematical Theory of Saving, *Economic Journal 38, 543-549.*

Samuelson P.A. (1965) Proof that Properly Anticipated prices Fluctuates Randomly, *Industrial Management Review 6, 41-49.*

Solow, R.M. (1956) A Contribution to the Theory of Economic Growth, *Quarterly Journal of Economics 70, 65-94.*

Stratonovich R.L. (1966) A New Representation for Stochastic Integrals and Equations, *Siam Journal of Control 4, 362-71.*

Swan, T.W. (1956) Economic Growth and Capital Accumulation, *Economic Record 32, 334-361.*

Weitzman, M.L. (1976) On the Welfare Significance of National Product in a Dynamic Economy, *Quarterly Journal of Economics 90, 156-162*.
Åström K. (1970) *Introduction to Stochastic Control Theory*, London: Academic Press.
Öksendal, B (2000, 2003) *Stochastic Differential Equations*, Heidelberg: Springer.