CHALMERS | GÖTEBORG UNIVERSITY

MASTER THESIS

Backward Stochastic Differential Equations without Pain

John Löfgren

Department of Mathematical Statistics
CHALMERS UNIVERSITY OF TECHNOLOGY
GÖTEBORG UNIVERSITY
Göteborg, Sweden 2012

Thesis for the Degree of Master of Science

Backward Stochastic Differential Equations without Pain

John Löfgren

CHALMERS | GÖTEBORG UNIVERSITY

Department of Mathematical Statistics
Chalmers University of Technology and Göteborg University
SE – 412 96 Göteborg, Sweden

Göteborg, June 2012

Abstract

Backward stochastic differential equations or BSDEs for short have been studied quite extensively the last two decades. The results have proved useful in stochastic PDEs, mathematical finance, stochastic controls etc.

The existing literature so far seems mostly written as correspondance between professionals with extensive experience and knowledge of both theory and applications while the novice reader will, at least more likely, be overwhelmed by the sometimes condensated arguments and frequent references that can occur in those publications.

We believe, however, that any reader who is familiar with basic measure theoretic probability, stochastic processes and stochastic integration can appreciate the ideas behind backward stochastic differential equations and become familiar with the basic results without to much hazzle.

The text then, is an attempt to present the most fundamental aspects of the theory to the senior undergraduate or beginning graduate student.

We will start out by discussing the most rudimentary problem. After this is done we will discuss more general problems and try to qualitatively explain the sometimes less understood fact that the solution to a BSDE consists of a unique pair of adapted processes.

Finally we will mention a bit about the history of BSDEs and suggest some further reading on the subject.

Acknowledgements

First and formost I would like to thank my supervisor Patrik Albin for great help, accessibility and patience.

I would also like to thank my friend Christian Alfredsson and his family for their hospitality and finally my parents, sister and my friend Tarek Gad for their support in general.

Contents

1 Introduction

We start this text by looking at the most elementary BSDE and the techniques for finding a solution to it.

Throughout this text we let $(\Omega, \mathcal{F}, \{\mathcal{F}\}_{t\geq 0}, P)$ be a complete filtered probability space. A Brownian motion $B(t)$ is defined on this space and $\{\mathcal{F}\}_{t\geq 0}$ is the natural filtration of $B(t)$ augmented by all the P-null sets.

A BSDE can be thought of as the stochastic analogue to an ODE with a terminal value condition. When solving an ODE one often ends up with some integral in the solution that should be computed. In the ODE case the integrand is deterministic and we can hence integrate in any direction in time.

1.1 Simplest case

However, in the stochastic case, one immediately encounters obstacles even in the most simple cases. For example, suppose we want to find some $\mathcal{F}_t$-adapted stochastic process $X(t)$ on the time interval $[0, T]$ such that $dX(t) = 0$ and $X(T) = \xi$ where ξ is some $\mathcal{F}_T$-measurable square-integrable random variable.

One might simply put $X(t) = \xi$ which is the only possible solution if the SDE is to be solved in the Itô fashion. The problem is however that we are only given that ξ is $\mathcal{F}_T$-adapted and not necessarily $\mathcal{F}_t$-adapted. Hence, when viewed as an Itô SDE, the problem does not have a solution in general.

As a first step to overcome this problem, we define $X(t)$ as

$$X(t) = \mathbb{E}[\xi|\mathcal{F}_t], \quad t \in [0, T]$$

Defined like this, $X(t)$ is $\mathcal{F}_t$-adapted and also a square-integrable martingale. That $X(t)$ is square-integrable is relatively clear so we only show that $X(t)$ is a martingale. So if $s < t$,

$$\begin{aligned}\mathbb{E}[X(t)|\mathcal{F}_s] &= \mathbb{E}[\mathbb{E}[\xi|\mathcal{F}_t]|\mathcal{F}_s] \\ &= \mathbb{E}[\xi|\mathcal{F}_s] \\ &= X(s).\end{aligned}$$

Now that we have established these properties of $X(t)$ we can use the following:

Theorem 1.1 (Martingale Representation Theorem)**.** *Let $X(t)$ be a square integrable local martingale adapted to the filtration $\mathcal{F}_t$. Then there exists a unique square integrable $\mathcal{F}_t$-adapted random process Z such that*

$$X(t) = X(0) + \int_0^t Z(s)dB(s).$$

So, we can now rewrite $X(t)$ as

$$X(t) = X(0) + \int_0^t Z(s)dB(s) \tag{1}$$

and

$$\begin{aligned} X(0) &= X(T) - \int_0^T Z(s)dB(s) \\ &= \xi - \int_0^T Z(s)dB(s). \end{aligned}$$

Rearranging a bit we get

$$X(t) = \xi - \int_t^T Z(s)dB(s). \tag{2}$$

If we rewrite (1) in differential form we get

$$\begin{cases} dX(t) &= Z(t)dB(t), \quad t \in [0,T] \\ X(T) &= \xi. \end{cases} \tag{3}$$

Thus, (2) is the solution to the BSDE (3) which is the simplest form of a BSDE. We notice that we had to add the term $Z(t)dB(t)$ in order to find an adapted solution.

It might be worth noting that if ξ is constant $Z(t)$ would be zero and we would get the ordinary ODE-form that we considered first.

1.2 Comments

The martingale representation theorem only guarantees the existence of a unique process Z but tells us nothing about its behaviour. However, in some cases it is possible to find expressions for Z. A paper that discusses these questions is [5].

2 Existence and Uniqueness of Solutions

2.1 A more general equation

In this section we will consider BSDEs having the following form:

$$\begin{cases} dX(t) &= f(X(t), Z(t))dt + Z(t)dB(t), \quad t \in [0, T] \\ X(T) &= \xi. \end{cases} \tag{4}$$

We will see that the solution to this kind of equation always consist of a unique pair of adapted processes $(X(t), Z(t))$. This differs from the deterministic case

$$\begin{cases} dx(t) &= f(x(t), z(t))dt, \quad t \in [0, T] \\ x(T) &= C \end{cases} \tag{5}$$

which can in general have infinitely many solutions.

So we continue now by finding solutions to (4). We will again use essentially the same method as we used in the previous section.

Equation (4) is basically equation (3) with an additional function $f(x, z)$ that describes the typical evolution of the system. We will assume that this function is Lipschitz continuous.

For (3) we defined $X(t)$ as

$$X(t) = \mathbb{E}[\xi | \mathcal{F}_t], \quad t \in [0, T].$$

With the function $f(x, z)$ added it seems reasonable that we shall define $X(t)$ as

$$X(t) = \mathbb{E}[\xi - \int_t^T f(s)ds | \mathcal{F}_t], \quad t \in [0, T] \tag{6}$$

where $f(s)$ is a shorthand notation for $f(x(s), z(s))$ where $(x(s), z(s))$ are two square integrable $\mathcal{F}_s$-adapted processes.

However, $X(t)$ is not a martingale but

$$M(t) = \mathbb{E}[\xi - \int_0^T f(s)ds | \mathcal{F}_t], \quad t \in [0, T] \tag{7}$$

is. Moreover we see that $M(0) = X(0)$.

We once again use the Martingale Representation Theorem to find a unique $Z(t)$ such that

$$M(t) = M(0) + \int_0^t Z(s)dB(s), \quad t \in [0, T]. \tag{8}$$

Another useful observation is that

$$\xi - \int_0^T f(s)ds = M(T) = X(0) + \int_0^T Z(s)dB(s). \tag{9}$$

Combining these relations we obtain

$$\begin{aligned}
X(t) &= M(t) + \int_0^t f(s)ds \qquad (10)\\
&= X(0) + \int_0^t Z(s)dB(s) + \int_0^t f(s)ds \qquad (11)\\
&= \xi - \int_0^T Z(s)dB(s) - \int_0^T f(s)ds + \int_0^t Z(s)dB(s) + \int_0^t f(s)ds\\
&= \xi - \int_t^T f(s)ds - \int_t^T Z(s)dB(s). \qquad (12)
\end{aligned}$$

Thus we have found an adapted solution $(X(t), Z(t))$ to the equation

$$\begin{cases} dX(t) &= f(x(t), z(t))dt + Z(t)dB(t), \quad t \in [0, T] \\ X(T) &= \xi. \end{cases} \tag{13}$$

The reason that we used the processes $x(s)$ and $z(s)$ instead of $X(s)$ and $Z(s)$ is that we will use a contractive map to prove the uniqueness of the solutions and the reasoning that led to (13) can be viewed as a mapping of $(x(s), z(s))$ to $(X(s), Z(s))$

We now turn to the proof that this solution is unique. The typical proof of this (that we will give later) is relatively mathematical and it may not be so easy to see what is going on. Therefore before giving that proof we will try to discuss what is happening without to much mathematical jargon.

We mentioned earlier that the deterministic analogue to (4) in general had infinitely many solutions. For example, if $f(x, z) = x + z$, then from (5) we have

$$x(t) = \int_0^t (x(s) + z(s))ds + D, \tag{14}$$

where D is some appropriate constant. Then any $x(s)$ and $z(s)$ such that $\int_0^T (x(s) + z(s))ds = C - D$ will be solutions and are infinitely many.

This is as we have mentioned different in the BSDE case. In the BSDE (4) we still have one equation and two unknown variables, yet we, as we shall see, obtain two unique solutions.

If we accept that we are looking for solutions that satisfy (4), then we can see the following:

From (10) above we see that

$$X(t) - \int_0^t f(s)ds = M(t) \tag{15}$$

where $M(t)$ is a $\mathcal{F}_t$-adapted martingale. So if we now use the assumption that we are looking at solutions that satisfy (4) we can write this as

$$X(t) - \int_0^t f(X(s), Z(s))ds = M(t). \tag{16}$$

It might be tempting to argue that we could have infinitely many pairs (X, Z) that would make (16) into a martingale. However, we also require that Z is such that (8) holds.

Intuitively and a bit informally perhaps, we argue that pushing X to much in one direction will make Z go to much in another direction (we must change Z to still make (16) into a martingale) and thus fail to make (8) hold.

Thus, one might expect some equilibrium where X and Z are such that both (8) and (16) are satisfied.

We will now prove that the solutions are unique and we start by recalling a definition and a theorem.

Definition 2.1. *Let R be a metric space with metric d. If ψ is a mapping from R into R and there is a number $c < 1$ such that*

$$d(\psi(x), \psi(y)) \leq c\, d(x, y)$$

for all $x, y \in R$, then ψ is said to be a contraction of R into R.

Theorem 2.1. *If R is a complete metric space, and if ψ is a contraction of R into R, then there exists one and only one $x \in R$ such that $\psi(x) = x$.*

A proof of Theorem 2.1. can be found in [6].

We recall that both $X(t)$ and $Z(t)$ are square integrable, $\mathcal{F}_t$-adapted processes. We will denote the space of such pairs by $N[0, T]$. It can be shown that the norm defined by

$$||(X(\cdot), Z(\cdot))||_{N[0,T]} = \left(\mathbb{E}_{\sup t\in[0,T]} |X(t)|^2 + \mathbb{E} \int_0^T |Z(t)|^2 dt \right)^{1/2} \tag{17}$$

makes $N[0,T]$ into a complete metric space.

We mentioned above that we require f to be Lipschitz continuous. By this we mean that there is some constant $L > 0$ such that

$$|f(x,z) - f(\bar{x},\bar{z})| \leq L\big(|x-\bar{x}| + |z-\bar{z}|\big) \tag{18}$$

for all $x, z, \bar{x}, \bar{z}$ such that $(x,z), (\bar{x},\bar{z}) \in N[0,T]$.

Before we begin our proof a comment is in order. In the proof we will consider the space $N[t,T]$ for some $t \in [0,T]$. We need the low end of the interval to be flexible in order to construct a contraction.

So now, if $(\bar{X}(\cdot), \bar{Z}(\cdot))$ is the solution that corresponds to $(\bar{x}(\cdot), \bar{z}(\cdot))$ in the above sense, then by Itô's formula applied to $(X(t) - \bar{X}(t))^2$ we have

$$\begin{aligned} d((X(t)-\bar{X}(t))^2) &= 2(X(t)-\bar{X}(t))d(X(t)-\bar{X}(t)) + d(X(t)-\bar{X}(t))d(X(t)-\bar{X}(t)) \\ &= 2(X(t)-\bar{X}(t))d(X(t)-\bar{X}(t)) + (Z(t)-\bar{Z}(t))^2 dt. \end{aligned} \tag{19}$$

Hence

$$(X(t)-\bar{X}(t))^2 = -2\int_t^T (X(s)-\bar{X}(s))d(X(s)-\bar{X}(s)) - \int_t^T (Z(s)-\bar{Z}(s))^2 ds. \tag{20}$$

Thus, we obtain

$$\begin{aligned} &\mathbb{E}(X(t)-\bar{X}(t))^2 + \mathbb{E}\int_t^T (Z(s)-\bar{Z}(s))^2 ds = -2\mathbb{E}\int_t^T (X(s)-\bar{X}(s))d(X(s)-\bar{X}(s)) \\ &= -2\mathbb{E}\int_t^T (X(s)-\bar{X}(s))\Big[(f(x(s),z(s)) - f(\bar{x}(s),\bar{z}(s)))ds + (Z(s)-\bar{Z}(s))dB(s)\Big]. \end{aligned} \tag{21}$$

The expectation of the Brownian integral is zero, so

$$\begin{aligned} &\mathbb{E}(X(t)-\bar{X}(t))^2 + \mathbb{E}\int_t^T (Z(s)-\bar{Z}(s))^2 ds \\ &= -2\mathbb{E}\int_t^T (X(s)-\bar{X}(s))(f(x(s),z(s)) - f(\bar{x}(s),\bar{z}(s)))ds. \end{aligned} \tag{22}$$

Then by using the Cauchy-Schwarz inequality, the Lipschitz condition and the two inequalities

$(x+y)^2 \le 2x^2+2y^2$ and $\sqrt{a+b} \le \sqrt{a}+\sqrt{b}$ for $a,b \ge 0$ we see

$$
\begin{aligned}
&\mathbb{E}(X(t)-\bar{X}(t))^2+\mathbb{E}\int_t^T (Z(s)-\bar{Z}(s))^2 ds \\
\le& 2\int_t^T \sqrt{\mathbb{E}(X(s)-\bar{X}(s))^2}\sqrt{\mathbb{E}(f(x(s),z(s))-f(\bar{x}(s),\bar{z}(s)))^2}ds \\
\le& 2L\int_t^T \sqrt{\mathbb{E}(X(s)-\bar{X}(s))^2}\sqrt{\mathbb{E}(|x(s)-\bar{x}(s)|+|z(s)-\bar{z}(s)|)^2}ds \\
\le& 2L\int_t^T \sqrt{\mathbb{E}(X(s)-\bar{X}(s))^2}\sqrt{\mathbb{E}(2|x(s)-\bar{x}(s)|^2+2|z(s)-\bar{z}(s)|^2)}ds \\
\le& \sqrt{8}L\int_t^T \sqrt{\mathbb{E}(X(s)-\bar{X}(s))^2}\Big(\sqrt{\mathbb{E}(x(s)-\bar{x}(s))^2}+\sqrt{\mathbb{E}(z(s)-\bar{z}(s))^2}\Big)ds. \qquad (23)
\end{aligned}
$$

We introduce the following quantities:

$$
\begin{cases}
\alpha(t)=\sqrt{\mathbb{E}(X(s)-\bar{X}(s))^2}, \\
\beta(t)=\sqrt{\mathbb{E}(x(s)-\bar{x}(s))^2}+\sqrt{\mathbb{E}(z(s)-\bar{z}(s))^2},
\end{cases}
$$

and rewrite (23) as

$$
\alpha(t)^2+\mathbb{E}\int_t^T (Z(s)-\bar{Z}(s))^2 ds \le \sqrt{8}L\int_t^T \alpha(s)\beta(s)ds. \qquad (24)
$$

The following which can be considered a version of Grönwalls Lemma is proved in the Appendix. The proof and Lemma both appear in [3].

Lemma 2.1. *Let* (24) *hold. Then,*

$$
\alpha(t)^2+\mathbb{E}\int_t^T (Z(s)-\bar{Z}(s))^2 ds \le 2L^2\Big(\int_t^T \beta(s)ds\Big)^2. \qquad (25)
$$

Using the Lemma, the inequality $(x+y)^2 = 2x^2 + 2y^2$ and Hölder's inequality we get

$$\begin{aligned}
&\mathbb{E}(X(t)-\bar{X}(t))^2 + \mathbb{E}\int_t^T (Z(s)-\bar{Z}(s))^2 ds \\
\leq & 2L^2\left[\int_t^T \big(\sqrt{\mathbb{E}(x(s)-\bar{x}(s))^2} + \sqrt{\mathbb{E}(z(s)-\bar{z}(s))^2}\big)ds\right]^2 \\
\leq & 4L^2\Big(\int_t^T \sqrt{\mathbb{E}(x(s)-\bar{x}(s))^2}ds\Big)^2 + 4L^2\Big(\int_t^T \sqrt{\mathbb{E}(z(s)-\bar{z}(s))^2}ds\Big)^2 \\
\leq & 4L^2(T-t)^2 \sup_{s\in[t,T]} \mathbb{E}(x(s)-\bar{x}(s))^2 + 4L^2 \int_t^T 1^2 ds \int_t^T \mathbb{E}(z(s)-\bar{z}(s))^2 ds \\
\leq & 4L^2(T-t)^2 \sup_{s\in[t,T]} \mathbb{E}(x(s)-\bar{x}(s))^2 + 4L^2(T-t) \int_t^T \mathbb{E}(z(s)-\bar{z}(s))^2 ds \\
\leq & 4L^2 \max\{(T-t)^2, (T-t)\} ||(x(\cdot), z(\cdot)) - (\bar{x}(\cdot), \bar{z}(\cdot))||^2_{N[t,T]}.
\end{aligned} \tag{26}$$

According to (10) we have

$$X(t) - \bar{X}(t) = M(t) - \bar{M}(t) + \int_0^t (f(s) - \bar{f}(s))ds. \tag{27}$$

So for $s \in [t,T]$,

$$X(s) - \bar{X}(s) - (X(t) - \bar{X}(t)) = M(s) - \bar{M}(s) - (M(t) - \bar{M}(t)) + \int_t^s (f(r) - \bar{f}(r))dr, \tag{28}$$

where $M(s) - \bar{M}(s) - (M(t) - \bar{M}(t))$ is a martingale. Further, using the inequality $(x+y)^2 \leq 2x^2 + 2y^2$

$$\begin{aligned}
&\mathbb{E}(\sup_{s\in[t,T]} (X(s) - \bar{X}(s))^2) \\
&\leq 2\mathbb{E}(\sup_{s\in[t,T]} (X(s) - \bar{X}(s) - (X(t) - \bar{X}(t)))^2) + 2\mathbb{E}((X(t) - \bar{X}(t))^2)
\end{aligned} \tag{29}$$

where the second term of the R.H.S. of (29) is less or equal to $4L^2(\int_t^T \beta(s)ds)^2$ while

$$\begin{aligned}
&2\mathbb{E}(\sup_{s\in[t,T]} (X(s) - \bar{X}(s) - (X(t) - \bar{X}(t)))^2) \\
&\leq 4\mathbb{E}(\sup_{s\in[t,T]} (M(s) - \bar{M}(s) - (M(t) - \bar{M}(t)))^2) + 4\mathbb{E}(\sup_{s\in[t,T]} (\int_t^s (f(r) - \bar{f}(r))dr)^2).
\end{aligned} \tag{30}$$

We will use the following which can be found in [1] on page 75:

Theorem 2.2 (Doob's Maximal Inequality)**.** *Pick a constant $p > 1$. For a right-continuous martingale or a non-negative submartingale $\{X(t)\}_{t\in[0,T]}$ such that the process $\{|X(t)|^p\}_{t\in[0,T]}$ is integrable, we have*

$$\mathbb{E}\big(\sup_{0\le t\le T}|X(t)|^p\big) \le q^p\mathbb{E}(|X(t)|^p), \tag{31}$$

where $\frac{1}{p}+\frac{1}{q}=1$.

So, using Doob's maximal inequality on the first term of (30) and Hölder's inequality on the second, we get

$$\begin{aligned}
&2\mathbb{E}(\sup_{s\in[t,T]}(X(s)-\bar{X}(s)-(X(t)-\bar{X}(t)))^2)\\
\le&16\mathbb{E}((M(T)-\bar{M}(T)-(M(t)-\bar{M}(t)))^2)+4\mathbb{E}\sup_{s\in[t,T]}\int_t^s 1^2dr\int_t^s(f(r)-\bar{f}(r))^2dr\\
\le&16\mathbb{E}\big[(X(T)-\bar{X}(T)-(X(t)-\bar{X}(t))-\int_t^T(f(s)-\bar{f}(s))ds)^2\big]+4(T-t)\mathbb{E}\int_t^T(f(r)-\bar{f}(r))^2dr\\
\le&32\mathbb{E}\big[(X(T)-\bar{X}(T)-(X(t)-\bar{X}(t)))^2\big]+32\mathbb{E}\big[(\int_t^T(f(s)-\bar{f}(s))ds)^2\big]+4(T-t)\mathbb{E}\int_t^T(f(r)-\bar{f}(r))^2dr.
\end{aligned} \tag{32}$$

We use the fact that $X(T)=\bar{X}(T)=\xi$ and Hölder's inequality to see

$$\begin{aligned}
&32\mathbb{E}\big[(X(T)-\bar{X}(T)-(X(t)-\bar{X}(t)))^2\big]+32\mathbb{E}\big[(\int_t^T(f(s)-\bar{f}(s))ds)^2\big]+4(T-t)\mathbb{E}\int_t^T(f(r)-\bar{f}(r))^2dr\\
&\quad\le 32\mathbb{E}\big[(X(t)-\bar{X}(t)))^2\big]+32(T-t)\mathbb{E}\big[\int_t^T(f(s)-\bar{f}(s))^2ds\big]+4(T-t)\mathbb{E}\int_t^T(f(r)-\bar{f}(r))^2dr.
\end{aligned} \tag{33}$$

The first term in the R.H.S. is smaller than $64L^2(\int_t^T\beta(s)ds)^2$ while for the second and third term we use the Lipschitz condition and the inequality $(x+y)^2\le 2x^2+2y^2$. That is,

$$\begin{aligned}
&32(T-t)\mathbb{E}\big[\int_t^T(f(s)-\bar{f}(s))^2ds\big]+4(T-t)\mathbb{E}\int_t^T(f(r)-\bar{f}(r))^2dr\\
\le&36(T-t)L^2\mathbb{E}\int_t^T(2(x(s)-\bar{x}(s))^2+2(z(s)-\bar{z}(s))^2)ds\\
\le&72(T-t)^2L^2\mathbb{E}\sup_{s\in[t,T]}(x(s)-\bar{x}(s))^2+72(T-t)L^2\mathbb{E}\int_t^T(z(s)-\bar{z}(s))^2ds\\
\le&72L^2\max\{(T-t)^2,(T-t)\}||(x(\cdot),z(\cdot))-(\bar{x}(\cdot),\bar{z}(\cdot))||^2_{N[t,T]}.
\end{aligned} \tag{34}$$

So from (26) we see that $\mathbb{E}\int_t^T (Z(s) - \bar{Z}(s))^2 ds \leq 4L^2 \max\{(T-t)^2, (T-t)\}||(x(\cdot), z(\cdot)) - (\bar{x}(\cdot), \bar{z}(\cdot))||^2_{N[t,T]}$. Similarly, from what we just have done we have $\mathbb{E}(\sup_{s\in[t,T]}(X(s)-\bar{X}(s))^2) \leq (4L^2 + 64L^2 + 72L^2)\max\{(T-t)^2, (T-t)\}||(x(\cdot), z(\cdot)) - (\bar{x}(\cdot), \bar{z}(\cdot))||^2_{N[t,T]}$. Thus we have the inequality

$$||(X(\cdot), Z(\cdot)) - (\bar{X}(\cdot), \bar{Z}(\cdot))||^2_{N[t,T]} \leq 144 \max\{(T-t)^2, (T-t)\}||(x(\cdot), z(\cdot)) - (\bar{x}(\cdot), \bar{z}(\cdot))||^2_{N[t,T]} \tag{35}$$

and since L and T are fixed we can choose t small enough to make $144\max\{(T-t)^2, (T-t)\} < 1$.

So, now we have finally found a contraction on $N[t,T]$ and thus by Theorem 2.1 also a unique solution to (4).

2.2 Comments

We hope that the discussion preceding the proof gives the reader some sort of visual assurance that the solutions are unique.

The author initially didn't intend to display a full proof of the uniqueness since the method used seemed well established and appears in numerous sources. However, most texts display very short outlines of the proof as did the one the author used when trying to understand it([3]).

When the author tried to fill in the gaps for himself he found many steps relatively hard and consulted his supervisor. It turned out that many details of the proof given in [3] was not stated and the authors also referred to Doob's Inequality instead of Doob's Maximal Inequality.

Thus we give a corrected version of the proof outlined in [3] and we have taken the time to fill in almost every detail of the calculations. We hope this will be of help to less experienced readers.

The author owes much of the presentation of the proof to his supervisor.

3 Some additional comments

As we have seen, the role of the random variable $Z(t)$ is to correct the behaviour of $X(t)$ in such a way that $X(T) = \xi$. There are of course many situations in real life where we want some goal to be met at a certain time in the future. Often random events interfere with our original plans and we need to take action along the way to make sure that we are on the best possible path to achieve our goal.

The theory of stochastic controls is often the mathematical way to deal with these questions. One of the cornerstones of stochastic controls is the stochastic maximum principle which is basically the stochastic analogue to the Pontryagin maximum principle. BSDEs first showed up as the adjoint equations in the stochastic maximum principle. Although this happened around 1970 a proof of uniqueness of solutions to more general equations were not given until 1990 in the paper [4] by S.Peng and E.Pardoux.

Since then BSDEs have been applied to a numerous branches of mathematical disciplines ranging from fluid dynamics to mathematical finance.

Any reader who wishes to continue studying BSDEs and their applications will most likely have to learn about stochastic controls. A good introductory book which discusses deterministic optimal controls is [2] and a more demanding book on stochastic controls is [7].

4 References

[1] Albin, M.P. *Graduate Course on Stochastic Differential Equations.* Lecture Notes, Chalmers University of Technology, 2009.

[2] Kirk, D. E. *Optimal Control Theory.* Dover, 2004.

[3] Ma, J. Yong, J. *Forward-Backward Stochastic Differential Equations and their Applications.* Springer, 1999.

[4] Peng, S. Pardoux, E. *Adapted Solution of a Backward Stochastic Differential Equation Systems and Controls, Letter 14*, 1990.

[5] Rudvik, A. *Stochastic Integral Representation of Functionals with Option-like Structure.* Master thesis, Chalmers University of Technology, 2007.

[6] Rudin, W. *Principles of Mathematical Analysis 3e.* McGraw-Hill, 1976.

[7] Yong, J. Zhou, X.Y. *Stochastic Controls.* Springer, 1999.

5 Appendix: A proof of Lemma 2.1

The R.H.S of equation (24) is $\sqrt{8}L\int_t^T \alpha(s)\beta(s)ds$. If we let $\theta(t) = \int_t^T \alpha(s)\beta(s)ds$ and recall the differentiation rule $\frac{d}{dt}\int_t^T f(s)ds = -f(t)$. Then since all the terms involved in (24) are positive we have

$$\theta'(t) = -\alpha(t)\beta(t) \geq= -\beta(t)\sqrt{\sqrt{8}L\theta(t)}. \tag{36}$$

We then use the following:

$$\begin{aligned}(\sqrt{\theta(t)})' &= \frac{\theta'(t)}{2\sqrt{\theta(t)}} \\ &\geq -\beta(t)\frac{\sqrt{2\sqrt{2}L\theta(t)}}{\sqrt{4\theta(t)}} \\ &= -\beta(t)\sqrt{\frac{L}{\sqrt{2}}}.\end{aligned} \tag{37}$$

Then, integrating the L.H.S. and R.H.S. of (37) from t to T we get

$$\sqrt{\theta(T)} - \sqrt{\theta(t)} \geq -\sqrt{\frac{L}{\sqrt{2}}}\int_t^T \beta(s)ds. \tag{38}$$

But $\theta(T) = 0$ so

$$-\sqrt{\theta(t)} \geq -\sqrt{\frac{L}{\sqrt{2}}}\int_t^T \beta(s)ds. \tag{39}$$

Squaring both sides of (39) then gives

$$\theta(t) \leq \frac{L}{\sqrt{2}}\Big(\int_t^T \beta(s)ds\Big)^2 \tag{40}$$

and we arrive at the desired inequality

$$\alpha(t)^2 + \mathbb{E}\int_t^T (Z(s) - \bar{Z}(s))^2 ds \leq 2L^2\Big(\int_t^T \beta(s)ds\Big)^2. \tag{41}$$