Z. Wahrscheinlichkeitstheorie verw. Gebiete
68, 315–329 (1985)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

Mean Stochastic Comparison of Diffusions

Bruce Hajek*

Department of Electrical and Computer Engineering and the Coordinated Science Laboratory
University of Illinois at Urbana-Champaign 1101 W. Springfield, Urbana, IL 61801, USA

Summary. Stochastic bounds are derived for one dimensional diffusions (and somewhat more general random processes) by dominating one process pathwise by a convex combination of other processes. The method permits comparison of diffusions with different diffusion coefficients. One interpretation of the bounds is that an optimal control is identified for certain diffusions with controlled drift and diffusion coefficients, when the reward function is convex. An example is given to show how the bounds and the Liapunov function technique can be applied to yield bounds for multidimensional diffusions.

I. Introduction

A well known comparison theorem for diffusions may be stated as follows (see [10, 3–5]) for proof and remarkable refinements):

Theorem 1.1. *Suppose that σ, b_1 and b_2 are Lipschitz continuous functions on $\mathbb{R}$ – i.e., for some constant K*

$$|\sigma(\theta)-\sigma(\theta')|+|b_1(\theta)-b_1(\theta')|+|b_2(\theta)-b_2(\theta')|\leqq K|\theta-\theta'|$$

and suppose that X^i, $i=1,2$ are the (pathwise unique) solutions to the stochastic differential equations driven by a Wiener process w

$$dX_t^i=\sigma(X_t^i)dw_t+b_i(X_t^i)dt;\qquad X^i(0)\ \text{given.}$$

Then the conditions $b_1\geqq b_2$ and $X^1(0)\geqq X^2(0)$ imply that

$$X_t^1\geqq X_t^2\qquad \text{for all } t,\ P \text{ a.s.}$$

* This work was supported by the Office of Naval Research under Contract N00014-82-K-0359 and the U.S. Army Research Office under Contract DAAG29-82-K-0091 (administered through the University of California at Berkeley).

Corollary 1.2. *Let x and y be one dimensional diffusions (on possibly distinct probability spaces) with respective generators*

$$\frac{\sigma^2(\theta)}{2}\frac{\partial^2}{\partial\theta^2}+b_1(\theta)\frac{\partial}{\partial\theta} \quad and \quad \frac{\sigma^2(\theta)}{2}\frac{\partial^2}{\partial\theta^2}+b_2(\theta)\frac{\partial}{\partial\theta},$$

where σ, b_1 and b_2 are Lipschitz continuous. Then $b_1 \geqq b_2$ and $x_0 \geqq y_0$ imply that

$$P[x_t \geqq c] \leqq P[y_t \geqq c] \qquad t \geqq 0,\ c \in \mathbb{R}.$$

Proof. The random processes x and X^1 have the same distribution – we denote this by $x \sim X^1$. Similarly $y \sim X^2$. Thus

$$P[x_t \geqq c] = P[X_t^1 \geqq c] \geqq P[X_t^2 \geqq c] = P[y_t \geqq c]. \quad \square$$

Results of this type have been applied to prove stability theorems for diffusions z in $\mathbb{R}^n$ by letting x be a process of the form $x_t = p(z_t)$ (often p is called a Liapunov function [12]). A major difficulty is that the diffusion term σ in Theorem 1.1 is the same for both diffusions. Thus, in order to compare processes with different diffusion terms, one must introduce a random time change to equalize the diffusion terms (see [3, 6]). This makes the method cumbersome and it is especially difficult to obtain accurate comparisons for large times (e.g., bounds on tails of invariant measures). An alternative method we explore in this paper is to dominate one process pathwise by a convex combination of other processes.

Throughout this paper we use the usual conventions of stochastic calculus. For example, each random process is assumed to be defined on a probability space $(\Omega, \mathbb{F}, P)$ equipped with an increasing family $(\mathbb{F}_t)$ of sub-σ-algebras of $\mathbb{F}$. The random processes are assumed to be adapted and Wiener processes are assumed to be $(\mathbb{F}_t)$ martingales. When a process x is said to be a semimartingale with representation

$$dx_t = \mu_t dt + \sigma_t dw_t$$

where w_t is a Wiener process, it is understood that μ (resp. σ) is locally integrable (square integrable) with probability one, and the above is shorthand notation for

$$x_t = x_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dw_s.$$

x is an "Ito process" in the terminology of [4].

Theorem 1.3. *Let x and y be semimartingales with representations*

$$dx_t = \mu_t dt + \sigma_t dw_t$$
$$dy_t = m dt + \rho dv_t$$

where w and v are Wiener processes and m and ρ are constants. Suppose that

$$\mu_t \leqq m \quad and \quad |\sigma_t| \leqq \rho \tag{1.1}$$

and that $x_0 \leqq y_0$. *Then*

$$P[x_t \geqq c] \leqq 2P[y_t \geqq c], \tag{1.2}$$

and for any nondecreasing convex function Φ *on* $\mathbb{R}$

$$E\Phi(x_t) \leqq E\Phi(y_t). \tag{1.3}$$

Proof. By using a standard extension [3, p. 72] of the probability space that x is defined on, we can assume without loss of generality that there exists a Wiener process $\tilde{w}$ on the same probability space which is independent of (x, μ, σ, w). Let Y^i, $i=1,2$ each be defined by

$$Y_t^i = y_0 + mt + \left[\int_0^t \sigma_s dw_s + (-1)^i \int_0^t (\rho^2 - \sigma_s^2)^{1/2} d\tilde{w}_s\right].$$

For each i, the process in square brackets is a continuous martingale with quadratic variation process $\rho^2 t$, and hence it has the same distribution as ρv [see 2, p. 384]. Thus, the processes Y^1 and Y^2 each have the same distribution as y:

$$Y^i \sim y \qquad \text{for } i=1,2. \tag{1.4}$$

Now define $\bar{Y}_t = (Y_t^1 + Y_t^2)/2$. Then,

$$d\bar{Y}_t = m\,dt + \sigma_t dw_t$$

so that

$$\bar{Y}_t - x_t = y_0 - x_0 + \int_0^t (m - \mu_s)\,ds.$$

By our assumptions, the right side is nonnegative. Therefore

$$\bar{Y}_t \geqq x_t \qquad \text{for all } t,\ P \text{ a.s.} \tag{1.5}$$

So, (since $\bar{Y}_t \leqq \max\{Y_t^1, Y_t^2\}$)

$$I\{x_t \geqq c\} \leqq I\{Y_t^1 \geqq c\} + I\{Y_t^2 \geqq c\} \tag{1.6}$$

and (since Φ is nondecreasing and convex)

$$\Phi(x_t) \leqq \Phi(\bar{Y}_t) \leqq \tfrac{1}{2}(\Phi(Y_t^1) + \Phi(Y_t^2)). \tag{1.7}$$

Taking the expectation of each term in inequalities (1.6) and (1.7) and using (1.4) yields the desired conclusions (1.2) and (1.3). □

Remarks (1) The conditions of Theorem 1.3 are satisfied when $x_t \equiv 0$ and $y_t \equiv v_t$. Then equality holds in (1.2) for $c=0$ so the factor two cannot be reduced.

(2) Since y itself satisfies the conditions placed on x, y may viewed as a semimartingale with given initial value which maximizes $E\Phi(x_t)$ over all semimartingales x satisfying (1.1). In other words, Theorem 1.3 identifies a solution to a certain stochastic optimal control problem. A similar interpretation holds for Theorem 1.1 (see [3, Sect. 6.2]) and for the other theorems of this paper.

(3) If y_0 is constant then Theorem 1.3 is valid (same proof) even if x and y are defined on different probability spaces.

(4) Since (1.4) and (1.5) are statements about processes (rather than just about random variables) inequality (1.3) may be strengthened to

$$E\Phi(x) \leqq E\Phi(y)$$

for any real valued function Φ defined on the space of continuous functions $C[0,t]$ such that Φ is (strongly continuous and bounded below) nondecreasing and convex. For example.

$$E \max_{0\leqq t\leqq T} x_t \leqq E \max_{0\leqq t\leqq T} y_t.$$

Theorem 2 and Theorem 4.1 in this paper can be strengthened similarly.

(5) In the proof of Theorem 1.3 and in similar proofs later, lower case letters are used for the random processes x and y being compared while upper case letters are used for auxiliary random processes and their convex combinations.

Theorem 1.3 is generalized in the following sections by relaxing condition (1.1) so that it need only hold when x_t is sufficiently large, and by allowing nonconstant m and ρ. An application is presented in Sect. V.

II. Domination on a Half Line – Reflecting Dominating Process

We now make the following assumptions:

u, ρ and y_0 are constants with $y_0 \geqq u$.

m is a convex, Lipschitz continuous function on $[u, +\infty)$.

v is a Wiener process.

Then by Skorohod's theory [9] (see [13] for a more general theory) a pair (l, y) is (strong sense) uniquely determined by (m, ρ, u, v, y_0) and the equations

$$\begin{aligned} &dy_t = m(y_t)dt + \rho dv_t + dl_t, \\ &y_t \geqq u, \quad I\{y_t > u\}\, dl_t = 0, \ l_0 = 0, \ y_0 \text{ given}, \end{aligned} \tag{2.1}$$

l is continuous, nondecreasing.

The process y is a Markov diffusion, and if the boundary u is reachable it is instantaneously reflecting. In the special case that m is a constant function we can write (l, y) explicitly as

$$y_t = y_0 + mt + \rho v_t + l_t$$

$$l_t = \max_{0\leqq s\leqq t} (u - y_0 - ms - \rho v_s)^+.$$

Theorem 2. *Suppose that x is a semimartingale with representation*

$$dx_t = \mu_t dt + \sigma_t dw_t + d\lambda_t$$

(here λ is a sample continuous process with finite variation over each bounded interval, a.s. and $\lambda_0=0$) and let (y, l) satisfy (2.1). *Suppose that*

$$\mu_t \leqq m(x_t) \quad \text{and} \quad |\sigma_t| \leqq \rho \qquad \text{whenever } x_t \geqq u,$$
$$I\{x_t > u\}\, d\lambda_t \leqq 0, \quad \text{and} \quad x_0 \leqq y_0.$$

Then

$$P[x_t \geqq c] \leqq 2P[y_t \geqq c]$$

and for any nondecreasing convex function Φ on $\mathbb{R}$

$$E\Phi(x_t) \leqq E\Phi(y_t).$$

Proof. Let $\tilde{w}$ be a Wiener process independent of $(x, \mu, \sigma, w, \lambda)$, set

$$\hat{\sigma}_t = \begin{cases} -\rho & \text{if } \sigma_t \leqq -\rho \\ \sigma_t & \text{if } -\rho < \sigma_t \leqq \rho \\ \rho & \text{if } \sigma_t \geqq \rho \end{cases}$$

and let (Y^i, l^i) be defined by (for $i=1, 2$)

$$\begin{aligned} &dY_t^i = m(Y_t^i)\,dt + [\hat{\sigma}_t\, dw_t + (-1)^i(\rho^2 - \hat{\sigma}_t^2)^{1/2}\, d\tilde{w}_t] + dl_t^i, \\ &Y_t^i \geqq u, \ I\{Y_t^i > u\}\, dl_t^i = 0, \ l_0^i = 0, \ Y_0^i = y_0, \\ &l^i \text{ is continuous, nondecreasing.} \end{aligned} \tag{2.2}$$

For each i, the term in square brackets in (2.2) is the differential of a Wiener process multiplied by ρ, so Y^1 and Y^2 each have the same distribution as y by uniqueness in law of the stochastic differential equations.

Define

$$\bar{Y}_t = \frac{Y_t^1 + Y_t^2}{2} \quad \text{and} \quad \bar{m}_t = \frac{m(Y_t^1) + m(Y_t^2)}{2}.$$

Then

$$d\bar{Y}_t = \bar{m}_t\, dt + \hat{\sigma}_t\, dw_t + \tfrac{1}{2}(dl_t^1 + dl_t^2)$$

and so

$$d\langle x - \bar{Y}, x - \bar{Y}\rangle_s = (\sigma_s - \hat{\sigma}_s)^2\, ds.$$

Now $\sigma_s = \hat{\sigma}_s$ whenever $x_s \geqq u$ (and hence whenever $x_s \geqq \bar{Y}_s$) so

$$\int_0^t I\{x_s - \bar{Y}_s \geqq 0\}\, d\langle x - \bar{Y}, x - \bar{Y}\rangle_s = 0. \tag{2.3}$$

Using (6.2) to reexpress the left side of (2.3), we conclude that

$$\int_0^\infty L_t^a(x - \bar{Y})\, da = 0$$

which, by the right continuity of L_t^a in a, implies that

$$L_t^0(x - \bar{Y}) = 0. \tag{2.4}$$

By Itô's formula (6.1) and by (2.4),

$$(x_t-\bar{Y}_t)^+=\int_0^t I\{x_s>\bar{Y}_s\}\,d(x_s-\bar{Y}_s).$$

Then, using the fact that $x_s\geq u$ whenever $x_s>\bar{Y}_s$, and using the fact (due to the convexity of m) that $\bar{m}_t\geq m(\bar{Y}_t)$, we have

$$\begin{aligned}(x_t-\bar{Y}_t)^+&=\int_0^t I\{x_s>\bar{Y}_s\}(\mu_s-\bar{m}_s)\,ds+\int_0^t I\{x_s>\bar{Y}_s\}\,d(\lambda_s-\tfrac{1}{2}(l_s^1+l_s^2))\\ &\leqq\int_0^t I\{x_s>\bar{Y}_s\}(m(x_s)-m(\bar{Y}_s))\,ds\\ &\leqq k\int_0^t (x_s-\bar{Y}_s)^+\,ds\end{aligned}$$

where k is the Lipschitz constant of m. Therefore $(x_t-\bar{Y}_t)^+=0$, or equivalently

$$\bar{Y}_t\geqq x_t \quad \text{for all } t, P \text{ a.s.}$$

The remainder of the proof follows that of Theorem 1.3.

Remark. If $u=0$ and $m(\theta)\equiv m$ then the transition density of y is given in terms of the standard Normal distribution function $\mathbb{N}$ by

$$p_t(\theta\,|\,y_0)=\frac{\partial}{\partial\theta}\left\{\mathbb{N}\left[\frac{\theta-y_0-mt}{(\rho t)^{1/2}}\right]-\exp\left[\frac{2m\theta}{\rho}\right]\mathbb{N}\left[\frac{-(\theta+y_0+\beta t)}{(\rho t)^{1/2}}\right]\right\} \tag{2.5}$$

and, if $m<0$, the invariant density is

$$p_{+\infty}(\theta)=\gamma e^{-\gamma\theta} \quad \text{where } \gamma=\frac{\rho^2}{-2m},$$

Thus under the conditions of Theorem 2, if $m<0$ then

$$\limsup_{t\to\infty} E\Phi(x_t)\leqq\int_0^\infty \gamma e^{-\gamma a}\Phi(a+u)\,da$$

and

$$\limsup_{t\to\infty} P[x_t\geqq c]\leqq 2e^{-\gamma(c-u)}.$$

Also, for example, if $P[x_0\geqq c]\leqq e^{-\gamma(c-u)}$ for all c then (take $x_0=y_0$) $P[x_t\geqq c]\leqq 2e^{-\gamma(c-u)}$ for all c and t.

III. Variable Diffusion and Zero Drift

An alternative technique, namely random time transformation, is applied in this section only. The technique is best suited to the case that the processes have zero drift.

Theorem 3. *Let x be a continuous martingale with representation*

$$x_t = x_0 + \int_0^t \sigma_s dw_s$$

such that for some Lipschitz continuous function ρ on $\mathbb{R}$

$$|\sigma_s| \leqq \rho(x_s).$$

and let y be the unique solution to the stochastic differential equation

$$y_t = x_0 + \int_0^t \rho(y_s) dw_s. \tag{3.1}$$

Then for any convex function Φ and any $t \geqq 0$

$$E\Phi(x_t) \leqq E\Phi(y_t). \tag{3.2}$$

Proof. Define a nondecreasing process δ by

$$\delta(t) = \int_0^t \frac{\sigma_u^2}{\rho(x_u)^2} du$$

with the convention that the integrand is one when $\sigma_u = \rho(x_u) = 0$. Then $\delta(t) \leqq t$ and, on the other hand, we can assume without loss of generality that for some constants a and b with $b > 0$,

$$\delta(t) \geqq a + bt \qquad \text{for all } t \geqq 0.$$

(If this is not already true, simply modify x by choosing $\sigma_u^2 = \rho(x_u)^2$ for all u larger than a given t for which inequality (3.2) is to be established.) Next, for each $s \geqq 0$ define the stopping time

$$\tau(s) = \inf\{t \colon \delta(t) > s\}.$$

Then

$$s \leqq \tau(s) \leqq (s-a)/b. \tag{3.3}$$

Define a random process z by

$$z_s = x_{\tau(s)}; \tag{3.4}$$

then, since x and δ have the same intervals of constancy,

$$z_{\delta(t)} = x_t. \tag{3.5}$$

Next, applying Lebesgue's formula for transformation from Lebesgue to Stieltjes integrals [1, p. 120]

$$\begin{aligned} z_s^2 - \int_0^s \rho(z_t)^2 dt &= z_s^2 - \int_0^{\tau(s)} \rho(z_{\delta(t)})^2 d\delta_t \\ &= x_{\tau(s)}^2 - \int_0^{\tau(s)} \sigma_t^2 dt. \end{aligned} \tag{3.6}$$

By (3.3), (3.4) and (3.6), the optional sampling theorem implies that

$$(z_t) \quad \text{and} \quad \left(z_t^2 - \int_0^t \rho(z_s)^2\, ds\right) \text{ are continuous martingales}$$

Thus z is a weak solution to the equation defining y, and under our assumptions the solution is unique. Thus

$$z \sim y \tag{3.7}$$

Now, if Φ is a convex function then $\Phi(z_s)$ is a submartingale. Therefore since $\delta(t) \leqq t$ the optional sampling theorem implies that

$$E\Phi(z_t) \geqq E\Phi(z_{\delta(t)}).$$

From (3.5) and (3.7) we conclude that

$$E\Phi(y_t) \geqq E\Phi(x_t). \quad \square$$

IV. Variable Diffusion and Drift

When the dominating diffusion y has a nonconstant diffusion coefficient it may no longer be possible to dominate x pathwise by *finite* convex combinations of processes with the same distribution as y (see remark at the end of the section). However, an infinite convex combination suffices:

Theorem 4.1. *Suppose that m and ρ are each convex Lipschitz continuous functions on $\mathbb{R}$, and suppose that μ and σ are Borel measurable functions on $\mathbb{R} \times \mathbb{R}_+$ such that for some constant K and all θ, θ' in $\mathbb{R}$ and $t \geqq 0$*

$$|\mu(\theta, t) - \mu(\theta', t)| + |\sigma(\theta, t) - \sigma(\theta', t)| \leqq K|\theta - \theta'|$$

and

$$|\sigma(\theta, t)| + |\mu(\theta, t)| \leqq K(1 + |\theta|).$$

Let x and y be solutions to the stochastic differential equations:

$$dx_t = \mu(x_t, t)dt + \sigma(x_t, t)dw_t$$
$$dy_t = m(y_t)dt + \rho(y_t)dv_t$$

where w and v are Wiener processes. Suppose that

$$\mu(\theta, t) \leqq m(\theta), \ 0 \leqq \sigma(\theta, t) \leqq \rho(\theta) \quad \text{for all } \theta, t$$

and that x_0 and y_0 are constants with $x_0 \leqq y_0$. Then

$$E\Phi(x_t) \leqq E\Phi(y_t)$$

for any nondecreasing convex function Φ on $\mathbb{R}$.

Proof. We will first prove the theorem under the following extra assumptions:

$$\Phi \text{ is Lipschitz continuous} \tag{4.1}$$

and for some ε with $0<\varepsilon<1$,

$$\varepsilon \leqq \rho(\theta) \quad \text{and} \quad \sigma(\theta, t) \leqq (1-\varepsilon)(\rho(\theta)-\varepsilon) \qquad \text{for all } \theta, t. \tag{4.2}$$

Given any Borel measurable function f on the interval $[0,1]$ we will abbreviate

$$\int_0^1 f(\alpha)d\alpha \quad \text{by } \int f(\alpha)d\alpha \quad \text{and} \quad \int_a^b f(\alpha)d\alpha \quad \text{by } \int_a^b f.$$

Let $L^2[0,1]$ denote the space of Borel measurable functions f on the interval $[0,1]$ such that $\|f\|_2$ is finite, where for any $p \geqq 1$,

$$\|f\|_p = (\int f(\alpha)^p d\alpha)^{1/p}.$$

Since ρ is Lipschitz continuous there is a constant K_ρ such that

$$\rho^2(\theta) \leqq K_\rho(1+\theta^2) \quad \text{and} \quad |\rho(\theta)-\rho(\theta')| \leqq K_\rho|\theta-\theta'|.$$

Choose a positive constant D so large that $K_\rho(1+D^{1/2})/D \leqq \varepsilon/2$ and define $\tilde{\rho}$ by $\tilde{\rho}(\theta) = \min(\rho(\theta), D)$.

Lemma 4.2. *For η in $L^2[0,1]$,*

$$\int_0^1 \rho(\eta_\alpha) - \tilde{\rho}(\eta_\alpha)d\alpha \leqq \varepsilon/2 \quad \textit{whenever} \quad \|\eta\|_2 \leqq D^{1/4}$$

Proof of Lemma 4.2.

$$\begin{aligned}\int_0^1 \rho(\eta_\alpha) - \tilde{\rho}(\eta_\alpha)d\alpha &\leqq \int_0^1 (\rho(\eta_\alpha) - \tilde{\rho}(\eta_\alpha))\rho(\eta_\alpha)d\alpha/D \\ &\leqq \|\rho(\eta)\|_2^2/D \\ &\leqq K_\rho(1+\|\eta\|_2^2)/D\end{aligned}$$

which implies the lemma. □

Lemma 4.3. *Let f and g Borel measurable functions on the interval $[0,1]$ such that for some constants D, a, and b,*

$$\begin{gathered}|f(\alpha)| \leqq D; \quad |g(\alpha)| \leqq D; \\ \int f(\alpha)d\alpha = \int g(\alpha)d\alpha = 0; \\ f(\alpha) \leqq 0 (\textit{resp. } f(\alpha) \geqq 0 \quad \textit{if} \quad \alpha \leqq a (\textit{resp. } \alpha > a); \quad \textit{and} \\ g(\alpha) \leqq 0 (\textit{resp. } g(\alpha) \geqq 0) \quad \textit{if} \quad \alpha \leqq b (\textit{resp. } \alpha > b);\end{gathered}$$

Then

$$\|f-g\|_2 \leqq (\||f|-|g|\|_2^2 + 6D\||f|-|g|\|_2)^{1/2}.$$

Proof of Lemma 4.3. Suppose without loss of generality that $a \leqq b$. Now

$$\|f-g\|_2^2 \leqq \||f|-|g|\|_2^2 + \int_a^b (|f|+|g|)^2$$
$$\leqq \||f|-|g|\|_2^2 + 2\int_a^b |f|^2 + 2\int_a^b |g|^2$$

and

$$2\int_a^b |f|^2 \leqq 2D\int_a^b |f| = D\left(2\int_0^b |f| - 2\int_0^a |f|\right) = D\left(2\int_0^b |f| - \int_0^1 |f|\right)$$
$$\leqq D\left(2\int_0^b |f| - 2\int_0^b |g| + \||f|-|g|\|_1\right) = D\left(2\int_0^b (|f|-|g|) + \||f|-|g|\|_1\right)$$
$$\leqq 3D\||f|-|g|\|_1 \leqq 3D\||f|-|g|\|_2.$$

Similarly

$$2\int_a^b |g|^2 \leqq 3D\||f|-|g|\|_2.$$

These inequalities easily imply the lemma. □

Given η in $L^2[0,1]$ and a nonnegative number θ define functions $B(\eta,\theta;\cdot)$ and $F(\eta,\theta;\cdot)$ on $[0,1]$ by

$$B(\eta,\theta;\alpha) = \lambda(\eta,\theta)\tilde{\rho}(\eta_\alpha)$$

where

$$\lambda(\eta,\theta) = \min\left(1-\varepsilon, \frac{\theta}{\int \tilde{\rho}(\eta_\alpha)d\alpha}\right)$$

and

$$F(\eta,\theta;\alpha) = \begin{cases} \tilde{\rho}(\eta_\alpha)\sqrt{1-\lambda(\eta,\theta)^2} & \text{if } \alpha \leqq \gamma(\eta,\theta) \\ -\tilde{\rho}(\eta_\alpha)\sqrt{1-\lambda(\eta,\theta)^2} & \text{if } \alpha > \gamma(\eta,\theta) \end{cases}$$

where $\gamma(\eta,\theta)$ is uniquely determined by the condition

$$\int F(\eta,\theta;\alpha)d\alpha = 0. \tag{4.3}$$

We seek a two parameter random process $(Y(t,\alpha): (\alpha,t)\in[0,1]\times\mathbb{R}_+)$ which is jointly measurable and adapted with

$$|||Y|||_t \triangleq E \sup_{0\leqq s\leqq t} \|Y(s,\cdot)\|_2^2 < +\infty$$

for all t, and processes x, w and $\tilde{w}$ such that

$$\begin{aligned} dY(t,\alpha) &= m(Y(t,\alpha))dt + B(Y(t,\cdot),\sigma(x_t,t);\alpha)dw_t \\ &\quad + F(Y(t,\cdot),\sigma(x_t,t);\alpha)d\tilde{w}_t, \text{ a.e. } \alpha, \\ dx_t &= \mu(x_t,t)dt + \sigma(x_t,t)dw_t, \\ Y(0,\alpha) &= y_0;\ x_0 \text{ is given}, \\ &w \text{ and } \tilde{w} \text{ are independent Wiener processes.} \end{aligned} \tag{4.4}$$

Using Lemma 4.3 it is easy to verify that the maps

$$(\eta, \theta) \rightsquigarrow B(\eta, \theta; \cdot)$$
$$(\eta, \theta) \rightsquigarrow F(\eta, \theta; \cdot)$$

are each bounded, uniformly continuous maps from $L^2[0,1] \times \mathbb{R}_+$ to $L^2[0,1]$. Hence, we can apply (a lemma-by-lemma generalization of) the Skorohod theory of existence of weak solutions as presented in [11, Sect. 6.1] or in [4] to deduce the following:

If we allow a change in probability space and, in particular, we allow substitution of new processes x and w (with the same distribution) for those originally given, then a collection $(x, Y, w, \tilde{w})$ which satisfies conditions (4.4) exists.

Define

$$\bar{m}_t = \int m(Y(t, \alpha))\,d\alpha$$
$$\bar{B}_t = \int B(Y(t, \cdot), \sigma(x_t, t); \alpha)\,d\alpha$$
$$\bar{Y}_t = \int Y(t, \alpha)\,d\alpha.$$

Then, using (4.3) and (4.4) we have

$$d\bar{Y}_t = \bar{m}_t\,dt + \bar{B}_t\,dw_t; \; \bar{Y}_0 = y_0. \tag{4.5}$$

For a.e. α, we have for all t

$$B(Y(t, \cdot), \sigma(x_t, t); \alpha)^2 + F(Y(t, \cdot), \sigma(x_t, t); \alpha)^2 = \tilde{\rho}(Y(t, \alpha))^2.$$

Therefore, for a.e. α,

$$dY(t, \alpha) = m(Y(t, \alpha))\,dt + \tilde{\rho}(Y(t, \alpha))\,dw_t^{\alpha}$$

where for each α, w^{α} is a Wiener process. Thus, for a.e. α, $Y(\cdot, \alpha)$ has the same distribution as the solution y^D to the stochastic differential equation

$$dy_t^D = m(y_t^D)\,dt + \tilde{\rho}(y_t^D)\,dv_t; \qquad y_0^D = y_0.$$

That is

$$y^D \sim Y(\cdot, \alpha) \qquad \text{a.e. } \alpha \tag{4.6}$$

Next, define a stopping time τ by

$$\tau = \inf\{t \colon \|Y(t, \cdot)\|_2 \geqq D^{1/4}\}.$$

If $t < \tau$ then by Lemma 4.2,

$$\int \rho(Y(t, \alpha))\,d\alpha \leqq \int \tilde{\rho}(Y(t, \alpha))\,d\alpha + \varepsilon/2.$$

Thus, if $t < \tau$ and $|x_t - \bar{Y}_t| \leqq \delta$, where $\delta = \varepsilon/2K_{\rho}$, then

$$\rho(x_t) \leqq \rho(\bar{Y}_t) + K_{\rho}|x_t - \bar{Y}_t| \leqq \int \rho(Y(t, \alpha))\,d\alpha + \varepsilon/2 \leqq \int \tilde{\rho}(Y(t, \alpha))\,d\alpha + \varepsilon.$$

So, if $t<\tau$ and $|x_t-\bar{Y}_t|\leqq\delta$, then (using (4.2))

$$\sigma_t\leqq(1-\varepsilon)(\rho(x_t)-\varepsilon)\leqq(1-\varepsilon)\int\tilde{\rho}(Y(t,\alpha))d\alpha$$

from which it follows that $\bar{B}_t=\sigma_t$. Consequently, if τ' is defined by

$$\tau'=\min\{t\geqq 0: x_t-\bar{Y}_t\geqq\delta\}$$

then

$$\bar{B}_t=\sigma_t \quad \text{if } (t\leqq\min(\tau,\tau') \quad\text{and}\quad x_t-\bar{Y}_t\geqq 0). \tag{4.7}$$

Compare the equation for x in (4.4) to equation (4.5) and use equation (4.7) to conclude by the same argument used in the proof of Theorem 2 that $x_t\leqq\bar{Y}_t$ for $t\leqq\min(\tau,\tau')$. Therefore $\tau'\geqq\tau$ and so conclude that

$$x_t\leqq\bar{Y}_t \quad \text{for } t\leqq\tau \quad P \text{ a.s.}$$

Using (4.6), note that

$$E\sup_{0\leqq s\leqq t}\|Y(s,\cdot)\|_2^2\leqq E\int_0^1\sup_{0\leqq s\leqq t}Y(s,\alpha)^2d\alpha=E\sup_{0\leqq s\leqq t}(y_s^D)^2.$$

The final term is bounded by a constant which does not depend on D. Thus, for any fixed t,

$$P[\tau>t]\to 1 \quad \text{as } D\to+\infty.$$

As D varies for t fixed, the random variables $\Phi(\bar{Y}(t))$ are uniformly integrable since (using the Lipschitz assumption on Φ)

$$\begin{aligned}E\Phi(\bar{Y}(t))^2&\leqq K_1+K_2E(\bar{Y}(t)^2)\\&\leqq K_1+K_2E\int Y(t,\alpha)^2d\alpha\\&=K_1+K_2E[y_t^D)^2]\end{aligned}$$

and the last term is bounded independently of D. Thus, if we write "$f\approx g$" to mean that $|f-g|$ can be made arbitrarily small by choosing D sufficiently large, we have

$$\begin{aligned}E\Phi(y_t)\approx E\Phi(y_t^D)&=E\int\Phi(Y(t,\alpha))d\alpha\geqq E\Phi(\bar{Y}(t))\\&\approx E[\Phi(\bar{Y}(t))I\{t\leqq\tau\}]\geqq E[\Phi(x_t)I\{t\leqq\tau\}]\\&\approx E\Phi(x_t).\end{aligned}$$

This establishes the theorem under the extra conditions (4.1) and (4.2).

We will now consider the general case in which (4.1) and (4.2) may no longer be true. First, we may assume that Φ is Lipschitz continuous without loss of generality since in general Φ is the limit of a monotone increasing sequence of nondecreasing, convex Lipschitz continuous functions Φ_n – thus we only need prove the theorem with Φ replaced by Φ_n and then let n tend to infinity to deduce the theorem with Φ replaced by Φ_n and then let n tend to infinity to deduce the general case by monotone convergence.

Next, suppose that $0<\varepsilon<1$ and let z^ε be the process defined by

$$dz_t^\varepsilon=m(z_t^\varepsilon)dt+\rho^\varepsilon(z_t^\varepsilon)dv_t; \quad z_t^\varepsilon=y_0$$

where

$$\rho^{\varepsilon}(\theta)=\varepsilon+\rho(\theta)/(1-\varepsilon).$$

Then when (ρ, y) is replaced by $(\rho^{\varepsilon}, z^{\varepsilon})$ the assumptions of the theorem as well as the stronger assumption (4.2) holds. Thus, we can apply the theorem to conclude that

$$E\Phi(x_t)\leqq E\Phi(z_t^{\varepsilon}). \tag{4.8}$$

Now by standard estimates used to prove existence and uniqueness for solutions to stochastic differential equations, we can easily show that

$$\lim_{\varepsilon\to 0} E \sup_{0\leqq s\leqq t} |z_s^{\varepsilon}-y_s|^2=0.$$

Therefore, the right side of inequality (4.8) converges to $E\Phi(y_t)$ as ε tends to zero, which proves that $E\Phi(x_t)\leqq E\Phi(y_t)$ as claimed.

Remark. The conditions of Theorem 4.1 are satisfied when $m=0$, $\rho(\theta)=\theta$, $y_0=1$ and $x_t=1$ for all t. Using Itô's formula it is easy to verify that then $y_t=\exp(v_t-t/2)$, and so $P[y_t\geqq 1]$ is equal to $P[v_t\geqq t/2]$ which tends to zero as t tends to infinity. Thus, inequality (1.2) is not true (even if 2 is replaced by a larger constant). Thus, x is not pathwise dominated by a convex combination of any *finite* number of processes with the same distribution as y.

V. Application to Semimartingales in n Dimensions

Let z be a semimartingale in $\mathbb{R}^n$ with representation

$$dz_t=\alpha_t dt+\sum_j \beta_t^j dw_t^j$$

where α, β^j are adapted, n-vector valued locally norm square integrable, and $(w_t^j\colon 1\leqq j\leqq p)$ is an adapted vector Wiener process. Use "$\|\ \|$" to denote the Euclidean norm in $\mathbb{R}^n$.

Theorem 5. *Suppose there is a function g and positive constants ρ and u such that*

$$\left.\begin{aligned}&\frac{z_t}{\|z_t\|}\cdot\alpha_t\leqq g(\|z_t\|)\\ \text{and}\qquad &\Big(\sum_j \|\beta_t^j\|^2\Big)^{1/2}\leqq\rho\end{aligned}\right\}\ \text{whenever}\ \|z_t\|\geqq u$$

Suppose also that m defined by

$$m(\theta)=g(\theta)+\rho^2/2\theta$$

is convex and Lipschitz continuous on $[u, +\infty)$ (for examples, g may satisfy these conditions). Let (y, l) be a (pathwise unique) solution to the stochastic differential equation

$$dy_t=m(y_t)dt+\rho dv_t+dl_t,$$

$$y_t\geqq u,\ I\{y_t>u\}dl_t=0,\ l_0=0,\ y_0=\max(\|z_0\|, u),$$

l is continuous, nondecreasing.

Then

$$P[\|z_t\| \geqq c] \leqq 2P[y_t \geqq c] \quad \text{for } c \text{ in } \mathbb{R}, \tag{5.1}$$

and for any nondecreasing convex function Φ on $\mathbb{R}$

$$E\Phi(\|z\|_t) \leqq E\Phi(y_t). \tag{5.2}$$

Proof. Let x be defined by $x_t = \max(u, \|z_t\|)$. Equivalently, $x_t = \max(u, s_t)$ where s_t is the one parameter semimartingale $s_t = h(z_t)$ where h is any twice continuously differentiable function such that $h(z) = \|z\|$ if $\|z\| \geqq u$ and $h(z) \leqq u$ if $\|z\| \leqq u$. Hence, we can apply Itô's formula (6.1) to yield that x_t is a semimartingale with representation

$$dx_t = \mu_t dt + \frac{1}{x_t} \sum_j z_t \cdot \beta_t dw_t^j + d\lambda_t = \mu_t dt + \sigma_t d\tilde{w}_t + d\lambda_t$$

where

$$\sigma_t^2 = \sum_j \left(\frac{z_t}{x_t} \cdot \beta_t^j\right)^2,$$

$$\mu_t = [z_t \cdot \alpha_t + \sum_j \|\beta_t^j\|^2 - \sigma_t^2]/x_t,$$

$\tilde{w}$ is a Wiener process (the construction of which may require enlarging the probability space if σ_t is sometimes zero), and λ is a continuous increasing process (twice the local time of $\|z_t\|$ at u) which is increasing only when $\|z_t\| = u$.

By the Schwarz inequality

$$\sigma_t^2 \leqq \sum_j \|\beta_t^j\|^2,$$

so that $|\sigma_t| \leqq \rho$ and $\mu_t \leqq m(x_t)$ whenever $x_t > u$. Therefore Theorem 2 can be applied to (x, ρ, m) to yield (5.1) and (5.2). □

Examples. If $g(\theta) = c - \rho^2/2\theta$ then $m(\theta) \equiv c$ and $y_t - u$ has transition density (2.5). If $g(\theta) = -x - \rho^2/2\theta$ then y is an Ornstien-Uhlenbeck process which is modified to reflect at u. If $g(\theta) = K/\theta$ then y becomes a Bessel process which is modified to reflect at u.

VI. Appendix – Local Time for Continuous Semimartingales

In our proof of Theorem 2 we use a method suggested by Perkins [8] which was used by Le Gall [5] for establishing a comparison theorem along the lines of Theorem 1.1. The method is based on local times – a topic we briefly review here.

The local time of a continuous real-valued semimartingale X at a is the nondecreasing process $L_t^a(X)$ defined by

$$L_t^a(X) = |X_t - a| - |X_0 - a| - \int_0^t I\{X_s > a\} dX_s.$$

One can show (see [14]) that there is a version of $L(X)$ which is jointly (continuous in t and right continuous in a). Itô's formula can be generalized: if f is the difference of two convex functions then (see [7, Sect. VI.II, especially (13.1) and (14.3)])

$$f(X_t)=f(X_0)+\int_0^t f'(X_s)\,dX_s+\int_{\mathbb{R}} f''(da)\,L_t^a(X) \tag{6.1}$$

where f' is the left derivative of f and f'' is the generalized second derivative (in general, a σ-finite signed measure) of f.

Finally, if g is a nonnegative Borel function on $\mathbb{R}$ [7],

$$\int_0^t g(X_s)\,d\langle X, X\rangle_s=\int_{\mathbb{R}} g(a)L_t^a(X)\,da \tag{6.2}$$

and for each a,

$$\int_0^t I\{x_s \neq a\}\,dL_s^a(x)=0.$$

References

1. Dellacherie, C., Meyer, P.A.: Probabilities and Potential *B*. New York: North-Holland 1982
2. Doob, J. L.: Stochastic Processes. New York: Wiley 1953
3. Ikeda, N., Watanabe, S.: A comparison theorem for solutions of stochastic differential equations and its applications. Osaka J. Math. **14**, 619–633 (1977)
4. Ikeda, N., Watanabe, S.: Stochastic Differential Equations and Diffusion Processes. New York: North-Holland 1981
5. Le Gall, J.F.: Application du temps local aux équations differentielles stochastiques unidimensionelles. Séminaire de Probabilités XVII, Lecture Notes in Mathematics **986**, 15–31. Berlin Heidelberg New York: Springer 1982
6. Malliavin, P.: Géometrie Differentielle Stochastique. Montréal: Les Presse de l'Université de Montréal 1978
7. Meyer, P.A.: Un cours sur les intégrales stochastiques. Séminaire de Probabilités *X*, Lecture Notes in Mathematics **511**, 245–400. Berlin Heidelberg New York: Springer 1976
8. Perkins, E.: Local time and pathwise uniqueness of solutions of stochastic differential equations. Seminaire de Probabilités XVI, Lecture Notes in Mathematics **920**, 201–208. Berlin Heidelberg New York: Springer 1982
9. Skorohod, A.V.: Stochastic equations for a diffusion process in a bounded region. Theory of Probability and its Applications **6**, 264–274 (1961)
10. Skorokhod, A.V.: Studies in the Theory of Random Processes. Reading, Massachusetts: Addison Wesley 1965
11. Stroock, D.W., Varadhan, S.R.S.: Multidimensional Diffusion Processes, Berlin Heidelberg New York: Springer 1979
12. Wonham, W.M.: Liapunov criteria for weak stochastic stability. J. of Differential Equations **2**, 195–207 (1966)
13. Yamada, T.: On the uniqueness of solution of stochastic differential equations with reflecting barrier conditions. Seminaire de Probabilités X, Lecture Notes in Mathematics **511**, 240–244. Berlin Heidelberg New York: Springer 1976
14. Yor M.: Sur la continuité des temps locaux associes à certaines semimartingales. Astérisque **52–53**, 23–35 (1978)

Received October 12, 1983