THEORY PROBAB. APPL.
Vol. 32, No. 2

Translated from Russian Journal

BOUNDS FOR THE MIXING RATE IN THE THEORY OF STOCHASTIC EQUATIONS*

A. YU. VERETENNIKOV

(Translated by Merle Ellis)

1. Introduction. Various theorems of probability theory have been established [1]-[3] under the condition that some mixing coefficient decreases fast enough. This condition is usually the convergence of some power of the mixing coefficient, perhaps with a weight function. For Markov processes, a thorough study [4] has been made of conditions for the exponential decrease of the uniformly strong mixing coefficient. However, in a number of situations, say for stochastic equations in a Euclidean space, these conditions are not satisfied and the processes themselves are not uniformly strongly mixed. In such a case one can hope to estimate weaker mixing coefficients. This is the situation examined below.

We study an Itô stochastic differential equation

$$dx_t = \sigma(t, x_t)\, dw_t + b(t, x_t)\, dt, \qquad t \geqq 0, \tag{1}$$

with initial value

$$x_0 = x. \tag{2}$$

Here $(w_t, \mathscr{F}_t)$ is a d-dimensional Wiener process on some complete probability space $(\Omega, \mathscr{F}, \mathbf{P})$ with an increasing family of complete σ-algebras, $\sigma(t, x)$ is a $d \times d$ matrix, $b(t, x)$ is a d-dimensional vector, σ and b are bounded and Borelian, and, for all x, $\lambda \in E^d$ and $t \geqq 0$,

$$\lambda^* \sigma\sigma^*(t, x)\lambda \geqq \nu|\lambda|^2, \qquad \nu > 0. \tag{3}$$

Recall that by a solution or weak solution of problem (1)-(2) one means a continuous process x_t that is measurable with respect to $\mathscr{F}_t$ for each $t \geqq 0$ and such that

$$\mathbf{P}\left(x_t = x + \int_0^t \sigma(s, x_s)\, dw_s + \int_0^t b(s, x_s)\, ds,\ t \geqq 0\right) = 1,$$

where $\int \sigma\, dw$ is the Itô stochastic integral and $\int b\, ds$ is the Lebesgue integral. According to known results of N. V. Krylov, the problem (1)-(2) has a solution on some probability space with a Wiener process. It is assumed that σ and b are such that this solution is unique in the sense of the distribution law and is a strong Markov process with respect to the flow of σ-algebras $(\mathscr{F}_t^x)$, where $\mathscr{F}_t^x$ is the $\mathbf{P}$-completion of the σ-algebra $\sigma(x_s: 0 \leqq s \leqq t)$ [5]-[7] (this condition is satisfied automatically for $d = 1$; for $d > 1$ it suffices for example that $\sigma\sigma^*(t, \cdot)$ be continuous uniformly with respect to $t \geqq 0$).

In Lemmas 1 as well as 2 below some estimates are established that characterize the renewal properties of a process. In the theorem of § 2 exponential bounds are derived from the strong mixing and complete regularity coefficients.

* Received by the editors December 20, 1984.

2. The main results. Let $\mathcal{F}^x_{a,b}$ be the $\mathbf{P}$-completion of the σ-algebra $\sigma(x_s\colon a \leqq s \leqq b)$. Below we examine the following mixing coefficients (see [4]):

(1) the strong mixing (s.m.) coefficient

$$\alpha(s) = \sup_t \sup_{\substack{A \in \mathcal{F}^x_{0,t} \\ B \in \mathcal{F}^x_{t+s,\infty}}} |\mathbf{P}(AB) - \mathbf{P}(A)\mathbf{P}(B)|;$$

(2) the complete regularity (c.r.) coefficient

$$\beta(s) = \sup_t \mathbf{E} \operatorname*{var}_{B \in \mathcal{F}^x_{t+s,\infty}} |\mathbf{P}(B \mid \mathcal{F}^x_{0,t}) - \mathbf{P}(B)|.$$

One has $\alpha(s) \leqq \beta(s)$.

THEOREM. *Suppose that, for some $r > 0$ and all $|z| \geqq r$, $t \geqq 0$,*

$$(b(t, z), z) \leqq -\gamma|z|, \qquad \gamma > 0. \tag{4}$$

Then for some $\lambda > 0$, for each x and all $s \geqq s(x)$,

$$\alpha(s) \leqq \beta(s) \leqq \exp(-\lambda s). \tag{5}$$

3. Auxiliary assertions. Put $\tau(x) = \inf(t \geqq 0\colon |x_t| \leqq R)$, $h_R(x) = \max(|x| - R, 0) = (|x| - R) \vee 0$, $R > 0$.

LEMMA 1. *Let condition* (4) *hold. Then there is an $R \geqq r$ such that, for any sufficiently small $\varepsilon > 0$, there is a $\delta > 0$ such that*

$$\mathbf{E} \exp(\delta\tau(x)) \leqq \exp(\varepsilon h_R(x)). \tag{6}$$

Proof. Consider the function $v(t, z) = \exp(\delta t + \varepsilon|x|)$. Let $|x| > R$. Then Itô's formula is applicable to the expression $\exp(\delta t + \varepsilon|x_t|)$ and, putting $T_N = \inf(t \geqq 0\colon |x_t| \geqq N)$, $a = \sigma\sigma^*/2$, we obtain

$$\begin{aligned}
&\exp(\delta(t \wedge \tau \wedge T_N) + \varepsilon|x_{t\wedge\tau\wedge T_N}|) - \exp(\varepsilon|x|) \\
&= \int_0^{t\wedge\tau\wedge T_N} \exp(\delta_s + \varepsilon|x_s|)\Bigg[\varepsilon(x_s, b(s, x_s))/|x_s| + \delta \\
&\qquad + \varepsilon^2 \sum_{i=1}^d \frac{a_{ii}(s, x_s)}{|x_s|} - \varepsilon^2 \sum_{j,k=1}^d \frac{x_s^j x_s^k}{|x_s|^3} a_{jk}(s, x_s)\Bigg] ds \\
&\quad + \int_0^{t\wedge\tau\wedge T_N} \exp(\delta s + \varepsilon|x_s|)\varepsilon \sum_{i,j=1}^d \frac{x_s^j}{|x_s|} \sigma^{ji}(s, x_s)\, dw_s^i.
\end{aligned}$$

Hence, for $\varepsilon \leqq \varepsilon_0$ and $\delta \leqq \varepsilon r/\gamma$,

$$\mathbf{E} \exp(\delta(t \wedge \tau \wedge T_N) + \varepsilon|x_{t\wedge\tau\wedge T_N}|) \leqq \exp(\varepsilon|x|).$$

Therefore, by Fatou's lemma, as $t, N \to \infty$,

$$\mathbf{E} \exp(\delta\tau(x)) \leqq \exp(\varepsilon(|x| - R)).$$

Lemma 1 is proved.

Let $\psi(t)$ be a random time substitution connected with some additive functional (see (8]), and suppose that, for all t and s,

$$C^{-1}|t - s| \leqq |\psi(t) - \psi(s)| \leqq C|t - s|, \qquad C > 0. \tag{7}$$

Then (6) is also valid in the substituted time, perhaps with a different δ (e.g., $\delta' = C^{-1}\delta$).

Let $z_t = |x_t| \vee |x| \vee R$ (x is the initial value of the process x_t). Let us calculate the stochastic differential of z_t ($a = \sigma\sigma^*/2$):

$$dz_t = d\varphi_t + \sum_{i,j=1}^{d} I(|x_t| > |x| \vee R) \frac{x_t^i}{|x_t|} \sigma^{ij}(t, x_t)\, dw_t^j + I(|x_t| > |x| \vee R)$$

$$\cdot \left[\sum_{k=1}^{d} \frac{x_t^k}{|x_t|} b^k(t, x_t) - \sum_{i,j=1}^{d} \frac{x_t^i x_t^i}{|x_t|^3} a_{ij}(t, x_t) + \sum_{i=1}^{d} \frac{a_{ii}(t, x_t)}{|x_t|} \right] dt,$$

where φ_t is the local time of z_t at the point $|x| \vee R$, i.e., $\varphi_0 = 0$, φ is continuous, increasing and $d\varphi_t = I(|x_t| = |x| \vee R)\, d\varphi_t$. Put

$$\beta_t^j = \sum_{i=1}^{d} \frac{x_t^i}{|x_t|} \sigma^{ij}(t, x_t).$$

Since a is nonsingular,

$$\begin{aligned}\sum_{j=1}^{d} (\beta_t^j)^2 &= \sum_{j=1}^{d} \left(\sum_{i=1}^{d} \frac{x_t^i}{|x_t|} \sigma^{ij}(t, x_t) \right)^2 \\ &= (\sigma(t, x_t) x_t/|x_t|, \sigma(t, x_t) x_t/|x_t|) \\ &= (\sigma^* \sigma(t, x_t) x_t/|x_t|, x_t/|x_t|) \geqq \nu > 0.\end{aligned}$$

Consider the functional $f(t) = \int_0^t |\beta_s|^2\, ds$. Put $\psi(t) = f^{-1}(t)$ (the inverse of f) and $\tilde{z}_t = z_{\psi(t)}$. Then $\tilde{z}_t$ satisfies the stochastic differential equation

$$d\tilde{z}_t = d\tilde{w}_t + \tilde{b}_t\, dt + d\tilde{\varphi}_t, \qquad \tilde{z}_0 = z_0, \tag{8}$$

where $\tilde{w}_t$ is some Wiener process and $\tilde{\varphi}_t$ is the local time (see [9]). The process $\tilde{b}_t$ satisfies

$$\tilde{b}_t \leqq -\tilde{\gamma}, \qquad \tilde{\gamma} > 0,$$

if R is sufficiently large (and we shall assume it to be such).

Along with (8) we consider the stochastic differential equation with instantaneous reflection at $|x| \vee R$,

$$dy_t = d\tilde{w}_t - (\tilde{\gamma}/2)\, dt + d\varphi_t^y, \qquad y_0 = z_0, \tag{9}$$

with the same Wiener process $\tilde{w}$. We have

$$\mathbf{P}(\tilde{z}_t \leqq y_t,\ t \geqq 0) = 1. \tag{10}$$

Indeed, let $v_t = y_t - \tilde{z}_t$. Then $v_0 = 0$. Let $v_t < 0$. Then $\tilde{z}_t > |x| \vee R$, $d\tilde{\varphi}_t = 0$, and

$$dv_t = (-\tilde{\gamma}/2 - \tilde{b}_t)\, dt + d\varphi_t^y - d\tilde{\varphi}_t \geqq (\tilde{\gamma}/2)\, dt + d\varphi_t^y > 0.$$

Hence the inequality $v_t < 0$ is impossible.

The process y_t has stationary distribution density

$$p(u) = \text{const} \exp(-\tilde{\gamma} u/2), \qquad u > |x| \vee R.$$

If we consider (9) with the initial condition distributed in accordance with the density p (let $\tilde{v}_t$ denote its solution), then by (10) we have $\tilde{z}_t \leqq \tilde{v}_t$ and thus

$$\mathbf{E} \exp(\varepsilon h_{R \vee |x|}(\tilde{z}_t)) \leqq \mathbf{E} \exp(\varepsilon h_{R \vee |x|}(\tilde{v}_t)) \leqq C(\varepsilon)$$

for $\varepsilon < \tilde{\gamma}/2$ and all $t \geqq 0$. As $\varepsilon \downarrow 0$ we have $C(\varepsilon) \downarrow 1$. Thus, the following lemma is valid:

LEMMA 2. *For some* $R > 0$ *and all* $\varepsilon \in (0, \tilde{\gamma}/2)$, $t \geqq 0$,

$$\mathbf{E} \exp(\varepsilon h_R(\tilde{x}_t)) \leqq \exp(\varepsilon h_R(x)) C(\varepsilon),$$

where $C(\varepsilon) \downarrow 1$ *as* $\varepsilon \downarrow 0$.

4. Proof of the theorem. It suffices to establish the estimate

$$\beta_s \leqq \exp(-\lambda s), \qquad \lambda > 0, \qquad s \geqq s(x),$$

for the solution x_t of equation (1)-(2) with some Wiener process w_t^1. It also suffices to prove such a bound for the process with substituted time $x_t^1 = x_{\psi(t)}$ (see § 3) since the substitution $\psi(t)$ satisfies (7) with some constant C.

Let $F = \{\xi \in C[0, \infty; E^d): \xi_{t_1} \in \Gamma_1, \dots, \xi_{t_m} \in \Gamma_m,\ t_i \geqq t+s,\ \Gamma_i \in \mathscr{B}^d\}$. We have

$$\mathbf{E} \operatorname*{var}_{B \in \mathscr{F}^{x^1}_{t+s,\infty}} |\mathbf{P}(B \mid \mathscr{F}^{x^1}_{0,t}) - \mathbf{P}(B)| = \mathbf{E} \operatorname*{var}_{\Gamma \subset F} |\mathbf{P}(x^1 \in \Gamma \mid \mathscr{F}^{x^1}_{0,t}) - \mathbf{P}(x^1 \in \Gamma)|.$$

Let x_t^2 be another copy of the solution of the initial equation with Wiener process $w^2 \perp\!\!\!\perp w^1$ (independent of w^1) and with a similar time substitution. Then also $x^2 \perp\!\!\!\perp x^1$. Since the distribution is Markovian and unique,

$$\mathbf{E} \operatorname*{var}_{\Gamma \subset F} |\mathbf{P}(x^1 \in \Gamma \mid \mathscr{F}^{x^1}_{0,t}) - \mathbf{P}(x^1 \in \Gamma)| = \mathbf{E} \operatorname*{var}_{\Gamma \subset F} |\mathbf{P}(x' \in \Gamma \mid \mathscr{F}^{x^1}_{0,t}) - \mathbf{P}(x^2 \in \Gamma)|.$$

Let us further suppose that there is a process x^3 with the same distribution as x^1 and x^2 which is therefore also obtained by a similar time substitution from the solution of the initial equation with some Wiener process (we assume that x^3 is continuous) such that $x_s^3 = x_s^2$ for $0 \leqq s \leqq t$ (so that $x^3 \perp\!\!\!\perp x^2$); moreover, for $s \geqq t$ the trajectory of x^3 is constructed in some nonanticipating way with respect to the trajectories of x^1 and x^2. Then

$$\mathbf{P}(x^1 \in \Gamma \mid \mathscr{F}^{x^1}_{0,t}) - \mathbf{P}(x^2 \in \Gamma) = \mathbf{E}_{t,x_t^1}(\mathbf{E}_{t,x_t^1,x_t^2} I(x^1 \in \Gamma) - \mathbf{E}_{t,x_t^1,x_t^2} I(x^2 \in \Gamma)),$$

and by the conditions $x_s^2 = x_s^3$, $0 \leqq s \leqq t$, $x^1 \perp\!\!\!\perp x^2$, the Markov property and uniqueness of the solution in the sense of the distribution law,

$$\mathbf{E}_{t,x_t^1,x_t^2} I(x^1 \in \Gamma) = \mathbf{E}(I(x^1 \in \Gamma) \mid \mathscr{F}^{x^1,x^2}_{0,t}) = \mathbf{E}_{t,x_t^1,x_t^3 \equiv x_t^2} I(x^1 \in \Gamma),$$

$$\mathbf{E}_{t,x_t^1,x_t^2} I(x^2 \in \Gamma) = \mathbf{E}(I(x^2 \in \Gamma) | x_t^2) = \mathbf{E}(I(x^3 \in \Gamma) \mid x_t^3 \equiv x_t^2) = \mathbf{E}_{t,x_t^1,x_t^3 \equiv x_t^2} I(x^3 \in \Gamma).$$

Therefore,

$$\begin{aligned} |\mathbf{P}(x^1 \in \Gamma \mid x_t^1) - \mathbf{P}(x^2 \in \Gamma)| &= |\mathbf{E}_{t,x_t^1,x_t^3} I(x^1 \in \Gamma) - \mathbf{E}_{t,x_t^1,x_t^3} I(x^3 \in \Gamma))| \\ &\leqq \mathbf{E}_{t,x_t^1} \mathbf{P}_{t,x_t^1,x_t^3}((x^1)_{t+s}^\infty \neq (x^3)_{t+s}^\infty) \\ &= \mathbf{P}_{t,x_t^1}((x^1)_{t+s}^\infty \neq (x^3)_{t+s}^\infty), \end{aligned} \tag{11}$$

where ξ_a^b is the trajectory of ξ on $[a, b]$. Thus, to obtain the desired bound it suffices to construct from x^2 and x^1 a process x^3 with the properties described above such that the probability on the right-hand side of (11) admits a good upper bound.

We shall assume that $R = 1$, $r \leqq \frac{1}{20}$; this can be achieved by a change of scale. Let $\tau_1 = \inf(s \geqq t: |x_t^1| \leqq \frac{1}{10})$. By Lemma 2 and the independence of x^1 and x^2,

$$\mathbf{E}_{t,x_t^2} \exp(\varepsilon h_R(x_{\tau_1}^2)) \leqq C(\varepsilon) \exp(\varepsilon h_R(x_t^2)),$$

$$\mathbf{E} \exp(\varepsilon h_R(x_{\tau_1}^2)) \leqq C(\varepsilon) \exp(\varepsilon h_R(x)).$$

Moreover,

$$\mathbf{E}_{t,x_t^1} \exp(\delta(\tau_1 - t)) \leqq \exp(\varepsilon h(x_t^1)).$$

Let $\tau_1' = \inf(s \geqq \tau_1 + 1: |x_s^2| \leqq \frac{1}{20}) \wedge (\tau_1 + 2)$.

Case A: $|x_{\tau_1'}^2| \leqq \frac{1}{20}$. Let us for the present leave off the time substitution (or assume it to be the identity) so that x^1 after the time τ_1 and x^2 after the time τ_1' are solutions of the equations

$$dx_s^1 = \sigma(s, x_s^1)\, dw_s^1 + b(s, x_s^1)\, ds, \qquad s \geqq \tau_1,$$

with initial values $x^1_{\tau_1}$ for $s=\tau_1$ and

$$dx^2_s=\sigma(s,x^2_s)\,dw^2_s+b(s,x^2_s)\,ds,\qquad s\geqq\tau'_1,$$

with initial value $x^2_{\tau'_1}$ for $s=\tau'_1$, respectively. Case A is the main case in the subsequent arguments.

Case B: $|x^2_{\tau'_1}|>\frac{1}{20}$. In this case we stop x^2 at the time $\tau'_1=\tau_1+2$. For the time being x^1 is constructed up to τ_1, and at this τ_1 it has the bound $|x'_{\tau_1}|\leqq\frac{1}{10}$. We further extend the process again in substituted time. Put

$$\tau_2=\inf\,(s\geqq\tau_1+2\colon |x^2_s|\leqq\tfrac{1}{10}),$$

$$\tau'_2=\inf\,(s\geqq\tau_2+1\colon |x^1_s|\leqq\tfrac{1}{20})\wedge(\tau_2+2).$$

Case B again breaks down into two cases: case BA: $|x^1_{\tau'_2}|\leqq\frac{1}{20}$ and case BB: $|x^1_{\tau'_2}|>\frac{1}{20}$. The probabilities of the corresponding events are separated from zero, as one can see, say, with the aid of the methods in [11].

Case BB again breaks down into two cases. Let $\tau_3=\inf\,(s\geqq\tau_2+2\colon |x^1_s|\leqq\frac{1}{10})$, $\tau'_3=\inf\,(s\geqq\tau_3+1\colon |x^2_s|\leqq\frac{1}{20})\wedge(\tau_3+2)$. Case BBA: $|x^2_{\tau'_3}|\leqq\frac{1}{20}$ and Case BBB: $|x^2_{\tau'_3}|>\frac{1}{20}$.

The procedure of decomposing into cases is continued in similar fashion.

Now consider case A (cases BA, BBA, etc., are considered similarly). Put (in unsubstituted time)

$$\varphi^1_1=\inf\,(s\geqq\tau_1\colon |x^1_s|\geqq 1)\wedge(\tau'_1+1),$$

$$\varphi^2_1=\inf\,(s\geqq\tau'_1\colon |x^2_s|\geqq 1)\wedge(\tau'_1+1),$$

and let $\mathscr{L}(\varphi^1_1,x^1_{\varphi^1_1})$ be the distribution of the pair $(\varphi^1_1,x^1_{\varphi^1_1})$, and $\mathscr{L}(\varphi^2_1,x^2_{\varphi^2_1})$ the distribution of the pair $(\varphi^2_1,x^2_{\varphi^2_1})$. By Theorem 1.1 in [12], Harnack's inequality, we have

$$\mathscr{L}(\varphi^2_1,x^2_{\varphi^2_1})\ll\mathscr{L}(\varphi^1_1,x^1_{\varphi^1_1}),$$

and moreover

$$p_1=d\mathscr{L}(\varphi^2_1,x^2_{\varphi^2_1})/d\mathscr{L}(\varphi^1_1,x^1_{\varphi^1_1})\leqq C \tag{12}$$

with some positive constant C (for all the conditions of the Krylov-Safonov theorem to hold it suffices to apply a routine transformation of the coordinates).

On some new independent probability space we realize a random throw of a point into the region bounded below by zero and above by the graph of the density p_1.

Case AN. The point falls into the region with $p_1>1$ (one is the graph of the density $d\mathscr{L}(\varphi^1_1,x^1_{\varphi^1_1})/d\mathscr{L}(\varphi^1_1,x^1_{\varphi^1_1})$). Then x' and x^2 are considered up to the time $\varphi_1=\varphi^1_1\wedge\varphi^2_1\wedge(\tau'_1+1)$ and then we again have to choose a variant of A or B, i.e., ANA or ANB.

Case AS. The point falls into the region with $p_1\leqq 1$; the probability of this is not less than $1/C$, where C is the constant in (12). Then we assume that $(x^3)^{\tau'_1}_0\equiv(x^2)^{\tau'_1}_0$, $x^3_{\varphi^0}=x^0$, where (φ^0,x^0) is the point where the distribution $\mathscr{L}(\varphi^1_1,x^1_{\varphi^1_1})$ was realized. On some new independent space we further construct the trajectory x^3 hitting the already-found point x^0 at φ^0 so that x^3 (in unsubstituted time) remains a solution of the corresponding stochastic differential equation. From the time $\varphi^0=\varphi^1_1$ the processes x^1 and x^3 are considered to be coincident or glued and both are solutions of one and the same stochastic equation with one and the same Wiener process.

Using the method of gluing from [13, Chap. 6], one can show that the constructed process x^3 is a solution of the initial equation (1)-(2) with some Wiener process transformed with the aid of a time substitution.

Thus, the trajectories of the constructed processes x^1 and x^3 generate sequences of "letters" or "words" *AS*, *BAS*, *BBAS*, $\cdots$, *ANAS*, *BANAS*, *BBANAS*, $\cdots$, *ANANAS*, *BANANAS*, *BBANANAS*, $\cdots$, *ANBANAS*, *ANBBANAS*, $\cdots$. Each word consists of symbols A, B, N, S and is almost surely finite, i.e., consists of a finite number of letters (this is formally shown below), and ends in the symbol S (in a gluing). Before S there must stand the symbol A with which every series of the form A, BA, BBA, $\cdots$ terminates. From A, passage to S is possible with probability depending on the trajectory but separated from zero, or to N. If a passage to N occurred, then again a series of the form A, BA, BBA, $\cdots$ is realized.

Let us define the length of each letter as the difference of the times of the ending and beginning of the letter. The time of the beginning of the first letter (A or B) is by definition equal to t. We shall assume that the time of its ending is equal to τ_1', and is the beginning of the following letter, etc. For example, let the first letter be B, and its length $|B| = \tau_1' - t$. Let the letter after it be A; then its beginning is τ_1' and its ending τ_2', and the length is equal to $|A| = \tau_2' - \tau_1'$. Suppose A is followed by N. Its ending is $\varphi_2 = \varphi_2^1 \wedge \varphi_2^2 \wedge (\tau_2' + 1)$, where $\varphi_2^1 = \inf(s \geqq \tau_2'\colon |x_s^1| \geqq 1)$, $\varphi_2^2 = \inf(s \geqq \tau_2'\colon |x_s^2| \geqq 1)$, and the length is $|N| = \varphi_2 - \tau_2'$, etc. We define the length of part of a word as the sum of the lengths of the letters in it. Let L denote the time from t to the gluing time, i.e., the length of the entire word (to the ending of the letter S). Then

$$\mathbf{P}(x_{t+s}^1 \neq x_{t+s}^3) \leqq \mathbf{P}(L > s).$$

Let us show that for sufficiently small $\delta > 0$ the exponential time $\mathbf{E} \exp(\delta L)$ is finite. Then by Chebyshev's inequality we obtain

$$\mathbf{P}(x_{t+s}^1 \neq x_{t+s}^3) \leqq \exp(-\delta s) \mathbf{E} \exp(\delta L), \tag{13}$$

which will prove the desired bound.

Let $\bar{B}$ denote the maximal series of letters B of the form $BB \cdots B$ in any part of a word; thus all words are representable in the form $\bar{B}AS$, $AN\bar{B}AS$, $\cdots$.

Below it is convenient to modify slightly the rule for constructing words and trajectories as follows. We shall assume that the gluing takes place fictitiously, and the whole procedure of constructing trajectories goes on as if gluing does not occur, as before, with two independent Wiener processes w^1 and w^2. Thus, words becomes infinite:

$$ANAS\bar{B}AS\bar{B}AN \cdots,$$

and so on. At the time of ending of A one determines, by a method independent of the preceding trajectory (namely, by throwing a point under the graph of the density p_1), the letter following A: S or N. The quantity L is defined as the time of ending of the first letter S minus t.

Let us make a further reduction. Let $\bar{A}$ denote the combination $\bar{B}A$ or the sole letter A if the "letter" $\bar{B}$ does not stand before it. Now words can be recorded thusly: $\bar{A}N\bar{A}N\bar{A}S\bar{A} \cdots$.

By the construction we have $|N| \leqq 2$, $|S| \leqq 2$. Let us prove that

$$\mathbf{E}(\exp(\delta |\bar{A}|) \mid \mathscr{F}_\theta) < \infty$$

for sufficiently small $\delta > 0$ (θ is the time of the beginning of the "letter" $\bar{A}$). First of all, $|A| \leqq 2$, so

$$\mathbf{E}(\exp(\delta |\bar{A}|) \mid \mathscr{F}_\theta) \leqq \exp(2\delta) \mathbf{E}(\exp(\delta |\bar{B}|) \mid \mathscr{F}_\theta).$$

Let $\theta_i \leqq \theta^i$ be the beginning and ending of the ith letter B, respectively. Let us suppose that the trajectory of x^1 is constructed up to θ^i and that of x^2 up to θ_i; for definiteness we consider the case of odd i (otherwise, switch places of x^1 and x^2). Recall that, for $i>1$, $|x^2_{\theta_i}|\leqq 1$. The event $|x^2_{\tau'_i}|\leqq\frac{1}{20}$ is defined independently of the trajectory of x^1 and its probability is separated from zero:

$$\mathbf{P}(|x^2_{\tau'_i}|\leqq\tfrac{1}{20})\geqq q>0 \tag{14}$$

for some constant q. If this event occurs, then the next letter is A, if not, then B again.

Let us make one further reduction. We shall assume that the trajectories are always constructed as described in Case B and the letter A is realized only fictitiously (depending on whether $|x^2_{\tau'_i}|\leqq\frac{1}{20}$ is satisfied or not), and independently of the trajectories constructed up to now. We have to estimate the length of a word up to the first appearance of the fictitious letter A, and we denote this quantity as before by $|\bar{B}|$.

Let the event Q_i denote the realization of the event A on the ith step and Q_i^c the complement. From the remarks made above (see (14)),

$$\begin{aligned}\mathbf{E}_{\mathscr{F}_\theta}\exp(\delta|\bar{B}|)\leqq\mathbf{E}_{\mathscr{F}_\theta}(&\exp(\delta|B_1|)+\exp(\delta|B_1|)I(Q_1^c)\exp(\delta|B_2|)\\ &+\exp(\delta|B_1|)I(Q_1^c)\exp(\delta|B_2|)I(Q_2^c)\exp(\delta|B_3|)+\cdots\\ &+\exp(\delta|B_1|)I(Q_1^c)\exp(\delta|B_2|)\cdots I(Q_{n-1}^c)\exp(\delta|B_n|)+\cdots).\end{aligned}$$

Let us estimate the nth term of this series. We have (let n be odd)

$$\begin{aligned}&\mathbf{E}_{\mathscr{F}_\theta}\exp(\delta|B_1|)I(Q_1^c)\exp(\delta|B_2|)\cdots I(Q_{n-1}^c)\exp(\delta|B_n|)\\ &\quad=\mathbf{E}_{\mathscr{F}_\theta}\exp(\delta|B_1|)I(Q_1^c)\exp(\delta|B_2|)\cdots I(Q_{n-2}^c)\exp(\delta|_{n-1}|)\\ &\quad\quad\cdot\mathbf{E}(I(Q_{n-1}^c)\exp(\delta|B_n|)\,|\,\mathscr{F}_{\theta_{n-1}}).\end{aligned}$$

Using the Cauchy inequality, we obtain

$$\begin{aligned}\mathbf{E}(I(Q_{n-1}^c)\exp(\delta|B_n|)\,|\,\mathscr{F}_{\theta_{n-1}})&=\mathbf{E}(I(Q_{n-1}^c)\mathbf{E}_{\theta_n,x^*}\exp(\delta|B_n|)\,|\,\mathscr{F}_{\theta_{n-1}})\\ &\leqq\mathbf{E}(I(Q_{n-1}^c)\exp(\varepsilon h_R(x^*))\,|\,\mathscr{F}_{\theta_{n-1}})\\ &\leqq(\mathbf{E}(I(Q_{n-1}^c)\,|\,\mathscr{F}_{\theta_{n-1}}))^{1/2}(\mathbf{E}(\exp(\varepsilon h_R(x^*))\,|\,\mathscr{F}_{\theta_{n-1}}))^{1/2}\\ &\leqq(1-q)^{1/2}(\mathbf{E}(\exp(\varepsilon h_R(x^*))\,|\,\mathscr{F}_{\theta_{n-1}}))^{1/2},\end{aligned}$$

where $x^*=x^2_{\theta_n}$ and since n is odd, $|x^2_{\theta_{n-1}}|\leqq 1$; whence, by Lemma 2,

$$\mathbf{E}(\exp(\varepsilon h_R(x^2_{\theta_n}))\,|\,\mathscr{F}_{\theta_{n-1}})\leqq C(\varepsilon)\exp(\varepsilon).$$

Thus

$$\begin{aligned}&\mathbf{E}_{\mathscr{F}_\theta}\exp(\delta|B_1|)I(Q_1^c)\exp(\delta|B_2|)\cdots I(Q_{n-1}^c)\exp(\delta|B_n|)\\ &\quad\leqq C(\varepsilon)\exp(\varepsilon)(1-q)^{1/2}\mathbf{E}_{\mathscr{F}_\theta}\exp(\delta|B_1|)I(Q_1^c)\cdots\exp(\delta|B_{n-1}|).\end{aligned}$$

By induction we conclude that

$$\begin{aligned}&\mathbf{E}_{\mathscr{F}_\theta}\exp(\delta|B_1|)I(Q_1^c)\exp(\delta|B_2|)\cdots I(Q_{n-1}^c)\exp(\delta|B_n|)\\ &\quad\leqq(C(\varepsilon)\exp(\varepsilon)(1-q)^{1/2})^{n-1}\mathbf{E}_{\mathscr{F}_\theta}\exp(\delta|B_1|).\end{aligned}$$

Therefore,

$$\mathbf{E}_{\mathscr{F}_\theta}\exp(\delta|\bar{B}|)\leqq\mathbf{E}_{\mathscr{F}_\theta}\exp(\delta|B_1|)\sum_{n=1}^{\infty}(C(\varepsilon)\exp(\varepsilon)(1-q)^{1/2})^{n-1}. \tag{15}$$

By Lemma 2, one can make, by the choice of $\varepsilon>0$, $C(\varepsilon)$ and $\exp(\varepsilon)$ so close to 1 that $C(\varepsilon)\exp(\varepsilon)(1-q)^{1/2}<1$, and then the series in the right-hand side of (15) converges and its sum tends to 1 as $\varepsilon\downarrow 0$. It remains to estimate $\mathbf{E}_{\mathscr{F}_\theta}\exp(\delta|B_1|)$.

If the beginning part of a word is considered, then by Lemmas 1 and 2

$$\mathbf{E}_{\mathcal{F}_\theta} \exp(\delta|B_1|) = \mathbf{E}_{t,x_t^1} \exp(\delta|B_1|) \leqq \exp(\varepsilon h_R(x_t^1)).$$

If a part of a word not at the beginning is considered, then by the construction $|x_\theta^1| \leqq 1$, $|x_\theta^2| \leqq 1$. Therefore, again

$$\mathbf{E}_{\mathcal{F}_\theta} \exp(\delta|B_1|) \leqq \mathbf{E}_{\mathcal{F}_\theta} \exp(\varepsilon h_R(x_\theta^1)) \leqq \exp(\varepsilon h_R((1, 0, \cdots, 0))).$$

Thus, for the first letter A (denote it by $\bar{A}_1$) we obtain

$$\mathbf{E}_{\mathcal{F}_\theta} \exp(\delta|\bar{A}_1|) \leqq C_\varepsilon \exp(\varepsilon h_R(x_t^1)),$$

and for all other $\bar{A}$, introducing the notation $\bar{A}_k$, $k = 2, 3, \cdots$, we find that

$$\mathbf{E}_{\mathcal{F}_\theta} \exp(\delta|\bar{A}_k|) \leqq C_\varepsilon,$$

where $C_\varepsilon \downarrow 1$ as $\varepsilon \downarrow 0$.

By Lemma 2,

$$\mathbf{E} \exp(\delta|\bar{A}_1|) \leqq C_\varepsilon \mathbf{E} \exp(\varepsilon h_R(x_t^1)) \leqq C_\varepsilon \exp(\varepsilon(|x| \vee 2)).$$

Now we can get a bound for $\mathbf{E} \exp(\delta L)$. Let N_i denote the realization of the letter N after $\bar{A}_i'$, $i = 1, 2, \cdots$. We have

$$\begin{aligned} \mathbf{E} \exp(\delta L) \leqq \mathbf{E}(&\exp(\delta(|\bar{A}_1|+2)) + \exp(\delta(|\bar{A}_1|+2)) \\ &\cdot I(N_1) \exp(\delta(|\bar{A}_2|+2)) + \cdots + \exp(\delta(|\bar{A}_1|+2)) \\ &\cdot I(N_1) \exp(\delta(|\bar{A}_2|+2)) I(N_2) \cdots \exp(\delta(|\bar{A}_n|+2)) + \cdots). \end{aligned}$$

Proceeding in exactly the same way as in estimating $\mathbf{E} \exp(\delta|\bar{B}|)$, we find that, for sufficiently small $\delta > 0$ and $\varepsilon > 0$,

$$\mathbf{E} \exp(\delta L) \leqq C_\varepsilon \sum_{n=0}^{\infty} (C_\varepsilon p)^n, \tag{16}$$

where $p \in (0, 1)$ is an upper bound of the probability of realization of the event N_i for each i. Here the constant C_ε is also close to 1 for sufficiently small ε and therefore the right-hand side of (16) is finite. By (13) we obtain the desired bound for the coefficient $\beta(s)$. The theorem is proved.

Remarks. (1) The gluing method is known to be equivalent to the classical technique (see [14]) and essentially reduces to checking some Doeblin type condition (see [14, Chap. V]). In the diffusion case one has to check such a condition at random times since the author does not know of an upper bound for the distribution densities of two diffusion processes of the form studied at a fixed time. A similar method of stopping a diffusion process in another problem is introduced in [15, Chap. IV].

(2) If one considers the stochastic equation (1)-(2) not on E^d but on a compact manifold, then an exponential bound similar to (5) also holds for the uniformly strong mixing coefficient (see [4]). The method of proof, at least for the homogeneous case, is well known [4]. For its realization one must also use the result of N. V. Krylov and M. V. Safonov, Harnack's inequality, in verifying the Doeblin type condition at random times.

REFERENCES

[1] I. A. IBRAGIMOV AND YU. V. LINNIK, *Independent and Stationary Sequence of Random Variables*, Wolters-Noordhoff, Groningen, 1971.

[2] A. D. VENT-TSEL' AND M. I. FRIEDLIN, *Fluctuations in Dynamical Systems under the Action of Small Random Perturbations*, Nauka, Moscow, 1979. (In Russian.)

[3] V. A. STATULYVAICHUS, *Theorems of large deviations for sums of dependent random variables*, I, Litovsk. Mat. Sb., 19 (1979), pp. 199-208. (In (Russian.).

[4] YU. A. DAVYDOV, *Mixing conditions for Markov chains*, Theory Probab. Appl., 18 (1973), pp. 312-329.

[5] N. V. KRYLOV, *On Itô stochastic integral equations*, Theory Probab. Appl., 14 (1969), pp. 330-336.

[6] ———, *On extracting a Markov process from a Markov system and construction of quasidiffusion processes*, Izv. Akad. Nauk SSSR Ser. Mat., 37 (1973), pp. 691-708. (In Russian.)

[7] ———, *Once more about the connection between elliptic operators and Itô's stochastic equations*, in Statistics and control of stochastic processes, Steklov seminar, 1984, N. V. Krylov, ed., Optimization Software, New York, 1985, pp. 214-229.

[8] E. B. DYNKIN, *Markov Processes*, Academic Press, New York, 1965.

[9] H. MCKEAN, *Stochastic Integrals*, Academic Press, New York, 1969.

[10] D. GRIFFEATH, *A maximal coupling for Markov chains*, Z. Wahrsch. Verw. Gebiete, 31 (1975), pp. ·95-106.

[11] D. STROOCK AND S. R. S. VARADHAN, *On the support of diffusion processes with applications to the strong maximal principle*, in Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, V, III, Univ. Calif. Press, Berkeley-Los Angeles, 1970, pp. 333-360.

[12] N. V. KRYLOV AND M. V. SAFONOV, *A certain property of the solutions of parabolic equations with measurable coefficients*, Izv. Akad. Nauk SSSR, 44 (1980), pp. 161-175. (In Russian.)

[13] D. STROOCK AND S. R. S. VARADHAN, *Multidimensional Diffusion Processes*, Springer, Berlin, 1979.

[14] J. L. DOOB, *Stochastic Processes*, John Wiley, New York, 1953.

[15] R. Z. KHAS'MINSKII, *Stability of Systems of Differential Equations under Random Perturbations of the Parameter*, Nauka, Moscow, 1969. (In Russian.)