Boletín de la Sociedad Matemática Mexicana Vol. 24 No. 1, 1979

LYAPUNOV CRITERIA FOR STABILITY OF DIFFERENTIAL EQUATIONS WITH MARKOV PARAMETERS

BY ONÉSIMO HERNÁNDEZ-LERMA

Introduction

In this paper we use Lyapunov-like techniques to study stability with probability 1 and weak stochastic stability of differential equations of the form

$$(*) \qquad \dot{x}(t) = F(x(t), y(t)), \qquad x \in R^n, y \in R^m$$

where $y(t)$ is a given time-homogeneous Markov process.

Lyapunov's direct method to study the stability of stochastic systems was introduced by Bertram and Sarachik [2] and, independently, by Kats and Krasovskii [9]. In [2] the type of stability considered is "stability in the mean," while in [9] Kats and Krasovskii considered a form of probabilistic stability weaker than stability w.p.1. In both of these works the noise process $y(t)$ in (*) is a Markov process with finite state space. Following the work of Kats and Krasovskii, Khasminskii [12] studied the stability w.p.1 and stability in the mean for stochastic differential equations of Ito type. In turn, Khasminskii's work for Ito equations initiated a fruitful area of research pursued by many other authors, notably Kushner, whose monograph [14] is still the standard reference on the subject. Recent results on stability for Ito equations are given in [6], [8].

Stability with probability 1 (defined in §2) is a type of "pathwise" stability with respect to an equilibrium solution of equation (*). A different type of stability is what Wonham [21] calls weak stochastic stability. Namely, the system (*) is said to be weakly stochastically stable if the solution process $x(t)$ has an invariant probability distribution. Wonham used Lyapunov techniques to study this form of stability for Ito equations. A similar approach was also followed by Kushner [15, 16], and Zakai [23].

For a short, readable description of the several types of stochastic stability and the work of other authors we refer to Kozin's survey [13]. For applications of systems of the form (*) in demography, engineering, physics, etc., see, for instance, [4], [18], [20].

The paper is divided in two parts. Part I(§§1–6) on stability w.p.1, can be described as follows: In §1 we state the basic assumptions on the noise process $y(t)$ and the system (*), and some of the properties of $x(t)$ and the joint process $(x(t), y(t))$ are reviewed. The definitions of stability w.p.1 and Lyapunov functions are given in §2. The results in §3 are the stochastic counterparts of the classical (deterministic) Lyapunov stability theorems. These results parallel the work of other authors [9], [12], [14], etc. for several types of stochastic equations and they are the natural extensions to stochastic systems of well-known results for deterministic differential equations [17]. In §§4–5 we state other instability and stability theorems, and then, in §6, we mention some examples.

Part II, on weak stochastic stability, consists of §§7–8. After some preliminaries (§7), we state in §8 the corresponding Lyapunov theorems. To prove these we use necessary and sufficient conditions for the existence of invariant probability measures given by Benes [1]. This approach was first followed by Zakai [23] for Ito equations.

Notation. Vectors are written as column matrices. If A is a matrix, A^* denotes its transpose and $\mathrm{Tr}(A)$ its trace. Given a function $v:R^n \to R^1$, $v_x(v_{xx})$ denotes the vector (matrix) of first (second) partial derivatives of v. That is, v_x and v_{xx} denote the gradient and the Hessian matrix of v, respectively. Random variables are tacitly referred to an underlying probability space $(\Omega, \mathfrak{F}, P)$. E denotes expectation; $E_{t,x}(P_{t,x})$ denotes expectation (probability) conditional on the event $x(t) = x$. If $t = 0$, we write $E_{0,x} = E_x$ and $P_{0,x} = P_x$. To minimize the numbering of formulas, and if no confusion may arise, some expressions which are referred to only within the same section or proof of a theorem are marked with asterisks (*).

Part I: Stability w.p.1

1. *Preliminaries and basic assumptions.* Throughout this work it is assumed that $y(t)$, $t \geq 0$, is a time-homogeneous Markov process with state space S in R^m. Furthermore, $y(t)$ is a Feller process and is right-continuous with finite limits from the left. Since $y(t)$ takes values in an Euclidean space, we can assume that $y(t)$ is a separable process [7]. The (weak infinitesimal) generator of $y(t)$ will be denoted by Q.

The right-continuity and Feller assumptions imply that $y(t)$ is strong Markov. On the other hand, if $y(t)$ is a jump process, the above hypotheses insure that w.p.1 on any finite interval of time only finitely many jumps occur. For definitions and results on Markov processes we refer to the books by Dynkin [5] or Gikhman and Skorokhod [7].

Given $y(t)$ as above, consider the n-dimensional process $x(t)$ defined by

$$\dot{x}(t) = F(x(t), y(t)), \qquad t \geq 0, x(0) = x. \tag{1.1}$$

We assume the following conditions on F and $y(t)$:

(1.2) *Assumptions.* The function $F: R^n \times S \to R^n$ is continuous in both variables $x \in R^n$, $y \in S$, and satisfies:

(a) there is a constant c such that

$$|F(x, y)| \leq c(1 + |x|) \quad \text{for all} \quad x \in R^n, y \in S.$$

(b) for any compact $K \subset R^n$ there is a constant $c = c_K$ such that

$$|F(x, y) - F(x', y)| \leq c|x - x'|$$

for all x, x' in K, and all $y \in S$.

In addition, the Markov process $y(t)$ is such that:

(c) the joint process $(x(t), y(t))$ satisfies the Feller property (see remark (1.3) below).

Conditions (1.2a, b) insure the existence of a unique solution of (1.1). On the other hand, even though $x(t)$ itself is not necessarily Markov, it is a well known fact (see, for instance, [22]) that the joint process $(x(t), y(t))$ is Markov. Furthermore, assumption (1.2c) implies that $(x(t), y(t))$ is strong Markov [7].

(1.3) *Remark.* Let us briefly recall what the Feller property means. For bounded measurable functions v on $R^n \times S$ and $t \geq 0$, define the operators T_t as

$$(T_t v)(x, y) = E_{x,y} v(x(t), y(t)).$$

The Markov process $(x(t), y(t))$ is said to be a Feller process if

$$v \in C \Rightarrow T_t v \in C,$$

where $C = C(R^n \times S)$ is the space of continuous bounded functions on $R^n \times S$. Assumption (1.2c) holds in many cases of interest. For instance, it can be easily verified in the case in which $y(t)$ is a Markov process with finitely many states, or more generally, when the state space S (finite or infinite) of $y(t)$ consists of isolated points (e.g. the integers). It also holds if $y(t)$ is a diffusion process with drift and covariance coefficients satisfying the usual growth and Lipschitz conditions [6, vol. 1], [8].

In analogy with stochastic differential equations of Ito type we have:

PROPOSITION 1.4. *Under assumptions* (1.2),

$$(a)\ E_{x,y} |x(t)|^{2p} \leq (1 + |x|^{2p}) e^{qt}$$

$$(b)\ E_{x,y} |x(t) - x|^{2p} \leq \bar{q}(1 + |x|^{2p}) t^p e^{qt}$$

for $t \geq 0, p = 1, 2, \cdots$, *where* q *and* $\bar{q}$ *are constants depending only on* p *and the constant* c *in* (1.2).

The proof of this result is essentially the same as for Ito equations; see, for instance [6] or [8].

The generator of $(x(t), y(t))$. Let B denote the Banach space of measurable bounded functions on $R^n \times S$ with uniform norm $\| v \| = \sup_{(x,y)} | v(x, y) |$, and let T_t be the operators defined in (1.3) for $v \in B$. (The operators T_t, $t \geq 0$, form a contraction semigroup: $T_{s+t} = T_s T_t$ for all $s, t \geq 0$, and $\| T_t v \| \leq \| v \|$ for all $v \in B$). If $v, v_k \in B$, we say that $v = \lim v_k$ if $v_k(x, y) \to v(x, y)$ for all $(x, y) \in R^n \times S$ and if $\| v_k \|$ are bounded. (That is, v is the *weak* limit of v_k; it is the "strong" limit if $\| v_k - v \| \to 0$.) Let B_0 be the set of functions $v \in B$ such that $\lim_{t \to 0} T_t v = v$. If for $v \in B_0$ the limit

$$Lv = \lim_{t \to 0} \frac{1}{t} (T_t v - v)$$

or more explicitly,

(1.5) $$Lv(x, y) = \lim_{t\to 0} \frac{1}{t} (E_{x,y}v(x(t), y(t)) - v(x, y))$$

exists and satisfies that $Lv \in B_0$, that is,

$$\lim_{t\to 0} T_t Lv = Lv,$$

we then say that "v is in the domain of L" and the operator L is called the (weak infinitesimal) generator of the process $(x(t), y(t))$; see Dynkin [5]. The domain of L includes the set of continuous bounded functions $v(x, y)$ on $R^n \times S$ such that v is of class C^1 in x, with bounded gradient v_x, and, as a function of y, $v(x, \cdot)$ is in the domain of the generator Q of $y(t)$. Henceforth, "v in the domain of L" means that v has at least these properties. In this case, Lv is given by

(1.6) $$Lv(x, y) = F(x, y)^* v_x(x, y) + Qv(x, y).$$

For instance, if $y(t)$ is a jump Markov process, then Lv is defined for continuous bounded functions $v(x, y)$ of class C^1 in x, with bounded gradient v_x. In particular, if $y(t)$ is a Markov chain with a finite state space $S = \{1, \cdots, N\}$ and infinitesimal matrix $Q = (q_{ij})$, we can write (1.6) as

(1.6a) $$Lv(x, y) = F(x, y)^* v_x(x, y) + \textstyle\sum_{j=1}^{N} q_{yj} v(x, j), \; y = 1, \cdots, N.$$

Here, Q is a $N \times N$ matrix such that $q_{yj} \geq 0$ for $j \neq y$ and $q_{yy} = -\sum_{j\neq y} q_{yj}$. If $y(t)$ is a diffusion process with drift vector $b(y)$ and diffusion matrix $a(y) = (a_{ij}(y))$, and $v(x, y)$ is a continuous bounded function of class C^1 in x, of class C^2 in y, and such that v_x, v_y, and v_{yy} are bounded, then Lv is given by

(1.6b) $$Lv(x, y) = F(x, y)^* v_x(x, y) + b(y)^* v_y(x, y) + \tfrac{1}{2} \operatorname{Tr}(a(y) v_{yy}).$$

Sometimes, it is convenient to interpret $Lv(x, y)$ as some sort of average value of the derivative of the function $v(x(t), y(t))$ along all the realizations of the process $(x(t), y(t))$ issuing from the point (x, y) at time $t = 0$. Consistent with this interpretation, we have the important formula [5]

(1.7) $$E_{x,y}v(x(T), y(T)) = v(x, y) + E_{x,y} \int_0^T Lv(x(s), y(s))\, ds,$$

which holds for any function v in the domain of L and any stopping time T for the process $(x(t), y(t))$, if $E_{x,y}(T) < \infty$.

2. Stability w.p.1 and Lyapunov functions

In addition to (1.2) we assume that

$$F(0, y) = 0 \quad \text{for all} \quad y \in S,$$

so that $x(t) = 0$ is an equilibrium solution of (1.1). Then, by the assumptions (1.2a,b), standard results for ordinary differential equations imply that, for any $y \in S$.

$$P_{x,y}(x(t) = 0 \quad \text{for all} \quad t > 0) = 1 \quad \text{if} \quad x = x(0) = 0$$

and

$$P_{x,y}(x(t) = 0 \quad \text{for some} \quad t > 0) = 0 \quad \text{if} \quad x = x(0) \neq 0.$$

(2.1) *Definition.* The system (1.1) (or the equilibrium solution $x(t) \equiv 0$) is said to be

(a) *stable* w.p.1 if, for any $\epsilon > 0$ and $y \in S$,

$$\lim_{x \to 0} P_{x,y}(\sup_{t \geq 0} |x(t)| \geq \epsilon) = 0$$

(b) *asymptotically stable* w.p.1 if it is stable w.p.1 and, in addition, for any $y \in S$,

$$\lim_{x \to 0} P_{x,y}(\lim_{t \to \infty} x(t) = 0) = 1$$

(c) *asymptotically stable in the large* w.p.1 if it is stable w.p.1 and, furthermore,

$$P_{x,y}(\lim_{t \to \infty} x(t) = 0) = 1$$

for any $x \in R^n$ and $y \in S$.

It may happen that some of the conditions (a), (b), or (c) in (2.1) hold only for values of y in a subset S_1 of S. In that case, we restrict the corresponding condition in (2.1) to values $y \in S_1$.

We introduce now the concepts of positive-definite function and Lyapunov function.

(2.2) *Definition.* (a) A continuous function $w(x)$ defined on an open sphere $B_h = \{x \in R^n : |x| < h\}$ is said to be *positive-definite* (in the sense of Lyapunov) if

$$w(0) = 0 \quad \text{and} \quad w(x) > 0 \quad \text{for} \quad x \neq 0.$$

(b) A continuous function $v(x, y)$ on $B_h \times S$ is said to be *positive-definite* if $v(0, y) = 0$ for all y, and if there exists a positive-definite function $w(x)$ such that

$$v(x, y) \geq w(x) \quad \text{for all} \quad x \in B_h, y \in S.$$

(c) A function v is *negative-definite* if $-v$ is positive-definite.

(2.3) *Definition.* A positive-definite function v on $B_h \times S$, for some $h \geq 0$, is called a *Lyapunov function* for the system (1.1) if v is in the domain of L and $Lv(x, y) \leq 0$ $(x \in B_h, y \in S)$.

Given that $(x(0), y(0)) = (x, y)$, with $x \in B_h$, let $\tau = \tau_{x,y}$ be the exit time from B_h of $x(t)$. Let $(\tilde{x}(t), \tilde{y}(t))$, $t \geq 0$, be the process defined by

$$\tilde{x}(t) = x(t \wedge \tau), \qquad \tilde{y}(t) = y(t \wedge \tau), \tag{2.4}$$

where $t \wedge \tau = \min(t, \tau)$, and let $v(x, y)$, $x \in B_h$, $y \in S$, be a Lyapunov function for system (1.1). Then, since $Lv \leq 0$, it follows from (1.7) that

$$E_{x,y}v(x(t \wedge \tau), y(t \wedge \tau)) \leq v(x, y)$$

for any $t \geq 0$, $x \in B_h$, $y \in S$. This proves the following important lemma originally due to Bucy [3]; see also Kushner [14].

LEMMA 2.5. *If $v(x, y)$ is a Lyapunov function for equation* (1.1), *then $v(\tilde{x}(t), \tilde{y}(t))$, $t \geq 0$, is a nonnegative supermartingale. Therefore, by the supermartingale inequality* (see, for instance, Dynkin [5], vol. 2)

$$P_{x,y}(\sup_{t\geq 0} v(\tilde{x}(t), \tilde{y}(t)) \geq \epsilon) \leq \frac{1}{\epsilon} v(x, y)$$

for any $\epsilon > 0$.

3. Lyapunov Theorems

The Lyapunov theorems for stability w.p.1 are the following (cf. LaSalle and Lefschetz [17] for deterministic systems).

THEOREM 3.1. *If there exists a Lyapunov function $v(x, y)$* for the system (1.1), *then the system is stable* w.p.1.

THEOREM 3.2. *Let $v(x, y)$ be a function defined on $B_h \times S$ for some $h > 0$, and assume that v is a Lyapunov function for the system* (1.1) *and such that*

(a) *Lv is negative-definite.*

If such a function v exists, then the system is asymptotically stable w.p.1. *If, in addition, v is defined for all $x \in R^n$ and $y \in S$, and satisfies:*

(b) *there exists a positive-definite function $w(x)$,*

$$x \in R^n, \quad \text{such that} \quad w(x) \to \infty \quad \text{as} \quad |x| \to \infty,$$

and

$$v(x, y) \geq w(x) \quad \text{for all} \quad x \in R^n, y \in S,$$

then the system (1.1) *is asymptotically stable in the large* w.p.1.

Remarks. (1) If $v(x, y)$ is a Lyapunov function, then in particular, v vanishes at $x = 0$ and has bounded gradient v_x. Therefore, by the mean value theorem,

$$(0 \leq)\ v(x, y) = x^* v_x(x', y),$$

which implies that $v(x, y) \to 0$ as $|x| \to 0$, uniformly on y. Thus following the European terminology for nonautonomous ODE's, we might refer to the latter property by saying that "v is decrescent," or that "v has an arbitrarily small upper bound." Likewise, we might refer to property (3.2b) by saying that "v is radially unbounded." (2) Condition (3.2b) implies, of course, that $v(x, y)$ is unbounded on $R^n \times S$. In this case, the requirement that v is in the domain of L is interpreted as follows: For each $h > 0 (h \to \infty)$, $v(x, y)$ is a bounded measurable function in the domain of L for $|x| < h$, $y \in S$, and the process $(x(t), y(t))$ in (1.5) is now the corresponding stopped process in (2.4); cf. [14,

§1.3]. A similar interpretation is assumed in all our results in which v is unbounded.

Proof of Theorem (3.1). Suppose that $v(x, y)$ is a Lyapunov function defined on $B_h \times S$ for some $h > 0$, and let $(\tilde{x}(t), \tilde{y}(t))$ be the process defined in (2.4). Then, for any $\epsilon > 0 (\epsilon < h)$, there exists a positive number $\epsilon_1 = \epsilon_1(\epsilon)$ such that, for any $y \in S$ and $|x| < h$,

$$
\begin{aligned}
P_{x,y}(\sup_{t\geq 0}|x(t)| \geq \epsilon) &= P_{x,y}(\sup_{t\geq 0}|\tilde{x}(t)| \geq \epsilon) \\
&\leq P_{x,y}(\sup_{t\geq 0} v(\tilde{x}(t), \tilde{y}(t)) \geq \epsilon_1) \\
&\leq \frac{1}{\epsilon_1} v(x, y)
\end{aligned}
$$

by the supermartingale inequality in Lemma (2.5). Letting $x \to 0$ we obtain Theorem (3.1).

We give next a slightly different proof of Theorem (3.1) in terms of the exit time of $x(t)$ from a neighborhood of the origin. This approach will be useful in the proof of later theorems.

Another Proof of Theorem (3.1). As in the previous proof, let $v(x, y)$ be a Lyapunov function defined on $B_h \times S$. Clearly, Definition (2.1a) of stability w.p.1 is equivalent to: For any $\epsilon > 0$ and $y \in S$,

$$
(*) \qquad \lim_{x\to 0} P_{x,y}(\tau^\epsilon < \infty) = 0,
$$

where $\tau^\epsilon = \inf\{t > 0 : |x(t)| \geq \epsilon\}$.

Let $I(A)$ denote the indicator function of the event A. Then, for $n = 1, 2, \cdots$, we have:

$$
\begin{aligned}
&E_{x,y}[v(x(\tau^\epsilon \wedge n), \quad y(\tau^\epsilon \wedge n))I(\tau^\epsilon < \infty)] \\
&\leq E_{x,y} v(x(\tau^\epsilon \wedge n), y(\tau^\epsilon \wedge n)) \\
&= v(x, y) + E_{x,y} \int_0^{\tau^\epsilon \wedge n} Lv(x(r), y(r))\, dr \quad \text{by (1.7)} \\
&\leq v(x, y).
\end{aligned}
$$

Letting $n \to \infty$, we obtain

$$
E_{x,y}[v(x(\tau^\epsilon), y(\tau^\epsilon))I(\tau^\epsilon < \infty)] \leq v(x, y),
$$

and therefore, if $\epsilon_1 = \epsilon_1(\epsilon) > 0$ denotes a lower bound for $v(x, y)$ when $|x| = \epsilon$, we have that

$$
\epsilon_1 P_{x,y}(\tau^\epsilon < \infty) \leq v(x, y).
$$

Finally, letting $x \to 0$, we obtain (*).

Proof of Theorem (3.2). First, we shall prove asymptotic stability. Let $v(x, y)$ be a Lyapunov function on $B_h \times S$ which satisfies conditions (3.2a), and let $(\tilde{x}(t), \tilde{y}(t))$, $t \geq 0$, be the process defined in (2.4). Since $v(\tilde{x}(t), \tilde{y}(t))$ is a nonnegative supermartingale, it follows from the martingale convergence theorem that there exists a nonnegative random variable $V(x, y)$, which may depend on the initial states $x(0) = x$, $y(0) = y$, $x \in B_h$, such that

$$v(\tilde{x}(t), \tilde{y}(t)) \rightarrow V(x, y) \quad \text{as} \quad t \rightarrow \infty, \text{ w.p.1},$$

and, therefore,

$$\begin{aligned} P_{x,y}(\lim_{t\rightarrow\infty} x(t) = 0) &\geq P_{x,y}(\lim_{t\rightarrow\infty} v(\tilde{x}(t), \tilde{y}(t)) = 0) \\ &= P(V(x, y) = 0), \end{aligned}$$

where $P(V(x, y) = 0)$ means $P_{x,y}(V = 0)$. By Theorem (3.1), the system (1.1) is stable w.p.1. Thus, to prove asymptotic stability it is enough to show that, for any $y \in S$,

$$(*) \qquad \lim_{x\rightarrow 0} P(V(x, y) = 0) = 1,$$

or equivalently, that or any $y \in S$ and $\eta > 0$, there exists a positive $\delta = \delta(y, \eta)$ such that

$$P(V(x, y) > 0) < \eta \qquad \text{if} \quad |x| < \delta,$$

Arguing by contradiction, let us assume that this is not true. Then, there exists $y_0 \in S$ and $\eta_0 > 0$ such that for any $\delta > 0$, $P(V(x_0, y_0) > c_0) > \eta_0$ for some $|x_0| < \delta$, and some $c_0 > 0$. Therefore, with probability at least η_0, there exists a random time T [5, §4.1] such that

$$v(\tilde{x}(t), \tilde{y}(t)) \geq \tfrac{1}{2} c_0 \qquad \text{for} \quad t > T,$$

where $(x(0), y(0)) = (x_0, y_0)$, which, in turn, implies that

$$|\tilde{x}(t)| \geq c_1 \qquad \text{for} \quad t \geq T,$$

for some $c_1 > 0$. On the other hand, by (3.2a), there exists $c_2 > 0$ such that

$$Lv(x, y) \leq -c_2 \qquad \text{if} \quad |x| \geq c_1.$$

Therefore, denoting by $I(A)$ the indicator function of an event A and using (1.7), we have:

$$\begin{aligned} 0 &\leq E_{x_0,y_0} v(\tilde{x}(t), \tilde{y}(t)) \\ &= v(x_0, y_0) + E_{x_0,y_0}[I(t < T) + I(t \geq T)] \\ &\qquad \cdot \int_0^t Lv(\tilde{x}(s), \tilde{y}(s))\, ds \\ &\leq v(x_0, y_0) + E_{x_0,y_0} I(t \geq T) \int_T^t Lv(\tilde{x}(s), \tilde{y}(s))\, ds \end{aligned}$$

$$\leq v(x_0, y_0) - c_2 E_{x_0,y_0} I(t \geq T)(t - T)$$

and letting t approach ∞, we get a contradiction. Therefore, (*) must be true. This proves asymptotic stability.

To prove the second part of Theorem (3.2) (asymptotic stability in the large), we assume now that $v(x, y)$ is a Lyapunov function defined for all $x \in R^n$ and $y \in S$, and that v satisfies (3.2a, b). The proof is reduced to showing (Lemma (3.3) below) that given *any* neighborhood $B_h = \{x \in R^n: |x| < h\}$, $h > 0$, of the origin, and *any* initial condition $x(0) = x \in R^n$, the trajectories $x(t)$ starting at x always reach B_h in a *finite* time w.p.1. Having that $x(t)$ has reached the neighborhood B_h, and choosing h sufficiently small, we can apply the first part of the theorem ("local" asymptotic stability) and the strong Markov property of the process $(x(t), y(t))$ to obtain the desired conclusion (asymptotic stability in the large).

LEMMA 3.3. *Let $v(x, y)$, $x \in R^n$, $y \in S$, be a Lyapunov function satisfying assumptions* (a) *and* (c) *of Theorem* (3.2). *Then for any $h > 0$, and any $x \notin B_h$, $y \in S$,*

$$P_{x,y}(x(t) \in B_h \quad \textit{for some} \quad t > 0) = 1.$$

Proof. Let r_1 and r_2 be positive numbers, with $r_1 < h < r_2$, such that $x(0) = x$ is in the set $A = \{x: r_1 < |x| < r_2\}$, and define $T = \inf\{t > 0: x(t) \notin A\}$. If $-c = \sup_{x \in A, y \in S} Lv(x, y)$, with $c > 0$ by (3.2a), we see from (1.7) that

$$(*) \qquad \begin{aligned} Ev(x(T \wedge n), y(T \wedge n)) &= v(x, y) + E\int_0^{T\wedge n} Lv(x(s), y(s))\, ds \\ &\leq v(x, y) - cE(T \wedge n), \end{aligned}$$

where $E = E_{x,y}$. Therefore, since $v \geq 0$,

$$E(T \wedge n) \leq c^{-1} v(x, y).$$

Letting $n \to \infty$, this shows that $E(T) < \infty$ and, hence, $P(T < \infty) = 1$. In turn, this implies that

$$P_{x,y}(|x(T)| = r_1) + P_{x,y}(|x(T)| = r_2) = 1.$$

Let

$$v_i = \inf\{v(x, y): |x| = r_i, y \in S\} \qquad i = 1, 2.$$

From (*),

$$v(x, y) \geq Ev(x(T \wedge n), y(T \wedge n))$$

and letting $n \to \infty$, we see that

$$\begin{aligned} v(x, y) &\geq Ev(x(T), y(T)) \\ &\geq v_1 P(|x(T)| = r_1) + v_2 P(|x(T)| = r_2) \\ &= v_1 + (v_2 - v_1) P(|x(T)| = r_2), \end{aligned}$$

where $P = P_{x,y}$. By assumption (3.2b), $v_2 > v_1$ for r_2 sufficiently large, so that

$$P(|x(T)| = r_2) \leq (v(x, y) - v_1)/(v_2 - v_1),$$

and the right side of this inequality tends to zero when $r_2 \to \infty$. Therefore, since

$$\{x(t) \notin B_h \quad \text{for all} \quad t\} \subset \bigcap_{r_2<\infty} \{|x(T)| = r_2\},$$

we obtain that $P(x(t) \notin B_h$ for all $t \geq 0) = 0$. This completes the proof of the Lemma and also (see the paragraph preceding the statement of Lemma (3.3)) completes the proof of Theorem (3.2).

By definition (cf. Wonham [21]), the process $x(t)$ is said to be *recurrent* if there exists a compact set $K \subset R^n$ such that for any $x \in R^n$, $y \in S$,

$$P_{x,y}(x(t) \in K \quad \text{for some} \quad t > 0) = 1.$$

Taking K as the closure of B_h in Lemma (3.3), we obtain the following corollary:

COROLLARY 3.4. *The hypotheses of Lemma* (3.3) *imply that the process* $x(t)$, $t \geq 0$, *is recurrent.*

This result is related with the existence of invariant distributions for the process $x(t)$ (§§7, 8).

4. Instability

A proof similar to that of Lemma (3.3) gives the following instability theorem; cf. Khasminski [12, Theor. 2.3].

THEOREM 4.1. *Let* $v(x, y)$ *be a nonnegative function defined for* $0 < |x| < h$, $y \in S$, (*where* $0 < h \leq \infty$ *is arbitrary*) *and such that*

(a) $v(x, y) \to \infty$ *as* $x \to 0$,

(b) v *is in the domain of* L *and* $Lv(x, y) \leq -c(0 < |x| < h, y \in S)$, *for some* $c > 0$.

Then the system (1.1) *is unstable; in fact, it is "uniformly unstable," in the sense that for any* $\epsilon > 0 (\epsilon < h)$

$$P_{x,y}(\tau^\epsilon < \infty) = 1 \quad \textit{for all} \quad 0 < |x| < \epsilon, y \in S,$$

where $\tau^\epsilon = \inf\{t > 0: |x(t)| \geq \epsilon\}$.

Proof. Let ϵ_1 be a positive number such that $0 < \epsilon_1 < \epsilon$ and $x = x(0)$ is in the set $A = \{x \in R^n: \epsilon_1 < |x| < \epsilon\}$. Let $T = \inf\{t > 0: x(t) \notin A\}$. Let us write $P_{x,y} = P$. By assumption (b), the same argument that we used in the proof of Lemma (3.3) gives that $P(T < \infty) = 1$, so that

$$(*) \qquad P(|x(T)| = \epsilon_1) + P(|x(T)| = \epsilon) = 1,$$

and

$$v(x, y) \geq v_1 P(|x(T)| = \epsilon_1) + v_0 P(|x(T)| = \epsilon)$$
$$= (v_1 - v_0)P(|x(T)| = \epsilon_1) + v_0,$$

where $v_1 = \inf(v(x, y): |x| = \epsilon_1, y \in S)$, and $v_0 = \inf(v(x, y): |x| = \epsilon, y \in S)$. Taking ϵ_1 sufficiently small, assumption (a) gives $v_1 - v_0 > 0$ and, therefore,

$$P(|x(T)| = \epsilon_1) \leq (v(x, y) - v_0)/(v_1 - v_0).$$

Using (a) once more, we see that the right side of this inequality tends to zero when $\epsilon_1 \to 0$. From this and (*), the theorem follows.

5. Other stability theorems

In Theorem (4.1) we showed (roughly) that if $v(x, y) \to +\infty$ as $x \to 0$ and $Lv \leq -c$, then system (1.1) is (uniformly) unstable. From this we might conjecture that changing the requirement $v(x, y) \to +\infty$ by $v(x, y) \to -\infty$ (as $x \to 0$) we would obtain stability, or perhaps asymptotic stability. With an additional assumption, this is indeed the case, as shown in (5.3) below. Definitions (5.2) and (5.4) are variations of the concepts of "S-function" and "G-function" introduced by Friedman and Pinsky (see Friedman [6], or Pinsky [19]).

For some $\lambda > 0$ and $h > 0$, let $V(x, y)$ be the function defined as

$$V(x, y) = e^{\lambda v(x,y)} \quad \text{if} \quad 0 < |x| < h, y \in S$$
$$= 0 \qquad \text{if} \quad x = 0, y \in S,$$

where $v(x, y)$ is a function defined for $0 < |x| < h, y \in S$, which satisfies:

(5.1) (a) There exist continuous functions $w_1(x)$, $w_2(x)$, $x \neq 0$, such that, for any $y \in S$,

$$w_1(x) \leq v(x, y) \leq w_2(x), \quad \text{and} \quad w_2(x) \to -\infty \quad \text{as} \quad x \to 0;$$

(b) v is the domain of L and $Lv(x, y) \leq -1$ $(0 < |x| < h, y \in S)$.

We want to show that, with an additional assumption, (5.1c), and λ sufficiently small, $V(x, y)$ is a Lyapunov function satisfying (3.2a). Clearly, $V(x, y)$ is a positive-definite function. Also, $LV(0, y) = 0$ (by definition of V and the fact that $x(0) = 0 \Rightarrow x(t) = 0$ for all $t > 0$, w.p.1), and for $0 < |x| < h, y \in S$,

$$LV(x, y) = F(x, y)^* V_x(x, y) + QV(x, y)$$
$$= \lambda V(x, y) F(x, y)^* V_x(x, y) + QV(x, y),$$

where Q is the generator of $y(t)$. On the other hand, using Taylor's theorem to expand the identity

$$V(x, y(t)) - V(x, y) = V(x, y)\{e^{\lambda[v(x,y(t)) - v(x,y)]} - 1\},$$

we see that, for $0 < |x| < h, y \in S$,

$$QV(x, y) = \lambda V(x, y)[Qv(x, y) + \lambda R_\lambda(x, y)], \tag{*}$$

where $R_\lambda(x, y)$ is a continuous nonnegative function of x, y, and which, in general, may also depend on λ. (For instance, if $y(t)$ is a diffusion process with drift vector $b(y)$ and diffusion matrix $a(y)$, then $R_\lambda(x, y) = \frac{1}{2} v_y{}^* a(y) v_y$.) Therefore,

$$\begin{aligned} LV(x, y) &= \lambda V(x, y)[F(x, y)^* v_x + Qv(x, y) + \lambda R_\lambda(x, y)] \\ &= \lambda V(x, y)[Lv(x, y) + \lambda R_\lambda(x, y)] \qquad (**) \\ &\leq \lambda V(x, y)(-1 + \lambda R_\lambda(x, y)), \quad \text{by (5.1b)}. \end{aligned}$$

Now, in addition to (5.1a, b), let us assume:

(5.1c) The "second order" term $R_\lambda(x, y)$ in (*) is bounded, say

$$0 \leq R_\lambda(x, y) \leq m, \; m > 0, \quad \text{for} \quad 0 < |x| < h, \; y \in S.$$

Then, from (**), $LV(x, y) < 0$ if $0 < \lambda < 1/m$. Before continuing let us give the following:

(5.2) *Definition.* A function $v(x, y)$ defined for $0 < |x| < h$, $y \in S$, is said to satisfy *condition A* if (5.1a–c) hold.

Hence, summing up, we have proved above that if $v(x, y)$ satisfies condition A, then for $\lambda > 0$ sufficiently small, the function $V(x, y)$ is a Lyapunov function which satisfies hypothesis (a) of Theorem (3.2). Therefore, we have proved:

THEOREM 5.3. *If there exists a function $v(x, y)$ which satisfies condition A, then the system* (1.1) *is asymptotically stable* w.p.1.

Comparing (5.3) with Theorem (3.2), we would expect that condition A plus some new condition imply asymptotic stability in the large. This new condition is what we call condition B.

(5.4) *Definition.* A nonnegative function $w(x, y)$ defined for all $x \neq 0$, $y \in S$, is said to satisfy *condition B* if

(a) w is in the domain of L and $Lw(x, y) < 0$ $(x \neq 0)$,

(b) $w(x, y) \to \infty$ as $|x| \to \infty$.

Compare this definition with (3.2a, b). An analysis of the proof of Lemma (3.3) shows immediately that the properties (of the function $v(x, y)$) that we used to prove that lemma are precisely conditions (a) and (b) in (5.4). Therefore, we have:

LEMMA 5.5. *If there exists a function $w(x, y)$ satisfying condition B, then for any neighborhood $U \subset R^n$ of the origin and any $x \notin U$, $y \in S$,*

$$P_{x,y}(x(t) \in U \quad \text{for some} \quad t > 0) = 1.$$

Proof. The same as for Lemma (3.3).

Also, as in (3.4), we have:

COROLLARY 5.6. *If there exists a function satisfying condition B, then* $x(t)$, $t \geq 0$, *is recurrent.*

Finally, combining Theorem (5.3) and Lemma (5.5), we obtain (cf. the paragraph preceding Lemma (3.3)):

THEOREM 5.7 *If there exists a function satisfying condition A and a function satisfying condition B, then the system* (1.1) *is asymptotically stable in the large w.p.*1.

6. Some examples

Let us consider the system

$$\ddot{z} + f(y)z + g(y)\dot{z} = 0, \qquad t \geq 0 \tag{6.1}$$

with given initial conditions $z(0)$ and $\dot{z}(0)$. Here, $y(t)$, $t \geq 0$, is a finite state Markov process, with state space $S = \{1, \cdots, N\}$ say, and f and g are bounded functions of $y(t)$. We can write the given equation as

$$\dot{x}(t) = A(y(t))x(t), \tag{6.2}$$

where $x(t)$ denotes the vector with components $x_1 = z$, $x_2 = \dot{z}$, and $A(y)$ is the matrix

$$A(y) = \begin{bmatrix} 0 & 1 \\ -g(y) & -f(y) \end{bmatrix}.$$

As usual, $(x(t), y(t))$ is a Markov process on $R^2 \times S$. Let $B(y)$ be a 2×2 symmetric and positive-definite matrix of functions of $y(t)$, so that

$$v(x, y) = \tfrac{1}{2}\, x^* B(y) x, \qquad x \in R^2, y \in S,$$

is a positive-definite function, and

$$Lv(x, y) = x^*[A(y)^* B(y) + \tfrac{1}{2} \textstyle\sum_{j=1}^N q_{yj} B(j)]x,$$

where $Q = (q_{ij})$ is the infinitesimal matrix of the process $y(t)$; see (1.6a). Therefore, $Lv(x, y)$ is such that $Lv \leq 0$ iff the matrix

$$Y(y) = A(y)^* B(y) + \tfrac{1}{2} \textstyle\sum_{j=1}^N q_{yj} B(j)$$

is nonpositive-definite. Correspondingly, Lv is negative-definite iff $Y(y)$ is negative-definite. This will depend, of course, on the values of $A(y)$, $B(y)$ and $Q = (q_{ij})$. For instance, let us assume that $g(y) > 0$, and that $y(t) = \pm \beta$, $\beta > 0$, is the random telegraph process (or two-state Markov process) with infinitesimal matrix

$$Q = \alpha \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}, \qquad \alpha > 0.$$

Hence, if $B(y)$ is the matrix

$$B(y) = \begin{bmatrix} 1 & 0 \\ 0 & b(y) \end{bmatrix}, \qquad \text{with} \qquad b(y) = 1/g(y),$$

we see that

$$Y(y) = \begin{bmatrix} 0 & -1 \\ 1 & -f(y)b(y) \end{bmatrix} + \tfrac{1}{2}\alpha \operatorname{sgn}(y) \left\{ -\begin{bmatrix} 1 & 0 \\ 0 & b(\beta) \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & b(-\beta) \end{bmatrix} \right\},$$
$$b(y) = 1/g(y),$$

which is nonpositive-definite (and therefore, (6.2) is stable) if α, $f(y)$ and $g(y)$ satisfy:

$$\text{(6.3)} \qquad \alpha g(y)\operatorname{sgn}(y)[g(\beta) - g(-\beta)] \le 2f(y)g(\beta)g(-\beta), \qquad y = \pm\beta.$$

More specifically, if, for example, (6.2) represents a damped harmonic oscillator with coefficients

$$\text{(6.4)} \qquad f(y) = 2\gamma, \qquad g(y) = w^2 + y, \qquad (g(y) > 0),$$

criterion (6.3) tells us (see Theorem (3.1)) that (6.2) is stable if the parameters α and β of the process $y(t)$ and the constants γ and w in (6.4) satisfy:

$$\alpha\beta(w^2 + y)\operatorname{sgn}(y) \le 2\gamma(w^2 + \beta)(w^2 - \beta), \qquad y = \pm\beta.$$

Thus, for instance, given α and β, and if $w^2 > \beta$, then (6.2) is stable for any $\gamma \ge 0$; but if $w^2 < \beta$, the same conclusion holds provided that $\gamma \le -\alpha/2$. Note that, in contrast with the deterministic case, the system (6.2) can be stable even if $f(y)$ takes negative values. As an example, take $f(y) = \alpha y/g(-y)$, $y = \pm\beta$, where $g(y)$ is as in (6.4); then (6.3) holds.

Now we want to show an example of a function $v(x, y)$ which satisfies condition A. Consider again the linear system (6.2), where $y(t)$ is the N-state Markov chain given above with infinitesimal matrix $Q = (q_{ij})$. For $x \ne 0$, $y \in S = \{1, \cdots, N\}$, define $v(x, y) = \log|xy|$, $x \in R^2$. Since $v_x = x/|x|^2$, we see from (1.6a) that

$$\text{(6.5)} \qquad \begin{aligned} Lv(x, y) &= |x|^{-2}x^*A(y)^*x + \textstyle\sum_{j=1}^N q_{yj}\log|xj| \\ &= |x|^{-2}x^*A(y)^*x + \textstyle\sum_{j=1}^N q_{yj}\log|j|. \end{aligned}$$

The nonnegative "second order" term $R_\lambda(x, y)$ in (5.1c) is given by

$$\lambda^2 R_\lambda(x, y) = \textstyle\sum_{j\ne y} q_{yj}[(j/y)^\lambda - \log(j/y)^\lambda - 1]$$

which is bounded. On the other hand, the function $v(x, y) = \log|xy| \to -\infty$ as $x \to 0$, and therefore, v satisfies condition A if $Lv \le -1$. In the special case of the damped harmonic oscillator with coefficients (6.4) and $y(t)$ the random telegraph process, the term $\sum_j q_{yj}\log|j|$ in (6.5) becomes

$$\alpha \operatorname{sgn}(y)(-\log|\beta| + \log|-\beta|) = 0$$

and therefore, the condition $Lv \le -1$ reduces to

$$Lv(x, y) = |x|^{-2}x^*A(y)^*x$$
$$= |x|^{-2}((1 - w^2 - y)x_1x_2 - 2\gamma x_2^2) \leq -1, \qquad x \neq 0, \qquad y = \pm\beta,$$

or equivalently,

$$(*) \quad x_1^2 + (1 - w^2 - y)x_1x_2 + (1 - 2\gamma)x_2^2 \leq 0, \qquad (x_1, x_2) \neq 0, \qquad y = \pm\beta.$$

If this holds, then by Theorem (5.3) the system is asymptotically stable. Thus for specific values of the parameters w, γ, β, we can use (*) to determine the region of asymptotic stability w.p.1.

To illustrate the instability Theorem (4.1), let $\dot{x} = A(y)x$ be an arbitrary linear system, and let $v(x, y) = |y||x|^{-2}$, $x \neq 0$. If $y(t)$ is a finite state jump process with infinitesimal matrix $Q = (q_{ij})$, then

$$Lv(x, y) = -2|y||x|^{-4}x^*A(y)^*x + |x|^{-2}\textstyle\sum_j q_{yj}|j|.$$

Therefore, since $v(x, y) \to \infty$ as $x \to 0$, the given system is unstable if $Lv \leq -c$, for some $c > 0$. In particular, if $y(t) = \pm\beta$ is the random telegraph process,

$$Lv(x, y) = -2\beta|x|^{-4}x^*A(y)^*x, \qquad y = \pm\beta,$$

and the system is unstable if $x^*A(y)x \geq c|x|^{24}$ for some constant $c > 0$.

Now, let $y(t)$, $t \geq 0$, be an m-dimensional diffusion process satisfying the stochastic differential equation

$$dy = b(y(t))\, dt + \sigma(y(t))\, dw, \tag{6.6}$$

where $w(t)$ is a Wiener process and $b(y)$, $\sigma(y)$ satisfy the so-called Ito conditions, which ensure existence and uniqueness of solutions of (6.6). (See, for instance, Gikhman and Skorokhod [8].) In this case, the process $x(t)$ defined by (1.1) can be seen as the first n components of the $(n + m)$-dimensional diffusion process $(x(t), y(t))$ satisfying jointly (1.1) and (6.6), namely,

$$\begin{aligned} dx &= F(x(t), y(t))\, dt \\ dy &= b(y(t))\, dt + \sigma(y(t))\, dw. \end{aligned} \tag{6.7}$$

Therefore, to study the stability of (1.1) we can analyze the joint system (6.7) and then restrict ourselves to the x-components. In particular, to construct Lyapunov functions we can try to use (at least in nice cases) the methods described by Kushner [14, 15]. Thus, for instance, if $x(t)$ (one-dimensional) satisfies

$$\dot{x}(t) = (a + y(t))x(t),$$

where $y(t)$ is the Ornstein-Uhlenbeck process

$$dy = -\alpha y(t)\, dt + \sigma\, dw,$$

with α and σ positive constants, it is easily verified that $v(x, y) = x^2e^{y/\alpha}$ is a Lyapunov function if $y \leq -(2a + \sigma^2/2\alpha^2)$. For other examples (including cases in which $y(t)$ is not a diffusion) we refer to Kushner [14] and its references.

We can also try to construct a function $v(x, y)$ satisfying condition A. To be specific, let us assume that $y(t)$ is 1-dimensional, so that

$$Lv(x, y) = F(x, y)^* v_x + b(y) v_y + \tfrac{1}{2} a(y) v_{yy},$$

where $a(y) = \sigma^2(y)$; see (1.6b). The simplest case we can think of is when (1.1) is linear and $y(t)$ is the standard Wiener process; in such a case $(x(t), y(t))$ satisfies (6.7) with

$$F(x, y) = A(y)x, \qquad b(y) \equiv 0, \qquad \sigma(y) \equiv 1,$$

for some matrix $A(y)$. If $v(x, y) = \log(|x\|y|)$, $|x\|y| > 0$, then

$$Lv(x, y) = |x|^{-2} x^* A(y)^* x - \tfrac{1}{2} y^{-2},$$

and the function

$$V(x, y) = e^{\lambda v(x,y)} = (|x\|y|)^{\lambda}, \qquad \lambda > 0,$$

satisfies

$$\begin{aligned} LV(x, y) &= \lambda(|x\|y|)^{\lambda}(|x|^{-2} x^* A(y)^* x - \tfrac{1}{2} y^{-2} + \tfrac{1}{2}\lambda y^{-2}) \\ &= \lambda V(x, y)(Lv(x, y) + \tfrac{1}{2}\lambda y^{-2}). \end{aligned}$$

Let us assume that $Lv(x, y) \leq -1$. (Clearly, a sufficient condition for $Lv \leq -1$ is that $A(y)$ is negative-definite and $y^2 \leq \frac{1}{2}$, $y \neq 0$.) Then

$$LV(x, y) \leq \lambda V(x, y)(-1 + \tfrac{1}{2}\lambda y^{-2})$$

and the right side is negative if $y^2 > \lambda/2$. Therefore, $V(x, y)$ is a Lyapunov function satisfying (3.2a) (and the system $\dot{x} = A(y)x$ is asymptotically stable) if V is defined for all x and y in a neighborhood of the origin and $y^2 > \lambda/2$, $\lambda > 0$.

Remark. If $v(x, y)$, $|x| < h$, $y \in S$, is a positive-definite function in the domain of L, then a sufficient condition for $Lv(x, y)$ to be negative-definite is that

$$Lv(x, y) = -\lambda v(x, y) \qquad (\text{or} \quad Lv \leq -\lambda v,\ Lv(0, y) = 0) \tag{6.8}$$

for some $\lambda > 0$. Thus, the existence of Lyapunov functions for the system (1.1) is in some way related to the existence of positive-definite solutions of equation (6.8). Unfortunately, however, to solve this equation does not appear to be an easy matter, not even in the simple case in which the system (1.1) is linear. For a certain class of diffusion Markov processes, the existence of positive solutions of (6.8) is equivalent to the finiteness of the expected value $E_{x,y}(e^{\lambda\tau})$, where τ is the exit time of $x(t)$ from an open ball $B_h = \{|x| < h\}$; see Khasminskii [10]

Part II. Weak stochastic stability

A stochastic process $x(t)$, $t \geq 0$, in R^n is said to be weakly stochastically stable if it admits an invariant probability distribution. This means that there is a (not necessarily unique) probability distribution $\bar{P}$ such that if $\bar{P}$ is the distribution of $x(0)$, then $\bar{P}$ is also the distribution of $x(t)$ for all $t > 0$; that is,

$$\bar{P}(x(t) \in A) = \bar{P}(x(0) \in A)$$

for any $t > 0$ and any Borel set A in R^n.

In this part we give Lyapunov-like criteria for weak stochastic stability of n-dimensional processes $x(t)$ defined by equations of the form (1.1).

As mentioned in the Introduction, the concept of weak stochastic stability was introduced by Wonham [21]. He used Lyapunov techniques to study that form of stability for stochastic differential equations of Ito type; see also Kushner [15, 16] and Zakai [23].

7. Preliminaries

We consider again the n-dimensional process $x(t)$ defined by (1.1). The process $y(t)$ and the function $F(x, y)$ are required to satisfy assumptions (1.2).

In addition, we shall now require that the state space S of $y(t)$ is σ-compact, and also:

(7.1) If $y = y(0) = (y_1, \cdots, y_m)^*$ is the initial state, then as $t \downarrow 0$,

$$E_y(y_i(t) - y_i) = b_i(y)t + o(t)$$

$$E_y[(y_i(t) - y_i)(y_j(t) - y_j)] = a_{ij}(y)t + o(t),$$

where $b_i(y)$ and $a_{ij}(y)$, $(i, j = 1, \cdots, m)$ are C^1 functions, with bounded derivatives, and such that

$$|b_i(y)| \leq C(1 + |y|)$$

$$|a_{ij}(y)| \leq C(1 + |y|^2),$$

for all $y \in S$; $i, j = 1, \cdots, m$, and some constant $c > 0$.

Condition (7.1) holds for all the "usual" jump or diffusion processes $y(t)$; see Dynkin [5]. Furthermore, if condition (7.1) holds, it can be shown (with the same proof as in Zakai [23, §2]) that:

(7.2) For any compact set $K \subset R^n \times S$, and $t > 0$,

$$P(t, x, y, K) = P_{x,y}((x(t), y(t)) \in K) \to 0 \quad \text{as} \quad |x| + |y| \to \infty.$$

Following Benes' terminology [1], we define a *moment* as a Borel measurable function $v(x, y)$ on $R^n \times S$ such that v is nonnegative and $v(x, y) \to \infty$ as $|x| + |y| \to \infty$. If μ is the probability distribution of $(x(0), y(0))$, let us denote by $U_t\mu$ the distribution of $(x(t), y(t))$; that is, for any Borel subset K of $R^n \times S$,

$$U_t\mu(K) = P((x(t), y(t)) \in K)$$

$$= \int_{R^n \times S} \mu(dx, dy) P(t, x, y, K).$$

With this terminology, μ is invariant iff $U_t\mu = \mu$ for all $t \geq 0$.

Benes [1] shows that for a Feller process $(x(t), y(t))$ (see (1.2c)) which satisfies (7.2) the following statements are equivalent:

(7.3) (a) $(x(t), y(t))$ has a nontrivial invariant probability distribution.

(b) There exists a Borel probability measure μ and a moment v such that

$$\sup_{t\geq 0} \int_{R^n \times S} v(x, y)\, U_t\mu(dx, dy) < \infty.$$

(c) There exists a Borel probability measure μ and a compact set K in $R^n \times S$ such that

$$\limsup_{t\to\infty} \frac{1}{t} \int_0^t U_s\mu(K)\, ds > 0.$$

Let us introduce the following:

(7.4) *Definition.* If $v(x, y)$, $x \in R^n$, $y \in S$, is a Lyapunov function which satisfies conditions (3.2a)–(3.2b), we then say that v satisfies *condition C.*

8. Lyapunov theorems

Comparison of the definition of moment with definitions (5.4) and (7.4) suggests the following result:

THEOREM 8.1. *Suppose that there exists a function $v(x, y)$ such that $v(x, y) \to \infty$ as $|x| + |y| \to \infty$, and such that either*

(a) v satisfies condition B, or

(b) v satisfies condition C.

Then the process $(x(t), y(t))$ has an invariant probability distribution, $\mu(dx, dy)$ say, and the corresponding marginal distribution

$$\bar{P}(\cdot) = \int_S \mu(\cdot, dy) = \mu(\cdot, S) \tag{8.2}$$

is an invariant probability distribution for $x(t)$, that is, $x(t)$ is weakly stochastically stable.

Proof. Let $v(x, y)$ be a function such that $v(x, y) \to \infty$ as $|x| + |y| \to \infty$. If, in addition, v satisfies (a) or (b), then it is a moment.

Now, let $v(x, y)$ be a function satisfying the hypotheses of the theorem, and let $\mu(dx, dy)$ be any Borel probability measure on $R^n \times S$ such that μ has compact support (that is, for some compact K, $\mu(K) = \mu(R^n \times S) = 1$) and $v(x, y)$ is integrable with respect to μ. Let us assume also that μ is the distribution of the initial state $(x(0), y(0))$. Next, since $Lv \leq 0$, we have from (1.7) that

$$E_{z,u}v(x(t), y(t)) \leq v(z, u)$$

and therefore,

$$\sup_{t\geq 0} \int v(x, y)\, U_t\mu(dx, dy) = \sup_{t>0} \int E_{z,u}v(x(t), y(t))\mu(dz, du)$$

$$(*) \qquad \leq \int v(z, u)\mu(dz, du) = \int_K v(z, u)\mu(dz, du) < \infty.$$

Here, the integrals $\int$ are taken over all $R^n \times S$. Therefore, by Benes' Theorem (7.3), $(x(t), y(t))$ has an invariant probability distribution. If we denote such distribution again by μ, we see that the measure $\bar{P}$ in (8.2) is an invariant probability distribution for $x(t)$, since

$$\begin{aligned}
\bar{P}(x(t) \in A) &= \mu(x(t) \in A, y(t) \in S) \\
&= U_t\mu(A \times S) \\
&= \mu(A \times S) \equiv \mu(x(0) \in A, y(0) \in S) \\
&= \bar{P}(x(0) \in A)
\end{aligned}$$

for any $t \geq 0$. This completes the proof of the theorem.

Remark. To obtain the first equality in (*) we used the fact that $U_t\mu$ is the "adjoint" of the operators $T_t v$ defined in Remark (1.3); that is, for any bounded measurable function v,

$$\int v d(U_t\mu) = \int T_t v \, d\mu,$$

where $T_t v(x, y) = E_{x,y}v(x(t), y(t))$; see Dynkin [5]. To prove that the relation holds for a function $v(x, y)$ as in Theorem (8.1), apply monotone convergence to a sequence of bounded measurable functions $v_n \uparrow v$.

In both parts (a) and (b) of Theorem (8.1) it is required that $v(x, y) \to \infty$ as $|(x, y)| \to \infty$. In the next theorem, which is modelled after one by Zakai [23] for Ito differential equations, that condition is dropped.

THEOREM 8.3. *The same conclusion of Theorem* (8.1) *holds if there exists a nonnegative function* $v(x, y)$, $x \in R^n$, $y \in S$, *such that*

(a) *v is in the domain of L,*

(b) *there exists $m > 0$ and a compact set $K \subset R^n \times S$ such that*

$$LV(x, y) \leq -m \quad \text{for} \quad (x, y) \notin K.$$

Proof. Let $M > 0$ be an upper bound of $Lv(x, y)$ on K. Then

$$\begin{aligned}
E_{x,y}Lv(x(s), y(s)) &\leq MP((x(s), y(s)) \in K) - mP((x(s), y(s)) \notin K) \\
&= -m + (m + M)P((x(s), y(s)) \in K),
\end{aligned}$$

where $P = P_{x,y}$. Then from Dynkin's formula (1.7),

$$\begin{aligned}
E_{x,y}v(x(t), y(t)) &= v(x, y) + E_{x,y} \int_0^t Lv(x(s), y(s)) \, ds \\
&\leq v(x, y) - mt + (m + M) \int_0^t P((x(s), y(s)) \in K) \, ds
\end{aligned}$$

and therefore, since $E_{x,y}v \geq 0$,

$$(*)\qquad (m + M)\,\frac{1}{t}\int_0^t P(x(s), y(s)) \in K)\, ds \geq m - \frac{1}{t}\, v(x, y),$$

where $P = P_{x,y}$. Now let $\mu(dx, dy)$ be a probability measure as in the proof of Theorem (8.1); in particular, recall that μ is the distribution of the initial state $(x(0), y(0)) = (x, y)$, has compact support, and v is integrable with respect to μ. Then taking expectations with respect to μ (denoted simply by E) in (*), we obtain

$$(m + M)\,\frac{1}{t}\int_0^t U_s\mu(K)\, ds \geq m - \frac{1}{t}\, Ev(x, y).$$

Letting $t \to \infty$, we have

$$\limsup_{t\to\infty} \frac{1}{t}\int_0^t U_s\mu(K)\, ds > 0.$$

Consequently, from (7.3c) the process $(x(t), y(t))$ has an invariant probability distribution.

Remarks. (a) In Theorems (8.1) and (8.3) we make no statement about the uniqueness of the invariant probability measure. For a class of diffusion processes described by Ito stochastic differential equations, Wonham [21] (see also Khasminskii [11]) has given sufficient conditions for (existence and) uniqueness of invariant distributions. In particular, one of the requirements in Wonham's paper is that the diffusion matrix, $a(y)$ say, of the process satisfies that

$$y^*a(y)y \geq c\, y^*y \quad \text{for all} \quad y,\ (c > 0),$$

so that the generator Q of $y(t)$ is strictly elliptic. In our case, however, if $x(t)$ satisfies (1.1), where $y(t)$ obeys the Ito equation

$$dy = b(y(t))\, dt + \sigma(y(t))\, dw$$

the joint process $(x(t), y(t))$, which satisfies the Ito equation

$$dx = F(x(t), y(t))\, dt$$

$$dy = b(y(t))\, dt + \sigma(y(t))\, dw,$$

has a diffusion matrix which is only nonnegative definite. Using a "controllability" assumption, Kushner [15, 16] gives sufficient conditions for the existence of a unique invariant distribution for a special class of Ito processes without the strict ellipticity condition and allowing certain types of discontinuities in the coeffficients.

(b) Let $P(t, x, y; dz, du)$ denote the transition probability measure of the process $(x(t), y(t))$, and let us assume that p has a density $p(t, x, y; z, u)$, that is

$$P(t, x, y; A, J) = P_{x,y}((x(t), y(t)) \in A \times J) = \int_{A\times J} zp(t, x, y; z, u)\, du\, dz,$$

where A and J are Borel subsets of R^n and S, respectively. Then $p = p(t, x, y; z, u)$ satisfies (at least formally) the forward equation

$$p_t = -\sum_{i=1}^n \frac{\partial}{\partial z_i}(pF_i(z, u)) + Q^*p, \tag{8.4}$$

where Q^* is the formal adjoint of the generator Q of $y(t)$ (see Gikhman and Skorokhod [7], or Dynkin [5]). For instance, if $y(t)$ is a Markov chain with a finite state space $S = \{1, \cdots, N\}$ and infinitestimal matrix $Q = (q_{ij})$, equation (8.4) becomes

$$p_t = -\sum_{i=1}^n \frac{\partial}{\partial z_i}(pF_i(z, u)) + \sum_{j=1}^n p(t, x, y; x, j)q_{ju}, \qquad u = 1, \cdots, N,$$

where $p = p(t, x, y; z, u)$. Similarly, if $y(t)$ is an m-dimensional diffusion process with drift vector $b(y) = (b_1(y), \cdots, b_m(y))^*$ and diffusion matrix $a(y) = (a_{ij}(y))$, the corresponding forward equation is (8.4), where

$$Q^*p = -\sum_{j=1}^m \frac{\partial}{\partial u_j}[pb_j(u)] + \tfrac{1}{2}\sum_{i,j=1}^m \frac{\partial^2}{\partial u_i \partial u_j}[pa_{ij}(u)].$$

If $(x(t), y(t))$ has an invariant probability distribution $\bar{P}$ with density $\bar{p} = \bar{p}(z, u)$, then $\bar{p}$ satisfies the "stationary" forward equation

$$0 = -\sum_{i=1}^n \frac{\partial}{\partial z_i}(\bar{p}(z, u)F_i(z, u)) + Q^*\bar{p}(z, u).$$

Acknowledgment. The author is grateful to professors W. H. Fleming, H. J. Kushner and S. Geman for many valuable suggestions.

Centro de Investigación del IPN, México 14, D.F.

References

[1] V. E. Benes, *Finite regular invariant measures for Feller Processes*, J. Appl. Probability, **5** (1968), 203–209.

[2] J. E. Bertram and P. E. Sarachik, *Stability of circuits with randomly time-varying parameters*, Trans. IRE-PGCT, **6** (1959), 260–270.

[3] R. S. Bucy, *Stability and positive supermartingales*, J. Differential Equations, **1** (1965), 151–155.

[4] W. G. Cumberland and C. A. Rohde, *A multivariate model for growth of populations*, Theor. Popul. Biol. **11** (1977), 127–139.

[5] E. B. Dynkin, Markov Processes, 2 vol., Springer-Verlag, Berlin-Heidelberg-New York, 1965.

[6] A. Friedman, Stochastic Differential Equations and Applications, 2 vol. Academic Press, New York, 1976.

[7] I. I. Gikhman and A. V. Skorokhod, Introduction to the Theory of Random Processes, Saunders, Philadelphia, 1969.

[8] ———, Stochastic Differential Equations, Springer-Verlag, Berlin-Heidelberg-New York, 1972.

[9] I. I. Kats and N. N. Krasovskii, *On the stability of systems with random parameters.* J. Appl. Math. Mech. (PMM) **24** (1960), 1225–1246.

[10] R. Z. Khasminskii, *On positive solutions of the equation $Uu + Vu = 0$.* Theory Probab. Appl. **4** (1959) 309–318.

[11] ———, *Ergodic properties of recurrent diffusion processes and stabilization of the solution*

to the Cauchy problem for parabolic equations, Theory Probab. Appl., **5** (1960), 179-196.

[12] ———, *On the stability of the trajectory of Markov processes*, J. Appl. Math. Mech (PMM), **26** (1962), 1554–1565.

[13] F. Kozin, *A survey of stability of stochastic systems*, Automatica, **5** (1969), 95–112.

[14] H. J. Kushner, Stochastic Stability and Control, Academic Press, New York, 1967.

[15] ———, *The Cauchy problem for a class of degenerate parabolic equations and asymptotic properties of the related diffusion processes*, J. Differential Equations, **6** (1969), 209-231.

[16] ———, *Stability and existence of diffusions with discontinuous or rapidly growing drift terms*, J. Differential Equations, **11** (1972), 156–168.

[17] J. P. LaSalle and S. Lefschetz, Stability by Lyapunov's Direct Method, with Applications. Academic Press, New York, 1961.

[18] J. A. Morrison and J. McKenna, *Analysis of some stochastic ordinary differential equations*, SIAM-Proc. Amer. Math. Soc. Vol. VI, 1973.

[19] M. A. Pinsky, *Stochastic stability and the Dirichlet problem*, Comm. Pure Appl. Math. **27** (1974), 311–350.

[20] N. G. Van Kampen, *Stochastic differential equations*, Phys. Reports, 24C(1976), 171–228.

[21] W. M. Wonham, *A Lyapunov criteria for weak stochastic stability*, J. Differential Equations, **2** (1966), 195–207.

[22] ———, *Random differential equations in control theory*, in Probabilistic Methods in Applied Mathematics, Vol. **2** (A. T. Bharucha-Reid, ed), Academic Press, New York, 1970.

[23] M. Zakai, *A Lyapunov criterion for the existence of stationary probability distributions for systems perturbed by noise*, SIAM J. Control, **7** (1969), 390–397.