Random Dynamical Systems

Zhisheng Shuai
Department of Mathematical and Statistical Science
University of Alberta

February 4, 2005

This lecture note is only used for my series of talks in Group meeting of Prof. James Muldowney and Prof. Michael Li. Most materials are chosen from books or articles in reference lists.

1 Introduction

What is randomness? Why randomness?
What is the difference between randomness and determinacy?

First, let us to look some examples in applications of mathematics:

Example 1.1. Consider a simple population growth model

$$\frac{dN}{dt} = (b(t) - d(t))N(t), \quad N(0) = A \tag{1.1}$$

where $N(t)$ is the size of the population at time t, $b(t)$ is the birth rate of population at time t, and $d(t)$ is the death rate of population at time t. It might happen that $b(t)$ and $d(t)$ are not completely known, but subject to some random environmental effects, so that we have

$$b(t) - d(t) + \text{“noise”},$$

where we do not know the exact behavior of the noise term, only its probability distribution. The functions $b(t), d(t)$ are assumed to be nonrandom. How do we solve (1.1) in this case?

Example 1.2. Consider the SIS epidemic model

$$\begin{cases} S' = \mu - \mu S - \lambda IS + \gamma I \\ I' = \lambda IS - \mu I - \gamma I \\ S(t) + I(t) = 1, \end{cases} \tag{1.2}$$

where S, I present respectively the fractions of the total population in susceptible class and infective class, μ is the death rate, λ is the average number of adequate contacts per day, and γ is the recovery rate. The basic reproduction number is

$$R_0 = \frac{\lambda}{\gamma + \mu}.$$

If $R_0 > 1$, $(\frac{1}{R_0}, 1 - \frac{1}{R_0})$ is globally stable; if $R_0 \leq 1$, then $(1, 0)$ is globally stable. Note that λ should be affected by many factors. It's difficult to be determined precisely. Instead, we often let λ is a random variable, which is more suitable. However, what's the asymptotic behavior of solutions under random λ? For example, $\lambda = \lambda_1$ or λ_2, where $R_0(\lambda_1) > 1$, $R_0(\lambda_2) \leq 1$, and $P(\lambda = \lambda_1) = P(\lambda = \lambda_2) = \frac{1}{2}$.

Example 1.3. A logistic population growth model is

$$x'(t) = rx(t)(1 - \frac{x(t)}{K}), \tag{1.3}$$

where x present the total population of X, r is a positive constant, called the *intrinsic growth rate*, and K, called environmental capacity, is a simple discrete random variable:

$$P(K = K_1) = P(K = K_2) = \frac{1}{2}, 0 < K_1 \leq K_2 < 0.$$

If $K_1 = K_2$, then $x = K$ is the globally asymptotic stable, which attracts all solutions with positive initial value. How about $K_1 < K_2$?

Example 1.4. Consider a nonautonomous logistic model,

$$x'(t) = r(t)x(t)(1 - \frac{x(t)}{K(t)}). \tag{1.4}$$

If $r(t)$, $K(t)$ are T-periodic functions, $K(t) > 0$ and $\int_0^T r(s)ds > 0$, then there exists a unique positive T-periodic solution of (1.4), $x^P(t)$, which is globally asymptotically stable for any solution $x(t)$ with positive initial value $x_0 > 0$. Now consider the random perturbation of (1.4) as follows

$$dx(t) = x(t)[(a(t) - b(t)x(t))dt + \alpha(t)dB(t)], \tag{1.5}$$

where $a(t) = r(t)$, $b(t) = \frac{r(t)}{K(t)}$ are T-periodic functions, $\dot{B}(t)$ is white noise, i.e., $B(t)$ is the 1-dimensional standard Brownian motion, and $\alpha(t)$ represents the intensity of the noise. If $a(t), b(t)$, and $\alpha(t)$ are continuous T-periodic functions, $a(t) > 0$, $b(t) > 0$ and $\int_0^T [a(s) - \alpha^2(s)]ds > 0$, then $E[1/x(t)]$ has a unique positive T-periodic solution $E[1/x^P(t)]$, which is also globally asymptotically stable. See D. Jiang's article*JMAA 2005*.

1.1 Some Mathematical Preliminaries

Having stated the problems we would like to solve, we now proceed to find reasonable mathematical notions corresponding to the quantities mentioned and mathematical models for the problem. In short, here is a first list of the notions that need a mathematical interpretation:

(1) Probability spaces
(2) Random variables
(3) Stochastic processes
(4) Brownian motion

In this subsection, we will discuss them briefly.

1.1.1 Probability Spaces

Definition 1.1. If Ω is a nonempty set, $\mathcal{F}$ is a $\sigma-$algebra ($\sigma-$field) of Ω, and P is a probability measure on the measurable space $(\Omega, \mathcal{F})$, then the triple $(\Omega, \mathcal{F}, P)$ is a *probability space*.

Here Ω is called the *sample space*. Its elements are called *sample points*. $\mathcal{F}$ is a σ-algebra of Ω means $\mathcal{F}$ is a collection of subsets of Ω that contains the empty set ϕ and is closed under the formation of complements and of countable unions of its members, i.e.

(a) $\phi \in \mathcal{F}$
(b) If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$
(c) $A_n \in \mathcal{F}, \forall n \in \mathbb{N}$, imply $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

Elements of $\mathcal{F}$ are called *events*, and $\mathcal{F}$ is also called *events field.* P is a probability measure on $(\Omega, \mathcal{F})$, i.e., a countably additive nonnegative measure with total mass 1. In details,

(a) countably additive: $P(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n)$ if $A_n \in \mathcal{F}, A_n \bigcap A_m = \phi, \forall n \neq m$.
(b) nonnegative: $P(A) \geq 0, \forall A \in \mathcal{F}$.
(c) $P(\Omega) = 1$.
Probability measure P is usually called *probability* for simpleness.

Proposition 1.1. *(1) $P(\phi) = 0, P(A^c) = 1 - P(A), \forall A \in \mathcal{F}$.*
(2) Finite additive: If $A_1, A_2, \ldots, A_n \in \mathcal{F}, A_i \bigcap A_j = \phi, \forall i \neq j$, then $P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$.
(3) $P(A \bigcup B) = P(A) + P(B) - P(AB), P(A - B) = P(A) - P(AB), \forall A, B \in \mathcal{F}$.
(4) If $A \subset B, \forall A, B \in \mathcal{F}$, then $P(A) \leq P(B)$.
(5) Jordan identity: $\forall A_1, A_2, \ldots, A_n \in \mathcal{F}$,
$P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i) - \sum_{1 \leq i < j \leq n} P(A_i A_j) + \sum_{1 \leq i < j < k \leq n} P(A_i A_j A_k) - \cdots + (-1)^n P(A_1 A_2 \cdots A_n)$.
Moreover, $P(\bigcup_{i=1}^{n} A_i) \leq \sum_{i=1}^{n} P(A_i)$.

Another important concept is independence. For $A, B \in \mathcal{F}$, A and B are *(stochastic) independent* if $P(AB) = P(A)P(B)$.

Now let us see a simple example as follows.

Example 1.5 (Throw A Coin). One tosses or throws a coin. There are two possible cases, namely, (i) head; (ii) tail. Denote 1 as head, -1 as tail. Then $\Omega = \{1, -1\}$, $\mathcal{F} = \{\phi, \{1\}, \{-1\}, \{1, -1\}\}$. Probability measure P:

$$P(\phi) = 0, P(\{1\}) = \frac{1}{2}, P(\{-1\}) = \frac{1}{2}, P(\{1, -1\}) = 1.$$

1.1.2 Random Variables

Definition 1.2. Suppose $(\Omega, \mathcal{F}, P)$ is a probability space, $X : \Omega \to \mathbb{R}$ is a *random variable* if $\{X(\omega) \leq a\} \in \mathcal{F}$ for $\forall a \in \mathbb{R}$, i.e., X is measurable function.

In general, a random variable is an $\mathcal{F}$-measurable function $X : \Omega \to E$, where $(\Omega, \mathcal{F}, P)$ is a probability space, (E, ξ) is a measure space. Here $\mathcal{F}$-measurable function means that $\forall B \in \xi$ implies $X^{-1}(B) \in \mathcal{F}$.

Suppose X is a random variable on $(\Omega, \mathcal{F}, P)$. Define $F(x) := P(X \leq x) = P(X \in (-\infty, x]), \forall x \in \mathbb{R}$. $F(x)$ is called the *distribution (function)* of X. If there exists $f(x)$ such that $F(x) = \int_{-\infty}^{x} f(u)du$, then $f(x)$ is called *probability density function* of X.

Suppose X is random variable on $(\Omega, \mathcal{F}, P)$, and $F(x)$ is the corresponding distribution function. If $\int_{-\infty}^{\infty} |x| dF(x) < \infty$, then the number

$$E[X] := \int_{-\infty}^{\infty} x dF(x)$$

is called the *expectation (or mean)* of X. Note that the above integral is Riemann-Stieltjes integral. We also can get the expectation of X by Lebesgue integral:

$$E[X] = \int_{\Omega} X(\omega) dP(\omega).$$

X and Y both are random variables on $(\Omega, \mathcal{F}, P)$. X and Y are *independent* if

$$F(x, y) = F_X(x) F_Y(y),$$

where $F(x, y) = P(X \leq x, Y \leq y)$, $F_X(x) = P(X \leq x) = \lim_{y \to +\infty} F(x, y) = F(x, +\infty)$.

Example 1.6 (Deterministic Variable). A deterministic variable $x \in \mathbb{R}$ is a special random variable. In fact, for $\forall x \in \mathbb{R}$, $(\Omega, \mathcal{F}, P)$, define:

$$\mathbf{x} : \Omega \to \mathbb{R}$$
$$\mathbf{x}(\omega) = x, \forall \omega \in \Omega.$$

Example 1.7 (Throw A Die). When a die is thrown, there are six possible cases. We take $\Omega = \{1, 2, 3, 4, 5, 6\}$, $\mathcal{F}$ is the collection of all subsets of Ω. For any $1 \leq i \leq 6$, $P(\{i\}) = \frac{1}{6}$, i.e., equality probability. Let $X : \Omega \to \mathbb{R}$ as follows

$$X(\omega) = \begin{cases} -1 & \text{if } \omega \text{ is odd} \\ 1 & \text{if } \omega \text{ is even.} \end{cases}$$

Then X is a random variable.

Example 1.8 (Normal Distribution). $X \sim N(\mu, \sigma^2)$, i.e., the distribution function $F(x)$ of X satisfies

$$F(x) = \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{x} \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right) dt,$$

where μ is the expectation or mean value of X, σ^2 is the variance of X. If $\mu = 0$ and $\sigma = 1$, i.e., $N(0, 1)$, then it is called standard *normal* (or *Gaussian*) distribution.

1.1.3 Stochastic Processes

Definition 1.3. A stochastic process is a parameterized collection of random variables $\{X_t, t \in T\}$ defined on a probability space $(\Omega, \mathcal{F}, P)$ and assuming values in $\mathbb{R}$.

Here the parameter space T is usually a subset of $\mathbb{R}$, called index set. For example, $T_1 = \mathbb{N}_0, T_2 = \mathbb{Z}, T_3 = \mathbb{R}^+$. Denote $X(t, \omega) = X_t(\omega)$, where $t \in T$, $\omega \in \Omega$. If we fix t, then we get $X_t : \Omega \to \mathbb{R}$, which is a random variable. If we fix ω, then we obtain $X_\omega : T \to \mathbb{R}$ is a *(sample) path.*

Example 1.9 (Simple Random Walk). Think of a particle moving randomly among the integers according to the following rules: At time $n = 0$, the particle is at the origin. At time $n = 1$, it moves either one unit left to -1 or one unit right to $+1$, with respective probabilities p and $q = 1 - p$ (e.g. $p = \frac{1}{2}$). At time n, the particle moves from its present position S_{n-1} by a unit distance left or right. Suppose X_n denote the displacement of the particle at the nth step from its position S_{n-1} at time $n - 1$, i.e., $P(X_n = -1) = p, P(X_n = 1) = 1 - p$ for each $n \geq 1$. $S_n = X_1 + \cdots + X_n$, $S_0 = 0$. Therefore, $S_n, n = 0, 1, \ldots$ is a stochastic process.

Example 1.10 (Real Value Function). Real value function $f : T \to \mathbb{R}$ is a special stochastic process. In fact, for any fixed $t \in T$, $f(t)$ is constant, by Example 1.6, further a random variable. Thus f is a stochastic process.

More special stochastic processes were studied in other books or articles, for example, process with independent increments(Poisson process, Brownian motion), Markov process, wide sense stationary process, finite second moments process, strictly stationary process, martingales, renewal process, point process, etc.

1.1.4 Brownian Motion

Under certain conditions, small particles (pollen grains) suspended in liquid can be observed to undergo a continual, irregular motion. This phenomenon is named after its nineteenth century discoverer, Scottish botanist Robert Brown (1828). The explanation for the motion lies in the random collisions with the molecules of the liquid.

A probabilistic theory for the Brownian motion, based on several simple assumptions was put forward by Albert Einstein in 1906. In 1923, Norbert Wiener gave a model describing the motion in the sense of mathematical point of view. For this reason, Brownian motion is also called the *Wiener process.* In the following, we will give the mathematical definition of 1-dimensional Brownian motion.

Definition 1.4. Suppose $\{B(t)\}_{t\geq 0}$ is a stochastic process on $(\Omega, \mathcal{F}, P)$. If

(1) independent increments: For any $t_0 < t_1 < \cdots < t_n$, $B(t_0), B(t_1)-B(t_2), \ldots, B(t_n)-B(t_{n-1})$ are independent.

(2) $\forall s,t > 0$, $B(s+t) - B(s) \sim N(0, c^2 t)$, i.e., $B(s+t) - B(s)$ has a normal distribution with mean value 0 and variance $c^2 t$.

(3) $B(t)$ is continuous in t, i.e., for fixed $\omega \in \Omega$, $B_\omega : T \to \mathbb{R}$ is continuous.

Then $\{B(t)\}_{t\geq 0}$ is called 1-dimensional *Brownian motion* or *Wiener process*.

If $c = 1$, then $B(t)$ is called *standard Brownian motion.* If $\{B_t^{(j)}\}, j = 1, 2, \ldots, k$ are k independent Brownian motions, then the vector-values process $\{\mathbf{B}_t\} = \{(B_t^{(1)}, B_t^{(2)}, \ldots, B_t^{(k)})\}$ is called a k-dimensional Brownian motion. Also k-dimensional Brownian motion can be defined by changing (2) in above Definition 1.4 to $\mathbf{B}(s+t) - \mathbf{B}(s)$ has a n-dimensional normal distribution.

Remark 1.1. 1. Brownian paths B_ω are Hölder continuous with exponent γ for any $\gamma < \frac{1}{2}$. However, if $\gamma > \frac{1}{2}$, then with probability 1, Brownian paths are not Hölder continuous with exponent γ.

2. With probability 1, Brownian paths are not Lipschitz continuous (and hence not differentiable) at any point.

3.With probability 1, Brownian paths are not bounded variation functions.

4. $B_\omega(t) - B_\omega(s) \approx \sqrt{t-s}$.

See R. Durrett's book(1996) for details.

Proposition 1.2. *Assume $\{B(t), t \geq 0\}$ is Brownian motion , then*

(1) $\{B(t+\tau) - B(\tau), t \geq 0\}, \forall \tau \geq 0$

(2) $\{\frac{1}{\sqrt{\lambda}} B(\lambda t), t \geq 0\}$, $\forall \lambda > 0$

(3) $\{tB(\frac{1}{t}), t \geq 0\}$ where $\{tB(\frac{1}{t})\}|_{t=0} := 0$

(4) $\{B(t_0 + s) - B(t_0), 0 \leq s \leq t_0\}$, $\forall t_0 > 0$

all of them are Brownian motion.

Suppose $\{B_t, t \geq 0\}$ be a Brownian motion. Fix $\triangle t > 0$, denote

$$\frac{\triangle B(t)}{\triangle t} := \frac{B(t + \triangle t) - B(t)}{\triangle t},$$

which is also a stochastic process. Let $\triangle t \to 0^+$, denote in the form $\{\frac{dB(t)}{dt}, t \geq 0\}$, which is called derivative of Brownian motion $\{B_t, t \geq 0\}$ in the sense of form or *white noise* with continuous parameter.

1.2 Ito Integrals

We now turn to the question of finding a reasonable mathematical interpretation of the "noise" term in the equation of Example 1.1:

$$\frac{dN}{dt} = (b(t) - d(t) + \text{“noise”})N(t), \tag{1.6}$$

namely,

$$\frac{dx}{dt} = f(t,x) + \sigma(t,x) \cdot \text{“noise”}, \tag{1.7}$$

where $x = N$, $f(t,x) = (b(t) - d(t))x(t)$, and $\sigma(t,x) = x(t)$. Represent "noise" as a generalized stochastic process called the *white noise* process, $\frac{dB(t)}{dt}$, where $B(t)$ is the 1-dimensional standard Brownian motion. To avoid the explanation of differential of a stochastic process, we apply the usual integration notation in the following:

$$x(t) = x_0 + \int_0^t f(s,x)ds + \int_0^t \sigma(s,x)dB_s, \;\; x(0) = x_0. \tag{1.8}$$

"The first integral of above equation is an ordinary Riemann integral by now". However, how about the second integral

$$\text{"}\int_0^t \sigma(s,x)dB_s\text{"?} \tag{1.9}$$

Note that $B_\omega : T \to \mathbb{R}$ is of unbounded variation on $[0,t]$, so Riemann-Stieltjes integral fails here. We should give appropriate mathematical interpretation about (1.9). This is our main aim of this subsection.

To define appropriate stochastic integral, we need to consider the limit operation of random variables. We present metric space of random variables.

Definition 1.5. $H := \{X : E|X|^2 < \infty\}$ is called *H-space*, namely, H is the collection of all random variables with finite second moment.

H-space is a linear space. Further, for $\forall X, Y \in H$, define $(X,Y) := E(XY)$. It can be shown that $(\cdot,\cdot)$ is an inner product on H. Define $||X|| = (X,X)^{\frac{1}{2}}$, for $\forall X \in H$, namely, $||X|| = (E|X|^2)^{\frac{1}{2}}$. Clearly, $||\cdot||$ is a norm on H. Let $d(X,Y) = ||X - Y||$, for any $\forall X, Y \in H$, and d is a metric on H.

Definition 1.6. Suppose $\{X, X_n, n \geq 1\} \subset H$. If $\lim_{n\to\infty} d(X_n, X) = 0$, namely, $\lim_{n\to\infty} ||X_n - X|| = \lim_{n\to\infty}(E|X_n - X|^2)^{\frac{1}{2}} = 0$, then X is called the *mean square limit* of $\{X_n, n \geq 1\}$, denoted by $\lim_{n\to\infty} X_n \overset{\text{m.s.}}{=} X$

H is complete, namely, if $\{X_n, n \geq 1\}$ is a Cauchy sequence of random variables in H, then there exists $X \in H$ such that $X_n \xrightarrow{\text{m.s.}} X$.

Example 1.11. Suppose $\{B(t), t \geq 0\}$ is a Brownian motion with $B(0) = 0$. For any $0 = t_0 < t_1 < t_2 < \cdots < t_n = t$, let $\triangle t_k = t_k - t_{k-1}, 1 \leq k \leq n$. Set $\lambda = \max\limits_{1\leq k\leq n} \triangle t_k$, $\triangle B_k = B(t_k) - B(t_{k-1})$. Then

$$\lim_{\lambda\to 0} \sum_{k=1}^n (B(t_k) - B(t_{k-1}))^2 \overset{\text{m.s.}}{=} t. \tag{1.10}$$

In fact, $\triangle B_k \sim N(0, t_k - t_{k-1})$. By properties of normal distribution, $E[\triangle B_k]^4 = 3(t_k - t_{k-1})^2$, $E[\triangle B_k]^2 = t_k - t_{k-1}$. Moreover, $(\triangle B_k)^2, 1 \leq k \leq n$ are independent each other, hence for $k \neq l$, $E[(\triangle B_k)^2(\triangle B_l)^2] = E[\triangle B_k]^2 E[\triangle B_l]^2 = (t_k - t_{k-1})(t_l - t_{l-1})$. Therefore,

$$\begin{aligned}
E[\sum_{k=1}^n (\triangle B_k)^2 - t]^2 &= E(\sum_{k=1}^n (\triangle B_k)^2)^2 - 2E(\sum_{k=1}^n (\triangle B_k)^2)t + t^2 \\
&= 3\sum_{k=1}^n (t_k - t_{k-1})^2 + 2\sum_{i<j}(t_i - t_{i-1})(t_j - t_{j-1}) - 2[\sum_{k=1}^n (t_k - t_{k-1})]t + t^2 \\
&= 2\sum_{k=1}^n (t_k - t_{k-1})^2 + [\sum_{k=1}^n (t_k - t_{k-1})^2 + 2\sum_{i<j}(t_i - t_{i-1})(t_j - t_{j-1}) \\
&\quad -(\sum_{k=1}^n (t_k - t_{k-1}))(\sum_{l=1}^n (t_l - t_{l-1}))] \\
&= 2\sum_{k=1}^n (t_k - t_{k-1})^2 \\
&\leq 2\lambda \sum_{k=1}^n (t_k - t_{k-1}) = 2\lambda t \to 0, \text{ as } \lambda \to 0.
\end{aligned}$$

Note that if we assume $g \in C^1$, then similar to (1.10), we have

$$\lim_{\lambda\to 0} \sum_{k=1}^n (g(t_k) - g(t_{k-1}))^2 = 0. \tag{1.11}$$

In fact,

$$\sum_{k=1}^{n}(g(t_k)-g(t_{k-1}))^2=\sum_{k=1}^{n}(g'(\xi_k)\frac{t}{k})^2=\frac{t}{k}\sum_{k=1}^{n}\left((g'(\xi_k))^2\frac{t}{k}\right)=\frac{t}{k}\int_0^t(g'(s))^2ds\to 0,\text{ as } k\to\infty.$$

From here, we can see the biggest difference between stochastic integral and deterministic integral. Also, it leads to the more term in so-called Ito formula later than chain rule of differentiate of deterministic calculus.

If assume g is only bounded variation function, not necessary continuous differentiable, is (1.11) right? That's still a problem for us.

Example 1.12. Under the same assumptions as last example, we further obtain

$$\lim_{\lambda\to 0}\sum_{k=1}^{n}B(t_{k-1})(B(t_k)-B(t_{k-1}))\overset{\text{m.s.}}{=}\frac{1}{2}B^2(t)-\frac{1}{2}t,\tag{1.12}$$

$$\lim_{\lambda\to 0}\sum_{k=1}^{n}B(t_k)(B(t_k)-B(t_{k-1}))\overset{\text{m.s.}}{=}\frac{1}{2}B^2(t)+\frac{1}{2}t.\tag{1.13}$$

In details, denote $L_n=\sum_{k=1}^{n}B(t_{k-1})(B(t_k)-B(t_{k-1}))$, $R_n=\sum_{k=1}^{n}B(t_{k-1})(B(t_k)-B(t_{k-1}))$. Then

$$L_n+R_n=\sum_{k=1}^{n}(B(t_k)+B(t_{k-1}))(B(t_k)-B(t_{k-1}))=\sum_{k=1}^{n}(B^2(t_k)-B^2(t_{k-1}))=B^2(t),$$

and by Example 1.11 we get

$$R_n-L_n=\sum_{k=1}^{n}(B(t_k)-B(t_{k-1}))^2\xrightarrow{\text{m.s.}}t\quad(\lambda\to 0).$$

Thus $\lim_{\lambda\to 0}R_n\overset{\text{m.s.}}{=}\frac{1}{2}B^2(t)+\frac{1}{2}t$, and $\lim_{\lambda\to 0}L_n\overset{\text{m.s.}}{=}\frac{1}{2}B^2(t)-\frac{1}{2}t$.

In general, we can show that

$$\lim_{\lambda\to 0}\sum_{k=1}^{n}B(t_{k-1}+\theta(t_k-t_{k-1}))(B(t_k)-B(t_{k-1}))\overset{\text{m.s.}}{=}\frac{1}{2}B^2(t)+\frac{1}{2}(2\theta-1)t,$$

where $0\le\theta\le 1$.

Remark 1.2. Note that (1.12) and (1.13) choose the left end point and right end point of every subinterval in partition of Riemann-Stieltjes integral, respectively. Unfortunately, they have different limit values as λ tends to zero. That is to say, the classical Riemann-Stieltjes doesn't work well. Instead of arbitrarily choosing point, K. Ito and R.L. Stratonovich considered the different definitions of stochastic integral; Ito used the left end point, while Stratonovich used the middle point.

Let B_t be Brownian motion. Then we define $\mathcal{F}_t$ to be the σ-algabra generated by the random variables $B_s(\cdot), s\le t$. In other words, $\mathcal{F}_t$ is the smallest σ-algebra containing all sets of the form

$$\{\omega: B_{t_1}(\omega)\in F_1,\ldots,B_{t_k}(\omega)\in F_k\},$$

where $t_j\le t$ and $F_j\subset\mathbb{R}$ are Borel sets, $j\le k=1,2,\ldots$.

First, introduce $\mathcal{L}_T^2$ space. Assume $\{g(t,\omega),t\ge 0\}$ be a stochastic process, satisfying

(1) $g(t,\omega)$ is $\mathcal{B}\times\mathcal{F}$-measurable, where $\mathcal{B}$ denotes the Borel σ-algebra on $[0,\infty)$.

(2) $g(t,\omega)$ is $\mathcal{F}_t$-adapted, i.e., for $\forall x\in\mathbb{R}$, $\{\omega: g(t,\omega)\le x\}\in\mathcal{F}_t$.

(3) $E[\int_0^t g^2(t,\omega)dt]<\infty$.

The collection of all $\{g(t,\omega),t\ge 0\}$ satisfying (1-3) is called $\mathcal{L}_T^2$ *space.*

Definition 1.7. Suppose $\{B(t), t \geq 0\}$ be 1-dimensional standard Brownian motion, $\{g(t,\omega), t \geq 0\}$ satisfies $g \in \mathcal{L}_T^2$, $[0,t] \subset [0,T]$, and B_s is $\mathcal{F}_t$-measurable, for $\forall 0 \leq s \leq t$. For any $0 = t_0 < t_1 < t_2 < \cdots < t_n = t$, let $\triangle t_k = t_k - t_{k-1}, 1 \leq k \leq n$, and $\lambda = \max_{1 \leq k \leq n} \triangle t_k$. If

$$\lim_{\lambda \to 0} \sum_{k=1}^{n} g(t_{k-1})(B(t_k) - B(t_{k-1})) \overset{\text{m.s.}}{=}: Ig(t), \tag{1.14}$$

then

$$Ig(t) = \int_0^t g(s,\omega) dB(s,\omega)$$

is called the *Ito integral* of $\{g(t,\omega), t \geq 0\}$ in $[0,t]$.

Also denote $Ig(t)$ as $I_g(t)$ and $I_g(0,t)$.

If replace $g(t_{k-1})$ by $g(\frac{t_{k-1}+t_k}{2})$ in (1.14), then we get the *Stratonovich integral*, denoted by

$$Sg(t) = \int_0^t g(s,\omega) \circ dB(s,\omega).$$

"In general one can say that the Stratonovich integral has the advantage of leading to ordinary chain rule formula under a transformation, i.e., there are no second order terms in the Stratonovich analogue of the Ito transformation formula. This property makes the Stratonovich integral natural to use for example in connection with stochastic differential equations on manifolds. However, Stratonovich integrals are not martingales, as we will see that Ito integrals are. This gives the Ito integral an important computation advantage, even though it does not behave so nicely under transformations. For our purposes the Ito integral will be most convenient, so we will base our discussion on that from now on." From Øksendal's SDE.

Example 1.13. Show $\int_0^t B(s)dB(s)$, where $B(0) = 0$.

By Example 1.12,

$$\int_0^t B(s)dB(s) = \frac{1}{2}B^2(t) - \frac{t}{2}.$$

In general,

$$\int_a^b B(s)dB(s) = \frac{1}{2}(B^2(b) - B^2(a)) - \frac{1}{2}(b-a).$$

Proposition 1.3. *For any $g_1, g_2, g_3 \in \mathcal{L}_T^2$, we have*

(1) $EI_g(t) = 0$;

(2) $EI_g^2(t) = \int_0^t Eg^2(s)ds$, $(I_g(s), I_g(t)) = \int_0^t Eg^2(u)du$;

(3) $I_{ag_1+bg_2}(t) = aI_{g_1}(t) + bI_{g_2}(t)$, for any $\forall a, b \in \mathbb{R}$;

(4) $I_g(0,t) = I_g(0,t_1) + I_g(t_1,t)$, $\forall 0 \leq t_1 \leq t$;

(5) $\{I_g(t), t \geq 0\}$ is a martingale with respect to $\mathcal{F}_t$. Furthermore, Doob's martingale inequality holds, i.e., for any $t \geq 0$, $\lambda > 0$, $p \geq 1$, we have

$$P(\max_{0 \leq u \leq t} |I_g(u)| > \lambda) \leq \frac{E(I_g(t))^p}{\lambda^p}.$$

Remark 1.3. A *filtration* is a family $\{\mathcal{F}_t\}_{t \geq 0}$ of σ-algebra $\mathcal{F}_t \subset \mathcal{F}$ such that

$$0 \leq s < t \Longrightarrow \mathcal{F}_s \subset \mathcal{F}_t.$$

An stochastic process $\{M_t\}_{t \geq 0}$ on $(\Omega, \mathcal{F}, P)$ is called a *martingale* with respect to a filtration $\{\mathcal{F}_t\}_{t \geq 0}$ if

(1) M_t is $\mathcal{F}_t$-measurable for all $t \geq 0$,

(2) $E|M_t| < \infty$ for all $t \geq 0$,

(3) $E[M_{t+s}|\mathcal{F}_t] = M_t$, for all $s, t \geq 0$.

1.3 Ito Formula

Recall (1.8) as

$$X(t) = X(0) + \int_0^t f(s,X)ds + \int_0^t \sigma(s,X)dB(s). \tag{1.15}$$

Note that if $X(t)$ is a stochastic process on $(\Omega, \mathcal{F}, P)$, then the first integral is a special Ito integral, in fact, a mean square integral. Namely, we can choose any point in every subinterval of partition instead of left end point of Ito integral.

Definition 1.8. Let $\{B(t), t \geq 0\}$ be 1-dimensional standard Brownian motion on $(\Omega, \mathcal{F}, P)$. A (1-dimensional) *Ito process* is a stochastic process X_t on $(\Omega, \mathcal{F}, P)$ of the form

$$X(t) = X(0) + \int_0^t f(s,X)ds + \int_0^t \sigma(s,X)dB(s), \tag{1.16}$$

or in the Ito differential form

$$dX(t) = f(t,X(t))dt + \sigma(t,X(t))dB(t), \tag{1.17}$$

where $f, \sigma \in \mathcal{L}_T^2$ are continuous functions. (1.16) is called (Ito) *stochastic integral equation*, while (1.17) is called (Ito) *stochastic differential equation.*

Theorem 1.1 (Ito Formula). *Suppose $\{X(t), t \geq 0\}$ be an Ito process given by (1.16) or (1.17), $h(t,x) \in C^2([0,\infty) \times \mathbb{R})$. Then*

$$Y(t) := h(t,X(t))$$

is again an Ito process, and $Y(t)$ satisfies the following stochastic integral equation:

$$Y(t) = Y(0) + \int_0^t [\frac{\partial h}{\partial t} + f\frac{\partial h}{\partial x} + \frac{\sigma^2}{2}\frac{\partial^2 h}{\partial x^2}](s,X(s))ds + \int_0^t \sigma\frac{\partial h}{\partial x}(s,X(s))dB(s), \tag{1.18}$$

or the equivalent stochastic differential equation

$$dY(t) = (\frac{\partial h}{\partial t} + f\frac{\partial h}{\partial x} + \frac{\sigma^2}{2}\frac{\partial^2 h}{\partial x^2})(t,X(t))dt + \sigma\frac{\partial h}{\partial x}(t,X(t))dB(t). \tag{1.19}$$

In other words,

$$dY(t) = \frac{\partial h}{\partial t}(t,X(t))dt + \frac{\partial h}{\partial x}(t,X(t))dX(t) + \frac{1}{2}\frac{\partial^2 h}{\partial x^2}(t,X(t))(dX(t))^2, \tag{1.20}$$

where $(dX(t))^2 = (dX(t)) \cdot (dX(t))$ is computed according to the rules

$$dt \cdot dt = dt \cdot dB(t) = dB(t) \cdot dt = 0, \;\; dB(t) \cdot dB(t) = dt. \tag{1.21}$$

Comparing Ito formula with chain rule of differentiate of deterministic calculus, we can find that the only difference is the term "$\frac{\partial^2 h}{\partial x^2}(t,X(t))dt$".

Example 1.14. Let us return to the integral in Example 1.13

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

Choose $X(t) = B(t)$, and $h(t,x) = \frac{1}{2}x^2$. Then $dX(t) = 0 \times dt + 1 \times dB(t) = dB(t)$, and $Y(t) = h(t,X(t)) = \frac{1}{2}B^2(t)$. Hence, $f = 0, \sigma = 1$, $\frac{\partial h}{\partial t} = 0$, $\frac{\partial h}{\partial x} = x$, $\frac{\partial^2 h}{\partial x^2} = 1$. Therefore,

$$\begin{aligned} dY(t) &= \sigma\tfrac{\partial h}{\partial x}dB(t) + [\tfrac{\partial h}{\partial t} + f\tfrac{\partial h}{\partial x} + \tfrac{\sigma^2}{2}\tfrac{\partial^2 h}{\partial x^2}]dt \\ &= B(t)dB(t) + (0 + 0 + \tfrac{1}{2})dt, \end{aligned}$$

namely,

$$d(\frac{1}{2}B^2(t)) = B(t)dB(t) + \frac{1}{2}dt.$$

Write in the form of integral

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

that is to say,

$$\int_0^t B(s)dB(s) = \frac{1}{2}B^2(t) - \frac{t}{2}.$$

Example 1.15. Consider $\int_0^t sdB(s)$. Here, $h(t,x) = tx$, $X(t) = B(t)$ and $Y(t) = h(t, X(t)) = tB(t)$. Then by Ito's formula,

$$dY(t) = B(t)dt + tdB(t) + 0,$$

namely,

$$d(tB(t)) = B(t)dt + tdB(t).$$

Therefore,

$$\int_0^t sdB(s) = tB(t) - \int_0^t B(s)ds,$$

which is reasonable from an integration-by-parts point of view.

Theorem 1.2 (Integration By Parts). *Suppose $\sigma(s)$ only depends on s and that σ is continuous and of bounded variation in $[0,t]$. Then*

$$\int_0^t \sigma(s)dB(s) = \sigma(t)B(t) - \int_0^t B(s)d\sigma(s). \tag{1.22}$$

We now turn to the situation in higher dimensions: Let $\mathbf{B}(t) = (B_1(t), \ldots, B_m(t))^T$ denote m-dimensional standard Brownian motion. Suppose $\mathbf{F}(t) = (f_1(t), \ldots, f_n(t))^T$,

$$\mathbf{\Sigma}(t) = (\sigma_{ij}(t)) = \begin{pmatrix} \sigma_{11}(t) & \cdots & \sigma_{1m}(t) \\ \vdots & & \vdots \\ \sigma_{n1}(t) & \cdots & \sigma_{nm}(t) \end{pmatrix},$$

for any $1 \le i \le n, 1 \le j \le m$, and $f_i(t), \sigma_{ij}(t) \in \mathcal{L}_T^2$ are $\mathcal{F}_t$-measurable, where $\mathcal{F}_t = \sigma(B_i(s_i), s_i \le t, 1 \le i \le m)$. If $\{\mathbf{X}(t) = (X_1(t), \ldots, X_n(t))^T, t \ge 0\}$ satisfies

$$\begin{cases} dX_1(t) &= f_1(t)dt + \sigma_{11}(t)dB_1(t) + \sigma_{12}(t)dB_2(t) + \cdots + \sigma_{1m}(t)dB_m(t), \\ dX_2(t) &= f_2(t)dt + \sigma_{21}(t)dB_1(t) + \sigma_{22}(t)dB_2(t) + \cdots + \sigma_{2m}(t)dB_m(t), \\ \cdots & \\ dX_n(t) &= f_n(t)dt + \sigma_{n1}(t)dB_1(t) + \sigma_{n2}(t)dB_2(t) + \cdots + \sigma_{nm}(t)dB_m(t), \end{cases} \tag{1.23}$$

or in matrix notation simply

$$d\mathbf{X}(t) = \mathbf{F}(t)dt + \mathbf{\Sigma}(t)d\mathbf{B}(t), \tag{1.24}$$

then $\{\mathbf{X}(t), t \ge 0\}$ is called n-dimensional Ito process.

Theorem 1.3 (Multi-dimensional Ito Formula). *Suppose $\{\mathbf{X}(t), t \ge 0\}$ is a n-dimensional Ito process, satisfying (1.24), and $\mathbf{H}(t, \mathbf{x}) = (h_1(t, \mathbf{x}), h_2(t, \mathbf{x}), \ldots, h_p(t, \mathbf{x}))^T$ be a C^2 map from*

$[0,\infty)\times\mathbb{R}^n$ *into* $\mathbb{R}^p$. *Then the process* $\{\mathbf{Y}(t)=\mathbf{H}(t,\mathbf{X}(t)), t\geq 0\}$ *is again a p-dimensional Ito process and satisfies*

$$\begin{aligned} dY_k(t) &= \left(\frac{\partial h_k}{\partial t}+\sum_{i=1}^n f_i\frac{\partial h_k}{\partial x_i}+\frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n\sum_{l=1}^m\frac{\partial^2 h_k}{\partial x_i\partial x_j}\sigma_{il}\sigma_{jl}\right)(t,\mathbf{X}(t))dt \\ &\quad +\sum_{l=1}^m\left(\sum_{i=1}^n\frac{\partial h_k}{\partial x_i}\sigma_{il}\right)(t,\mathbf{X}(t))dB_l(t), \quad 1\leq k\leq p, \end{aligned} \tag{1.25}$$

in other words,

$$dY_k(t)=\frac{\partial h_k}{\partial t}(t,\mathbf{X})dt+\sum_{i=1}^n\frac{\partial h_k}{\partial x_i}(t,\mathbf{X})dX_i+\frac{1}{2}\sum_{i,j=1}^n\frac{\partial^2 h_k}{\partial x_i\partial x_j}(t,\mathbf{X})dX_idX_j, \tag{1.26}$$

where $dB_idB_j=\delta_{ij}dt, dB_idt=dtdB_i=0$, δ *is Kronecker Delta.*

Example 1.16 (Special Two Dimensional Ito Formula). Suppose $\mathbf{X}(t)=(X_1(t),X_2(t))^T$ satisfies

$$\begin{pmatrix} dX_1(t) \\ dX_2(t)\end{pmatrix}=\begin{pmatrix} f_1(t) \\ f_2(t)\end{pmatrix}dt+\begin{pmatrix} f_1(t) & 0 \\ 0 & f_2(t)\end{pmatrix}\begin{pmatrix} dB_1(t) \\ dB_2(t)\end{pmatrix}. \tag{1.27}$$

Set $h(t,\mathbf{x})=x_1x_2$, then $Y(t)=X_1(t)X_2(t)$ satisfies

$$dY(t)=[f_1(t)X_2(t)+f_2(t)X_1(t)]dt+f_1(t)X_2(t)dB_1(t)+f_2(t)X_1(t)dB_2(t). \tag{1.28}$$

1.4 Stochastic Differential Equations

We now return to the possible solutions $X_t(\omega)$ of the stochastic differential equation

$$dX(t)=f(t,X(t))dt+\sigma(t,X(t))dB(t). \tag{1.29}$$

It is natural to ask: Can one obtain existence and uniqueness theorems for such equations like fundamental theorems of deterministic ordinary differential equations? How can one solves a given such equation?

First, we will consider the second question by looking at some simple examples. Then discuss the existence and uniqueness of solutions.

Example 1.17. Recall Example 1.1

$$dN(t)=(b(t)-d(t))N(t)dt+\alpha N(t)dB(t), \tag{1.30}$$

where α is a constant. Assume that $b(t)-d(t)=r$ is a positive constant. Then we get

$$dN(t)=rN(t)dt+\alpha N(t)dB(t), \tag{1.31}$$

thus, if $B(0)=0$,

$$\int_0^t\frac{dN(t)}{N(t)}=rt+\alpha B(t). \tag{1.32}$$

Let $h(t,x)=\ln x$, $Y(t)=\ln N(t)$, then $\frac{\partial h}{\partial t}=0$, $\frac{\partial h}{\partial x}=\frac{1}{x}$, $\frac{\partial^2 h}{\partial x^2}=-\frac{1}{x^2}$. Therefore

$$dY(t)=d\ln N(t)=\left[0+rN(t)\frac{1}{N(t)}+\frac{\alpha^2N^2(t)}{-2N^2(t)}\right]dt+\alpha N(t)\frac{1}{N(t)}dB(t),$$

namely,

$$d(\ln N(t)) = (r - \frac{\alpha^2}{2})dt + \alpha dB(t).$$

Integrating both sides from 0 to t, we obtain

$$\ln \frac{N(t)}{N(0)} = (r - \frac{\alpha^2}{2})t + \alpha B(t),$$

therefore,

$$N(t) = N(0) \exp\left((r - \frac{\alpha^2}{2})t + \alpha B(t) \right). \tag{1.33}$$

Remark 1.4. (1) If let $\alpha = 0$, then (1.33) changes to $N(t) = N(0)\exp(rt)$, which is just the solution of deterministic population growth model.

(2) (a) If $r > \frac{\alpha^2}{2}$, then $N(t) \to \infty$ as $t \to \infty$, a.s..

(b) If $r < \frac{\alpha^2}{2}$, then $N(t) \to 0$ as $t \to \infty$, a.s..

(c) If $r = \frac{\alpha^2}{2}$, then $N(t)$ will fluctuate between arbitrary large and arbitrary small values as $t \to \infty$, a.s..

(3) Random perturbation is “harmful” since the population doesn,t extinct if $r > 0$ when $\alpha = 0$; the population extincts if $r < \frac{\alpha^2}{2}$ when $\alpha \neq 0$. Namely, noise is beneficial for unstable and harmful for stable.

“Noise $\Longrightarrow$ unstable.”

(4) Suppose we change λ in Example 1.2 to λ plus white noise without other changing, i.e., $\lambda + \alpha W(t)$, where $W(t) = \dot{B}(t)$ formally.

Question: If $\alpha = 0$, $R_0 > 1$, then endemic equilibrium $(\frac{1}{R_0}, 1 - \frac{1}{R_0})$ is globally stable. Is noise helpful for control the infectious disease? How does we describe it in mathematics?

Example 1.18 (Black-Scholes Model). Suppose $S(t)$ present the price of some stock, and satisfy

$$\begin{cases} dS(t) = S(t)(\mu dt + \sigma dB(t)), \\ S(0) = S_0, \end{cases} \tag{1.34}$$

where μ, σ are constants. Similarly to Example 1.17, we can get

$$S(t) = S_0 \exp\left((\mu - \frac{\sigma^2}{2})t + \sigma B(t) \right). \tag{1.35}$$

Example 1.19 (Ornstein-Uhlenbeck Equation). Solve the Ornstein-Uhlenbeck equation (or Langevin equation)

$$dX(t) = -\mu X(t)dt + \sigma dB(t), \tag{1.36}$$

where μ, σ are real constants. From (1.36), we get

$$e^{\mu t}[dX(t) + \mu X(t)dt] = \sigma e^{\mu t} dB(t),$$

namely,

$$d(X(t)e^{\mu t}) = \sigma e^{\mu t} dB(t).$$

Integrating both sides from t_0 to t, $\forall\ 0 \leq t_0 < t \leq T$, we have

$$X(t)e^{\mu t} - X(t_0)e^{\mu t_0} = \int_{t_0}^{t} \sigma e^{\mu s} dB(s),$$

therefore,

$$X(t) = X(t_0)\exp(-\mu(t - t_0)) + \int_{t_0}^{t} \sigma \exp(-\mu(t - s)) dB(s). \tag{1.37}$$

Even for most deterministic ordinary differential equations, it is difficult and even impossible to obtain the explicit expression of solutions. Instead, one can prove the existence and uniqueness of solutions. Similar conclusion is given for stochastic differential equations.

Theorem 1.4 (Existence and Uniqueness Theorem for Stochastic Differential Equations). *Suppose* $\{X(t), t \geq 0\}$ *be an Ito process, satisfying*

$$dX(t) = f(t, X(t))dt + \sigma(t, X(t))dB(t), \quad 0 \leq t \leq T, X(0) = Z. \tag{1.38}$$

If

A1 [measurable] $f(t,x), \sigma(t,x) : [0,T] \times \mathbb{R} \to \mathbb{R}$ *are measurable functions and belong to* $\mathcal{L}^2_{[T\times\mathbb{R}]}$*;*

A2 [Lipschitz] *there exists some constant* K *such that for any* $\forall t \in [0,T]$*,* $\forall x, y \in \mathbb{R}$*, we have*

$$|f(t,x) - f(t,y)| + |\sigma(t,x) - \sigma(t,y)| < K|x-y|;$$

A3 [bounded] *there exists some constant* $C > 0$ *such that*

$$|f(t,x)| + |\sigma(t,x)| \leq C(1 + |x|), \quad \forall t \in [0,T], x \in \mathbb{R};$$

A4 [initial conditions] Z *is* $\mathcal{F}_0$*-measurable, and* $E|Z|^2 < \infty$*.*

Then the above stochastic differential equation (1.38) has a unique t-continuous solution $\{X(t), t \geq 0\}$*, and* $X(t)$ *is* $\mathcal{F}_t$*-measurable and* $EX^2(t) < \infty$ *for* $\forall t \in [0,T]$*.*

Theorem 1.5. *Under the same assumptions as Theorem 1.4, for above stochastic differential equation (1.38) and its unique solution* $\{X(t), t \geq 0\}$*, we have*

$$\lim_{h\to 0} \frac{1}{h} E[(X(t+h) - x)|X(t) = x] = f(t,x), \tag{1.39}$$

$$\lim_{h\to 0} \frac{1}{h} E[(X(t+h) - x)^2|X(t) = x] = \sigma^2(t,x). \tag{1.40}$$

Remark 1.5. If we change "$dB(t)$" to "$\circ\, dB(t)$" in (1.29), then we get the stochastic differential equations in the sense of Stratonovich stochastic integral. Both of them describe system with random forcing functions "$\sigma(t, X(t))dB(t)$" or "$\sigma(t, X(t)) \circ dB(t)$". In next subsection, other randomness will be discussed instead of "white noise", e.g., "real noise".

1.5 Random Differential Equations

Last subsection, stochastic differential equations, a simple class of random differential equations, are introduced. In T. Saaty's book(1981), R. Syski classified random differential equations into three basic types:

(a) Random initial conditions;
(b) Random forcing functions;
(c) Random coefficients.

In the first, also the simplest case, the initial conditions are random variables. This case only requires a minimum of probabilistic concepts, and the treatment follows very closely that for deterministic differential equations.

Example 1.20. Consider the IVP

$$\begin{cases} X' = aX \\ X(0) = v, \end{cases} \tag{1.41}$$

where $X(t,\omega), a \in \mathbb{R}^1$, and v is a random variable on $(\Omega, \mathcal{F}, P)$, namely, $v : \Omega \to \mathbb{R}$ is $\mathcal{F}$-measurable. The solution of IVP is

$$X(t) = ve^{at}, \quad t \geq 0,$$

which is a stochastic process on $(\Omega, \mathcal{F}, P)$. Simply, let $\Omega = \{-1, 1\}$, $P(1) = P(-1) = \frac{1}{2}$, and $v : \Omega \to \mathbb{R}$ satisfying $v(-1) = -1$ and $v(1) = 1$. Thus, for any given t, $X_t : \Omega \to \mathbb{R}$ satisfies $X_t(-1) = -e^{at}$ and $X_t(1) = e^{at}$. Therefore, $\{X(t), t \geq 0\}$ is a stochastic process on $(\Omega, \mathcal{F}, P)$.

The second class if often called stochastic differential equations, which have the form of

$$dX(t) = f(t, X(t))dt + \sigma(t, X(t))dZ(t), \tag{1.42}$$

where $Z(t)$ is some stochastic process. The initial condition $X(0)$ can be taken to be random or deterministic. In most cases, $Z(t)$ is chosen by $B(t)$, the standard Brownian motion, and $dB(t)$ is so-called "white noise". More details can be seen in last subsection.

In the third case, the parameters of the system we regarded as the random variables. For example, Example 1.2 and Example 1.3. The study of equations of this type has a rather recent origin, and is less developed than the previous two. We can change Example 1.20 to get the following example:

Example 1.21. Consider the IVP

$$\begin{cases} X' = vX \\ X(0, \omega) \equiv X_0, \end{cases} \tag{1.43}$$

where $X(t, \omega), X_0 \in \mathbb{R}^1$, and v is a random variable on $(\Omega, \mathcal{F}, P)$. The solution of IVP is

$$X(t) = X_0 e^{vt}, \quad t \geq 0$$

which is a stochastic process on $(\Omega, \mathcal{F}, P)$. Simply, let $\Omega = \{-1, 1\}$, $P(1) = P(-1) = \frac{1}{2}$, and $v : \Omega \to \mathbb{R}$ satisfying $v(-1) = -1$ and $v(1) = 1$. Thus, for any given t, $X_t : \Omega \to \mathbb{R}$ satisfies $X_t(-1) = X_0 e^{-t}$ and $X_t(1) = X_0 e^{t}$. Therefore, $\{X(t), t \geq 0\}$ is a stochastic process on $(\Omega, \mathcal{F}, P)$.

These three types are not mutually exclusive, and most of the discussion is concerned with equations of mixed type. For example, change the initial condition X_0 to a random variable in Example 1.21.

In general, stochastic differential equations, i.e., the second type, can be regarded as case of the third case. Consider

$$X'(t) = F(t, X, \omega). \tag{1.44}$$

Let $F = f(t, X) + \lambda\sigma(t, X)$, where $\lambda = W(t)$ is white noise. Then we get

$$\begin{aligned} dX(t) &= (f(t, X) + W(t)\sigma(t, x))dt \\ &= f(t, X(t))dt + \sigma(t, X(t))W(t)dt \\ &= f(t, X(t))dt + \sigma(t, X(t))dB(t). \end{aligned}$$

For stochastic differential equations, Theorem 1.4 describes the existence and uniqueness of solutions. For random differential equations, we also have similar theorem, but with stronger conditions.

Consider initial value problem of random differential equation

$$\begin{cases} X'(t) = F(X(t), t) \\ X(t_0) = X_0, \end{cases} \tag{1.45}$$

Theorem 1.6. *If $F : H \times T \to H$ satisfies the mean square Lipschitz condition*

$$||F(X, t) - F(Y, t)|| \leq K(t)||X - Y|| \tag{1.46}$$

where $\int_{t_0}^{a} K(t)dt < \infty$, then there exists a unique mean square solution on $[t_0, a)$ for any initial condition $X_0 \in H$.

The conclusion holds also for vector case. Proof see J.L. Strand's article (JDE 1970) or thesis (UC 1968).

Unfortunately, this theorem has very limited applicability even in the linear case.

Example 1.22. Consider the random differential equation

$$X'(t) = A(\omega)X(t), \tag{1.47}$$

where $A(\omega)$ is a random variable. This equation has a unique sample solution

$$X(t, \omega) = X_0(\omega)e^{tA(\omega)}$$

for initial condition $X(0) = X_0$. If A is bounded almost surely, then mean square Lipschitz is satisfied, thus mean square solution exists uniquely. However, since A is normal random variable, A fails essential bounded. J.L. Strand shows that the mean square solution exists for all X_0 with finite moments of all orders if and only if A is essential bounded.

Therefore, the better existence and uniqueness of random differential equations are still problems. Not like the research of stochastic differential equations, that of random differential equations is still less. However, such as Example (1.2) and Example (1.3), the problems associating random differential equations are natural consideration. As we know, since assume the coefficient be very simple random variable, like discrete random variable, the conclusion is also little.

From next section, the stability and asymptotic behavior of stochastic differential equations and random differential equations will be considered for entering the field of random dynamical systems.

2 Stochastic Stability

What are equilibria of random differential equations?
How to say stability in the sense of randomness?
Is Lyapunov's direct method still useful?

Before we start to discuss the stochastic stability, four kinds of stochastic limit need to be introduced in first subsection. In second subsection, the definitions of stability associating different kinds of limit are given. Further, some methods to show stochastic stability are shown.

2.1 Stochastic Limits

For completely understanding the derivative of stochastic process, we need to understand the stochastic limit and stochastic convergence of sequence of random variables. Since random variables are $\mathcal{F}$-measurable from sample space Ω to $\mathbb{R}$, the stochastic convergence looks like the convergence of sequence of functions. As we know, there are some kinds of convergence about sequence of functions, such as pointwise convergence, uniform convergence, L^p convergence, and pointwise convergence almost everywhere. Uniform convergence implies pointwise convergence; L^∞ convergence implies pointwise convergence almost everywhere; but neither $L^p(1 \leq p < \infty)$ convergence implies pointwise convergence almost everywhere nor pointwise convergence almost everywhere implies L^p convergence.

Definition 2.1 (Mean Square Convergence). A sequence of random variables $\{X_n\}$ *converges in mean square* to a random variable X as $n \to \infty$ if

$$\lim_{n\to\infty} X_n \overset{\text{m.s.}}{=} X, \tag{2.1}$$

namely,

$$\lim_{n\to\infty} E[(X_n - X)^2] = 0. \tag{2.2}$$

Example 2.1. Let $\{X_n\}$ be a sequence of random variable and distribution according to

$$X_n = \begin{cases} 1 & P(1) = \frac{1}{n} \\ 0 & P(0) = 1 - \frac{1}{n}. \end{cases} \tag{2.3}$$

Then $\{X_n\}$ converges in mean square to zero identically. In fact,

$$\lim_{n\to\infty} E[(X_n - X)^2] = \lim_{n\to\infty} E[X_n^2] = \lim_{n\to\infty} \frac{1}{n} = 0.$$

Example 2.2. If we keep all statements given in Example 2.1 unchanged except changing (2.3) to

$$X_n = \begin{cases} n & P(n) = \frac{1}{n^2} \\ 0 & P(0) = 1 - \frac{1}{n^2}. \end{cases} \tag{2.4}$$

Then $\{X_n\}$ doesn't converge in mean square. In fact, since X_m, X_n are independent, we have

$$\begin{aligned} \lim_{m,n\to\infty} E[(X_m - X_n)^2] &= \lim_{m,n\to\infty} E[X_m^2 - 2X_mX_n + X_n^2] \\ &= 2\lim_{m\to\infty} E[X_m^2] - 2\lim_{m,n\to\infty} E[X_m]E[X_n] \\ &= 2 - 2\lim_{m,n\to\infty} \frac{1}{m}\frac{1}{n} = 2. \end{aligned}$$

Definition 2.2 (Convergence in Probability). A sequence of random variables $\{X_n\}$ *converges in probability* to a random variable X as $n \to \infty$ if for any $\epsilon > 0$,

$$\lim_{n\to\infty} P\{|X_n - X| > \epsilon\} = 0. \tag{2.5}$$

Denote as $\lim_{n\to\infty} X_n \overset{\mathrm{P}}{=} X$ or $X_n \xrightarrow{\mathrm{P}} X$ as $n \to \infty$.

Example 2.3. Revised Example 2.2.

$$X_n = \begin{cases} n & P(n) = \frac{1}{n^2} \\ 0 & P(0) = 1 - \frac{1}{n^2}. \end{cases}$$

$\{X_n\}$ converges in probability to zero identically. In fact,

$$\begin{aligned} \lim_{n\to\infty} P\{|X_n - X| > \epsilon\} &= \lim_{n\to\infty} P\{|X_n| > \epsilon\} \\ &= \lim_{n\to\infty} P\{X_n = n\} \\ &= \lim_{n\to\infty} \frac{1}{n^2} = 0. \end{aligned}$$

This example points out an important feature which separates mean square convergence from convergence in probability. Mean square convergence is dependent on both the values that the random variables can take and the probability is associated with them. Convergence in probability, on the other hand, is only concerned with the probability of an event.

Proposition 2.1. *Convergence in mean square implies convergence in probability.*

Proof. Using Tchebychef inequality

$$P\{|X_n - X| > \epsilon\} \leq \frac{E[(X_n - X)^2]}{\epsilon^2}, \tag{2.6}$$

we can get the conclusion. □

Example 2.3 shows that the converse of Proposition 2.1 is not true.

Remark 2.1. Two important inequalities:
Tchebychef inequatlity
If random variable X has finite second moment, then

$$P\{|X| > \epsilon\} \leq \frac{E[X^2]}{\epsilon^2}. \tag{2.7}$$

Schwartz inequality

$$|E[XY]|^2 \leq E[X^2]E[Y^2]. \tag{2.8}$$

Question: If $X_n \xrightarrow{\mathrm{m.s.}} X$, as $n \to \infty$, then $\lim_{n\to\infty} E[X_n] = E[x]$, since by Schwartz inequality

$$|E[X_n] - E[X]| = |E[X_n - X]| \leq E[|X_n - X|] \leq \left(E[(X_n - X)^2]\right)^{1/2}.$$

Is it still true if $X_n \xrightarrow{\mathrm{P}} X$, as $n \to \infty$? The following example gives the answer, no!

Example 2.4. Consider a sequence $\{X_n\}$ of random variables defined by

$$X_n = \begin{cases} n & P(n) = \frac{1}{n} \\ -1 & P(-1) = 1 - \frac{1}{n}. \end{cases} \tag{2.9}$$

Clearly, $X_n \xrightarrow{\mathrm{P}} X \equiv -1$, $E[X] = -1$, and $E[X_n] = n\frac{1}{n} - (1 - \frac{1}{n}) = \frac{1}{n} \to 0$. Hence $\lim_{n\to\infty} E[X_n] \neq E[X]$.

Definition 2.3 (Almost Sure Convergence). A sequence $\{X_n\}$ of random variables is said to *converge almost surely* (or *with probability* 1) to a random variable X as $n \to \infty$ if

$$P\{\lim_{n\to\infty} X_n = X\} = 1. \tag{2.10}$$

Denote as $\lim_{n\to\infty} X_n \overset{\text{a.s.}}{=} X$ or $X_n \xrightarrow{\text{a.s.}} X$ as $n \to \infty$.

Proposition 2.2. *$\{X_n\}$ converges almost surely to X, i.e.,*

$$P\{\lim_{n\to\infty} X_n = X\} = 1, \tag{2.11}$$

if and only if

$$P\{\lim_{n\to\infty} X_n \neq X\} = 0, \tag{2.12}$$

if and only if

$$\lim_{n\to\infty} P\{\bigcup_{j\geq n}(|X_j - X| > \epsilon)\} = 0, \tag{2.13}$$

if and only if

$$\lim_{n\to\infty} P\{\sup_{j\geq n}|X_j - X| > \epsilon\} = 0, \tag{2.14}$$

if and only if

$$\lim_{n\to\infty} P\{\bigcap_{j\geq n}(|X_j - X| \leq \epsilon)\} = 1, \tag{2.15}$$

if and only if

$$\lim_{n\to\infty} P\{\inf_{j\geq n}|X_j - X| \leq \epsilon\} = 1. \tag{2.16}$$

Example 2.5. Revised Example 2.1.

$$X_n = \begin{cases} 1 & P(1) = \frac{1}{n} \\ 0 & P(0) = 1 - \frac{1}{n}. \end{cases}$$

$X_n \xrightarrow{\text{P}} 0, X_n \xrightarrow{\text{m.s.}} 0$, but$X_n \not\xrightarrow{\text{a.s.}} 0$, as $n \to \infty$. In fact, note that X_n are independent for any n, then we have

$$\begin{aligned}
\lim_{n\to\infty} P\{\bigcap_{j\geq n}(|X_j - X| \leq \epsilon)\} &= \lim_{n\to\infty} P\{\bigcap_{j\geq n}(X_j = 0)\} \\
&= \lim_{n\to\infty}(1 - \frac{1}{n})(1 - \frac{1}{n+1})\cdots \\
&= \lim_{n\to\infty} \exp\left(-\sum_{j=0}^{\infty}\frac{1}{n+j}\right) \\
&= 0 \neq 1,
\end{aligned}$$

since for $\forall j$, $(1 - \frac{1}{n+j}) \sim \exp(-\frac{1}{n+j})$ as $n \to \infty$.

Example 2.6. Revised Example 2.2.

$$X_n = \begin{cases} n & P(n) = \frac{1}{n^2} \\ 0 & P(0) = 1 - \frac{1}{n^2}. \end{cases}$$

We have $X_n \not\xrightarrow{\text{m.s.}} 0$, $X_n \xrightarrow{\text{P}} 0$, and $X_n \xrightarrow{\text{a.s.}} 0$, as $n \to \infty$. In fact,

$$\lim_{n\to\infty} P\{\bigcup_{j\geq n}(|X_j - X| > \epsilon)\} = \lim_{n\to\infty} P\{\bigcup_{j\geq n}(X_j = j)\} \leq \lim_{n\to\infty}\sum_{j\geq n}\frac{1}{j^2} = 0.$$

Proposition 2.3. *Convergence almost surely implies convergence in probability.*

Proof. Comparing (2.13) and (2.5), since

$$\{|X_n - X| < \epsilon\} \subset \bigcup_{j \geq n} \{|X_j - X| < \epsilon\},$$

we can conclude it. $\square$

Definition 2.4 (Convergence in Distribution). A sequence $\{X_n\}$ of random variables is said to *converge in distribution* to a random variable X as $n \to \infty$ if

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x), \tag{2.17}$$

where F_{X_n} and F_X are probability distribution functions of X_n and X, respectively. Denote as $\lim_{n\to\infty} X_n \overset{\mathrm{D}}{=} X$ or $X_n \xrightarrow{\mathrm{D}} X$ as $n \to \infty$.

Proposition 2.4. *Convergence in probability implies convergence in distribution. Moreover, $X_n \xrightarrow{P} X$ as $n \to \infty$ if and only if the distribution function $F_{Y_n}(y)$ of $Y_n = X_n - X$ tends to that of $Y \equiv 0$ as $n \to \infty$, i.e.,*

$$F_Y(y) = \begin{cases} 1 & y \geq 0 \\ 0 & y < 0. \end{cases}$$

Proof see Soong's book.

Remark 2.2. In sum,
convergence almost surely $\Longrightarrow$ convergence in probability $\Longrightarrow$ convergence in distribution,
mean square convergence $\Longrightarrow$ convergence in probability $\Longrightarrow$ convergence in distribution.

2.2 Definition of Stochastic Stability

For random differential equation

$$X' = F(t, X, \omega), \tag{2.18}$$

if $F(t, 0, \omega) \equiv 0$, then $X = 0$ is called an *equilibrium* of (2.18). Specially, for stochastic differential equation

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

if $f(t, 0) \equiv 0$, $\sigma(t, 0) \equiv 0$, then $X = 0$ is an equilibrium of (2.19). Note that the equilibrium is also a trivial solution of the deterministic system without random perturbation "$\sigma(t, X(t))dB(t)$". Even in the deterministic case, the concept of stability of the equilibrium can be given various meanings. The diversity is even greater in the presence of "randomness". We shall confine ourselves to those which is our view of greatest practice interest. Accordingly to different stochastic convergence in last subsection, we introduce the different stochastic stability.

In the following, we assume the $X(t) = 0$ is equilibrium of (2.18) or (2.19) unless stated otherwise. Suppose that the initial value problem of (2.18) or (2.19) with initial value $X(t_0) = X_0$ exists a unique solution for any $t \geq t_0$. Now we only consider the case that X_0 is a deterministic constant.

Definition 2.5 (Mean Square Stable). If for each $\epsilon > 0$, there exists $\delta > 0$ such that $|X_0| < \delta$ implies

$$E[|X(t, t_0, X_0)|^2] < \epsilon, \ \forall t \geq t_0, \tag{2.20}$$

then $X(t) = 0$ is said to be *mean square stable.*

Definition 2.6 (Asymptotically Mean Square Stable). If $X(t) = 0$ is mean square stable, and for sufficiently small δ, $|X_0| < \delta$ implies for any $\epsilon > 0$, there exists $T = T(\epsilon, t_0, X_0)$ such that for any $t > t_0 + T$ we have

$$E[|X(t, t_0, X_0)|^2] < \epsilon, \tag{2.21}$$

then $X(t) = 0$ is said to be *asymptotically mean square stable.*

Definition 2.7 (Exponentially Mean Square Stable). If there exist constants $A > 0$ and $\alpha > 0$ such that

$$E[|X(t, t_0, X_0)|^2] \leq A|X_0|^2 e^{-\alpha(t-t_0)}, \tag{2.22}$$

then $X(t) = 0$ is said to be *exponentially mean square stable.*

Definition 2.8 (Almost Surely Stable). If for each $\epsilon > 0$, there exists $\delta > 0$ such that $|X_0| < \delta$ implies

$$P\left\{ \bigcup_{|X_0|<\delta} \left(\sup_{t_0 \leq t < \infty} |X(t, t_0, X_0)| > \eta \right) \right\} < \epsilon, \ \forall \eta, \tag{2.23}$$

then $X(t) = 0$ is said to be *almost surely stable* or *stable with probability* 1 or *sample stable.*

Definition 2.9 (Asymptotically Almost Surely Stable). If $X(t) = 0$ is almost surely stable, and for sufficiently small δ, $|X_0| < \delta$ implies for any $\epsilon > 0$, there exists $T = T(\epsilon, t_0, X_0)$ such that for any $t > t_0 + T$ we have

$$P\{ \sup_{t > t_0 + T} |X(t, t_0, X_0)| > \eta\} < \epsilon, \forall \eta, \tag{2.24}$$

then $X(t) = 0$ is said to be *asymptotically almost surely stable.*

Definition 2.10 (Stable in Probability). If for each $\epsilon > 0$, there exists $\delta > 0$ such that $|X_0| < \delta$ implies

$$P\{|X(t, t_0, X_0)| > \eta\} < \epsilon, \ \forall t \geq t_0, \forall \eta, \tag{2.25}$$

then $X(t) = 0$ is said to be *stable in probability.*

Definition 2.11 (Asymptotically Stable in Probability). If $X(t) = 0$ is stable in probability, and for sufficiently small δ, $|X_0| < \delta$ implies for any $\epsilon > 0$, there exists $T = T(\epsilon, t_0, X_0)$ such that for any $t > t_0 + T$ and any η we have

$$P\{|X(t, t_0, X_0)| > \eta\} < \epsilon, \tag{2.26}$$

then $X(t) = 0$ is said to be *asymptotically stable in probability.*

Definition 2.12 (Asymptotically Stable in the Large). If $X(t) = 0$ is asymptotically stable in any of above sense, and δ in the definition of asymptotically stable can be chosen, namely, for any X_0, the definition still holds, then $X(t) = 0$ is said to be *asymptotically stable in the large.*

Remark 2.3. From discussion of the relation of different convergence, it is clear that both of mean square stable and almost surely stable is stronger than stable in probability.

Example 2.7. Revisit Example 1.17

$$dN(t) = rN(t)dt + \alpha N(t)dB(t), \tag{2.27}$$

and the solution

$$N(t) = N(0)\exp((r - \frac{\alpha^2}{2})t + \alpha B(t)). \tag{2.28}$$

For the second moment, we have

$$E|(N(t))^2| = |N(0)|^2 \exp\left((2r+\alpha^2)t\right), \tag{2.29}$$

therefore, $N(t) = 0$ is exponential mean square stable if and only if

$$r < -\frac{\alpha^2}{2}. \tag{2.30}$$

On the other hand, $N(t) = 0$ is asymptotically almost surely stable in the large for $r < \frac{\alpha^2}{2}$, and almost surely unstable for $r \geq \frac{\alpha^2}{2} \geq 0$ since $\lim_{t\to\infty} \frac{B(t)}{t} \overset{\text{a.s.}}{=} 0$ by strong law of large numbers.

Remark 2.4. If X is normal, i.e., $X \sim N(\mu, \sigma^2)$, then for every $p > 0$,

$$E[(e^X)^p] = \exp\left(p\mu + \frac{p^2\sigma^2}{2}\right). \tag{2.31}$$

Remark 2.5. Suppose $\{X_n\}$ be a independent sequence of real-valued identically distribution random variables defined on $(\Omega, \mathcal{F}, P)$, all with the same distribution function F, and consider the limiting behavior of the partial sum $S_n = X_1 + \cdots + X_n$. Then the strong law of large number shows that

$$\lim_{n\to\infty} \frac{S_n}{n} \overset{\text{a.s.}}{=} E[X_1]$$

if and only if $E[X_1] < \infty$.

2.3 Stability of Linear Random Differential Equation

Consider the system

$$X'(t) = A(t,\omega)X + b(t,\omega), \tag{2.32}$$

where $A : [t_0, \infty) \times \Omega \to \mathbb{R}^n \times \mathbb{R}^n$ and $b : [t_0, \infty) \times \Omega \to \mathbb{R}^n$. As in the deterministic case, a solution of (2.32) possesses one of some type of stability if and only if the trivial solution to the homogeneous equation

$$X' = A(t,\omega)X \tag{2.33}$$

possesses the type of stability. For this reason we need only investigate stability properties of the null solution of (2.33).

Theorem 2.1 (R.W. Edsinger). *Let $Y(t,\omega)$ be a fundamental matrix solution of (2.33) with $Y(t_0,\omega)$ equal to the identity matrix. Then (2.33) is:*

(a) mean square stable if and only if there exists a positive constant K such that

$$||Y(t,\cdot)||_\infty \leq K, t \geq t_0. \tag{2.34}$$

(b) uniformly mean square stable if and only if there exists a positive constant K such that

$$||Y(t,\cdot)Y^{-1}(s,\cdot)||_\infty \leq K, \quad \text{for} \ \ t_0 \leq s \leq t < \infty. \tag{2.35}$$

(c) asymptotically mean square stable if and only if

$$||Y(t,\cdot)||_\infty \to 0 \quad \text{as } t \to \infty. \tag{2.36}$$

(d) uniformly asymptotically mean square stable if and only if there exists $K > 0$, $\alpha > 0$ such that

$$||Y(t,\cdot)Y^{-1}(s,\cdot)||_\infty < Ke^{-\alpha(t-s)}, \quad \text{for} \ \ t_0 \leq s \leq t < \infty. \tag{2.37}$$

Here for any random variable R, $||R||_\infty$ is the essential supremum of $|R|$.

Remark 2.6. R.W. Edsinger prove the analogous theorem for the mean stability of linear systems with random coefficients to the famous W. Coppel's theorem, namely, "stability" is equivalent to the "boundedness" of a fundamental matrix solution for linear systems. See JDE 1970.

Remark 2.7. In fact, R.W. Edsinger prove a more generalized conclusion than before. Mean square stability can be changed to $L(p,q)$ stable, where $p \geq q \geq 1$, associated $||\cdot||_\infty$ to $||\cdot||_r$, where $r = \frac{pq}{p-q}$ ($r = \infty$ if $p = q$), then the conclusion still holds. Here $L(p,q)$ stable means that $||X_0||_p < \delta$ implies $||X(t,t_0,X_0)||_q < \epsilon$, namely, for any random initial condition X_0, changing $|X_0|$ to $||X_0||_p$ and changing $||X(t,t_0,X_0)||_2$ to $||X(t,t_0,X_0)||_q$.

In the following, consider the "autonomous" linear system of (2.33),

$$X' = A(\omega)X, \tag{2.38}$$

where we assume $A(\omega)$ is essentially bounded, i.e., there exists $L > 0$ and $S \subset \Omega$ such that $P(S) = 1$, and $|A(\omega)| \leq L$ for all $\omega \in S$.

Theorem 2.2 (T. Morozan). *The following asserations are equivalent:*

(1) The matrix $A(\omega)$ is P-stable, i.e., there exists $\alpha > 0$ such that $P\{\omega : \max_{\lambda \in S(A(\omega))} Re\lambda < -\alpha\} = 1$ where $S(A(\omega))$ is the spectrum of $A(\omega)$.

(2) There exists $K > 0$, $\alpha > 0$, $D \subset \Omega$, $P(D) = 1$ such that

$$|e^{A(\omega)t}| \leq Ke^{-\alpha t} \quad \textit{for all} \quad t \in \mathbb{R}^n, \omega \in D. \tag{2.39}$$

(3) The null solution of system (2.38) is exponentially mean square stable.

The proof see JDE 1967.

Just in the deterministic case, stability of a nonlinear equation is in general difficult to prove. However, the proof is facilitated by the fact the equation linearized by means of a Taylor expansion usually exhibits in the neighborhood of the equilibrium position the same behavior as regards stability as does the original equation. We mention the following theorem without proof.

Theorem 2.3 (R.Z. Khas'minskiǐ). *Suppose that*

$$dX(t) = f(t, X(t))dt + \sigma(t, X(t))dB(t), X(t_0) = X_0 \tag{2.40}$$

is a Ito stochastic differential equation with $f(t,0) \equiv 0$ and $\sigma(t,0) \equiv 0$. Suppose that

$$|f(t,x) - A(t)x| = o(|x|)$$

and

$$|\sigma(t,x) - L(t)x| = o(|x|),$$

uniformly in $t \geq t_0$ as $|x| \to 0$. Let $A(t)$ and $L(t)$ denote $n \times n$ matrices that are bounded functions of t. Consider the linear equation

$$dX(t) = A(t)X(t)dt + L(t)X(t)dB(t), X(t_0) = X_0. \tag{2.41}$$

If the equilibrium position of (2.41) is asymptotically almost surely stable (or in probability), then the solution $X(t) = 0$ of the system (2.40) is asymptotically stable in probability.

2.4 Lyapunov Direct Method

For theory of deterministic ordinary differential equations, stability of the system, i.e. the behavior of "perturbed" orbits relative to original orbits, is an important research field. M.A. Lyapunov proposed two methods to study stability in his seminal thesis in 1892, namely, first method and second method.

Lyapunov's first method: The method of linearization of the nonlinear equation along an orbit, and the transfer of stability from the linear to the nonlinear equation.

Lyapunov's second method: The method of Lyapunov functions, i.e. of scalar functions on the state space which decrease along orbits. The biggest advantage of this method is that it is "direct", hence one does not need to solve the equation explicitly. The biggest drawback is that there is no general method for obtaining Lyapunov functions.

Consider stochastic differential equation

$$dX(t) = f(t, X(t))dt + \sigma(t, X(t))dB(t). \tag{2.42}$$

Set $v(t,x)$ be a positive definite function continuously differentiable with respect to t and twice continuously differentiable with respect to the x. By Ito's formula, we have

$$dV(t) = dv(t, X(t)) = [\frac{\partial v}{\partial t} + f\frac{\partial v}{\partial x} + \frac{\sigma^2}{2}\frac{\partial^2 v}{\partial x^2}]dt + \sigma\frac{\partial v}{\partial x}dB(t). \tag{2.43}$$

This would mean that the ordinary definition of stability holds for each single path $X(\omega)$. However, due to the presence of the fluctuational term in (2.43), the condition $dV(t) \le 0$ can be satisfied only in degenerate cases. Therefore, it makes sense to require instead that $X(t)$ not run "up hill" on the average, that is,

$$E[dV(t)] \le 0.$$

Since

$$E[dV(t)] = E[Lv(t, X(t))dt],$$

where

$$L := \frac{\partial}{\partial t} + f\frac{\partial}{\partial x} + \frac{\sigma^2}{2}\frac{\partial^2}{\partial x^2}, \tag{2.44}$$

the requirement will be satisfied if

$$Lv(t,x) \le 0, \quad \text{for all } t \ge 0, x \in \mathbb{R}. \tag{2.45}$$

We shall refer to the function $v(t,x)$ used here as a *Lyapunov function* corresponding to the stochastic differential equation (2.42).

Theorem 2.4. *Suppose that there exists a positive definite function $v(t,x)$ defined on a half-cylinder $[t_0,\infty) \times U_h, U_h = \{x \in \mathbb{R}^n : |x| < h\}$, where $h > 0$, that is everywhere, with the possible exception of the point $x = 0$, continuously differentiable with respect to t and twice continuously differentiable with respect to the component x_i of x.*

(1) If

$$Lv(t,x) \le 0 \quad \textit{for all } t \ge t_0, 0 < |x| \le h, \tag{2.46}$$

where

$$L = \frac{\partial}{\partial t} + \sum_{i=1}^{n} f_i(t,x)\frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(\sigma(t,x)(\sigma(t,x))^T)_{ij}\frac{\partial^2}{\partial x_i \partial x_j}. \tag{2.47}$$

Then, the equilibrium $X = 0$ of (2.42) is almost surely stable.

(2) If $v(t,x)$ is decrescent and $Lv(t,x)$ is negative definite, then the equilibrium $X = 0$ is asymptotically almost surely stable.

Theorem 2.5. *Suppose that there exists a function $v(t,x)$ defined on a half-cylinder $[t_0,\infty)\times U_h, U_h=\{x\in\mathbb{R}^n : |x|<h\}$, where $h>0$, that is everywhere, with the possible exception of the point $x=0$, continuously differentiable with respect to t and twice continuously differentiable with respect to the component x_i of x. If*

$$\lim_{x\to 0}\inf_{t\geq t_0} v(t,x)=\infty \tag{2.48}$$

and

$$\sup_{t_0\leq t,\epsilon<|x|\leq h} Lv(t,x)<0 \quad \text{for all } 0<\epsilon<h, \tag{2.49}$$

or, instead of (2.49), only $Lv(t,x)\leq 0$ and

$$y\sigma(t,x)(\sigma(t,x))^T y^T \geq m(x)|y|^2 \quad \text{for all } y\in\mathbb{R}^n, |x|\leq h, t\geq t_0, \tag{2.50}$$

where m is a positive definite function, then the equilibrium $X=0$ of (2.42) is almost surely unstable.

Theorem 2.6. *Suppose $v(t,x)$ be continuous on $[t_0,\infty)\times\mathbb{R}^n$, that for all $x\neq 0$ is once continuously differentiable with respect to t and twice continuously differentiable with respect to x_i. If $v(t,x)$ satisfies the inequalities*

$$c_1|x|^2\leq v(t,x)\leq c_2|x|^2 \tag{2.51}$$

and

$$Lv(t,x)\leq -c_3|x|^2 \tag{2.52}$$

for certain positive constant c_1, c_2, and c_3. Then the equilibrium $X=0$ of (2.42) is exponentially mean square stable.

Example 2.8. Consider the stochastic stability of growth model using Lyapunov direct method again,

$$dN(t)=rN(t)dt+\alpha N(t)dB(t). \tag{2.53}$$

Let $v(t,x)=|x|^\lambda$, where $\lambda>0$. Thus

$$Lv(t,x)=L(|x|^\lambda)=(r+\frac{1}{2}\alpha^2(\lambda-1))\lambda|x|^\lambda.$$

If $r<\frac{\alpha^2}{2}$, then we can choose λ such that $0<\lambda<1-\frac{2r}{\alpha^2}$ and hence

$$Lv\leq -kv.$$

Therefore, by Theorem 2.4, $X=0$ is asymptotically almost surely stable. If $r\geq\frac{\alpha^2}{2}$, then we can choose $v(t,x)=-\log|x|$ and we get $Lv=-r+\frac{\alpha^2}{2}\leq 0$. Since condition (2.50) is satisfied for $\alpha\neq 0$ with $m(x)=\alpha^2|x|^2$, by Theorem 2.5, the equilibrium $X=0$ is almost surely unstable.

Theorem 2.4, 2.5, and 2.6 still hold for multi-dimensional system except changing different definition of L. In fact, from n-dimensional Ito formula, i.e., Theorem 1.1.3, it is clear that define L as follows:

$$L:\frac{\partial}{\partial t}+\sum_{i=1}^n f_i\frac{\partial}{\partial x_i}+\frac{1}{2}\sum_{i,j=1}^n\sigma_i\sigma_j\frac{\partial^2}{\partial x_i\partial x_j}, \tag{2.54}$$

if we take $p=1$ and $m=1$.

Example 2.9. Consider the second order ordinary differential equation

$$x''(t)+3x'(t)+2x(t)=0, \tag{2.55}$$

namely, in the form of system,

$$\begin{cases} x'(t) = y(t) \\ y'(t) = -2x(t) - 3y(t). \end{cases} \tag{2.56}$$

Clearly, the equilibrium $(0,0)$ of system (2.56) is asymptotically stable. Now consider the new system with stochastic perturbation,

$$X''(t) + 3X'(t) + 2X(t) = \sigma(X, X', t)B'(t), \tag{2.57}$$

or in the form of system

$$\begin{cases} dX(t) = Y(t)dt \\ dY(t) = -2X(t)dt - 3Y(t)dt + \sigma(X,Y,t)dB(t), \end{cases} \tag{2.58}$$

where $B(t)$ is standard Brownian motion, $|\sigma(X,Y,t)| \leq \epsilon(|X| + |Y|)$, $0 < \epsilon \ll 1$. Now want to show the stability of $(0,0)$. Let $v(x,y) = 8x^2 + 4xy + 2y^2 \geq x^2 + y^2$. Then we have

$$\begin{aligned} Lv(x,y) &= 0 + y(t)\tfrac{\partial v}{\partial x} + (-2x - 3y)\tfrac{\partial v}{\partial y} + \tfrac{1}{2}\sigma^2 \tfrac{\partial^2 v}{\partial y^2} \\ &= -8x^2 - 4xy - 8y^2 + 2\sigma^2 \\ &\leq -v(x,y) + 2\epsilon^2(|x| + |y|)^2 \\ &\leq -(1 - 4\epsilon^2)v(x,y). \end{aligned}$$

Thus $(0,0)$ is asymptotic almost surely stable.

2.5 Stochastic stabilization and destabilization

As we know, most of random differential equations can be looked as perturbation system of deterministic system with noise. Then, how does the stability of system change under noise? In his book, X. Mao(1994) gave some examples and theorems to show that it is possible either to stabilize or to destabilize under appropriate noise.

First, it has been observed that noise can have a stabilizing effect. For instance, consider an unstable system

$$x'(t) = x(t), t \geq 0, x(0) = x_0 \in \mathbb{R}. \tag{2.59}$$

Perturb this system by noise and say the perturbed system has the form

$$dX(t) = X(t)dt + 2X(t)dB(t), t \geq 0, X(0) = X_0 \in \mathbb{R}. \tag{2.60}$$

From Example 2.8, since $r = 1$, $\alpha = 2$, $r < \frac{\alpha^2}{2}$, $X = 0$ is asymptotically almost surely stable.

On the other hand, it has also been observed that noise can destabilize a stable system. For example, suppose that a given exponentially stable system

$$x'(t) = -x(t), t \geq 0, x(0) = x_0 \in \mathbb{R} \tag{2.61}$$

is perturbed by noise and the stochastically perturbed system is described by a one-dimensional Ito equation

$$dX(t) = -X(t)dt + \epsilon dB(t), t \geq 0, X(0) = X_0, \tag{2.62}$$

where $\epsilon > 0$. It can be shown that the solution satisfies

$$\limsup_{t \to \infty} |X(t)| \overset{\text{a.s.}}{=} \infty. \tag{2.63}$$

Proof see Mao's book. Moreover, in his book, there are more conclusions about stochastic stabilization and destabilization.

References

[1] Ludwig Arnold, *Stochastic differential equations: Theory and Application*, John Wiley & Sons, New York, 1974.

[2] Ludwig Arnold, *Random dynamical systems*, Springer, New York, 1998.

[3] Rabi N. Bhattacharga, Edward C. Waymire, *Stochatic processes with applications*, John Wiley & Sons, New York, 1990.

[4] Vivek S. Borkar, *Probability theory: An advanced course*, Springer, New York, 1995.

[5] Yuan Shih Chow, Henry Teicher, *Probability theory: Independence interchangeability martingales*, Springer, New York, 1978.

[6] I. Chueshov, *Monotone random system*, Lecture Notes in Mathematics, Vol. 1779, Springer, New York, 2002.

[7] Kai Lai Chung, *Elementary probability theory with stochastic processes*, Springer, New York, 1979.

[8] Nguyen Dinh Cong, *Topological dynamics of random dynamical systems*, Oxford Univ. Press, Oxford, 1997.

[9] Ruth F. Curtain(Edited), Stability of stochastic dynamical systems, *Lecture Notes in Mathematics*, Vol. **294**, Springer, Berlin, 1972.

[10] Richard Durrett, *Stochastic calculus*, CRC Press, Boca Raton, 1996.

[11] K. D. Elworthy, *Stochastic differential equations on manifolds*, Cambridge Univ. Press, Cambridge, 1982.

[12] Avner Friedman, *Stochastic differential equations and applications*, Volume 1, 2, Academic Press, New York, 1975.

[13] Kiyosi Itô(Edited), *Proceedings of the internatial symposium on stochastic differential equations, Kyoto, 1976*, John Wiley & Sons, New York, 1978.

[14] Gopinath Kallianpur, Jie Xiong, *Stochastic differential equations in infinite dimensional spaces*, Institute of Mathematical Statistics, USA, 1995.

[15] R. Z. Khas'minskiĭ, *Stochastic stability of differential equations*, Sijthoff & Noordhoff, Netherlands, 1980.

[16] Hiroshi Kunita, *Stochastic flows and stochastic differential equations*, Cambridge Univ. Press, Cambridge, 1990.

[17] Harold J. Kushner, *Stochastic stability and control*, Academic Press, New York, 1967.

[18] Harold J. Kushner, *Introduction to stochastic control*, Holt, Rinehart & Winston, New York, 1971.

[19] Harold J. Kushner, *Weak convergence methods and singularly perturbed stochastic control and filtering problems*, Birkhäuser, Boston, 1990.

[20] John W. Lamperti, *Probability: A survey of the mathematical theory*, Second Edition, John Wiley & Sons, New York, 1996.

[21] Pei-Dong Liu, Min Qian, *Smooth ergodic theory of random dynamical systems*, Lecture Notes in Mathematics, Vol. 1606, Springer, New York, 1995.

[22] Xuerong Mao, *Stability of stochastic differential equations with respect to semimartingales*, Longman, 1991.

[23] Xuerong Mao, *Exponential stability of stochastic differential equations*, Marcel Dekker, New York, 1994.

[24] Bernt Øksendal, *Stochastic differential equations: An introduction with applications*, First Edition, 1985, Fourth Edition, Springer, Berlin, 1995.

[25] Yu A. Rozanov, *Probability theory, random processes and mathematical statistics*, Kluwer Academic Publishers, Dordrecht, 1995.

[26] Thomas L. Saaty, *Modern nonlinear equations*, McGraw-Hill Book Company, New York, 1967. Revision by Dover Publications, New York, 1981.

[27] A.V. Skorokhod, *Asymptotic methods in the theory of stochastic differential equations*, "Translations of mathematicsal monographs" Vol. **78**, AMS, Providence, Rhode Island, 1989.

[28] T.T. Soong, *Random differential equations in science and egineering*, Academic Press, New York, 1973.

[29] S.K. Srinivasan, R. Vasuderan, *Introduction to random differential equations and their applications*, Elsevier, New York, 1971.

[30] J.L. Strand, *Stochastic ordinary differential equations*, Ph.D. Thesis, Univ. of California (Berkeley), 1968.

[31] Rangquan Wu, *Stochastic differential equations*, Pitman Advanced Publishing, Boston, 1985.

[32] Ludwig Arnold, A formula connecting sample and moment stability of linear stochastic systems, *SIAM J. Appl. Math.*, **44**, 793–802, 1984.

[33] Ludwig Arnold, *Random dynamical systems*, 1–43, in "Dynamical systems", edited by R. Johnson, Lecture Notes in Mathematics, Vol. 1609, Springer, New York, 1994.

[34] Ludwig Arnold, Björn Schmalfuss, Lyapunov's second method for random dynamical systems, *J. Diff. Eq.*, **177**, 235–265, 2001.

[35] K. Benjelloun, E.K. Boukas, Mean square stochastic stability of linear time-delay ststem with Markovian jumping parameters, *IEEE Transactions on Automatic Control*, **43**, No. 10, 1456–1460, 1998.

[36] Arno Berger, Stefan Siegmund, On the gap between random dynamical systems and continuous skew products, *J. Dynamics and Diff. Eq.*, **15**, 237–279, 2003.

[37] Nils Berglund, Barbara Gentz, Geometric singular perturbation theory for stochastic differential equations, *J. Diff. Eq.*, **191**, 1–54, 2003.

[38] Rabi Bhattacharya, Mukul Majumdar, Random dynamical systems: a review, *Economic Theory*, **23**, 13–38, 2004.

[39] Christian Bucher, Stochastic stability of systems with random properties, 39–49, in "Nonlinear dynamical and stochastic mechanics", edited by W.H. Kliemann etc., *Fields Institute Communications*, Vol. **9**, AMS, Providence, Rhode Island, 1996.

[40] R.W. Edsinger, Mean stability of linear ordinary differential equations with random coefficients, *J. Diff. Eq.*, **8**, 448–456, 1970.

[41] Neil Fenichel, Geometric singular perturbation theory for ordinary differential equations, *J. Diff. Eq.*, **31**, 53–98, 1979.

[42] Patrick Florchinger, Lyapunov-like techniques for stochastic stability, *SIAM J. Control and Optimization*, **33**, No. 4, 1151–1169, 1995.

[43] Tetsuo Furumochi, Mean stability of quasi-linear systems with random coefficients, 298–301, in "Japan-United States seminar on ordinary differential and functional equations", *Lecture Notes in Mathematics*, Vol. **243**, Springer, Berlin, 1971.

[44] Jerome A. Goldstein, An existence theorem for linear stochastic differential equations, *J. Diff. Eq.*, **3**, 78–87, 1967.

[45] John A. Jacquez, Carl P. Simon, The stochastic SI model with recruitment and deaths: I. Comparison with the closed SIS model, *Math. Biosci.*, **117**, 77–125, 1993.

[46] Daqing Jiang, Ningzhong Shi, A note on nonautonomous logistic equation with random pertubation, *J. Math. Anal. Appl.*, **303**, 164–172, 2005.

[47] R.Z. Khas'minskii, Necessary and sufficient conditions for the asymptotic stability of linear stochastic systems, 144–147,??

[48] Yuri Kifer, Principal eigenvalues, topological pressure, and stochastic stability of equilibrium states, *Israel J. Math.*, **70**, No. 1, 1–47, 1990.

[49] F. Kozin, S. Prodromou, Necessary and sufficient conditions for almost sure sample stability of linear Ito equations, *SIAM J. Appl. Math.*, **21**, 413–424, 1971.

[50] Y.K. Lin, G.Q. Cai, Stochastic stability of non-linear systems, *Int. J. Non-Linear Mechanics*, **29**, No. 4, 539–553, 1994.

[51] Xuerong Mao, Glenn Marion, Eric Renshaw, Environmental Brownian noise suppresses explosious in population dynamics, *Stochastic Processes and Their Applications*, **97**, 95–110, 2002.

[52] Xuerong Mao, Sotirios Sabanis, Eric Renshaw, Asymptotic behavior of the stochastic Lotka-Volterra model, *J. Math. Anal. Appl.*, **287**, 141–156, 2003.

[53] Xuerong Mao, Chenggui Yuan, Jiezhong Zou, Stochastic differential delay equations of population dynamics, *J. Math. Anal. Appl.*, **304**, 296–320, 2005.

[54] Janusz Mierczyinski, Wenxian Shen, Xiao-Qiang Zhao, Uniform persistence for nonautonomous and random parabolic Kolmogorov systems, *J. Diff. Eq.*, **204**, 471–510, 2004.

[55] Richard R. Mitchell, Frank Kozin, Sample stability of second order linear differential equations with wide band noise coefficients, *SIAM J. Appl. Math.*, **27**, 571–605, 1974.

[56] Salah-Eldin A. Mohammed, Michael K. R. Schentzow, The stable manifold theorem for stochastic differential equations, *The Annals of Probability*, **27**, No. 2, 615–652, 1999.

[57] T. Morozan, Stability of some linear stochastic systems, *J. Diff. Eq.*, **3**, 153–169, 1967.

[58] T. Morozan, Stability of linear systems with random parameters, *J. Diff. Eq.*, **3**, 170–178, 1967.

[59] Mark A. Pinsky, Stochastic stability and boundary problems, 105–109, in "Probability: Proceeding of the symposium in pure mathematics of the American Mathematical Society", edited by Joseph L. Doob, *Proceedings of Symposia in Pure Mathematics*, Vol. **31**, AMS, Providence, Rhode Island, 1977.

[60] J.L. Strand, Random ordinary differential equations, *J. Diff. Eq.*, **7**, 553–583, 1970.

[61] Toufic M. Suidan, Stationary measures for a randomly forced Burgers equation, *Communications on Pure and Applied Mathematics*, **LVIII**, 0620–0638, 2005.

[62] E.C. Zeeman, Stochastic stability and classification of dynamical systems, 76–82, in "Workshop on dynamical systems" edited by Z. Coelho, E. Shiels, *Pitman research notes in mathematics series*, Vol. **221**, Longman Scientific & Technical, Harlow, Essex, UK, 1990.